Комп'ютинг

В. А. Андруник, В. А. Висоцька,
В.В. Пасічник, Л.Б. Чирун, Л.В. Чирун

ЧИСЕЛЬНІ МЕТОДИ В КОМП'ЮТЕРНИХ НАУКАХ

Том 2



omputing

Навчальний посібник

Бібліотека
Галицького коледжу
ім. В.Чорновола
38932
0
Чисельні методи в комп'ютерні

004.93
Ч-66

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ "ЛЬВІВСЬКА ПОЛІТЕХНІКА"

В.А. Андруник, В.А. Висоцька,
В.В. Пасічник, Л.Б. Чирун, Л.В. Чирун

ЧИСЕЛЬНІ МЕТОДИ В КОМП'ЮТЕРНИХ НАУКАХ

Том 2

Навчальний посібник

за науковою редакцією д.т.н., професора,
Лауреата державної премії України у галузі науки та техніки
В.В. Пасічника

№ 38952

Львів – 2020
Видавництво «Новий Світ – 2000»

УДК 81:004.93
ББК 22.19я73
А 64

Гриф надано Науково-методичною радою
Національного університету «Львівська політехніка»
(Протокол № 28 від 27.04.2017р.)

Рецензенти:

Гожий О.П. – доктор технічних наук, професор кафедри комп'ютерної інженерії Чорноморського національного університету імені Петра Могили;

Дияк І.І. – кандидат фізико-математичних наук, доцент, декан факультету прикладної математики та інформатики Львівського національного університету імені Івана Франка;

Дронюк І.М. – кандидат фіз.-мат. наук, доцент кафедри автоматизованих систем управління Національного університету «Львівська політехніка»;

Квєтний Р.Н. – доктор технічних наук, професор, завідувач кафедри автоматики та інформаційно-вимірювальної техніки Вінницького національного технічного університету, член-кореспондент Національної академії педагогічних наук України;

Шаронова Н.В. – доктор технічних наук, професор, завідувач кафедри інтелектуальних комп'ютерних систем Національного технічного університету «Харківський політехнічний інститут».

Андрунник В.А., Висоцька В.А., Пасічник В.В., Чирун Л.Б., Чирун Л.В. Чисельні методи в комп'ютерних науках: навчальний посібник, Том 2 за ред. В.В. Пасічника – Львів: Видавництво «Новий Світ – 2000», 2020 – 805 с. розділ 1-3, комп'ютерний проект 1-2 - 536 с.

ISBN 978-617-7519-12-5

Навчальний посібник містить матеріал для вивчення основних теоретичних засад, функціональних можливостей та практичного застосування теорії чисельних методів, розроблення прикладних засобів та інформаційних систем аналізу та опрацювання інформації за допомогою чисельних методів. Теоретичний та практичний матеріал викладено у доступній формі. Виклад матеріалу супроводжується значною кількістю прикладів, що полегшує його сприйняття та засвоєння. Подається перелік питань та тестів для самоконтролю, а також завдання для самостійного виконання трьох рівнів складності та довідкова інформація для розв'язування задач. Навчальний посібник призначається для студентів, що навчаються за спеціальностями 122 «Комп'ютерні науки» та 124 «Системний аналіз» і споріднених спеціальностей, які пов'язані з вивченням чисельних методів в інформатиці та інформаційних технологіях. Може бути використаний аспірантами в якості підгрунтя для наукових досліджень та викладачами в якості дидактичного матеріалу, а також для самостійного вивчення та підвищення кваліфікації. Книга призначена для спеціалістів із проектування, розроблення та впровадження інтелектуальних систем опрацювання інформаційних ресурсів, науковців в галузі глобальних інформаційних системи, систем штучного інтелекту, Інтернет-технологій, фахівців з електронної комерції, Інтернет-маркетингу та Інтернет-реклами, менеджерів комплексних Web-проектів, а також для здобувачів 3-ого (освітньо-наукового) рівня вищої освіти в галузі знань 12 «Інформаційні технології».

ISBN 978-617-7519-12-5

© Андрунник В.А, Висоцька В.А., Пасічник В.В., Чирун Л.Б., Чирун Л.В., 2020
© ФОП Піча С.В., Видавництво «Новий Світ – 2000», 2020

Зміст

Передмова наукового редактора серії підручників та навчальних посібників «КОМП'ЮТИНГ» ... 10
Вступне слово авторів ... 15
Розділ 1. Чисельні методи розв'язування звичайних диференціальних рівнянь ... 24
1.1. Загальна постановка задачі ... 25
1.2. Метод Ейлера ... 27
1.2.1. Загальна постановка задачі ... 27
1.2.2. Графічна інтерпретація ... 28
1.3. Метод Рунге-Кутта ... 30
1.3.1. Загальна постановка задачі ... 30
1.3.2. Метод Рунге-Кутта другого порядку точності ... 31
1.4. Багатокрокові методи розв'язування звичайних диференціальних рівнянь ... 33
1.4.1. Метод Рунге-Кутта третього порядку точності ... 33
1.4.2. Метод Рунге-Кутта четвертого порядку точності ... 35
1.5. Особливості методу Рунге-Кутта ... 36
1.5.1. Ідея методу Рунге-Кутта ... 36
1.5.2. Загальний підхід отримання формул Рунге-Кутта ... 38
1.5.3. Часткові випадки методу Рунге-Кутта ... 39
1.6 .Різницеві методи розв'язування задачі Коші для звичайних диференціальних рівнянь (ЗДР) ... 44
1.6.1. Екстраполяційні та інтерполяційні різницеві формули ... 44
1.6.2. Часткові випадки різницевих формул. Екстраполяційні формули Адамса ... 46
1.6.3. Часткові випадки різницевих формул. Інтерполяційні формули Адамса ... 48
1.7. Методи скінчених різниць розв'язування крайових задач для диференціальних рівнянь ... 49
1.7.1. Загальна постановка задачі ... 49
1.7.2. Основні різницеві співвідношення ... 50
1.7.3 .Метод скінченних різниць розв'язування крайової задачі для диференціальних рівнянь ... 52
1.7.4. Метод прогонки розв'язування апроксимуючої СЛАР для крайової задачі диференціальних рівнянь з точністю $o(h)$ ... 52
1.7.5. Метод прогонки розв'язування апроксимуючої СЛАР з точністю $o(h^2)$ ... 55
1.8. Чисельні методи розв'язування крайових задач для диференціальних рівнянь в частинних похідних. Метод сіток розв'язування крайових задач для диференціальних рівнянь еліптичного типу ... 55
1.8.1. Метод розв'язування задачі Діріхле ... 55
1.8.2. Розв'язування рівняння Пуассона ... 58
1.9. Метод невизначених коефіцієнтів розв'язування задачі Діріхле (метод Рейхенбаха) ... 59
1.10. Часткові випадки методу Рейхенбаха ... 60
1.10.1. Перша різницева схема розв'язування рівняння Пуассона ... 60
1.10.2. Друга різницева схема розв'язування рівняння Пуассона ... 61
1.10.3. Третя різницева схема розв'язування рівняння Пуассона ... 62
1.11 Апроксимація крайових умов методом Коллатца ... 63

1.12. Розв'язування різницевих рівнянь ... 64
1.12.1. Різницевий принцип максимуму ... 64
1.12.2. Розв'язування різницевої системи ... 65
1.13. Розв'язування рівнянь гіперболічного типу ... 66
1.13.1. Метод сіток розв'язування задачі Коші для рівнянь гіперболічного типу ... 66
1.13.2. Вибір сітки для рівнянь гіперболічного типу ... 68
1.14. Метод сіток розв'язування лінійного рівняння параболічного типу (для розв'язування задачі Коші) ... 69
Контрольні питання ... 71
Задачі для самостійної роботи ... 71
Завдання до лабораторної роботи ... 72

Розділ 2. Чисельні методи розв'язування задач оптимізації ... 73
2.1. Загальна постановка задачі оптимізації ... 75
2.1.1. Визначення цільової функції ... 75
2.1.2. Формальне визначення задачі оптимізації ... 77
2.1.3. Класифікація задач оптимізації ... 78
2.1.4. Метод переведення критеріїв в обмеження ... 80
2.1.5. Метод контрольних показників ... 80
2.1.6. Задача багатокритеріального математичного програмування ... 81
2.1.7. Деякі методи багатокритеріальної оптимізації ... 84
2.1.8. Метод квазіоптимізації локальних критеріїв (метод послідовних поступок) ... 88
2.1.9. Метод згортання векторного критерію в суперкритерій ... 90
2.2. Безумовна оптимізація ... 90
2.1. Метод штрафних функцій ... 91
2.2. Геометрична інтерпретація штрафних функцій ... 93
2.3. Аналіз алгоритмів розв'язування задачі методом штрафних функцій ... 95
2.4. Обчислювальні складнощі, що пов'язані зі штрафними функціями ... 99
2.5. Аналіз та інтерпретація отриманих результатів ... 104
2.3. Багатокритеріальні задачі оптимізації та основні підходи до їх розв'язування ... 105
2.3.1. Основні поняття багатокритеріальної задачі оптимізації ... 105
2.3.2. Геометрична ілюстрація багатокритеріальної задачі ... 105
2.3.3. Постановка задачі багатокритеріальної оптимізації ... 108
2.3.4. Еталонні точки ... 108
2.3.5. Критерії оптимальності. Критерій Парето ... 108
2.3.6. Лексикографічний порядок ... 109
2.3.7. Скаляризація ... 109
2.3.8. Зважена сума ... 110
2.3.9. Функція скаляризації Чебишева ... 110
2.3.10. Метод зміни обмежень (ε-обмеження) ... 110
2.3.11. Методи розв'язування. Інтерактивність. Еволюційні методи ... 111
2.4. Лінійне програмування ... 111
2.4.1. Місце лінійного програмування в системі задач математичного програмування ... 111
2.4.2. Формальна постановка задачі лінійного програмування ... 112
2.4.3. Чисельні методи оптимізації. Методи прямого пошуку ... 113
2.4.4. Лінійний пошук без використання похідних ... 116
2.4.5. Інтервал невизначеності ... 117
2.4.6. Функції однієї змінної ... 118

2.4.7. Функції *n* змінних ... 119
2.4.8. Методи прямого пошуку для функції однієї змінної ... 120
2.4.9. Задача оптимізації методом Ньютона ... 120
2.5. Порівняння методів лінійного пошуку без розрахунку похідної ... 122
2.5.1. Класифікація методів лінійного пошуку без розрахунку похідної ... 122
2.5.2. Рівномірний пошук ... 123
2.5.3. Послідовний пошук ... 123
2.5.4. Задача оптимізації методом дихотомічного пошуку (ділення пополам) ... 123
2.5.5. Задача оптимізації за методом золотого січення ... 126
2.5.6. Задача оптимізації за методом Фібоначчі ... 130
2.5.7. Методи апроксимації. Квадратична інтерполяція ... 136
2.6. Аналіз методів та алгоритмів розв'язування задачі нелінійного програмування ... 138
2.6.1. Класифікація методів нелінійного програмування ... 138
2.6.2. Формальна постановка задачі з використанням методів прямого пошуку для оптимізації унімодальних функцій якості без обмежень ... 139
2.6.3. Визначення та основні властивості опуклих функцій ... 142
2.7. Методи прямого пошуку для функцій *n* змінних ... 145
2.7.1. Метод покоординатного спуску ... 147
2.7.2. Метод Хука – Дживса ... 147
2.7.3. Метод Нелдера – Міда ... 149
2.7.4. Метод циклічного покоординатного спуску ... 156
2.7.5. Метод Розенброка ... 160
2.7.6. Метод спряжених градієнтів Флетчера та Рівса ... 166
Контрольні питання ... 176
Задачі для самостійної роботи ... 176

Розділ 3. Основи статистичного моделювання ... 187
3.1. Моделювання випадкових величин на комп'ютері ... 190
3.1.1. Моделювання на комп'ютері рівномірно розподіленої величини α на [0,1], [a,b] ... 190
3.1.2. Моделювання на комп'ютері випадкової величини з довільним законом розподілу ... 194
3.1.3. Моделювання нормально-розподіленої випадкової величини ξ з $M\xi=0$, $D\xi=1$ ... 197
3.2. Обчислення інтегралів методом статистичного моделювання ... 197
3.2.1. Аналітичний метод обчислення однократного інтеграла ... 197
3.2.2. Метод площ обчислення визначеного інтеграла ... 198
3.2.3. Обчислення кратних інтегралів ... 198
3.2.3.1. Метод *n*-кратних об'ємів ... 199
3.2.3.2. Метод середнього значення (аналітичний метод) ... 200
3.3. Розв'язання систем лінійних алгебраїчних рівнянь ... 200
3.3.1. Метод Дж. фон Неймана-Улама ... 200
3.3.2. Побудова імітаційної моделі розв'язання системи лінійних алгебраїчних рівнянь (СЛАР) ... 204
3.3.3. Випадкове блукання частинки з поглинаючим екраном ... 206
3.3.4. Теорема про побудову $X_t(\omega)$ ... 208
3.3.5. Обчислення оберненої матриці методом статистичного моделювання ... 209
3.3.6. Алгоритм методу статистичного моделювання розв'язку СЛАР ... 210
3.4. Розв'язок першої крайової задачі для рівняння $\Delta u=f$ ... 211

3.5. Марківські ланцюги та інтегральні рівняння ... 214
3.5.1. Загальна постановка задачі ... 214
3.5.2. Лінійні оцінки та умови неусувності ... 215
3.6. Методи статистичного моделювання для розв'язування стохастичних диференціальних рівнянь ... 218
3.6.1. Загальна постановка задачі ... 218
3.6.2. Збіжність методу Ейлера в середньому квадратичному ... 221
3.6.3. Збіжність методу Ейлера як збіжність розподілів ... 224
3.6.4. Оптимальні різницеві схеми розв'язку стохастичних рівнянь ... 225
3.7. Наближене інтегрування стохастичних диференціальних рівнянь ... 227
3.7.1. Однокрокова апроксимація ... 228
3.7.2. Апроксимація на скінченному проміжку ... 230
3.7.3. Узагальнення на системи рівнянь ... 232
3.8. Методи побудови емпіричних формул за вибірковими значеннями ... 233
3.8.1. Загальна постановка задачі ... 233
3.8.2. Лінійна залежність ... 235
3.8.3. Метод вирівнювання ... 236
3.8.4. Квадратична (параболічна) залежність ... 239
3.8.5. Визначення параметрів емпіричної формули ... 241
3.8.6. Метод вибраних точок ... 242
3.8.7. Метод середніх ... 243
3.8.8. Метод найменших квадратів ... 245
3.8.9. Деякі міркування про вибір емпіричної формули з двома параметрами ... 249
3.8.10. Емпіричні формули, які містять три параметри ... 254
3.8.11. Уточнення отриманої емпіричної формули ... 258
3.8.12. Загальний метод визначення параметрів емпіричної формули ... 260
3.9. Метод статистичного аналізу економетрики ... 264
3.9.1. Метод кореляційного аналізу ... 264
3.9.2. Метод регресійного аналізу ... 275
3.9.3. Задача оптимізації статистичного аналізу складних процесів ... 280
3.9.4. Алгебра планування експерименту ... 281
3.9.5. Метод Бокса-Вілсона ... 289
3.10. Приклади математичного опису складних об'єктів експериментально-статистичними методами ... 299
3.10.1. Математичний опис промислового об'єкту методом пасивного експерименту ... 299
3.10.2. Математичний опис напіввиробничого об`єкту методом активного експерименту ... 300
3.11. Кількість спостережень ... 302
3.12. Квантування за часом ... 306
3.13. Метод експертних оцінок перевірки адекватності математичних моделей ... 311
3.14. Комп'ютерні технології ідентифікації виробничих агрегатів ... 315
3.15. Методичні прийоми дослідження впливу факторів на виробничі та економічні процеси ... 320
3.16. Контрольні питання ... 325
3.17. Задачі для самостійної роботи ... 325
3.18. Завдання до лабораторної роботи ... 326

Комп'ютерний проект 1. Аналіз статистичних даних 330
1.Форми та методи подання і попереднє статистичне опрацювання числових даних часових послідовностей 332
1.1. Попереднє опрацювання даних та подання результатів 332
1.2. Формування файлу даних у формі Excel та Word таблиць 332
1.3. Графічне подання даних 334
1.4 .Описова статистика – кількісні характеристики даних 335
1.5. Побудова гістограми 337
1.6. Побудова кутуляти 338
2. Виявлення тенденції часового ряду методами згладжування 339
2.1. Методи згладжування часових рядів 340
2.2. Метод рухомого середнього 341
2.3. Метод зваженого рухомого середнього 342
2.4 .Властивості рухомого середнього 343
2.5. Медіанна фільтрація 347
2.6. Нормування часових послідовностей 349
2.7. Критерії ефективності згладжування часових рядів 349
2.8. Формули для зваженого рухомого середнього 349
3. Кореляційний аналіз часових послідовностей 352
3.1. Кореляційне поле 353
3.2. Коефіцієнт кореляції 356
3.3. Кореляційне відношення 359
3.4. Властивості кореляційного відношення 360
3.5. Кореляційна матриця 360
3.6. Автокореляція 361
3.7. Автокореляція в часових рядах 361
3.8. Розрахунок автокореляції 362
4 .Ієрархічний агломеративний кластерний аналіз багатовимірних даних 365
5 .Порядок роботи 378
6. Хід роботи 379
7. Форма звітності 381

Комп'ютерний проект 2. Розрахункова робота реалізації чисельних методів 383
1. Алгоритми розв'язування нелінійних рівнянь 286
1.1. Основні поняття та визначення 386
1.2. Метод половинного ділення 387
1.2.1 .Постановка задачі 387
1.2.2. Алгоритм методу 388
1.3. Метод хорд 389
1.3.1. Постановка задачі 389
1.3.2. Алгоритм методу 391
1.4. Метод Ньютона (метод дотичних) 391
1.4.1. Постановка задачі 392
1.4.2. Правила визначення рухомого кінця для методу Ньютона 395
1.5. Метод ітерацій (метод послідовних наближень) 396
2. Система лінійних алгебраїчних рівнянь 400
2.1. Основні поняття та визначення 400
2.1.1. Поняття матриці 400
2.1.2. Приклади матриць 401
2.1.3. Рівність матриць 402

2.1.4. Додавання матриць ... 402
2.1.5. Множення матриці на число ... 403
2.1.6. Віднімання матриць ... 403
2.1.7. Множення матриць ... 403
2.1.8. Транспонування матриць ... 404
2.1.9. Обернена матриця ... 405
2.1.10. Властивості оберненої матриці ... 405
2.2. Класифікація методів розв'язання СЛАР на ПК ... 405
2.3. Особливості методів Гауса ... 407
2.4. Метод Гауса з послідовним виключенням невідомих ... 408
2.5. Метод Крамера ... 411
2.6. Метод оберненої матриці ... 413
2.7. Матричний метод ... 414
3. Інтерполювання функцій ... 422
3.1. Математична постановка задачі інтерполювання ... 424
3.2. Інтерполяційний многочлен Лагранжа ... 426
3.3. Перша інтерполяційна формула Ньютона для рівновіддалених вузлів інтерполяції ... 429
3.4. Метод Гауса ... 432
3.5. Метод Стірлінга ... 432
3.6. Метод Бесселя ... 432
4. Розв'язування нелінійних та трансцендентних рівнянь ... 433
4.1. Метод половинного ділення ... 434
4.2. Метод пропорційних частин (хорд) ... 435
4.3. Метод Рибакова ... 437
4.4. Метод простих ітерацій ... 438
5. Точність чисельного інтегрування ... 439
6. Методи розв'язку крайових задач ... 440
6.1. Метод прогону ... 440
6.2. Проекційні методи (на прикладі методу Гальоркіна) ... 442
7. Методи пошуку максимуму функції двох змінних ... 447
7.1. Метод попокоординатного спуску ... 447
7.2. Метод спірального попокоординатного спуску ... 448
8. Розв'язування чисельними методами диференціальних рівнянь ... 449
8.1. Основні поняття та визначення ... 449
8.2. Класифікація численних методів розв'язування задачі Коші ... 452
8.3. Одноточкові методи розв'язування задачі Коші на ПК ... 454
8.3.1. Метод Ейлера ... 454
8.3.2.Метод Рунге-Кутта ... 456
9. Градієнтні методи пошуку екстремуму ... 459
9.1. Лінії рівня ... 459
9.2. Найпростіший градієнтний метод оптимізації функцій ... 460
9.2.1. Властивості градієнтів ... 460
9.2.2. Алгоритм градієнтного методу ... 462
9.2.3. Вплив величини кроку на градієнтний пошук ... 462
9.2.4. Критерій закінчення пошуку ... 463
9.2.5 .Недоліки градієнтного методу ... 464
9.2.6. Обчислення похідних ... 464
9.2.7. Модифікації алгоритмів градієнтного методу ... 464

9.3. Метод найшвидшого спуску ... 465
9.4. Метод Ньютона ... 465
9.5. Модифікований метод Ньютона ... 466
9.6. Методи змінної метрики ... 466
9.7. Недоліки та переваги методу ДФП ... 467
10. Диференціальні рівняння в частинних похідних ... 468
10.1. Різницеві методи розв'язування диференціальних рівнянь у частинних похідних ... 469
10.2 .Етапи чисельного розв'язування диференціальних рівнянь (ДР) у частинних похідних на ПК ... 470
10.3. Гіперболічні рівняння ... 475
10.4. Параболічні рівняння ... 475
10.5. Приклади розв'язування диференціальних рівнянь в частинних похідних в середовищі MathCAD ... 476
11. Контрольні питання ... 478
12. Задачі для самостійної роботи ... 480
13. Завдання до розрахункової роботи ... 481

Список літератури ... 513

Передмова наукового редактора серії підручників та навчальних посібників «КОМП'ЮТИНҐ»

Шановний Читачу!

Розпочинаючи роботу над вступною статтею до другої черги крупного ініціативного освітньо-наукового проекту щодо видання серії україномовних посібників, підручників та монографій, які виходять у світ під узагальненим гаслом «COMPUTING», число книг якої вже перекрочило чотири десятки найменувань хочу щиро подякувати Тобі Вельмишановний Читачу, авторам, видавцю, фінансовим меценатам та спонсорам, котрі в результаті творчої, синергетичної взаємодії забезпечили успішність цього унікального вітчизняного інноваційного проекту.

Пройшло майже десятиліття як у 2008-му році до друку було подано рукопис видання першого підручника серії «КОМП'ЮТИНГ» - «Сховища даних» за авторством професорів Н.Б. Шаховської та В.В. Пасічника. Нині, наприкінці 2017 року завершився перший етап проекту, світ побачили понад чотири десятки найменувань навчальних посібників, підручників та наукових монографій, що загалом стало не пересічною подією в українському освітньому, науковому та професійно-бізнесовому IT середовищі.

Нині можна впевнено стверджувати, що започаткований не чисельною групою ініціативних високо фахових викладачів та науковців кафедри інформаційних систем та мереж Львівської політехніки проект серії україномовних посібників, підручників та наукових монографій, об'єднаних проблематикою комп'ютингу – успішно реалізований.

Автори та видавці «видали на гора» перших декілька десятків оригінальних україномовних видань, а це Вельмишановний Читачу в наш час, зовсім не мало.

Започаткування наступного (другого) етапу реалізації видавничого освітньо-наукового проекту співпадає з імплементацією сутнісних змін в системі середньої та вищої освіти України загалом і її інформаційно-технологічного сектору зокрема.

В українській середній школі реалізовується амбітний проект «Нова українська школа». Ми стаємо учасниками процесів запровадження викладання у школі предмету «Інформатика», починаючи з другого класу і вихід на профільну інформатично-технологічну освіту в старшій школі (10-12 –і класи).

При цьому ми уважно відслідковуємо базові тенденції розвитку шкільної інформативної освіти у світовому вимірі. У цьому контексті не можемо обійти увагою роботу групи освітян в рамках зазначеного проекту з розроблення системи сучасних вітчизняних шкільних освітніх стандартів з інформатики. Проект реалізовувався у 2016-2017 роках на замовлення МОН України. Не буде зайвим навести ряд відомостей щодо зарубіжного досвіду запровадження шкільних освітніх стандартів з інформатики, які подані у підготовленому у вересні 2017 р. Оксаною Володимирівною Пасічник підсумковому аналітичному матеріалі.

«Шкільний предмет інформатика у різних країнах має як різне змістове наповнення, так і різне місце у навчальних планах та стандартах. Це пов'язано із динамічною історією викладання інформатики, а також зі змінами, які спостерігаються останніми роками. На початку впровадження цього навчального предмету в освітніх системах різних країн наприкінці 1980-х років, основна увага приділялась алгоритмічному мисленню та розвитку вмінь складати програми. З середини 1990х до початку 2000х років відбувалось зміщення у напрямку вивчення прикладних програм та опанування курсу інформаційних та комунікаційних технологій. У зв'язку з тим, що зазначені технології набували все більшого поширення та ставали інтуїтивно зрозумілішими, вивчення окремого предмету в школі

втрачало актуальність, тож на початок 2010-х, в освітніх системах багатьох країн цей предмет був відсутнім, інтегрованим, або викладався факультативно.

Водночас, у переліку ключових компетентностей переважної більшості національних стандартів шкільної освіти присутня цифрова грамотність чи ІКТ-компетентність, як наскрізне вміння користуватися цифровими технологіями. Все очевиднішою стає необхідність вивчення не стільки конкретних програмних засобів чи технологій, як принципів їх функціонування та вміння створювати власне цифрове середовище, в тому числі програмуючи його під свої потреби.

За результатами дослідження, проведеного Вутт та Брумельгуса, інтеграція інформаційної грамотності в інші предмети в Нідерландах призвела до зникнення предмету на початку 2000х років, а у 2013 році знову почало розглядатись питання щодо повернення предмету в шкільні програми, надаючи йому оновленого змісту та ролі.

Слід зауважити, що у контексті шкільного предмету, інформатика у різних системах освіти також зустрічається під назвами ІКТ (інформаційні та комунікаційні технології у Чехії, Таїланді, Ісландії), комп'ютерні науки (США, Канада, програма міжнародного бакалаврату International Baccalaureate) та комп'ютинг (Великобританія, Сингапур). У деяких країнах виділяється окремий предмет під назвою програмування або кодування (Бельгія, Данія, Естонія, Іспанія, Фінляндія, Угорщина, Нідерланди, Норвегія, Польща, Португалія, Словаччина). У ряді країн також використовуються такі терміни, як алгоритмічні застосунки (Греція), алгоритмічне розв'язування задач (Словаччина), розроблення алгоритмів та моделей даних (Угорщина), алгоритми та робототехніка (Tecnología, Programación y Robótica, Іспанія), інформатика та цифрові науки (Informatique et sciences du numérique, Франція). В Австралії та Новій Зеландії предмет запроваджується із назвою цифрові технології.

До прикладу, в Хорватії предметна галузь об'єднана із технологіями під загальною назвою технології та інформатика, в межах якої виділяється 4 складові: технічна творчість, світ техніки, інформаційні та комунікаційні технології та комп'ютерне розв'язування задач, а у Словаччині та Болгарії цей предмет об'єднано в одну групу із математикою. У Швейцарії інтегрований модуль має назву медіа та інформатика.

У національному звіті Королівської академії інженерії "Шлях інформатики у школах Великобританії: вимкнути чи перезавантажити?" окремою рекомендацією є уникнення використання терміну ІКТ (ICT), і розділення предметної області на такі аспекти, як цифрова грамотність (Digital Literacy), інформаційні технології (Information Technology) та інформатика (Computer Science). Наводиться цікава аналогія вивчення інформатики з мовною освітою, у якій можна виділити читання та письмо (базова грамотність), вивчення мови (принципи функціонування) та літератури (приклади застосування).

Також хотілося би згадати термін, для якого, на жаль, наразі не вдається знайти україномовний відповідник – це computational thinking: розв'язання задач, проектування систем та розуміння людської поведінки, виходячи з базових понять інформатики. Вперше цей термін використав Сеймур Пеперт (математик, інформатик, педагог) ще у 1996 році, але актуалізація його дослідження та використання відбулася після виходу у світ статті Джанет Вінг у 2006.

Описове визначення та пояснення цього терміну виглядає так: процеси мислення, пов'язані з формулюванням проблеми та представленням її вирішення у такий спосіб, щоб його можна було реалізувати за допомогою комп'ютера.

Це поняття інкапсулює алгоритмічне мислення, декомпозицію, виявлення шаблонів та абстракцію. Декомпозицією є вміння розкладати задачу на складові елементи, оперуючи якими можна чітко пояснити процес виконавцеві, або записати його для власного подальшого використання. Виявлення шаблонів передбачає вміння знаходити тенденції,

схожі елементи або відмінності, котрі дозволять робити прогнози. Узагальнення шаблонів та абстрагування означає вміння відкидати відомості, які не є значущими для розв'язання певної задачі, а також підсумовувати необхідну інформацію. Розроблення алгоритму є вмінням створювати покрокову стратегію для розв'язання задачі.

Часто до цих основних елементів додають такі характеристики, як творчість, експериментування, відлагодження, наполегливість та співпраця. Таке складене синтетичне означення обчислювального мислення виходить на ключову позицію при визначенні цілей вивчення предмету інформатики в освітніх системах різних країн.

У сучасному інформаційному суспільстві цифрова компетентність не обмежується базовими навичками роботи з типовими програмами, а проникнення інформаційних, комунікаційних та цифрових технологій у всі сфери діяльності людини означає, що розуміння принципів функціонування комп'ютерних систем та програм є основоположною складовою грамотності у 21 столітті. Це, а також прогнозований дефіцит ІТ-фахівців, який оцінюється у 756 тисяч осіб до 2020 року, спричинили до того, що зміст, роль та місце предмету інформатика, нині активно трансформується, і це відображається у відповідних стандартах освіти.

Зокрема, у 2010 було розпочато європейську кампанію з реалізації маніфесту е-навичок, одним із кроків якої є курикулярна реформа, що полягає в інтегруванні програмування у шкільні програми в 15 країнах, оновленні програм з інформатики, розробці стратегій оцифрування процесів функціонування шкіл, впровадженні у навчальні програми питань ІТ-безпеки та авторського права, тіснішій співпраці з професійною спільнотою.

Навесні 2009 року університет Альберти (Канада) додав інформатику як один з вибіркових вступних предметів, цей крок підтримали ряд університетів: Калгарі, Летбриджа, МакЮен. Це спричинило підвищення інтересу до предмету інформатики, як вибіркового на шкільному рівні.

З 2013 року відзначається європейський тиждень кодування, набула значного розмаху глобальна акція Години коду на основі ініціативи Code.Org, яка має на меті популяризацію інформатики у середніх школах.

Відповідно до даних звіту «Обчислювальне майбутнє. Комп'ютерне програмування та кодування – пріоритети, навчальні програми та Європейські ініціативи», що оновлений наприкінці 2015 року на основі опитувань 21 міністерства освіти європейських країн, у 16 країнах інформатика та кодування включені у навчальні стандарти на національному, регіональному чи локальному рівні. Спостерігається тенденція до виділення окремого предмету (якщо його не було раніше) та його вивчення на раніших освітніх етапах (в тому числі у початковій школі). При цьому 12 країн виділяють окремий предмет кодування/комп'ютингу на регіональному чи шкільному рівні, 13 країн інтегрують кодування у загальний курс інформаційних та комунікаційних технологій.

Сталий досвід присутності окремого інформативного предмету в навчальних планах з кінця 1980-х років мають такі країни, як Україна, Ізраїль, Польща, Литва, Словаччина, Кіпр. У школах багатьох країн протягом тривалого часу інформатика або не викладалась взагалі, або мала вигляд вибіркового предмету, спрямованого на оволодіння базовими інформаційними та комунікаційними технологіями. У деяких країнах паралельно існувала можливість обрати поглиблений курс програмування.

Підсумовуючи аналітичне опрацювання широкого спектру матеріалів національних стандартів шкільної освіти щодо стану та перспектив вивчення предмету інформатики можна сформулювати наступні висновки.

- Залежно від особливостей національних систем шкільної освіти, предмет може бути окремим чи інтегрованим, обов'язковим або вибірковим, імовірно з різною реалізацією цієї варіативності на різних рівнях освіти.

- У більшості країн, які оновлювали стандарти освіти останніми роками, спостерігаємо зміщення від суто прикладних вмінь використовувати програмні засоби до глибшого розуміння принципів роботи та розробки програм та систем.
- Робиться важливий наголос на відповідальне та безпечне використання інформаційних технологій, а також розуміння потенційних переваг та ризиків їх поширення та використання з різними цілями.

Випускник школи XXI століття повинен бути компетентним учасником інформаційного суспільства, усвідомлюючи його особливості, принципи функціонування, переваги та ризики швидкозмінних інформаційних технологій, а також вміти творчо розв'язувати проблемні ситуації за допомогою цифрових пристроїв та програм».

Наша увага як ініціаторів видавничого освітньо-наукового проекту до шкільної складової у системі цифрової грамотності членів сучасного інформаційного суспільства є цілком закономірною. Адже, шкільна освіта закладає фундамент як для здобуття вищої освіти, так і загалом подальшого суспільного буття.

В умовах становлення та стрімкого розвитку суспільства знань і цифрових навичок невідворотною є якісна зміна підходів до визначення змісту вищої освіти, методів і форм оцінювання якості підготовки фахівців, набуття кваліфікації, яка відповідала б вимогам стейкхолдерів. Основне протиріччя сучасної системи освіти слід трактувати як протиріччя між високими темпами приросту обсягів знань і обмеженими можливостями їх сприйняття та засвоєння особою, членом цього ж суспільства знань. Це протиріччя невідворотно примушує як теоретиків, так і педагогів-практиків відмовитися від ідеального образу усебічно розвиненої особистості і перейти до нового ідеального образу особи з максимально розвинутими здібностями, яка вдосконалюється шляхом саморегуляції та самоосвіти впродовж всього життя. Інтеграція вищої освіти України у світову освітню систему є об'єктивним процесом, який сутнісно трансформує національну систему вищої освіти, ключовими складовими якої стають освітні програми та комплекси відповідних ступеневих компетентностей. Компетентнісний підхід до побудови освітніх програм у вищій школі, що реалізується в Україні відповідно до стратегії реформування національного освітнього простору складає основу студентоцентрованого навчання, базовим і системо твірним елементом якого є студент, слухач, учень, що прагнуть набути нових знань, умінь та навичок.

Інформатична освіта університетського рівня в цьому контексті не є виключенням. Саме таку парадигму ми прагнемо реалізувати в рамках другого етапу проекту видання книг освітньо-наукової серії КОМП'ЮТИНГ.

У вищій школі України з 2016 року імплементується нова номенклатура галузей знань та спеціальностей. В контексті наших професійних інтересів мова йде про галузь знань 12- «Інформаційні Технології» з шістьма спеціальностями: 121- програмна інженерія; 122 – комп'ютерні науки; 123 – комп'ютерна інженерія; 124 – системний аналіз; 125 – кібербезпека; 126 – інформаційні системи та технології.

Часовий відтинок, що пройшов з моменту виходу у світ першої книги освітньо-наукової серії КОМП'ЮТИНГ характеризується бурхливими процесами становлення, формування та стрімкого зростання інформаційного суспільства у всіх без винятку напрямах, профілях та аспектах. Не є виключенням в цьому потоці і освітня галузь, у якій відбувається активний пошук нових парадигм, концепцій, підходів. Якщо наприкінці 2010 року в україномовній «освітній» літературі та «наукових» публікаціях набув популярності феномен ІКТ (англ. ICT), який не зовсім коректно був привнесений в україномовне термінологічне середовище як інформаційно-комунікаційні технології замість оригінального подання, яке слід перекладати як «інформаційні та комунікаційні технології» (див. першоджерело, в якому вперше використано концепт), то для кінця 2017 року характерною

ознакою та особливістю європейського освітнього простору є нове «народження» та запровадження до освітніх стандартів різних освітніх та освітньо-кваліфікаційних рівнів галузі комп'ютингу термінологічного концепту TIC (англ.IST), який трактується як технології інформаційного суспільства (information societies technology).

Навіть перший побіжний погляд на здавалося б незначні «косметичні» уточнення трактування понять та концептів наштовхує на думку щодо їх глибинної тектонічної природи та сутнісних методологічних змін, які за цим невідворотно послідують. Зверну увагу Шановного Читача на принципову інноваційну ознаку концепту TIC (англ. IST), в якому інформаційні технології переводяться до розряду загальнолюдських та загально цивілізаційних. Пророче бачення видатного українського вченого В.І. Вернадського щодо високотехнологічного підґрунтя ноосферного шару сучасного світу знаходять своє *методологічне, наукове та виробничо-технологічне втілення.*

Основним завданням, яке ставиться перед виданнями освітньо-наукової серії КОМП'ЮТИНГ на другому (нинішньому) етапі її розгортання, є якомога повніше, змістовніше та яскравіше подання нашої професійної галузі КОМП'ЮТИНГу в трактуванні ТЕХНОЛОГІЙ ІНФОРМАЦІЙНОГО СУСПІЛЬСТВА, розуміючи, що такими технологіями у сучасному світі є саме комп'ютерні інформаційні технології. Цей процес цілком відповідає нинішньому тренду, який фіксується появою таких інформаційно-технологічних інновацій, якими є хмарні обчислення (Cloud Computing), інтернет речей (IoT), розумні міста (SC), розумні регіони (SR), системи штучного інтелекту, технології кібербезпеки та великих за обсягом даних (Big Data), а також нового освітньо-наукового напрямку «Науки про дані» (Data Science), який покликаний опікатися дослідженням процесів аналізу даних та знань.

Нове «дихання», яке отримує ініційований групою небайдужих IT фахівців, - освітян, бізнесменів, науковців видавничий проект серії КОМП'ЮТИНГ має надійне творче підґрунтя та добру верифіковану часом традицію. Бажаю Тобі Шановний Читачу ще тривалий час мати потребу та можливість успішно розвиватись в галузі КОМП'ЮТИНГу, активно опрацьовуючи все нові і нові видання цієї без перебільшення унікальної освітньо-наукової серії.

З глибокою повагою науковий редактор серії підручників «КОМП'ЮТИНГ», д.т.н., професор Володимир Пасічник.

Вступне слово авторів

Суспільство розпочало важливий період свого розвитку – еру інформатизації. Для вирішення теоретичних і практичних задач, що виникають під час діяльності людини у різних галузях науки, техніки та виробництва з метою звільнення людини від надмірного інтелектуального навантаження, значний ефект отримуємо за рахунок використання обчислювальної техніки при умові достатнього програмного забезпечення та ефективного його використання. Тому дисципліни «Методи обчислень» та «Чисельні методи» набувають особливо великого значення у підготовці фахівців високої кваліфікації.

Сучасний розвиток науки та обчислювальної техніки характеризується все більш зростаючим рівнем використання комп'ютерних моделей як для дослідження поведінки явищ та процесів, що оточують людину, так і для розв'язування практичних задач, що пов'язані з управлінням та прогнозуванням [1-93, 103-109, 112-114, 125-136, 293-302]. Вивчення навчальної дисципліни «Чисельні методи в комп'ютерних науках» дозволяє студентам оволодіти знаннями в галузі практичних методів вирішення математичних проблем, що виникають у процесі інженерної діяльності та моделювання фізичних систем: засвоїти способи обчислення на сучасних комп'ютерах із застосуванням пакетів спеціальних прикладних програм. Об'єктом вивчення навчальної дисципліни є типові математичні задачі, до яких зводиться розв'язування практичних проблем, що виникають в процесі розроблення інформаційних систем та систем моделювання. Предметом вивчення навчальної дисципліни є чисельні методи розв'язування типових математичних задач. Мета навчального посібника – ознайомлення студентів із постановками основних математичних задач і чисельних методів їх розв'язування, набуття студентами навичок реалізації на комп'ютері чисельних методів, навичок роботи з відомими комп'ютерними математичними пакетами.

Чисельні методи – методи наближеного або точного розв'язування задач прикладної математики, які грунтуються на побудові послідовності дій над скінченною множиною чисел [94-102, 110-111, 137-292, 303-320]. Згідно основних вимог чисельні методи мають бути стійкими та збіжними [234]. Основне питання теорії чисельних методів: отримання чисельних методів, які задовольняють вимоги високої точності, стійкості та економічності. Знаходження чисельних методів, що задовольняють ці вимоги, є складною задачею оптимізації чисельних методів. Статистичне опрацювання експериментальних даних зазвичай грунтується на граничних теоремах теорії ймовірностей та обчисленні порівняльних оцінок. Однак, для підвищення якості оцінок необхідна велика кількість даних, обсяг обчислень може виявитися дуже великим. Чисельні методи націлені на скорочення кількості обчислень при збереженні якості результатів. До найбільш ефективних чисельних методів в цій галузі відносяться методи, які застосовують швидке перетворення Фур'є. Для розв'язування задач апроксимації та обчислення функцій різних класів застосовують чисельні методи інтерполювання, найменших квадратів, ортогоналізації, врівноваження значень, умовної мінімізації тощо. Найбільш актуальними є методи кусково-многочленної та раціональної сплайнової апроксимації, а також адаптивної апроксимації та нелінійної за параметром апроксимації [91-93].

Інженеру часто доводиться вирішувати алгебраїчні та трансцендентні рівняння і системи рівнянь, що можуть бути самостійною задачею (наприклад, аналіз рівноваги сил в жорсткій системі балок або дослідження умов та параметрів рівноваги хімічної реакції, тощо) або частиною більш складних задач. В обох випадках практична цінність чисельного методу в значній мірі визначається швидкістю і ефективністю отримання розв'язку. Розглянемо найбільш відомі чисельні методи та ефективні алгоритми розв'язування систем лінійних алгебраїчних рівнянь.

Чисельне інтегрування та диференціювання починається із визначення відповідних операцій. Однак, з урахуванням необхідності економії обсягу обчислень та з урахуванням

некоректності задачі диференціювання з'являється велика кількість чисельних методів для різних класів функцій та різного роду вихідних даних.

Інженери часто зіштовхуються з диференційними рівняннями та системами диференційних рівнянь при розробці нових виробів чи технологічних процесів, так як більша частина законів фізики формалізується саме у вигляді диференційних рівнянь. Будь-яка задача проектування, яка пов'язана з розрахунком потоків енергії чи руху тіл, в кінцевому результаті зводиться до розв'язку диференційних рівнянь. Нажаль, лише дуже малу частину з них можливо вирішити без використання обчислювальних машин. Тому чисельні методи розв'язку диференційних рівнянь відіграють важливу роль у практиці інженерних розрахунків.

Основою чисельних методів розв'язування багатьох класів рівнянь є дискретизація задачі з наступним приведенням отриманих нелінійних рівнянь до послідовності систем алгебраїчних рівнянь. У зв'язку з цим чисельні методи можна поділити за способом дискретизації на проекційні, скінченно-різницеві та проекційно-різницеві, а за способом розв'язування лінійної системи – на прямі методи, ітераційні методи та комбіновані.

Розв'язування різних класів рівнянь та багатьох інших задач зводиться до задач мінімізації функцій та функціоналів за наявності або відсутності обмежень. Чисельні методи розв'язування задач мінімізації випливають із методів швидкого спуску по поверхні (мінімізація функції мети). Наприклад, методи швидкого спуску, градієнтного, загального градієнтного та найшвидшого спуску, методу можливих та спряжених напрямів тощо. Чисельні методи використовують в обчислювальній математиці для розв'язування відповідного типу задач.

Обчислювальна математика – розділ математики, що включає коло питань, які пов'язані із виконанням обчислень та використанням комп'ютерів. Точніше, обчислювальна математика – теорія чисельних методів розв'язування типових математичних задач. Класи задач чисельних методів та обчислювальної математики поділяють на:

1) розв'язування лінійних рівнянь;
2) знаходження власних значень та векторів матриці;
3) знаходження сингулярних значень і векторів матриці;
4) чисельне розв'язування нелінійних алгебраїчних рівнянь та їх систем;
5) чисельне розв'язування систем нелінійних алгебраїчних рівнянь;
6) чисельне розв'язування диференціальних рівнянь та систем (як звичайних диференціальних рівнянь, так і рівнянь з частинними похідними);
7) чисельне розв'язування систем диференціальних рівнянь;
8) чисельне розв'язування інтегральних рівнянь;
9) задачі апроксимації функцій;
10) задачі інтерполяції функцій;
11) чисельне інтегрування та обчислення похідної;
12) задачі екстраполяції;
13) задачі оптимізації;
14) обернені задачі.

Чисельні методи використовувалися ще за часів Ньютона (1642-1727рр.) для розв'язування задач з астрономії, геодезії та обчислення механічних конструкцій. На той час обчислення з використанням чисельних методів виконувалися з доволі високою точністю (до восьми знаків після коми). Наприклад, французький математик і астроном Урбен Левер'є (1811-1878 рр.), уточнюючи траєкторію руху планети Уран, виявив відхилення від обчисленої траєкторії. Він припустив, що ці відхилення спричиняє інша планета, яка до того не спостерігалась астрономами. Використовуючи чисельні методи, він за півроку обчислив масу і орбіту невідомої планети, що справляє дію на Уран і виводить планету із рівноваги. Один примірник своїх розрахунків Левер'є відразу ж послав Йогану Галле з Берлінської обсерваторії, який отримавши лист 23 вересня 1846 року, негайно почав спостереження і в ту ж ніч, дуже близько від місця вказаного Левер'є, знайшов невідому планету, яку пізніше назвали Нептуном.

Статистичне опрацювання експериментальних даних зазвичай грунтується на граничних теоремах теорії ймовірностей та вимагає обчислення оцінок в порівнянні з простими формулами. Однак, для підвищення якості оцінок необхідна велика кількість даних, і обсяг обчислень може виявитися дуже великим. Тому чисельні методи тут націлені на скорочення обсягу обчислень при збереженні якості результатів. Найефективнішими чисельними методами в цій галузі є методи, які застосовують швидке перетворення Фур'є.

Для розв'язування задач апроксимації та обчислення функцій різних класів застосовують чисельні методи інтерполювання, найменших квадратів, ортогоналізації, врівноваження значень, умовної мінімізації тощо. Найактуальнішими є методи кусково-многочленної та раціональної сплайн-апроксимації, а також адаптивної апроксимації та нелінійної за параметром апроксимації.

Чисельне інтегрування та диференціювання здійснюється на основі означення відповідних операцій. Однак, через необхідність економії обсягу обчислень та некоректність задачі диференціювання розроблено велику кількість чисельних методів для різних класів функцій та різного роду вихідних даних.

Основою чисельних методів розв'язування багатьох класів рівнянь є дискретизація задачі з наступним зведенням отриманих, взагалі кажучи, нелінійних рівнянь до послідовності систем алгебраїчних рівнянь. У зв'язку з цим чисельні методи можна поділити за способом дискретизації на проекційні, скінченно-різницеві та проекційно-різницеві, а за способом розв'язування лінійної системи — на прямі, ітераційні та комбіновані методи.

- У прямих наближених методах використовують заміну вихідних даних простішою функцією (наприклад, з використанням методів інтерполяції та апроксимації) або заміну способу обчислень (наприклад, замінюючи інтеграл сумою простих числових доданків, похідну – різницею) для спрощення обчислення.
- В ітераційних методах можна знайти розв'язок задачі, використовуючи послідовність формул, де кожен новий уточнений наближений розв'язок змінної обчислюється через попереднє її значення. Ітераційний процес пошуку наближеного розв'язку завершується тоді, коли виконується умова: різниця за модулем двох останніх знайдених значень менше-рівна значенню попередньо заданої сталої.

Розв'язування різних класів рівнянь та багатьох інших задач зводиться до задач мінімізації функцій та функціоналів за наявності або відсутності обмежень. Чисельні методи розв'язування задач мінімізації випливають із різних ідей швидкого спуску поверхнею, яка відповідає мінімізованій функції. До них належать методи швидкого спуску, градієнтного, загального градієнтного та найшвидшого спуску, методи можливих та спряжених напрямів тощо. Для оцінки чисельних методів, вводять такі основні характеристики: складність; порядок методу; збіжність; швидкість збіжності; стійкість до похибок обчислень; стійкість до похибок вихідних даних.

Складність методу оцінюється кількістю та якістю обчислень, які необхідні для досягнення достатньо близького наближення розв'язку задачі.

Порядок методу – вимоги до знань про функції, що входять до математичного формулювання задачі (наприклад, використання в методі похідних цих функцій):

- метод нульового порядку – використовує тільки значення цих функцій;
- метод першого порядку – використовує значення функцій та їх перших похідних;
- метод другого порядку – використовує значення функцій, їх перших та других похідних тощо.

Чисельні методи називають *збіжними*, якщо результати прямують до точного розв'язку задачі при прямуванні параметрів чисельних методів до певних граничних значень. Існує лінійна збіжність, нелінійна збіжність та квадратична збіжність.

Стійкість до похибок обчислень характеризує числовий метод, застосування якого приводить до розв'язку задачі, незважаючи на помилки заокруглень і обчислень. Для цього в

№ 38932

чисельних методах, якщо потрібно, передбачаються додаткові операції, що не змінюють суть методу, але забезпечують його стійкість до помилок обчислень.

Стійкість до похибок вихідних даних – характеристика чисельного методу, що при невеликих похибках вихідних даних забезпечує отримання наближеного розв'язку задачі з незначною похибкою. Стійкість до похибок вихідних даних досягається, як правило, шляхом модифікації чисельного методу, тобто внесенням змін до суті методу.

Чисельні методи називаються *стійкими*, якщо результати неперервно залежать від вихідних даних задачі або якщо похибка заокруглення, що пов'язана з реалізацією чисельних методів на комп'ютері, залишається обмеженою при заданих межах зміни параметрів чисельних методів.

Чисельні методи називаються *збіжними*, якщо результати прямують до точного розв'язку задачі при прямуванні параметрів чисельних методів до певних граничних значень.

Основне питання теорії чисельних методів: отримання чисельних методів, які задовольняють вимоги високої точності, стійкості та економічності. Розробка чисельних методів, що задовольняють ці вимоги, є складною задачею оптимізації чисельних методів.

Основна відмінність обчислювальної математики полягає в тому, що при розв'язуванні обчислювальних задач людина оперує машинними числами, що є дискретною проекцією дійсних чисел на конкретну архітектуру комп'ютера. Тому важливу роль в обчислювальній математиці відіграють оцінки точності алгоритмів та їх стійкість до представлення чисел у пам'яті комп'ютера. Наприклад, для розв'язування лінійної системи алгебричних рівнянь рідко використовують обчислення оберненої матриці, так як цей метод може привести до помилкового розв'язування у випадку зі сингулярною матрицею. А розповсюджений у лінійній алгебрі метод, який заснований на обчисленні визначника матриці та її доповнення, вимагає набагато більше арифметичних операцій, ніж будь-який стійкий метод розв'язування лінійної системи рівнянь. Чисельні методи називають стійкими, якщо результати неперервно залежать від вхідних даних задачі або якщо похибка заокруглення, що пов'язана з реалізацією чисельних методів на комп'ютері, залишається обмеженою при заданих межах зміни параметрів чисельних методів. Чисельні методи поділяються на наступні класи задач.

1. Чисельні методи алгебри.
2. Чисельні методи аналізу.
3. Чисельні методи розв'язування звичайних диференціальних рівнянь.
4. Чисельні методи розв'язування диференціальних рівнянь з частинними похідними.
5. Чисельні методи математичної статистики.
6. Математичне програмування: методи оптимізації, дослідження операцій та теорія ігор.

Спочатку для розв'язування математичних задач існували та набули значного розвитку аналітичні методи, які реалізувати на комп'ютері не було можливості. А всі результати, які можна було отримати на комп'ютері, були числовими. З появою систем символьної математики (Eureka, Reduce, SciCalculator, MathCad, MatLab, Mathematica, Maple, Derive) стала можлива комп'ютерна реалізація символьних методів. Всі числові методи поділяються на 2 класи – аналітичні та чисельні. Як одні, так і другі поділяють на точні та наближені. Основну частину всіх обчислювальних методів складають наближені чисельні методи.

Досвід розв'язування науково-дослідних і прикладних задач показує, що, незалежно від складності, кінцевої мети досягають постановкою експерименту або методом математичного моделювання. Кожен з цих методів має переваги й недоліки. За допомогою експерименту розв'язують навіть дуже складні задачі, при цьому достовірність результатів тим вища, чим ретельніше відпрацьована методика експерименту. Водночас здобуті результати відносяться лише до умов проведення експерименту, внаслідок чого узагальнення результатів на інші умови є некоректним. Також враховують економічний важіль постановки складного експерименту. В цьому випадку кращі можливості має метод математичного моделювання за допомогою комп'ютера, коли аналізують не реальну задачу, а її модельне представлення. Процес математичного моделювання подають у такій послідовності:

- фізична постановка задачі;

- математична постановка задачі;
- математичне дослідження задачі;
- аналіз та осмислення математичного розв'язку;
- порівняння розв'язку з експериментом.

Розглянемо докладніше математичну постановку та математичне дослідження задачі. Математична постановка полягає у формуванні математичної моделі досліджуваної задачі, яка звичайно є системою рівнянь математичної фізики (диференціальних, інтегральних, інтегрально-диференціальних). Математичне дослідження задачі зводиться до розв'язування системи рівнянь і аналізу здобутих результатів. Для порівняно простих задач вдається розв'язати систему рівнянь і розв'язок подати у вигляді залежностей, які виражені через елементарні та інші відомі функції. Якщо це можливо, то говорять, що знайдено аналітичний (точний) розв'язок задачі. Однак, переважна більшість практично важливих задач аналітичних розв'язків не має. До таких належать, наприклад, задачі будівництва: визначення напружено-деформованого стану пластин, плит, фундаментів; задачі стійкості, теплопровідності для твердих тіл; напрямленої дифузії тощо. У цих випадках використовують чисельні методи, які, оперуючи системою алгебраїчних рівнянь (аналогів рівнянь математичної фізики), дають можливість побудувати деяку послідовність арифметичних операцій, збільшення кількості яких до нескінченності дає точний розв'язок. Оскільки на практиці здійснюють скінченне число кроків (операцій), то знайдений розв'язок є наближеним. А тому що обчислювальні операції виконують над числами, то відповідні методи дістали назву чисельних. Найбільшого розвитку чисельні методи набули останнім часом завдяки застосуванню комп'ютерів, що мають високу швидкість обчислень і великий об'єм оперативної пам'яті. Проте основна роль при цьому відводиться, звичайно, людині, яка повинна вміти сформулювати і поставити задачу, описати її математичними залежностями (створити математичну модель об'єкта), скласти алгоритм розв'язування задачі на комп'ютері, написати програму на алгоритмічній мові, розв'язати задачу й оцінити результати. При оцінюванні результатів розрахунку поєднання чисельних методів та комп'ютера отримують оперативно ефективний результат, варіюючи найсуттєвіші параметри розрахункової схеми задачі з наступним чисельним аналізом впливу їх на кінцевий результат. Це чисельний експеримент, оскільки умови задачі можна змінювати багато разів. Незважаючи на відмінності в методології, до чисельного експерименту дотичний фізичний експеримент/дослідження, де необхідна оцінка достовірності здобутих результатів.

Математична модель об'єкта – це та сукупність рівнянь, за допомогою якої досліджують реальні фізичні об'єкти (процеси, явища). Математична модель не тотожна досліджуваному об'єкту, а є лише його наближеним описом, оскільки її будують з деякими спрощеннями та ідеалізацією. У моделі враховують найважливіші моменти і взаємозв'язки, найхарактерніші для досліджуваного реального об'єкта. Разом з тим, внаслідок заміни реального об'єкта відповідною математичною моделлю, стало можливим сформулювати задачу як математичну і скористатися для її розв'язування тим чи іншим математичним апаратом.

Алгоритм – це зрозумілий і точний припис (вказівка) виконавцеві здійснювати послідовність дій, що спрямовані на досягнення зазначеної мети або розв'язання поставленої задачі. Точність розв'язку – це міра близькості чисельного розв'язку до аналітичного. Збіжність розв'язку – це поступове наближення його до точного. Після вибору математичної моделі об'єкта й опису її на алгоритмічній машинній мові здійснюють чисельну реалізацію задачі на комп'ютері. При реалізації практичних задач здебільшого застосовують комп'ютери, що виконують від кількох сотень до мільйонів операцій за секунду. Найбільшого застосування в інженерних розрахунках набули комп'ютери, які мають не тільки високу швидкість обчислень, сучасне програмне забезпечення, але й розвинуту сервісну частину, яка дає можливість оперативно діагностувати похибки, графічно відображати результати обчислень, здійснювати розрахунки в режимі діалогу.

Розширення можливостей прикладної математики обумовило математизацію різних галузей науки: хімії, біології, економіки, геології, психології, медицини, техніки тощо. Дослідження

реальних фізичних процесів та явищ полягає у побудові відповідних математичних моделей і розробленні методів їх дослідження. У зв'язку з цим студентам технічних спеціальностей вищих закладів освіти читають курси, пов'язані із чисельними методами математики. У результаті вивчення курсів «Чисельні методи» та «Методи обчислень» студенти повинні знати чисельні методи та вміти їх використовувати при розв'язування наукових та науково-технічних задач із таких галузей як комп'ютерні науки та інформаційні технології. Досвід роботи на кафедрі «Інформаційні системи та мережі» Національного університету «Львівська політехніка» показує, що при вивченні цих курсів найбільші складнощі у студентів виникають при побудові та дослідженні математичних моделей в технічних науках, вивченні та реалізації чисельних методів їх розв'язування. В посібнику викладено курс лекцій у відповідності до програми із математичних методів для технічних спеціальностей. Навчальний посібник містить матеріал для вивчення основних теоретичних засад, функціональних можливостей та практичного застосування теорії чисельних методів, розроблення прикладних засобів та інформаційних систем аналізу та опрацювання інформації за допомогою чисельних методів. Теоретичний та практичний матеріал викладено у доступній формі. Викладення матеріалу супроводжується значною кількістю прикладів, що полегшує його сприйняття та засвоєння. Подається перелік питань та тестів для самоконтролю, а також завдання для самостійного виконання трьох рівнів складності та довідкова інформація для розв'язування задач. Навчальний посібник призначений для студентів, що навчаються за спеціальностями 122 «Комп'ютерні науки та інформаційні технології» та 124 «Системний аналіз» і споріднених спеціальностей, які пов'язані з вивченням чисельних методів в інформатиці та інформаційних технологій. Може бути використаний аспірантами як підгрунтя для наукових досліджень та викладачами як дидактичний матеріал, а також для самостійного вивчення та підвищення кваліфікації. Книга призначена для фахівців із проектування, розроблення та впровадження інтелектуальних систем опрацювання інформаційних ресурсів, науковців у галузі глобальних інформаційних системи, систем штучного інтелекту, Інтернет-технологій, фахівців з електронної комерції, Інтернет-маркетингу та Інтернет-реклами, менеджерів комплексних Web-проектів, а також для здобувачів 3-ого (освітньо-наукового) рівня вищої освіти в галузі знань 12 «Інформаційні технології». Весь матеріал розбито на 10 тем у 2-х томах. Кожна тема розрахована на 2-4 аудиторних години (1-2 лекції), і, таким чином, на весь курс відводиться 16 лекцій. До кожної теми подано контрольні питання та завдання для самоконтролю. Основну увагу в лекціях звернено на роз'яснення принципових моментів та аспектів теоретичного матеріалу, які часто викликають у студентів труднощі. Складні доведення в посібнику не подаються, оскільки їх розглядають під час аудиторних занять та на консультаціях із самостійного опрацювання дисципліни за рекомендованою літературою. Вивчення чисельних методів базується на розв'язуванні практичних задач, тому теоретичний матеріал супроводжується прикладами, що полегшує його розуміння. Згідно з вимогами освітньо-професійної програми підготовки бакалаврів напряму 12 «Інформаційні системи» студенти після вивчення навчальної дисципліни «Чисельні методи» або «Методи обчислень» повинні оволодіти такими компетентностями:

- знати: загальні поняття, пов'язані з чисельними методами; постановки типових математичних задач; чисельні методи лінійної та нелінійної алгебри; чисельні методи наближення функцій; методи чисельного диференціювання та інтегрування функцій; чисельні методи розв'язування звичайних диференціальних рівнянь; чисельні методи розв'язування інтегральних рівнянь; чисельні методи розв'язування задач математичної фізики; теоретичні особливості чисельних методів та можливості їх адаптації до інженерних задач;
- вміти: проектувати, програмувати, тестувати й відлагоджувати програми, що реалізують чисельні методи; розв'язувати математичні задачі з використанням математичних пакетів; здійснювати обґрунтований вибір чисельного методу при вирішенні практичної задачі.

Одним із способів розв'язування практичної задачі є експеримент. Наприклад, будують ракету, запускають і перевіряють характеристики, які цікавлять. Якщо вони не задовольняють, то будують нову ракету тощо. Зрозуміло, що потрібний результат (при необмежених ресурсах) буде, врешті-решт, досягнутий, проте занадто дорогою ціною. Інший спосіб – побудова математичної (комп'ютерної) моделі об'єкта або явища, що вивчається, і проведення всіх розрахунків на комп'ютері. Для фізичних процесів математична модель зазвичай записується у вигляді набору рівнянь, в які як коефіцієнти входять характеристики тіл або речовин, що беруть участь у процесі. Для того щоб зрозуміти місце чисельних методів у процесі розв'язування задач, що виникають у практичній діяльності людини, з використанням комп'ютера, слід навести основні етапи цього процесу. Етапи розв'язування практичних задач на комп'ютері.

1. Постановка задачі: формулювання задачі та визначення кінцевої мети розв'язку.
2. Побудова математичної моделі, тобто математичне формулювання задачі.
3. Вибір чисельного методу для розв'язування математичної задачі.
4. Розроблення алгоритму.
5. Програмна реалізація алгоритму.
6. Тестування програми (налагодження на тестових задачах).
7. Проведення розрахунків на реальних даних.
8. Аналіз результатів.

Найскладнішим із перерахованих етапів є другий етап, а вивчення методів його реалізації є предметом інших навчальних дисциплін (наприклад, моделювання систем й інших). У подальшому викладенні матеріалу цього навчального посібника передбачено, що математичне формулювання задачі вже є, потрібно тільки навчитися її розв'язувати на комп'ютері з використанням чисельних методів. Варто зазначити, що якщо математична модель вибрана недостатньо коректно, то які б методи не застосовувалися для розрахунків з її використанням, отримані висновки будуть ненадійні, або й зовсім неправильні. Так, для рівняння розрахунків запуску ракети з поверхні Землі, бо в ньому може бути не враховано опір повітря. Розв'язок, отриманий за допомогою чисельного методу, зазвичай є наближеним, тобто містить деяку похибку. Джерелами похибок (у порядку значимості) є:

- невідповідність математичної постановки задачі досліджуваному реальному явищу;
- похибка початкових даних;
- похибка чисельного методу розв'язування;
- помилки заокруглення та розрахунків.

У навчальному посібнику висвітлюються питання, які пов'язані з розкриттям основних понять предметної області: математичної постановки задачі, чисельного методу, ітерації, характеристик чисельних методів, абсолютної та відносної похибок розв'язку, джерел похибок та інших; розглянуті постановки основних математичних задач та найбільш відомі чисельні методи їх розв'язання; наведені приклади програмної реалізації чисельних методів. У ході розв'язання конкретної практичної задачі спеціаліст повинен, перш за все, визначити, до якого типу математичної задачі належить ця практична задача, вибрати чисельний метод для її розв'язання, розробити програмну реалізацію методу самостійно чи вміти застосувати для її розв'язання один із відомих комп'ютерних математичних пакетів. Підведемо підсумки. Мета вивчення дисципліни «Методи обчислень» та «Чисельні методи»:

– надання основних знань з методів обчислень, а також практичних навичок використання методів та засобів сучасних інформаційних технологій у повсякденній практичній діяльності;

– підготувати студентів до ефективного використання сучасних комп'ютерних технологій при розв'язуванні фахових завдань.

Прослухавши курс "«Методи обчислень» та «Чисельні методи», студент повинен вміти обґрунтувати вибір чисельного методу розв'язування математичної задачі, знати особливості його реалізації на комп'ютері, володіти алгоритмом методу, провести необхідні обчислення та аналіз отриманих результатів, а також мати навички практичного використання програмного забезпечення для розв'язування математичних задач.

В результаті вивчення курсу студент повинен знати:

– етапи розв'язування задач з використанням комп'ютера;

– суть математичного моделювання; схему обчислювального експерименту;

– основні групи методів, які використовуються для розв'язування математичних задач;

– вимоги до чисельних методів;

– джерела похибок, їх класифікацію;

– означення абсолютної і відносної похибок, правильної, сумнівної, значущої цифр наближеного числа;

– правила заокруглення;

– загальні формули для похибок;

– правила підрахунку цифр;

– формули подання і способи заокруглення чисел в комп'ютері;

– способи зменшення обчислювальних похибок;

– суть методів, особливості реалізації на комп'ютері, швидкість і умови збіжності, оцінки для наближень до кореня;

– суть методу Жордана-Гауса і його модифікацій, особливості реалізації на комп'ютері;

– методи простих ітерацій і Зейделя, умови збіжності методів;

– суть методу квадратного кореня й особливості його реалізації;

– особливості розв'язування систем із погано обумовленими матрицями;

– постановку задачі наближення функцій, суть методів наближення (інтерполювання, середньоквадратичне наближення, рівномірне), як оптимально вибрати вузли інтерполювання, найпростіші інтерполяційні методи для розв'язування рівнянь з одним невідомим, особливості реалізації методів на комп'ютері;

– суть задач чисельного диференціювання та інтегрування, різні підходи до побудови формул чисельного інтегрування, особливості реалізації на комп'ютері диференціювання та інтегрування;

– постановку задачі, класифікацію методів і суть методів Рунге-Кутта та Ейлера;

– геометричну інтерпретацію різновидів методу Ейлера;

– підходи до оцінки точності методів;

– методи розв'язування задач про власні значення;

– метод скінчених різниць;

– особливості розв'язування крайових задач;

– метод скінчених елементів;

– чисельні методи розв'язування інтегральних рівнянь.

В результаті вивчення цього курсу студент повинен вміти:

– будувати математичні моделі найпростіших об'єктів;

– вивести загальну формулу для обчислення похибок (абсолютної, відносної) функції;

– виконувати дії з наближеними числами;

– оцінювати похибки результатів та обґрунтовувати правила підрахунку цифр;

– виконувати обчислення без точного врахування похибок і розв'язувати пряму і обернену задачі теорії похибок;

– оцінювати похибки заокруглень при виконанні арифметичних операцій на комп'ютері;

– навести приклади задач, які чутливі до похибок вхідних даних; приклади стійких і нестійких методів;

– обґрунтувати збіжність методів, дати їх геометричну інтерпретацію;

– написати алгоритми і програми, застосовувати їх для знаходження із заданою точністю коренів нелінійних рівнянь,

– користуватися бібліотечними програмами та прикладними програмними засобами;

– розв'язувати системи лінійних алгебраїчних рівнянь методами виключення, обчислювати визначники, ранги матриць, обернені матриці, підрахувати число арифметичних дій, що

необхідні для розв'язування системи методами виключення, використовувати бібліотечні програми методів типу Жордана-Гауса;

– застосовувати методи простих ітерацій та Зейделя для розв'язування систем лінійних рівнянь, записувати відповідні алгоритми і програми для методів, обґрунтовувати збіжність, оцінити похибку наближення до розв'язку; розв'язувати системи лінійних рівнянь із симетричними матрицями методом квадратного кореня;

– обґрунтовувати існування та єдиність розв'язку задачі інтерполювання;

– виводити формули інтерполяційних многочленів Лагранжа і Ньютона;

– оцінювати похибку інтерполювання,

– будувати інтерполяційні многочлени і кубічні сплайни;

– обчислювати значення функцій за допомогою інтерполяційних многочленів;

– застосовувати інтерполяційні многочлени для обчислення значень функцій і розв'язування рівнянь;

– обґрунтувати умову застосовності лінійної та квадратичної інтерполяції;

– знаходити найкращу середньоквадратичну апроксимацію функції, що задана на відрізку;

– знаходити методом найменших квадратів наближення таблично заданих функцій;

– будувати емпіричні формули, виконувати згладжування таблично заданих функцій;

– будувати формули чисельного диференціювання та інтегрування, давати їх геометричну інтерпретацію, оцінювати похибки, обчислювати похідні й означені інтеграли, записувати відповідні алгоритми і програми, використовувати бібліотечні програми і прикладні програмні засоби;

– обґрунтовувати методи типу Ейлера, Рунге-Кутта;

– розв'язувати задачу Коші (для одного рівняння та системи першого і вищих порядків) за допомогою формули Тейлора, методами типу Ейлера, Рунге-Кутта;

– розв'язувати задачі на власні значення;

– застосовувати різницеві методи та метод скінчених елементів до розв'язування крайових задач;

– розв'язувати інтегральні рівняння чисельними методами;

– записувати відповідні алгоритми і програми;

– використовувати бібліотечні програми.

Предмети «Методи обчислень» та «Чисельні методи» є логічним продовженням курсів «Математичний аналіз», «Лінійна алгебра», «Диференціальні рівняння» та «Алгоритмізація та програмування» і змістовно пов'язані з базовими дисциплінами. Засвоєння студентами основних положень цих дисциплін має науково-прикладне значення на початковому етапі навчання і формування фахівця загалом. Навчальним планом передбачається: вивчення дисципліни на лекційних та лабораторних заняттях, самостійна робота студентів; перевірка основних теоретичних знань та практичних умінь студентів за допомогою тестових завдань та контрольної роботи; складання заліку. Основними труднощами при вивченні цієї дисципліни є багатоплановість матеріалу, який розглядається, і його великий обсяг. Тому успішне засвоєння курсу неможливе без регулярної самостійної роботи з літературою і творчого відношення до виконання практичних і лабораторних робіт. Під час викладання дисципліни використовуються комп'ютерні навчальні програми та практичні завдання і вправи. Всі лабораторні роботи виконуються на комп'ютері. При проведенні лабораторних робіт на комп'ютері звертається увага на охорону здоров'я студентів.

Розділ 1.
Чисельні методи розв'язування звичайних диференціальних рівнянь

- Загальна постановка задачі
- Метод Ейлера
  - Загальна постановка задачі
  - Графічна інтерпретація
- Метод Рунге-Кутта
  - Загальна постановка задачі
  - Метод Рунге-Кутта другого порядку точності
- Багатокрокові методи розв'язування звичайних диференціальних рівнянь
  - Метод Рунге-Кутта третього порядку точності
  - Метод Рунге-Кутта четвертого порядку точності
- Особливості методу Рунге-Кутта
  - Ідея методу Рунге-Кутта
  - Загальний підхід отримання формул Рунге-Кутта
  - Часткові випадки методу Рунге-Кутта
- Різницеві методи розв'язування задачі Коші для звичайних диференціальних рівнянь (ЗДР)
  - Екстраполяційні та інтерполяційні різницеві формули
  - Часткові випадки різницевих формул. Екстраполяційні формули Адамса
  - Часткові випадки різницевих формул. Інтерполяційні формули Адамса
- Методи скінчених різниць розв'язування крайових задач для диференціальних рівнянь
  - Загальна постановка задачі
  - Основні різницеві співвідношення
  - Метод скінченних різниць розв'язування крайової задачі для диференціальних рівнянь
  - Метод прогонки розв'язування апроксимуючої СЛАР для крайової задачі диференціальних рівнянь з точністю o(h)
  - Метод прогонки розв'язування апроксимуючої СЛАР з точністю $o(h^2)$

- Чисельні методи розв'язування крайових задач для диференціальних рівнянь в частинних похідних. Метод сіток розв'язування крайових задач для диференціальних рівнянь еліптичного типу
  - Метод розв'язування задачі Діріхле
  - Розв'язування рівняння Пуассона
- Метод невизначених коефіцієнтів розв'язування задачі Діріхле (метод Рейхенбаха)
- Часткові випадки методу Рейхенбаха
  - Перша різницева схема розв'язування рівняння Пуассона
  - Друга різницева схема розв'язування рівняння Пуассона
  - Третя різницева схема розв'язування рівняння Пуассона
- Апроксимація крайових умов методом Коллатца
- Розв'язування різницевих рівнянь
  - Різницевий принцип максимуму
  - Розв'язування різницевої системи
- Розв'язування рівнянь гіперболічного типу
  - Метод сіток розв'язування задачі Коші для рівнянь гіперболічного типу
  - Вибір сітки для рівнянь гіперболічного типу
- Метод сіток розв'язування лінійного рівняння параболічного типу (для розв'язування задачі Коші)
- Контрольні питання
- Задачі для самостійної роботи
- Завдання до лабораторної роботи

Звичайні диференційні рівняння (ЗДР) широко застосовуються для математичного моделювання процесів та явищ в різних галузях науки та техніки. Перехідні процеси в радіотехніці, кінетика хімічних реакцій, динаміка біологічних популяцій, рух космічних об'єктів, моделі економічного розвитку досліджуються за допомогою ЗДР [94-102, 110-111, 137-292, 303-320].

При дослідженні різних процесів та проблем в області механіки, фізики, електромеханіки, теорії управління, біології, медицини, економіки, соціології приходять до різних типів лінійних та нелінійних крайових задач для звичайних диференціальних рівнянь. Теорія крайових задач для нелінійних диференціальних рівнянь є одним із актуальних розділів сучасної математики, оскільки запити практики та багатьох теоретичних питань, які ще не повністю розв'язані, багато в чому стимулюють бурхливий розвиток цієї області математики.

1.1. Загальна постановка задачі

До диференційного рівняння n-го порядку як невідомі величини входять функція $y(x)$ та її перші n похідних по аргументу x:

$$\varphi(x, y, y', ..., y^{(n)}) = 0. \quad (1.1)$$

Рівняння (1.1) еквівалентне системі n рівнянь першого порядку

$$\varphi_k(x, y_1, y_1', y_2, y_2', ..., y_n, y_n') = 0, \quad (1.2)$$

де k=1, 2,… , n.

Рівняння (1.1) та еквівалентна йому система (1.2) мають множину розв'язків. Єдині розв'язки відокремлюють за допомогою додаткових умов, які повинні задовольняти шукані розв'язки. В залежності від виду таких умов розглядають три типи задач, для яких доведено існування та єдність розв'язків.

Перший тип – це задачі Коші, або задачі з початковими умовами. Для таких задач крім початкового рівняння (1.1) в будь-якої точці x_0 повинні бути задані початкові умови, тобто значення функції $y(x)$ та її похідних

$$y(x_0) = y_0, \ y'(x_0) = y_{10}, \ ..., \ y^{(n-1)}(x_0) = y_{n-1,0}. \quad (1.3)$$

Для системи ЗДР типа (1.2) початкові умови задаються у вигляді

$$y_1(x_0) = y_{10}, \ y_2(x_0) = y_{20}, \ ..., \ y_n(x_0) = y_{n0}. \quad (1.4)$$

До другого типу задач відносяться так звані граничні або крайові задачі, в яких додаткові умови задаються у вигляді функціональних співвідношень між шуканими розв'язками. Кількість умов повинна співпадати з порядком n-го рівняння або системи. Якщо розв'язок задачі визначається в інтервалі $x \in [x_0, x_k]$, то такі умови можуть бути задані як на межах, так і в інтервалі. Мінімальний порядок ЗДР, для яких може бути сформульована гранична задача, дорівнює двом.

Третій тип задач для ЗДР – це задачі на власні значення. Такі задачі відрізняються тим, що окрім шуканих функцій $y(x)$ та їх похідних до рівняння входять додатково m невідомих параметрів $\lambda_1, \lambda_2, ..., \lambda_m$, які називаються *власними значеннями*. Для розв'язку на інтервалі $[x_0, x_k]$ необхідно задати $m+n$ граничних умов. Як приклад можна назвати задачі визначення власних частот, коефіцієнтів, структури електромагнітних полів і механічних напружень в коливальних системах, задачі знаходження фазових коефіцієнтів, коефіцієнтів затухання, розподілення напруженості полів хвильових процесів, тощо.

Методи наближеного інтегрування диференційних рівнянь можна умовно поділити на три групи:

- *аналітичні*, які дозволяють одержати розв'язок у вигляді аналітичного виразу;
- *графічні*, які дають наближений розв'язок у вигляді графіка;
- *чисельні*, які дають наближений розв'язок у вигляді таблиці.

До чисельного розв'язку ЗДР потрібно звертатися, коли не вдається побудувати аналітичний розв'язок задачі через відомі функції. Хоча для деяких задач чисельні методи є більш ефективними навіть при наявності аналітичних розв'язків. Чисельні методи розв'язування звичайних диференціальних рівнянь задають у явному (1.5) та неявному вигляді (1.6):

$$y' = f(x), \tag{1.5}$$

$$f(x, y, y') = 0. \tag{1.6}$$

Для отримання конкретного розв'язку необхідно мати початкову умову

$$y(x_0) = y_0 = const. \tag{1.7}$$

В теорії диференціальних рівнянь відома **теорема про існування та єдиність розв'язку** задачі (1.5)-(1.7): якщо функція $f(x,y)$ неперервна в деякій обмеженій області $R = \{|x - x_0| < a, |y - y_0| < b\}$ або в деякій частині смуги, де обмеження поставлені лише на деяку обмежену смугу $D = \{|x - x_0| < a, |y| < \infty\}$ і функція задовольняє умову Ліпшіца $|f(x,\widetilde{y}) - f(x,\widetilde{\widetilde{y}})| \le L|\widetilde{y} - \widetilde{\widetilde{y}}|$, то задача (1.5)-(1.7) має розв'язок і цей розв'язок єдиний. Умова Ліпшіца надто жорстка і тому її можна замінити наступною умовою: функція $f(x,y)$ є неперервно-диференціойованою.

1.2. Метод Ейлера

1.2.1. Загальна постановка задачі

Маємо звичайне диференційне рівняння (ЗДР) першого порядку

$$y' = f(x, y). \tag{1.8}$$

Необхідно знайти розв'язок цього рівняння $y = y(x)$, яке задовольняє початковій умові

$$y(x_0) = y_0. \tag{1.9}$$

Така задача називається *задачею Коші*. Чисельний розв'язок задачі Коші полягає у находженні значень $y_1, y_2, ..., y_n$ в точках $x_1 = x_0 + h; x_2 = x_0 + 2h; ...; x_n = x_0 + nh$ відрізка $[a; b]$, де h – крок інтегрування; $x_0 = a; x_n = b$. Якщо нанести точки $(x_0, y_0), (x_1, y_1), ..., (x_n, y_n)$ на площину XOY і з'єднати їх відрізками прямої, то можна одержати ломану лінію, яка називається ламаною Ейлера – наближене зображення кривої. Позначимо $\Delta y_0 = y_1 - y_0$; $\Delta y_1 = y_2 - y_1$; $\Delta y_{n-1} = y_n - y_{n-1}$; $\Delta x_i = x_{i+1} - x_i = h$, $i = 1,2,...,n-1$.

Ейлер Леонард (27.11.1734 - 18.09.1783) – математик, фізик, механік та астроном. Народився в Швейцарії. З 1756 працював у Петербурзькому університеті (Росія). У 1768 він стає незрячим, але продовжуючи роботу, стає автором понад 865 досліджень. Поховано Л. Ейлера у Петербурзі.

Замінимо похідну в (1.8) $y' = \dfrac{dy}{dx} = \dfrac{\Delta y}{\Delta x}$ – відношенням кінцевих різниць

$$\frac{\Delta y}{\Delta x} = f(x, y). \tag{1.10}$$

$$\Delta y = f(x, y)\Delta x. \tag{1.11}$$

При $x = x_0$ маємо $\frac{\Delta y_0}{\Delta x} = f(x_0, y_0), \Delta y_0 = f(x_0, y_0)\Delta x$ або $y_1 = y_0 + hf(x_0, y_0)$. При $x = x_1$ маємо $\Delta y_1 = f(x_1, y_1)\Delta x$ або $y_2 = y_1 + hf(x_1, y_1)$. Для наступних значень x маємо

$$\begin{aligned} y_{i+1} &= y_i + hf(x_i, y_i) \\ &\cdots\cdots\cdots\cdots\cdots \\ y_n &= y_{n-1} + hf(x_{n-1}, y_{n-1}) \end{aligned} \tag{1.12}$$

Таким чином, одержимо таблицю значень $y_1, y_2, ..., y_n$ у відповідних точках $x_1, x_2, ..., x_n$, яка є грубим наближенням розв'язку ЗДР. Метод Ейлера найпростіший і порівняно чисельний метод інтегрування з великою похибкою.

Приклад 1.1. Розв'язати диференційне рівняння $\frac{dy}{dx} = 0{,}05e^{-0{,}05x} - 0{,}0065y$ на інтервалі $[0{,}10]$ $y(0) = 0$.

Розв'язок. $f(x, y) = 0{,}05e^{-0{,}05x} - 0{,}0065y$. Нехай $h = 1$. Розв'язок задачі подамо в вигляді табл. 1.1

Таблиця 1.1

Розв'язок задачі прикладу 1.1

i	x_i	y_i	$F(x_i,y_i)$
0	0	0	0,05
1	1	0,05	$0{,}05e^{-0{,}05\cdot1} - 0{,}0065\cdot0{,}05 = 0{,}0473$
2	2	$0{,}05 + 1\cdot0{,}0473 = 0{,}0973$	$0{,}05e^{-0{,}05\cdot2} - 0{,}0065\cdot0{,}0973 = 0{,}0446$
3	3	$0{,}0973 + 1\cdot0{,}0446 = 0{,}1418$	$0{,}05e^{-0{,}05\cdot3} - 0{,}0065\cdot0{,}1418 = 0{,}0421$
4	4	$0{,}1418 + 1\cdot0{,}0421 = 0{,}184$	$0{,}05e^{-0{,}05\cdot4} - 0{,}0065\cdot0{,}184 = 0{,}0397$
5	5	$0{,}184 + 1\cdot0.0397 = 0{,}2237$	$0{,}05e^{-0{,}05\cdot5} - 0{,}0065\cdot0{,}2237 = 0{,}0375$
6	6	$0{,}2237 + 1\cdot0{,}0375 = 0{,}2612$	$0{,}05e^{-0{,}05\cdot6} - 0{,}0065\cdot0{,}2612 = 0{,}0353$
7	7	$0{,}2612 + 1\cdot0{,}0353 = 0{,}2965$	$0{,}05e^{-0{,}05\cdot7} - 0{,}0065\cdot0{,}2965 = 0{,}0333$
8	8	$0{,}2965 + 1\cdot0{,}0333 = 0{,}3298$	$0{,}05e^{-0{,}05\cdot8} - 0{,}0065\cdot0{,}3298 = 0{,}0314$
9	9	$0{,}3298 + 1\cdot0{,}0314 = 0{,}3612$	$0{,}05e^{-0{,}05\cdot9} - 0{,}0065\cdot0{,}3612 = 0{,}0295$
10	10	$0{,}3612 + 1\cdot0.0295 = 0{,}391$	$0{,}05e^{-0{,}05\cdot10} - 0{,}0065\cdot0{,}391 = 0{,}0278$

1.2.2. Графічна інтерпретація

Нехай задано $y' = f(x, y)$ і $y(x_0) = y_0$. Виберемо достатньо малий крок h і побудуємо систему рівновіддалених вузлів $x_i = x_0 + ih$, $i = \overline{0, \infty}$. Шукану інтегральну криву $y = y(x)$ замінимо ламаною $y = y(x)$ і точкою $M_0(x_0, y_0)$ (рис. 1.1). Отримана ламана лінія M_0, M_1, M_2,... обов'язково повинна проходити через точку M_0 і точку $M_i(x_i, y_i)$ Кожна ланка кривої M_iM_{i+1} є прямолінійною і знаходиться між прямими $x = x_i$ та $x = x_{i+1}$.

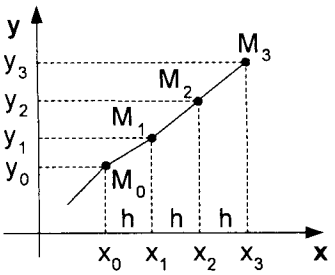


Рис. 1.1. Приклад шуканої інтегральної кривої y =y(x)

Якщо функція $y = f(x)$ в точці x_0 є диференційована, то її графік має дотичну в точці $(x_0, f(x_0))$, кутовий коефіцієнт якої співпадає з першою похідною $f'(x_0)$: $y - y_0 = k(x - x_0)$.

$$y_{i+1} - y_i = k(x_{i+1} - x_i), \quad \frac{y_{i+1} - y_i}{h} = f'(x_i, y_i). \tag{1.13}$$

Ділянка ламаної Ейлера в кожній точці M_i має напрям, що співпадає з напрямом першої похідної $y_i = f(x_i, y_i)$ інтегральної кривої (1.5), яка проходить через точку M_i. На основі формули (1.7) запишемо:

$$y_{i+1} = y_i + \Delta y_i, \tag{1.14}$$

де $\Delta y_i = h \cdot f(x_i, y_i)$, $i = \overline{0,\infty}$.

Формула (1.14) називається *формулою Ейлера*. Метод Ейлера є дуже простим числовим методом інтегрування диференціальних рівнянь і має такі недоліки:

1) мала точність (дає практично задовільний результат при малих значеннях h);

2) систематичне накопичення помилок, тому частіше використовують *модифікований метод Ейлера*:

$$y' = f(x, y). \tag{1.15}$$

$$y'(x_0) = y_0, \; x_i = x_0 + ih, \; i = \overline{0,\infty}. \tag{1.16}$$

На основі методу Ейлера маємо:

$$y_{i+1} = y_i + hf_i, \tag{1.17}$$

де $f_i = f(x_i, y_i)$.

Розглядають якусь проміжну точку $x_{i+\frac{1}{2}} = x_i + \frac{h}{2}$ та шукають

$$y_{i+\frac{1}{2}} = y_i + \frac{h}{2} f_i . \tag{1.18}$$

Знаходимо значення напряму поля інтегральних кривих в точці $\left(x_{i+\frac{1}{2}}, y_{i+\frac{1}{2}} \right)$, тобто $f_{i+\frac{1}{2}} = f\left(x_{i+\frac{1}{2}}, y_{i+\frac{1}{2}} \right)$. На основі цього маємо **модифікований метод Ейлера**:

$$y_{i+1} = y_i + h f_{i+\frac{1}{2}} . \tag{1.19}$$

Іншою модифікацією методу Ейлера є більш удосконалений метод Ейлера-Коші, при якому спочатку визначають грубе наближення розв'язку:

$$\tilde{y}_{i+1} = y_i + h f_i . \tag{1.20}$$

Виходячи з цього знаходять напрям поля інтегральних кривих $\tilde{f}_{i+1} = f(x_{i+1}, \tilde{y}_{i+1})$, а тоді вважають, що

$$y_{i+1} = y_i + h \cdot \frac{f_i + \tilde{f}_{i+1}}{2} . \tag{1.21}$$

Формула (1.21) називається **формулою Ейлера-Коші**.

1.3. Метод Рунге-Кутта

1.3.1. Загальна постановка задачі

Рунге Карл Давід Тольме (30.08.1856 – 03.01.1927) – німецький фізик і математик. Спеціаліст з методів обчислень розв'язування диференціальних рівнянь.

Кутта Мартін Вільгельм (1867–1944) – німецький математик. Спеціаліст з обчислювальної математики.

Маємо диференційне рівняння $y' = f(x, y)$, початкові умови $y(x_0) = y_0$. Знайти $y = y(x)$ у вигляді таблиці на відрізку $[a; b]$.

1. Розіб'ємо відрізок інтегрування $[a, b]$ на n рівних частин системою точок $x_i = x_0 + i \cdot h$ ($i = 0, 1, 2, \ldots, n$), $x_0 = a$, $x_n = b$, $h = \frac{b-a}{n}$.

2. Знайдемо для кожного значення

$$k_1^{(i)} = h f(x_i, y_i) ; \tag{1.22}$$

$$k_2^{(i)} = h f\left(x_i + \frac{h}{2}, y_i + \frac{k_1^{(i)}}{2}\right) ; \tag{1.23}$$

$$k_3^{(i)} = h f\left(x_i + \frac{h}{2}, y_i + \frac{k_2^{(i)}}{2}\right) ; \tag{1.24}$$

$$k_4^{(i)} = hf(x_i + h, y_i + k_3^{(i)}). \quad (1.25)$$

3. Обчислюємо $\Delta y_i = \frac{1}{6}(k_1^{(i)} + 2k_2^{(i)} + 2k_3^{(i)} + k_4^{(i)})$, $(\Delta y_i = y_{i+1} - y_i)$.

4. Обчислюємо послідовно y_i, $i = 1,2...n$, або

$$y_{i+1} = y_i + \frac{1}{6}(k_1^{(i)} + 2k_2^{(i)} + 2k_3^{(i)} + k_4^{(i)}). \quad (1.26)$$

1.3.2. Метод Рунге-Кутта другого порядку точності

На практиці використовують метод Рунге-Кутта від другого до четвертого порядку точності. Метод вищих порядків точності дуже громіздкий, тому його застосовують дуже рідко.

$$y' = f(x, y). \quad (1.27)$$

$$y(x_0) = y_0. \quad (1.28)$$

Задачу (1.27)-(1.28) розв'язуємо, будуючи ітераційну формулу

$$y_{i+1} = y_i + \Delta y_i, \quad (1.29)$$

де Δy_i – приріст функції y, який обчислюється за формулою:

$$\Delta y_i = c_1 k_1 + c_2 k_2, \quad (1.30)$$

де c_1, c_2 – деякі невідомі константи

$$k_1 = hf(x_i, y_i). \quad (1.31)$$

$$k_2 = hf(x_i + \alpha_1 h, y_i + \beta_i k_1). \quad (1.32)$$

де h – деяка стала величина.

Для того, щоб метод (1.29)-(1.32) був другого порядку точності необхідно, щоб розклад функції y в ряд Тейлора в околі точки x_i до другого доданка включно узгоджувався з формулою (1.29). Запишемо розклад функції y в ряд Тейлора:

$$y_{i+1} = y(x_i + h) = y(x_i) + h \cdot y'(x_i) + \frac{h^2}{2!} y''(x_i) + O(h^2). \quad (1.33)$$

Розкладемо в ряд Тейлора приріст Δy_i з формули (1.30) із врахуванням (1.31)-(1.32)

$$\Delta y_i = h[c_1 f(x_i, y_i) + c_2 f(x_i + \alpha_1 h, y_i + \beta_1 hf(x_i, y_i))]. \quad (1.34)$$

Врахуємо, що f є функцією двох змінних при розкладі в ряд Тейлора, тоді

$$\Delta y_i = h\left\{c_1 f(x_i, y_i) + c_2\left[f(x_i, y_i) + \alpha_1 h \frac{\partial f}{\partial x}(x_i, y_i) + \beta_1 hf(x_i, y_i)\frac{\partial f}{\partial y}(x_i, y_i)\right]\right\}. \quad (1.35)$$

Згрупуємо доданки з однаковими степенями h, тоді

$$\Delta y_i = hf(x_i, y_i)\cdot[c_1 + c_2] + h^2\left[c_2\alpha_1\frac{\partial f(x_i, y_i)}{\partial x} + c_2\beta_1 f(x_i, y_i)\frac{\partial f(x_i, y_i)}{\partial y}\right]. \quad (1.36)$$

Введемо оператор

$$D = \frac{\partial}{\partial x} + f\frac{\partial}{\partial y}. \quad (1.37)$$

Враховуючи (1.27) маємо, що

$$y' = \frac{\partial y}{\partial x} = f. \quad (1.38)$$

$$y'' = \frac{d}{dx}(y') = \frac{d}{dx}f(x, y) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial x} = Df. \quad (1.39)$$

Формулу (1.29) можна переписати так:

$$y_{i+1} = y_i + h(c_1 + c_2)\cdot f_i + h^2\left[c_2\alpha_1\frac{\partial f_i}{\partial x} + c_2\beta_1 f_i\frac{\partial f_i}{\partial y}\right]. \quad (1.40)$$

У формулі (1.40) розглянемо частковий випадок, коли $\alpha_1 = \beta_1 = 1$. Коефіцієнти вибираються рівними для того, щоб вираз при h^2 відповідав формулі (1.37), тобто щоб утворився диференціал D.

$$y_{i+1} = y_i + h(c_1 + c_2)f_i + h^2 c_2 Df_i. \quad (1.41)$$

Порівнюючи коефіцієнти при однакових степенях у формулах (1.40) і (1.33), отримуємо систему рівнянь

$$\begin{cases} c_1 + c_2 = 1, \\ c_2 = 0.5 \end{cases} \Rightarrow c_1 = c_2 = 0.5. \quad (1.42)$$

Тоді формули (1.29)-(1.32) будуть мати вигляд:

$$y_{i+1} = y_i + \Delta y_i. \quad (1.43)$$

$$\Delta y_i = \frac{1}{2}(k_1 + k_2). \quad (1.44)$$

$$k_1 = hf(x_i, y_i). \quad (1.45)$$

$$k_2 = hf(x_{i-1}, y_i + k_1). \quad (1.46)$$

Метод Рунге- Кутта має значно більш високу точність ніж матод Ейлера, що дозволяє збільшити крок розв'язання. Його максимальну величину визначає припустима похибка. Такий вибір часто здійснюється автоматично і включається як складова частина в алгоритм методу Рунге- Кутта.

Будь-яку з формул однокрокових методів використовують для розв'язання систем диференціальних рівнянь і диференціальних рівнянь вищих порядків.

1.4. Багатокрокові методи розв'язування звичайних диференціальних рівнянь

1.4.1. Метод Рунге-Кутта третього порядку точності

Методи, які для обчислення наступної точки використовують інформацію про попередню точку, називаються **однокроковими**. Методи, які для обчислення наступної точки використовують інформацію про декілька попередніх точок, називаються **багатокроковими**. Серед багатокрокових методів найпопулярнішими є *методи Адамса* до 12-го порядку включно та *Гіра* до 6-го порядку включно. Розглянемо метод Рунге-Кутта третього порядку точності.

$$y' = f(x, y). \tag{1.47}$$

$$y(x_0) = y_0. \tag{1.48}$$

Задачу (1.47)-(1.48) розв'язуємо, будуючи ітераційну формулу

$$y_{i+1} = y_i + \Delta y_i. \tag{1.49}$$

де Δy_i – приріст функції y, який обчислюється за формулою:

$$\Delta y_i = c_1 k_1 + c_2 k_2 + c_3 k_3. \tag{1.50}$$

де c_1, c_2, c_3 – деякі невідомі константи

$$k_1 = hf(x_i, y_i). \tag{1.51}$$

$$k_2 = hf(x_i + \alpha_2 h, y_i + \beta_{21} k_1). \tag{1.52}$$

$$k_3 = hf(x_i + \alpha_3 h, y_i + \beta_{31} k_1 + \beta_{32} k_2). \tag{1.53}$$

де h – деяка стала величина.

Коефіцієнти α, β, c вибираємо так, щоб формули (1.49)-(1.53) узгоджувались з розкладом функції y в ряд Тейлора в околі точки x_i до доданка, що містить h^3 включно. Введемо оператор диференціала

$$D = \frac{\partial}{\partial x} + f\frac{\partial}{\partial y}. \tag{1.54}$$

$$\begin{aligned} D(Df) &= D\left(\frac{\partial f}{\partial x} + f\frac{\partial f}{\partial y}\right) = D\left(\frac{\partial f}{\partial x}\right) + D(f)\frac{\partial f}{\partial y} + fD\left(\frac{\partial f}{\partial y}\right) = \\ &= \frac{\partial}{\partial x}\cdot\left(\frac{\partial f}{\partial x}\right) + f\frac{\partial}{\partial y}\cdot\left(\frac{\partial f}{\partial x}\right) + \frac{\partial f}{\partial x}\cdot\frac{\partial f}{\partial y} + f\frac{\partial f}{\partial y}\cdot\frac{\partial f}{\partial y} + f\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) + f^2\frac{\partial}{\partial y}\cdot\frac{\partial f}{\partial y} = \\ &= \frac{\partial^2 f}{\partial x^2} + 2f\frac{\partial^2 f}{\partial x \partial y} + f^2\frac{\partial^2 f}{\partial y^2} + \frac{\partial f}{\partial y}\cdot Df. \end{aligned} \tag{1.55}$$

$$D^2 = \frac{\partial^2}{\partial x^2} + 2f\frac{\partial^2}{\partial x \partial y} + f^2\frac{\partial^2}{\partial y^2}. \tag{1.56}$$

Тоді формулу (1.54) запишемо із врахуванням формули (1.56)

$$D(Df) = D^2 f + \frac{\partial f}{\partial y} \cdot Df \,. \tag{1.57}$$

Аналогічно можна вивести оператор D^3:

$$D^3 = \frac{\partial^3}{\partial x^3} + 3f \frac{\partial^3}{\partial x^2 \partial y} + 3f^2 \frac{\partial^3}{\partial x \partial y^2} + f^3 \frac{\partial^3}{\partial y^3} \,. \tag{1.58}$$

Запишемо розклад функції y в ряд Тейлора:

$$y(x_i + h) = y(x_i) + hy'(x_i) + \frac{h^2}{2!} y''(x_i) + \frac{h^3}{3!} y'''(x_i) + O(h^4) \,. \tag{1.59}$$

У формулах (1.49)-(1.53) розкладаємо в ряд Тейлора коефіцієнти k_2, k_3 в околі точки (x_i, y_i). Коефіцієнти в отриманій формулі та формулі (1.59) повинні узгоджуватись. При цьому враховуємо, що

$$y' = f(x,y) \,.\ y'' = Df \,,\ y''' = D^2 f + \frac{\partial f}{\partial y} \cdot Df \,. \tag{1.60}$$

Введемо такі позначення:

$$f_x = \frac{\partial f}{\partial x} \,,\ f_y = \frac{\partial f}{\partial y} \,,\ f_{x^2} = \frac{\partial^2 f}{\partial x^2} \,. \tag{1.61}$$

Тоді формули (1.54)-(1.57) матимуть вигляд:

$$Df = f_x + f \cdot f_y \,. \tag{1.62}$$

$$D^2 f = f_{x^2} + 2f \cdot f_{xy} + f^2 \cdot f_{y^2} \,. \tag{1.63}$$

$$D^3 f = f_{x^3} + 3f \cdot f_{x^2 y} + 3f^2 \cdot f_{xy^2} + f^3 \cdot f_{y^3} \,. \tag{1.64}$$

Після розкладу в ряд Тейлора коефіцієнтів k_2, k_3 отримаємо, що

$$y_{i+1} = y_i + \Delta y_i \,,\ i = \overline{0, \infty} \,. \tag{1.65}$$

$$\Delta y_i = \frac{1}{6}(k_1 + 4k_2 + k_3) \,. \tag{1.66}$$

$$k_1 = h \cdot f(x_i, y_i) \,. \tag{1.67}$$

$$k_2 = h \cdot f\left(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}\right) \,. \tag{1.68}$$

$$k_3 = h \cdot f(x_{i+1} y_i + k_1 + 2k_2) \,. \tag{1.69}$$

1.4.2. Метод Рунге-Кутта четвертого порядку точності

Тепер розглянемо метод Рунге-Кутта четвертого порядку точності.

$$y_{i+1} = y_i + \Delta y_i,\ i = \overline{0,\infty}. \tag{1.70}$$

$$\Delta y_i = \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4). \tag{1.71}$$

$$k_1 = h \cdot f(x_i, y_i). \tag{1.72}$$

$$k_2 = h \cdot f\left(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}\right). \tag{1.73}$$

$$k_3 = h \cdot f\left(x_i + \frac{h}{2}, y_i + \frac{k_2}{2}\right). \tag{1.74}$$

$$k_3 = h \cdot f\left(x_i + h, y_i + k_3\right). \tag{1.75}$$

У таблиці 1.2 наведено схему Рунге-Кутта

Таблиця 1.2

Схема Рунге–Кутта

i	x	y	$f(x,y)$	$k=hf(x,y)$	Δy
0	x_0	y_0	$f(x_0, y_0)$	$k_1^{(0)}$	$k_1^{(0)}$
	$x_0 + \frac{h}{2}$	$y_0 + \frac{k_1^{(0)}}{2}$	$f(x_0 + \frac{h}{2}, y_0 + \frac{k_1^{(0)}}{2})$	$k_2^{(0)}$	$2k_2^{(0)}$
	$x_0 + \frac{h}{2}$	$y_0 + \frac{k_2^{(0)}}{2}$	$f(x_0 + \frac{h}{2}, y_0 + \frac{k_2^{(0)}}{2})$	$k_3^{(0)}$	$2k_3^{(0)}$
	$x_0 + h$	$y_0 + k_3^{(0)}$	$f(x_0 + h, y_0 + k_3^{(0)})$	$k_4^{(0)}$	$k_4^{(0)}$
	$\Delta y_0 = \frac{1}{6}(k_1^{(0)} + 2k_2^{(0)} + 2k_3^{(0)} + k_4^{(0)})$				
1	x_1	$y_1 = y_0 + \Delta y_0$	$f(x_1, y_1)$	$k_1^{(1)}$	$k_1^{(1)}$
	$x_1 + \frac{h}{2}$	$y_1 + \frac{k_1^{(1)}}{2}$	$f(x_1 + \frac{h}{2}, y_1 + \frac{k_1^{(1)}}{2})$	$k_2^{(1)}$	$2k_2^{(1)}$
	$x_1 + \frac{h}{2}$	$y_1 + \frac{k_2^{(1)}}{2}$	$f(x_1 + \frac{h}{2}, y_1 + \frac{k_2^{(1)}}{2})$	$k_3^{(1)}$	$2k_2^{(1)}$
	$x_1 + h$	$y_1 + k_3^{(1)}$	$f(x_1 + h, y_1 + k_3^{(1)})$	$k_4^{(1)}$	$k_4^{(1)}$
	$\Delta y_1 = \frac{1}{6}(k_1^{(1)} + 2k_2^{(1)} + 2k_3^{(1)} + k_4^{(1)})$				
2	x_2	$y_2 = y_1 + \Delta y_1$	$f(x_2, y_2)$	$k_1^{(2)}$	$k_1^{(2)}$
	$x_2 + \frac{h}{2}$	$y_2 + \frac{k_1^{(2)}}{2}$	$f(x_2 + \frac{h}{2}, y_2 + \frac{k_1^{(2)}}{2})$	$k_2^{(2)}$	$2k_2^{(2)}$

Продовження таблиці 1.2

	$x_2+\frac{h}{2}$	$y_2+\frac{k_2^{(2)}}{2}$	$f(x_2+\frac{h}{2}, y_2+\frac{k_2^{(2)}}{2})$	$k_3^{(2)}$	$2k_3^{(2)}$
	x_2+h	$y_2+k_3^{(2)}$	$f(x_2+h, y_2+k_3^{(2)})$	$k_4^{(2)}$	$k_4^{(2)}$
	$\Delta y_2=\frac{1}{6}(k_1^{(2)}+2k_2^{(2)}+2k_3^{(2)}+k_4^{(2)})$				
	.....				
n	x_n	$y_n=y_{n-1}+\Delta y_{n-1}$	$f(x_n, y_n)$	$k_1^{(n)}$	$k_1^{(n)}$
	$x_n+\frac{h}{2}$	$y_n+\frac{k_1^{(n)}}{2}$	$f(x_n+\frac{h}{2}, y_n+\frac{k_1^{(n)}}{2})$	$k_2^{(n)}$	$2k_2^{(n)}$
	$x_n+\frac{h}{2}$	$y_n+\frac{k_2^{(n)}}{2}$	$f(x_n+\frac{h}{2}, y_n+\frac{k_2^{(n)}}{2})$	$k_3^{(n)}$	$2k_3^{(n)}$
	x_n+h	$y_n+k_3^{(n)}$	$f(x_n+h, y_n+k_3^{(n)})$	$k_4^{(n)}$	$k_4^{(n)}$
	$\Delta y_n=\frac{1}{6}(k_1^{(n)}+2k_2^{(n)}+2k_3^{(n)}+k_4^{(n)})$				

Приклад 1.2. Маємо диференційне рівняння $y'=0{,}25y^2+x^2$, початкові умови $y(0)=-1$. Знайти $y=y(x)$ у вигляді таблиці на відрізку $[0;0{,}2]$, $h=0{,}1$.

Розв'язок. $f(x,y)=0{,}25y^2+x^2$, $x_0=0$, $y(x_0)=y_0=-1$.

Таблиця 1.3

Розв'язок прикладу 1.2

i	x	y	$f(x,y)$	$k=hf(x,y)$	Δy
0	0	-1	$0{,}25\cdot(-1)^2+0^2=$ $=0{,}25$	$0{,}1\cdot 0{,}25=0{,}025$	0,025
	$0+\frac{0{,}1}{2}=0{,}05$	$-1+\frac{0{,}025}{2}=$ $=-0{,}9875$	....	....	....

1.5. Особливості методу Рунге-Кутта

1.5.1. Ідея методу Рунге-Кутта

Нехай для задачі Коші

$$y'(x)=f(x,y). \tag{1.76}$$

$$y(x_0)=y_0. \tag{1.77}$$

Відомо значення розв'язку $y(x) \in \mathbf{R}^1$. Як знайти $y(x+h)$, $h>0$?

Справедлива рівність

$$y(x+h) = y(x) + \int_0^h y'(x+t)dt . \tag{1.78}$$

Випадок 1.1. Якщо обчислити $I = \int_0^h y'(x+t)dt$ за допомогою формули прямокутників, то

$$I + h \cdot y'(x) + o(h^2) . \tag{1.79}$$

З (1.76) маємо $I + h \cdot \int (x, y(x)) + o(h^2)$, тоді (1.78) дає формулу

$$y(x+h) = y(x) + h \cdot \int (x, y(x)) + o(h^2) . \tag{1.80}$$

Зауваження 1.1. Алгоритм розрахунків для комп'ютерів. Позначимо

$$x_i = x_0 + i \cdot h ;\ i = \overline{1,n} ;\ x_n = x_0 + X ;\ y_i \equiv y(x_i) . \tag{1.81}$$

$$y_{j+1} = y_j + h \cdot f(x_j, y_j) + o(h^2) . \tag{1.82}$$

Випадок 1.2. Наближаємо інтеграл I формулою трапецій

$$I = \frac{h}{2}[y'(x) + y'(x+h)] + o(h^3) . \tag{1.83}$$

Підставимо (1.83) у (1.78), маємо

$$y(x+h) = y(x) + \frac{h}{2}[f(x, y(x)) + f(x+h, y(x+h))] + o(h^3) . \tag{1.84}$$

Зауваження 1.2. Для розрахунків будемо мати формулу Адамса 2-го порядку:

$$y_{j+1} = y_j + \frac{h}{2}[f(x_j, y_j) + f(x_{j+1}, y_{j+1})] + o(h^3) . \tag{1.85}$$

Ця формула є неявною, бо треба визначити y_{j+1} як корені трансцендентного рівняння (1.85).

Адамс Джон Куг (05.06.1819 – 21.01.1892) – англійський астроном і математик. Спеціаліст з чисельного інтегрування.

Щоб формула (1.85) стала явною і, зручною для розрахунків, потрібно в праву частину (1.84) підставити формулу Ейлера (1.80). Тоді позначивши $k(h) \equiv h \cdot f(x_j, y_j)$, одержимо

$$y(x+h) = y(x) + \frac{h}{2}[f(x, y) + f(x+h, y + k(h))] + o(h^3) . \tag{1.86}$$

Розрахункова формула, щоб одержати таблицю значень для розв'язків задачі Коші (1.76), (1.77), має вигляд:

$$\begin{cases} k(h) \equiv h \cdot f(x_j, y_j), \\ y_{j+1} = y_j + \dfrac{h}{2}[f(x_j, y_j) + f(x_{j+1} + h_j, y_j + k(h))]. \end{cases} \tag{1.87}$$

Якщо (1.84) задовольняє умови оператора, то (1.85) можна розв'язати методом простої ітерації:

$$y_{j+1}^{(k+1)} = y_j^{(k)} + \frac{h}{2}[f(x_j, y_j^{(k)}) + f(x_{j+1}, y_{j+1}^{(k)})], \quad k = 0,1,2,\ldots. \tag{1.89}$$

при $y_{j+1}^{(0)} = y_j + h \cdot f(x_j, y_j)$ знаходимо методом Ейлера.

Випадок 1.3. Якщо

$$I = h \cdot y'(x + \frac{h}{2}) + o(h^2). \tag{1.90}$$

$$y(x+h) = y(x) + h \cdot f(x + \frac{h}{2}, y(x + \frac{h}{2})) + o(h^2). \tag{1.91}$$

Для розрахунків зручні формули

$$\begin{cases} k(h) \equiv h \cdot f(x, y), \\ y(x+h) = y(x) + h \cdot f(x + \dfrac{h}{2}, y(x) + \dfrac{k(h)}{2}) + o(h^2). \end{cases} \tag{1.92}$$

де $y(x + \frac{h}{2}) \approx y(x) + \frac{h}{2} f(x, y(x))$ обчислюємо за формулою Ейлера (1.78) з кроком h. Формули (1.86) та (1.92) вигідні тим, що для обчислення треба вести розрахунки правої частини рівняння (1.76) в різних точках. Можна обчислення $f(x, y(x))$ оформити у вигляді оператора-функції у комп'ютерному середовищі.

1.5.2. Загальний підхід отримання формул Рунге-Кутта

Приведемо загальний підхід Рунге-Кутта наближеного обчислення розв'язків задачі Коші (1.76), (1.77). Введемо параметри

$$a_1 = 0; a_2, \ldots, a_r \neq 0; \quad p_1, p_2, \ldots, p_r; \{\beta_{ij}\}, 0 < j < i \leq r. \tag{1.93}$$

Якщо порушуються нерівності (1.93) у індексах, то завжди $\beta_{ij} = 0$.

Обчислимо допоміжні конструкції

$$k_1(h) \equiv h \cdot f(x, y); \tag{1.94}$$

$$k_2(h) \equiv h \cdot f(x + a_2 h; y + \beta_{21} k_1(h)); \tag{1.95}$$

$$k_3(h) \equiv h \cdot f(x + a_3 h; y + \beta_{31} k_1(h) + \beta_{32} k_2(h)); \tag{1.96}$$

$$k_r(h) \equiv h \cdot f(x + a_r h; y + \beta_{r1} k_1(h) + ... + \beta_{r,r-1} k_{r-1}(h)) . \quad (1.97)$$

Тоді можна наближено обчислити за такою формулою

$$y(x+h) \approx y(x) + \sum_{i=1}^{r} p_i k_i(h) \equiv z(h) . \quad (1.98)$$

Розв'яжемо проблему визначення параметрів a_i, β_{ij}, p_i. Позначимо через

$$\varphi(h) \equiv y(x+h) - z(h) . \quad (1.99)$$

похибку методу Рунге-Кутта. Якщо $f(x,y)$ аналітична функція, то $k_1(h), k_2(h), ..., k_r(h)$ аналітичні за параметром h.

Нехай для $f(x,y)$ існують похідні, наприклад, $\varphi'(h), \varphi''(h), ..., \varphi^{(s-1)}(h)$. Тоді a_i, β_{ij}, p_i вибираємо за умови, що

$$\varphi(0) = \varphi'(0) = \varphi''(0) = ... = \varphi^{(s)}(0) = 0 , \ \varphi^{(s+1)}(0) \neq 0 . \quad (1.100)$$

За формулою Тейлора з врахуванням (1.100) маємо

$$\varphi(h) = \sum_{i=1}^{s} \frac{\varphi^{(i)}(0)}{i!} h^i + \frac{\varphi^{(s+1)}(\theta)}{(s+1)!} h^{s+1} = \frac{\varphi^{(s+1)}(\theta)}{(s+1)!} h^{s+1} , \ 0 < \theta < 1 . \quad (1.101)$$

Означення 1.1. Функцію $\varphi(h)$ назвемо *похибкою методу Рунге-Кутта* на кроці h; s – порядок похибки: r – порядок методу Рунге-Кутта. Точність методу тим вища, чим більший порядок похибки методу.

1.5.3. Часткові випадки методу Рунге-Кутта

Розглянемо часткові випадки, які містяться у наступних теоремах.

Теорема 1.1. (r=1) Нехай $f \in C^{2,2}(\mathbf{R}^1, \mathbf{R}^1)$. Тоді:

1) для r=1, s=1 існує єдина формула Рунге-Кутта з точністю $o(h^2)$:

$$y(x+h) = y(x) + h \cdot f(x,y) + o(h^2) . \quad (1.102)$$

2) для r=1, $s \geq 2$ формули Рунге-Кутта не існують.

Доведення.

1. Запишемо похибку методу (1.99) для r=1, s=1:

$$\varphi = y(x+h) - y(x) - p_1 h \cdot f(x,y) \quad (1.103)$$

Зрозуміло, що

$$\varphi(0) = 0; \varphi'(0) = (y'(x+h) - p_1 f(x,y))\big|_{h=0} = f(x,y)(1 = p_1) . \quad (1.104)$$

Тоді $\varphi'(0) = 0$ для $p_1 = 1$. Таким чином. (1.98) дає при $p_1 = 1$ та r=1 єдину формулу (1.103).

2. Знайдемо другу похідну від $\varphi(h)$ по h, вводячи позначення

$\bar{x} = x + h, \bar{y} = y(x+h)$:

$$\varphi''(h) = y''_{hh}(x+h) = \frac{d}{dh}(y_h(x+h)) = \frac{d}{dh}(f(x+h, y(x+h))) =$$
$$= \frac{df}{d\bar{x}}\frac{d\bar{x}}{dh} + \frac{df}{d\bar{y}}\frac{d\bar{y}}{dh} + \frac{df}{dh}\frac{dh}{dh} = f'_{\bar{x}} + f'_{\bar{y}}\bar{y}' + f'_h = f'_{\bar{x}} + f_{\bar{y}} f \neq 0. \tag{1.105}$$

Взагалі кажучи $\varphi''(0) \neq 0$ для будь-якої $f(x, y)$. Дійсно, якщо, наприклад, $f(x, y) = y$, тоді $\varphi''(0) = 1 \cdot y$, що дає $\varphi''(0) \neq 0$. Звідси випливає, що для $r=1$, $s \geq 2$ формули Рунге-Кутта не існують.

Теорема 1.2. ($r=2$) Нехай $f \in C^{3,3}(\mathbf{R}^1, \mathbf{R}^1)$. Тоді

1) Існує безліч формул Рунге-Кутта для $r=2$; $s=2$ з точністю $o(h^3)$;
2) Для $r=2$, $s \geq 3$ не існує формули Рунге-Кутта.

Доведення. Похибка методу (1.99) для $r=2$, $s=2$ матиме вигляд

$$\varphi(h) = y(x+h) - y(x) - p_1 \cdot h \cdot f(x, y) - p_2 \cdot h \cdot f(x, y). \tag{1.106}$$

$$\bar{x} = x + a_2 \cdot h; \tag{1.107}$$

$$\bar{y} = y + \beta_{21} \cdot h \cdot f(x, y(x)). \tag{1.108}$$

Обчислимо похідні від похибки $\varphi(h)$

$$\varphi'(h) = y'(x+h) - p_1 \cdot f(x, y) - p_2 \cdot f(\bar{x}, \bar{y}) -$$
$$- p_2 \cdot h \cdot [a_2 \cdot f'_{\bar{x}}(\bar{x}, \bar{y}) + \beta_{21} \cdot f'_{\bar{y}}(\bar{x}, \bar{y}) \cdot f(x, y)]_1; \tag{1.109}$$

$$\varphi''(h) = y''(x+h) - 2p_2[a_2 \cdot f'_{\bar{x}}(\bar{x}, \bar{y}) + \beta_{21} \cdot f'_{\bar{y}}(\bar{x}, \bar{y}) \cdot f(x, y)]_1 -$$
$$- p_2 h[a_2^2 f''_{\bar{x}\bar{x}}(\bar{x}, \bar{y}) + 2a_2\beta_{21} f''_{\bar{x}\bar{y}}(\bar{x}, \bar{y}) f(x, y) + \beta_{21}^2 f''_{\bar{y}\bar{y}}(\bar{x}, \bar{y}) f^2(x, y)]_2; \tag{1.110}$$

$$\varphi'''(h) = y'''(x+h) - 3p_2[a_2^2 f''_{\bar{x}\bar{x}}(\bar{x}, \bar{y}) +$$
$$+ 2a_2\beta_{21} f''_{\bar{x}\bar{y}}(\bar{x}, \bar{y}) f(x, y) + \beta_{21}^2 f''_{\bar{y}\bar{y}}(\bar{x}, \bar{y}) f^2(x, y)]_2 + o(h); \tag{1.111}$$

$$y' = f; y'' = \frac{d}{dh}(f) = f'_{\bar{x}} + f'_{\bar{y}} f; y'''(x+h) = \frac{d}{dh}(y''_{hh}(x+h)) =$$
$$= \frac{d}{dh}(f'_{\bar{x}} + f'_{\bar{y}} y' + f'_h) = f''_{\bar{x}\bar{x}} + 2f''_{\bar{x}\bar{y}} f + f''_{\bar{y}\bar{y}} f^2 + f'_{\bar{y}} y''. \tag{1.112}$$

Обчислимо y', y'' та y''' при $h=0$ за формулами (1.112) та підставимо у (1.109)-(1.111). В результаті одержимо $\varphi(0) = 0$; $\varphi'(0) = (1 - p_1 - p_2) f(x,)$; $\varphi''(0) = (1 - 2p_2 a_2) f'_x(x, y) + (1 - 2p_2\beta_{21}) f'_y(x, y)$:

$$\varphi'''(0) = (1 - 3p_2 a_2^2 f''_{xx}(x, y) + (2 - 6p_2 a_2 \beta_{21}) f''_{xy}(x, y) f(x, y) +$$
$$+ (1 - 3p_2\beta_{21}^2) f''_{yy}(x, y) f^2(x, y) + f'_y(x, y) y''(x). \tag{1.113}$$

Таким чином, $\varphi(0)=\varphi'(0)=\varphi''(0)=0$ для $\forall f(x,y)$, якщо

$$\left.\begin{aligned} 1-p_1-p_2&=0,\\ 1-2p_2a_2&=0,\\ 1-2p_2\beta_{21}&=0. \end{aligned}\right\}. \tag{1.114}$$

Нелінійна система (1.114) має безліч розв'язків. Далі $\varphi'''(0)\neq 0$ для $\forall f(x,y)$. Наприклад, для $f(x,y)=y$, тоді $\varphi'''(0)\neq 0$, бо $\varphi'''(0)$ містить член з $y\cdot 1\neq 0$ в другій похідній $y''(x)=f_x'+f_y''\cdot f$. Значить,

1) для r=2; s=2 існує безліч формул Рунге-Кутта, бо система (1.114) має безліч дійсних розв'язків $\{p_1,p_2,\alpha_2,\beta_{21}\}\subset \mathbf{R}^1$:

$$y(\mathrm{x}+\mathrm{h})=\mathrm{y}(\mathrm{x})+\mathrm{p}_1\cdot h\cdot f(x,y)+p_2\cdot h\cdot f(\bar{\mathrm{x}},\bar{\mathrm{y}})+\mathrm{o}(\mathrm{h}^3); \tag{1.115}$$

2) для r=2; $s\geq 3$ формул не існує.

Наслідок 1.1. Параметри p_1=1; p_2=1; $a_2=\frac{1}{2};\beta_{21}=\frac{1}{2}$ є розв'язками системи (1.114), тоді формули Рунге-Кутта (1.98) можна подати у вигляді

$$\left.\begin{aligned} &k_1(\mathrm{h})\equiv h\cdot f(x,y);\\ &y(x+h)=y(x)+h\cdot f(\mathrm{x}+\frac{h}{2};\mathrm{y}+\frac{1}{2}\cdot \mathrm{k}_1(\mathrm{h}))+\mathrm{o}(\mathrm{h}^3). \end{aligned}\right\}. \tag{1.116}$$

Це формула (1.87).

Наслідок 1.2. Параметри $p_1=p_2=\frac{1}{2};\alpha_2=\beta_{21}=1$ є розв'язком (1.114), тоді формула (1.98) прийме вигляд

$$k_1(h)\equiv h\cdot f(x,y);k_2(h)\equiv h\cdot f(x+h;y+k_1(h)); \tag{1.117}$$

$$y(x+h)=y(x)+\frac{1}{2}hf(x,y)+\frac{1}{2}hf(\mathrm{x}+\mathrm{h};\mathrm{y}+\mathrm{h}\,f(x,y))+o(\mathrm{h}^3). \tag{1.118}$$

Це формула (1.92).

Теорема 1.3. (r=3) Нехай $f\in C^{4,4}(\mathbf{R}^1\times\mathbf{R}^1)$ Тоді

1) існує безліч формул Рунге-Кутта для r=3; s=3;
2) не існує формул Рунге-Кутта для r=3; $s\geq 4$, причому

$$\varphi(h)=\frac{\varphi^{(4)}(\theta\cdot \mathrm{h})}{4p}h^4;0<\theta<1. \tag{1.119}$$

Доведення. Для r=3 запишемо формулу (1.92):

$$y(x+h)\approx y(x)+p_1\cdot k_1(h)+p_2\cdot k_2(h)+p_3\cdot k_3(h), \tag{1.120}$$

а похибка

$$\varphi(h) = y(x+h) - y(x) - p_1 \cdot h \cdot f(x,y) - p_2 \cdot h \cdot f(x + a_2 h; y + \\ + \beta_{21}\,\mathrm{h} \cdot f(\mathrm{x},\mathrm{y})) - p_3 \cdot h \cdot f(x + a_3 h; y + \beta_{31}\,\mathrm{h} \cdot f(\mathrm{x},\mathrm{y}) + \beta_{31} \cdot \mathrm{h} \times \\ \times f(x + a_2 h; y + \beta_{21}\,\mathrm{h} \cdot f(\mathrm{x},\mathrm{y}))) \tag{1.121}$$

Якщо обчислити

$$\varphi'(h), \varphi''(h), \varphi'''(x), \varphi^{(4)}(h), \tag{1.122}$$

то $\varphi(0) = \varphi'(0) = \varphi''(0) = \varphi'''(0) = 0$, для $\forall\{f(x,y)\}$.

Параметри $\{p_1, p_2, p_3, \alpha_2, \alpha_3, \beta_{21}, \beta_{31}, \beta_{32}\} \subset \mathbf{R}^1$ задовольняють систему 6 рівнянь з 8-ма невідомими

$$\left.\begin{array}{l} 1 - p_1 - p_2 - p_3 = 0; \\ 1 - 2p_1 a_2 - 2p_3 a_3 = 0; \\ 1 - 3p_1 a_2^2 - 3p_3 a_3^2 = 0; \\ 1 - 6p_2 \beta_{32} a_2 = 0; \\ a_2 - \beta_{21} = 0; \\ a_3 - \beta_{21} - \beta_{32} = 0. \end{array}\right\} \tag{1.123}$$

Система (1.123) має *множину* дійсних розв'язків, а значить і множину формул Рунге-Кутта (1.120). Але $\varphi^{(4)} \neq 0$ для $\forall f(x,y)$, тобто $\varphi^{(4)} = o(h^4)$.

Наслідок 1.3. Найбільш поширену формулу Рунге-Кутта (1.120) одержимо для $a_2 = a_3 = \beta_{21} = \frac{1}{2}$; $\beta_{31} = -1; \beta_{32} = 2; p_1 = p_3 = \frac{1}{6}; p_2 = \frac{2}{3}$ (розв'язки (1.123)).

Тоді формула (1.120) має вигляд

$$y(x+h) = y(x) + \frac{1}{6}[k_1(h) + 4k_2(h)] + o(h^4), \tag{1.124}$$

де $k_1(h) \equiv h \cdot f(x,y)$, $k_2(h) \equiv h \cdot f(x + \frac{h}{2}, y + \frac{k_1(h)}{2})$, $k_3(h) \equiv h \cdot f(x + \frac{h}{2}, y - k_1(h) + 2k_2(h))$.

Зауваження 1.3. Якщо права частина $f(x,y) \equiv f(x)$, то (1.124) є *формулою Сімпсона*

$$y(x+h) - y(x) = \frac{h}{6}[f(x) + 4f(x + \frac{h}{2}) + f(x+h)] + o(h^4). \tag{1.125}$$

Теорема 1.4. (r=4, r=5). Нехай $f \in C^{5,5}(\mathbf{R}^1 \times \mathbf{R}^1)$, тоді

1) для r=4, r=5 і s=4 існує безліч формул Рунге-Кутта;
2) для r=4, r=5 і $s \geq 5$ не існують формули Рунге-Кутта.

Доведення. Для r=4 формула (1.92) прийме вигляд (аналогічно і для r=5):

$$y(x+h) \approx y(x) + \sum_{i=1}^{4} p_i \cdot k_i(h), \tag{1.126}$$

$$\varphi(h) = y(x+h) - \sum_{i=1}^{4} p_i \cdot k_i(h). \tag{1.127}$$

Обчислимо $\varphi'(h), \varphi''(h), \varphi'''(x), \varphi^{(5)}(h)$. Умова (1.100) при s=4 дає для $\forall f(x,y): \varphi'(0) = \varphi''(0) = \varphi'''(0) = \varphi^{(4)}(0) = 0$, для бо $\varphi^{(5)}(0) \neq 0$. Наприклад, $f(x,y) \equiv y$. Умова буде виконуватися, якщо невідомі параметри $\{p_1, p_2, p_3, p_4, \alpha_3, \alpha_4, \beta_{32}, \beta_{42}, \beta_{43}\} \subset \mathbf{R}^1$ знаходяться з нелінійної системи 11 рівнянь з 13-ма невідомими:

$$\left.\begin{array}{l} 1 - p_1 - p_2 - p_3 - p_4 = 0; \alpha_2 = \beta_{21}; \alpha_5 = \beta_{31} + \beta_{32}; \\ \alpha_4 = \beta_{41} + \beta_{42} + \beta_{43}; \mathrm{p}_2\, a_2 + \mathrm{p}_3\, a_3 + \mathrm{p}_4\, a_4 = \dfrac{1}{2}; \\ \mathrm{p}_2\, a_2^2 + \mathrm{p}_2\, a_3^2 + \mathrm{p}_4\, a_4^2 = \dfrac{1}{3}; \mathrm{p}_2\, a_2^3 + \mathrm{p}_3\, a_3^3 + \mathrm{p}_4\, a_4^3 = \dfrac{1}{4}; \\ p_3\beta_{32}\alpha_2 + p_4\beta_{42}\alpha_2\alpha_3 + p_4\beta_{43}\alpha_2\alpha_3 = \dfrac{1}{6}; \\ p_3\beta_{32}\alpha_2\alpha_3 + p_4\beta_{42}\alpha_2\alpha_4 + p_3\beta_{45}\alpha_3\alpha_4 = \dfrac{1}{8}; \\ p_3\beta_{32}\alpha_2^2 + p_4\beta_{42}\alpha_3^2 + p_4\beta_{43}\alpha_3^3 = \dfrac{1}{12}; \\ p_4\beta_{43}\alpha_3 = \dfrac{1}{12}; \\ p_4 > 0; \beta_{42} \neq 0; \beta_{32} \neq 0; \alpha_2 \neq 0. \end{array}\right\} \tag{1.128}$$

Теорема 1.4 доведена.

Наслідок 1.4. Один із розв'язків (1.128) $a_2 = a_3 = \beta_{32} = \beta_{21} = 0.5$; $\beta_{31} = \beta_{42} = \beta_{41} = 0$; $a_4 = \beta_{33} = 1$; $p_1 = p_4 = \dfrac{1}{6}$; $p_2 = p_3 = \dfrac{1}{3}$ дає формулу Рунге-Кутта (1.126) з точністю $\varphi(h) = o(h^5)$:

$$y(x+h) = y(x) + \frac{1}{6}[k_1(h) + 2k_2(h) + 2k_3(h) + k_4(h)] + o(h^{(5)}), \tag{1.129}$$

де для методу 4-го порядку похибка на кожному кроці становить $o(h^{(5)})$, а сумарна похибка на кінцевому інтервалі інтегрування є величиною $o(h^{(4)})$. В свою чергу, відповідні коефіцієнти $k_1(h)$, $k_2(h)$, $k_3(h)$, $k_4(h)$, враховуючи крок інтегрування h, знаходимо за такою системою рівнянь:

$$\left.\begin{aligned} k_1(h) &\equiv h\cdot f(x,y);\\ k_2(h) &\equiv h\cdot f(x+\frac{h}{2};y+\frac{k_1(h)}{2});\\ k_3(h) &\equiv h\cdot f(x+\frac{h}{2};y+\frac{k_2(h)}{2}).\\ k_4(h) &\equiv h\cdot f(x+h;y+k_3(h)); \end{aligned}\right\} \qquad (1.130)$$

Наслідок 1.5. Існують безліч формул Рунге-Кутта (1.92) для r=5, s=5. Доведення аналогічно теоремі 1.4, але вони мають точність, що і в випадку r=4.

Зауваження 1.4. На практиці відрізок $[x_0; x_0+X]$, де існує розв'язок задачі Коші (1.76), (1.77), ділять на n частин точками $\{x_0, x_1, x_2, ..., x_n\} \subset [x_0; x_0+X]$ з постійним кроком $h \equiv x_{i+1} - x_i; i = 0,1,...,n-1$ і на кожному елементарному відрізку застосовуємо (1.139), (1.130). Тоді маємо розрахункові формули Рунге-Кутта:

$$y_{i+1} = y_i + \frac{1}{6}[k_2^{(i)} + 2k_2^{(i)}(h) + 2k_3^{(i)}(h) + k_4^{(i)}(h)] + o(h^5), i = 0,1,...,n-1. \qquad (1.131)$$

Похибка формули (1.131) на всьому відрізку $[x_0; x_0+X]$ має порядок $o(h^4)$. Можна продовжити будувати аналогічно формули Рунге-Кутта (1.92) для $s \geq 6$, якщо r=6, але найбільш розповсюдженими формулами Рунге-Кутта є формули (1.131).

1.6. Різницеві методи розв'язування задачі Коші для звичайних диференціальних рівнянь (ЗДР)

1.6.1. Екстраполяційні та інтерполяційні різницеві формули

Однією з переваг методу Рунге-Кутта є те, що на кожному кроці треба обчислити в різних точках праву частину рівняння:

$$y'(x) = f(x,y), \qquad (1.132)$$

$$y(x_0) = y_0. \qquad (1.133)$$

Але для складних $f(x,y)$, це приводить до накопичення похибки обчислень та веде до витрат часу на комп'ютері.

Наприклад, для r=1; s=1 використаємо формулу Ейлера

$$y_{i+1} = y_i + hf(x_i, y_i), \qquad (1.134)$$

яка з точністю $o(h^2)$ дає таблицю розв'язку (1.132), (1.133). Або для r=4; s=4 маємо:

$$y_{i+1} = y_i + \frac{1}{6}(k_1^{(i)}(h) + 2k_2^{(2)} + 2k_3^{(i)}(h) + k_4^{(i)}), \qquad (1.135)$$

яка з точністю $o(h^5)$ теж дає таблицю розв'язку (1.132), (1.133).

Тому можна йти іншим шляхом: використати табличні значення розв'язку, що одержані за формулою (1.134) або за формулою (1.135)

Таблиця 1.4

Для обчислення наступного значення y_{m+1} і т.д.

y_{m-j}	y_{m-j+1}	...	y_{m-1}	y_m	y_{m+1}	...
x_{m-j}	x_{m-j+1}	...	x_{m-1}	x_m	x_{m+1}	...

Значення функції $y_{m-j},\ldots,y_m$ називають входом в таблицю, їх можна порахувати за однією з формул Рунге-Кутта.

Якщо знайшли значення $y_{m-j},\ldots,y_m$, то слід обчислення продовжити так:

$$\left.\begin{array}{c} f_{m-j} \equiv f(x_{m-j}, y_{m-j}); \\ f_{m-j+1} \equiv f(x_{m-j+1}, y_{m-j+1}); \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f_m \equiv f(x_m, y_m); \end{array}\right\}. \tag{1.136}$$

Тоді на відрізку $[x_{m-j}, x_m)$ функцію $\varphi(x) \equiv f(x, y(x))$ можна наблизити інтерполяційним многочленом Лагранжа. Маємо очевидну рівність:

$$y_{m+1} = y_{m-j+k} + \int_{x_{m-j+k}}^{x_{m+1}} \varphi(x)dx, \tag{1.137}$$

де k – параметр формули (1.137). Далі візьмемо:

$$\varphi(x) = L_{m,j}(x) + R_{m,j}(x), \tag{1.138}$$

де j – індекс многочлена Лагранжа

$$L_{m,j}(x) = \sum_{i=0}^{j} f_{m-i}\Phi_i(x), \tag{1.139}$$

де m – початковий індекс точок від x_m до x_{m-j};

$$\Phi_i(x) = \prod_{j\neq i} \frac{x - x_j}{x_i - x_j}, \tag{1.140}$$

$$R_{m,j}(x) = \frac{f^{(j+1)}(\xi)}{(j+1)!}(x - x_m)(x - x_{x-1})\ldots(x - x_{x-j}). \tag{1.141}$$

Підставимо (1.138) у (1.137):

$$y_{m+1} = y_{m-j+k} + \int_{x_{m-j+k}}^{x_{m+1}} L_{m,j}(x)dx - o(h^{j-i}). \tag{1.142}$$

Використовуючи (1.142) можна продовжити таблицю 1.4. значень шуканої функції $y(x)$ задачі Коші (1.132), (1.133) на один крок, тобто одержати y_{m+1}, тощо.

Підставимо (1.139) у (1.142), тоді матимемо

$$y_{m+1} = y_{m-j+k} + \sum_{i=0}^{j} f_{m-i}\Phi_i(x) + o(h^{j+1}), \tag{1.143}$$

$$\beta_i \equiv \int_{x_{m-j+k}}^{x_{m+1}} \Phi_i(x)dx = h\int_{-j+k}^{1} Q_i(t)dt, \tag{1.144}$$

де $Q_i(t)$ – многочлен степеня j в результаті заміни:

$$Q_i(t) \equiv \Phi_i(x_m + th); x - x_m = th. \tag{1.145}$$

Відзначимо, що $\{\beta_i\}$ можна порахувати та створити відповідні таблиці. Якщо вибрати різні k, а також використати різні інтерполяційні формули, то можна отримати різні формули (1.143).

Означення 1.2. Формули (1.143) називаються *екстраполяційними різницевими формулами* (вони отримані шляхом інтегрування інтерполяційного многочлена екстрапольованого (продовженого) на $[x_m, x_{m-1}]$).

Означення 1.3. Формули

$$y_{m+1} = y_{m-j+k} + \sum_{i=-1}^{j} f_{m-1}\beta_i + o(h^{j+2}). \tag{1.146}$$

назвемо *інтерполяційною різницевою формулою* (використовується значення f_{m+1}, яке невідоме, $f_{m-j+k} \equiv f(x_{m+1}, y_{m+1})$). Зауважимо, що інтерполяційна формула (1.146) складніша екстраполяційної формули, бо інтерполяційна є неявною формулою, а екстраполяційна є явною формулою. Дуже часто інтерполяційна формула (1.146) використовується для контролю точності.

1.6.2 Часткові випадки різницевих формул. Екстраполяційні формули Адамса

Використаємо інтерполяційний многочлен Ньютона назад, починаючи з точки x_m. Тоді, якщо перейти до нової змінної t, а саме: $x = x_m + t \cdot h$, він прийме вигляд:

$$L_{m,j(x)} = f_m + tf^1_{m-\frac{1}{2}} + \frac{t(t+1)}{2!}f^2_{m-1} + \ldots + \frac{t(t+1)\ldots(t+j-1)}{j!}f^j_{m-\frac{j}{2}}, \tag{1.147}$$

$$t = \frac{x - x_m}{h}. \tag{1.148}$$

Випадок 1.4. Беремо $k = j$ у формулі (1.142) та враховуючи (1.147), отримаємо відповідну різницеву формулу (1.149) розв'язування задачі Коші для звичайних диференційних рівнянь:

$$y_{m+1} = y_m + \int_{x_m}^{x_{m+1}} L_{m,j}(x)dx + o(h^{j+1}) =$$

$$= y_m + \left[h\int_0^1 \left[f_m + tf^1_{m-\frac{1}{2}} + \ldots + \frac{t(t+1)\ldots(t+j-1)}{j!} f^j_{m-\frac{j}{2}} \right] \right] dt + \tag{1.149}$$

$$+ o(h^{j+1}) \equiv y_m + h\sum_{i=0}^{j} a_i f^i_{m-\frac{i}{2}} + o(h^{j+1}).$$

де $a_0 = \int_0^1 dt = 1; a_1 = \int_0^1 tdt = \frac{1}{2}; a_2 = \int_0^1 \frac{t(t+1)}{2!}dt = \frac{5}{12}; a_3 = \frac{3}{8}; a_4 = \frac{251}{720}; a_5 = \frac{95}{288}; a_6 = \frac{19087}{60490}$; тощо.

Якщо $g \equiv h \cdot f$, то формула (1.149) прийме вигляд

$$y_{m+1} = y_m + g_m + \frac{1}{2}g^1_{m-\frac{1}{2}} + \frac{5}{12}g^3_{m-\frac{3}{2}} + \frac{251}{720}g^4_{m-2} + o(h^5), \tag{1.150}$$

яку назвемо *екстраполяційною формулою Адамса.*

Зауваження 1.5. Якщо f – гладка функція, то скінченні різниці f_i^3 майже постійні, тоді з формули (1.150) маємо конструкцію:

$$y_{m+1} = y_m + g_m + \frac{1}{2}g^1_{m-\frac{1}{2}} + \frac{5}{12}g^2_{m-1} + \frac{3}{8}g^3_{m-\frac{3}{2}} + o(h^4). \tag{1.151}$$

*Зауваження 1.*6. Іноді вигідно подати y_{m+1} через $y'_i \equiv f_i$. Для цього скінченні різниці запишемо через похідні:

$$\begin{aligned} f^1_{m-\frac{1}{2}} &= f_m - f_{m-1} = y'_m - y'_{m-1}; \\ f^2_{m-1} &= y'_m - 2y'_{m-1} + y'_{m-2}; \\ f^3_{m-\frac{3}{2}} &= y'_m - 3y'_{m-1} + 3y'_{m-2} - y'_{m-3}; \end{aligned} \tag{1.152}$$

Якщо у (1.151) обмежитися одним членом, тоді $y_{m+1} = y_m + h \cdot f_m + o(h^2) = y_m + hy'_m + o(h^2)$ – формула Ейлера. Якщо членів 4, то

$$y_{m+1} = y_m + \frac{h}{24}\left[55y'_m - 59y'_{m-1} + 37y'_{m-2} - 9y'_{m-3}\right] + o(h^5). \tag{1.153}$$

екстраполяційна формула Адамса.

Метод Адамса – група методів чисельного інтегрування звичайних диференційних рівнянь, які дозволяють обчислювати таблицю наближених значень розв'язку за даними в початкових точках. В однокрокових методах для обчислення значення y_{m+1} використовується значення тільки y_m і для підвищення точності при фіксованому кроці необхідно проводити обчислення великої

кількості допоміжних величин. Це є причиною того, що для багатьох задач застосування формул Рунге-Кутта неможливе внаслідок надто великого обсягу обчислень. Тому часто раціональніше переходити до багатокрокових методів, які дають можливість, використовуючи значення $f(x_i,y_i)$, обчислені на попередніх кроках, отримати прийнятну точність. Серед k-крокових методів найчастіше використовують методи інтегрування на сітці з постійним кроком, які називаються скінченно-різницевими схемами.

Випадок 1.5. Нехай $k=j$-1. Отримаємо екстраполяційну формулу Адамса. З формули (1.142) при $k=j$-1 маємо:

$$y_{m+1} = y_{m-j+j-1} + \int_{x_{m-j}}^{x_{m+1}} \left[f_m + xf^1_{m-\frac{1}{2}} + \frac{x(x+1)}{2!} f^2_{m-1} + \ldots + \frac{x(x+1)\ldots(x-j+1)}{j!} f^j_{m-\frac{j}{2}} \right] dx =$$

$$= y_{m-1} + \int_{-1}^{1} \left[g_m + tg^1_{m-\frac{1}{2}} + \ldots + \frac{t(t+1)\ldots(t-j+1)}{j!} g^j_{m-\frac{j}{2}} \right] dt \equiv \qquad (1.154)$$

$$\equiv y_{m-1} + \sum_{i=0}^{j} a_i g^2_{m-\frac{i}{2}} + o(h^{j+1})$$

де $a_0 = \int_{-1}^{1} dt = 2; a_1 = \int_{-1}^{1} tdt = 0; a_2 = \int_{-1}^{1} \frac{t(t+1)}{2!} dt = \frac{1}{3}; a_3 = \frac{1}{3}; a_4 = \frac{28}{90}; a_5 = \frac{78}{90}.$

Таким чином, шукана екстраполяційна формула Адамса для $k=j$-1 має вигляд:

$$y_{m+1} = y_{m-1} + 2g_m + \frac{1}{3}g^2_{m-1} + \frac{1}{3}g^3_{m-\frac{3}{2}} + \frac{29}{90}g^4_{m-2} + \frac{78}{90}g^5_{m-\frac{5}{2}} + \ldots + a_j g^j_{m-\frac{j}{2}} + o(h^{j+1}). \qquad (1.155)$$

Зауваження 1.7. Якщо обмежитися різницями четвертого порядку та врахувати, що $\frac{28}{90} \approx \frac{30}{90} = \frac{1}{3}$, то формула прийме вигляд:

$$y_{m+1} \approx y_{m-1} + 2g_m + \frac{1}{3}\left[g^2_{m-1} + g^3_{m-\frac{2}{3}} + g^4_{m-2} \right] + (h^5). \qquad (1.156)$$

1.6.3. Часткові випадки різницевих формул. Інтерполяційні формули Адамса

Інтерполяційні формули Адамса отримаємо з формули

$$y_{m+1} = y_m + \int_{x_{m-j}}^{x_{m+1}} L_{m,j}(x)dx + o(h^{j+1}), \qquad (1.157)$$

якщо у інтерполяційному многочлені $L_{m,j}(x)$ Ньютона взяти вузол x_{m+1} за початкову точку $x - x_{m+1} = ht$:

$$L_{m,j}(x) = f_{m+1} + tf^1_{m+\frac{1}{2}} + \frac{t(t+1)}{2!} f^2_m + \frac{t(t+1)(t+2)}{2!} f^3_{m-\frac{1}{2}} + \ldots + \\ + \frac{t(t+1)\ldots(t+j-1)}{j!} f^j_{m-\frac{j}{2}+1} \tag{1.158}$$

Якщо у (1.158) взяти $k=j$, то

$$y_{m+1} = y_{m-j+j-1} + \int_{x_{m-1}}^{x_{m+1}} \left[f_m + xf^1_{m-\frac{1}{2}} + \frac{x(x+1)}{2!} f^2_{m-1} + \ldots + \right. \\ \left. + \frac{x(x+1)\ldots(x-j+1)}{j!} f^j_{m-\frac{j}{2}} \right] dx =$$

$$= y_{m-1} + \int_{-1}^{1} \left[g_m + tg^1_{m-\frac{1}{2}} + \ldots + \frac{t(t+1)\ldots(t-j+1)}{j!} g^j_{m-\frac{j}{2}} \right] dt \equiv$$

$$\equiv y_{m-1} + \sum_{i=0}^{j} a_i g^2_{m-\frac{i}{2}} + o(h^{j+1})$$

де $a_0 = \int_{-1}^{0} dt = 1; a_1 = -\frac{1}{2}; a_2 = -\frac{1}{12}; a_3 = -\frac{1}{24}; a_5 = -\frac{9}{160}$; а, отже,

$$y_{m+1} = y_m + g_{m+1} - \frac{1}{2} g^1_{m+\frac{1}{2}} - \frac{1}{12} g^2_m - \frac{1}{24} g^3_{m-\frac{1}{2}} - \\ - \frac{17}{720} g^4_{m-1} - \frac{9}{160} g^5_{m-\frac{1}{2}} - \ldots \tag{1.159}$$

Формула (1.159) через похідні матиме вигляд:

$$y_{m+1} = y_m + \frac{h}{24}\left[9y'_{m+1} + 19y'_m - 5y'_{m-1} + y'_{m-2}\right] + o(h^5). \tag{1.160}$$

Аналогічно можна побудувати інтерполяційну формулу Адамса для $k=y$-1.

1.7. Методи скінчених різниць розв'язування крайових задач для диференціальних рівнянь

1.7.1. Загальна постановка задачі

Розглянемо найпростіші методи розв'язування крайових задач для звичайних ДР. Цей метод буде поширений на розв'язування рівнянь з частинними похідними. Нехай задана крайова задача для ДР 2-го порядку:

$$y''(x) + p(x)y'(x) + g(x)y(x) = f(x) \tag{1.161}$$

на відрізку $[a,b]$ з умовами на кінцях

$$\left.\begin{aligned}\alpha_0 y(a)+\alpha_1 y'(a)=A;\\ \beta_0 y(b)+\beta_1 y'(b)=B,\end{aligned}\right\}$$

$$\begin{aligned}&\alpha_i^2+\beta_i^2>0; i=0,1.\\ &p,q,f\in C^2([a,b])\\ &\exists y\in C^4([a,b])\end{aligned} \qquad (1.162)$$

$$\alpha_i^2+\beta_i^2>0; i=0,1.$$

Якщо $p,q,f\in C^2([a,b])$, то $p,q,f\in C^2([a,b])$. Як правило, задачу (1.161), (1.162) розв'язати аналітично неможливо. Чисельний метод розв'язування (1.161), (1.162) полягає в заміні $y'(x)$ та $y''(x)$ різницевими відношеннями. Ідея:

1. Розбиваємо відрізок $[a,b]$ на n рівних частин точками: $a=x_0<x_1<\ldots<x_{n-1}<x_n=b$, тобто $x_i=a+ih, i=0,1,\ldots,n$, $h=\dfrac{b-a}{n}.\{x_i\}\in[a,b]$ назвемо вузлами, які, як правило, вибираємо рівновіддаленими : $x_{i+1}-x_i=h$.
2. У кожному з вузлів $x_1,x_2,\ldots,x_{n-1}$ замінимо 1-у, 2-у похідну відповідними різницевими співвідношеннями у рівнянні (1.161) у цих точках:

$$y''(x_i)+p(x_i)y'(x_i)+q(x_i)y(x_i)=f(x_i). \qquad (1.163)$$

Одержимо лінійну алгебраїчну систему (n-1) рівнянь відносно (n+1)-го невідомого $y_i\equiv y(x_i)$. Зробити систему замкненою допомагають крайові умови, тобто одержати (n+1) рівняння з (n+1) невідомими $y_0,y_1,\ldots,y_n$. Якщо система СЛАР розв'язана, то її розв'язки дадуть таблицю наближених значень розв'язків крайової задачі у вузлах $\{x, i=0,1,\ldots,n\}$.

1.7.2. Основні різницеві співвідношення

Лема 1.1. Нехай $y\in C^2([a,b])$. Тоді

$$\frac{y_{i+1}-y_i}{h}=y_i'+o(h), y_i\equiv y(x_i). \qquad (1.164)$$

Доведення. Формула Тейлора з залишковим членом, який містить 2-у похідну:

$$y(x+h)=y(x)+\frac{y'(x)}{1!}h+\frac{y''(x+\Theta h)}{2}h, 0<\Theta<1. \qquad (1.165)$$

Розділимо (1.165) на $h>0$ і одержимо (1.164).

Лема 1.2. Нехай $y\in C^3([a,b])$, то

$$\frac{y_{i+1}-y_{i-1}}{2h}=y_i'+o(h^2). \qquad (1.166)$$

Доведення. Застосуємо формула Тейлора в околі точки x_i, із залишковим членом, який містить 3-ю похідну.

$$y(x_i+h)=y(x_i)+\frac{y'(x_i)}{1!}h+\frac{y''(x_i)}{2!}h^2+\frac{y'''(x_i+\Theta_1 h)}{3!}h^3,\ 0<\Theta_1<1. \tag{1.167}$$

$$y(x_i-h)=y(x_i)-\frac{y'(x_i)}{1!}h+\frac{y''(x_i)}{2!}h^2-\frac{y'''(x_i+\Theta_1 h)}{3!}h^3,\ 0<\Theta_2<1. \tag{1.168}$$

Якщо відняти $y(x_i+h)-y(x_i-h)$ та поділити на $2h$, то

$$\frac{y_{i+1}-y_{i-1}}{2h}=y_i'+\frac{h^2}{2\cdot 3!}\left[y'''(\tilde{x}_i)+y'''(\tilde{\tilde{x}}_i)\right], \tag{1.169}$$

що еквівалентне (1.166).

Лема 1.3. Нехай $y\in C^4([a,b])$. Тоді

$$\frac{y_{i+1}-2y_i+y_{i-1}}{h^2}=y_i''+o(h^2). \tag{1.170}$$

Доведення. Застосуємо формулу Тейлора з залишковим членом, який містить четверту похідну у околі точки x_i:

$$y_{i+1}=y_i+hy_i'+\frac{h^2}{2}y_i''+\frac{h^3}{3!}y_i'''+\frac{h^4}{4!}y^{(4)}(x_i+\Theta_i h), \tag{1.171}$$

$$y_{i-1}=y_i-hy_i'+\frac{h^2}{2}y_i''-\frac{h^3}{3!}y_i'''+\frac{h^4}{4!}y^{(4)}(x_i-\Theta_i h). \tag{1.172}$$

Якщо додати та поділити на h^2, то

$$\frac{y_{i+1}-2y_i+y_{i-1}}{h^2}=y_i''+\frac{h^2}{2\cdot 3!}\left[y^{(4)}(\tilde{x}_i)+y^{(4)}(\tilde{\tilde{x}}_i)\right], \tag{1.173}$$

що еквівалентне (1.170).

Лема 1.4. Якщо $y\in C^3([a,b])$, то

$$\frac{y_{i+2}-y_{i+1}+y_i}{h^2}=y_i''+o(h). \tag{1.174}$$

Доведення. Розкладемо в ряд Тейлора в околі точки x_i з кроком $2h$ функцію $y(x)$:

$$y_{i+1}=y_i+2hy_i'+\frac{4h^2}{2}y_i''+\frac{8h^3}{3!}y'''(x_i+\Theta_i 2h),0<\Theta_i<1. \tag{1.175}$$

Тоді (1.176) дає (1.174).

$$\frac{y_{i+2}-2y_{i+1}+y_i}{h^2}=\frac{1}{h^2}\left[\begin{array}{l} y_i+2hy_i'+2h^2y_i''+\frac{4}{3}h^3y'''(\tilde{x}_i)- \\ -2y_i-2hy_i'-\frac{2h^2}{2!}y_i''-\frac{2h^3}{3!}y'''(\tilde{\tilde{x}}_i)+y_i \end{array}\right]= \tag{1.176}$$

$$=\frac{1}{h^2}\left[h^2y_i''+\frac{8}{31}h^3y'''(\tilde{x}_i)-\frac{2h^3}{3!}y'''(\tilde{\tilde{x}}_i)\right]$$

1.7.3. Метод скінченних різниць розв'язування крайової задачі для диференціальних рівнянь

У рівнянні (1.161) в точці x_i замінимо другу і першу похідні з точністю $o(h^2)$, ввівши позначення

$$y_i\equiv y(x_i),\ p_i\equiv p(x_i),\ g_i\equiv g(x_i),\ f_i\equiv p(f_i):$$
$$\frac{y_{i+1}-2y_i+y_{i-1}}{h^2}+p_i\frac{y_{i+1}-2y_i}{2h}+g_iy_i=f_i. \tag{1.177}$$

Це СЛАР (n-1) рівнянь з (n+1) невідомими $y_0,y_1,...,y_n$. Щоб отримати єдиний розв'язок потрібно підключити крайові умови з точністю $o(h^2)$.

$$\left.\begin{array}{l} \alpha_0y_0+\alpha_1\frac{y_2-y_0}{2h}=A; \\ \beta_0y_n+\beta_1\frac{y_n-y_{n-2}}{2h}=B. \end{array}\right\}. \tag{1.178}$$

Таким чином, система (1.161), (1.162) апроксимується СЛАР (1.177), (1.178) з точністю $o(h^2)$. Виникають задачі:

I. Існування єдиного розв'язку системи (1.177), (1.178).

II. Оптимального розв'язування СЛАР (1.177), (1.178).

III. Співвідношення точного розв'язку $y^*(x_i)$ задачі (1.161), (1.162) і наближення $y_i\equiv y(x_i)$.

1.7.4. Метод прогонки розв'язування апроксимуючої СЛАР для крайової задачі диференціальних рівнянь з точністю $o(h)$

Запишемо ДР (1.161) в точці x_i, та замінимо другу і першу похідні з точністю $o(h)$ на різницеві відношення (1.164), (1.176):

$$\frac{y_{i+2}-2y_{i+1}+y_i}{h^2}+p_i\frac{y_{i+1}-y_i}{h}+g_iy_i=f_i, \tag{1.179}$$
$$i=0,1,2,...,n-2;$$

$$\alpha_0 y_0 + \alpha_1 \frac{y_1 - y_0}{h} = A, \tag{1.180}$$

$$\beta_0 y_n + \beta_1 \frac{y_n - y_{n-1}}{2h} = B. \tag{1.181}$$

Формулу (1.179) запишемо у зручному вигляді

$$y_{i+2} + m_i y_{i+1} + k_i y_i = k^2 f_i, \tag{1.182}$$

$$m_i \equiv -2 + hp_i; k_i \equiv 1 - hp_i + h^2 g_i, \ i = 0, 1, ..., n-2. \tag{1.183}$$

Знайдемо y_1 через y_2, для цього в (1.182) підставимо i=0:

$$y_2 + m_0 y_1 + k_0 y_0 = h^2 f_0. \tag{1.184}$$

З (1.180) знайдемо y_0:

$$y_0 = \frac{Ah - \alpha_1 y_1}{\alpha_0 h - \alpha} = -\frac{Ah}{\alpha_1 - \alpha_0 h} + \frac{\alpha_1}{\alpha_1 - \alpha_0 h} y_1. \tag{1.185}$$

Підставимо (1.185) у (1.184):

$$y_2 + m_0 y_1 + k_0 \left[-\frac{Ah}{\alpha_1 - \alpha_0 h} + \frac{\alpha_1}{\alpha_1 - \alpha_0 h} y_1 \right] = h^2 f_0. \tag{1.186}$$

$$\left(m_0 + \frac{k_0 \alpha_1}{\alpha_1 - \alpha_0 h} \right) y_1 = h^2 f_0 + \frac{k_0 Ah}{\alpha_1 - \alpha_0 h} - y_2.$$

$$y_1 = \frac{\alpha_1 - \alpha_0 h}{m_0(\alpha_1 - \alpha_0 h) + k_0 \alpha_1} \left[h^2 f_0 + \frac{k_0 Ah}{\alpha_1 - \alpha_0 h} - y_2 \right], \tag{1.187}$$

$$y_1 = c_0 (d_1 - y_2), \tag{1.188}$$

$$\left. \begin{aligned} c_0 &= \frac{\alpha_1 - \alpha_0 h}{m_0(\alpha_1 - \alpha_0 h) + k_0 \alpha_1}; \\ d_0 &= h^2 f_0 + \frac{k_0 Ah}{\alpha_1 - \alpha_0 h} - y_2. \end{aligned} \right\} \tag{1.189}$$

Далі систему (1.182) запишемо при i=1:

$$y_3 + m_1 y_2 + k_1 c_0 (\alpha_0 - y_2) = h^2 f_{,}. \tag{1.190}$$

що еквівалентне

$$y_2 = \frac{1}{m_0 - k_1 c_0} \left[h^2 f_1 - k_1 c_0 d_0 - y_3 \right]. \tag{1.191}$$

Остаточно

$$y_2 = c_1(d_1 - y_3), \tag{1.192}$$

$$c_1 \equiv \frac{1}{m_0 - k_1 c_0}; d_1 \equiv h^2 f_1 - k_1 c_0 \alpha_0 .. \tag{1.193}$$

Подамо y_{i-1} через y_{i+1}. Тут працює метод математичної індукції. Нехай знаходимо y_i за формулою:

$$y_i = c_{i-1}(d_{i-1} - y_{i+1}). \tag{1.194}$$

Підставимо (1.194) у (1.182):

$$y_{i+2} + m_i y_{i+1} + k_i c_{i-1}(d_{i-1} - y_{i+1}) = h^2 f_i; y_{i+1} = c_i(d_i - y_{i+2}), \tag{1.195}$$

$$c_i = \frac{1}{m_i - k_i c_i - 1}; d_i \equiv h^2 f_i - k_i c_{i-1} d_{i-1}; i = 0,1,...,n-2. \tag{1.196}$$

Таким чином, система (1.182) записується у вигляді (1.195) для i=0,1,2,…,n-2, де c_0, d_0 обчислюють за формулою (1.185), c_i, d_i - за (1.196). Одержали твердження.

Теорема 1.5. Нехай $p(x)$, $q(x)$, $f(x) \in \mathbf{C}^1([a,b])$ у (1.161), (1.162). Тоді апроксимуюча система для (1.179)-(1.181) з точністю $o(h)$ має єдиний розв'язок, який знаходиться методом прогонки: прямим і оберненим ходом (1.194) - (1.201).

Алгоритм розв'язання апроксимуючої системи (1.179) - (1.181) з точністю $o(h)$.

I. Прямий хід методу прогонки:

1) обчислюємо m_i, k_i за формулою (1.183), i = 0,1,..., n-2;

2) c_0, d_0 – за формулою (1.189), c_i, d_i – за формулою (1.196), i = 1,… , n-2.

II. Обернений хід.

3) На 1-му кроці оберненого ходу узгоджуємо числа c_{n-2}, d_{n-2} з крайовими умовами на правому кінці (1.181). Дійсно, для $i = n-2$ рівняння (1.181) і (1.195) приймуть вигляд

$$y_{n-1} = c_{n-2}(d_{n-2} - y_n), \tag{1.197}$$

$$\beta_0 y_n + \beta_1 \frac{y_n - y_{n-1}}{h} = B. \tag{1.198}$$

Підставимо y_{n-1} з (1.197) у (1.198):

$$y_n = \frac{B y_n + \beta_1 c_{n-2} d_{n-2}}{\beta_0 h + \beta_1 (1 + c_{n-2})}. \tag{1.199}$$

4) c_{n-2}, d_{n-2} знайдені у I. Далі з (1.195) маємо:

$$\begin{aligned} y_{n-1} &= c_{n-2}(d_{n-2} - y_n); \\ y_{n-2} &= c_{n-3}(d_{n-3} - y_{n-1}); \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_1 &= c_0(d_0 - y_2). \end{aligned} \tag{1.200}$$

5) y_0 знаходимо з (1.180):

$$y_0 = \frac{\alpha_1 y_1 - Ah}{\alpha_1 - \alpha_0 h}. \tag{1.201}$$

Розглянемо апроксимацію задачі (1.161), (1.162) з точністю $o(h)$.

1.7.5. Метод прогонки розв'язування апроксимуючої СЛАР з точністю $o(h^2)$

Теорема 1.6. Нехай $p(x)$, $q(x)$, $f(x) \in \mathbf{C}^1([a,b])$ у (1.161), (1.162). Тоді апроксимуюча система для (1.202), (1.203) має єдиний розв'язок, який знаходиться методом прогонки для апроксимуючої системи з точністю $o(h^2)$.

Доведення. З точністю $o(h^2)$ у (1.161), (1.162) підставимо різницеві відношення (1.166), (1.170):

$$\frac{y_{i+1} - 2y_i + y_{i-1}}{h^2} + p_i \frac{y_{i+1} - y_{i-1}}{2h} + g_i y_i = hf_i, \; i = 0,1,..., n\text{-}1. \tag{1.202}$$

$$y'(x_0) \approx \frac{y_2 - y_0}{2h}, \tag{1.203}$$

$$y'(x_n) \approx \frac{y_n - y_{n-2}}{2h}, \tag{1.204}$$

$$\left.\begin{aligned} \alpha_0 y_0 + \alpha_1 \frac{y_2 - y_0}{2h} &= A; \\ \beta_0 y_0 + \beta_1 \frac{y_n - y_{n-2}}{2h} &= B. \end{aligned}\right\}. \tag{1.205}$$

Стиковка потребує підключення до (1.202) двох рівнянь крайової умови (1.205).

1.9. Чисельні методи розв'язування крайових задач для диференціальних рівнянь в частинних похідних. Метод сіток розв'язування крайових задач для диференціальних рівнянь еліптичного типу

1.9.1. Метод розв'язування задачі Діріхле

Діріхле Петер Густав Лежан (13.02.1805–05.05.1859) – німецький математик. Спеціаліст з математичного аналізу, механіки, математичної фізики, теорії чисел.

Розглянемо диференціальне рівняння еліптичного типу

$$Lu \equiv a(x,y)u''_{xx}(x,u) + b(x,y)u''_{yy}(x,u) + c(x,u)u'_x(x,u) + \tag{1.206}$$

$$+d(x,u)u'_y(x,u) + g(x,u)u(x,u) = f(x,u).$$

де $\{a,b,c,d,g,f\} \subset \mathbf{C}(Q)$; $\{u \equiv u(x,y)\} \subset \mathbf{C}^{2,2}(Q)$, Q – однозв'язна область з межею ∂Q, причому

$$a,b \geq 0; g \leq 0, \tag{1.207}$$

з крайовою умовою

$$u\big|_{\partial Q} = \varphi(x,y). \tag{1.208}$$

де $\varphi(x,y) \in \mathbf{C}(\partial Q)$.

Метод сіток є одним з чисельних методів розв'язування крайової задачі (1.206), (1.208). Розглянемо зліченну кількість паралельних прямих

$$\left.\begin{aligned} x &= x_0 + ih, \quad i = 0,\pm1,\pm2,... \\ y &= y_0 + jl, \quad j = 0,\pm1,\pm2.... \end{aligned}\right\}. \tag{1.209}$$

Перетин прямих (1.209) назвемо вузлами, які належать $\overline{Q} \equiv Q \cup \partial Q$, і позначимо

$$(i,j) \equiv (x_i, y_j) = (x_0 + ih, y_0 + jl). \tag{1.210}$$

Означення 1.5. Два вузли $P_i \equiv P(x_i, y_i)$ і $P_j \equiv P(x_j, y_j)$ суміжні, якщо вони віддалені один від одного в напрямку осей Ox і Oy не більше, ніж на величину кроків h і l (очевидно, що $(i,j) \in \overline{Q}$):

$$\rho(P_i, P_j) \leq \sqrt{h^2 + l^2}. \tag{1.211}$$

Означення 1.5. Сукупність вузлів, у яких всі суміжні вузли належать $\overline{Q}$, назвемо внутрішньою і позначатимемо Q^*. Всі інші вузли називаються граничними і позначатимемо ∂Q^*. $L(u)$ розглядається у $\overline{Q}$. Це означає, що a,b,c,d,g,f визначені у вузлах (i, j). Розв'язок в $\overline{Q}$ аналітично, як правило знайти неможливо. Тому виникає задача знайти розв'язок у вузлах $\overline{Q^*}$. при цьому виникають три проблеми:

I) Заміна задачі (1.206), (1.208) в області $\overline{Q}$ на задачу в $\overline{Q^*}$ – це задача дискретизації.

II) Задачі в $\overline{Q}$ і в $\overline{Q^*}$ не еквівалентні, тобто розв'язки (1.206), (1.208) і розв'язки дискретизованої задачі не співпадають. Надалі розв'язки (1.206), (1.208) будемо позначати u_{ij}^* у вузлах (i, j), а нової дискретної задачі – u_{ij}. Будемо розглядати чисельні методи, для яких $u_{ij} \xrightarrow[h,l \to 0]{} u_{ij}^*$.

III) Проблеми розробки алгоритмів розв'язування дискретизованої задачі. Введемо позначення:

$$\begin{aligned} u_{ij} &\equiv u(x_0+ih, y_0+jl); & a_{ij} &\equiv a(x_0+ih, y_0+jl);\\ c_{ij} &\equiv c(x_0+ih, y_0+jl); & b_{ij} &\equiv b(x_0+ih, y_0+jl);\\ d_{ij} &\equiv d(x_0+ih, y_0+jl); & g_{ij} &\equiv g(x_0+ih, y_0+jl). \end{aligned} \tag{1.212}$$

Тоді в кожному вузлі $f_{ij} \equiv f(x_0+ih, y_0+jl), (i,j) \in Q^*$ записуємо рівняння (1.206).

$$a_{ij}(u''_{xx})_{ij} + b_{ij}(u''_{yy})_{ij} + c_{ij}(u'_x)_{ij} + d_{ij}(u'_y)_{ij} + g_{ij}u_{ij} = f_{ij}. \tag{1.213}$$

Лема 1.5. Нехай функції $\{a,b,c,d,g,f\} \in \mathbf{C}^{2,2}(\overline{Q})$. Тоді $\{u \equiv u(x,y)\} \in \mathbf{C}^{4,4}(\overline{Q})$ і мають місце леми 1.1-1.4:

$$\left.\begin{aligned} (u''_{xx})_{ij} &= \frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h^2} + o(h^2),\\ (u''_{yy})_{ij} &= \frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{l^2} + o(l^2),\\ (u'_x)_{ij} &= \frac{u_{i+1,j} - u_{i-1,j}}{2h} + o(h^2),\\ (u'_y)_{ij} &= \frac{u_{i+1,j} - u_{i-1,j}}{2l} + o(l^2), \end{aligned}\right\}. \tag{1.214}$$

Крайова задача (1.213) породжує апроксимуючу задачу з точністю $o(h^2+l^2)$

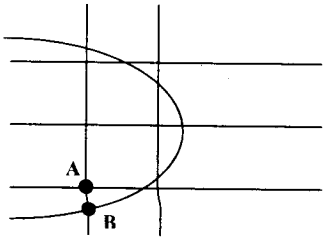
$$\begin{aligned} (lu)_{ij} \equiv a_{ij}\frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h^2} + b_{ij}\frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{l^2} +\\ + c_{ij}\frac{u_{i+1,j} - u_{i-1,j}}{2h} + d_{ij}\frac{u_{i+1,j} - u_{i-1,j}}{2l} + g_{ij}u_{ij} = f_{ij}, (i,j). \end{aligned} \tag{1.215}$$

Рівнянь стільки ж, скільки внутрішніх вузлів у Q^*. У (1.215) ввійдуть і граничні вузли (рис. 1.1а). Тому підключаємо граничні умови (1.208). Нехай $A \in \partial Q^*$, тоді

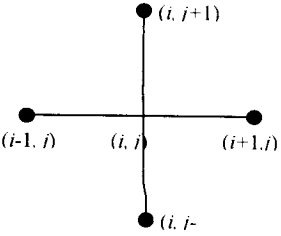
$$u\big|_A = \varphi_B, \ B \in \partial Q \tag{1.216}$$

Отже, система (1.215), (1.216) замкнена, тобто кількість рівнянь співпадає з кількістю невідомих, причому це система лінійних алгебраїчних рівнянь.

На рис. 1.1 подані приклади дискретизованих систем розв'язування крайових задач для диференціальних рівнянь в частинних похідних. Вони мають використовувати метод сіток розв'язування крайових задач для диференціальних рівнянь еліптичного типу.



а)



б)

Рис. 1.1. Приклади дискретизованих систем

1.9.2. Розв'язування рівняння Пуассона

Розглянемо рівняння Пуассона

$$\Delta u \equiv u''_{xx} + u''_{yy} = f\,, \tag{1.217}$$

за початковою функцією

$$u\Big|_{\partial Q} = \varphi\,. \tag{1.218}$$

Тоді з точністю $o(h^2 + l^2)$ маємо апроксимуючу СЛАР

$$\frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h^2} + \frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{l^2} = f_{ij}\,. \tag{1.219}$$

з похибкою $o(h^2 + l^2)$ методу сіток. Позначимо $\alpha \equiv \frac{h}{l}$, тоді

$$u_{i+1,j} + u_{i-1,j} + a^2 u_{i,j+1} + a^2 u_{i,j-1} - \left(2 + 2a^2\right) u_{ij} = h^2 f_{ij}\,. \tag{1.220}$$

якщо $h=l$, то

$$u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{ij} = h^2 f_{ij}\,. \tag{1.221}$$

за умови (рис. 1.1а)

$$u_A = \varphi_B\,. \tag{1.222}$$

Говорять, що дискретизована система (1.219) або (1.221) побудована за першим шаблоном (рис 1.1б).

Зауваження 1.8. Якщо $f \equiv 0$, то $u_{ij} = \frac{1}{4}(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1})$, то працює метод простої ітерації $u_{ij}^{(n+1)} = \frac{1}{4}\left(u_{i+1,j}^{(n)} + u_{i-1,j}^{(n)} + u_{i,j+1}^{(n)} + u_{i,j-1}^{(n)}\right)$. Наприклад, за методом сіток будувати дискретизовану систему за допомогою вузлів шаблону 2 (рис. 1.2а) чи шаблону 3 (рис. 1.2б) не можна.

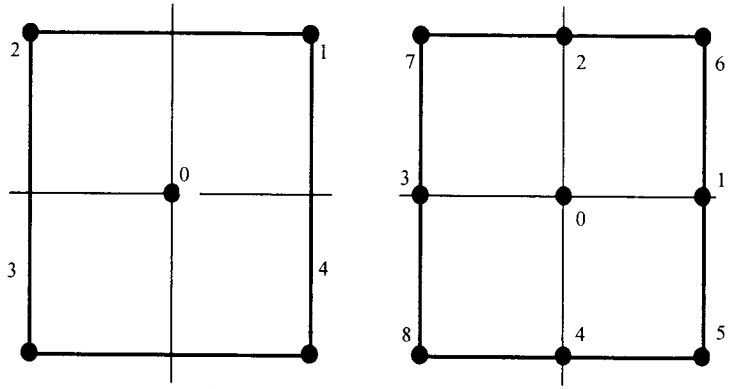


Рис. 1.2. Приклади дискретизованих систем

1.10. Метод невизначених коефіцієнтів розв'язування задачі Діріхле (метод Рейхенбаха)

Рейхенбах Ганс Фрідріх Герберт Гунтер (26.09.1891–09.04.1953) – німецький філософ і логік.

Позначимо $0 \equiv (i,j); 1 \equiv (i+1,j); \ldots; 8 \equiv (i-1, j-1)$ (рис. 1.2б). Розглянемо лінійну комбінацію

$$(lu)_0 \equiv \sum_{i=0}^{N} \sigma_i u_i , \tag{1.223}$$

де $\{\sigma_i\}$ – невизначені коефіцієнти, $\{u_i\}$ – значення $\{u \equiv u(x,y)\}$ у вузлах i=0,1,...,N.

Теорема 1.7. Нехай $u \in C^{(n+1,n+1)}(\overline{Q})$. Тоді з точністю $o(H^{n+1})$ маємо

$$(Lu)_0 = (lu)_0 + o(H^{n+1}). \tag{1.224}$$

$$H = \inf_{1 \le j \le N} \rho(P_0, P_j). \tag{1.225}$$

Доведення. Розкладемо за формулою Тейлора функцію $u(x, y)$ в околі вузла $0 \equiv (i, j)$, тоді (1.223) матиме вигляд

$$(lu)_0 \equiv \sum_{i=0}^{N} \sigma_i u_i = \sum_{i+k \le n} \gamma_{ik} \left(\left(\frac{\partial^{i+k} u}{\partial x^i \partial y^k} \right)_0 + R \right). \tag{1.226}$$

$$\gamma_{ik} \equiv L_{ik}(\sigma_0, \sigma_1, \ldots, \sigma_N) \tag{1.227}$$

лінійна функція коефіцієнтів $\{\sigma_j : 0 \le j \le N\}$. Похибка методу дискретизації $R = \theta H^{k+1} K M_{n+1}$, де $0 < \theta < 1$, K – константа, яка не залежить від H;

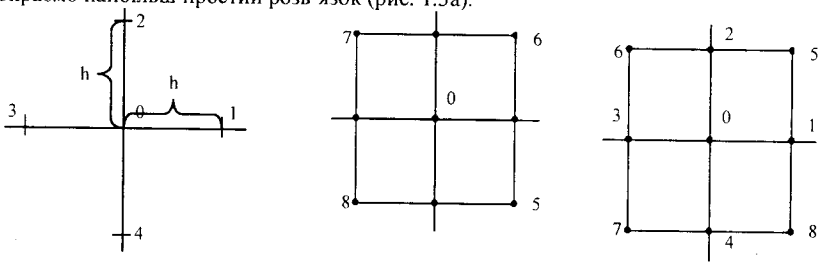
$$M_{n+1} = \max_{i+k\le n+1}\left\{\sup_{\overline{Q}}\left\{\left|\frac{\partial^{i+k}u}{\partial x^i \partial y^k}\right|\right\}\right\};\ H = \inf_{1\le j\le N}\rho(P_0, P_j)\cdot \tag{1.228}$$

Записуємо (1.226) у вигляді (1.224). Це означає, що коефіцієнти зліва у оператора Lu та відповідні коефіцієнти справа у оператора lu в точці 0 рівні, а решта коефіцієнтів $\gamma_{ik} = 0$:

$$\left.\begin{aligned} &a_0 = \gamma_{20};\ b_0 = \gamma_{02};\ 0 = \gamma_{11};\\ &c_0 = \gamma_{10};\ d_0 = \gamma_{01};\ g_0 = \gamma_{00};\\ &\gamma_{ik} = 0;\ 2 < i + k \le n. \end{aligned}\right\}. \tag{1.229}$$

Система залежить від $\sigma_0, \sigma_1, \ldots, \sigma_N$, бо $L_{ik}(\sigma_0, \sigma_1, \ldots, \sigma_N)$. Із системи (1.229) знаходимо $\sigma_0, \sigma_1, \ldots, \sigma_N$ за Рейхенбахом. Ідея полягає в тому, що від $(Lu)_0 = f_0$ переходимо до $\sum_{i=0}^{N}\sigma_i u_i = f_0$.

Зауваження 1.9. Якщо система (1.229) має декілька розв'язків або ∞, то вибираємо найбільш простий розв'язок (рис. 1.3а).



а) б) в)

Рис. 1.3. Приклади розв'язування дискретизованих систем

1.11. Часткові випадки методу Рейхенбаха

1.11.1. Перша різницева схема розв'язування рівняння Пуассона

Нехай задано рівняння Пуассона

$$\Delta u \equiv u''_{xx} + u''_{yy} = f(x, y). \tag{1.230}$$

Якщо $f \in C^2(\overline{Q})$, то $f \in C^4(\overline{Q})$. За Рейхенбахом маємо для $\sigma_1 = \sigma_2 = \sigma_3 = \sigma_4$ співвідношення

$$(\Delta u)_0 \approx \sum_{i=0}^{4}\sigma_i u_i = (lu_0). \tag{1.231}$$

Розкладемо u_1, u_2, u_3, u_4 в ряди

$$\begin{aligned} u_1 &\equiv u(x_0+h, y_0) = u_0 + \left[u'_x h + u''_{xx}\frac{h^2}{2!} + u'''_{xxx}\frac{h^3}{3!} + \dots \right]_0; \\ u_3 &\equiv u(x_0-h, y_0) = u_0 + \left[-u'_x h + u''_{xx}\frac{h^2}{2!} - u'''_{xxx}\frac{h^3}{3!} + \dots \right]_0; \\ u_2 &\equiv u(x_0, y_0+h) = u_0 + \left[u'_y h + u''_{yy}\frac{h^2}{2!} + u'''_{yyy}\frac{h^3}{3!} + \dots \right]_0; \\ u_4 &\equiv u(x_0, y_0-h) = u_0 + \left[-u'_y h + u''_{yy}\frac{h^2}{2!} - u'''_{yyy}\frac{h^3}{3!} + \dots \right]_0. \end{aligned} \tag{1.232}$$

Підставимо ці розклади у (1.231), в результаті одержимо

$$\begin{aligned} (lu)_0 &= \sigma_0 u_0 + \sigma_1 (u_1 + u_2 + u_3 + u_4) = (\sigma_0 + 4\sigma_1) u_0 + \\ &+ \sigma_1 h^2 (\Delta u)_0 + \sigma_1 \frac{h^4}{12} (\Delta(\Delta u))_0 + \dots \end{aligned} \tag{1.233}$$

Умова (1.231) буде виконуватись, якщо $h^2\sigma_1 = 1$. Звідси $\sigma_1 = \frac{1}{h^2}$, $\sigma_0 = -\frac{4}{h^2}$.

$$\begin{aligned} (lu)_0 &= -\frac{4}{h^2}u_0 + \frac{1}{h^2}(u_1 + u_2 + u_3 + u_4) + o(h^2) = \frac{1}{h^2}[u_1 + u_2 + u_3 + u_4 - 4u_0] + \\ &+ o(h^2) = (\Delta)u_0 + \frac{h^2}{12}(\Delta\Delta u)_0 + \dots = (\Delta u)_0 + R \end{aligned} \tag{1.234}$$

де $|R| \le \frac{h^2}{\sigma} M_4$. У (1.234) замінимо $(\Delta u)_0 = f_0$ і, відкидаючи R, маємо $u_1 + u_2 + u_3 + + u_4 - 4u_4 = h^2 f_0$. Звідки одержимо СЛАР для знаходження $\{u_{ij}\}$ з точністю $o(h^2)$

$$u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{ij} = f_{ij} h^2. \tag{1.235}$$

1.11.2. Друга різницева схема розв'язування рівняння Пуассона

Розглянемо, так звану, другу різницеву схему (рис. 1.3б), тобто $\sigma_2 = \sigma_5 = \sigma_6 = \sigma_7 = \sigma_8$. Тоді

$$(lu)_0 \equiv \sigma_0 u_0 + \sigma_2 (u_5 + u_6 + u_7 + u_8). \tag{1.236}$$

Розкладемо в ряди функцію $u(x,y)$ у відповідних вузлах за умови існування частинних похідних:

$$u_5 \equiv u(x_0+h, y_0+h) = u_0 + \left[h\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}\right)u + \frac{h^2}{2!}\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}\right)^2 u + \right.$$
$$\left. + \frac{h^3}{3!}\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}\right)^3 u + \frac{h^4}{4!}\left(\frac{\partial}{\partial x}+\frac{\partial}{\partial y}\right)^4 u + \ldots \right]_0; \quad (1.237)$$

$$u_6 \equiv u(x_0-h, y_0+h) = u_0 + \left[h\left(-\frac{\partial}{\partial x}-\frac{\partial}{\partial y}\right)u + \ldots \right]_0;$$

$$u_7 \equiv u(x_0-h, y_0-h) = u_0 + \left[h\left(-\frac{\partial}{\partial x}+\frac{\partial}{\partial y}\right)u + \ldots \right]_0;$$

$$u_8 \equiv u(x_0+h, y_0-h) = u_0 + \left[h\left(\frac{\partial}{\partial x}-\frac{\partial}{\partial y}\right)u + \ldots \right]_0.$$

Тоді різницевий оператор $(lu)_0$, визначений (1.236), прийме вигляд

$$(lu)_0 = (\sigma_0 + 4\sigma_2)u_0 + 2\sigma_2 h^2 (\Delta u)_0 + \sigma_2 \frac{h^4}{6}\left(u^{IV}_{x^4} + u^{IV}_{y^4} + 6u^{IV}_{x^2y^2}\right)_0. \quad (1.238)$$

Потрібно, щоб $(\Delta u)_0 \approx (lu)_0$. Значить $\sigma_0 + 4\sigma_2 = 0$; $2\sigma_2 h^2 = 1$ звідки

$$\sigma_0 = -\frac{2}{h^2};\ \sigma_2 = \frac{1}{2h^2}. \quad (1.239)$$

$$(lu)_0 = -\frac{2}{h^2}u_0 + \frac{1}{2h^2}(u_5+u_6+u_7+u_8) + R. \quad (1.240)$$

$$|R| \le \frac{1}{2h^2}\frac{h^4}{6}(M_4+M_4+6M_4) = \frac{2h^2}{3}M_4. \quad (1.241)$$

Отже, *другою різницевою схемою* назвемо СЛАР вигляду:

$$u_{i+1,j+1} + u_{i-1,j+1} + u_{i-1,j-1} + u_{i+1,j-1} - 4u_{ij} = 2h^2 f_{ij}. \quad (1.242)$$

1.11.3. Третя різницева схема розв'язування рівняння Пуассона

Схемою, що зображена на рис. 1.3в назвемо третьою різницевою схемою. У цьому випадку різницевий оператор запишеться так

$$(lu)_0 \equiv c_0 u_0 + c_1(u_1+u_2+u_3+u_4) +$$
$$+c_2(u_5+u_6+u_7+u_8) = (c_0+4c_1+4c_2)u_0) +$$
$$+h^2(c_1+2c_2)\times\Delta u_0 + \frac{h^2}{12}(c_1+2c_2)(\Delta(\Delta u))_0 + \quad (1.243)$$
$$+c_2 h^2 u^{m'}_{x^2y^2} + o(h^2).$$

Таким чином $(\Delta u)_0 = (lu)_0 + R(h)$. Звідки

$$\begin{cases} c_0 + 4c_1 + 4c_2 = 0, \\ h^2(c_1 + 2c_2) = 1. \end{cases} \tag{1.244}$$

Підберемо c_2 так , щоб члени з добутком 4-го порядку можна було б одержати шляхом диференціювання 2 рази оператора Лапласа. Це приводить до $6c_2h^2 = 1$. $\Delta u = u''_{xx} + u''_{yy}$. Значить

$$\Delta(\Delta u) = u^{IV}_{x^4} + u^{IV}_{y^4} + 2u^{IV}_{x^2y^2} \cdot \frac{h^2}{12}\left[h^2(c_1 + 2c_2)(\Delta(\Delta u)) + 12c_2h^2u^{IV}_{x^2y^2}\right]. \tag{1.245}$$

Звідки $12c_2h^2 = 2$, тобто $c_2 = \frac{1}{6h^2}$, а, отже, $c_0 = -\frac{10}{3h^2}, c_1\frac{2}{3h^2}$. Маємо *третю різницеву схему Пуассона*

$$\frac{4\sum_{i=1}^{4} u_i + \sum_{j=5}^{8} u_i - 2u_0}{6} \equiv h^2 f_0 + \frac{h^2}{12}(\Delta f)_0. \tag{1.246}$$

Отже, розглянули найбільш розповсюджені три схеми (1.235), (1.242), (1.246).

1.12. Апроксимація крайових умов методом Коллатца

Розглянемо апроксимацію крайових умов на випадок рис. 1.4. Позначимо: A – граничний вузол; C – внутрішній вузол; B – точка на межі ∂Q області Q.

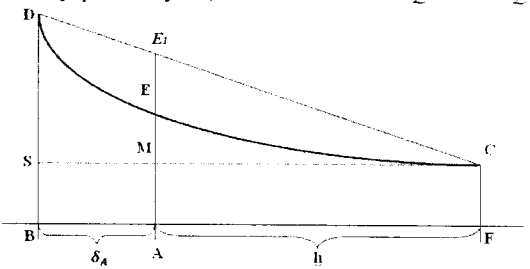


Рис. 1.4. Приклад апроксимації крайових точок

Зрозуміла геометрична інтерпретація відповідних значень розв'язків $u(x,y)$ крайової задачі: $u(B) = BD$; $u(A) = AE$; $u(C) = CF$; $\delta_A = BA < h = MC = AF$. Зрозуміле відношення з подібності прямокутних трикутників.

$$\frac{SD}{BF} = \frac{ME_1}{MC}. \tag{1.247}$$

Враховуючи, що $SB = u(B) - u(C)$; $BF = \delta_A + h$; $ME_1 = u(A) - u(C)$; $MC=AF=h$ відношення (1.247) прийме вигляд

$$\frac{u(B)-u(C)}{\delta_A + h} = \frac{u(A)-u(C)}{h}. \tag{1.248}$$

$$(h+\delta_A)u_A = (h+\delta_A)u(C) + h\left[u(B)-u(C)\right]. \tag{1.249}$$

$$u_A = \frac{\delta_A u(C) + h\varphi(B)}{h+\delta_A} + o(h), \tag{1.250}$$

яку назвемо *формулою Коллатца*. Зауважимо, що можна одержати аналогічну формулу з точністю $o(h^2)$.

1.13. Розв'язування різницевих рівнянь

1.13.1. Різницевий принцип максимуму

Розглянемо задачу Діріхле

$$Lu = f, \tag{1.251}$$

$$u\big|_{\partial Q} = \phi(x,y) \neq 0, \tag{1.252}$$

$$a > 0, b > 0, g < 0, \varphi \in C(\partial Q). \tag{1.253}$$

Відповідна апроксимуюча задача має вигляд:

$$(lu)_{ij} = A_{ij}u_{i+1,j} + B_{ij}u_{i,j+1} + C_{ij}u_{i-1,j} + D_{ij}u_{i,j-1} - G_{ij}u_{ij} = f_{ij}. \tag{1.254}$$

$$u\big|_A = \varphi_B, \tag{1.255}$$

де введено позначення

$$\begin{aligned} &A_{ij} \equiv \frac{a_{ij}}{h^2} + \frac{c_{ij}}{2h}; B_{ij} \equiv \frac{b_{ij}}{h^2} + \frac{d_{ij}}{2l}; C_{ij} \equiv \frac{a_{ij}}{h^2} - \frac{c_{ij}}{2h}; \\ &D_{ij} \equiv \frac{b_{ij}}{l^2} + \frac{d_{ij}}{2h}; G_{ij} \equiv -g_{ij} + \frac{2a_{ij}}{h^2} + \frac{2b_{ij}}{l^2}. \end{aligned} \tag{1.256}$$

Зауваження 1.10. Кроки $h>0$ і $l>0$ можна вибрати так, щоб $A_{ij} > 0, B_{ij} > 0, C_{ij} > 0, D_{ij} > 0, G_{ij} > 0$.

Лема 1.6. Зв'язок між коефіцієнтами (1.256) задається формулою

$$g_{ij} = A_{ij} + B_{ij} + C_{ij} + D_{ij} - G_{ij}. \tag{1.257}$$

Доведення. Складемо лінійну комбінацію $A_{ij}, B_{ij}, C_{ij}, D_{ij}, G_{ij}$ за формулами (1.256), що стоять у правій частині (1.257); в результаті одержимо g_{ij}.

Теорема 1.8 (дискретний принцип максимуму). Нехай на множині Q^* задана система чисел $\{u_{ij}\} \subset \mathbf{R}^1$. Якщо на внутрішніх вузлах Q^*

$$(lu)_{ij} \geq 0, \forall (i, y) \in Q^*. \tag{1.258}$$

то додатній максимум досягається на межі ∂Q^*. Якщо $(lu)_{ij} \leq 0, \forall (i, j) \in Q^*$ то від'ємний мінімум досягається на ∂Q^*. Якщо $u_{ij} = const$, тоді $(lu)_{ij} = 0$.

Доведення.

Випадок 1.6. Нехай $(lu)_{ij} \geq 0, \forall (i, j) \in Q^*$. Доведення проведемо від супротивного. Нехай знайдеться точка $(i_0, j_0) \in Q^*$ де досягається максимум $u_{i_0 \cdot j_0} \equiv M > 0$ Тоді хоча б у одному з найближчих вузлів значення u менше M, що дає можливість провести оцінку

$$\begin{aligned} 0 \leq (lu)_{i_0 \cdot j_0} = A_{i_0 \cdot j_0} u_{i_0+1 \cdot j_0} + B_{i_0 \cdot j_0} u_{i_0, j_0+1} + C_{i_0-1, j_0} u_{i_0 \cdot j_0} + D_{i_0 \cdot j_0} u_{i_0 \cdot j_0-1} - G_{i_0 \cdot j_0} M < \\ < (A_{i_0 \cdot j_0} + B_{i_0 \cdot j_0} + C_{i_0 \cdot j_0} + D_{i_0 \cdot j_0} - G_{i_0 \cdot j_0}) M = g_{i_0, j_0} M < 0, \end{aligned} \tag{1.259}$$

бо $g_{i_0 \cdot j_0} \leq 0$. Отже, за умовою $(lu)_{i_0 \cdot j_0} \geq 0$, а одержали $(lu)_{i_0 \cdot j_0} < 0$ що суперечить умові теореми 1.8.

Випадок 1.7. Нехай $(lu)_{i_0 \cdot j_0} \leq 0$. Доведення знову від супротивного. Нехай у точці $(i_0, j_0) \in Q^*$ досягається мінімум m<0. Тоді $0 \geq (lu)_{i_0 \cdot j_0} = g_{i_0 \cdot j_0} m > 0$ що означає $g_{i_0 j_0} < 0$. Отже, від'ємний мінімум m<0 досягається на межі ∂Q^*

Теорема 1.9 (існування і єдиність). Система лінійних алгебраїчних рівнянь $(lu)_{ij} = f_{ij}$. яка апроксимує $(lu)_{ij} = f_{ij}$ має єдиний розв'язок.

Доведення. Розглянемо однорідну систему

$$(lu)_{ij} = 0, \tag{1.260}$$

$$u\big|_A = 0, \tag{1.261}$$

тобто $f_{ij} = 0, \varphi_B = 0$. Система (1.260), (1.261) має єдиний розв'язок $u_{ij} = 0$. Дійсно, якщо $u_{ij} \neq 0$, то за теоремою 1.8 максимум (або мінімум) M>0 (або m<0) досягається на межі. Але $\varphi_B = 0$. що неможливо. Значить (1.260), (1.261) має розв'язок $u_{ij} = 0$. А це означає, що $|A| \neq 0$, ця система $(lu)_{ij} = f_{ij}$ має єдиний розв'язок $u_{ij} \neq 0$, так як $f_{ij} \neq 0$.

1.13.2. Розв'язування різницевої системи

Перепишемо систему $(lu)_{ij} = f_{ij}$ у вигляді

$$u_{ij} = \frac{A_{ij}}{G_{ij}} u_{i+1,j} + \frac{B_{ij}}{G_{ij}} u_{i,j+1} + \frac{C_{ij}}{G_{ij}} u_{i-1,j} + \frac{D_{ij}}{G_{ij}} u_{i,j+1} - \frac{f_{ij}}{G_{ij}}, \tag{1.262}$$

$$u_A = \frac{\delta_A u_C + h\varphi_B}{h + \delta_A}, \tag{1.263}$$

яку можна розв'язати методом простої ітерації (МПІ).

Теорема 1.10. Якщо $h>0$ і $l>0$ вибрані так, що $G_{ij} \neq 0$, то МПІ для (1.262), (1.263) збігається.

Доведення. Нехай $g<0$. Вибираємо $h>0$ і $l>0$ так, щоб $A_{ij}, A_{ij}, ..., G_{ij} > 0$. Тоді

$$u_{ij}^{(n+1)} = \frac{A_{ij}}{G_{ij}} u_{i+1,j}^{(n)} + \frac{B_{ij}}{G_{ij}} u_{i,j+1}^{(n)} + \frac{C_{ij}}{G_{ij}} u_{i-1,j}^{(n)} + \frac{D_{ij}}{G_{ij}} u_{i,j-1}^{(n)} - \frac{f_{ij}}{G_{ij}}. \tag{1.264}$$

$u_{ij}^{(n+1)} \underset{n\to\infty}{\to} u_{ij}$, так як сума коефіцієнтів <1.

Дійсно, так як $A_{ij} + B_{ij} + C_{ij} + D_{ij} - G_{ij} = g_{ij} < 0$, то $A_{ij} + B_{ij} + C_{ij} + D_{ij} < G_{ij}$, а значить сума коефіцієнтів <1. Для граничних умов (1.263) МПІ.

$$u_A^{(n+1)} = \frac{\delta_A}{h + \delta_A} u_C^{(n)} + \frac{h}{h + \delta_A} \varphi_B, \tag{1.265}$$

де $\frac{\delta_A}{h + \delta_A} < 1$. Швидкість збіжності визначається

$$q = \min_{(i,j)} \left\{ \frac{\delta_A}{\delta_A + h} + \frac{A_{ij} + B_{ij} + C_{ij} + D_{ij}}{G_{ij}} \right\} < 1. \tag{1.266}$$

1.14. Розв'язування рівнянь гіперболічного типу

1.14.1. Метод сіток розв'язування задачі Коші для рівнянь гіперболічного типу

Маємо рівняння

$$Lu \equiv au''_{xx} - bu''_{yy} + cu'_x + du'_y + gu = f, \tag{1.267}$$

де $a \cdot b > 0$, $a, b, c, d, g, f \in \mathbf{C}(\mathbf{R} \times G)$, $G = [y \geq 0]$, $u(x, y) \in G$, з крайовими умовами

$$u\big|_{y=0} = \varphi(x); u'_y\big|_{y=0} = \psi(x). \tag{1.268}$$

$$(i, j) : x = ih, y = jl; (i, j = 0, \pm 1, \pm 2, \pm 3, ...). \tag{1.269}$$

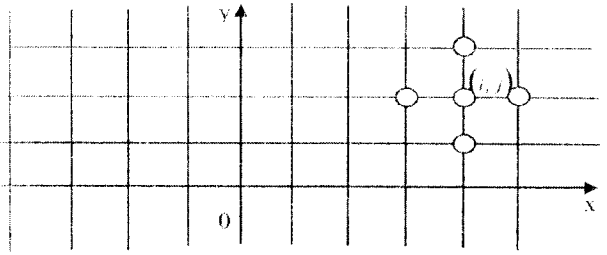


Рис. 1.5. Приклад методу сіток розв'язування задачі Коші для рівнянь гіперболічного типу

Введемо систему паралельних прямих (рис. 1.5), перетин яких дає вузли

Тоді значення шуканої функції $u(x,y)$ в цих вузлах позначимо $u_{ij} \equiv u(ih, jl)$.

Різницевий вираз з точністю $o(h^2 + l^2)$ матиме вигляд

$$(lu)_{ij} \equiv a_{ij}\frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h^2} - b_{ij}\frac{u_{i,j+1} - 2u_{ij} + u_{i,j-1}}{l^2} + c_{ij}\frac{u_{i+1,j} - u_{i-1,j}}{2h} + \\ + d_{ij}\frac{u_{i,j+1} - u_{i,j-1}}{2l} + g_{ij}u_{ij} = f_{ij}, \qquad (1.270)$$

$$(lu)_{ij} \equiv A_{ij}u_{i+1,j} - B_{ij}u_{i,j+1} + C_{ij}u_{i-1,j} + D_{ij}u_{i,j-1} + G_{ij}u_{ij} = f_{ij}, \qquad (1.271)$$

$$A_{ij} \equiv \frac{a_{ij}}{h^2} + \frac{c_{ij}}{2h}; B_{ij} \equiv \frac{b_{ij}}{h^2} - \frac{d_{ij}}{2l}; C_{ij} \equiv \frac{a_{ij}}{h^2} + \frac{c_{ij}}{2h}; \\ D_{ij} \equiv -\frac{b_{ij}}{l^2} - \frac{d_{ij}}{2h}; G_{ij} \equiv g_{ij} - \frac{2a_{ij}}{h^2} + \frac{2b_{ij}}{l^2}. \qquad (1.272)$$

Вибором $l>0$ можна добитися того, щоб $B_{ij} \neq 0$, тоді

$$u_{i,j+1} = \frac{1}{B_{ij}}[A_{ij}u_{i+1,j} + C_{ij}u_{i-1,j} + D_{ij}u_{i,j-1} + G_{ij}u_{ij} - f_{ij}]. \qquad (1.273)$$

З (1.243) випливає, що, знаючи розв'язки у вузлах $(i,j-1)$ та (i,j), можна знайти розв'язки у всіх вузлах $(i,j+1)$ горизонтального ряду. Схема (1.243) називається явною. При $j=0$ за першою умовою

$$u_{i,0} = \varphi(x_i). \qquad (1.274)$$

Розв'язок на першій горизонтальній можна знайти з умови (1.268)

1-й спосіб. Якщо $(u'_y)_{i,0} = \frac{u_{i,1} - u_{i,0}}{l} + o(l)$, то $\frac{u_{i,1} - u_{i,0}}{l} = \psi'(x_i)$. Звідки

$$u_{i,1} = \varphi_i + l\psi_i + o(l). \qquad (1.275)$$

2-й спосіб. Одержати $u_{i,1}$ з точністю $o(l^2)$. Нехай розв'язок (1.267) за умови (1.268) можна продовжити в області від'ємних $y<0$. Якщо

$$u'_y\Big|_{y=0} = \frac{u_{i,1} - u_{i,-1}}{2l} + o(l^2); u_{i,0} = \varphi_i,. \tag{1.276}$$

то $u_{i,1} = u_{i,-1} + 2l\psi(x_i)$. Запишемо рівняння (1.270) для j=0:

$$\begin{aligned} &A_{i,0}u_{i+1,0} - B_{i,0}u_{i,+1} + C_{i,0}u_{i-1,0} + D_{i,0}u_{i,-1} + G_{i,0}u_{i,0} = f_{i,0};\\ &u_{i+1,0} = \varphi_{i+1} = \varphi(x_{i+1}); u_{i-1,0} = \varphi_{i-1} = \varphi(x_{i-1});\\ &A_{i,0}\varphi(x_{i+1}) - B_{i,0}u_{i,1} + C_{i,0}\varphi(x_{i-1}) + D_{i,0}(u_{i,1} - 2l\psi(x_i)) + G_{i,0}\varphi(x_i) = f_{i,0}. \end{aligned} \tag{1.277}$$

$$u_{0,j} = \varphi_i; \tag{1.278}$$

$$u_{i,1} = \frac{1}{B_{i,0} - D_{i,0}}[A_{i,0}\varphi(x_{i+1}) + C_{i,0}\varphi(x_{i-1}) + 2lD_{i,0}\psi(x_i) + G_{i,0}\varphi(x_i) - f_{i,0}]. \tag{1.279}$$

Формули (1.279) і (1.274) дають $u_{i,0}, u_{i,1}$ для формули (1.273), за якими визначається $u_{i,2}, u_{i,3}$ тощо.

1.14.2. Вибір сітки для рівнянь гіперболічного типу

Нехай u_{ij} – розв'язок різницевої задачі, u^*_{ij} – відповідне значення розв'язку диференціальної задачі у відповідній точці (рис. 1.6).

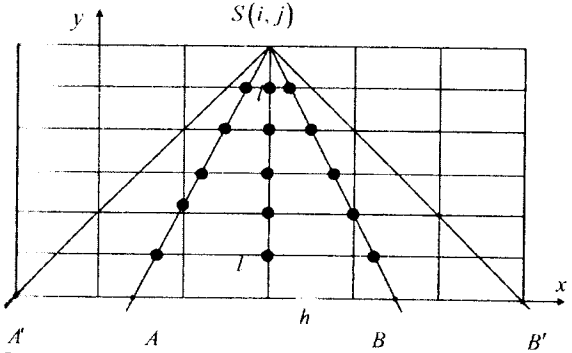


Рис. 1.6. Приклад вибору сітки для рівнянь гіперболічного типу

Потрібно, щоб

$$u_{ij} \underset{h,l\to 0}{\to} u^*_{ij}. \tag{1.280}$$

Розглянемо хвильове рівняння

$$u''_{xx} - u''_{yy} = 0, \tag{1.281}$$

за умови

$$u\Big|_{y=0} = \varphi(x); u'_y\Big|_{y=0} = \psi(x), -\infty < x + \infty. \tag{1.282}$$

Теорема 1.11. Необхідною умовою (1.280) збіжності методу сітки для задачі (1.281), (1.282) є умова $\left|\frac{l}{h}\right| = a \le 1$.

Доведення. Нехай $x_i = ih, y_i = jl, i, j = 0, \pm1, \pm2, \pm, 3, \ldots$ Зафіксуємо точку S. Доведемо, що $u_{ij} \to u_{ij}^*$. Розглянемо однорідне рівняння коливання

$$\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} - \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{l^2} = 0. \qquad (1.283)$$

Зрозуміло, що означення ΔSAB – є характеристичний трикутник системи. Розв'язок у точці $S(i,j)$ визначається значенням на AB, де $a = tg\varphi$ – це кутовий коефіцієнт сторони SB. Відомо, що розв'язок диференціальної задачі в точці S повністю визначається його значенням на основі характеристичного трикутника ΔSAB. Для коливного рівняння коливний коефіцієнт рівний ±1. Якщо $a>1$, то характеристичний трикутник різницевої системи знаходиться всередині характеристичного трикутника диференціального рівняння.

Нехай a>1. На відрізку $A'B'$ змінюємо початкову умову диференціальної задачі при $l,h \to 0$ так, щоб $a = const$. Умова на $A'B'$ будемо змінювати, залишаючи їх незмінними на AB (змінюючи їх тільки на AA', $B'B$). При цьому розв'язок диференціального рівняння в точці S можна як завгодно змінити, але розв'язок різницевої задачі змінюватися не буде. При такому розміщенні трикутників збіжності може не бути. Отже необхідною умовою збіжності є $a \to 1$. Для збіжності різницевої системи характеристичний трикутник диференціального рівняння повинен міститися у характеристичному трикутнику різницевої системи.

1.15. Метод сіток розв'язування лінійного рівняння параболічного типу (для розв'язування задачі Коші)

Розглянемо розв'язування задачі Коші та змішаних задач для лінійного диференціального рівняння параболічного типу вигляду

$$Lu = \frac{\partial u}{\partial t} - a\frac{\partial^2 u}{\partial x^2} - b\frac{\partial u}{\partial x} - cu = f \qquad (1.284)$$

де a,b,c,d – задані функції змінних x і t. і $a>0$.

Нехай потрібно знайти розв'язок $u(x,t)$ рівняння (1.284) у півплощині $t>0$, який задовольняє початковій умові

$$u(x,0) = \varphi(x), -\infty < x + \infty, \qquad (1.285)$$

де $\varphi(x)$ – задана функція.

Для знаходження наближеного розв'язку цієї задачі методом сіток розглянемо прямокутну сітку вузлів (рис. 1.7а). яка утворюється точками перетину двох сімей паралельних прямих $x = ih, i = 0, \pm1, \pm2, \ldots$, $t = jl$, $j = 1, 2, 3, \ldots$ Для кожного вузла

$(i, j), j \geq 1$ запишемо різницеве рівняння, яке апроксимує з деякою точністю рівняння (1.284).

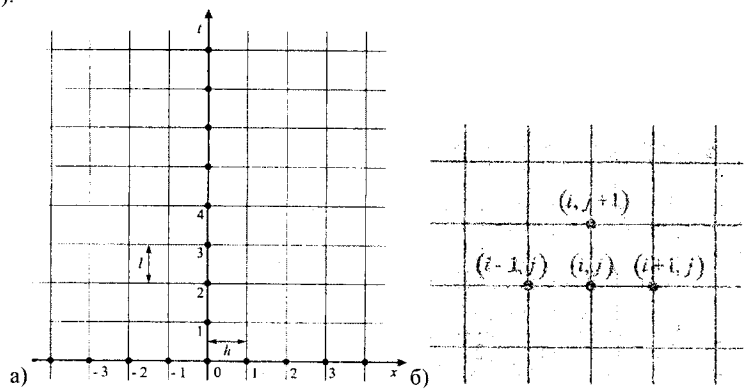


Рис. 1.7. Приклад розв'язання лінійного рівняння параболічного типу

Для цього замінимо похідні $\frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2}$ у вузлі (i,j) відповідно різницевими відношеннями

$$\frac{u_{i,j+1} - u_{ij}}{l}; \frac{u_{i,j} - u_{i,j-1}}{l}; \frac{u_{i,j+1} - u_{i,j-1}}{l}. \tag{1.286}$$

У відповідності до цих способів апроксимації похідних отримуємо три типи різницевих апроксимацій диференціального рівняння (1.284):

$$l^{(1)}u_{ij} = \frac{u_{i,j+1} - u_{ij}}{l} - a_{ij}\frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h^2} - \\ -b_{ij}\frac{u_{i+1,j} - u_{i-1,j}}{2h} - c_{ij}u_{ij} = f_{ij}, \tag{1.287}$$

$$l^{(2)}u_{ij} = \frac{u_{i,j} - u_{i,j-1}}{l} - a_{ij}\frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h^2} - \\ -b_{ij}\frac{u_{i+1,j} - u_{i-1,j}}{2h} - c_{ij}u_{ij} = f_{ij}, \tag{1.288}$$

$$l^{(3)}u_{ij} = \frac{u_{i,j+1} - u_{i,j-1}}{l} - a_{ij}\frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h^2} - \\ -b_{ij}\frac{u_{i+1,j} - u_{i-1,j}}{2h} - c_{ij}u_{ij} = f_{ij}. \tag{1.289}$$

Різницеве рівняння (1.287) містить значення розв`язку у чотирьох вузлах, зображених на рис. 1.7б. і апроксимує рівняння (1.284) з точністю до $o(l + h^2)$;

різницеве рівняння (1.288) містить значення розв'язку у чотирьох вузлах зображених на рис. 1.8а, і апроксимує рівняння (1.284) також з точністю до $o(l+h^2)$; до різницевого рівняння (1.289) входять значення розв'язку у п'яти вузлах (рис. 1.8б), і апроксимація рівняння (1.284) в цьому випадку буде $o(l^2+h^2)$. Для вузлів нульового горизонтального ряду $i=0$ з початковою умовою (1.285) маємо:

$$u_{i0}=\varphi(ih)=\varphi_i, i=0,\pm1,\pm2,.... \qquad (1.290)$$

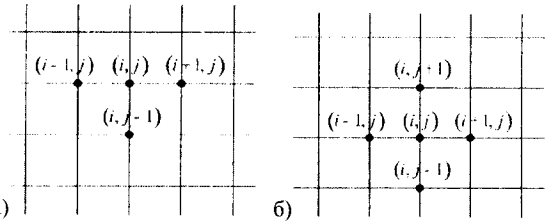


Рис. 1.8. Приклад розв'язання лінійного рівняння параболічного типу

1.16. Контрольні питання

1. Назвіть типи задач, для яких доведено існування та єдність розв'язків.
2. Який мінімальний порядок звичайних диференціальних рівнянь, для яких може бути сформульована гранична задача?
3. На які групи можна умовно поділити методи наближеного інтегрування диференціальних рівнянь?
4. В чому полягає чисельний розв'язок задачі Коші?
5. Графічна інтерпретація методу Ейлера.
6. Які методи називаються однокроковими?
7. Які методи називаються багатокроковими?
8. Яким шляхом отримані *екстраполяційні різницеві формули*?
9. Для чого використовується інтерполяційна різницева формула?
10. Опишіть прямий хід методу прогонки.
11. Опишіть обернений хід методу прогонки.
12. В чому суть методу метод Рейхенбаха?
13. Сформулюйте часткові випадки методу Рейхенбаха.
14. Опишіть методи розв'язування рівнянь гіперболічного типу.
15. Вибір сітки для рівнянь гіперболічного типу.
16. В чому суть різницевого принципу максимуму?

1.17. Задачі для самостійної роботи

Задача 1.17.1. Побудувати формулу знаходження $y(x+h)$ для розв'язування задачі Коші (1.76), (1.77), застосувавши формулу Сімпсона для обчислення інтеграла у правій частині рівності (1.78).

Задача 1.17.2. Одержати декілька нових формул Рунге-Кутта для r=2: s=1; 2, розв'язавши систему нелінійних рівнянь (1.114).

Задача 1.17.3. Одержати нові формули для r=3, s=3 шляхом знаходження параметрів $\{p_1, p_2, p_3, \alpha_2, \alpha_3, \beta_{21}, \beta_{31}, \beta_{32}\} \subset \mathbf{R}^1$ з системи нелінійних рівнянь (1.123).

Задача 1.17.4. Побудувати декілька нових пробних формул Рунге-Кутта з точністю $o(h^5)$ шляхом знаходження відповідних параметрів $p_i, \alpha_i, \beta_{ij}$ для r=4, s=4.

Задача 1.17.5. Виписати нелінійну систему для побудови формул Рунге-Кутта точності $o(h^6)$ (r=5, s=5;6).

Задача 1.17.6. Побудувати екстраполяційну формулу Адамса.

Задача 1.17.7. Записати одержану екстраполяційну формулу Адамса задачі 1.17.6 через похідні для $k=j-2$, $k=j-3,\ldots$, $k=j-m$.

Задача 1.17.8. Побудувати інтерполяційні формули Адамса для $L_{m,j}(x) = y_{m+1}, y_{m+2}, \ldots, y_{m-n}$.

Задача 1.17.9. Записати одержані інтерполяційні формули Адамса вправи 1.17.8 через похідні.

Задача 1.17.10. Подати екстраполяційну формулу Адамса (1.156) через похідні.

1.18. Завдання до лабораторної роботи

Дано задачу Коші (див. варіант завдання). На відрізку $[0;1]$ з кроком $h=0{,}1$ розв'язати методом Ейлера; методом Рунге-Кутта 4-го порядку.

№ варіанта	Задача Коші	№ варіанта	Задача Коші
1.	$y'=2x+3y, y(0)=4$	10.	$y'=5x+y, y(0)=5$
2.	$y'=2x+3y, y(0)=3$	11.	$y'=3x-2y, y(0)=6$
3.	$y'=4x-y, y(0)=2$	12.	$y'=3x+2y, y(0)=-2$
4.	$y'=3x-2y, y(0)=1$	13.	$y'=6x-y, y(0)=-1$
5.	$y'=5x-3y, y(0)=2$	14.	$y'=x+y, y(0)=0$
6.	$y'=4x+y, y(0)=3$	15.	$y'=7x+y, y(0)=3$
7.	$y'=6x-4y, y(0)=1$	16.	$y'=3x-4y, y(0)=2$
8.	$y'=2x-y, y(0)=5$	17.	$y'=2x+5y, y(0)=-3$
9.	$y'=4x-5y, y(0)=5$	18.	$y'=7x-6y, y(0)=1$

Розділ 2.
Чисельні методи розв'язування задач оптимізації

- Загальна постановка задачі оптимізації
  - Визначення цільової функції
  - Формальне визначення задачі оптимізації
  - Класифікація задач оптимізації
  - Метод переведення критеріїв в обмеження
  - Метод контрольних показників
  - Задача багатокритеріального математичного програмування
  - Деякі методи багатокритеріальної оптимізації
  - Метод квазіоптимізації локальних критеріїв (метод послідовних поступок)
  - Метод згортання векторного критерію в суперкритерій
- Безумовна оптимізація
  - Метод штрафних функцій
  - Геометрична інтерпретація штрафних функцій
  - Аналіз алгоритмів розв'язування задачі методом штрафних функцій
  - Обчислювальні складнощі, що пов'язані зі штрафними функціями
  - Аналіз та інтерпретація отриманих результатів
- Багатокритерійні задачі оптимізації та основні підходи до їх розв'язування
  - Основні поняття багатокритерійної задачі оптимізації
  - Геометрична ілюстрація багатокритерійної задачі
  - Постановка задачі багатокритерійної оптимізації
  - Еталонні точки
  - Критерії оптимальності. Критерій Парето
  - Лексикографічний порядок
  - Скаляризація
  - Зважена сума
  - Функція скаляризації Чебишева
  - Метод зміни обмежень (ε-обмеження)
  - Методи розв'язування. Інтерактивність. Еволюційні методи
- Лінійне програмування
  - Місце лінійного програмування в системі задач математичного програмування
  - Формальна постановка задачі лінійного програмування
  - Чисельні методи оптимізації. Методи прямого пошуку

- Лінійний пошук без використання похідних
- Інтервал невизначеності
- Функції однієї змінної
- Функції *n* змінних
- Методи прямого пошуку для функції однієї змінної
- Задача оптимізації методом Ньютона

- Порівняння методів лінійного пошуку без розрахунку похідної
  - Класифікація методів лінійного пошуку без розрахунку похідної
  - Рівномірний пошук
  - Послідовний пошук
  - Задача оптимізації методом дихотомічного пошуку (ділення пополам)
  - Задача оптимізації за методом золотого січення
  - Задача оптимізації за методом Фібоначчі
  - Методи апроксимації. Квадратична інтерполяція
- Аналіз методів та алгоритмів розв'язування задачі нелінійного програмування
  - Класифікація методів нелінійного програмування
  - Формальна постановка задачі з використанням методів прямого пошуку для оптимізації унімодальних функцій якості без обмежень
  - Визначення та основні властивості опуклих функцій
- Методи прямого пошуку для функцій *n* змінних
  - Метод покоординатного спуску
  - Метод Хука – Дживса
  - Метод Нелдера – Міда
  - Метод циклічного покоординатного спуску
  - Метод Розенброка
  - Метод спряжних градієнтів Флетчера та Рівса
- Контрольні питання
- Задачі для самостійної роботи
- Завдання до лабораторної роботи

В техніці, економіці та природничих науках, як і в інших областях, часто виникають задачі оптимізації складної сукупності обладнання, операцій, ланцюгів або процесів [1-93, 103-109, 112-114, 125-136, 293-302]. Потрібно мінімізувати або максимізувати деяку функцію, яка називається *цільовою* функцією. Вона може характеризувати ціну, вагу, загальну ефективність або щось подібне при визначених обмеженнях. Такі задачі оптимізації, що сформульовані математично, можуть бути об'єднані під загальною назвою *задачі нелінійного програмування*. Методи розв'язування таких задач лежать в основі нелінійного програмування. В найширшому розумінні задача нелінійного програмування полягає в пошуку екстремуму цільової функції при заданих обмеженнях у вигляді рівностей і (або) нерівностей. Обмеження можуть бути лінійними і (або) нелінійними. Однак, загально прийнятою є більш вузька постановка загальної задачі

нелінійного програмування, в якій виключаються з розгляду наступні часткові випадки [94-102, 110-111, 137-292, 303-320]:

а) змінні можуть приймати тільки цілочисельні значення (нелінійне цілочисельне програмування);

б) обмеження містять як параметр час (оптимальне управління, динамічна оптимізація).

Змінні задачі можуть бути конструктивними параметрами, показниками регулятора, показами вимірювальних приладів, тощо, тоді як цільова функція може являти собою вартість, вагу, річний прибуток тощо, а обмеження – технічні вимоги, умови роботи, пропускну здатність або фактори безпеки, притаманні цьому процесу. Як часткові випадки нелінійного програмування можуть розглядатись задачі лінійного та квадратичного програмування. Розгляд тільки математичної постановки задачі нелінійного програмування не дозволяє виявити всі фактори, від яких залежить оптимізація реального процесу. Доцільним є розгляд взаємозв'язку фізичної задачі та її математичної постановки [1-93, 103-109, 112-114, 125-136, 293-302].

2.1. Загальна постановка задачі оптимізації

2.1.1. Визначення цільової функції

Задача оптимізації – задача знаходження точки (точок) мінімуму, або декількох мінімумів заданої функції. У задачах математичного програмування з одним критерієм потрібно визначити значення цільовий функція, що відповідає, наприклад, мінімальним витратам або максимальному прибуткові. Однак, ми практично в будь-якій реальній ситуації можемо знайте кілька цілей, що суперечать одна одній. Покажемо, наскільки широкий діапазон проблем, що можуть бути адекватно сформульовані як багатокритеріальні, і які характеристики варто використовувати як критерії.

1. Планування очищення нафти:
   - min {витрати},
   - min {кількість імпортованої сирої нафти},
   - min {кількість сировини з високим вмістом сірки},
   - min {відхилення від заданого складу},
   - min {згоряння газів}.
2. Планування виробництва:
   - max {сумарний чистий доход},
   - max {мінімальний чистий доход за будь-який період},
   - min {кількість невиконаних замовлень},
   - min {понаднормовий час},
   - min {запаси готової продукції}.
3. Вибір портфеля цінних паперів:
   - max {дохід},
   - min {ризик},
   - max {дивіденди},
   - min {відхилення від бажаного рівня розмаїтості паперів}.
4. Складання кошторису капіталовкладень:
   - max {наявність засобів},
   - min {попит на капітальні вкладення},

- min {щорічні експлуатаційні витрати},
- max {інвестиції в проекти, пов'язані з охороною навколишнього середовища},
- max {інвестиції в проекти в заданому регіоні},
- max {інвестиції в проекти по заданій товарній спеціалізації}.

5. Управління лісовим господарством:
   - max {стійкий врожай деревини},
   - max {людино-дні відпочинку в лісі},
   - max {людино-дні полювання в лісі},
   - max {ареал поширення диких тварин},
   - max {число місяців випасу домашніх тварин},
   - min {перевищення бюджету}.
6. Управління спусками водоймищ:
   - max {вигоди від рекреації на водоймищі},
   - max {вигоди від регулювання стоку нижче водоймища},
   - max {кількість енергії, що виробляється в басейні річки},
   - min {недопостачання води на комунальні потреби в басейні річки},
   - max {прибуток від зрошення земель нижче водоймища}.
7. Формування ревізійної служби у фірмі:
   - max {дохід},
   - min {ріст чисельності персоналу служби},
   - min {зменшення чисельності персоналу служби},
   - min {надлишкові понаднормові},
   - min {недовикористання кваліфікації кадрів},
   - max {час, відведений на професійний ріст}.
8. Транспортування:
   - min {вартість},
   - min {середній час доставки вантажів пріоритетним клієнтам},
   - max {виробництво за певною технологією},
   - min {витрата палива}.
9. Виготовлення суміші для сосисок, копченої ковбаси і салямі:
   - min {вартість},
   - min {жирність},
   - min {відхилення від необхідного кольору},
   - max {протеїн},
   - min {відхилення від необхідного змісту вологи},
   - min {відхилення від необхідної пропорції свинини та яловичини}.

Таким чином, для ефективного вирішення кожної з цих задач необхідно в першу чергу побудувати багатокритеріальну математичну модель, яку потім потрібно оптимізувати, попередньо обравши метод, що найкрще підходить для цього.

Необхідність використання ПК пояснюється складністю тих задач, які розв'язуються. Для знаходження оптимального розв'язку використовують різні методи оптимізації, які в загальному випадку можна розділити на наступні класи: *аналітичні, числові, графічні, експериментальні* тощо. *Математичне програмування* включає в

себе *лінійне, нелінійне, динамічне, стохастичне* та ін. програмування. Як частковий випадок, розглядається тут нелінійне програмування, постановка задачі якого представлена нижче. Воно має справу з оптимізацією нелінійних функцій при лінійних і (або) нелінійних обмеженнях. Типовими областями його застосування є прогнозування, планування промислового виробництва, управління ресурсами, контроль якості продукції, проектування технологічних ліній (процесів), облік та планування капіталовкладень. Поки що не існує загального методу розв'язування нелінійних задач оптимізації. Нелінійне програмування при вирішенні задач включає в себе елементи експериментування, розвиток його до цих пір зводився до пропозиції часткових алгоритмів, їх програмування, перевірки результатів застосування цих алгоритмів в конкретних задачах, які представляють практичний інтерес, та побудова кращих алгоритмів на основі набутого досвіду.

Визначати стратегію поведінки і приймати рішення нам постійно потрібно в повсякденному житті. Ці питання стають як перед окремими людьми, так і цілими колективами, організаціями, підприємствами (вживатимо термін "складні системи"), які у своїй діяльності повинні брати до уваги різні обставини, що можуть вплинути на досягнення мети. Для прийняття більшості рішень на побутовому рівні науковий підхід не є вкрай необхідний, тут людині достатньо керуватися досвідом та здоровим глуздом. Але для прийняття відповідних рішень, особливо коли потрібно передбачити вплив багатьох чинників або ціна помилки може виявитися надто високою, наукове обгрунтування діяльності стає не тільки бажаним, але й необхідним.

Нас будуть цікавити математичні методи, які застосовуються для планування діяльності, прийняття рішень і оцінки ефективності функціонування складних систем. Хоча поодинокі випадки розробок і застосувань подібних методів зустрічалися в працях вчених минулого (наприклад, Й. Бернуллі, Л. Ейлера, Ж. Л. Лагранжа), проте лише у 30-40-х роках XX століття дослідження у галузі прийняття рішень оформилися в окремий науковий напрям, який отримав назву "дослідження операцій".

2.1.2. Формальне визначення задачі оптимізації

Нехай задано деяку множину X із n-вимірного евклідового простору і функцію $f(x)$, визначену на X. Необхідно знайти точки мінімуму значень функції $f(x)$ на X. Або:

$$f(x) \to \min,\ x \in X. \tag{2.1}$$

Тут $f(x)$ – цільова функція, X – допустима множина, кожна точка x цієї множини – допустима точка задачі. Також, задачу оптимізації можна сформулювати як пошук максимуму (максимумів) цільової функції:

$$f(x) \to \max,\ x \in X. \tag{2.2}$$

Ця задача еквівалентна попередній задачі мінімізації цільової функції зі знаком мінус, в тому сенсі, що множини їхніх розв'язків збігаються.

Розв'язки задачі можна розділити на дві множини:

- *глобальні* (глобального мінімуму), це такі допустимі точки x^*, в яких цільова функція має найменше значення на всій допустимій області:

$$f(x^*) \le f(x),\ \forall\ x \in X; \tag{2.3}$$

- *локальні* (локального мінімуму), це такі допустимі точки x^*, в яких цільова функція приймає найменше значення в деякому околі:

$$f(x^*) \le f(x), \forall\, x \in X \cap U_\varepsilon(x^*), \tag{2.4}$$

де $U_\varepsilon(x^*) = \{x \in R^n \mid \|x - x^*\| \le \varepsilon\}$ – куля радіусу ε з центром x^*.

2.1.3. Класифікація задач оптимізації

- *Задача безумовної оптимізації* – задача оптимізації, допустимою множиною якої є весь евклідів простір (R^n):

$$f(x) \to \min,\ x \in R^n. \tag{2.5}$$

Одним із методів розв'язування задач оптимізації є метод найшвидшого спуску.

- *Задача умовної оптимізації* – задача оптимізації, допустима множина X якої є власною підмножиною R^n. Способи розв'язування задач умовної оптимізації значною мірою залежать від того, як задається допустима множина.
- *Задача математичного програмування* (або задача з рівностями та нерівностями) – задача умовної оптимізації, допустима множина якої має вигляд:

$$X = \{x \in R^n \mid g_i(x),\ i = \overline{1,k};\ g_i(x) = 0,\ i = \overline{k+1,m};\ x \in P\}. \tag{2.6}$$

Формально задача математичного програмування записується так:

$$\begin{cases} f(x) \to \min, \\ g_i(x) \le 0,\ i = \overline{1,k}, \\ g_i(x) = 0,\ i = \overline{k+1,m}, \\ x \in P. \end{cases} \tag{2.7}$$

Обмеження g_i називаються функціональними обмеженнями, а $x \in P$ – прямим.

- *Задача опуклого програмування* – задача оптимізації, цільова функція та допустима множина якої – опуклі:

$$f(x) \to \min,\ x \in X. \tag{2.8}$$

де $f(x)$ – опукла функція, X – опукла множина.

- *Багатокритеріальна оптимізація або програмування* (англ. *Multi-objective optimization*) – це процес одночасної оптимізації двох або більше конфліктуючих цільових функцій в заданій області визначення. Задача багатокритеріальної оптимізації зустрічаються в багатьох галузях науки та техніки.

На практиці задачі, що не мають невизначеностей, скоріше є винятком, ніж правилом. Поряд з розглянутими вище існує ще один важливий вид невизначеності – *невизначеність мети*, яка виявляється у наявності декількох, в більшості випадків незбіжних аспектів оцінки якості того чи іншого розв'язку з множини допустимих. У формальному вигляді аспекти оцінки якості відображаються за допомогою множини критеріїв. Таким чином виникає багатокритерійна задача дослідження операцій,

загальний вигляд якої наступний: $Q_1(a,x) \Rightarrow Max$, $Q_2(a,x) \Rightarrow Max$, ..., $Q_n(a,x) \Rightarrow Max$, $x \in X$. Знайти розв'язок, який одночасно був би найкращим за всіма критеріями, неможливо, тому що в загальному випадку покращення значення одного з критеріїв приводить до його погіршення.

Приклади. *Купівля автомобілю.* Важливими критеріями в цьому випадку є: питомі витрати пального, максимальна швидкість, комфортність, прохідність, зовнішній вигляд, вартість, гарантія, наявність сервісного обслуговування, ціна, і звичайно, придбати авто, яке б водночас витрачало найменше пального, було б при цьому найдешевшим, найкомфортнішим та найшвидшим неможливо.

Вибір місця зупинки для мандрівників у горах. Важливими критеріями є: наявність джерела або річки, наявність палива, можливість купання, наявність комарів, близькість до населених пунктів, відсутність хижих звірів.

Одним з найбільш розповсюджених способів є приведення множини критеріїв до одного глобального та розв'язування класичної однокритерійної задачі. Однак застосування цього підходу має суттєві вади, і однією з них є те, що отриманий розв'язок для деяких специфічних задач може навіть не належати до множини Парето-оптимальних. Методи згортання критеріїв приводять первісну задачу до однокритерійної задачі такого вигляду: $Q(Q_1(a,x),...,Q_n(a,x)) \Rightarrow Max$, $x \in X$. Найбільш вживаними є:

1. *Лінійна згортка*:

$$Q = \sum_{i=1}^{n} (c_i \times Q_i(a,x)) \Rightarrow Max,\ x \in X,\ \sum_{i=1}^{n} c_i = 1,\ c_i > 0. \cdot \tag{2.9}$$

2. *Згортка нормованих критеріїв*:

$$Q = \sum_{i=1}^{n} (c_i \times \frac{Q_i(a,x) - Q_i^{\min}}{Q_i^{\max} - Q_i^{\min}}) \Rightarrow Max,\ x \in X,\ \sum_{i=1}^{n} c_i = 1,\ c_i > 0. \cdot \tag{2.10}$$

В цих методах c_i — вагові коефіцієнти критеріїв, які повинні відображати їх важливість, $Q_i^{\min}, Q_i^{\max}$ — мінімальне та максимальне значення i-го критерію.

Основною проблемою цих методів є проблема виявлення точних значень вагових коефіцієнтів – ця процедура в більшості випадків є суб'єктивною. Окрім того, коефіцієнти в методі лінійної згортки повинні бути розмірними величинами, тому що критерії в більшості випадків мають різну розмірність. З метою позбавлення від цього недоліку в згортці нормованих критеріїв окремі критерії спочатку нормуються (нормовані критерії є безрозмірними та змінюються в інтервалі від 0 до 1). Але внаслідок такого "вдосконалення" зважуються нормовані критерії, які не мають змістовного навантаження, і тому об'єктивне визначення вагових коефіцієнтів ще більше ускладнюється. Як відзначив один з фахівців в галузі ДО – довільність (що викликана багатокритерійністю) переноситься в іншу інстанцію (визначення значень вагових коефіцієнтів).

Інший підхід до розв'язування проблеми багатокритерійності – аксіоматичний, що полягає в формулюванні множини аксіом з наступним формальним виведенням виду функції корисності (глобального критерію), за допомогою якого й здійснюватиметься остаточний вибір. У цьому випадку виявляються всі обмеження, які

побічно накладаються при евристичному застосуванні того чи іншого методу. Виявляється, що лінійна згортка обгрунтована при достатньо жорстких умовах, які в багатьох випадках не виконуються.

Існують дещо більш обгрунтовані методи згортання, такі як метод ідеальної точки. *Метод ідеальної точки* базується на існуванні "ідеальної точки" для розв'язку задачі, у якій досягається екстремум всіх критеріїв (*принцип Джофріона*). Оскільки ідеальна точка, звичайно ж, не знаходиться серед допустимих, виникає проблема знаходження точки, що найменш віддалена від ідеальної, і належить до множини допустимих. Все було б добре, якщо б існувало єдине поняття "віддалі". Таким чином, для розв'язування задачі за допомогою методу "ідеальної точки" необхідно насамперед визначити її координати, і надалі визначити метрику, за допомогою якої можна було б виміряти віддаль до оптимальної точки. Для визначення координат "ідеальної точки" розв'язуємо n однокритерійних задач за кожним з критеріїв оптимізації $Q_i(a,x) \Rightarrow Max,\ x \in X,\ i = \overline{1,n}$. Сукупність оптимальних значень критеріїв кожної з однокритерійних задач Q_i^* і визначить координати ідеальної точки $Q^* = (Q_1^*,...,Q_n^*)$ в критерійному просторі. Якщо "ідеальна точка" належить до множини допустимих (що зустрічається вкрай рідко), то розв'язок отриманий. В іншому випадку визначаємо "віддаль" до ідеальної точки, вводячи метрику, і розв'язуємо однокритерійну задачу знаходження точки з числа допустимих, яка найменш віддалена від ідеальної. Задача має вигляд $Q(x) = \rho(Q(a,x) - Q^*) \Rightarrow Min,\ x \in X$. Якщо обрана Евклідова метрика, то критерій буде мати вигляд

$$Q(x) = \sqrt{\sum_{i=1}^{n} (Q_i(x) - Q_i^*)^2} \Rightarrow Min. \qquad (2.11)$$

2.1.4. Метод переведення критеріїв в обмеження

Метод полягає у виділенні головного критерію $Q_1(x)$, за яким проводитиметься оптимізація нормативних значень Q_i^N для кожного з критеріїв, що залишилися (значення критерію не може бути меншим за нормативне), та розв'язуванні отриманої таким чином однокритерійної задачі оптимізації

$$Q_1(x) \Rightarrow Max,\ Q_2(x) \geq Q_2^N,...,Q_n(x) \geq Q_n^N,\ x \in X. \qquad (2.12)$$

Основними проблемами при застосуванні цього методу є складнощі з визначенням головного критерію та нормативних значень для інших критеріїв. Якщо нормативні значення обрані недостатньо великими, то не всі резерви покращення їх значень будуть використані. Якщо ж ці значення будуть завеликими, то задача взагалі не буде мати розв'язків, оскільки область допустимих розв'язків виявиться порожньою.

2.1.5. Метод контрольних показників

Метод дозволяє позбутися деяких проблем, притаманих методу переведення критеріїв в обмеження. Система нормативів задається для всіх критеріїв, і критерій якості представляється у вигляді $Q(x) = \min_i \frac{Q_i}{Q_i^N} \Rightarrow Max$. Залишається проблема

обгрунтування значень нормативів і додається проблема знаходженння розв'язку мінімаксної задачі. Метод послідовних поступок є одним з найбільш змістовно обгрунтованих, і за умови відсутності суперечностей в перевагах особи, що приймає рішення, може дати добрий результат. Насамперед впорядкуємо критерії за важливістю в порядку її спадання: $Q_1 \succ Q_2 \succ \ldots \succ Q_n$. Після цього на кожному i-му кроці алгоритму розв'язується задача оптимізації за критерієм Q_i та призначається поступка ΔQ_i, на яку ми готові зменшити отримане оптимальне значення критерію Q_i^*, щоб покращити значення інших критеріїв, що менш важливі, ніж Q_i. Значення цих критеріїв розраховуються за відомими координатами оптимуму x^*. Призначення поступки означає введення на кожному кроці ще одного додаткового обмеження $Q_i \geq Q_i^* - \Delta Q_i$, і, таким чином, на $(i+1)$-му кроці розв'язуватиметься задача $Q_{i+1}(x) \Rightarrow Max$, $x \in X$, $Q_1(x) \geq Q_1^* - \Delta Q_1, \ldots, Q_i(x) \geq Q_i^* - \Delta Q_i$. Процес розв'язування закінчується у випадку, коли або досягнуто останнього критерію, або ж призначення поступки недоцільне. При необхідності процес можна повторити, здійснивши аналіз попередніх результатів. Таким чином метод послідовних поступок є достатньо гнучким і дозволяє уникнути багатьох проблем, що властиві іншим методам. Для його реалізації достатньо взяти за основу метод розв'язування однокритерійної задачі.

2.1.6. Задача багатокритеріального математичного програмування

Задача багатокритеріального математичного програмування має вигляд:

$$\begin{aligned} &\max\{q_1(x) = Q_1\} \\ &\max\{q_2(x) = Q_2\} \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &\max\{q_k(x) = Q_k\} \\ &x \in X \end{aligned}, \qquad (2.13)$$

де X – множина допустимих значень змінних x; k – кількість цільових функцій (критеріїв); Q_i – значення i-го критерію (цільової функції), max – означає, що цей критерій потрібно максимізувати. Зауважимо, що, власне кажучи, багатокритеріальна задача відрізняється від звичайної задачі оптимізації тільки наявністю декількох цільових функцій замість однієї. При наявності в багатокритеріалній задачі критеріїв з різною розмірністю з метою усунення цієї проблеми використовують *нормалізацію критеріїв*. Способи нормалізації представлені в таблиці 2.1.

Таблиця 2.1

Способи нормалізації

Нормалізація	Математичний вираз
Зведення до безрозмірних величин	$q(x\|y)/\rho(q(x\|y))$
Приведення до однієї розмірності $\rho[\varphi]$	$q(x\|y)/\alpha(y)$ $\rho[q(x\|y)] = \rho[\alpha(y)] \times \rho[\varphi]$

Нормалізація	Математичний вираз
Зміна інгрідієнту	$-q(x\mid y)$ $1/q(x\mid y)$
Природна нормалізація $\lambda_i(x\mid y)$	$\dfrac{q(x\mid y)-\min\limits_{x\in X}(q(x\mid y))}{\max\limits_{x\in X}(q(x\mid y))-\min\limits_{x\in X}(q(x\mid y))}$
Нормалізація порівняння	$\max\limits_{x\in X}(q(x\mid y))-q(x\mid y)$
Нормалізація Савіджа	$q(x\mid y)/\max\limits_{x\in X}(q(x\mid y))$
Нормалізація усереднення	$q(x\mid y)/\sum\limits_{y\in Y}q(x\mid y)$

У цій таблиці y – елемент простору G. G – простір елементів довільної природи, названих *цільовими термами* (у конкретних інтерпретаціях це сукупність, перелік або нумерація якісних властивостей) елементів $x\in X$.

Згорткою компонент багатоцільового показника q_i називається відображення $g_i\{Q\to R_1\}$, що перетворить сукупність компонентів багатоцільового показника q, що відповідають цільовим термам $y\in Y$, у скалярний цільовий показник $g(q(x\mid y))=g\left[\{q(x\mid y)\}y\in Y\right]\in R_1$. Основними видами згорток є лінійні, мінімізаційні, максимізаційні добутки і функції Кобба-Дугласа вигляду:

$$
\begin{aligned}
g(q(x)) &= \sum_{y\in Y}\alpha(y)q(x\mid y)\\
g(q(x)) &= \min_{y\in Y}\left[\alpha(y)q(x\mid y)+\beta(y)\right]\\
g(q(x)) &= \max_{y\in Y}\left[\alpha(y)q(x\mid y)+\beta(y)\right]\\
g(q(x)) &= \prod_{y\in Y}\alpha(y)q(x\mid y)\\
g(q(x)) &= \prod_{y\in Y}\left[\alpha(y)q(x\mid y)\right]^{\beta(y)}
\end{aligned}
\qquad (2.14)
$$

Проблеми одержання та обгрунтування вибору згорток є основним напрямом теорії корисності. До теперішнього часу сформульовані основні принципи вибору, що наведені в таблиці 2.2.

Таблиця 2.2.

Принципи вибору

Принцип вибору	Умова оптимальності
Домінантності R^{dom}	$q(x^0\mid y)\geq q(x\mid y)\forall x\in X$
Часткової домінантності	$\exists Y_1\subseteq Y: q(x^0\mid y)\geq q(x\mid y)\forall x\in X, y\in Y_1$
Парето $\bar{R}^{dom}$	$\exists x\in X:\left\{q(x\mid y)\geq q(x^0\mid y)\forall y\in Y\right\}\cap$ $\cap\left\{\exists y^0\in Y: q(x^0\mid y)>q(x^0\mid y^0)\right\}$

Принцип вибору	Умова оптимальності
Слейтера	$\exists x \in X : q(x \mid y) \geq q\left(x^0 \mid y\right) \forall y \in Y$
Часткової ефективності Джеррі	$\exists x \in X : q(x \mid y) \geq q\left(x^0 \mid y\right) \forall y \in Y$ $\exists M > 0 : \dfrac{q(x \mid y) - q\left(x^0 \mid y\right)}{q\left(x^0 \mid y\right) - q(x \mid y')} \leq M$ $\forall y \in Y, \forall x \in \left\{x \in X : q(x \mid y) > q\left(x^0 \mid y\right)\right\}$ $\forall y' \in \left\{y \in Y : q(x \mid y) < q\left(x^0 \mid y\right)\right\}$
Невласної ефективності Джеррі	$\exists x \in X : q(x \mid y) \geq q\left(x^0 \mid y\right) \forall y \in Y$ $\forall M > 0, \exists y'' \in Y, \exists x \in \left\{x \in X : q(x \mid y) > q\left(x^0 \mid y''\right)\right\}$ $\dfrac{q(x \mid y'') - q\left(x^0 \mid y''\right)}{q\left(x^0 \mid y'\right) - q(x \mid y')} < M$ $\forall y' \in \left\{y \in Y : q(x \mid y') < q\left(x^0 \mid y'\right)\right\}$
Рівності	$q\left(x^0 \mid y\right) \geq q\left(x^0 \mid y''\right) \forall y, y'' \in Y$
Сумарної ефективності	$\sum_{y \in Y} q\left(x^0 \mid y\right) \geq \sum_{y \in Y} q(x \mid y) \forall x \in X$
Неша	$\prod_{y \in Y} q\left(x^0 \mid y\right) \geq \prod_{y \in Y} q(x \mid y) \forall x \in X$
Компромісу	$\exists \alpha(y) \in A : \sum_{y \in Y} \alpha(y) q\left(x^0 \mid y\right) = \max_{x \in X} \sum_{y \in Y} \alpha(y) q(x \mid y)$
Домінуючого результату	$\max_{y \in Y} q\left(x^0 \mid y\right) \geq \max_{y \in Y} q(x \mid y) \forall x \in X$
Гарантованого результату	$\min_{y \in Y} q\left(x^0 \mid y\right) \geq \min_{y \in Y} q(x \mid y) \forall x \in X$
Найменшого відхилення	$\left\|q\left(x^0 - q^*\right)\right\| \leq \left\|q\left(x - q^*\right)\right\| \forall x \in X$ $q^* = \left\{q^*(y)\right\}_{y \in Y}$ $q^*(y) = \max_{x \in X} q(x \mid y) \forall y \in Y$
Лямбда-критерій	$x^0 \in X_{\lambda^0} : \lambda^0 = \max\left(\lambda \mid X_\lambda \neq \Lambda\right)$ $X_\lambda = \left\{x \in X : \lambda_q(x \mid \lambda) \geq \lambda, \forall y \in Y\right\}$
Альфа-критерій Гурвіца $R_\lambda^\alpha(q)$	$x_0 \in X : \left[\alpha \min_{y \in Y} \lambda_q\left(x^0 \mid y\right) + (1-\alpha) \max_{y \in Y} \lambda_q\left(x^0 \mid y\right)\right] \geq$ $\geq \left[\alpha \min_{y \in Y} \lambda_q(x \mid y) + (1-\alpha) \max_{y \in Y} \lambda_q(x \mid y)\right] \forall x \in X$
Максимум функції невизначеності	$x_0 \in X : H\left[q\left(x^0\right)\right] \geq H\left[q(x)\right] \forall x \in X$

У задачах вибору розв'язку, які формалізовані у вигляді моделі векторної оптимізації, першим природним кроком варто вважати виділення області компромісів (або розв'язків, оптимальних по Парето). Вектор називається оптимальним по Парето розв'язком, якщо не існує $x \in X$ такого, що виконуються нерівності

$$q(x \mid y) \geq q\left(x^0 \mid y\right) \forall y \in Y \cdot \tag{2.15}$$

$$q(x \mid y) \neq q\left(x^0 \mid y\right) \cdot \tag{2.16}$$

Областю компромісів G_x називається підмножина допустимої множини розв'язків X, яка володіє властивістю, що всі розв'язки, що належать їй не можуть бути покращені одночасно по всіх локальних критеріях — компонентах вектора ефективності. Отже, для будь-яких двох розв'язків, що належать області $G_x\left(x', x'' \in G_x\right)$, обов'язково має місце протиріччя хоча б з одним з локальних критеріїв. Це автоматично приводить до необхідності проводити вибір розв'язку в G_x на основі деякої схеми компромісу, що і є причиною назви цієї підмножини – *областю компромісів*. Оптимальний розв'язок, який обраний на основі багатокритеріального підходу незалежно від принципу оптимальності, що обирається, завжди має належати області компромісів. Інакше він може бути покращений і, отже, не є оптимальним. Таким чином, область компромісів є областю потенційно оптимальних компромісів. Звідси випливає, що при виборі розв'язку за векторним критерієм ефективності можна обмежити пошук оптимального розв'язку областю компромісів G_x, що, як правило, значно вужче всієї області можливих розв'язків X.

2.1.7. Деякі методи багатокритеріальної оптимізації

Принцип справедливого компромісу. Нехай усі локальні критерії, що утворять вектор ефективності, мають однакову важливість.

Справедливим будемо вважати такий компроміс, при якому відносний рівень зниження якості по одному або декількох критеріях не перевершує відносного рівня підвищення якості за іншими критеріями (менше або дорівнює).

Цьому принципові можна дати наступну математичну інтерпретацію. Нехай в області компромісів G_x дано два розв'язки x' і x'', якість яких оцінюється критеріями $Q_1(x)$ і $Q_2(x)$. Розв'язок x' перевершує розв'язок x'' за критерієм Q_1, але поступається йому за критерієм Q_2. Необхідно порівняти ці розв'язки й вибрати найкращі на основі принципу справедливого компромісу. Для порівняння цих розв'язків на основі принципу справедливого компромісу введемо міру відносного зниження якості розв'язку по кожному з критеріїв – ціну поступки x:

$$W_1 = \frac{\Delta Q_1\left(x', x''\right)}{\max\limits_{x', x''} Q_1(x)} \cdot \qquad W_2 = \frac{\Delta Q_2\left(x', x''\right)}{\max\limits_{x', x''} Q_2(x)} \tag{2.17}$$

де ΔQ_1 і ΔQ_2 — абсолютні зниження рівня критеріїв при переході від розв'язку x' до розв'язку x'' (для критерію Q_1) і при зворотному переході (для критерію Q_2).

Якщо відносне зниження критерію Q_1 більше, ніж критерію Q_2, то варто віддати перевагу розв'язку x'. Це випливає з порівняння ціни поступки за кожним критерієм. Алгоритм розв'язку задачі векторної оптимізації, що базується на принципі справедливого компромісу, включає наступні кроки.

Крок 0. Вибираємо x' і $x'' \in D_x$.

Крок 1. Обчислюємо за формулами (2.17) x_1 і x_2.

Крок 2. Якщо $x_1 > x_2$, то вибираємо x', якщо $x_1 < x_2$, те вибираємо x''.

Крок 3. Якщо не існує вектора $x \in X$ (переважно x' або x''), то обчислення зупиняється, інакше вибираємо новий вектор x''' і переходимо до кроку 1.

Модель визначення області компромісів, а також модель справедливого компромісу інваріантні до масштабу виміру критеріїв. Однак, перша з них є допоміжною при виборі розв'язку, а інша може бути використана тільки в тих ситуаціях вибору розв'язку, для яких ідея справедливого компромісу може бути виправдана. Тому в більшості випадків доводиться використовувати інші принципи оптимальності, що доцільно тільки у випадку нормалізованого простору критеріїв, коли всі локальні критерії мають єдиний масштаб виміру. Однак, у більшості випадків масштаби виміру критеріїв неоднакові, і виникає необхідність проводити нормалізацію критеріїв, тобто штучно зводити їх до єдиної міри.

Більшість принципів нормалізації ґрунтується на введенні ідеального розв'язку, тобто розв'язку, що володіє ідеальним вектором ефективності Q_i. Цей ідеальний розв'язок можна апріорно припустити виходячи з інформації про об'єкт. А можна вирішити задачу оптимізації для кожного локального критерію та значення вектора ефективності, що відповідає цьому розв'язку, прийняти за ідеальний вектор ефективності Q_i. Тоді вибір оптимального розв'язку стає рівнозначним найкращому наближенню до цього ідеального вектора $Q = (Q_{i1}, ..., Q_{in})$ У цьому випадку замість дійсного значення критеріїв розглядають їх відхилення від ідеального значення

$$dQ_j = Q_{ij} - Q_j, \tag{2.18}$$

або їх безрозмірне відносне значення

$$\overline{Q}_j = Q_j / Q_{ij}. \tag{2.19}$$

При розв'язуванні цієї проблеми використовуються обидва способи нормалізації. Таким чином, успішний розв'язок проблеми нормалізації багато в чому залежить від того, наскільки точно й об'єктивно вдається визначити ідеальну якість розв'язку. Після нормалізації критеріїв ефективності задача вибору розв'язку набуває математичного змісту. У теоретико-множинному відношенні вона стає задачею впорядкування обмежених векторних множин, а з погляду теорії наближень – задачею наближення в метричному скінченно-вимірному просторі. Це дає можливість проводити обґрунтований вибір принципів оптимальності і виявляти їх логічний зміст. Отже, для цього випадку принцип оптимальності ідентичний принципу наближення, а узагальнений скалярний критерій оптимізації – критерієві наближення, що є деякою функцією відхилення від ідеальної функції.

Принцип слабкої оптимальності по Парето. Вектор $x_1 \in X$ називається слабко оптимальним по Парето розв'язком (оптимальним по Слейтеру), якщо не існує вектора $x_1 \in X$, такого, що

$$q(x \mid y) \geq q\left(x^1 \mid y\right) \forall y \in Y. \tag{2.20}$$

Нехай $x_{oj} \quad j = \overline{1,m}$ є оптимальні розв'язки для звичайних скалярних оптимізаційних задач, кожна з яких максимізує компоненту $Q_i(x)$ вектора $Q(x)$:

$$\max_{x \in X} Q_j(x), j = \overline{1,m}. \tag{2.21}$$

Якщо вони є максимальними розв'язками для кожної i, то вважаємо, що $Q_j\left(x_{oj}\right) > Q_i\left(x_j\right) \quad j = \overline{1,m}$, де x_{oj} — оптимальний розв'язок задачі. Покладемо, що S_{oj} відображає множину розв'язків, кожен з яких відповідає компоненті Q_j, і

$$S_{oj} = \left\{x \mid Q_j(x) \leq Q_j\left(x_{oj}\right) + a_j\right\}, \tag{2.22}$$

де a_j – допустима кількість обмежень відповідної області стосовно Q_j. Тоді оптимальний достатній розв'язок – це такий розв'язок, при якому мінімальний компонент (найгірший компонент) максимізується на множині, що задовольняє достатній умові $x \in X$ та $x \in S_{o_1 n}, S_{o_2 n}, \ldots, S_{o_m n} \ldots$. Вона може бути сформульована як

$$\max z. \tag{2.23}$$

$$Q_j(x) \geq z, x \in X \tag{2.24}$$

$$Q_i(x) \geq Q_j\left(x_{oj}\right) - a_j, \; i = \overline{1,m}. \tag{2.25}$$

Тут задача (2.23)–(2.25) немає розв'язку, якщо a_j не настільки велике, що перетин $\{S_j\}$ непорожній. Величини a_j повинна бути визначена на основі значень $Q_j(x_{oj})$ або аналізу точності. Можна довести, що оптимальний розв'язок задачі (2.23) – (2.25) є оптимальний розв'язок по Парето. Алгоритм розв'язування задачі має наступні етапи.

Крок 1. Беремо l=1 і розв'язуємо задачу (2.23) – (2.25).

Викликаємо початковий розв'язок x_l і оцінюємо цільову функцію $Q(x_l)$.

Крок 2. Коли x_l задано, розкладаємо $Q(x_l)$ на задовільні і незадовільні компоненти. Позначимо їх відповідно через S_l і $\overline{S}_l$.

Якщо $S_l \neq \varnothing$, тоді ця задача вважається нерозв'язною, а якщо $S_l \equiv \varnothing$, то x_l – оптимальний, що відповідає вимогам рішення. Якщо $\overline{S}_l \neq \varnothing$ і $S_l \neq \varnothing$, то для $j \in S_l$ визначається $a_{lj}>0$, допустиме стосовно $Q_j(x_l)$ (a_{lj}=0 означає, що j-а цільова функція $q_j(x)$ не може приймати значення, відмінне від $q_j(x_l)$).

Крок 3. Розв'язуємо задачу max z за умови

$$Q_j(x) \geq z, \; j \in \overline{S}_l. \tag{2.26}$$

$$Q_i(x) \geq Q_j\left(x_{oj}\right) - a_j, \; i = \overline{1,m}, \; j \in S_l, \; x \in X. \tag{2.27}$$

Викликаємо вихідний розв'язок x_{l+1}. Якщо $x_{l+1}= x_l$, то задача немає розв'язку; якщо $x_{l+1} \neq x_l$, то беремо $l= l+1$ і повертаємося до кроку 2.

При цьому алгоритм закінчується.

Принцип наближення по всіх локальних критеріях до ідеального розв'язку. В основу цього підходу покладена ідея наближення за всіма критеріями.

Нехай ця задача багатокритеріального програмування

$$\begin{aligned} &\max\{q_1(x) = Q_1\} \\ &\max\{q_2(x) = Q_2\} \\ &\ldots\ldots\ldots\ldots\ldots \\ &\max\{q_k(x) = Q_k\} \end{aligned} \qquad (2.28)$$

і задані граничні умови

$$\alpha_{\alpha_1} x_1 + \alpha_{\alpha_2} x_2 + \ldots + \alpha_{\alpha_n} x_n = b_\alpha \left(\alpha = \overline{1,r}\right). \qquad (2.29)$$

$$\alpha_{r+\beta,1} x_1 + \alpha_{r+\beta,2} x_2 + \ldots + \alpha_{r+\beta,n} x_n \le b_{r+\beta} \left(\beta = \overline{1,m-r}\right). \qquad (2.30)$$

$$x_1 \ge 0, x_2 \ge 0, \ldots, x_n \ge 0, \ldots. \qquad (2.31)$$

Серед розв'язків системи (2.28) – (2.31) потрібно відшукати таке значення вектора $x^*\left(x_1^*, \ldots, x_n^*\right)$, при якому локальні критерії приймуть по можливості максимальне (мінімальне) значення одночасно. Розглянемо кожну окрему функцію $q_i(x)$ і припустимо, що для кожного фіксованого i ($i = \overline{1,m}$) розв'язана задача максимізації. Нехай відповідні оптимальні значення характеризуються векторами

$$x_{oi}\left(x_{o1}, x_{o2}, \ldots, x_{on}\right), i = \overline{1,m}. \qquad (2.32)$$

На цих оптимальних значеннях визначимо значення критеріїв відповідно

$$Q_{oi}\left(x_{oi}\right) = \left(Q_{o1}\left(x_{o1}\right), Q_{o2}\left(x_{o2}\right), \ldots, Q_{on}\left(x_{on}\right)\right). \qquad (2.33)$$

Зрозуміло, що вектори (2.32), що визначають вектори точки в просторі змінних $\left(x_1, x_2, \ldots, x_n\right) \in \Omega$, будуть різними: деякі з них можуть збігатися один з одним. Розглянемо вектор $Q(x)$ з компонентами $Q(x) | F_{oi}$ з (2.32) і побудуємо квадрат евклідової норми

$$R(x) = \left\| Q(x) - Q^0 \right\|^2. \qquad (2.34)$$

вектора $Q(x) - Q^0$, визначеного для всіх $x \in \Omega$.

Зауважимо, що Q^0 буде являти собою одиничний вектор у просторі вектора $Q(x)$. Назвемо його ідеальним значенням вектора $Q(x)$. Поставлена задача тепер сформулюється так: ця система цільових функцій (2.28) і ці умови задачі (2.29) – (2.31). Потрібно визначити точку $x \in \Omega$, у якій функція $R(x)$ досягає мінімуму.

Таким чином, пошук векторно-оптимального плану $x \in \Omega$ в цій задачі зведено до оптимізації виразу (2.34) на розв'язах системи лінійних нерівностей (2.29) – (2.31). Оскільки вираз (2.34) являє собою квадратичну функцію змінних $(x_1, x_2, ..., x_n) \in \Omega$, то задача знаходження векторно-оптимального значення звелася до задачі опуклого програмування: задана опукла функція $R(x)$, визначена на множині $x \in \Omega$. Потрібно відшукати точку $x \in \Omega$, що забезпечує виконання умови $R(x^*) = \min(x)$, $x \in \Omega$. Таким чином, алгоритм розв'язку задачі (2.28) – (2.31) складається з двох основних етапів:

етап 1: $\max Q_i(x), i = \overline{1,m}$;

етап 2: $\min R(x)$.

2.1.8. Метод квазіоптимізації локальних критеріїв (метод послідовних поступок)

У цьому випадку здійснюється пошук не єдиного точного оптимуму, а деякої області розв'язків, близьких до оптимального – *квазіоптимальної множини*. При цьому рівень допустимого відхилення від точного оптимуму визначається з урахуванням точності постановки задачі (наприклад, в залежності від точності обчислення величини критеріїв), а також деяких практичних понять (наприклад, вимог точності розв'язку задачі). Спочатку виконується якісний аналіз відносної важливості критеріїв; на підставі такого аналізу критерії розташовуються і нумеруються в порядку зменшення важливості, так що головним вважається критерій Q_1, менш важливий Q_2, потім випливають інші локальні критерії Q_3, Q_4, ..., Q_m. Максимізується перший по важливості критерій Q_1 і визначається його найбільше значення W_1. Потім визначається допустиме зниження (поступка) $d_1 \geq 0$ критерію Q_1. Визначимо нову допустиму область $X(1)$, як підобласть X вигляду:

$$X(1) = X_n\{x \mid Q_1(x) \geq W_1 - d_1\}. \tag{2.35}$$

Такий підхід дозволяє значно звузити первісну допустиму область X, коли переходимо до наступного по важливості критерію. Після цього знаходимо найбільше значення другого критерію Q_2 на множині $X(1)$ за умови, що значення першого критерію повинно бути не менше, ніж W_1-d_1. Знов визначається значення поступки $d_2 \geq 0$, але вже за другим критерієм, що разом з першим використовується при знаходженні умовного максимуму третього критерію, тощо. Нарешті, максимізується останній по важливості критерій Q_m за умови, що значення кожного критерію Q_r з m-1 попередніх повинно бути не менше відповідної величини W_r-d_r. Стратегії, що одержані, вважаються оптимальними:

$$X(i) = X_n(i-1)\{x \mid Q_i(x) \geq W_i - d_i\}. \tag{2.36}$$

Таким чином, оптимальною вважається будь-яка стратегія, що є розв'язком останньої задачі з наступної послідовності задач: знайти

$$1)W_1 = \sup_1(x), x \in X$$
$$2)W_2 = \sup_2(x), x \in X \qquad (2.37)$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$m)W_m = \sup_m(x), x \in X$$

Якщо критерій Q_m на множині стратегій, що задовольняють обмеженням задачі m) з (2.37) не досягає свого найбільшого значення W_m, то розв'язком багатокритеріальної задачі вважають максимізуючу послідовність $\{x_k\}$ з послідовності множин

$$X_{m-1} \subset X_{m-2} \subset \ldots \subset X_1 \subset X. \qquad (2.38)$$

Практично подібні максимізуючі послідовності є зміст розглядати і для того випадку, коли верхня грань у задачі m) досягається, так як для розв'язування екстремальних задач широко застосовуються ітеративні методи.

Алгоритм розв'язування задачі векторної оптимізації включає наступні кроки.

Крок 1. Нехай x_{o1} – розв'язок задачі (2.37)

$$\max Q_1(x), x \in X. \qquad (2.39)$$

Крок 2. Нехай x_{ok} – розв'язок задачі

$$\max Q_k(x), x \in X(k-1), \qquad (2.40)$$

де X_k визначається з (2.36)-(2.37).

Крок 3. Якщо $k<m$, то беремо $k=k+1$ і повторюємо крок 2. Якщо $k=m$, то x_{om} вважаємо оптимальним розв'язком. Алгоритм завершений. Значення поступок d_i $\left(i=\overline{1,m}\right)$ послідовно визначаються при вивченні взаємозв'язку приватних критеріїв.

Спочатку вирішується питання про визначення допустимого зниження d_1 першого критерію від найбільшого значення W_1. Практично для цього задають кілька величин поступок $d_{11},d_{21},d_{31},\ldots$ і шляхом розв'язування задачі (2.37) визначають відповідні максимальні значення $W_2(d_{11}),W_2(d_{21}),W_2(d_{31}),\ldots$. Далі розглядають пари критеріїв Q_2 і Q_3. Знову визначають пробні значення поступок $d_{12},d_{22},d_{32},\ldots$ і, вирішуючи задачу, шукають найбільші значення $W_3(d_{12}),W_3(d_{22}),W_3(d_{32}),\ldots$. Отримані дані аналізують, призначають d_2, переходять до наступної пари критеріїв Q_2 і Q_3. тощо. Нарешті, в результаті аналізу взаємного впливу критеріїв Q_{m-1} і Q_m вибирають значення останньої поступки d_{m-1} і шукають оптимальні стратегії, розв'язуючи задачу m) з (2.37) (звичайно обмежуються однією такою стратегією). Таким чином, хоча формально при використанні методу послідовних поступок досить розв'язати лише задачу (2.37), однак, для визначення значення поступок з метою з'ясування взаємозв'язку приватних критеріїв фактично доводиться розв'язувати більше таких задач.

Для розв'язування багатокритеріальної задачі потрібно так ранжирувати критерії. щоб потім зручніше було вибирати значення поступок. З огляду на зазначене вище, можна зробити наступний висновок. Метод послідовних поступок доцільно застосовувати для розв'язування тих багатокритеріальних задач, в яких усі приватні критерії природнім чином впорядковані по важливості, причому кожен критерій

настільки істотно більш важливий, ніж наступний, що можна обмежитися обліком тільки попарного зв'язку критеріїв і вибирати допустиме зниження чергового критерію з урахуванням поводження лише одного наступного критерію.

Особливо зручним є випадок, коли в результаті попереднього аналізу багатокритеріальної задачі з'ясовується, що можна допустити поступки лише в межах "інженерної" точності (5-10% від найбільшого значення критерію).

2.1.9. Метод згортання векторного критерію в суперкритерій

Одним з розповсюджених методів розв'язування багатокритеріальних задач є метод зведення багатокритеріальної задачі до однокритеріальної шляхом згортання векторного критерію в суперкритерій. При цьому кожен критерій збільшується на відповідний йому ваговий коефіцієнт (коефіцієнт важливості).

$$F(x)=\sum_{i=1}^{m}\alpha_i Q_i(x),\ \alpha_i \geq 0. \tag{2.41}$$

В цьому випадку виникають труднощі із *правильним* підбором вагових коефіцієнтів a_i. Існують різні способи вибору коефіцієнтів a_i. Одним із них є знаходження a_i в залежності від відносної важливості критеріїв. Такий підбір зазначених коефіцієнтів можна виконувати відповідно до таблиці:

Таблиця 2.3

Шкала відносної важливості

Інтенсивність відносної важливості	Визначення
1	Рівна важливість порівнюваних вимог
3	Помірковано (слабка) перевага одного над іншим
5	Сильна (істотна) перевага
7	Очевидна перевага
9	Абсолютна (гнітюча) перевага
2,4,6,8	Проміжні роз'язки між двома сусідніми оцінками

2.2. Безумовна оптимізація

Задача *безумовної оптимізації* – це знаходження мінімуму чи максимуму функції при відсутності будь-яких обмежень. Незалежно від того, що більшість практичних задач оптимізації мають обмеження, вивчення методів безумовної оптимізації важливе з декількох точок зору. Велика кількість алгоритмів розв'язку задачі з обмеженнями припускає зведення її до послідовності задач безумовної оптимізації за допомогою множників Лагранжа чи за допомогою штрафних та бар'єрних функцій. Інший клас методів заснований на пошуку відповідного напряму та наступної мінімізації паралельно до цього напряму. Лінійний пошук за напрямом еквівалентний мінімізації функції однієї змінної без обмежень чи з простими обмеженнями, такими, як двохсторонні обмеження на змінну. Нарешті, пояснення методів безумовної оптимізації може бути природним чином поширене на пояснення процедур розв'язування задач з обмеженнями.

2.2.1. Метод штрафних функцій

Метод штрафних функцій оснований на перетворені вихідної задачі з обмеженнями в одну чи в послідовність задач безумовної оптимізації. За допомогою функцій, що задають обмеження, будується так званий штраф, який додається до цільової функції вихідної задачі так, що порушення будь-якого із обмежень стає невигідним з точки зору отриманої задачі безумовної оптимізації. Для пояснення цього методу розглянемо наступну задачу з єдиним обмеженням $h(x)=0$:

1. Мінімізувати $f(x)$ при умові $h(x)=0$.
2. Нехай вона перетворена в задачу безумовної оптимізації:
3. Мінімізувати $f(x)+\mu h^2(x)$ при умові $x\in E_n$, де $\mu>0$ - деяке велике число. Інтуїтивно зрозуміло, що на оптимальному розв'язку $\bar{x}$ останньої задачі значення $h^2(\bar{x})$ повинно бути наближеним до нуля, так як в іншому випадку завжди можна змістить в іншу точку x', в якій приріст $f(x)$ буде при достатньо великому μ меншим, ніж $\mu h^2(x)$.

Розглянемо тепер задачу з єдиним обмеженням-нерівністю: мінімізувати $f(x)$ при умові $g(x)\le 0$. Очевидно, що форма $f(x)+\mu g^2(x)$ у цьому випадку недоцільна, так як при штраф буде зніматись незалежно від знаку $g(x)$. Бажано щоб штраф знімався лише в недопустимих точках, тобто, там де $g(x)>0$. Тому допустимою еквівалентною задачею безумовної оптимізації є наступна задача: мінімізувати $f(x)+\mu\max\{0,g(x)\}$ при умові $x\in E_n$. Якщо $g(x)\le 0$, то $\max\{0,g(x)\}=0$ і штраф не береться. Якщо ж $g(x)>0$, то $\max\{0,g(x)\}>0$, тобто знімається штраф $\mu g(x)$. Зазвичай підходяща штрафна функція повинна визначати додатній штраф в недопустимих точках і не штрафувати допустимі точки. Якщо обмеження мають форму $g_i(x)\le 0, i=1,\dots,m$ і $h_i(x)\le 0, i=1,\dots,l$, то доцільною є штрафна функція наступного вигляду:

$$\alpha(x)=\sum_{i=1}^{m}\varphi[g_i(x)]+\sum_{i=1}^{l}\psi[h_i(x)], \tag{2.42}$$

де φ і ψ – неперервні функції, що задовільняють умовам:

$\varphi(y)=0$, якщо $y\le 0$, і $\varphi(y)>0$, якщо $y>0$,

$\psi(y)=0$, якщо $y=0$, і $\psi(y)>0$, якщо $y\ne 0$.

Типовими формами функцій φ і ψ є наступні форми:

$$\varphi(y)=[\max\{0,y\}]^p, \tag{2.43}$$
$$\psi(y)=|y|^p,$$

де p – ціле додатнє число. Таким чином, штрафна функція $\alpha=\alpha(x)$ зазвичай має вигляд:

$$\alpha(x)=\sum_{i=1}^{m}[\max\{0,g_i(x)\}]^p+\sum_{i=1}^{l}|h_i(x)|^p. \tag{2.44}$$

Функцію $f(x)+\mu\alpha(x)$ будемо називати *допоміжною*, відповідну задачу з штрафом будемо називати *допоміжною задачею*.

Приклад 2.1. Розглянемо задачу: мінімізувати x при умові $-x+2\le 0$.

Розв'язок. Покладемо $\alpha(x)=[\max\{0,g(x)\}]^2$. Тоді

$$\alpha(x)=\begin{cases}0, & x\ge 2,\\ (-x+2)^2, & x<2.\end{cases}$$

На рис. 2.1 зображені штрафна функція $\alpha(x)$ та допоміжна функція $f(x)+\mu\alpha(x)$. Зауважимо, що мінімум $f+\mu\alpha$ досягається у точці $2-\dfrac{1}{2\mu}$. При $\mu\to\infty$ послідовність таких точок збігається до $\bar{x}=2$, яка є точкою мінімума цільової функції вихідної задачі.

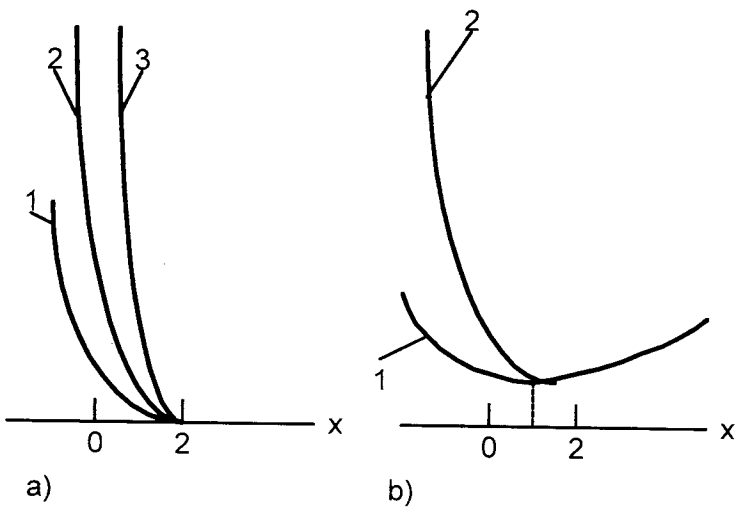


Рис. 2.1. Штрафні та допоміжні функції при а) 1 - $\mu_1\alpha$, 2 - α, 3 - $\mu_2\alpha$; b) 1 - $f+\mu_1\alpha$, 2 - $f+\mu_2\alpha$, причому $\mu_1=0.5,\mu_2=1.5$

Приклад 2.2. Розглянемо задачу: мінімізувати $x_1^2+x_2^2$ при умові $x_1+x_2-1=0$.

Розв'язок. Оптимальне значення досягається в точці $\left(\dfrac{1}{2},\dfrac{1}{2}\right)$ і воно рівне $\dfrac{1}{2}$.

Тепер побудуємо наступну задачу зі штрафом при достатньо великому μ: мінімізувати $x_1^2+x_2^2+\mu(x_1+x_2-1)^2$ при умові $(x_1,x_2)\in E_2$. Зауважимо, що при довільному $\mu\ge 0$ цільова функція цієї задачі є випуклою. Тоді необхідною і достатньою умовою

оптимальності є рівність нулю градієнта функції $x_1^2 + x_2^2 + \mu(x_1 + x_2 - 1)^2$, тобто $x_1 + \mu(x_1 + x_2 - 1) = 0$, $x_2^2 + \mu(x_1 + x_2 - 1) = 0$.

Розв'язуючи цю систему з двох рівнянь, отримуємо $x_1 = x_2 = \dfrac{\mu}{1 + 2\mu}$. Звідси випливає, що розв'язок допоміжної задачі може бути знайдений як завгодно близьким до розв'язку вихідної задачі при виборі достатньо великого μ.

2.2.2. Геометрична інтерпретація штрафних функцій

Використаємо приклад 2.2 для того, щоб подати геометричну інтерпретацію штрафних функцій. Припустимо, що обмеження $h(x)=0$ порушене так, що $h(x) = x_1 + x_2 - 1 = \varepsilon$, тобто отримана наступна задача: мінімізувати $x_1^2 + x_2^2$ при умові $x_1+x_2-1=\varepsilon$. Якщо підставити $x_2 = 1 + \varepsilon - x_1$ в цільову функцію, то задача зводиться до мінімізації $x_1^2 + (1+\varepsilon - x_1)^2$. Оптимум знаходиться там, де похідна обернеться в нуль, тобто оптимальне значення x_1 задовільняє рівняння $2x_1 - 2(1+\varepsilon - x_1) = 0$. Тому при будь-якому фіксованому значенні ε оптимальним розв'язком сформульованої вище задачі є точка з координатами $x_1 = x_2 = (1+\varepsilon)/2$, а оптимальне значення цільової функції рівне $\frac{1}{2}(1+\varepsilon)^2$. Крім того, при будь-якому ε верхня межа значень $x_1^2 + x_2^2$ при умові $x_1 + x_2 - 1 = \varepsilon$ рівна ∞. Тому для будь-якої точки $(x_1, x_2) \in E_2$, для якої $x_1 + x_2 - 1 = \varepsilon$, значення цільової функції лежить в інтервалі $\left[\frac{1}{2}(1+\varepsilon)^2, \infty\right]$. Множина $\left\{\left[h(x), f(x)\right] : x \in E_2\right\}$ зображена на рис.2.2. Ця множина огинається знизу параболою $\frac{1}{2}(1+h)^2 = \frac{1}{2}(1+\varepsilon)^2$. При фіксованому значенні $\mu > 0$ допоміжна задача складається із мінімізації $f(x) + \mu h^2(x)$ на E_2. Графік $f(x) + \mu h^2(x) = k$ в просторі (h,f) зображений на рис. 2.2 пунктирною параболою. Ця парабола пересікається з віссю f в точці $f=k$. При мінімізації $f + \mu h^2$ парабола повинна опускатися як можно нижче до тих пір, поки залишиться хоча би одна загальна точка з заштрихованою множиною, тобто поки парабола не стане лише дотикатися до заштрихованої області, як зображено на рис. 2.2. Це значить, що для данного значення μ оптимальне значення цільової функції допоміжної задачі рівне відповідному відрізку на вісі f. Зауважимо, що оптимальний розв'язок є недопустимою точкою вихідної задачі, так як $h \neq 0$ в точці дотику. Більш того, оптимальне значення цільової функції допоміжної задачі менше, ніж вихідної. Зауважимо також, що зі зростанням μ парабола $f + \mu h^2$ стає гострішою і точка дотику наближується до оптимального розв'язку вихідної задачі.

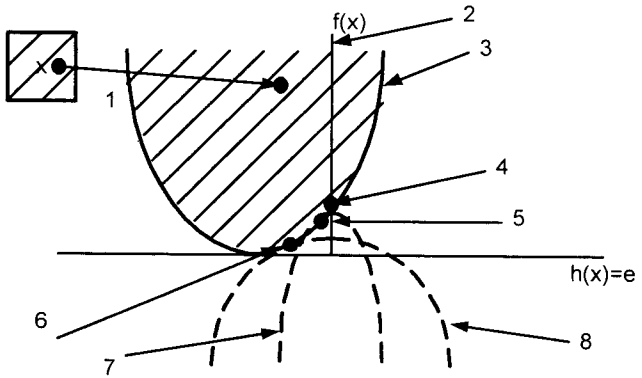


Рис. 2.2. Штрафні функції в просторі (h,f), де 1 – (h,f) - відображення; 2 – допустимі точки вихідної задачі; 3 – нижча парабола, що проходить через точки вигляду $\left[\varepsilon, \frac{(1+\varepsilon)^2}{2}\right]$; 4 – оптимальний розв'язок вихідної задачі; 5 – оптимальний розв'язок допоміжної задачі при $\mu' > \mu$; 6 – оптимальний розв'язок допоміжної задачі при значенні параметру, що рівне μ; 7 - $f + \mu' h^2$; 8 - $f + \mu h^2$

Ідея градієнтних методів базується на тому, що напрямок градієнта є напрямком найшвидшого зростання функції. Отже, протилежний напрямок є напрямком найшвидшого спадання функції. Напрямок градієнта перпендикулярний в будь-якій точці лінії постійного рівня, оскільки уздовж цієї лінії значення функції є сталим. В методі найшвидшого спуску пошук мінімуму функції здійснюється вздовж напрямку, оберненого до напрямку градієнту. Як критерій завершення обчислень, найчастіше використовується близкість до нуля градієнту. Основною складністю застосування методу найшвидшого спуску є його залежність від вибору масштабу змінних, що оптимізуються. В методах, що використовують другі похідні, використовується матриця Гессе, яку необхідно інвертувати на кожному кроці, як в методі Ньютона-Рафсона. Оскільки ця процедура може займати основний час обчислень, використовуються наближення до неї, і позитивно визначена симетрична матриця при виконанні певних умов стає рівною асиптотично матриці Гессе, як у методі Давідсона-Флетчера-Павела. Основна ідея методів штрафів для розв'язування задачі оптимізації з обмеженнями полягає у перетворені задачі мінімізації функції з обмеженнями у задачу пошуку мінімуму без обмежень модифікованої функції, в якій при порушенні обмежень додаткова складова, яка сформована на основі обмежень, "штрафувала" б модифіковану функцію, збільшуючи її значення. На цій ідеї побудовані метод Фіако-Маккорміка, який як крок алгоритму використовує метод Девідсона-Флетчера-Павела, та метод Бокса, який, по суті, є модифікацією симплексного методу Нелдера-Міда, що дозволяє враховувати обмеження. Для задач оптимізації функцій однієї змінної без обмежень умови екстремуму задаються у вигляді рівності нулеві похідної від функції, а характер екстремуму (мінімум, максимум, точка перегину) визначається знаком другої похідної.

У випадку функцій n змінних у результаті прирівнювання часткових похідних до нуля виникає система нелінійних рівнянь, а характер екстремуму визначається негативною чи позитивною визначеністю матриці Гессе в точці екстремуму. За будь-яких умов аналітичний розв'язок цих задач можливий лише в найпростіших випадках.

У випадку пошуку екстремумів функції n змінних з обмеженнями у вигляді рівнянь первинна функція заміняється функцією Лагранжа, і система рівнянь отримується шляхом знаходження часткових похідних від змінних та множників Лагранжа і прирівнюванням їх до нуля. Умови Куна-Такера є розширенням ідей методу множників Лагранжа на задачі з обмеженнями загального вигляду – як у вигляді рівнянь, так і у вигляді нерівностей. Ці умови є необхідними умовами, які повинні виконуватись в стаціонарній точці. Розв'язування отриманої системи рівнянь та обмежень можливе лише для найпростіших задач, і в загальному для розв'язування використовуються чисельні методи.

Методи прямого пошуку є методами, у яких використовуються лише значення функції, і які орієнтовані на пошук екстремумів унімодальних функцій. Ідеї та складові методів прямого пошуку використовуються і при побудові складніших методів пошуку оптимальних розв'язків. Найефективнішим є метод, що використовує числа Фібоначі, однак, з точки зору зручності реалізації, доцільно використовувати метод золотого перетину, який незначно поступається методу Фібоначі. Окрім цих методів, використовують також апроксимаційні, що використовують ідею квадратичної інтерполяції. Методи прямого пошуку застосовуються також і для пошуку екстремумів функцій n змінних. Найпростішим з них є метод покоординатного спуску (у випадку пошуку мінімуму), який почергово по кожній зі змінних реалізує одновимірний прямий пошук, і в якості процедури одновимірного пошуку найчастіше використовується метод золотого перетину. Однак, з практичної точки зору цей метод є доволі повільним, а тому були розроблені ефективніші методи, одним з найвідоміших серед яких є метод Хука-Дживса. Цей метод використовує інформацію про локальну поведінку функції для обрання найвідповідальнішого напрямку пошуку за зразком. Метод Нелдера–Міда (або пошук по деформованому многограннику) є розвитком симплексного методу. Ідея методу полягає в порівнянні значень функції в $(n+1)$ вершинах симплекса і переміщенні симплекса в напрямку оптимальної точки за допомогою ітераційної процедури. Нелдер і Мід запропонували декілька модифікацій цього методу з неправильними симплексами. У результаті утворився дуже надійний метод прямого пошуку, що є одним із найефективніших, якщо $n<6$.

2.2.3. Аналіз алгоритмів розв'язування задачі методом штрафних функцій

Основна ідея методу полягає у перетворенні задачі мінімізації функції $z = f(x)$ з відповідними обмеженнями, накладеними на x у задачу пошуку мінімуму без обмежень функції $Z = f(x) + P(x)$. Функція $P(x)$ є *штрафною*. Необхідно, щоб при порушенні обмежень вона "штрафувала" функцію Z, тобто збільшувала її значення. У цьому випадку мінімум Z знаходиться всередині області обмежень. Функція $P(x)$, що задовольняє цій умові, може біти не єдиною. Задачу мінімізації можна сформулювати наступним чином: мінімізувати функцію

$$z = f(x) \tag{2.45}$$

$$c_j(x) \geq 0,\ j = 1,2,...,m. \tag{2.46}$$

Функцію $P(x)$ зручно записати наступним чином:

$$P(x) = r\sum_{j=1}^{m}\frac{1}{c_j(x)}. \tag{2.47}$$

де r – додатня величина. Тоді функція $Z = \varphi(x,r)$ буде мати вигляд

$$Z = \varphi(x,r) = f(x) + r\sum_{j=1}^{m}\frac{1}{c_j(x)}. \tag{2.48}$$

Якщо x приймає допустимі значення, тобто значення, для яких $c_j(x) \geq o$, то Z приймає значення, які більше відповідних значень $f(x)$ (істинної функції мети задачі), і різницю можна зменшити за рахунок того, що r може бути дуже малою величиною. Але якщо x приймає значення, які хоч і є допустимими, але близькі до границі області обмежень, і щонайменш одна з функції $c_j(x)$ близька до нуля, тоді значення функції $P(x)$ і, відповідно, значення функції Z стануть дуже великими. Таким чином, вплив функції $P(x)$ полягає в створенні "хвилі з крутими межами" вздовж кожної границі області обмежень. Отже, якщо пошук починається з допустимої точки і здійснюється пошук мінімуму функції $\varphi(x,r)$ без обмежень, то мінімум, звісно, буде досягатись в середині допустимої області для задачі з обмеженнями. Приймаючи r достатньо малою величиною, для того щоб вплив $P(x)$ був малим у точці мінімуму, ми можемо зробити точку мінімуму функції $\varphi(x,r)$ *без обмежень*, щоб вона співпадала з точкою мінімуму функції $f(x)$ з обмеженнями.

Розглянемо простий приклад для ілюстрації наведеного вище.

Приклад 2.3. Використовуючи штрафну функцію, яка задана рівнянням (2.47), мінімізувати функцію $f(x) = x$ при обмеженнях $x \geq 2$, тобто $x - 2 \geq 0$. Мінімальним значенням функції є 2 при x=2. Як за допомогою штрафної функції можна знайти розв'язок? Розглянемо функцію $\varphi(x,r) = x + \frac{r}{x-2}$.

Розв'язок. Область обмежень лежить справа від вертикальної прямої x=2. Не важко помітити, що послідовність точок Q_1, Q_2, Q_3 прямує до точки Q – мінімуму функції при наявності обмежень. Дійсно, знайдемо мінімум функції $\varphi(x,r)$ методом штрафних функцій: $\frac{d\varphi}{dx} = 1 - \frac{r}{(x-2)^2}$. Якщо $d_\varphi / dx = 0$, $(x-2)^2 = r$, то $x = 2 \pm \sqrt{r}$. Тоді $\frac{d^2\varphi}{dx^2} = \frac{2r}{(x-2)^3}$ і мінімум досягається при $x = 2 + \sqrt{r}$ всередині області обмежень. Отже, функція $\varphi(x,r)$ має мінімум, що рівний $2 + 2\sqrt{r}$ при $x = 2 + \sqrt{r}$. Тоді Q_1 є точкою з координатами (3;4), Q_2 - точка з координатами (2,5;3), Q_3 - точка з координатами (2,1;2.2). При $r \to 0$ мінімум без обмежень функції $\varphi(x,r)$ наближається до значення 2 і

мінімальною точкою є точка x=2. В загальному випадку неможливо аналітично визначити положення мінімуму функції $\varphi(x,r)$, якщо розглядати її як звичайну функцію від r. Для його визначення необхідно звернутись до чисельних методів. Слід відзначити, що якщо функція мети $f(x)$ опукла, а функція $c_j(x)$ увігнута, то функція $\varphi(x,r)$, задана рівнянням (2.48), також є опуклою функцією у області обмежень, яка сама є опуклою. Отже, $\varphi(x,r)$ має для цього значення r єдиний мінімум. Якщо x_1 і x_2 – точки, які належать допустимій області, тобто $c_j(x_1) \ge 0$ і $c_j(x_2) \ge 0$ для $j = 1,2,...,m$, то при $0 < \theta < 1$ справедлива нерівність $c_j(\theta x_2 + (1-\theta)x_1) \ge \theta c_j(x_2) + (1-\theta)c_j(x_1) \ge 0$, так як функція $c_j(x)$ опукла. Отже, допустима область є опуклою.

Таким чином, точка $x_2 + (1-\theta)x_1$ при $0 < \theta < 1$ також є допустимою.

Крім того, функція $1/c_j(x)$ є опуклою для усіх x, які задовольняють нерівність $c_j(x) \ge 0$. Якщо $h(x) = 1/c_j(x)$, то $\nabla h(x) = \dfrac{-\nabla c_j(x)}{[c_j(x)]^2}$. Отже, гессіан функції $h(x)$ має вигляд $H(x) = -\dfrac{C(x)}{[c_j(x)]^2} + \dfrac{2\nabla c(x)\nabla c(x)^{\mathrm{T}}}{[c_j(x)]^3}$, де $C(x)_{ik} = \partial^2 c_j(x)/\partial x_i \partial x_k$ є гессіан функції $c_j(x)$. Тоді, якщо p – довільний вектор, то справедлива рівність $p^{\mathrm{T}}H(x)p = -\dfrac{p^{\mathrm{T}}C(x)p}{[c_j(x)]^2} + \dfrac{2[p^{\mathrm{T}}\nabla c_j(x)]^2}{[c_j(x)]^3}$, де завжди $p^{\mathrm{T}}H(x)p > 0$, так як $C(x)$ – від'ємно визначена матриця, так як $c_j(x)$ – опукла функція і $c_j(x) \ge 0$. Тоді матриця $H(x)$ позитивно визначена і $1/c_j(x)$ опукла по всій області. Якщо $r > 0$, то функція $P(x)$, що задана рівнянням (2.47), і функція $\varphi(x,r)$, що задана рівнянням (2.48), також опуклі. Тепер можна узагальнити результат, отриманий в прикладі 2.1 цього розділу на випадок загальної задачі з обмеженнями (2.45) і (2.46). Нехай, $x_1^*, x_2^*, ..., x_k^*$ – мінімальні точки функції $\varphi(x,r)$ для спадаючої послідовності значень $r_1, r_2, ..., r_k, ...$, що прямує до нуля. Тоді послідовність точок $x_1^*, x_2^*, ..., x_k^*, ...$ збігається до оптимального розв'язку задачі з обмеженнями (2.45) і (2.46) при $r_k \to 0$. Отже,

$$\lim \; x_k^* = x^*. \tag{2.49}$$

$$\lim_{rk \to 0}[\min \varphi(x, r_k)] = f(x^*), \tag{2.50}$$

де x^* – мінімальна точка функції $f(x)$ при наявності обмежень.

Отриманий результат можна довести наступним чином. Оскільки, $f(x)$ неперервна функція і $f(x^*) \le f(x)$ для усіх допустимих точок, то, задавши довільне достатньо мале значення ε, можна знайти допустиму точку x', таку, що

$$f(x') < f(x^*) + \varepsilon/2. \tag{2.51}$$

Оскільки r_k – спадна послідовність, яка прямує до нуля, тоді можна знайти таке значення K, що для $k \geq K$ справедлива нерівність

$$r_k \leq \left\{ \frac{\varepsilon}{2m} \min_j \left[\frac{1}{c_j(x)} \right] \right\}. \tag{2.52}$$

Оскільки $P(x) > o$ з визначення функції $\varphi(x,r)$, маємо

$$f(x^*) \leq \min \varphi(x, r_k) = \varphi(x_k^*, r_k), \tag{2.53}$$

де x_k^* – мінімальна точка функції $\varphi(x,r_k)$ для задачі без обмежень.

Крім того, якщо $k>K$, то $r_k<r_K$ і справедлива нерівність

$$\varphi(x, r_k) \leq \varphi(x_K^*, r_K). \tag{2.54}$$

Це випливає з того, що x_k^* мінімізує функцію $\varphi(x,r_k)$ і в будь-якій іншій точці області x, зокрема, в точці x_K^*, функція буде приймати значення, більше ніж $\varphi(x_K^*, r_k)$. Тому при $r_k < r_K$

$$\varphi(x_K^*, r_k) = f(x_K^*) + r_K \sum_{j=1}^{m} \frac{1}{c_j(x_K^*)} > f(x_K^*) + r_k \sum_{j=1}^{m} \frac{1}{c_j(x_K^*)}, \tag{2.55}$$

Отже, $\varphi(x_K^*, r_K) > \varphi(x_K^*, r_k)$. Тоді

$$f(x^*) \leq \varphi(x_k^*, r_k) \leq \varphi(x_K^*, r_k) < \varphi(x_K^*, r_K). \tag{2.56}$$

Оскільки значення x_K^* мінімізує функцію $\varphi(x_K^*, r_K)$, то

$$\varphi(x_K^*, r_K) \leq \varphi(x^*, r_K) = f(x^*) + r_K \sum_{j=1}^{m} \frac{1}{c_j(x^*)}. \tag{2.57}$$

Отже, з рівнянь (2.56) і (2.57) отримаємо

$$f(x^*) \leq \varphi(x_k^*, r_k) \leq f(x^*) + r_K \sum_{j=1}^{m} \frac{1}{c_j(x^*)}. \tag{2.58}$$

З рівняння (2.52) випливає, що

$$f(x^*) + r_k \sum_{j=1}^{m} \frac{1}{c_j(x^*)} \leq f(x^*) + \frac{\varepsilon}{2}. \tag{2.59}$$

Тоді з рівняння (2.53) випливає, що $f(x^*) \leq \varphi(x_k^*, r_k) < f(x^*) + \frac{\varepsilon}{2} + \frac{\varepsilon}{2}$ і

$$\varphi(x_K^*, r_k) - f(x^*) < \varepsilon. \tag{2.60}$$

Оскільки ε може бути обране як завгодно малим, завжди можна знайти таке значення k, при якому $f(x^*) < \varphi(x_k^*, r_k) < f(x^*) + \varepsilon$. Таким чином, при $k \to \infty$ $(r_k \to 0)$

$$\lim_{r_k \to 0} \varphi(x_k^*, r_k) = f(x^*). \tag{2.61}$$

З цього доведення випливає, що при $r_k \to 0$

$$f(x_k^*) \to f(x^*),\ r_k \sum_{j=1}^{m} \frac{1}{c_j(x_k^*)} \to 0. \tag{2.62}$$

Як приклад наведемо доведення того, що $f(x_1^*), f(x_2^*), ..., f(x_k^*)$ утворюють спадаючу послідовність, таку, що

$$f(x_{k+1}^*) < f(x_k^*). \tag{2.63}$$

Очевидно, що якщо функція $f(x)$ опукла, а функція $c_j(x)$ при $j = 1,...,n$ увігнута, то функція $f(x)$ при наявності обмежень має єдиний мінімум.

Приклад 2.4. Розглянемо наступну задачу: мінімізувати функцію $f(x_1, x_2) = \frac{1}{3}(x_1 + 1)^3 + x_2$ при обмеженнях $x_1 - 1 \ge 0, x_2 \ge 0$.

Розв'язок. Подібно до рівняння (2.48) запишемо $\varphi(x, r) = \frac{1}{3}(x_1 + 1)^3 + x_2 + r\left(\frac{1}{x_1 - 1} + \frac{1}{x_2}\right)$. Необхідні умови мінімуму функції φ записуються у вигляді рівнянь $(x_1 + 1)^2 - \frac{r}{(x_1 - 1)^2} = 0$, $1 - \frac{r}{x_2^2} = 0$, що мають наступний розв'язок: $x_1(r) = (1 + \sqrt{r})^{1/2}$, $x_2(r) = \sqrt{r}$. Тоді мінімальним значенням функції $\varphi(x, r)$ буде

$$\varphi^*(r) = \left\{\frac{1}{3}[(1 + \sqrt{r})^{1/2} + 1]^3 + \sqrt{r} + r\left[\frac{1}{\sqrt{r}} + \frac{1}{(1 + \sqrt{r})^{1/2} - 1}\right]\right\} =$$

$$= \left\{\frac{1}{3}[(1 + \sqrt{r})^{1/2} + 1]^3 + \sqrt{r} + r\left[\frac{1}{\sqrt{r}} + \frac{(1 + \sqrt{r})^{1/2} + 1}{\sqrt{r}}\right]\right\} =$$

$$= \{\frac{1}{3}[(1 + \sqrt{r})^{1/2} + 1]^3 + \sqrt{r} + \sqrt{r}[1 + 1 + (1 + \sqrt{r})^{1/2}]\}.$$

Таким чином, при $r \to 0$, $x_1(r) \to 1, x_2(r) \to 0$, $\varphi^*(r) \to f(1;0) = 8/3$.

2.2.4. Обчислювальні складності, що пов'язані зі штрафними функціями

Розв'язок допоміжної задачі буде наближеним до оптимального розв'язку вихідної задачі при виборі μ достатньо великим. Але, якщо вибрати дуже велике μ та спробувати розв'язати допоміжну задачу, то можна зустрітися з деякими обчислювальними труднощами. При великому значенні μ найбільша увага приділяється допустимості біжучої точки, і більшість процедур безумовної оптимізації приведуть до швидкого руху у напрямку до допустимої точки. Однак, навіть якщо ця точка виявляється далеко від оптимальної, може відбутись передчасна зупинка процесу. Для прикладу припустимо, що в процесі оптимізації досягнута допустима точка, в якій

$\alpha(x)=0$. Рух з x вздовж напрямку d може привести в недопустимі точки або в допустимі точки з більшим значенням цільової функції, особливо при наявності нелінійних обмежень-рівностей. В обох випадках значення допоміжної функції $f(x+\lambda d)+\mu\alpha(x+\lambda d)$ більше, ніж $f(x)+\mu\alpha(x)$ для не дуже малих значень довжини кроку λ. В останньому випадку це очевидно. В першому випадку $\alpha(x+\lambda d)>0$ і, так як μ дуже велике, то будь-яке наближене значення f від $f(x)$ до $f(x+\lambda d)$ буде, як правило, компенсоване значенням $\mu\alpha(x+\lambda d)$. Таким чином, покращення можливе лише тоді, коли довжина кроку λ є дуже мала, так щоб член $\mu\alpha(x+\lambda d)$ був достатньо малий, незважаючи на те, що μ дуже велике. В цьому випадку покращення $f(x+\lambda d)$ у порівнянні з $f(x)$ може компенсувати те, що $\mu\alpha(x+\lambda d)>0$. Необхідність використання дуже малих кроків може привести до повільної збіжності та передчасної зупинки.

Алгоритм методу штрафних функцій. Враховуючи складності, пов'язані з використанням великого параметру штрафу, більшість алгоритмів штрафних функцій застосовують послідовність зростаючих параметрів. Для кожного нового значення параметру штрафу використовується алгоритм оптимізації, який починається із оптимального розв'язку, що відповідає попередньому значенню параметра. Нижче приведений алгоритм методу штрафних функці для розв'язування задачі мінімізації $f(x)$ при умовах $g(x)\le 0$, $h(x)=0, x\in X$. Використовується штравна функція α у вигляді:

$$\alpha(x)=\sum_{i=1}^{m}\varphi[g_i(x)]+\sum_{i=1}^{l}\psi[h_i(x)], \tag{2.64}$$

де φ та ψ – неперервні функции, які задовольняють умовам:

$$\varphi(y)=0, \text{ якщо } y\le 0, \text{ та } \varphi>0, \text{ якщо } y>0,$$
$$\psi(x)=0, \text{ якщо } y=0, \text{ та } \psi(y)>0, \text{ якщо } y\ne 0.$$

Цей метод не накладає будь-яких обмежень на f, g та h, окрім неперервності. Але він може ефективно використовуватися лише в тих випадках, коли є ефективна процедура розв'язування задачі на кроці 1 основного етапу.

Початковий етап. Обрати $\varepsilon>0$, як критерій зупинки. Обрати початкову точку x_1, штрафний параметр $\mu_1>0$ та число $\beta>0$. Взяти k=1 та перейти до основного етапу.

Основний етап.

Крок 1. При початковій точці x_k розв'язати наступну задачу: мінімізувати $f(x)+\mu_k\alpha(x)$ при умові $x\in X$. Взяти x_{k+1} рівне оптимальному розв'язку цієї задачі та перейти до кроку 2.

Крок 2. Якщо $\mu_k\alpha(x_{k+1})<\varepsilon$, то зупинитися, в іншому випадку взяти $\mu_{k+1}=\beta\mu_k$, замінити k на k+1 та перейти до кроку 1.

Запишемо наш алгоритм на мові програмування С:

```
#include <iostream.h>
#include <graphics.h>
#include <conio.h>
#include <stdlib.h>
#include <stdio.h>
#include <math.h>
int n=3;
void main(void){int i,j,k,s;n=3;
float r[3]={1,0.25,0.01};
float x,v,p,d[3],y1[3],y2[3],x1[3],x2[3];
 for(;;){ clrscr();
 cout<<"Vykorystanja schtrafnyh funkcij dla optymizaciji kryterijiv z
obmegennjamy!!!"<<endl<<endl;
cout<<"1. Funkcia f(x)=x pry x-2>=0.\n";
cout<<"2. Funkcia f(x1,x2)=1/3(x1+1)^3+x2 pry x1-1>=0, x2>=0.\n";
cout<<"\n\nInput number: ";
cin>>s;switch(s){
```

Приклад 2.5. Використовуючи штрафну функцію, яка задана рівнянням (2.47), мінімізувати функцію $f(x)=x$ при обмеженнях $x\geq 2$, тобто $x-2\geq 0$. Мінімальним значенням функції є 2 при $x=2$. Як за допомогою штрафної функції можна знайти розв'язок? Розглянемо функцію $\varphi(x,r)=x+\dfrac{r}{x-2}$.

```
case 1:{clrscr();
cout<<"Funkcia F(x)-x pry x-2>=0.\n\n";
cout<<"v(x,r)=x+r/(x-2)\n";
for(i=0;i<n;i++){
d[i]=4-4*r[i];
y1[i]=(2+sqrt(r[i])*2)/1;
cout<<"\n "<<i+1<<". pry r="<<r[i];
   if(y1[i]>=0) cout<<"   v(x)="<<y1[i];
}
cout<<"\n\n\nPoshchatkove x pry f(x)=0 - bilogo koljoru!!";
getch();getch();
int gdriver = DETECT, gmode, errorcode; /* request auto detection */
int xmax, ymax;
initgraph(&gdriver, &gmode, "");/* initialize graphics and local variables */
errorcode = graphresult();/* read result of initialization */
if (errorcode != grOk){ /* an error occurred */
   printf("Graphics error: %s\n", grapherrormsg(errorcode));
   printf("Press any key to halt:");
   getch();   exit(1);
}
setcolor(getmaxcolor());
xmax = getmaxx();
ymax = getmaxy();
line(0, ymax, xmax,ymax); /* draw a diagonal line */
line(0,0,0,ymax);
line (200,0,200,ymax);
  for(i=0;i<n;i++){
  if(y1[i]>0){setcolor(i+1);line(100*y1[i],0,100*y1[i],ymax);}
  }
getch();closegraph();/* clean up */
```

Програма подає графік функції $\varphi(x,r)$ і визначає положення точок функції мінімуму для різних r $(1;0{,}25;i\,0{,}01)$. Область обмежень лежить справа від вертикальної прямої $x=2$. Послідовність точок Q_1,Q_2,Q_3 прямує до точки Q – мінімуму функції при наявності обмежень. Дійсно, знайдемо мінімум функції $\varphi(x,r)$ методом штрафних функцій: $\dfrac{d\varphi}{dx}=1-\dfrac{r}{(x-2)^2}$. Відповідно, якщо $d_\varphi/dx=0$, $(x-2)^2=r$,

то $x = 2^{+}\sqrt{r}$. Тоді $\frac{d^2\varphi}{dx^2} = \frac{2r}{(x-2)^3}$ і мінімум досягається при $x = 2+\sqrt{r}$ всередині області обмежень.

```
cout<<"dv/dx=1-r/(x-2)^2 and d^2v/dx^2=2r/(x-2)^3\n\n";
cout<<"REZULTAT";
for(i=0;i<n;i++){
x1[i]=2+sqrt(r[i]);
cout<<"\n "<<i+1<<". pry r="<<r[i];
x=2+sqrt(r[i]);
cout<<"   x="<<x;
 }
 cout<<"\n\n\nPoshchatkove x pry f(x)=0 - bilogo koljoru!!";getch();  getch();
 /* request auto detection */ /* initialize graphics and local variables */
initgraph(&gdriver, &gmode, "");
errorcode = graphresult();/* read result of initialization */
if (errorcode != grOk){ /* an error occurred */
   printf("Graphics error: %s\n", grapherrormsg(errorcode));
   printf("Press any key to halt:");
   getch();   exit(1);
}
setcolor(getmaxcolor());
xmax = getmaxx();ymax = getmaxy();
line(0, ymax, xmax,ymax); /* draw a diagonal line */
line(0,0,0,ymax);line (200,0,200,ymax);
  for(i=0;i<n;i++){
  if(x1[i]>0){setcolor(i+1);line(100*x1[i],0,100*x1[i],ymax);}
  }
getch();closegraph();/* clean up */
```

Отже, функція $\varphi(x,r)$ має мінімум, що рівний $2+2\sqrt{r}$ при $x = 2+\sqrt{r}$. Тоді Q_1 є точкою з координатами (3;4); Q_2 – точка з координатами (2,5;3); Q_3 – точка з координатами (2,1;2,2). Зрозуміло, що при $r \to 0$ мінімум без обмежень функції $\varphi(x,r)$ наближається до значення 2 і мінімальною точкою є точка x=2.

В загальному випадку неможливо аналітично визначити положення мінімуму функції $\varphi(x,r)$, якщо розглядати її як звичайну функцію від r. Для його визначення необхідно звернутись до чисельних методів. Слід відзначити, що якщо функція мети $f(x)$ опукла, а функція $c_j(x)$ увігнута, то функція $\varphi(x,r)$, яка задана рівнянням (2.48), також є опуклою функцією у області обмежень, яка сама є опуклою. Отже, $\varphi(x,r)$ має для даного значення r єдиний мінімум.

```
cout<<"REZULTAT:"<<endl;
for(i=0;i<n;i++){
cout<<endl<<"v"<<i+1<<"(x)="<<y1[i];
cout<<"   pry x="<<x1[i];
}
   for(i=0;i<n;i++)cout<<endl<<"Q"<<i+1<<"("<<x1[i]<<";"<<y1[i]<<")";
break;
}
```

Приклад 2.6. Мінімізувати функцію $f(x_1,x_2) = \frac{1}{3}(x_1+1)^3 + x_2$ при $x_1 - 1 \ge 0, x_2 \ge 0$.

Розв'язок. По аналогії з рівнянням (2.48) запишемо $\varphi(x,r) = \frac{1}{3}(x_1+1)^3 + x_2 + r\left(\frac{1}{x_1-1} + \frac{1}{x_2}\right)$. Необхідні умови мінімуму функції φ записуються у

вигляді рівнянь $(x_1+1)^2-\frac{r}{(x_1-1)^2}=0,\ 1-\frac{r}{x_2^2}=0,$ що мають наступний розв'язок: $x_1(r)=(1+\sqrt{r})^{1/2},\ x_2(r)=\sqrt{r}$. Тоді мінімальним значенням функції $\varphi(x,r)$ буде

$$\varphi^*(r)=\left\{\frac{1}{3}[(1+\sqrt{r})^{1/2}+1]^3+\sqrt{r}+r\left[\frac{1}{\sqrt{r}}+\frac{1}{(1+\sqrt{r})^{1/2}-1}\right]\right\}=$$

$$=\left\{\frac{1}{3}[(1+\sqrt{r})^{1/2}+1]^3+\sqrt{r}+r\left[\frac{1}{\sqrt{r}}+\frac{(1+\sqrt{r})^{1/2}+1}{\sqrt{r}}\right]\right\}=$$

$$=\{\frac{1}{3}[(1+\sqrt{r})^{1/2}+1]^3+\sqrt{r}+\sqrt{r}[1+1+(1+\sqrt{r})^{1/2}]\}.$$

Таким чином, при $r\to 0$, $x_1(r)\to 1, x_2(r)\to 0$, $\varphi^*(r)\to f(1;0)=8/3$.

```
case 2:{clrscr();
cout<<"Funkcia f(x1,x2)=1/3(x1+1)^3+x2 pry x1-1>=0, x2>=0.\n\n";
cout<<"v(x,r)=1/3(x1+1)^3+x2+r(1/(x1-1)+1/x2)\n";
for(i=0;i<n;i++){
d[i]=4-4*r[i];
y1[i]=pow((1+sqrt(r[i])),1/2);//(2+sqrt(r[i])*2)/1;
y2[i]=sqrt(r[i]);
cout<<"\n "<<i+1<<". pry r="<<r[i];
cout<<"   v(x1)="<<y1[i]<<"   v(x2)="<<y2[i];
}
cout<<"\n\n\nPoshchatkovi x1 nf x2 pry f(x)=0 - bilogo koljoru!!";
getch();getch();
int gdriver = DETECT, gmode, errorcode; /* request auto detection */
int xmax, ymax;
initgraph(&gdriver, &gmode, "");/* initialize graphics and local variables */
errorcode = graphresult();/* read result of initialization */
if (errorcode != grOk){ /* an error occurred */
   printf("Graphics error: %s\n", grapherrormsg(errorcode));
   printf("Press any key to halt:");
   getch(); exit(1);
}
setcolor(getmaxcolor());
xmax = getmaxx();
ymax = getmaxy();
line(0, ymax, xmax,ymax); /* draw a diagonal line */
line(0,0,0,ymax);
line (100,0,100,ymax);
  for(i=0;i<n;i++){
  if(y1[i]>0){setcolor(i+1);line(100*y1[i],0,100*y1[i],ymax);}
  if(y2[i]>0){setcolor(i+5);line(100*y2[i],0,100*y2[i],ymax);}
  }
getch();
closegraph();/* clean up */
cout<<"dv(x)/dx={1/3*[(1+r^(1/2))^(1/2)+1]^3+r^(1/2)+r^(1/2)*(1+1+(1+r^(1/2))^(1/2))}\
n\n";
cout<<"REZULTAT: r->0  x1(r)->1  x2(r)->0  v(r)->f(0;1)=8/3";
for(i=0;i<n;i++){
p=pow(r[i],1/2);
v=pow((1+p),1/2);
x1[i]=1/3*pow((v+1),3)+p+p*(1+1+v);
cout<<"\n "<<i+1<<". pry r="<<r[i];
cout<<"   x="<<x1[i];
}
 cout<<"\n\n\nPoshchatkove x pry f(x)=0 - bilogo koljoru!!";getch();  getch();
 /* request auto detection *//* initialize graphics and local variables */
initgraph(&gdriver, &gmode, "");
```

```
errorcode = graphresult();/* read result of initialization */
if (errorcode != grOk){ /* an error occurred */
   printf("Graphics error: %s\n", grapherrormsg(errorcode));
   printf("Press any key to halt:");
   getch();   exit(1);
}
setcolor(getmaxcolor());
xmax = getmaxx();
ymax = getmaxy();
line(0, ymax, xmax,ymax); /* draw a diagonal line */
line(0,0,0,ymax);
line (100,0,100,ymax);
  for(i=0;i<n;i++){
  if(x1[i]>0){setcolor(i+1);line(100*x1[i],0,100*x1[i],ymax);}
  }
getch();closegraph();/* clean up */
cout<<"REZULTAT:"<<endl;
for(i=0;i<n;i++){
cout<<endl<<"v"<<i+1<<"(x1)="<<y1[i]<<"  v"<<i+1<<"(x2)="<<y2[i];
cout<<"   pry F(x1,x2)="<<x1[i];
}
for(i=0;i<n;i++)cout<<endl<<"Q"<<i+1<<"{"<<x1[i]<<";"<<y1[i]<<";"<<y2[i]<<"}";
getch(); break;}
default :{ cout <<"\nError!!!\n";break;}
    }
cout<<"\n\n\n1.Continue"<<endl<<"2.Exit"<<endl<<endl<<"Input number:";
cin>>s;
if(s==2)exit(1);
    }
}
```

2.2.5. Аналіз та інтерпретація отриманих результатів

Приклад 2.7. Мінімізувати $(x_1-2)^4+(x_1-2x_2)^2$ при $x_1^2-x_2=0$, $x\in X=E_2$.

Розв'язок. Зауважимо, що на k-ій ітерації при заданому значенні параметру штрафу μ_k для отримання x_{μ_k} повинна розв'язуватись наступна задача: мінімізувати $(x_1-2)^4+(x_1-2x_2)^2+\mu_k(x_1^2-x_2)^2$ при умові $x\in E_2$. В таблиці 2.4 наведені результати розрахунків за методом штрафних функцій. За початкову взяли точку x=(2.0,1.0), в якій значення цільової функції рівне 0.0. За початкове значення параметру штрафу взято число $\mu_1=0.1$, а число β вибране рівним 10.0. Зауважимо, що $f(x_{\mu_k})$ і $\theta(\mu_k)$ – зростаючі функції, а $\alpha(x_{\mu_k})$ – не зростаюча. Процес був зупинений після 5-ти ітерацій, де $\alpha(x_{\mu_k})=0.0000028$. Але, щоб показати , що $\mu_k\alpha(x_{\mu_k})$ збігається до нуля, була виконана ще одна ітерація (рис. 2.3).

Таблиця 2.4

Результати розрахунків методом штрафних функцій

k	μ_k	$x_{k+1}=x_{\mu_k}$	$f(x_{k+1})$	$\alpha(x_{\mu_k})=h^2(x_{\mu_k})$	$\theta(\mu_k)$	$\mu_k\alpha(x_{\mu_k})$
1	0.1	(1.4539,0.7608)	0.0935	1.8307	0.2766	0.1831
2	1.0	(1.1687,0.7407)	0.5753	0.3908	0.9661	0.3908
3	10.0	(0.9906,0.8425)	1.5203	0.01926	1.7129	0.1926
4	100.0	(0.9507,0.8875)	1.8917	0.000267	1.9184	0.0267
5	10^4	(0.9461,0.8934)	1.9405	0.0000028	1.9433	0.0028
6	10^5	(0.9231,09124)	1.9745	0.00000016	1.9723	0.00024

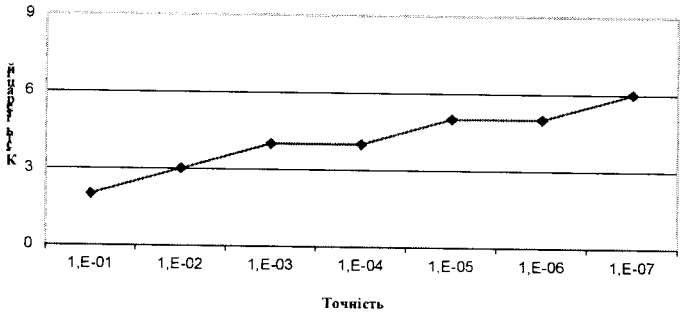


Рис. 2.3. Графічна інтерпретація отриманих результатів

2.3. Багатокритерійні задачі оптимізації та основні підходи до їх розв'язування

2.3.1. Основні поняття багатокритерійної задачі оптимізації

На практиці задачі, які не мають невизначеностей, швидше є винятком, ніж правилом. Поряд із розглянутими раніше існує ще один важливий тип невизначеності – невизначеність мети, що виявляється у наявності декількох, в більшості випадків незбіжних аспектів оцінки якості того чи іншого розв'язку з множини допустимих. У формальному вигляді аспекти оцінки якості відображаються за допомогою множини критеріїв. Таким чином виникає багатокритерійна задача дослідження операцій, загальний вигляд якої наступний: $Q_1(a,x) \Rightarrow Max$, $Q_2(a,x) \Rightarrow Max, ...,$ $Q_n(a,x) \Rightarrow Max$, $x \in X$. Знайти розв`язок, який одночасно був би найкращим за всіма критеріями, неможливо, тому що в загальному випадку покращення значення одного з критеріїв приводить до погіршення значення іншого.

Приклад 2.8. Купівля автомобіля. Важливими критеріями в цьому випадку є питомі витрати пального, максимальна швидкість, комфортність, прохідність, зовнішній вигляд, вартість, гарантія, наявність сервісного обслуговування, ціна. І звичайно, придбати авто, яке б витрачало найменше пального, було б при цьому найдешевшим, найкомфортнішим та найшвидшим водночас неможливо.

Приклад 2.9. Вибір місця зупинки для мандрівників у горах. Важливими критеріями є: наявність джерела або річки, наявність галявини, можливість купання, відсутність комарів, близькість до населених пунктів, відсутність хижих звірів.

2.3.2. Геометрична ілюстрація багатокритерійної задачі

Розглянемо загальну задачу оптимізації за двома критеріями з двома змінними:

$$Q_1(x_1,x_2) \Rightarrow Max,\ Q_2(x_1,x_2) \Rightarrow Max,\ x=(x_1,x_2),\ x \in X. \qquad (2.65)$$

Зобразимо область допустимих розв'язків у просторі змінних (x_1, x_2). Значення критеріїв Q_1, Q_2, відображатимемо у просторі критеріїв (Q_1, Q_2). Кожній конкретній точці множини допустимих розв'язків $\left(x_1^{(k)}, x_2^{(k)}\right)$ відповідатиме одне і лише одне значення кожного з критеріїв $Q_1\left(x_1^{(k)}, x_2^{(k)}\right)$, $Q_2\left(x_1^{(k)}, x_2^{(k)}\right)$, хоча обернене твердження не завжди буде відповідати дійсності (декілька розв'язків можуть бути рівноцінними з точки зору значень критеріїв), тобто відповідне відображення буде гомоморфним. Здійснивши таку операцію для всіх точок допустимої області в просторі змінних, отримаємо її образ в просторі критеріїв. На рис. 2.4 розв'язки 4 та 5 відображаються в одну й ту ж саму точку в просторі критеріїв, тобто є ідентичними з точки зору їх якості. Крім того, вони є гіршими, ніж розв'язки 2 та 3, у яких значення кожного з критеріїв є більші, ніж у 2 та 3. Розв'язки 1, 2, та 3 не є порівняльними, тобто без додаткової інформації неможливо визначити, який з них є кращим – значення за одним з критеріїв для них є більші, а за іншим – менші. В той же час, аналізуючи розв'язки, що знаходяться на кривій A-B-C можна зробити висновок, що вони є множиною *найкращих* розв'язків: для будь-якого іншого розв'язку з множини допустимих завжди знайдеться хоча б один з розв'язків, що знаходяться на A-B-C, та кращий за нього (тобто такий, що домінує). Таким чином розв'язки, що лежать на A-B-C, не домінують над ніякими іншими розв'язками, що належать до допустимої області.

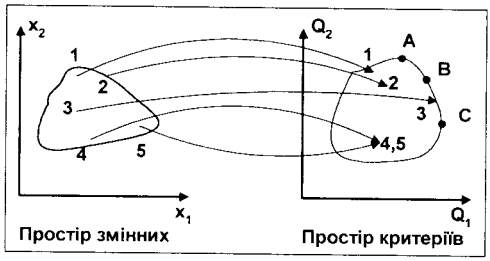


Рис. 2.4. Відображення допустимої області в простір критеріїв.

Множина недомінованих розв'язків багатокритерійної задачі називається *множиною Парето-оптимальних розв'язків* (Вільфредо Парето одним з перших досліджував задачі такого типу) і є таким чином в загальному випадку розв'язком багатокритеріальної задачі. В свою чергу розв'язок належить до множини Парето-оптимальних, якщо він не домінується ніяким іншим. Побудова множини Парето-оптимальних розв'язків для задач ДО в більшості випадків є неможливою внаслідок значних обчислювальних труднощів. Крім того, в більшості випадків завдання полягає в знаходженні одного розв'язку. Такий розв'язок повинен належати до множини Парето-оптимальних, а ось яким він повинен бути, може виявитися лише в процесі його побудови. Тому були розроблені та застосовуються на практиці цілий ряд методів, деякі з них розглядаються далі.

При оптимізації реального процесу параметри і (або) змінні пов'язані фізичними законами, такими, як закон збереження маси чи енергії, котрі повинні бути включені в задачу нелінійного програмування як обмеження у вигляді рівностей. Таким чином, одна група обмежень складається з функціональних зв'язків, які слід врахувати, щоб оптимізація була фізично здійсненною. Друга група обмежень включає задані граничні

діапазони значень змінних або параметрів, що забезпечують їх фізичну можливість реалізації і сумісності з даним процесом. До цієї групи відносяться обмеження у вигляді нерівностей. Замінити ці обмеження або додаватись до них можуть емпіричні зв'язки, як правило, виражені рівностями. Для спрощення рівнянь моделі процесу часто вводять нові змінні, які утворюють додаткові обмеження у вигляді рівностей в тому випадку, коли вихідні рівняння не можуть бути розв'язані в явному вигляді відносно визначеної змінної. На рис. 2.5 показані зв'язки між різними частинами задачі нелінійного програмування і алгоритмом розв'язку. Кожне обмеження у вигляді рівності зменшує на одиницю число степенів свободи в моделі процесу і приводить до появи ще однієї залежної змінної. Переважно передбачається, що рівняння моделі процесу записуються достатньо акуратно, так що всі рівності є незалежними; якщо ж через неуважність чи помилки в задачу включаються два надлишкових, або, іншими словами, залежних, рівняння, то число степенів свободи буде відрізнятися від дійсного. Число залишкових степенів свободи повинно відповідати числу незалежних змінних. Воно є важливим поняттям в задачі оптимізації довільного типу при обмеженнях у вигляді рівностей, бо якщо число змінних рівне числу незалежних обмежень у вигляді рівностей, то оптимізувати немає що – значення всіх змінних можна безпосередньо визначити з розв'язку системи обмежень у вигляді рівностей. Якщо число змінних перевищує число незалежних обмежень у вигляді рівностей, то єдиний спосіб розв'язування, прийнятний для даної моделі процесу – це варіювання незалежних змінних, поки цільова функція не досягне свого оптимального значення. Слід відзначити, що змінні являють собою той механізм, за допомогою якого здійснюється передача інформації між цільовою функцією та заданими обмеженнями. В цільову функцію можуть входити як залежні, так і оптимізуючи змінні.

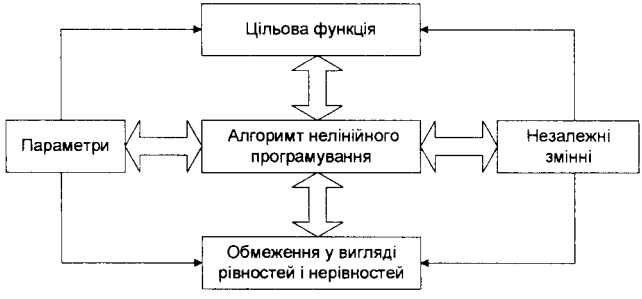


Рис.2.5. Співвідношення між окремими частинами задачі нелінійного програмування і алгоритмом її розв'язування

Для розв'язування наведеної задачі існує багато методів, вибір кожного з яких визначається конкретним змістом задачі та досвідом дослідника. Одним із цих методів є метод Хука-Джівса, який придатний для розв'язування задачі нелінійного програмування в її загальній постановці, тобто у випадку обмежень як у вигляді рівнянь, так і у вигляді нерівностей. Саме цей метод реалізований програмно. Детальний опис цього методу з теоретичними обґрунтуваннями його процедур і кроків подано нижче.

2.3.3. Постановка задачі багатокритерійної оптимізації

Задача багатокритеріальної оптимізації формулюється таким чином:

$$\min_{\vec{x}}\{f_1(\vec{x}), f_2(\vec{x}),..., f_k(\vec{x})\}, \vec{x} \in S. \tag{2.66}$$

де $f_i : R^n \to R$ – це k $(k \geq 2)$ *цільових функцій*. Вектори розв'язків $\vec{x} = (x_1, x_2,..., x_n)^T$ належать до непорожньої області визначення S.

Задача багатокритеріальної оптимізації полягає у пошуку вектора цільових змінних, який задовольняє накладені обмеження та оптимізує векторну функцію, елементи якої відповідають цільовим функціям. Ці функції утворюють математичний опис критерію задовільності та, зазвичай, взаємно конфліктують. Звідси, *оптимізувати* означає знайти такий розв'язок, коли значення цільових функцій були б прийнятними для постановника задачі.

2.3.4. Еталонні точки

Для можливості оцінки якості знайдених розв'язків, зазвичай розглядають такі точки в області значення цільової функції: ідеальна точка (y^I), утопічна точка (y^U) та надір (надир, y^N). У деяких випадках ці точки можуть бути розв'язками.

Ідеальна точка визначається як вектор $y^I = (y_1^I,..., y_p^I)$, кожна з координат якого має оптимальне значення відповідної складової цільової функції:

$$y_k^I = \min_{x \in X} f_k(x) = \min_{y \in Y} y_k. \tag{2.67}$$

Точка *надіру* $y^N = (y_1^N,..., y_p^N)$ визначається як вектор:

$$y_k^N = \max_{x \in X_E} y_k(x) = \max_{y \in Y_N} y_k, \; k = \overline{1,p}. \tag{2.68}$$

Утопічну точку y^U обчислюють на основі ідеальної:

$$y^U = y^I - eU, \tag{2.69}$$

де $e > 0$, U – одиничний вектор.

2.3.5. Критерії оптимальності. Критерій Парето

Вектор розв'язку $\vec{x} \in S$ називається *оптимальним за Парето*, якщо не існує $\vec{x}' \in S$ такого, що $f_i(\vec{x}) \leq f_i(\vec{x}')$ для всіх $i = \overline{1,k}$ та $f_i(\vec{x}) < f_i(\vec{x}')$ хоч для одного i. Множину оптимальних за Парето розв'язків можна позначити як $P(S)$. Цільовий вектор є оптимальним за Парето, якщо відповідний йому вектор з області визначення також оптимальний за Парето. Множину оптимальних за Парето цільових векторів можна позначити як $P(Z)$. Множина оптимальних за Парето векторів є підмножиною оптимальних за Парето в слабкому сенсі векторів. Вектор $\vec{x}' \in S$ є слабким оптимумом

за Парето тоді, коли не існує вектора $\vec{x} \in S$ такого, що $f_i(\vec{x}) < f_i(\vec{x}')$ для всіх $i = \overline{1,k}$. Діапазон значень оптимальних за Парето розв'язків в області допустимих значень дає корисну інформацію про досліджувану задачу, якщо цільові функції обмежено областю визначення. Нижні границі оптимальної за Парето множини представлено в «ідеальному цільовому векторі» $\vec{z} \in R^k$. Його компоненти z_i отримані шляхом мінімізації кожної цільової функції у межах області визначення. Множину оптимальних за Парето розв'язків також називають *Парето-фронтом* (англ. *Pareto-frontier*).

2.3.6. Лексикографічний порядок

Якщо одні цільові функції важливіші за інші, критерій оптимальності можна визначити за лексикографічним порядком.

Відношення лексикографічного порядку $<_{lex}$ між векторами $\vec{a}$ та $\vec{b}$ виконується, якщо $a_q < b_q$, де $q = \min\{k : a_k \neq b_k\}$. Тобто, перші q компонент вектора $\vec{a}$ менші за компоненти вектора $\vec{b}$, а компоненти $q+1$ – рівні (якщо такі є). Лексикографічний порядок для випадку дійсних чисел є лінійним. Вектор $\vec{x} \in X$ є лексикографічним розв'язком, якщо не існує вектора $\vec{x}' \in X$ такого, що $f(\vec{x}') <_{lex} f(\vec{x})$. Оскільки відношення лексикографічного порядку є лінійним, можна довести, що вектор $\vec{x}$ є лексикографічним розв'язком, якщо для всіх $\vec{x}' \in X$ виконується:

$$f(\vec{x}) <_{lex} f(\vec{x}'). \tag{2.70}$$

Основною особливістю розв'язків за лексикографічним порядком є існування вибору між критеріями. Лексикографічна впорядкованість вимагає ранжування критеріїв в тому сенсі, що оптимізація за критерієм f_k можлива лише тоді, коли було досягнуто оптимуму для попередніх критеріїв. Це означає, що перший критерій має найбільший пріоритет, і лише у випадку існування декількох розв'язків за цим критерієм, буде пошук розв'язків за другим та рештою критеріїв.

Існування ієрархії серед критеріїв, дозволяє розв'язувати лексикографічні задачі послідовно, крок за кроком мінімізуючи за кожним наступним критерієм, та використовуючи оптимальні значення попередніх критеріїв як обмеження.

2.3.7. Скаляризація

Для отримання оптимальних за Парето розв'язків часто використовують *методи скаляризації*. Оскільки цільова функція задачі багатокритеріальної оптимізації має векторні значення, її перетворюють на функцію зі скалярним значенням. Таким чином, задача багатокритеріальної оптимізації зводиться до задачі оптимізації з однією скалярною цільовою функцією. Функція скаляризації має задовольняти наступним умовам. Нехай F – функція скаляризації, що перетворює векторну функцію $\vec{y} = f(\vec{x})$ на скалярну. Якщо F зберігає впорядкованість за Парето $\vec{y}$, тобто, якщо для довільних $\vec{y}^1, \vec{y}^2 \subset \vec{f}(X)$ виконується:

$$\vec{y}^1 \leq \vec{y}^2 \Rightarrow F(\vec{y}^1) < F(\vec{y}). \tag{2.71}$$

тоді розв'язок $\vec{x}^0$, що мінімізує F на X є розв'язком за Парето.

Якщо F зберігає відношення порядку $<$ в $\vec{y}$, тобто, якщо для довільних $\vec{y}^1, \vec{y}^2 \in \vec{f}(X)$ виконується:

$$\vec{y}^1 < \vec{y}^2 \Rightarrow F(\vec{y}^1) < F(\vec{y}), \tag{2.72}$$

тоді розв'язок $\vec{x}^0$, що мінімізує F на X є слабким за Парето. Якщо F неперервна на $\vec{y}$, та $\vec{x}^0$ єдина точка мінімуму F на X, тоді $\vec{x}^0$ є розв'язком за Парето.

2.3.8. Зважена сума

Наведена в (2.73) функція F_1 зберігає впорядкованість за Парето для $\omega > 0$.

$$F_1(\vec{f}(\vec{x})) = \omega_1 f_1(\vec{x}) + \ldots + \omega_\gamma f_\gamma(\vec{x}). \tag{2.73}$$

Тому розв'язки, що мінімізують F_1 на X для довільних $\omega > 0$ є оптимальними за Парето. Однак, F_1 не зберігає впорядкованість за Парето для $\omega \geq 0$, а зберігає лише відношення $<$, і тому розв'язки, що мінімізують F_1 на X для $\omega \geq 0$ є слабкими за Парето. *Недоліком методу зважених сум у випадку не опуклої множини значень цільових функцій є неможливість охопити всі оптимальні за Парето точки з множини Парето-фронту. В задачах комбінаторної багатокритеріальної оптимізації множина* цільових значень не є опуклою, тому метод зважених сум не підходить для скаляризації цільових функцій для цих задач.

2.3.9. Функція скаляризації Чебишева

Зважена функція скаляризації Чебишева зберігає відношення $<$, і тому мінімум F_∞ є слабким за Парето.

$$F_\infty(\vec{f}(\vec{x})) = \max_{1 \leq i \leq r} \omega_i f_i(\vec{x}). \tag{2.74}$$

2.3.10. Метод зміни обмежень (ε-обмеження)

За методом зміни обмежень одну з цільових функцій залишають як цільову, а решту перетворюють на обмеження. Тобто, нехай f_r буде цільовою, а решта $f_1, \ldots, f_{r-1}$ як обмеження нерівності:

$$\min_x f_r(\vec{x}), \tag{2.75}$$

за умов $f_i(\vec{x}) \leq \varepsilon_i$, $i = \overline{1, r-1}$, $\vec{x} \in X$.

Значення $\varepsilon_1, \ldots, \varepsilon_{r-1}$ можуть розглядатись як *допустимі рівні* для $f_1, \ldots, f_{r-1}$.

2.3.11. Методи розв'язування. Інтерактивність. Еволюційні методи

Часто, розв'язування задачі багатокритеріальної оптимізації відбувається за участю експерта – людини, яка обирає та ухвалює рішення на основі інформації, яка представлена системою підтримки прийняття рішень. Можлива участь групи з декількох експертів. У випадку участі людини у пошуку розв'язку алгоритми та методи називають інтерактивними. Згадки про застосування генетичних алгоритмів для розв'язування задачі багатокритеріальної оптимізації стоуються кінця 1960-тих років.

2.4. Лінійне програмування

2.4.1. Місце лінійного програмування в системі задач математичного програмування

Термін *лінійне програмування* (ЛП) з'явився вперше в 1951р. в працях американських вчених Дж.Данціга та Т.Купманса. Задачі ЛП входять як важлива складова до більш широкого класу задач *математичного програмування* (МП, рис. 2.6).

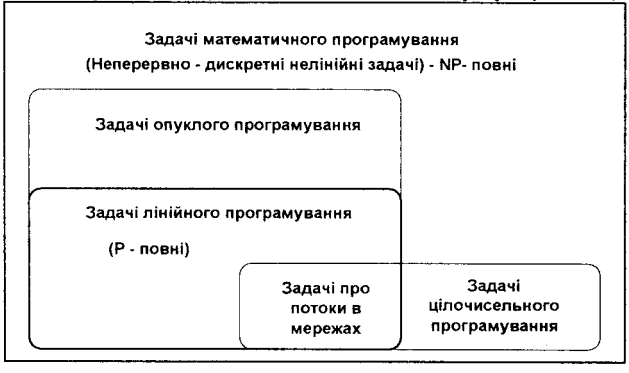


Рис.2.6. Класифікація задач математичного програмування

В загальному задача МП подається наступним чином:

$$f(x) \Rightarrow Max,\ g(x) \le 0,\ h(x) = 0\ . \tag{2.76}$$

$$x = (x_1,\ldots,x_n),\ g = (g_1(x),\ldots,g_m(x)),\ h = (h_1(x),\ldots,h_l(x))\ . \tag{2.77}$$

Задачі максимізації та мінімізації еквівалентні, також це ж має місце і щодо знаків $\le, \ge$. Задачі МП за складністю поділяються на P- та NP- повні. Природнім є прагнення розвинути більш ефективні методи розв'язування якомога ширших класів задач. Однак, як доводить багаторічний досвід, загальність методу та його ефективність знаходяться в певному антагонізмі. Разом з тим дуже важливо знати, чи можна сподіватись на створення достатньо загальних та ефективних методів, чи необхідно свідомо розбивати задачі на більш вузькі класи та, використовуючи їх специфіку, розробляти для них ефективні алгоритми.

Більшість з дискретних та комбінаторних задач МП можуть бути в принципі розв'язані за допомогою повного перебору варіантів. Однак, число кроків методу перебору зростає експоненційно в залежності від розміру задачі. Для деяких задач такого типу можна побудувати ефективні методи розв'язування, однак кількість їх невелика. Аналіз труднощів синтезу ефективних алгоритмів розв'язування задач привів до постановки центральної проблеми дискретної математики – пошуку відповіді на запитання: "Чи можна виключити перебір при розв'язуванні дискретних задач?". Зараз ця проблема є відкритою.

Аналогом алгоритмічної нерозв'язуваності в скінченій області є перебір експоненційного числа варіантів, а аналогом алгоритмічної розв'язуваності – існування суттєво більш економічного алгоритму, ніж перебір. Загально прийнятним вважається, що задача перебірного типу розв'язується ефективно, якщо наявний алгоритм, що розв'язує її за час, що обмежений поліномом від "розміру задачі".

Вважають, що перебірна задача Π_1 зводиться до перебірної задачі Π_2, якщо метод розв'язування задачі Π_2 можна перетворити в метод розв'язування задачі Π_1. Приведення є поліноміальним, якщо його можна здійснити за поліноміальний час. Головними об'єктами є клас NP – всіх перебірних задач та клас P – перебірних задач, які можуть бути розв'язані за поліноміальний час машиною Т'юринга. Очевидно, що $P \subseteq NP$. Центральним є питання про те, чи тотожні ці класи. В класі NP виявлені, так звані, універсальні (NP-повні) задачі, до яких поліноміально зводиться довільна задача з NP, тобто вони є еталоном складності. Якщо б вдалося довести, що хоча б одна з NP-повних задач належить до класу P, то це означало б, що $P = NP$, і можна було б сподіватися на побудову ефективних алгоритмів для різних класів дискретних задач. Хоча практичний досвід розв'язування дискретних задач дозволяє вважати, що задачі цих двох класів значною мірою розрізняються за складністю, в строгому сенсі ця різниця не доведена. Важливий результат математика Л. Хачіяна – доведення факту P-повноти задач лінійного програмування.

2.4.2. Формальна постановка задачі лінійного програмування

В загальному вигляді задача ЛП формулюється наступним чином:

$$Q(x) = \sum_{j=1}^{n} c_j x_j \Rightarrow Max\,(Min). \tag{2.78}$$

$$\sum_{j=1}^{n} a_{ij} x_j (\leq, =, \geq) b_i,\ i = \overline{1,m},\ j = \overline{1,n},\ x_j (\leq, =, \geq) 0. \tag{2.79}$$

Задача ЛП в канонічній формі є наступною:

$$Q(x) = \sum_{j=1}^{n} c_j x_j \Rightarrow Max\ . \tag{2.80}$$

$$\sum_{j=1}^{n} a_{ij} x_j = b_i,\ i = \overline{1,m},\ j = \overline{1,n},\ x_j \geq 0. \tag{2.81}$$

Задачу ЛП в загальній формі завжди можна перетворити в задачу в канонічній формі за допомогою наступних підстановок:

1) $Q(x) \Rightarrow Min$, $\quad -Q(x) \Rightarrow Max$,

2) $\sum_{j=1}^{n} a_{ij} x_j \le b_i$, $\quad \sum_{j=1}^{n} a_{ij} x_j + s_i = b_i$,

3) $\sum_{j=1}^{n} a_{ij} x_j \ge b_i$, $\quad \sum_{j=1}^{n} a_{ij} x_j - s_i = b_i$,

4) $x_j \le 0$, $\quad x_j^{\circlearrowleft} = -x_j,\ x_j^{\circlearrowleft} \ge 0$,

5) $x_j >< 0$, $\quad x_j = x_j^{+} - x_j^{-},\ x_j^{+}, x_j^{-} \ge 0$.

2.4.3. Чисельні методи оптимізації. Методи прямого пошуку

Необхідність використання ПК пояснюється складністю тих задач, які розв'язуються. Для знаходження оптимального розв'язку використовують різні методи оптимізації, які в загальному випадку можна поділити на наступні класи: *аналітичні, числові, графічні, експериментальні* і ін. Математичне програмування включає в себе лінійне, нелінійне, динамічне, стохастичне тощо. Як частковий випадок, розглядаємо нелінійне програмування, постановка задачі якого пов'язана з оптимізацією нелінійних функцій при лінійних і (або) нелінійних обмеженнях. Типовими областями застосування є прогнозування, планування промислового виробництва, управління ресурсами, контроль якості продукції, проектування технологічних ліній (процесів), облік та планування капіталовкладень. Поки що не існує загального методу розв'язування нелінійних задач оптимізації. Нелінійне програмування при розв'язуванні задач включає в себе елементи експериментування. Його розвиток до цих пір зводився до пропозиції часткових алгоритмів, їх програмування, перевірки результатів застосування цих алгоритмів в конкретних задачах, які представляють практичний інтерес, та побудова кращих алгоритмів на основі набутого досвіду.

Інженери та спеціалісти з дослідження операцій часто стикаються з необхідністю розв'язування оптимізаційних задач. На практиці зустрічаються різні в змістовному сенсі задачі оптимізації. Це можуть бути задачі проектування, задачі розподілу обмежених ресурсів, задачі розрахунку траєкторії польоту ракети тощо. В минулому в задачах такого типу зовсім прийнятими вважались розв'язки із досить широкої області. В інженерному проектуванні, наприклад, задавався великий коефіцієнт надійності. Але з розвитком виробництва все більше зростають вимоги до проектів. В інших галузях діяльності, таких як проектування космічних кораблів, умови функціонування проектованих систем пропонують екстремальні вимоги до характеристик проекту. Таким чином, існує необхідність в розв'язуванні наступних важливих питань. Яке найбільш ефективніше використання наявних ресурсів? Чи можна отримати більш економний в тому чи іншому сенсі проект? В яких межах можна рахувати ризик допустимим? Важливість та актуальність цих проблем викликали останнім часом інтенсивні розробки моделей та методів оптимізації.

З іншого боку, розвиток моделей та методів оптимізації стимулювався значним збільшенням розмірності та складності оптимізаційних задач, що викликані суттєвим технологічним прогресом після другої світової війни. Інженери та керівники виробництва були змушені враховувати всі існуючі фактори та їх взаємозв'язки, які впливають на якість прийнятих рішень. Деякі з цих зв'язків не завжди навіть могли бути зрозумілими та потребували, таким чином, додаткового аналізу та перевірених гіпотез. Успіх методів вимірів і статистичних методів перевірки гіпотез суттєво допомогли вновити взаємодію між різними компонентами досліджуваних систем.

Зараз методи дослідження операцій досить широко застосовуються інженерами, економістами, спеціалістами, які відповідають за прийняття рішень у військовій, культурній, адміністративній та інших галузях. Це пояснюється вже накопиченим позитивним досвідом практичного використання моделей та методів оптимізації. На ранньому етапі застосування дослідження операцій в плануванні та управлінні виробництвом, як правило, використовувались методи лінійного програмування та статистичного аналізу. Тепер добре відомі ефективні методи та програми для розв'язування такого типу задач – це нелінійне програмування.

Математичне програмування має справу із задачею оптимізації значень деякої цільової функції при обмеженнях типу рівнянь. Задача, в якій всі функції лінійні, називається *задачею лінійного програмування*. В іншому випадку має місце задача нелінійного програмування. Розробка симплекс-методу та поява швидкодіючих обчислювальних машин зробили лінійне програмування важливим інструментом розв'язування багатьох проблем, які виникають в самих різних областях. Але більшість реальних задач не може бути адекватно описана за допомогою моделей лінійного програмування через нелінійність цільової функції чи деяких обмежень. Останнім часом значного прогресу досягли дослідження нелінійних задач.

Розв'язування задач нелінійної оптимізації без обмежень базується на результатах математичного аналізу. Але знаходження екстремумів функції та її аналітичне розв'язування можливе лише для найпростіших випадків. Для задач з обмеженнями можливе застосування методу множників Лагранжа, а для обмежень загального типу формулюються фундаментальні умови Куна-Такера. Однак, це є в більшості випадків інакшим формулюванням первісної задачі, так як розв'язування системи нелінійних рівнянь за складністю еквівалентне їй. Тому, значного поширення для таких задач набули чисельні методи, важливе місце серед яких посідають *методи прямого пошуку*. Ці методи, як до речі і всі інші, не гарантують знаходження глобального оптимуму. Підвищення їх надійності досягається вибором декількох достатньо віддалених початкових точок та визначенням відповідних інтервалів пошуку.

В техніці, економіці і природничих науках, як, і в інших областях, часто виникають задачі оптимізації складної сукупності обладнання, операцій, ланцюгів або процесів. Потрібно мінімізувати або максимізувати деяку функцію, яка називається *цільовою функцією*. Вона може характеризувати ціну, вагу, загальну ефективність або щось подібне при визначених обмеженнях.

Подібні задачі оптимізації, які сформульовані математично, можуть бути об'єднані під загальною назвою *задача нелінійного програмування*; методи розв'язування таких задач лежать в основі нелінійного програмування. В широкому розумінні задача нелінійного програмування полягає в пошуку екстремуму цільової функції при заданих обмеженнях у вигляді рівностей і (або) нерівностей.

Обмеження можуть бути лінійними і (або) нелінійними. Однак, загально прийнятою є більш вузька постановка загальної задачі нелінійного програмування, в якій виключаються з розгляду наступні часткові випадки:

а) змінні можуть приймати тільки цілочисельні значення (нелінійне цілочисельне програмування);

б) обмеження містять як параметр час (оптимальне управління, динамічна оптимізація).

Змінні задачі можуть бути конструктивними параметрами, установками регулятора, показами вимірювальних приладів тощо. Цільова функція може представляти собою вартість, вагу, річний прибуток тощо., а обмеження – технічні вимоги, умови роботи, пропускну здатність або фактори безпеки, що притаманні цьому процесу.

Як часткові випадки нелінійного програмування можуть розглядатись задачі лінійного та квадратичного програмування. Розгляд тільки математичної постановки задачі нелінійного програмування не дозволяє виявити всі фактори, від яких залежить оптимізація реального процесу. Доцільним є розгляд взаємозв'язку фізичної задачі та її математичної постановки.

При оптимізації реального процесу параметри і (або) змінні пов'язані фізичними законами, такими, як закон збереження маси чи енергії, котрі повинні бути включені в задачу нелінійного програмування як обмеження у вигляді рівностей. Таким чином, одна група обмежень складається з функціональних зв'язків, які слід врахувати, щоб оптимізація була фізично здійсненною.

Друга група обмежень включає задані граничні діапазони значень змінних або параметрів, що забезпечують їх фізичну можливість реалізації і сумісності з процесом; до цієї групи відносяться обмеження у вигляді нерівностей. Замінити ці обмеження або додаватись до них можуть емпіричні зв'язки, як правило, виражені рівностями.

Загальна задача нелінійного програмування має вигляд: мінімізувати $f(x)$, при умовах $g_i(x) \le 0$, $i = \overline{1,m}$, $h_j(x) = 0$, $j = \overline{1,l}$, $x \in X$. Тут $f, g_1, ..., g_m, h_1, ..., h_l$ – визначені на E_n функції, X – множина із E_n, x – вектор з компонентами $x_1, ..., x_n$. Задача полягає в знаходженні змінних $x_1, ..., x_n$, що задовольняють обмеження та відповідають при цьому мінімальному значенню функції $f(x)$.

Функцію $f(x)$ називають *цільовою функцією* чи критерієм оптимальності. Кожну умову $g_i(x) \le 0$, $i = \overline{1,m}$ називають обмеженням-нерівністю чи обмеженням у формі нерівності, а умова вигляду $h_j(x) = 0$, $j = \overline{1,l}$ – обмеженням-рівністю чи обмеженням у формі рівності. Вектор $x \in X$, що задовольняє всім обмеженням, називають допустимим розв'язком, чи допустимою точкою. Сукупність всіх допустимих точок утворює допустиму область. Таким чином, задача нелінійного програмування полягає в знаходженні такої допустимої точки $\bar{x}$, для якої $f(x) \ge f(\bar{x})$ при всіх допустимих розв'язках x. Точка $\bar{x}$ називається оптимальним розв'язком чи просто розв'язком задачі. Відповідно, задача нелінійного програмування може бути сформульована як задача максимізації $f(x)$, а обмеження – нерівності записані у вигляді $g_i(x) \le 0$, $i = \overline{1,m}$. В спеціальному випадку, коли цільова функція $f(x)$ лінійна і всі обмеження, включаючи відношення, що описують множину X, можуть бути представлені у вигляді

лінійних рівнянь і/чи нерівностей, сформульована вище задача називається *задачею лінійного програмування.*

Для задач оптимізації функцій однієї змінної без обмежень умови екстремуму задаються у вигляді рівності похідної від функції нулеві, а характер екстремуму (мінімум, максимум, точка перегину) визначається знаком другої похідної. У випадку функцій n змінних у результаті прирівнювання часткових похідних до нуля виникає система нелінійних рівнянь, а характер екстремуму визначається негативною чи позитивною визначеністю матриці Гессе в точці екстремуму. За будь-яких умов аналітичний розв'язок цих задач можливий лише в найпростіших випадках.

У випадку пошуку екстремумів функції n змінних з обмеженнями у вигляді рівнянь первинна функція заміняється функцією Лагранжа, і система рівнянь отримується шляхом знаходження часткових похідних від змінних та множників Лагранжа і прирівнюванням їх до нуля. Умови Куна-Такера є розширенням ідей методу множників Лагранжа на задачі з обмеженнями загального вигляду – як у вигляді рівнянь, так і у вигляді нерівностей. Ці умови є необхідними умовами, які повинні виконуватись в стаціонарній точці. Розв'язування отриманої системи рівнянь та обмежень можливе лише для найпростіших задач, і в загальному для розв'язування використовуються чисельні методи.

Методи прямого пошуку є методами, у яких використовуються лише значення функції і які орієнтовані на пошук екстремумів унімодальних функцій. Ідеї та складові методів прямого пошуку використовуються і при побудові складніших методів пошуку оптимальних розв'язків. Найефективнішим є метод, що використовує числа Фібоначі. Однак, з точки зору зручності реалізації, доцільно використовувати *метод золотого перетину*, який незначно поступається методу Фібоначі. Окрім цих методів, використовують також апроксимаційні, які використовують ідею квадратичної інтерполяції. Методи прямого пошуку застосовуються також і для пошуку екстремумів функцій n змінних. Найпростішим з них є *метод покоординатного спуску* (у випадку пошуку мінімуму), який почергово по кожній зі змінних реалізує одновимірний прямий пошук, і в якості процедури одновимірного пошуку найчастіше використовується метод золотого перетину. Однак, з практичної точки зору цей метод є доволі повільним, а тому були розроблені ефективніші методи, найвідомішим серед яких є *метод Хука-Дживса*. Цей метод використовує інформацію про локальну поведінку функції для обрання найвідповідальнішого напрямку пошуку за зразком.

Метод Нелдера-Міда (або пошук по деформованому многограннику) є розвитком симплексного методу. Ідея методу полягає в порівнянні значень функції в $(n+1)$ вершинах симплекса і переміщенні симплекса в напрямку оптимальної точки за допомогою ітераційної процедури. Нелдер і Мід запропонували декілька модифікацій цього методу з неправильними симплексами. У результаті утворився дуже надійний *метод прямого пошуку*, що є одним із найефективніших, якщо $n<6$.

2.4.4. Лінійний пошук без використання похідних

Одновимірний пошук є основою багатьох алгоритмів для розв'язування задач нелінійного програмування. Звичайно, алгоритми нелінійного програмування представляють собою наступну процедуру. Задається точка x_k, визначається вектор напрямку d_k та відповідна довжина кроку λ_k, після чого розраховується нова точка

$x_{k+1} = x_k + \lambda_k d_k$. Потім цей процес повторюється. Визначення довжини кроку λ_k досягається розв'язуванням задачі мінімізації функції $f(x_k + \lambda_k d_k)$, що залежить від змінної λ. Це – задача одномірного пошуку. Мінімізація може проводиться при всіх дійсних значеннях λ, невід'ємних λ чи таких λ, при яких $x_k + \lambda d_k$ допустима.

Розглянемо задачу мінімізації функції Q, яка залежить від однієї змінної λ. Один із підходів до мінімізації Q полягає в тому, щоб знайти похідну Q', прирівняти її до нуля та розв'язати отримане рівняння відносно λ. Зауважимо, що Q визначена неявно в межах функції f декількох змінних. Зокрема, при заданих векторах x та d маємо $Q(\lambda) = f(x + \lambda d)$. Якщо f - не диференційна, то і Q - не диференційна. Якщо f - диференційна, то і $Q'(\lambda) = d^T \nabla f(x + \lambda d)$. Відповідно, щоб знайти λ, для якого $Q'(\lambda) = 0$, необхідно розв'язати рівняння $d^T \nabla f(x + \lambda d) = 0$, яке звичайно нелінійне по λ. Значення λ, яке задовольняє рівнянню $Q'(\lambda) = 0$, не обов'язково визначає мінімум функції $Q(\lambda)$. Це може бути локальний мінімум, локальний максимум чи сідлова точка. У зв'язку з цим, крім деяких спеціальних випадків, звичайно уникають мінімізації Q прирівнюванням до нуля її похідної. Замість цього застосовують деякі чисельні процедури для мінімізації функції Q. Тут розглядаються деякі методи, що не використовують розрахунок похідних для мінімізації функції Q однієї змінної на замкненому обмеженому інтервалі. Ці методи попадають під категорії методів одночасного (паралельного) чи послідовного лінійного пошуку. В першому випадку вихідні дані для всіх розрахунків відомі *apriori*, в той же час в послідовному пошуку значення функції на попередніх ітераціях використовують для визначення наступних значень.

2.4.5. Інтервал невизначеності

Розглянемо задачу одновимірної мінімізації функції $Q(\lambda)$ при умові $a \le \lambda \le b$. Так як точний локальний мінімум Q на $[a,b]$ невідомий, то цей інтервал називається *інтервалом невизначеності*. Якщо в процесі пошуку можна виключити з цього інтервалу будь-які його частини, які не мають мінімуму, то інтервал невизначеності скорочується. Взагалі $[a,b]$ називається інтервалом невизначеності, якщо точка мінімуму $\overline{\lambda}$ належить $[a,b]$, хоча її точне значення невідоме.

В наступній теоремі доводиться, що якщо функція Q строго квазіопукла в інтервалі $[a,b]$, то інтервал невизначеності може бути скорочений за допомогою розрахунків значень функції Q в двох точках, що належать цьому інтервалу.

Теорема 2.1. Нехай $Q: E_1 \to E_1$ строго квазіопукла на інтервалі $[a,b]$. Нехай $\lambda, \mu \in [a,b]$, такі, що $\lambda < \mu$. Якщо $Q(\lambda) > Q(\mu)$, то $Q(z) \ge Q(\mu)$ для всіх $z \in [a,\lambda)$. Якщо $Q(\lambda) \le Q(\mu)$, то $Q(z) \ge Q(\lambda)$ для всіх $z \in (\mu,b]$.

Доведення. Нехай $Q(\lambda) > Q(\mu)$ та $z \in [a,\lambda)$. Припустімо, що твердження теореми невірне, тобто нехай $Q(z) < Q(\mu)$. Так як точка λ може бути представлена у вигляді опуклої комбінації точок z та μ, Q – строго квазіопукла, тоді

$$Q(\lambda) < \max\{Q(z), Q(\mu)\} = Q(\mu). \qquad (2.82)$$

Це протирічить припущенню про те, що $Q(\lambda) > Q(\mu)$. Отже, $Q(z) \geq Q(\mu)$. Аналогічно доводиться друга частина теореми 2.1. Як випливає із теореми, при умові строгої квазіопуклості Q із $Q(\lambda) > Q(\mu)$ випливає, що новим інтервалом невизначеності є $[\lambda, b]$. З іншого боку, якщо $Q(\lambda) \leq Q(\mu)$, то новим інтервалом невизначеності буде $[a, \lambda]$. Ці два випадки проілюстровані на рис. 2.7.

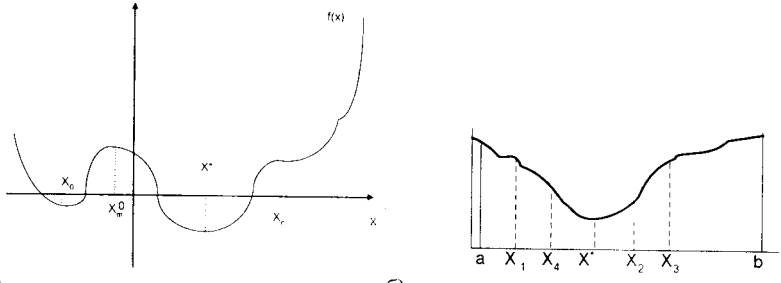


Рис.2.7. Скорочення інтервалу невизначеності, де 1 - $Q(\lambda)$, 2 - $Q(\mu)$, 3 - λ та 4 - μ

В літературі по нелінійному програмуванню часто використовується поняття *унімодальності функції* Q. Тут використовується еквівалентне поняття строгої квазіопуклості. Тепер розглянемо декілька процедур мінімізації строгої квазіопуклої функції на замкненому обмеженому інтервалі завдяки ітеративному скороченню інтервалу обмеженості.

2.4.6. Функції однієї змінної

Функція $f(x)$ має локальний мінімум в точці x_0, якщо існує деяка додатня величина δ, така, що якщо $|x - x_0| < \delta$, то $f(x) \geq f(x_0)$, тобто якщо існує окіл точки x_0, такий, що для всіх значень x в цьому околі $f(x)$ більше $f(x_0)$. Функція $f(x)$ має глобальний мінімум у точці x^*, якщо для всіх x справедлива нерівність $f(x) \geq f(x^*)$. На рис. 2.8а – графічне представлення функції $f(x)$, що має локальний мінімум у точці x_0 та глобальний мінімум у точці x^*.



а) б)

Рис.2.8. Графічне подання функції $f(x)$

В точках x_0 і x^* похідна $f'(x)$ (градієнт функції) рівна нулю. Отже, x_0 і x^* будуть розв'язками рівняння $f'(x)=0$. Точка x_m, в якій досягається локальний максимум, і точка x_c, у якій міститься точка горизонтального перегину функції, також задовільняють цю умову, яка є **необхідною**, але не є **достатньою** умовою існування екстремуму. Розглянемо розклад функції $f(x)$ у ряд Тейлора в околі точки x_0, (вважаючи, що $f(x)$ та її похідні неперервні):

$$f(x_0+h)-f(x)=hf'(x_0)+\frac{h^2}{2!}f''(x_0)+\frac{h^3}{3!}f'''(x_0)+\frac{h^4}{4!}f''''(x_0)+\ldots. \qquad (2.83)$$

Якщо в точці x_0 досягається мінімум, то ліва частина співвідношення буде невід'ємною для будь-якого достатньо малого h ($|h|<\delta$). Отже, перша похідна $f'(x_0)$ повинна бути рівною нулю. Оскільки в наступній складовій $h^2>0$ завжди, коли $f''(x_0)>0$, то у точці x_0 досягається мінімум. Якщо $f'(x_m)=0$ і $f''(x_m)<0$, то з аналогічних міркувань у точці x_m досягається максимум. Таким чином, у загальному випадку, якщо функція $f(x)$ і її похідні неперервні, то точка x_0 є точкою екстремуму (максимуму або мінімуму) тоді, і лише тоді, коли n парне, де n - порядок першої похідної. що не перетворюється в нуль у точці x_0. Якщо $f^n(x_0)<0$, то в точці x_0 досягається максимум, якщо $f^n(x_0)>0$, то в точці x_0 досягається мінімум.

2.4.7. Функції *n* змінних

Розглянемо функцію n дійсних змінних $f(x_1,x_2,x_3,\ldots,x_n)=f(x)$. Точка в n-вимірному евклідовому просторі з координатами $x_1,x_2,x_3,\ldots,x_n$, є вектором-стовпчиком x. Градієнт функції, тобто вектор із компонентами $\partial f/\partial x_1,\partial f/\partial x_2,\ldots,\partial f/\partial x_n$, позначається $\nabla f(x)$ або g(x). Матриця Гессе (гессіан) функції $f(x)$ позначається як G(x) і являє собою симетричну матрицю $n\times n$ елементів виду

$$G_{ij}=\frac{\partial^2 f}{\partial x_i \partial x_j}. \qquad (2.84)$$

Функція $f(x)$ має локальний мінімум у точці x_0, якщо існує окіл точки x_0, такий, що $f(x)$ більше $f(x_0)$ у всіх точках цього околу, тобто існує додатня величина δ, така, що для $|x-x_0|<\delta$ справедлива нерівність $f(x)\ge f(x_0)$. У випадку глобального мінімуму в точці x^* для всіх x справедлива нерівність $f(x)\ge f(x^*)$. Розкладаючи, отримаємо

$$\begin{aligned} f(x_0+h)-f(x_0)&=\sum_{i=1}^{n}h_j\frac{\partial f}{\partial x_i}(x_1,\ldots,x_n)+\frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n}h_ih_j\frac{\partial^2 f}{\partial x_i\partial x_j}(x_1,\ldots,x_n)+\ldots \\ &=h^T\nabla f(x_0)+\frac{1}{2}h^T G(x_0)h+\ldots \end{aligned} \qquad (2.85)$$

Якщо x_0 є точкою мінімуму функції $f(x)$, то кожна перша окрема похідна $\partial f / \partial x_i \quad i = (1,...,n)$ повинна перетворюватись в нуль у точці x_0. Якщо це не так, то відповідним вибором h_i можна добитися того, що різниця $f(x_0 + h) - f(x_0)$ буде від'ємною. Отже, необхідною умовою мінімуму в точці x_0 є рівняння $\nabla f(x_0) = 0$, тобто $\frac{\partial f(x_0)}{\partial x_i} = 0 \ (i = 1,...,n)$. Тоді знак різниці $f(x_0 + h) - f(x_0)$ визначається членом $\frac{1}{2} h^T G(x_0) h$. Якщо матриця $G(x_0)$ додатньо визначена, те цей член додатній для всіх h. Отже, необхідними і достатніми умовами мінімуму є $\nabla f(x_0) = 0, G(x_0)$ - додатньо визначена матриця. Відповідно необхідними і достатніми умовами максимуму є $\nabla f(x_m) = 0, \ G(x_m)$ - від'ємно визначена матриця.

2.4.8. Методи прямого пошуку для функції однієї змінної

Методи прямого пошуку є методами, у яких використовуються тільки значення функції. У більшості випадків рівняння $f'(x) = 0$ не розв'язується простим способом. Тому розглянемо декілька простих чисельних процедур, що дозволяють знайти екстремум (мінімум) функції $f(x)$. За допомогою чисельних методів безпосередньо шукаємо мінімум функції $f(x)$ у деякому інтервалі $a < x < b$ (рис. 2.8б), де, як передбачається його місцезнаходження, розраховуючи значення функції в обраних точках цього інтервалу. Іноді це єдина можлива стратегія пошуку. Наприклад, вартість проведення хімічного процесу може залежати від температури процесу. Відомо, що вартість є функцією від T, хоча й невідомий явний вид функції. Проте, можна поставити експеримент і провести процес при різних температурах, визначивши ту, для якої вартість мінімальна.

Можна спробувати знайти положення мінімуму в точці, що апроксимує його з потрібною точністю, або визначити малий інтервал, у якому знаходиться мінімум. Спробуємо досягти поставленої мети найбільш ефективним способом, тобто здійснюючи найменшу кількість обчислень функції. У наведеному вище прикладі, неможливо точно регулювати температуру процесу, тому точність у $1°C$ або навіть у $10°C$ є цілком задовільною. Проте, оскільки проведення експериментів потребує певних витрат, їх кількість повинна бути мінімальною.

Припустимо, що точки a та b визначають (можливо, дуже грубо) інтервал, що містить дійсну точку мінімуму, і усередині цього інтервалу функція *унімодальна*, тобто має один мінімум у точці x^*. Отже, функція має форму, близьку до тієї, що зображена на рис. 2.8. Якщо відомі значення функції такого виду в трьох точках x_1, x_2, x_3, таких, що $a < x_1 < x_2 < x_3 < b$, а $f(x_2) < f(x_1)$ і $f(x_2) < f(x_3)$, то $x_1 < x^* < x_3$. Тоді точка x^* буде лежати усередині інтервалу $(x_1; x_3)$, який менший за розміром, ніж інтервал $(a; b)$.

2.4.9. Задача оптимізації методом Ньютона

Для функцій однієї змінної класичний підхід при пошуку значень x в точках перегину функції $f(x)$ полягає в розв'язуванні рівняння $f'(x) = 0$.

Розв'язати таке рівняння не завжди просто. Тому коротко розглянемо чисельний

метод його розв'язування. Приблизний ескіз кривої $y = f`(x)$ дозволить одержати наближений розв'язок. Якщо можна знайти два значення $a$ і $b$, такі, що $f`(a)$ і $f`(b)$ мають протилежні знаки, то тоді, в силу очевидних припущень про неперервність, буде існувати корінь η цього рівняння, причому $a < \eta < b$ (рис. 2.9.а).

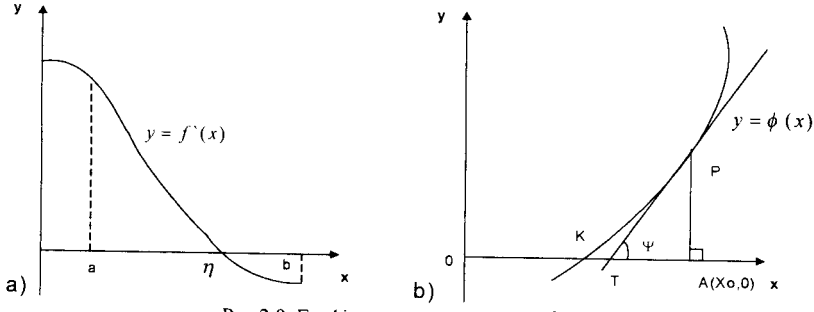


Рис.2.9. Графічне подання функції $f(x)$

Метод Ньютона дозволяє поліпшити відносно грубу апроксимацію, щоб одержати корінь рівняння $\varphi(x) \equiv 0$. [У цій задачі $\varphi(x) \equiv f`(x)$.] На рис. 2.9.b точка x_0, що є координатою x точки P, являє собою апроксимацію кореня рівняння $\varphi(x) = 0$. Нехай PT – дотична до кривої у точці P, а T – точка, в якій дотична перетинає вісь x. Тоді, в загальному випадку, OT є кращою апроксимацією кореня, що лежить у точці K. Тепер OT=OA-TA= x_0 -TA . Крім того,

$$\frac{PA}{TA} = tg\psi = \varphi`(x_0), \tag{2.86}$$

$$TA = \frac{PA}{\varphi`(x_0)} = \frac{\varphi(x_0)}{\varphi`(x_0)}. \tag{2.87}$$

$$x_1 = x_0 - \frac{\varphi(x_0)}{\varphi`(x_0)}. \tag{2.88}$$

Аналогічно можна отримати краще значення для

$$x_2 = x_1 - \frac{\varphi(x_0)}{\varphi`(x_0)} \tag{2.89}$$

і в загальному випадку

$$x_{r+1} = x_r - \frac{\varphi(x_r)}{\varphi`(x_r)}. \tag{2.90}$$

Ітерації можуть бути продовжені доти, поки для двох наступних апроксимацій не буде досягнута необхідна точність.

2.5. Порівняння методів лінійного пошуку без розрахунку похідної

2.5.1. Класифікація методів лінійного пошуку без розрахунку похідної

Великою перевагою **прямих методів** є те, що від **цільової функції** не потрібно диференціювання і, більше того, вона може бути не задана в аналітичному вигляді. Єдине, на чому грунтуються **алгоритми прямих методів мінімізації**, це можливість визначення значень функції $f(x)$ в заданих точках. До найбільш поширених прямих методів можна віднести **метод рівномірного пошуку, метод ділення відрізка навпіл (дихотомічний пошук), метод золотого перетину (січення)** та **метод Фібоначі**.

Очевидно, що при заданій строго квазиопуклої на інтервалі $[a_1, b_1]$ функції Q кожний із розглянутих в тут методів дає за скінчене число кроків таку точку λ, що $|\lambda - \bar{\lambda}| \le l$, де l – довжина кінцевого інтервалу невизначеності, а λ – точка мінімуму на заданому інтервалі. Зокрема, при заданій довжині кінцевого інтервалу невизначеності, котра задовільняє необхідну ступінь точності, необхідне число розрахунків функції n може бути визначено як найменше позитивне ціле, яке задовільняє наступні відношення:

1. Метод рівномірного пошуку $n \ge \dfrac{2(b_1 - a_1)}{l} - 1$,
2. Дихотомічний пошук $\left(\dfrac{1}{2}\right)^{\frac{n}{2}} \ge \dfrac{l}{b_1 - a_1}$,
3. Метод золотого січення $(0.618)^{n-1} \ge \dfrac{1}{b_1 - a_1}$,
4. Метод Фібоначчі $F_n \ge \dfrac{b_1 - a_1}{l}$.

Із цих виразів видно, що необхідна кількість розрахунків є функцією відношення $(b_1 - a_1)/l$. Для фіксованого значення $(b_1 - a_1)/l$ найменша кількість необхідних розрахунків функції відповідає більш ефективному алгоритму. Зрозуміло, що з цієї точки зору найбільш ефективним алгоритмом є метод Фібоначчі, далі – метод золотого січення, дихотомічний пошук і, нарешті, метод рівномірного пошуку.

Зауважимо, що для достатньо великих n значення $1/F_n$ прямує до $(0.618)^{n-1}$, так що методи Фібоначчі та золотого січення є майже ідентичними.

Необхідно підкреслити, що серед методів без використання похідної для мінімізації строго квазиопуклих функцій на замкненому обмеженому інтервалі метод Фібоначчі є найбільш ефективним в тому сенсі, що вимагає найменшу кількість розрахунків функції при заданих вимогах до коефіцієнта стиску довжини інтервалу невизначеності. Наведені характеристики всіх розглянутих вище процедур базуються на припущені про строгу квазиопуклість функції, що мінімізується. В багатьох задачах це припущення не виконується чи, можливо, не може бути достатньо легко перевірене. Одним із шляхів усунення цієї складності, головним чином якщо початковий інтервал невизначеності великий, є розбиття його на маленькі інтервали, знаходження мінімуму

на кожному із підінтервалів та наступний вибір найменшого значення з мінімумів на підінтервалах. Метод поділу пополам та метод Ньютона вимагають інформації про похідну функції, яку ми маємо мінімізувати.

2.5.2. Рівномірний пошук

Рівномірний пошук є прикладом одночасного пошуку, коли точки, в яких обчислюються значення функції, обираються заздалегідь. Інтервал невизначеності $[a_1, b_1]$ ділиться на частини сіткою з точок $a_1 + k\delta$ для $k = \overline{1,n}$, де $b_1 = a_1 + (n+1)\delta$, як показано на рис. 2.10:

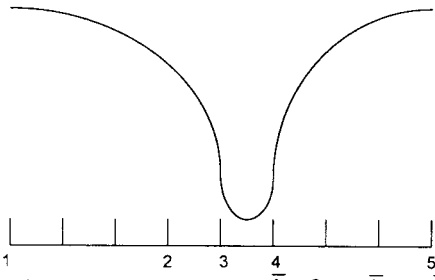


Рис.2.10. Рівномірний пошук, де 1 - a_1, 2 - $\overline{\lambda} - \delta$, 3 - $\overline{\lambda}$, 4 - $\overline{\lambda} + \delta$ та 5 - b_1

Функція Q обчислюється в кожній із n точок сітки. Нехай $\overline{\lambda}$ – точка сітки з найменшим значенням Q. Якщо функція Q строго квазиопукла, то точка її мінімуму належить інтервалу $[\overline{\lambda} - \delta, \overline{\lambda} + \delta]$.

Вибір величини комірки δ. Після n розрахунків функції інтервал невизначеності скорочений до довжини 2δ. Так як $n = [(b_1 - a_1)/\delta] - 1$, то для подальшого скорочення інтервалу невизначеності необхідно збільшувати кількість обчислень функції. Часто метод скорочення обчислювальної похибки складається із використання на початку процедури сітки з великою коміркою та з переходу потім до сітки з малим розміром комірки.

2.5.3. Послідовний пошук

Очевидно, що можна запропонувати більш ефективну процедуру, що використовує інформацію, яка отримана на попередніх ітераціях, для визначення місця наступних вимірів. Розглянемо наступні процедури послідовного пошуку: дихотомічний пошук, метод золотого січення та метод Фібоначчі.

2.5.4. Задача оптимізації методом дихотомічного пошуку (поділу пополам)

Розглянемо функцію $Q: E_1 \to E_1$, яку необхідно мінімізувати на інтервалі $[a_1, b_1]$. Припустимо, що Q строго квазиопукла. Очевидно, що найменше число розрахунків значень функції, які необхідні для скорочення інтервалу невизначеності, рівне двом. На рис. 2.11а локалізовані дві точки λ_1 та μ_1. $Q(\lambda_1) < Q(\mu_1)$ для $Q = Q_1$, і,

відповідно за теоремою, новим інтервалом невизначеності є $[a_1, \mu_1]$. Але для $Q = Q_2$ зауважимо, що $Q(\lambda_1) > Q(\mu_1)$, і, відповідно, за теоремою новим інтервалом невизначеності є $[\lambda_1, b_1]$. Таким чином, в залежності від функції довжина нового інтервалу невизначеності рівна $\mu_1 - a_1$ чи $b_1 - \lambda_1$.

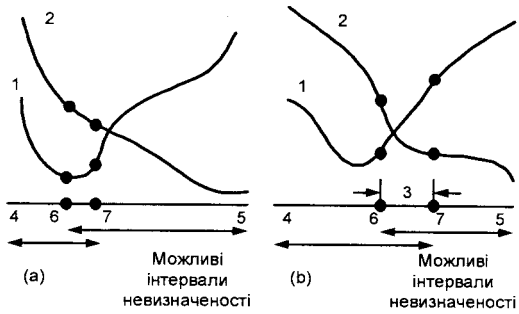


Рис. 2.11. Можливі інтервали невизначеності, де 1-Q_1, 2-Q_2, 3-2ε, 4-a_1, 5-b_1, 6-λ_1, 7-μ_1

Зауважимо, що аргіогі невідомо, чи буде $Q(\lambda_1) < Q(\mu_1)$ чи $Q(\lambda_1) > Q(\mu_1)$. Якщо виконується рівність $Q(\lambda_1) = Q(\mu_1)$, то інтервал невизначеності можна скоротити до $[\lambda_1, \mu_1]$. Необхідно зауважити, що точна рівність досить малоймовірна в практичній ситуації. Тому оптимальна стратегія вибору точок λ_1 та μ_1 є в прагненні застерегти себе від найгіршого результату, тобто мінімізувати максимум з $\mu_1 - a_1$ чи $b_1 - \lambda_1$. Це може бути досягнуте вибором λ_1 та μ_1 як середини інтервалу $[a_1, b_1]$. Але в цьому випадку будемо мати тільки одну точку та не зможемо скоротити далі інтервал невизначеності. Тому, як показано на рис. 2.11b. λ_1 та μ_1 вибрані симетрично на відстані $\varepsilon > 0$ від середини інтервалу кожна. Тут число $\varepsilon > 0$ настільки мале, щоб довжина нового інтервалу невизначеності $\varepsilon + (b_1 - a_1)/2$ була достатньо близькою до теоретично оптимального значення $(b_1 - a_1)/2$, і в той же час таке, щоби значення функції $Q(\lambda_1)$ та $Q(\mu_1)$ були різними. В дихотомічному пошуку місце кожного із перших двох спостережень λ_1 та μ_1 обирається симетрично на відстані ε від середини $(b_1 - a_1)/2$. В залежності від значення функції Q в точках λ_1 та μ_1 визначається новий інтервал невизначеності. Потім процес повторюється.

Нижче подано алгоритм дихотомічного методу для мінімізації строго квазиопуклої функції на інтервалі $[a_1, b_1]$.

Алгоритм дихотомічного пошуку

Початковий етап. Вибрати константу різниці $2\varepsilon > 0$ та допустиму кінцеву довжину інтервалу невизначеності $l > 0$. Нехай $[a_1, b_1]$ – початковий інтервал невизначеності. Покласти $k = 1$ та перейти до основного етапу.

Основний етап.

Крок 1. Якщо $b_k \to a_k < l$, то зупинитися; точка мінімуму належить інтервалу $[a_k, b_k]$. В іншому випадку обчислити

$$\lambda_k = \frac{a_k + b_k}{2} - \varepsilon\,,\ \mu_k = \frac{a_k + b_k}{2} + \varepsilon\,. \tag{2.91}$$

та перейти до кроку 2.

Крок 2. Якщо $Q(\lambda_k) < Q(\mu_k)$, взяти $a_{k+1} = a_k$ та $b_{k+1} = \mu_k$. В іншому випадку беремо $a_{k+1} = \lambda_k$ та $b_{k+1} = b_k$. Замінити k на $k+1$ та перейти до кроку 1.

Довжина інтервалу невизначеності на початку $(k+1)$-ї ітерації рівна

$$(b_{k+1} - a_{k+1}) = \frac{1}{2^k}(b_1 - a_1) + 2\varepsilon(1 - \frac{1}{2^k})\,. \tag{2.92}$$

Ця формула може бути використана для визначення числа ітерацій, які необхідні для досягнення бажаної точності. Так як кожна ітерація вимагає двох вимірів, то ця формула може бути також використана для визначення кількості обчислень функції Q.

Приклад 2.10. Нехай необхідно мінімізувати функцію Q на замкненому та обмеженому інтервалі. Крім того, функція Q псевдоопукла і, відповідно, диференційовна. На k-ій ітерації інтервал невизначеності рівний $[a_k, b_k]$. Похідна $Q'(\lambda_k)$ відома та розглянемо наступні три можливих випадки:

1. Якщо $Q'(\lambda_k) = 0$, то із псевдоопуклості Q виплаває, що точка λ_k – точка мінімуму.
2. Якщо $Q'(\lambda_k) > 0$, то для $\lambda > \lambda_k$ маємо $Q'(\lambda_k)(\lambda - \lambda_k) > 0$ та в силу псевдоопупуклості Q це означає, що $Q(\lambda) \geq Q(\lambda_k)$. Іншими словами, мінімум лежить зліва від λ_k, так що новим інтервалом невизначеності $[a_{k+1}, b_{k+1}]$ буде відрізок $[a_k, \lambda_k]$.
3. Якщо $Q'(\lambda_k) < 0$, то $Q'(\lambda_k)(\lambda - \lambda_k) < 0$ для $\lambda < \lambda_k$, так що $Q(\lambda) \leq Q(\lambda_k)$. Таким чином, мінімум лежить праворуч від λ_k, так що новим інтервалом невизначеності $[a_{k+1}, b_{k+1}]$ буде відрізок $[\lambda_k, b_k]$.

Положення λ_k на інтервалі $[a_k, b_k]$ повинно бути обране таким чином, щоб мінімізувати максимально можливу довжину нового інтервалу невизначеності, тобто щоби мінімізувати більшу із величин $\lambda_k - a_k$ та $b_k - \lambda_k$. Очевидно, що оптимальним положенням λ_k є середина відрізку $\frac{1}{2}(a_k + b_k)$. Таким чином, на будь-якій ітерації k похідна Q' обчислюється в середній точці інтервалу невизначеності. В залежності від значення Q' процес або зупиняється, або будується новий інтервал невизначеності, довжина якого рівна половині довжини попереднього інтервалу. Зауважимо, що ця процедура дуже схожа на метод дихотомічного пошуку, за винятком того, що на кожній ітерації вимагається лише одне обчислення похідної, в той же час як в методі дихотомічного пошуку необхідними є обчислення функції двічі.

Програма 2.1. Програма методу половинного поділу

```
#include <iostream.h>
#include <stdlib.h>
#include <stdio.h>
#include <math.h>
```

```
#include <conio.h>
void main(void){ int n=5,j; double x1,x2,x3,x4,L,R,f,x; x1=-3; x3=5; j=0;
  clrscr(); cout<<"Metod polovynogo podilu";
  A: j++;cout<<endl<<j<<": "; cout<<"["<<x1<<";"<<x3<<"]";
  x=(x1+x3)/2;  f=2*x+2;
  if(f==0){n=0;cout<<"  f=0   xmin="<<x;getch();}
  if(f>0){n--;x3=x;cout<<"  f>0   x1="<<x;getch();}
  if(f<0){n--;x1=x;cout<<"  f<0   x1="<<x;getch();}
  if(n>0)goto A; getch();}
```

Тепер опрацюємо наші результати та намалюємо наступний графік для різних методів оптимізації одновимірного унімодального критерію – пошук найменшого можливого інтервалу невизначеності, враховуючи різну точність (рис. 2.12).

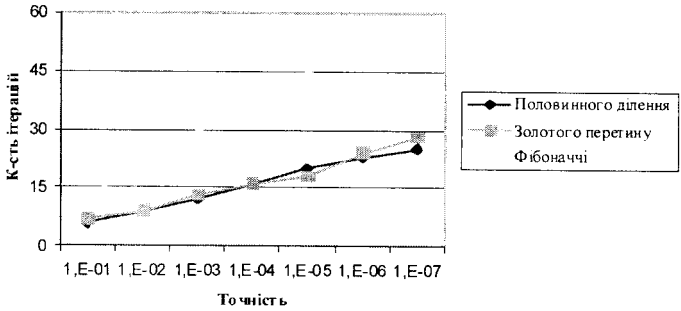


Рис. 2.12. Порівняння результатів оптимізації різними методами

2.5.5. Задача оптимізації за методом золотого січення

Порівняння різних процедур лінійного пошуку природно проводити у відповідності до наступного коефіцієнту стиску:

$$\frac{\textit{довжина інтервалу невизначеності після } v \textit{ виконаних спостережень}}{\textit{довжина інтервалу невизначеності перед виконанням спостереження}}$$

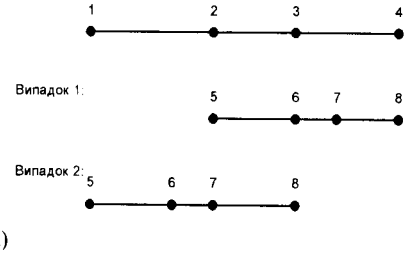
Очевидно, що більш ефективні схеми відповідають меншим значенням коефіцієнта стиску. В дихотомічному пошуку значення цього коефіцієнта приблизно рівне $(0.5)^{v/2}$. Тепер опишемо більш ефективний метод золотого січення, для якого значення коефіцієнта стиску рівне $(0.618)^{v-1}$. Метод може бути використаний для мінімізації строгоквазіопуклих функцій.

Нехай на k-ій ітерації методу золотого січення інтервал невизначеності рівний $[a_k, b_k]$. За теоремою новий інтервал невизначеності $[a_{k+1}, b_{k+1}]$ рівний $[\lambda_k, b_k]$, якщо $Q(\lambda_k) > Q(\mu_k)$, та $[a_k, \mu_k]$, якщо $Q(\lambda_k) \le Q(\mu_k)$. Точки λ_k та μ_k обираються, виходячи із наступних умов:

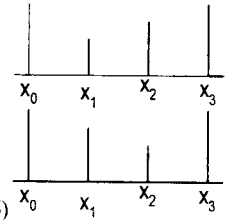
1. Довжина нового інтервалу невизначеності $b_{k+1} - a_{k+1}$ не залежить від результату на k-ій ітерації. тобто від того, чи виконується нерівність $Q(\lambda_k) > Q(\mu_k)$ або $Q(\lambda_k) \le Q(\mu_k)$. Крім того, повинна виконуватись рівність

$b_k - \lambda_k = \mu_k - a_k$. Таким чином, якщо $\lambda_k = a_k + (1-\alpha)(b_k - a_k)$, де $\alpha \in (0,1)$, то для μ_k повинно бути $\mu_k = a_k + \alpha(b_k - a_k)$, так що $b_{k+1} - a_{k+1} = \alpha(b_k - a_k)$.

2. Для нової ітерації λ_{k+1} та μ_{k+1} обираються так, що λ_{k+1} співпадає з μ_k або μ_{k+1} співпадає з λ_k. Якщо цього досягти, то на $(k+1)$-ій ітерації необхідно виконати лише один новий розрахунок функції. Щоб показати це, розглянемо рисунок та два наступні випадки (рис. 2.13а).



а)



б)

а). $f_1 < f_2$, Новий інтервал (x_0, x_2)

а). $f_1 > f_2$, Новий інтервал (x_1, x_3)

Рис. 2.13. а) Правило золотого січення, де $1 - a_k$, $2 - \lambda_k$, $3 - \mu_k$, $4 - b_k$, $5 - a_{k+1}$, $6 - \lambda_{k+1}$, $7 - \mu_{k+1}$ та $8 - b_{k+1}$ та б) інтервали для методу золотого січення

Випадок 1. $Q(\lambda_k) > Q(\mu_k)$. В цьому випадку $a_{k+1} = \lambda_k$ та $b_{k+1} = b_k$. Замінимо k на $k+1$. При $\lambda_{k+1} = \mu_k$ маємо

$$\mu_k = \lambda_{k+1} = a_{k+1} + (1-\alpha)(b_{k+1} - a_{k+1}) = \lambda_k + (1-\alpha)(b_k - \lambda_k). \tag{2.93}$$

Підставляючи вираз для λ_k та μ_k

$$\lambda_k = a_k + (1-\alpha)(b_k - a_k), \tag{2.94}$$

$$\mu_k = a_k + \alpha(b_k - a_k) \tag{2.95}$$

в це рівняння, отримаємо, що $\alpha^2 + \alpha - 1 = 0$.

Випадок 2. $Q(\lambda_k) \le Q(\mu_k)$. В цьому випадку $a_{k+1} = a_k$ та $b_{k+1} = \mu_k$. Замінимо k на $k+1$. При $\mu_{k+1} = \lambda_k$ маємо

$$\lambda_k = \mu_{k+1} = a_{k+1} + \alpha(b_{k+1} - a_{k+1}) = a_k + \alpha(\mu_k - a_k). \tag{2.96}$$

Підставляючи вираз для λ_k та μ_k

$$\lambda_k = a_k + (1-\alpha)(b_k - a_k), \tag{2.97}$$

$$\mu_k = a_k + \alpha(b_k - a_k) \tag{2.98}$$

в це рівняння, отримаємо, що $\alpha^2 + \alpha - 1 = 0$. Коренями рівняння $\alpha^2 + \alpha - 1 = 0$ є $\alpha \approx 0.618$ та $\alpha \approx -1.618$. Так як α повинно бути з інтервалу $(0,1)$, то $\alpha \approx 0.618$. Таким чином, якщо на k-ій ітерації λ_k та μ_k обираються у відповідності з

$$\lambda_k = a_k + (1-\alpha)(b_k - a_k), \tag{2.99}$$

$$\mu_k = a_k + \alpha(b_k - a_k), \tag{2.100}$$

де $\alpha \cong 0.618$, то довжина інтервалу невизначеності стискається з коефіцієнтом 0.618. На першій ітерації необхідні два обчислення функції в точках λ_1 та μ_1, але на кожній наступній необхідне лише одне обчислення, так як $\lambda_{k+1} = \mu_k$ або $\mu_{k+1} = \lambda_k$.

Нижче поданий алгоритм золотого січення для мінімізації строго квазіопуклої функції на інтервалі $[a_1, b_1]$.

Алгоритм методу золотого січення

Початковий етап. Обрати допустиму кінцеву довжину інтервалу невизначеності $l > 0$. Нехай $[a_1, b_1]$ – початковий інтервал невизначеності. Взяти $\lambda_1 = a_1 + (1-\alpha)(b_1 - a_1)$ та $\mu_1 = a_1 + \alpha(b_1 - a_1)$, де $\alpha \cong 0.618$. Обчислити $Q(\lambda_1)$ та $Q(\mu_1)$. Взяти $k = 1$ та перейти до основного етапу.

Основний етап.

Крок 1. Якщо $b_k - a_k < l$, то зупинитися; оптимальна точка належить інтервалу $[a_k, b_k]$. В іншому випадку, якщо $Q(\lambda_k) > Q(\mu_k)$, то перейти до кроку 2, а якщо $Q(\lambda_k) \le Q(\mu_k)$, то до кроку 3.

Крок 2. Взяти $a_{k+1} = \lambda_k$, $b_{k+1} = b_k$, $\lambda_{k+1} = \mu_k$, $\mu_{k+1} = a_{k+1} + \alpha(b_{k+1} - a_{k+1})$. Обчислити $Q(\mu_{k+1})$ та перейти до кроку 4.

Крок 3. Взяти $a_{k+1} = a_k$, $b_{k+1} = \mu_k$, $\mu_{k+1} = \lambda_k$, $\lambda_{k+1} = a_{k+1} + (1-\alpha)(b_{k+1} - a_{k+1})$. Обчислити $Q(\lambda_{k+1})$ та перейти до кроку 4.

Крок 4. Замінити k на $k+1$ та перейти до кроку 1.

Приклад 2.11. Розглянемо наступну задачу: мінімізувати $\lambda^2 + 2\lambda$ при умові $-3 \le \lambda \le 5$.

Розв'язок. Очевидно, що $Q(\lambda_1) < Q(\mu_1)$. Відповідно, новий інтервал невизначеності рівний [-3.0,1.944]. Цей процес повторюється, результати розрахунків наведені в таблиці нижче. Значення функції Q, які обчислені на кожній ітерації, *позначені зірочкою.*

Таблиця 2.5

Розрахунки за методом золотого січення

k	a_k	b_k	λ_k	μ_k	$Q(\lambda_k)$	$Q(\mu_k)$
1	-3.000	5.000	0.056	1.944	0.115*	7.667*
2	-3.000	1.944	-1.112	0.056	-0.987*	0.115
3	-3.000	0.056	-1.832	-1.112	-0.308*	-0.987
4	-1.832	0.056	-1.112	-0.664	-0.987	-0.887*
5	-1.832	-0.664	-1.384	-1.112	0.853*	-0.987
6	-1.384	-0.664	-1.112	-0.936	-0.987	-0.996*
7	-1.384	-0.936	-1.208	-1.112	-0.957*	-0.987
8	-1.208	-0.936	-1.112	-1.032	-0.987	-0.999*
9	-1.112	-0.936				

Після вісьми ітерацій, які мають дев'ять обчислень функції, інтервал невизначеності рівний [-1.112,-0.936], так що, як точку мінімуму, можна взяти, наприклад, середину цього інтервалу -1.024. Зауважимо, що точкою точного мінімуму є -1.0.

Пошук методом золотого січення. Не завжди можна заздалегідь визначити, скільки разів потрібно обчислювати функцію. У методі Фібоначчі це потрібно знати для визначення L_2, тобто положення початкової точки. Метод "золотого січення" майже настільки ж ефективний, як і метод Фібоначчі. Проте, при цьому не потрібно знати n – кількість обчислень функції, обумовлену спочатку. Після того, як виконано j обчислень, виходячи з тих же міркувань, що і раніше, записуємо $L_{j-1} = L_j + L_{j+1}$.

Проте, якщо n невідомо, то ми не можемо використовувати умову $L_{n-1} = 2L_n - \varepsilon$. Якщо відношення довжин наступних інтервалів буде постійним, $\frac{L_{j-1}}{L_j} = \frac{L_j}{L_{j+1}} = \frac{L_{j+1}}{L_{j+2}} = \ldots = \tau$, то $\frac{L_{j-1}}{L_j} = 1 + \frac{L_{j+1}}{L_j}$, тобто $\tau = 1 + 1/\tau$. Таким чином, $\tau^2 - \tau - 1 = 0$, звідки $\tau = (1+\sqrt{5})/2 \approx 1{,}618033989$. Тоді $\frac{L_{j-1}}{L_{j+1}} = \tau^2$, $\frac{L_{j-2}}{L_{j+1}} = \tau^3$ і т.д. Отже, $\frac{L_1}{L_n} = \tau^{n-1}$, тобто $L_n = \frac{L_1}{\tau^{n-1}}$. Проаналізувавши ці два значення функції, визначимо той інтервал, який повинен досліджуватись надалі. Цей інтервал буде містити одну з попередніх точок і наступну точку, яку розташовуємо симетрично до неї. Перша точка знаходиться на відстані L_1/τ від одного кінця інтервалу, інша – на такій же відстані від іншого. Оскільки $\lim_{n\to\infty} F_{n-1}/F_n = 1/n$, то з рівняння видно, що пошук методом "золотого січення" є граничною формою пошуку методом Фібоначчі. Назва "золоте січення" пішла від назви відношення в рівнянні. L_{j-1} ділиться на дві частини так, що відношення цілого до більшої частини рівне відношенню більшої частини до меншої, тобто дорівнює так званому "золотому відношенню".

Таким чином, якщо шукаємо інтервал (x_0, x_3) і є два значення функції f_1 і f_2 в точках x_1 і x_2, то слід розглянути два випадки, поданих на рис. 2.13б.

Програма 2.2. Метод золотого січення

```
#include <iostream.h>
#include <stdlib.h>
#include <stdio.h>
#include <math.h>
#include <conio.h>
void main(void){  int n=35,j;
  double a,b,l,m,fl,fm,al=0.618;
  a=-3; b=5; j=0;
  clrscr();
  l=a+(1-al)*(b-a);
  m=a+al*(b-a);
  fl=pow(l,2)+2*l;
  fm=pow(m,2)+2*m;
  cout<<"Metod zolotogo peretynu";
  A: j++;cout<<endl<<j<<": ";
  cout<<"["<<a<<";"<<b<<"] ";  getch();//  fl=pow(l,2)+2*l;//  fm=pow(m,2)+2*m;
  if(fl>fm){n--;a=l;l=m;m=a+al*(b-a);fm=pow(m,2)+2*m;}
  if(fl<=fm){n--;b=m;m=l;l=a+(1-al)*(b-a);fl=pow(l,2)+2*l;} /*  x2=(3*x3+x1)/4;
```

```
  cout<<"  x2="<<x2;
  x4=x1-x2+x3;     cout<<"  x4="<<x4;
  getch();
  L=x2-x1;
  R=x3-x2;
  if(L>R){  n--;
  f2=pow(x2,2)+2*x2;
  f4=pow(x4,2)+2*x4;
  if(f4<f2){x1=x1;x3=x2;}
  if(f4>f2){x1=x4;x3=x3;}  }*/
 if(n>0)goto A;getch();}
```

2.5.6. Задача оптимізації за методом Фібоначчі

Метод Фібоначчі є процедурою лінійного пошуку для мінімізації строго квазиопуклої функції на замкненому обмеженому інтервалі. Подібно до методу золотого січення процедура пошуку Фібоначчі вимагає два обчислення функції на першій ітерації, а на кожній наступній – тільки одне. Але ця процедура відрізняється від методу золотого січення тим, що скорочення інтервалу невизначеності змінюється від ітерації до ітерації. Процедура базується на послідовності Фібоначчі $\{F_v\}$, яка визначається наступним чином:

$$F_{v+1} = F_v + F_{v-1},\ v = 1,2,\ldots,\ F_0 = F_1 = 1. \tag{2.101}$$

Таким чином, послідовність має вигляд 1,1,2,3,5,8,13,21,34,55,89,144,233,...

Припустимо, що на k-ій ітерації інтервал невизначеності рівний $[a_k, b_k]$. Розглянемо дві точки λ_k та μ_k, що визначаються наступним чином:

$$\lambda_k = a_k + \frac{F_{n-k-1}}{F_{n-k+1}}(b_k - a_k),\ k = \overline{1, n-1}, \tag{2.102}$$

$$\mu_k = a_k + \frac{F_{n-k}}{F_{n-k+1}}(b_k - a_k),\ k = \overline{1, n-1}, \tag{2.103}$$

де n – задана загальна кількість обчислень функції.

За теоремою новий інтервал невизначеності $[a_{k+1}, b_{k+1}]$ рівний $[\lambda_k, b_k]$, якщо $Q(\lambda_k) > Q(\mu_k)$, і $[a_k, \mu_k]$, якщо $Q(\lambda_k) \le Q(\mu_k)$. В першому випадку, враховуючи попередні формули та $v = n - k$, отримаємо

$$b_{k+1} - a_{k+1} = b_k - \lambda_k = b_k - a_k - \frac{F_{n-k-1}}{F_{n-k+1}}(b_k - a_k) = \frac{F_{n-k}}{F_{n-k+1}}(b_k - a_k). \tag{2.104}$$

В другому випадку маємо:

$$b_{k-1} - a_{k+1} = \mu_k - a_k = \frac{F_{n-k}}{F_{n-k+1}}(b_k - a_k). \tag{2.105}$$

Таким чином, в цих двох випадках довжина інтервалу стискається з коефіцієнтом F_{n-k} / F_{n-k-1}. Тепер покажемо, що на $(k+1)$-ій ітерації $\lambda_{k+1} = \mu_k$ або $\mu_{k+1} = \lambda_k$, так що вимагається лише одне обчислення функції. Припустимо, що

$Q(\lambda_k) > Q(\mu_k)$. Тоді $a_{k+1} = \lambda_k$ та $b_{k+1} = b_k$. Таким чином, враховуючи заміну k на $k+1$, отримаємо

$$\lambda_{k+1} = a_{k+1} + \frac{F_{n-k-2}}{F_{n-k}}(b_{k+1} - a_{k+1}) = \lambda_k + \frac{F_{n-k-2}}{F_{n-k}}(b_k - \lambda_k). \tag{2.106}$$

Підставляючи вираз для λ_k, отримаємо

$$\lambda_{k+1} = a_k + \frac{F_{n-k-1}}{F_{n-k+1}}(b_k - a_k) + \frac{F_{n-k-2}}{F_{n-k}}(1 - \frac{F_{n-k-1}}{F_{n-k+1}})(b_k - a_k). \tag{2.107}$$

Якщо взяти $v = n-k$, то $1-(F_{n-k-1}/F_{n-k+1}) = F_{n-k}/F_{n-k+1}$. Підставляючи цю нерівність у вираз для λ_{k+1}, отримаємо

$$\lambda_{k+1} = a_k + (\frac{F_{n-k-1} - F_{n-k-2}}{F_{n-k+1}})(b_k - a_k). \tag{2.108}$$

Тепер припустимо, що $v = n-k-1$, тоді маємо

$$\lambda_{k+1} = a_k + \frac{F_{n-k}}{F_{n-k+1}}(b_k - a_k) = \mu_k. \tag{2.109}$$

Аналогічно, якщо $Q(\lambda_k) \le Q(\mu_k)$, то $\mu_{k+1} = \lambda_k$. Таким чином, в цих двох випадках на $(k+1)$-ій ітерації вимагається лише одне обчислення. Нарешті, на першій ітерації вимагаються два обчислення функції, а на кожній наступній – лише одне. Відповідно, в кінці $(n-2)$-ї ітерації буде виконано $(n-1)$ обчислення функції. Далі, для $k = n-1$

$$\lambda_{n-1} = \mu_{n-1} = \frac{1}{2}(a_{n-1} + b_{n-1}). \tag{2.110}$$

Відповідно, або $\lambda_{n-1} = \mu_{n-2}$, або $\mu_{n-1} = \lambda_{n-2}$, тобто теоретично не повинні виконуватись нові обчислення функції на цій стадії. Але, щоб забезпечити наступне скорочення інтервалу невизначеності, точка останнього розрахунку трохи переміщається вправо чи вліво від точки $\lambda_{n-1} = \mu_{n-1}$, так що $\frac{1}{2}(b_{n-1} - a_{n-1})$ є довжина кінцевого інтервалу невизначеності $[a_n, b_n]$.

Вибір кількості обчислень функції. На відмінну від дихотомічного пошуку та процедури золотого січення в методі Фібоначчі необхідно, щоб загальне число n обчислень функції було обране заздалегідь. Це пояснюється тим, що точки, в яких проводяться обчислення, визначаються за формулами

$$\lambda_k = a_k + \frac{F_{n-k-1}}{F_{n-k+1}}(b_k - a_k),\ k = \overline{1, n-1}, \tag{2.111}$$

$$\mu_k = a_k + \frac{F_{n-k}}{F_{n-k+1}}(b_k - a_k),\ k = \overline{1, n-1}. \tag{2.112}$$

Відповідно, ці розрахунки залежать від n. Із формул

$$b_{k+1} - a_{k+1} = b_k - \lambda_k = b_k - a_k - \frac{F_{n-k-1}}{F_{n-k+1}}(b_k - a_k) = \frac{F_{n-k}}{F_{n-k+1}}(b_k - a_k). \qquad (2.113)$$

$$b_{k+1} - a_{k+1} = \mu_k - a_k = \frac{F_{n-k}}{F_{n-k+1}}(b_k - a_k). \qquad (2.114)$$

робимо висновок, що довжина інтервалу невизначеності стискається на k-ій ітерації з коефіцієнтом F_{n-k} / F_{n-k+1}. Відповідно, після $(n-1)$ ітерацій, де n – задане загальне число обчислень функції, довжина інтервалу невизначеності скоротиться від $b_1 - a_1$ до $b_n - a_n = (b_1 - a_1)/F_n$. *Звідси n повинно бути вибране так, щоб величина* $(b_1 - a_1)/F_n$ збігалась з відповідною точністю. Нижче наведений алгоритм методу Фібоначчі для мінімізації строго квазіопуклої функції на інтервалі $[a_1, b_1]$.

Алгоритм методу Фібоначчі

Початковий етап. Обрати допустиму кінцеву довжину інтервалу невизначеності $l > 0$ та константу різниці $\varepsilon > 0$. Задати початковий інтервал невизначеності $[a_1, b_1]$ та обрати число n обчислень функції так, щоби $F_n > (b_1 - a_1)/l$. Взяти

$$\lambda_1 = a_1 + (F_{n-2} / F_n)(b_1 - a_1) \text{ та } \mu_1 = a_1 + (F_{n-1} / F_n)(b_1 - a_1). \qquad (2.115)$$

Розрахувати $Q(\lambda_1)$ та $Q(\mu_1)$. Взяти $k = 1$ та перейти до основного етапу.

Основний етап.

Крок 1. Якщо $Q(\lambda_k) > Q(\mu_k)$, то перейти до кроку 2, а якщо $Q(\lambda_k) \le Q(\mu_k)$, то до кроку 3.

Крок 2. Взяти $a_{k+1} = \lambda_k$, $b_{k+1} = b_k$. Потім беремо $\lambda_{k+1} = \mu_k$, $\mu_{k+1} = a_{k+1} + (F_{n-k-1} / F_{n-k})(b_{k+1} - a_{k+1})$. Якщо $k = n-2$, то перейти до кроку 5; в іншому випадку обчислити $Q(\mu_{k+1})$ та перейти до кроку 4.

Крок 3. Беремо $a_{k+1} = a_k$, $b_{k+1} = \mu_k$, $\mu_{k+1} = \lambda_k$, $\lambda_{k+1} = a_{k+1} + (F_{n-k-2} / F_{n-k})(b_{k+1} - a_{k+1})$. Якщо $k = n-2$, то перейти до кроку 5; в іншому випадку обчислити $Q(\lambda_{k+1})$ та перейти до кроку 4.

Крок 4. Замінити k на $k+1$ та перейти до кроку 1.

Крок 5. Взяти $\lambda_n = \lambda_{n-1}$ та $\mu_n = \lambda_n + \varepsilon$. Якщо $Q(\lambda_n) > Q(\mu_n)$, то беремо $a_n = \lambda_n$, $b_n = b_{n-1}$. В іншому випадку, якщо $Q(\lambda_n) \le Q(\mu_n)$, взяти $a_n = a_{n-1}$, $b_n = \lambda_n$. Зупинитися; оптимальний розв'язок знаходиться в інтервалі $[a_n, b_n]$.

Приклад 2.12. Розглянемо наступну задачу: мінімізувати $\lambda^2 + 2\lambda$ при умові $-3 \le \lambda \le 5$.

Розв'язок. Результати подані в табл. 2.6.

Таблиця 2.6

Розрахунки за методом Фібоначчі

k	a_k	b_k	λ_k	μ_k	$Q(\lambda_k)$	$Q(\mu_k)$
1	-3.000	5.000	0.056	1.944	0.115*	7.667*
2	-3.000	1.944	-1.112	0.056	-0.987*	0.115
3	-3.000	0.056	-1.832	-1.112	-0.308*	-0.987
4	-1.832	0.056	-1.112	-0.664	-0.987	-0.887*
5	-1.832	-0.664	-1.384	-1.112	0.853*	-0.987
6	-1.384	-0.664	-1.112	-0.936	-0.987	-0.996*
7	-1.384	-0.936	-1.208	-1.112	-0.957*	-0.987
8	-1.208	-0.936	-1.112	-1.032	-0.987	-0.999*
9	-1.112	-0.936				

Пошук методом Фібоначчі. Припустимо, що потрібно визначити мінімум як можна точніше, тобто з найменшим можливим інтервалом невизначеності, але при цьому можна виконати тільки n обчислень функції. Яким чином потрібно обрати n точок, у яких обчислюється функція? З першого погляду зрозуміло, що не варто шукати розв'язок для всіх точок, одержаних у результаті експерименту. Навпаки, треба спробувати зробити так, щоб значення функції, що отримані в попередніх експериментах, визначали положення наступних точок. Дійсно, знаючи значення функції, ми маємо інформацію про саму функцію і положення її мінімуму і використовуємо цю інформацію у подальшому пошуку. Припустимо, що є інтервал невизначеності (x_1, x_3) і відомо значення функції $f(x_2)$ усередині цього інтервалу (рис. 2.14). Якщо можна обчислити функцію усього один раз у точці x_4, тоді де потрібно помістити точку x_4 для того, щоб одержати найменший можливий інтервал невизначеності?

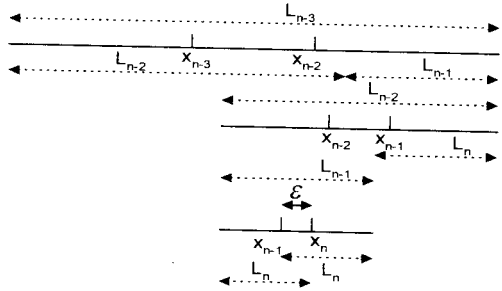


Рис. 2.14. Інтервали для методу Фібоначчі

Нехай $x_2 - x_1 = L$ і $x_3 - x_2 = R$, причому $L > R$, як показано на рис. 2.14, і ці значення будуть фіксовані, якщо відомі x_1, x_2 та x_3. Якщо x_4 знаходиться в інтервалі, то:

1) якщо $f(x_4) < f(x_2)$, то новим інтервалом невизначеності буде (x_1, x_2) довжиною $x_2 - x_1 = L$;

2) якщо $f(x_4) > f(x_2)$, то новим інтервалом невизначеності буде $(x_4; x_3)$ довжиною $x_3 - x_4$.

Оскільки невідомо, яка з цих ситуацій буде мати місце, оберемо x_4 таким чином, щоб мінімізувати найбільшу з довжин $x_3 - x_4$ і $x_2 - x_1$. Досягнути цього можна, зробивши довжини $x_3 - x_4$ і $x_2 - x_1$ рівними, тобто помістивши x_4 всередині інтервалу симетрично щодо точки x_2, що вже лежить усередині інтервалу. Будь-яке інше положення точки x_4 може призвести до того, що отриманий інтервал буде більшим за L. Переміщуючи x_4 симетрично щодо x_2, ми в будь-якому випадку нічим не ризикуємо. Якщо виявиться, що можна виконати ще одне обчислення функції, то варто застосувати описану процедуру до інтервалу (x_1, x_2), у якому вже є значення функції, обчислене в точці x_4, або до інтервалу (x_4, x_3), у якому вже є значення функції, обчислене в точці x_2. Отже, стратегія зрозуміла із самого початку: потрібно помістити таку точку у середині інтервалу невизначеності симетрично щодо точки, яка там вже знаходиться. Парадоксально, але, щоб зрозуміти, як потрібно починати обчислення, необхідно розібратися в тому, як його варто закінчувати.

На n-му обчисленні n-у точку варто розташувати симетрично відносно $(n-1)$-ї точки. Положення цієї точки в принципі залежить від нас. Для того, щоб одержати найбільше зменшення інтервалу на даному етапі, слід розділити навпіл попередній інтервал. Тоді точка x_n буде збігатися з точкою x_{n-1}. Однак, при цьому ми не одержуємо ніякої нової інформації. Звичайно точки x_{n-1} і x_n розміщуються на достатній відстані одна від одної, щоб визначити, у якій половині, лівій або правій, знаходиться інтервал невизначеності. Вони розташовуються на відстані $\varepsilon/2$ з обох сторін від середини відрізка L_{n-1}: можна самим задати величину ε або обрати цю величину рівною мінімально можливій відстані між двома точками. (Припустимо, що в нашому прикладі регулювання температури можливе з інтервалом у $1 \circ C$, тому $\varepsilon = 1$.). Інтервал невизначеності буде мати довжину L_n отже, $L_{n-1} = 2L_n - \varepsilon$ (рис. 2.15).

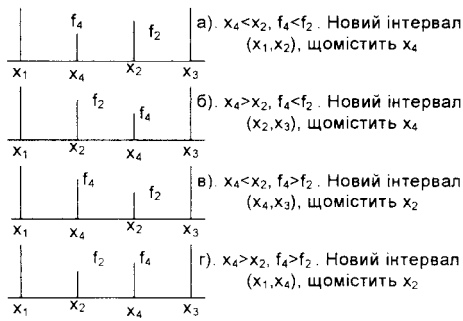


Рис. 2.15. Нові інтервали для методу Фібоначчі

На попередньому етапі точки x_{n-1} і x_{n-2} повинні бути розміщені симетрично усередині інтервалу L_{n-2} на відстані L_{n-1} від кінців цього інтервалу. Отже, $L_{n-2} = L_{n-1} + L_n$ (рис. 2.15, середня частина). З рис. 2.15 зрозуміло, що на передостанньому етапі x_{n-2} залишається як внутрішня точка. Аналогічно

$L_{n-3} = L_n + L_{n-1}$ (рис. 2.15, верхня частина). У загальному випадку $L_{j-1} = L_j + L_{j+1}$ при $1 < j < n$. Отже,

$$L_{n-1} = 2L_n - \varepsilon, \tag{2.116}$$

$$L_{n-2} = L_{n-1} + L_n = 3L_n - \varepsilon, \tag{2.117}$$

$$L_{n-3} = L_{n-2} + L_{n-1} = 5L_n - 2\varepsilon, \tag{2.118}$$

$$L_{n-4} = L_{n-3} + L_{n-2} = 8L_n - 3\varepsilon \quad \text{і т.д.} \tag{2.119}$$

Якщо визначити послідовність чисел Фібоначчі наступним чином: $F_o = 1$, $F_1 = 1$, і $F_k = F_{k-1} + F_{k-2}$, $k = 2,3,...,$ то $L_{n-j} = F_{j+1}L_n - F_{j-1}\varepsilon$, $j = 1,2,...,n-1$. Якщо початковий інтервал (a,b) має довжину $L_1 = (b-a)$, то $L_1 = F_n L_n - \varepsilon F_{n-2}$, тобто $L_n = \frac{L_1}{F_n} + \varepsilon\frac{F_{n-2}}{F_n}$. Отже, провівши n обчислень функції, ми зменшимо початковий інтервал невизначеності в $1/F_n$ разів у порівнянні з його початковою довжиною (незважаючи на ε), і це — найкращий результат. Якщо пошук почався, то його нескладно продовжити, використовуючи описане вище правило симетрії. Отже, необхідно знайти положення першої точки, що на відстані L_2 від одного з кінців початкового інтервалу, причому неважливо, від якого кінця, оскільки друга точка розміщується відповідно до правила симетрії на відстані L_2 від другого кінця інтервалу:

$$L_2 = F_{n-1}L_n - \varepsilon F_{n-3} = F_{n-1}\frac{L_1}{F_n} + \varepsilon\frac{(F_{n-1}F_{n-2} - F_n F_{n-3})}{F_n} = \frac{F_{n-1}}{F_n}L_1 + \frac{(-1)^n\varepsilon}{F_n}. \tag{2.120}$$

Після того як знайдене положення першої точки, числа Фібоначчі більше не потрібні. Використовуване нами значення ε, може визначатися з практичних міркувань. Воно повинно бути менше L_1/F_{n+1}, у іншому випадку ми будемо дарма витрачати час на обчислення функції. Отже, пошук методом Фібоначчі, названий так через появу при пошуку чисел Фібоначчі, є ітераційною процедурою. В процесі пошуку інтервалу (x_1, x_2) з точкою x_2, що знаходиться в цьому інтервалі, наступна точка x_4 завжди обирається такою, що $x_3 - x_4 = x_2 - x_1$ або $x_4 - x_1 = x_3 - x_2$, тобто $x_4 = x_1 - x_2 + x_3$. Якщо $f(x_2) = f_2$ і $f(x_4) = f_4$, то існують чотири випадки (рис. 2.15).

Для розв'язку запропонованої задачі методом Фібоначі була реалізована програма. роздрук якої подано в додатку. Програма реалізована на мові C++ з використанням можливостей цієї мови. Зокрема, в програмі використано такий потужний засіб. як вказівники на функції та масиви фукцій. Це дає можливість легкого формування задач для розв'язування програмою за допомогою комбінації різних наборів обмежень як увигляді рівностей, так і нерівностей. При запуску програми виводяться відомості про метод і задачу, яку цей метод розв'язує, про автора програми. Після цього пропонується вибір конкретної задачі (введення номеру задачі). Розв'язок обраної задачі супроводжується коментарями і повідомленнями про поточний стан

параметрів алгоритму на кожному кроці оптимізації. Користувач має можливість вибору кроку для перегляду, а також можливість детальнішого перегляду процедури пошуку допустимих (майже допустимих) точок. Робота програми завершується відповідним повідомленням з виводом на екран результату: отриманої точки і значення функції мети в ній (у випадку вдалого завершення процесу).

У випадку невдалого пошуку маємо результат, отриманий на останньому виконаному кроці оптимізації. Загалом в програмі реалізовано набір функцій та процедур, кожна з яких реалізує ті чи інші дії в алгоритмі. Так окремо реалізовано функцію оптимізації, яка слугує основою алгоритму. Окремо реалізована функція пошуку допустимих точок, яка в укрупненій схемі співпадає з попередньою, оскільки базується на пошуку методом Фібоначчі. Реалізовано також функцію побудови вершин вихідного багатогранника, яка використовується як при оптимізації, так і при побудові многогранника для пошуку допустимих або майже допустимих точок. Відповідно реалізовано набір функцій, що являють собою обмеження, і різні варіанти функції мети для формування різних задач шляхом утворення різноманітних наборів останніх. Програму реалізовано згідно з вимогами проектування програмного забезпечення. Робота програми перевірена на контрольному прикладі, який з відповідними поясненнями отриманих результатів представлений далі.

Програма 2.3. Метод Фібоначчі

```
#include <iostream.h>
#include <stdlib.h>
#include <stdio.h>
#include <math.h>
#include <conio.h>
void main(void){ int n=35,j;
  double x1,x2,x3,x4,L,R,f2,f4;
  x1=-3; x3=5; j=0;
  clrscr();
  cout<<"Metod Fibonachchi";
  A: j++;cout<<endl<<j<<": ";
  cout<<"["<<x1<<";"<<x3<<"]";
  x2=(3*x3+x1)/4; cout<<"  x2="<<x2;
  x4=x1-x2+x3;    cout<<"  x4="<<x4;
  getch();
  L=x2-x1;
  R=x3-x2;
  if(L>R){ n--;
  f2=pow(x2,2)+2*x2;
  f4=pow(x4,2)+2*x4;
  if(f4<f2){x1=x1;x3=x2;}
  if(f4>f2){x1=x4;x3=x3;}  }
  if(n>0)goto A;
getch();}
```

2.5.7. Методи апроксимації. Квадратична інтерполяція

Попередні методи побудовані на ідеї пошуку малого інтервалу, у якому знаходиться екстремум (мінімум) функції. Тепер застосуємо інший підхід, що ґрунтується на використанні декількох значень функції в певних точках для апроксимації її звичайним поліномом хоча б в невеличкій області значень. Після цього положення мінімуму функції апроксимується положенням мінімуму полінома, оскільки останній обчислити простіше.

Квадратична інтерполяція. Якщо відомі значення функції $f(x)$ у трьох різних точках α,β,γ, рівні відповідно $f_\alpha, f_\beta, f_\gamma$, то функція $f(x)$ може бути апроксимована квадратичною функцією $\varphi(x) = Ax^2 + Bx + C$, де A, B і C визначаються з рівнянь $A\alpha^2 + B\alpha + C = f_\alpha$, $A\beta^2 + B\beta + C = f_\beta$, $A\gamma^2 + B\gamma + C = f_\gamma$.

Після перетворень цих рівнянь отримуємо

$$A = [(\gamma-\beta)f_\alpha + (\alpha-\gamma)f_\beta + (\beta-\alpha)f_\gamma]/\Delta, \tag{2.121}$$

$$B = [(\beta^2-\gamma^2)f_\alpha + (\gamma^2-\alpha^2)f_\beta + (\alpha^2-\beta^2)f_\gamma]/\Delta, \tag{2.122}$$

$$C = [\beta\gamma(\gamma-\beta)f_\alpha + \gamma\alpha(\alpha-\gamma)f_\beta + \alpha\beta(\beta-\alpha)f_\gamma]/\Delta, \tag{2.123}$$

де $\Delta = (\alpha-\beta)(\beta-\gamma)(\gamma-\alpha)$. Зрозуміло, що $\varphi(x)$ буде мати мінімум у точці $x = -B/2A$, якщо $A>0$. Отже, можна апроксимувати точку мінімуму функції $f(x)$ значенням

$$\delta = \frac{1}{2}\left[\frac{(\beta^2-\gamma^2)f_\alpha + (\gamma^2-\alpha^2)f_\beta + (\alpha^2-\beta^2)f_\gamma}{(\beta-\gamma)f_\alpha + (\gamma-\alpha)f_\beta + (\alpha-\beta)f_\gamma}\right]. \tag{2.124}$$

Цей метод може безпосередньо застосовуватись до функцій однієї змінної. Він може бути дуже корисний для здійснення лінійного пошуку в процедурах. В цих процедурах потрібно знайти мінімум функції $f(x)$ в точках прямої $x_0 + \lambda d$, де x_0 — задана точка, а d визначає заданий напрямок. Значення функції $f(x_0 + \lambda d)$ на цій прямій є значеннями функції однієї змінної λ:

$$\varphi(\lambda) = f(x_0 + \lambda d). \tag{2.125}$$

Ідеї та результати, викладені вище, перетворюються в обчислювальні процедури, що описані далі. Припустимо, що задані: унімодальна функція однієї змінної $f(x)$, початкова апроксимація положення мінімуму і довжина кроку H, що є величиною того ж порядку, що і відстань від точки A до точки мінімуму x^* (умова, яку не завжди легко задовольнити). Обчислювальна процедура містить наступні кроки:

1. Обчислити $f(A)$ та $f(A+H)$.

2. Якщо $f(A) < f(A+H)$, то взяти як третю точку $A-H$ і обчислити $f(A-H)$. У протилежному випадку як третю точку взяти $A+2H$ і знайти $f(A+2H)$ (рис. 2.16).

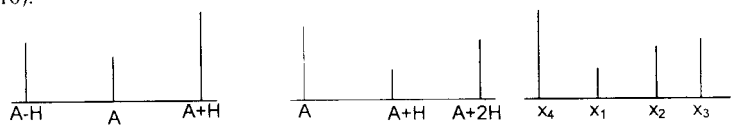


Рис. 2.16. Інтервали для квадратичної інтерполяції

3. Використовуючи координати цих трьох точок, знайти δ із рівняння і обчислити $f(\delta)$.

4. Якщо різниця між найменшим значенням функції і наступним найменшим значенням функції менше заданої точності, то процедура завершується.

Якщо процедура не завершилася на кроці 4, то точка з найбільшим значенням звичайно відкидається, і ми повертаємося на крок 3. Але якщо, залишивши точку з найбільшим значенням функції, ми визначимо кінцеві границі інтервалу, у якому лежить мінімум, то слід дійсно залишити це значення і потім повернутися на крок 3. Наприклад, на рис. 2.16 залишені точки x_1, x_2 і x_4, а не точки x_1, x_2 і x_3. Зауважимо, що якщо точність E задана занадто малою, то α, β, γ а також $f_\alpha, f_\beta, f_\gamma$ будуть дуже близькі один до одного і значення δ (2.124) може стати взагалі недосяжним. Щоб подолати цю проблему, перепишемо рівняння для другої і наступних інтерполяцій:

$$\delta = \frac{1}{2}(\alpha + \beta) + \frac{\frac{1}{2}(f_\alpha - f_\beta)(\beta - \gamma)(\gamma - \alpha)}{(\beta - \gamma)f_\alpha + (\gamma - \alpha)f_\beta + (\alpha - \beta)f_\gamma}. \qquad (2.126)$$

2.6. Аналіз методів та алгоритмів розв'язування задачі нелінійного програмування

Як уже зазначалось для розв'язування НП-задачі існує багато методів, кожен з яких має свої переваги і недоліки. Особливістю є те, що для розв'язку певних різних типів НП-задач доцільно використовувати різні методи й алгоритми. Існує багато ознак та способів класифікації методів нелінійного програмування. Однією з загальних є класифікація запропонована в [76-80, 88, 92-93], яку розглянемо далі.

2.9.1. Класифікація методів нелінійного програмування

1. *Класифікація за деякими аспектами постановки задачі*

1.1 . Характер функції мети (з обмеженнями, без обмежень):

(а) без обмежень;

(б) обмеження у вигляді рівностей;

(в) обмеження у вигляді нерівностей;

(г) обмеження як у вигляді рівностей, так і нерівностей.

1.2 . Дискретні (цілочисельні) змінні; неперервні змінні.

1.3 . Випукле, квадратичне, сепарабельне тощо програмування.

2. *Класифікація за характерними рисами методів розв'язування*

2.1. Методи, які використовують похідні; методи, які не використовують похідні (пошук)

2.2. Аналітичне визначення похідних; чисельне їх визначення

2.3. Методи, які використовують перші похідні; методи, які використовують другі похідні

2.4. Градієнтні методи; методи, які не використовують градієнт

2.5. Градієнтний метод з малим кроком; градієнтний метод з великим кроком

2.6. Одночасна ітерація за всіма змінними в процесі пошуку; релаксація (послідовний пошук кожен раз по одній змінній)

2.7. Методи внутрішньої точки; методи зовнішньої точки

2.8. Детермінований пошук; випадковий пошук

2.9. Початковий вектор знаходиться в допустимій або в недопустимій області

Зупинимося детальніше на деяких запропонованих вище класах і типах алгоритмів, з метою їх порівняння й обґрунтування вибору реалізованого алгоритму методу ковзаючого припущення. Важливим базовим класом методів є методи пошуку екстремуму нелінійних функцій без обмежень. Всі відомі методи пошуку екстремуму функції в залежності від інформації, що використовується для отримання наступної точки, можна розділити на три групи:

1) Прямі методи, які використовують лише значення функції.
2) Методи першого порядку, які використовують додатково значення перших похідних.
3) Методи другого порядку, які потребують визначення других похідних.

Серед методів цього класу можна назвати градієнтні методи (метод найшвидшого спуску), метод других похідних (метод Ньютона), метод спряжених напрямків, методи змінної метрики, метод паралельних дотичних (партан-метод) і т. д. Важливу в цьому класі групу становлять методи, які не використовують похідних, оскільки до таких належить метод ковзаючого припущення. Ефективними методами одномірного пошуку є дихотомічний пошук, метод золотого січення, метод Фібоначі, метод Пауела. Серед методів багатомірного пошуку, що не використовують похідних, можна відзначити прямий пошук (запропонований Хуком і Дживсом), методи Розенброка і Девіса, Свенна, Кемпі, Пауела, метод випадкового пошуку, а також запропонований Нелдером і Мідом метод пошуку по деформованому багатограннику, на основі якого побудований метод ковзаючого припущення.

2.6.2. Формальна постановка задачі з використанням методів прямого пошуку для оптимізації унімодальних функцій якості без обмежень

Формально задача нелінійного програмування (НП-задача) може бути сформульована в загальному вигляді наступним чином: мінімізувати

$$f(x), \quad x \in E^n, . \tag{2.127}$$

при m лінійних і (або) нелінійних обмежень у вигляді рівностей

$$h_j(x) = 0, \quad j = 1,\ldots,m, \tag{2.128}$$

і $(p-m)$ лінійних і (або) нелінійних обмежень у вигляді нерівностей

$$g_j(x) \geq 0, \quad j = m+1,\ldots,p. \tag{2.129}$$

Хоча в деяких часткових випадках обмеження у вигляді рівностей можуть бути явно розв'язані відносно вибраних змінних, які потім виключаються з задачі як незалежні змінні, і в задачі залишаються тільки обмеження у вигляді нерівностей, в більшості випадків обмеження у вигляді рівностей не можуть бути явно розв'язані і тому зберігаються. Інколи зустрічається інше представлення виразів (2.127)-(2.129): мінімізувати

$$\{f(x) \mid x \in R\}, \tag{2.130}$$

де R – область простору змінної x, для якої справедливі умови (2.127)-(2.129). Наприклад

$$R = \{x \mid h_j(x) = 0, \quad g_j(x) \geq 0 \quad \text{для всіх } j\}. \tag{2.131}$$

Знак нерівності у виразі $g_j(x) \geq 0$ може бути змінений на протилежний шляхом множення на -1, що не змінить математичної постановки задачі. Так само задачу мінімізації можна перетворити на задачу максимізації, помноживши на -1 функцію мети $f(x)$. В кожній точці x n–вимірного простору змінних $x_1,\ldots,x_n$ функція $f(x)$ має визначене значення, і, відповідно, цей n-вимірний простір являє собою скалярне поле для критерію оптимальності. В цьому просторі можна накреслити сімейство ліній рівня (еквіпотенціальних гіперповерхонь) для вибраних значень функції $f(x)$. Простір змінних $x_1,\ldots,x_n$ є також скалярним полем для функцій обмежень, для яких також можна зобразити графічно еквіпотенціальні гіперповерхні. За допомогою класичних методів аналізу в загальному випадку неможливо наперед виявити розташування точки x^*, яка відповідає мінімальному (або максимальному) значенню функції $f(x)$, оскільки вона може знаходитися на перетині обмежуючих поверхонь або між ними.

В загальному вигляді процес знаходження розв'язку НП–задачі (2.130)-(2.131) з використанням її геометричної інтерпретації включає наступні етапи:

1. Знаходження області допустимих розв'язків задачі, що визначена співвідношенням (2.131) (якщо вона пуста, то задача не має розв'язку).
2. Будується гіперповерхня $f(x) = H$.
3. Визначається гіперповерхня найнижчого (найвищого) рівня або встановлюється нероз'язність задачі через необмеженість функції (2.130) знизу (зверху).
4. Знаходиться точка області допустимих розв'язків, через яку проходить гіперповерхня найнижчого (найвищого) рівня, і визначається в ній значення функції (2.130).

Задачі лінійного і квадратичного програмування можуть розглядатися як два часткові випадки загальної НП-задачі. Якщо і функція (2.127), і рівняння (2.128), і нерівності (2.129) лінійні, то ми маємо справу з задачею лінійного програмування. Якщо ж цільова функція квадратична, а обмеження лінійні, то має місце задача квадратичного програмування. На відміну від задач лінійного програмування (ЛП), для НП-задач, як вже зазначалось, немає загального методу розв'язку. В ЛП-задачі множина допустимих розв'язків R завжди є випуклою з скінченною кількістю крайніх точок. Використовуючи апарат симплекс-методу і перебираючи тільки крайні точки, завжди за скінченне число кроків можливо знайти оптимальний розв'язок. На противагу, в НП-задачах, якщо в обмеженнях є нелінійність, то випуклість допустимої множини і скінченність числа її крайніх точок необов'язкові. Через ці особливості й виникають основні складності розв'язування НП-задач.

В області нелінійного програмування велика увага приділяється визначенню необхідних й достатніх умов екстремальності. Вектор $x^* = [x^*_1, \ldots, x^*_n]^T$, який задовольняє умовам (2.127)-(2.129) або (2.130), (2.131), називається оптимальною

точкою, а відповідне значення $f(x^*)$ – оптимальне значення цільової функції. Пара x^* та $f(x^*)$ складають оптимальний розв'язок. Можуть існувати різні типи оптимальних розв'язків, якщо функція мети має декілька екстремумів (т.зв. мультимодальна). Глобальний оптимальний розв'язок – це найменше значення $f(x)$, тоді як локальний (або відносний) оптимальний розв'язок – найменше значення $f(x)$ в деякому околі вектора x (рис. 2.17). Як для глобального, так і для локального мінімуму справедливе відношення: $f(x^*) \le f(x)$, але для глобального розв'язку це співвідношення виконується для всіх $x \in E^n$, тоді як для локального мінімуму це має місце тільки для деякої малої області ξ, де $\|x - x^*\| < \zeta$.



Рис. 2.17. Локальні та глобальні оптимуми при відсутності обмежень

Більшість алгоритмів дають тільки локальні оптимальні розв'язки, так як на кожному етапі розв'язування при русі до x^* вони залежать в основному від локальних властивостей функції мети і обмежень. Щоправда, деякі алгоритми, в тому числі й ті, які розглянуті тут, найімовірніше, закінчаться в точці глобального оптимуму, через те що тут при пошуку розглядається широкий діапазон зміни x, однак збіжність до глобального оптимуму не може бути гарантована без додаткових відомостей про природу функції мети і обмежень. Структура НП-задачі в загальному випадку така, що повністю критерій оптимальності ще не розроблений. Внаслідок цього важливим є розгляд окремих випадків і умов, при яких вектор x є розв'язком задачі нелінійного програмування. Результатами проведення досліджень в цьому напрямку є цілий ряд сформульованих понять і теорем класичного математичного аналізу, які лежать в основі диференційних методів пошуку екстремуму. Окремо слід розглянути теорему Куна-Такера, яка має виключно важливе значення для теорії нелінійного програмування. В ній сформульовані умови оптимальності для задачі нелінійного програмування заданої наступним чином:

$$f(x) \Rightarrow \min, \quad g_i(x) \le 0, i = \overline{1,m} \tag{2.132}$$

Теорема Куна-Такера може бути узагальнена й для інших випадків НП-задачі. Тут наведений її варіант для задачі випуклого програмування.

Терема 2.1(Теорема Куна-Такера). Нехай $f(x), g_i(x), i=\overline{1,m}$, мають неперервні часткові похідні на деякій відкритій множині R^n, яка містить x^*. Якщо x^* є точкою мінімуму функції $f(x)$ при обмеженнях $g_i(x) \le 0, i=\overline{1,m}$, які відповідають умовам регулярності у вигляді лінійної незалежності векторів $\nabla g_i(x^*)$, то існують такі невід'ємні множники Лагранжа $\lambda_1, \lambda_2, \ldots, \lambda_m$, такі що

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g(x^*) = 0; \quad \sum_{i=1}^{m} \lambda_i g(x^*) = 0, \quad \lambda_i \ge 0. \tag{2.133}$$

Доведення цієї теореми з використанням наслідків і понять інших базових лем і теорем з усіма необхідними поясненнями і обґрунтуваннями можна знайти в книгах [76-80, 88, 92-93].

2.6.3. Визначення та основні властивості опуклих функцій

Нехай $f: S \to E_1$, де S – непорожня *опукла* множина в E_n. Тоді функція f опукла на S, якщо

$$f[\lambda x_1 + (1-\lambda)x_2] \le \lambda f(x_1) + (1-\lambda) f(x_2) \tag{2.134}$$

для будь-яких $x_1, x_2 \in S$ і $\lambda \in (0,1)$. Функція f *строго опукла* на S, якщо

$$f[\lambda x_1 + (1-\lambda)x_2] < \lambda f(x_1) + (1-\lambda) f(x_2). \tag{2.135}$$

Функція $f: S \to E_1$ називається ввігнутою (строго ввігнутою), якщо $-f$ опукла (строго опукла) на S. Розглянемо геометричну інтерпретацію сформульованих понять. Нехай x_1 та x_2 – дві різні точки із області ефективності f. Розглянемо точку $\lambda x_1 + (1-\lambda)x_2$, де $\lambda \in (0,1)$. Зауважимо, що $\lambda f(x_1) + (1-\lambda) f(x_2)$ дає середньозважене значення функцій $f(x_1)$ та $f(x_2)$, а $f[\lambda x_1 + (1-\lambda)x_2]$ дає значення f в середній точці $\lambda x_1 + (1-\lambda)x_2$. Таким чином, для опуклої функції значення f в точках відрізку, що з'єднює x_1 та x_2, не перевищує середньозважене (з тим же λ) значення величин $f(x_1)$ та $f(x_2)$. Наступні функції є опуклими:

$$f_1(x) = 3x + 4, \; f_2(x) = |x|, \; f_3(x) = x^2 - 2x, \; f_4(x) = -x^{\frac{1}{2}}, \; x \ge 0,$$

$$f_5(x_1, x_2) = 2x_1^2 + x_2^2 - 2x_1x_2, \; f_6(x_1, x_2, x_3) = x_1^4 + 2x_2^2 + 3x_3^2 - 4x_1 - 4x_2x_3.$$

Неперервність опуклих функцій. Важливою властивістю опуклих функцій є їх неперервність всередині області ефективності.

Терема 2.2. Нехай S – непорожня опукла множина в E_n, $f: S \to E_1$ – опукла функція. Тоді f неперервна у внутрішніх точках множини S.

В багатьох випадках припущення про опуклість функції може бути замінено більш слабкими припущеннями про квазіопуклість чи псевдоопуклість функції.

Нехай S – непорожня опукла множина в E_n (рис. 2.18). Кажуть, що функція $f: S \to E_1$ квазіопукла, якщо для всіх $x_1, x_2 \in S$ виконується нерівність

$$f[\lambda x_1 + (1-\lambda)x_2] \le \max\{f(x_1), f(x_2)\} \tag{2.136}$$

при будь-якому $\lambda \in (0,1)$.

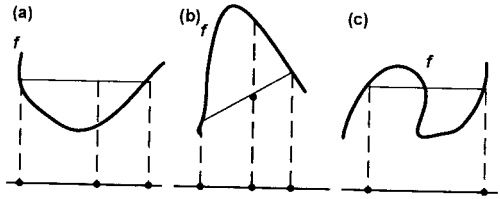
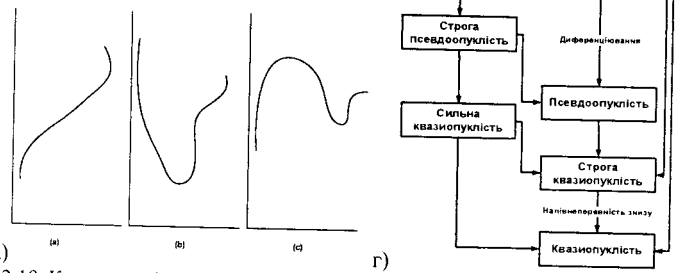


Рис. 2.18. Опуклі та ввігнуті функції. a - опукла функція;b - ввігнута функція;c - функція, що не є ні опуклою, ні ввігнутою.

Функція f строго квазиопукла (рис. 2.19а), якщо нерівність виконується як строга при $f(x_1) \neq f(x_2)$. Якщо ж нерівність виконується як строга для будь-яких $x_1 \neq x_2$, то кажуть, що f – сильно квазиопукла функція.



а) г)

Рис. 2.19. Квазиопуклість та псевдоопуклість. а - функція квазиопукла та псевдоопукла; б - функція квазиопукла, але не псевдоопукла; в - функція не квазиопукла і не псевдоопукла г - зв'язок різних типів опуклості

Нехай S – непорожня відкрита опукла множина в E_n. Кажуть, що функція $f: S \to E_1$ – псевдоопукла функція, якщо для будь-яких $x_1, x_2 \in S$, таких, що $\nabla f(x_1)^T (x_2 - x_1) \geq 0$, виконується нерівність $f(x_2) \geq f(x_1)$. Якщо для будь-яких різних x_1 та x_2 із S, таких, що $\nabla f(x_1)^T (x_2 - x_1) \geq 0$, справедлива нерівність $f(x_2) > f(x_1)$, то f – строго псевдоопукла функція.

Наведемо короткі відомості важливих властивостей різних типів опуклих функцій (рис. 2.19б). Всюди тут $f: S \to E_1$ та S – непорожня опукла множина в E_n.

Строго опуклі функції

- Функція f неперервна в $\operatorname{int} S$.
- Множина $\{(x, y): x \in S, y \geq f(x)\}$ опукла.
- Множина $\{x \in S: f(x) \leq \alpha\}$ опукла для кожного дійсного α.
- Диференційовна функція f строго опукла тоді і тільки тоді, коли $f(x) > f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x})$ при всіх $x, \bar{x} \in S$.
- Нехай f двічі диференційована. Тоді, якщо матриця Гессе $H(x)$ позитивно визначена для кожного $x \in S$, тоді f – строго опукла функція. Крім того, якщо f

строго опукла, тоді матриця Гессе $H(x)$ позитивно визначена для кожного $x \in S$.

- Кожний локальний мінімум функції f на опуклій множині $X \subset S$ є єдиним глобальним мінімумом.
- Якщо $\nabla f(\bar{x}) = 0$, то $\bar{x}$ – єдина точка глобального мінімуму функції f на множині S.
- *Функція f досягає максимуму на компактній багатограній множині $X \subset S$ в* екстремальній точці цієї множини.

Опуклі функції

- Функція f неперервна в $\operatorname{int} S$.
- Функція f опукла тоді і лише тоді, коли множина $\{(x,y): x \in S, y \geq f(x)\}$ опукла.
- Множина $\{x \in S: f(x) \leq \alpha\}$ опукла для кожного дійсного α.
- Диференційовна функція f опукла тоді і тільки тоді, коли $f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x})$ при всіх $x, \bar{x} \in S$.
- Двічі диференційована функція f опукла тоді і лише тоді, коли матриця Гессе $H(x)$ позитивно визначена для всіх $x \in S$.
- Кожний локальний мінімум функції f на опуклій множині $X \subset S$ є глобальним мінімумом.
- Якщо $\nabla f(\bar{x}) = 0$, то $\bar{x}$ – точка глобального мінімуму функції f на множині S.
- Функція f досягає максимуму на компактній багатограній множині $X \subset S$ в екстремальній точці цієї множини.

Псевдоопуклі функції

- Множина $\{x \in S: f(x) \leq \alpha\}$ опукла для кожного дійсного α.
- Кожний локальний мінімум функції f на опуклій множині $X \subset S$ є глобальним мінімумом.
- Якщо $\nabla f(\bar{x}) = 0$, то $\bar{x}$ – точка глобального мінімуму функції f на множині S.
- Функція f досягає максимуму на компактній багатогранній множині $X \subset S$ в екстремальній точці цієї множини.

Квазіопуклі функції

- Функція f квазіопукла тоді і тільки тоді, коли множина $\{x \in S: f(x) \leq \alpha\}$ опукла для будь-якого дійсного α.
- Функція f досягає максимуму на компактній багатограній множині $X \subset S$ в екстремальній точці цієї множини.
- Диференційовна функція f на множині S квазіопукла тоді і тільки тоді, коли із $x_1, x_2 \in S$ та $f(x_1) \leq f(x_2)$ випливає, що $\nabla f(x_2)^T (x_1 - x_2) \leq 0$.
- Локальним мінімумом строго квазіопуклої функції на опуклій множині $X \subset S$ є також глобальним. Якщо, крім того, функція сильно квазіопукла, то мінімум досягається в єдиній точці. Якщо функція f є і строго квазіопуклою і напівнеперервною знизу, то вона квазіопукла, так що вона володіє всіма перерахованими властивостями.

2.7. Методи прямого пошуку для функцій n змінних

На розробку методів прямого пошуку для визначення мінімуму функцій n змінних було витрачено багато зусиль. Методи прямого пошуку є методами, у яких використовуються тільки значення функції. Розглянемо докладно два з них, які ефективні і застосовуються в багатьох програмах. Розглянемо функцію двох змінних. Її лінії постійного рівня подані на рис. 2.20, а мінімум лежить уточці (x_1^*, x_2^*) Лінією постійного рівня називається крива в двовимірному перерізі простору параметрів (в цьому випадку – в площині (x_1,x_2)), значення функції на якій – константа (рис. 2.20).

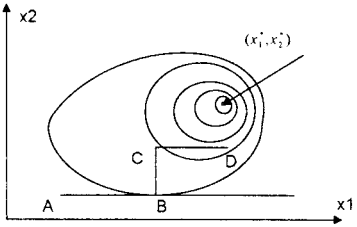


Рис. 2.20. Метод прямого пошуку для функцій n змінних

Для перевірки працездатності запропоновано декілька функцій, що через свої властивості є достатньо "незручними", а тому – тестовими для методів оптимізації. Нижче наведено декілька прикладів таких функцій.

1. Функція Розенброка (рис. 2.21):

$$f(x_1, x_2) = 100\,(x_2 - x_1^2)^2 + (1 - x_1)^2; \quad x^* = (1;1)\,. \tag{2.137}$$

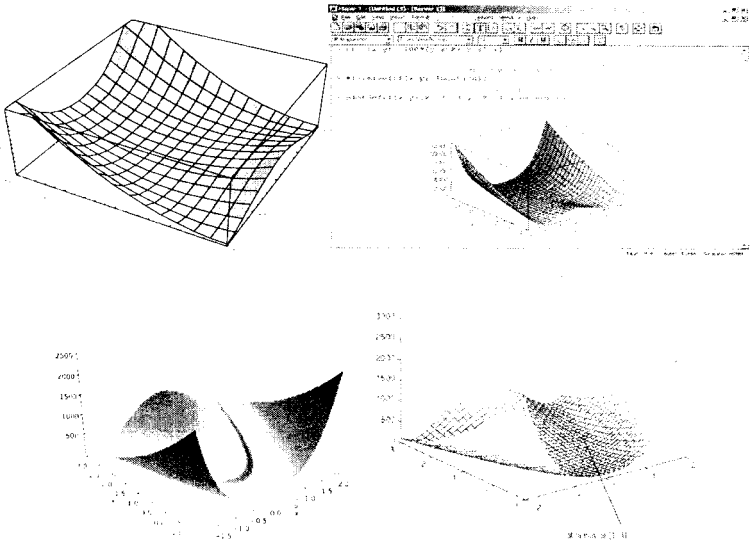


Рис. 2.21. Функція Розенброка

2. Функція Пауела (рис. 2.22):

$$f(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4;\ x^* = (0;0;0;0). \qquad (2.138)$$

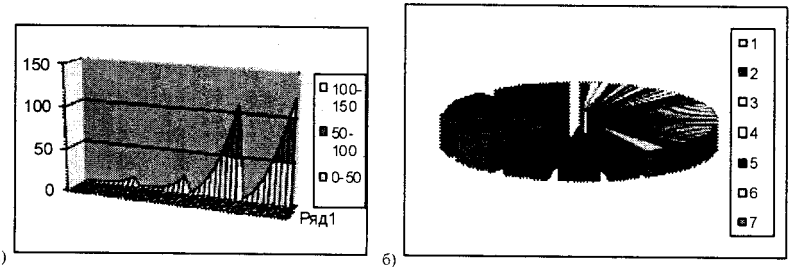


Рис. 2.22. Функція Пауела

3. Двовимірна експоненційна функція (рис. 2.23):

$$f(x_1, x_2) = \sum_a [(e^{-ax_1} - e^{-ax_2}) - (e^{-a} - e^{-10a})]^2, \qquad (2.139)$$

де $a = 0{,}1(0{,}1)1^*$; $x^* = (1;10)$.

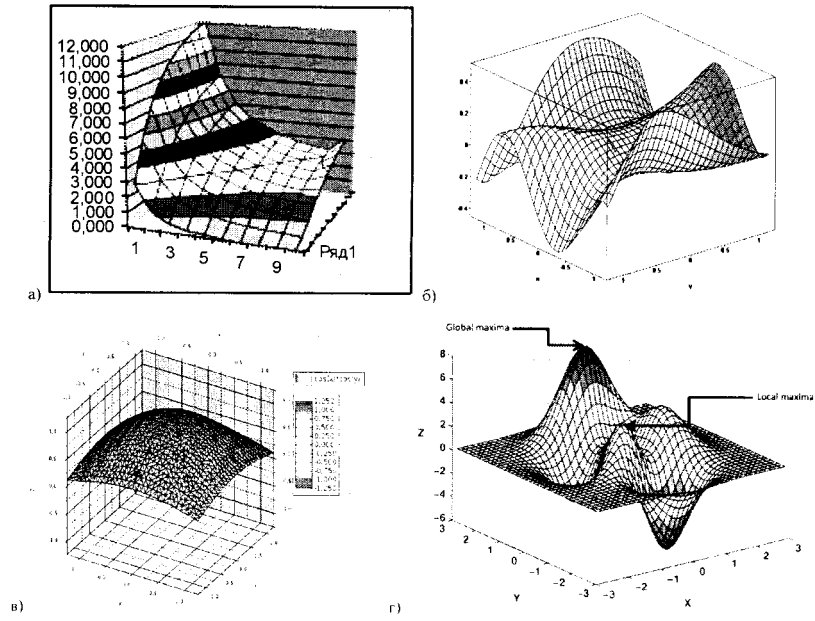


Рис. 2.23. Двовимірна експоненційна функція

Будь-яка серйозна оптимізаційна процедура повинна ефективно вирішувати ці та інші тестові задачі.

2.7.1. Метод покоординатного спуску

Найпростішим методом пошуку є метод покоординатного спуску. З точки A ми здійснюємо пошук мінімуму уздовж напрямку осі x_1 і, таким чином, знаходимо точку B, в якої дотична до лінії постійного рівня паралельна до осі x_1. Потім, здійснюючи пошук із точки B в напрямку осі x_2, отримуємо точку C. Здійснюючи пошук паралельно осі x_1, одержуємо точку D, тощо. Таким чином, ми приходимо до оптимальної точки. Будь-який з одновимірних методів, які описані у попередньому розділі, може бути використаний для пошуку екстремуму вздовж осі. Очевидно, що цю ідею можна застосувати для функцій n змінних.

Теоретично цей метод ефективний у випадку єдиного мінімуму функції, але на практиці він здається занадто повільним. Тому були розроблені більш складні методи, що використовують більше інформації на основі вже отриманих значень функції.

2.7.2. Метод Хука - Дживса

Цей метод був розроблений у 1961 році, але і дотепер він є дуже ефективним та оригінальним. Пошук складається з послідовності досліджуваних кроків навколо базової точки, за яким у випадку успіху реалізується пошук за зразком.

Алгоритм методу Хука – Дживса утворює наступну послідовність кроків:

1. Обрати початкову базову точку b_1 і крок довжиною h_j для кожної змінної $x_j, j = 1,2,...,n$. У деяких програмах для кожної змінної використовується крок h, проте зазначена вище модифікація теж може виявитися корисною.

2. Обчислити $f(x)$ у базовій точці b_1 з метою одержання даних про локальну поведінку функції $f(x)$. Ці дані будуть використовуватися для знаходження найвідповіднішого напрямку пошуку за зразком, за допомогою якого можна сподіватися досягти більшого зменшення значення функції (рис. 2.24).

Функцюя $f(x)$ у базовій точці b_1 знаходимо наступним чином:

- Обчислюється значення функції $f(b_1)$ у базовій точці b_1
- Значення кожної змінної (по черзі) змінюється додаванням довжини кроку.

Таким чином, ми обчислюємо значення функції $f(b_1 + h_1 e_1)$, де e_1 – одиничний вектор у напрямку осі x_1. Якщо це призводить до зменшення значення функції, то b_1 змінюємо на $b_1 + h_1 e_1$. У протилежному випадку обчислюється значення функції $f(b_1 - h_1 e_1)$, і якщо її значення зменшилось, то b_1 заміняємо $b_1 - h_1 e_1$. Якщо ні один із пророблених кроків не призводить до зменшення значення функції, то точка b_1 залишається незмінною і розглядаються зміни в напрямку осі x_2, тобто знаходиться значення функції $f(b_1 + h_2 e_2)$ тощо. (рис. 2.25). Коли будуть розглянуті усі n змінних, отримаємо нову базову точку b_2.

- Якщо $b_2 = b_1$, тобто зменшення функції не було досягнуто, таке ж дослідження повторюється навколо тієї ж базової точки b_1, але зі

зменшеною довжиною кроку. На практиці задовільним є зменшення кроку (кроків) у десять разів від початкової довжини.

- Якщо $b_2 \neq b_1$ то виконується пошук за зразком.

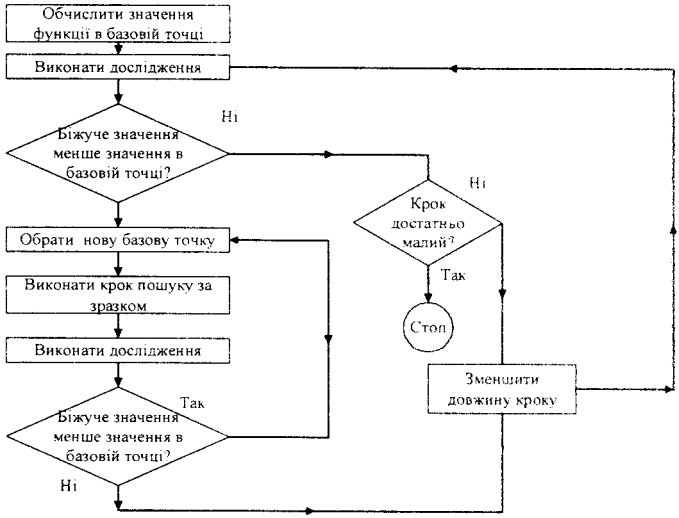


Рис. 2.24. Блок-схема методу Хука-Дживса

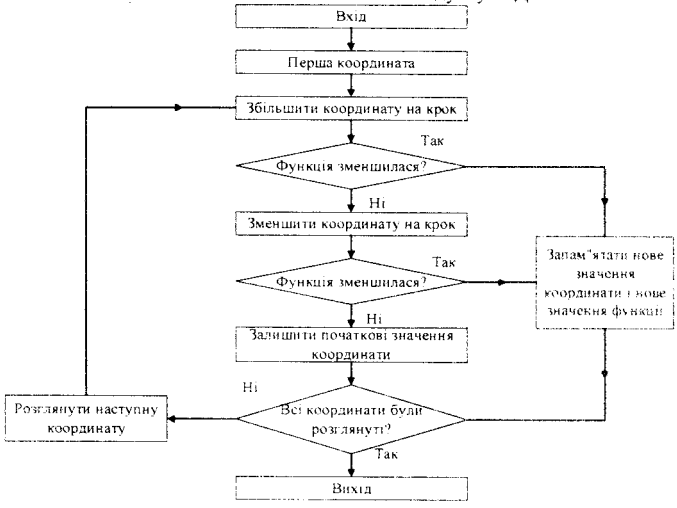


Рис. 2.25. Блок-схема методу Хука-Дживса

3. Під час пошуку за зразком використовується інформація, що отримана в процесі дослідження, і мінімізація функції завершується пошуком у напрямку, що задається зразком. Ця процедура здійснюється в такий спосіб:

а) Доцільно рухатися з базової точки b_2 в напрямку $b_2 - b_1$, оскільки пошук у цьому напрямку вже привів до зменшення значення функції. Тому обчислимо функцію в точці зразка $P_1 = b_1 + 2(b_2 - b_1)$. У загальному випадку $P_i = b_i + 2(b_{i+1} - b_i)$.

б) Далі дослідження слід продовжувати довкола точки $P_1(P_i)$.

в) Якщо найменше значення на кроці 3б) менше значення в базовій точці b_2 (у загальному випадку b_{i+1}), то одержуємо нову базову точку b_3 (b_{i+2}), після чого повторюємо крок 3а. У протилежному випадку пошук за зразком із точки $b_2(b_{i+1})$ не здійснюємо, а продовжуємо дослідження в точці $b_2(b_{i+1})$.

4. Завершуємо процес, коли довжина кроку (довжини кроків) буде зменшена до заданого малого значення.

2.7.3. Метод Нелдера - Міда

Цей метод виявився досить ефективним і легко реалізовуваним на ПК. В його основі лежить симплексний пошук Спендлі, Хекста і Хімсворта, розроблений для статичного планування експерименту. Відомо що регулярні многогранники в E^n є симплексами. Наприклад, у випадку двох змінних регулярний симплекс являє собою трикутник; у випадку трьох змінних він буде являти собою тетраедр. При пошуку мінімуму функції мети $f(x)$ пробні вектори x можуть бути вибрані в точках $f(x)$, які знаходяться в вершинах симплексу. З аналітичної геометрії відомо, що координати регулярного симплексу визначаються наступною матрицею D, в якій стовпцями є вершини, пронумеровані від 1 до n+1, а рядки, які пронумеровані від 1 до n – координати.

$$D = \begin{bmatrix} 0 & d_1 & d_2 & \dots & d_2 \\ 0 & d_2 & d_1 & \dots & d_2 \\ 0 & d_2 & d_2 & \dots & d_2 \\ 0 & \dots & \dots & \dots & \dots \\ 0 & d_2 & d_2 & \dots & d_1 \end{bmatrix}, \qquad (2.140)$$

де D – матриця $n\times(n+1)$ при

$$d_1 = \frac{t}{n\sqrt{2}}(\sqrt{n+1} + n - 1), \quad d_2 = \frac{t}{n\sqrt{2}}(\sqrt{n+1} - 1), \qquad (2.141)$$

де t – відстань між двома вершинами.

Функція мети може бути обчислена в кожній з вершин симплексу. З вершини, де функція мети максимальна, проводиться проектуюча пряма через центр ваги симплексу. Далі ця точка замінюється на спроектовану, що приводить до побудови нового т.з. відображеного симплексу. Продовження цієї процедури, в якій кожен раз вилучається вершина, в якій функція мети максимальна, а також використання правил зменшення розміру симплексу і попередження циклічного руху в околиці екстремуму

дозволяють здійснити пошук, котрий не використовує похідні і в якому величина кроку на довільному етапі фіксована, а напрямок пошуку можна змінювати. Практичні труднощі при використанні регулярних симплексів, а саме: відсутність прискорення пошуку і труднощі при проведенні пошуку на викривлених "ярах" та "хребтах", привели до необхідності деяких покращень, які запропонували Нелдер і Мід в своєму методі. Він передбачає, що симплекс може міняти форму й розміри. В цьому методі мінімізується функція n незалежних змінних з використанням $n+1$ вершини многогранника в E^n. Детально цей алгоритм може бути описаний наступним чином.

Нехай $x_i^{(k)} = \left[x_{i1}^{(k)}, \ldots, x_{ij}^{(k)}, \ldots, x_{in}^{(k)}\right]^T$, $i=1, \ldots, n+1$, є i-ю вершиною (точкою) $f(x)$ на k-му етапі пошуку, $k=0,1, \ldots$, і значення функції мети в ній $f(x_i^{(k)})$. Відзначимо вектори многогранника, які дають максимальне й мінімальне значення $f(x)$. Визначимо

$$f(x_h^{(k)}) = \max\{f(x_1^{(k)}), \ldots, f(x_{n+1}^{(k)})\}, \tag{2.142}$$

$$f(x_l^{(k)}) = \min\{f(x_1^{(k)}), \ldots, f(x_{n+1}^{(k)})\}, \tag{2.143}$$

де $x_h^{(k)} = x_j^{(k)}$ і $x_l^{(k)} = x_i^{(k)}$. Визначимо координати центра ваги $x_{n+2}^{(k)}$ всіх вершин, окрім $x_h^{(k)}$, за наступною формулою

$$x_{n+2,j}^{(k)} = \frac{1}{n}\left[\left(\sum_{i=1}^{n+1} x_{ij}^{(k)}\right) - x_{hj}^{(k)}\right], \tag{2.144}$$

де $j=1, \ldots, n$, – координатні напрямки.

Початковий многогранник вибирається у вигляді регулярного симплексу (хоча це не обов'язково). Процедура пошуку оптимальної вершини складається з наступних операцій:

1. **Відображення** – проектування $x_h^{(k)}$ через центр ваги у відповідності до співвідношення

$$x_{n+3}^{(k)} = x_{n+2}^{(k)} + \alpha(x_{n+2}^{(k)} - x_h^{(k)}), \tag{2.145}$$

де $\alpha>0$ є коефіцієнтом відображення; $x_{n+2}^{(k)}$ – центр ваги, $x_h^{(k)}$ – вершина з максимальним значенням функції мети на k-му етапі.

2. **Розтяг**. Ця операція полягає в наступному: якщо $f(x_{n+3}^{(k)}) \le f(x_l^{(k)})$, то вектор $(x_{n+3}^{(k)} - x_{n+2}^{(k)})$ розтягується відповідно до співвідношення

$$x_{n+4}^{(k)} = x_{n+2}^{(k)} + \gamma(x_{n+3}^{(k)} - x_{n+2}^{(k)}), \tag{2.146}$$

де $\gamma>1$ являє собою коефіцієнт розтягу. Якщо $f(x_{n+4}^{(k)}) < f(x_l^{(k)})$, то $x_h^{(k)}$ замінюється на $x_{n+4}^{(k)}$ і процедура продовжується знову з опреції 1 при $k=k+1$. В протилежному випадку $x_h^{(k)}$ заміняється на $x_{n+3}^{(k)}$ і також здійснюється перехід до операції 1 при $k=k+1$.

3. **Стиснення**. Якщо $f(x_{n+3}^{(k)}) > f(x_i^{(k)})$ для всіх $i \ne h$, то вектор $(x_h^{(k)} - x_{n+2}^{(k)})$ стискається у відповідності до формули

$$x_{n+5}^{(k)} = x_{n+2}^{(k)} + \beta(x_h^{(k)} - x_{n+2}^{(k)}), \qquad (2.147)$$

де $0<\beta<1$ являє собою коефіцієнт стиснення. Потім $x_h^{(k)}$ замінюється на $x_{n+5}^{(k)}$ і відбувається повернення до операції 1.

4. **Редукція**. Якщо $f(x_{n+3}^{(k)}) > f(x_h^{(k)})$, вектори $(x_i^{(k)} - x_l^{(k)}), i = \overline{1, n+1}$ зменшуються в 2 рази з відліком від $x_l^{(k)}$ за формулою

$$x_i^{(k)} = x_l^{(k)} + 0.5(x_i^{(k)} - x_l^{(k)}), \quad i = 1,...,n+1. \qquad (2.148)$$

Після чого повертаємося до опереції 1 для продовження пошуку. Критерієм закінчення пошуку є виконання наступної умови

$$\left\{\frac{1}{n+1}\sum_{i=1}^{n+1}\left[f(x_i^{(k)}) - f(x_{n+2}^{(k)})\right]^2\right\}^{1/2} \le \varepsilon, \qquad (2.149)$$

де ε – довільне мале число.

Таким чином, за допомогою операцій розтягу й стиснення розміри і форма многогранника маштабуються так, щоб вони задовольняли топологію розв'язуваної задачі. Щодо параметрів α,β, та γ, Нелдер і Мід рекомендують наступні їх значення для оптимізації без обмежень: $\alpha=1, \beta=0.5$ та $\gamma=2$.

Оскільки в найбільш загальній постановці НП-задачі передбачено наявність обмежень, то потрібно розглянути відповідні методи. При наявності обмежень в НП-задачі виникають нові труднощі. Відомі обчислювальні процедури для розв'язування задач з обмеженнями грунтуються на одному з наступних підходів.

Поширення апарату лінійного програмування на нелінійні умови шляхом багатократного використання процедури лінійної апроксимації (методи апроксимуючого програмування). Перетворення НП-задачі з обмеженнями в еквівалентну послідовність задач безумовної оптимізації шляхом введення до розгляду штрафних функцій (методи штрафних функцій). Використання ковзаючих допусків, які дозволяють оперувати в процесі розв'язуваннія задачі як з допустимими, так і з недопустимим векторами в просторі рішень.

До методів, які використовують першу процедуру, можна віднести методи апроксимуючого лінійного програмування (МАП, ПОП НЛП), проективні методи (Розена, Заутендайка), метод зведеного градієнту. Другий підхід використовують різні варіанти методу штрафних функцій (параметричні – МПБМ, непараметричні – методи внутрішньої точки, методи зовнішньої точки, комбіновані).

Використання третього підходу реалізовано в методі ковзаючого допуску, який розглянутий і реалізований програмно в цій роботі. Для обгрунтування вибору саме цього методу для роз'язування НП-задачі з обмеженнями як у вигляді рівностей, так і у вигляді нерівностей, наведемо порівняння й оцінки, отримані практично при аналізі різноманітних методів розв'язування такого типу задач, що подані в [76-80, 88, 92-93]. Доцільність вибору того чи іншого зі згаданих вище алгоритмів з метою його практичного використання визначається ефективністю цього алгоритму при розв'язуванні конкретного класу задач нелінійного програмування за допомогою ПК. Як і очікувалось, ні один з алгоритмів нелінійного програмування не може вважатися

найкращим в порівнянні з іншими алгоритмами при розв'язуванні довільних НП-задач і при довільних обставинах. Найкращим критерієм оцінки можна вважати "комплексний критерій". До числа важливих критеріїв відносяться наступні:

1) час, який необхідний для реалізації серії обчислювальних процедур;
2) ступінь складності задачі;
3) точність розв'язку;
4) простота практичного використання алгоритму;
5) простота програми, яка реалізує алгоритм;
6) можливість розв'язування задач, що становлять практичний інтерес.

Тестові задачі показали, що алгоритм ковзаючого допуску придатний для розв'язування широкого класу задач, за винятком деяких окремих задач. Слід зауважити, що при використанні алгоритму ковзаючого допуску було затрачено дуже багато часу на пошук точки, яка задовільняє обмеження-рівності. Оскільки в цьому випадку ми маємо справу з прямим пошуком, то для реалізації алгоритму на ПК знаходити похідні в аналітичному вигляді не потрібно. Тому, часові затрати на підготовку для розв'язування всіх задач виявляються незначними. Більш того, в процесі виконання робіт по підготовці помилки майже виключені. З іншого боку, затрати машинного часу при розв'язуванні задач великої розмірності виявляються більш значними, ніж у випадку оптимізації за допомогою алгоритмів МОПГ та НЛП. Щодо програмної реалізації алгоритму ковзаючого припуску (з власного досвіду): слід зазначити значні затрати. Деякі інші переваги й особливості цього методу будуть згадані нижче, в описі самого алгоритму і стратегії.

Методи прямого пошуку є методами, у яких використовуються лише значення функції і орієнтовані на пошук екстремумів унімодальних функцій. Ідеї та складові методів прямого пошуку використовуються і при побудові складніших методів пошуку оптимальних рішень. Найефективнішим є метод, що використовує числа Фібоначі, однак з точки зору зручності реалізації доцільно використовувати метод золотого перетину, який незначно поступається методу Фібоначі. Окрім цих методів, використовують також апроксимаційні, що використовують ідею квадратичної інтерполяції. Методи прямого пошуку застосовуються також і для пошуку екстремумів функцій n змінних. Найпростішим з них є метод покоординатного спуску (у випадку пошуку мінімуму), який почергово по кожній зі змінних реалізує одновимірний прямий пошук, і в якості процедури одновимірного пошуку найчастіше використовується метод золотого перетину. Однак з практичної точки зору цей метод є доволі повільним, а тому були розроблені ефективніші методи, одним з найвідоміших серед яких є метод Хука-Дживса. Цей метод використовує інформацію про локальну поведінку функції для обрання найвідповідальнішого напрямку пошуку за зразком.

Метод Нелдера-Міда (або пошук по деформованому многограннику) є розвитком симплексного методу Спендлі, Хекста і Хімсворта. Множина $(n+1)$ рівновіддалених точок в n-вимірному просторі називається регулярним симплексом. Ця конфігурація розглядається в методі Спендлі, Хекста і Хімсворта. Отже, у двовимірному просторі симплексом є рівносторонній трикутник, а в тривимірному просторі— правильний тетраедр. Ідея методу полягає в порівнянні значень функції в $(n+1)$ вершинах симплекса і переміщенні симплекса в напрямку оптимальної точки за допомогою ітераційної процедури. У симплексному методі, запропонованому спочатку, регулярний симплекс використовувався на кожному етапі. Нелдер і Мід запропонували декілька

модифікацій цього методу, з неправильними симплексами. У результаті утворився дуже надійний метод прямого пошуку, що є одним з найефективніших, якщо n <6. У методі Спендлі, Хекста і Хімсворта симплекс переміщається за допомогою трьох основних операцій: відбивання, розтягнення і стискання. Зміст цих операцій стане зрозумілим при розгляді кроків процедури.

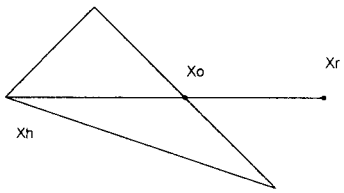
Алгоритм методу Нелдера-Міда.

А. Знайдемо значення функції $f_1 = f(x_1), f_2 = f(x_2) ... f_{n+1} = f(x_{n+1})$ у вершинах симплекса.

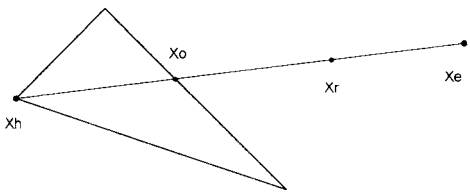
Б. Знайдемо найбільше значення функції f_h, наступне за найбільшим значенням функції f_g, найменше значення функції f_l та x_h, x_g і x_l, які відповідають їм.

В. Знайдемо центр ваги всіх точок, за винятком точки x_h. Нехай центром ваги буде $x_0 = \frac{1}{n}\sum_{i \neq h} x_i$, обчислимо $f(x_0) = f_0$.

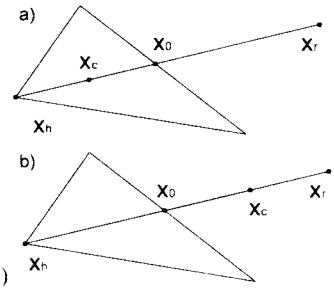
Г. Зручніше всього почати переміщення від точки x_h. Відбиваючи точку x_h відносно точки x_0, одержимо точку x_r і знайдемо $f(x_r) = f_r$. Операція відбивання ілюструється рис. 2.26а. Якщо $\alpha > 0$ – коефіцієнт відбивання, то положення точки x_r визначається в такий спосіб: $x_r - x_0 = a(x_0 - x_h)$, тобто $x_r = (1+a)x_0 - ax_h$. ($a = | x_r - x_0 | / | x_0 - x_h |$.)



Операція відбивання



а) Операція розтягання



б)

Рис. 2.26. Операції відбивання та розтягування

Д. Порівняємо значення функцій f_r і f_l.

1. Якщо $f_r < f_l$, те ми одержали найменше значення функції. Напрям з точки x_0 в точку x_r найбільш зручний для переміщення. Таким чином, ми робимо розтяг у цьому

напрямку і знаходимо точку x_e і значення функції $f_e = f(x_e)$. Рисунок вище ілюструє операцію розтягування симплексу. Коефіцієнт розтягування $\gamma > 1$ визначається з наступних співвідношень:

$x_e - x_0 = \gamma(x_r - x_0)$, тобто $x_e = \gamma x_r + (1-\gamma)x_0$, де $\gamma = |x_e - x_0| / |x_r - x_0|$.

а) Якщо $f_e < f_l$, то замінюємо точку x_h на точку x_e і перевіряємо $(n+1)$-шу точку симплекса на збіжність до мінімуму (див. крок 3). Якщо збіжність досягнута, то процес зупиняється; у протилежному випадку повертаємося на крок Б.

б) Якщо $f_e \geq f_l$, то відкидаємо точку x_e. Очевидно, ми перемістилися занадто далеко від точки x_0 до точки x_r. Тому варто замінити точку x_h на точку x_r, у якій було отримане покращення (крок Д, 1), перевірити збіжність і, якщо вона не досягнута, повернутися на крок В.

2. Якщо $f_r > f_l$, але $f_r \leq f_g$, то x_r є кращою точкою в порівнянні з іншими двома точками симплекса, заміняємо точку x_h, на точку x_h і, якщо збіжність не досягнута, повертаємося на крок Б, тобто виконуємо пункт 1,б, описаний вище.

3. Якщо $f_r > f_l$ і $f_r > f_g$, то перейдемо до кроку Е.

Е. Порівняємо значення функцій f_r і f_h.

1. Якщо $f_r > f_h$, то переходимо безпосередньо до кроку стиску Е. 2.

Якщо $f_r > f_h$, то заміняємо точку x_h на точку x_r і значення функції f_h на значення функції f_r. Запам`ятовуємо значення $f_r > f_g$ із кроку Д,2, наведеного вище. Потім переходимо до кроку Е, 2.

2. У цьому випадку $f_r > f_h$, тому зрозуміло, що ми перемістилися занадто далеко від точки x_h до точки x_0. Спробуємо виправити це, знайшовши точку x_c (а потім f_c) за допомогою кроку стискання, зображеного на рис. 2.26б.

Якщо $f_r > f_h$, то відразу переходимо до кроку стискання і знаходимо точку x_c із співвідношення $x_c - x_0 = \beta(x_h - x_0)$, де β $(0 < \beta < 1)$ – коефіцієнт стискання. Тоді $x_c = \beta x_h + (1-\beta)x_0$.

Якщо $f_r < f_h$, то спочатку замінимо точку x_h на точку x_r, а потім виконаємо стискання. Тоді точку x_c знайдемо зі співвідношення $x_c - x_0 = \beta(x_r - x_0)$, тобто $x_c = \beta x_r + (1-\beta)x_0$

Ж. Порівняємо значення функцій f_c і f_h.

1. Якщо $f_c < f_h$, то заміняємо точку x_h, на точку x_c, і якщо збіжність не досягнута, то повертаємося на крок Б.

2. Якщо $f_c > f_h$, то очевидно, що всі наші спроби знайти значення менше f_h, закінчилися невдачею, тому ми переходимо на крок 3.

3. На цьому кроці ми зменшуємо розмірність симплекса діленням навпіл відстані від кожної точки симплекса до x_l – точки, що визначає найменше значення функції.

Таким чином, точка x_i заміняється на точку $x_i + \frac{1}{2}(x_i - x_l)$, тобто змінюємо точку x_i точкою $\frac{1}{2}(x_i + x_l)$.

Потім обчислюємо f_i для $i = 1,2,....,(n+1)$, перевіряємо збіжність і, якщо вона не досягнута, повертаємося на крок В.

І. Перевірка збіжності ґрунтується на тому, щоб стандартне відхилення $(n+1)$-го значення функції було менше деякого заданого малого значення ε. У цьому випадку обчислюється $o^2 = \sum_{i=1}^{n+1}(f_i - \bar{f})^2/(n+1)$, де $\bar{f} = \Sigma f_i/n+1$.

Якщо $\sigma < \varepsilon$, то всі значення функції дуже близькі одне до одного, і тому вони, можливо, лежать поблизу точки мінімуму функції x_l. Виходячи з цього, такий критерій збіжності є розумним, хоча Бокс, Дэвіс і Свенн пропонують те, що вони вважають більш "безпечною" перевіркою.

Кроки цієї процедури подані у вигляді блок-схеми на рис. нижче.

Коефіцієнти α, β, γ у вищенаведеній процедурі є відповідно коефіцієнтами відображення, стискання і розтягування. Нелдер і Мід рекомендують брати $\alpha = 1, \beta = 0{,}5, \gamma = 2$. Рекомендація ґрунтується на результатах експериментів із різноманітними комбінаціями значень. Ці значення параметрів дозволяють методу бути ефективним і працювати в різних складних ситуаціях.

Початковий симплекс вибирається на наш розсуд. У цій програмі точка x_1 є початковою точкою, далі у програмі формуються точки $x_2 = x_1 + ke_1$, $x_3 = x_1 + ke_2$, $k_{n+1} = x_1 + ke_n$, де k – довільна довжина кроку, а e_j – одиничний вектор.

Блок-схема алгоритму подана на рис. 2.27.

В цій роботі розглянуті методи мінімізації функції декількох змінних *f*, які не використовують розрахунок похідних. В основі описаних методів лежить наступне. При заданому векторі *x* визначається допустимий напрям *d*. Потім, виходячи із точки *x*, функція *f* мінімізується паралельно напрямку *d* одним із методів, які розглянуті далі.

Припустимо, що задача лінійного пошуку складається з мінімізації $f(x + \lambda d)$ при умові, що $\lambda \in L$, де L звичайно задається у формі $L = E_1$, $L = \{\lambda : \lambda \geq 0\}$ або $L = \{\lambda : a \leq \lambda \leq b\}$. У формулюванні алгоритмів для спрощення припустимо, що точка мінімуму $\bar{\lambda}$ існує. Але в реальних задачах це припущення може не виконуватися. Оптимальне значення цільової функції в задачі лінійного пошуку може бути необмеженим або оптимальне значення функції кінцеве, але не досягається ні при якому λ. В першому випадку цільова функція початкової задачі необмежена і розрахунок зупиняється. В другому випадку можна обрати таке $\bar{\lambda}$, що $f(x + \bar{\lambda} d)$ буде достатньо наближеним до $\inf\{f(x+\lambda d) : \lambda \in L\}$.

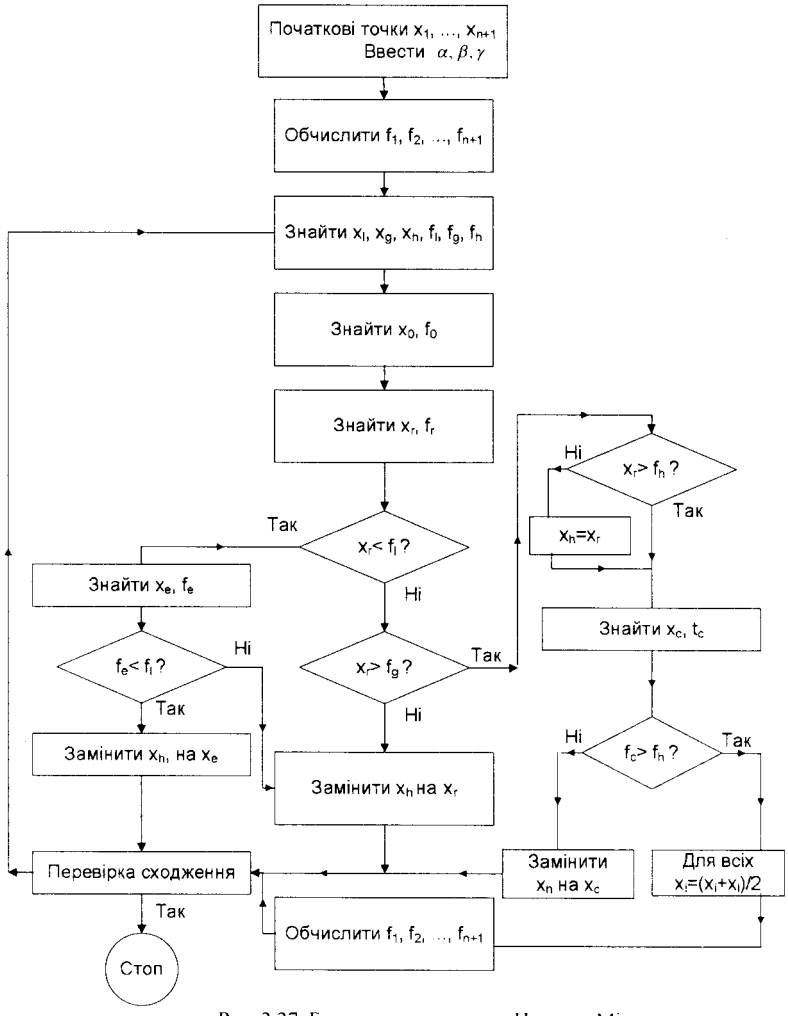


Рис. 2.27. Блок-схема алгоритму Нелдера-Міда

2.7.4. Метод циклічного покоординатного спуску

В цьому методі як напрямок пошуку використовуються координатні вектори. Точніше, метод здійснює пошук паралельно напрямків $d_1, \ldots, d_n$, де d_j – вектор, всі компоненти якого, за винятком j-го, мають значення 0. Таким чином, при пошуку за напрямком d_j змінюється лише змінна x_j, в той час як всі інші змінні залишаються зафіксованими. Схематично цей метод представлений на рис. 2.28.

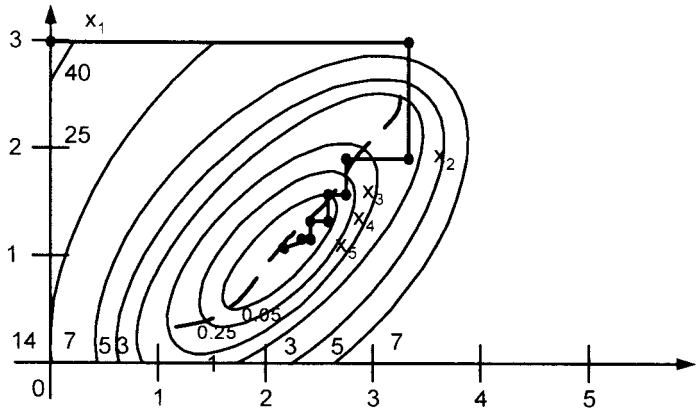


Рис. 2.28. Ілюстрація до методу циклічного покоординатного спуску

Алгоритм циклічного покоординатного спуску. Нижче приводиться алгоритм методу циклічного покоординатного спуску для мінімізації функції декількох змінних, який не вимагає використання похідних. Коротко показано, що якщо функція диференціюється, то метод збігається до стандартної точки. Для зупинки алгоритму можуть бути використані декілька критеріїв. В наведеному нижче алгоритмі процес зупиняється, якщо $\|x_{k+1} - x_k\| < \varepsilon$. Очевидно, що можна застосувати і будь-який інший критерій зупинки.

Початковий етап. Вибрати число $\varepsilon > 0$, яке буде використовуватися для зупинки алгоритму, і обрати в якості $d_1,...,d_n$ координатні напрямки. Обрати початкову точку x_1, покласти $y_1 = x_1$, $k=j=1$ та перейти до основного етапу.

Основний етап.

Крок 1.

Покласти λ_j рівне оптимальному розв'язку задачі мінімізації $f(y_j + \lambda d_j)$ при умові $\lambda \in E_1$. Покласти $y_{j+1} = y_j + \lambda_j d_j$. Якщо $j < n$, то замінити j на $j+1$ та повернутися до кроку 1. Якщо $j=n$, то перейти до кроку 2.

Крок 2.

Покласти $x_{k+1}=y_{n+1}$. Якщо $\|x_{k+1} - x_k\| < \varepsilon$, то зупинитися. В іншому випадку покласти $y_1 = x_{k+1}$, $j=1$, замінити k на $k+1$ та перейти до кроку 1.

Приклад 2.13. Розглянемо наступну задачу: мінімізувати $(x_1 - 2)^4 + (x_1 - 2x_2)^2$.

Зауважимо, що оптимальним розв'язком цієї задачі є точка (2,1), в якій значення функції рівне 0. В табл.2.7 приведені результати розрахунків за методом циклічного покоординатного спуску для початкової точки (0,3). Зауважимо, що на кожній ітерації вектори y_2 та y_3 отримані за допомогою одномірної мінімізації за напрямками (1,0) та (0,1) відповідно. Зауважимо також, що помітний спад функції отриманий під час перших декількох ітерацій, тоді як на останніх ітераціях процес очевидно

сповільнюється. Після семи ітерацій отримана точка (2.22,1.11), значення функції в якій рівне 0.0023 (табл. 2.7).

Таблиця 2.7

Результати розрахунку за методом циклічного по координатного спуску

k	x_k $f(x_k)$	j	d_j	y_j	λ_j	y_{j+1}
1	(0.00,3.00)	1	(1.0,0.0)	(0.00,3.00)	3.13	(3.13,3.00)
	52.00	2	(0.0,1.0)	(3.13,3.00)	-1.44	(3.13,1.56)
2	(3.13,1.56)	1	(1.0,0.0)	(3.13,1.56)	-0.50	(2.63,1.56)
	1.63	2	(0.0,1.0)	(2.63,1.56)	-0.25	(2.63,1.31)
3	(2.63,1.31)	1	(1.0,0.0)	(2.63,1.31)	-0.19	(2.44,1.31)
	0.16	2	(0.0,1.0)	(2.44,1.31)	-0.09	(2.44,1.22)
4	(2.44,1.22)	1	(1.0,0.0)	(2.44,1.22)	-0.09	(2.35,1.22)
	0.04	2	(0.0,1.0)	(2.35,1.22)	-0.05	(2.35,1.17)
5	(2.35,1.17)	1	(1.0,0.0)	(2.35,1.17)	-0.06	(2.29,1.17)
	0.015	2	(0.0,1.0)	(2.29,1.17)	-0.03	(2.29,1.14)
6	(2.29,1.14)	1	(1.0,0.0)	(2.29,1.14)	-0.04	(2.25,1.14)
	0.007	2	(0.0,1.0)	(2.25,1.14)	-0.02	(2.25,1.12)
7	(2.25,1.12)	1	(1.0,0.0)	(2.25,1.12)	-0.03	(2.22,1.12)
	0.004	2	(0.0,1.0)	(2.22,1.12)	-0.01	(2.22,1.121)

На рис. 2.28 показані лише лінії рівня цільової функції та точки, що отримані методом циклічного покоординатного спуску. Уповільнення на останніх ітераціях пояснюється тим, що паралельно яру, показаного штриховою лінією, робляться дуже малі кроки за ортогональними напрямками.

Теорема 2.3. Мінімізація паралельно незалежних напрямків. Нехай задана функція, що диференціюється

$$f: E_n \to E_1. \tag{2.150}$$

Розглянемо задачу мінімізації $f(x)$ при умові, що $x \in E_n$ та алгоритм, відображення якого A визначено наступним чином. Вектор $y \in A(x)$, якщо він отримується послідовною мінімізацією функції f паралельно напрямків $d_1,\ldots,d_n$, починаючи з точки x. Напрямки пошуку $d_1,\ldots,d_n$ можуть залежати від x, а норма кожного із них рівна 1. Припустимо, що виконуються наступні умови:

1. Існує $\varepsilon > 0$, таке, що $\det\left[D(x)\right] \geq \varepsilon$ для кожного $x \in E_n$. Тут $D(x)$ – матриця порядку $n \times n$, стовпцями якої є побудовані алгоритмом напрямки. $\det\left[D(x)\right]$ – визначник $D(x)$.
2. Мінімум функції f, паралельний до будь-якої прямої в E_n – єдиний.

Збіжність методу циклічного покоординатного спуску до стаціонарної точки випливає безпосередньо із теореми 2.3 при наступних припущеннях:

1. Мінімум функції f, паралельний до будь-якої прямої в E_n – єдиний.
2. Послідовність точок, які генеруються алгоритмом, утримуються в компактній множині простору E_n.

Зауважимо, що напрямками пошуку, які використовуються на кожній ітерації, є координатні вектори, так що матриця напрямків $D=I$. Очевидно, що припущення (1) теореми 2.3 виконується. Як інший підхід – можна скористатися для доведення збіжності теоремою 2.4. Для цього треба показати, що повне алгоритмічне відображення замкнуте в кожній точці x, що задовольняє умові $\nabla f(x) \neq 0$. В цьому випадку, як функція спуску α береться сама f, а множина розв'язків $\Omega = \{x : \nabla f(x) = 0\}$

Теорема 2.4. Теорема збіжності. Нехай X – не порожня замкнена множина в E_n, $\Omega \subset X$ – не порожня множина розв'язків. Розглянемо точково-множинне відображення $A: X \to X$. При заданій початковій точці $x_1 \in X$ послідовність будується у відповідності з наступною ітераційною процедурою. Якщо $x_k \in \Omega$, то процес зупиняється. В іншому випадку допускається $x_{k+1} \in A(x_k)$, k присвоюється значення $k+1$ та процедура повторюється. Припустимо, що побудована алгоритмом послідовність $x_1, x_2, \ldots$ утримується в компактній підмножині множини X, і нехай існує неперервна функція α, котру назвемо *функцією спуску*, така, що $\alpha(y) < \alpha(x)$, якщо $x \notin \Omega$, і $y \in A(x)$. Якщо відображення A замкнуте на доповненні до Ω, то або алгоритм зупиниться через скінчену кількість кроків в точці із Ω, або він породить безупинну послідовність $\{x_k\}$, таку, що:

1. Будь-яка збіжна підпослідовність послідовності $\{x_k\}$ має межу в Ω, тобто всі граничні точки $\{x_k\}$ належать Ω.
2. $\alpha(x_k) \to \alpha(x)$ для деякого $x \in \Omega$.

Прискорений крок. Встановили, що метод циклічного покоординатного спуску при мінімізації диференціної функції збігається до точки з нульовим значенням градієнта. При відсутності диференційовності метод може зупинитися в неоптимальній точці.

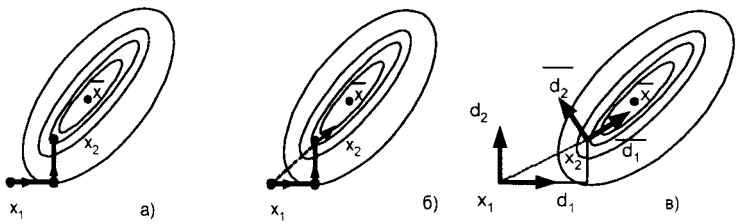


Рис. 2.29. Ефект ярності. а – зупинка в точці x_2; б – пошук продовжується паралельно до напрямку x_2-x_1, в – побудова нових напрямків пошуку в методі Розенброка

Як показано на рис. 2.29а, пошук паралельний до будь-якої координатної осі в точці x_2 не приводить до поліпшення цільової функції, і в результаті виконання алгоритму зупиняється. Це пояснюється наявністю яру, викликаного недиференційовністю f. Як показано на рис. 2.29б, ця складність може бути ліквідована паралельно до напрямку x_2-x_1. Пошук паралельно до напрямку x_{k+1}-x_k часто використовується в процедурах циклічного покоординатного спуску, навіть, якщо

функція не диференціюється. Звичайно, емпіричним шляхом встановлюється, що такий крок виконується на кожній p-й ітерації. Ця модифікація методу циклічного покоординатного спуску часто прискорює збіжність, власне, коли послідовність точок утворює зигзагоподібну траєкторію паралельно до дна яру. Такий крок звичайно називають прискореним кроком.

2.7.5. Метод Розенброка

В першопочатковому варіанті метода Розенброка не використовувалась одномірна мінімізація за напрямком, а застосовувались дискретні кроки паралельно до напрямків пошуку (рис. 2.29в). Тут подано неперервний варіант методу із використанням одномірної мінімізації. На кожній ітерації процедура здійснює ітеративний пошук паралельно n лінійно незалежних та ортогональних напрямків. Коли отримана нова точка в кінці ітерації, будується нова множина ортогональних векторів. На рис. 2.29в нові напрямки позначені через $\bar{d}_1$ та $\bar{d}_2$.

Побудова напрямків пошуку. Нехай $d_1,\ldots, d_n$ – лінійно-незалежні вектори, за нормою рівні 1. Припустимо, що ці вектори взаємно-ортогональні, тобто $d_i^T d_j = 0$ для $i \neq j$. Починаючи з біжучої точки x_k, цільова функція послідовно мінімізується паралельно до кожного із напрямків, в результаті отримуємо точку x_{k+1}. Власне $x_{k+1} - x_k = \sum_{j=1}^{n} \lambda_j d_j$, де λ_j – довжина кроку за напрямком d_j. Новий набір напрямків $\bar{d}_1,\ldots,\bar{d}_n$ будується за допомогою процедури Грама-Шмідта наступним чином:

$$a_j = \begin{cases} d_j, & якщо \quad \lambda_j = 0, \\ \sum_{i=j}^{n} \lambda_i d_i, & якщо \quad \lambda_j \neq 0, \end{cases}$$

$$b_j = \begin{cases} a_j, & при \quad j = 1, \\ a_j - \sum_{i=1}^{j-1} \left(a_j^T \bar{d}_i\right) \bar{d}_i, & при \quad j \geq 2, \end{cases}$$

$$\bar{d}_j = \frac{b_j}{\|b_j\|} \qquad (2.151)$$

Лема 2.1 показує, що нові напрямки лінійно-незалежні та ортогональні.

Лема 2.1. Припустимо, що вектори $d_1,\ldots, d_n$ лінійно незалежні та взаємно ортогональні. Тоді напрямки $\bar{d}_1,\ldots,\bar{d}_n$, які визначені відношенням (2.151), також лінійно незалежні та взаємно-ортогональні для будь-якої множини $\lambda_1,\ldots,\lambda_n$. Крім того, якщо $\lambda_j = 0$, то $\bar{d}_j = d_j$.

Доведення. Перед тим як доводити лему, покажемо, що $a_1,\ldots,\ a_n$ лінійно незалежні. Припустимо, що $\sum_{j=1}^{n}\mu_j a_j = 0$. Нехай $I=\{j:\lambda_j=0\}$ та $J(j)=\{i: i\notin I, i\geq j\}$. Із (2.151) отримаємо, що

$$0=\sum_{j=1}^{n}\mu_j a_j=\sum_{j\in I}\mu_j d_j+\sum_{j\notin I}\mu_j\left(\sum_{i=j}^{n}\lambda_i d_i\right)=\sum_{j\in I}\mu_j d_j+\sum_{j\notin I}\left(\lambda_j\sum_{i\in J(j)}^{n}\mu_i\right)d_j \qquad (2.152)$$

Так як $d_1,\ldots,d_n$ лінійно незалежні, маємо $\mu_j=0$ для $j\in I$ та $\lambda_j\sum_{i\in J(j)}\mu_i=0$ для $j\notin I$. Але $\lambda_j\neq 0$ для $j\notin I$ і, відповідно, $\sum_{i\in J(j)}\mu_i=0$. Таким чином, $\mu_1=\ldots=\mu_n=0$, і, значить, $a_1,\ldots,\ a_n$ лінійно-незалежні.

Доведемо за індукцією, що $b_1,\ldots,\ b_n$ лінійно незалежні. Так як $b_1=a_1\neq 0$, то достатньо показати, що якщо $b_1,\ldots,\ b_k$ лінійно незалежні, то $b_1,\ldots,b_k,b_{k+1}$ також лінійно незалежні. Припустимо, що $\sum_{j=1}^{k+1}\alpha_j b_j=0$. Із визначення b_{k+1} в (2.151) отримаємо, що

$$0=\sum_{j=1}^{k}\alpha_j b_j+\alpha_{k+1}b_{k+1}=\sum_{k=1}^{k}\left[\alpha_j-\frac{\alpha_{k+1}\left(a_{k+1}^T\overline{d}_j\right)}{\left\|b_j\right\|}\right]b_j+\alpha_{k+1}a_{k+1} \qquad (2.153)$$

Із (2.151) випливає, що вектор b_j є лінійною комбінацією векторів $a_1,\ldots,\ a_j$. Так як $a_1,\ldots,\ a_{k-1}$ лінійно незалежні, то з (2.153) випливає, що $\alpha_{k+1}=0$. Так як $b_1,\ldots,\ b_k$, за припущенням індукції, лінійно незалежні, то з (2.153) отримуємо, що $\alpha_j-\alpha_{k+1}a_{k+1}^T d_j/\left\|b_j\right\|=0$, $j=1,\ \ldots,\ k$. Звідси випливає, що, якщо $\alpha_{k+1}=0$, то $a_j=0$ для всіх j. Це означає, що вектори лінійно незалежні. Із визначення $\overline{d}_j$ тут же випливає їх лінійна незалежність. Тепер покажемо ортогональність $b_1,\ldots,\ b_n$ і, відповідно, ортогональність $\overline{d}_1,\ldots,\overline{d}_n$. Із (2.151) очевидно, що $b_1^T b_2=0$ і, таким чином, достатньо показати, що якщо взаємно ортогональні $b_1,\ldots,\ b_k$, то і $b_1,\ldots,\ b_k,\ b_{k+1}$ також взаємно ортогональні. Враховуючи, що $b_j^T\overline{d}_i=0$ для $i\neq j$, із (2.153) випливає, що

$$b_j^T b_{k+1}=b_j^T\left[a_{k+1}-\sum_{i=1}^{k}\left(a_{k+1}^T\overline{d}_i\right)\overline{d}_i\right]=b_j^T a_{k+1}-\left(a_{k+1}^T\overline{d}_j\right)b_j^T\overline{d}_j=0 \qquad (2.154)$$

Таким чином, $b_1,\ldots,b_k,b_{k+1}$ взаємно ортогональні. Завершуючи доведення, покажемо, що, якщо $\lambda_j=0$, то з (2.151) отримуємо, що

$$b_j = d_j - \sum_{i=1}^{j-1} \frac{1}{\|b_i\|}\left(d_j^T b_i\right)\overline{d}_i \quad . \tag{2.155}$$

Зауважимо, що b_i є лінійною комбінацією векторів $a_1,\ldots,a_i$, так як $b_i = \sum_{r=1}^{i} \beta_{ir} a_r$. Таким чином, із (2.151) випливає, що

$$b_i = \sum_{r\in R} \beta_{ir} d_r + \sum_{r\in \overline{R}} \beta_{ir}\left(\sum_{s=r}^{n} \lambda_s d_s\right), \tag{2.156}$$

де $R = \{r : r \le i, \lambda_r = 0\}$, $\overline{R} = \{r : r \le i, \lambda_r \ne 0\}$. Розглянемо $i < j$ і зауважимо, що $d_j^T d_v = 0$ для $v \ne i$. Для $r \in R$ очевидно, що $r \le i < j$, і, відповідно, $d_j^T d_r = 0$. Для $r \notin R$ маємо $d_j^T\left(\sum_{s=r}^{n} \lambda_s d_s\right) = \lambda_j d_j^T d_j = \lambda_j$. За припущенням λ_j=0. Помноживши (2.156) на d_j^T, отримуємо, що $d_j^T b_i = 0$ для $i < j$. З (2.155) випливає, що $b_j = d_j$ і, отже, $\overline{d}_j = d_j$.

Із цієї леми випливає, що новий напрямок $\overline{d}_j$ співпадає зі старим d_j, якщо λ_j=0. Отже, необхідно обчислювати нові напрямки тільки для тих індексів, для яких $\lambda_j \ne 0$.

Алгоритм Розенброка з дискретним кроком. Запропонований Розенброком метод не використовує одновимірну мінімізацію. Замість цього за ортогональними напрямками виконуються дискретні кроки, довжина яких змінюється в залежності від значення функції в обчислювальній точці. Наведемо алгоритм цього варіанту методу.

Початковий етап. Вибрати число $\varepsilon > 0$ для зупинки алгоритму, коефіцієнт розтягування $\alpha > 1$ та коефіцієнт стиску $\beta \in (-1,0)$. Взяти в якості $d_1,\ldots,d_n$ координатні напрямки та вибрати $\overline{\Delta}_1,\ldots,\overline{\Delta}_n > 0$ початкову довжину кроку паралельно до кожного із напрямків. Вибрати початкову точку x_1, покласти $y_1 = x_1, k = j = 1, \Delta_j = \overline{\Delta}_j$ для всіх j та перейти до основного етапу.

Основний етап.

Крок 1. Якщо $f(y_j + \Delta_j d_j) < f(y_j)$, то крок по j-му напрямку вважається успішним. Покласти $y_{j+1} = y_j + \Delta_j d_j$ та замінити Δ_j на $\alpha\Delta_j$. Якщо ж $f(y_j + \Delta_j d_j) \ge f(y_j)$, то крок вважається неуспішним. Покласти $y_{j+1} = y_j$ та Δ_j замінити на $\beta\Delta_j$. Якщо $j < n$, то замінити j на j+1 та повторити крок 1. В іншому випадку, тобто при j=n, перейти до кроку 2.

Крок 2. Якщо $f(y_{n+1}) < f(y_1)$, тобто, якщо хоча би один спуск за напрямком на кроці 1 виявився успішним, покласти $y_1 = y_{n+1}, j = 1$ та повторити крок 1. Нехай $f(y_{n+1}) = f(y_1)$, тобто кожний із n останніх спусків за напрямком на кроці 1 був

неуспішним. Якщо $f(y_{n+1}) < f(x_k)$, тобто хоча б один успішний спуск зустрівся під час k-ої ітерації на кроці 1, то перейти до кроку 3. Якщо $f(y_{n+1}) = f(x_k)$, тобто не було жодного успішного спуску за напрямком, то зупинитися. При цьому x_k є приблизним оптимальним розв'язком, якщо $|\Delta_j| < \varepsilon$ для всіх j. В іншому випадку покласти $y_1 = y_{n+1}, j = 1$ та перейти до кроку 1.

Крок 3. Покласти $x_{k+1} = y_{n+1}$. Якщо $\|x_{k+1} - x_k\| < \varepsilon$, то зупинитися; x_{k+1} – приблизний оптимальний розв'язок. В іншому випадку обчислити $\lambda_1,...,\lambda_n$ із відношення $x_{k+1} - x_k = \sum_{j+1}^{n} \lambda_j d_j$, побудувати нові напрямки у відповідності з (1), позначити їх через $d_1,\ldots, d_n$, покласти $\Delta_j = \overline{\Delta}_j$ для всіх j, покласти $y_1 = x_{k+1}$, замінити k на $k+1$, покласти $j=1$ та перейти до кроку 1. Зауважимо, що дискретні кроки вибираються паралельно до n напрямків спуску на кроці 1. Якщо рух паралельно d_j є успішним, то Δ_j заміняється на $\alpha\Delta_j$. Якщо ж цей напрямок є неуспішним, то Δ_j заміняється на $\beta\Delta_j$. Так як $\beta < 0$, то неуспішний напрямок призводить до зсуву в зворотньому напрямку паралельно j-го вектора на наступній реалізації кроку 1. Зауважимо, що крок 1 повторюється до тих пір, поки неуспішний напрямок буде мати місце при спускові по кожному із напрямків пошуку. В цьому випадку будується нові напрямки пошуку у відповідності з процедурой Грама – Шмідта.

Програмна реалізація порівняння методів Хука - Дживса та Розенброка з поясненнями. Цей метод був розроблений у 1961 році, але й дотепер він є дуже ефективним та оригінальним. Пошук складається з послідовності досліджувальних кроків навколо базової точки, за яким у випадку успіху реалізується пошук за зразком. Дослідимо методи прямого пошуку для функцій n змінних. Розв'яжемо програмним чином за методом Хука-Дживса з дискретним кроком наступну задачу $(x_1 - 2)^2 + (x_1 - 2x_2)^2 \Rightarrow Min$. Значення координат початкової точки x_1=2, x_2=3, параметрів методу становлять $\alpha = 1.0$, $h = 0.2$. Виконати 4 ітерації методу та визначити координати точки оптимуму після цих ітерацій.

Алгоритм методу Хука – Дживса утворює наступну послідовність кроків:

1. Обрати початкову базову точку b_1 і крок довжиною h_j для кожної змінної $x_j, j = 1,2,...,n$. У деяких програмах для кожної змінної використовується крок h, проте зазначена вище модифікація теж може виявитися корисною.

```
#include <iostream.h>
#include <stdlib.h>
#include <stdio.h>
#include <conio.h>
#include <math.h>
double x1,x2,r,p1,p2;
double fun(double , double , double , double , int);
void main(void){double h1,h2;int i=0;x1=2.0;x2=3.0;
clrscr();
h1=0.1;h2=0.2;
```

2. Обчислити $f(x)$ у базовій точці b_1 з метою одержання даних про локальну поведінку функції $f(x)$. Ці дані будуть використовуватися для знаходження найвідповіднішого напрямку пошуку за зразком, за допомогою якого можна сподіватися досягнути більшого зменшення значення функції.

Функція $f(x)$ у базовій точці b_1 знаходиться наступним чином:

- Обчислюється значення функції $f(b_1)$ у базовій точці b_1:

```
f=pow((x1-2),2)+pow((x1-2*x2),2);
```

- Значення кожної змінної почергово змінюється додаванням довжини кроку:

```
f=func(x1,x2,f,h1,h2,i);
```

Обчислюємо значення функції $f(b_1 + h_1 e_1)$, де e_1 – одиничний вектор у напрямку осі x_1. Якщо це призводить до зменшення значення функції, то b_1 заміняється на $b_1 + h_1 e_1$. У протилежному випадку обчислюється значення функції $f(b_1 - h_1 e_1)$, і, якщо її значення зменшилось, то b_1 заміняємо на $b_1 - h_1 e_1$. Якщо, ні один із пророблених кроків не призводить до зменшення значення функції, то точка b_1 залишається незмінною і розглядаються зміни в напрямку осі x_2, тобто знаходиться значення функції $f(b_1 + h_2 e_2)$ тощо. Коли будуть розглянуті усі n змінних, отримуємо нову базову точку b_2.

```
double func(double xx1, double xx2, double ff, double h1, double h2, int i){
            double f1,f2,f3,f4;
            double b1,b2;
            b1=xx1;
            b2=xx2;
            f1=pow(((xx1+h1)-2),2)+pow(((xx1+h1)-2*(xx2+h2)),2);
            f2=pow(((xx1-h1)-2),2)+pow(((xx1-h1)-2*(xx2-h2)),2);
            f3=pow(((xx1+h2)-2),2)+pow(((xx1+h2)-2*(xx2+h1)),2);
            f4=pow(((xx1-h2)-2),2)+pow(((xx1-h2)-2*(xx2-h1)),2);
if(i<5){
      if(ff>f1){
            xx1=xx1+h1;
            xx2=xx2-h2;
            cout<<"1!  "<<f1<<endl;
            ff=func(xx1,xx2,f1,h1,h2,i);}
      else {
            if(ff>f2){
                        xx1=xx1-h1;
                        xx2=xx2-h2;
                        cout<<"2!  "<<f2<<endl;
                        ff=func(xx1,xx2,f2,h1,h2,i);}
            else {
                  if(ff>f3){
                              xx1=xx1-h1;
                              xx2=xx2-h2;
                              cout<<"3!  "<<f3<<endl;
                              ff=func(xx1,xx2,f3,h1,h2,i);}
                  else {
                        if(ff>f4){
                                    xx1=xx1-h1;
                                    xx2=xx2-h2;
                                    cout<<"4!  "<<f4<<endl;
                                    ff=func(xx1,xx2,f4,h1,h2,i);}};};};
```

- Якщо $b_2 = b_1$, тобто зменшення функції не було досягнуто, таке ж дослідження повторюється навколо тієї ж базової точки b_1, але зі зменшеною довжиною кроку. На практиці задовільним є зменшення кроку (кроків) у десять разів від початкової довжини.

```
if(xx1==b1 && xx2==b2){
                    i++;
                    cout<<" p1 ";
                    h1=h1/10;
                    h2=h2/10;
                    ff=func(xx1,xx2,ff,h1,h2,i);}
```

- Якщо $b_2 \neq b_1$ то виконується пошук за зразком.

```
else{               i++;
                    cout<<" p2 ";
                    p1=b1+2*(xx1-b1);
                    p2=b2+2*(xx2-b2);
                    ff=func(p1,p2,ff,h1,h2,i);}
```

3. При пошуку за зразком використовується інформація, що отримана в процесі дослідження, і мінімізація функції завершується пошуком у напрямку, який заданий зразком. Ця процедура здійснюється в такий спосіб:

а). Доцільно рухатися з базової точки b_2 в напрямку b_2-b_1 , оскільки пошук у цьому напрямку вже привів до зменшення значення функції. Обчислимо функцію в точці зразка $P_1 = b_1 + 2(b_2 - b_1)$. У загальному випадку $P_i = b_i + 2(b_{i+1} - b_i)$.

б). Далі дослідження слід продовжувати навколо точки $P_1(P_i)$.

в). Якщо найменше значення на кроці 3б менше значення в базовій точці b_2 (у загальному випадку b_{i+1}), то одержуємо нову базову точку b_3 (b_{i+2}). після чого повторюємо крок 3а. У протилежному випадку пошук за зразком із точки b_2 (b_{i+1}) не здійснюємо, а продовжуємо дослідження в точці b_2 (b_{i+1})

4. Завершуємо процес, коли довжина кроку (довжини кроків) буде зменшена до заданого малого значення.

Тепер проаналізуємо отримані результати відповідно до різної точності та намалюємо діаграму – залежність точності результату від кількості ітерацій для методу Хука-Дживса та методу Розенброка (ряд1 – метод Хука-Дживса, ряд2 – метод Розенброка):

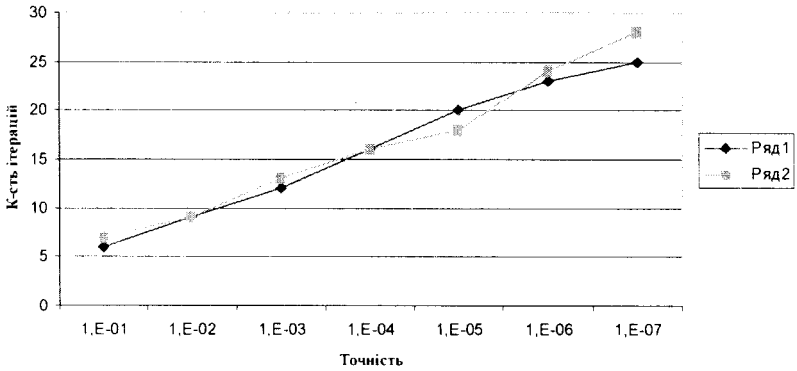


Рис. 2.30. Графічна інтерпретація отриманих результатів

Як бачимо, для нашої задачі $(x_1-2)^2+(x_1-2x_2)^2 \Rightarrow Min$ (значення координат початкової точки x_1=2, x_2=3 параметрів методу становлять $\alpha=1.0$, $h=0.2$) незалежно від методу розв'язування, графіки майже співпадають.

2.7.6. Метод спряжних градієнтів Флетчера та Рівса

За своєю змістовною постановкою множина задач дослідження операцій може бути розбита на ряд класів, основними з яких є:

- задачі керування запасами ;
- задачі розподілу ресурсів ;
- задачі ремонту та заміни обладнання (технічного обслуговування);
- задачі масового обслуговування;
- задачі теорії розкладу (календарного планування);
- задачі вибору маршруту (транспортні задачі);
- задачі мережевого планування та керування ;
- задачі планування та розміщення;
- комбіновані задачі.

Інженери та фахівці з дослідження операцій часто стикаються з необхідністю розв'язування оптимізацій них задач. На практиці зустрічаються різноманітні в змістовному сенсі задачі оптимізації. Це можуть бути задачі проектування, задачі розподілу обмежених ресурсів, задачі розрахунку траєкторії польоту ракети тощо. В недалекому минулому в задачах такого типу прийнятими вважались розв'язки із досить широкої області. В інженерному проектуванні, наприклад, звичайно задавався великий коефіцієнт надійності. Але з розвитком виробництва все більше зростають вимоги до проектів. В інших галузях діяльності, таких, як проектуваня космічних апаратів, умови функціонування проектованих систем пропонують екстремальні вимоги до характеристик проекту. Таким чином, існує необхідність в розв'язуванні наступних важливих питань. Яке найбільш ефективне використання наявних ресурсів? Чи можна отримати більш економний в тому чи іншому сенсі проект? В яких межах можна вважати ризик допустимим? Важливість та актуальність цих проблем викликали в останній час інтенсивні розробки моделей та методів оптимізації.

З іншого боку, розвиток моделей та методів оптимізації стимулювався значним збільшенням розмірності та складності оптимізаційних задач, що викликані суттєвим технологічним прогресом після другої світової війни. Інженери та керівники виробництва були змушені враховувати всі існуючі фактори та їх взаємозв'язки, які впливають на якість прийнятих рішень. Деякі з цих зв'язків не завжди могли бути зрозумілі та потребували, таким чином, додаткового аналізу та перевірених гіпотез. Успіх в методах вимірів і статистичних методах перевірки гіпотез суттєво допомогли в налагодженні взаємодії між різними компонентами досліджуваних систем.

В наш час методи дослідження операцій досить широко застосовуються інженерами, економістами, спеціалістами, які відповідають за прийняття рішень у військовій, культурній, адміністративній та інших галузях. Це пояснюється вже накопиченим позитивним досвідом практичного використання моделей та методів оптимізації. На ранньому етапі застосування дослідження операцій в плануванні та управлінні виробництвом, як правило, використовувались методи лінійного

програмування та статистичного аналізу. Тепер добре відомі ефективні методи та машинні програми для розв'язування такого роду задач – це нелінійне програмування.

Ідея градієнтних методів базується на тому, що напрямок градієнта є напрямком найшвидшого зростання функції. Отже, протилежний напрямок є напрямком найшвидшого спадання функції. Напрямок градієнта перпендикулярний в будь-якій точці лінії постійного рівня, оскільки уздовж цієї лінії значення функції є сталим. В методі найшвидшого спуску пошук мінімуму функції здійснюється вздовж напрямку, оберненого до напрямку градієнту. Як критерій завершення обчислень, найчастіше використовується близькість до нуля градієнту. Основною складністю застосування методу найшвидшого спуску є залежність його від вибору масштабу змінних, що оптимізуються. В методах, що використовують другі похідні, використовується матриця Гессе, яку необхідно інвертувати на кожному кроці, як в методі Ньютона-Рафсона. Оскільки ця процедура може займати основний час обчислень, використовуються наближення до неї, і позитивно визначена симетрична матриця в кінцевому результаті при виконанні певних умов стає рівною асиптотично матриці Гессе, як у методі Давідсона-Флетчера-Павела. Основна ідея методів штрафів для розв'язування задачі оптимізації з обмеженнями полягає у перетворені задачі мінімізації функції з обмеженнями у задачу пошуку мінімуму без обмежень модифікованої функції, в якій при порушенні обмежень додаткова складова, яка сформована на основі обмежень, "штрафувала" б модифіковану функцію, збільшуючи її значення. На цій ідеї побудовані метод Фіако-Маккорміка, який як крок алгоритму використовує метод Девідсона-Флетчера-Павела, та метод Бокса, який, по суті, є модифікацією симплексного методу Нелдера-Міда, що дозволяє враховувати обмеження. Для задач оптимізації функцій однієї змінної без обмежень умови екстремуму задаються у вигляді рівності нулеві похідної від функції, а характер екстремуму (мінімум, максимум, точка перегину) визначається знаком другої похідної. У випадку функцій n змінних у результаті прирівнювання часткових похідних до нуля виникає система нелінійних рівнянь, а характер екстремуму визначається негативною чи позитивною визначеністю матриці Гессе в точці екстремуму. За будь-яких умов аналітичний розв'язок цих задач можливий лише в найпростіших випадках.

У випадку пошуку екстремумів функції n змінних з обмеженнями у вигляді рівнянь первинна функція заміняється функцією Лагранжа, і система рівнянь отримується шляхом знаходження часткових похідних від змінних та множників Лагранжа та прирівнюванням їх до нуля.

Умови Куна-Такера є розширенням ідеї методу множників Лагранжа на задачі з обмеженнями загального вигляду – як у вигляді рівнянь, так і у вигляді нерівностей. Ці умови є необхідними умовами, які повинні виконуватись в стаціонарній точці. Розв'язок отриманої системи рівнянь та обмежень можливий лише для найпростіших задач, і в загальному для розв'язування використовуються чисельні методи.

Методи прямого пошуку є методами, у яких використовуються лише значення функції, і які орієнтовані на пошук екстремумів унімодальних функцій. Ідеї та складові методів прямого пошуку використовуються й при побудові складніших методів пошуку оптимальних рішень. Найефективнішим є метод, що використовує числа Фібоначі, однак з точки зору зручності реалізації доцільно використовувати метод золотого перетину, який незначно поступається методу Фібоначі. Окрім цих методів, використовують також апроксимаційні, що використовують ідею квадратичної

інтерполяції. Методи прямого пошуку застосовуються також і для пошуку екстремумів функцій n змінних. Найпростішим з них є метод покоординатного спуску (у випадку пошуку мінімуму), який почергово по кожній зі змінних реалізує одновимірний прямий пошук, і в якості процедури одновимірного пошуку найчастіше використовується метод золотого перетину. Однак з практичної точки зору цей метод є доволі повільним, і тому були розроблені ефективніші методи, одним з найвідоміших серед яких є метод Хука-Дживса. Цей метод використовує інформацію про локальну поведінку функції для обрання найвідповідальнішого напрямку пошуку за зразком.

Метод Нелдера-Міда (або пошук по деформованому многограннику) є розвитком симплексного методу. Ідея методу полягає в порівнянні значень функції в (n+1) вершинах симплекса та переміщенні симплекса в напрямку оптимальної точки за допомогою ітераційної процедури. Нелдер і Мід запропонували декілька модифікацій цього методу з неправильними симплексами. У результаті утворився дуже надійний метод прямого пошуку, що є одним із найефективніших, якщо n<6. Метод спряжних градієнтів, який запропонували Флетчер та Рівс, відхиляє напрямок найшвидшого спуску шляхом додавання до нього з позитивним коефіцієнтом напрямку, що використовується на попередньому етапі. Для квадратичного випадку, як це буде показано пізніше, таке відхилення напрямку найшвидшого спуску породжує множину спряжних напрямків.

Алгоритм метода спряжних градієнтів, який запропонували Флетчер та Рівс. Розглянемо алгоритм метода спряжних градієнтів для мінімізації функції, що диференціюється.

Початковий етап. Обрати число $\varepsilon > 0$ для зупинки алгоритму та початкову точку x_1. Покласти $y_1 = x_1$, $d_1 = -\nabla f(x_1)$, $k = j = 1$ та перейти до основного етапу.

Основной етап.

Крок 1. Якщо $\left\|\nabla f(y_j)\right\| < \varepsilon$, то зупинитися. В іншому випадку взяти в якості λ_j оптимальний розв'язок задачі мінімізації $f(y_j + \lambda d_j)$ при $\lambda \geq 0$ та покласти $y_{j+1} = y_j + \lambda_j d_j$. Якщо $j < n$, то перейти до кроку 2: в іншому випадку перейти до кроку 3.

Крок 2. Покласти $d_{j+1} = -\nabla f(y_{j+1}) + \alpha_j d_j$, де $\alpha_j = \dfrac{\left\|\nabla f(y_{j+1})\right\|^2}{\left\|\nabla f(y_j)\right\|^2}$. Замінити j на $j+1$ та перейти до кроку 1.

Крок 3. Покласти $y_1 = x_{k+1} = y_{n+1}$, $d_1 = -\nabla f(y_1)$, $j = 1$, замінтити k на k+1та перейти до кроку 1. Зауважимо, що $d_{j+1} = \dfrac{1}{\mu_1}\left[\mu_2 d_j - \mu_1 \nabla f(y_{j+1})\right]$, де $\mu_1 = \dfrac{\left\|\nabla f(y_j)\right\|^2}{\left\|\nabla f(y_j)\right\|^2 + \left\|\nabla f(y_{j+1})\right\|^2}$, $\mu_2 = \dfrac{\left\|\nabla f(y_{j+1})\right\|^2}{\left\|\nabla f(y_j)\right\|^2 + \left\|\nabla f(y_{j+1})\right\|^2}$, так що d_{j+1}, по суті, є опуклою комбінацією біжучого напрямку найшвидшого спуску та напрямку, що використовується на попередній ітерації.

Пошук екстремальних значень критерію за допомогою методів спряжних напрямків. Спочатку метод був запропонований Девідоном (Davidon [1959]), а потім розвитий Флетчером і Пауеллом (Fletcher, Powell [1963]). Метод Девідона-Флетчера-Пауелла називають також і методом змінної метрики. Він попадає в загальний клас квазиньютонівських процедур, у яких напрямки пошуку задаються у вигляді $-D_j \nabla f(y)$. Напрямок градієнта є, таким чином, відхиленим у результаті множення на $-D_j$, де D_j – позитивно визначена симетрична матриця порядку $n \times n$, що апроксимує обернену матрицю Гессе. На наступному кроці матриця D_j+1 представляється у вигляді суми D_j і двох симетричних матриць рангу один кожна. У зв'язку з цим схема іноді називається схемою корекції рангу два.

Алгоритм Девідона-Флетчера-Пауелла. Розглянемо алгоритм Девідона-Флетчера-Пауелла мінімізації диференційовної функції декількох змінних. Зокрема, якщо функція квадратична, то, як буде показано пізніше, метод генерує з'єднані напрямки і зупиняється після виконання однієї ітерації, тобто після пошуку вздовж кожного зі з'єднаних напрямків.

Початковий етап. Нехай $\varepsilon > 0$ – константа для зупинки. Вибрати точку x_1 і початкову симетричну позитивно визначену матрицю D_1. Покласти $y_1 = x_1$, $k = j = 1$ і перейти до основного етапу.

Основний етап.

Крок 1. Якщо $\| \nabla f(y_j) \| < \varepsilon$, то зупинитися; у протилежному випадку покласти $d_j = -D_j \nabla f(y_j)$ і взяти в якості λ_j оптимальний розв'язок задачі мінімізації $f(y_j + \lambda d_j)$ при $\lambda \geq 0$. Покласти $y_j+1 = y_j + \lambda_j d_j$. Якщо $j < n$, то перейти до кроку 2. Якщо $j = n$, то покласти $y_1 = x_k+1 = y_n+1$, замінити k на $k+1$, покласти $j=1$ і повторити крок 1.

Крок 2. Побудувати D_j+1 у такий спосіб :

$$D_{j+1} = D_j + \frac{p_j p_j^T}{p_j^T p_j} - \frac{D_j q_j q_j^T D_j}{q_j^T D_j q_j}, \tag{2.157}$$

$$pj = \lambda_j d_j, \tag{2.158}$$

$$q_j = \nabla f(y_j+1) - \nabla f(y_j). \tag{2.159}$$

Замінити j на $j + 1$ і перейти до кроку 1.

Приклад 2.14. Розглянемо наступну задачу: мінімізувати $(x_1 - 2)^4 + (x_1 - 2x_2)^2$.

Розв'язок. Результати обчислень методом Девідона-Флетчера-Пауелла приведені в табл. 2.8.

Таблиця 2.8

Результати обчислень за методом Девідона-Флетчера-Пауелла

k	x_k $f(x_k)$	j	y_j $f(y_j)$	$\nabla f(y_j)$	$\| \nabla f(y_j) \|$	D	d_j	λ_j	y_{j+1}
1	(0.00, 3.00) (52.00)	1	(0.00, 3.00) (52.00)	(-44.00, 24.00)	50.12	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	(44.00, - 24.00)	0.062	(2.70, 1.51)
		2	(2.70, 1.51) (0.34)	(0.73, 1.28)	1.47	$\begin{bmatrix} 0.25 & 0.38 \\ 0.38 & 0.81 \end{bmatrix}$	(-0.67, - 1.31)	0.22	(2.55, 1.22)

Продовження таблиці 2.8

2	(2.55, 1.22) (0.1036)	1	(2.55, 1.22) (0.1036)	(0.89, -0.44)	0.99	$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$	(-0.89, 0.44)	0.11	(2.45, 1.27)
		2	(2.45, 1.27) (0.0490)	(0.18, 0.36)	0.40	$\begin{bmatrix}0.65 & 0.45\\ 0.45 & 0.46\end{bmatrix}$	(-0.28, -0.25)	0.64	(2.27, 1.11)
3	(2.27, 1.11) (0.008)	1	(2.27, 1.11) (0.008)	(0.18, -0.20)	0.27	$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$	(-0.18, 0.20)	0.10	(2.25, 1.13)
		2	(2.25, 1.13) (0.004)	(0.04, 0.04)	0.06	$\begin{bmatrix}0.80 & 0.38\\ 0.38 & 0.31\end{bmatrix}$	(-0.05, -0.03)	2.64	(2.12, 1.05)
4	(2.12, 1.05) (0.0005)	1	(2.12, 1.05) (0.0005)	(0.05, -0.08)	0.09	$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$	(-0.05, 0.08)	0.10	(2.115, 1.058)
		2	(2.115, 1.058) (0.0002)	(0.004, 0.004)	0.006				

На кожній ітерації вектор d_j для $j = 1, 2$ визначається у вигляді $-D_j \nabla f(y_j)$, де D_1 – одинична матриця, а D_2 обчислюється за формулами (2.157)–(2.159). При $k = 1$ маємо $p_1 = (2.7, -1.49)^T$, $q_1 = (44.73, -22.72)^T$. На другій ітерації $p_1 = (-0.1, 0.05)^T$, $q_1 = (-0.7, 0.8)^T$ і, нарешті, на третій ітерації $p_1 = (-0.02, 0.02)^T$, $q_1 = (-0.14, 0.24)T$. Точка y_{j+1} обчислюється оптимізацією уздовж напрямку d_j при початковій точці y_j для $j = 1, 2$. Процедура зупинена в точці $y_2 = (2.115, 1.058)^T$ на четвертій ітерації, тому що норма $||f(y_2)|| = 0.006$ досить мала. Траєкторія руху, отримана методом, показана на рис. 2.31.

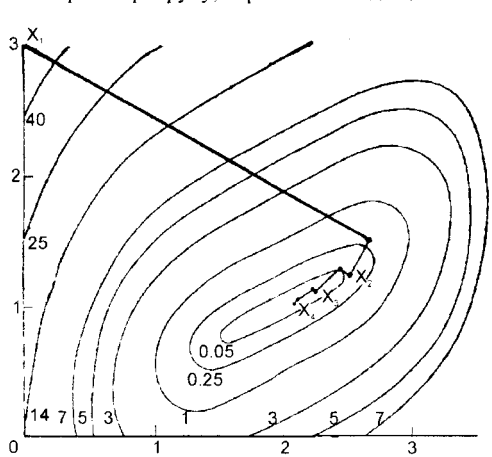


Рис. 2.31. Метод Девідона-Флетчера-Пауелла

Лема 2.2 показує, що кожна матриця D_j позитивно визначена і d_j є напрямком спуску. Для доказу леми знадобиться:

Теорема 2.5. Нехай S – непорожня множина в E_n, точка $x \in \mathbf{C}^1$. Конусом можливих напрямків у точці x називається множина $D = \{d : d \neq 0, x + \lambda d \in S$ при всіх $\lambda \in (0, \delta)$ для деякого $\delta > 0\}$.

Визначення 2.1. Нехай x і y – вектори з E_n та $|x^T y|$ – абсолютне значення скалярного добутку $x^T y$. Тоді виконується наступна нерівність, називана нерівністю Шварца : $| x^T y | \leq \|x\| \, \|y\|$.

Лема 2.2. Нехай $y_1 \in E_n$, а D_1 – початкова позитивно-визначена симетрична матриця. Для $j = 1, ..., n$ покладемо $y_{j+1} = y_j + \lambda_j d_j$, де $d_j = -D_j \nabla f(y_j)$, а $\lambda_j \in$ оптимальним розв'язком задачі мінімізації $f(y_j + \lambda d_j)$ при $\lambda \geq 0$. Нехай, крім того, для $j = 1, ..., n-1$ матриця D_{j+1} визначається за формулами (2.157)-(2.159). Якщо $\nabla f(y_j) \neq 0$ для $j = 1, ..., n$, то матриці $D_1, ..., D_n$ симетричні й позитивно-визначені, так що $d_1, ..., d_n$ – напрямки спуску.

Доведення. Проведемо доказ за індукцією. При $j = 1$ матриця D_1 симетрична і позитивно-визначена за умовою леми. Крім того, $\nabla f(y_1)^T d_1 = -\nabla f(y_1)^T D_1 \nabla f(y_1) < 0$, тому що D_1 позитивно-визначена. Тоді за теоремою 2.5 вектор d_1 визначає напрямок спуску. Припустимо, що твердження леми справедливе для деякого $j \leq n-1$. Покажемо, що воно справедливо для $j+1$. Нехай x – ненульовий вектор з E_n, тоді з (2.157) маємо

$$x^T D_{j+1} x = x^T D_j x + \frac{(x^T p_j)^2}{p_j^T q_j} - \frac{(x^T D_j q_j)^2}{q_j^T D_j q_j} \,. \qquad (2.160)$$

Так як D_j – симетрична позитивно-визначена матриця, то існує позитивно-визначена матриця $D_j 1/2$, така, що $D_j = D_j 1/2 \; D_j 1/2$. Нехай $a = D_j \, 1/2x$ і $b = D_j \, 1/2 q_j$. Тоді $x^T D_j x = a^T a$, $q_j^T D_j q_j = b^T b$ і $x^T D_j q_j = a^T b$. Підставляючи ці вирази в (2.160), одержуємо :

$$x^T D_{j+1} x = \frac{(a^T a)(b^T b) - (a^T b)^2}{b^T b} + \frac{(x^T p_j)^2}{p_j^T q_j} \,. \qquad (2.161)$$

За нерівністю Шварца маємо: $(a^T a)(b^T b) \geq (a^T b)^2$. Таким чином, щоб довести, що $x^T D_{j+1} x \geq 0$, досить показати, що $p_j^T q_j > 0$ та $b^T b > 0$. З (2.158) і (2.159) випливає, що

$$p_j^T q_j = \lambda_j d_j^T [\nabla f(y_{j+1}) - \nabla f(y_j)]. \qquad (2.162)$$

За припущенням $\nabla f(y_j) \neq 0$, і D_j позитивно-визначена, так що $\nabla f(y_j)^T D_j \nabla f(y_j) > 0$. Крім того, d_j – напрямок спуску, і, отже, $\lambda j > 0$. Тоді з (2.162) випливає, що $p_j^T q_j > 0$. Крім того, $q_j \neq 0$, і , отже, $b^T b = q_j^T D_j q_j > 0$.

Покажемо тепер, що $x^T D_{j+1} x > 0$. Припустимо, що $x^T D_{j+1} x = 0$. Це можливо тільки в тому випадку, якщо $(a^T a)(b^T b) = (a^T b)^2$ і $p_j^T x = 0$. Насамперед відзначимо, що $(a^T a)(b^T b) = (a^T b)^2$ тільки при $a = \lambda b$, тобто $D_j 1/2 x = \lambda D_j 1/2 q_j$. Таким чином, $x = \lambda q_j$. Тому що $x \neq 0$, те $\lambda \neq 0$. Далі, $0 = p_j^T x = \lambda \, p_j^T q_j$ суперечить тому, що $p_j^T q_j > 0$ і $\lambda \neq 0$. Отже, $x^T D_{j+1} x > 0$, тобто матриця D_j+1 позитивно-визначена. Оскільки $\nabla f(y_{j+1}) \neq 0$ і D_{j+1} позитивно-визначена, маємо $\nabla f(y_{j+1})^T d_{j+1} = -\nabla f(y_{j+1})^T \; D_{j+1} \nabla f(y_{j+1}) < 0$. Звідси за теоремою 2.5 випливає, що d_{j+1} – напрямок спуска. Лема доведена.

Квадратичний випадок. Надалі знадобиться:

Теорема 2.6. Нехай $f(x) = c^T x + \Delta x^T H x$, де H – симетрична матриця порядку n x n. Розглянемо H – з'єднані вектори d_1, …, d_n і довільну точку x_1. Нехай λ_k для $k = 1, \ldots, n$ – оптимальний розв'язок задачі мінімізації $f(x_k + \lambda d_k)$ при $\lambda \in E_1$ і $x_{k+1} = x_k + \lambda d_k$. Тоді для $k = 1, \ldots, n$ справедливі наступні твердження :

$$\nabla f(x_{k+1})^T d_j = 0, j = 1, \ldots, k; \tag{2.163}$$

$$\nabla f(x_1)^T d_k = \nabla f(x_k)^T d_k; \tag{2.164}$$

де x_{k+1} є оптимальним розв'язком задачі мінімізації $f(x)$ за умови $x - x_1 \in L(d_1, \ldots, d_k)$, де $L(d_1, \ldots, d_k)$ – лінійний підпростір, натягнутий на вектори $d_1, \ldots, d_k$, тобто

$$L(d_1 - d_k) = \left\{ \sum_{j=1}^{k} \mu_j d_j : \mu_j \in E_1 \ \forall \ j \right\}. \tag{2.165}$$

Зокрема, x_{n+1} – точка мінімуму функції f на E_n.

Якщо цільова функція f квадратична, то у відповідності зі сформульованою нижче теоремою 2.7 напрямки d_1, …, d_k, що генеруються методом Девідона-Флетчера-Пауелла, є сполученими. Отже, відповідно до твердження 3 теореми 2.6 метод зупиняється після завершення однієї ітерації в оптимальній точці. Крім того, матриця D_{n+1}, отримана наприкінці ітерації, збігається зі зворотньою до матриці Гессе H.

Теорема 2.7. Нехай H – симетрична позитивно визначена матриця порядку n x n. Розглянемо задачу мінімізації $f(x) = c^T x + \Delta x^T H x$ за умови $x \in E_n$. Припустимо, що задача вирішена методом Девідона-Флетчера-Пауелла при початковій точці y_1 і початковій позитивно визначеній матриці D_1. Зокрема, нехай λ_j, $j = 1, \ldots, n$ – оптимальний розв'язок задачі мінімізації $f(y_j + \lambda d_j)$ при $\lambda \geq 0$ і $y_{j+1} = y_j + \lambda_j d_j$, де $d_j = -D_j \nabla f(y_j)$, а D_j визначається за формулами (2.157)–(2.159). Якщо $\nabla f(y_j) \neq 0$ для всіх j, то напрямкb d_1, …, d_n є H-сполученими і $D_{n+1} = H^{-1}$. Крім того, y_{n+1} є оптимальним розв'язком задачі.

Доведення. Насамперед покажемо, що для j, такого, що $1 \leq j \leq n$, справедливі наступні твердження :

1. $d_1, \ldots, d_j$ лінійно-незалежні.
2. $d_i^T H d_k = 0$ для $i \neq k$; і, $k \leq j$.
3. $D_{j+1} H p_k$ або (що є еквівалентним) $D_{j+1} H d_k = d_k$ для $1 \leq k \leq j$, $p_k = \lambda_k d_k$.

Проведемо доведення за індукцією. Для $j = 1$ твердження 1 і 2 очевидні. Щоб довести твердження 3, відзначимо насамперед, що для будь-якого k справедливі рівності

$$H p_k = H(\lambda_k d_k) = H(y_{k+1} - y_k) = \nabla f(y_{k+1}) - \nabla f(y_k) = q_k. \tag{2.166}$$

Зокрема, $H p_1 = q_1$. Таким чином, припускаючи $j = 1$ у (2.157), одержуємо

$$D_2 H p_1 = \left[D_1 + \frac{p_1 p_1^T}{p_1^T q_1} - \frac{D_1 q_1 q_1^T D_1}{q_1^T D_1 q_1} \right] q_1 = p_1, \tag{2.167}$$

тобто твердження 3 справедливе при $j = 1$.

Тепер припустимо, що твердження 1, 2 і 3 справедливі для $j \leq n - 1$. Покажемо, що вони також справедливі й для $j + 1$. Нагадаємо, що за твердженням 1 теореми 2.6

$d_i \nabla f(y_{j+1}) = 0$ для $i \le j$. За індуктивним припущенням $d_i = D_{j+1} H d_i$, $i \le j$. Таким чином, для $i \le j$ маємо

$$0 = d_i \nabla f(y_{j+1}) = d_i^T H D_{j+1} \nabla f(y_{j+1}) = -d_i^T H d_{j+1}. \tag{2.168}$$

Через припущення щодо індукції ця рівність показує, що твердження 2 також справедливе для j+1. Тепер покажемо, що твердження 3 справедливе для j+1.

Припускаючи $k \le j$+1, маємо

$$D_{j+2} H p_k = \left[D_{j+1} + \frac{p_{j+1} p_{j+1}^T}{p_{j+1}^T q_{j+1}} - \frac{D_{j+1} q_{j+1} q_{j+1}^T D_{j+1}}{q_{j+1}^T D_{j+1} q_{j+1}} \right] H p_k . \tag{2.169}$$

З огляду на (2.166) і припускаючи $k = j + 1$ у (2.169), одержимо, що D_j+2$Hp_{j+1} = p_{j+1}$. Тепер нехай $k \le j$. Тому, що твердження 2 справедливе для $j + 1$, то

$$p_{j+1}^T H p_k = \lambda_k \lambda_{j+1} d_{j+1}^T H d_k = 0. \tag{2.170}$$

За припущенням індукції з (2.166) і внаслідок того, що твердження 2 справедливе для $j + 1$, одержуємо

$$q_{j+1}^T D_{j+1} H p_k = q_{j+1}^T p_k = p_{j+1}^T H p_k = \lambda_{j+1} \lambda_k d_{j+1}^T H d_k = 0 . \tag{2.171}$$

Підставляючи (2.170) і (2.171) у (2.169) та з огляду на припущення індукції, одержуємо

$$D_{j+2} H p_k = D_{j+1} H p_k = p_k . \tag{2.172}$$

Таким чином, твердження 3 справедливо для j+1.

Залишилося показати, що твердження 1 справедливе для j+1. Припустимо, що $\sum_{i=1}^{j+1} \alpha_i d_i = 0$. Помноживши цю рівність на $d_{j+1}^T H$ та з огляду на те, що твердження 2 справедливе для j+1, одержуємо, що $\alpha_{j+1} d_{j+1}^T H d_{j+1} = 0$. За умовою теореми $\nabla f(y_{j+1}) \ne 0$, а за лемою 2.2 матриця D_{j+1} позитивно визначена, так що $d_{j+1} = -D_{j+1} \nabla f(y_{j+1}) \ne 0$. Тому, що H позитивно визначена, то $d_{j+1}^T H d_{j+1} \ne 0$ і, отже, $\alpha_{j+1} = 0$. Звідси випливає, що $\sum_{i=1}^{j} \alpha_i d_i = 0$, і тому що $d_1, \ldots, d_j$ лінійно-незалежні за припущенням індукції, то $\alpha_i = 0$ для $i = 1, \ldots, j$... Таким чином, $d_1, \ldots, d_{j+1}$ лінійно незалежні й твердження 1 справедливе для j+1. Отже, твердження 1, 2 і 3 виконуються. Зокрема спряженість $d_1, \ldots, d_n$ випливає з тверджень 1 і 2, якщо покласти $j = n$.

Нехай тепер $j = n$ у твердженні 3. Тоді $D_{n+1} H d_k = d_k$ для $k = 1, \ldots, n$... Якщо за D взяти матрицю, стовпцями якої є вектори $d_1, \ldots, d_n$, то $D_{n+1} H D = D$. Тому що D має обернену, то $D_{n+1} H = I$, що можливе тільки в тому випадку, якщо $D_{n+1} = H^{-1}$. Нарешті, y_{n+1} є оптимальним розв'язком за теоремою 2.6. Теорема доведена.

Таблиця 2.9

Результати обчислень методом спряжених напрямків Девідона-Флетчера-Пауелла

k	r_k	$x_{k+1} = x_{r_k}$	$f(x_{k+1})$	$P(x_{r_k})$	$\varphi(r_k)$	$r_k P(x_{r_k})$
1	0.1	(1.4539,0.7608)	0.0935	1.8307	0.2766	0.1831
2	1.0	(1.1687,0.7407)	0.5753	0.3908	0.9661	0.3908
3	10.0	(0.9906,0.8425)	1.5203	0.01926	1.7129	0.1926
4	100.0	(0.9507,0.8875)	1.8917	0.000267	1.9184	0.0267
5	10^3	(0.9461,0.8934)	1.9405	0.0000028	1.9433	0.0028
6	10^5	(0.9231,09124)	1.9745	0.00000016	1.9723	0.00024

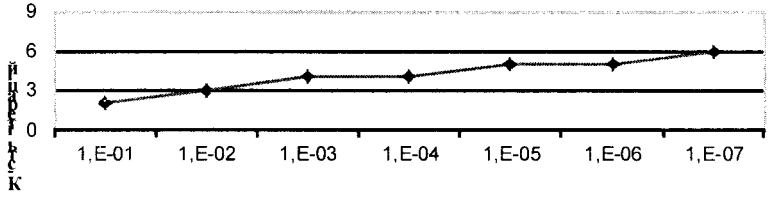


Рис. 2.32. Графічна інтерпретація отриманих результатів

Метод найшвидшого спуску грунтується на використанні значення функції та її градієнта, тобто інформації про напрямок пошуку. Напрямок градієнта перпендикулярний в будь-якій точці до лінії постійного рівня, оскільки вздовж цієї лінії значення функції є сталим. В методі найшвидшого спуску використовується ця властивість напрямку градієнта. Основною складністю застосування методу найшвидшого спуску є залежність його від вибору масштабу змінних, що оптимізуються. Метод Ньютона-Рафсона грунтується на використанні квадратичних функцій й інвертує матрицю Гессе на кожному кроці, що є працемісткою процедурою. З метою уникнення цієї складності у методі Девідона-Флетчера-Пауелла пошук здійснюється в напрямку позитивно визначеної симетричної матриці, що гранично наближається до матриці Гессе. Пошук здійснюється за взаємноспряженими напрямками з накопиченням попередньої інформації. За цих умов мінімум квадратичної функції n змінних буде знайдений не більш ніж за n кроків. Метод ДФП є досить потужною оптимізаційною процедурою, дуже ефективною при оптимізації більшості функцій незалежно від того, квадратичні вони чи ні.

Основна ідея методу розв'язування задач з обмеженнями полягає у перетворенні задачі мінімізації функції з обмеженнями у задачу пошуку мінімуму без обмежень функції зі "штрафним" додатком, що формується на основі обмежень і різко збільшує значення функції мети при наближенні до межі. Вплив штрафної функції полягає в створенні "гребеня з крутими межами" вздовж кожної границі області обмежень. Метод Фіако-Маккорміка у якості ітераційного кроку використовує метод Давідона-Флетчера-Пауелла. При програмній реалізації методу Фіако-Маккорміка потрібно звернути увагу на те, щоб в процесі одновимірного пошуку не вийти за межі області обмежень. Метод Бокса, по суті, є модифікацією симплексного методу Нелдера-Міда, що дозволяє враховувати обмеження. Ідеального методу пошуку оптимуму для багато екстремальних функцій навіть без обмежень, не кажучи про задачі з

обмеженнями, не існує. Тому з метою підвищення достовірності отриманих результатів процедуру слід повторити декілька разів з різними початковими значеннями, а також порівняти результати роботи декількох методів.

Метод спряжних градієнтів, який запропонували Флетчер та Рівс, відхиляє напрямок найскоршого спуску шляхом додавання до нього з позитивним коефіцієнтом напрямку, що використовується на попередньому етапі. Для квадратичного випадку, як це буде показано пізніше, таке відхилення напрямку найскоршого спуску породжує множину спряжних напрямків.

Для задач оптимізації функцій однієї змінної без обмежень умови екстремуму задаються у вигляді рівності похідної від функції нулеві, а характер екстремуму (мінімум, максимум, точка перегину) визначається знаком другої похідної. У випадку функцій n змінних у результаті прирівнювання часткових похідних до нуля виникає система нелінійних рівнянь, а характер екстремуму визначається негативною чи позитивною визначеністю матриці Гессе в точці екстремуму. За будь-яких умов аналітичний розв'язок цих задач можливий лише в найпростіших випадках.

У випадку пошуку екстремумів функції n змінних з обмеженнями у вигляді рівнянь первинна функція заміняється функцією Лагранжа, і система рівнянь отримується шляхом знаходження часткових похідних від змінних та множників Лагранжа і прирівнюванням їх до нуля.

Умови Куна-Такера є розширенням ідей методу множників Лагранжа на задачі з обмеженнями загального вигляду – як у вигляді рівнянь, так і у вигляді нерівностей. Ці умови є необхідними умовами, які повинні виконуватись в стаціонарній точці. Розв'язування отриманої системи рівнянь та обмежень можливе лише для найпростіших задач, і в загальному для розв'язування використовуються чисельні методи.

Методи прямого пошуку є методами, у яких використовуються лише значення функції й орієнтовані на пошук екстремумів унімодальних функцій. Ідеї та складові методів прямого пошуку використовуються і при побудові складніших методів пошуку оптимальних розв'язків. Найефективнішим є метод, що використовує числа Фібоначі. Однак, з точки зору зручності реалізації доцільно використовувати метод золотого січення, який незначно поступається методу Фібоначі. Окрім цих методів, використовують також апроксимаційні, що використовують ідею квадратичної інтерполяції.

Методи прямого пошуку застосовуються також і для пошуку екстремумів функцій n змінних. Найпростішим з них є метод покоординатного спуску (у випадку пошуку мінімуму), який почергово по кожній зі змінних реалізує одновимірний прямий пошук, і в якості процедури одновимірного пошуку найчастіше використовується метод золотого січення. Однак з практичної точки зору цей метод є доволі повільним, а тому були розроблені ефективніші методи, одним з найвідоміших серед яких є метод Хука-Дживса. Цей метод використовує інформацію про локальну поведінку функції для обрання найвідповідальнішого напрямку пошуку за зразком.

Метод Нелдера–Міда (або пошук по деформованому многограннику) є розвитком симплексного методу. Ідея методу полягає в порівнянні значень функції в $(n+1)$ вершинах симплекса і переміщенні симплекса в напрямку оптимальної точки за допомогою ітераційної процедури. Нелдер і Мід запропонували декілька модифікацій цього методу з неправильними симплексами. У результаті утворився дуже надійний метод прямого пошуку, що є одним із найефективніших, якщо $n<6$.

2.8. Контрольні питання

1. Сформулювати постановку одно- та багатокритеріальної задач ПР.
2. Поясніть сутність класичних критеріїв ПР?
3. Поясніть сутність похідніх критеріїв ПР?
4. Сформулюйте суть методу МАІ, назвіть переваги та недоліки.
5. Наведіть приклади одно- та багатокритеріальної задач ПР.
6. Що таке цільова функція і проектні параметри?
7. Що таке унімодальна функція?
8. Особливості алгоритмів оптимізації одновимірних функцій.
9. Суть алгоритму методу половинного ділення.
10. Як можна підвищити ефективність методу оптимізації?
11. Суть алгоритму методу золотого перерізу та Фібоначчі?
12. Класифікація методів безумовної оптимізації функцій багатьох змінних.
13. Особливості методів нульового порядку оптимізації багатомірних функцій.
14. Суть методів покоординатного спуску, Хука-Дживса та Нелдера-Міда.
15. Особливості алгоритмів оптимізації багатовимірних функцій.
16. Що таке градієнт та антиградієнт?
17. В чому полягає суть градієнтних методів пошуку оптимуму?
18. Чим градієнтний метод відрізняється від методу найшвидшого спуску?
19. Суть методу Флетчера-Рівса.
20. Особливості алгоритмів оптимізації багатовимірних функцій.
21. Класифікація методів безумовної оптимізації функцій багатьох змінних.
22. Як залежить час оптимізації від розмірності задачі?
23. Чи завжди при збільшенні розмірності задачі збільшується час оптимізації?
24. Які фактори впливають на час розв'язування задачі оптимізації?
25. Постановка задачі апроксимації.
26. Постановка задачі лінійного програмування.
27. До якого класу задач оптимізації відноситься задача лінійного програмування?
28. Методи розв'язування задач лінійного програмування.
29. Основі етапи симплекс-методу розв'язування задачі лінійного програмування.
30. З якою метою проводять аналіз чутливості оптимального розв'язку?

2.9. Задачі для самостійної роботи

Задача 2.9.1. Знаходження оптимальних розв'язків для розв'язування одно- та багатокритеріальних задач прийняття рішень.

Задача прийняття рішень (класична постановка задачі):

При заданих значеннях детермінованих факторів, з урахуванням стохастичних та невизначених факторів, знайти оптимальні значення керованих факторів із областей їх допустимих значень, які б оптимізували заданий критерій.

$$1. \quad \mathbf{Q} = < Q, A, I, B, D >,$$

де Q – критерій; A – множина альтернатив; I – наявна інформація; B – особливості особи, яка приймає рішення; D – правила вибору (прийняття рішення - ПР).

Класичні критерії прийняття рішення

Мінімаксний (MM) $Z_{mm} = \max_i (\min_j f_{ij})$

Байєса-Лапласа (BL) $Z_{BL.} = \max_i \sum_{j=1}^{n} f_{ij} \cdot p_j$

Критерій Севіджа (S) $Z_S = \min_i \max_j (\max_i f_{ij} - f_{ij})$

Похідні критерії прийняття рішення

Критерій Гурвиця (HW) $Z_{HW} = \max_i \left(c \cdot \min_j f_{ij} + (1-c) \cdot \max_j f_{ij} \right)$

Критерій Гермейєра (G) $Z_G = \max_i \left(\min_j f_{ij} p_j \right)$

2. Багатокритеріальна задача має вигляд: $< Q, A, I, B, D >$, $\mathbf{Q}=\{Q_1, Q_2, \ldots, Q_N\}$, де N – кількість критеріїв; Q – скінчена множина (вектор) критеріїв.

Для розв'язування багатокритеріальних задач використовується метод аналізу ієрархій (Analytic Hierarchy Process).

*Основні етапи **AHP**:*

1. Структуризувати задачу у вигляді ієрархічної структури: цілі; критерії; альтернативи.
2. Заповнити матрицю попарних порівнянь за перевагою критеріїв.
3. Заповнити матрицю попарних порівнянь для альтернатив за кожним із критеріїв.
4. Обчислити коефіцієнти важливості для критеріїв та альтернатив.
5. Обчислити кількісний індикатор якості кожної альтернативи за формулою:

$$Q^{ГП}(a_j) = \sum_{i=1}^{N} w_i V_{ij} ,$$

де $Q^{ГП}(a_j)$ – глобальний критерій для альтернативи a_j; w_i – ВКВК окремого критерію Q_i: V_{ij} – коефіцієнт важливості альтернативи a_j за критерієм Q_i.

Найкращою вважається альтернатива з максимальним значенням $Q^{ГП}$.

Порядок виконання роботи

1. Скласти та розв'язати задачу прийняття рішення у відповідності до варіанту, використовуючи необхідні критерії ПР або метод аналізу ієрархій (MAI).

Таблиця 2.10

Варіанти завдань для задачі 2.9.1

Варі-ант	Клас задач ПР	Використати	Кількість альтернатив	Стани природи	Кількість критеріїв
1	однокритеріальна	класичні критерії	10	10	
2	багатокритеріальна	MAI	5		2/3
3	однокритеріальна	класичні та похідні	8	12	
4	багатокритеріальна	MAI	3		5
5	однокритеріальна	похідні	12	8	
6	багатокритеріальна	MAI	4		4

2. Для варіантів 1,3,5 необхідно обов'язково використати декілька критеріїв; для решти – розрахувати індекс узгодженості та відношення узгодженості. Результати представити у вигляді таблиць.

Задача 2.9.2. Методи пошукової оптимізації функцій однієї змінної.

Оптимізація – це пошук найкращого рішення. В кожній задачі оптимізації використовуються такі поняття як критерій, керовані змінні та цільова функція. 7Критерій – це показник, який дозволяє визначити якість отриманого розв'язку задачі. Керовані змінні – це такі параметри задачі, значення яких можна змінювати. Цільова функція – це функція, що пов'язує керовані змінні та критерії таким чином, що дозволяє обчислити значення критерію при довільних значеннях керованих змінних.

Методи звуження інтервалу використовують таку теорему: якщо цільова функція *f(x)* унімодальна, неперервна і в точці x^* має оптимум, то для точок x_1, x_2, визначених за умовою $a < x_1 < x_2 < b$ існує два правила:

1) якщо функція $f(x_1) > f(x_2)$, то оптимум x^* належить відрізку $[x_1; b]$, $a \to x_1$;
2) якщо $f(x_1) < f(x_2)$, то оптимум x^* належить відрізку $[x_2; a]$, $b \to x_2$;
3) якщо не виконуються умови першого та другого правила, то оптимум належить відрізку $[x_1; x_2]$, $a \to x_1$, $b \to x_2$.

З використанням цієї теореми розроблено декілька алгоритмів пошуку оптимуму: алгоритм половинного ділення; метод дихотомії; алгоритм методу золотого перерізу; Фібоначчі.

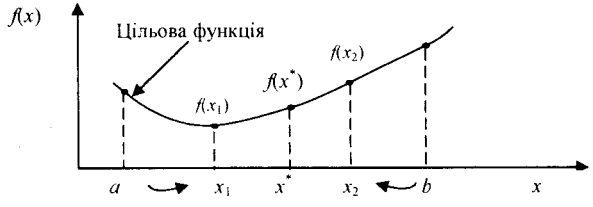


Рис. 2.33. Метод половинного ділення

Порядок виконання роботи

1. Скласти блок-схему алгоритму та розробити програму оптимізації функції $y = f(x)$, унімодальної на інтервалі $[a; b]$. Початкові дані брати з табл. 2.11 у відповідності до варіанту.

Таблиця 2.11

Варіанти завдань для задачі 2.9.2

Варіант	Цільова функція $f(x)$	Інтервал невизначеності $[a; b]$	Тип шуканого оптимуму	Точність рішення ε	Метод оптимізації
1	$y = 2^x/(0{,}5 - x)$	$[0{,}6; 3{,}6]$	max	0,01	половинного ділення
2	$y = (x-5)^2 + e^{x/2}$	$[-10; 10]$	min	0,1	дихотомії
3	$y = -e^{2+0.1x}$	$[-1; 1]$	max	0,01	золотого перерізу
4	$y = \sqrt{x^2 - 5x + 12}$	$[-20; 20]$	min	0,1	Фібоначчі

Продовження таблиці 2.11

5	$y = \sin(x)\ln(x)$	$[13; 17]$	max	0,01	половинного ділення
6	$y = -x \cdot e^{-0.1x}$	$[0; 50]$	min	0,1	дихотомії
7	$y = -x^2 + 3 \cdot x - 12$	$[-3; 3]$	max	0,01	золотого перерізу
8	$y = -\lvert x^3 - 50 \cdot x \rvert$	$[0; 7]$	min	0,1	Фібоначчі

2. Результати пошуку оптимуму функції представити:

а) для методу половинного ділення таблицею вигляду:

k	$a^{(k)}$	$b^{(k)}$	$L^{(k)}$	$x_1^{(k)}$	$x_m^{(k)}$	$x_2^{(k)}$	$y_1^{(k)}$	$y_m^{(k)}$	$y_2^{(k)}$
1									
...									
k									
$k+1$									

б) для методу дихотомії таблицею вигляду:

k	$a^{(k)}$	$b^{(k)}$	$L^{(k)}$	$x_1^{(k)}$	$x_m^{(k)}$	$x_2^{(k)}$	$y_1^{(k)}$	$y_2^{(k)}$
1								
...								
k								
$k+1$								

в) для методів золотого перерізу і Фібоначчі таблицею вигляду:

k	$a^{(k)}$	b	$L^{(k)}$	$x_1^{(k)}$	$x_2^{(k)}$	$y_1^{(k)}$	$y_2^{(k)}$
1							
...							
k							
$k+1$							

де k – номер кроку пошуку; L – довжина інтервалу невизначеності.

3. Достовірність отриманих результатів перевірити, використовуючи математичний пакет MathCAD. Зробити висновки. Оформити звіт.

Задача 2.9.3. Методи нульового порядку функції багатьох змінних.

До **методів нульового порядку** відноситься пошук за симплексом (не плутати з симплекс-методом в лінійному програмуванні), метод покоординатного спуску Гауса-Зейделя, випадковий пошук тощо. Розглянемо основні ідеї означених методів.

Пошук за симплексом полягає в тому, що при пошуку оптимуму в просторі незалежних змінних будується регулярний симплекс і обчислюється значення цільової функції в його вершинах. Регулярний симплекс в n-вимірному просторі представляє собою багатогранник, який має $n+1$ рівновіддалених вершин. Наприклад, у випадку двох змінних симплексом є рівнобічний трикутник; в трьохвимірному просторі симплекс представляє собою тетраедр. Після побудови симплекса визначається вершина, яка має найбільше значення цільової функції. На рис. 2.34 – це вершина з номером 1. Далі знайдена вершина проектується через центр ваги інших вершин симплекса (точки 2 і 3) в нову вершину (точка 4). Точки 2, 3 та 4 використовуються як вершини нового симплекса. Таким чином, трикутник ніби перевертається через сторону з найменшим значенням цільової функції. При пошуку мінімуму використовуються наступні два правила.

Правило "накриття" точки мінімуму. Якщо вершина, якій відповідає найбільше значення цільової функції, побудована на попередній ітерації, то замість неї береться вершина з меншим значенням цільової функції.

Правило циклічного руху. Якщо деяка вершина симплекса не виключається на протязі багатьох ітерацій, то необхідно зменшити розмір симплекса. Пошук завершується, коли розмір симплекса або різниця між значеннями функції в вершинах симплекса стають достатньо малими. Недоліком цього методу є велика кількість ітерацій; крім того, він не завжди забезпечує розв'язок задачі.

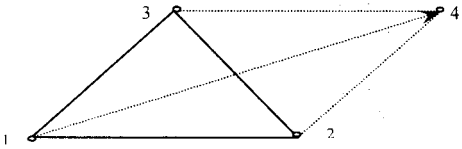


Рис. 2.34. Метод пошуку за симплексом

Ідея методу **покоординатного спуску** полягає в тому, що спочатку робиться крок в напрямку, який паралельний до однієї координатної осі, та обчислюється значення цільової функції. Якщо значення цільової функції зменшується, то рух продовжується далі в цьому ж напрямку, а якщо функція збільшується, то повертаємося назад і робимо крок в іншому напрямку. І так продовжується до тих пір, поки не буде знайдена оптимальна точка (рис. 2.35).

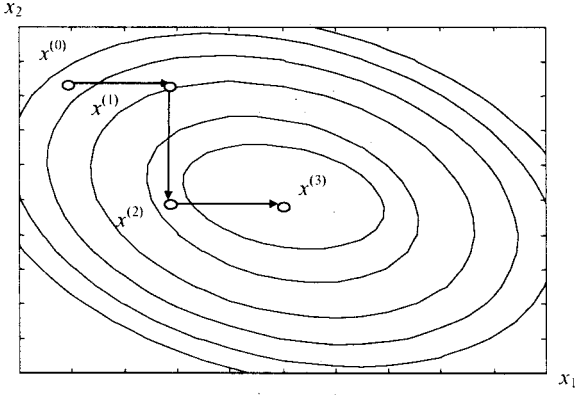


Рис. 2.35. Ілюстрація методу покоординатного спуску.

Особливість цього методу полягає в тому, що пошук оптимуму проводиться виключно паралельно до координатних осей. На початку пошуку вибирається великий крок і перевіряється значення функції за всіма напрямками. Якщо буде зростання функції за всіма напрямками, то необхідно зменшити крок.

Недоліком цього методу є обмежені можливості при пошуку оптимуму. Метод застосовують тоді, коли залежність між змінними $x_1, x_2, \ldots x_n$ практично відсутня.

Ідея **випадкового пошуку** полягає в тому, що вибір напрямку руху здійснюється випадково. Якщо цільова функція зменшується, то рух у вибраному напрямку продовжується. В протилежному випадку – необхідно повернутися на один крок назад

та знову випадково обрати напрямок пошуку тощо. Усі випадкові методи пошуку реалізуються за ітераційною формулою: $X^{(k+1)} = X^{(k)} + \xi^{(k)}$, де k – номер ітерації; $\xi^{(k)}$ – випадкова величина. Популярність методів випадкового пошуку пояснюється їх простотою та можливостями для користувачів модифікувати методи.

Порядок виконання роботи

1. Скласти схему алгоритму та програму оптимізації функції $y = f(x_1, x_2)$ методом нульового порядку. Початкові дані є в табл. 2.12 у відповідності до варіанту.
2. Достовірність отриманих результатів перевірити, використовуючи математичний пакет MathCAD.
3. Зробити висновки. Оформити звіт.

Таблиця 2.12

Варіанти завдань для задачі 2.9.3

Варіант	Цільова функція y	Точність розв'язку ε	Метод оптимізації
1	$y = (x_1 - 2)^2 + (x_2 - 4)^2$	0,001	Хука-Дживса
2	$y = x_1^2 + x_2^2 - x_1 x_2$	0,0001	Покоординатного спуску
3	$y = e^{x_1^2 + x_2^2} + 2x_1 - 3{,}5x_2$	0,001	Нелдера-Міда
4	$y = (x_1 + 3)^2 + (x_2 + 5)^2$	0,01	Хука-Дживса
5	$y = e^{2x_1^2 + x_2^2} + 1{,}1x_1 + 3{,}6x_2$	0,001	Покоординатного спуску
6	$y = e^{x_1^2 + 0{,}8x_2^2} + 1{,}2x_1 + 2x_2$	0,01	Нелдера-Міда

Задача 2.9.4. Градієнтні методи оптимізації функцій багатьох змінних.

Градієнтні методи відносяться до методів першого порядку. Особливість цих методів – це використання градієнта. *Градієнтом функції* f(X) називається вектор, величина якого визначає швидкість зміни функції f(X), а напрямок співпадає з напрямком найбільшого зростання цієї функції. Вектор, який протилежний за напрямком градієнту називається *антиградієнтом*. Всі методи першого порядку, базуються на ітераційній процедурі, яка реалізується за формулою:

$$X^{(k+1)} = X^{(k)} + \lambda^{(k)} \cdot \nabla f(X^{(k)}),$$

де k – номер ітерації; $X^{(k)}$ – поточне наближення до розв'язку; $\lambda^{(k)}$ – параметр, який характеризує довжину кроку; $\nabla f(X^{(k)})$ – напрямок пошуку в n-мірному просторі керованих змінних.

Градієнтний метод є послідовністю кроків, кожен з яких складається з 2 операцій:

1) визначення напрямку найбільшої крутизни спуску, тобто напрямок антиградієнта функції f(X).
2) переміщення в вибраному напрямку на задану відстань.

Градієнтний метод має свої недоліки. Для того, щоб рухатися завжди в напрямку антиградієнта, крок повинен бути невеликим. Тому їх буде багато. Але, оскільки, на кожному кроці необхідно постійно обчислювати градієнт, то наближення до оптимуму буде повільним. Метод **найшвидшого спуску** є модифікацією градієнтного метода. Він відрізняється від градієнтного тим, що градієнт обчислюється тільки в початковій точці і рух в напрямку антиградієнта продовжується до тих пір, поки зменшується значення цільової функції f(X). Вибір кроку λ при переході від точки $X^{(k)}$ до $X^{(k+1)}$ визначається

згідно умови $\lambda^{(k)} = \min_{\lambda \geq 0} f(X^{(k)} + \lambda^{(k)} \cdot \nabla f(X^{(k)}))$, тобто на кожному кроці розв'язується одномірна задача мінімізації.

Метод найшвидшого спуску має два недоліки: по-перше, методу властива поступова збіжність до точки мінімуму внаслідок малого $\nabla f(X)$ в околі цієї точки, по-друге, необхідно розв'язувати задачу одномірної оптимізації – обирати на кожному кроці оптимальне значення $\lambda^{(k)}$. Головна перевага цього метода полягає в тому, що йому властива стійкість та простота обчислень. Використовують цей метод тоді, коли цільову функцію можна добре апроксимувати лінійною залежністю.

Порядок виконання роботи:

1. Скласти схему алгоритму та розробити програму пошуку локального мінімуму функції $y = f(x_1, x_2)$ градієнтним методом. Вихідні дані брати з табл. 2.13 у відповідності до варіанту.

Таблиця 2.13

Варіанти завдань для задачі 2.9.4

Варіант	Цільова функція y	Початкова точка x^0	Початкова довжина кроку λ	Точність розв'язку ε	Метод оптимізації
1	$y = (x_1 - 2)^2 + (x_2 - 4)^2$	$[1{,}2; 3{,}4]$	0,1	0,001	Флетчера-Рівса
2	$y = x_1^2 + x_2^2 - x_1 x_2$	$[0{,}1; 0]$	0,01	0,0001	Найшвидшого спуску
3	$y = e^{x_1^2 + x_2^2} + 2x_1 - 3{,}5x_2$	$[0; 0]$	0,1	0,001	Градієнтного спуску
4	$y = (x_1 + 3)^2 + (x_2 + 5)^2$	$[-4; -9]$	0,1	0,01	Флетчера-Рівса
5	$y = e^{2x_1^2 + x_2^2} + 1{,}1x_1 + 3{,}6x_2$	$[-1; 0{,}5]$	0,1	0,001	Найшвидшого спуску
6	$y = e^{x_1^2 + 0{,}8x_2^2} + 1{,}2x_1 + 2x_2$	$[1; -0{,}2]$	0,1	0,01	Градієнтного спуску

2. Результати пошуку оптимуму функції представити таблицею вигляду:

Крок пошуку k	Поточна точка x	Значення функції y	Поточна довжина кроку λ	Градієнт G	Норма градієнта $\|G\|$
1					
...					
k					
...					
N					

3. Достовірність отриманих результатів перевірити, використовуючи математичний пакет MathCAD. Зробити висновки. Оформити звіт.

Приклад програми пошуку оптимуму функції двох змінних у MathCAD.

1. Записуємо цільову фнкцію

$$f(x1,x2) := e^{\left(x1^2+x2^2\right)} + 2\cdot x1 - 3.5\cdot x2$$

2. Визначаємо і виводимо частинні похідні цільової функції f'x1(x1.x2) і f'x2(x1.x2)

$$fx1(x1,x2) := \frac{d}{dx1} f(x1,x2) \quad fx1(x1,x2) \rightarrow 2\cdot x1\cdot\exp\left(x1^2 + x2^2\right) + 2$$

$$fx2(x1,x2) := \frac{d}{dx2} f(x1,x2) \quad fx2(x1,x2) \rightarrow 2\cdot x2\cdot\exp\left(x1^2 + x2^2\right) - 3.5$$

3. Розвязуємо систему рівнянь f'x1(x1.x2)=0 f'x2(x1.x2)=0, використовуючи процедуру given-find

Given

$$fx1(x1,x2) = 0$$

$$fx2(x1,x2) = 0$$

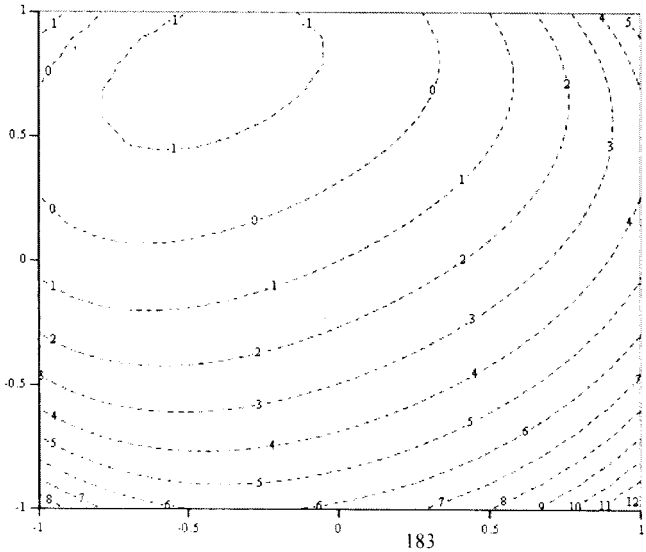
$$\text{Find}(x1,x2) \rightarrow \begin{pmatrix} -.4458874951226988477 \\ .7803031164647229834 \end{pmatrix}$$

4. Виводимо значення цільової функції в точці мінімуму x=-0.4459 y=0.7803

$$f(-0.4459, 0.7803) = -1.38$$

5. Будуємо лінії рівня цільової функції, використовуючи функцію CreateMesh

$$Z := \text{CreateMesh}(f, -1, 1, -1, 1)$$



Z

Задача 2.9.5. Дослідження залежності часу оптимізації від розмірності задачі.

Збільшення кількості керованих змінних (розмірності) істотно ускладнює розв'язування задачі оптимізації. А при деякому значені починається різке зростання часу обчислення оптимуму (рис. 2.36). Таке явище в оптимізації називається проблемою "прокляття" розмірності. Однією з задач сучасних методів оптимізації є розробка ефективних алгоритмів, що дозволяють відсунути стіну складності. *Ефективним алгоритмом* вважається такий алгоритм, складність якого (кількість ітерацій) описується поліноміальною функцією від розмірності задачі.

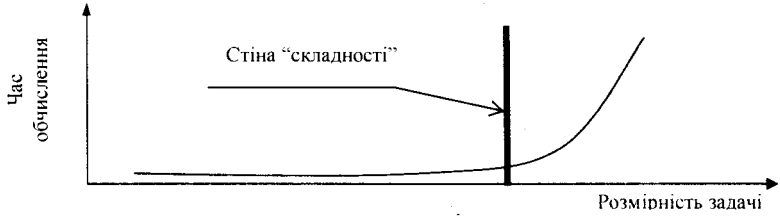


Рис. 2.36. Проблема "прокляття" розмірності

Порядок виконання роботи

1. Розробити програму для дослідження залежності часу розв`язування задачі безумовної оптимізації від кількості керованих змінних.
2. Вихідні дані брати з табл. 2.14 у відповідності до варіанту.

Таблиця 2.14

Варіанти завдань для задачі 2.9.5

Варіант	Цільова функція $f(x)$	Початкова точка x^0	Варіант	Цільова функція $f(x)$	Початкова точка x^0
1	$y=\sum_{i=1}^{n}\frac{(x_i-n)^2}{i^2}$	$x_i^0=2i,\ i=\overline{1,n}$	4	$y=\sum_{i=1}^{n}\frac{(x_i+i)^2}{0,1i^2}$	$x_i^0=i,\ i=\overline{1,n}$
2	$y=\sum_{i=1}^{n}\frac{(x_i+7i)^2}{i}$	$x_i^0=0,\ i=\overline{1,n}$	5	$y=\sum_{i=1}^{n}\frac{(x_i+i)^4}{i}$	$x_i^0=5i,\ i=\overline{1,n}$
3	$y=\sum_{i=1}^{n}(x_i-i)^4$	$x_i^0=\frac{i}{2},\ i=\overline{1,n}$	6	$y=\sum_{i=1}^{n}(0,1x_i+2i)^2$	$x_i^0=3,\ i=\overline{1,n}$

3. Для пошуку мінімуму функції використати будь-який градієнтний метод.
4. Експеримент провести в діапазоні від n=2 до n=30.
5. Результати експерименту представити табл 2.15.

Таблиця 2.15

Варіанти завдань для задачі 2.9.5

Кількість керованих змінних n	Час пошуку оптимуму t, сек
2	
...	
30	

6. Достовірність отриманих результатів перевірити, використовуючи математичний пакет MathCAD. За допомогою методу найменших квадратів

апроксимувати отримані експериментальні дані функцією $a_0 + a_1 e^{a_2 n}$ та вивести графік цієї функції. Зробити висновки. Оформити звіт.

Приклад програми опрацювання результатів експерименту у MathCAD.

1. Задасмо значення змінних, необхідних для роботи з масивами

ORIGIN := 1 N := 20 i := 2.. N

2. Вводимо значення часу оптимізації, отримані в результаті експерименту

$T := (2\ 2.5\ 3.2\ 4.5\ 6.2\ 8\ 10.6\ 14\ 19\ 26\ 38\ 52\ 76\ 106\ 140\ 200\ 280\ 400\ 600)^T$

3. Записуємо цільову функцію

$$g(a0, a1, a2) := \left| \begin{array}{l} z \leftarrow 0 \\ \text{for } j \in 1..N-1 \\ \quad z \leftarrow z + \left[T_j - \left(a0 + a1 \cdot e^{a2 \cdot j} \right) \right]^2 \end{array} \right.$$

4. Формуємо початкове наближення

a0 := 2 a1 := 1 a2 := 1

5. Знаходимо мінімум функції g(a0, a1, a2) за допомогою функції Minimize

$$P := \text{Minimize}(g, a0, a1, a2) \qquad P = \begin{pmatrix} 0.503 \\ 0.613 \\ 0.361 \end{pmatrix}$$

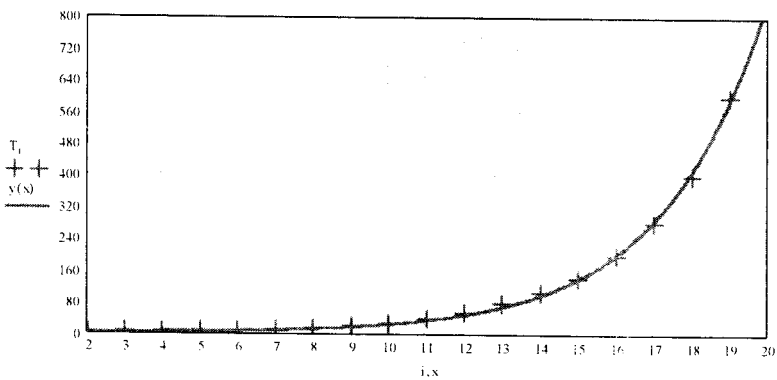
6. Визначаємо значення функції g(a0, a1, a2) в точці мінімуму

$$gmin := \left| \begin{array}{l} y \leftarrow 0 \\ \text{for } j \in 1..N-1 \\ \quad y \leftarrow y + \left[T_j - \left(P_1 + P_2 \cdot e^{P_3 \cdot j} \right) \right]^2 \end{array} \right. \qquad gmin = 495.88$$

7. Задасмо отриману апроксимувальну функцію

$$y(x) := P_1 + P_2 \cdot e^{P_3 \cdot x}$$

8. Виводимо значення експериментальних даних та графік апроксимувальної функції



Задача 2.9.6. Аналіз чутливості оптимального розв'язку задачі лінійного програмування, використовуючи програмні пакети MathCAD та MS Excel.

Аналіз чутливості – це процедура, яка дозволяє встановити залежність оптимального розв'язку від варіації початкових даних. Аналіз чутливості відіграє велику роль в задачах оптимізації та проводиться за 2 причинами:

1. Деякі параметри задач лінійного програмування, такі, як фінанси, запаси ресурсів можна регулювати. Аналіз чутливості дозволяє оцінити вплив зміни цих параметрів на оптимальний розв'язок. Якщо виявитися, що оптимальний розв'язок (наприклад, відношення прибутку до витрат) можна значно покращити за рахунок невеликих змін параметрів, то необхідно провести ці зміни.

2. В багатьох випадках оцінки параметрів отримуються шляхом статистичної обробки експериментальних даних. Тому такі оцінки не можуть бути точними. Якщо вдається визначити, які параметри більшою мірою впливають на значення цільової функції, то необхідно збільшити точність їх оцінок.

Важливу роль при аналізі чутливості виробничих задач відіграють тіньові ціни та маргінальні оцінки. Для цього використовують значення тіньових цін та маргінальні оцінки. Тіньові ціни визначають приріст максимального прибутку при використані додаткової одиниці деякого ресурсу. Значення маргінальної оцінки показує наскільки знижується максимальний прибуток при випуску одиниці цієї продукції.

Порядок виконання роботи

1. Самостійно придумати та розв`язати задачу лінійного програмування (кількість керованих змінних повинна бути не менша 4, кількість обмежень на значення керованих змінних – не менша 3).

2. Визначити зону нечутливості оптимального розв'язку задачі лінійного програмування до варіації початкових даних. Експерименти проводити у MathCAD.

3. Ознайомитися з надбудовою MS Excel „пошук розв'язку" та отримати за її допомогою розв'язок задачі. Зробити висновки. Оформити звіт.

Приклад програми опрацювання результатів експерименту у MathCAD.

1. Записуємо цільову функцію

$f(x1, x2, x3, x4) := x1 + 2 \cdot x2 - x3 + x4$

2. Формуємо початкове наближення (можна задати будь-які числа)

$x1 := 0 \quad x2 := 1 \quad x3 := 0 \quad x4 := 5$

3. Описуємо обмеження

Given

$x1 + 2x2 + x4 = 4$

$x1 + x2 + x3 = 8$

$x1 \geq 0 \quad x2 \geq 0 \quad x3 \geq 0 \quad x4 \geq 0$

4. Знаходимо мінімум цільової функції за допомогою функції Minimize (для пошуку максимуму використовують функцію Maximize)

Reshenie := Minimize (f, x1, x2, x3, x4)

5. Виводимо точку мінімуму та значення цільової функції в цій точці

$Reshenie^{T} = (0 \quad 0 \quad 8 \quad 4)$

$f(0, 0, 8, 4) = -4$

Розділ 3. Основи статистичного моделювання

- Моделювання випадкових величин на комп'ютері
  - Моделювання на комп'ютері рівномірно розподіленої величини α на [0,1], [a,b]
  - Моделювання на комп'ютері випадкової величини з довільним законом розподілу
  - Моделювання нормально-розподіленої випадкової величини ξ з $M\xi=0$, $D\xi=1$
- Обчислення інтегралів методом статистичного моделювання
  - Аналітичний метод обчислення однократного інтеграла
  - Метод площ обчислення визначеного інтеграла
  - Обчислення кратних інтегралів
    - Метод *n*-кратних об'ємів
    - Метод середнього значення (аналітичний метод)
- Розв'язання систем лінійних алгебраїчних рівнянь
  - Метод Дж. фон Неймана-Улама
  - Побудова імітаційної моделі розв'язання системи лінійних алгебраїчних рівнянь (СЛАР)
  - Випадкове блукання частинки з поглинаючим екраном
  - Теорема про побудову $X_r(\omega)$
  - Обчислення оберненої матриці методом статистичного моделювання
  - Алгоритм методу статистичного моделювання розв'язку СЛАР
- Розв'язок першої крайової задачі для рівняння $\Delta u=f$
- Марківські ланцюги та інтегральні рівняння
  - Загальна постановка задачі
  - Лінійні оцінки та умови незміщеності
- Методи статистичного моделювання для розв'язування стохастичних диференціальних рівнянь
  - Загальна постановка задачі
  - Збіжність методу Ейлера в середньому квадратичному
  - Збіжність методу Ейлера як збіжність розподілів
  - Оптимальні різницеві схеми розв'язку стохастичних рівнянь
- Наближене інтегрування стохастичних диференціальних рівнянь
  - Однокрокова апроксимація
  - Апроксимація на скінченному проміжку

- Узагальнення на системи рівнянь

- Методи побудови емпіричних формул за вибірковими значеннями
  - Загальна постановка задачі
  - Лінійна залежність
  - Метод вирівнювання
  - Квадратична (параболічна) залежність
  - Визначення параметрів емпіричної формули
  - Метод вибраних точок
  - Метод середніх
  - Метод найменших квадратів
  - Деякі міркування про вибір емпіричної формули з двома параметрами
  - Емпіричні формули, які містять три параметри
  - Уточнення отриманої емпіричної формули
  - Загальний метод визначення параметрів емпіричної формули
- Метод статистичного аналізу економетрики
  - Метод кореляційного аналізу
  - Метод регресійного аналізу
  - Задача оптимізації статистичного аналізу складних процесів
  - Алгебра планування експерименту
  - Метод Бокса-Вілсона
- Приклади математичного опису складних об'єктів експериментально-статистичними методами
  - Математичний опис промислового об'єкту методом пасивного експерименту
  - Математичний опис напіввиробничого об'єкту методом активного експерименту
- Кількість спостережень
- Квантування за часом
- Метод експертних оцінок перевірки адекватності математичних моделей
- Комп'ютерні технології ідентифікації виробничих агрегатів
- Методичні прийоми дослідження впливу факторів на виробничі і економічні процеси
- Контрольні питання
- Задачі для самостійної роботи
- Завдання до лабораторної роботи

Статистичне моделювання – це чисельний метод вирішення математичних завдань, при якому шукані величини подають ймовірнісними характеристиками якого-небудь випадкового явища. Це явище моделюється, після чого необхідні характеристики наближено визначають шляхом статистичного опрацювання результатів спостережень цієї моделі.

Статистичне моделювання сформувалося як самостійний науковий напрям в середині XIX – на початку XX ст., хоча виникли значно раніше. Вони базуються на теорії ймовірностей, яка має досить тривалу історію становлення й сформувалася в самостійний напрям в XVI ст. (Б. Паскаль, П. Ферма), а незабаром ймовірнісні моделі стали використовуватися при опрацюванні статистичних даних [1-93, 103-109, 112-114, 125-136, 293-302]. Важливий внесок у теорію ймовірності Я.Бернуллі: він довів закон великих чисел в простому випадку незалежних випробувань. У першій половині XIX ст. теорія ймовірностей починає застосовуватися до аналізу помилок спостережень; П. Лаплас і С. Пуассон довели перші граничні теореми. У другій половині XIX ст. – основний внесок П. Л. Чебишева, А. А. Маркова і А. М. Ляпунова. У цей час були доведені закон великих чисел, центральна гранична теорема, а також розроблена теорія ланцюгів Маркова. Сучасний вигляд теорія ймовірностей отримала завдяки аксіоматизації, що була запропонована А. Н. Колмогоровим [94-102, 110-111, 137-292, 303-320]. Основу статистичних методів складає відображення явищ і процесів за допомогою випадкових (стохастичних) подій та їх динаміки, які описуються відповідними імовірнісними (статистичними) характеристиками і статистичними закономірностями [94-102, 110-111, 137-292, 303-320].

Термін "стохастичні" уточнює поняття "випадковий", яке в повсякденному сенсі прийнято пов'язувати з відсутністю причин появи подій, з появою не тільки повторюваних і підлеглістю деяким закономірностям, але і одиничних подій. Процеси ж, які відображаються статистичними закономірностями, повинні бути жорстко пов'язані з наперед заданими, певними причинами, а «випадковість» означає, що вони можуть з'явитися або не з'явитися при наявності заданого комплексу причин.

Статистичні відображення системи в загальному випадку (за аналогією з аналітичними) можна подати символічним чином, ніби у вигляді «розмитої» точки (розмитої області) в деякому просторі, в яку переводить враховані в моделі властивості системи [94-102, 110-111, 137-292, 303-320]. Межі області задані з деякою ймовірністю $p(x)$ («розмиті»), і рух точки описується деякою випадковою функцією.

Метод статистичного моделювання (МСМ) полягає у розв'язуванні різноманітних математичних задач шляхом побудови для шуканої детермінованої невідомої X такої випадкової величини (ВВ) $\xi \equiv \xi(\omega): \Omega \to \mathbf{R}^1(\mathbf{R}^n)$ (ВВ ξ задана на імовірнісному просторі $(\Omega, \Im, P)$), щоб $X = M\xi(\omega)$). Таким чином, при розв'язуванні задач МСМ можна виділити три етапи. Нехай X – невідоме математичної задачі (розв'язок СЛАР, значення визначеного інтегралу, розв'язок задачі Діріхле тощо).

I. Будують випадкову величину $\xi \equiv \xi(\omega)$ з законом розподілу $F_\xi(x)(F'_\xi(x) = p_\xi(x))$ таким чином, щоб $X = M\xi$.

II. За допомогою комп'ютера знаходимо N значень ВВ ξ, тобто $\xi_1, \xi_2, ..., \xi_n$ за законом розподілу $p_\xi(x)$.

III. Обчислюємо статистичне математичне сподівання для ВВ зі значеннями $\xi_1, \xi_2, ..., \xi_n$ для $\overline{\xi} = \frac{1}{N}\sum_{i=1}^{N} \xi_i \approx M\xi = X$. Точність $\varepsilon > 0$ для $\overline{\xi}$ визначається за теоремою А.М. Колмогорова $\lim P\left\{\omega : \left|\frac{1}{N}\sum_{i=1}^{N} \xi_i - M\xi\right| \le \varepsilon\right\} = 1$, де $M\xi < \infty$.

3.1. Моделювання випадкових величин на комп'ютері

3.1.1. Моделювання на комп'ютері рівномірно розподіленої величини α на [0,1], [a,b]

Нехай побудовано ймовірнісний простір $(\Omega, \Im, P)$ для *CE*, на ньому задано α=α(ω) – рівномірно розподілена випадкова величина (*p.p.BB*) на [0,1], тобто щільність $p_\alpha(x)$ або функція розподілу $F_\alpha(x)$ задана відповідно

$$p_\alpha(x) = \begin{cases} 1, & x \in [0,1], \\ 0, & x \notin [0,1]. \end{cases} \tag{3.1}$$

$$F_\alpha(x) = \begin{cases} x, & x \in [0,1], \\ 1, & x > 1, \\ 0, & x < 0. \end{cases} \tag{3.2}$$

Відомо, що $M\alpha=1/2$; $D\alpha=1/12$. Найбільш розповсюджений метод – це метод одержання так званих псевдовипадкових значень *BB* (випадкової величини) за допомогою

$$X_{n+1} = K \cdot x_n \,(\mathrm{mod} 2^m), \tag{3.3}$$

де $x_n=1$, K – просте число, x_{n+1} дорівнює залишку від ділення $K \cdot x_n$ на 2^m(n=0,1,2,…); m – кількість бітів у представленні числа на комп'ютері.

Далі значення *p.p.BB* α(ω) одержуємо, якщо поділити значення x_n за формулою (3.3) на 2^m ($0 \le x_n \le 2^m$):

$$0 < \alpha_i = \frac{x_i}{2^m} < 1. \tag{3.4}$$

Таким чином, маємо числову послідовність

$$\{\alpha_i\} \subset (0,1). \tag{3.5}$$

У математичному забезпеченні комп'ютера можна знайти функції обчислення за формулою (3.4) числових значень *p.p.BB* α.

Лема 3.1 (про моделювання *p.p.BB* на [*a*,*b*]).

Якщо α=α(ω) – *p.p.BB* з [0,1], то $\alpha^* = (b-a)\alpha + a$ є *p.p.BB* на [*a*,*b*], тобто

$$p_{\alpha^*}(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a,b], \\ 0, & x \notin [a,b]. \end{cases} \tag{3.6}$$

Доведення. Обчислимо F_α^* за значенням. Тоді

$$F_{\alpha^*}(x) = P\{\omega : (b-a)\alpha(\omega) + a < x\} =$$

$$= P\{\omega : \alpha(\omega) < \frac{x-a}{b-a}\} = F_\alpha(\frac{x-a}{b-a}) = \left\{ \begin{array}{l} 0, якщо \frac{x-a}{b-a} < 0, \\ \frac{x-a}{b-a}, якщо 0 \le \frac{x-a}{b-a} \le 1, \\ 1, якщо \frac{x-a}{b-a} > 1 \end{array} \right\} =$$

$$= \left\{ \begin{array}{l} 0, якщо x < a \\ \frac{x-a}{b-a}, якщо x \in [a,b] \\ 0, якщо x > b \end{array} \right\}, \quad P_\alpha = \frac{dF_\alpha \cdot(x)}{dx} = \left\{ \begin{array}{l} 0, для x \notin [a,b] \\ \frac{1}{b-a}, для x \in [a,b] \end{array} \right\}$$

Отже, довели (3.6).

Теорема 3.1. Нехай $F_\alpha(x)$ функція розподілу неперервної *р.р.ВВ* α на [0,1];

$$\{\beta_i = \beta_i(\omega)\}, i = 1,2,...,N-1. \tag{3.7}$$

Незалежні в сукупності бернулеві дискретні *ВВ*;

$$P(\beta_i = 1) = P(\beta_i = 0) = \frac{1}{2}. \tag{3.8}$$

Тоді функція розподілу *ВВ* $\sum_{i=1}^{N-1} \frac{\beta_i}{2^i}$ відрізняється від функції розподілу *р.р.ВВ* α не більше ніж на $\frac{1}{2^{N-1}}$

$$P\left\{\omega : \sum_{i=1}^{N-1} \frac{\beta_i}{2^i} \le x\right\} - P\{\omega : \alpha(\omega) < x\} \le \frac{1}{2^{N-1}}. \tag{3.9}$$

Лема 3.2. Нехай $\{\beta_i\}$, i=1,2,...,N-1, незалежні в сукупності бернулеві дискретні *ВВ* зі знаком розподілу (3.8) і

$$\beta_N = \beta_N(\omega) - р.р.ВВна[0,1]. \tag{3.10}$$

$$\alpha^{**}(\omega) = \sum_{i=1}^{N-1} \frac{\beta_i}{2^i} + \frac{\beta_N}{2^{N-1}}. \tag{3.11}$$

є *р.р.ВВ* на [0,1].

Доведення проведемо методом математичної індукції. *ВВ* незалежні, якщо за означенням шукана функція дорівнює добутку одновимірних функцій розподілу

$$F_{\beta_1\beta_2...\beta_N}(x_1, x_2, ... x_N) = F_{\beta_1}(x_1)F_{\beta_2}(x_2)...F_{\beta_N}(x_N). \tag{3.12}$$

Використаємо формулу повної ймовірності.

Нехай A – подія H_i – несумісні гіпотези: $H_i \cap H_j = \varnothing, i \ne j, \bigcup H_i = \Omega$. Тоді

$$P(A) = \sum_{i=1}^{N} P(H_i)P(A/H_i). \tag{3.13}$$

Доведемо твердження для N=2; тобто $\alpha^{**}(\omega) = \beta_1/2 + \beta_2/2$ є *p.p.BB* на [0,1], тобто її закон розподілу є рівномірно розподіленим:

$$F_{\frac{\beta_1}{2}+\frac{\beta_2}{2}}(x) = \left\{ \begin{array}{c} 0, x < 0 \\ x, x \in [0,1] \\ 1, x > 1 \end{array} \right\}. \tag{3.14}$$

$$P\left\{\omega : \frac{\beta_1}{2} + \frac{\beta_2}{2} \le x\right\} = P\{\omega : \beta_1 = 0\} P\left\{\omega : \frac{\beta_1}{2} + \frac{\beta_2}{2} \le x / \beta_1 = 0\right\} +$$

$$+ P\{\omega : \beta_1 = 1\} P\left\{\omega : \frac{\beta_1}{2} + \frac{\beta_2}{2} \le x / \beta_1 = 1\right\} = \frac{1}{2} P\left\{\omega : \frac{\beta_2}{2} \le x\right\} + \tag{3.15}$$

$$+ \frac{1}{2} P\left\{\omega : \frac{1}{2} + \frac{\beta_2}{2} \le x\right\} = \frac{1}{2} P\{\omega : \beta_2 \le 2x\} + \frac{1}{2} P\left\{\omega : \frac{\beta_2}{2} \le x - \frac{1}{2}\right\}.$$

Обчислимо згідно $\beta_2 = \beta_2(\omega) \in p.p.BB$ на [0,1] з законом розподілу (3.1):

$$P\{\omega : \beta_2 \le 2x\} = \left\{ \begin{array}{c} 0 для 2x < 0, \\ 2x для 0 \le 2x \le 1, \\ 1 для 2x > 1 \end{array} \right\} = \left\{ \begin{array}{c} 0, якщо x < 0, \\ 2x, якщо 0 \le x \le \frac{1}{2}, \\ 1, якщо x > \frac{1}{2}, \end{array} \right\}. \tag{3.16}$$

$$P\{\omega : \beta_2 \le 2x - 1\} = \left\{ \begin{array}{c} 0 для 2x - 1 < 0, \\ 2x - 1 для 0 \le 2x - 1 \le 1, \\ 1 для 2x - 1 > 1 \end{array} \right\} = \left\{ \begin{array}{c} 0, якщо x < \frac{1}{2}, \\ 2x - 1, якщо \frac{1}{2} \le x \le 1, \\ 1, якщо x > 1, \end{array} \right\}. \tag{3.17}$$

Тоді (3.15) з врахуванням (3.16), (3.17) має вигляд

$$P\left\{\omega : \frac{\beta_1}{2} + \frac{\beta_2}{2} \le x\right\} = P\left\{\omega : \frac{\beta_1}{2} + \frac{\beta_2}{2} < x\right\} =$$

$$= \left\{ \begin{array}{c} \frac{1}{2}0 + \frac{1}{2}0, якщо (x < 0) \cap (x < \frac{1}{2}) = (x < 0,) \\ \frac{1}{2}2x, якщо 0 < x \le \frac{1}{2}, \\ \frac{1}{2}[(2x - 1) + 1], якщо \frac{1}{2} < x \le 1, \\ \frac{1}{2}[1 + 1], якщо \left(x > \frac{1}{2}\right) \cap (x > 1) = x > 1 \end{array} \right\} = \left\{ \begin{array}{c} 0, x < 0, \\ x, 0 \le x \le 1, \\ 1, x > 1 \end{array} \right\}$$

Що і доводить твердження леми 3.2 при N=2. Далі нехай лема 3.2 має місце для N=k, тобто $\beta_1, \beta_2, \ldots, \beta_{k-1}$ бернулеві *BB*, β_k – *p.p.BB* α на [0,1]), а значить

$$P\left\{\omega:\frac{\beta_1}{2}+\frac{\beta_2}{2^2}+\ldots\frac{\beta_{k-1}}{2^{k-1}}+\frac{\beta_{k-2}}{2^{k-1}}<x\right\}=\left\{\begin{array}{l}0,x<0\\x,0\le x\le 1\\1,x>1\end{array}\right\}. \tag{3.18}$$

Доведемо, що для $N=k+1$ лема 3.2 має місце, де $\{\beta_1,\beta_2,\ldots,\beta_k\}$ – бернулеві випадкові величини незалежні в сукупності β_{k+1}– *p.p.BB* на [0,1]. Беремо

$$\begin{aligned}a^{***}&=\frac{\beta_1}{2}+\frac{\beta_2}{2^2}+\ldots+\frac{\beta_k}{2^k}+\frac{\beta_{k+1}}{2^{k+1}}=\frac{\beta_1}{2}+\frac{1}{2}\left[\frac{\beta_2}{2}+\frac{\beta_3}{2^2}+\ldots+\frac{\beta_k}{2^{k-1}}+\frac{\beta_{k+1}}{2^k}\right]=\\&=\frac{\beta_1}{2}+\frac{1}{2}a^{**}(\omega)\end{aligned} \tag{3.19}$$

яка за лемою 3.2 є *p.p.BB* на [0,1]. Аналогічно до випадку $N=2$ можемо довести, що a^{***} є *p.p.BB* на [0,1]. Проведемо ***доведення*** теореми 3.1. Нехай β_N – *p.p.BB* на [0,1] і не залежить від бернулевих *BB* $\{\beta_i\}(i=1,2,\ldots,N)$. Тоді

$$\begin{aligned}P\left\{\omega:\sum_{i=1}^{N-1}\frac{\beta_i}{2^i}\le x\right\}&=P\left\{\omega:\alpha^{**}(\omega)-\frac{\beta_N(\omega)}{2^{N-1}}\le x\right\}=\\&=P\left\{\omega:\alpha^{**}(\omega)\le x+\frac{\beta_N(\omega)}{2^{N-1}}\right\},\end{aligned} \tag{3.20}$$

де $\alpha^{**}=\sum_{i=1}^{N-1}\frac{\beta_i}{2^i}+\frac{\beta_N}{2^{N-1}}$ - *p.p.BB* за лемою 3.2.

За формулою повної ймовірності $H_1=\{\omega:\beta_N=0\}$; $H_2=\{\omega:0<\beta_N\le 1\}$.

$$P\left\{\omega:\alpha^{**}\le x+\frac{\beta_N(\omega)}{2^{N-1}}\right\}=P\{\omega:\beta_N=0\}P\left\{\omega:\alpha^{**}\le x+\frac{\beta_N}{2^{N-1}}\Big|_{\beta_N=0}\right\}+$$

$$+P\{\omega:0<\beta_N\le 1\}P\left\{\omega:\alpha^{**}\le x+\frac{\beta_N}{2^{N-1}}\Big|_{0<\beta_N\le 1}\right\}\le$$

$$\le 1\cdot\left\{\omega:\alpha^{**}\le x+\frac{1}{2^{N-1}}\right\}\le P\{\omega:\alpha^{**}<x\}+P\left\{\omega:x\le\alpha^{**}\le x+\frac{1}{2^{N-1}}\right\}=F_{\alpha^{**}}(x)+\frac{1}{2^{N-1}}.$$

Тобто маємо твердження теореми 3.1.

Наслідок 3.1. На $(\Omega,\Im,P)$ з ймовірністю 1 сума ряда дискретних випадкових бернулевих велечин з вагою $1/2^i$ дорівнює *p.p.BB* α на [0,1], тобто

$$P\left\{\omega:\sum_{i=1}^{\infty}\frac{\beta_i}{2^i}=\alpha\right\}=1. \tag{3.21}$$

Доведення. З доведення теореми 3.1 випливає

$$0\le P\left\{\omega:\sum_{i=1}^{N-1}\frac{\beta_i}{2^i}<x\right\}\le F_\alpha(x)+\frac{1}{2^{N-1}}. \tag{3.22}$$

яке еквівалентне

$$0 \le P\left\{\omega : \sum_{i=1}^{N-1} \frac{\beta_i}{2^i} < x\right\} - F_\alpha(x) \le \frac{1}{2^{N-1}}. \tag{3.23}$$

З неперервності ймовірності $(\lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n))$ випливає, що

$$P\left\{\omega : \sum_{i=1}^{\infty} \frac{\beta_i}{2^i} < x\right\} = F_\alpha(x) = P\{\omega : \alpha < x\}, \tag{3.24}$$

$$\sum_{i=1}^{\infty} \frac{\beta_i(\omega)}{2} = \alpha(\omega) \tag{3.25}$$

є *p.p.BB* на [0,1] з ймовірністю 1.

Формула подання *p.p.BB* на [0,1] представлена у вигляді $\alpha(\omega) = \frac{\beta_1}{2} + \frac{\beta_2}{2^2} + ... + \frac{\beta_N}{2^N} + ..$ Тоді у двійковій системі можна записати:

$$\alpha(\omega) = 0, \beta_1\beta_2\beta_3...\beta_N ....; \beta_N = \beta_n \in \{0,1\}, n \ge 1. \tag{3.26}$$

Наслідок 3.2. Обираючи ряд (3.26) можемо значення *p.p.BB* на [0,1] записати у *k*-бітовому представленні числа на комп'ютері з деякою точністю.

Зауваження 3.1. Одержана послідовність значень $\left\{\alpha_n \equiv \sum_{i=1}^{n} \frac{1}{2^i}\beta_i\right\}$ перевіряється за допомогою критерію χ^2 на узгодженість з *p.p.BB* α(ω) на [0,1].

3.1.2. Моделювання на комп'ютері випадкової величини з довільним законом розподілу

Теорема 3.2. (перетворення М.В.Смірнова). На ймовірнісному просторі $(\Omega, \Im, P)$ задано

1) ξ=ξ(ω) – *BB* з законом розподілу $F_\xi(x)$;

2) α=α(ω) – *p.p.BB* α(ω) на [0,1], тобто щільність $p_\alpha = \begin{cases} 1, x \in [0,1] \\ 0, x \notin [0,1] \end{cases}$

3) ξ задовільняє рівняння Смірнова

$$\int_{-\infty}^{\xi} dF_\xi(x) = \int_{-\infty}^{\xi} p_\xi(x)dx = \alpha. \tag{3.27}$$

Смірнов Микола Васильович (17.10.1900–02.06.1966) – математик, фахівець з теорії ймовірності та математичної статистики.

Тоді закон розподілу для ξ можна одержати у формі закону розподілу для *p.p.BB* α на [0,1], тобто

$$F_\xi(x) = F_\alpha(F_\xi(x)). \tag{3.28}$$

Доведення. Умови 1), 2) вимагають знання таких фактів. За означенням *ВВ* $\alpha=\alpha(\omega)$, $\xi=\xi(\omega) \in \Im$ – *вимірні функції* тобто

$$\{\omega:\alpha(\omega)<x\} \in \Im;\ \{\omega:\xi(\omega)<x\} \in \Im,\ \forall x \in \mathbf{R}^1. \tag{3.29}$$

ВВ задана, якщо

$$\begin{aligned} F_a(x) &= P\{\omega : a < x\}, \\ F_\xi(x) &= P\{\omega : \xi < x\}, \end{aligned} \tag{3.30}$$

Умова 3) означає, що ліву частину (3.27) можна записати

$$\int_{-\infty}^{\xi} dF_\xi(x) = F_\xi(\xi) - F_\xi(-\infty) = F_\xi(\xi). \tag{3.31}$$

Має місце твердження з теорії вимірних функцій: якщо $\xi \in \Im$–*вимірною*, f – боролева функція, тоді $f(\xi) \in \Im$-*вимірною*, тобто є *ВВ*. Функція $F_\xi(x)$ – боролева функція, тобто $F_\xi(x)$ – $\Im$-*вимірна*. Тобто є випадковою величиною, тоді рівняння М.В.Смірнова (3.27) з врахуванням (3.31) прийме вигляд

$$F_\xi(\xi) = \alpha(\omega). \tag{3.32}$$

За означенням функції розподілу та її властивості, що вона є не спадною, маємо для $\xi(\omega)<x$ в значеннях функції розподілу

$$F_\xi(\xi(\omega)) \le F_\xi(x). \tag{3.33}$$

Зауважимо, що функція $F_\xi(x)$ перетворює нерівність $\xi<x$ в еквівалентну нерівність (3.33), тобто розв'язки відповідних нерівностей однакові

$$\{\omega : \xi < x\} = \{\omega : F_\xi(\xi) < F_\xi(x)\}. \tag{3.34}$$

Тоді рівність (3.31) перетвориться у ланцюжок рівностей з врахуванням властивостей ймовірності для рівних подій (3.34)

$$\begin{aligned} F_\xi(x) &= P\{\omega : \xi \le x\} = P\{\omega : F_\xi(\xi) \le F_\xi(x)\} = \\ &= P\left\{\omega : \int_{-\infty}^{\xi} dF_\xi(x) \le \int_{-\infty}^{x} dF_\xi(x)\right\} = P\{\omega : a(\omega) \le F_\xi(x)\} = \\ &= P\{\omega : \alpha(\omega) < F_\xi(x)\} = F_a(F_\xi(x)). \end{aligned} \tag{3.35}$$

Залишилося довести, що

$$\{\omega : \alpha(\omega) \le F_\xi(x)\} \in \Im \tag{3.36}$$

а на $\Im$ задана міра P.

Дійсно. $\forall x \in \mathbf{R}^1$ за означенням *ВВ*, як $\Im$-вимірної функції

$$\{\omega : \alpha < x\} \in \Im. \tag{3.37}$$

але $F_\xi(x) \in [0,1] \subset \mathbf{R}^1$, тобто множина $\{\omega : \alpha < F_\xi(x)\} \in \Im$ міститься в множинах (3.37), тобто існує ймовірність події (3.36). Теорема 3.2 доведена.

Наслідок 3.3. За допомогою комп'ютера одержуємо послідовність значень *p.p.BB* на [0,1] { α_1, α_2,..., α_N}. Підставляємо у (3.27) і одержимо N трансцидентних рівнянь у загальному випадку

$$\int_{-\infty}^{\xi} dF_\xi(x) = \alpha_i \ , \ i = \overline{1,N} \ / \tag{3.38}$$

N розв'язків цих рівнянь, якщо існують, дають послідовність $\{\xi_1,\xi_2,\ldots,\xi_N\}$.

Рівняння (3.38) еквівалентне $F_\xi(\xi) = \alpha_i$. Якщо існує F_ξ^{-1}, тоді $\xi_i = F_\xi^{-1}(\alpha_i)$. Метод статичного моделювання реалізовано, тобто $X = M\xi \approx \frac{1}{N}\sum_{i=1}^{N} \xi_i$ з певною точністю.

Лема 3.3. Якщо на $(\Omega, \Im, P)$ задана неперервна *p.p.BB* $\alpha=\alpha(\omega)$ на [0,1], тоді 1- α – *p.p.BB* на [0,1].

Доведення. Здійснимо перетворення

$$F_{1-\alpha}(x) \equiv P\{\omega : 1-\alpha \le x\} = P\{\omega : -\alpha \le x-1\} = P\{\omega : \alpha \ge 1-x\} = 1 - P\{\omega : \alpha < 1-x\} =$$

$$= \left\{ \begin{array}{l} 0, якщо 1-x<0, \\ 1-x, якщо 0 \le 1-x \le 1, \\ 1, якщо 1-x>1. \end{array} \right\} = \left\{ \begin{array}{l} 1, якщо x>1 \\ x, якщо 0 \le x \le 1 \\ 0, якщо x<0 \end{array} \right\}$$

Приклад 3.1. Довжина вільного шляху нейтрону в однорідному середовищі є *BB* $\xi=\xi(\omega)$ зі щільністю

$$p_\xi(x) = \left\{ \begin{array}{c} \sigma e^{-\sigma x}, x \ge 0, \sigma > 0 \\ 0, x < 0. \end{array} \right\}.$$

Змоделювати на комп'ютері ξ.

Розв'язок. Запишемо рівняння Смірнова (3.27) $\int_{-\infty}^{\xi} p_\xi(x)dx = \int_{-\infty}^{0} 0 \cdot dx + \int_{0}^{\xi} p_\xi(x)dx = \alpha$,

еквівалентне $\int_0^\xi \sigma e^{-\sigma x} dx = -e^{-\sigma x}\Big|_0^\xi = -e^{-\xi\sigma} + 1 = \alpha$; $e^{-\xi\sigma} = 1-\alpha$; $-\sigma\xi = \ln(1-\alpha)$. Тоді $\xi = -\frac{1}{\sigma}\ln(1-\alpha) = -\frac{1}{\sigma}\ln\alpha$. І остаточно $\xi = -\frac{1}{\sigma}\ln\alpha$. Значить, змоделювавши значення *p.p.BB* { $\alpha(\omega)$}: $\alpha_1,\alpha_2,\ldots,\alpha_N$, маємо відповідні значення $\xi_1,\xi_2,\ldots,\xi_N$.

Приклад 3.2. Одержати на комп'ютері значення нормально розподіленої *BB* ξ зі щільністю $p_\xi(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{x^2}{2}}$.

Розв'язок. Рівняння Смірнова (3.27) набуває вигляду

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\xi} e^{-\frac{x^2}{2}} dx = \alpha_i . \tag{3.39}$$

Відомо, що не існує первісної $F(x)$ в елементарних функціях для $f(x)=e^{-\frac{x^2}{2}}$, а розв'язати (3.39), тобто знайти ξ через α надзвичайно складно навіть за допомогою комп'ютера будь-якої потужності. Тому існують алгоритми моделювання нормально-розподіленої *ВВ* за допомогою *р.р.ВВ* α на [0,1].

3.1.3. Моделювання нормально-розподіленої випадкової величини ξ з $M\xi=0$, $D\xi=1$

На $(\Omega,\acute{p},P)$ беремо $\{a_i(\omega),\ \omega\in\Omega,\ i=1,2,\ldots,N\}$ - *р.р.ВВ* на [0,1], які є незалежні у сукупності: $M\{a_i(\omega)\}=\frac{1}{2}, D\{a_i(\omega)\}=\frac{1}{12}$. Розглянемо випадкову величину $a(\omega)=\sum_{i=1}^{N}a_i(\omega)$. Легко обчислити $M\{a(\omega)\}=\frac{N}{2}, D\{a(\omega)\}=\frac{N}{12}$.

Теорема 3.3. Нехай на $(\Omega,\acute{p},P)$ задана випадкова величина $\eta_N(\omega)=\sqrt{\frac{12}{N}}\sum_{i=1}^{N}\left(a_i(\omega)-\frac{1}{2}\right)$. Тоді $\lim_{N\to\infty}\eta_N(\omega)=\eta(\omega)$ є нормально розподіленою *ВВ* на $(\Omega,\acute{p},P)$, якщо $\{a_1,a_2,\ldots,a_N\}$ *р.р.ВВ* на [0,1] і незалежні у сукупності.

Доведення. Зауважимо, що $\eta_N(\omega)=f\{a_1(\omega),a_2(\omega),\ldots,a_N(\omega)\}$, де f – борелева функція, а значить $\lim_{N\to\infty}\eta_N(\omega)=\eta(\omega)$ - *ṕ- вимірною* тобто *ВВ* на $(\Omega,\acute{p},P)$. Якщо обчислити за означенням (3.41) $M\{\eta_N(\omega)\}=0, D\{\eta_N(\omega)\}=1$. Гранична теорема ймовірностей дає $\lim_{N\to\infty}\eta_N(\omega)=\eta(\omega)$. Вона є нормально розподіленою *ВВ* з $M\{\eta_N(\omega)\}=0, D\{\eta_N(\omega)\}=1.$, якщо $\{a_1,a_2,\ldots,a_N\}$ мають одинаковий закон розподілу.

Наслідок 3.4. Для N=12 маємо формулу Янсона (шведський математик) моделювання нормально розподіленої функції

$$\eta_{12}=\sum_{i=1}^{12}a_i(\omega)-6. \tag{3.40}$$

Формула (3.40) є основою функції та процесу у математичному забезпеченні комп'ютерів для моделювання нормально розподілених *ВВ*.

3.2. Обчислення інтегралів методом статистичного моделювання

3.2.1. Аналітичний метод обчислення однократного інтегралу

Нехай потрібно обчислити $I(f)=\int_0^1 f(x)dx$.

І. Будуємо на $(\Omega,\acute{p},P)$ *р.р.ВВ* беремо $\{a(\omega)\}$ на [0,1] з законом розподілу

$$p_a(x)=\left\{\begin{matrix}1, якщо\ x\in[0,1]\\ 0, якщо\ x\notin[0,1]\end{matrix}\right\}$$

Як відомо, нова $ВВ$ $\xi(\omega) = f(\alpha(\omega))$ має математичне сподівання

$$M\{\xi(\omega)\} = \int_{-\infty}^{\infty} p_0(f(x))dx = \int_{-\infty}^{0} 0 \cdot f(x)dx + \int_{x}^{1} 1 \cdot f(x)dx + \int_{1}^{\infty} 0 \cdot f(x)dx =$$

$$= \int_{0}^{1} f(x)dx = I(f). \quad (3.41)$$

II. За допомогою комп'ютера моделюємо $N>>10^5$ значень $p.p.ВВ$ $\{a(\omega)\}$: $a_1,a_2,\ldots,a_N$, а значить і N значень $\xi_i=f(a_i)$:

$$\xi_1,\xi_2,\ldots,\xi_N. \quad (3.42)$$

III. Статистичне середнє з (4.42) дорівнює $\overline{\xi} = \frac{1}{N}\sum_{i=1}^{N} f(a_i) \approx \int_{0}^{1} f(x)dx.$

3.2.2. Метод площ обчислення визначеного інтегралу

За означенням визначений інтеграл дорівнює площі криволінійної трапеції

$$S = I(f) = \int_{0}^{1} f(x)dx. \quad (3.43)$$

I. Розглянемо стохастичний експеримент: попадання точки у квадрат, який породжує ймовірнісний простір $(\Omega,\acute{p},P)$, A – подія. Що точка попадає під криву $\{f(x)\}$. Ймовірність події A: $P(A) = \frac{S}{S_a} = \int_{0}^{1} f(x)dx.$

II. Обчислення $P(A)$. Розглянемо дві $p.p.ВВ$ $\{a(\omega)\}$ та $\{\xi(\omega)\}$ на $[0,1]$. Змоделюємо (згенеруємо) N значень на комп'ютері: $a_1,a_2,\ldots,a_N$, $\xi_1,\xi_2,\ldots,\xi_N$.

Одержуємо N точок з координатами на комп'ютері: (a_i, ξ_i), $i=1,2,\ldots,N$.

Перевіряємо умову

$$\xi(\omega) \le f(a_i), \quad (3.44)$$

Якщо виконується умова (3.44), тобто $\xi_i \le f(a_i)$, то точка попадає в криволінійну трапецію. Нехай в N експериментах точка L разів попала в криволінійну трапецію. Тоді $P(A) = \lim_{N\to\infty} V_N(A) = \frac{L}{N}$, $\int_{0}^{1} f(x)dx \approx \frac{L}{N}$.

3.2.3. Обчислення кратних інтегралів

Для обчислення кратних інтегралів застосовуються МСМ. Нехай треба обчислити n-кратний інтеграл $I(f) = \iint_G \ldots \int f(x_1, x_2, \ldots, x_n)dx_1dx_2\ldots dx_n$, де

$$G = \begin{cases} (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n \mid 0 \le x_1 \le 1; 0 \le \varphi_i(x_1, \ldots, x_{i-1}) \le x_i \le \\ \le \psi_i(x_1, \ldots, x_{i-1}) \le 1; i = 1,2,\ldots,n; 0 \le f(x_1, x_2, \ldots, x_n) \le 1. \end{cases}$$

3.2.3.1. Метод *n*-кратних об'ємів

Беремо N незалежних у сукупності випадкових точок $P_k = (\xi_1^{(k)}, \xi_2^{(k)}, ..., \xi_n^{(k)}, \xi^{(k)}); k = \overline{1, N}$, де $\xi_1^{(k)}, \xi_2^{(k)}, ..., \xi_n^{(k)}, \xi^{(k)}$ – незалежні р.р.В.В на [0,1].

Стохастичний експеримент: кидаємо точку у n-вимірну трапецію. Подія A – попадання під графік $y = f(x_1, x_2, ..., x_n)$. Повторюємо N разів цей експеримент. Нехай L разів з N точка $(x_1, x_2, ..., x_n)$ попадає під цю криву, тобто виконуються умови

$$\left.\begin{array}{l} 0 \le \xi_1^{(k)} \le 1 \\ \varphi_2(\xi_1^{(k)}) \le \xi_2^{(k)} \le \psi_2(\xi_1^{(k)}) \\ \varphi_3(\xi_1^{(k)}, \xi_2^{(k)}) \le \xi_3^{(k)} \le \psi_3(\xi_1^{(k)}, \xi_2^{(k)}) \\ \\ \varphi_n(\xi_1^{(k)}, ..., \xi_{n-1}^{(k)}) \le \xi_n^{(k)} \le \psi_n(\xi_1^{(k)}, ..., \xi_{n-1}^{(k)}) \\ 0 \le \xi^{(k)} \le f(\xi_1^{(k)}, ..., \xi_n^{(k)}) \le 1 \end{array}\right\}$$

При великих N»10^5: $I(f) \approx \dfrac{L}{N} = P(A)$. Має місце твердження.

Теорема 3.4. Якщо задати число $\varepsilon > 0$ та $0 < \beta < 1$, то $P\left\{\omega : \left|I(f) - \dfrac{L}{N}\right| < \varepsilon\right\} = \beta$, де β - ймовірність надійності (вона близька до 1), число випробувань (кількість повторення експериментів) задовольняє нерівність

$$N \ge \frac{I\{1 - I\}}{\varepsilon^2} t_\beta^2, \tag{3.45}$$

t_β визначається за β з рівності

$$\Phi(t_\beta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t_\beta} e^{-\frac{x^2}{2}} dx = \frac{1 + \beta}{2}. \tag{3.46}$$

Зауважимо, що (3.45) доводиться за допомогою теореми Бернуллі та перетворення Чебишева.

Алгоритм знаходження N.

Задаємо $N_0 > 0$ та обчислюємо $I^{(0)} \approx \dfrac{L_0}{N_0}$. Далі знаходимо $N_1 = \dfrac{I^{(0)}(1 - I^{(0)})}{\varepsilon^2} t_\beta^2$.

Якщо $N_1 > N_0$, то збільшуємо число випробувань до $N_1 = 2N_0$. знаходимо $I^{(1)} \approx \dfrac{L_1}{N_1}$ та обчислюємо $N_2 = \dfrac{I^{(1)}(1 - I^{(1)})}{\varepsilon^2} t_\beta^2$ і т. д. до N_k.

3.2.3.2. Метод середнього значення (аналітичний метод).

Теорема 3.5. Нехай

1) $f(P) \in \mathbf{R}^1, P \in G \subset \mathbf{R}^n$;

2) $m < |f(P)| < M, P = (x_1, x_2, ..., x_n)$;

3) $P(\xi_1^{(k)}, ..., \xi_n^{(k)})$ – р.р.ВВ на $[0,1]\times[0,1]\times...\times[0.1]$.

Тоді $P\left\{\omega : \left| I(f) - \frac{1}{N}\sum_{k=1}^{N} f(P_k) \right| < \varepsilon \right\} = \beta$, якщо

$$N \geq \frac{\iint ... \int_G |f(P) - I(f)|^2 dP}{\varepsilon^2} t_\beta^2 \qquad (3.47)$$

t_β^2 визначається з (3.46); $P_k \in G$, звідки $f(P_k)$. Якщо $P_k \in \overline{G}$, то $f(P_k) = 0$.

Алгоритм реалізації на комп'ютері.

На практиці вибір N проводять таким чином, бо у формулі (3.47) чисельник невідомий.

1) Задаємо $\forall N_0$ та обчислюємо

$$\delta_{N_0} = \frac{1}{N_0}\sum_{i=1}^{N_0}[f(P_i)]^2 - I_{N_0}^2 \approx \frac{\iint ... \int_G |f(P) - I(f)|^2 dP}{\varepsilon^2} t_\beta^2.$$

2) Визначаємо $N_1 = \dfrac{\left(1 + 4\sqrt{\dfrac{2}{N_0}}\right)\delta_{N_0}}{\varepsilon^2} t_\beta^2$.

3) Якщо $N_1 > N_0$, то збільшуємо число випробувань до N_1, та обчислюємо

$\delta_{N_1} = \dfrac{1}{N_1}\sum_{i=1}^{N_1}[f(P_i)]^2 - I_{N_1}^2$ і $N_2 = \dfrac{\left(1 + 4\sqrt{\dfrac{2}{N_1}}\right)\delta_{N_0}}{\varepsilon^2} t_\beta^2$. Порівняємо $N_2 > N_1$ і т. д. до N_k.

3.3. Розв'язання систем лінійних алгебраїчних рівнянь

3.3.1. Метод Дж. фон Неймана-Улама

При викладенні методу статистичного моделювання для розв'язання систем лінійних алгебраїчних рівнянь будемо спиратися на підхід, вперше сформульований Дж. фон Нейманом та Уламом. Нехай система лінійних алгебраїчних рівнянь задана у вигляді

$$x = \alpha x + \beta, \qquad (3.48)$$

де $x \equiv (x_1, x_2, ...x_n)^T$ – вектор-стовпчик невідомих, β – вектор правих частин $\beta \equiv (\beta_1, \beta_2, ..., \beta_n)^T$ і $\alpha \equiv \|\alpha_{i,j}\|_l^n$ – матриця системи.

Припустимо, що найбільше за модулем власне значення A менше за одиницю, тобто збігається метод послідовних наближень $x^{(k)} = \alpha x^{(k-1)} + \beta$. Якщо покласти $x^{(0)} = \beta$, то

$$x(k) = (\alpha^k + \alpha^{k+1} + ... + \alpha + E)\beta \tag{3.49}$$

і точним розв'язком системи є $x = \lim_{k\to\infty}(E + \alpha + ... + \alpha^k)\beta = (E - \alpha)^{-1}\beta$.

Поставимо у відповідність системі (3.48) ланцюг Маркова (не обов'язково однорідний) з n станами. На початковий розподіл π та матрицю переходу $\hbar$ (k – номер переходу) накладемо додаткові умови, пов'язані з системою (3.48). які надалі сформулюємо, і розглянемо задачу про обчислення скалярного добутку (h,x), де h – заданий вектор. Очевидно, що задача про обчислення однієї з компонент розв'язку x є частинним випадком цієї задачі.

Якщо одночасно обчислюється деяка кількість m скалярних добутків (h_j, x) ($j = 1,...,m$), то розв'язок системи x може бути також одержаний при відповідному виборі m і h_i. Будемо пов'язувати системою (3.48) і вектором h деякий фіксований ланцюг Маркова з множини ланцюгів $\{\pi, \hbar_k\}$, $\pi = (\pi_1, ..., \pi_n)$, $\hbar_k = \|p_{i,j}^{(k)}\|_l^n$, $\sum_{i=1}^{n}\pi_i = 1$, $\pi_i \geq 0$, $\sum_{i=1}^{n} p_{i,j}^{(k)} = 1$, $p_{i,j}^{(k)} \geq 0$ ($i = \overline{1,n}; k = 1,2,...$), для яких виконуються умови

1. $\pi_i > 0$, якщо $h_i > 0$;
2. $p_{i,j}^{(k)} > 0$, якщо $\alpha_{i,j} > 0$ ($i, j = 1,2,...,n$).

Покладемо

$$q_i^{(0)} = \begin{cases} h_i/\pi_i, & \pi_i > 0, \\ 0, & \pi_i = 0, \end{cases} \quad q_{i,j}^{(k)} = \begin{cases} \alpha_{i,j}/p_{i,j}^{(k)}, & p_{i,j}^{(k)} > 0, \\ 0, & p_{i,j}^{(k)} = 0, \end{cases} \tag{3.50}$$

Візьмемо деяке ціле $N>0$ і будемо розглядати траєкторії ланцюга Маркова довжини N. Об'єкт, який описується ланцюгом Маркова, надалі будемо називати частинкою і вважати, що ця частинка змінює свої стани (рухається) у відповідності з траєкторією ланцюга. Частинка, яка рухається, має вагу Q_k, яка змінюється при русі частинки за траєкторією $i_0 \to i_1 \to ... \to i_N$ наступним чином. У початковий момент, коли частинка знаходиться у стані i_0, вона має вагу $Q_0 = q_{i_0}^0$; при переході зі стану i_0 до стану i_1, її вага дорівнює $Q_1 = q_{i_0}^0 q_{i_0,i_1}$, тощо тобто

$$Q_0 = q^0_{i_0};\ Q_m = Q_{m-1} q^0_{i_{m-1} i_m}. \tag{3.51}$$

Введемо випадкову величину $\xi_N \equiv \xi_N(\omega) \in \mathbf{R}^1$ на (Ω, Ξ, P), визначену на траєкторіях марковського ланцюга довжини N: $\xi_N = \sum_{m=0}^{N} Q_m \beta_{i_m}$. Кожній траєкторії довжини N ставиться у відповідність ймовірність $\pi_{i_0} p_{i_0 i_1} \ldots p_{i_{N-1} i_N}$, і математичне сподівання ВВ ξ_N, обчислюється наступним чином

$$M\xi_N = \sum_{i_0=1}^{n} \ldots \sum_{i_N=1}^{n} \xi_N \pi_{i_0} p_{i_0 i_1} \ldots p_{i_{N-1} i_N} =$$

$$= \sum_{i_0=1}^{n} \ldots \sum_{i_N=1}^{n} \xi_N \pi_{i_0} p_{i_0 i_1} \ldots p_{i_{N-1} i_N} \sum_{m=0}^{N} q^0_{i_0} q_{i_0 i_1} \ldots q_{i_{m-1} i_m} \beta_{i_m}.$$

Підставляючи замість $q^0_{i_0}$ та $q_{i_{m-1} i_m}$ їхні значення з (3.51) і змінюючи порядок сумування, одержимо $M\xi_N = \sum_{m=0}^{N} \sum_{i_0=1}^{n} \ldots \sum_{i_N=1}^{n} h_{i_0} \alpha_{i_0 i_1} \ldots \alpha_{i_{m-1} i_m} \beta_{i_m}$. За означенням добутку матриць $\alpha^2 = \left\| \alpha_{i_0 i_1} \right\|_1^n \times \left\| \alpha_{i_1 i_2} \right\|_1^n = \left\| \sum_{i_1=1}^{n} \alpha_{i_0 i_1} \alpha_{i_1 i_2} \right\|_{i_0 i_2=1}^{n}$ та $\alpha^m = \left\| \sum_{i_1=1}^{n} \ldots \sum_{i_{m-1}=1}^{n} \alpha_{i_0 i_1} \ldots \alpha_{i_{m-1} i_m} \right\|_{i_0 i_m=1}^{n}$.

Звідси випливає $M\xi = \left(h, \sum_{m=0}^{N} \alpha^m \beta \right)$. При $N \to \infty$ математичне сподівання $M\xi_N$ збігається до $(h, \bar{x})$. Такою є одна з можливих схем методу статистичного моделювання, пов'язаних із системами лінійних алгебраїчних рівнянь, яка полягає в обчисленні середніх значень випадкової величини ξ_N за траєкторіями ланцюга Маркова. Зробимо декілька зауважень, які стосуються наведеної схеми.

1. Як відомо, $(h, x) = (x^*, \beta)$, де x^* є розв'язком рівняння

$$x^* = \alpha^T x^* + h, \tag{3.52}$$

а α^T – матриця, транспонована до A. Дійсно, знайдемо h і β за допомогою x^* та x з рівностей (3.48) та (3.52) відповідно. Тоді одержимо, що $(x^* - \alpha^T x^*, x) = (x^*, x - \alpha x)$, що є рівносильним до відомої рівності $(\alpha^T x^*, x) = (x^*, \alpha x)$. Звідси випливає, що при обчисленні скалярного добутку (h, x) можна скористатися рівнянням $x^* = \alpha^T x^* + h$.

Марковські ланцюги, які відповідають обом рівнянням, можуть співпадати, якщо для них одночасно виконаються умови (3.50). Схема, пов'язана з рівнянням (3.52) (спряжена схема), є більш зручною у випадку, коли необхідно обчислювати одночасно декілька скалярних добутків (h_j, x), $j=1,\ldots,m$, оскільки в цьому випадку марковський ланцюг лишається незмінним для всіх j, а змінюється лише випадкова величина ξ_N і є можливою оцінка її середнього за одними й тими ж траєкторіями.

2. Для одержання розв'язку системи (3.48) досить взяти за h_j n векторів, у яких j-та компонента дорівнює одиниці, а решта – нулю.

Як відомо, r-й стовпчик матриці $(E-\alpha)^{-1}$ може бути одержаний як розв'язок системи $x=\alpha x+h_r$, де h_r – вектор-стовпчик з одиничною r-тою компонентою і нульовими іншими компонентами. тому при $\beta=h_r$. та $h=h_j$ схема, пов'язана з рівнянням (3.48), дає як середнє значення ВВ ξ_N елемент матриці $(E-\alpha)^{-1}$ з номером стовпчика r та номером рядка j, а схема, пов'язана з рівнянням (3.52), – елемент тієї ж матриці з номером рядка r і стовпчика j. Різниця $M\xi_{N-1}-M\xi_N$ дорівнює $(h,\alpha^{N+1}\beta)$ для першої схеми і $(\beta,(\alpha^T)^{N+1}h)$ – для другої. Оцінки цих скалярних добутків можуть бути використані для визначення власних чисел матриці A, коефіцієнтів характеристичного многочлена тощо у відповідності з відомими обчислювальними схемами лінійної алгебри. Однак у цьому випадку необхідно додатково дослідити питання оцінки статистичної похибки.

3. З рівності $\lim\limits_{N\to\infty} M\xi_N=(h,x)$ випливає, що неусувність оцінки ξ_N має місце для траєкторій марківського ланцюга нескінченної довжини. Дійсно, зі збіжності (мається на увазі збіжність за деякою нормою) ряду $\sum\limits_{m=0}^{\infty}\alpha^m$ випливає, що, починаючи з деякого номера $N=N_0$, внесок додаткових доданків у $M\xi_{N_0}$, які з'являються в результаті збільшення N, становиться дуже малим. Практично N_0 – номер ланки траєкторії, на якій її необхідно переривати. Він може обиратися в результаті оцінки внеску наступних членів ряду (3.49). При цьому необхідна певна обережність.

Якщо $\alpha_{i,j}=0$ хоча б для однієї пари номерів i,j, то, оскільки $q_{i,j}=0$, вага частинки після переходу її зі стану з номером i до стану з номером j перетворюється на нуль. Очевидно, що всі її внески при наступних переходах будуть нульовими (це випливає з (3.50)). Можна вважати що частинка при переході зі стану i до стану j поглинається і відбувається обрив її траєкторії. У цьому випадку можна не робити штучне переривання траєкторії.

Таким чином, інтегрування за мірою, яка визначається обраним нами ланцюгом Маркова, принципово дозволяє знаходити розв'язок системи лінійних алгебраїчних рівнянь, і природно приходимо до задачі обчислення інтеграла, яка обговорювалася раніше. При цьому вибір різних ланцюгів Маркова відповідає різним варіантам методу істотної вибірки. Разом з тим спеціальна структура міри, що пов'язана з ланцюгом Маркова, дозволяє одержати і деякі спеціальні результати стосовно дисперсії оцінок.

Зупинимося далі коротко на застосуваннях методу статистичного моделювання до розв'язання задач лінійної алгебри. Практично метод статистичного моделювання у вищеописаній формі може конкурувати з класичними обчислювальними методами при виконанні наступних умов.

1. Порядок n матриці α дуже великий, елементи її можуть буї и досить просто

обчислені й (необхідно оцінити розв'язок системи або елементи оберненої матриці з невеликою точністю. Величина n практично залежить від характеристик комп'ютера.

2. Використання методу статистичного моделювання є особливо зручним у випадку, коли необхідно оцінити один або невелику кількість скалярних добутків вигляду $(h, \bar{x})$ з невеликою кількістю значущих цифр, бо при цьому немає необхідності у попередньому знаходженні розв'язку $\bar{x}$. Прикладом системи рівнянь, де процедура методу статистичного моделювання є особливо простою, є різницевий аналог крайової задачі для рівняння Пуассона.

3.3.2. Побудова імітаційної моделі розв'язання системи лінійних алгебраїчних рівнянь (СЛАР)

Нехай нам необхідно знайти розв'язок наступної СЛАР:

$$Ax = b, \tag{3.53}$$

де $A \equiv \{a_{ij}, i, j = \overline{1,n}\} \subset \mathbf{R}^1$; $\det A \neq 0$; $x \equiv (x_1,...,x_n)^T \in \mathbf{R}^n$; $b \equiv (b_1,...,b_n) \in \mathbf{R}^n$.

Спочатку наведемо детермінований випадок побудови розв'язку (3.53). Запишемо (3.53) у вигляді:

$$x = \alpha x + \beta, \tag{3.54}$$

де $\alpha = E - A$, $\beta = b$.

Якщо $\|\alpha\| < 1$, тоді послідовність $\{x_{(n)} = \alpha x^{(n-1)} + \beta; n \geq 1; x_0 = \beta\}$ збігається до x^* при $n \to \infty$, де x^* – розв'язок (3.53). Якщо переписати (3.54) у вигляді $(E - \alpha)x = \beta$ та скористатися відомою формулою (при існуванні $(E-\alpha)^{-1}$): $(E-\alpha)^{-1} = E + \alpha + \alpha^2 + ... + \alpha^n + ...$, то можна записати розв'язок (3.53) у вигляді наступного ряду:

$$x = (E-\alpha)^{-1}\beta = \beta + \alpha\beta + \alpha^2\beta + ... + \alpha^n\beta + ... \tag{3.55}$$

Якщо вираз (3.55) записати покоординатно, то одержимо:

$$x_m = \beta_m + \sum_{i=1}^{n} \alpha_{m,i}\beta_i + \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} \alpha_{m,i_1}\alpha_{i_1,i_2}\beta_{i_2} + ... + \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}...\sum_{i_k=1}^{n} \alpha_{m,i_1}\alpha_{i_1,i_2}...\alpha_{i_{k-1},i_k}\beta_{i_k} + ... \tag{3.56}$$

Реалізовуючи одержану формулу на комп'ютері, досягти бажаної точності при великих значеннях n (n>10^2) можна за значний проміжок часу. Можна також побудувати ітераційний процес розв'язання (3.53) таким чином:

$x^{(0)} = \beta$; $x^{(1)} = \alpha x^{(0)} + \beta$; $x^{(2)} = \alpha x^{(1)} + \beta$; ...; $x^{(k)} = \alpha x^{(k-1)} + \beta = \beta + \alpha\beta + ... + \alpha^k\beta$.

Однак у цьому випадку для обчислення розв'язку необхідно на кожному кроці використовувати всі компоненти попередніх ітерацій. Тому при великих n більш придатним для використання є метод статистичного моделювання розв'язання СЛАР.

Побудуємо ймовірнісну модель, яка дала би можливість обчислювати координати x_m, $m=\overline{1,n}$. Будемо вважати, що: 1) $\alpha_{ij}>0$: 2) $\sum_{i=1}^{n}\alpha_{m,i}=1$. Проведемо наступний стохастичний експеримент; будемо кидати точку на відрізок [0,1]. При цьому $\Omega=[0,1]$; $\Im$ – мінімальна σ-алгебра борелевих множин; $P\{\omega:[\alpha,\beta)\subset[0,1]\}=|\alpha-\beta|$. Розглянемо на $(\Omega,\Im,P)$ рівномірно розподілену неперервну випадкову величину (р.р.НВВ) $\xi\equiv\xi(\omega)$ на [o,1], тобто

$$p_{\xi}(x)=\begin{cases}1; & x\in[0,1],\\ 0; & x\notin[0,1].\end{cases}$$

Вибираємо n реалізацій р.р.НВВ $\xi:\xi_1,...,\xi_n$. Відрізок [0,1] можна записати у вигляді: $[0,1]=\bigcup_{i=1}^{n}J_i=\bigcup_{i=1}^{n}[\alpha_{i,m},\alpha_{m,i-1})$, де $\alpha_{m,0}=0$; $\alpha_{m,n}=1$. Будемо вважати, що

$$P\{\omega:\xi=\xi_m\in J_k\}=\alpha_{m,k},\ k,m=\overline{1,n}.$$

Розглянемо на $(\Omega,\Im,P)$ нову дискретну випадкову величину (ДВВ) $Y_m\equiv Y_m(\omega)$ таку, що $P\{\omega:Y_m=\beta_i\}=\alpha_{m,i}$, $i,m=\overline{1,n}$. Тоді за означенням математичного сподівання одержимо $MY_m=\sum_{i=1}^{n}\alpha_{m,i}\beta_i$. Отже, MY_m дорівнює другому доданку формули (3.56).

Аналогічно введемо нову ДВВ $Z_m\equiv Z_m(\omega)$ таким чином, щоб $MZ_m=\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\alpha_{m,i_1}\alpha_{i_1,i_2}\beta_{i_2}$. Для цього розглянемо подію $\{\omega:\xi_m\in J_{i_1}\}\cap\{\omega:\xi_{i_1}\in J_{i_2}\}$ та реалізацію ДВВ $Z_m(\omega)=\beta_{i_2}$. Для неї одержуємо

$$P\{\omega:Z_m=\beta_{i_2}\}=P\left\{\omega:\bigcup_{i_1}\{\xi_m\in J_{i_1}\}\cap\{\xi_{i_1}\in J_{i_2}\}\right\}=$$

$$=\sum_{i_1}P\{\omega:\{\xi_m\in J_{i_1}\}\cap\{\xi_{i_1}\in J_{i_2}\}\}=\sum_{i_1}\alpha_{m,i_1}\alpha_{i_1,i_2}.$$

За означенням математичного сподівання отримаємо $MZ_m=\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\alpha_{m,i_1}\alpha_{i_1,i_2}\beta_{i_2}$. Отже, MZ_m дорівнює третьому доданку формули (3.56). Аналогічно будуємо наступні ДВВ. Отже, можна побудувати ДВВ $X_m\equiv X_m(\omega)$, яка визначається формулою $X_m=\underbrace{\beta_m+Y_m+Z_m+...+V_m}_{k}$. При цьому $MX_m\approx x_m$. Повторимо вищенаведений експеримент побудови X_m $N>>10^5$ разів. Одержимо послідовність реалізацій ДВВ X_m: $X_m^{(1)},X_m^{(2)},...,X_m^{(N)}$. Тоді за теоремою Чебишева одержимо, що

$\overline{X}_m = \frac{1}{N}\sum_{i=1}^{N} X_m^{(i)} \approx MX_m \approx x_m$, оскільки $P\left\{\omega : \left|\overline{X}_m - x_m\right| \le \sigma\sqrt{\frac{1}{\varepsilon N}}\right\} \ge 1-\varepsilon$, де $\sigma^2 = DX_m$, $\varepsilon > 0$ – точність.

3.3.3. Випадкове блукання частинки з поглинаючим екраном

Розглянемо стохастичний експеримент випадкового блукання частинки, яка має скінченну кількість станів: $s_1, s_2, ..., s_n, s_{n+1}, t_1, t_2, ..., t_n, t_{n+1}$. Перехід зі стану s_i у s_j позначимо $(s_i \to s_j)$ і будемо вважати це елементарною подією. Створимо математичну модель стохастичного експерименту $(\Omega, \Im, P)$,

$$\Omega = \left\{\omega_{ij} \equiv (s_i \to s_j); i = \overline{1, n+1}, \overline{j = 1, n}\right\};$$

$$\Im(\mathbf{R}) = \left\{\Omega, \varnothing, \bigcup_{i,j \le n+1}(s_i \to s_j), \overline{\bigcup_{i,j \le n+1}(s_i \to s_j)}\right\}.$$

Нехай визначено також $1 \ge p_{ij} \equiv P\{\omega_{ij} \equiv (s_i \to s_j)\} \ge 0$. Стан $s_{n+1} \equiv \Gamma$, який відповідає повній зупинці частинки, назвемо поглинаючим екраном $p_{n+1,j} \equiv P\{\omega_{ij} \equiv (s_{n+1} \to s_j), j = \overline{1, n}\} = 0$. Процес блукання частинки припиняється, коли вона вперше попаде на екран. Описана схема станів частинки $\{s_j, j = \overline{1, n+1}\}$ з p_{ij} називається дискретним ланцюгом Маркова. Матрицею переходів ймовірностей, або перехідною матрицею ймовірностей, назвемо матрицю

$$\Pi = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} & p_{1,n+1} \\ p_{21} & p_{22} & \cdots & p_{2n} & p_{2,n+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} & p_{n,n+1} \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Зупиненим дискретним ланцюгом Маркова назвемо дискретний ланцюг Маркова, який зв'язаний з повною зупинкою на екрані Γ, тобто $p_{n+1,j} = 0$, $j = \overline{1,n}$. Нехай маємо зупинений дискретний ланцюг Маркова з матрицею Π, $s_i \ne \Gamma (i < n+1)$, а випадкове блукання частинки починається зі стану $s_i = s_{i_0}$, але після проміжних станів $s_{i_1}, s_{i_2}, ..., s_{i_m}$ закінчуються на екрані $s_{i_{m+1}} \equiv \Gamma$. Сукупність переходів

$$\left\{(s_{i_0} \to s_{i_1}), (s_{i_1} \to s_{i_2}), ..., (s_{i_m} \to s_{i_{m+1}} \equiv \Gamma)\right\} \equiv \left\{s_{i_0}, s_{i_1}, s_{i_2}, ..., s_{i_m}, s_{i_{m+1}} \equiv \Gamma)\right\} \equiv T_i$$

назвемо траєкторією частинки. Вона має випадкову структуру.

Розглянемо ДВВ $X_i \equiv X(T_i)$ – функціонал траєкторії. Поставимо у відповідність X_i її значення $\xi(X_i)$:

$$\xi(X_i) = \beta_{i_0} + v_{i_0 i_1}\beta_{i_1} + v_{i_0 i_1}v_{i_1 i_2}\beta_{i_2} + ... + v_{i_0 i_1}...v_{i_{m-1}i_m}\beta_{i_m} + v_{i_0 i_1}...v_{i_m i_{m+1}}\beta_{i_{m+1}}, \quad \text{де} \quad \beta_i,$$

$j = i_0, i_1, ..., i_m$ – відповідні вільні члени системи (3.54). Значення v_{ij} знаходимо з рівностей $\alpha_{ij} = p_{ij} v_{ij}$,

$$p_{ij} = \begin{cases} \alpha_{ij}, & \alpha_{ij} \geq 0, \\ -\alpha_{ij}, & \alpha_{ij} < 0, \end{cases} \qquad v_{ij} = \begin{cases} 1, & \alpha_{ij} \geq 0, \\ -1, & \alpha_{ij} < 0, \end{cases}$$

$$P(X_i) = P\{(s_{i_0} \to s_{i_1}) \cap ... \cap (s_{i_m} \to s_{i_{m+1}})\} = p_{i_0 i_1} p_{i_1 i_2} ... p_{i_m i_{m+1}}$$

Лема 3.4. Якщо дискретний ланцюг Маркова зупинений, то випадково блукаюча частинка перейде на екран за скінченну кількість переходів з ймовірністю I.

Доведення. Позначимо: $q_i(t) \equiv P\left\{\left(s_i \xrightarrow{t<\infty} \Gamma\right)\right\}$, причому послідовність $\{q_i(t)\}$ монотонно зростає. З означення зупиненого дискретного ланцюга Маркова випливає, що існує $t_0 < \infty$ такс, що для довільного i маємо $0 < q_i(t_0) \leq 1$, тобто з ненульовою ймовірністю за $t_0 < \infty$ частинка зі стану s_i перейде до $s_{n+1} \equiv \Gamma$. Нехай $q \equiv \min_i q_i(t_0) > 0$, тоді

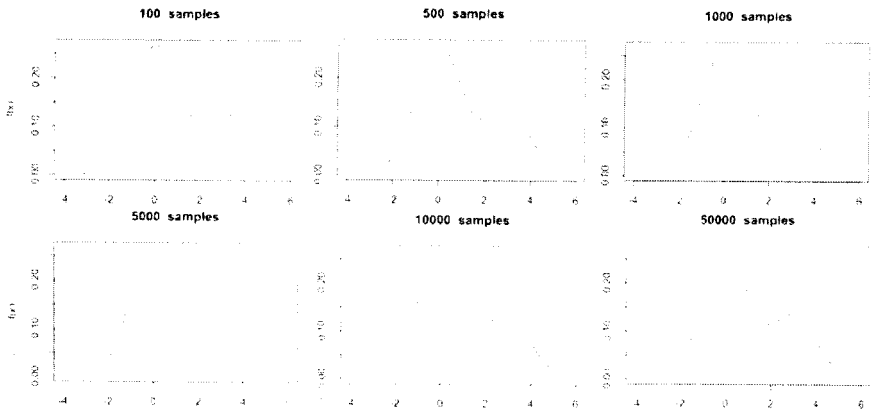
$$P\left\{\left(s_i \overset{t_0<\infty}{\not\to} \Gamma\right)\right\} = 1 - P\left\{\left(s_i \xrightarrow{t_0<\infty} \Gamma\right)\right\} \leq 1 - q.$$

Однак p_{ij} не залежить від минулого, тому

$$P\left\{\left(s_i \overset{vt_0<\infty}{\not\to} \Gamma\right)\right\} = P\left\{\bigcap_v \left(s_i \overset{t_0<\infty}{\not\to} \Gamma\right)\right\} = \left[P\left(s_i \overset{t_0<\infty}{\not\to} \Gamma\right)\right]^v \leq (1-q)^v \underset{v\to\infty}{\to} 0.$$

Тоді $P\left\{\left(s_i \xrightarrow{vt_0<\infty} \Gamma\right)\right\} = 1 - P\left\{\left(s_i \overset{vt_0<\infty}{\not\to} \Gamma\right)\right\} \underset{v\to\infty}{\to} 1.$

Приклад імітаційного моделювання на основі ланцюгів Маркова:



3.3.4. Теорема про побудову $X_i(\omega)$

Теорема 3.6. Нехай задано СЛАР, причому $\|\alpha\| < 1$. Тоді існують дискретні випадкові величини $X_i \equiv X_i(\omega) \in R^1$, для яких

$$MX_i = x_i;\ i = 1,\ldots,n. \tag{3.57}$$

Доведення. Довільну траєкторію T_i, яка починається зі стану s_i, в залежності від першого кроку можна розбити на $(n+1)$ траєкторій:

$$\begin{aligned}
T_{i1} &= \left\{s_i, s_1, s_{i2}, \ldots \Gamma = s_{i_{m-1}}\right\},\\
T_{i2} &= \left\{s_i, s_2, s_{i2}, \ldots, \Gamma = s_{i_{m-1}}\right\},\\
&\ldots\ldots\ldots

$$MX_i = \sum_{T_i} \xi(X_i) P\{X_i\} = \sum_{j=1}^{n-1} \sum_{T_{ij}} \xi(X_{ij}) P(X_{ij}) =$$

$$= \sum_{j=1}^{n} \sum_{T_j} \left[\beta_i + \upsilon_{ij} \xi(X_j) \right] p_{ij} P(X_j) + \beta_i p_{i,n+1} =$$

$$= \sum_{j=1}^{n} p_{ij} \upsilon_{ij} \sum_{T_j} \xi(X_j) P\{X_j\} + \beta_i \left[\sum_{j=1}^{n} p_{ij} \sum_{T_j} P(X_j) + p_{i,n+1} \right] =$$

$$= \sum_{j=1}^{n} \alpha_{ij} MX_j + \beta_i,$$

$$MX_i = \sum_{j=1}^{n} \alpha_{ij} MX_j + \beta_i. \tag{3.65}$$

Але рівняння (3.57) для i-ої компоненти запишеться у вигляді

$$x_i = \sum_{j=1}^{n} \alpha_{ij} x_j + \beta_i, i = 1, 2, ..., n. \tag{3.66}$$

Порівнюючи (3.65) і (3.66), одержимо:

$$x_i = MX_i. \tag{3.67}$$

3.3.5. Обчислення оберненої матриці методом статистичного моделювання

Теорема 3.7. Нехай $A=E-\alpha$ та $\|\alpha\| < 1$. Тоді існують дискретні випадкові величини $X_{ij} \equiv X_{ij}(\omega)$ такі, що

$$A^{-1} = \{x_{ij}\}_{i,j=1}^{n} = \{MX_{ij}\}_{i,j=1}^{n}. \tag{3.68}$$

Доведення. Домножимо вираз

$$(E-\alpha)^{-1} = \{x_{ij}\}_{i,j=1}^{n}. \tag{3.69}$$

зліва на (E-α). Одержимо

$$(E-\alpha)\{x_{ij}\}_{i,j=1}^{n} = E. \tag{3.70}$$

$$\sum_{k=1}^{n} (\delta_{ik} - \alpha_{ik}) x_{ik} = \delta_{ij}, 1 \le i, j \le n. \tag{3.71}$$

Звідки отримаємо

$$x_{ij} = \sum_{k=1}^{n} (\alpha_{ik} x_{ik}) + \delta_{ij}, 1 \le i, j \le n. \tag{3.72}$$

тобто елементи кожного стовпчика $(x_{1j}, x_{2j}, ..., x_{nj})^T \quad j = 1, ..., n$ матриці A^{-1} визначаються з відповідної СЛАР. З теореми 3.6 одержимо, що при фіксованому j частинка перейде зі стану $s_i = s_{i0}$ і отримаємо *ДВВ* X_{ij} зі значеннями

$$\xi(X_{ij}) = \delta_{i0j} + v_{i0i1}\delta_{i1j} + ... + v_{i0i1}v_{i1i2}...v_{i_{m-1}i_m}\delta_{i_m j}, \tag{3.73}$$

$$T_{ij} = \{s_{i_0}, s_{i_1}, ..., s_{i_m}, \Gamma \equiv s_{i_{m+1}}\}, \tag{3.74}$$

де v_{ij} визначається з $\alpha_{ij} = p_{ij}v_{ij}, p_{ij} = P\{(s_i \to s_j)\}$. Тоді $MX_{ij} = x_{ij}, i, j = 1, 2, ..., n.$

3.3.6. Алгоритм методу статистичного моделювання розв'язку СЛАР

На основі вищенаведених теоретичних відомостей сформулюємо алгоритм методу статистичного моделювання розв'язку СЛАР:

I. Реалізація алгоритму випадкового шляху частинки на комп'ютері

1) $p_{ij} = \dfrac{t_{ij}}{10^q}, i = 1, ..., n, j = 1, ..., n+1, \sum_{j=1}^{n+1} t_{ij} = 10^q, q$ – число бітів подання слова у пам'яті комп'ютера, де зберігається число t_{ij}.

2) Розглянемо частинку, яка має початковий стан $s_i, i = 1, ..., n$. Реалізуємо рівномірно розподілену *ВВ* $\gamma \equiv \gamma(\omega)$ на $[0,1]$ за допомогою стандартної підпрограми генерації псевдовипадкових чисел та відповідних алгоритмів. Це можливо, оскільки

$$\sum_{j=1}^{n+1} p_{ij} = 1; \; i = 1, 2, ..., n. \tag{3.75}$$

Якщо $0 \le \gamma^* < p_{i1}$, то це відповідає переходу $(s_i \to s_1)$; якщо $p_{i1} \le \gamma^* < p_{i1} + p_{i2}$, то це відповідає стану $(s_i \to s_2)$; якщо $\sum_{j=1}^{n-1} p_{ij} \le \gamma^* < \sum_{j=1}^{n} p_{ij}$, то це відповідає стану $(s_i \to s_n)$; якщо $\sum_{j=1}^{n} p_{ij} \le \gamma^* < \sum_{j=1}^{n+1} p_{ij} = 1$, то це відповідає стану $(s_i \to \Gamma = s_{n+1})$; На підставі цього зрозуміло, що кількість випадків, які сприяють переходам

$$s_i \to s_1, s_i \to s_2, ..., s_i \to s_n, s_i \to s_{n+1} = \Gamma, \tag{3.76}$$

пропорційна числам $t_{i1}, t_{i2}, ..., t_{in}, t_{in+1}$, причому всіх випадків 10^q. Тоді за класичним визначенням ймовірності

$$p_{ij} \equiv p(s_i \to s_j) := \frac{t_{ij}}{10^q} \; i = 1, ..., n, \; j = 1, ..., n+1. \tag{3.77}$$

Вибираючи нові значення γ, отримаємо випадкові траєкторії шляху частинки з початковим станом s_i та ймовірностями p_{ij}. Відомо, що частинка потрапить на

поглинаючий екран за скінченне число переходів до стану $\Gamma \equiv s_{n+1}$. Будуємо траєкторію T_j і значення випадкової величини $x_i \equiv X(T_j)$.

II. Алгоритм одержання x_i, $i = 1,\ldots,n$.

1) Отримаємо N випадкових шляхів частинок з випадковими траєкторіями $T_i^{(k)}$, $k = 1,2,\ldots,N$ $(n \geq 10^5)$ та обчислюємо

$$\xi(X_i^{(k)}) = \beta_i^{(k)} + v_{i,i_1}^{(k)}\beta_{i_1}^{(k)} + \ldots + v_{i,i_1}^{(k)}v_{i_1,i_2}^{(k)}\ldots v_{i_{m-1},i_m}^{(k)}\beta_{i_m}^{(k)},\ k = 1,2,\ldots,N. \tag{3.78}$$

2) Візьмемо наближено

$$x_i \approx \frac{1}{N}\sum_{k=1}^{N}\xi(X_i^{(k)}). \tag{3.79}$$

3) Кількість N вибираємо, де $N \leq \dfrac{9\sigma^2}{\varepsilon^2}$, – точність, з якою необхідно знайти розв'язок СЛАР; $\sigma^2 \approx \dfrac{1}{N-1}\sum_{i=1}^{N}\left(\xi_i - \bar{\xi}\right)^2$.

3.4. Розв'язок першої крайової задачі для рівняння $\Delta u = f$

Розглянемо дуже простий випадок першої крайової задачі для рівняння

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x,y). \tag{3.80}$$

в одиничному квадраті. Без обмеження загальності вважають, що крайова умова задається у вигляді $u|_\Gamma = 0$, де Γ – межа квадрата. Побудуємо у квадраті сітку з кроком $l = \dfrac{1}{L}$. Замінюючи наближено частинні похідні в рівнянні (3.80) іншими з врахуванням крайових умов, одержимо систему $(L\text{-}1)^2$ лінійних алгебраїчних рівнянь вигляду

$$u_{i,j} = a_{i,j}u_{i-1,j} + b_{i,j}u_{i,j-1} + c_{i,j}u_{i+1,j} + d_{i,j}u_{i,j+1} - l^2 f_{i,j}, \tag{3.81}$$

де $u_{i,j}$ – наближене значення u в точці сітки,

$$f_{i,j} \equiv f(il, jl),\ a_{i,j} = \begin{cases} 1/4, & i \neq 1, \\ 0, & i = 1, \end{cases}\ b_{i,j} = \begin{cases} 1/4, & j \neq 1, \\ 0, & j = 1, \end{cases}$$

$$c_{i,j} = \begin{cases} 1/4, & i \neq L-1, \\ 0, & i = L-1, \end{cases}\ d_{i,j} = \begin{cases} 1/4, & j \neq L-1, \\ 0, & j = L-1. \end{cases} \tag{3.82}$$

Таблиця 3.1

Вигляд матриці A (L=4)

c	1	2	3	4	5	6	7	8	9	10(v)
1		1/4		1/4						
2	1/4		1/4		1/4					
3		1/4				1/4				
4	1/4				1/4		1/4			

Продовження таблиці 3.1

5		1/4		1/4		1/4		1/4		
6			1/4		1/4				1/4	
7				1/4				1/4		
8					1/4		1/4		1/4	
9							1/4		1/4	
10										1

У тих випадках, коли коефіцієнти дорівнюють нулеві, нам байдуже до яких невідомих вони відносяться. Можна вважати, що вони відносяться до деякого одного фіктивного невідомого v, що відповідає крайовим точкам сітки. При суцільній нумерації невідомих за правилом $u_k \equiv u_{i,j}$, де $k=j+(L-1)(i-1)$, матриця A має вигляд, поданий у табл. 3.1. Матриці A природньо співставити стохастичну матрицю P (див. табл. 3.2). Величини $q_{i,j}$ в цьому випадку дорівнюють одиниці (крім випадку переходу до стану 10, коли вони дорівнюють нулеві). Конструкція вектора h залежить від того, яку величину потрібно визначити в результаті розв'язку задачі.

Використовуючи спряжену схему для знаходження розв'язку системи рівнянь (3.80) і обираючи за π вектор, всі компоненти якого дорівнюють $\frac{1}{(L-1)^2}$, приходимо до наступного алгоритму. Частинка народжується з однаковою ймовірністю у будь-якій внутрішній точці сітки і з однаковою ймовірністю переходить у будь-яку сусідню. При переході до граничної точки траєкторія переривається. Для одержання розв'язку відводиться $(L\text{-}1)^2$ комірок пам'яті. Якщо частинка народилась в точці з номером k_0, то при її переході до точки з номером k (народження точок також вважається переходом) у k-ту комірку пам'яті додається величина $l^2 f_{k0}$.

Таблиця 3.2

Вигляд матриці *P*

c	1	2	3	4	5	6	7	8	9	10(v)
1		1/4		1/4						1/2
2	1/4		1/4		1/4					1/4
3		1/4				1/4				1/2
4	1/4				1/4		1/4			1/4
5		1/4		1/4		1/4		1/4		0
6			1/4		1/4				1/4	1/4
7				1/4				1/4		1/2
8					1/4		1/4		1/4	1/4
9							1/4		1/4	1/2
10										1

Досліджується M незалежних траєкторій частинки. Середнє траєкторій наявних елементів у кожній з комірок, що відведені для компонент вектора розв'язків системи, служить оцінкою відповідної компоненти. Ця проста схема була спеціально представлена з загальної точки зору. Можна без особливих складностей розглянути також схеми розв'язків різницевих рівнянь, які відповідають іншим задачам для рівнянь у частинних похідних. У загальному випадку, зрозуміло, частинка може переходити не тільки до сусідніх точок сітки, обчислюється вага і може змінювати знак при переходах частинки. З вищевикладеного випливає, що при конструюванні схеми методу статистичного моделювання є досить багато довільних припущень. Зокрема, будемо виходити з довільного вибору марковського ланцюга $(\pi.P)$. Очевидно, що можна вибрати цей ланцюг непридатним для оцінки тієї величини, яка нас цікавить. Якщо нас

цікавить компонента розв'язку системи x_{k_0} з номером k_0, а ймовірності переходу до стану k_0 вибрані дуже малими, то частинка зрідка буде переходити у точку з номером k_0, але при кожному такому переході внесок, який задається її вагою, може бути дуже великим. Це приводить до того, що дисперсія Dx_{k_0} оцінки x_{k_0} буде великою. Міркування такого роду можуть бути корисними при виборі ланцюга (π,P).

Якщо мова йде про знаходження вектора розв'язків системи лінійних алгебраїчних рівнянь, то можуть бути корисними наступні результати.

Припустимо, що матриця α системи $x=\alpha x+\beta$ симетричною. При використанні прямої схеми розв'язок можна одержати, поставивши у відповідність системі для n ланцюгів Маркова $\{\pi_j,P\}$, $j=1,2,\ldots,n$, де π_j – вектор, який має всі нульові компоненти, за винятком компоненти з номером j, яка дорівнює одиниці, а P не залежить від j.

Однак, більш економним є не конструювання незалежних траєкторій, які відповідають кожному з ланцюгів при оцінці кожного невідомого, а використання однієї, вважаючи що частинка з'явилася у стані з номером i, коли вона переходить до цього стану на якому-небудь етапі блукань. Це можна зробити завдяки спеціальному вигляду початкових розподілів π_j. Якщо при заданому способі обчислень $\xi_N^{(i)}$ означає оцінку x_i, i-ої компоненти розв'язку системи (3.81), що відповідає траєкторіям ланцюга довжини N, то має зміст співвідношення

$$\xi_N^{(i)} = \beta_i + (\alpha_{i,j} / p_{i,j})\xi_{N-1}^{(j)}, \tag{3.83}$$

коли остання ланка траєкторії – перехід зі стану з номером j до стану з номером i. Покладемо $\xi^{(i)} = \lim\limits_{N\to\infty} \xi_N^{(i)}$.

Лема 3.5. Якщо норма матриці $\left\| \dfrac{\alpha_{i,j}^2}{p_{i,j}} \right\|^n$ менша за одиницю, то дисперсії $D\xi^{(i)}$, $i=1,\ldots,n$ скінченні й задовольняють умову:

$$D\xi^{(i)} = \sum_{j=1}^{n} \frac{\alpha_{i,j}^2}{p_{i,j}} (D\xi^{(j)} + x_j^2) - \sum_{j=1}^{n} a_{i,j} x_j^2. \tag{3.84}$$

Доведення. За формулою обчислення дисперсії

$$D\xi^{(i)} = M(\xi^{(i)})^2 - x_i^2 = \lim_{N\to\infty} (\xi_N^{(i)})^2 - x_i^2. \tag{3.85}$$

Згідно з (3.83), одержимо

$$D\xi^{(i)} = \lim_{N\to\infty} \left\{ \beta_i^2 + 2\beta_i M \left\{ \frac{\alpha_{i,j}}{p_{i,j}} \right\} \xi_{N-1}^{(j)} \right\} +$$

$$+ M \left(\frac{\alpha_{i,j}}{p_{i,j}} \xi_{N-1}^{(j)} \right)^2 - x_i^2 =$$

$$= \beta_i^2 + 2\beta_i \sum_{j=1}^{n} a_{i,j} x_j + \sum_{j=1}^{n} \frac{\alpha_{i,j}^2}{p_{i,j}} M(\xi^{(j)})^2 - x_i^2. \tag{3.86}$$

Справедливість (3.84) випливає з останньої рівності та $x=\alpha x+\beta$. Якщо покласти формально

$$p_{i,j} = p^*_{i,j} = \frac{a_{i,j}\overline{x}_i}{\sum_{k=1}^{n} a_{i,k}x_k} = \frac{a_{i,j}\overline{x}_j}{\overline{x}_i - \beta_i}, \tag{3.87}$$

то права частина рівняння (3.84) перетворюється на нуль. Звідси випливає

Теорема 3.8. Якщо матриця $\mathrm{P}^* = \left\| p^*_{i,j} \right\|$, яка визначається рівністю (3.87), є стохастичною, то виконуються умови симетричності $x=\alpha x+\beta$, які пов'язують P^* з α, і норма матриці $R = \left\| r_{i,j} \right\|$, де $r_{i,j}$ визначається рівністю

$$r_{i,j} = \alpha_{i,j}\sum_{k=1}^{n}\frac{\alpha_{i,j}\overline{x}_k}{\overline{x}_j} = \frac{\alpha_{i,j}(\overline{x}_i - \beta_i)}{\overline{x}_j}, \tag{3.88}$$

менша за одиницю , тоді виконується рівність

$$D\xi^{(i)} = 0,\ i = 1,\dots,n. \tag{3.89}$$

Якщо ж для кожного i

$$\left| 1 - \frac{\beta_i}{\overline{x}_i} \right| < 1, \tag{3.90}$$

то $\xi^{(i)} \equiv \overline{x}_i$ на всіх траєкторіях ланцюга.

Умова (3.90) виконується, якщо всі β_i та $\alpha_{i,j}$ є додатніми.

Доведення. Перше з тверджень теореми 3.8 є очевидним. Позначимо $\xi^{(i)} \equiv 1 - \frac{a_i}{\overline{x}_i}$. Тоді, згідно з (3.84) та (3.87) одержимо

$$\xi^{(i)}_N = \sum_{k=0}^{N}\left\{\prod_{m=1}^{k}\frac{x_{i_{m-1}} - \beta_{i_{m-1}}}{x_{i_m}}\right\}\beta_{i_k} =$$

$$= \overline{x}_i\sum_{k=0}^{N}\left(\prod_{m=1}^{k}\varepsilon_{i_{m-1}}\right)\left(1-\varepsilon_{i_m}\right) = \overline{x}_i\left(1 - \prod_{k=0}^{N}\varepsilon_{i_k}\right).$$

Звідси й випливає друге твердження теореми 3.8. Одержаний результат дозволяє, зокрема, враховувати апріорні відомості про розв'язок для зменшення дисперсії оцінок $\xi^{(i)}$.

3.5. Марковські ланцюги та інтегральні рівняння

3.5.1. Загальна постановка задачі

Зауважимо перш за все, що прийоми розв'язування СЛАР, що описані у попередніх пунктах, можуть бути узагальнені на випадок інтегральних рівнянь 2-го роду. Для цього, однак, треба використати ланцюги Маркова з довільною кількістю станів. Надалі будемо розглядати інтегральні рівняння вигляду

$$\varphi(x) = \int_D K(y,x)\varphi(y)dy + f(x), \tag{3.91}$$

де D – область n-вимірного евклідового простору $x \in D$; φ та f функції, що визначені на D; ядро $K(y,x)$ визначено на декартовому добутку D на себе й інтеграл у (3.91) розуміємо в сенсі Лебега. Хоча розгляд рівнянь більш загального вигляду не складає особливих труднощів, зупинимося на вказаному простому випадку, оскільки він дозволяє навести всі принципово важливі факти.

Нехай збігається метод послідовних наближень для розв'язування рівняння (3.91), тобто розв'язок $\varphi(x)$ може бути поданий у вигляді збіжного ряду

$$\varphi(x) = \sum_{i=0}^{\infty} \psi_i(x), \tag{3.92}$$

$$\psi_0(x) = f(x), \ \psi_i(x) = \int_D K(y,x)\psi_{i-1}(y)dy, \ i = 1,2,\ldots. \tag{3.93}$$

При цьому будемо вважати функції φ та f елементами деякого Банахового простору F і збіжність ряду (3.92) вважати в сенсі збіжності за нормою цього простору. Норму функції φ з F будемо позначати, як завжди $\|\varphi\|_F$ або просто $\|\varphi\|$, якщо зрозуміло, про який простір йдеться. Рівняння (3.91) зручно також записувати у вигляді

$$\varphi = K\varphi + f, \tag{3.94}$$

позначаючи через K інтегральний оператор з ядром $K(y,x)$. Надалі розглядатиметься задача про обчислення методом статистичного моделювання функціонала

$$J = \int_D \varphi(x)h(x)dx = (\varphi, h), \tag{3.95}$$

де $h \in F^*$, де F^* - функціональний простір, спряжений до F. Функціонал J можна також подати через розв'язок φ^* спряженого до (3.91) рівняння

$$\varphi^*(x) = \int_D K(y,x)\varphi^*(y)dy + h(x), \tag{3.96}$$

яке по аналогії до (3.93) буде записуватися також у формі

$$\varphi^* = K^*\varphi^* + h. \tag{3.97}$$

$$J = (\varphi, h) = (f, \varphi^*), \tag{3.98}$$

тобто задача є симетричною стосовно вихідного рівняння (3.91) та спряженого до нього (3.94).

3.5.2. Лінійні оцінки та умови незміщеності

Як правило, з рівнянням (3.91) або (3.94) будемо пов'язувати однорідний ланцюг Маркова, якщо задано щільністю початкового розподілу $\pi(x)$ та перехідною щільністю $p(y \to x)$. Випадок неоднорідного ланцюга недостатньо досліджений. Стосовно однорідного ланцюга $\{\pi, p\}$ необхідно зробити природні припущення, які пов'язані з абсолютною неперервністю введених надалі мір.

1. На множині $\{x : f(x) \neq 0\}$ функція $\pi(x)$ є додатньою.
2. На множині $\{x, y : K(y,x) \neq 0$ функція $p(y \to x)$ є додатньою.

Крім того, припускаємо, що

$$\int p(y \to x)dx = 1 - g(y), \ 0 \le g(y) < 1. \tag{3.99}$$

і будемо трактувати $g(y)$ як імовірність поглинання при переході з точки y до точки x.

У відповідності з загальними методами моделювання, вибіркові траєкторії ланцюга будуємо наступним чином. Моделюючи випадкову величину зі щільністю розподілу $\{\pi(x)\}$, знаходимо точку x_0. Її розглядаємо як початкову координату частинки, що рухається вздовж траєкторії ланцюга. Далі з ймовірністю $g(x_0)$ частинка поглинається. Якщо поглинання не відбулося, то моделюється випадкова величина x_1 щільність розподілу якої дорівнюй $p(x_0 \to x_1)/(1 - g)(x_0)$. Взагалі, якщо відомо x_{i-1}, то з імовірністю $g(x_{i-1})$ траєкторія переривається й з імовірністю $1\text{-}g(x_{i-1})$ вона продовжується, і визначається x_i — реалізація випадкової величини, що розподілена зі щільністю $p(x_{i-1} \to x_i)/(1 - g)(x_{i-1})$. Таким чином, якщо $g > 0$, то траєкторії довжини k- $x_0 \to x_1 \to x_k$ відповідає значення щільності ймовірності, яка дорівнює

$$\pi(x_0)p(x_0 \to x_1)\ldots p(x_{x-1} \to x_k)g(x_k). \tag{3.100}$$

У цьому випадку майже всі траєкторії є скінченними, задання щільності на скінченних траєкторіях визначає міру. Якщо g=0, то майже всі траєкторії мають нескінченну кількість ланок. Міра визначається на всіх скінченних відрізках траєкторії $x_0 \to x_1 \to \ldots x_n$ (циліндричних множинах) і, згідно теореми Колмогорова, може бути розповсюджена на множину всіх траєкторій. Функціонал J, який нас цікавить, оцінюється далі як середнє за вибірковими траєкторіями деякої випадкової величини (оцінки функціоналу), конструкція якої може бути в значній мірі довільною. Як правило, припускається, що ця оцінка має властивість незміщеності.

Будемо розглядати при $g(x) > 0$ оцінки вигляду

$$\tilde{J}[k] = \sum_{i=0}^{k} Q_i^{(k)}(x_0,\ldots,x_k)h(x_i). \tag{3.101}$$

і при $g(x)$=0 оцінки вигляду

$$\tilde{J}[k] = \sum_{i=0}^{\infty} Q_i(x_0,x_1,\ldots)h(x_i). \tag{3.102}$$

де k - номер, який відповідає точці переривання траєкторії, а $Q_i^{(k)}$ та Q_i – деякі задані функції траєкторії, які за припущенням інтегровні за мірою, що визначена на множині траєкторій. Оцінки вигляду (3.101) та (3.102) будемо називати *лінійними оцінками* і припускати скінченність їхніх дисперсій. Вимога незміщеності накладає на функції $Q_i^{(k)}$ та Q_i певні умови, які обговоримо нижче. Для одержаання цих умов представимо функціонал $J=(\varphi,h)$ у зручному для нас вигляді:

$$\begin{aligned}(\varphi,h) &= \sum_{i=0}^{\infty}\int H(x)\psi_i(x)dx = \int H(x)\{f(x) + \\ &\sum_{i=0}^{\infty}\int dx_1 \ldots \int dx_i k(x_i, x)k(x_{i-1}, x_i)\ldots k(x_1, x_2)f(x_1)\}dx\end{aligned} \tag{3.103}$$

Остання рівність при деякій зміні позначень рівносильна рівності

$$(\varphi,h) = \sum_{i=0}^{\infty}\int f(x_0)k(x_0, x_1)\ldots k(x_{i-1}, x_i)H(x_i)d\Xi_i. \tag{3.104}$$

де $d\Xi_t = dx_0, d\Xi_t = dx_0 dx_1 \ldots dx_t$ – математичне сподівання оцінки $\tilde{J}[k]$ дорівнює (припускається, що $g(x)$ відмінна від нуля):

$$M\tilde{J}[k] = \sum_{k=0}^{\infty} \int_{\Re_t} \tilde{J}[k]\pi(x_0)p(x_0 \to x_1)\ldots p(x_{k-1} \to x_k)g(x_k)d\Xi_k. \tag{3.105}$$

Необхідною умовою незміщеності оцінки є виконання рівності

$$(\varphi, h) = M\tilde{J}[k]. \tag{3.106}$$

За наявності деяких додаткових припущень про абсолютну збіжність нескінченних рядів можуть бути одержані достатні умови незміщеності. Найпростішою з них є наступна умова:

$$\sum_{i=0}^{\infty}[\int_{\Re_i} f(x_0)k(x_0, x_1)\ldots k(x_{k-1}, x_k)H(x_k) - \sum_{i=0}^{k} Q_i^{(k)}(\Xi_k)(x_i)\pi(x_0)p(x_0 \to x_1)\ldots p(x_{k-1} \to x_k)g(x_k)]\,d\Xi_k = 0 \tag{3.107}$$

Покладемо далі

$$Q_i^{(k)} = q_i \omega_i^{(k)}, \tag{3.108}$$

$$q_0 = \frac{f(x_0)}{\pi(x_0)}, q_i = q_0 = \frac{k(x_0, x_1)}{p(x_0 \to x_1)} \cdots \frac{k(x_{i-1}, x_i)}{p(x_{i-1} \to x_i)}. \tag{3.109}$$

Якщо врахувати, що $g(x_k) = 1 - \int_0 p(x_k \to x_{k+1})dx_{k+1}$, то формально рівність (3.107) можна записати у вигляді:

$$\int_0 \cdot\left[1 - \omega_0^{(0)}\right] f(x_0 dx_0) + \sum_{k=1}^{\infty} \int_{\Re_k} f(x_0)[\{1 - \omega_k^{(k)}\}k(x_0 x_1)\ldots k(x_{k-1}, x_k)\times$$

$$H(k)\sum_{i=1}^{k-1}\left\{\omega_i^{(k)} - \omega_i^{(k+1)}\right\}k(x_0 x_1)\ldots k(x_{i-1}, x_i)H(i)p(x_i \to x_{i+1})\ldots p(x_{k-1} \to x_k)]\,d\Xi_k = 0$$

Якщо провести зміну порядку інтегрування та сумування, то можна одержати наступний результат.

Теорема 3.9. Для незсуненості оцінки (3.101) при $g(x)>0$ ($x \in \Omega$) і при виконанні умов

1. $\sum_{k=0}^{\infty} \int_{\Re_k} |\omega_k^{(k)} f(x_0)K(x_0, x_1)\ldots K(x_{i-1}, x_i)h(x_i)|\,d\Xi_i < \infty.$

2. $\sum_{i=0}^{\infty}\sum_{k=0}^{\infty} \int_{\Re_k} |(\omega_i^{(k)} - \omega_i^{(k+1)})f(x_0)K(x_0, x_1)\times\ldots\times$

$\times K(x_{i-1}, x_i)h(x_i)p(x_i \to x_{i+1})\ldots p(x_{k-1} - x_k)|\,d\Xi_k < \infty$

досить виконання рівності

$$\sum_{i=0}^{\infty}\int_{R_i} f(x_0)K(x_0,x_1)...K(x_{i-1},x_i)h(x_i)\times$$
$$\times\{1-\omega_i^{(i)}-\sum_{k=i}^{\infty}\int_{\Re_{k-1}}[\omega_i^{(k)}-\omega_i^{(k+1)}]p(x_i\to x_{i+1})\times...\times \qquad (3.110)$$
$$\times p(x_k\to x_{k+1})dx_{i+1}...dx_{k+1}\}d\Xi_i=0.$$

Може бути цікавим випадок, коли $Q_i^{(k)}$ залежать тільки від перших i+1 змінних $x_0,\ldots, x_i$, бо обчислення при цьому, як правило, зміщуються. Рівність (3.110) у цьому випадку також спрощується і може бути записана у вигляді

$$\sum_{i=0}^{\infty}\int_{R_i} f(x_0)K(x_0,x_1)...K(x_{i-1},x_i)h(x_i)\{1-\omega_i^{(i)}$$
$$-\sum_{k=i}^{\infty}\omega_i^{(k+1)}-\omega_i^{(k)}]\int_d dx_{i+1}...\int_d dx_{k+1}p(x_i\to x_{i+1})\times...\times \qquad (3.111)$$
$$\times p(x_k\to x_{k+1})\}d\Xi_i=0.$$

Шляхом нескладних перетворень рівність (3.110) можна подати у вигляді

$$\sum_{i=0}^{\infty}\int_{R_i} p(x_0\to x_1)...p(x_i\to x_{i+1})q_i h(x_i)\{1-\sum_{k=i}^{\infty}\int_d dx_{i+1}\times...\times$$
$$\times\int_d dx_k\, p(x_i\to x_{i+1})...p(x_{k-1}\to x_k)g(x_k)\omega_i^{(k)}\}d\Xi_i=0. \qquad (3.112)$$

який іноді є більш зручним для застосування.

3.6. Методи статистичного моделювання для розв'язування стохастичних диференціальних рівнянь

При дослідженні на комп'ютерах динамічних систем часто використовується метод статистичного моделювання. Застосування цього методу для дослідження систем, які задані стохастичними диференціальними рівняннями, вимагає заміни цих рівнянь різницевими співвідношеннями. Як правило, використовують різницеві схеми та Рунге-Кутта. Однак, відомі оцінки похибки різницевих методів розв'язування детермінованих рівнянь не можуть бути використані у цифровому моделюванні стохастичних рівнянь, оскільки їх розв'язки майже всюди недиференційовні. У цьому розділі для оцінки похибок використовується стохастичний аналог ряду Тейлора, який дозволяє розкласти розв'язки стохастичного диференціального рівняння в ряд за нелінійними функціоналами вінерівського процесу. Вивчено збіжність розподілів, дано оцінки точності методів Ейлера та Рунге-Кутта, запропоновано модифікації методів Ейлера та Рунге-Кутта для підвищення швидкості збіжності.

3.6.1. Загальна постановка задачі

Нехай динамічна система описується стохастичним диференціальним рівнянням

$$dx(t)=a(x(t),t)dt+\sigma(x(t),t)dw(t), \qquad (3.113)$$

де $w(t)=w(t,\omega) \in \mathbf{R}^1$ – стандартний вінерів процес, $a(x,t)$, $\sigma(x(t),t)$ – гладкі за обидвома аргументами функції, які зростають на нескінченності не швидше за лінійні:

$$|a(x,t)^2| + |\sigma(x,t)^2| \le c^2(1+|x^2|), \tag{3.114}$$

де $0<C<\infty$. і ця умова виконується рівномірно для всіх $tm[t_0, t_0+T]$. Під σ-алгеброю $\mathsf{F}_{t_0}^t$ будемо розуміти мінімальну σ алгебру, відносно якої величини $x(t_0)$ та $w(s)$-$w(t_0)$ є вимірні для всіх $s \in [t, t_0]$.

Відомо, що умови регулярності, які вимагаються від $a(x,t)$ та $b(x,t)$ є достатніми для того, щоб $\mathsf{F}_{t_0}^t$-вимірний розв'язок $x(t)$ існував і був єдиним в сенсі стохастичної еквівалентності. На відрізку $[t_0, t_0+T]$. виберемо n точок $t_\kappa = t_0 + kh$, k=0.1,..., h=T/n – крок дискретизації. За допомогою процедури числового інтегрування рівняння (3.113) за початковою умовою $x_0=x(t_0)$ за визначенням приростів вінерового процесу $\Delta w_k = w(t_{k+1}) - w(t_k)$ необхідно одержати оцінки $\hat{x}_k$ точного розв'язку $x_k=x(t_k)$. Величини $\hat{x}_k$ є функціями від вибіркових значень процесу. Тому $\hat{x}_k$ визначені на тому самому ймовірнісному просторі, що й вибіркові значення точного розв'язку, і вимірні відносно σ-алгебри $\mathsf{F}_{t_0}^t$. Дослідження на комп'ютері збіжності різницевих схем у середньому квадратичному при різних h повинно відбуватися для однієї й тої самої реалізації вінерового процесу $w(t)$ (з наступним усередненням за множиною реалізацій). Тому датчик випадкових приростів Δw_k організовується таким чином, щоб значення процесу $w(t)$ при послідовному зменшенні кроку інтегрування h у два рази лишалися однаковими в одні й ті самі моменти часу. Для цього попередньо вибирається мінімальне значення кроку $h_{min} = T/2^m$ і при інтегруванні рівнянь (3.113) з більшим кроком $h = T/2^N$, $N \ge M$, прирости вінерового процесу утворюються за допомогою процедури

$$\Delta w_k = \left(\frac{T}{2^M}\right)^{1/2} \sum_{i=1}^{r} \zeta_{(k-1)l+i}, \tag{3.115}$$

де $l=2^{M-N}$; $\zeta_i = \zeta_i(\omega) m \mathbf{R}^1$ – центровані незалежні Гаусові випадкові величини з одиничною дисперсією. Генеруючи відрізки послідовностей псевдовипадкових чисел $\{\zeta_i\}$ з одним і тим самим початковим значенням, одержуємо, таким чином, для різних $h = T\backslash 2^N$ вибірки з однієї й тієї самої реалізації процесу $w(t)$.

Надалі розглядатимемо різницеві схеми вигляду

$$\tilde{x}_{k+1} = \tilde{x}_k + Q_k(\tilde{x}_k, H, \Delta w_k), \tag{3.116}$$

де Q_k $(x,H,\Delta w)$ визначається конкретним виглядом різницевої схеми розв'язку рівняння (3.113) в залежності від області застосування рівняння (3.113) будемо розрізняти збіжності двох типів: у середньому квадратичному і збіжність розподілів наближених розв'язків до точних.

При вивченні збіжності в середньому квадратичному *якість* апроксимації точного розв'язку x_n різницевою схемою, за якою обчислюються оцінки $\tilde{x}_n$, характеризується середньою квадратичною похибкою R, яка визначається за формулою

$$R = \max_{1\le k\le n}(x_k - \tilde{x}_k)^2 / x_0\})^{\frac{1}{2}}, \tag{3.117}$$

де початкові умови точного та наближеного розв'язків співпадають ($\tilde{x}_0 = x_0$). Введемо локальну похибку різницевої схеми розв'язку стохастичного рівняння

$$\rho_k = x_{k+1} - \tilde{x}_{k+1}, \tag{3.118}$$

де різниця $x_{k+1} - \tilde{x}_{k+1}$ обчислена за умови, що на k-му кроці початкові умови співпадають: $x_k = \tilde{x}_k$. Знайдемо зв'язок між глобальною похибкою на k-му кроці $x_k - \tilde{x}_k$ та локальною похибкою ρ_k.

Лема 3.6. Якщо в різницевій схемі (3.114) та диференціальному рівнянні (3.113) початкові умови співпадають та $Q_k(x,h,\Delta w)$ задовольняють умову

$$M\{(Q_k(x,h,\Delta w_k) - Q_k(y,h,\Delta w_k))^2 / x, y\} \le K(h)|x-y|^2, \tag{3.119}$$

де

$$\lim_{h\to 0} K(h) = 0, \tag{3.120}$$

то

$$x_k - \tilde{x}_k = \sum_{j=0}^{k-1}\rho_j + \varepsilon, \tag{3.121}$$

де

$$\lim_{h\to 0}(M\{\frac{\varepsilon^2}{x_0}\} / M\{(x_k - \tilde{x}_k)^2 / x_0\}) = 0. \tag{3.122}$$

Доведення. Проведемо методом індукції. При k=1 лема є справедливою. Припустимо, що вона є вірною на k-му кроці. Тоді з означення похибки ρ_k (див. (3.118) та (3.114)) отримаємо, що вибірка з точного розв'язку задовільняє рекурентне співвідношення

$$x_{k+1} = x_k + Q(x_k, h, w_k) + \rho_k. \tag{3.123}$$

Підставляючи x_k з (3.121) до цієї рівності та враховуючи умову (3.116), одержимо твердження леми на $(k+1)$-му кроці. Збіжність розподілів послідовності $\{\tilde{x}_k\}$ до розподілів послідовності $\{x_k\}$ випливає зі збіжності умовних моментів

$$\tilde{m}_k(x) = M\{\frac{f(x_k)}{x_0} = x\}. \tag{3.124}$$

для відповідних умовних моментів точного розв'язку для функції $f(x)$ класу гладких та обмежених $\forall x \in (+\infty;-\infty)$. При вивченні збіжності розподілів якість апроксимації характерезуємо за допомогою критерію

$$q = \max_{1\le k\le n} m_k - \tilde{m}_k. \tag{3.125}$$

$$m = ess\sup|m(x)|. \tag{3.126}$$

Введемо на k-му кроці

$$\delta_k(x) = M\left\{\frac{m_k(x_{k+1})}{x_k} = x\right\} - M\left\{\frac{m_k(\tilde{x}_{k+1})}{x_k} = x\right\}. \tag{3.127}$$

та доведемо наступне твердження.

Лема 3.7. Якщо існує $0<A<\infty$ таке, що $|\delta_k(x)|< A$ для довільного $k\le n$ рівномірно по x, то для критерію точності q (див. (3.125)), справджується оцінка $q = o(A/h)$.

Доведення проводиться за допомогою методу оператора переводу, який потрібно застосувати до операторного рівняння

$$\hat{m}_{k+1}(x) = M\left\{\frac{\hat{m}_k(\tilde{x}_{k+1})}{\tilde{x}_{k+1}} = x\right\}. \tag{3.128}$$

3.6.2. Збіжність методу Ейлера в середньому квадратичному

Різницева схема Ейлера розв'язування рівняння (3.113) набуває вигляду

$$\tilde{x}_{k+1} = \tilde{x}_k + a_k(\tilde{x}_k)h + \sigma_k(\tilde{x}_k)\Delta w_k,\ k = 0,1,\ldots. \tag{3.129}$$

де $a_k(\tilde{x}_k) = a_k(\tilde{x}_k, t_k)$, $\sigma_k(\tilde{x}_k) = \sigma_k(\tilde{x}_k, t_k)$, $\Delta w_k = w(t_{k+1}) - w(t)$ – приріст вінерового процесу на проміжку t_{k+1}, t_k. Вважають, що початкова умова $\tilde{x}_0$ є відомою.

Теорема 3.10. Нехай коефіцієнти $a(x,t), \sigma(x,t)$ регулярні та мають всі похідні по t до другого порядку, а по x – до четвертого. Тоді мають місце наступні оцінки для R: якщо $\sigma'_x\sigma(x,t) \neq 0$, то

$$R \le h^{\frac{1}{2}}\left(\int_{t_0}^{t_0+T} M\left\{\left[\frac{\sigma'_x\sigma(x(t),t)\right]^2}{x_0}\right\}dt\right)^{\frac{1}{2}}, \tag{3.130}$$

якщо $\sigma'_x\sigma(x,t) = 0$, то

$$R \le 2h\left(\int_{t_0}^{t_0+T}\left\{[a_x\sigma(x(t),t)]^2 + \frac{[L_t\sigma(x(t),t)]^2}{x_0}\right\}dt\right)^{\frac{1}{2}} + \\ + hM\left\{\int_{t_0}^{t_0+T}\left(\frac{[L_t\sigma(x(t),t)dt]^2}{x_0}\right)^{\frac{1}{2}}\right\}. \tag{3.131}$$

$$L_t = \frac{\partial}{\partial t} + L,\ L = a(x,t)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2}{\partial x^2} \tag{3.132}$$

є твірним оператором марківського процесу $x(t)$.

Для доведення теореми 3.8 отримаємо стохастичний аналог ряду Тейлора.

Теорема 3.11. Нехай у рівнянні (3.113) коефіцієнти $a(x,t).\sigma(x,t)$ задовільняють умови регулярності, які сформульовані в теоремі 3.10. Тоді розв'язок рівняння Іто (3.113) за початковою умовою $x(t_0)=x_0$ можна подати у вигляді стохастичного аналога ряду Тейлора – Іто

$$x(t) = x_0 + a\Delta t + \sigma\Delta w + \frac{1}{2}\sigma'_x\sigma\left[(\Delta w)^2 - \Delta t\right] + a'_x\sigma\int_{t_0}^{t}\Delta w(s)ds + \\ + L_f\sigma\int_{t_0}^{t} sw(s) + \frac{1}{6}(\sigma'_x,\sigma)'_x\sigma\left[2(\Delta w)^3 - 3\Delta w\Delta t\right] + L_f a\frac{(\Delta t)^2}{2} + C(\Delta t), \tag{3.133}$$

де $\Delta t = t - t_0$, $\Delta w = \Delta w(t) = w(t) - w(t_0)$, a'_x – похідна по x, всі похідні в (3.133) обчислені в точці (x_0,t_0). Для залишкового члена C в (3.133) справедлива оцінка

$$M\left\{C^2(\Delta t)\Big/x_0\right\} = K(\Delta t)^4, \tag{3.134}$$

де $K < \infty$.

Доведення. Позначимо через α скінченну послідовність символів 0, 1. Через $q_0(\alpha)$ позначимо кількість нулів у цій послідовності, через $q_1(t)$ – кількість одиниць. Нехай $z(t)$ – випадковий процес, невипереджуючий відносно вінерового процесу $w(t)$. Визначимо твірну сім'ю $z(t)$ випадкових процесів $C_\alpha^z(t)$, де a відіграє роль параметра сім'ї за допомогою наступних рівностей:

$$dz(t) = C_1^z(t)dt + C_0^z(t)dw(t), \tag{3.135}$$

$$dC_a^z(t) = C_{a_1}{}^z(t)dt + C_{a_0}{}^z(t)dw(t). \tag{3.136}$$

Тут $a_0(a_1)$ – послідовність, утворена з a та приписаного праворуч символу 0(1). Наведені вище рівності рекурентно визначають сім'ю C_α^z, якщо існують відповідні стохастичні диференціали.

Лема 3.8. Нехай для процесів $x(t)$ та $y(t)$ існують всі C_α^x та C_2^x при a, яке задовольняє умову

$$q_0(a) + 2q_1(a) \le n+1, \tag{3.137}$$

де $n \ge 0$ ціле, та виконуються рівності

$$x(t_0) = y(t_0) = x_0, \quad C_a^x(t_0) = C_a^y(t_0) = C_a^y(t_0). \tag{3.138}$$

Тоді функція $\varphi(t) = M\{(x(t) - y(t))^2 / x_0\}$, де $t \ge t_0$, є диференційованою $(n+1)$ разів по t та

$$\varphi(t) = o((t+t_0)^n). \tag{3.139}$$

Доведення леми 3.8 полягає в диференціюванні $\varphi(t)$ з використанням формули Іто. Для доведення теореми 3.11 залишається перевірити, що якщо процес $x(t) \equiv x(t,\omega)$ є розв'язком вихідного рівняння (3.113), а процес $y(t) \equiv y(t,\omega)$ визначається правою частиною рівності (3.133) без залишкового члена, то умови леми 3.8 виконуються, якщо покласти $n=3$. При цьому умови існування $C_a^x(t)$ переходять у відповідні умови гладкості коефіцієнтів $a(x,t)$, $\sigma(x,t)$, які вказані у формулюванні теореми 3.11. Існує

конструктивний спосіб побудови розкладів виду (3.133) із залишковим членом в інтегральній формі. Цей спосіб порівняно громіздкий, тому тут не поданий.

Доведення теореми 3.10. Порівнюючи (3.129) та (3.133) при $x_0=x_1$, $\Delta t=h$, $\Delta w=\Delta w_k$, $t_0=t_k$ згідно формули (3.118) отримаємо

$$\rho_k=\frac{1}{2}\sigma_x'\sigma\left[\Delta w_k^2-h\right]+a_x'\sigma\int_{t_k}^{t_{k+1}}\Delta w_k(s)ds+ \\ +L_f\sigma\int_{t_k}^{t_{k+1}}sdw(s)+\frac{1}{6}(\sigma_x'\sigma)_x'\sigma(2\Delta w_k^3-3\Delta w_k h), \tag{3.140}$$

де всі функції в правій частині обчислюються в точці $\{x_k,t_k\}$. Згідно з лемою 3.5, спільна похибка $x_k-\hat{x}_k$ співпадає з сумою локальних похибок ρ_k. Розглянемо спочатку випадок $\sigma_x'\sigma\neq 0$. Тоді в (3.134) всі доданки після першого можна відкинути, оскільки вони мають більш високий порядок малості по h в середньому квадратичному. Тому (3.119) набуде вигляду

$$x_k-\hat{x}_k=\sum_{j=0}^{k-1}\rho_j=\frac{1}{2}\sum_{j=0}^{k-1}\sigma_x'\sigma(x_j,t_j)\left[(\Delta w_j)^2-h\right]. \tag{3.141}$$

Величина $\left[(\Delta w_j)^2-h\right]/2$ має нульове середнє значення та дисперсію $h/2$. Якщо винести з-під знака суми множник $\frac{1}{2}h^{1/2}$ отримуємо

$$x_k-\hat{x}_k-\frac{1}{2}h^{1/2}\sum_{j=0}^{k-1}\sigma_x'\sigma(x_j,t_j)\xi_j h^{1/2}, \tag{3.142}$$

де $\xi_j=\frac{1}{\sqrt{2h}}\left[(\Delta w_j)^2-h\right]$ – нульовим середнім та одиничною дисперсією, які не залежать від x_j та незалежні при різних j. Можна довести, що незалежно від розподілу величин ξ_j

$$\sum_{j=0}^{n-1}\sigma_x'\sigma(x_j,t_j)\xi_j h^{1/2}=\int_{t_0}^{t_0+T}\sigma_x'\sigma(x(t),t)dw_1(t)+o(1). \tag{3.143}$$

де $o(1)$ - величини, які збігаються до нуля при $h\to 0$, $nh=T$, $w_1(t)$ стандартний вінерів процес, який визначається приростами $\xi_j h^{\frac{1}{2}}$. Процес $w_1(t)$ є невипереджуючим стосовно σ-алгебри $\Im_{t_0}'$. З вищенаведених міркувань маємо

$$x_k-\hat{x}_k-\frac{1}{2}h^{1/2}\int_{t_0}^{t_k}\sigma_x'\sigma(x_j,t_j)sw_1(t)+o(1). \tag{3.144}$$

Проводячи оцінку цього інтеграла за допомогою формули для середнього

квадрату стохастичного інтеграла Ito, приходимо до шуканого результату (3.130). Застосовуючи аналогічний прийом при підсумовуванні локальних похибок у випадку $\sigma_x\sigma = 0$, одержуємо оцінку (3.131).

Наслідок 3.5. Якщо коефіцієнти рівняння (3.113) задовільняють умови теореми 3.8, то різницева схема

$$\hat{x}_{k+1} = \hat{x}_k + a_k(\hat{x}_k)h + \frac{1}{2}\frac{\partial\sigma_k}{\partial x}\sigma_k(\hat{x}_k)\left[(\Delta w_k)^2 - h + b_k(\hat{x}_k)\Delta w_k\right] \tag{3.145}$$

має точність $R=O(h)$. Зауважимо, що використання у різницевій схемі (3.145) відомого наближенго співвідношення $(\Delta w_k)^2 \cong h$ приводить до менш точної (у розумінні збіжності в середньому квадратичному) схеми Ейлера (3.129).

3.6.3. Збіжність методу Ейлера як збіжність розподілів

Теорема 3.12. Якщо коефіцієнти рівняння (3.113) мають чотири неперервні похідні по x, дві неперервні похідні по t, введена у пункті 3.2 функція $f(x)$ має шість неперервних похідних по x, то оцінка похибки (3.125) методу Ейлера дорівнює $q=O(h)$.

Доведення. Для розрахунку похибки (3.125) обчислимо

$$\varphi_k(x,h) = M\{\frac{f\left(x(t_k+h)\right)}{x_k} = x\}. \tag{3.146}$$

Розкладемо функцію $\varphi_k(x,h)$ у ряд Тейлора за степенями h. З формули Ito одержимо співвідношення

$$d_h\varphi_k(x,h) = L_t f(x)dt, \tag{3.147}$$

де L_t визначається рівністю (3.132). Застосувавши це співвідношення, одержимо розклад виразу (3.145) у ряд Тейлора за змінною h в околі точки $t = t_k, x = x_k$

$$\begin{aligned}
&\varphi_k(x,h) = f(x) + L_t f(x)h + L_t^2 f(x)\frac{h^2}{2} + o(h^2) = \\
&= f + f_x'\left[ah + \left(a_x'a + \frac{1}{2}\sigma^2 a_x''\right)\frac{h^2}{2}\right] + \\
&+ \frac{1}{2}f_x''\left[\sigma^2 h + \left(a^2 + a\sigma_x'\sigma + \sigma^2 a_x' + \frac{1}{4}\sigma^2(\sigma^2)_x''\right)h^2\right] + \\
&+ \frac{1}{2}f'''[a\sigma^2 + \sigma^3\sigma']h^2 + \frac{1}{8}f^{IV}\sigma^4 h^2 + \frac{\partial}{\partial t}\left(af_x' + \frac{1}{2}\sigma f_x''\right)\frac{h^2}{2} + o(h^2),
\end{aligned} \tag{3.148}$$

де штрихи позначають похідні по x, аргументи x,t та індекси k тут і надалі у цьому пункті не писатимемо. Розклад (3.148) справедливий, якщо функції $a(x,t), \sigma(x,t)$ мають обмежену четверту, а $f(x)$ – обмежену шосту похідні по x. То такий запас гладкості необхідний для того, щоб залишковий член у розкладі (3.148) мав порядок $o(h^2)$. Далі обчислимо оцінку

$$\hat{\varphi}_k(x,t) = M\left\{ f(\hat{x}_{k+1}) \Big/ \hat{x}_k = x \right\}. \tag{3.149}$$

при використанні різницевої схеми Ейлера (3.129). Розкладаючи $\hat{\varphi}_k(x,t)$ у ряд Тейлора в околі точки x, одержимо

$$\begin{aligned} \hat{\varphi}_k(x,t) &= f(x) + f'_x M\{\Delta x / x\} + \frac{1}{2} f''_x M\{(\Delta x)^2 / x\} + \\ &+ \frac{1}{6} f'''_x M\{(\Delta x)^3 / x\} + \frac{1}{2} f^{IV}_x M\{(\Delta x)^4 / x\} + o(h^2), \end{aligned} \tag{3.150}$$

де $\Delta x = \hat{x}_{k+1} + \hat{x}_k = a_k(\hat{x}_k)h + \sigma_k(\hat{x}_k)\Delta w_k$. Завдяки гаусовості приростів вінерового процесу Δw_k моменту приросту Δx, легко обчислити з (3.129):

$$\begin{aligned} &M\{\Delta x / x\} = a_k h, \\ &M\{(\Delta x)^2 / x\} = a_k^2 h^2 + \sigma_k^2 h, \\ &M\{(\Delta x)^3 / x\} = 3\sigma_k^2 a_k h + o(h^2), \end{aligned} \tag{3.151}$$

$$M\{(\Delta x)^{2/4} / x\} = 3\sigma_k^4 h^2 + o(h^2). \tag{3.152}$$

Підставляючи (3.151) у (3.150), (3.148) та (3.150) у (3.127), одержимо нев'язку методу (3.129)

$$\boldsymbol{\delta}_k(\mathbf{h}) = \mathrm{o}(\mathrm{h}^2). \tag{3.153}$$

Застосувавши до (3.153) лему 3.6, одержуємо $q=O(h)$, що й треба було довести. Аналогічним чином можна показати, що модифікована схема Ейлера (3.145) також має похибку $q=o(h)$ і тому має однакову з (3.129) швидкість збіжності у сенсі збіжності розподілів.

3.6.4. Оптимальні різницеві схеми розв'язку стохастичних рівнянь

Наведемо означення оптимальної в середньому квадратичному різницевої схеми.

Означення 3.1. Різницева схема вигляду (3.116) називається оптимальною, якщо вона на кожному кроці має мінімальну середню квадратичну похибку

$$S_k = M\left\{ \left(x_{k+1} - \hat{x}_{k+1} \right)^2 / x_k = \hat{x}_k, \Delta w_k \right\}, \tag{3.154}$$

де x_{k+1} – поточний розв'язок рівняння (3.113) у точці t_{k+1} з початковою умовою x_k, $\hat{x}_{k+1}$ – наближений розв'язок, який апроксимує x_{k+1}. Розв'язок задачі мінімізації S_k задається оператором умовного математичного сподівання

$$\hat{x}_{k+1} = M\left\{ x_{k+1} / x_k = \hat{x}_k, \Delta w_k \right\}. \tag{3.155}$$

Для одержання оптимальної різницевої схеми скористаємося розкладом (3.133), де $x_0 = x_0$, $\Delta w(t) = \Delta w_k$, $\Delta t = h$:

$$M\left\{x_{k+1} / x_k = \hat{x}, \Delta w_k\right\} = \hat{x} + ah + \sigma w_k +$$
$$+\frac{1}{2}\sigma_x'\sigma\left[\left(\Delta w_k\right)^2 - h\right] + a_x'\sigma I\left\{\int_{t_k}^{t_{k+1}} \Delta w(s)ds / \Delta w_k\right\} +$$
$$+L_1\sigma M\left\{\int_{t_k}^{t_{k+1}} sdw(s) / \Delta w_k\right\} + L_t a\frac{h^2}{2} + \tag{3.156}$$
$$+\frac{1}{6}\left(\sigma_x'\sigma\right)_x'\sigma\left[2\left(\Delta w_k\right)^3 - 3\Delta w_k h\right],$$

де всі похідні обчислюються у точці $\left(\hat{x}, t_k\right)$. Математичні сподівання у правій частині цієї рівності обчислюються за допомогою наступної леми.

Лема 3.9. Нехай $w(t)$ – стандартний вінерів процес, $t \in \left[t_k, t_{k+1}\right]$. Тоді

$$M\left\{w(t) - w(t_k) / \Delta w_k\right\} = \frac{t - t_k}{h}\Delta w_k. \tag{3.157}$$

Доведення. Величини $w(t)$–$w(t_k)$ та $\Delta w_k = w(t_{k+1}) - w(t_k)$ мають гаусів розподіл з дисперсіями $\sigma_1^2 = t - t_k, \sigma_2^2 = h$ відповідно й коефіцієнтом кореляції

$$p = \left(\frac{t - t_k}{h}\right)^{1/2}. \tag{3.158}$$

Підставивши σ_1^2, σ_2^2 та p у формулу для умовного розподілу Гуаса, одержимо шукані результати. Враховуючи, що $M\left\{dw(s) / \Delta w_k\right\} = dt\Delta w_k / h$, застосуємо лему для обчислення математичних сподівань у (3.156) і одержимо

$$M\left\{\int_{t_k}^{t_{k+1}} \Delta w(s)ds / \Delta w_k\right\} = M\left\{\int_{t_k}^{t_{k+1}} sdw(s) / \Delta w_k\right\} = \frac{\Delta w_k h}{2}. \tag{3.159}$$

Підставивши цей вираз у (3.156), одержимо твердження наступної теореми.

Теорема 3.13. Якщо коефіцієнти рівняння (3.113) задовільняють умови регулярності, які сформульовані в умові теореми 3.10, то оптимальна різницева схема (3.155) розв'язку рівняння Іто (I) має вигляд

$$\hat{x}_{k+1} = \hat{x} + ah + \sigma\Delta w_k + \frac{1}{2}\sigma_x'\sigma\left[\left(\Delta w_k\right)^2 - h\right] +$$
$$+\left(a_x'\sigma + L_1\sigma\right)\frac{\Delta w_k h}{2} + L_t a\frac{h^2}{2}. \tag{3.160}$$

де всі функції обчислюються у точці $\hat{x}_k$, t_k.

Різницева схема, одержана в цій теоремі, може бути використана як для порівняння точності різних різницевих схем з класу (3.116), так і для одержання більш точних, ніж схема Ейлера (3.129), методів розв'язку стохастичних диференціальних рівнянь. Розглянемо точність власне оптимальної різницевої схеми (3.160). Порівнявши (3.133) та (3.160), одержимо локальну похибку p_k оптимальної різницевої схеми:

$$p_k = a_x'\sigma\left[\int_{t_k}^{t_{k+1}} \Delta w(s)\,ds - \frac{1}{2}\Delta w_k h\right] + L_t\sigma\left[s\,dw(s) - \frac{1}{2}\Delta w_k h\right]. \qquad (3.161)$$

Оптимальна різницева схема відноситься до типу (3.116), тому для обчислення глобальної похибки $x_k - \hat{x}_k$ застосуємо лему 3.5. Використовуючи для підсумовування похибок той же прийом, що й при доведенні теореми 3.8, одержимо результат,який сформулюємо у вигляді наступної теореми.

Теорема 3.14. Для точності (3.119) оптимальної різницевої схеми (3.160) справедлива оцінка

$$R \le h\left[\int_{t_0}^{t_0+T} M\left\{\left[a_x'\sigma\left(x(t),t\right)\right]^2 + \left[L_t b\left(x(t),t\right)\right]^2 / x_0\right\} dt\right]^{1/2}. \qquad (3.162)$$

Теорема 3.14 встановлює потенційну границю точності розв'язку рівняння (3.113) з кроком h при коефіцієнтах a і b загального вигляду. Оптимальна різницева схема є незручною для розрахунків, оскільки вимагає великої кількості похідних від коефіцієнтів на кожному кроці. Апроксимуючи вказані похідні з різними степенями точності, можна одержати низку різницевих схем для практичних розрахунків.

3.7. Наближене інтегрування стохастичних диференціальних рівнянь

Розглянемо стохастичне диференціальне рівняння

$$dx = a(t,x)dt + \sigma(t,x)dw(t),\ \ x(f_0) = x_0, \qquad (3.163)$$

де $w(t)=w(t,w)$ – стандартний вінерів процес. Наведемо приклад чисельної апроксимації у середньому квадратичному випадкової величини $\{x(t_0 \mid T) =)x(t_0 \mid T,\omega)\}$.

Якщо розбити проміжок $[t_0, t_0+T]$ на m рівних частин та покласти $h = T/m$, $t_{k+1} = t_k + h,\ \ k = 0,1,\ldots,m-1$, то відома рекурентна апроксимація

$$\bar{x}(t_0) = x(t_0); \qquad (3.164)$$

$$\bar{x}(t_{k+1}) = \bar{x}(t_k) + a(t_k, \bar{x}(t_k))h + \sigma(t_k, \bar{x}(t_k))w_{k+1}, \qquad (3.165)$$

де w_{k+1} – незалежні за сукупністю випадкові величини, які розподілені нормально з нульовим математичним сподівання та дисперсією h, дає для квадрата середньоквадратичного відхилення $M_{t_0,x_0}(x(t_0+T) - \bar{x}(t_0+T))^2$ значення, яке дорівнює

$o(h)$. Розглянемо метод, який використовує на кожному кроці як і (3.165), лише величину w_{k+1}, але точність якого дорівнює $o(h)^2$.

3.7.1. Однокрокова апроксимація

Сконструюємо випадкову величину

$$\bar{x}(t_0+T) = x_0 + c_1 h + c_2 w(h) + c_3 w^2(h), \tag{3.166}$$

де $w(h) = w(t_0+h) - w(t_0)$, і підберемо c_1, c_2, c_3 таким чином, що квадрат середньоквадратичного відхилення $M\{(x(t_0+h) - \bar{x}(t_0+h))^2\}$ мав би якнайвищий порядок малості по h. Для цього введемо випадковий процес $\{x,w\}$ та його оператор L:

$$Lf(h,x,w)\frac{\delta f}{\delta h} + a(h,x)\frac{\delta f}{\delta h} + \frac{1}{2}\sigma^2(h,x)\frac{\partial^2 f}{\partial x^2} + \sigma(h,x)\frac{\partial^2 f}{\partial x \partial w} + 1\frac{\partial^2 f}{2\partial w^2}. \tag{3.167}$$

Наведемо важливу формулу, що пов'язана з тейлорівським розкладом півгруп

$$M_{t_0,x_0,w_0} f(t_0+h, x(t_0+h), w(t_0+h)) = f(t_0,x_0,w_0) + Lf(t_0,x_0,w_0)h + \ldots + \frac{1}{k!}L^k f(t_0,x_0,w_0)h^2 + o(h^{k+1}). \tag{3.168}$$

Формулу (3.168) можна одержати при відповідних обмеженнях на функції a, σ та f з відомих теорем теорії півгруп. Призначимо $\varphi(x,h,w) = x - \bar{x}(t_0+h) = x - x_0 - c_1 h - c_2 w - c_3 w^2$. З (3.168) випливає, що якщо $L^s\varphi^2(0,x_0,0) = 0$, $s = 1,\ldots,k$, то $M_{t_0,x_0,w_0}(x(t_0+h) - \bar{x}(t_0+h))^2 = o(h^{k+1})$. Шляхом простих перетворень одержимо формулу

$$L(f\cdot g) = Lf\cdot g + f\cdot Lg + Sf\cdot Sg, \tag{3.169}$$

$$Sf(h,x,w) = \sigma(h,x)\frac{\delta f}{\delta x} + \frac{\delta f}{\delta w}. \tag{3.170}$$

Використовуючи (3.169), знаходимо

$$L\varphi^2 = 2\varphi\cdot L\varphi + (S\varphi)^2, \tag{3.171}$$

$$L^2\varphi^2 = 2(L\varphi)^2 + 2\varphi L^2\varphi + 2S\varphi(SL\varphi + LS\varphi) + (S^2\varphi)^2. \tag{3.172}$$

Позначивши (f_0) значення функції f при $h=0, x=x_0, w=0$. зі співвідношень $(L\varphi^2)_0 = 0$ та $(L^2\varphi^2)_0 = 0$ одержимо

$$(S\varphi)_0 = 0,\ (L\varphi)_0 = 0,\ (S^2\varphi)_0 = 0. \tag{3.173}$$

$$c_1 = a(t_0,x_0) - \frac{1}{2}\sigma(t_0,x_0)\frac{\delta\sigma}{\delta x}(t_0,x_0),\ c_2 = \sigma(t_0,x_0),\ c_3 = \frac{1}{2}\sigma(t_0,x_0)\frac{\delta\sigma}{\delta x}(t_0,x_0). \tag{3.174}$$

Отже, однократна апроксимація (3.166) з коефіцієнтами (3.174) має порядок h^3 для квадрата середньоквадратичної похибки. Відзначимо, що звичайна апроксимація (3.165) має на одному кроці порядок h^2. Якщо ввести випадкову величину $y \equiv \int_0^h w dt$ та сконструювати $\bar{x}(t_0 + h)$ за формулою

$$\bar{x}(t_0 + h) = x_0 + c_1 h + c_2 w + c_3 w^2 + c_4 hw + c_5 w^3 + c_6 y, \tag{3.175}$$

то можна підібрати коефіцієнти $c_1, \ldots, c_6$ таким чином, що квадрат середньоквадратичної похибки на одному кроці дорівнюватиме $o(h^4)$. Розглянемо для цього випадковий процес $\{x, w, y\}$, який породжений системою

$$\begin{aligned} x &= a dt + \sigma dw, \\ dw &= 1 \cdot dw, \\ dy &= w dt. \end{aligned} \tag{3.176}$$

Оператор L цього процесу дорівнює

$$Lf(h, x, w, y) = \frac{\delta f}{\delta h} + a\frac{\delta f}{\delta x} + w\frac{\delta f}{\delta y} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2} + \frac{1}{2}\frac{\partial^2 f}{\partial w^2} + \frac{\partial^2 f}{\partial x \partial w}. \tag{3.177}$$

Формули (3.169)-(3.172) з врахуванням нової змінної y зберігаються і для процесу $\{x, w, y\}$. При цьому функція $\varphi(h, x, w, y)$ дорівнює

$$\varphi = x - x_0 - c_1 h - c_2 w - c_3 w^2 - c_4 hw - c_5 w^3 - c_6 y. \tag{3.178}$$

Знаходячи з (3.169), (3.172) $L^3\varphi^2$ і прирівнюючи до нуля $(L^3\varphi^2)_0$, одержимо поряд з (3.173)

$$(SL\varphi)_0 = 0, \quad (LS\varphi)_0 = 0, \quad (S^3\varphi)_0 = 0. \tag{3.179}$$

З формул (3.174) та (3.179) можна одержати (3.175) та наступні вирази для коефіцієнтів c_4, c_5, c_6

$$c_4 = \left(\frac{\delta\sigma}{\delta t} + a\frac{\delta\sigma}{\delta x} - \frac{1}{2}\sigma\left(\frac{\delta\sigma}{\delta x}\right)^2 \right)_0, \quad c_5 = \left(\frac{1}{6}\frac{\delta}{\delta x}\left(\sigma\frac{\delta\sigma}{\delta x}\right) \right)_0 \tag{3.180}$$

$$c_6 = \left(\sigma\frac{\delta a}{\delta x} - \frac{\delta\sigma}{\delta t} - a\frac{\delta\sigma}{\delta x} - \frac{1}{2}\sigma^2\frac{\partial^2\sigma}{\partial x^2} \right)_0. \tag{3.181}$$

Апроксимація (3.175) з коефіцієнтами (3.174), (3.180)-(3.181) має порядок h^4 квадрата середньоквадратичної похибки на одному кроці. Запровадження випадкової величини $z \equiv \int_0^h w^2 dt$ і формування $\bar{x} = (t_0 + h)$ за формулою

$$\begin{aligned} \bar{x} = (t_0 + h) = x_0 + c_1 h + c_2 w + c_3 w^2 + c_4 hw + c_5 w^3 + \\ + c_6 y + c_7 h^2 + c_8 hw^2 + c_9 w^4 + c_{10} wy + c_{11} z \end{aligned} \tag{3.182}$$

дозволяє підібрати коефіцієнти $c_1,\ldots,c_{11}$, які неважко обчислити, використовуючи вищенаведені міркування. При цьому порядок квадрата середньоквадратичної похибки на одному кроці дорівнює h_5.

Зауваження 3.2. У даному пінкті детально не обговорюються всі умови, при яких справедливі вищевикладені міркування. Строге обґрунтування може бути одержане при відповідних умовах, що пов'язані з гладкістю і ростом коефіцієнтів a,σ та їх похідних. Наприклад, умови обмеженості та неперервності коефіцієнтів a,σ та їх частинних похідних по t,x до четвертого порядку включно є достатніми для справедливості формули

$$M(x(t_0+h)-\bar{x}(t_0+h))^2 = o(h^3). \tag{3.183}$$

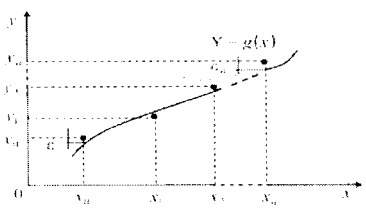
3.7.2. Апроксимація на скінченному проміжку

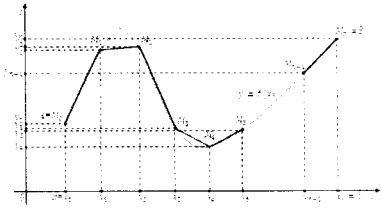
Розглянемо послідовність

$$\bar{x}(t_0)=x_0,\ \ \bar{x}(t_0+(k+1)h)=\bar{x}(t_0+kh)+\bar{\sigma}w_{k+1}+\left(\bar{a}-\frac{1}{2}\bar{\sigma}\frac{\overline{\partial\sigma}}{\partial x}\right)h+\frac{1}{2}\bar{\sigma}\frac{\overline{\partial\sigma}}{\partial x}w_{k+1}^2. \tag{3.184}$$

де $h=\dfrac{T}{m}$; $k=\overline{0,m-1}$; $w_1,\ldots,w_m$ – незалежні за сукупністю випадкові величини, які розподілені нормально з нульовим математичним сподіванням та дисперсією h; риска означає, що функція обчислюється у точці $(t_0+kh,\bar{x}(t_0+kh))$. Позначимо

$$\varepsilon_k^2 = M_{t_0,x_0}(x(t_0+kh)-\bar{x}(t_0+kh))^2. \tag{3.185}$$





Оцінимо ε_{k+1}^2:

$$\varepsilon_{k+1}^2 = M_{t_0,x_0}(x(t_0+(k+1)h)-\bar{x}(t_0+(k+1)h))^2 = \\ = M_{t_0,x_0}\left(M_{t_0+kh,\bar{x}(t_0+kh),\bar{x}(t_0+kh)}\times\varphi_{k+1}^2(h,x(t_0+(k+1)h),w_{k+1}(h);\bar{x}(t_0+kh))\right) \tag{3.186}$$

$$\varphi_{k+1}(h,x,w_{k+1};\bar{x}(t_0+kh)) = x-\bar{x}(t_0+kh)-\left(\bar{a}-\frac{1}{2}\bar{\sigma}\frac{\overline{\partial\sigma}}{\delta x}\right)h-\bar{\sigma}w_{k+1}-\frac{1}{2}\bar{\sigma}\frac{\overline{\partial\sigma}}{\delta x}w_{k+1}^2. \tag{3.187}$$

Скориставшись формулою (3.168), знайдемо

$$\left(M_{t_0+kh,x(t_0+kh),\overline{x}(t_0+kh)}\varphi_{k+1}^2 = (x(t_0+kh)-\overline{x}(t_0+kh)\right)^2 + (L\varphi_{k+1}^2)_0 h + \frac{1}{2}(L^2\varphi_{k+1}^2)_0 h^2 + 0(h^3) \quad (3.188)$$

$$(L\varphi_{k+1}^2)_0 = 2(x_k - \overline{x}_k)(a-\overline{a}) + (\sigma - \sigma^2). \quad (3.189)$$

$$(L^2\varphi_{k+1}^2)_0 = 2(a-\overline{a})^2 + (x_k-\overline{x}_k)\left(\frac{\delta a}{\delta t} + a\frac{\delta a}{\delta x} + \frac{1}{2}\sigma^2\frac{\partial^2 a}{\partial x^2}\right) + 2(\sigma-\sigma^2)\left(\sigma\frac{\delta a}{\delta x} + \frac{\delta\sigma}{\delta t} + a\frac{\delta\sigma}{\delta x} + \frac{1}{2}\sigma^2\frac{\partial^2\sigma}{\partial x^2}\right) + \left(\sigma\frac{\delta\sigma}{\delta x} - \sigma\frac{\overline{\delta\sigma}}{\delta x}\right) \quad (3.190)$$

У формулах (3.189),(3.190) $x_k = x(t_0+kh), \overline{x}_k = \overline{x}(t_0+kh)$; функції без риски обчислюються у точці (t_0+kh, x_k); функції з рискою – у точці $(t_0+kh, \overline{x}_k)$.

Припускаючи обмеженість відповідних похідних від коефіцієнтів x та σ, одержимо з (3.185)-(3.190) наступну нерівність

$$\varepsilon_{k+1}^2 \le \varepsilon_k^2 + A\varepsilon_k^2 h + B\varepsilon_k h^2 + Ch^3. \quad (3.191)$$

A,B,C –деякі додатні сталі. Розглянемо рівність

$$\eta_{k+1}^2 \le \eta_k^2 + A\eta_k^2 h + B\eta_k h^2 + Ch^3,\ \eta_0 = 0. \quad (3.192)$$

Зрозуміло, що $\varepsilon_k^2 \le \eta_k^2$ і, отже, якщо $\eta_m^2 = o(h^2)$, то й $\varepsilon_m^2 = o(h^2)$.

Доведемо, що $\eta_m^2 = o(h^2)$. Зауважимо, що η_k^2 монотонно зростає і що для малих k значення η_k^2 менше, ніж h^2. Якщо це виконується для всіх $k = 0,1,\ldots m-1$, то $\eta_m^2 = o(h^2)$. Нехай $\eta_{k_0}^2 \le h^2$, а для $k > k_0, \eta_{k_0}^2 > h^2$. Для $k > k_0$, з (3.192) одержимо

$$\eta_{k+1}^2 \le \eta_k^2(1+(A+B+C)h). \quad (3.193)$$

$$\eta_{k_0}^2(1+(A+B+C)h)^m \le D\cdot\eta_{k_0}^2 \le Dh^2,\ D = const. \quad (3.194)$$

Таким чином, доведено, що порядок квадрату середньоквадратичної похибка на проміжку $[t_0, t_0+T]$ при використанні однокрокової апроксимації (3.166), (3.174) довіює h^2. Аналогічні міркування показують, що при використанні звичайної апроксимації (3.165) цей порядок дорівнює h. Формули (3.166), (3.174) можуть бути незручними у використанні, що пов'язано з необхідністю обчисляти на кожному кроці похідну $\frac{\delta\sigma}{\delta x}$. Тоді можна скористатися ідеєю методу Рунге-Кутта і шукати однокрокову апроксимацію у вигляді:

$$\overline{x}(t_0+h) = x_0 a_0 h + p_1\sigma_0 w + p_2\sigma\left(t_0, x_0 + a_1\sigma_0 w + a_2\sigma_0\frac{h}{w}\right)w, \quad (3.195)$$

де $a_0 = a(t_0, x_0)$, $\sigma_0 = \sigma(t_0, x_0)$. Якщо $\sigma\left(t_0, x_0 + a_1\sigma_0 w + a_2\sigma_0 \frac{h}{w}\right)w$ замінити виразом $\sigma_0 w + \frac{\delta\sigma}{\delta x}(t_0, x_0)(a_1\sigma_0 w^2 + a_2\sigma_0 h)$ і порівняти (3.195) з (3.166), (3.174), то одержимо наступні співвідношення для коефіцієнтів p_1, p_2, a_1, a_2:

$$p_1 + p_2 = 1,\ p_2 a_1 = -p_2 a_2 = \frac{1}{2}. \tag{3.196}$$

Порядок квадрата середньоквадратичної похибки на одному кроці апроксимації (3.195),(3.196) дорівнює h^3.

3.7.3. Узагальнення на системи рівнянь

Розглянемо систему стохастичних рівнянь

$$dx = a(t,x)dt + \sigma(t,x)dw. \tag{3.197}$$

У (3.197) $x(t)$, $a(t,x)$, $\sigma(t,x)$ – n-вектори, а $w(t)$ – стандартний вінерів процес. Усі вищевикладені результати легко перенести на системи вигляду (3.197). Рекурентні формули апроксимації набувають вигляду:

$$\overline{x}(t_0) = x_0, \tag{3.198}$$

$$\overline{x}(t_0 + (k+1)h) = \overline{x}(t_0 + kh) + \overline{\sigma} w_{k+1} + \left(\overline{a} - \frac{1}{2}\overline{\left(\sigma, \frac{\delta}{\delta x}\right)\sigma}\right)h + \frac{1}{2}\overline{\left(\sigma, \frac{\delta}{\delta x}\right)\sigma} w_{k+1}^2. \tag{3.199}$$

У (3.198)-(3.199) риска означає, що функція обчислюється в точці $\left(t_0 + kh, \overline{x}(t_0 + kh)\right)$; i-та координата вектора $\left(\sigma, \frac{\delta}{\delta x}\right)\sigma$ дорівнює $\sum_{j=1}^{n} \sigma_j \frac{\delta\sigma_i}{\delta x_j}$.

Розглянемо більш загальну схему стохастичних рівнянь

$$dx(t) = a(t,x)dt + \sum_{r=1}^{q} \sigma_r(t,x)dw_r(t). \tag{3.200}$$

Тут x, a, σ_r – n-вектори а $w_r(t)$, r=1,…,q – незалежні вінерові процеси. Для системи (3.200) можна скористатись апроксимацією вигляду

$$\overline{x}(t_0 + h) = x_0 + c_0 h + \sum_{r=1}^{q} c_r w_r + \sum_{i,j=1}^{q} c_{ij} w_i w_j. \tag{3.201}$$

Необхідно зауважити, що в цьому випадку неможливо підібрати вектори c_0, c_r, c_{ij} таким чином, щоб порядок квадрата середньоквадратичної похибки на одному кроці був h^3. Це пов'язано з не комутативністю операторів S_r, які є аналогом оператора

S. Однак, якщо ввести випадкові величини $\left\{ z_{ij} = \int_0^h w_i dw_j \right\}$, то за допомогою випадкового вектора вигляду

$$\overline{x}(t_0 + h) = x_0 + c_0 h + \sum_{r=1}^{q} c_r w_r + \sum_{i,j=1}^{q} c_{ij} z_{ij}, \tag{3.202}$$

Вже можна відшукати такі вектори c_0, c_r, c_{ij}:

$$c_0 = a,\ c_r = \sigma_r,\ c_{ij} = \left(\sigma_i, \frac{\delta}{\delta x} \right) \sigma_j. \tag{3.203}$$

Усі функції в (3.203) обчислюються в точці $(t_0, x(t_0))$.

3.8. Методи побудови емпіричних формул за вибірковими значеннями

3.8.1. Загальна постановка задачі

Нехай, вивчаючи функціональну залежність $y = f(x)$, проведено вимірювання величин x та y і в результаті отримано таблицю значень

x	x_1	x_2	...	x_n
y	y_1	y_2	...	y_n

Якщо аналітичний вигляд функції $f(x)$ невідомий або дуже складний, то виникає практично важлива задача: знайти емпіричну формулу

$$\widetilde{y} = \widetilde{f}(x), \tag{3.204}$$

значення якої мало б відрізнялися від експериментальних даних y_i. В такій постановці наша задача дуже невизначена: тому вказують досить вузький клас функцій F (наприклад лінійні, квадратичні), якому повинна належати шукана функція $\widetilde{f}(x)$, і тоді задача зводиться до знаходження значень найкращих параметрів. В багатьох випадках клас F визначається потребою простоти емпіричної формули; інколи цей клас підказаний самою природою явища. Геометрично задача побудови емпіричної формули полягає в проведенні кривої Γ вигляду (3.204) з деякого класу K, яка «тісно прилягає» до системи точок $M_i(x_i, y_i)$, $i=1,2,...,n$, (рис. 3.1а). При цьому повинен бути введений математичний сенс поняття «тісно прилягає».

Підкреслимо, що задача побудови емпіричної формули відрізняється від задачі інтерполювання. Як правило, початкового матеріалу дуже багато, і є досить проста аналітична залежність між змінними x та y. Ця залежність звичайно не зводиться до інтерполяційних формул (які дають значення, що співпадають з заданими значеннями), так як графік емпіричної функції $\widetilde{y} = \widetilde{f}(x)$ не проходить через систему точок $M_i(x_i, y_i)$, $i=1,2,...,n$. Крім того, ці дані $M_i(x_i, y_i)$, $i=1,2,...,n$, як правило, є наближеними й містять помилки. Тому інтерполяційна формула, яка повторює ці помилки, не кажучи про її складність, не є ідеальним розв'язком поставленої задачі. Можливо, проста емпірична формула згладжує похибки і краще відображає дійсність.

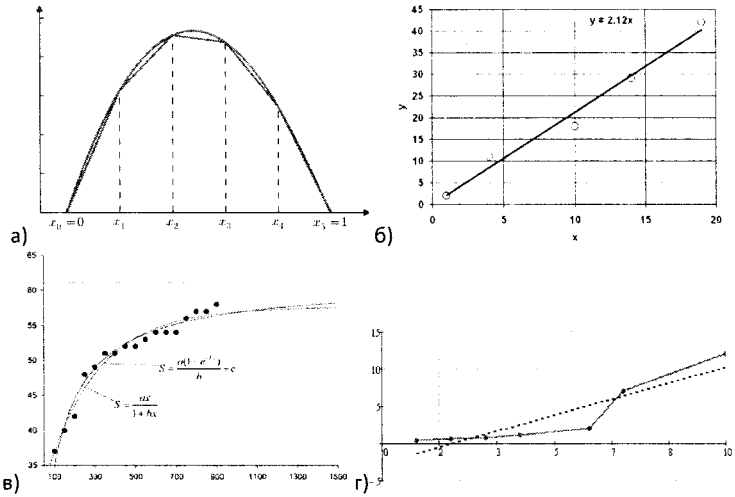


Рис. 3.1. Приклади графіків емпіричної функції $\widetilde{y} = \widetilde{f}(x)$

Побудова емпіричної формули складається з двох етапів:

1) встановлення загального вигляду формули;

2) визначення найкращих її параметрів.

Якщо невідомо характер залежності між невідомими x та y, то вигляд емпіричної формули можна вибрати довільним. Перевага надається простим формулам, які володіють доброю точністю. Якщо відсутні відомості про проміжкові дані, то припускається, що емпірична функція аналітична без точок розриву.

Вдалий підбір емпіричної формули в значній мірі залежить від професійності дослідника-аналітика. Велике значення має геометричне представлення одержаних даних в декартовій або в спеціальній (напівлогарифмічних, логарифмічних) системах координат. При певному досвіді за заданими точками можна наближено вгадати загальний вигляд залежності між змінними.

У багатьох випадках можна обмежитися поліномом

$$y = \sum_{i=0}^{m} a_i x^i \,. \tag{3.205}$$

Часто використовують інші елементарні функції (дробово-лінійну, степеневу, показникову, логарифмічну,...). Далі будуть наведено приклади, які полегшують вибір вигляду емпіричної формули. Що стосується найкращих значень параметрів, які входять до формули, то ця задача є більш легкою і розв'язується відомими методами.

Емпіричні формули не претендують на роль законів природи, а є лише гіпотезами, які більш чи менш задовільняють досліджувані дані. Однак, вони мають дуже велике значення. В історії науки відомі приклади, як визначення важливої емпіричної формули призводило до важливих наукових відкриттів.

Зауваження 3.3. При побудові емпіричної формули можна припусти, що вихідні дані (x_i, y_i) додатні. Якщо це не так: наприклад всі $x_i < 0$, то достатньо

розглянути таблицю значень $(-x_i, y_i)$. Розглянемо загальний випадок, коли знак значень x_i змінний. Тоді очевидно, що можна підібрати такі значення m та n, що

$$\xi_i = x_i + m > 0, \tag{3.206}$$

$$\eta_i = y_i + n > 0. \tag{3.207}$$

Звідси отримаємо, що поставлену задачу можна вирішити для набору значень (ξ_i, η_i). Тому надалі будемо розглядати випадок таблиці з додатними елементами.

Зауважимо, що надалі не будемо торкатися питань, які пов'язані з оцінкою якості вихідних даних. Зокрема, не будемо враховувати випадкові похибки цих даних.

3.8.2. Лінійна залежність

Нехай для невідомих x та y відомі їх значення (x_i,y_i), $i=1,2,..,n$, які розміщені в порядку зростання першої змінної $0 < x_1 < x_2 < .. < x_n$. На координатній площині xOy побудуємо систему точок $M_i(x_i, y_i)$, $i=1,2,..,n$ (для зручності масштаби по осях Ox і Oy виберемо різними). Якщо виявиться, що ці точки розміщуються приблизно на деякій прямій L, то природньо припустити, що залежність між x та y є лінійна:

$$y = ax + b, \tag{3.208}$$

де a та b - сталі. Для знаходження параметрів a та b вимірюються координати двох точок, які лежать на прямій $(N_1(\tilde{x}_1, \tilde{y}_1)$ та $N_2(\tilde{x}_2, \tilde{y}_2))$. Тоді параметри a та b можуть бути визначені з системи рівнянь (спосіб вибраних точок):

$$\begin{cases} \tilde{y}_1 = a\tilde{x}_1 + b, \\ \tilde{y}_2 = a\tilde{x}_2 + b. \end{cases} \tag{3.209}$$

Результат виключення коефіцієнтів a та b може бути записаний у вигляді визначника

$$\begin{vmatrix} x & y & 1 \\ \tilde{x}_1 & \tilde{y}_1 & 1 \\ \tilde{x}_2 & \tilde{y}_2 & 1 \end{vmatrix} = 0. \tag{3.210}$$

Геометрично рівняння (3.209) – це рівняння прямої, яка проходить через точки $N_1(\tilde{x}_1, \tilde{y}_1)$ та $N_2(\tilde{x}_2, \tilde{y}_2)$.

Приклад 3.3. Підібрати емпіричну формулу для табличних даних

x	2	4	5	8
y	0.350	0.573	0.725	0.947

Розв'язок. Побудувавши відповідні точки $M_i(x_i,y_i)$, $i=1,2,3,4$, в зручному масштабі, переконуємося, що вони лежать на одній лінії $y=ax+b$ (рис. 3.1б). По графіку знаходимо відрізок, який відсікає пряма на осі ординат b=0.147. Крім того, при $x = 10$ маємо $y = 1.138$. Тому $a = \dfrac{1.138 - 0.147}{10} = 0.099$. Можна прийняти

$$\tilde{y} = 0.099x + 0.147. \tag{3.211}$$

Таблиця 3.3

Значення емпіричної функції $\widetilde{y}$

x	2	4	6	8
$\widetilde{y}$	0,345	0,543	0,741	0,939
$y-\widetilde{y}$	0,005	0,05	–0,016	0,008

Легко задати аналітичний критерій для прямолінійності ряду точок $M_i(x_i,y_i)$. Покладемо

$$\Delta x_i = x_{i+1} - x_i, \quad \Delta y_i = y_{i+1} - y_i. \tag{3.212}$$

$$k_i = \frac{\Delta x_i}{\Delta y_i}, \quad i=1,2,..,n-1. \tag{3.213}$$

Якщо $k_i = const i$, то точки $M_i(x_i,y_i)$ точно лежать на одній прямій. Якщо

$$k_1 \approx k_2 \approx ... \approx k_{n-1}, \tag{3.214}$$

то точки $M_i(x_i,y_i)$ приблизно лежать на одній прямій. В залежності від виконання (3.214) вирішується питання чи потрібно шукати емпіричну функцію у лінійному представленні чи ні. Зокрема, якщо значення x_i рівновіддалені, тобто $\Delta x_i = const$, то достатньо пересвідчитись, що значення Δy_i також є постійні (або майже постійні).

Приклад 3.4. Перевірити на прямолінійність наступний ряд точок

i	1	2	3	4	5	6
x_i	0,5	1	1,5	2	2,5	3
y_i	0,62	1,64	2,58	3,7	5,02	6,04

Розв'язок. Так, як значення x_i – рівновіддалені, то складемо піблицю Δy_i:

Таблиця 3.4. Перевірка на прямолінійність

i	1	2	3	4	5
Δy_i	1,02	0,94	1,12	1,32	1,02

Як бачимо, значення Δy_i досить близькі між собою, тому як наближену емпіричну формулу можна вибрати лінійну.

3.8.3. Метод вирівнювання

Нехай для змінних x та y їх відповідні значення (x_i,y_i), $i=1,2,...,n$, такі, що точки $M_i(x_i,y_i)$ не розміщені на прямій лінії. Тоді, в багатьох випадках, вводять нові змінні

$$X=\varphi(x,y); \quad Y=\phi(x,y). \tag{3.215}$$

Можна добитися того, що утворенні точки $N_i(X_i, Y_i)$, де $X_i = \varphi(x_i,y_i)$ і $Y_i=\phi(x_i,y_i)$, лежать на деякій прямій лінії площини OXY *(метод вирівнювання)*. Обов'язковою вимогою при цьому є відповідна однозначність перетворення (3.215).

Розглянемо, наприклад, нелінійну залежність вигляду

$$\varphi(y)=a\,\varphi(x)+b, \tag{3.216}$$

де a і b – сталі, $\varphi(x)$ і $\phi(y)$ – строго монотонні функції. На площині OXY функція (3.216) є відображенням деякої кривої.

Вводячи нові змінні $X=\varphi(x)$ і $Y=\phi(y)$, маємо

$$Y = aX + b \tag{3.217}$$

і, отже, при наявності залежності (3.216) точки $N_i(\varphi(x_i),\ \phi(y_i))$, i=1,2,...,n, на новій координатній площині OXY розміщені на прямій лінії. Навпаки, якщо при побудові на площині OXY виявиться, що точки N_i практично лежать на прямій лінії, то між змінними x і y має місце залежність (3.216). Функції $\varphi(x)$ та $\phi(y)$, зазвичай, знаходяться методом проб на основі знайденої схожості лінії (3.216) з відомими кривими або перевіркою виконання відповідних аналітичних критеріїв.

До вигляду (3.216) зводиться, наприклад, степенева залежність

$$y = cx^a, \tag{3.218}$$

де a і c – сталі, причому $x > 0$ і $y > 0$. Логарифмуючи формулу (3.218), отримаємо:

$$\lg y = a \lg x + \lg c,\ X = \lg x;\ \ Y = \lg y,\ Y = aX + b, \tag{3.219}$$

де $b = \lg c$. Таким чином, степенева залежність (3.218) між змінними x і y знаходиться, якщо точки $N_i(\lg x_i,\ \lg y_i)$, i = 1,2,...,n, лежать на одній прямій в площині OXY.

Зауважимо, що в логарифмічній координатній сітці OXY початком координат є точка x=1 і y=1. Тому вихідні дані x_i і y_i вигідно змінити так, щоб мало місце співвідношення $x_i \geq 1$ і $y_i \geq 1$. Якщо покласти x = 1 в співвідношенні (3.218), то будемо мати y=c. Таким чином, стала c являє собою ординату точки перетину з віссю OY відповідного прямолінійного графіка в площині OXY.

Приклад 3.5. Табличні дані відповідають формулі вигляду

$$y = cx^a. \tag{3.220}$$

Знайти параметри a та c.

x	10	20	30	40	50	60	70	80
y	1.06	1.33	1.52	1.68	1.81	1.91	2.01	2.11

Розв'язок. Наносячи ці точки на логарифмічний папір, побачимо, що вони з достатньою точністю розміщені на прямій лінії (рис. 3.2). Тим самим підтвердивши наше припущення (3.220) про характер залежності між x та y. Для визначення параметрів a і c скористаємось крайніми точками A(10;1,06) і B(80;2,11).

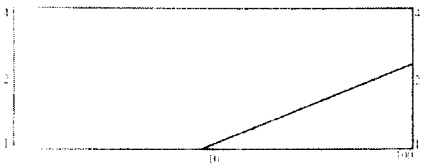


Рис. 3.2. Графік емпіричної функції $\tilde{y} = \tilde{f}(x)$ для прикладу 3.5

Логарифмуючи співвідношення (3.220), отримаємо $\lg y = a \lg x + \lg c$. Звідси, підставляючи координати точок A і B, знаходимо $\lg 1{,}06 = \lg c + a \lg 10$, а потім

$$\lg 2{,}11 = \lg c + a \lg 80$$

$$\left.\begin{aligned} \lg c + a &= 0.0253, \\ \lg c + 1.9031a &= 0.3243; \end{aligned}\right\} \text{ та } a = \frac{0.3243 - 0.0253}{1{,}9031 - 1} = \frac{0{,}2990}{0.9031} = 0.3311$$

$$\lg c = 0{,}0253 - 0{,}3311 = -0{,}3058 = 1{,}6942;\ c = 0{,}4945$$

$$\tilde{y} = 0.4945x^{0.3311}. \tag{3.221}$$

В таблиці 3.5 наведені значення $\tilde{y} = \tilde{y}(x)$, що обчислені за формулою (3.221), а також їх відхилення $\Delta = y - \tilde{y}$ від табличних даних.

Таблиця 3.5

Значення $\tilde{y} = \tilde{y}(x)$ та відхилення $\Delta = y - \tilde{y}$

x	10	20	30	40	50	60	70	80
$\tilde{y}$	1.060	1.320	1.525	1.677	1.806	1.919	2.012	2.110
$\Delta = y - \tilde{y}$	0	0.010	-0.005	0.003	0.004	-0.009	-0.002	0

Інший важливим випадком є показникова залежність

$$y = ce^{ax}, \tag{3.222}$$

де $c > 0$. Логарифмуючи, отримаємо $\lg y = ax + \lg c$, $X = x$, $Y = \lg y$, $Y = aX + b$, де $b = \lg c$.

Якщо є підозра на показникову залежність (3.222), то для побудови в площині OXY добре використовувати напівлогарифмічний папір з рівномірною шкалою по осі OX і логарифмічною шкалою по осі OY. Початок координат в напівлогарифмічній сітці відповідає точці $x = 0$ та $y = 1$; тому краще брати $y_i \geq 1$. Для прямолінійного графіка в площині OXY, що відповідає формулі (3.222), стала c представляє собою ординату точки перетину цього графіка з віссю OY ($x = 0$, $y = c$).

Приклад 3.6. Результати вимірювання представлені таблицею:

x	0	0.2	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
y	1	0.833	0.667	0.540	0.400	0.33	0.286	0.25	0.222	0.20	0.182	0.167

Скласти емпіричну формулу.

Розв'язок. Наносимо точки $M_i(x_i, y_i)$ на координатну сітку OXY і бачимо, що вони розміщуються приблизно на кривій, яка схожа на гілку гіперболи, асимптотично наближену до осі Ox (рис. 3.3а). Тому вважаємо, що

$$y = \frac{1}{ax + b} \text{ або } \frac{1}{y} = ax + b. \tag{3.223}$$

тобто отримуємо залежність вигляду (3.216). Ввівши нові змінні $X = x$; $Y = \frac{1}{y}$, отримуємо лінійну залежність:

$$Y = aX + b. \tag{3.224}$$

Відповідні дані наведені в табл. 3.6.

Таблиця 3.6.

X	0	0.2	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Y	1	1.2	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6

Побудувавши в площині OXY точки $N_i(X_i, Y_i)$, переконуємось, що вони лежать приблизно на прямій лінії L (рис. 3.3б). Отже, гіпотеза про характер залежності виправдалась, і вигляд формули (3.223) вибрано правильно. За рис. 3.3б визначаємо параметри прямої L: $b = 1$; $a = \frac{6-1}{5-0} = 1$.

Таким чином, отримали шукану емпіричну формулу $y = \frac{1}{x+1}$.

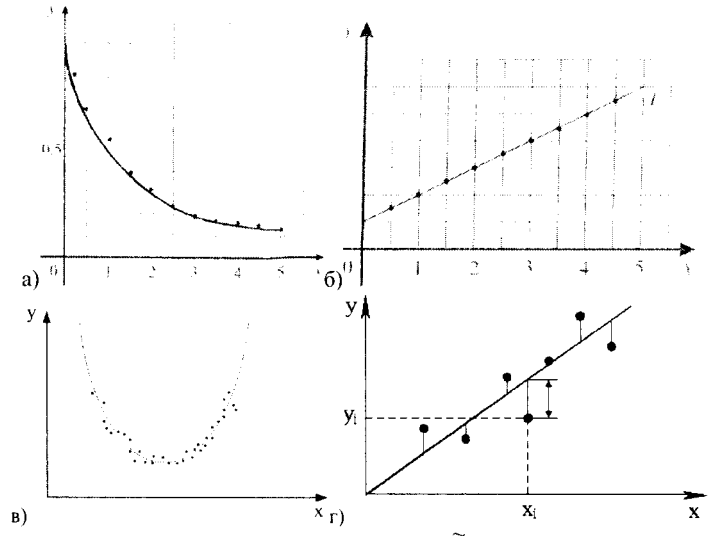


Рис. 3.3. Графіки емпіричної функції $\tilde{y} = \tilde{f}(x)$ для прикладу 3.6

3.8.4. Квадратична (параболічна) залежність

Якщо для цих значень (x_i, y_i), i=1,2,...,n, не справджується лінійна залежність, то можна спробувати більш загальну квадратичну залежність

$$y = ax^2 + bx + c, \quad a \neq 0. \tag{3.225}$$

Формула (3.225) має місце, коли точки $M_i(x_i, y_i)$ розміщенні на відрізку параболи з вертикальною віссю (рис. 3.4а). Для підтвердження цього факту з відповідною точністю проводимо криву Γ, поблизу якої групуються дані точки $M_i(x_i, y_i)$ і вибираємо на ній точку $N(x_0, y_0)$, по можливості, що співпадає з однією з точок M_k (x_k, y_k), $1 \le k \le n$. Припустимо, що крива Γ є параболою (3.225), отримаємо

$$y_0 = ax_0^2 + bx_0 + c. \tag{3.226}$$

Віднімаючи з рівняння (3.225) рівність (3.226), знаходимо

$$y - y_0 = a(x^2 - x_0^2) + b(x - x_0), \text{ або } y - y_0 = a(x - x_0)^2 + b_1(x - x_0), \tag{3.227}$$

де $b_1 = b + 2ax_0$. Якщо тепер ввести нові змінні

$$X = x - x_0, \quad Y = \frac{y - y_0}{x - x_0}, \tag{3.228}$$

то з рівняння (3.227) отримаємо лінійну залежність

$$Y = aX + b_1. \tag{3.229}$$

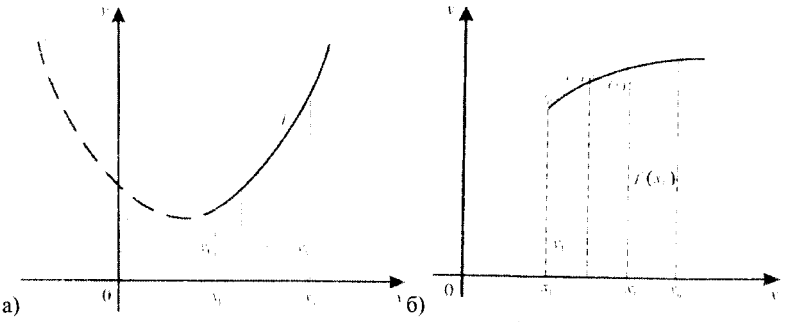


Рис. 3.4. Приклади графіків емпіричної функції $\widetilde{y} = \widetilde{f}(x)$ параболічної залежності

Таким чином, параболі (3.225) на площині OXY відповідає пряма (3.229) на площині OXY. Навпаки, якщо буде виявлено, що точки $N_i(X_i, Y_i)$, $i = \overline{1, n-1}$

$$X_i = x_i - x_0, \quad Y_i = \frac{y_i - y_0}{x_i - x_0}, \tag{3.230}$$

лежать на прямій (3.229), то точки $M_i\ (x_i y_i)$ розміщені на параболі (3.225), причому

$$\left.\begin{aligned} b &= b_1 - 2ax_0, \\ c &= y_0 + ax_0^2 - b_1 x_0. \end{aligned}\right\} \tag{3.231}$$

Введемо аналітичний критерій для квадратичної залежності (3.225). Нехай $M_i\ (x_i, y_i)$ – дані таблиці значень, причому $\Delta x_i = x_{i+1} - x_i > 0$, $i = \overline{1, n-1}$. При наявності залежності (3.225) послідовність $y_1, y_2, ..., y_n$ – або монотонна, тобто $\Delta y_i = y_{i+1} - y_i$, $i = \overline{1, n-1}$ зберігає сталий знак; або ж ця послідовність має єдиний екстремум, тобто різниця Δy_i лише один раз змінює знак.

Введемо поділені різниці першого і другого порядків

$$y[x_i; x_{i+1}] \equiv \frac{\Delta y_i}{\Delta x_i}, \quad y[x_i; x_{i+1}; x_{i+2}] \equiv \frac{y[x_{i+1}; x_{i+2}] - y[x_i; x_{i+1}]}{x_{i+2} - x_i} \tag{3.232}$$

$$y[x_i; x_{i+1}; x_{i+2}] = \frac{\Delta\left(\dfrac{\Delta y_i}{\Delta x_i}\right)}{\Delta_1 x_i}, \tag{3.233}$$

де $\Delta_1 x_i = x_{i+2} - x_i = \Delta x_i + \Delta x_{i+1}$. Точки $M_i(x_i, y_i)$ розміщені на параболі (3.225) тоді і тільки тоді, коли зберігають постійні значення поділені різниці другого порядку. Зокрема, якщо значення $x_1, x_2, ..., x_n$ рівновіддалені, тобто $\Delta x_i = h = const$, то для існування емпіричної квадратичної залежності (3.225) необхідно і достатньо, щоб була постійна друга різниця $\Delta^2 y_i \equiv y_{i+2} - 2y_{i+1} + y_i \equiv c$, $i = \overline{1, n-1}$, тоді $\Delta^2 y_i = 2h^2 a$. Зауважимо, що друга різниця досить чутлива до відхилення від параболічного закону.

Приклад 3.7. У таблиці 3.7 подана система значень. Дослідити на квадратичну залежність.

Таблиця 3.7

Вхідні дані для прикладу 3.7

X	0	0.5	1	2	4	5	6
Y	0	-1.76	-3.00	-3.96	0.24	5.40	13.34

Розв'язок. Складемо за формулою (3.233) таблицю відповідних різниць (табл. 3.8). Так як друга поділена різниця приблизно постійна, то можна вважати, що між змінними x та y має місце наближена квадратична залежність.

Таблиця 3.8

Таблиця відповідни різниць для вхідних даних прикладу 3.7

x	Δx	y	Δy	$\frac{\Delta y}{\Delta x}$	$\Delta\left(\frac{\Delta y}{\Delta x}\right)$	$\Delta_1 x$	$\frac{\Delta\left(\frac{\Delta y}{\Delta x}\right)}{\Delta_1 x}$
0		0					
	0,5		-1,76	-3,52			
0,5		-1,76			1,04	1	1,04
	0,5		-1,24	-2,48			
1		-3,00			1,52	1,5	1,01
	1		-0,96	-0,96			
2		-3,96			2,92	3	0,97
	2		3,72	1,86			
4		-0,24			3,78	3	1,26
	1		5,64	5,64			
5		5,40			2,30	2	1,15
	1		7,94	7,94			
6		13,34					

3.8.5. Визначення параметрів емпіричної формули

Якщо вибрано вигляд емпіричної формули, то виникає задача визначення найкращих коефіцієнтів (параметрів), які входять в цю формулу. В загальному випадку ця задача ставиться наступним чином: нехай задана система значень M_i (x_i, y_i), i=1,2, ... n, наближено описується формулою вигляду

$$y = \tilde{f}(x; a_1, a_2, ..., a_m), \tag{3.234}$$

де $\tilde{f}$ – відома функція і $a_1, a_2, ..., a_m$ – невідомі сталі, кількість яких m зазвичай менша числа точок M_i, тобто $m < n$. Необхідно визначення цих сталих.

Якщо значення (x_i, y_i) точно зв'язані залежністю (3.234), то параметри $a_1, a_2, ..., a_m$ можуть бути знайдені з системи рівнянь

$$y_j = \tilde{f}(x_j; a_1, a_2, ..., a_m), \ j = \overline{1, n}. \tag{3.235}$$

Однак на практиці значення (x_i, y_i) мають неусувні похибки і кількість рівнянь системи (3.235) більша ніж кількість невідомих; тому система (3.235), як правило, є

несумісною. Доводиться шукати найкращі значення $\bar{a}_1, \bar{a}_2, ..., \bar{a}_m$, наближено задовільняючи систему (3.235), тобто такі, що відхилення

$$y_j - \tilde{f}(x_j; \bar{a}_1, \bar{a}_2, ..., \bar{a}_m) = s_j, \quad j = \overline{1,n}. \tag{3.236}$$

є можливо малими за абсолютною величиною.

Геометрична задача зводиться до проведення кривої вигляду (3.234), яка найбільш тісно прилягає до даної системи точок. Найбільш поширеними є емпіричні формули, лінійно залежні від параметрів, тобто формули вигляду

$$y = \varphi_0(x) + a_1\varphi_1(x) + a_2\varphi_2(x) + ... + a_m\varphi_m(x). \tag{3.237}$$

У цьому випадку система (3.234) лінійна і дослідження її порівняно просте. При нелінійній залежності в (3.234) від параметрів $a_1, a_2, ..., a_m$ система (3.235) також нелінійна і знаходження точних або наближених розв'язків є складною задачею. Тоді таку систему наближено замінюють лінійною. У наступних параграфах розглянемо три найбільш вживані методи визначення параметрів емпіричних формул: 1) метод вибраних точок; 2) метод середніх; 3) метод найменших квадратів.

3.8.6. Метод вибраних точок

Нехай для системи експериментальних даних M_i (x_i, y_i), $i = \overline{1,n}$, побудована емпірична формула

$$y = \tilde{f}(x; a_1, a_2, ..., a_m). \tag{3.238}$$

яка містить $m (m \leq n)$ вільних параметрів $a_1, a_2, ..., a_m$, де $\tilde{f}$ – відома функція.

На координатній площині XOY з можливою обережністю проводимо криву Γ, найбільш близьку до точок M_i.

На кривій Γ вибираємо систему m (за кількістю параметрів) точок $N_j(\tilde{x}_j, \tilde{y}_j)$, $j = \overline{1,m}$, не обов'язково співпадаючу з точками M_i. При цьому бажано, щоб вибрані точки N_j були по можливості рівномірно розподілені по всій робочій області кривої Γ і можливо далі розміщались одна від одної, і в той же час не знаходились дуже близько до мало надійних кінцевих точок M_1 і M_n. Для зручності беруть абсцису $\tilde{x}_j$ цих точок, що співпадає з відповідними значеннями осі Ox координатної сітки. Після цього з усією ретельністю замірюють координати $\tilde{x}_j, \tilde{y}_j$, $j = \overline{1,n}$. Тоді параметри $a_1, a_2, ..., a_m$, в загальному випадку, можуть бути визначені з системи m рівнянь

$$\tilde{y}_j = \tilde{f}_j(\tilde{x}_j; a_1, a_2, ..., a_m), \quad j = \overline{1,n}. \tag{3.239}$$

Застосування цього методу для лінійної залежності $y = ax + b$ – в попередніх розділах. Для випадку квадратичної залежності $y = ax^2 + bx + c$ коефіцієнти a, b і c визначаються з системи трьох рівнянь

$$\left.\begin{aligned}\widetilde{y}_1 &= a\widetilde{x}_1^2 + b\widetilde{x}_1 + c,\\ \widetilde{y}_2 &= a\widetilde{x}_2^2 + b\widetilde{x}_2 + c,\\ \widetilde{y}_3 &= a\widetilde{x}_3^2 + b\widetilde{x}_3 + c.\end{aligned}\right\}. \tag{3.240}$$

Зауважимо, що метод вибраних точок містить геометричні побудови, які допускають певну довільність, і тому є грубим. Його слід застосовувати в тих випадках, коли точність вихідних даних відносно невелика. Для збільшення точності метода рекомендується користуватися сіткою з дрібним поділом. Достоїнство методу – простота використання й наочність. У наступних параграфах вкажемо більш точний аналітичний метод визначення коефіцієнтів емпіричної формули (3.238).

3.8.7. Метод середніх

Якщо в емпіричну формулу

$$y = \widetilde{f}(x; a_1, a_2, ..., a_m). \tag{3.241}$$

підставити дані $M_i(x_i, y_i)$, то ліва частина формули, взагалі кажучи, не буде рівна правій. Різниця (нев'язка)

$$\widetilde{f}(x; a_1, a_2, ..., a_m) - y_i = \varepsilon_i,\ i = \overline{1,n}, \tag{3.242}$$

називаються відхиленнями і представляють собою відстань по вертикалі від точок M_i (x_i, y_i) до графіка емпіричної функції (3.241), які взяті зі знаком (+) або (-) (рис. 3.4б).

Згідно з методом середніх за найкраще положення емпіричної кривої приймається те, для якого рівна нулеві алгебраїчна сума E всіх відхилень ε_i, тобто повинна виконуватися рівність

$$E = \sum_{i=1}^{m} \varepsilon_i = 0. \tag{3.243}$$

Для визначення за методом середніх сталих $a_1, a_2, ..., a_m$, де $m \le n$, всі відхилення ε_i розбиваються на m груп, які місгять приблизно одинакову кількість відхилень. Прирівнюючи до 0 алгебраїчну суму E_j, $j = \overline{1,m}$, відхилень, які входять в кожну з даних груп одержуємо систему, яка містить стільки ж рівнянь, скільки і невідомих коефіцієнтів. Розв'язавши цю систему, знайдемо коефіцієнти $a_1, a_2, ..., a_m$. Зауважимо, що так як сума E_j відхилень для кожної групи рівна 0, то рівна 0 також і сума всіх відхилень (E=0 всіх відхилень), тобто для нашої системи рівність (3.243) виконується.

Приклад 3.8. Кількість Q речовини в %, яка залишається в системі через t хвилин від початку хімічної реакції, задається таблицею:

t	7	12	17	22	27	32	37
Q	83,7	72,9	63,2	54,7	47,5	41,4	36,3

Скласти емпіричну формулу залежності величини від часу t.

Розв'язок. На площині OtQ задана система точок (t_i, Q_i), яка приблизно розміщується на параболі з вертикальною віссю. Тому будемо шукати емпіричну формулу у вигляді

$$Q = at^2 + bt + c. \tag{3.244}$$

Для визначення коефіцієнтів a, b і c застосуємо метод середніх. Підставляючи табличні дані в формулу (3.244), отримуємо рівняння для відповідних відхилень:

$$\varepsilon_1 = 49a + 7b + c - 83.7; \qquad \varepsilon_2 = 144a + 12b + c - 72.9;$$
$$\varepsilon_3 = 289a + 17b + c - 63.2; \qquad \varepsilon_4 = 484a + 22b + c - 54.7;$$
$$\varepsilon_5 = 729a + 27b + c - 47.5; \qquad \varepsilon_6 = 1024a + 32b + c - 41.4;$$
$$\varepsilon_7 = 1369a + 37b + c - 36.3.$$

Для визначення коефіцієнтів a, b і c за методом середніх відхилень ε_i потрібно розбити на три групи. Об'єднаємо, наприклад, в групу 1 відхилення ε_1, ε_2, ε_3, в групу 2 – ε_4 та ε_5, а в групу 3 – ε_6 та ε_7. Тоді отримаємо систему

$$\begin{cases} \varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 0 \\ \varepsilon_4 + \varepsilon_5 = 0 \\ \varepsilon_6 + \varepsilon_7 = 0 \end{cases} \tag{3.245}$$

$$\begin{cases} 482a + 36b + 3c = 219.8 \\ 1213a + 49b + 2c = 102.2 \\ 239a + 69b + 2c = 77.7 \end{cases}. \tag{3.246}$$

Розв'язком системи (3.246), будуть a=0.0235; b=–2.6115; c=100.8295. Звідси одержимо, що шукана емпірична формула має вигляд

$$\widetilde{Q} = 0.0235t^2 - 2.6115t + 100.8295. \tag{3.247}$$

Для порівняння в таблиці 3.9 наведені результати розрахункіи за формулою (3.247). Результат методу середніх залежить від способу групування відхилень. Найбільш добрі емпіричні формули одержуються, якщо відхилення групуються в порядку їх номерів індексів, і кожна група містить, по можливості, однакове число членів.

Таблиця 3.9

Оцінка точності емпіричної формули (3.247)

i	Q_i	$\widetilde{Q}_i$	$\varepsilon_i = Q_i - \widetilde{Q}_i$
1	83,7	83,7	0
2	72,9	72,88	0,02
3	63,2	63,23	-0,03
4	54,7	54,75	-0,05
5	47,5	47,45	0,05
6	41,4	41,33	0,07
7	36,3	36,38	0,08

3.8.8. Метод найменших квадратів

Нехай відомо вигляд емпіричної функції

$$y = \tilde{f}(x;a_1,a_2,...,a_m). \tag{3.248}$$

$$\varepsilon_i = f(x;a_1,a_2,...,a_m) - y_i,\ i = \overline{1,n}, \tag{3.249}$$

відхилення емпіричної формули (3.248) від вихідних даних (x_i, y_i).

Відповідно до методу найменших квадратів найкращими коефіцієнтами $a_1,a_2,...,a_m$ вважаються ті, для яких сума квадратів відхилень

$$S(a_1,a_2,...,a_m) = \sum_{i=1}^{n}[f(x;a_1,a_2,...,a_m) - y_i]^2. \tag{3.250}$$

буде найменшою. Звідси, використовуючи необхідні умови екстремуму функції багатьох змінних, одержуємо, так звану, нормальну систему для визначення коефіцієнтів $a_1,a_2,...,a_m$

$$\frac{\partial S}{\partial a_1} = 0,\ \frac{\partial S}{\partial a_2} = 0,\ ...,\ \frac{\partial S}{\partial a_m} = 0. \tag{3.251}$$

Якщо система (3.250) має єдиний розв'язок, то він буде відповідно знайденим розв'язком. Система (3.251) спрощується, якщо емпірична функція $\tilde{f}(x;a_1,a_2,...,a_m)$ лінійна відносно параметрів $a_1,a_2,...,a_m$. При

$$\tilde{f}(x;a_1,a_2,...,a_m) = \varphi_0(x) + a_1\varphi_1(x) + a_2\varphi_2(x) + ... + a_m\varphi_m(x), \tag{3.252}$$

будемо мати $S(a_1,a_2,...,a_m) = \sum_{i=1}^{n}[a_1\varphi_1(x_i) + a_2\varphi_2(x_i) + ... + a_m\varphi_m(x_i) - Y_i]^2$ при

$$Y_i = y_i - \varphi_0(x_i). \tag{3.253}$$

$$\begin{cases} \frac{1}{2}\frac{\partial S}{\partial a_1} = \sum_{i=1}^{n}\varphi_1(x_i)[a_1\varphi_1(x_i) + a_2\varphi_2(x_i) + ... + a_m\varphi_m(x_i) - Y_i] = 0 \\ ..... \\ \frac{1}{2}\frac{\partial S}{\partial a_m} = \sum_{i=1}^{n}\varphi_m(x_i)[a_1\varphi_1(x_i) + a_2\varphi_2(x_i) + ... + a_m\varphi_m(x_i) - Y_i] = 0 \end{cases}. \tag{3.254}$$

Вводимо позначення

$$(\varphi_j,\varphi_k) \equiv \sum_{i=1}^{n}\varphi_j(x_i)\varphi_k(x_i) \text{ і } (\varphi_j,Y) \equiv \sum_{i=1}^{n}\varphi_j(x_i)Y_i. \tag{3.255}$$

Систему (3.254) можна записати у вигляді нормальної системи

$$\begin{cases} a_1(\varphi_1,\varphi_1)+a_1(\varphi_1,\varphi_2)+\ldots+a_1(\varphi_1,\varphi_m)=(\varphi_1,Y) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_m(\varphi_m,\varphi_1)+a_m(\varphi_m,\varphi_2)+\ldots+a_m(\varphi_m,\varphi_m)=(\varphi_m,Y) \end{cases} \tag{3.256}$$

Теорема 3.15. Параметри $\{a_1,a_2,\ldots,a_m\}\subset \mathrm{R}^1$ емпіричної формули (3.248) у випадку лінійності (3.252) за методом найменших квадратів визначаються з системи лінійних алгебраїчних рівнянь (3.256). В частковому випадку, якщо емпірична функція є поліномом

$$y=a_0+a_1x+\ldots+a_mx^m. \tag{3.257}$$

$$\varphi_j(x^j),\ j=\overline{0,m}. \tag{3.258}$$

$$(\varphi_j,\varphi_k)=\sum_{i=1}^{n}x_i^{j+k}\equiv[x^{j+k}]. \tag{3.259}$$

$$(\varphi_j,Y)=(\varphi_j,y)=\sum_{i=1}^{n}x_j^iy_i\equiv[x^jy]. \tag{3.260}$$

Тобто нормальна система (3.256) буде мати вигляд

$$\begin{cases} a_0n+a_1[x]+a_2[x^2]+\ldots+a_m[x^m]=[y] \\ a_0[x]+a_1[x^2]+a_2[x^3]+\ldots+a_m[x^{m+1}]=[xy] \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0[x^m]+a_1[x^{m+1}]+a_2[x^{m+2}]+\ldots+a_m[x^{2m}]=[x^my] \end{cases} \tag{3.261}$$

Наслідок 3.6. У випадку емпіричної функції вигляду полінома (3.257) параметри $\{\{a_1,a_2,\ldots,a_m\}\subset \mathrm{R}^1$ визначаються з системи лінійних алгебраїчних рівнянь (3.261). Метод найменших квадратів має ту перевагу, що якщо сума S квадратів відхилень ε_i мала, то самі відхилення також малі за абсолютною величиною. Для методу середніх, де утворюється алгебраїчна сума відхилень, такого висновку зробити не можна. Недоліком методу найменших квадратів є громіздкість обчислень. Тому його застосовують при опрацюванні, коли треба отримати досить точні значення параметрів. В цьому випадку всі обчислення потрібно проводити з потрібною кількістю десяткових знаків, бо коефіцієнти можуть сильно відрізнятися від дійсних значень. Але грубі наближення можуть бути отримані значно простіше, тобто застосування методу не буде виправдано. Наприклад, якщо виникає втрата цифр при відніманні, то обчислення повинні бути приведені з достатньою кількістю запасних правильних значущих цифр. Тут можна користуватися наступним правилом: якщо числові значення коефіцієнтів бажано мати з m вірними значущими цифрами, і якщо попередні обчислення показують, що перші p цифр зникнуть при обчисленні, то обчислення повинні бути проведені з $m+p+1$ вірними значущими цифрами на всіх стадіях роботи.

Приклад 3.9. Використовуючи метод найменших квадратів, вивести емпіричну формулу для табличної функції $Q = f(t)$. Покладемо

$$y = ax^2 + bx + c. \tag{3.262}$$

Для обчислення нормальної системи складаємо таблицю 3.10. Тоді будемо мати систему нормальних рівнянь:

$$\begin{cases} 4088a + 154b + 7c = 399.7 \\ 120736a + 4088b + 154c = 7688.9 \\ 3795092a + 120736b + 4088c = 186054.3 \end{cases}. \tag{3.263}$$

Розв'язок. Розв'язавши цю систему отримаємо a = 0.023381, b = –2.6066, c = 100.791.

Таблиця 3.10

Визначення параметрів згідно (3.262) методом найменших квадратів

t°	t	t^2	t^3	t^4	Q	Qt	Qt^2
1	7	49	343	2401	83,7	585,9	4101,3
1	12	144	1728	20736	72,9	874,8	10497,6
1	17	289	4913	83521	63,2	1074,4	18264,8
1	22	484	10648	234256	54,7	1203,4	26474,8
1	27	729	19683	531441	47,5	1282,5	34627,5
1	32	1024	32768	1048576	41,4	1324,8	42393,6
1	37	1369	50653	1874161	36,3	1343,1	49694,7
7	154	4088	120736	3795092	399,7	7688,9	186054

Звідки шукана емпірична формула буде мати вигляд

$$\widetilde{Q} = 0.02338t^2 - 2.60066t + 100.79. \tag{3.264}$$

Таблиця 3.11

Оцінка точності емпіричної формули (3.264)

t	Q	$\widetilde{Q}$	$\varepsilon = Q - \widetilde{Q}$
1	83,7	83,69	0,01
2	72,9	72,88	0,02
3	63,2	63,24	-0,04
4	54,7	54,76	-0,06
5	47,4	47,46	0,04
6	41,4	41,32	0,08
7	36,3	36,36	-0,06

Таблиця 3.11 показує узгодженість отриманої формули з дослідними даними. Маємо $\sum_{i=1}^{7} \varepsilon_i^2 = 0.0173$. Бачимо, що формула (3.264) узгоджується краще з експериментальними даними, ніж формула (3.247), яка обчислена методом середніх.

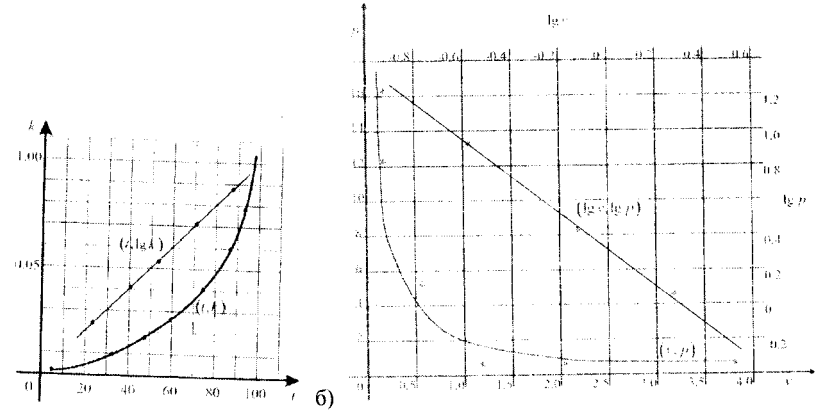
Приклад 3.10. Наступна таблиця показує значення граничної електропровідності k скла в залежності від температури t в градусах C.

t	14,5	30	64,5	74,5	86,7	94,5	98,9
k	0	0,004	0,018	0,029	0,051	0,073	0.09

Підібрати емпіричну формулу для функції $k = f(t)$.

Розв'язок. Точки $M_i(t_i, \lg k_i)$, $i = \overline{1,7}$, за виключенням першої, приблизно розташовані на одній прямій (рис. 3.5а). Тому емпіричну формулу вибираємо у вигляді показникової функції $k = be^{at}$. Для зручності покладемо

$$1000k = ce^{at}, \tag{3.265}$$



а) б)

Рис. 3.5. Приклади графіків емпіричної функції $\widetilde{y} = \widetilde{f}(x)$

де $c = 1000b$. Логарифмуючи (3.265), маємо

$$y = \lg 1000k = \lg c + aMt, \tag{3.266}$$

де $M = \lg e = 0.43429$. Перетворені дані розміщені в таблиці 3.12. Коефіцієнти $\lg c$ та aM в формулі (3.266) визначаємо двома способами, ігноруючи першу точку.

Таблиця 3.12

Заміна змінних t та y

t	$1000k$	y=lg1000k	k_1	$\Delta_1 = k - k_1$	k_2	$\Delta_2 = k - k_2$
14,5	0	—	0,0019	-0,0019	0,0025	-0,0025
30	4	0,6021	0,0039	-0,0001	0,0048	-0,0008
64,5	18	1,2553	0,0185	-0,0005	0,02	-0,002
74,5	29	1,4624	0,0293	-0,0003	0,0302	-0,0012
86,7	51	1,7076	0,0512	-0,0002	0,0499	0,0011
94,5	73	1,8633	0,0731	-0,0001	0,0688	0,0042
98,9	90	1,9542	0,0893	-0,0007	0,0825	0,0075

1. **Метод середніх**. Визначаємо відхилення

$$\varepsilon_1 = 0.6021 - (\lg c + 30aM); \quad \varepsilon_2 = 1.2553 - (\lg c + 64.5aM);$$

$$\varepsilon_3 = 1.4624 - (\lg c + 74.5aM); \quad \varepsilon_4 = 1.7076 - (\lg c + 86.7aM);$$

$$\varepsilon_5 = 1.8633 - (\lg c + 94.5aM).$$

Припустивши, що $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 0$, $\varepsilon_4 + \varepsilon_5 + \varepsilon_6 = 0$, одержуємо систему

$$\begin{cases} 3\lg c + 169aM = 3.3198 \\ 3\lg c + 280.1aM = 5.5251 \end{cases} \quad (3.267)$$

Розв'язуючи систему (3.267), знаходимо $aM = \frac{2.2053}{111.1} = 0.01985$; $a = \frac{0.01985}{M} = 0.0457$ та $\lg c = \frac{1}{3}(3.3198 - 169 \cdot 0.01985) = 1.9884$. Звідки

$$k = 0.9737 \cdot 10^{-3} e^{0.0457t}. \quad (3.268)$$

2. **Метод найменших квадратів**. Проміжкові обчислення наведені в таблиці 3.13.

Таблиця 3.13

Визначення параметрів формули (3.266) методом найменших квадратів

t^0	t	t^2	y	yt
1	30	900	0,6021	18,063
1	64,5	4160,25	1,2553	80,967
1	74,5	5550,25	1,4624	108,949
1	86,7	7516,89	1,7076	148,049
1	94,5	8930,25	1,8633	176,082
1	98,9	9781,21	1,9542	193,7
6	4491	36838,85	8,8449	725,38

Звідси отримуємо нормальну систему

$$\begin{cases} 6\lg c + 449.1aM = 8.8449 \\ 449.1\lg c + 36839aM = 725.38 \end{cases} \text{ та } \begin{cases} \lg c + 74.183aM = 1.4742 \\ \lg c + 82.02aM = 1.6145 \end{cases}$$

Розв'язуючи останню систему, знаходимо aM= 0.01789; a=0.0412 і $\lg c$ =0.1471; c=1.403. Таким чином,

$$k = 1.403 * 10^{-3} * e^{0.0412t}. \quad (3.269)$$

Результати обчислення за формулами (3.268) і (3.269) приведені у табл. 3.13.

3.8.9. Деякі міркування про вибір емпіричної формули з двома параметрами

Нехай для заданої системи значень (x_i, y_i), $i = \overline{1,n}$, де $n \le 3$ та $x_1 < x_2 < \ldots < x_n$, потрібно знайти емпіричну формулу вигляду

$$y = \tilde{f}(x; a, b), \quad (3.270)$$

яка містить лише два параметри a і b. Якщо виявиться, що

$$\frac{\Delta y_i}{\Delta x_i} = const, \quad (3.271)$$

то шукана залежність лінійна:

$$y = ax + b. \quad (3.272)$$

і задача, таким чином, легко розв'язується. Іншим простим випадком є наявність квадратичної залежності

$$y - y_i = a(x - x_i)^2 + b(x - x_i), \tag{3.273}$$

яка розв'язується відомими методами.

Розглянемо загальний випадок, коли співвідношення (3.270), взагалі кажучи, не зводиться до формул (3.272) і (3.273). Достатньою умовою існування емпіричної формули вигляду (3.270), де $\tilde{f}$ – відома функція, є сумісність (із заданою точністю) системи рівнянь $y_i = \tilde{f}(x_i; a, b)$, $i = \overline{1, n}$. Виключаючи звідси невідомі a та b, одержуємо систему умов для точок (x_i, y_i), що гарантує існування залежності (3.270). Але такий підхід є досить складним. Виведемо необхідну умову існування емпіричної залежності вигляду (3.270) для заданої системи точок (x_i, y_i). Нехай M_i (x_i, y_i), M_j (x_i, y_i), M_k (x_i, y_i) – три системи значень із нашої сукупності. Вважаючи, що крива (3.270) проходить через точки M_i, M_j, M_k, будемо мати

$$y_i = \tilde{f}(x_i; a, b), \; y_j = \tilde{f}(x_j; a, b), \; y_k = \tilde{f}(x_k; a, b). \tag{3.274}$$

Виключаючи із системи (3.274) параметри a та b, отримаємо співвідношення вигляду

$$\Phi(x_i, x_j, x_k, y_i, y_j, y_k) = 0. \tag{3.275}$$

Виконання рівності (3.275) для довільних значень i, j, k $(1 \le i < j < k \le n)$ є необхідною умовою для існування залежності (3.270). Так як перевірка співвідношення (3.275) пов'язана з громіздкими обчисленнями, то на практиці звичайно обмежуються однією трійкою точок: початковою (x_1, y_1), проміжковою (x_s, y_s) та кінцевою (x_n, y_n) (з метою досягнення найбільшого діапазону), тобто вважають $i = 1$, $j = s$, $1 < s < n$, $k = n$. Точку M_s вибирають так, щоб співвідношення (3.275) було якомога простішим. Зауважимо, що інколи замість проміжної точки M_s вигідно брати точку $\tilde{M}_S(\tilde{x}_S, \tilde{y}_S)$, що не належить до вибраних точок M_i. Тоді координати x_s та y_s визначаються інтерполюванням.

Приклад 3.11. Отримати умову для існування степенемої залежності

$$y = ax^b, \tag{3.276}$$

$$x_i > 0;\; y_i > 0, \;\; i = 1, 2, \ldots n. \tag{3.277}$$

Розв'язок. Виберемо

$$x_S = \sqrt{x_1 x_n}. \tag{3.278}$$

З формули (3.276) маємо

$$y_1 = ax_1^b; \; y_S = ax_S^b = ax_1^{\frac{b}{2}} x_n^{\frac{b}{2}}; \; y_n = ax_n^b. \tag{3.279}$$

Виключаючи із співвідношень (3.279) параметри a та b, одержимо

$$y_1 y_n = y_S^2 \;\text{і}\; y_S = \sqrt{y_1 y_n}. \tag{3.280}$$

Таким чином, для існування степеневої залежності (3.276) необхідно, щоб середньому геометричному x_S значень x_1 та x_n відповідало середнє геометричне y_S значень y_1 та y_n. Взагалі, якщо має місце степенева залежність (3.276) і значення x_i утворюють геометричну прогресію, то значення y_i також утворюють геометричну прогресію. Якщо значення $x_S = \sqrt{x_1 x_n}$ не є табличним, то відповідне значення y_S визначається за допомогою інтерполювання. Далі в цьому параграфі будемо розглядати залежності, які найбільш часто зустрічаються:

I. $y = ax + b$; II. $y = ax^b$; III. $y = ab^x$; IV. $y = a + \frac{b}{x}$;

V. $y = \frac{1}{ax + b}$; VI. $y = \frac{x}{ax + b}$; VII. $y = a \ln x + b$.

Аналогічно до того, як це зроблено в прикладі 3.11, для існування залежностей I-VII легко вивести прості потрібні умови вигляду

$$x_S = \bar{x}_S, \; y_S = \bar{y}_S. \tag{3.281}$$

де $\bar{x}_S = \varphi(x_1 x_n)$ і $\bar{y}_S = \psi(y_1 y_n)$, причому вважається, що $x_i > 0$ і $y_i > 0$, $i = \overline{1,n}$. Вирази для $\bar{x}_S$ і $\bar{y}_S$ наведені в таблиці 3.14. Таблиця 3.15 полегшує вибір вигляду емпіричної формули серед вказаних.

Алгоритм вибору емпіричної формули. 1. Для перевірки придатності певної емпіричної формули, користуючись вихідними даними, знаходимо значення $x_S = \bar{x}_S$, $y_S = \bar{y}_S$ та порівнюємо їх зі значенням $\psi(y_1 y_n) = \bar{y}_S$ наведеним у таблиці.

2. Вибираємо ту емпіричну функцію, для якої відхилення $|y_S - \bar{y}_S|$ мале.

3.Для кінцевого вибору потрібно врахувати також проміжкові значення. Якщо величина $|y_S - \bar{y}_S|$ велика, то відповідна емпірична формула відкидається.

Зауваження 3.4. Слід розуміти, що такий підхід є дещо грубо орієнтовним, оскільки не враховуємо поведінку всіх проміжкових даних (x_1, y_1). Крім того, таблиця 3.14 охоплює невелику кількість залежностей і може статися, що змінні x та y підлягають деякій закономірності, яка не входить в наш перелік.

Зауваження 3.5. Всі залежності в таблиці 3.14 методом вирівнювання можуть бути просто перетворені в лінійні. Тому тут можна використовувати критерій прямолінійності (формула (3.214)) для перетворених вихідних даних (x_1, y_1).

Таблиця 3.14

Необхідні умови для наявності емпіричних залежностей I-VII

№	$\bar{x}_S$	$\bar{y}_S$	Емпіричні формули	Спосіб вирівнювання
I	$\frac{x_1 + x_n}{2}$ (середнє арифметичне)	$\frac{y_1 + y_n}{2}$ (середнє арифметичне)	$y = ax + b$	$y = ax + b$
II	$\sqrt{x_1 x_n}$ (середнє геометричне)	$\sqrt{y_1 y_n}$ (середнє геометричне)	$y = ax^b$	$Y = a + bX$, де $X = \lg x$, $Y = \lg y$, $a = \lg a$

Продовження таблиці 3.14

III	$\frac{x_1+x_n}{2}$ (середнє арифметичне)	$\sqrt{y_1 y_n}$ (середнє геометричне)	$y = ab^x$ або $y = ae^{\beta x}$, де $\beta = \ln b$	$Y = \alpha + \beta x$, де $Y = \lg y$, $\alpha = \lg a$, $\beta = \lg b$
IV	$\frac{2x_1x_n}{x_1+x_n}$ (середнє гармонійне)	$\frac{y_1+y_n}{2}$ (середнє арифметичне)	$y = a + \frac{b}{x}$	$Y = ax + b$, де $Y = xy$
V	$\frac{x_1+x_n}{2}$ (середнє арифметичне)	$\frac{2y_1y_n}{y_1+y_n}$ (середнє гармонійне)	$y = \frac{1}{ax+b}$	$Y = ax + b$, де $Y \equiv \frac{1}{y}$
VI	$\frac{2x_1x_n}{x_1+x_n}$ (середнє гармонійне)	$\frac{2y_1y_n}{y_1+y_n}$ (середнє гармонійне)	$y = \frac{x}{ax+b}$	$Y = ax + b$, де $Y \equiv \frac{x}{y}$
VII	$\sqrt{x_1x_n}$ (середнє геометричне)	$\frac{y_1+y_n}{2}$ (середнє арифметичне)	$y = a \lg x + b$	$y = aX + b$, де $X = \lg x$

Таблиця 3.15

Підбір емпіричної формули на основі критеріїв № I-IV (див. табл. 3.14)

№	$\bar{x}_S$	$\bar{y}_S$	$\hat{y}_S$	$\lvert \hat{y}_S - \bar{y}_S \rvert$	Вигляд формули
I	$\frac{x_1+x_n}{2} = \frac{273+373}{2} = 323$	$\frac{y_1+y_n}{2} = \frac{29.4+77.3}{2} = 53.35$	50.5	2.85	$y = ax + b$ - мало підходить
II	$\sqrt{x_1x_n} = \sqrt{273 \cdot 373} = 319.1$	$\sqrt{y_1y_n} = \sqrt{29.4 \cdot 77.3} = 47.7$	48.7	1.0	$y = ax^b$ - найкраще підходить
III	$\frac{x_1+x_n}{2} = \frac{273+373}{2} = 323$	$\sqrt{y_1y_n} = \sqrt{29.4 \cdot 77.3} = 47.7$	50.5	2.8	$y = ab^x$ - мало підходить
IV	$\frac{2x_1x_n}{x_1+x_n} = 2\frac{273 \cdot 373}{273+373} = 315.3$	$\frac{y_1+y_n}{2} = \frac{29.4+77.3}{2} = 53.35$	46.9	6.45	$y = a + \frac{b}{x}$ не підходить
V	$\frac{x_1+x_n}{2} = \frac{273+373}{2} = 323$	$\frac{2y_1y_n}{y_1+y_n} = 2\frac{29.4 \cdot 77.3}{29.4+77.3} = 42.6$	50.5	7.9	$y = \frac{1}{ax+b}$ не підходить
VI	$\frac{2x_1x_n}{x_1+x_n} = 2\frac{273 \cdot 373}{273+373} = 315.3$	$\frac{2y_1y_n}{y_1+y_n} = 2\frac{29.4 \cdot 77.3}{29.4+77.3} = 42.6$	46.9	4.3	$y = \frac{x}{ax+b}$ не підходить
VII	$\sqrt{x_1x_n} = \sqrt{273 \cdot 373} = 319.1$	$\frac{y_1+y_n}{2} = \frac{29.4+77.3}{2} = 53.35$	48.7	4.65	$y = a \lg x + b$ не підходить

Функції I-VII монотонні, тоді впорядковані дані (x_1, y_1) при

$$\Delta x_i = x_{i+1} - x_i > 0,\ \overline{i = 1, n-1}, \tag{3.282}$$

повинні володіти сталим знаком приросту

$$\Delta y_i = y_{i+1} - y_i,\ \overline{i = 1, n-1}, \tag{3.283}$$

Якщо ця обставина не має місця, то залежності I-VII не підходять, і слід інтуїтивно з постановки задачі пропонувати вигляд емпіричної формули.

Приклад 3.12. Визначити вигляд емпіричної формули, яка відповідає наступній таблиці

x	273	283	288	293	313	333	353	373
y	29,4	33,3	35,2	37,2	45,8	55,2	65,6	77,3

Розв'язок. Будемо шукати емпіричну формулу серед залежностей I-VII, наведених в таблиці 3.14, згідно з вказаним алгоритмом. Результати обчислень наведені в таблиці 3.15, причому в необхідних випадках застосована лінійна інтерполяція. З таблиці 3.15 видно, що згідно з необхідним критерієм потрібно зупинитися на степеневій залежності $y = ax^b$.

Приклад 3.13. Наступна таблиця задає тиск p насиченого пару в кг/см^2, що відповідає питомому об'єму v в м^3/кг.

V	3,334	1,630	0,8657	0,4323	0,2646	0,1699	0,1146
p	0,482	1,034	2,027	4,247	7,164	11,48	17,60

Знайти емпіричну формулу для залежності $F(v, p) = 0$.

Розв'язок. Підберемо вигляд формули, користуючись таблицею 3.14. За формулою II маємо $\overline{v}_S = \sqrt{v_1 v_n} = \sqrt{3.334 \cdot 0.1146} = 0.618$ і $\overline{p}_S = \sqrt{p_1 p_n} = \sqrt{0.482 \cdot 17.60} = 2.92$.

Значення $\hat{p}_S$, яке відповідає $\overline{v}_S = 0.618$, знайдемо за допомогою лінійної інтерполяції: $\hat{p}_S = 2.027 + \dfrac{4.247 - 2.027}{0.4323 - 0.8657} \cdot (0.618 - 0.8657) = 3.295$. Так як відхилення $|\overline{p}_S - \hat{p}_S| = 0.375$ порівняно незначне, то можна вибрати емпіричну формулу вигляду

$$p = av^b. \tag{3.284}$$

Це узгоджується також з природою газових законів. Логарифмуючи формулу (3.284), маємо $\lg p = \lg a + b \lg v$. Звідси, якщо візьмемо $X = \lg v$, $Y = \lg p$, отримаємо лінійну залежність

$$Y = bX + \lg a. \tag{3.285}$$

Значення $X_i = \lg v_i$, і $Y_i = \lg p_i$, (i=1,2, ..., 7) наведені в таблиці 3.16. Побудувавши точки $M_i(\lg v_i, \lg p_i)$ переконуємось, що вони розміщені на прямій лінії (рис. 3.5б). Отже, вибір емпіричної формули проведено правильно.

Коефіцієнти формули (3.285) обчислюємо методом середніх. Запишемо відхилення $\varepsilon_i = \lg a + b \lg v_i - \lg p_i$, $\overline{i = 1,7}$, і, розіб'ємо їх на дві групи:

$$\varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 0 \text{ і } \varepsilon_4 + \varepsilon_5 + \varepsilon_6 + \varepsilon_7 = 0$$

Отримаємо систему

$$\left.\begin{aligned} 3\lg a + 0.6725b &= 0.0043 \\ 4\lg a - 2.6522b &= 3.7886 \end{aligned}\right\}$$

Звідки одержуємо $b = -1.066$; $\lg a = 0.2402$; $a = 1.740$. Остаточно маємо:

$$pv^{1.066} = 1.740. \tag{3.286}$$

В таблиці 3.16 задані розбіжності, які одержані за формулою (3.286) значень p_e, з табличними даними p.

Таблиця 3.16

Оцінка точності емпіричної формули (3.286)

v	p	$\lg v$	$\lg p$	p_e	$\lvert p - p_e \rvert$	Δ^2
3,34	0,482	0,5229	-0,3170	0,482	0,000	0
1,630	1,034	0,2122	0,0145	1,033	0,001	$1 \cdot 10^6$
0,8657	2,027	-0,0626	0,3068	2,028	-0,001	$1 \cdot 10^6$
0,4323	4,247	-0,3642	0,6281	4,241	0,006	$36 \cdot 10^6$
0,2646	7,164	-0,5774	0,8551	7,176	-0,012	$144 \cdot 10^6$
0,1699	11,48	-0,7698	1,0599	11,48	0,00	0
0,1146	17,60	-0,9408	1,2455	17,51	1,090	$8100 \cdot 10^6$
					$\sum$=0,084	$\sum$=0,0083

3.8.10. Емпіричні формули, які містять три параметри

Розглянемо для заданої системи значень (x_i, y_i), $i = \overline{1,n}$ табличних даних, найважливіших представників емпіричних формул вигляду

$$y = \tilde{f}(x; a, b, c), \tag{3.287}$$

де a, b і c – деякі сталі.

1. Квадратична залежність. Нехай

$$y = ax^2 + bx + c. \tag{3.288}$$

2. Степенева залежність. Покладемо

$$y = ax^b + c, \tag{3.289}$$

звідси $y - c = ax^b$. Логарифмуючи цей вираз, одержуємо $\lg(y - c) = \lg a + b \lg x$. Звідси, покладаючи $\lg(y - c) = Y$ і $\lg x = X$, отримаємо лінійну залежність

$$Y = \lg a + bX. \tag{3.290}$$

Визначення параметрів формули (3.289) слід почати зі знаходження значення c. Для цього запишемо середнє геометричне

$$x_S = \sqrt{x_1 x_n}, \tag{3.291}$$

де x_1 і x_n – крайні значення змінної x, і, користуючись методом лінійної інтерполяції для x_S, знайдемо відповідне значення y_S. Вважаючи, що точки $M_1(x_1, y_1)$, $M_S(x_S, y_S)$, $M_n(x_n, y_n)$ розміщені на кривій (3.289), маємо три рівності: $y_1 = c + ax_1^b$, $y_S = c + ax_S^b$ і $y_n = c + ax_n^b$. Піднесемо $x_S = \sqrt{x_1 x_n}$ до степеня b і, помноживши на a, отримаємо $ax_S^b = \sqrt{ax_1^b ax_n^b}$ або $y_S - c = \sqrt{(y_1 - c)(y_n - c)}$.

Розв'язуючи останню рівність відносно c, знаходимо

$$c = \frac{y_1 y_n - y_S^2}{y_1 + y_n - 2y_S}. \qquad (3.292)$$

Коли c визначено, будуємо точки $N_i(X_i, Y_i)$, де $X_i = \lg x_i$, $Y_i = \lg(y_i - c)$, $(i = \overline{1,n})$. Якщо ці точки розміщені на прямій (або майже на прямій), то справедлива залежність (3.289), причому сталі a і b знаходяться звичайним способом.

Приклад 3.14. Для змінних x та y задана таблиця значень:

x	250	500	900	1200	1600	2000
y	0,10	0,28	0,80	1,38	2,56	4,10

Знайти емпіричну формулу, яка пов'язує ці змінні.

Розв'язок. Побудуємо емпіричну формулу вигляду $y = ax^b + c$. Знаходимо

$$x_S = \sqrt{x_1 x_n} = \sqrt{250 \cdot 2000} = 707.$$

На графіку цьому значенню x_S відповідає $y_S = 0.507$. Звідки

$$c = \frac{0.10 \cdot 4.10 - (0.507)^2}{0.10 + 4.10 - 2 \cdot 0.507} = 0.048.$$

Інші параметри a і b знайдемо методом середніх. Складемо початкові рівняння:

$$\lg(y_i - 0.048) = \lg a + b \lg x_i,\ i = \overline{1,6},\ \text{тобто}$$

$$-1.2840 = \lg a + 2.3979b; \quad -0.6345 = \lg a + 2.6990b; \quad -0.1238 = \lg a + 2.9542b;$$

$$0.1245 = \lg a + 3.0792b; \quad 0.4000 = \lg a + 3.2041b; \quad 0.6077 = \lg a + 3.3010b.$$

Групуючи ці рівняння по три, отримаємо

$$\left.\begin{aligned} -2.0423 &= 3\lg a + 8.0511b \\ 1.1322 &= 3\lg a + 9.5843b \end{aligned}\right\}. \qquad (3.293)$$

Розв'язуючи систему (3.293), знаходимо a=5.789·10^{-7} та b=2.071. Звідси випливає, що шукана емпірична функція буде мати вигляд

$$\tilde{y} = 5.789 \cdot x^{2.071} + 0.048. \qquad (3.294)$$

Порівняння значень $\tilde{y}$, отриманих за формулою (3.294), з табличними даними y наведено в таблиці 3.17.

Таблиця 3.17

Оцінка точності емпіричної формули (3.294)

x	y	$\tilde{y}$	$\varepsilon = y - \tilde{y}$
250	0,10	0,102	-0,002
500	0,28	0,273	+0,007
900	0,80	0,808	-0,008
1200	1,38	1,426	-0,046
1600	2,56	2,552	+0,008
2000	4,10	4,020	+0,080

3. Показникова залежність. Нехай

$$y = ae^{bx} + c. \qquad (3.295)$$

Перенесемо c вліво і, логарифмуючи, одержимо $\lg(y-c)=\lg a+(bM)x$, де M = 0,43429. Таким чином,

$$Y=\lg a+bMx. \tag{3.296}$$

де $Y=\lg(y-c)$. Спочатку визначимо параметр c. Для цього, як і в попередньому випадку, виберемо крайні точки $M_1(x_1,y_1)$, $M_n(x_n,y_n)$ та складемо середнє арифметичне $x_S=\frac{x_1+x_n}{2}$. Для значення x_S знайдемо відповідне значення y_S за лінійною інтерполяцією. Підставляючи ці значення в емпіричну формулу (3.295), маємо $y_1=ae^{bx_1}+c$; $y_n=ae^{bx_n}+c$; $y_S=ae^{\frac{b}{2}(x_1+x_n)}+c$. Звідки $y_1-c=ae^{bx_1}$; $y_n-c=ae^{bx_n}$; і, відповідно $(y_1-c)(y_n-c)=a^2e^{b(x_1+x_n)}$, тобто $(y_1-c)(y_n-c)=(y_s-c)^2$.

Якщо розв'язати останнє рівняння відносно c, то одержимо

$$c=\frac{y_1y_n-y_S^2}{y_1+y_n-2y_S}.$$

Якщо знайдена лінійна залежність (3.296), то решта параметрів a і b знаходяться звичайними способами.

Виведемо аналітичний критерій для показникової залежності (3.295), вважаючи, що значення x_i рівновіддалені, тобто $\Delta x_i=h=const$, $i=\overline{1,n-1}$. З формули (3.295) маємо $y_i=ae^{bx_i}+c$ і $y_{i+1}=ae^{b(x_i+h)}+c$. Звідси

$$y_{i+1}-y_i\equiv\Delta y_i=a_1e^{bx_i}, \tag{3.297}$$

де $a_1=a(e^{bh}-1)$. Прологарифмувавши рівність (3.297), маємо $\lg\Delta y_i=\lg a_1+bMx_i$. При наявності залежності (3.295) точки $N_i(x_i,\lg\Delta y_i)$, $(i=\overline{1,n-1})$ розміщені на прямій лінії. Отже, враховуючи, що Δx_i сталі, отримуємо шуканий критерій

$$\Delta(\lg\Delta y_i)=const,\ i=\overline{1,n-2}. \tag{3.298}$$

Приклад 3.15. Для змінних x та y задана таблиця значень:

x	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
y	1,30	1,44	1,59	1,78	1,97	2,19	2,46	2,74	3,06	3,42	3,84

Знайти емпіричну формулу залежності x від y.

Розв'язок. Складемо таблицю різниць Δy_i (таблиця 3.18).

Таблиця 3.18

Перевірка на показникову залежність

x	y	Δy	$\lg\Delta y$	$\Delta(\lg\Delta y)$
0	1,30	0,14	-0,854	0,030
0,1	1,44	0,15	-0,824	0,103
0,2	1,59	0,19	-0,721	0,000
0,3	1,78	0,19	-0,721	0,137
0,4	1,97	0,22	-0,658	0,089

Продовження таблиці 3.18

0,5	2,19	0,27	-0,569	0,016
0,6	2,46	0,28	-0,553	0,058
0,7	2,74	0,32	-0,495	0,051
0,8	3,06	0,36	-0,444	0,067
0,9	3,42	0,42	-0,377	
1,0	3,84			

Оскільки, різниця Δy монотонно зростає, то залежність x і y не є лінійною. Перевіряємо критерій (3.298) для показникової залежності (3.295). З таблиці 3.18 видно, що має місце значне відхилення для $\Delta(\lg \Delta y)$. Тому формулу $y = ae^{bx} + c$ можна прийняти тільки як наближення. Знайдемо $x_S = \frac{1}{2}(x_1 + x_n) = \frac{1}{2}(0 + 1.0) = 0.5$. Звідси

$$c = \frac{3.84 \cdot 1.30 - (2.19)^2}{1.30 + 3.84 - 2 \cdot 2.19} = 0.258.$$

Параметри a та b визначимо за методом середніх. Маємо $\lg(y_i - 0.258) = \lg a + (bM)x_i$, $i = \overline{0,10}$. Складемо два рівняння:

$$\sum_{i=0}^{5} \lg(y_i - 0.258) = 6\lg a + bM\sum_{i=0}^{5} x_i \quad \text{і} \quad \sum_{i=0}^{10} \lg(y_i - 0.258) = 6\lg a + bM\sum_{i=0}^{10} x_i.$$

Підставивши значення x_i та y_i, отримаємо систему

$$\left.\begin{aligned} 0.9169 &= 6\lg a + 1.5bM \\ 2.2392 &= 5\lg a + 4.0bM \end{aligned}\right\}$$

звідки a=1,044; b=1,234. Таким чином, шукана емпірична формула має вигляд

$$\tilde{y} = 1.044e^{1.234x} + 0.258. \qquad (3.299)$$

Узгодженість отриманої формули з початковими даними показана в таблиці 3.19.

Таблиця 3.19

Відхилення емпіричної формули (3.299) від табличних даних

x	y	$\tilde{y} = 1.044e^{1.234x} + 0.258$	$\Delta = y - \tilde{y}$
0	1,30	1,302	-0,002
од	1,44	1,439	+0,001
0,2	1,59	1,594	-0,004
0,3	1,78	1,770	+0,010
0,4	1,97	1,968	+0,002
0,5	2,19	2,193	-0,003
0,6	2,46	2,447	+0,013
0,7	2,74	2,735	+0,005
0,8	3,06	3,048	+0,012
0,9	3,42	3,428	-0,008
1,0	3,84	3,844	-0,004

Зауваження 3.6. Що стосується емпіричних формул, що містять більше трьох параметрів, то вони рідко зустрічаються на практиці, тому детально розглядати їх не будемо.

3.8.11. Уточнення отриманої емпіричної формули

Вкажемо простий прийом уточнення отриманої емпіричної формули у випадку, коли вона дає суму квадратів відхилень більш значну, ніж бажана. Нехай для заданої системи значень (x_i, y_i), $i = \overline{1,n}$ знайдена емпірична формула

$$y = f(x). \tag{3.300}$$

Потрібно зменшити суму квадратів відхилень

$$s = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} [y_i - f(x_i)]^2, \tag{3.301}$$

де $\varepsilon_i = y_i - f(x_i)$, $i = \overline{1,n}$. Розглянемо функцію

$$\tilde{y} = f(x) + c, \tag{3.302}$$

де c – деяка стала величина.

Підберемо число c так, щоб сума квадратів нових відхилень

(3.303)

була мінімальною. Очевидно, отримаємо $\tilde{s} = \sum_{i=1}^{n} (c - \varepsilon_i)^2$. Для мінімуму функцій $\tilde{s} = \tilde{s}(c)$ необхідно, щоб $\frac{\partial \tilde{s}}{\partial c} = 2\sum_{i=1}^{n} (c - \varepsilon_i) = 0$. Звідси $nc = \sum_{i=1}^{n} \varepsilon_i$. Значить

$$c = \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i. \tag{3.304}$$

Оскільки $\frac{\partial^2 \tilde{s}}{\partial c^2} = 2n > 0$, то значення (3.304) дає найменше значення для суми квадратів $\tilde{s}$.

Теорема 3.16. Найкраща стала c являє собою середнє арифметичне відхилень ε_i (3.304). Таким чином, якщо $c \neq 0$, то додавання до правої частини емпіричної формули (3.300) сталої c, що визначається формулою (3.304), призводить до уточнення цієї формули в розумінні суми квадратів відхилень, оскільки $\tilde{s} < s$. Якщо c=0 або близьке до нуля, то вказаний метод не дає потрібного ефекту. В цьому випадку можна покласти $\tilde{y} = f(x) + c\varphi(x)$, де $\varphi(x)$ – відома функція, яка не перетворюється в нуль у всіх точках x_i, тобто така, що

$$\sum_{i=1}^{n} \varphi(x_i) > 0. \tag{3.305}$$

Звідси $\tilde{s} = \sum_{i=1}^{n}[y_i - f(x_i) - c\varphi(x_i)]^2 = \sum_{i=1}^{n}[c\varphi(x_i) - \varepsilon_i]^2$. Необхідна умова мінімуму функції $\tilde{s}$ дає $\frac{1}{2}\frac{\partial \tilde{s}}{\partial c} = \sum_{i=1}^{n}\varphi(x_i)[c\varphi(x_i) - \varepsilon_i] = 0$. Звідки

$$c = \frac{\sum_{i=1}^{n}\varepsilon_i\varphi(x_i)}{\sum_{i=1}^{n}[\varphi(x_i)]^2}, \tag{3.306}$$

оскільки $\frac{1}{2}\frac{\partial^2 \tilde{s}}{\partial^2 c} = \sum_{i=1}^{n}\varphi^2(x_i) > 0$.

Теорема 3.17. Для функції $\varphi(x)$ за умови (3.305) найкращою сталою для зменшення суми квадратів відхилень є "середнє вагове" відхилень ε_i, яке визначається формулою (3.306). Зокрема, якщо $\varphi(x) = x$, то маємо

$$c = \frac{\sum_{i=1}^{n}\varepsilon_i x_i}{\sum_{i=1}^{n} x_i^2}. \tag{3.307}$$

Приклад 3.16. Для табличних даних (таблиця 3.20) одержана емпірична формула

$$\bar{y} = \frac{x}{0.051x + 0.209}. \tag{3.308}$$

Шляхом додавання сталої, уточнити цю формулу.

Таблиця 3.20

Вхідні дані для прикладу 3.16

x	8	10	15	20	30	40	50	80
y	13,0	14,0	15,4	16,3	17,2	17,8	18,5	18,8

Розв'язок. Вважатимемо, що $\tilde{y} = \frac{x}{0.051x + 0.209} + c$, де $c = \frac{\sum_{i=1}^{8}\varepsilon_i}{8} = \frac{0.47}{8} = 0.06$ (таблиця 3.21). Тому

$$\tilde{y} = \frac{x}{0.051x + 0.209} + 0.06 = \frac{1.0031x + 0.0125}{0.051x + 0.209}. \tag{3.309}$$

Результати уточнення наведені в таблиця 3.21.

Таблиця 3.21

Порівняння емпіричних формул (3.308) і (3.309)

x	y	$\overline{y}$ за формулою (7)	$\widetilde{y}$ за уточ. формулою (8)	$\varepsilon = y - \overline{y}$	ε^2	$\widetilde{\varepsilon} = y - \widetilde{y}$	$\widetilde{\varepsilon}^2$
8	13,0	12,97	13,02	+0,03	$9\cdot10^{-4}$	-0,02	$4\cdot10^{-4}$
10	14,0	13,91	13,96	+0,09	$81\cdot10^{-4}$	+0,04	$16\cdot10^{-4}$
15	15,4	15,40	15,45	0,00	0	-0,05	$25\cdot10^{-4}$
20	16,3	16,27	16,32	+0,03	$9\cdot10^{-4}$	-0,02	$4\cdot10^{-4}$
30	17,2	17,25	17,30	-0,05	$25\cdot10^{-4}$	-0,10	$100\cdot10^{-4}$
40	17,8	17,79	17,84	+0,01	$1\cdot10^{-4}$	-0,04	$16\cdot10^{-4}$
60	18,5	18,39	18,44	+0,11	$121\cdot10^{-4}$	+0,06	$36\cdot10^{-4}$
80	18,8	18,65	18,70	+0,15	$225\cdot10^{-4}$	+0,10	$100\cdot10^{-4}$
Σ				0,47	0,0471	-0,03	0,0301

3.8.12. Загальний метод визначення параметрів емпіричної формули

До цього розглядалися, в основному, емпіричні формули, які лінійно залежать від параметрів (або зводяться до такого вигляду), і для цього часткового випадку давались ефективні методи визначення параметрів. Нехай для сукупності впорядкованих значень (x_i, y_i), $i = \overline{1,n}$, побудована емпірична формула

$$y = \widetilde{f}(x; a_1, a_2, ..., a_m), \tag{3.310}$$

яка містить m параметрів ($m < n$), причому функція $\widetilde{f}$ має неперервні частинні похідні за всіма своїми параметрами. Потрібно визначити параметри $a_1, a_2, ..., a_m$ так, щоб формула (3.310) була узгодженою з вихідними даними, тобто відхилення

$$\varepsilon_i = y_i - \widetilde{f}(x; a_1, a_2, ..., a_m), \tag{3.311}$$

повинні бути найменшими за абсолютною величиною.

Якщо вихідні дані (x_i, y_i) не містять помилок і залежність $\widetilde{f}$ є точною, то задача знаходження параметрів a_i, $i = \overline{1,m}$, зводиться до розв'язку m рівнянь ($m < n$) з наступної системи n рівнянь:

$$\left.\begin{aligned} y_1 &= \widetilde{f}(x_1; a_1, a_2, ..., a_m) \\ y_2 &= \widetilde{f}(x_2; a_1, a_2, ..., a_m) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n &= \widetilde{f}(x_n; a_1, a_2, ..., a_m) \end{aligned}\right\}. \tag{3.312}$$

Але на практиці, через відсутність вказаних властивостей, система (3.312), як правило, є несумісною, тобто значення a_i, $i = \overline{1,m}$, знайдені з деяких m рівнянь системи (3.312), не задовольняють решту n-m рівнянь. Для наближеного розв'язку системи (3.312) роблять так: яким-небудь способом (наприклад, графічним або шляхом

розв'язування вибраних m рівнянь системи (3.312)) знаходять наближені значення параметрів $a_1^{(0)}, a_2^{(0)}, ..., a_m^{(0)}$. Нехай

$$a_i = a_1^{(0)} + \alpha_i, \; i = \overline{1,m}, \tag{3.313}$$

де a_i – поправки, які припускають «малими», і

$$\varepsilon^{(0)} = y_j - \widetilde{f}(x_j; a_1^{(0)}, a_2^{(0)} ..., a_m^{(0)}), \;\; j = 1,2,...,n \tag{3.314}$$

j – відповідні нев'язки. Підставляючи значення (3.313) у рівняння системи (3.312), розглядаючи праві частини одержаних рівнянь за степенями поправок a_i і утримуючи лише члени першого порядку відносно цих поправок, одержимо

$$\begin{aligned} y_j &= \widetilde{f}(x_j; a_1^{(0)}, a_2^{(0)}, ..., a_m^{(0)}) + \\ &+ \sum_{k=1}^{m} \widetilde{f}'_{a_k}(x_j; a_1^{(0)}, a_2^{(0)}, ..., a_m^{(0)}) \alpha_k, \;\; j = 1,2,...,n. \end{aligned} \tag{3.315}$$

Ввівши позначення

$$\widetilde{f}'_{a_k}(x_j; a_1^{(0)}, a_2^{(0)}, ..., a_m^{(0)}) = b_{jk}, \;\; j = 1,2,...,n, \; k = 1,2,...,m, . \tag{3.316}$$

знаходимо

$$\sum_{k=1}^{m} b_{jk} a_k = \varepsilon_j^{(0)}, \;\; j = 1,2,..., n. \tag{3.317}$$

Система (3.317) лінійна відносно невідомих поправок a_i, k=1,2,...,m, і, взагалі кажучи, є несумісною, так як кількість рівнянь, більше за кількість невідомих.

Рівняння системи (3.317) називаються *умовними*, а сама система *системою умовних рівнянь*. Систему умовних рівнянь (3.317) можна «розв'язати», в певному розумінні, описаним вище методом середніх або методом найменших квадратів. Підставляючи в нелінійну систему (3.312) знайдені значення $a_i^{(1)} = a_i^{(0)} + \alpha_i$, $i = 1,2,...,m$, можна визначити нові нев'язки $\varepsilon_i^{(1)}$ і, якщо необхідно, повторити процес.

Приклад 3. 17. Результати експерименту характеризуються наступною таблицею

x	0	1	2	3
y	2.01	1.21	0.74	0.45

Вважаючи, що x та y пов'язані показниковою залежністю

$$y = a \cdot e^{-b \cdot x}, \tag{3.318}$$

визначити найкращі значення параметрів a і b.

Розв'язок. Підставляючи табличні дані у формулу (3.318), одержимо систему

$$\left.\begin{aligned} a &= 2.01 \\ a \cdot e^{-b} &= 1.21 \\ a \cdot e^{-2b} &= 0.74 \\ a \cdot e^{-3b} &= 0.45 \end{aligned}\right\} \tag{3.319}$$

Розв'язуючи, наприклад, перші два рівняння цієї системи одержимо наближені значення параметрів $a^{(0)}=2{,}01$; $b^{(0)}=0{,}51$ Для знаходження поправок $\alpha = a - a^{(0)}$; $\beta = b - b^{(0)}$ складемо систему умовних рівнянь (3.312). Відповідні значення похідних $\left(\frac{\partial y}{\partial a}\right)_0 = e^{-b^{(0)}x}$; $\left(\frac{\partial y}{\partial b}\right)_0 = -a^{(0)}x \cdot e^{-b^{(0)}x}$ і нев'язок $\varepsilon^{(0)} = y - a^{(0)}e^{-b^{(0)}x}$ знаходяться в таблиці 3.22.

Таблиця 3.22

Коефіцієнти умовних рівнянь

x	$\left(\frac{\partial y}{\partial a}\right)_0$	$\left(\frac{\partial y}{\partial b}\right)_0$	$\varepsilon^{(0)}$
0	1	0	0
1	0,600	-1,206	0,004
2	0,361	-1,451	0,014
3	0,216	-1,302	0,016

Звідки отримаємо систему умовних рівнянь

$$\left.\begin{aligned} &\alpha = 0; \\ &0.600\alpha - 1.206\beta = 0.004 \\ &0.361\alpha - 1.451\beta = 0.014 \\ &0.216\alpha - 1.302\beta = 0.016 \end{aligned}\right\} \qquad (3.320)$$

Для розв'язку системи (3.320) застосуємо метод найменших квадратів. Проміжкові обчислення приведемо в таблиці 3.23, де c_α і c_β позначають коефіцієнти при α і β в системі (3.320) і c_0 — відповідні вільні члени.

Отже, нормальна система рівнянь має вигляд

$$\left.\begin{aligned} &1.5370\alpha - 1.5286\beta = 0.01090; \\ &-1.5286\alpha + 5.2550\beta = -0.04596 \end{aligned}\right\}$$

Таблиця 3.23

Розв'язок системи (3.320) методом найменших квадратів

c_α	c_β	c_0	c_α^2	$c_\alpha c_\beta$	c_β^2	$c_\alpha c_0$	$c_\beta c_0$
1	0	0	1	0	0	0	0
0,600	-1,206	0,004	0,3600	-0,7236	1,4544	0,00240	-0,00482
0,361	-1,451	0,014	0,1303	-0,5238	2,1054	0,00505	-0,02031
0,216	-1,302	0,016	0,0467	-0,2812	1,6952	0,00345	-0,02083
$\sum$			1,5370	-1,5286	5.2550	0,01090	-0,04596

Розв'язавши цю систему, знайдемо α=–0.004; β=–0.009. Звідси отримуємо виправлені значення параметрів α = 2.01–0.004 = 2.006; β = 0.51–0.009 = 0.501. Отже, шукана емпірична формула має вигляд

$$\tilde{y} = 2.006 \cdot e^{-0.501x}. \qquad (3.321)$$

У таблиці 3.24 подано порівняння результатів, які дає емпірична формула (3.321) з табличними даними, де $\sum \varepsilon_i^2 = 7.3 \cdot 10^{-5}$.

Таблиця 3.24

Порівняння значень емпіричної формули (3.321) з табличними даними

x	y	$\tilde{y}$	$\varepsilon = y - \tilde{y}$	ε^2
0	2.01	2.006	0.004	$1.6\cdot10^{-5}$
1	1.21	1.215	-0.005	$2.5\cdot10^{-5}$
2	0.74	0.736	0.004	$1.6\cdot10^{-5}$
3	0.45	0.446	0.004	$1.6\cdot10^{-5}$

Зауважимо, що за допомогою логарифмування формули (3.318) легко отримати еквівалентну формулу, яка лінійно залежить від вхідних параметрів.

Приклад 3.18. Для змінних x і y, значення яких задані наступною таблицею знайдена емпірична залежність

$$\tilde{y} = e^{cx}(a\sin x + b\cos x). \tag{3.322}$$

x	0	0.2	0.6	1.2	1.6	2
y	2.05	1.944	1.638	0.907	0.423	0.028

Визначити значення параметрів a, b і c у формулі (3.322).

Розв'язок. Підставимо у емпіричну формулу (3.322) ці значення (x_i, y_i), отримаємо наступну систему рівнянь:

$$\left.\begin{aligned} 2.05 &= b \\ 1.944 &= e^{0.2c}(0.1987a + 09801b) \\ 1.638 &= e^{0.6c}(0.5646a + 0.8253b) \\ 0.907 &= e^{1.2c}(0.9320a + 0.3624b) \\ 0.423 &= e^{1.6c}(0.9996a + 0.0292b) \\ 0.028 &= e^{2c}(0.9090a + 0.4169b) \end{aligned}\right\} \tag{3.323}$$

Розв'язуючи наближено, наприклад перші 3 рівняння системи (3.323), знайдемо $a_0 = 1$; $b_0 = 2.05$; $c_0 = -0.495$. Покладемо $a = a_0 + \alpha$, $b = b_0 + \beta$, $c = c_0 + \gamma$.

Таблиця 3.25

Коефіцієнти умовних рівнянь

x	$\left(\frac{\partial\tilde{y}}{\partial a}\right)_0$	$\left(\frac{\partial\tilde{y}}{\partial b}\right)_0$	$\left(\frac{\partial\tilde{y}}{\partial c}\right)_0$	$\varepsilon^{(0)}$
0	0	1	0	0
0,2	0,1795	0,0886	0,3992	-0,0540
0,6	0,4183	0,6114	1,0030	-0,0336
1,2	0,5167	0,2009	1,1141	-0,0214
1,6	0,4536	-0,0132	0,6832	-0,0034
2	0,3378	-0,1549	0,0404	-0,0072

Для визначення поправок α, β та γ складемо лінійну систему умовних рівнянь (3.317). Значення похідних

$$\left(\frac{\partial\tilde{y}}{\partial a}\right)_0 = e^{c_0 x}\sin x;\quad \left(\frac{\partial\tilde{y}}{\partial b}\right)_0 = e^{c_0 x}\cos x;\quad \left(\frac{\partial\tilde{y}}{\partial c}\right)_0 = xe^{c_0 x}(a_0\sin x + b_0\cos x)$$

і незв'язок $\varepsilon^{(0)} = y - e^{c_0 x}(a_0\sin x + b_0\cos x)$ зведено у таблиці 3.25.

Звідси система рівнянь для поправок має вигляд

$$\left.\begin{aligned}
&\beta = 0\\
&0.1795\alpha + 0.0886\beta + 0.3992\gamma = -0.0540;\\
&0.4183\alpha + 0.6114\beta + 1.0030\gamma = -0.0336;\\
&0.5167\alpha + 0.2009\beta + 1.1141\gamma = -0.0214;\\
&0.4536\alpha - 0.0132\beta + 0.6823\gamma = -0.0034;\\
&0.3378\alpha - 0.1549\beta + 0.0404\gamma = -0.0072.
\end{aligned}\right\} \qquad (3.324)$$

Розв'язок системи (3.324) будемо шукати методом середніх. Групуючи по два рівняння системи, отримаємо

$$\left.\begin{aligned}
&0.1795\alpha + 1.0886\beta + 0.3992\gamma = -0.0540;\\
&0.9350\alpha + 0.8123\beta + 2.1171\gamma = -0.0550;\\
&0.7914\alpha - 0.1681\beta + 0.7227\gamma = -0.0106
\end{aligned}\right\}. \qquad (3.325)$$

Розв'язуючи звичайним методом систему (3.325), знаходимо α=–0.0266; β=–0.0466; γ=0.0036. Значить,

a=1–0.0266 = 0.9734; b= 2.05–0.0466 = 2.0034; c = –0.495 + 0.0036 =–0.4914.

Таким чином, шукана емпірична формула має вигляд

$$\widetilde{y} = e^{-0.4914x}(0.9734\sin x + 2.0034\cos x). \qquad (3.326)$$

У таблиці 3.26 приведемо дані розходження між значеннями функції $\widetilde{y}$, які знайдені за емпіричною формулою (3.326), і табличними значеннями y.

Таблиця 3.26

Характеристика точності емпіричної формули (3.326)

x	y	$\widetilde{y}$	$\Delta = y - \widetilde{y}$
	,050	,003	0, 047
,2	,944	,960	- 0,016
,6	,638	,640	- 0,002
,2	,907	,906	0, 001
,6	,423	,417	0, 006
	,028	,019	0, 007

Зауваження 3.7. Якщо значення y_i у системі (3.312) одержані в результаті вимірювань, то слід потурбуватися, щоб ці значення були записані з однаковим числом значущих цифр. У протилежному випадку – ввести відповідні ваги.

3.9. Метод статистичного аналізу економетрики

3.9.1. Метод кореляційного аналізу

Нехай проведено стохастичний експеримент, математичною моделлю якого є ймовірнісний простір $(\Omega, \Im, P)$. На ньому задано випадкову величину $\Theta \equiv \Theta(\omega) \in \mathbf{R}^1$ і випадковий вектор $\xi \equiv (\xi_1(\omega), \xi_2(\omega), \ldots, \xi_n(\omega)) \in \mathbf{R}^n$.

Означення 3.2. Залежність випадкового параметра Θ від випадкових змінних $\xi \equiv (\xi_1, \xi_2, ..., \xi_n)$ називається *кореляційною*, якщо кожній точці з області зміни змінних ξ відповідає ряд розподілу значень параметра Θ, і якщо зі змінами змінних ξ ці ряди закономірно змінюють свое положення. Закономірність зміни положень цих рядів у просторі змінних ξ описується рівнянням регресії, яке є геометричним місцем точок деякого середнього показника рядів розподілу. Для пояснення цього, без обмежень загальності, розглянемо кореляційну залежність деякої випадкової змінної величини $\Theta \equiv \Theta(\omega) \in \mathbf{R}^1$ від іншої випадкової змінної $\xi_1 \equiv \xi_1(\omega) \in \mathbf{R}^1$.

Нехай дана таблиця (3.27) спостережень значень Θ і ξ_1.

Таблиця 3.27

Спостереження значень Θ і ξ_1

№ досл. / Параметри	1	2	3	...	k	...	N-1	N
Θ	Θ_1	Θ_2	Θ_3	...	Θ_k	...	Θ_{N-1}	Θ_N
ξ_1	ξ_1	ξ_{12}	ξ_{13}	...	ξ_{1k}	...	ξ_{1N-1}	ξ_{1N}

Перенесемо дані таблиці 3.25 на координатну площину з осями Θ і ξ_1. Отримаємо, так зване, поле кореляції (рис. 3.6а). Кожному спостереженню з таблиці 1 буде відповідати деяка точка на полі кореляції.

Далі прослідкуємо за зміною положень статистичних рядів розподілу значення Θ за значеннями ξ_1. Для цього весь діапазон зміни ξ_1 розіб'ємо на ряд рівних інтервалів $\Delta\xi_1$. Всі точки, які потрапили в даний інтервал $\Delta\xi_{1i}$, віднесемо до середини даного інтервалу $\Delta\xi_{1i}$. Отримаємо трансформоване поле кореляції (рис. 3.6б). Тепер підрахуємо частинні середні арифметичні $\overline{\Theta}_i$ для кожного значення $\Delta\xi_{1i}$.

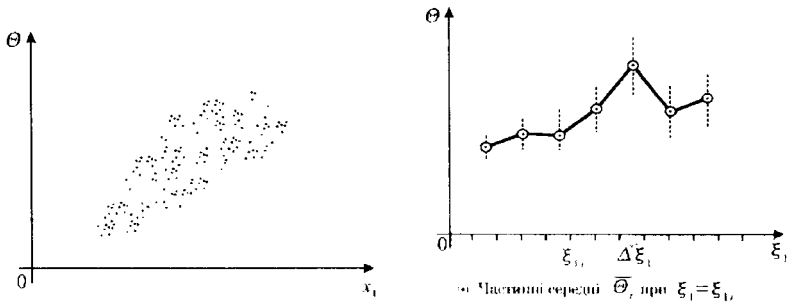
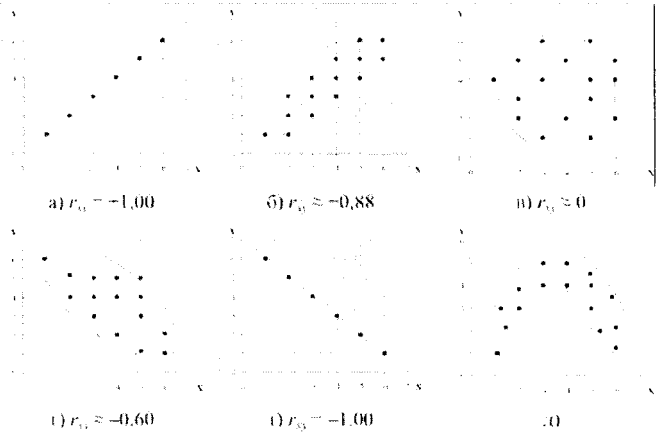


Рис. 3.6. а) Поле кореляції та б) трансформоване поле кореляції

$$\overline{\Theta}_i = \frac{\sum_{k=1}^{r_i} \Theta_{ik}}{r_i} \quad \text{при } \xi_i = \xi_{1i}. \tag{3.327}$$

де r_i – число точок, які знаходяться в інтервалі $\Delta\xi_{1i}$, причому $\sum r_i = N$, N – загальне число спостережень. Потім послідовно об'єднуємо точки $\overline{\Theta}_i$ відрізками прямих.



а) $r_{\xi\Theta} = -1{,}00$ б) $r_{\xi\Theta} \approx -0{,}88$ в) $r_{\xi\Theta} \approx 0$

г) $r_{\xi\Theta} \approx -0{,}60$ д) $r_{\xi\Theta} = -1{,}00$ е) 0

Означення 3.3. Ламана лінія, що утворена точками $\overline{\Theta}_i$ (формула (3.327)), називається *емпіричною регресією* Θ на ξ_1. Вона показує, як зміщується центр ваги рядів розподілу Θ зі збільшенням ξ_1, або як в середньому змінюється Θ зі збільшенням ξ_1. Очевидно, що більшу довіру будуть мати ті точки емпіричної лінії регресії, котрі належать інтервалам з більшою кількістю спостережень. Зі збільшенням числа спостережень емпірична лінія регресії буде звільнятися від випадкових флуктуацій, приймаючи більш правильний, закономірний вигляд.

Означення 3.4. Граничне положення емпіричної лінії регресії, до якого вона прямує при необмеженому збільшенні числа спостережень (і одночасному, але повільнішому зменшенні інтервалів $\Delta\xi$) називається *граничною теоретичною лінією регресії параметра* Θ *на* ξ_1 або просто *лінією регресії* Θ на ξ_1.

Знаходження за результатами скінченої кількості спостережень лінії регресії Θ на ξ_1 складають задачу кореляційного аналізу. Процес знаходження лінії регресії зводиться до розрахунку параметрів її рівняння. В принципі, кореляційну залежність між Θ і ξ_1 можна було б виділити через параметри щільності двовимірного розподілу. Вираз для щільності двовимірного нормального розподілу має вигляд

$$p(\xi_1;\Theta) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-R^2}} \exp\left\{-\frac{1}{2(1-R^2)}\left[\frac{(\xi_1-\overline{\xi}_1)^2}{\sigma_1^2} - \right.\right.$$
$$\left.\left. - 2R\frac{(\xi_1-\overline{\xi}_1)(\Theta-\overline{\Theta})}{\sigma_1\sigma_2} + \frac{(\Theta-\overline{\Theta})^2}{\sigma_2^2}\right]\right\}, \qquad (3.328)$$

де $\overline{\xi}_1$ і $\overline{\Theta}$ – середні арифметичні значення параметрів ξ_1 і Θ: σ_1 і σ_2 – середні квадратичні значення відхилень для параметрів ξ_1 і Θ; R – коефіцієнт кореляції, який характеризує зв'язок між параметрами ξ_1 та Θ.

У випадку, якщо R=0 , що означає відсутність кореляційного зв'язку між ξ_1 та Θ, вираз (3.328) перетворюється в добуток двох нормально розподілених незалежних випадкових величин. Якщо шляхом перетворення (3.328) знайти вираз для умовної ймовірності розподілу величини, тобто при фіксованих значеннях величини можна побачити, що умовні щільності розподілу нормальні з середнім значенням

$$\Theta\Big|_{\xi_1=\xi_{1i}} = \overline{\Theta} + R\frac{\sigma_2}{\sigma_1}(\xi_{1i} - \overline{\xi_1}) \tag{3.329}$$

і дисперсією

$$\sigma_2^2\Big|_{\xi_1=\xi_{1i}} = \sigma_2^2(1-R^2). \tag{3.330}$$

Вираз (3.329) є ні чим іншим, як рівнянням лінії регресії Θ на ξ_1, яке показує, як будуть змінюватися центри ваг рядів розподілу Θ при зміні ξ_1. Таким чином, знаючи параметри щільності двовимірного нормального розподілу σ_1, σ_2, R, $\overline{\xi_1}$ і $\overline{\Theta}$ – можна визначити параметри рівняння, яке описує кореляційну залежність між ξ_1 і Θ. Аналогічний результат для параметрів лінії регресії дає метод найменших квадратів, зміст якого полягає в наступному. Якщо для кожного значення ξ_{1i} величина Θ нормально розподілена, то найкращі оцінки для коефіцієнтів рівняння лінії регресії одержуються при виконанні умови

$$\lambda = \sum_{i=1}^{N}(\Theta_{ie} - \Theta_{ip})^2 = \min, \tag{3.331}$$

тобто, сума квадратів відхилень спостережуваних значень Θ_{ie} від значень Θ_{ip}, обчислених за розрахунковим рівнянням

$$\Theta = a_0 + a_1\xi_1 \tag{3.332}$$

повинна бути найменшою. Рівняння (3.332) – це видозмінене рівняння (3.329).

Оцінки для коефіцієнтів a_0 і a_1 знайдені цим методом, будуть найкращими у тому розумінні, що вони розподілені нормально, з середнім, яке рівне отриманим значенням, і з найменшими можливими дисперсіями. Більш детальне застосування методу найменших квадратів для знаходження коефіцієнтів рівняння зв'язків розглянемо далі. Викладені принципи справедливі й тоді, коли досліджується залежність Θ від двох або більшої кількості змінних

$$\Theta = a_0 + a_1\xi_1 + a_2\xi_2 + \ldots + a_n\xi_n. \tag{3.333}$$

У випадку множинної кореляції розглядається уже не лінія кореляції, а відповідна площина або гіперплощина регресії й щільність багатовимірного розподілу. Як було сказано, коефіцієнт кореляції R виражає міцність досліджуваних зв'язків. Повна дисперсія σ_Θ^2 значень Θ навколо середнього значення $\overline{\Theta}$ складається з частини $\sigma_\Theta^2 R^2$, яка обумовлена заміною змінних в рівнянні (3.333), і залишку $\sigma_\Theta^2(1-R^2)$, який не залежить від цих змінних і визначається дією невизначених факторів. Фізично коефіцієнт кореляції характеризує ту долю коливань параметра Θ, яка викликана дією параметрів, які можна контролювати.

Висновок 3.1. Чим більше значення R, тобто тісніший кореляційний зв'язок, тим значніше знайдена залежність проявляється серед зовнішніх дій, тим точніше за даними значеннями $\xi \in \mathbf{R}^n$ можна передбачити значення $\Theta \in \mathbf{R}^1$.

Висновок 3.2. Значення R знаходиться в межах $-1 \le R \le +1$. Якщо $R = |1|$, то зв'язок є функціональний, тобто враховано ті з параметрів, від яких у тій чи іншій мірі залежить Θ. Якщо R=0, то кореляційної залежності між досліджуваними параметрами немає. Якщо ж $0 \le R \le |1|$, то говорять про наявність сильнішої або слабшої кореляційної залежності. У розпорядженні дослідника є цілий ряд формул для обчислення R, якими він може скористатися. Найпростіше оцінюється сила парної кореляційної залежності. Найбільш вживаною при цьому є формула

$$R_{\Theta\xi} = \frac{\sum_{k=1}^{N}(\Theta_k - \overline{\Theta})(\xi_k - \overline{\xi})}{N\sigma_\Theta\sigma_\xi}, \qquad (3.334)$$

де $\overline{\Theta}$ – вибіркове середнє значення параметра Θ; ξ – вибіркове середнє значення параметра ξ; N – число спостережень; $\sigma_\Theta, \sigma_\xi$ – середнє квадратичне значення відповідних змінних.

Мірою сили зв'язку у випадку множинної кореляції є множинний коефіцієнт кореляції $R_{\Theta,\xi_1,\xi_2,\ldots,\xi_n}$, який може бути визначений через коефіцієнти парної кореляції параметрів, які присутні у рівнянні. Наприклад, для рівняння з двома змінними ξ_1 і ξ_2 коефіцієнт множинної кореляції визначається за формулою

$$R_{\Theta\xi_1\xi_2} = \sqrt{\frac{r_{\Theta\xi_1}^2 + r_{\Theta\xi_2}^2 - 2r_{\Theta\xi_1}\cdot r_{\Theta\xi_2}\cdot r_{\xi_1\xi_2}}{1 - r_{\xi_1\xi_2}^2}}, \qquad (3.335)$$

де $r_{\Theta\xi_1}; r_{\Theta\xi_2}; r_{\xi_1\xi_2}$ – коефіцієнти парної кореляції для відповідних параметрів.

Очевидно, що коефіцієнт кореляції, який оцінює силу зв'язку функції з m змінними, не може бути більшим ніж коефіцієнт кореляції, який оцінює силу зв'язку функції з m+1 змінними, бо враховуючи ще одну змінну є нагода більш повніше характеризувати функцію залежності. Використання коефіцієнта кореляції як міри сили зв'язку можливе лише у тому випадку, коли досліджувані випадкові змінні мають багатовимірний нормальний розподіл. У протилежному випадку інтерпретація коефіцієнта кореляції викликає труднощі і є ненадійною. Ця вимога значно звужує сферу застосування методу кореляційного аналізу для практичних застосувань економетрики, обмежуючи тільки одержанням рівнянь ліній регресій. Але саме випадок лінійної кореляційної залежності є найбільш важливим і поширеним у практичних дослідженнях, тому що за існуючими технічними обмеженнями параметри об'єкту можуть змінюватися у порівняно вузьких межах і на цих ділянках криволінійні залежності можуть бути з достатньою точністю апроксимовані залежностями прямолінійними. Прослідкуємо більш детально, як використовується метод множинної лінійної кореляції для отримання математичного опису складного об'єкту, який характеризується великим числом змінних параметрів.

Нехай маємо таблицю даних, яка відображає поведінку незалежних контрольованих змінних ξ і відповідні спостереження деякої змінної Θ, яка характеризує ефективність роботи об'єкту (таблиця 3.28). Допустимо, що між Θ і $\xi_1,\xi_2,...,\xi_n$ існує лінійна кореляційна залежність вигляду

$$\Theta = a_0 + a_1\xi_1 + a_2\xi_2 + ... + a_n\xi_n . \tag{3.336}$$

Потрібно визначити коефіцієнти $a_0,a_1,a_2,...,a_n$ рівняння (3.336). Спочатку всі змінні переведемо у стандартизований масштаб за формулами

$$t_{\Theta_i} = \frac{\Theta_i - \overline{\Theta}}{\sigma_\Theta};\ t_{ji} = \frac{\xi_{ji} - \overline{\xi}_j}{\sigma_{\xi_j}}. \tag{3.337}$$

де $t_{\Theta_i}; t_{ji}$ – значення відповідної змінної у стандартизованому масштабі; $\overline{\Theta}$ і $\overline{\xi}$ – середні значення відповідних змінних; σ_Θ і σ_{ξ_j} – середні квадратичні відхилення відповідних змінних. Неважко показати, що для всіх досліджуваних змінних у стандартизованому масштабі:

$$\overline{t} = 0 \text{ і } \sigma_t = 1 . \tag{3.338}$$

Таблиця 3.28

Поведінка незалежних контрольованих змінних ξ

№ спостережень	Значення параметрів, які спостерігаються							
	Θ	ξ_1	ξ_2	ξ_3	...	ξ_k	...	ξ_n
1	Θ_1	ξ_{11}	ξ_{21}	ξ_{31}	...	ξ_{k1}	...	ξ_{n1}
2	Θ_2	ξ_{12}	ξ_{22}	ξ_{32}	...	ξ_{k2}	...	ξ_{n2}
3	Θ_3	ξ_{13}	ξ_{23}	ξ_{33}	...	ξ_{k3}	...	ξ_{n3}
...	...	...	...	...	...	...	...	...
k	Θ_k	ξ_{1k}	ξ_{2k}	ξ_{3k}	...	ξ_{kk}	...	ξ_{nk}
...	...	...	...	...	...	...	...	...
N	Θ_N	ξ_{1N}	ξ_{2N}	ξ_{3N}	...	ξ_{kN}	...	ξ_{nN}

У зв'язку з цим спрощуються формули основних показників лінійної кореляції. Коефіцієнти кореляції між стандартизованими параметрами визначаються формулою

$$r_{\xi_j\xi_k} = r_{jk} = \frac{1}{N}\sum_{i=1}^{N} t_{jk}t_{ki}, \quad r_{\Theta\xi_j} = r_{\Theta j} = \frac{1}{N}\sum_{i=1}^{N} t_{\Theta i}t_{ji}. \tag{3.339}$$

Можна показати, що рівняння регресії у стандартизованому масштабі не має вільного члена і приймає вигляд

$$t_\Theta = \beta_1 t_1 + \beta_2 t_2 + ... + \beta_j t_j + ... + \beta_n t_n , \tag{3.340}$$

де $t_1,t_2,...,t_j,...,t_n$ – стандартизовані значення змінних $\xi_1,\xi_2,...,\xi_j,...,\xi_n$; t_Θ – значення величини Θ у стандартизованому масштабі. Коефіцієнти рівняння (3.340) знаходимо методом найменших квадратів за виконання умови

$$\lambda = \sum_{i=1}^{N}(t_{\Theta i_e} - t_{\Theta i_p})^2 \Rightarrow \min . \tag{3.341}$$

Умова мінімуму для співвідношення (3.341) визначається наступними виразами

$$\frac{\partial\lambda}{\partial\beta_1}=0;\ \frac{\partial\lambda}{\partial\beta_2}=0;\ \ldots;\ \frac{\partial\lambda}{\partial\beta_n}=0\,. \tag{3.342}$$

Підставляючи у (3.341) замість $t_{\Theta p}$ його значення з (3.340), одержимо:

$$\lambda=\sum_{i=1}^{N}(t_{\Theta ic}-\beta_1 t_{1i}-\beta_2 t_{2i}-\ldots-\beta_j t_{ji}-\ldots-\beta_n t_{ni})^2\,. \tag{3.343}$$

Використовуючи (3.342), знаходимо

$$\left.\begin{aligned}
&2\sum_{i=1}^{N}(t_{\Theta ic}-\beta_1 t_{1i}-\beta_2 t_{2i}-\ldots-\beta_j t_{ji}-\ldots-\beta_n t_{ni})t_{1i}=0\\
&2\sum_{i=1}^{N}(t_{\Theta ic}-\beta_1 t_{1i}-\beta_2 t_{2i}-\ldots-\beta_j t_{ji}-\ldots-\beta_n t_{ni})t_{2i}=0\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&2\sum_{i=1}^{N}(t_{\Theta ic}-\beta_1 t_{1i}-\beta_2 t_{2i}-\ldots-\beta_j t_{ji}-\ldots-\beta_n t_{ni})t_{ni}=0
\end{aligned}\right\}. \tag{3.344}$$

Розкриємо дужки у кожному рівнянні системи (3.344) і помножимо ліву і праву частини на відповідне $\frac{1}{N}t_{jk}$. При кожному коефіцієнті β_i отримаємо за (3.339) коефіцієнти парної кореляції r_{jk}. Беручи до уваги, що коефіцієнти

$$r_{jj}=\frac{1}{N}\sum_{i=1}^{N}t_{ji}^2=\sigma_t^2=1\,. \tag{3.345}$$

отримаємо наступну СЛАР відносно невідомих $\beta_1.\beta_2.\ldots.\beta_j.\ldots.\beta_n$

$$\left.\begin{aligned}
r_{\Theta 1}&=\beta_1+\beta_2 r_{12}+\beta_3 r_{13}+\ldots+\beta_j r_{1j}+\ldots+\beta_n r_{1n}\\
r_{\Theta 2}&=\beta_1 r_{21}+\beta_2+\beta_3 r_{23}+\ldots+\beta_j r_{2j}+\ldots+\beta_n r_{2n}\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
r_{\Theta j}&=\beta_1 r_{j1}+\beta_2 r_{j2}+\beta_3 r_{j3}+\ldots+\beta_j+\ldots+\beta_n r_{jn}\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
r_{\Theta n}&=\beta_1 r_{n1}+\beta_2 r_{n2}+\beta_3 r_{n3}+\ldots+\beta_j r_{nj}+\ldots+\beta_n
\end{aligned}\right\} \tag{3.346}$$

Для складних виробничих об'єктів СЛАР (3.346), як правило, є системою досить високого порядку і для її розв'язування необхідно використовувати сучасні комп'ютери з потужним програмними засобами. При цьому можуть використовуватися як точні так і наближені методи. СЛАР (3.346) у матричному вигляді записують так

$$A=B\cdot C\,. \tag{3.347}$$

$$A=\begin{Vmatrix} r_{\Theta 1} \\ r_{\Theta 2} \\ \dots \\ r_{\Theta n} \end{Vmatrix};\ B=\begin{Vmatrix} 1 & r_{12} & r_{13} & \dots & \dots & r_{1n} \\ r_{21} & 1 & r_{23} & \dots & \dots & r_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ r_{j1} & r_{j2} & r_{j3} & \dots & 1 & r_{jn} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & r_{n3} & \dots & \dots & 1 \end{Vmatrix};\ C=\begin{Vmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_j \\ \dots \\ \beta_n \end{Vmatrix}. \qquad (3.348)$$

Розв'язування системи (3.347) – це знаходження компонент матриці C у (3.348). Помножимо обидві частини матричного рівняння на обернену матрицю B^{-1}

$$B^{-1}A = B^{-1}BC\,. \qquad (3.349)$$

Оскільки $B^{-1}B = 1$, то розв'язок системи (3.347) можна записати так

$$C = B^{-1}A\,. \qquad (3.350)$$

Операція знаходження оберненої матриці полягає в тому, щоб кожен елемент матриці B, розміщений у i-тому рядку й j-тому стовпці, замінити на інший елемент, рівний $\frac{d_{ij}}{\det B}$, де $\det B$ – визначник матриці B, а d_{ij} – алгебраїчне доповнення даного елемента. Для виконання операції знаходження обернених матриць високого порядку комп'ютери мають спеціальні програмні продукти, які використовують методи статистичного моделювання. У результаті розв'язку СЛАР (3.346) визначаються коефіцієнти $\beta_1,\beta_2,\dots,\beta_j,\dots,\beta_n$ стандартизованого рівняння регресії. Знайдені коефіцієнти рівняння регресії необхідно перевірити на значущість їх відмінності від нуля. Оскільки в (3.346) $\beta_1,\beta_2,\dots,\beta_j,\dots,\beta_n$ входить як лінійна комбінація Θ. то з нормальності розподілу Θ випливає нормальність розподілу $\beta_1,\beta_2,\dots,\beta_j,\dots,\beta_n$. Це дозволяє використовувати для визначення залежності коефіцієнтів рівняння відомий критерій Ст'юдента. Необхідні для застосування критерію Ст'юдента величини середніх квадратичних значень для коефіцієнтів $\beta_1,\beta_2,\dots,\beta_j,\dots,\beta_n$ отримують із співвідношення

$$\sigma^2_{\beta_i} = \sigma^2_{\Theta}\alpha_j\,. \qquad (3.351)$$

де σ^2_{Θ} – дисперсія вихідного параметра Θ; α_j – діагональні члени оберненої матриці B^{-1}. Якщо в результаті перевірки якийсь коефіцієнт виявиться статистично не значимим, його замінюють у рівнянні нулем.

Із розв'язку системи (3.346) одержимо коефіцієнти $\beta_1,\beta_2,\dots,\beta_j,\dots,\beta_n$ стандартизованого рівняння регресії (3.340). За цими коефіцієнтами можна якісно оцінити вплив окремих змінних факторів. Оскільки всі змінні величини виражені в стандартизованому масштабі, то знайдені коефіцієнти $\beta_1,\beta_2,\dots,\beta_j,\dots,\beta_n$ показують порівняльну силу впливу зміни кожної змінної на зміну вихідного параметра. Вивчаючи

знаки та співвідношення між величиною коефіцієнтів $\beta_1, \beta_2, ..., \beta_j, ..., \beta_n$, можна визначити, які змінні істотно впливають на процес і у якому напрямку. Для практичного використання одержаного рівняння (3.340) необхідно повернутися до звичайного масштабу. Коефіцієнти рівняння (3.333) визначаються за формулою

$$a_j = \beta_j \frac{\sigma_\Theta}{\sigma_j}; \; a_\Theta = \overline{\Theta} - \sum_{i=1}^{N} a_j \xi_{ji} . \tag{3.352}$$

Підставивши знайдені коефіцієнти у рівняння (3.333), отримаємо лінійну математичну модель складного об'єкту. Визначивши параметри зв'язку, необхідно оцінити їх силу. Для обчислення коефіцієнта множинної кореляції $R_{\Theta,\xi_1,\xi_2,...\xi_n}$ скористаємося формулою

$$R_{\Theta,\xi_1,\xi_2,...\xi_n} = \sqrt{\beta_1 r_{\Theta 1} + \beta_2 r_{\Theta 2} + ... \beta_n r_{\Theta n}} . \tag{3.353}$$

В теорії кореляційного аналізу розроблені прийоми, які дозволяють оцінити точність і достовірність одержаного значення коефіцієнта кореляції в залежності від числа проведених спостережень. По-перше, обчислене значення коефіцієнта кореляції підлягає поправці на систематичну похибку. Розглянемо її походження.

Нехай дослідник, маючи певне число спостережень, збільшує число незалежних змінних у рівнянні регресії. При цьому коефіцієнт кореляції, очевидно, буде зростати. Розглядаючи крайній випадок, легко бачити, що якщо б число параметрів n було рівним числу спостережень N, то гіперплощина регресії, побудована методом найменших квадратів, пройшла б через всі N точок і коефіцієнт кореляції був би рівний 1, хоча в дійсності зв'язок може бути слабким, що одразу виявляється при збільшенні числа спостережень. Таким чином, чим ближче n до N, тим сильніше перебільшується тіснота зв'язку. Це перебільшення виправляється коректуванням.

Формула коректування має вигляд

$$R'_{\Theta,\xi_1,\xi_2,...\xi_n} = \sqrt{1 - (1 - R^2)\frac{N-1}{N-n}} . \tag{3.354}$$

де $R'_{\Theta,\xi_1,\xi_2,...\xi_n}$ – скоректоване значення R; N – число спостережень; n – число змінних параметрів у рівнянні регресії.

Означення 3.5. Величину

$$\alpha = N - n . \tag{3.355}$$

назвемо *числом степенів вільності*, яке характеризує достатність зібраного експериментального матеріалу. Вважається, що зв'язок не буде перебільшений, якщо

$$\frac{N-n}{n} = \frac{\alpha}{n} \geq 20 . \tag{3.356}$$

З формули (3.354) видно, що виправлення є сенс застосовувати, коли кількість невідомих порівняна з кількістю спостережень. Реальність виведеного зв'язку встановлюється шляхом порівняння коефіцієнта кореляції R і його середнього квадратичного значення σ_R. Коли справджується рівність R=0, то справедливе співвідношення

$$\sigma_R = \frac{1}{\sqrt{N-1}}. \tag{3.357}$$

Висновок 3.3. Можна показати, що якщо виконується нерівність

$$\frac{|R|}{\sigma_R} > 3, \tag{3.358}$$

то з імовірністю $P=0.99$ можна вважати R значимим і зв'язок реальним. Протилежний випадок говорить про те, що R відрізняється від «0» випадково, і реальний зв'язок не підтверджується. Встановивши реальність досліджуваного зв'язку, можна перейти до оцінки похибки у визначенні R, які допускаємо, коли базуємося на обмеженому числі спостережень. Для оцінки довірчих інтервалів коефіцієнта кореляції Фішером був запропонований наступний критерій. Коефіцієнт кореляції R прирівнюється до гіперболічного тангенса деякої величини z

$$R = thz, \tag{3.359}$$

$$z = \frac{1}{2}\ln\frac{1+R}{1-R}. \tag{3.360}$$

Щільність розподілу величини z є майже незмінною за формою для змінного числа спостережень N і зі збільшенням N швидко наближається до нормального розподілу з середнім значенням

$$\bar{z} = \frac{1}{2}\ln\frac{1+R}{1-R} + \frac{R}{2(N-1)} \tag{3.361}$$

і дисперсією

$$\sigma_z^2 = \frac{1}{N-3}. \tag{3.362}$$

Щоб позбутися необхідності обчислювати z у кожному конкретному випадку, складені таблиці, які дають можливість знаходити значення R за даними значеннями z і навпаки, або одержувати z методами статистичного моделювання.

Оскільки розподіл z нормальний, то можна записати

$$P(-\alpha\sigma_z < z - z_0 < \alpha\sigma_z) = \Phi(\alpha), \tag{3.363}$$

$$\Phi(\alpha) = \frac{1}{2\pi}\int_0^{\alpha} e^{-\frac{u^2}{2}}\,du \tag{3.364}$$

інтеграл Лапласа, табульовані значення якого можна знайти у будь-якій книзі з теорії ймовірності.

Співвідношення (3.361) – це ймовірність того, що модуль різниці $(z - z_0)$ не перевищить величини $\alpha\sigma_z$, де z_0 відповідає справжньому значенню коефіцієнта кореляції R_0. Використання критерію Фішера для оцінки меж істинного значення коефіцієнта кореляції зводиться до наступного алгоритму.

1. Відповідно до обчисленого значення R знаходимо z.

2. Одночасно за формулою (3.362) обчислюємо дисперсію σ_z^2.
3. Задаючи ймовірність $P(-\alpha\sigma_z < z - z_0 < \alpha\sigma_z) = \Phi(\alpha) \geq 0.95$ за таблицями $\Phi(\alpha)$ (див. додаток) знаходимо відповідне σ. Тоді, за (3.363), з прийнятою ймовірністю P можна стверджувати, що значення z_0 знаходиться в межах

$$z - \alpha\sigma_z < z_0 < z + \alpha\sigma_z. \tag{3.365}$$

4. Обчислюючи $z - \alpha\sigma_z = z_1$ і $z + \alpha\sigma_z = z_2$ за таблицями, здійснюємо зворотний перехід і знаходимо відповідно R_1 і R_2. Тоді з ймовірністю P можна стверджувати, що істинне значення коефіцієнта кореляції R_0 знаходиться в межах

$$R_1 < R_0 < R_2. \tag{3.366}$$

Важливо підкреслити той факт, що викладений метод оцінки інтервалу зміни істинного коефіцієнта кореляції справедливий лише в тому випадку, якщо досліджувані змінні або їх перетворені значення нормально розподілені. На цьому перевірку достовірності знайденої кореляційної залежності можна вважати завершеною. У випадку, коли наперед відомо, що досліджуваний показник тісно пов'язаний з деякими змінними, включеними у лінійне рівняння, але коефіцієнт кореляції виявився малим, це може служити ознакою того, що лінійна форма вибрана невірно. Таким чином, причину малості значення R можна пояснити двояким чином:

- Вибрана лінійна форма зв'язку відповідає дійсно існуючій, але самий зв'язок слабкий і у значній мірі порушується дією неврахованих факторів. Збільшення тісноти зв'язку у цьому випадку можна добитися введенням у рівняння регресії додаткових змінних.

- Справжня закономірність має нелінійний характер і заміна її прямолінійною залежністю вносить надто велику похибку. У цьому випадку слід перейти до методів регресійного аналізу.

Важливо мати на увазі те, що кореляційний аналіз не дає інформацію про реально існуючі кореляційні зв'язки. Форму зв'язку, тобто порядок рівняння і вигляд кожного її члена дослідник визначає самостійно на основі різних припущень і гіпотез.

Про тип теоретичної ліній регресії можна судити за зовнішнім виглядом емпіричного графіку (рис. 3.6б). Цей прийом базується па законі великих чисел.

У думці позбавляючись від тих стрибків, які можуть бути випадковими на емпіричному графіку, дослідник може визначити порядок кривої регресії. Потім підбираються декілька рівнянь підходящих типів кривих. З них перевагу віддаємо тій, котра має менше параметрів. Однак цей шлях можливий практично тільки при дослідженій кореляційної залежності від однієї змінної величини.

Вигляд теоретичної лінії регресії можна деколи визначити, виходячи з теоретичних міркувань про матеріальну природу ознак Θ і $\xi=(\xi_1, \xi_2, \ldots, \xi_n)$.

В цьому випадку на допомогу беруть фізичні і хімічні закони, які лежать в основі досліджуваного процесу. Підсумовуючи викладені вище принципи, методи і прийоми використання кореляційного аналізу для одержання математичної моделі складних об'єктів, можна зробити наступні висновки:

1. Метод кореляційного аналізу є потужним засобом дослідження складних об'єктів управління, який дозволяє встановити існування кореляційного зв'язку між досліджуваними параметрами, оцінити силу зв'язку і визначити параметри рівняння, яке описує знайдені зв'язки.

2. Ефективність використання методу кореляційного аналізу істотно залежить від виконання ряду умов, які в обстановці реального виробництва іноді важко виконати. Основними і найсильнішими вимогами є існування багатовимірного нормального закону розподілу для досліджуваних величин (n-вимірного, якщо число змінних n). Тільки при виконанні цієї умови одержані коефіцієнти кореляції допускають надійну інтерпретацію, як показники сили зв'язку, і можуть бути визначені їх довірчі інтервали.

3. З вищесказаного випливає, що, практично, прийоми і методи кореляційного аналізу можуть плідно використовуватися для одержання лінійних моделей складних реальних об'єктів, тому що при нелінійній формі рівняння поняття багатовимірного розподілу змінних втрачає сенс.

Як уже було сказано, в більшості практичних випадків не лінійність реальних об'єктів може задовільно апроксимуватися лінійною залежністю. Однак якщо навіть у вузьких діапазонах зміни змінних об'єкт виявляється істотно нелінійним, то застосувати для його дослідження методи кореляційного аналізу досить важко. В цьому випадку використовують методи регресійного аналізу, який розглянемо нижче.

3.9.2. Метод регресійного аналізу

Будемо вважати, що всі досліджувані випадкові величини $\xi_1, \xi_2, \ldots, \xi_n$, надалі ξ мають деяку довільну щільність розподілу, але вихідна випадкова величина Θ має нормальну щільність розподілу. Це рівне вимозі одновимірного нормального розподілу. Умови появи одновимірного нормального розподілу менш жорсткі, ніж для багатовимірного розподілу. Їх виконання у виробничій практиці є досить реальним. Для використання критерію, який оцінює значимість коефіцієнтів знайденого рівняння регресії, потрібно, щоб виконувалися наступна умова: всі розподіли Θ для різних фіксованих значень змінних ξ повинні мати однакову дисперсію. Це припущення не таке жорстке, як може здатися на перший погляд, і практично завжди справджується. Вказані припущення справедливі як для лінійних, так і для нелінійних задач. Особливо ефективні методи регресійного аналізу при побудові нелінійних математичних моделей складних реальних об'єктів. Розглянемо, як використовується прийоми регресійного аналізу для побудови математичної моделі складного об'єкту. Припустимо, що між вихідною величиною Θ та змінними параметрами ξ вигляд зв'язку відомий.

Припустимо, що складний процес може бути описаний рівнянням 2-го степеня

$$\Theta = b_0 + \sum_{i=1}^{n} b_i \xi_i + \sum_{i=1}^{n} b_{ii} \xi_i^2 + \sum_{i,j=1}^{C_n^2} b_{ij} \xi_i \xi_j \,. \qquad (3.367)$$

де N – число експериментальних спостережень, а n – число змінних.

Розглядаючи в рівнянні (3.367) члени другого порядку як деякі нові незалежні змінні ε_{ij}, рівняння (3.367) можна звести до лінійного вигляду

$$\Theta = b_0 + \sum_{i=1}^{n} b_i \xi_i + \sum_{i=1}^{n} b_{ii} \varepsilon_{ii} + \sum_{i,j=1}^{C_n^2} b_{ij} \varepsilon_{ij} \,. \qquad (3.368)$$

Рівняння (3.368) має m змінних, де

$$m = 2n + C_n^2 . \tag{3.369}$$

Коефіцієнти рівняння (3.368) знаходяться методом найменших квадратів. Попередньо всі змінні ε_{ij} стандартизуються. Система нормальних рівнянь має вигляд

$$\left.\begin{aligned} \sum tt_1 &= \beta_1 \sum t_1^2 + \beta_2 \sum t_1 t_2 + \ldots + \beta_m \sum t_1 t_m \\ \sum tt_2 &= \beta_1 \sum t_2 t_1 + \beta_2 \sum t_2^2 + \ldots + \beta_m \sum t_2 t_m \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum tt_m &= \beta_1 \sum t_m t_1 + \beta_2 \sum t_m t_2 + \ldots + \beta_m \sum t_m^2 \end{aligned}\right\}. \tag{3.370}$$

Тут t – значення виходу Θ у стандартизованому масштабі; t_i, t_j – значення незалежних змінних рівняння (3.368) у стандартизованому масштабі, а межі сумування такі ж, як і в (3.368). Як бачимо, регресійний аналіз не потребує знаходження парних коефіцієнтів кореляції, котрі при заданих припущеннях не мають сенсу. Для розв'язку системи (3.370) та перевірки значущості отриманих коефіцієнтів використовуються ті ж самі прийоми й методи, що і для кореляційного аналізу. Для практичного використання знайденого рівняння необхідно перевести його в натуральний масштаб за формулами (3.352). Для перевірки адекватності, з якою узгоджується рівняння (3.367) з досліджуваним складним об'єктом, обчислюється середній квадрат відхилення спостережених значень Θ_{ic} від значень Θ_{ip}, обчислених за формулою (3.367) підстановкою ξ_k і ε_k

$$\sigma_{\Theta_p}^2 = \frac{\sum_{i=1}^{N} (\Theta_{ic} - \Theta_{ip})^2}{N} . \tag{3.371}$$

Значення σ_{Θ_p} порівнюється зі значенням дисперсії спостережуваних Θ_{ic} відносно середнього $\overline{\Theta}$. Адекватність відтворення є задовільною, якщо $\sigma_{\Theta_p}^2$ виявиться хоча б на порядок менше дисперсії спостережених значень виходу σ_{Θ}^2 . В цьому випадку передбачення за знайденим рівнянням (3.367) буде набагато кращим ніж за середнім $\overline{\Theta}$. Різновидом регресійного аналізу є метод одержання математичного опису, запропонований у 1959 році Брандоном. В його основі, крім загальних припущень регресійного аналізу, лежить припущення, що досліджувані змінні слабо пов'язані між собою, і шукане вихідне рівняння зв'язку можна подати у вигляді

$$\Theta = \overline{\Theta} \cdot f_1(\xi_1) \cdot f_2(\xi_2) \cdot f_3(\xi_3) \cdot \ldots \cdot f_n(\xi_n) . \tag{3.372}$$

де $\overline{\Theta}$ – середнє значення виходу. Для ілюстрації методу розглянемо приклад, де Θ залежить від дії 4-х змінних параметрів

$$\Theta = \overline{\Theta} \cdot f_1(\xi_1) \cdot f_2(\xi_2) \cdot f_3(\xi_3) \cdot f_4(\xi_4) . \tag{3.373}$$

Для простоти обчислення береться 6 виробничих станів (табл. 3.29) і алгоритм обчислення буде таким.

1 Спочатку визначається величина $\overline{\Theta}$

$$\overline{\Theta} = \frac{1}{N}\sum_{i=1}^{N}\Theta_i, \tag{3.374}$$

де N – число експериментальних точок.

2. Одержані експериментально значення виходу Θ_i діляться на $\overline{\Theta}$, утворивши параметр φ_{1i}

$$\varphi_{1i} = \frac{\Theta_i}{\overline{\Theta}}. \tag{3.375}$$

3. Будується кореляційне поле для залежності $\varphi_1 = f_1(\xi_1)$ і визначається вигляд цієї залежності. Часто для спрощення розрахунків приймається лінійне наближення

$$f_1(\xi_1) = a_1 + b_1\xi_1. \tag{3.376}$$

Таблиця 3.29

Вхідні дані для обчислення кореляційного поля

Виробничі стани	Θ_i	ξ_1	ξ_2	ξ_3	ξ_4
1	0,87	0,2	26	4,2	10
2	0,93	0,6	22	4,1	15
3	0,96	0,9	21	4,3	35
4	0,94	1,0	19	3,7	40
5	0,93	0,8	25	4,4	50
6	0,89	0,5	23	4,5	75

Звичайно, можливе використання і більш складної криволінійної залежності.

Методом найменших квадратів обчислюються параметри a_1 і b_1 та будується лінія регресії $f_1(\xi_1)$ (рис. 3.7а).

4. Далі знаходять всі значення функції $f_1(\xi_1)$ для прийнятих значень змінної ξ_1 і визначається новий параметр

$$\varphi_2 = \frac{\varphi_1}{f_1(\xi_1)} = f_2(\xi_2)\cdot f_3(\xi_3)\cdot f_4(\xi_4), \tag{3.377}$$

який залежить тепер тільки від 3 змінних.

5. Потім будується графік залежності $\varphi_2 = f_2(\xi_2)$, за яким визначається рівняння лінії регресії $f_2(\xi_2)$ (рис. 3.7б)

6. Після цього обчислюється значення наступного параметра φ_3

$$\varphi_3 = \frac{\varphi_2}{f_2(\xi_2)} = f_3(\xi_3)\cdot f_4(\xi_4), \tag{3.378}$$

який залежить уже від 2-х змінних. Будується кореляційне поле $\varphi_3 = f_3(\xi_3)$ і визначається вигляд функції як краще наближення до експериментальних точок (рис. 3.7в). Такий розрахунок продовжується до тих пір, поки не будуть визначені всі функції $f_i(\xi_i)$. У нашому прикладі останнім обчислюється параметр φ_4

$$\varphi_4 = \frac{\varphi_3}{f_3(\xi_3)} = f_4(\xi_4). \tag{3.379}$$

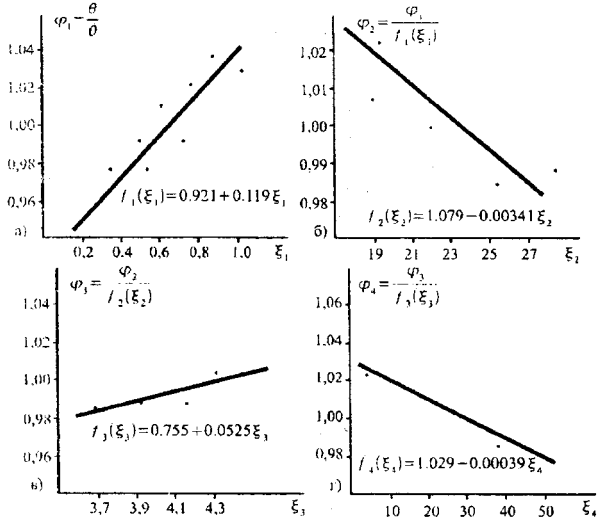


Рис. 3.7. Метод Брандона

Зауваження 3.8. Для оцінки правильності вибраної форми подання об'єкту можна обчислити значення функції

$$\varphi_{5i} = \frac{\Theta_i}{\bar{\Theta} \cdot f_1(\xi_1) \cdot f_2(\xi_2) \cdot f_3(\xi_3) \cdot f_4(\xi_4)}. \qquad (3.380)$$

Відхилення φ_{5i} від одиниці є оцінкою якості вибраної апроксимації. У таблиці 3.30 подані окремі етапи розв'язку задачі для випадку 4-х змінних. Результат подають у вигляді нелінійного рівняння, перетвореного до зручного вигляду

$$\Theta = 0.728(1+0.129\xi_1)(1-0.0136\xi_2)(1+0.0678\xi_3)(1-0.00037\xi_4).. \qquad (3.381)$$

Зрозуміло, що на практиці знаходження рівняння множинної регресії пов'язане з опрацюванням за викладеним вище методом значно більшого вихідного матеріалу. Об'єм обчислень збільшується, однак це успішно долають сучасні комп'ютери.

Таблиця 3.30

Результат застосування методу Брандона

Виробничі стани	φ_1	$f_1(\xi_1)$	φ_2	$f_2(\xi_2)$	φ_3	$f_3(\xi_3)$	φ_4	$f_4(\xi_4)$	φ_5
1	0,949	0,943	1,006	0.990	1,016	0,996	1,02	1,024	0,996
2	1,007	0,992	1,015	1,004	1,011	0,991	1,02	1,019	1,001
3	1,044	1,028	1,028	0,008	1,008	1,001	1,007	1,005	1,002
4	1,026	1,040	1,040	0,014	0,972	0,970	1,002	1,001	1,001
5	1,007	1,016	1,016	0.994	0,998	1,006	0,992	0,944	0,998
6	0,958	0,980	0,980	1,001	0,987	0,012	0,975	0,997	0,998

При використанні методу Брандона істотним є питання про степінь впливу окремих змінних на хід процесу. Місцезнаходження кожної функції $f_i(\xi_i)$ у формулі (3.373), а відповідно, і послідовність розв'язування задачі визначається важливістю відповідної змінної ξ_i: чим важливіша вона в процесі, тим меншим повинен бути порядковий номер f_i. Якщо ж порівняльний вплив змінних $\xi_1, \xi_2, ..., \xi_n$ залишився невідомим або невірно виявленим і, отже, неправильно задана структура рівняння (3.373), використання описаного методу може привести до хибних висновків.

Приклад 3.19. Розглянемо процес полімеризації етилену в трубчастому реакторі під високим тиском. Знайти математичну залежність продуктивності реактора від параметрів, які визначають хід процесу. Нехай розглядаються такі параметри: тиск p, температура t, концентрація C і витрата етилену V.

Розв'язок. Вихідні дані для розрахунків одержано в результаті спеціальних розрахунків. Продуктивність реактора встановлювалася зважуванням продукту через кожну годину. Тиск, температура і витрати газу усереднювалися за 1 годину за діаграмами реєструючи приладів. Концентрація кисню визначалася двічі на годину. Було зроблено 200 вимірювань. Залежність продуктивності реактора від основних параметрів процесу подана у вигляді

$$Q = Q_0 \cdot f_1(p) \cdot f_2(t) \cdot f_3(C) \cdot f_4(V). \qquad (3.382)$$

В результаті побудови кореляційних полів було встановлено, що лінія регресії $f_1(p)$ представляється кривою параболічного типу

$$f_1(p) = a_0 + a_1 p + a_2 p^2. \qquad (3.383)$$

Інші функції $f_2(t); f_3(C); f_4(V)$ були вибрані лінійними

$$f_i(\xi_i) = a_i + b_i \xi_i. \qquad (3.384)$$

Обчислення проведено у середовищі MATLAB. Одержана формула має вигляд

$$Q = 13 \cdot 10^{-13}(211 + 0.33p - 1.16 \cdot 10^{-4} p^2)(t - 112)(C + 55)(V + 587). \qquad (3.385)$$

Порівняння результатів розрахунку продуктивності за формулою (3.385) зі значеннями продуктивності, яка спостерігалася на практиці дає задовільні результати. Все викладене про метод регресійного аналізу можна підсумувати наступним чином:

1. Основні припущення регресійного аналізу мають досить загальний характер і зводяться до виконання вимог одновимірного нормального розподілу параметра Θ та однаковості дисперсій для різних фіксованих значень змінних ($\xi_1, \xi_2, ..., \xi_n$). Практично ці вимоги виконуються в умовах реальних об'єктів.

2. Метод регресійного аналізу дозволяє одержати математичний опис нелінійних об'єктів з відповідною оцінкою значущості отриманих результатів.

3. До недоліків регресійного аналізу слід віднести відсутність обґрунтованого і строгого критерію для оцінки адекватності отриманої математичної моделі.

3.9.3. Задача оптимізації статистичного аналізу складних процесів

Будь-який активний експеримент передбачає цільову зміну вхідних параметрів, тобто, в основі лежить планування експерименту. Існує два підходи до планування експерименту: класичний і факторний. При *класичному* підході одночасно варіюють тільки одним фактором. Сюди відносять метод Зейделя-Гаусса, метод градієнта та інші методи. *Факторним* будемо називати планування, коли одночасно варіюють всіма факторами. Задача оптимізації розбивається на два етапи: вихід в район оптимуму з найменшими витратами сировини і часу знаходження математичного опису поверхні функції

$$E = f(z_1, z_2, \ldots, z_n). \tag{3.386}$$

поблизу оптимуму, де E оптимізуючий параметр (наприклад: мінімум вартості, максимум продуктивності установки тощо); $z_1, z_2, \ldots, z_n$ вхідні параметри, які впливають на процес (наприклад: концентрація, хімічний склад, втрата тощо).

На рис. 3.8 представлена поверхня функції $E = f(z_1, z_2)$, яка подана у горизонталях, де показані різні методи руху до оптимуму. Рух починається від точки M. За методом Зейделя-Гаусса відбувається почергова зміна кожного параметру до досягнення частинного максимуму. Рухаючись по прямій MN, фіксуємо точку N. Далі рухаємося вздовж параметра z_1, по прямій NK і т. д. до досягнення оптимуму. Недоліки методу:

1. Довгий шлях руху до оптимуму.
2. Необхідність фіксування всіх параметрів, крім одного на тривалий час.
3. При складній конфігурації поверхні функції $E=f(z_1,z_2)$ метод взагалі не приводить до мети. Так точка A фіксується як оптимальна, так як рух вздовж z_1, (пряма I-III) і вздовж z_2 (пряма II-IV) не дає результатів.

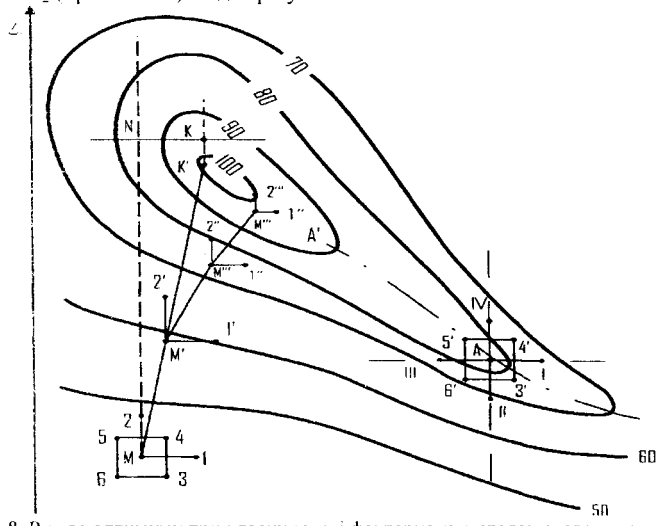


Рис. 3.8. Рух до оптимуму при класичному і факторному методах експериментування

Метод градієнта передбачає рух до оптимуму за найкоротшим шляхом. Для цього в початковій точці M робиться два пробних кроки у точку 1 і в точку 2 та обчислюється напрямок градієнта MM'. Виконується рух в цьому напрямку з деяким кроком. Величина кроку руху і пробного кроку зменшується за часом просування до оптимуму. Після кожного кроку здійснюється поправка напряму градієнта за допомогою пробних кроків (1′; 2′ в точці M'; 1″; 2″ в точці M'' тощо). До недоліків методу градієнта слід віднести відсутність інформації про взаємозв'язок параметрів між собою та неефективність методу при наявності випадкових збурень у порівнянні з факторним методом руху до оптимуму. Крім того, при складній конфігурації функції $E=f(z_1,z_2)$ метод не приводить до цілі. Дійсно, пробні кроки з точки A в будь-який бік (I, II, Ш, IV) дають зменшення E і точка A фіксується як оптимальна.

Факторним методом пошуку оптимальних умов є *метод Бокса-Уілсона*, при якому як і в методі градієнта обчислюється напрямок градієнта, але рух проводиться до досягнення локального екстремуму, як в методі Зейделя-Гаусса, чим досягається деяка економія кількості проведених експериментів, хоча шлях руху до оптимуму може дещо розтягнутися. Градієнт обчислюється уже за 4-ма точками (3, 4, 5, 6 в точці M), що дозволяє визначити його більш точно. Так як на кожному з рівнів фактор знаходиться двічі (для z_1 на нижньому рівні точки 5 і 6) і є можливість усереднити результати двох кроків (для z_1 крок з точки 6 в точку 3 і з точки 5 в точку 4). Крім того, перехід з точки 6 в точку 5 і з точки 4 в точку 3 дозволяє визначити зв'язок параметрів між собою та вплив їх добутку на E. Чотири досліджуваних точки для двох параметрів дають можливість знайти математичну модель поверхні в околі цих точок у вигляді

$$E = b_0 + b_1 z_1 + b_2 z_2 + b_{12} z_1 z_2.$$

Таким чином, у випадку попадання у точку A, вдається розібратися у ситуації і знайти напрямок AA', що дозволяє вийти в район оптимуму. В 3.8.4 викладено правила та рекомендації, які дозволяють економно спланувати експеримент та приведені розрахункові формули для обчислення коефіцієнтів математичної моделі. В 3.8.5 викладено метод Бокса-Вілсона.

3.9.4. Алгебра планування експерименту

Приведемо факторне планування експерименту. Назвемо факторним простором простір, на вісях якого відкладаються значення варійованих факторів. Нехай кожен фактор має свої інтервали зміни, які обумовлюються технологією процесу

$$z_{i\min} \le z_i \le z_{i\max}.$$

Для кожного параметра вибирається крок варіювання λ_i. Очевидно, що $\lambda_i \le z_{i\max} - z_{i\min}$. Далі проводиться нормалізація. Пояснимо це на прикладі.

Приклад 3.20. $z_{i\ min}$ =1600; $z_{i\ max}$ = 2000. Нехай $\lambda_1 = 50$ і $z_{10} = 1800$ (початкова точка).

Розв'язок. Тоді нормалізоване z_1 буде рівне (далі нормалізовані параметри будемо позначати просто z_1, бо скрізь при розгляді маємо на увазі нормалізовані змінні) $\|z_1\| = \dfrac{z_1 - z_{1,0}}{\lambda_1} = \dfrac{z_1 - 1800}{50}$. При факторному плануванні вибирається певне число рівнів по кожному фактору і реалізуються комбінації цих рівнів. Так, рівням -1: 0; +1 для

фактора z_1 будуть відповідати величини 1750; 1700 і 1850. Для якісних параметрів нормалізація полягає у присвоєнні кожному стану свого рівня. Пояснимо це на прикладах.

Приклад 3.21. Можуть бути застосовані 1, 2 або 3 фільтра. За нульову точку приймаємо застосування 2-х фільтрів. Тоді рівні будуть -1; 0; +1.

Приклад 3.22. Нехай використовуємо стару і нову сировину. Приймаючи ці стани за крайні, можемо присвоїти їм рівні -1 та + 1 .

Всі можливі комбінації рівнів факторів назвемо «повним факторним експериментом» (ПФЕ). При умові рівності числа рівнів K для кожного фактора число комбінацій ПФЕ рівне $N = K^n$, де n – кількість факторів.

Якщо реалізується 1/2; 1/3 тощо частина ПФЕ, то це буде відповідний «дробовий частковий факторний експеримент» (скорочено ДФЕ).

Таблиця 3.31

ПФЕ для трьох факторів *A*, *B* і *C* для двох рівнів варіювання +1 та -1

№ дослідів	*A*	*B*	*C*	*E*
1	-1	-1	-1	E_1
2	+1	-1	-1	E_2 } ½ ПФЕ
3	-1	+1	-1	E_3
4	+1	+1	-1	E_4 } ½ ПФЕ
5	-1	-1	+1	E_5
6	+1	-1	+1	E_6 } ½ ПФЕ
7	-1	+1	+1	E_7
8	+1	+1	+1	E_8 } ½ ПФЕ

У таблиці 3.31 подані ПФЕ для трьох факторів *A*, *B* і *C* для двох рівнів варіювання +1 та -1. Таблицю планувань записують іноді більш компактно, приймаючи для досліду в точці (-1; -1; -1) позначення (1), а для досліду (+1; +1; +1) позначення *abc*, тоді всю таблицю можна записати рядком (1), *a*, *b*, *ab*, *c*, *ac*, *bc*, *abc*. Цей спосіб позначення також буде застосований нижче. В таблиці E – значення виходу E_i при відповідній комбінації вхідних параметрів. Матриця планування у випадку ПФЕ для великої кількості факторів складається аналогічно, при цьому частота зміни рівня кожного наступного фактора вдвічі менша за попередній (табл. 3.31).

Приклад 3.23. Планування на трьох рівнях (-1; 0; +1) для трьох незалежних змінних *D*, *F* і *G*:

Розв'язок. ПФЕ $N = 3^3 = 27$. ДФЕ: 1) $\frac{1}{3}N = 9$. 2) $\frac{1}{9}N = 3$. Матриця планування подана таблицею 3.32. Якщо математичну модель об'єкту шукати у вигляді рівняння

$$E = f(z_1, z_2, \ldots, z_n) = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \ldots \beta_n z_n + \beta_{12} z_1 z_2 + \beta_{13} z_1 z_3 + \ldots + \\ + \beta_{n-1,n} z_{n-1} z_n + \beta_{11} z_1^2 + \beta_{22} z_2^2 + \ldots + \beta_{nn} z_n^2 + \ldots. \qquad (3.387)$$

$$\beta_0 = f(z_{10}, z_{20}, z_{30}, \ldots, z_{n0}); \ \beta_i = \frac{\partial f}{\partial z_i}; \ \beta_{ij} = \frac{\partial^2 f}{\partial z_i z_j}; \ \beta_{ii} = \frac{1}{2}\frac{\partial^2 f}{\partial z_i^2};$$

і т. д.: n – загальне число факторів; i, j – поточні номери факторів, причому $i \neq j$, представляє собою ряд Тейлора, то задача буде зводитися до знаходження коефіцієнтів при невідомих β_i, β_{ij} тощо.

У процесі обчислення коефіцієнтів при невідомих у рівнянні (3.387) необхідна перевірка їх значущості. Для цього потрібно знайти похибку експерименту, що виступає у формі середнього квадратичного відхилення $\sigma_{\overline{E}}$. Для отримання $\sigma_{\overline{E}}$ здійснюють дублювання дослідів. Число паралельних дослідів γ вибирають як правило від 2 до 4.

Таблиця 3.32

Вхідні дані для прикладу 3.23

№ дослідів	D	F	G	E	№ дослідів	D	F	G	E	№ дослідів	D	F	G	E
1	-1	-1	-1	E_1	10	-1	-1	0	E_{10}	19	-1	-1	+1	E_{19}
2	0	-1	-1	E_2	11	0	-1	0	E_{11}	20	0	-1	+1	E_{20}
3	+1	-1	-1	E_3	12	+1	-1	0	E_{12}	21	+1	-1	+1	E_{21}
4	-1	0	-1	E_4	13	-1	0	0	E_{13}	22	-1	0	+1	E_{22}
5	0	0	-1	E_5	14	0	0	0	E_{14}	23	0	0	+1	E_{23}
6	+1	0	-1	E_6	15	+1	0	0	E_{15}	24	+1	0	+1	E_{24}
7	-1	+1	-1	E_7	16	-1	+1	0	E_{16}	25	-1	+1	+1	E_{25}
8	0	+1	-1	E_8	17	0	+1	0	E_{17}	26	0	+1	+1	E_{26}
9	+1	+1	-1	E_9	18	+1	+1	0	E_{18}	27	+1	+1	+1	E_{27}

Нехай маємо N дослідів при γ паралельних

$$
\begin{array}{cccccc}
E_{11}; & E_{12}; & \ldots & E_{1k}; & \ldots & E_{1\gamma}; \\
E_{21}; & E_{22}; & \ldots & E_{2k}; & \ldots & E_{2\gamma}; \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
E_{j1}; & E_{j2}; & \ldots & E_{jk}; & \ldots & E_{j\gamma}; \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
E_{N1}; & E_{N2}; & \ldots & E_{Nk}; & \ldots & E_{N\gamma}.
\end{array}
$$

Обчислюється середнє значення E для кожного j-ого рядка

$$\overline{E}_j = \frac{E_{j1} + E_{j2} + \ldots + E_{jk} + \ldots + E_{j\gamma}}{\gamma}. \qquad (3.388)$$

і визначається оцінка дисперсії виміру σ_ε^2

$$\sigma_\varepsilon^2 = \frac{\sum_{j=1}^{N} \frac{\sum_{k=1}^{\gamma} (E_{ij} - \overline{E}_j)^2}{\gamma - 1}}{N}. \qquad (3.389)$$

Тоді дисперсія середнього

$$\sigma_{\bar{E}}^2 = \frac{\sigma_\varepsilon^2}{\gamma}. \tag{3.390}$$

і похибка експерименту

$$\sigma_{\bar{E}} = \sqrt{\sigma_{\bar{E}}^2}. \tag{3.391}$$

Для перевірки рівнозначності вимірів використаємо критерій Кохрена, згідно з яким, якщо маємо дисперсії по рядках і є їх сума

$$\sigma_j^2 = \frac{\sum_{k=1}^{\gamma}(E_{ij} - \bar{E}_j)^2}{\gamma - 1} \tag{3.392}$$

$$\sigma_\Sigma^2 = \sum_{j=1}^{N}\sigma_j^2. \tag{3.393}$$

то для перевірки рівнозначності вимірів необхідно вибрати найбільшу з рядкових дисперсій $\sigma_{j\max}^2$ і знайти G-критерій

$$G = \frac{\sigma_{j\max}^2}{\sigma_\varepsilon^2}. \tag{3.394}$$

Якщо $G < G_m$, G_m береться з таблиць, то досліди рівнозначні. При рівнозначних дослідах слід збільшити число паралельних дослідів у рядку, де $\sigma_{j\max}^2$, або збільшити точність вимірів, застосувавши кращий прилад. Якщо ж немає можливості збільшити точність вимірів або число паралельних дослідів, то слід здійснювати розрахунки, відмовившись від дублювання дослідів ($\gamma = 1$). При плануванні експерименту завжди слід прагнути до варіювання на двох рівнях. Це найбільш просте планування при максимумі наочності, воно не обмежує діапазон вивчення факторів, так як серії дослідів можуть бути заплановані для різних рівнів факторів. Це легко виконати, коли фактори кількісні й експериментатор має можливість вільно варіювати їх рівнями. Практично визначають коефіцієнти ряду Тейлора (3.387), обмежившись поліномом деякого степеня. Якщо віднести розсіяння точок до похибок експерименту, то можлива апроксимація функції виходу більш низького степеня. Зазвичай обмежуються 1 або 2 степенями наближення. Подібна апроксимація записується через рівняння регресії, яка для другого степеня наближення має вигляд

$$\begin{aligned} E_p = {} & b_0 + b_1 z_1 + b_2 z_2 + \ldots b_n z_n + b_{12} z_1 z_2 + b_{13} z_1 z_3 + \ldots + \\ & + b_{n-1} z_{n-1} z_n + b_{11} z_1^2 + b_{22} z_2^2 + \ldots + b_{nn} z_n^2. \end{aligned} \tag{3.395}$$

Тут і надалі будемо позначати: E – істинне значення виходу; E_p – значення виходу, що обчислене за тією чи іншою формулою. Можна показати, що у випадку, якщо вихідний експериментальний матеріал відповідає ортогональній матриці планування, котра задовольняє наступним умовам:

$$1)\ \sum_{j=1}^{N} z_{ij} = 0\ ;\ 2)\ \sum_{j=1}^{N} z_{ij}^2 = N\ ;\ 3)\ \sum_{j=1}^{N} z_{ij} z_{lj} = 0\ ,\ i \neq j\ , \tag{3.396}$$

де j – поточний номер дослідної точки; i і l – поточні номери факторів, причому $i \neq l$.

Справедливі наступні формули для обчислення коефіцієнтів рівняння (3.396)

$$b_0 = \frac{1}{N}\sum_{j=1}^{N} E_j\ ;\quad b_i = \frac{\sum_{j=1}^{N} z_{ij} E_j}{\sum_{j=1}^{N} z_{ij}^2} = \frac{1}{N}\sum_{i=1}^{N} z_{ij} E_j\ ;\quad b_{ij} = \frac{1}{N}\sum_{i=1}^{N} z_{ij} z_{lj} E_j\ ,\ i \neq j\ . \tag{3.397}$$

Коефіцієнти при квадратах параметрів типу b_{ij} не можуть бути визначені окремо і входять в b_0. Легко перевірити, що матриця планування для двох рівнів (табл. 3.31) задовольняє вимозі ортогональності, а будь-яка матриця планування з числом рівнянь більше двох (табл. 3.32) не є ортогональною. Знайдені за формулами (3.397) коефіцієнти будуть тільки оцінками відповідних коефіцієнтів рівняння (3.387). Дійсно, для знаходження, наприклад, коефіцієнтів рівняння (3.396) достатньо планування на двох рівнях, так як воно дозволяє роздільно визначити коефіцієнти типу b_i і b_{ij}, але всі коефіцієнти більш високих степенів рівняння (3.387) будуть присутніми у знайдених оцінках. Так, з табл. 3.33 випливає, що одержано β_0, разом з β_{11}, β_{22}, β_{1111}, β_{2222}, тощо, у вигляді b_0, так як стовпці для z_0, z_1^2, z_2^2, z_1^4, z_2^4 тощо повністю співпадають. Стовпець z_0 вводиться для ілюстрації і не змінює формули для обчислення b_0

$$b_0 = \frac{1}{N}\sum_{j=1}^{N} z_{0j} E_j = \frac{1}{N}\sum_{j=1}^{N} E_j\ . \tag{3.398}$$

так як $z_0 = +1$.

Таблиця 3.33

Матриця планування

№ дослідів	z_0	z_1	z_2	z_1z_2	z_1^2	z_2^2	z_1^3	z_2^3	E
1	+1	-1	+1	+1	+1	+1	-1	-1	E_1
2	+1	+1	-1	-1	+1	+1	+1	-1	E_2
3	+1	1	+1	1	+1	+1	1	+1	E_3
4	+1	+1	+1	+1	+1	+1	+1	+1	E_4

Коефіцієнт β_i отримаємо разом з β_{111}, β_{11111} і т. д., так як стовпці для z_1, z_1^3, z_1^5 і т.д. повністю співпадають. Для перевірки значущості отриманих коефіцієнтів використаємо критерій Ст`юдента t. Для визначеного рівня значущості q (часто вибирають q = 5%) і числа степенів вільності $k = N(\gamma - 1)$ знаходять з таблиць відповідне t. Для ортогонального планування дисперсії похибок визначення коефіцієнтів рівні між собою і розраховуються за формулою

$$\sigma_b^2 = \frac{\sigma_E^2}{N}\ . \tag{3.399}$$

Коефіцієнт значущий, якщо

$$b_i > \sqrt{\sigma_b^2 \cdot t}\,. \tag{3.400}$$

Введемо поняття «ступеня вільності». Числом ступенів вільності будемо називати число окремо визначених оцінок для оцінок коефіцієнтів рівняння (3.387).Тоді число ступенів вільності f при N дослідах буде рівне $f = N$-1. Наприклад, з табл. 3.33 випливає, що f=4-1=3. Незалежно один від одного визначаються b_1, b_2, b_{12} (якщо припустити, що всі коефіцієнти другого і вище порядків досить малі). В силу того, що кожен коефіцієнт має лише одну ступінь вільності, має місце наступна рівність

$$\sum_{i=1}^{N}\frac{1}{N}\left(\sum_{j=1}^{N} z_{ij}E_j\right)^2 + \sum_{i=1}^{N}\frac{1}{N}\left(\sum_{\substack{j=1\\ i\neq l}}^{N} z_{ij}z_{lj}E_j\right)^2 = \sum_{j=1}^{N}\left(E_j\right)^2 - \frac{1}{N}\left(\sum_{j=1}^{N}E_j\right)^2. \tag{3.401}$$

Цю рівність можна використовувати для перевірки правильності арифметичних операцій. Для табл. 3.33 маємо, використовуючи (3.401)

$$\frac{1}{4}(-E_1+E_2-E_3+E_4)^2+\frac{1}{4}(-E_1-E_2+E_3+E_4)^2+\frac{1}{4}(E_1-E_2-E_3+E_4)^2 =$$

$$=(E_1^2+E_2^2+E_3^2+E_4^2)+\frac{1}{4}(E_1+E_2+E_3+E_4)^2.$$

Використовуючи формули (3.397), обчислимо коефіцієнти типу b_i і b_{ij} для матриці планування, поданій у табл. 3.33

$$b_0=\frac{E_1+E_2+E_3+E_4}{4};\quad b_1=\frac{-E_1+E_2-E_3+E_4}{4}; \tag{3.402}$$

$$b_2=\frac{-E_1-E_2+E_3+E_4}{4};\quad b_{12}=\frac{E_1-E_2-E_3+E_4}{4}.$$

Вважаючи коефіцієнти при вищих степенях рівняння (3.387) незначними, можна вважати, що b_0 є оцінкою для β_0, β_{11} і β_{22}; b_1 є оцінкою для β_1; b_2 є оцінкою для β_2 і b_{12} є оцінкою для β_1. Кожен з коефіцієнтів визначається з точністю $\pm t\sigma_b$, де

$$\sigma_b=\sqrt{\sigma_b^2}\,. \tag{3.403}$$

Зауважимо, що не завжди можна застосовувати ПФЕ. Іноді застосування ДФЕ дає таку ж ефективність, що й ПФЕ, але при цьому не зменшується здатність до розв'язування. ДФЕ вигідно застосовувати, коли:

1) частина коефіцієнтів у добутку трьох і більше факторів складає малу частку порівняно з коефіцієнтами типу b_i і b_{ij};

2) діапазон зміни досліджуваних змінних малий;

3) ситуація невизначена, тобто тільки для деяких факторів очікуються значимі коефіцієнти.

Якщо виявиться, що на процес впливають багато факторів, то слід доповнити досліди до ПФЕ. Таким чином, планування ДФЕ потрібно проводити враховуючи фізику процесу. Нехай, наприклад, $\beta_{1,2}$=0 (табл. 3.33). Звісно, що можна

використовувати це для включення в експеримент ще одного фактора z_3 і змінювати його стосовно z_1z_2. Тоді оцінка для x_3 буде сумісною оцінкою $\beta_{1,2}$ і β_3, але так як $\beta_{1,2}=0$, то $\beta_3=b_3$. Таким чином, для цієї ситуації можна застосувати ДФЕ для 4-х дослідів (тобто половину ПФЕ). Це називається плануванням зі змішуванням. Для кращого розуміння принципу побудови матриці планування ДФЕ введемо поняття визначаючого контрасту. Визначаючий контраст дозволяє визначити, що з чим змішано, не вивчаючи матрицю планування для визначення співпадаючих стовпців. Для цього вводиться таке правило

$$z_i^2 = I,\ z_i \cdot I = z_i. \tag{3.404}$$

Принцип отримання визначаючого контрасту розглянемо на прикладі.

Приклад 3.24. Нехай маємо 5 факторів: *A*, *B*, *C*, *D*, *F*. Відомо, що $\beta_{ABC}=0$ і $\beta_{BC}=0$.

Розв'язок. У такій ситуації замість ПФЕ ($N=2^5=32$) можна застосувати ДФЕ з кількістю дослідів $\frac{1}{4}N=8$, тобто прийняти фактори *A*, *B* і *C* за основні, побудувавши для них ПФЕ ($2^3=8$), а фактори *D* і *F* – не основними. Тоді матрицю планування можна отримати з умов відповідно

$$D = ABC \text{ і } F = BC. \tag{3.405}$$

Таким чином, матриця планування набуде вигляду табл. 3.34. Знайдемо визначальний контраст цієї табл. 3.34. Позначимо $A^2 = I$; $B^2 = I$; тощо. Випишемо всі включені неосновні фактори (3.404). Помножимо ці рівності відповідно на *D* і *F*, отримаємо

$$D^2 = ABCD = I \text{ і } F^2 = DCF = I. \tag{3.406}$$

Отримані рівняння необхідно перемножити, використовуючи всі комбінації. У нашому випадку є всього одна комбінація $I = ABCD * BCF = ADF$.

Таблиця 3.34

Матриця планування

№ дослідів	*A*	*B*	*C*	*D*	*E*	*F*	Позн. комбін. рівняння
1	-	-	-	-	+	E_1	*f*
2	+	-	-	+	+	E_2	*adf*
3	-	+	-	+	-	E_3	*bd*
4	+	+	-	-	-	E_4	*ab*
5	-	-	+	+	-	E_5	*cd*
6	+	-	+	-	-	E_6	*ac*
7	-	+	+	-	+	E_7	*bcf*
8	+	+	+	+	+	E_8	*abcdf*

Якщо до першого блоку додати другий блок, де b_A є оцінкою для β_A; β_{BCD}; $-\beta_{ABCF}-\beta_{DF}$, то b_A буде оцінкою для двох коефіцієнтів: b_A для β_A; β_{BCD} (при

додаванні блоків); b_A для $\beta_{ABCF}; \beta_{DF}$ (при відніманні блоків). Таким чином, для оцінки β_A досить, щоб

$$BCD = 0. \tag{3.407}$$

Таким чином, видно, що ефективність ДФЕ зростає при відсутності взаємодії: чим менше є взаємодій, тим менша кількість дослідів, яка необхідна для успішного експерименту. Але, з іншого боку, чим менше ДФЕ, тим нижче точність. Дійсно, якщо припустити і в одному, і в двох, і в трьох блоках незалежне визначення β_A (тобто визначення β_A з найбільшою точністю), то дисперсії відповідно будуть

$$\frac{\sigma_E^2}{8}; \frac{\sigma_E^2}{16}; \frac{\sigma_E^2}{24}. \tag{3.408}$$

При реалізації першого і третього блоків, коли $DF \neq 0$, $ABCF = BCD = 0$, у кожному з блоків відповідно

$$b_A^1 = \beta_A + \beta_{DF}; \; b_A^2 = \beta_A - \beta_{DF}; \; b_A^3 = \beta_A - \beta_{DF}. \tag{3.409}$$

Тоді для роздільної оцінки β_A необхідно додати другий і третій блоки, після чого знайти

$$\beta_A - \beta_{DF} = \frac{1}{2}(b_A^2 + b_A^3). \tag{3.410}$$

і додати перший блок

$$\beta_A = \frac{1}{2}\left[\frac{1}{2}(b_A^2 + b_A^3) + b_A'\right] = \frac{1}{4}b_A^2 + \frac{1}{4}b_A^3 + \frac{1}{2}b_A'. \tag{3.411}$$

Дисперсія цієї оцінки буде

$$\sigma_{\beta A}^2 = \frac{1}{2}\sigma_E^2\left[\left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{2}\right)^2\right] = \frac{3}{16}\sigma_E^2. \tag{3.412}$$

Динамікою змішування є змішування комбінацій коефіцієнтів в ДФЕ. Найменший ДФЕ, який дозволяє застосувати змішування, складається з 8 спостережень. 9 спостережень діляться на два блоки, в кожному з яких будь-яка незначна взаємодія може бути змішана. Розглянемо це на прикладі.

Приклад 3.25. Необхідно провести експерименти з процесором, який визначається трьома факторами: *A*, *B*, *C*, але стандартної сировини на проведення восьми незалежних спостережень (ПФЕ) не вистачає і доводиться застосовувати іншу сировину. Щоб зміна сировини не спотворювала результати експериментів, вирішили додати коефіцієнт, який визначає зміну сировини із взаємодією другого порядку (*ABC*), вважаючи його досить малим. Для цього всі досліди були розділені на два блоки (для першої (+1) і другої (-1) партії сировини відповідно).

У першому блоці були зібрані комбінації рівнів, де $ABC = +1(a;b;c;abc)$, а у другому, де $ABC = -1[(1);ac;bc;ab]$. Подібним чином можна здійснити змішування з

великою кількістю якісних станів. Матриця планування подана у таблиці 3.35. Тут тільки b_{ABC} буде викривлене зміною кількості сировини (позначимо x)

$$b_{ABC} = \frac{1}{8}[(a+x)+(b+x)+(c+x)+(abc+x)-(1)-ac-bc-ab], \qquad (3.413)$$

Таблиця 3.35

Матриця планування для прикладу 3.25

№ дослідів	A	B	C	ABC	Комбін. рівняння	E
1	+	-	-	+	a	E_1
2	-	+	-	+	b	E_2
3	-	-	+	+	c	E_3
4	+	+	+	+	abc	E_4
5	-	-	-	-	(1)	E_5
6	+	-	+	-	ac	E_6
7	-	+	+	-	bc	E_7
8	+	+	-	-	ab	E_8

а всі інші оцінки спотворені не будуть

$$b_{AB} = \frac{1}{8}[(1)+ab+(c+x)+(abc+x)-(a+x)-(b+x)-ac-bc] = \\ = \frac{1}{8}[(1)+ab+c+abc-a-b-ac-bc]. \qquad (3.414)$$

Отже, докладно розглянули алгебру планування експерименту при факторному аналізі.

3.9.5. Метод Бокса-Вілсона

Використовуючи факторне планування експерименту, Бокс та Вілсон у 1951 році запропонували новий *метод пошуку оптимальних умов для складних процесів*. При цьому рух до оптимуму здійснюється по найкоротшому шляху (рис. 3.8). Області, далекі від оптимуму, апроксимуються гіперплощиною

$$E_p = b_0 + b_1 z_1 + \ldots + b_n z_n, \qquad (3.415)$$

тобто тут визначаємо лише напрям руху, не описуючи поверхню більш точно.

Здійснюючи покроковий рух від точок з меншим виходом E до точок з більшим виходом E, реалізується ДФЕ у кожній точці, згідно алгебри планування експерименту. Цю частину експериментальних робіт назвемо покроковим рухом, так як рух тут співпадає з рухом у напрямку градієнта

$$\nabla E_p = \frac{\partial E_p}{\partial z_1}\cdot\vec{i} + \frac{\partial E_p}{\partial z_2}\cdot\vec{j} + \frac{\partial E_p}{\partial z_3}\cdot\vec{k} + \ldots. \qquad (3.416)$$

де $\vec{i}, \vec{j}, \vec{k}$ – одиничні вектори у напрямку координатних осей.

Друга частина робіт містить більш громіздкі експерименти, так як в області, близькій до оптимуму або в області з великою кривизною, апроксимація гіперплощиною стає неадекватною і потрібно визначати, крім коефіцієнтів типу b_i, всі взаємодії, щоб отримати рівняння у вигляді (3.387).

1. *Покроковий рух.* Покроковий рух об'єднує метод руху у напрямку градієнта з методом ДФЕ або ПФЕ. Задачі, які розв'язуються цим методом, полягають у досягненні області, близької до оптимуму, або області з великою кривизною при мінімальних затратах. Метод оптимальний у розумінні наступних умов:

а) дисперсія $\sigma^2_{\bar{E}}$ однакова для всіх точок, які знаходяться на однаковій відстані від центру;

б) дисперсія $\sigma^2_{b_i}$ для коефіцієнтів рівняння регресії $b_0, b_1, ..., b_n$ мінімальна;

в) оцінки коефіцієнтів рівняння регресії статистично незалежні. Дійсно, нехай у ПФЕ для n=2 z_1 встановлюється на рівні ±1 з похибкою $\pm\varepsilon$ (табл. 3.35,3.36);

Таблиця 3.36

Матриця планування

№ дослідів	z_0	z_1	z_2	z_3	E
1	+1	-1	-1	+1	$E_1 - \varepsilon$
2	+1	+1	-1	-1	$E_2 + \varepsilon$
3	+1	-1	+1	-1	$E_3 - \varepsilon$
4	+1	+1	+1	+1	$E_4 + \varepsilon$

b_i' – коефіцієнти рівняння регресії для досліду з помилками;

b_i – коефіцієнти рівняння регресії для досліду без помилок.

$$b_0' = \frac{(E_1 - \varepsilon) + (E_2 + \varepsilon) + (E_3 - \varepsilon) + (E_4 + \varepsilon)}{4} = b_0;$$

$$b_1' = \frac{(E_2 + \varepsilon) + (E_4 + \varepsilon) - (E_1 - \varepsilon) - (E_3 + \varepsilon)}{4} = b_1 + \varepsilon;\ b_2' = b_2;\ b_{12}' = b1_2.$$

Тільки b_1 визначений з помилкою. При розв'язуванні задачі оптимізації покроковий рух є початковим етапом розв'язку. Приступаючи до дослідження процесу, експериментатор повинен грамотно сформулювати умови його протікання:

а) Визначити точність контрольно-вимірювальної апаратури і по можливості використовувати для експерименту найбільш точну апаратуру. Тут слід також враховувати швидкодію вимірювальної апаратури. Очевидно, що швидкодія апаратури повинна бути не менше швидкості перехідних процесів при становленні рівнів.

б) Визначити кількість параметрів, які визначають процес, і спробувати розмістити їх по рангу впливу, базуючись на фізиці процесу. Якщо кількість параметрів більша 7-10, то слід застосувати спочатку відсівні елементи. Для кожного параметра необхідно встановити інтервали варіювання, які зумовлені технологією процесу: $z_{i\min} \le z_i \le z_{i\max}$ і вибрати крок варіювання λ_i. Крок варіювання повинен бути досить великим у порівняння з похибками експерименту і досить малим, щоб, по можливості, точно вловити напрямок до оптимуму. Спочатку λ_i вибирають інтуїтивно, а потім корегують.

в) Правильно вибрати параметр оптимізації. Необхідно, щоб вибраний вихід системи відповідав фізиці процесу. Іноді може бути декілька оптимізуючи величин, наприклад якість і ціна. У цьому випадку слід шукати компроміс.

г) Врахувати всю статистику процесу, тобто:

- дисперсію σ_E^2;
- найбільш сприятливі рівні z_i, які відповідають точці з найбільшим виходом;
- розрахувати час запізнення, який необхідний для правильного зняття показів.

Вважаємо, що процесом можна керувати і він є відтворюваним (якщо це можна визначити за статистичними даними), тому що у протилежному випадку метод Бокса-Вілсона незастосовний. Зазвичай, немає відомостей про відтворюваність процесу. Тоді, після постановки деякої серії дублюючих дослідів в одній точці (серії у різні проміжки часу), обчислюють відповідні дисперсії і, якщо вони чуже відрізняються, роблять висновок про невідтворюваність процесу. Отже, маємо кількість факторів, що рівна n, їх інтервали варіювання і кроки $\lambda_1, \lambda_2, ..., \lambda_n$.

Маємо дисперсію, що одержана зі статистичних даних $(\sigma_E^2)_{CT}$, і основний (базисний) рівень для всіх факторів. Початкова точка у факторному просторі вибирається з максимальним виходом. Маємо апріорні або інтуїтивні відомості про степінь впливу на процес кожного з n факторів і про характер їх взаємодій. Далі, вибирається матриця планування. Якщо $n = 2 \div 3$, то вибирається ПФЕ. При $n > 2 \div 3$ реалізують спочатку ДФЕ, вибравши відповідний визначальний контраст. Якщо про характер взаємодій нічого не відомо, то слід вибрати визначальний контраст, у якому нехтують вищими взаємодіями. Складається карта проведення експерименту (табл. 3.40). У табл. 3.36 у першому рядку записуються фактори в позначеннях z_i, а в дужках – експериментальна відповідність фактору. Записується матриця планування й здійснюється її рандомізація (встановлюється порядок реалізації).

Суть рандомізації – виключити появу систематичних похибок за рахунок строгої реалізації матриці планування. Це досягається тим, що за таблицею випадкових чисел вибирається порядок реалізації випробувань. Якщо здійснено N випробувань при трьох паралельних ($\gamma = 3$), то загальна кількість дослідів рівна $3N$. У клітинці номери рядків матриці планування проставлені у порядку їх реалізації. При $\gamma = 3$ будуть три такі стовпці. Після реалізації матриці планування обчислюються порядкові середні $\overline{E}_j$ за формулою (3.388) і порядкові дисперсії

$$\sigma_j^2 = \frac{\sum_{k=1}^{\gamma} (E_{jk} - \overline{E}_j)^2}{\gamma - 1}. \qquad (3.417)$$

Перевіряється точність вимірів і за формулою (3.391) обчислюється дисперсія σ_E^2. Якщо $\sigma_E^2 > (\sigma_E^2)_{CT}$, то слід знайти причину «поганого» експерименту. У кращому випадку – це неточність вимірів. Друга причина – невідтворюваність дослідів тягне за собою відмові від методу Бокса-Вілсона. Далі підраховуються суми типу $\sum_{j=1}^{N} z_{ij} \overline{E}_j$ і

$\sum_{j=1}^{N} z_{ij} z_{lj} \overline{E}_j$, $l \neq j$. Оцінки повинні бути незалежними і їх кількість повинна бути рівна $f=n-1$. За формулою (3.408) перевіряється правильність арифметичних операцій, а за формулою (3.404) підраховуються коефіцієнти рівняння регресії і перевіряється їх значущість. Якщо майже всі добутки незначущі, то знаходяться причини незначущості коефіцієнтів. Можуть бути наступні причини незначущості коефіцієнтів:

а) основний рівень для факторів з малими коефіцієнтами близький до максимального (частковий екстремум);

б) одиниці зміни вибрані дуже малими;

в) вихід системи не залежить від варіювання цими факторами.

Якщо інтервал зміни за деяким фактором збільшити не можна, керуючись міркуваннями природного ходу процесу, вартості матеріалу тощо., то приймається, що цей фактор не впливає на процес, і ним ми надалі нехтуємо. За іншими факторами збільшується інтервал варіювання і експерименти повторюються.

Якщо має місце причина а) або б), то при цьому відповідний коефіцієнт збільшується; якщо діє причина в), то відповідно коефіцієнт не зміниться. Далі, необхідно повторити можливість апроксимації рівняння $E = \varphi(z_1, z_2, ..., z_n)$ гіперплощиною. Для цього необхідно підрахувати критерій Фішера

$$F = \frac{\sigma^2_{Ead}}{\sigma^2_{\overline{E}}}, \tag{3.418}$$

і порівняти його з табличним.

Обчислення σ^2_{Ead} бажано провести двома методами для перевірки правильності арифметичних операцій. Перший спосіб:

$$(\sigma^2_{Ead})_1 = \frac{\sum_{j=1}^{N} \overline{E}_j^2 - N \sum_{i=0}^{n} b_i^2}{N-(n+1)}. \tag{3.419}$$

Другий спосіб:

$$(\sigma^2_{Ead})_2 = \frac{\sum_{j=1}^{N} \left[E_j - (b_0 + b_1 z_1 + b_2 z_2 + ... + b_n z_n)\right]^2}{N-(n+1)}. \tag{3.420}$$

У формулі (3.420) член $(b_0 + b_1 z_1 + b_2 z_2 + ... + b_n z_n) = E_r$ є розрахунковим значенням виходу при апроксимації гіперплощиною. Замість z_i слід підставити відповідно «+1» або «-1». Якщо $(\sigma^2_{Ead})_1 = (\sigma^2_{Ead})_2$, то всі операції виконані правильно не тільки у обчисленні σ^2_{Ead}, але й b_i. Табличний критерій знаходиться для ступенів свободи $f_1 = N-(n+1)$ та $f_2 = n\gamma - 1$, де γ – число паралельних дослідів.

Якщо $F_{table} \geq F$ – апроксимація адекватна і можна приступати до покрокового руху (табл. 3.36). Для руху по градієнту незалежні змінні необхідно замінити

пропорційно до коефіцієнтів нахилу b_i. Для цього обчислюються добутки $b_i\lambda_i$ і фактор, для якого цей добуток максимальний, приймається за базовий. Для нього вибирають крок варіювання (λ_{iKB}), залишаючи старий крок або вводячи новий для більш дрібного покрокового руху. Кроки варіювання за іншими факторами визначаються пропорційно базовому. Обчислюються відношення

$$\Delta = \frac{\lambda_{iKB}}{b_i\lambda_i}, \text{ або } \lambda_{iKB} = b_i\lambda_i\Delta,\ \lambda_{iKB} < 0, \text{ якщо } b_i < 0. \qquad (3.421)$$

У табл. 3.36 при покроковому русі у стовпці для першого уявного досліду (не поставлений експериментально, а прогнозований) будемо мати $z_{i0} + \lambda_{iKB}$, для другого уявного досліду $z_{i0} + 2\lambda_{iKB}$, тощо.

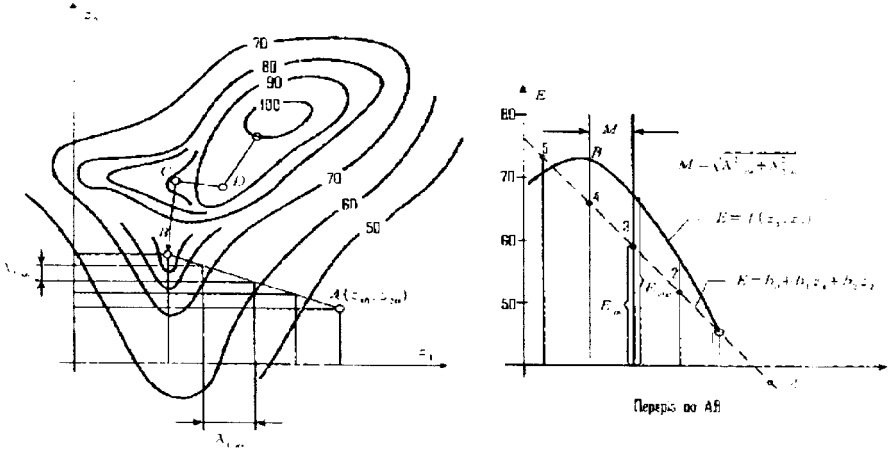


Рис. 3.9. Стрімке сходження

Для кожного уявного досліду обчислюється E_{pr} Для першого уявного досліду

$$E_{pr1} = b_0 + b_1\frac{\lambda_{1KB}}{\lambda_1} + b_2\frac{\lambda_{2KB}}{\lambda_2} + \ldots + b_n\frac{\lambda_{nKB}}{\lambda_n}.$$

Для другого $E_{pr2} = 2E_{pr1} - b_0$ і т.д. Уявні досліди продовжують до тих пір, поки виконується нерівність $E_{prj} \le (1 \div 2)E_{\max}$, де $E_{\max}$ – максимально можливий вихід (з фізичних міркувань). Деякі з уявних дослідів реалізуються для перевірки відповідності апроксимації, і краща реалізація береться за нову початкову точку. На рис. 3.9 перший покроковий рух відповідає прямій AB. На перегині по AB показані: реальна крива $E = f(z_1, z_2)$, її апроксимація і уявні досліди (точки 1, 2, ...). Реалізуючи деякі з уявних дослідів, вибирають найкращий і приймають таку точку за нову базову точку (точка B). Покроковий рух дозволяє пройти повз «горбів» і «впадин» характеристики та зупинитися на найвищій вершині (рис. 3.10). Для цього не слід зупинятися на першому горбі, а рухатися далі, якщо всі фактори залишаються в межах інтервалів варіювання.

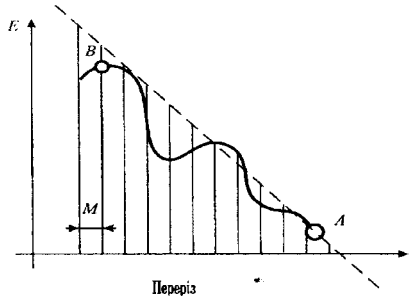


Рис. 3.10. Стрімке сходження при наявності кількох максимумів і мінімумів

У точці B виконується наступна серія експериментів і вибирається новий напрям руху. Так повторюємо до тих пір, поки можлива апроксимація гіперплощиною. При цьому слід від серії до серії дослідів зменшувати λ_i, оскільки, по мірі наближення до оптимуму збільшуються коефіцієнти b_{ii} і зменшуються b_i та b_{ij}. В результаті покрокового руху можливе попадання у одну з чотирьох ситуацій (рис. 3.11).

Апроксимація у вигляді рівняння (3.393) дозволяє визначити ситуацію і намітити найкоротший шлях виходу з несприятливої області (II, III, IV).

Зауважимо, що покроковий рух вдалині від оптимуму не слід робити з великою точністю, тобто проводити досліди при великій кількості паралельних дослідів, оскільки це лише метод руху до оптимуму, а не метод опису функції виходу.

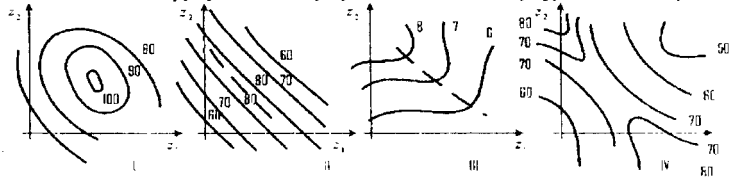


Рис. 3.11. Можливі ситуації після стрімкого сходження

При дуже малому σ_E^2 опис гіперплощиною може виявитися неадекватним там, де така апроксимація недоцільна. Тому при покроковому русі обмежуються зазвичай двома паралельними дослідами. Однак, якщо точність вимірів надто низька, то при $\gamma = 2$ не вдається отримати значущі коефіцієнти і тоді слід збільшити число паралельних дослідів.

2. *Досліди в областях з великою кривизною.*

Як зазначалося, сигналом закінчення покрокового руху є неадекватність подання рівняння зв'язку $E = f(z_1, z_2, ..., z_n)$ гіперплощиною, тобто при цьому слід переходити до многочленна другого степеня

$$E_p = b_0 + b_1 z_1 + b_2 z_2 + ... b_n z_n + b_{12} z_1 z_2 + b_{13} z_1 z_3 + ... + \\ + b_{n-1} z_{n-1} z_n + b_{11} z_1^2 + b_{22} z_2^2 + ... + b_{nn} z_n^2. \qquad (3.422)$$

«Відчути» перебування в одному з чотирьох варіантів (рис. 3.11) можна, поглянувши на результати експерименту. Якщо одинакові E_i повторюються для різних комбінацій рівнів, то це можливо по одній з двох причин:

а) є помилки у вимірюваннях;

б) коефіцієнти типу b_{ij} і b_{ii} значущі. Останнє й означає, що робоча точка знаходиться в області з великою кривизною.

Якщо апроксимація гіперплощиною неадекватна, а матриця планування представляє собою малий ДФЕ і незалежне визначення коефіцієнтів типу b_{ij} неможливе, слід доповнити ДФЕ до ПФЕ. В області близькій до оптимуму, постають нові задачі по опрацюванню результатів. Основним питанням не є питання про менші затрати при русі до оптимуму, як це було при покроковому русі, а найбільш точний математичний опис функції зв'язку. При цьому окремої оцінки потребують всі коефіцієнти типу b_i; b_{ii}; b_{il}, $i \neq l$.

ПФЕ не дозволяє розділити оцінки для β_0; β_{11}; β_{22}; ...; β_{nn}, оскільки $\sum_{i=1}^{N} z_{ij}^2 z_{lj}^2 \neq 0$, тобто планування не ортогональне до добутку квадратів.

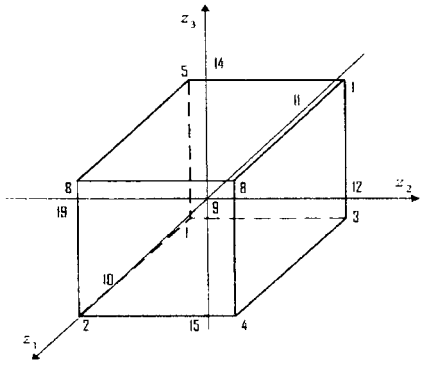


Рис. 3.12. Модель ортогонального центрального композиційного планування

Для роздільної оцінки всіх коефіцієнтів многочленна другого степеня застосовують так зване *центральне композиційне планування*. Центральне композиційне планування (скорочено ЦКП) включає в себе ГПФЕ та деяку кількість додаткових точок, яка залежна від числа факторів. Розрізняють два види ЦКП: ортогональне ЦКП; рототабельне ЦКП.

1. *Ортогональне ЦКП* історично виникло раніше, ніж рототабельне ЦКП і завдяки своїй простоті довгий час вважалося оптимальним. Загальна кількість дослідів при ортогональному ЦКП обчислюється за формулою $N = 2^n + 2n + 1$, де 2^n – відповідає центральній точці з координатами (0.0,....0). Модель ортогонального ЦКП для $n = 3$ подана на рис. 3.12, а матриця планування – у таблиці 3.37.

Номери точок на рис. 3.12 відповідають номерам точок у матриці планування.

Для забезпечення ортогональності другого порядку, тобто щоб

$$\sum_{i=1}^{N} z_{ij}^2 z_{lj}^2 = 0, \qquad (3.423)$$

необхідно провести лінеаризацію рівняння регресії. При цьому рівняння регресії записується у вигляді

$$E = b_0 z_0 + b_1 z_1 + b_2 z_2 + \ldots + b_n z_n + b_{n+1} z_{n+1} + b_{n+2} z_{n+2} + \ldots + \\ + b_{n+C_n^2} z_{n+C_n^2} + b_{n+C_n^2+1} z_{n+C_n^2+1} + \ldots + b_{2n+C_n^2} z_{2n+C_n^2}. \qquad (3.424)$$

де $C_n^2 = \dfrac{n!}{2!(n-2)!}$ – число сполучень з n до 2; $z_{n+1}=z_1 z_2$; $z_{n+2}=z_1 z_2$; …; $z_{n+C_n^2} = z_{n-1} z_n$;

$z_{n+C_n^2+i} = \bar{z}_i^2 - \dfrac{\sum\limits_{j=1}^{N} z_{ij}^2}{N}$; тобто виконується заміна тільки тих членів, які стоять при квадратичних коефіцієнтах. Нові коефіцієнти рівняння регресії обчислюються за формулами:

$$b_i' = \frac{\sum\limits_{j=1}^{N} z_{ij} E_j}{\sum\limits_{j=1}^{N} (z_{ij})^2};\ b_l = \frac{\sum\limits_{j=1}^{N} z_{ij} z_{kj} E_j}{\sum\limits_{j=1}^{N} (z_{ij} z_{kj})^2},\ i \neq k;\ b_t = \frac{\sum\limits_{j=1}^{N} (z_{ij}^2 - \dfrac{\sum\limits_{j=1}^{N} z_{ij}^2}{N}) \bar{E}_j}{\sum\limits_{j=1}^{N} (z_{ij}^2 - \dfrac{\sum\limits_{j=1}^{N} z_{ij}^2}{N}) \bar{E}_j}. \qquad (3.425)$$

де $n+1 \le l \le n+C_n^2$, $n+C_n^2+1 \le t \le 2n+C_n^2$.

Таблиця 3.37

Матриця планування

№ дослідів	z_1	z_2	z_3	z_1^2	z_2^2	z_3^2	$z_1^2 - \dfrac{\sum z_1^2}{N}$	$z_2^2 - \dfrac{\sum z_2^2}{N}$	$z_3^2 - \dfrac{\sum z_3^2}{N}$	E
1	-1	-1	-1	+1	+1	+1	0,27	0,27	0,27	E_1
2	+1	-1	-1	+1	+1	+1	0,27	0,27	0,27	E_2
3	-1	+1	-1	+1	+1	+1	0,27	0,27	0,27	E_3
4	+1	+1	-1	+1	+1	+1	0,27	0,27	0,27	E_4
5	-1	-1	+1	+1	+1	+1	0,27	0,27	0,27	E_5
6	+1	-1	+1	+1	+1	+1	0,27	0,27	0,27	E_6
7	-1	+1	+1	+1	+1	+1	0,27	0,27	0,27	E_7
8	+1	+1	+1	+1	+1	+1	0,27	0,27	0,27	E_8
9	0	0	0	0	0	0	-0,73	-0,73	-0,73	E_9
10	$+\alpha$	0	0	α^2	0	0	0,75	-0,73	-0,73	E_{10}
11	$-\alpha$	0	0	α^2	0	0	0,75	-0,73	-0,73	E_{11}
12	0	$+\alpha$	0	0	α^2	0	-0,73	0,75	-0,73	E_{12}
13	0	$-\alpha$	0	0	α^2	0	-0,73	0,75	-0,73	E_{13}
14	0	0	$+\alpha$	0	0	α^2	-0,73	-0,73	0,75	E_{14}
15	0	0	$-\alpha$	0	0	α^2	-0,73	-0,73	0,75	E_{15}

Вибір величини α здійснюється за табл. 3.38. Перевірка значущості одержаних коефіцієнтів здійснюється за допомогою критерію Стьюдента: b_i – значущі, якщо $b_i' > \sqrt{\sigma_{b_i'}^2} \cdot t$.

Таблиця 3.38

Вибір величини α

Число змінних	2	3	4	5
ПФЕ або ДФЕ	2^2	2^3	2^4	$\frac{1}{2}2^5$
$2m+1$	5	7	9	11
α	1,000	1,215	1,414	1,520

Дисперсії обчислюються за формулами

$$\sigma_{b'}^2 = \frac{\sigma_E^2}{\sum_{j=1}^{N}(z_{ij})^2};\ \sigma_{b_i}^2 = \frac{\sigma_E^2}{\sum_{j=1}^{N}(z_{ij}z_{kj})^2},\ i \neq k;\ \sigma_{b_i}^2 = \frac{\sigma_E^2}{\sum_{j=1}^{N}\left(z_{ij} - \frac{\sum z_{ij}^2}{N}\right)^2}. \qquad (3.426)$$

2. *Ротобалельне ЦКП.* При ротобалельному ЦКП експеримент планується таким чином, що дисперсія на однакових відстанях від центру постійна, тобто $\sigma_E^2 = const$ при $\sum_{j=1}^{N} z_{ij}^2$. Це одноманітне планування, при якому дисперсія залишається постійною при $0 \le p \le 1$. Матриця планування створюється наступним чином, щоб матриця нормальних рівнянь була інваріантна до ортогональної поверхні. При цьому збільшується (у порівняні з ортогональним ЦКП) число точок у центрі і змінюється величина α. Величина α для ПФЕ рівна: $\alpha = 2^{n/4}$; для ДФЕ:

а) при 0.5 ГІФЕ $\alpha = 2^{(n-1)/4}$; б) при 0.25 ПФЕ $\alpha = 2^{(n-2)/4}$.

Вибір величини α ікількості центральних точок проводиться згідно табл. 3.39.

Таблиця 3.39

Вибір величини α ікількості центральних точок

Число факторів n		Число точок факторного планування	Число заданих точок	Число центральних точок	Загальна кількість точок	α
о		4	4	5	13	1,414
3		8	6	6	20	1,680
4		16	8	7	31	2,000
5	ПФЕ	32	10	10	52	2,378
	ДФЕ	16	10	6	32	2,000

Для обчислення коефіцієнтів рівняння регресії та дисперсії обчислюють постійні для матриці планування величини

$$C = \frac{N}{\sum_{j=1}^{N} z_{ij}^2};\ \chi = \frac{n \cdot N}{(n+2)N_q};\ A = \frac{1}{2\chi[(n+2)\chi - n]}, \qquad (3.427)$$

де N_q – загальна кількість точок (без центральних).

Після реалізації матриці планування та обчислення середніх значень виходу $\overline{E}_j$ у кожному j-му рядку обчислюються суми

$$S_0 = \sum_{j=1}^{N} \overline{E}_j \; ; \; S_i = \sum_{j=1}^{N} z_{ij} \overline{E}_j \; ; \; S_{ii} = \sum_{j=1}^{N} z_{ij}^2 \overline{E}_j \; ; \; S_{il} = \sum_{j=1}^{N} z_{ij} z_{lj} \overline{E}_j \; , \; i \neq l \; . \qquad (3.428)$$

Після обчислення всіх допоміжних величин обчислюються коефіцієнти рівняння регресії

$$b_0 = \frac{A}{N}\left[2\chi^2 (n+2) S_0 - 2\chi C \sum_{i=1}^{n} S_i \right] ; \; b_i = S_i \frac{C}{N} ;$$

$$b_{ii} = \frac{A}{N}\left[C^2 \left[(n+2)\chi - n \right] + C^2 (1-\chi) C \sum_{i=1}^{n} S_{ii} - 2\chi C S_0 \right] ; \; b_{il} = \frac{C}{\chi N} S_{il} \; , \; i \neq l \; . \qquad (3.429)$$

Для визначення значущості коефіцієнтів рівняння регресії обчислюються відповідні дисперсії

$$\sigma_{b_0}^2 = 2A\chi^2 (n+2) \frac{\sigma_{\overline{E}}^2}{N} = \frac{1}{N-n} \sigma_{\overline{E}}^2 \; ; \; \sigma_{b_i}^2 = \frac{\sigma_{\overline{E}}^2}{N} = \frac{\sigma_{\overline{E}}^2}{\sum_{j=1}^{N} z_{ij}^2} \; ; \qquad (3.430)$$

$$\sigma_{b_{ii}}^2 = A\left[(n+1)\chi - (n-1) \right] C^2 \frac{\sigma_{\overline{E}}^2}{N} \; ; \; \sigma_{b_{il}}^2 = C^2 \frac{\sigma_{\overline{E}}^2}{N\chi} \; .$$

Тоді b_i значущий, якщо $b_i > \sqrt{\sigma_{b_i}^2} \cdot t$. Для перевірки адекватності подання функції виходу многочленом другого степеня обчислюються суми

$$S_R = \sum_{j=1}^{N} (E_{jpr} - \overline{E}_j)^2 \; , \; S_E = \sigma_{\overline{E}}^2 (N_0 - 1) \; , \; S_{LF} = S_R - S_E \; . \qquad (3.431)$$

$$E_{jpr} = b_0 + b_1 z_{1j} + b_2 z_{2j} + \ldots + b_n z_{nj} + b_{1,2} z_{1j} z_{2j} + \ldots + \\ + b_{(n-1),n} z_{(n-1)j} z_{nj} + b_{1,1} z_{1j}^2 + \ldots + b_{nn} z_{nj}^2 , \qquad (3.432)$$

обчислюється за експериментальними точками; N_0 – кількість точок у центрі експерименту.

Критерій Фішера обчислюється за формулою

$$F = \frac{S_{LF}}{f_{LF}} \frac{1}{\sigma_{\overline{E}}^2} \; ; \; f_{LF} = N - \frac{(n-2)(n+1)}{2} - (N_0 - 1) \; . \qquad (3.433)$$

Табличне значення F_{tab} знаходиться для $f_1 = f_{LF}$ і $f_2 = N_0$-1 для відповідного рівня значущості q. Порівняння ортогонального ЦКП і рототабельного ЦКП показує, що при однакових умовах проведення експерименту рототабельне планування дає більш точний розв'язок, але число дослідів при рототабельному плануванні більше. Якщо необхідно одержати більш точний опис процесу, завжди слід застосовувати

рототабельне ЦКП. Ортогональне ЦКП потрібно застосовувати у тих випадках, коли маємо обмежену кількість сировини для проведення дослідів або час проведення експерименту істотно впливає на процес, а кожний дослід пов'язаний з новими затратами часу.

3.10. Приклади математичного опису складних об'єктів експериментально-статистичними методами

В цьому параграфі розглядаються практичні приклади математичного опису складних об'єктів з використанням методів пасивного і активного експерименту.

3.10.1. Математичний опис промислового об'єкту методом пасивного експерименту

Об'єктом дослідження є цех формування та сушіння каучукової стрічки на заводі синтетичного каучуку. У процесі попереднього дослідження було встановлено, що цей цех за своєю структурою є типовим складним об'єктом, у якому було виділено 19 вхідних параметрів ($x_1, x_2, ..., x_{19}$), 5 керуючих (z_1, ..., z_5) та 1 вихідний параметр, який оцінює якість каучукової стрічки. Крім того, на процес впливає збурення неконтрольованих зовнішніх факторів. Було встановлено також, що за своїми динамічними властивостями цей об'єкт відповідає ланці з чистим запізненням. Безпосереднім вимірюванням кожного каналу були визначені $\tau_{зап}$, їх величини були в межах від 40 до 60 хвилин. Після закінчення попередніх досліджень був проведений експеримент по вимірюванню статистичних даних. В умовах нормального протікання процесу через однакові інтервали з врахуванням всіх часових зсувів реєструвалися значення 25 контрольованих параметрів. При цьому використовувалися покази реєструючих і вимірювальних приладів та дані спеціально організованих хімічних аналізів. Інтервал між сусідніми замірами складає 1,5 години.

Реєстрації підлягають наступні параметри:

v_1 - час сушіння каучукової стрічки, *хв.*
z_1 - витрата латексу, $м^3/год.$
z_2 - витрати електроліту, *л/год.*
z_3 - витрати оцтової кислоти, *л/год.*
z_4 - витрата II води, $м^3/год.$
z_5 - витрата I води, $м^3/год.$
x_1 - температура зкоагульованої маси, *град.*
x_2 - товщина стрічки, *мм.*
x_3 - витрата III води, $м^3/год.$
x_4 - вміст каучуку в латексі, %.
x_5 - вміст некалю в латексі, %.
x_6 - поверхневий натяг латексу.
x_7 - концентрація оцтової кислоти.
x_8 - вакуум I коробки.
x_9 - вакуум II коробки.
x_{10} - дефо латексу.
x_{11} - продуктивність, *рулон/год.*
x_{12} - розхід латексу (цеховий), *кг/год.*
x_{13} - розхід масляної емульсії, *кг/год.*
x_{14} - температура латексу, *град.*
x_{15} - температура масляної емульсії, *град.*
x_{16} - температура II ємності, *град.*
x_{17} - вміст масла у масляній емульсії.
x_{18} - температура промивної води, *град.*
x_{19} - дефо стрічки.

На протязі експерименту було проведено 82 досліди. В результаті дослідження була отримана таблиця статистичних даних, яка містить 82 компоненти інформації про 25 технологічних параметрів процесу. На основі зібраного статистичного матеріалу методом кореляційного аналізу одержано рівняння, яке описує лінійну залежність вихідного параметра v_1 (час сушіння стрічки) від 24 контролюючих змінних. Усі

необхідні обчислення проводилися на обчислювальній машині «Еліот». Машинний час склав 20 годин. Скоректоване значення коефіцієнта множинної кореляції, який оцінює щільність знайденої залежності, рівне 0,62. Одержане рівняння має дві форми.

1. У стандартизованому масштабі:

$$t_{v_1} = 0.136t_{z_1} - 0.033t_{z_2} - 0.186t_{z_3} + 0.27t_{z_4} - 0.059t_{z_5} + 0.308t_{x_1} - 0.008t_{x_2} + \\ + 0.06t_{x_3} - 0.027t_{x_4} + 0.351t_{x_5} - 0.256t_{x_6} + 0.096t_{x_7} + 0.06t_{x_8} + 0.107t_{x_9} - \\ - 0.033t_{x_{10}} - 0.064t_{x_{11}} - 0.432t_{x_{12}} + 0.245t_{x_{13}} - 0.190t_{x_{14}} + 0.104t_{x_{15}} - 0.425t_{x_{16}} - \\ - 0.067t_{x_{17}} - 0.008t_{x_{18}} - 0.124t_{x_{19}} \qquad (3.434)$$

Введення стандартизованого масштабу дозволяє оцінити порівняльний вплив технологічних параметрів на хід процесу. Величина коефіцієнта при аргументі у рівнянні (3.434) характеризує представництво цього параметра у загальній сукупності змінних, які описують процес, а знак коефіцієнта визначає напрямок ефекту.

2. У натуральному масштабі:

$$v_1 = 16.785 + 0.133z_1 - 0.006z_2 - 0.08z_3 + 0.216z_4 - 0.05z_5 + 4.038x_1 - \\ - 0.007x_2 + 0.311x_3 - 0.319x_4 + 0.432x_5 - 2.728x_6 + 0.062x_7 + 0.004x_8 + \\ + 0.007x_9 - 0.00093x_{10} - 0.087x_{11} - 0.106x_{12} + 0.062x_{13} - 0.161x_{14} + 0.141x_{15} - \\ - 0.302x_{16} - 0.298x_{17} - 0.005x_{18} - 0.001x_{19}. \qquad (3.435)$$

Рівняння (3.435), подане у натуральному масштабі виміру змінних, дозволяє шляхом підстановки поточних значень параметрів обчислити значення вихідного параметра. Наведений приклад знаходження математичного опису процесу формування і сушіння каучуку показує, що метод кореляційного аналізу, який спирається на можливості сучасних обчислювальних машин, дозволяє побачити якісні і кількісні залежності між контролюючими параметрами складних процесів.

3.10.2. Математичний опис напіввиробничого об'єкту методом активного експерименту

Вивчалася напіввиробнича установка, яка характеризується двома управляючими факторами Z_1 і z_2 і одним вихідним параметром E^*. Фізично Z_1 представляє температуру реактора, яка вимірюється у градусах за Цельсієм, Z_2 – концентрація реагенту, яка вимірюється у процентах, і E – кількісний вихід метакрилової кислоти. Вихідний продукт вимірюється у грамах річного виходу E і виражається у процентах по відношенню до загального виходу. Інтервали варіювання: $20°C \le z_1 \le 170°C$. $5\% \le z_2 \le 50\%$. Точність вимірювання: $\Delta z_1 = \pm 1°C$; $\Delta z_2 = \pm 0.05\%$; $\Delta E = \pm 1.5\%$. Згідно зі статистикою процесу, дисперсія одиничного виміру є $\sigma^2_{E\,cm} = 7.9$.

Початкова точка була вибрана як найбільш оптимальна, згідно статистичних даних z_{10}=80°C. z_{20}=32.6%. E=86.8%. Кроки варіювання були вибрані таким чином, щоб стати вище над похибками вимірювання: $\lambda_1 = 5°C$; $\lambda_2 = 1\%$. Карта проведення досліду подана у табл. 3.40. З табл. 3.40 видно, що спочатку були поставлене ПФЕ при двох

дублюючих дослідах, при відповідній рандомізації дослідів. Була проведена перевірка рівноточності вимірювань за допомогою критерію Кохрена і перевірка адекватності за допомогою критерію Фішера. У табл. 3.40 вказані всі результати обчислення. Була отримана дисперсія одиничного вимірювання $\sigma_E^2 = 7.0$, яка показує, що досліди поставлені добре, а експерименти відтворювані, бо $\sigma_E^2 \approx \sigma_{Ecm}^2$.

Таблиця 3.40

Приклад карти проведення досліду

Позначення фактора				$\xi_1(T)$	$\xi_2(C)$	$E(\%)$			$\overline{E}_j^2$	σ_j^2
Основний рівень				80	32,6	86,6				
Крок варіювання				5	1	E_j	E_{j_1}	E_{j_2}		
Порядок реалізації	3	2	1	-1	-1	84,0	82,0	86,0	7056,00	8,0
	1	4	2	+1	-1	87,5	89,0	86,0	7656,01	4,5
	4	3	3	-1	+1	81,7	83,4	80,0	5674,90	5,8
	2	1	4	+1	+1	85,0	82,8	87,2	7225,00	9,7
Порядковий номер										
b_i				1,7	-1,2	b_0=84.5		1 =	28612,0	28,0
$b_i \times \lambda_i$				8,5	-1,2	E_{j_1}	$E_{j_{pr}}$	G = 0.346/$G_{tab\ 1\%}$ =0.968 F = 33.5/3.5 = 9.6 $F_{tab\ 1\%}$= 12.75; f_1=1; f_2= 7 σ_E^2= 3.5; σ_b =0.93		
Новий крок варіювання				5	-0,7					
Новий основний рівень				85	31,6	87,5	-			
Уявні досліди			1	90	30,9	89,5	87,04	Перший кроковий рух $E_{j_{pr}}$ = 84.5 + 2.54j		
			2	95	30,2	91,7	89,58			
			3	100	29,5	92,0	92,12			
			4	105	28,8	90,5	94,66			
			5	110	28,1	81,3	97,20			
Основний рівень				100	29.5	92.0	$\overline{E}_j^2$	$\sum_j \overline{E}_j^2$ = 33236.2 $N\sum_i b_i^2$ = 33218.0 $F = \frac{33236.2 - 33218.0}{3.5} = 5.2$ σ_b = 0.93 $F < F_{tab}$		
Крок варіювання				5	1	-				
Порядок реалізації	3		1	-1	-1	88,0	7744,0			
	1		2	+1	-1	90,5	8100,0			
	4		3	-1	+1	92,7	8556,0			
	2		4	+1	+1	94,0	8836,0			
Порядковий номер										
b_i				0,88	2,12	b_0=91.1				
$b_i \times \lambda_i$				4,40	2,12	E_{j_1}	$E_{j_{pr}}$			
Новий крок варіювання				5	2,4					
Новий основний рівень				105	30,5	94,0	-			
Уявні досліди			1	110	32,9	97,2	97,0	Другий кроковий рух $E_{j_{pr}}$ = 91.1 + 5.9j		
			2	115	35,3	94,2	102,9			
			3	120	37,7	78,0	108,8			
Основний рівень				110	32,9	97,0	$\overline{E}_j^2$	$\sum_j \overline{E}_j^2$ = 34869.9 $N\sum_i b_i^2$ = 34944.0 $F = \frac{71.1}{3.5} = 20.3$ $F > F_{tab}$		
Крок варіювання				5	2,5	-				
Порядок реалізації	2		1	-1	-1	93,4	8723,5			
	4		2	+1	-1	88,0	7744,0			
	1		3	-1	+1	98,0	9604,0			
	3		4	+1	+1	98,0	8798,4			
Порядковий номер										
b_i				-2,35	2,65	b_0=93.35				
Порядок реалізації	12		5	0	0	96,8	α = 1.414; C = 1.63 χ = 0.81; A = 0.5 S_0 =1226.9; S_1=18.6; S_2=10.6 S_{11}=751.4; S_{22}=733.4; S_{21}=1.4				
	7		6	0	0	97,0				
	9		7	0	0	98,0				
	8		8	0	0	96,0				
Порядковий номер										

							b_0 =96.9; b_1 =-2.3; b_2=1.3
	10		9	0	0	72,2	b_{11}= -0.8; b_{22}=-3.0; b_{12}=0.35
	5		10	+1,414	0	91,0	
	13		11	-1,414	0	97,5	
	6		12	0	+1,414	90,0	
	11		13	0	-1,414	90,0	

Після першої серії дослідів були обчислені коефіцієнти рівняння регресії і перевірена їх значущість, тому оскільки обидва коефіцієнти значущі, і апроксимація гіперплощиною адекватна, було вирішено здійснювати покроковий рух. Для цього була вибрана точка з найкращим виходом і прийнята за основну. Кроки варіювання були одержані за загальним правилом: $\Delta = \frac{1}{1.7}$, звідки $\lambda_{2kv} = -1.2 \cdot \frac{1}{1.7} = -0.7$.

Рівняння гіперплощини $E_{jp} = 84.5 + j(1 \cdot 1.7 - 1.2(-0.7)) = 84.5 + 2.54\,j$.

В результаті першого покрокового руху була вибрана базова точка для проведення другої серії експериментів. Надалі не проводилося дублювання дослідів, а використовувалася дисперсія, що одержана з першої серії експериментів. Після опрацювання результатів другої серії експериментів виявилося, що $b_1 = 0.88 < \sigma_b = 0.93$ b. Оскільки має місце наближена рівність $b_1 \approx \sigma_b$, то будемо вважати b_1 значущим.

Після перевірки адекватності був проведений другий покроковий рух і підставлена третя серія експериментів. Оскільки апроксимація рівняння регресії неадекватна гіперплощині в околі цієї точки, експерименти були доповнені до рототабельного і результати опрацьовані. Для визначення значущості одержаних коефіцієнтів рівняння регресії були обчислені дисперсії $\sigma^2_{b_0} = 0.423$; $\sigma^2_{b_i} = 0.26$; $\sigma^2_{b_{ij}} = 0.1$. Перевірка значущості:

$(b_0) = 96.9 > \sqrt{0.423} = 0.65$ – значущий; $(b_1) = 2.3 > \sqrt{0.26} = 0.51$ – значущий;

$(b_2) = 1.3 > 0.51$ – значущий; $(b_{11}) = 0.8 > \sqrt{0.36} = 0.62$ – значущий;

$(b_{22}) = 3.0 > 0.62$ – значущий; $(b_{12}) = 0.35 > \sqrt{0.1} = 0.316$ – значущий.

Таким чином, многочлен другого порядку записується у вигляді:

$$E = 96.6 - 2.3z_2 - 0.8z_1^2 - 3.0z_2^2 - 0.35z_1z_2.$$

Адекватність перевіряється згідно формул. Обчислені $\sigma_R = 17.84$; $\sigma_E = 3.5 \cdot 4 = 14$; $\sigma_{LF} = 17.84 - 14 = 3.84$; $F = \frac{3.84/3}{3.5} = 0.366$. З таблиць для $f_1 = 13 - \frac{4 \cdot 3}{2} - (5-1) = 3$; $f_2 = 5 - 1 = 4$; $q = 5\%$. Знаходимо $F_{табл} = 6.59$. Звідки $F < F_{табл}$.

Отже, апроксимація многочленом другого степеня адекватна.

3.11. Кількість спостережень

Дослідника, який приступає до статистичного аналізу складного об'єкту регулювання, перш за все цікавить питання, яку кількість спостережень (експериментів) слід здійснити, щоб за одержаними даними можна надійно судити про властивості

об'єкту, який досліджується. На мові статистики задача формулюється наступним чином: який повинен бути об'єм вибіркової сукупності, щоб з певною ймовірністю P можна було б вважати, що відхилення від справжнього значення досліджуваної властивості не перевищує допустиму похибку ε. У першому наближенні розв'язку задачі про число спостережень дає таблиця достатньо великих чисел (табл. 3.41). Ця таблиця складена за формулами, які випливають з теореми Бернуллі.

Таблиця 3.41

Перше наближення розв'язку задачі про кількість спостережень

P / ε	0.85	0.90	0.95	0.99	0.995	0.999
0.05	207	270	384	663	787	1082
0.04	323	422	600	1036	1231	1691
0.03	575	751	1067	1843	2188	3007
0.02	1295	1621	2400	4146	4924	6767

Відносно величини ймовірності P немає якогось загального алгоритму, однакового при всіх дослідженнях. Чим ближче до 1 буде ймовірність, тим надійнішим буде висновок. На практиці приймають P=0.95. Допустима похибка ε при дослідженнях встановлюється в залежності від природи досліджуваного явища. В більшості випадків допустима похибка приймається рівною ε =0.95. Результати, які дає таблиця достатньо великих чисел, дуже завищені. Число спостережень, які необхідні для дослідження, можна значно зменшити, якщо мати додаткову інформацію про характер зміни технологічного параметра. Припустимо, що дослідник знає так звану «міру розкиду», тобто виражене у процентах відношення σ середнього квадратичного відхилення а до середнього значення параметра $\bar{x}$

$$h = \frac{\sigma}{\bar{x}}. \qquad (3.436)$$

Якщо відома міра розкиду h, то досить велике число (при ймовірності P і допустимій похибці ε) може бути знайдено за допомогою теореми Ляпунова для незалежних, однаково розподілених випадкових величин зі скінченою дисперсією. Складена номограма для визначення числа необхідних спостережень N при заданій ймовірності P і мірі зміни h. В табл. 3.42 приводиться деяке значення N, визначені за цією номограмою для випадку P = 0.99. З порівняння даних табл. 3.41 (стовпець з P=0.99) і табл. 3.35 видно, що значення «міри розкиду» параметру дозволяє набагато зменшити число необхідних дослідів.

Таблиця 3.42

Друге наближення розв'язку задачі про кількість спостережень

ε / h	0.02	0.03	0.04	0.05
5	40	18	11	7
10	160	70	40	25
15	350	160	90	55
20	650	300	160	100
25	1000	450	250	160
50	4000	1700	1000	650

Викладені методи визначення числа необхідних спостережень, однак, не відповідають на важливе для дослідника питання, за який проміжок часу слід провести вказані спостереження. Зрозуміло, що якщо ці спостереження проводяться за короткий

проміжок часу, на протязі якого технологічний параметр змінюється лише на незначну частину свого діапазону, то одержана інформація не відображає його істотні статистичні властивості. Висновки, зроблені на основі таких спостережень, будуть справедливі лише для вузької області зафіксованих змін і не можуть бути поширені на весь робочий діапазон зміни змінної. Таким чином, при розв'язанні питання про число спостережень виконання однієї тільки вимоги про достовірність отриманих результатів недостатньо, необхідно враховувати характер зміни параметра у часі.

Розглянемо питання про визначення числа спостережень, орієнтуючись на тимчасові характеристики досліджуваного параметра. Конкретизуємо задачу так. Відомо: діапазон зміни технологічного параметра x (від x_A до x_B).

Ймовірності попадання технологічного параметра на нижню межу діапазону p_1 і верхню межу p_2.

Визначити час T, необхідний для визначення статистичних властивостей параметра при дискретних вимірюваннях з інтервалом Δt. Будемо вважати час спостереження достатнім, якщо за серію $N = \dfrac{T}{\Delta t}$ вимірювань досліджуваний параметр хоча б один раз попаде на верхню і на нижню межі діапазону. Природно очікувати, що при виконання цієї вимоги буде охоплений цілий діапазон зміни параметра, так як ймовірність попадання в середину діапазону значно більша, ніж на її границі (рис. 3.13а). Зазвичай технологічні параметри добре налагоджених виробництв мають невеликі зміни коливань, і, відповідно, ймовірності попадання на межі діапазону p_1 і p_2 будуть малими. В цьому випадку можна використовувати формулу Пуассона. Тоді ймовірність P_m того, що технологічний параметр з'явиться на межі діапазону рівно m разів, можна розраховувати за формулою:

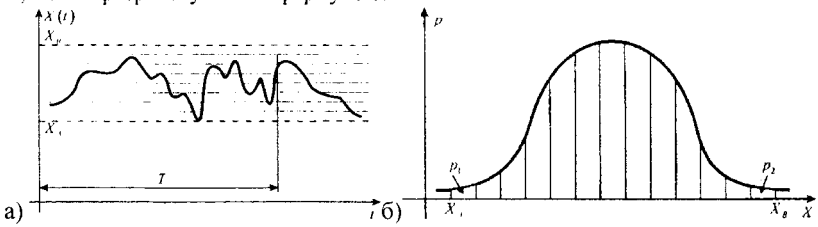


Рис. 3.13. а) зміна досліджуваного параметра у часі; б) визначення ймовірностей p_1 та p_2 попадання на верхню і на нижню границі діапазону

$$P_m = \frac{\lambda^m}{m!} e^{-\lambda}. \tag{3.437}$$

$$\lambda = aT. \tag{3.438}$$

де a – середнє число попадань на межу діапазону в одиницю часу, а T – проміжок часу, на протязі якого проводяться вимірювання. Отже, λ рівне середньому числу попадань параметра на границю діапазону за весь час вимірювання. Ймовірність того, що параметр хоча б раз попаде на границю, можна визначити таким чином

$$P_1 = \sum_{m=1}^{N} P_m = 1 - P(0). \tag{3.439}$$

де $P(0)$ визначає ймовірність того, що параметр жодного разу не попаде на границю діапазону ($m = 0$). Згідно (3.437)

$$P(0) = e^{-\lambda}. \tag{3.440}$$

Таким чином, ймовірність того, що технологічний параметр хоча б один раз попаде на нижню границю, визначається за формулою

$$P_1' = 1 - e^{-\lambda_1}. \tag{3.441}$$

Аналогічно для верхньої границі

$$P_1'' = 1 - e^{-\lambda_2}. \tag{3.442}$$

За серію N вимірювань повинно відбутися співпадіння цих подій. Тому, загальна ймовірність попадання параметра хоча б один раз і на верхню і на нижню границю діапазону визначається добутком цих ймовірностей

$$P_z = P_1'P_1'' = (1 - e^{-\lambda_1})(1 - e^{-\lambda_2}). \tag{3.443}$$

Розглянемо випадок, коли ймовірність попадання на нижню і верхню межі одинакові

$$p_1 = p_2 = p. \tag{3.444}$$

Це означає, що розподіл ймовірності всередині діапазону симетричний (нормальний, біноміальний та інші симетричні розподіли).

Тоді $\lambda_1 = \lambda_2 = \lambda$ і формула (3.443) має вигляд

$$P_z = (1 - e^{-\lambda})^2. \tag{3.445}$$

Прийнявши потрібне P_z, з виразу (3.445) можна знайти λ, а потім за формулою (3.438), визначити потрібний час виміру. Розглянемо наступні прийняті в наукових дослідженнях значення P_z: 0,94; 0,95; 0,96; 0,97; 0,98; 0,99. Підставляючи перше значення у вираз (3.445), маємо $0.94 = (1 - e^{-\lambda})^2$, звідки $\lambda = 3.5$. Проводячи аналогічні розрахунки для інших прийнятих значень P_z, отримаємо такі значення λ (табл. 3.43).

Таблиця 3.43

Середнє число попадань параметра на границю діапазону за весь час досліду

p	0,94	0,95	0,96	0,97	0,98	0,99
λ	3,52	3,68	3,90	4,19	4,60	5,30

Визначимо число попадань параметра на границі діапазону за одиницю часу a. Якщо Δt – час між вимірюваннями, то $\frac{1}{\Delta t}$ – є число t вимірів за одиницю часу. Тоді величину a знайдемо, помноживши ймовірність попадання p на частоту вимірювань

$$a = p\frac{1}{\Delta t}. \tag{3.446}$$

$$\lambda = \frac{p}{\Delta t}T. \tag{3.447}$$

$$T = \frac{\Delta t}{p}\lambda . \qquad (3.448)$$

Маючи щільність розподілу досліджуваної змінної, величину p можна визначити як ймовірність попадання значень змінної в крайній квадрат діапазону (рис. 3.13б). Інтервал часу між вимірами Δt визначається на основі міркувань, викладених у наступному параграфі. Визначивши відношення $\frac{p}{\Delta t}$ і вибравши з табл. 3.43 значення λ, яке відповідає прийнятій загальній ймовірності P_z, але за формулою (3.448) визначимо шукану величину T. Вона одержується у тих самих одиницях, що й Δt (сек., хв., год. тощо). Число необхідних спостережень можна визначити знаючи T і Δt

$$N = \frac{T}{\Delta t} . \qquad (3.449)$$

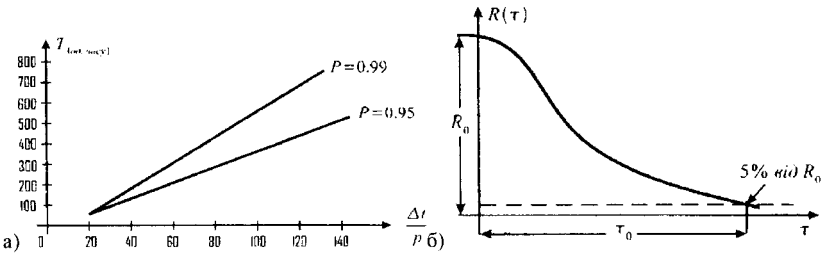


Рис. 3.14. а) залежності часу спостереження T від відношення $\frac{\Delta t}{p}$;

б) автокореляційна функція досліджуваної змінної

На рис. 3.14а подані залежності T від $\frac{\Delta t}{p}$ при двох найбільш вживаних значеннях P. Ці залежності дозволяють легко визначити час спостереження T.

3.12. Квантування за часом

Як випливає з розгляду кореляційного і регресійного методів, при статистичному аналізі складних об'єктів немає необхідності реєструвати неперервні заміни змінних, вся необхідна інформація одержується в результаті опрацювання даних окремих дослідів. Ця обставина дозволяє експериментатору у широких межах варіювати частоту знімання даних, або, що те ж саме, величину інтервалу часу Δt між двома сусідніми вимірами. Свобода у виборі інтервалу Δt практично обмежена наступними міркуваннями. Завищена частота зміни веде до того, що результати двох сусідніх експериментів виявляються між собою взаємопов'язаними (корельованими). У цьому випадку результати вимірів не можна розглядати як сукупність випадкових величин. З іншого боку, якщо інтервал відліків Δt вибраний дуже великим, результати експерименту виявляться під впливом повільних, нестаціонарних змін технологічних параметрів об'єкта. При цьому інформація, яка міститься в даних початкових дослідів, буде поступово старіти, втрачати свою цінність через умови експерименту, які

змінилися. Крім того, збільшення Δt призводить до затягування експерименту по зборі статистичних даних і до його подорожчання. Практично при виборі інтервалу експериментатор часто обмежений знизу роздільною здатністю контрольно-вимірювальної апаратури. Наприклад, у цей час у більшості випадків якість продуктів контролюється шляхом лабораторних хіманалізів, тривалість яких у ряді випадків становить кілька годин. Експериментатор не може приступати до відбору нових проб, поки не готові результати попереднього досліду, а організація паралельних аналізів приводить до залучення додаткового устаткування й фахівців.

Впровадження на виробництві систем контролю автоматичних аналізуючих приладів дозволить значно збільшити швидкодію процесу попереднього опрацювання даних і, тим самим, підвищити частоту проведення досвідів. У цьому випадку завдання дослідника полягає в тому, щоб синхронізувати в часі роботу всіх автоматичних контрольно-вимірювальних приладів, тобто необхідно вибрати такий інтервал знімання даних Δt, що був би кратний дискретності роботи всіх використовуваних автоматичних пристроїв. Зупинимося більш докладно на кількісній оцінці величини Δt. Вибір інтервалу часу між вимірами оснований на наступних міркуваннях. Як відомо, методи кореляційного й регресійнюго аналізу використовують дані незалежних спостережень. При дискретній реєстрації безперервної випадкової змінної два сусідніх спостереження неправильно вважати незалежними, якщо вони розділені інтервалом часу $\Delta t = \tau_0$, де τ_0 – час кореляції, який визначається за автокореляційною функцією досліджуваної змінної. Таким чином, нижньою межею для величини Δt можна вважати час кореляції та (рис. 3.14б). При визначенні верхньої межі до уваги беруться такі фактори, як можлива нестаціонарність досліджуваного процесу, що викликає старіння накопичуваної інформації; порушення умов, у яких здійснюється експеримент, і матеріальні витрати, пов'язані з його проведенням. Цікаво розглянути у якій мірі зібраний матеріал може бути використаний для відновлення безперервної вихідної функції. З ростом Δt за рахунок втрати інформації зростає похибка відтворення шуканої функції. Ефект квантування безперервного сигналу будемо розглядати як дію деякої еквівалентної перешкоди квантування за часом. Для кількісного опису цієї «перешкоди» скористаємося середньоквадратичною оцінкою.

При дискретному контролі здійснюється вимірювання значень контрольованого параметра $x(t)$ у дискретні моменти часу $t_i = i \cdot \Delta t$, де Δt – період дискретизації, i – ціле число й інтерполяція цих значень на інші моменти часу. Різниця між сигналом $x^*(t)$, отриманим у результаті інтерполяції, і кривою зміни контрольованого параметра $x(t)$ становить похибку інтерполяції $e(t)$. Процес утворення похибки інтерполяції проілюстрований умовною блок-схемою рис. 3.15а. Відбір дискретних значень вхідного сигналу здійснюється ідеальним імпульсним елементом (ІЕ), що формує миттєві імпульси, які відстоять один під одного на проміжок часу Δt, площі яких дорівнюють значенням вхідного сигналу в моменти часу, які збігаються з моментами їхньої появи. Інтерполяція здійснюється безперервним лінійним фільтром з передаточною функцією $I(j\omega)$. Для кривої, зображеної на рис. 3.15б, $I(j\omega) = \dfrac{1 - e^{-j\omega\Delta t}}{j\omega}$, що відповідає екстраполятору нульового порядку. Труднощі обчислення вихідного сигналу $\varepsilon(t)$ пов'язані з тим, що система рис. 3.15б не приводиться безпосередньо до звичайної імпульсної системи, оскільки вона реагує на вхідний сигнал не тільки в дискретні

моменти часу. Однак, цю систему розглядають як систему з періодично змінними параметрами. Дійсно, ідеальний імпульсний елемент можна вважати безінерційним підсилювачем, коефіцієнт підсилення якого $u(t)$ є періодичною функцією часу

$$u(t) = \sum_{n=-\infty}^{+\infty} e^{j\frac{2\pi n}{\Delta t}t}, \tag{3.450}$$

де Δt – період дискретизації; $\delta(t)$ – дельта-функція Дірака.

Звідси для імпульсної перехідної функції системи $I_\varepsilon(t;\tau)$ одержимо

$$I_\varepsilon(t;\tau) = u(\tau)\cdot I(t-\tau) - \delta(t-\tau) = \frac{1}{\Delta t}\sum_{n=-\infty}^{+\infty} e^{j\frac{2\pi n}{\Delta t}t}\cdot I(t-\tau) - \delta(t-\tau). \tag{3.451}$$

де $I(t)$ – імпульсна перехідна функція інтерполюючого фільтра.

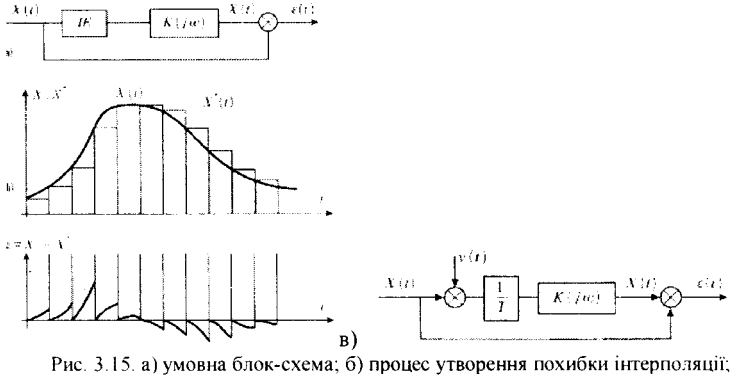


Рис. 3.15. а) умовна блок-схема; б) процес утворення похибки інтерполяції; в) еквівалентна умовна неперервна система для визначення середнього квадрату похибки інтерполювання

Застосовуючи до формули (3.451) перетворення Фур'є за аргументом m, одержимо вираз для передаточної функції $I_\varepsilon(jw;t)$ системи

$$I_\varepsilon(j\omega;t) = \int_{-\infty}^{+\infty} I_\varepsilon(t;\tau)\cdot e^{j\omega\tau}\cdot d\tau = \frac{1}{\Delta t}\sum_{n=-\infty}^{+\infty} I\left(j\omega + j\frac{2\pi n}{\Delta t}\right)\cdot e^{j\frac{2\pi n}{\Delta t}t} - 1. \tag{3.452}$$

Нехай тепер вхідний сигнал $x(t)$ є стаціонарним випадковим процесом із відомою спектральною щільністю потужності $S_x(\omega)$. Сигнал похибки $\varepsilon(t)$ буде при цьому періодично нестаціонарним випадковим процесом, і його середній квадрат визначається в результаті подвійного усереднення за часом і за реалізаціями. Відповідно до цього для середнього квадрату σ_ε^2 сигналу похибки маємо

$$\sigma_\varepsilon^2 = \frac{1}{\Delta t}\int_0^{\Delta t}\frac{1}{2\pi}\int_{-\infty}^{+\infty} I(j\omega;t)\cdot I(-j\omega;t)\cdot S_x(\omega)d\omega. \tag{3.453}$$

Змінюючи в цій формулі порядок інтегрування й з врахуванням того, що

$$\frac{1}{\Delta t}\int_0^{\Delta t} e^{j\frac{2\pi}{\Delta t}} dt = \begin{cases} 1, & i = 0; \\ 0, & i \neq 0; \end{cases} \tag{3.454}$$

одержимо

$$\sigma_\varepsilon^2 = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left\{-1+\frac{1}{\Delta t}[I(j\omega)\cdot I(-j\omega)]+\frac{1}{(\Delta t)^2}\sum_{i=-\infty}^{+\infty}\left[I\left(j\omega+j\frac{2\pi i}{\Delta t}\right)\right]^2\right\}S_x(\omega)d\omega =$$
$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|-1+\frac{1}{\Delta t}I(j\omega)\right|^2 S_x(\omega)d\omega+\frac{1}{2\pi}\int_{-\infty}^{+\infty}\sum_{\substack{i=-\infty\\ i\neq 0}}^{+\infty}\frac{1}{(\Delta t)^2}\left|I\left(j\omega+j\frac{2\pi i}{\Delta t}\right)\right|^2 S_x(\omega)d\omega. \tag{3.455}$$

Виконуючи в останньому інтегралі заміну змінних ω на $\left(\omega^* + \frac{2\pi i}{\Delta t}\right)$, одержимо

$$\sigma_\varepsilon^2 = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|-1+\frac{1}{\Delta t}I(j\omega)\right|^2 S_x(\omega)d\omega+\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|\frac{I(j\omega)}{\Delta t}\right|^2\sum_{\substack{i=-\infty\\ i\neq 0}}^{+\infty}S_x\left(\omega+\frac{2\pi i}{\Delta i}\right)d\omega. \tag{3.456}$$

З виразу (3.456) видно, що середній квадрат вихідного сигналу розглянутої системи дорівнює середньому квадрату вихідного сигналу умовної неперервної системи, що зображена на рис. 3.15в, на яку, крім вхідного сигналу, діє незалежний шум $y(t)$ зі спектральною щільністю $S_y(\omega)$, рівною

$$S_y(\omega) = \sum_{\substack{i=-\infty\\ i\neq 0}}^{+\infty} S_x\left(\omega+\frac{2\pi i}{\Delta i}\right). \tag{3.457}$$

Цю похибку будемо називати *похибкою квантування за часом*. Для середнього квадрата похибки інтерполяції маємо

$$\sigma_\varepsilon^2 = \frac{1}{2\pi}\int_{-\infty}^{+\infty} S_x(\omega)d\omega+\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left|\frac{I(j\omega)}{\Delta t}\right|^2 S_x^*(\omega)d\omega-\frac{1}{\pi}\int_{-\infty}^{+\infty}\mathrm{Re}\left\{\frac{I(j\omega)}{\Delta t}\right\}S_x(\omega)d\omega, \tag{3.458}$$

де $S_x^*(\omega) = \sum_{i=-\infty}^{+\infty} S_x\left(\omega+\frac{2\pi i}{\Delta i}\right)$ – періодична парна функція частоти.

Інтегруючи рівняння (3.457) у загальному вигляді, одержимо вираз для відносного середньоквадратичного значення похибки квантування за часом

$$\sigma^2 = \frac{\sigma_\varepsilon^2}{\sigma_x^2} = 1-\frac{4}{\pi\sigma_x^2\cdot\Delta t}\int\frac{S_x(\omega)}{\omega}\cdot\sin\left(\frac{\omega\cdot\Delta t}{2}\right)d\omega, \tag{3.459}$$

де σ_x^2 – середньоквадратичне значення вимірюваної величини.

Розглянемо тепер кількісну залежність середньоквадратичного значення похибки квантування за часом від періоду квантування Δt для двох випадків:

а) Вхідний сигнал має спектральну щільність у вигляді обмеженого рівномірного спектра

$$S_x(\omega)=\begin{cases}\dfrac{\pi\sigma_x^2}{\omega_{\max}}, & |\omega|\le\omega_{\max},\\ 0, & |\omega|>\omega_{\max},\end{cases} \tag{3.460}$$

$$R_{xx}(\tau)=\frac{\sigma_x^2\cdot\sin(\omega_{\max}\cdot\tau)}{\omega_{\max}\cdot\tau}. \tag{3.461}$$

У цьому випадку вираз для відносного середньоквадратичного значення похибки квантування буде мати вигляд

$$\sigma^2=\frac{\sigma_\varepsilon^2}{\sigma_x^2}=2-\frac{2}{\omega_{\max}\cdot\Delta t}\int_0^{\omega_{\max}}\frac{\sin(\omega\cdot\Delta t)}{\omega}d\omega. \tag{3.462}$$

Для значення $\omega_{\max}\cdot\Delta t\le 1$ можна користуватися наближеною формулою

$$\sigma^2=\frac{\omega_{\max}^2(\Delta t)^2}{9},. \tag{3.463}$$

яка дає розбіжність, що не перевищує 3%.

На рис. 3.16а показана залежність $\sigma^2=f(\omega_{\max}\cdot\Delta t)$, обчислена за (3.420).

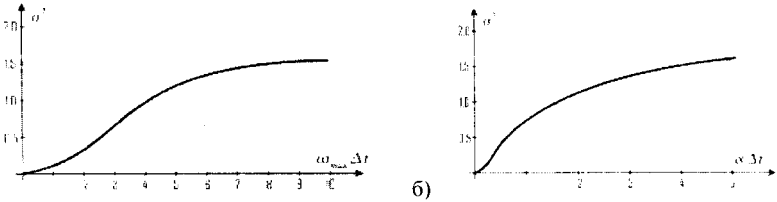
б) Вимірювана величина має спектральну щільність і кореляційну функцію

$$S_x(\omega)=\frac{2\alpha\cdot\sigma_x^2}{\alpha^2+\omega^2},\quad R_{xx}(\tau)=\sigma_x^2\cdot e^{-\alpha|\tau|}. \tag{3.464}$$

Для такого сигналу середньоквадратичне значення перешкоди має вигляд

$$\sigma^2=2-\frac{2}{\alpha\cdot\Delta t}(1-e^{-\alpha\Delta t}). \tag{3.465}$$

Залежність $\sigma^2=\varphi(\alpha;\Delta t)$ зображена на рис. 3.16б.



а) б)

Рис. 3.16. Залежність середнього квадрата помилки інтерполяції від добутку:

а) $\omega_{\max}\cdot\Delta t$ та б) $\alpha\cdot\Delta t$

Використовуючи графіки рис. 3.16 можна визначити співвідношення дисперсії корисного сигналу й похибки, та обрати період квантування за часом.

Цікаво відзначити, що відношення похибки до шуму рівне одиниці для сигналу з

рівномірною обмеженою спектральною щільністю дає співвідношення

$$\Delta t = \frac{3.6}{\omega_{\max}}, \tag{3.466}$$

де $\omega_{\max}$ – максимальна частота спектра, а для сигналу зі спектральною щільністю вигляду $\frac{2\alpha\sigma_x^2}{\alpha^2+\omega^2}$ будемо мати $\Delta t = \frac{1.6}{\alpha} = 1.6T_k$, де T_k – стала часу автокореляційної функції вхідного сигналу.

3.13. Метод експертних оцінок перевірки адекватності математичних моделей

Відомо, що основні задачі економічних і технологічних систем можуть бути успішно розв'язані за допомогою різноманітних методів математичного моделювання, якщо адекватність цієї моделі задовільняє умови її використання на практиці.

Як наслідки цього, в промисловій та непромисловій сфері, в технологічних, економічних та інших процесах відсутні автоматизовані системи управління з функціями, які притаманні експертним системам по забезпеченню видачі рекомендацій на підтримку оптимальних рішень і проведення моделювання економічних, соціологічних, екологічних та інших реальних систем.

В зв'язку з цим проблема перевірки адекватності математичних моделей, отриманих з використанням теоретичних або експериментальних методів залишається актуальною. Тому мають право на життя математичні методи або методичні прийоми, які дозволяють вирішити цю задачу хоча б частково.

Перевірка адекватності математичної моделі методом Фішера, яка отримана за допомогою методів планування активного експерименту або методів кореляційно-регресивного аналізу, як правило практично зводиться до визначення адекватності перетворення табличної моделі (матриці спостереження) в алгебраїчне рівняння, яке виступає математичною моделлю. Це означає перевірку похибки розрахунків коефіцієнтів моделі відносно матриці спостережень. Зауважимо, що в умовах комп'ютерного опрацювання інформації похибка цих перетворень завжди задовільна. Але це не може бути перевіркою адекватності отриманої математичної моделі відповідним реальним властивостям об'єкта дослідження, бо неможливо визначити похибку даних статистичної матриці спостереження параметрів об'єкта дослідження. Таким чином, перевірка адекватності отриманої математичної моделі відповідному реальному об'єкту дослідження існуючими методами неможлива. Зауважимо, що отримані математичні моделі за допомогою теоретичних положень неможливо перевірити на адекватність. Тому рекомендують після отримання математичної моделі перевіряти експериментально умови, для яких в експериментальних дослідженнях існують методи перевірки адекватності одержаних математичних моделей реального процесу. А це означатиме, що дослідник завжди використовує побудовані математичні моделі, не знаючи їх адекватності реальному процесу.

Запропонований метод отримання експертних оцінок адекватності математичних моделей реальним процесам в певній мірі дає дослідникам кількісну відносну оцінку адекватності без визначення абсолютної похибки або оцінок параметрів за вибіркою.

Запропонований метод експертних оцінок адекватності математичних моделей реальному процесу рекомендується використовувати, на нашу думку, за таких умов: *по-перше*, існуючі методи не забезпечують розв'язок цієї задачі; *по-друге*, існує достатня кількість експертів (фахівців), які без використання математичних моделей інтуїтивно вирішують задачі прогнозування і визначення оптимальних (раціональних) режимів функціонування процесу, що вивчається; *по-третє*, існують умови проведення експериментів перевірки адекватності математичних моделей або є додаткова інформація про об'єкт дослідження після отримання математичної моделі, яка може бути використана для додаткової перевірки адекватності моделі. Крім цього визначено, що питання доцільності та використання цієї математичної моделі в практичній діяльності або в проведенні математичного моделювання актуальне вже зараз, а не тільки в майбутньому.

Тут закладена гіпотеза, що із суб'єктивних знань фахівців високої кваліфікації можливо одержати кількісну оцінку явищ з відповідною вірогідністю та малою величиною похибки. Це дало можливість запропонувати методичні прийоми перевірки адекватності математичних моделей. Алгоритм *методу експертних оцінок* перевірки адекватності математичної моделі складається з наступних етапів.

I. Визначається група експертів з провідних фахівців відповідної галузі об'єкта досліджень в кількості $S = 1,2,...,k$, де $k > 5$. Не рекомендується збільшувати кількість експертів за рахунок малокомпетентних фахівців. Значні похибки їх прогнозу погіршать показники загальної оцінки такої групи експертів, і, отже, будуть знижені вимоги до перевірки моделі. Дослідник – автор математичної моделі в число експертів не включається. Після чого експерти і дослідник визначають план проведення додаткових експериментів з перевірки адекватності моделі. Керівником колективу проведення експериментів слід призначити фахівця цієї галузі, але не з числа експертів і того, хто не приймав участі в отриманні математичної моделі.

II. Дослідник разом з групою експертів заносять значення рівнів вхідних параметрів в таблицю 1, згідно погодженого ними плану експериментів, яка складається з рядків у запланованій кількості $1,2,...i,...,m$ проведення експериментів і стовпчиків у кількості, що рівна числу вхідних $1,2,...,j,...,n$ та вихідної (або вихідних) параметрів об'єкта дослідження. Далі, до початку проведення експерименту кожен експерт незалежно один від одного, без пояснень з яких міркувань і з використанням якої додаткової інформації та додаткових знань, заносить ті значення вихідної змінної (або змінних), які він вважає будуть досягнуті після проведення всіх запланованих експериментів. Зауважимо, що всі заповнені експертами таблиці залишаються нікому невідомими до кінця проведення експерименту. Дослідник заповнює стовпчик таблиці з прогнозованим значенням вихідної змінної тільки за умов виконаних розрахунків за математичною моделлю, що перевіряється шляхом почергової зміни рівнів значень вхідних параметрів в моделі і результати кожного розрахунку послідовно заносить у стовпчик вихідної змінної (або вихідних) своєї таблиці. Ця таблиця також залишається невідомою до кінця експериментів. Заповнені таким чином таблиці з підписами експертів і дослідника в закритому вигляді передаються керівникові групи експерименту.

III. Наступним етапом відповідно до планів експериментів проводяться досліди безпосередньо на об'єкті дослідження в присутності всіх експертів, дослідника і керівника. Результат експериментів також заноситься в стовпчик вихідної змінної в окрему таблицю. Процедура заповнення таблиці виконується в присутності всіх

учасників експерименту. Ця таблиця з результатами експериментів підписується всіма експертами, дослідником й керівником експериментів. Для підвищення якості перевірки моделі можна запланувати проведення за кожним експериментом декількох повторних серій. Цього разу стовпчик таблиці результатів експериментів буде заповнений усередненим значенням серій кожного окремого експерименту.

IV. Наступний етап методу включає безпосередньо процедуру перевірки адекватності моделі й прийняття рішень по її використанню. В загальному вигляді з використанням експериментально-статистичних методів перевіряється математична модель вигляду:

$$\hat{Y} = a_0 + \sum_{j=1}^{n} a_j x_j + \sum_{j,k=1}^{C_n^2} b_{jk} x_j x_k + \sum_{j=1}^{n} c_j x_j^2 + \lambda, \qquad (3.467)$$

де $x_1, x_2, \ldots, x_j, \ldots, x_n$ – вхідні змінні, $\hat{Y}$ – вихідний параметр об'єкта дослідження, λ – похибка моделі. Але можуть бути й інші структури математичної моделі.

Приймається гіпотеза, що та математична модель найближче описує властивості об'єкту дослідження впливу всіх суттєвих факторів вхідних змінних, яка досягає мінімуму суми квадратів різниці відхилень дійсних вихідних значень об'єкта дослідження (отриманих під час проведення експериментів) й розрахункових даних, які одержані з математичної моделі (3.467) при однакових значеннях вхідних параметрів кожного експерименту або прогнозованих значеннях вихідного параметру різних експертів (або найкращої оцінки експертів). У запропонованому методі експертних оцінок перевірки адекватності математичного моделювання рівняння (3.467) використовується для визначення похибки одержаної математичної моделі (3.467) відносно її показників порівняно з показниками різних експертів, а також з усередненою оцінкою групи (підгрупи) експертів.

V. Процедура перевірки адекватності моделі полягає в наступному. Для наглядності і зручності виконання робіт рекомендується заповнити значення вхідних даних в таблицю 2, яка складається з m рядків і k стовпчиків, в які занесені з таблиці 1 всі прогнозовані значення вихідної змінної, значення експертів Y_{eki} стовпчика k+1, куди заносять значення вихідної змінної дослідника, які отримані після розрахунків з використанням моделі Y_{mi} і стовпчика k+2 дійсних значень вихідного параметру, які отримані в результаті проведення експерименту Y_{di}. Далі за формулами

$$\lambda_{eki} = (Y_{eki} - Y_{di})^2, \qquad (3.468)$$

$$\lambda_{mi} = (Y_{mi} - Y_{di})^2, \qquad (3.469)$$

необхідно провести відповідні розрахунки похибки прогнозу всіх експертів λ_{eki} і похибки математичної моделі λ_{mi} для кожного i-го експерименту відповідно з планами експериментів, відносно дійсних значень вихідної змінної Y_{di}.

Результати розрахунків значень λ_{eki} і λ_{mi} занести в таблицю 3, яка містить m+1 рядків та k+1 стовпчиків: де в (m+1)-й рядок заносять значення середньої похибки кожного експерта і математичної моделі λ_m. Далі за значеннями 1,2,...,k стовпчиків визначають середню похибку кожного експерта λ_s, а за значеннями k-го стовпчика – середню похибку математичної моделі λ_{mi}.

Рекомендується визначати три рівні значень адекватності моделі з оцінками: перший рівень – "відмінно", другий рівень – "добре", третій рівень – "задовільно".

Перевірка умов для третього рівня визначається формулою

$$\overline{\lambda}_m \le \frac{\sum_{s=1}^{k} \overline{\lambda}_s}{k}. \qquad (3.470)$$

Якщо умови (3.470) не виконуються, то математична модель приймається неадекватною, а значить не рекомендується для використання її на об'єкті дослідження у виробничій або економічній діяльності підприємства. Якщо умови (3.470) виконуються, то оцінка адекватності математичної моделі є позитивною на рівні "задовільно".

VI. Перевірку якості моделі з другим рівнем рекомендується продовжити наступними чином: зі значень похибки всіх k експертів рядків m таблиці 3 необхідно визначити з групи експертів в кількості $k_1 = k-h$, які мають значення похибки менше від середнього значення похибки всіх експертів, де h – кількість експертів з похибкою більшою за середню. Перевірка умов для другого рівня визначається за формулою:

$$\overline{\lambda}_m \le \frac{\sum_{s=1}^{k_1} \overline{\lambda}_s}{k_1}. \qquad (3.471)$$

Якщо умови (3.471) виконуються, оцінка адекватності моделі визнається на "добре". В цьому випадку адекватність моделі вважається точнішою завдяки тому, що гірші прогнози експертів вилучені з перевірки. Перевірку математичної моделі за визначенням оцінки адекватності на "відмінно", як найкращої оцінки адекватності моделі, рекомендується здійснювати таким чином. Зі значень $\overline{\lambda}_{ek}$ похибки кожного експерта рядка m таблиці 3 вибирається найменше значення $\overline{\lambda}_{ek_{\min}}$ та перевіряється умова

$$\overline{\lambda}_m \le \overline{\lambda}_{ek_{\min}}. \qquad (3.472)$$

Якщо умови (3.472) виконуються, оцінка адекватності математичної моделі, що перевіряється, визнається "відмінно". Виконання умови (3.472) підвищує вірогідність адекватності моделі, яка перевіряється, і надає більше впевненості в можливості її використання. Оцінки адекватності моделей "добре" та "відмінно" можуть свідчити про певну якість математичної моделі, і тому ці моделі можуть бути використані в промисловій та економічній діяльності відповідних галузей у вирішенні задач прогнозування і управління, а також моделюванні економічних і технологічних процесів. Якщо оцінка математичної моделі визначена тільки на "задовільно" буде доцільним залучитися додатковою рекомендацією групи експертів на можливість її використання в практичній діяльності. Можливо ця рекомендація може бути визначена більшістю голосів експертів або слід продовжити додаткові експерименти.

Запропонований метод відрізняється від методу максимального балу тим, що потребує від експертів прогнозованих значень вихідної змінної, які після проведення експериментів перевіряються відносно дійсних значень, а тому потребує від них більшої відповідальності та мобілізує експерта на повну віддачу своїх наукових знань.

практичного досвіду і максимального інтересу до отримання мінімальної похибки свого прогнозу. Особливо ефективним може бути використання цього методу в створенні комплексного кваліфікаційного тренінгу в моделюванні різних виробничих і економічних систем для проведення практичних учбових занять зі студентами, а також виконання самостійних досліджень з цих питань аспірантами і пошукувачами.

3.14. Комп'ютерні технології ідентифікації виробничих агрегатів

Однією з найбільш важливих задач АСУ ТП є оптимізація технологічних режимів, яка без адекватної математичної моделі практично не може бути вирішена. Тому не менш важливою задачею в системах управління є ідентифікація статичних характеристик технологічних процесів. Розв'язанню цієї задачі присвячена значна кількість теоретичних робіт й досліджень з використанням експериментально-статичних методів, які мають ряд істотних переваг у порівнянні з іншими методами. Однак і вони не позбавлені недоліків.

При використанні методу планування активних експериментів, незважаючи на його постійне вдосконалювання, необхідно урахувати ряд умов. *По-перше*, при їх реалізації на промислових об'єктах необхідно вносити завдання на зміну параметрів за планом, який визначено одним з методів планування, що не передбачено виробничим планом роботи промислових об'єктів. *По-друге*, математична модель описує досліджуваний об'єкт лише за вхідними керованими змінними, а некеровані вхідні параметри (які теж мають суттєвий вплив на вихідні параметри) практично залишаються не врахованими. Це, як відомо, істотно знижує ефективність їхнього використання навіть при невеликій кількості змінних параметрів, бо отримана математична модель не враховує всі фактори впливу на об'єкт дослідження.

Недоліком методів ідентифікації статичних характеристик технологічних процесів, які базуються на пасивному експерименті, є відсутність помилки експерименту. Внаслідок чого серед занесених значень параметрів, є значення параметрів, які характеризують не тільки статичний режим, а також перехідний режим. Крім того, якщо промисловий агрегат значний час знаходився у визначеному статичному режимі або в різний час кілька разів знаходився в статичних режимах, то в матриці спостереження поряд з рядками, що характеризують динамічний режим, є декілька рядків зі значенням параметрів, що практично не відрізняються один від одного. Тому отримана матриця спостереження з заданою кількістю рядків N^{zd} може містити, з одного боку, рядки з неякісною інформацією (рядки, що характеризують перехідний режим) або значну похибку i-го виміру. З іншого боку, рядки, що характеризують однакові статичні режими. Це означає, що фактично матриця спостереження містить не N^{zd} рядків, а набагато менше, що значно знижує адекватність отриманої математичної моделі. Тому потрібно задовольнити вимоги, щоб за тривалий період роботи промислових агрегатів у матриці спостереження було занесено задану кількість статичних режимів (N^{zd}), які визначені з достовірною ймовірністю і необхідною точністю одержання математичного опису цього процесу.

Тут запропонована комп'ютерна технологія (метод) ідентифікації неперервних технологічних процесів за допомогою керуючих комп'ютерів (КК), яка дозволяє в багатьох випадках розв'язати задачу одержання математичної моделі промислових об'єктів. Цей метод особливо ефективний, коли застосування відомих методів

ідентифікацій на діючих промислових об'єктах не є реально можливим. Але обов'язковими умовами, які забезпечують одержання моделі, є присутність на об'єкті діючої АСУ з функціонуванням її в режимі збирання і контролю значень всіх вхідних і вихідних параметрів, які необхідні для побудови математичних моделей об'єкта дослідження. Включення в ідентифікацію керуючих КК насамперед дозволяє оперативно з заданою точністю визначити дійсні значення параметрів і тим самим забезпечити необхідну якість достовірності вхідної й вихідної інформації.

Основною перевагою запропонованого методу ідентифікації є наявність ряду процедур, що дозволяють без проведення на промисловому об'єкті планових експериментів одержати якісну матрицю спостереження, рядки якої забезпечують неперевищення заданого граничного значення похибки вимірювань кожного досліду. Завдяки цьому у матрицю заносяться тільки ті значення вхідних і вихідних параметрів, що дійсно відображають істотні властивості статичних режимів досліджуваного промислового об'єкта. Введення в комп'ютерний метод ідентифікації цих процедур дозволяє цілком виключити ті значення, які виміряні зі значною похибкою, а також значень параметрів, що характеризують динамічний режим. Крім того, у матрицю спостереження заносяться не миттєві дійсні значення параметрів, що належать до статичних режимів, а усереднені дійсні значення параметрів за визначений час перебування досліджуваного процесу у кожному статичному режимі. Метод включає обов'язкові процедури, що забезпечують одержання достатньої кількості необхідної інформації для визначення адекватної математичної моделі у вигляді полінома заданого порядку (Z), який описує досліджуваний промисловий об'єкт із заданим числом контрольованих й керованих змінних параметрів при заданому допустимому інтервалі їх варіювання. Сюди включено також процедури перевірки наявності в матриці по кожному вхідному параметру значень, які близькі до їх мінімального і максимального значень інтервалу варіювання. Реалізація запропонованого методу ідентифікації здійснюється в режимі нормальної експлуатації промислового об'єкта і функціонування АСУ ТП розв'язку задач контролю й інших задач.

Для того, щоб можливо було використати комп'ютерні технології ідентифікації технологічних об'єктів, досліднику-аналітику потрібно визначити завдання і умови отримання за допомогою комп'ютера достовірної інформації, яка в повній мірі відображає властивості об'єкта дослідження.

Цей етап роботи дослідника по суті не відрізняється від проведення активного експерименту на об'єкті для отримання математичної моделі з використанням методів планування багатофакторних експериментів. Досліднику необхідно:

1) Використовуючи відомі знання про об'єкт дослідження, визначити всі керовані і некеровані вхідні параметри суттєвого впливу й вихідні параметри, за якими планується отримання математичних моделей.

2) Визначити, яким поліномом першого чи другого ступеню буде ідентифіковано об'єкт.

3) Встановити умови перевірки присутності ідентичних рядків матриці.

4) Визначити похибку кожного досліду, мінімальну кількість дослідів планованих досліджень $N^{\alpha l}$ та кількість мінімальних різних значень по кожному параметру в кожному стовпчику вхідних параметрів.

5) Встановити для кожного вхідного параметру ті значення, які будуть вважатися задовільними відносно мінімальних і максимальних значень кожного вхідного параметра.

Матриця спостереження, у яку заносяться значення параметрів зберігається на носіях інформації комп'ютера. Ця матриця містить кількість рядків N^{zd}, як мінімум на одиницю більше коефіцієнтів математичної моделі, і кількість стовпчиків $n + m$, де n – кількість вхідних параметрів, m – кількість вихідних параметрів досліджуваного об'єкта. Реалізація комп'ютерного методу ідентифікації здійснюється наступним чином.

I. В умовно заданому першому циклі роботи комп'ютера спостережені значення технологічних параметрів, після визначення дійсних значень вхідних параметрів $x_{11}x_{21},...,x_{i1}...,x_{n1}$ і вихідних параметрів $y_{11}y_{21},...,y_{k1},...,y_{m1}$ досліджуваного технологічного об'єкта, заносяться у нульовий рядок матриці спостереження для визначення належності базових і нових вимірів параметрів до одного і того ж статичного режиму. Вони визначаються як базові значення $x^{\delta}_{11}x^{\delta}_{21},...,x^{\delta}_{i1},...,x^{\delta}_{n1}$, $y^{\delta}_{11}y^{\delta}_{21},...,y^{\delta}_{k1},...,y^{\delta}_{m1}$, і заносяться в перший рядок матриці спостереження як усереднені значення, що визначають перший статичний режим досліджуваного промислового об'єкта.

II. Починаючи з другого циклу роботи комп'ютера, після визначення дійсних значень параметрів $x_{12}x_{22},...,x_{i2},...,x_{n2}$ та $y_{12}y_{22},...,y_{k2},...,y_{m2}$ необхідно з заданою точністю x_i^{zd} і y_k^{zd} визначити належність біжучих значень параметрів і значень параметрів, отриманих в умовно першому попередньому циклі, до одного і того ж статичного режиму. Перевіряються умови:

$$\left|x_j^{\delta} - x_{j1}\right| \le \Delta x_j^{zd}, \tag{3.473}$$

$$\left|y_k^{\delta} - y_{k1}\right| \le \Delta y_k^{zd}, \tag{3.474}$$

де x_i^{zd} і y_k^{zd} – задана дослідником похибка експерименту, виходячи з умов планування активного i-го експерименту. При цьому кожен дослід необхідно провести в заданій кількості (h^{zd}) за визначеними усередненими значеннями кожного статичного режиму.

Якщо умови (3.473) або (3.474) не виконуються, то параметри біжучого і попереднього базового циклів роботи комп'ютера не належать до одного і того ж статичного режиму, і біжучі значення параметрів в цьому циклі визначаються як значення, які одержані в умовно першому циклі роботи комп'ютера і перезаписуються, як нові базові значення у нульовий рядок матриці, і в перший рядок матриці спостереження як усереднені значення, що визначають перший статичний режим об'єкта дослідження. Якщо умови (3.473) і (3.474) виконуються, то об'єкт дослідження за цей період часу знаходиться в статичному режимі, значення параметрів поточного циклу роботи комп'ютера усереднюються з попередніми значеннями параметрів першого рядка матриці спостереження і, як нові усереднені значення переписуються в перший рядок матриці. Після чого перевіряються умови

$$h_{\xi} \ge h^{zd}, \tag{3.475}$$

де h_{ξ} – біжуча кількість циклів, при якій виконувалися умови (3.473) і (3.474). Ці умови (3.475) визначають достатню кількість часу (або кількість вимірів h^{zd}), коли об'єкт дослідження знаходиться в одному і тому ж статичному режимі.

Якщо умова (3.475) не виконується, то значення параметрів $i+1$ наступного циклу роботи комп'ютера перевіряються умовами (3.473) і (3.474). Якщо умови (3.473), (3.474) не виконуються, то тоді $i+1$ цикл приймається як умовно перший цикл роботи

комп'ютера і виконуються описані вище відповідні дії зі значеннями параметрів. Якщо умови (3.473) і (3.474) виконуються, то значення параметрів j+1 циклу усереднюються з усередненими значеннями першого рядку матриці за h_{ε} циклів і записуються в рядок першого статичного режиму функціонування об'єкта дослідження. Після чого знову перевіряються умови (3.475). Якщо умови (3.475) виконуються, то об'єкт необхідний час з заданою похибкою дослідження знаходився в статичному режимі і перший рядок усереднених значень вхідних x_{1l} і вихідних y_{1k} параметрів об'єкта дослідження можна використовувати для визначення математичної моделі. Виконання умов (3.473)-(3.474) свідчить, що всі процедури отримання значень першого рядка матриці спостереження виконані. Виконання умов (3.473)-(3.475) є обов'язковими для визначення значень всіх статичних режимів.

III. Для визначення базового рядка другого статичного режиму функціонування об'єкта, необхідно, щоб виконувалися умови

$$\left|\bar{x}_{1j} - x_{i+1,j}\right| \geq \Delta x_j'^{zd}, \tag{3.476}$$

де $x_{i+1,j}$ – значення вимірювань технологічних параметрів для визначення другого, умови (3.476) визначають, що об'єкт можливо знаходиться в іншому статичному режимі, який відрізняється від першого, і значення другого рядка матриці спостереження буде відрізняється від попереднього хоча б за одним з вхідних параметрів на величину $\Delta x_j'^{zd}$. Тому, достатньою умовою виконання умов (3.476) є те, що ці умови повинні виконуватися хоча б за одним з n вхідних параметрів.

Якщо умови (3.476) не виконуються, то продовжується робота комп'ютера по виміру поточних значень параметрів і визначенню базових значень параметрів другого статичного режиму до того часу, коли умови (3.476) будуть виконані. Якщо умови (3.476) виконуються, значення параметрів цього циклу заносяться в нульовий рядок, як базові значення другого статичного режиму, і в другий рядок матриці, як усереднені значення другого статичного режиму.

IV. Починаючи з наступного циклу роботи комп'ютера по визначенню значень другого статичного режиму, перевіряються умови (3.473), (3.474) і (3.475) відповідно до алгоритму отримання усереднених значень першого статичного режиму. Отримання усереднених значень третього, а значить N-го статичного режиму виконуються за процедурами аналогічними до тих, які виконувались для отримання другого статичного режиму. Таким чином, для отримання значень параметрів в кількості N рядків матриці спостереження, процедури, починаючи з визначення другого рядка до N-го рядка ідентичні. Але набір даних матриці спостереження може бути не закінчений, так як ці умови є необхідні, але не достатні.

V. Далі треба виконати перевірку:

$$\left|\bar{x}_{ij} - x_{i\min}\right| \leq \Delta x_i'^{zd}, \tag{3.477}$$

$$\left|\bar{x}_{ij} - x_{i\max}\right| \leq \Delta x_i'^{zd}. \tag{3.478}$$

де $\bar{x}_{ij}$ – середні значення i-го параметра j-го рядка: x_{imin} і x_{imax} – задані допустимі значення вхідних параметри при яких функціонує об'єкт дослідження. Ці умови забезпечують отримання адекватної математичної моделі на заданому інтервалі

функціонуванні об'єкта дослідження.

VI. Після цього для визначення кількості різних усереднених значень за кожним вхідним параметром, які знаходяться в кожному стовпчику матриці, необхідно *перевірити виконання умов (3.478)*:

$$\left|\overline{x}_{ij} - x_{i(1,2,\ldots,j-1,j+1,\ldots n)}\right| \geq \Delta x_i^{nzd}, \tag{3.479}$$

які визначать кількість разів ξ_i значення параметрів в кожному стовпчику, які відрізняються один від одного на Δx_i^{nzd}, потім перевіряються умови (3.480):

$$\left|\xi_i - Z\right| \geq 1, \tag{3.480}$$

де Z – степінь поліному математичної моделі об'єкту дослідження.

Умови (3.480) дозволяють стверджувати: якщо поліном першого ступеня, то таких значень повинно бути не менше двох; якщо поліном другого ступеня таких значень – не менше трьох і так далі. Якщо умови (3.479) і (3.480) не виконуються, то комп'ютер продовжує накопичувати значення нових статичних режимів за визначеними вище процедурами. Якщо виконуються умови (3.477), (3.478), (3.479) і (3.480), то матриця спостереження заповнена усередненими значеннями всіх вхідних і вихідних параметрів, які відображають істотні риси об'єкта дослідження. Щоб отримати необхідну матрицю спостереження статичних режимів користувачу (досліднику) необхідно через діалог вікна операторської станції (відеотерміналу) задати наступні необхідні дані: кількість вхідних n і вихідних m параметрів об'єкта дослідження, за якими збір інформації про роботу агрегату здійснює комп'ютер; діапазон варіювання по кожному вхідному параметру x_{imin} і x_{imax}; допустимі значення відхилення поточних значень параметрів Δx_i^{zd}, за якими визначають, що об'єкт дослідження знаходиться в статичному режимі; мінімальні значення по кожному вхідному параметру, які відрізняють рядки матриці спостереження, що характеризують різні статичні режими Δx_i^{zd}; максимально допустимі значення Δx_i^{nzd} і Δx_i^{mzd}, відхилення усереднених значень вхідних параметрів, які спостерігаються при визначенні статичних режимів, до допустимих границь варіювання параметрів x_{imin} і x_{imax}; h^{zd} – необхідна мінімальна кількість поточних значень параметрів для визначення усереднених значень, які характеризують статичні режими; N – кількість коефіцієнтів математичної моделі, які потрібно визначити; Z – порядок полінома математичної моделі. Перетворення усереднених значень параметрів статичних режимів об'єкта дослідження, які занесені в матрицю спостереження в коефіцієнти математичної моделі виконувались із використанням пакетів прикладних програм статистичної опрацювання інформації [STATISTIKA, MATHCAD]. Теоретична перевірка математичної моделі виконувалась методом Фішера (3.473). Практична перевірка адекватності математичних моделей виконувалась з використанням експертного метода.

При проведенні випробувань комп'ютерних технологій ідентифікації різних технологічних агрегатів з кількістю вхідних параметрів від 4 до 7 із отриманням математичної моделі у вигляді полінома другого степеня КК на протязі двох-трьох тижнів, отримувала в необхідній кількості рядків матриці спостереження від 25 до 55.

Запропонований метод був випробуваний на технологічних агрегатах: аміачне виробництво, реактор виробництва полістиролу; доменна піч та інші агрегати. З

використанням комп'ютерів позитивних результатів випробувань комп'ютерних технологій в ідентифікації виробничих умов досягнуто практично на всіх агрегатах. Тому реалізація комп'ютерних технологій ідентифікації на сучасних комп'ютерах значно збільшить ефективність використання цієї технології.

3.15. Методичні прийоми дослідження впливу факторів на виробничі та економічні процеси

Вирішення питань моделювання та прогнозування будь-яких явищ природи, технологічних і економічних процесів розвитку підприємства та інших процесів, можливе тільки у випадку створення математичних моделей цих процесів. При цьому вони повинні враховувати одночасний вплив всіх важливих факторів. Ці математичні моделі з достатньою точністю повинні відображати властивості реальних об'єктів дослідження. У зв'язку з тим, що задача одержання адекватних математичних моделей відповідним реально діючим процесам є дуже складною, то дуже важливою стає проблема розв'язку задач ідентифікації цих об'єктів дослідження.

Практика підтвердила, що найбільш ефективними методами знаходження математичних моделей на цей час є експериментально-статистичні методи, які в загальному, стосовно інших методів, мають значний перелік позитивних рис. А серед експериментально-статистичних методів найбільш ефективними з точки зору створення більш достовірних математичних моделей є *методи активних експериментів*. Однак при проведенні активних експериментів, коли водночас необхідно виконувати значну кількість дослідів, витрати на виконання цих робіт у декілька разів будуть перевищувати витрати при одержанні математичних моделей з використанням пасивних методів кореляційного або регресійного аналізу. Але також відомо, що ефективність математичних моделей практично дуже низька, бо математичні передумови, які закладені в основу використання цих методів, у реальних умовах виробництва, економіки і природніх явищ дуже рідко виконуються. Слід зауважити, що необхідність вдосконалення методів ідентифікації технологічних, економічних та інших процесів постійно продовжуються. Крім того, можуть бути дуже корисними такі методичні прийоми наукових досліджень, при яких із відомих методів будуються науково-дослідні процедури, які забезпечать розв'язок задач ідентифікації з достатньою точністю. Звичайно, якщо іншими методами на даний час скористатися неможливо. Слід також зазначити наступне, якщо багатопараметричні процеси описуються набором однопараметричмих моделей різних типів з високим рівнем адекватності, то використати набір таких моделей при розв'язуванні задач прогнозування або пошуку оптимальних розв'язків практично дуже складно, навіть при застосуванні сучасних комп'ютерів. При цьому пошук глобального оптимуму за одновимірними моделями багатофакторних процесів взагалі не можна здійснити при наявності в об'єкті навіть незначного суттєвого взаємозв'язку між вхідними параметрами процесу. Тому завжди, як правило, одержують багатовимірну математичну модель лінійної або нелінійної форми залежно від властивостей об'єктів. Але одержання таких математичних моделей, нажаль, все ще лишається великою проблемою. Тут пропонується методологічний прийом дослідження одержання математичних моделей у формі поліному першого чи другого порядку з n вхідними змінними параметрами тоді, коли не можуть бути використані інші методи. При цьому реальний процес дослідження повинен задовольняти такі умови і обмеження:

- наявність обсягу статистичних даних про процес, не дає змоги отримати адекватну модель з заданою похибкою;
- на час випробувань недостатньо коштів або інших можливостей провести роботи з використанням методів активних експериментів за методикою повного, факторного експерименту, або композиційного чи іншого плану;
- існує інформація про раніше проведені однофакторні активні експерименти за n суттєвими змінними параметрами процесу, причому ці результати подані у формі графіків, таблиць або одномірних математичних виразів, джерелами інформації можуть виступати: довідкова література, монографії, наукові статті, звіти наукових досліджень;
- на час прийняття рішення необхідно виконати роботи з моделювання, прогнозування або пошуку глобального оптимума процесу, як з найменшими затратами часу. Зрозуміло, що при наявності у майбутньому коштів на проведення нових досліджень з використанням багатофакторного активного експерименту можна буде одержати більш адекватну математичну модель.

Алгоритм запропонованого методологічного прийому дослідження складається з наступних кроків:

1. Вибір об'єкта з обов'язковим виділенням суттєвих факторів, критерія якості управління. При цьому методи опрацювання інформації об'єкту дослідження та інформаційне забезпечення розрахунків для отримання багатовимірних математичних моделей повинні відповідати процедурам методів планування багатофакторного аналізу.
2. Відповідно до обраного методу проведення дослідження створюється матриця проведеного експерименту. За планами експериментів, які записані в цій матриці проводяться уявні "експерименти". Виконання експериментів здійснюється не на об'єкті дослідження, а на отриманих раніше адекватних одновимірних моделях цього об'єкту, а саме, на одномірних графіках або таблицях чи математичних моделях. Процедура проведення експериментів виконується наступним чином: за значеннями рівнів вхідних параметрів першого рядка матриці на одновимірних моделях визначаються значення вихідних параметрів, і всі ці значення заносяться в незаповнену частину першого рядка матриці планування експериментів у відповідні стовпчики значень вихідних параметрів. Потім за значенням рівнів вхідних параметрів другого, третього і останнього рядку матриці виконуються послідовно всі уявні "експерименти", які заплановані відповідно до обраного дослідником методом дослідження. Таким чином, не проводячи безпосередньо на об'єкті дослідження нових активних експериментів, отримується інформація, яка буде задовільняти умови проведення активних експериментів.
3. В залежності від виду моделі та методів, які використовувалися, обчислюються коефіцієнти рівнянь, які дозволяють з певною точністю описати технологічний або економічний процес, що вивчається. Оскільки похибки експерименту визначалися і перевірялися дослідниками, які отримували одновимірні моделі, і адекватність цих математичних моделей їх задовільняла. Тоді можна вважати, що і адекватність отриманих багатовимірних математичних моделей задовільна, хоча і частково зменшена.

Зауваження 3.9. Перевірки ефективності запропонованого методу здійснювались на джерелах інформації дуже відомих і поширених технологічних процесів. Виявилось, що одержані результати завжди були позитивні.

Розглянемо методологію для одержання багатовимірних математичних моделей синтезу аміаку на технологічному агрегаті. Інформація використана з графіків (рис. 3.17а) залежності значення відсотків виходу аміаку в газі (вихід аміаку) від температури x_2, об'ємної швидкості x_3 при значенні тиску x_1 = 300 атм, та інформація графіка (рис. 3.17б) залежності значення відсотків виходу аміаку від тих же рівнів значень параметрів тиску x_1, і об'ємної швидкості x_3 при другому рівні значення тиску x_1 = 500 атм.

а) залежність від температури: 1 – при 400°С; 2 – при 450°С; 3 – при 500°С; 4 – при 550°С;

б) залежність від тиску: 1 – при 300 атм; 2 – при 450 атм; 3 – при 600 атм.

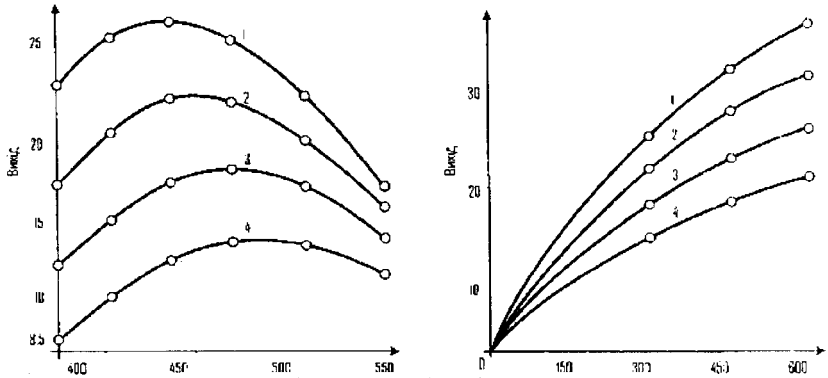


Рис. 3.17. Залежність виходу аміаку від температури та тиску.

Математична модель визначена у вигляді рівняння:

$$\hat{Y} = a_0 + \sum_{j=1}^{n} a_j x_j + \sum_{j,k=1}^{C_n^2} b_{jk} x_j x_k + C x_1 x_2 x_3, \qquad (3.481)$$

коефіцієнти якого можна одержати за допомогою методу повного факторного експерименту (ПФЕ).

Таблиця 3.44 містить матрицю планування ПФЕ для трьох факторів x_1, x_2, x_3 при двох рівнях варіювання - "+1" та "-1", заповнена на підставі використання графіків.

Рівню "+1" відповідають: x_1 = 500агм, x_2 =525°С, x_3 = 12000 Γ^{-1}

Рівню "-1" відповідають: x_1 = 300атм, x_2 = 400°С, x_3 = 15000 Γ^{-1}

За відповідними формулами, які використовуються в методі ПФЕ були визначені коефіцієнти рівняння (3.471), яке описує процес синтезу аміаку

$$Y = 18.11 + 3.21x_1 + 1.16x_2 - 5.76x_3 + 0.43x_1x_2 + 1.37x_1x_3 - 0.83x_2x_3 + 0.04x_1x_2x_3. \qquad (3.482)$$

Таблиця 3.44

Матриця планування ПФЕ для трьох факторів x_1, x_2, x_3 при двох рівнях варіювання - "+1" та "-1"

№	x_0	x_1	x_2	x_3	$x_1 x_2$	$x_2 x_3$	$x_1 x_3$	$x_1 x_2 x_3$	Y_{exp}	Y_{rozr}
1	2	3	4	5	6	7	8	9	10	11
1	+	-	-	-	+	+	+	-	21.8	21.791
2	+	-	-	+	+	-	-	+	9.3	9.301
3	+	-	+	-	-	+	-	+	20.6	20.601
4	+	-	+	+	-	-	+	-	13.4	13.411
5	+	+	-	-	-	-	+	+	28.85	28.837
6	+	+	-	+	-	+	-	-	12.85	12.855
7	+	+	+	-	+	-	-	-	29.2	29.095
8	+	+	+	+	+	+	+	+	18.85	18.857

Обчислені значення відсотків виходу аміаку Y_{rozr}, які одержано з рівняння (3.474), знаходяться в стовпчику 11 табл. 3.44. Як видно з таблиці 3.44, значення відсотків виходу аміаку, які одержано розрахунковим шляхом, з незначною похибкою співпадають зі значеннями відсотка виходу аміаку Y_{exp}, які взяли з графіків рис. 3.17. Це вказує на задовільну адекватність багатовимірної моделі трьохвимірного об'єкту дослідження. Для побудови математичної моделі процесу синтезу метанолу використовувалась таблиця залежності виходу метанолу у (в об'ємних %) при різних значеннях температури x_1 та об'ємній швидкості x_2 тисяч м^3 за годину (Γ^{-1}), при тиску 250 атм та співвідношення CO:H^2=1:2. Ця залежність наведена в табл. 3.45.

Таблиця 3.45

Залежність виходу металону при різних значеннях температури та швидкості

Об'ємна швидкість Γ^{-1}	Температура°, С				
	340	360	370	380	400
	Вихід метанолу, %				
40000	4,2	5,4	5,7	6,0	5,1
60000	3,4	4,7	5,1	5,4	4,7
80000	3,0	4,3	4,6	5,0	4,4
100000	2,7	4,0	4,4	4,8	4,2
120000	2,5	3,8	4,2	4,7	4,0

Математична модель одержання метанолу визначалась у вигляді рівняння:

$$Y = a_0 + a_1 x_1 + a_2 x_2 + b x_1 x_2 + c_1 x_1^2 + c_2 x_2^2, \qquad (3.483)$$

коефіцієнти якого можна одержати за допомогою методу композиційного планування. В табл. 3.46 дана матриця композиційного планування для двох факторів x_1 і x_2 з плечем, рівним 1 при трьох рівнях варіювання: "+1", "0", "-1".

Рівню "+1" відповідають $x_1 = 400°C$, $x_2 = 120000\ \Gamma^{-1}$

Рівню "0" відповідають $x_1 = 370°C$, $x_2 = 80000\ \Gamma^{-1}$

Рівню "-1" відповідають $x_1 = 340°C$, $x_2 = 40000\ \Gamma^{-1}$

За відповідними формулам були визначені коефіцієнти рівня, які описують процес одержання метанолу. Розрахункові значення відсотка виходу метанолу, які одержані за рівнянням, наведені у стовпчику 8 табл. 3.46. Як бачимо з табл. 3.46, значення відсотка виходу метанолу, одержані розрахунковим способом, співпадають зі значеннями відсотка виходу метанолу, що взяті з таблиці.

Таблиця 3.46

Матриця композиційного планування для двох факторів x_1 і x_2 з плечем, рівним 1 при трьох рівнях варіювання: "+1", "0", "-1"

№	x_1	x_2	$x_1 x_2$	x_1^2	x_2^2	Y_{exp}	Y_{rozr}
1	2	3	4	5	6	7	8
1	-1	-1	+1	+1	+1	4,2	4,1547
2	+1	-1	-1	+1	+1	5,1	5,1147
3	-1	+1	-1	+1	+1	2,5	2,4547
4	+1	+1	+1	+1	+1	4	4,0147
5	+1	0	0	+1	0	4	4,3527
6	-1	0	0	+1	0	3	3,0918
7	0	+1	0	0	+1	4,25	4,2506
8	0	-1	0	0	+1	5,7	5,6506
9	0	0	0	0	0	4,65	4,7387

З одержаних результатів використання цього методу для двох технологічних процесів бачимо, що цей метод може бути широко використаний для вирішення задач ідентифікації складних технологічних процесів. Використання цих даних дозволяє вирішувати задачу ідентифікації цих промислових об'єктів і дає можливість в значній мірі скоротити час та витрати на проведення нових експериментів.

Зауваження 3.10. У випадку, якщо даних про об'єкт дослідження з таблиць або графіків недостатньо для заповнення матриці планування експериментів відповідно до обраного методу, необхідно піти на спрощення виду моделі і, відповідно, вибрати інший метод активного експерименту. Наприклад, модель нелінійного виду замінити на лінійну, а метод композиційного планування замінити методом повнофакторного експерименту. Або доповнити результатами, які одержані безпосередньо на промисловому об'єкті при проведенні додаткових експериментів у кількості, яка забезпечує можливість мати заповнену матрицю в необхідному обсязі планування експерименту. Але, і в такому випадку, витрати на експеримент та час на розв'язування задач ідентифікації об'єктів дослідження будуть також скорочені.

Зауваження 3.11. Зараз є велика кількість даних про однофакторні активні експерименти для різних промислових об'єктів, а також даних про економічну і фінансову діяльність різних підприємств непромислової сфери. І тому є впевненість, що із залученням запропонованого методологічного прийому дослідження можливо суттєво збільшити ефективність використання одержаної цінної інформації, на здобуття якої вже витрачені великі кошти.

Висновок 3.4. Одержані результати використання цього методу для технологічних процесів синтезу аміаку та метанолу підтверджують, що цей методологічний прийом може широко використовуватися для розв'язування задач ідентифікації складних багатофакторних технологічних й економічних процесів за умов, які викладені вище. Запропонована методологія дозволяє використовувати інформацію (дані) раніше проведених експериментів для розв'язку задач ідентифікації цих об'єктів. та надає можливість в значній мірі скоротити час та витрати на проведення нових експериментів. І не зважаючи на те, що адекватність отриманих багатовимірних математичних моделей буде нижчою ніж адекватність окремо взятих одномірних моделей об'єкта дослідження, такі багатовимірні моделі дозволяють в певній мірі вирішувати задачі прогнозування і оптимального управління, і тим самим забезпечувати оператора промислового агрегату або фахівців підприємства непромислової сфери, отримувати підтримку у прийнятті ними оптимальних рішень.

3.16. Контрольні питання

1. Моделювання на комп'ютері рівномірно розподіленої величини.
2. Суть аналітичного методу обчислення однократного інтегралу.
3. Поясніть *метод площ обчислення визначеного інтегралу*.
4. Обчислення оберненої матриці методом статистичного моделювання.
5. Алгоритм методу статистичного моделювання розв'язку СЛАР.
6. Для чого використовуються ланцюги Маркова?
7. Для яких систем використовується метод статистичного моделювання?
8. Що таке квадратична (параболічна) залежність?
9. Визначення параметрів емпіричної формули.
10. Сформулюйте суть *методу вибраних точок*?
11. Суть *методу середніх*.
12. Загальний метод визначення параметрів емпіричної формули.
13. Особливості *методу кореляційного аналізу*.
14. Використання прийомів регресійного аналізу для побудови математичної моделі складного об'єкту.
15. Суть алгебри планування експерименту.
16. Особливості методу пошуку оптимальних умов для складних процесів.
17. Математичний опис промислового об'єкту методом пасивного експерименту.
18. Суть методу експертних оцінок перевірки адекватності математичних моделей.
19. Особливості комп'ютерних технологій ідентифікації виробничих агрегатів.
20. Реалізація комп'ютерного методу ідентифікації.
21. Які умови і обмеження повинен задовольняти реальний процес дослідження?

3.17. Задачі для самостійної роботи

Задача 3.17.1. Побудувати алгоритм методу статистичного моделювання розв'язку інтегрального рівняння

$$\varphi(x) = \lambda \int_{-\infty}^{+\infty} e^{-|x-y|}\varphi(y)dy + \frac{1}{\sqrt{2\pi}} e^{\frac{x^2}{2}}.$$

Задача 3.17.2. Написати алгоритм методу статистичного моделювання розв'язку інтегрального рівняння

$$\varphi(x) = q \int_{0}^{\pi} e^{-|x-y|}\varphi(y)dy + e^{-x}$$

Задача 3.17.3. Побудувати алгоритм методу статистичного моделювання для розв'язку двовимірного різницевого рівняння Пуассона, виходячи з загальної методики розв'язку СЛАР.

Задача 3.17.4. Записати алгоритм розв'язку задачі обчислення

$$\int_{0}^{1} (1-x)\varphi(x)dx,$$

де $\varphi(x)$ є розв'язком інтегрального рівняння

$$\varphi(x)=\frac{1}{2}\int_0^1 e^{-(x-y)^2\cdot\varphi^2(y)}\varphi^2(y)dy+1$$

Вказівка. Застосувати метод послідовних наближень:

$$\varphi_n(x)=\frac{1}{2}\int_0^1 \exp\left\{-(x-y)^2/\varphi^2_{n-1}(y)\right\}\varphi_n(y)\,dy+1.$$

Задача 3.17.5. Методом статистичного моделювання розв'язати задачу Діріхле у кубі $0\le x,y,z\le 1$ для рівняння $\Delta u=-3\pi^2\sin\pi x\sin\pi y\sin\pi z$ за умови на грані z=1: $u(x,y,z)\big|_{z=1}=\sin\pi x\sin\pi y$; $u(x,y,z)$ на інших гранях.

3.18. Завдання до лабораторної роботи

Задача 3.18.1. Обчислити наближено за допомогою стохастичних методів Ейлера, Рунге-Кутта та Мільштейна розв'язок стохастичного диференціального рівняння вигляду $dx(t,\omega)=a(x(t,\omega);t)dt+\sigma(x(t,\omega);t)dw(t)$, з початковою умовою $x(t_0)=x_0$. На відрізку [0,T], T=30, h=0.5, x_0=0. Дані для обчислень взяти з табл. 3.47, де n_1,n_2,n_3,n_4 – чотири останні цифри номера залікової книжки студента.

Таблиця 3.47

Вхідні дані для задачі 3.18.1

№	$a(t,x)$	$\sigma(t,x)$	θ	ε
1	$-x^3+\theta x$	x	$0, n_1n_2,n_3,n_4$	-
2	$e^{-x}+(\theta\sin(2x))/2$	$\sin x$	$0, n_1n_2,n_3,n_4$	-
3	$-\varepsilon(1-2\theta\varepsilon)x$	θ	$0, n_1n_2,n_3,n_4$	$0, n_1n_2$
4	$\theta x+\varepsilon$	x^2	$0, n_1n_2,n_3,n_4$	n_1, n_2n_3
5	$1-\sin(\varepsilon t+x)$	θt	$0, n_1n_2,n_3,n_4$	n_1, n_4n_3
6	$1+\cos(\varepsilon\theta t+x)$	θ^3	$0, n_1n_2,n_3,n_4$	n_4, n_4n_3
7	$\sin t+x\varepsilon$	θt^2	$0, n_1n_2,n_3,n_4$	n_4, n_1n_2

Задача 3.18.2. Методом статистичного моделювання обчислити інтеграл за одиничним n-вимірним кубом. Дані для обчислення брати з табл. 3.48.

Таблиця 3.48

Індивідуальні завдання для обчислення задачі 3.18.2

N	n	$f(x)$	N	n	$f(x)$	N	n	$f(x)$
1	5	$\sin\left(\sum_{i=1}^{n}x_i^2\right)$	11	6	$\frac{\sin(x_1+x_2+x_3)}{1+x_4^2+x_5^2+x_6^2}$	21	7	$1/\left(1+arctg(\sum_{i=1}^{n}x_i^2)\right)$
2	6	$\cos\left(\sum_{i=1}^{n}x_i^3\right)$	12	5	$\exp\left\{-\sum_{i=1}^{n}x_i^2\right\}$	22	8	$\sin\left(\sum_{i=1}^{n}x_i\right)/\left(1+\sum_{i=1}^{n}\cos x_i^2\right)$
3	6	$1/\left(1+\sum_{i=1}^{n}x_i\right)$	13	7	$sh\left(\prod_{i=1}^{n}x_i\right)$	23	7	$\prod_{i=1}^{n}\sin x_i^2$
4	5	$\frac{\cos\left(\sum_{i=1}^{4}x_i\right)}{\sin\left(\sum_{i=1}^{6}x_i\right)}$	14	6	$\frac{\cos\left(\sum_{i=1}^{4}x_i\right)}{\cos\left(\sum_{i=1}^{n}x_i\right)}$	24	5	$\frac{\sin\left(\sum_{i=1}^{n}x_i\right)}{1+\prod_{i=1}^{n}x_i}$

Продовження таблиці 3.48

5	4	$\sum_{i=1}^{n}\cos x_i^2$	15	7	$\prod_{i=1}^{n}\left(1+\cos x_i^2\right)$	25	4	$\sin\left(\prod_{i=1}^{n} x_i\right)$
6	7	$1/\sqrt{1+2\sum_{i=1}^{n} x_i^2}$	16	5	$arctg\left(\sum_{i=1}^{n} x_i^2\right)$	26	5	$\cos\left(\prod_{i=1}^{n} x_i\right)$
7	6	$1/\left(1+\sum_{i=1}^{n} x_i^2\right)$	17	7	$arcctg\left(\sum_{i=1}^{n} x_i^2\right)$	27	6	$arcctg\left(\sum_{i=1}^{n} x_i\right)$
8	8	$\ln^{-1}\left(2+\sum_{i=1}^{n} x_1\right)$	18	6	$ch\left(\prod_{i=1}^{n} x_i\right)$	28	8	$arctg\left(\sum_{i=1}^{n} x_i\right)$
9	4	$\ln\left(2+\sum_{i=1}^{n} x_1\right)$	19	5	$\frac{\sin x_i^2 \cos x_i^2}{1+\sum_{i=1}^{n}\sin^2 x_i}$	29	7	$\sqrt{1+\sum_{i=1}^{n}\sin^2\left(x_i^2\right)}$
10	5	$\frac{x_1+x_2+x_3}{5+x_3+x_4+x_5^2}$	20	8	$arctg\left(\prod_{i=1}^{n} x_i\right)$	30	4	$\frac{x_1+x_2^2+x_3^2}{2+x_1+x_2+x_3+x_4}$

Задача 3.18.3. За вибірковими даними визначити вигляд емпіричної формули і визначити її параметри:

Таблиця 3.49

Індивідуальні завдання для обчислення задачі 3.18.3

1	x	1,01	1,51	2,02	2,51	3,01	3,49	3,98	4,48	4,99	5,49
	y	12,03	8,84	6,99	6,03	5,55	5,12	4,63	4,41	4,01	3,96
2	x	1,02	1,59	2,12	2,61	3,05	3,56	4,00	4,50	5,03	5,58
	y	3,02	2,94	2,71	3,32	2,69	2,99	3,22	3,42	4,00	3,49
3	x	1,10	1,55	2,09	2,52	3,07	3,57	4,05	4,56	5,06	5,53
	y	1,84	2,22	3,54	3,50	5,30	5,42	7,12	8,27	8,71	10,09
4	x	1,08	1,53	2,05	2,58	3,02	3,58	4,06	4,56	5,01	5,51
	y	4,37	4,01	3,29	3,10	3,22	2,99	2,90	2,37	1,87	1,82
5	x	1,10	1,33	1,58	1,81	2,09	2,32	2,59	2,85	3,14	3,43
	y	2,72	2,91	3,18	3,50	3,71	3,88	4,06	4,18	4,39	4,44
6	x	5,12	10,59	15,07	21,42	27,33	32,76	37,07	45,20	49,34	55,13
	y	6,50	7,17	7,39	7,85	8,01	8,20	8,23	8,49	8,57	8,53
7	x	9,35	15,26	20,67	26,49	33,14	39,39	44,49	50,71	56,01	63,00
	y	8,68	9,86	10,65	11,20	11,87	12,32	12,59	12,85	13,18	13,40
8	x	0,21	0,69	1,03	1,38	1,71	2,07	2,47	2,92	3,38	3,87
	y	0,19	0,47	0,67	0,66	0,92	1,00	1,15	1,39	1,48	1,72
9	x	0,32	0,93	1,59	2,29	2,92	3,62	4,30	4,99	5,63	6,25
	y	2,44	3,62	4,98	5,43	5,59	5,94	6,67	9,62	11,39	12,60
10	x	0,32	0,93	1,59	2,29	2,92	3,62	4,30	4,99	5,63	6,25
	y	45,71	5,97	3,82	3,27	3,06	2,86	2,79	2,63	2,54	2,60
11	x	0,38	0,93	1,61	2,35	3,02	3,63	4,35	5,02	5,72	6,25
	y	48,33	7,97	4,60	4,02	3,75	3,72	3,33	3,27	2,67	3,26
12	x	0,11	0,39	0,66	0,89	1,15	1,43	1,67	1,95	2,23	2,45
	y	1,78	2,20	4,30	5,49	6,57	7,15	10,48	12,52	17,53	24,23
13	x	1,03	1,63	2,16	2,71	3,26	3,77	4,35	4,91	5,50	6,01
	y	1,27	1,42	1,93	2,35	2,73	3,93	5,12	6,55	9,05	12,24
14	x	1,15	1,64	2,28	2,77	3,42	3,78	4,49	5,08	5,63	6,11
	y	5,44	7,01	10,62	13,40	17,42	21,70	28,60	35,68	41,37	49,45

Продовження таблиці 3.49

15	x	1,01	1,35	1,71	2,09	2,49	2,85	3,28	3,64	4,04	4,47
	y	5,02	5,51	6,52	7,92	9,64	10,06	14,06	19,17	25,58	32,56
16	x	9,33	15,72	20,87	26,74	31,95	37,97	44,61	51,32	58,23	64,57
	y	9,67	11,82	13,76	13,48	14,99	16,74	16,41	17,88	19,88	20,50

Задача 3.18.4. За вибірковими даними побудувати множинну лінійну регресію

Таблиця 3.50

Індивідуальні завдання для обчислення задачі 3.18.4

Варіант 1				Варіант 2				Варіант 3			
X_1	X_2	X_3	Y	X_1	X_2	X_3	Y	X_1	X_2	X_3	Y
2.31	10.1	6.315	7.627	2.12	9.97	6.28	7.45	2.41	10.25	6.32	7.73
4.67	1.7	7.729	10.7	4.4	11.43	7.64	10.48	4.78	11.95	7.79	10.94
6.17	13.9	8.479	11.53	6.16	13.61	8.25	11.25	6.26	13.87	8.57	11.82
8.7	14.4	8.691	13.40	8.69	14.28	8.61	13.33	8.95	14.52	8.83	13.59
10.7	15.1	10.5	17.02	10.47	14.93	10.21	16.9	10.75	15.18	10.55	17.5
13.5	17.1	10.52	18.75	13.41	17.04	10.43	18.57	13.57	17.41	10.59	18.81
16.2	18.9	11.68	21.14	15.98	18.82	11.54	20.91	16.43	19.01	11.83	21.26
18.3	20.3	13.77	23.37	18.24	20.3	13.73	23.32	18.55	20.52	13.83	23.38
21.2	21.7	13.7	27.45	20.89	21.48	13.64	27.16	21.45	22.02	13.97	27.62
22.7	22.4	14.43	27.13	22.66	22.22	14.41	27	22.75	22.53	14.55	27.18
25.1	22.5	14.07	29.61	24.91	22.41	13.98	29.59	25.19	22.64	14.17	29.87
26.1	24.7	16.46	32.52	26.03	24.45	16.45	32.24	26.25	24.89	16.64	32.64
27.5	24.8	15.02	31.8	27.25	24.75	14.83	31.71	27.63	25.02	15.07	32.01
29.9	25	15.27	35.18	29.74	24.89	15.06	35.1	30.16	25.9	15.38	35.25
32.1	26	15.66	37.07	31.80	25.95	15.61	36.77	32.20	26.15	15.72	37.14
33.7	27.4	17.21	38.85	33.66	27.27	17.14	38.7	33.94	27.60	17.27	38.96
35.8	28.9	17.47	?	35.57	28.8	17.39	?	36	29.02	17.56	?
Варіант 4				Варіант 5				Варіант 6			
X_1	X_2	X_3	Y	X_1	X_2	X_3	Y	X_1	X_2	X_3	Y
2.37	10.27	6.4	7.73	2.25	9.9	6.09	7.49	2.61	10.35	6.61	7.72
4.77	11.87	7.88	10.85	4.42	11.54	7.49	10.64	4.89	11.78	7.94	1077
6.24	13.88	8.5	11.54	6.08	13.73	8.46	11.44	6.24	14.09	8.62	11.86
8.7	14.62	8.86	13.52	8.65	14.26	8.59	13.24	9.01	14.64	8.83	13.73
10.79	15.26	10.51	17.13	10.64	14.91	10.43	16.99	10.79	15.17	10.68	17.04
13.6	17.29	10.53	18.75	13.29	17.02	10.52	18.57	13.53	17.42	10.66	18.8
16.31	19.04	11.74	21.15	15.95	18.84	11.65	21.07	16.32	19.24	11.78	21.28
18.4	20.45	13.96	23.49	18.25	20.06	13.55	23.23	18.6	20.6	13.78	23.7
21.25	21.94	13.86	27.5	21.1	21.71	13.67	27.37	21.48	22.04	13.74	27.63
22.87	22.55	14.6	27.16	22.67	22.31	14.33	27.12	23.02	22.69	14.56	27.45
25.15	22.56	14.24	29.73	24.99	22.39	13.95	29.42	25.17	22.65	14.09	29.71
26.27	24.79	16.59	32.71	26	24.5	16.34	32.43	26.4	24.83	16.65	32.6
27.7	24.82	15.03	31.83	27.34	24.76	14.81	31.74	27.62	24.82	15.12	31.81
30	25.11	15.34	35.18	29.75	24.99	15.13	35.16	30.19	25.17	15.42	35.22
32.25	26.11	5.84	37.12	31.87	25.94	15.46	37.07	32.25	26.22	15.77	37.26
33.85	27.58	17.3	38.97	33.55	27.35	16.96	38.74	33.76	27.72	17.4	39.20
35.96	29.09	17.48	?	35.71	28.83	17.25	?	35.97	29.15	17.77	?
Варіант 7				Варіант 8				Варіант 9			
X_1	X_2	X_3	Y	X_1	X_2	X_3	Y	X_1	X_2	X_3	Y
2.61	10.32	6.71	7.95	2.4	11.27	6.78	8.61	1043	9.18	6.17	7.29
4.97	11.96	8.14	10.75	5.66	12.48	8.49	11.33	3092	10.94	7.7	9.79
6.55	13.96	8.95	11.66	6.87	14.11	9.46	12.48	5.49	13.3	8.43	11.26
9.16	14.62	8.83	13.49	9.52	14.64	9.44	14.37	8.17	13.89	8.12	12.42

Продовження таблиці 3.50

10.68	15.18	10.95	17.44	11.57	16.17	10.82	17.43	9.68	14.49	10.41	16.04
13.93	17.34	10.92	19.05	13.73	17.47	11.01	19.5	13.42	16.73	10.39	18.34
16.27	19.21	11.75	21.42	16.87	19.85	12.16	21.22	15.92	17.97	11.36	20.94
18.74	20.36	14.05	23.85	18.73	21.43	14.46	23.84	18.04	19.45	13.32	22.74
21.31	22.21	14.09	27.92	22.16	22.12	13.98	27.8	20.69	21.49	12.72	27.09
22.88	22.84	14.6	27.27	23.33	22.98	14.72	27.82	22.68	21.8	14.22	26.43
25.13	22.63	14.38	30.04	25.84	22.61	14.58	30.01	24.33	21.64	13.51	29.39
26.28	24.93	16.57	32.83	26.69	25.03	17.07	33.32	25.64	24.48	15.83	32.37
27.71	24.94	15.51	31.89	27.78	25.16	15.39	31.87	27.14	24.02	14.52	3.52
30.01	25.27	15.44	35.21	30.86	26.05	15.95	35.35	29.22	24.42	15.06	34.89
32.08	26.18	16	37.27	32.87	26.44	16.37	37.5	31.09	25.96	15.32	36.33
33.74	27.61	17.59	38.99	34.65	28.09	17.94	39.91	33.34	27.05	16.65	38.35
36.31	29.23	17.97	?	36.25	29.76	18.44	?	35.64	28.12	16.55	?
Варіант 10				Варіант 11				Варіант 12			
X_1	X_2	X_3	Y	X_1	X_2	X_3	Y	X_1	X_2	X_3	Y
3.18	10.95	6.94	8.5	1.49	10.12	5.97	7.54	0.95	8.71	4.67	5.94
5.76	12.66	8.2	11.67	3.65	10.54	6.62	10.36	2.32	10.89	6.16	10.2
7.26	14.35	9.03	12.11	6.11	13.09	7.92	11.46	6.09	13.54	7.18	10.22
8.95	15.04	9.87	14.09	8.59	14.03	8.17	12.38	7.36	12.56	6.81	12.4
11.44	16.26	10.65	18.01	9.73	15.13	10.01	15.88	10.28	13.5	9.87	15.64
14.59	18.13	10.57	19.21	12.49	16.51	10.02	16.21	11.97	16.69	8.58	18.23
16.91	19.73	12.18	21.33	15.58	17.73	11.61	20.34	15.08	18.84	10.22	19.34
18.4	21.09	14.02	23.54	18.03	19.8	13.43	23.34	16.43	19.59	12.72	23.28
21.84	22.47	13.77	27.72	20.43	20.55	13.33	27.34	19.59	20.15	12.39	25.8
23.88	22.58	15.01	27.18	21.71	22.08	13.78	26.3	22.44	20.79	13.12	25.47
25.98	22.68	14.51	30.26	23.88	22.36	13.64	29.51	24.53	20.57	2.58	27.72
26.85	25.76	17.62	33.09	26	23.84	15.47	32.25	24.56	23.28	16.15	31.4
28.71	25.63	15.59	32.22	26.52	24.29	14.52	31.78	27.39	24.47	13.3	31.67
30.38	25	16.23	35.42	29.26	24.33	14.75	34.71	29.41	23	14.73	33.68
32.66	26.36	16.63	37.21	31.55	25.73	14.59	35.96	31.82	24.87	14.94	36.97
33.88	28.54	17.28	39.64	32.61	27.38	17.02	38.33	33.13	26.17	16.81	37.22
36.54	29.13	18.59	?	35.66	28.14	16.45	?	35.2	28.3	16.57	?

Комп'ютерний проект 1. Аналіз статистичних даних

- Форми та методи подання і попереднє статистичне опрацювання числових даних часових послідовностей
  - Попереднє опрацювання даних та подання результатів
  - Формування файлу даних у формі Excel та Word таблиць
  - Графічне подання даних
  - Описова статистика – кількісні характеристики даних
  - Побудова гістограми
  - Побудова кумуляти
- Виявлення тенденції часового ряду методами згладжування
  - Методи згладжування часових рядів
  - Метод рухомого середнього
  - Метод зваженого рухомого середнього
  - Властивості рухомого середнього
  - Медіанна фільтрація
  - Нормування часових послідовностей
  - Критерії ефективності згладжування часових рядів
  - Формули для зваженого рухомого середнього
- Кореляційний аналіз часових послідовностей
  - Кореляційне поле
  - Коефіцієнт кореляції
  - Кореляційне відношення
  - Властивості кореляційного відношення
  - Кореляційна матриця
  - Автокореляція
  - Автокореляція в часових рядах
  - Розрахунок автокореляції
- Ієрархічний агломеративний кластерний аналіз багатовимірних даних
- Порядок роботи
- Хід роботи
- Форма звітності

Метою роботи є ознайомлення з основними методами візуалізації, графічного відображення та первинного статистичного опрацювання числових даних, які представлені вибірковою сукупністю або часовим рядом; ознайомлення з основними методами висвітлення поведінки досліджуваного показника, яка представлена характером його тренду, за допомогою методів згладжування часових рядів та подання отриманих результатів засобами табличного процесора MS Excel. Метою роботи також є ознайомлення з методами кореляційного аналізу експериментальних даних, які представлені часовими послідовностями. Для цього потрібно:

– побудувати кореляційне поле;

– визначити значення коефіцієнта кореляції;
– обчислити кореляційне відношення;
– побудувати графіки автокореляційних функцій;
– розбити одну з послідовностей на три рівні частини;
– побудувати для них кореляційну матрицю;
– знайти коефіцієнти множинної кореляції.

Необхідно розділити задану множину об'єктів, кожен з яких характеризується однаковою сукупністю конкретних ознак, на окремі групи, використовуючи ієрархічний агломеративний кластерний аналіз[94-102, 110-111, 137-292, 303-320].

Більшість користувачів не знайомі зі всіма можливостями Excel. Excel потрібен перш за все тим людям, які у своїй роботі оперують числами (наприклад, бухгалтерам та науковим працівникам). Але цим не вичерпується потенціал електронних таблиць. Практично у всіх випадках, коли інформація може бути подана в табличному вигляді, Excel є незмінним помічником. Будь-який користувач, який знає можливості Excel, завжди може знайти застосування у своїй роботі. За допомогою Excel можна створювати, редагувати і друкувати оформлені хорошим дизайном таблиці. Завдяки вбудованим математичним і логічним функціям, можна дуже швидко виконувати різні операції, як над окремими цифрами, так і над текстами, проводити статистичний аналіз та оптимізацію даних. Можна створювати всілякі форми графічного подання змісту таблиць, створювати бази даних та працювати з ними. Excel має величезні можливості і, безсумнівно, є однією з провідних програм свого класу.

Починаючи працювати на будь-якому комп'ютері, можна бути певним, що на ньому встановлена ця програма. Це надзвичайно важливий момент, оскільки пошук потрібного програмного забезпечення як правило займає дні, а його освоєння місяці. Саме тому, вміння використовувати Excel є дуже перспективним, важливим з точки зору освоєння надзвичайно широкого спектру засобів опрацювання й подання результатів практичних досліджень.

Перші версії Excel були створені ще до появи операційних систем сімейства Windows. Але й після цього Excel не зупинився у своєму розвитку: були створені версії 7.0, 97, 2000, 2002, 2003, в кожній з яких з'являлися нові можливості та вдосконалювалися наявні, усувалися помилки попередніх версій.

Excel входить до складу пакету офісних програм Microsoft Office і, як правило, встановлюється разом з усіма іншими програмами пакету. Незважаючи на те, що суттєвих відмінностей між версіями Excel немає, користувач, що освоїв одну з них, легко переходить на іншу. Початкове вивчення все ж краще починати з тієї версії, яка встановлена на комп'ютері. Разом з тим, якщо потужність комп'ютера дозволяє, слід встановити найпоширенішу версію програми, а саме MS Excel 2003.

Електронні таблиці Excel мають величезний набір можливостей. Проте, зазвичай, користувач у своїй роботі застосовує тільки їх невелику частину. Відбувається це не тільки від незнання, але і від відсутності в цьому гострої необхідності. І дійсно, всі функції, закладені в Excel, можна умовно розділити на ті, знання яких дуже бажане або просто необхідне, і ті, вивчати які можна у міру потреби.

Для успішного аналізу й прийняття ефективних рішень в економічних та екологічних ситуаціях відповідна інформація про їх стан повинна бути якомога більш повно представлена. Зокрема, у випадках, коли числові дані є пов'язані з моментами часу – надані у формі часових рядів або впорядкованих в часі вибірок.

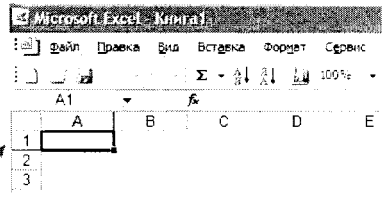
КП.1. Форми і методи подання та попереднє статистичне опрацювання числових даних часових послідовностей

КП.1.1. Попереднє опрацювання даних та подання результатів

Отримані в різних дослідженнях дані переважно характеризують спостережуваний об'єкт в двох аспектах – статичному й динамічному. Статичний аспект дає характеристику об'єкта значеннями конкретних величин, які можуть бути як реальними, тобто характеризують об'єкт таким як він є (кількість елементів конкретної схеми, значення елемента за даною шкалою) або випадковими – з відомим або невідомим законом розподілу їх ймовірності. Динамічний аспект характеризує дані про об'єкт протягом часу, регулярно або нерегулярно, але так, щоб кожне значення було прив'язане до моменту часу його спостереження чи реєстрації. Ці два аспекти виражаються з точки зору їх опрацювання двома класами: вибірковими сукупностями – вибірками і часовими послідовностями або часовими рядами. Для першого класу результатом опрацювання є визначення закону розподілу випадкових значень елементів вибірки – варіант. Зауважимо, що невипадкові значення не потребують знаходження виду та параметрів закону розподілу, а використовуються безпосередньо в розрахунках або їх перед цим усереднюють. Для другого класу, який є різновидом випадкових процесів, результатом опрацювання є аналітичне подання тенденції розвитку досліджуваного показника в часі. Завершенням таких досліджень є математична модель, в якій ці дві компоненти одного і того ж самого набору даних (вважаємо, що дані прив'язані до часу або до іншого показника) об'єднані аналітично, а сама модель, з точністю до прийнятого критерію адекватності, описує їхню поведінку. Попереднє опрацювання даних фактично дає підстави для побудови такої математичної моделі, проте для цього мають бути реалізовані певні етапи. Для досягнення поставленої мети в цій роботі необхідно забезпечити поетапне виконання низки завдань. Тобто, вирішення цих завдань здійснюється наступними етапами.

КП.1.2. Формування файлу даних у формі Excel та Word таблиць

Для цього, отримані дані формуються у файл даних, тобто файл, в якому дані відповідним чином впорядковані (наприклад: в послідовності їх отримання, у відповідності з часом отримання). Такий файл має мати форму таблиці «об'єкт - властивість», «порядковий номер – значення показника». Як правило такий файл є звичайним комп'ютерним файлом, який може бути сформований за допомогою Word або навіть побудований в «Блокноті», але він має мати відповідне розширення, наприклад, *.dat або *.txt. Після формування такого файлу його необхідно внести в книгу Excel. Для цього необхідно у відкритій книзі Excel відкрити доступ до віддалених файлів.



Далі, за допомогою

вказавши формат вихідних даних і тип розділювача, виконати відповідні кроки. В результаті в книзі буде введено вказаний файл. Для отриманих експериментальних даних може виявитися значний розмір стовпчиків, що у випадку внесення такої таблиці у сторінку Word спричинить незручність сприйняття даних, оскільки дані можуть вимагати декілька незаповнених сторінок. Тоді, отриману таблицю варто «стиснути» в Excel, подавши її у такій формі. Подання даних у формі стиснутої таблиці здійснюють за допомогою розбиття оригінальної на кількість частин так, щоб в таблиці було як найменше вільних комірок, а сама таблиця не виходила за межі, що визначені текстовим форматом. Не слід дуже ущільнювати стовпчики таблиці, оскільки дані повинні бути читабельними за значеннями, номери та їх значення не повинні бути дуже близько. Це має важливе практичне значення, особливо, коли дані представляються для аналізу в роздрукованому, але без електронного супроводу, представленні. Такі дані для здійснення аналізу сканують і порушення цієї вимоги може призвести до суттєвих помилок.

Таблиця КП.1.

Подання значень даних у формі стисненої таблиці

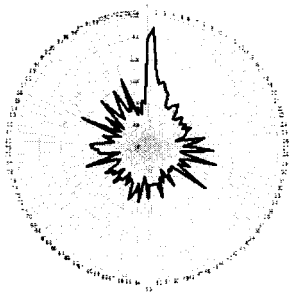
РЕЗУЛЬТАТИ РЕЄСТРАЦІЇ ДИНАМІКИ ПОКАЗНИКА									
№ п/п	Значення показника	№ п/п	Значення показника	№ п/п	Значення показника	№ п/п	Значення показника	№ п/п	Значення показника
1	**2465**	21	**933**	41	**690**	61	**800**	81	**871**
2	**2738**	22	**885**	42	**748**	62	**971**	82	**803**
3	**1698**	23	**663**	43	**1042**	63	**711**	83	**1043**
4	**1482**	24	**1321**	44	**609**	64	**934**	84	**819**
5	**1544**	25	**1148**	45	**717**	65	**937**	85	**651**
6	**1386**	26	**826**	46	**856**	66	**1111**	86	**1654**
7	**1315**	27	**1118**	47	**741**	67	**733**	87	**806**
8	**1032**	28	**1497**	48	**1183**	68	**1006**	88	**725**
9	**998**	29	**757**	49	**836**	69	**921**	89	**1160**
10	**1119**	30	**865**	50	**836**	70	**1140**	90	**695**
11	**1203**	31	**1605**	51	**843**	71	**809**	91	**1039**
12	**1107**	32	**872**	52	**841**	72	**686**	92	**1482**
13	**889**	33	**784**	53	**1150**	73	**744**	93	**1001**
14	**907**	34	**1055**	54	**819**	74	**1462**	94	**835**
15	**1043**	35	**786**	55	**683**	75	**876**	95	**1098**
16	**968**	36	**913**	56	**1265**	76	**683**	96	**1489**
17	**895**	37	**1597**	57	**841**	77	**1108**	97	**935**
18	**1030**	38	**695**	58	**1047**	78	**1146**	98	**915**
19	**893**	39	**716**	59	**1271**	79	**1295**	99	**1675**
20	**1089**	40	**756**	60	**882**	80	**1320**	100	**779**

КП1.3. Графічне подання даних

Основним типом візуалізації даних у звітах експериментальних та науково-практичних досліджень є графіки. Графік відображає відношення між двома величинами, одна з яких є незалежною змінною і її значення, зазвичай, відкладають вздовж горизонтальної осі – *абсциси*; друга змінна є залежною і її значення відкладають вздовж вертикальної осі – *ординати*. Фактично, будь-який графік складається з декількох основних елементів: осей з надписами, експериментальних точок, ліній, що з'єднують ці точки (їх ще називають *кривими*), пояснюючих написів на рисунку та підпису під рисунком.



а)



б)

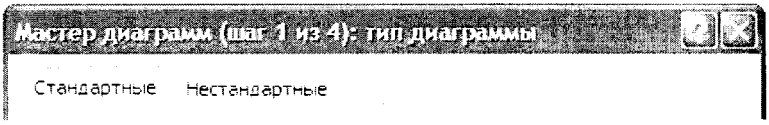
Рис. КП.1. Графічне подання динаміки показника: *а* – в декартовій системі координат, *б* – в полярній системі координат.

В загальному, щоб побудувати графік необхідно:

– підготувати таблицю з відображуваними даними;
– визначити розміри осей з реперними штрихами і цифрами;
– нанести експериментальні точки і провести криві;
– зробити відповідні написи для осей та шкал на рисунку;
– підготувати підпис до рисунка.

Наявність табличного процесора (електронної таблиці MS Excel) суттєво спрощує процедуру побудови графіка, переважно з геометричної точки зору, тобто креслення його за точками, в той час, як сама підготовка та редакція й корегування його зовнішнього вигляду залишаються прерогативою фахівця.

На графіку мають бути позначені осі, виділені маркери, має бути назва графіка, назви осей, значення поділок повинно відповідати значенням варіант. Для побудови графіка в полярній системі координат використовуємо «Майстер діаграм» і вибираємо на вкладці «Стандартні» опцію «Пелюсткова»



Редагування графіка – вибір розміру шрифту, параметрів та кольору осей здійснюємо звичайним способом.

КП.1.4. Описова статистика – кількісні характеристики даних

Крім табличного та графічного представлення даних, в їх супровід включають загальні числові та статистичні характеристики, які відносять до описової або дескриптивної статистики.

Описова статистика дає підстави для формування компетенцій щодо вибору шкали вимірювань, автоматизації опрацювання даних при застосуванні різних форматів на етапі їх збору, подання результатів у різних формах, графічного подання результатів, обчислення статистичних параметрів розподілу та оцінки параметрів генеральної сукупності з використанням інформаційних технологій. Вона займається вибором кількісної інформації, яка необхідна (або цікава) для різних людей. Великі масиви даних, перш ніж вони вивчатимуться людиною, мають узагальнюватися або згортатися. Саме це робить описова статистика, яка описує, узагальнює або зводить до бажаного вигляду властивості масивів даних. Описова статистика застосовується для аналізу та інтерпретації статистичних даних, побудови статистичних розподілів та обчислення відповідних числових параметрів, що характеризують досліджувану сукупність. Її використовують для організації збирання інформації, перевірки якості даних та їхньої інтерпретації, зображення статистичного матеріалу.

Описова (дескриптивна) статистика – це найбільш загальні статистичні показники, що описують розподіл даних, приймаючи за норму нормальний розподіл. Це тому, що характерною властивістю нормального розподілу є те, що 68% всіх його спостережень лежать в діапазоні $\bar{x} \pm$ одне стандартне відхилення від середнього арифметичного, а в діапазоні $\bar{x} \pm$ два стандартних відхилення міститься 95% значень вибірки з нормальним розподілом. Ці два показники – середнє арифметичне і стандартне відхилення є основними параметрами нормального розподілу.

Середнє арифметичне є мірою центральної тенденції, що відображає найбільш характерне для даної вибірки значення. Його визначають за формулою

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i , \qquad \text{(КП.1)}$$

де n – обсяг вибірки.

Оманливість цього показника, проілюструємо таким прикладом: в одному купе вагона розмістилася бабуся 60 років з чотирма онуками: одному – 4 роки, двом – по 5 років й одному – 6 років. Середнє арифметичне віку всіх пасажирів цього купе 80/5=16. В іншому купе розташувалася компанія молоді: двоє –15-тилітніх, один – 16-річний та двоє – 17-річних. Середній вік пасажирів цього купе так само дорівнює 80/5=16. Таким чином, за середнім арифметичним пасажири цих купе не відрізняються. Але, якщо звернутися до показника стандартного відхилення, то виявиться, що середній розкид щодо середнього віку в першому випадку виявиться 24,6, а в другому випадку – 1.

Мода (позначається «*Mo*») – це значення, яке найбільш часто зустрічається серед вибірки змінних. Часто застосовується для непараметричних даних і для рангових шкал.

Медіана (позначається «*Me*») – значення, яке ділить навпіл впорядковану множину змінних, тобто для визначення медіани необхідно впорядкувати дані, наприклад, за зростанням. Способи визначення значення медіани для парної і непарної кількості даних відрізняються. Для непарної кількості даних визначають її номер у впорядкованій сукупності за такою формулою:

$$Me(n)=\frac{n+1}{2}. \qquad (КП.2)$$

У випадку парної кількості даних визначають номери двох сусідніх серединних значень $\frac{n}{2}$ і $\frac{n+2}{2}$. Середнє арифметичне цих двох значень вибірки (з номерами $x\left(\frac{n}{2}\right)$ і $x\left(\frac{n+2}{2}\right)$) і буде значенням медіани для парного n, тобто:

$$Me(n)=\frac{Me\left(\frac{n}{2}\right)+Me\left(\frac{n+2}{2}\right)}{2}. \qquad (КП.3)$$

Значення середнього, моди та медіани є близькими один до одного. В ідеальному нормальному розподілі вони рівні, оскільки мають однаковий зміст: середина розподілу.

Розмах (інтервал) – показник, який вказує на ширину діапазону значень. Він дорівнює різниці між максимальним і мінімальним значеннями.

Стандартне відхилення (σ, читається «сигма») – є мірою мінливості (варіації) ознаки, яка відображає величину його розкиду відносно середнього арифметичного

$$\sigma=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}. \qquad (КП.4)$$

Для більш точного та наочного представлення про варіацію значень показника відносно середнього використовується коефіцієнт варіації:

$$v=\frac{\sigma}{\bar{x}}\cdot 100\%. \qquad (КП.5)$$

Коефіцієнт варіації виражає міру мінливості ознаки у відсотках.

Асиметрія – показник, що відображає перекіс розподілу відносно моди вліво або вправо. Це має місце в тих випадках, коли якісь причини сприяють більш частій появі значень, які є більшими або, навпаки, меншими за середнє арифметичне. Для лівосторонньої або додатньої асиметрії в розподілі частіше зустрічаються більш низькі значення, а для правосторонньої або від'ємної – вищі.

Ексцес – показник, що відображає висоту розподілу. У тих випадках, коли якісь причини сприяють появі близьких до середніх значень, утворюється розподіл з додатнім ексцесом. Якщо ж в розподілі переважають крайні значення, причому одночасно і більш низькі, і більш високі, то такий розподіл характеризується від'ємним ексцесом, і в центрі розподілу може утворитися впадина, яка перетворює його в двохвершинний.

Описова статистика включає в себе табулювання (складання таблиць), подання та опис сукупності даних. Засіб аналізу «Описова статистика» використовується для створення одномірного статистичного звіту, який містить інформацію про центральну тенденцію та мінливість даних початкового діапазону.

За допомогою опції «Описової статистики» отримують такі кількісні дані: середнє арифметичне, стандартну помилку, медіану, моду, стандартне відхилення, дисперсію, ексцес, асиметрію, розмах, мінімальне та максимальне значення, суму значень, їх кількість та рівень надійності.

Ці характеристики як правило використовують при первинній ідентифікації даних, тобто їхнього порівняння і виявлення змін. Табличний процесор MS Excel враховує необхідність такого представлення характеристик даних і забезпечує процедури їх визначення за допомогою операторів, що зведені нижче в таблицю.

Результати описової статистики	
Показники	Значення
Середнє	932,7837838
Стандартна помилка	41,34955245
Медіана	841
Мода	836
Стандартне відхилення	251,5195083
Дисперсія вибірки	63262,06306
Ексцес	1,371916976
Асиметричність	1,384445279
Інтервал	996
Мінімум	609
Максимум	1605
Сума	34513
Обсяг	37
Рівень надійності (95,0%)	83,86077869

а

Результати описової статистики	
Показники	Значення
Середнє	932,78
Стандартна помилка	41,35
Медіана	841,00
Мода	836,00
Стандартне відхилення	251,52
Дисперсія вибірки	63262,06
Ексцес	1,37
Асиметричність	1,38
Інтервал	996,00
Мінімум	609,00
Максимум	1605,00
Сума	34513,00
Обсяг	37,00
Рівень надійності (95,0%)	83,86

б

Рис. КП.2. Результати описової статистики: *а* – безпосередньо отримані, *б* – підготовлені до звіту (кількісні дані подані з двома розрядами).

Подання числових значень має бути приведено до відповідної розрядної сітки.

КП.1.5. Побудова гістограми

Якщо отримані дані утворюють звичайну репрезентативну вибірку, тобто вибірку, обсяг якої є достатнім для визначення їхнього розподілу, традиційно, для визначення його вигляду будують гістограму. Гістограма є дуже наближеним відображенням графіка функції щільності закону розподілу даних цієї вибірки. Вона також є діаграмою, яка наочно відображає метод групування даних за деякою істотною ознакою. Методи групування широко використовуються для початкового опрацювання даних, оскільки вони суттєво зменшують обсяг даних і виявляють найбільш характерну для них структуру – частоту значень, розподіл в групах, а в більшості випадків характер розподілу: одно чи багато модальність, положення та наближене значення моди як середини модового інтервалу, наближений вигляд модового інтервалу. Найбільш проблематичним у побудові гістограми є вибір кількості інтервалів групування, тобто кількості груп, на які розбивається вибірка. Для вибору чи визначення кількості інтервалів розбиття існує декілька десятків формул, проте найбільш поширеними є:

формула Стерджеса $k = 1 + \log_2 n$, k – кількість інтервалів;

формула Скотта $h = 3.5 \cdot s \cdot n^{-1/3} = \dfrac{3.5 \cdot s}{\sqrt[3]{n}}$,

де h – ширина інтервалу, s – стандартне відхилення значень ряду.

Побудова гістограми полягає в реалізації наступних кроків:

1. За обсягом вибірки (кількістю даних) n та стандартним відхиленням s визначають кількість або ширину інтервалів. Далі визначають значення меж інтервалів (за термінологією Excel – «Інтервал кишень»).

2. Послідовно «книга Excel → Сервіс → Аналіз даних → Гістограма» заповнюють дані в Excel комірках Майстра гістограми,



Рис. КП.3. Майстер побудови гістограми.

а саме:

- у «Вхідний інтервал» вносять дані вибірки;
- в «Інтервал кишень» вносять значення меж інтервалів;
- вказують комірку для «Вихідного інтервалу»;
- відмічають потрібні результати: «Інтегральний процент» та «Вивід графіку» і «ОК».

За даними стовпчика «Частота» в таблиці отриманого результату будують, використовуючи «Майстер діаграм» і опцію «Гістограма», гістограму отриманих даних. На рис. зображені двомірне та трьохмірне подання гістограми засобами Excel.





На підставі зображення гістограми здійснюють аналіз частот значень на предмет апроксимації емпіричної щільності закону розподілу вихідних даних – отриманої гістограми деякою аналітичною функцією. Як таку функцію часто вибирають функцію закону розподілу серед відомих функцій розподілів випадкових величин або використовують відповідні інші математичні функції.

КП.1.6. Побудова кумуляти

Основними недоліками гістограми є такі: суб'єктивність у виборі кількості інтервалів групування, значення середнього арифметичного, моди і медіани подаються з точністю до розміру інтервалу, обмежена кількість вузлів апроксимації. Проте, деякі з

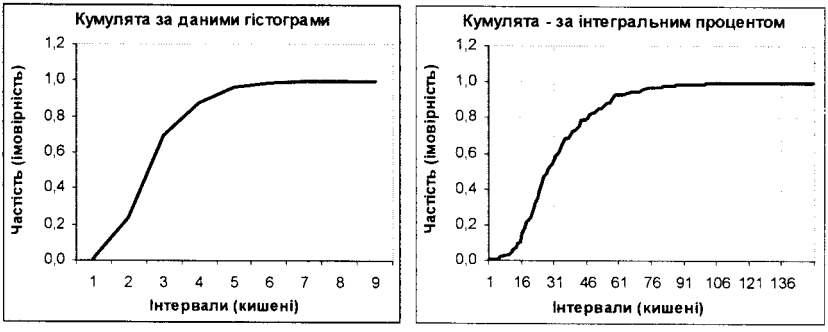
цих недоліків можна усунути використовуючи метод побудови гістограми для побудови *кумуляти* (емпіричного графіка функції закону розподілу), якщо зробити такі зміни.

1. Встановити кількість інтервалів рівною $n-1$, що виключає суб'єктивність у виборі потрібної формули.

2. Побудувати гістограму для такої кількості інтервалів (будуть присутні порожні інтервали, а тому інтерпретація такої гістограми є специфічною).

3. За даними стовпчика «Інтегральний процент» в таблиці отриманого результату будують, використовуючи «Майстер діаграм» і опцію «Графік», графік емпіричної кумуляти (емпіричної функції закону розподілу отриманих даних).

Емпіричну кумуляту можна побудувати і на підставі гістограми, оскільки, оскільки в таблиці її побудови також є стовпчик з даними про інтегральний процент. Відмінність цих двох кумулят в тому, що кумулята побудована за даними гістограми має k вузлів апроксимації і є ломаною кривою, а кумулята побудована за інтегральним процентом має $n-1$ вузлів апроксимації і є більш плавною. На рис. зображені обидві кумуляти.



а *б*

Рис. КП.4. Кумуляти побудовані: *а* – за даними гістограми;
б – за інтегральним процентом.

Розглянутий підхід до попереднього опрацювання результатів досліджень дає підстави для подання, як самих результатів, так і їхніх основних характеристик, в достатньо інформативній наочній формі – графіки, діаграми та кількісній формі – показниках описової статистики. Результати такого підходу можуть бути достатніми для приведення їх у наукових звітах, дипломних та магістерських кваліфікаційних робота, а також в дисертаційних дослідженнях.

КП.2. Виявлення тенденції часового ряду методами згладжування

При побудові математичних моделей часових рядів, переважно з метою визначення динаміки показника, необхідно в першу чергу виділити тенденцію та

відокремити її від випадкових відхилень, зумовлених різними перешкоджаючими факторами. Дослідження часового ряду починається з його графічного представлення. При візуальному способі будується графік часового ряду, на основі якого висувається гіпотеза про його структуру, в першу чергу про форму тренду. Цей підхід дає задовільні результати при відносно монотонних тенденціях. Однак, у випадку значних флуктуацій модельованого процесу можливість помилок у виборі виду функції тренда при цьому підході зростає. Методи виявлення основної тенденції розвитку досліджуваного об'єкта переважно визначаються на основі докладного вивчення фактичного розвитку його динаміки. Вони повинні узгоджуватись з результатами спостережень і статистикою емпіричного матеріалу. Ці методи мають різну логічну змістовність і тому застосовуються до часових рядів в залежності від цілей дослідження. Основна мета полягає в тому, щоб розкривати загальні закономірності розвитку, затушовані окремими, іноді випадковими обставинами. Проте кожен з них має свої особливості. Для виявленні тенденції – характеру розвитку використовують процедуру згладжування часового ряду. Суть її зводиться до заміни фактичних рівнів часового ряду розрахунковими, але з меншими коливаннями, що сприяє більш чіткому проявленню тенденції та її характеру. Саме в цьому випадку, тенденцію зображають гладкою неперервною функцією, яку або її графік називають ***трендом*** часового ряду.

КП.2.1. Методи згладжування часових рядів

Методи згладжування можна умовно розділити на два класи, в основі яких лежать різні підходи: аналітичний та алгоритмічний.

Аналітичний підхід оснований на припущенні, що дослідник може на підставі візуального аналізу задати загальний вигляд функції, вважаючи що її графік відповідає характеру тенденції. Наприклад, на основі візуального та змістовного аналізу властивостей об'єкта та динаміки часового ряду, поведінку якого він описує, як функція може бути використана: експонента, гіпербола, парабола, степенева функція, поліноми вищих степенів та інші функції.

Наступний етап передбачає аналітичне або статистичне оцінювання статистичних невідомих параметрів вибраної функції, яка в цьому випадку стає математичною моделлю тенденції даного часового ряду.

Іншими словами, аналітичний підхід означає заміну значень рівнів часового ряду значеннями, що теоретично розраховані на підставі явного аналітичного вигляду функції, якою апроксимують візуально визначений тренд.

Цей підхід успішно реалізований в Excel, використовуючи поліноми до 6-го степеня включно, а також степеневу, експонентну та логарифмічну функції.

В **алгоритмічному підході** вигляд тренду отримують за рахунок різних алгоритмів, які практично реалізують процедури згладжування. Ці процедури надають дослідникові лише алгоритм розрахунку нового значення часового ряду в будь-який заданий момент часу t.

Ці методи можна класифікувати так:

- просте або звичайне ковзне (рухоме) середнє;
- зважене ковзне (рухоме) середнє;
- експоненціальне згладжування;
- медіанне згладжування.

Найбільш вживаними є методи згладжування часових рядів за допомогою ковзних середніх.

Метод ковзних (рухомих) середніх є одним із найстаріших відомих способів згладжування часового ряду. Він базується на переході від початкових значень ряду до їх середніх значень на інтервалі часу, довжина якого обрана заздалегідь. При цьому сам вибраний інтервал часу ковзає вздовж ряду.

Суть цього методу зводиться до заміни фактичних рівнів ряду послідовностями рівнів, що мають, як правило, значно менші коливання, ніж вихідні дані. Зменшення флуктуації дає можливість наочно виявити основну тенденцію. Часто таку операцію над вихідними даними називають фільтруванням, а оператор її здійснення фільтром. У зарубіжних статистичних пакетах для цих процесів вживають абревіатуру MA – від англійського ковзної середньої (рухоме середнє).

Ковзні середні дозволяють згладити як випадкові, так і періодичні коливання, виявити наявну тенденцію в розвитку процесу і є важливим інструментом при фільтрації компонент часового ряду. Вони можуть бути визначені за допомогою простих ковзних або зважених середніх.

Вибір методу виявлення основної тенденції розвитку залежить від технічних можливостей обчислень та від уміння застосовувати відповідні методи, а також від завдань, які стоять перед дослідженням. Якщо треба дати загальну картину розвитку, його грубу модель, що основана на механічному повторенні одних і тих же дій крок за кроком до послідовності рівнів, то можна обмежитися методом ковзної середньої. Якщо ж мета дослідження полягає в розробці математичної моделі тренду, то сам метод ковзної середньої буде недостатнім. Тоді треба буде використовувати метод кінцевих різниць або метод найменших квадратів.

КП.2.2. Метод рухомого середнього

Метод рухомих (ковзних) середніх дає оцінку середнього рівня за деякий період часу. Чим більше інтервал часу, до якого належить середня, тим більше згладжуватиме рівень, але тим менш точно буде описана тенденція вихідного ряду динаміки. Одержуваний таким чином ряд ковзних середніх поводиться набагато більш гладко, ніж вихідний ряд, за рахунок усереднення відхилень вихідного ряду. Таким чином, ця процедура дає уявлення про загальну тенденцію поведінки ряду. Її застосування особливо корисне для рядів з сезонними коливаннями і незрозумілим характером тренду. Зокрема, перехід до ряду ковзних середніх може бути використаний для виявлення коливальної компоненти часового ряду.

Згладжування ряду динаміки за допомогою ковзної середньої полягає в тому, що обчислюється середній рівень з певного числа перших за порядком рівнів ряду. Ці перші рівні ряду утворюють інтервал постійної величини, в якому рівні розташовані в тому ж порядку, що й у часовому ряді. Заміна першого і останнього наступними сусідніми рівнями за незмінної величини інтервалу створює враження руху цього інтервалу вздовж часового ряду. Оскільки, після кожного кроку, для рівнів охоплених цим інтервалом обчислюється середнє значення, то його використовують для утворення нового ряду. Причому, це обчислене значення в новому ряді відповідає тому моменту часу, якому відповідає значення рівня в середині ковзаючого інтервалу. Для забезпечення такої часової відповідності інтервал має охоплювати непарну кількість рівнів, тобто $w = 2n - 1$, де $n = 1, 2, 3, \ldots$. Таким чином, при обчисленнях середнього рівня межі інтервалу ніби «ковзають» по рівнях часового ряду від його початку до кінця, щоразу відкидаючи один рівень на початку інтервалу і додаючи в кінці наступний. Звідси назва – ***ковзна середня***.

Алгоритм для обчислення простої ковзної середньої є такий

$$\widetilde{y}_i = y_1^* + y_2^* + \ldots + y_k^* + \sum_{j=k+1}^{N-2k} \left[\frac{1}{w} \sum_{i=j}^{j+2k+1} y_i \right] + y_{N-k}^* + \ldots + y_{N-1}^* + y_N^* \quad . \qquad \text{(КП.6)}$$

Визначення ковзної середньої у випадку парного числа рівнів в ковзаючому інтервалі ускладнюється тим, що тоді середня має бути віднесена тільки до середини між двома моментами часу, що знаходяться всередині інтервалу згладжування, а в такий момент часу спостереження не проводились.

Якщо графічне представлення часового ряду нагадує пряму лінію, то в цьому випадку ковзна середня не спотворює динаміку досліджуваного явища.

У випадку, коли тенденція вихідного ряду, що характеризує досліджуваний процес, не може бути описана лінійним трендом, більш надійним є використання зваженої ковзної середньої.

КП.2.3. Метод зваженого рухомого середнього

Поряд з простими ковзними середніми застосовуються також поліноміальні або зважені середні. Вони дозволяють точніше описати початок форми основну тенденцію ряду, оскільки при обчисленні зваженої середньої кожному рівню ряду в межах інтервалу згладжування приписується певна вагу, що залежить від відстані до середини інтервалу. При побудові зваженої ковзної середньої на кожній активній ділянці значення центрального рівня замінюється на обчислене, що визначене за формулою зваженої середньої арифметичної. Іншими словами, зважена змінна середня відрізняється від простої ковзної середньої тим, що рівні, що входять в інтервал усереднення, підсумовуються з різними вагами.

Проста ковзна середня враховує всі рівні ряду, що входять до інтервалу згладжування, з рівними вагами, а зважена середня приписує кожному рівню вагу, що залежить від відстані цього рівня до рівня, що стоїть всередині цього інтервалу. Це зумовлене тим, що для простої ковзної середньої в інтервалі згладжування обчислення здійснюють на підставі прямої – поліному першого порядку, а для згладжуванні зваженою ковзною середньою використовують поліноми більш високих порядків, переважно другого або третього. Тому метод простої ковзної середньої можна розглядати як окремий випадок методу зваженої ковзної середньої.

Вагові коефіцієнти визначаються за допомогою методу найменших квадратів (МНК), причому немає необхідності щоразу обчислювати їх заново для рівнів ряду, що входять в інтервал згладжування, так як вони будуть однаковими для кожного його положення. Так як ваги симетричні відносно центрального рівня, то в існуючих таблицях використано символічний запис: наведено ваги лише для половини рівнів інтервалу згладжування, причому виділено вагу, що відноситься до рівня, який стоїть у центрі ділянки згладжування. Для решти рівнів ваги не наводяться, тому що вони можуть бути симетрично відображені.

Основними властивостями вагових коефіцієнтів є такі:

1) ваги завжди симетричні відносно центрального рівня;

2) сума ваг в інтервалі згладжування дорівнює одиниці;

3) додатні та від'ємні значення ваг забезпечують для згладженої кривої можливість відтворювати різні вигини кривої тренду.

Алгоритм згладжування зваженим ковзним середнім з розміром „вікна" – інтервалу згладжування $w = 2k + 1$, яке послідовно зміщується вздовж рівнів ряду і усереднює охоплені ним рівні y_i

$$\widetilde{y}_i = y_1^* + y_2^* + \ldots + y_k^* + \sum_{j=k+1}^{N-2k} \left[\frac{1}{w} \sum_{i=j}^{j+2k+1} \alpha_i y_i \right] + y_{N-k}^* + \ldots + y_{N-1}^* + y_N^* \quad . \qquad \text{(КП.7)}$$

В цих формулах значення y_q^*, де $q = 1, 2, \ldots, k, N-k, \ldots, N-1, N$, на початку та в кінці ряду, обчислюються за відповідними інтерполяційними формулами, α_i – ваги, для яких $\sum_{i=1}^{2k+1} \alpha_i = 1$.

КП.2.4. Властивості рухомого середнього

Ковзне середнє, згладжуючи вихідний ряд, дає уявлення про загальну тенденцію поведінки ряду: його тренд і циклічну компоненту, а тому потрібно знати його властивості, до яких відносять наступні.

1. При застосуванні методу ковзних середніх вибір величини інтервалу згладжування повинен бути зроблений на основі змістовних міркувань і бути прив'язаним до періоду можливо існуючих коливальних процесів. Якщо процедура ковзного середнього використовується для згладжування часового ряду за відсутності будь яких коливань, то найчастіше величину інтервалу згладжування вибирають рівною трьом, п'яти або семи. Чим більшим є інтервал усереднення (згладжування), тим більш гладким виглядає графік тренду.

2. Сусідні члени ряду ковзних середніх сильно корельовані, так як в їх формуванні беруть участь одні й ті ж члени вихідного ряду. Це може призвести, до того, що ряд ковзних середніх може містити циклічні компоненти, які відсутні у вихідному ряді. Це явище носить назву *ефекту Слуцького-Юла*.

3. Як метод усереднення, крім згаданого вище звичайного середнього арифметичного, можна розглядати і зважені ковзні середні, тобто коли значення вихідного ряду в інтервалі згладжування підсумовується з певними вагами. Подібні процедури доцільні, якщо зміна часового ряду в часі носить явно нелінійний характер.

Зараз існують різні таблиці для значень ваг зваженого ковзного середнього. Алгоритми звичайного і зваженого ковзного середнього з врахуванням втрат рівнів на кінцях ряду наведені нижче. Хоча, готових процедур для реалізації цих алгоритмів в середовищі MS Excel немає, проте самі алгоритми досить просто можуть бути реалізовані математичними операціями **Майстра функцій.** Під час реалізації алгоритмів необхідно звернути увагу на індексацію варіант. Приклади застосування простих алгоритмів ковзних середніх наведено нижче.

Лінійне згладжування, коли $\mathbf{w = 3}$.

Наведений алгоритм реалізує ковзне середнє, використовуючи мінімальний інтервал згладжування, причому перша і третя формули обчислюють найбільш відповідні значення втрачені на початку і в кінці ряду.

$$\bar{y}_1 = \frac{5y_1 + 2y_2 - y_3}{6}; \qquad \bar{y}_i = \frac{y_{i-1} + y_i + y_{i+1}}{3}, \qquad \bar{y}_n = \frac{-y_{n-2} + 2y_{n-1} + 5y_n}{6}, \quad \text{(КП.8)}$$

$$i = 2, 3, \ldots, n-1.$$

Результат роботи цього алгоритму відображено графіком на рис. КП.5. для підвищення ефективності згладжування процедуру можна застосувати до вже згладжених рівнів.



Рис. КП.5. Згладжування простим ковзним середнім для $w = 3$.

***Лінійне згладжування для* w = 5.**

В цьому випадку, в результаті згладжування втрачається по два рівні на початку і в кінці часового ряду. Для усунення цього недоліку в алгоритмі використовуються спеціальні формули перерахунку втрачених значень на підставі реальних, тобто включаючи і самі втрачені значення. Для обчислення втрачених значень на початку часового ряду використовують першу і другу формулу, а для обчислення втрачених в кінці – четверту і п'яту формули.

$$\bar{y}_1 = \frac{3y_1 + 2y_2 + y_3 - y_5}{5}, \qquad \bar{y}_2 = \frac{4y_1 + 3y_2 + 2y_3 + y_4}{10}, \qquad \text{(КП.9)}$$

$$\bar{y}_i = \frac{y_{i-2} + y_{i-1} + y_i + y_{i+1} + y_{i+2}}{5}, \qquad \bar{y}_{n-1} = \frac{y_{n-3} + 2y_{n-2} + 3y_{n-1} + 4y_n}{10},$$

$$\bar{y}_n = \frac{-y_{n-4} + y_{n-2} + 2y_{n-1} + 3y_n}{5}, \qquad \textbf{де} \quad i = 3, 4, \ldots, n-2$$

Результат роботи цього алгоритму зображено наступним графіком на рис. КП.6.

Ковзне середнє дозволяє проводити операцію згладжування, як лінійними, так і нелінійними алгоритмами різних модифікацій. В літературі з опрацювання результатів експериментальних досліджень саме ці формули широко використовуються.

Ковзне середнє має найбільш просту інтерпретацію і алгоритм, проте його ефективність з точки зору згладжування не є вельми високою, особливо за малих значень розміру інтервалу w. Тому в багатьох випадках застосовують нелінійне згладжування.

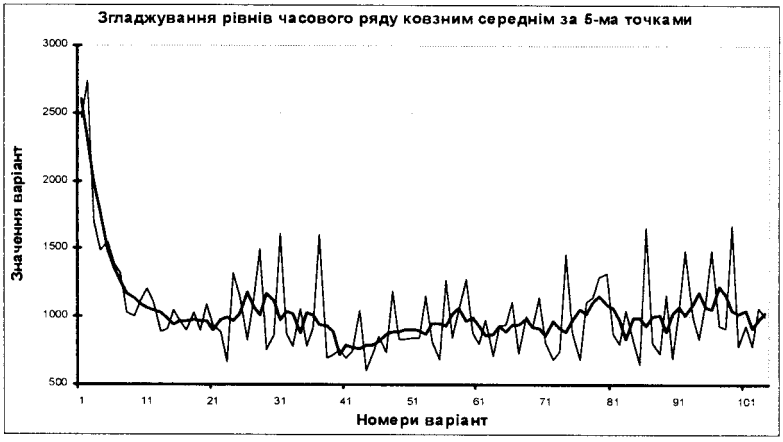


Рис. КП.6. Згладжування ковзним середнім для $w = 5, \quad k = 2$.

Нелінійне згладжування **w = 7.**

У випадку нелінійних змін тренду загальної тенденції використовують нелінійні формули. Наприклад, алгоритм нелінійного згладжування з врахуванням втрачених рівнів на початку і в кінці ряду з величиною інтервалу згладжування має такий вигляд.

$$\bar{y}_1 = \frac{39y_1 + 8y_2 - 4y_3 - 4y_4 + y_5 + 4y_6 - 2y_7}{42},$$

$$\bar{y}_2 = \frac{8y_1 + 19y_2 + 16y_3 + 6y_4 - 4y_5 - 7y_6 + 4y_7}{42},$$

$$\bar{y}_1 = \frac{39y_1 + 8y_2 - 4y_3 - 4y_4 + y_5 + 4y_6 - 2y_7}{42},$$

$$\bar{y}_2 = \frac{8y_1 + 19y_2 + 16y_3 + 6y_4 - 4y_5 - 7y_6 + 4y_7}{42},$$

$$\bar{y}_{n-2} = \frac{y_{n-6} - 4y_{n-5} + 2y_{n-4} + 12y_{n-3} + 19y_{n-2} + 16y_{n-1} - 4y_n}{42},$$

$$\bar{y}_{n-1} = \frac{4y_{n-6} - 7y_{n-5} - 4y_{n-4} + 6y_{n-3} + 16y_{n-2} + 19y_{n-1} + 8y_n}{42},$$

$$\bar{y}_n = \frac{-2y_{n-6} + 4y_{n-5} + y_{n-4} - 4y_{n-3} - 4y_{n-2} + 8y_{n-1} + 39y_n}{42},$$

$$i = 4, 5, \ldots, n-3.$$

Експоненціальне згладжування.

Поряд з простими і зваженими (поліноміальними) середніми для виявлення тенденцій у динаміці часових рядів використовуються експоненціальні середні. Вони формуються під впливом усіх попередніх рівнів ряду таким чином, що більш пізня інформація набуває більшого значення в силу застосування спеціальної системи ваг. В цьому методі згладжене значення визначається лише двома значеннями – біжучим і останнім згладженим рівнем та співвідношенням їхніх ваг, а саме α і $(1-\alpha)$. Важливим параметром і значним недоліком цього методу є оптимальний вибір величини α цього співвідношення.



Рис. КП.7. Результат нелінійного згладжування для $w = 7$.

Суттєвою ознакою застосування експоненціальних середніх є обґрунтування величини параметра згладжування α. Чим він менший, тим більше згладжуються рівні в ряді, який аналізується. Це означає зростання питомої α. Цей метод згладжування має досить широке застосування в прогнозуванні економічних часових рядів. Метод є однопараметричним, а його єдиний параметр α вибирається за наступної умови:

$$\begin{aligned} &0.1 \le \alpha \le 0.3, \\ &\widetilde{y}_0 = y_1, \\ &\widetilde{y}_1 = \alpha y_1 + (1-\alpha)\widetilde{y}_0, \\ &\widetilde{y}_2 = \alpha y_2 + (1-\alpha)\widetilde{y}_1, \\ &\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\ &y_n = \alpha y_n + (1-\alpha)\widetilde{y}_{n-1}. \end{aligned} \qquad \text{(КП.10)}$$

Початкове $\widetilde{y}_0$ значення можна також розрахувати за формулами для ковзного середнього. Загалом алгоритм медіанного згладжування є таким:

$$\widetilde{y}_i = \widetilde{y}_0^* + \sum_{i=1}^{N} \left[\alpha\, y_i + (1-\alpha)\widetilde{y}_{i-1} \right], \qquad \text{(КП.11)}$$

де $\widetilde{y}_0^*$ – екстрапольоване значення, α – параметр згладжування;

Результати цього типу згладжування відображені нижче графічно на рис.КП.8. Аналіз графіка експоненціального згладжування показує, що експоненціальне згладжування дає сильне «запізнення», а тому для часових рядів зі значною дисперсією рівнів воно є мало ефективним. Крім того, результат згладжування сильно залежить від параметра згладжування α, для якого на цей час відсутня методика визначення його величини, і хоча існують деякі підходи, проте усі вони є емпіричними.



Рис. КП.8. Результат експоненціального згладжування.

КП.2.5. Медіанна фільтрація

Медіана відноситься до розподіленого середнього, тобто є значенням ознаки, яка займає місце всередині варіаційного ряду і на відміну від середньої арифметичної, що узагальнює величину показника, залишає значення того показника, який відповідає медіані. Вона є найбільш стійкою характеристикою варіаційного ряду. Як параметр вибірки, вона є однією з числових характеристик її структури та іноді використовується як альтернатива середньому арифметичному значенню елементів вибірки.

Медіанне згладжування є не обчислювальною нелінійною процедурою, оскільки для нього не виконується одна з аксіом лінійності, а саме: медіана суми двох довільних послідовностей не рівна сумі їхніх медіан. Характерною особливістю медіанного згладжування є те, що монотонні ділянки послідовності даних та різкі перепади вона залишає без змін, а для немонотонних ділянок в межах розміру ковзного інтервалу згладжування залишає лише центроване значення рівне їхній медіані, тобто ефективно усуває ті рівні, які порушують монотонність.

Зміст алгоритму медіанного згладжування часових рядів полягає у визначені значення медіани для рівнів інтервалу згладжування. Далі значенням медіани заміняють значення того рівня часового ряду, якому відповідає середина інтервалу згладжування. Іншими словами, медіана відповідає моменту часу t за умови, що межі інтервалу згладжування відповідають моментам часу $[t-k,\ t+k]$, тобто, коли інтервал згладжування є рівним $w = 2k + 1$. Медіанне згладжування повністю усуває поодинокі екстремальні або аномальні значення рівнів, які віддалені один від одного як мінімум на половину величини інтервалу згладжування; зберігає різкі перепади в тенденції

(ковзне середнє та експоненціальне згладжування їх змазує); ефективно усуває поодинокі рівні з дуже великими або дуже малими значеннями, які мають випадковий характер і різко виділяються серед інших рівнів.

Медіанна фільтрація використовується досить рідко, хоча вона є вельми ефективною для часових рядів, рівні яких мають суттєво асиметричний розподіл. Загальним алгоритмом медіанної фільтрації є такий

$$\bar{y}_i = \max\{\min(y_{i-1}, y_i), \min(y_i, y_{i+1}), \min(y_{i-1}, y_{i+1})\}, \qquad \text{(КП.12)}$$

де індекс i приймає значення $i = 2, 3, 4, \ldots, n-1$.

Результати застосування медіанної фільтрації відображені на наступному графіку (рис.КП.9).



Рис. КП.9. Ефективність застосування медіанної фільтрації.

При повторному застосуванні медіанного згладжування (за постійного розміру інтервалу вікна) перепади в поведінці рівнів зберігаються до тих пір, поки є зміни в згладжуваному ряді. Основними властивостями медіани є такі: кількість додатних відхилень від медіани є рівна кількості від'ємних та сума абсолютних відхилень варіант вибірки відносно медіани є мінімальною. Вперше ковзна оцінка медіани, тобто медіанне згладжування, для аналізу часових рядів була запропонована і застосована Тьюки в 1971 р. Він також вказував, що медіанне згладжування зберігає в часових рядах більш різкі зміни їх рівня (тобто перепади). Початкове $\bar{y}_1$ і кінцеве $\bar{y}_n$ значення обчислюють за формулами для ковзного середнього. Такий розрахунок можна провести як до виконання операції медіанної фільтрації, так і після неї. Крім того, як початковий і кінцевий рівні можна використати початкове y_1 і кінцеве y_n оригінальні значення. В середовищі пакету MS Excel значення для медіанної фільтрації, тобто реалізацію цього алгоритму можна здійснити, використовуючи процедуру знаходження медіани, тобто

$$\bar{y}_i = \textbf{МЕДІАНА}(y_{i-1}, y_i, y_{i+1}). \qquad \text{(КП.13)}$$

Медіанну фільтрацію переважно здійснюють за трьома рівнями, тобто інтервал згладжування має $w = 3$, проте можна використовувати будь-які непарні додатні цілі числа.

КП.2.6. Нормування часових послідовностей

Якщо, деяку якість досліджуваного об'єкта необхідно описати числом, то таке число x формується як сума балів. Нехай, це число змінюється від деякого мінімального значення $x_{\min}$ (відображення відсутності цієї якості) до деякого максимального значення $x_{\max}$ (найвищий ступінь прояву її наявності, прояву тощо). Отримання такого числа дає підстави для вирішення проблеми порівняння двох об'єктів, але тільки за показником, якому відповідає це число. Проте, треба завжди пам'ятати, в яких межах змінюється показник. Найбільш часто застосовується лінійне перетворення, яке полягає в тому, що значення рівнів часового ряду приводять до інтервалу значень $[0, 1]$, використовуючи такий алгоритм

$$y_i'' = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}}, \qquad \text{(КП.14)}$$

де y_i'' – нормоване значення, y_i – біжуче значення рівня, $y_{\min}$ та $y_{\max}$ – найменше та найбільше значення рівнів цього часового ряду.

Нормування часових послідовностей в першу чергу дає можливість порівняти показники, що отримані для різних об'єктів, оскільки при такому нормуванні структура рядів (пропорційність між рівнями в рядах) залишається незмінною. Це дає можливість порівнювати обчислені показники та побудовані на таких даних моделі.

КП.2.7. Критерії ефективності згладжування часових рядів

Використання того чи іншого методу згладжування залежить, з одного боку від забезпечення вимог самої задачі, а з іншого – від структури самого ряду. Тому, важливе значення має визначення ефективності існуючих методів. Для цього використовують різні критерії, серед яких найбільш адекватними цій задачі є: критерій поворотних точок та коефіцієнт кореляції між вихідним та згладженим рядами.

Критерії поворотних точок. Для оцінювання ефекту згладжування пропонується використати критерій поворотних точок, зміст якого полягає в звичайному підрахунку рівнів, значення яких є більшими або меншими ніж два сусідні. Цей критерій легко реалізується в Excel.

Коефіцієнт кореляції. Для оцінювання тісноти зв'язку між вихідним (оригінальним рядом) та згладженим використовують коефіцієнт кореляції. Для його знаходження в Excel є спеціальна опція в надбудові «Аналіз даних».

Формула для обчислення поворотних точок

=ЕСЛИ((I3>I2);(I3>I4);ИЛИ(ЕСЛИ((I3<I2);(I3<I4))))

КП.2.8. Формули для зваженого ковзного середнього

В задачах з використанням простого ковзного середнього обчислюється середнє значення рівнів, що входять в інтервал згладжування. Ефект в цьому випадку, навіть при використанні великих інтервалів згладжування, є не дуже значним. Тому застосовують повторне згладжування, збільшуючи для кожного повторного підходу розмір вікна. В цьому випадку ефект є суттєвим.

В таблицях наведені значення ваг для ковзних середніх. Таблиці розроблені різними авторами і дають різний ефект згладжування. Наприклад, для простої ковзної середньої з розміром інтервалу згладжування $w=5$ формули для розрахунку втрачених рівнів мають вигляд, поданий наступними виразами (КП.9). Іншими словами, приведені таблиці з формулами з Кендела реалізують прості ковзні середні, оскільки для інтервалу згладжування всередині часового ряду, що відповідає середньому стовпчику таблиці, всі ваги рівні 1.

Згладжування формулами з Кендела

N=3

1	2	n
5	1	-1
2	1	2
-1	1	5
6	3	6
0,913	0,577	0,913

N=5

1	2	3	n-1	n
3	4	1	0	-1
2	3	1	1	0
1	2	1	2	1
0	1	1	3	2
-1	0	1	4	3
5	10	5	10	5
0,775	0,548	0,447	0,548	0,775

N=7

1	2	3	4	n-2	n-1	n
13	5	7	1	1	-1	-5
10	4	6	1	2	0	-2
7	3	5	1	3	1	1
4	2	4	1	4	2	4
1	1	3	1	5	3	7
-2	0	2	1	6	4	10
-5	-1	1	1	7	5	13
28	14	28	7	28	14	28
0,681	0,535	0,423	0,378	0,423	0,535	0,681

N=9

1	2	3	4	5	n-3	n-2	n-1	n
17	56	22	32	1	8	-2	-16	-7
14	47	19	29	1	11	1	-7	-4
11	38	16	26	1	14	4	2	-1
8	29	13	23	1	17	7	11	2
5	20	10	20	1	20	10	20	5
2	11	7	17	1	23	13	29	8
-1	2	4	14	1	26	16	38	11
-4	-7	1	11	1	29	19	47	14
-7	-16	-2	8	1	32	22	56	17
45	180	90	180	9	180	90	180	45
0,615	0,511	0,422	0,357	0,333	0,357	0,422	0,511	0,615

N=11

1	2	3	4	5	6	5	4	3	2	1
7	15	25	10	15	1	5	0	-5	-5	-3
6	13	22	9	14	1	6	1	-2	-3	-2
5	11	19	8	13	1	7	2	1	-1	-1
4	9	16	7	12	1	8	3	4	1	0
3	7	13	6	11	1	9	4	7	3	1
2	5	10	5	10	1	10	5	10	5	2
1	3	7	4	9	1	11	6	13	7	3
0	1	4	3	8	1	12	7	16	9	4
-1	-1	1	2	7	1	13	8	19	11	5
-2	-3	-2	1	6	1	14	9	22	13	6
-3	-5	-5	0	5	1	15	10	25	15	7
22	55	110	55	110	11	110	55	110	55	22
0,564	0,486	0,416	0,357	0,316	0,302	0,316	0,357	0,416	0,486	0,564

N=13

1	2	3	4	5	6	7	6	5	4	3	2	1
25	44	19	32	13	20	1	8	1	-4	-5	-16	-11
22	39	17	29	12	19	1	9	2	-1	-3	-11	-8
19	34	15	26	11	18	1	10	3	2	-1	-6	-5
16	29	13	23	10	17	1	11	4	5	1	-1	-2
13	24	11	20	9	16	1	12	5	8	3	4	1
10	19	9	17	8	15	1	13	6	11	5	9	4
7	14	7	14	7	14	1	14	7	14	7	14	7
4	9	5	11	6	13	1	15	8	17	9	19	10
1	4	3	8	5	12	1	16	9	20	11	24	13
-2	-1	1	5	4	11	1	17	10	23	13	29	16
-5	-6	-1	2	3	10	1	18	11	26	15	34	19
-8	-11	-3	-1	2	9	1	19	12	29	17	39	22
-11	-16	-5	-4	1	8	1	20	13	32	19	44	25
91	182	91	182	91	182	13	182	91	182	91	182	91
0,524	0,406	0,406	0,355	0,314	0,287	0,277	0,287	0,314	0,355	0,406	0,463	0,524

N=15

1	2	3	4	5	6	7	8	7	6	5	4	3	2	1
29	91	161	35	119	49	77	1	35	7	-7	-7	-49	-35	-13
26	82	146	32	110	46	74	1	38	10	2	-4	-34	-26	-10
23	73	131	29	101	43	71	1	41	13	11	-1	-19	-17	-7
20	64	116	26	92	40	68	1	44	16	20	2	-4	-8	-4
17	55	101	23	83	37	65	1	47	19	29	5	11	1	-1
14	46	86	20	74	34	62	1	50	22	38	8	26	10	2
11	37	71	17	65	31	59	1	53	25	47	11	41	19	5
8	28	56	14	56	28	56	1	56	28	56	14	56	28	8
5	19	41	11	47	25	53	1	59	31	65	17	71	37	11
2	10	26	8	38	22	50	1	62	34	74	20	86	46	14
-1	1	11	5	29	19	47	1	65	37	83	23	101	55	17
-4	-8	-4	2	20	16	44	1	68	40	92	26	116	64	20
-7	-17	-19	-1	11	13	41	1	71	43	101	29	131	73	23
-10	-26	-34	-4	2	10	38	1	74	46	110	32	146	82	26
-13	-35	-49	-7	-7	7	35	1	77	49	119	35	161	91	29
120	420	840	210	840	420	840	15	840	420	840	210	840	420	120
0,492	0,442	0,395	0,352	0,314	0,285	0,265	0,258	0,265	0,285	0,314	0,352	0,395	0,442	0,492

Згладжування формулами з Полларда. Застосування цих формул продемонструємо таким прикладом. Нехай маємо фрагмент таблиці з рівнями деякого часового ряду. Обчислимо для рівня №20 його згладжене значення за формулою, ваги якої приведені в стовпчику 9, тобто для величини інтервалу згладжування $w = 9$.

Номер рівня	15	16	17	18	19	**20**	21	22	23	24	25
Значення рівня	859	590	538	823	619	**564**	708	569	504	976	567

В межі такого інтервалу включені рівні з такими номерами: 16, 17, 18, 19, 20, 21, 22, 23, 24. Тоді ваги для стовпчика 9 розподіляться в такий спосіб:

w_0^9 = 0,33114 для рівня № 20,

w_1^9 = 0,266557 для рівнів № 19 і № 21,

w_2^9 = 0,11847 для рівнів № 18 і № 22,

w_3^9 = -0,00987 для рівнів № 17 і № 23,

w_4^9 = -0,04072 для рівнів № 16 і № 24,

а сама формула матиме такий вигляд

$$y_{20} = (-0{,}04072) \times 590 + (-0{,}00987) \times 538 + 0{,}11847 \times 823 + 0{,}266557 \times 619 + 0{,}33114 \times \mathbf{564} + 0{,}266557 \times 708 + 0{,}11847 \times 569 + (-0{,}00987) \times 504 + (-0{,}04072) \times 976 = \tilde{y}_{20}.$$

	3	**5**	**7**	**9**	**11**	**13**	**15**
w_0	0,78082	0,559441	0,412587	0,33114	0,277945	0,240057	0,211541
w_1	0,10959	0,293706	0,293706	0,266557	0,238693	0,214337	0,193742
w_2		-0,07343	0,058741	0,11847	0,141267	0,147356	0,145904
w_3			-0,05874	-0,00987	0,035723	0,065492	0,082918
w_4				-0,04072	-0,02679	0	0,024027
w_5					-0,02786	-0,02786	-0,01413
w_6						-0,01935	-0,0245
w_7							-0,01373

Як видно цієї формули, ваги розташовані симетрично відносно середини інтервалу згладжування.

КП.3. Кореляційний аналіз часових послідовностей

У практичній діяльності системного аналітика часто зустрічаються ситуації, в яких оцінку однієї з властивостей об'єкта необхідно здійснювати з врахуванням оцінки іншої властивості. У цьому випадку виникає необхідність визначення взаємного впливу властивостей. Закономірності такого впливу досить складно описати математичними моделями. В подібних ситуаціях використовують кореляційну оцінку показників, встановлюючи елемент якісної, експертної оцінки впливу одного показника на інший. Метою дослідника при вирішенні зазначеної задачі є не тільки знаходження кореляційної залежності між двома властивостями об'єкта, а й отримання якісної (експертної) оцінки впливу однієї властивості на іншу.

Під кореляційним аналізом розуміють групу методів, що дозволяють виявляти наявність і ступінь взаємозв'язку між кількома параметрами, що змінюються

випадковим чином. Міра такого взаємозв'язку оцінюється спеціальними числовими характеристиками, а також їх статистиками, що визначають ступінь близькості цього взаємозв'язку до функціонального, який може існувати між параметрами, що володіють детермінованим характером зміни.

Кореляційний зв'язок з'являється, коли одному і тому ж значенню аргументу (незалежної змінної) відповідає низка значень функції (залежної змінної). Тоді зв'язок виявляється у вигляді тенденції зміни середніх значень функції залежно від змін аргументу. Цим кореляційний зв'язок відрізняється від функціонального, який виникає у разі, коли заданому значенню аргументу відповідає цілком певне значення функції. Кореляційний зв'язок є неповним, оскільки залежність між функцією і аргументом в кожному конкретному випадку схильна до впливу з боку інших чинників (які найчастіше мають мінливий характер).

Найбільш повно в статистиці розроблена методологія парної кореляції, що розглядає вплив варіації однієї факторної ознаки на результатну.

Дослідження парної кореляції здійснюється на основі кореляційного аналізу, передбачає послідовне вирішення низки завдань:

- виявлення зв'язку;
- опис зв'язку в табличній і графічній формах;
- вимірювання тісноти зв'язку;
- формулювання висновків про характер існуючого зв'язку.

Основні завдання кореляційного аналізу – це визначення і вираження форми аналітичної залежності результативної ознаки у від факторних ознак x_i.

Відмінною рисою кореляційного аналізу є вимірювання тісноти зв'язку між y та x. Його основними числовими характеристиками є коефіцієнт кореляції і кореляційне відношення.

Виділяють такі етапи кореляційного аналізу:

– виявлення взаємозв'язку між ознаками;

– визначення форми зв'язку;

- визначення сили (тісноти) і напрямку зв'язку .

Інтерпретуючи результати кореляційного аналізу потрібно врахувати, що коефіцієнт кореляції є статистичним показником, який не вказує на те, що досліджувані величини знаходяться в причинно-наслідковому зв'язку. Тому, будь-яке трактування кореляційної залежності повинно ґрунтуватися на інформації про суть і характер досліджуваних експериментальних даних та процесів, яким вони відповідають. До переваг кореляційного аналізу можна віднести можливість створення нового правила взаємодії функцій одна з одною, а також оцінку взаємодії функцій, які отримані невідомим шляхом. Недоліками є те, що всі результати, отримані за допомогою цієї методики можна використовувати лише в області цього дослідження або близької до нього. Після виявлення стохастичних зв'язків між досліджуваними змінними величинами дослідник приступає до математичного опису виявленої ним і цікавої для нього залежності.

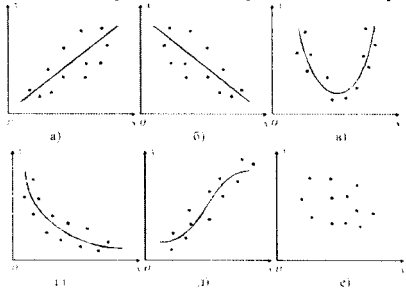
КП.3.1. Кореляційне поле

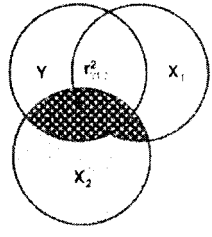
У своїй практиці дослідник часто стикається з необхідністю встановлення факту існування функціональних чи інших залежностей між отриманими експериментальними даними. Проте, вже на підставі візуального аналізу поля кореляції

можна висунути гіпотезу щодо існуючого зв'язку між усіма можливими значеннями X та Y: лінійний або нелінійний, сильний, слабкий або відсутній. Переважно такий зв'язок є випадковим. Графічне подання взаємозв'язку між двома досліджуваними послідовностями називається *кореляційним полем* або *полем кореляції* або *діаграмою розсіювання.* Графічний метод забезпечує наочне зображення форми зв'язку між цими послідовностями. Для цього, в прямокутній системі координат будують графік – по осі ординат відкладають індивідуальні значення однієї послідовності, вибраної в якості результативної ознаки Y, а по осі абсцис – індивідуальні значення іншої – факторної ознаки X. Саме сукупність точок результативної і факторної ознак називається *полем кореляції.*

Отже, кореляційне поле – це сукупність точок у прямокутній системі координат, абсциса кожної з яких відповідає значенню факторної ознаки (x), а ордината – значенню результативної ознаки (y) певної одиниці спостереження. Кількість точок на графіку відповідає кількості одиниць спостереження. Використовується для аналізу наявності та характеру (напряму) зв'язку між результатами двох вибірок спостережень. Розміщення точок на графіку свідчить про наявність і напрям зв'язку. Загалом, локалізація точок кореляційного поля вказує на наявність прямого, оберненого зв'язку між ознаками або його відсутність, а також на форму лінії регресії (рис. КП.10). Розміщення точок на кореляційному полі дозволяє судити про характер залежності, наприклад: лінійна, параболічна, гіперболічна, логічна, логарифмічна, експонентна, показникова або відсутність залежності.

Кореляційні зв'язки можна вивчати на якісному рівні з діаграм розсіяння емпіричних значень змінних X та Y і відповідним чином їх інтерпретувати. Так, наприклад, якщо підвищення рівня однієї змінною супроводжується підвищенням рівня іншої, то йдеться про *додатну* кореляцію або прямий зв'язок (рис. КП.11,). Якщо ж зростання однієї змінної супроводжується зниженням значень іншої, то маємо справу з *від'ємною* кореляцією або оберненим зв'язком (рис. КП.12). Нульовою або відсутньою називається кореляція за відсутності зв'язку змінних (рис. КП.10). Проте нульова загальна кореляція може свідчити лише про відсутність *лінійної* залежності, а не взагалі про відсутність будь якого *статистичного* зв'язку. При цьому функцію, графік якої відповідає розміщенню точок називають теоретичною лінією регресії. Для вибору тієї чи іншої форми кореляційної залежності, треба порівняти уявну емпіричну лінію регресії з графіками відомих функцій. Методи кореляційного аналізу широко застосовуються для виявлення та опису стохастичних залежностей між випадковими величинами, якими переважно є зібрані або експериментальні дані.





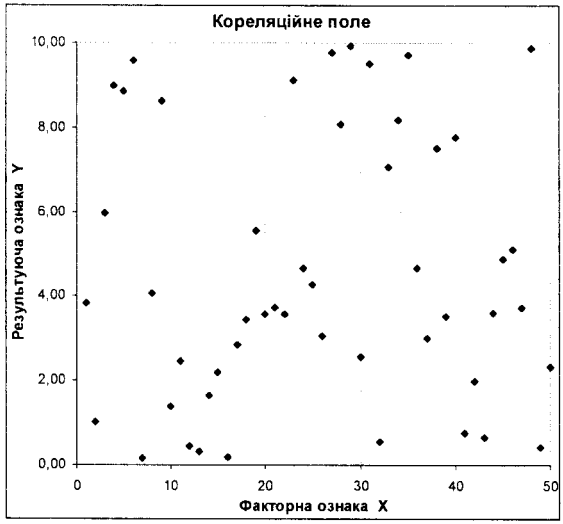


Рис. КП.10. Візуальна оцінка характеру зв'язку вказує на його відсутність.

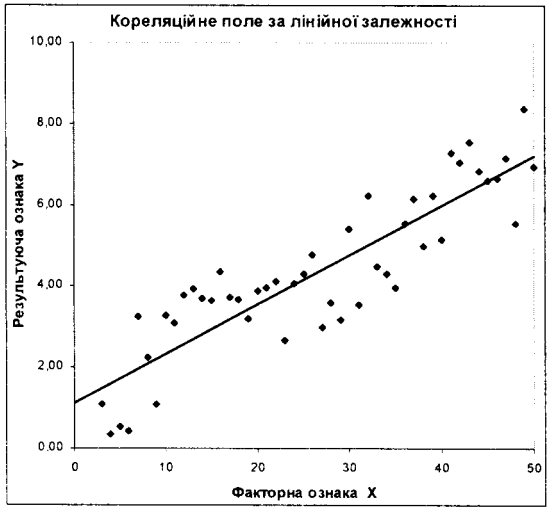


Рис. КП.11. Прямий лінійний зв'язок.

Для експериментального вивчення залежності між випадковими величинами Y і X проводять деяку кількість незалежних дослідів. Результат i-го досліду дає пару значень (x_i, y_i), де $i = 1, 2, \ldots, n$. Отже, досліджувані послідовності можна подати так: $X = \langle x_1, x_2, \ldots, x_n \rangle$; $Y = \langle y_1, y_2, \ldots, y_n \rangle$. Якщо послідовності подати у вигляді функцій, що залежать від одного аргументу, то, провівши кореляційний аналіз, можна

встановити взаємний вигляд зв'язку між ними та його величину, при цьому обсяг дани
має бути однаковий. Про наявність чи відсутність кореляції між двома випадковим
величинами якісно можна судити з вигляду поля кореляції, відобразивш
експериментальні пари точок на координатну площину.

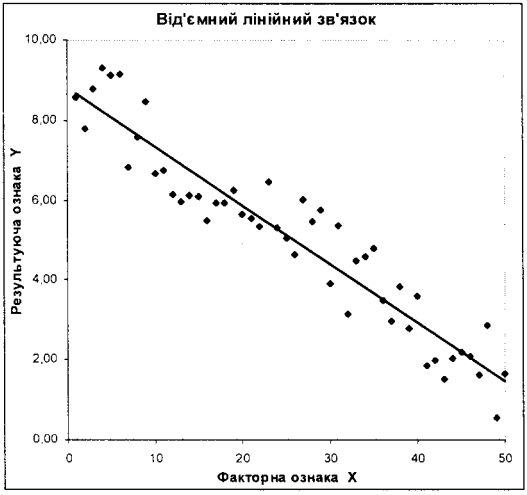


Рис. КП.12. Обернений лінійний зв'язок.

КП.3.2. Коефіцієнт кореляції

Для кількісної оцінки тісноти зв'язку служить вибірковий *коефіцієнт кореляц*
Вибірковий коефіцієнт кореляції r за абсолютною величиною не перевищує одиниц
Для незалежних випадкових величин коефіцієнт кореляції дорівнює нулю, але він мож
бути рівний нулю для деяких залежних величин, які при цьому називаютьс
некорельованими.

Для випадкових величин, що мають нормальний розподіл, відсутність кореляц
означає і відсутність будь-якої залежності.

Вибірковий коефіцієнт кореляції не змінюється при зміні початку відліку
масштабу величин . Коефіцієнт кореляції характеризує не довільну залежність, а тільк
лінійну. Лінійна ймовірнісна залежність випадкових величин полягає в тому , що пр
зростанні однієї випадкової величини інша має тенденцію зростати (або спадати)
лінійним законом.

Коефіцієнт кореляції характеризує ступінь тісноти лінійної залежності.
загальному випадку, коли величини X та Y пов'язані деякою стохастичною
залежністю, коефіцієнт кореляції може мати значення в межах $-1 \le r \le +1$.

Відзначимо властивості коефіцієнта кореляції:

- коефіцієнт парної кореляції обчислюється для кількісних ознак;
- коефіцієнт кореляції симетричний, тобто не змінюється, якщо X та
  поміняти місцями;
- коефіцієнт кореляції є величиною безрозмірною;
- коефіцієнт кореляції не змінюється при зміні одиниць виміру ознак X та Y

- величина коефіцієнта кореляції не змінюється від додавання до X та Y невипадкових доданків;
- величина коефіцієнта кореляції не змінюється від множення X та Y на додатні числа;
- якщо одну з величин, не змінюючи іншу, помножити на –1, то на –1 треба помножити і коефіцієнт кореляції.

Схема застосування кореляційного аналізу з практичною метою приблизно така: є кілька параметрів, які спостерігаються протягом деякого проміжку часу, про які, за результатами спостережень (або з будь-яких апріорних міркувань), можна припустити, що вони можуть бути взаємопов'язані будь-яким чином.

Обчислення коефіцієнта кореляції здійснюють за такою формулою

$$r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i-1}^{n} (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i-1}^{n} (y_i - \bar{y})^2}}. \qquad \text{(КП.15)}$$

В статистичній літературі рекомендують використовувати для обчислення коефіцієнта кореляції наступний вираз

$$r_{xy} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2 \right] \left[n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i \right)^2 \right]}}. \qquad \text{(КП.16)}$$

В цьому випадку немає потреби обчислювати відхилення біжучих значень від середньої величини, а це виключає помилки в розрахунках при заокругленні середніх величин.

Коефіцієнт кореляції r_{xy} є випадковою величиною, оскільки обчислюється для випадкових величин. Стосовно нього можна висувати і перевіряти такі гіпотези:

1. Коефіцієнт кореляції значимо відрізняється від нуля, тобто між величинами є взаємний зв`язок. Для перевірки цієї гіпотези обчислюють тестову статистику за такою формулою

$$\xi = \left(0.5 \cdot \ln\left(\frac{1 + r_{xy}}{1 - r_{xy}} \right) - \frac{|r_{xy}|}{2(n-1)} \right) \sqrt{n-3}. \qquad \text{(КП.17)}$$

Обчислене значення ξ порівнюється з табличним значенням коефіцієнта Стьюдента $t(p = 0.95,\ f = \infty) = 1.96$. Якщо тестова статистика є більшою за табличне значення, то коефіцієнт кореляції значимо відрізняється від нуля. З формули випливає, що, чим більше вимірів n, тим більшою є тестова статистика, а, отже, коефіцієнт кореляції значимо відрізняється від нуля.

2. Значення коефіцієнта кореляції є значимим, якщо обчислений коефіцієнт детермінації перевірити за допомогою критерію Стьюдента

$$t_{розр} = r_{xy} \cdot \sqrt{\frac{n-k-1}{1-r_{xy}^2}} \,. \qquad \text{(КП.18)}$$

Табличне значення t-критерію Стьюдента за довірчої ймовірності 0,95 і для числа ступенів свободи $\gamma = (n-k-1)$ порівнюється з обчисленим значенням. Якщо обчислене значення перевищує табличне, то коефіцієнт кореляції визнається значимим. Оцінка значущості коефіцієнта парної кореляції з використанням t-критерію Стьюдента може бути обчислена за такою формулою

$$t_{розр} = \sqrt{\frac{r_{xy}^2}{1-r_{xy}^2} \cdot (n-2)} \,. \qquad \text{(КП.19)}$$

Обчислене за цією формулою значення t порівнюється з критичним значенням t-критерію, яке береться з таблиці значень t Стьюдента з врахуванням заданого рівня значущості і числа ступенів свободи (n-2).

3. Відмінність між двома коефіцієнтами кореляції є значимою, якщо тестова статистика

$$\zeta = 0.5 \cdot \ln\left(\frac{(1+r_1)(1-r_2)}{(1-r_1)(1+r_2)}\right) \cdot \frac{1}{\sqrt{\frac{1}{n_1-3}+\frac{1}{n_2-3}}}, \qquad \text{(КП.20)}$$

яку також порівнюють з табличним значенням $t(p,\infty) = 1.96$

Сам по собі коефіцієнт кореляції не має змістовної інтерпретації. Проте його квадрат $R = r^2$, який називають *коефіцієнтом детермінації* (позначається R і зазвичай виражається у %), має простий зміст – це показник того, наскільки зміни залежної ознаки пояснюються змінами незалежної.

З визначення коефіцієнта детермінації випливає, що він приймає значення в діапазоні від 0 % до 100 %.

Якщо дві змінні функціонально лінійно залежні (точки на кореляційному полі лежать на одній прямій), то можна сказати, що зміна однієї з них повністю пояснюється зміною іншої, а це якраз той випадок, коли коефіцієнт детермінації дорівнює 100 % (при цьому коефіцієнт кореляції може дорівнювати як +1, так і – 1).

Чим вище за модулем (за абсолютною величиною) значення коефіцієнта кореляції, тим сильніший зв'язок між ознаками.

Прийнято вважати, що коефіцієнти кореляції, які за модулем більше 0.7, вказують на сильний зв'язок (при цьому коефіцієнти детермінації > 50%, тобто одна ознака визначає іншу більш, ніж наполовину).

Коефіцієнти кореляції, які по модулю менше 0,7 , але більше 0.5 , говорять про зв'язок середньої сили (при цьому коефіцієнти детермінації менше 50 % , але більше 25%). Нарешті, коефіцієнти кореляції, які по модулю менше 0,5 , говорять про слабкий зв'язок (при цьому коефіцієнти детермінації менше 25%).

КП.3.3. Кореляційне відношення

При відхиленні парної статистичної залежності від лінійної коефіцієнт кореляції втрачає свій сенс як характеристика ступеня тісноти зв'язку. У цьому випадку користуються таким виміром зв'язку як ***кореляційне відношення***. Кореляційне відношення застосовують в тих випадках, коли:

- між парою досліджуваних ознак відзначається нелінійна залежність;
- характер вибіркових даних (кількість, щільність розташування на кореляційному полі) допускає, по-перше, їх групування по осі ординат, по-друге, можливість підрахунку "окремих" математичних очікувань всередині кожного інтервалу групування.

Послідовність методики обчислення кореляційного відношення. Нехай між X та Y існує нелінійна залежність, зображена на рис. КП.13.

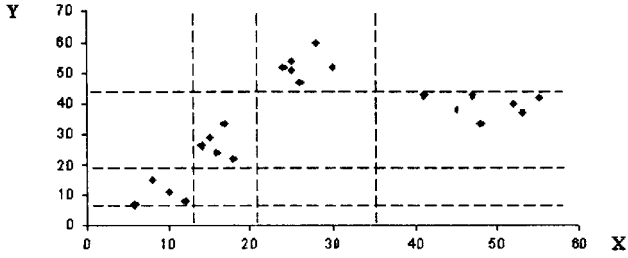


Рис. КП.13. Нелінійна залежність між компонентами X і Y.

1. Розіб'ємо кореляційне поле за змінною X на L інтервалів групування, що не перетинаються, які можуть мати різну довжину.

2. Знайдемо "частинні" математичні сподівання відгуку Y в кожній з L виділених груп

$$\overline{m}_{Y_j} = \frac{1}{n_j} \cdot \sum_{k=1}^{n_j} y_j \,, \qquad \text{(КП.21)}$$

де $j = \overline{1, L}$, $k = \overline{1, n_j}$, n_j – кількість елементів вибірки в j-му інтервалі групування.

3. Знайдемо математичне сподівання частинних групувань відгуків, використовуючи "частинні" $\overline{m}_{Y_j}$

$$\overline{m}_Y = \frac{1}{n} \cdot \sum_{j=1}^{L} n_j \cdot \overline{m}_{Y_j} \,. \qquad \text{(КП.22)}$$

4. Обчислимо групову дисперсію вихідної змінної Y

$$\sigma^2_{\overline{m}_Y} = \frac{1}{n} \cdot \sum_{j=1}^{L} n_j \cdot \left(\overline{m}_{Y_j} - \overline{m}_Y\right)^2 , \qquad \text{(КП.23)}$$

і дисперсію, яка отримана за не згрупованим відгуком

$$\overline{\sigma}^2_Y = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(Y_i - \overline{m}_Y\right)^2 . \qquad \text{(КП.24)}$$

5. Кореляційне відношення для залежної змінної Y і незалежної змінної X може бути отримане з відношення

$$\overline{\rho}_{Y\cdot X} = \frac{\overline{\sigma}_{\overline{m}_Y}}{\overline{\sigma}_Y}. \qquad \text{(КП.25)}$$

КП.3.4. Властивості кореляційного відношення

Кореляційне відношення не має властивості симетрії, тобто $\overline{\rho}_{Y\cdot X} \neq \overline{\rho}_{X\cdot Y}$. Крім того, $\overline{\rho}_{Y\cdot X}$ не від'ємне, оскільки вважається, що воно є результатом добування кореня квадратного з $\rho^2_{Y\cdot X}$. Кореляційне відношення $0 \le \rho_{Y\cdot X} \le 1$.

З $\rho_{Y\cdot X} = 1$ випливає, що між Y та X існує однозначна функціональна залежність. Протилежне твердження в загальному випадку не вірне.

Відсутність кореляційного зв'язку між Y та X означає, що умовні середні $\overline{m}_{Y_j}$ зберігають від групи до групи постійне значення, що дорівнює загальному середньому $\overline{m}_Y$. Тому, необхідно також відзначити, що між $\rho_{Y\cdot X}$ і $\rho_{X\cdot Y}$ немає будь-якої певної залежності. Некорельованість Y від X не означає некорельованість X від Y. І, нарешті, дослідження показали, що $\rho_{YX} \ge |r_{YX}|$. Умова рівності виконується у випадку лінійної залежності X та Y. Всі зауваження щодо змістовної інтерпретації $\rho_{Y\cdot X}$ аналогічні інтерпретації значень отриманих даних.

КП.3.5. Кореляційна матриця

У випадку великої кількості спостережень, коли коефіцієнти кореляції необхідно послідовно обчислювати для декількох вибірок, для зручності отримані коефіцієнти зводять в таблиці, які називають *кореляційними матрицями*.

Кореляційна матриця – це квадратна таблиця, в якій на перетині відповідних рядка і стовпця знаходиться коефіцієнт кореляції між відповідними параметрами.

У MS Excel для обчислення кореляційних матриць використовується процедура **КОРЕЛЯЦІЯ** з пакету **Аналіз даних**. Процедура дозволяє отримати кореляційну матрицю, що містить коефіцієнти кореляції між різними параметрами.

Для реалізації процедури необхідно:

– виконати команду **Сервіс** → **Аналіз даних** ;

– в списку **Інструменти аналізу** вибрати рядок **КОРЕЛЯЦІЯ** і натиснути кнопку ОК;

– в діалоговому вікні вказати **Вхідний інтервал**, тобто ввести посилання на клітинки, які містять аналізовані дані. Вхідний інтервал повинен містити не менше двох стовпців;

– в розділі **Групування** перемикач встановити відповідно до введених даних, тобто за стовпцями чи за рядками;

– вказати **Вихідний інтервал**, тобто ввести посилання на клітинку, для якої будуть виведені результати аналізу. Розмір вихідного діапазону буде визначений автоматично, і на екран буде виведене повідомлення у разі можливого накладення вихідного діапазону на вхідні дані. Натиснути кнопку ОК .

В результаті у вихідний діапазон буде виведена кореляційний матриця, в якій на перетині кожних рядка і стовпця знаходиться коефіцієнт кореляції між відповідними

параметрами. Значення коефіцієнтів кореляції рівне +1, що розміщені вздовж діагоналі, вказує на те, що кожен стовпець у вхідному діапазоні повністю корелює сам з собою.

В процесі інтерпретації кожен коефіцієнт кореляції між відповідними параметрами розглядається окремо. Зазначимо, що хоча в результаті буде отримана трикутна матриця, кореляційна матриця є симетричною, оскільки в порожніх клітинках в правій верхній половині таблиці знаходяться ті ж самі коефіцієнти кореляції, що і в нижній лівій (симетрично розташовані відносно діагоналі.

КП.3.6. Автокореляція

Для вивчення природи динаміки рівнів, які відповідають різним часовим інтервалам часто використовується *поняття автокореляції*, яка характеризує не тільки взаємозалежність рівнів одного й того ж ряду, що відносяться до різних моментів спостережень, але і ступінь стійкості розвитку процесу в часі. Корельованість рівнів часових послідовностей із застосуванням парного коефіцієнта кореляції правильно показує тісноту зв'язку лише в тому випадку, якщо в кожній з них відсутня автокореляція. Існування залежності між попередніми і наступними рівнями часової послідовності в статистичній літературі називають *автокореляцією*.

Застосування методів класичної теорії кореляції в часових послідовностях пов'язано з тим, що для більшості часових послідовностей має місце залежність наступних рівнів від попередніх. Так як методика кореляційного аналізу грунтується на принципі статистичної незалежності даних, наявність автокореляції може призвести до помилкового визначення суттєвості та довірчих меж коефіцієнтів регресії і до інших наслідків, що ставить під сумнів результати аналізу. Тому, якщо аналіз проводиться за даними за різні періоди, необхідно переконатися у відсутності автокореляції в досліджуваних рядах динаміки.

Тому, перш ніж корелювати такі послідовності за рівнями, необхідно перевірити кожну з них на наявність або відсутність в них автокореляції.

КП.3.7. Автокореляція в часових рядах

Явище автокореляції має місце у тих випадках, коли кореляційний аналіз проводиться за даними за певні періоди, може виявитися явище автокореляції, тобто зв'язок між даними за попередні і подальші періоди. За наявності тенденції і циклічних коливань значення кожного наступного рівня ряду залежать від попередніх значень. Кореляційну залежність між послідовними рівнями часового ряду називають *автокореляцією* рівнів ряду. Кількісно її можна виміряти за допомогою лінійного коефіцієнта кореляції між рівнями вихідного часового ряду y_t і рівнями цього ряду, зсунутими на кілька кроків у часі $y_{t-\tau}$.

Ступінь тісноти статистичного зв'язку між рівнями часового ряду, зсунутими на τ одиниць часу визначається величиною коефіцієнта кореляції $r(\tau)$. Оскільки, $r(\tau)$ вимірює тісноту зв'язку між рівнями одного і того ж часового ряду, його прийнято називати *коефіцієнтом автокореляції*. При цьому довжину часового зсуву називають зазвичай лагом (τ). Коефіцієнт автокореляції обчислюється за безпосередніми даними рядів динаміки, коли фактичні рівні одного ряду розглядають як значення факторної ознаки, а рівні цього ж ряду зсунуті на один період, приймають за результативну ознаку (цей зсув називається лагом).

Число періодів, за якими розраховується коефіцієнт автокореляції, називають

лагом. Із збільшенням лага число пар значень, за якими розраховується коефіцієнт автокореляції, зменшується. Максимальний лаг повинен бути не більше (n/4). Примітка: щоб уникнути плутанини, слід звернути увагу на порядок, за яким буде проводитися зміна рівнів, а саме, вниз або вгору. Відповідно і в формулах за різними джерелами, ряд із зсувом відображають y_{t-1} або y_{t+1}.

Коефіцієнт автокореляції характеризує тісноту тільки лінійного зв'язку поточного й аналізованого рівнів ряду. Тому за коефіцієнтом автокореляції можна судити про наявність лінійної (або близькою до лінійної) тенденції. Для деяких часових рядів, які мають сильну нелінійну тенденцію (наприклад, параболу або експоненту), коефіцієнт автокореляції рівнів вихідного ряду може наближатися до нуля. Послідовність коефіцієнтів автокореляції першого, другого тощо порядків називають автокореляційною функцією часової послідовності. Графік залежності значень коефіцієнтів автокореляції від величини лагу (порядку коефіцієнта автокореляції) називають корелограмою. За допомогою аналізу автокореляційної функції і корелограми можна виявити структуру послідовності, тобто визначити присутність у ній тієї чи іншої компоненти.

Аналіз структури ряду можна проводити таким чином:

- якщо найбільш високим виявився коефіцієнт автокореляції першого порядку, то досліджуваний ряд містить тільки тенденцію;
- якщо найвищим виявився коефіцієнт автокореляції порядку τ, то ряд містить циклічні коливання з періодичністю в τ моментів часу;
- якщо жоден з коефіцієнтів автокореляції не є значущим, можна зробити одне з двох припущень щодо структури ряду:
  - ряд не містить тенденції і циклічних коливань, а включає тільки випадкову компоненту;
  - ряд містить сильну нелінійну тенденцію, для виявлення якої потрібно провести додатковий аналіз.

Для судження про наявність або відсутність автокореляції в досліджуваній послідовності значень, фактичне значення коефіцієнта автокореляції зіставляють з табличним для 5% або 1% рівня значимості (тобто за величиною ймовірності допустити помилку при прийнятті гіпотези про незалежність рівнів ряду). Якщо обчислене значення менше ніж табличне, то гіпотеза про відсутність автокореляції приймається і, навпаки, в протилежному випадку, відкидається. Необхідно підкреслити, що лінійні коефіцієнти автокореляції характеризують тісноту тільки лінійного зв'язку поточного і попередніх рівнів ряду. Тому, за коефіцієнтами автокореляції можна судити тільки про наявність чи відсутність лінійної залежності (або близькою до лінійної). Для деяких часових рядів, які мають сильну нелінійну тенденцію (наприклад, параболу другого порядку або експоненту), коефіцієнт автокореляції рівнів вихідного ряду може наближатися до нуля. За знаком коефіцієнта автокореляції не можна робити висновок про зростаючу або спадну тенденції в рівнях ряду.

КП.3.8. Розрахунок автокореляції

Загальною формулою для обчислення коефіцієнта автокореляції є така

$$r(\tau)=\frac{(n-\tau)\sum_{t=1}^{n-\tau} y_t y_{t+\tau}-\sum_{t=1}^{n-\tau} y_t \sum_{t=1}^{n-\tau} y_{t+\tau}}{\sqrt{[(n-\tau)\sum_{t=1}^{n-\tau} y_t^2-(\sum_{t=1}^{n-\tau} y_t)^2][(n-\tau)\sum_{t=1}^{n-\tau} y_{t+\tau}^2-(\sum_{t=1}^{n-\tau} y_{t+\tau})^2]}} \quad \text{(КП.26)}$$

Порядок коефіцієнтів автокореляції визначає часовий лаг: першого порядку (при $\tau = 1$), другого порядку (при $\tau = 2$) тощо.

Послідовність коефіцієнтів автокореляції рівнів першого, другого, третього і т.д. порядків називають автокореляційною функцією. Значення автокореляційної функції можуть коливатися від -1 до +1, але зі стаціонарності випливає, що $r(\tau) = -r(\tau)$. Графік автокореляційної функції називається корелограмою.

Вибірковий коефіцієнт автокореляції обчислюється за формулою:

$$r(\tau) = \frac{\frac{1}{n-\tau}\sum_{t=1}^{n-\tau}(y_t - \bar{y})\cdot(y_{t+\tau} - \bar{y})}{\frac{1}{n-1}\sum_{t=1}^{n}(y_t - \bar{y})^2}. \qquad \text{(КП.27)}$$

Коефіцієнт автокореляції рівнів ряду першого порядку, що вимірює залежність між сусідніми рівнями ряду y_t та y_{t-1}, тобто при лагу 1, розраховується за формулою:

$$r_1 = \frac{\sum_{t=2}^{n}(y_t - \bar{y}_1)\cdot(y_{t-1} - \bar{y}_2)}{\sqrt{\sum_{t=2}^{n}(y_t - \bar{y}_1)^2 \cdot \sum_{t=2}^{n}(y_{t-1} - \bar{y}_2)^2}}, \qquad \text{(КП.28)}$$

де $\bar{y}_1 = \frac{\sum_{t=2}^{n} y_t}{n-1}$; $\bar{y}_2 = \frac{\sum_{t=2}^{n} y_{t-1}}{n-1}$.

Аналогічно визначаються коефіцієнти автокореляції другого і вищих порядків. Так, коефіцієнт автокореляції другого порядку характеризує тісноту зв'язку між рівнями y_t та y_{t-2} і визначається за формулою:

$$r_2 = \frac{\sum_{t=3}^{n}(y_t - \bar{y}_3)\cdot(y_{t-2} - \bar{y}_4)}{\sqrt{\sum_{t=3}^{n}(y_t - \bar{y}_3)^2 \cdot \sum_{t=3}^{n}(y_{t-2} - \bar{y}_4)^2}}, \qquad \text{(КП.29)}$$

де $\bar{y}_3 = \frac{\sum_{t=3}^{n} y_t}{n-2}$; $\bar{y}_4 = \frac{\sum_{t=3}^{n} y_{t-2}}{n-2}$.

Для обчислення коефіцієнта автокореляції за формулою (КП.26) в Excel можна скористатися функцією **КОРРЕЛ**. Припустимо, що базова змінна включає діапазон A1 : A34. Тоді коефіцієнт автокореляції дорівнює:

= **КОРРЕЛ** (A1 : A33 ; A2 : A34).

На практиці, як правило, при обчисленні автокореляції використовується формула (КП.27). Аналіз автокореляційної функції і корелограми дозволяє визначити лаг, за якого автокореляція є найбільш високою, тобто за допомогою аналізу автокореляційної функції і корелограми можна виявити структуру ряду.

Тому коефіцієнт автокореляції рівнів і автокореляційну функцію доцільно використовувати для виявлення в часовому ряді наявності або відсутності трендової компоненти і сезонної компоненти.

Приклад. *Аналіз часового ряду валового внутрішнього продукту*

Валовий внутрішній продукт є на стадії виробництва сумою доданих вартостей галузей економіки, а на стадії використання – вартістю товарів і послуг, призначених для кінцевого споживання, накопичення й експорту. Як вихідна інформація використовуються дані: номінальний обсяг валового внутрішнього продукту, млрд. крб. (з 1998 р млн. крб.) – квартальні дані з 1994:1 по 2003:1 (табл. КП.2).

Таблиця КП.2

ВВП

Дата	4кв.1994	1кв.1995	2кв.1995	3кв.1995	4кв.1995	1кв.1996	2кв.1996	3кв.1996	4кв.1996	1кв.1997
ВВП	225.00	235.00	325.00	421.00	448.00	425.00	469.00	549.00	565.00	513.00
№	1	2	3	4	5	6	7	8	9	10

Дата	2кв.1997	3кв.1997	4кв.1997	1кв.1998	2кв.1998	3кв.1998	4кв.1998	1кв.1999	2кв.1999	3кв.1999
ВВП	555.00	634.00	641.00	551.00	602.00	676.00	801.00	901.00	1102.00	1373.00
№	11	12	13	14	15	16	17	18	19	20

Дата	4кв.1999	1кв.2000	2кв.2000	3кв.2000	4кв.2000	1кв.2001	2кв.2001	3кв.2001	4кв.2001	1кв.2002
ВВП	1447.00	1527.00	1697.00	2038.00	2044.00	1922.00	2120.00	2536.00	2461.00	2268.00
№	21	22	23	24	25	26	27	28	29	30

Дата	2кв.2002	3кв.2002	4кв.2002	1кв.2003
ВВП	2523.00	3074.00	2998.00	2893.10
№	31	32	33	34

Графік цього ряду наведено на рис. КП.14.

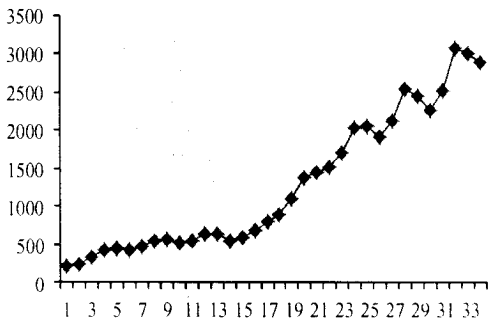


Рис. КП.14.

Бачимо, що дані мають зростаючий тренд. Таким чином, вже візуальний аналіз дозволяє зробити висновок про нестаціонарність вихідного часового ряду. Перевіримо це припущення, обчислимо коефіцієнти автокореляції (табл. КП.3) і побудуємо графік автокореляційної функції часового ряду ВВП, тобто його (корелограму) (рис. КП.15).

Таблиця КП.3

Автокореляційна функція								
Лаг	1	2	3	4	5	6	7	8
Коефіцієнти автокореляції	0.914	0.811	0.717	0.651	0.576	0.480	0.387	0.315

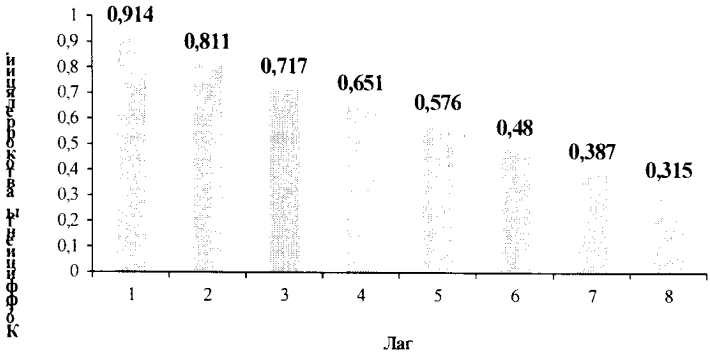


Рис. КП.15. Корелограма.

Корелограма автокореляційної функції у разі стаціонарного часового ряду повинна швидко спадати з ростом t після кількох перших значень. З рисунка видно, що досліджуваний ряд не є стаціонарним. Часовий ряд валового внутрішнього продукту містить трендовую компоненту

КП.4. Ієрархічний агломеративний кластерний аналіз багатовимірних даних

В практичних дослідженнях різноманітних явищ, процесів, ситуацій, об'єктів дані спостережень використовують для отримання корисної і застосовної на практиці інформації переважно для прийняття рішень. Для цього, в першу чергу, здійснюють попереднє опрацювання існуючих даних, яке полягає у поділі даних на однорідні групи. В результаті такого поділу пошук закономірностей в такій групі стає більш коректним. Основний метод, яким здійснюється такий поділ, є процедура кластерного аналізу. Ідея класифікації отриманих даних щодо деяких, подібних між собою об'єктів, за допомогою кластерного аналізу є за своїм змістом інтуїтивною, якщо припустити, що в n-вимірному просторі ознак існує деяка метрика, за якою ці об'єкти можна згрупувати в окремі групи – кластери.

Кластерний аналіз є одним із методів багатовимірного статистичного аналізу, тобто аналізу даних, коли кожне спостереження подається не одним деяким показником, а сукупністю значень різних показників. Він включає низку алгоритмів, за допомогою яких здійснюється і формування самих кластерів і розподіл об'єктів за кластерами. Кластерний аналіз, перш за все, розв'язує задачу внесення структури в дані, тобто їх групової однорідності, а також забезпечує виділення компактних,

віддалених одна від одної груп об'єктів, тобто шукає «природне» розбиття сукупності на області скупчення об'єктів.

Методи багатовимірного аналізу – найбільш діючий кількісний інструмент дослідження процесів, які описані більшою кількістю характеристик. До них відносяться: кластерний аналіз, таксономія, розпізнавання образів.

Кластерний аналіз найбільше яскраво відбиває риси багатовимірного аналізу в сенсі класифікації конкретних об'єктів. Його процедури, а існує багато різновидностей кластерного аналізу, необхідні в тому випадку, коли структуру даних неможливо виявити ні візуально, ні за допомогою експертів. Велика перевага кластерного аналізу в тому, що він дозволяє здійснювати розбиття об'єктів не за одним параметром, а за цілим набором ознак. Крім того, на відміну від більшості математико-статистичних методів він не накладає будь-яких обмежень на вид об'єктів, що підлягають кластеризації.

Кластерний аналіз дозволяє розглядати досить значні обсяги даних, різко скорочувати і стискати їх, робити їх компактними та наочними. Важливе значення метод має стосовно сукупностей часових рядів, які характеризують динаміку розвитку. Іншими словами, стосовно часових рядів, він дозволяє виділяти періоди, в яких значення відповідних показників були достатньо близькими, а також визначати групи часових рядів з найбільш подібною динамікою. Іноді підхід кластерного аналізу називають у літературі чисельною таксономією, чисельною класифікацією, класифікацією із самонавчанням. Перше застосування кластерний аналіз знайшов у соціології. Назва кластерний аналіз походить від англійського слова cluster (гроно, скупчення). Вперше в 1939 році був визначений предмет кластерного аналізу та зроблено його опис дослідником Тріоном. Кластерний аналіз можна застосовувати до інтервальних даних, частот та бінарних даних.

Головне призначення кластерного аналізу – вирішення таких основних задач:

- розробка типології або класифікації;
- виявлення відповідної структури в даних;
- дослідження корисних концептуальних схем групування об'єктів;
- висунення гіпотез на основі дослідження даних;
- перевірка гіпотез або дослідження з метою визначити, чи дійсно типи (групи), виділені тим або іншим способом присутні в існуючих даних;
- розбиття множини досліджуваних об'єктів і ознак на однорідні, у відповідному розумінні, групи або кластери.

Незалежно від предмета дослідження застосування кластерного аналізу включає наступні етапи:

- відбір даних для кластеризації та подання їх у вигляді таблиці «об'єкт-властивість»;
- нормування даних, які подіх таблицею «об'єкт - властивість»;
- вибір та обґрунтування метрики для формування матриці близькостей;
- побудова матриці близькостей на основі нормованої таблиці «об'єкт - властивість»;
- вибір стратегії об'єднання для процедури кластерного аналізу;
- реалізація процедури кластерного аналізу стосовно матриці близькості;

- побудова дендрограми та виділення за відповідними критеріями потрібних кластерів.

Кластерний аналіз ставить такі вимоги до даних:

- показники не повинні корелювати між собою;
- показники властивостей мають бути безрозмірними;
- вплив будь-яких факторів на значення показників має бути виключений.

Методи кластерного аналізу можна застосовувати в різних випадках, у тому разі, коли мова йде про просте угруповання, в якому усе зводиться до утворення груп по кількісній подібності. В залежності від конкретної прикладної задачі мета кластерного аналізу може бути різна, наприклад:

– зрозуміти структуру множини об'єктів, розбивши їх на однорідні, в тому чи іншому сенсі, групи і, тим самим, спростити подальше опрацювання даних для прийняття рішень, працюючи з кожним кластером окремо;

– виділити нетипові об'єкти, які не належать до жодного з кластерів. Цю задачу називають одноласовою класифікацією виявлення нетиповості або новизни;

– зменшити, у випадку надвеликих вибірок X, обсяг даних для збереження, залишивши по одному, найбільш характерному представнику від кожного кластера;

– дослідити динаміку об'єктів в процесі їх експлуатації за зміною відстаней всередині класів та між класами.

Переваги кластерного аналізу:

а) дозволяє робити розбиття об'єктів не по одному параметру, а за цілим набором ознак;

б) кластерний аналіз, на відміну від більшості математико-статистичних методів, не накладає ніяких обмежень на вид об'єктів, що розглядаються, і дозволяє досліджувати множину вихідних даних практично довільної природи;

в) дозволяє розглядати досить великий обсяг інформації й різко скорочувати, стискати більші масиви інформації, робити їх компактними й наочними;

г) кластерний аналіз можна використати циклічно: у цьому випадку дослідження проводиться доти, поки не будуть досягнуті необхідні результати. При цьому кожен цикл тут може надавати інформацію, що здатна сильно змінити спрямованість і підходи подальшого застосування кластерного аналізу. Цей процес можна представити системою зі зворотним зв'язком;

д) в задачах прогнозування досить перспективним є поєднання кластерного аналізу з іншими кількісними методами (наприклад, з регресійним аналізом).

Недоліки й обмеження кластерного аналізу:

а) склад і кількість кластерів залежить від обраних критеріїв розбиття;

б) при приведенні вихідного масиву даних до більш компактного вигляду можуть виникати певні перекручення, а також можуть губитися індивідуальні риси окремих об'єктів через заміну їхніми характеристиками узагальнених значень параметрів кластера;

в) при проведенні класифікації об'єктів часто ігнорується можливість відсутності в розглянутій сукупності яких-небудь значень кластерів.

У кластерному аналізі вважається, що:

а) обрані характеристики допускають, в принципі, бажане розбиття на кластери;

б) одиниці виміру (масштаб) обрані правильно.

Вибір масштабу відіграє велику роль. Як правило, дані нормалізують (обчисленням середнього й діленням на стандартне відхилення) так, що дисперсія виявляється рівною одиниці.

Методи кластерного аналізу можна застосовувати навіть тоді, коли необхідно здійснити звичайний поділ множини об'єктів на групи лише за кількісною подібністю. Алгоритми кластерного аналізу мають розроблену програмну реалізацію, що дозволяє розв'язувати задачі великої розмірності. Метод ієрархічного кластерного аналізу (інколи – чисельної таксономії) існуючих даних здійснює класифікацію, яка раніше не існувала, або створює нову, ігноруючи попередню, переглядаючи дані знову.

Вихідні дані. Для кластерного аналізу дані, що зібрані в процесі дослідження чи отримані експериментально подають у формі таблиці «об'єкт - властивість», в такий спосіб, що першим стовпчиком є назви об'єктів, які підлягають групуванню, а решта стовпчиків відповідають конкретним властивостям (ознакам, характеристикам, показникам тощо) і містять їх конкретні значення.

Задача. Полягає в тому, щоб на підставі даних представлених таблицею «об'єкт - властивість», розбити множину $G = \{ g_i : g \in G,\ i = 1,\ 2,\ \ldots,\ m \}$ цих об'єктів на k (k – ціле) кластерів – неперетинних підмножин $Q_1, Q_2, \ldots, Q_k$ так, щоб $\bigcup_{j=1}^{k} Q_j = G$, а $\bigcap_{j=1}^{k} Q_j = \varnothing$, тобто:

а) кожен об'єкт g_i має належати одній і тільки одній підмножині розбиття;

б) об'єкти, що належать одному і тому самому кластеру, були подібними;

в) об'єкти, що належать до різних кластерів були відмінними.

Результат. За отриманими даними має бути побудована дендрограма і приведена її інтерпретація.

***Виконання даної роботи полягає** у реалізації її двох частин, а саме, побудови матриці близькості на підставі таблиці об'єкт-властивість та проведення самого кластерного аналізу на основі побудованої матриці близькостей.*

Частина І. Побудова матриці близькості.

Ілюстрацію процедури агломеративного ієрархічного кластерного аналізу доцільно провести на конкретному прикладі, з відповідними поясненнями та обґрунтуваннями, використовуючи для розрахунків табличний процесор Ms Excel 2003.

Як приклад використано результати тренажерної підготовки операторського персоналу систем опрацювання аерокосмічних зображень, які використовуються в задачах виявлення на них об'єктів заданого класу. В результаті опрацювання одного і того ж сценарію, який полягав у поданні на монітор послідовності тестових зображень, для кожного оператора отримані дані про час опрацювання кожного тесту, тобто пошуку і виявлення та ідентифікації об'єкта на тестовому прикладі цього сценарію. Значення часу є випадковими величинами з асиметричним одномодальним розподілом і поданими у вигляді індивідуальних часових рядів.

Формування таблиці «об'єкт- властивість». В результаті первинного опрацювання даних з використанням описової статистики в межах показників визначених для цього в Ms Excel 2003 побудована таблиця, яку називають таблицею «об'єкт-властивість». В цьому випадку – це таблиця «оператор-індивідуальні

показники» розміром $n \times m$, де $n = \overline{1, 13}$ – кількість операторів, а $m = \overline{1, 14}$ – кількість використаних показників описової статистики. Отже, для кластерного аналізу подається множина G, яка включає m об'єктів, кожен з яких характеризується n ознаками. Дані представлені нижче в таблиці КП.4.

Таблиця КП.4

Таблиця «оператор - показник» – значень показників об'єктів за описовою статистикою

	Показники													
Об'єкти	Середнє	Стнд. пом.	Медіана	Мода	Стнд. відх.	Дисперсія	Ексцес	Асиметрія	Розмах	Мінімум	Максимум	Сума	К-сть даних	Надійність
Bur	1886,2	26,68	1839,5	1697	322,4	103921	3,837	1,257	2290	1196	3486	275378	146	52,7
Cub	678,9	17,57	636,0	708	187,6	35189	0,790	0,995	976	376	1352	77395	114	34,8
Cup	1506,5	22,74	1484,0	1663	245,9	60482	-0,267	0,516	1130	1050	2180	176263	117	45,0
Hav	753,7	16,74	726,5	787	169,1	28600	1,714	1,208	892	483	1375	76878	102	33,2
Hod	709,5	16,62	674,0	714	171,1	29272	1,742	1,261	900	462	1362	75204	106	32,9
Kol	430,2	8,96	415,0	429	105,3	11089	6,887	2,169	722	241	963	59372	138	17,7
Lob	665,0	18,36	622,5	601	179,9	32354	0,828	1,037	845	355	1200	63840	96	36,4
Lot	608,2	15,81	593,0	623	178,2	31751	2,029	1,163	981	347	1328	77246	127	31,3
Oli	655,3	16,69	619,0	609	166,0	27563	3,005	1,503	890	396	1286	64872	99	33,1
Per	728,3	8,19	717,0	705	96,5	9313	7,528	1,662	769	504	1273	101229	139	16,2
Pon	675,9	14,43	651,5	518	165,8	27489	0,451	0,754	833	395	1228	89213	132	28,5
Sol	1633,1	28,83	1588,0	1446	311,9	97268	0,590	0,771	1524	1117	2641	191070	117	57,1
Syr	704,4	19,03	645,5	606	215,3	46362	1,952	1,325	1176	318	1494	90165	128	37,7

Наведені в табл. КП.4 показники за величиною і розмірністю є дуже неоднорідними, а це означає неможливість обґрунтованої інтерпретації отриманого результату кластерного аналізу. Це приводить до того, що величина, отриманих в результаті кластеризації відстаней між точками, що відображають положення об'єктів у просторі їхніх властивостей, визначатиметься довільно обраним масштабом. Щоб усунути неоднорідність виміру вихідних даних, всі їх значення попередньо нормуються, тобто виражаються через відношення цих значень до деякої величини, що відображає певні властивості цього показника. Важливо, щоб усі значення змінних змінювалися в порівняльних шкалах. Крім того, за наявності значного розкиду величин значень необхідно звести їх до інтервалу [0,1], тобто здійснити їх нормування.

Нормування таблиці «об'єкт-властивість». Основні труднощі полягають у виборі способу нормування, оскільки, необхідно максимально врахувати якісну специфіку показників. Нормування використовують в тому випадку, коли ознаки представлені різними шкалами і мають різний фізичний зміст (розмірність). Нормування – це перехід до деякого однакового опису для всіх ознак і введення нової умовної одиниці вимірювання, яка допускає формальне зіставлення об'єктів чи їх ознак. Нормування приводить до безрозмірних величин, зберігаючи за ними відповідність ознакам. Перехід від початкових значень x до нормованих z здійснюють за формулами:

$$z = \frac{x - x_{\text{сер}}}{s}, \qquad z = \frac{x}{x_{\text{етал}}}, \qquad z = \frac{x}{x_{\text{сер}}}, \qquad \text{(КП.30)}$$

$$z = \frac{x}{x_{\max}}, \qquad z = \frac{x - x_{\text{сер}}}{x_{\max} - x_{\min}}, \qquad z = \frac{x - x_{\min}}{x_{\max} - x_{\min}},$$

де s – середньоквадратичне відхилення.

Найбільш поширеним нормуванням є лінійне перетворення, яке проводять за модифікованою останньою формулою.

Визнання рівноцінності різних показників не завжди є виправданим. Очевидно, що великі значення дисперсії, суми та їх варіації будуть суттєво впливати на результат кластерного аналізу.

В результаті нормування за цією формулою всі дані табл. КП.4 є приведені до одиничного інтервалу, тобто ця таблиця прийме вигляд табл. КП.5. Нормування здійснюється по кожному стовпчику окремо, тобто нормуються ознаки одного типу.

Таблиця КП.5.

Таблиця «оператор - показник» – після нормування

	Показники													
Оператори	**Середнє**	**Стнд. пом.**	**Медіана**	**Мода**	**Стнд. Відх.**	**Дисперсія**	**Ексцес**	**Асиметрія**	**Розмах**	**Мінімум**	**Максимум**	**Сумма**	**К-сть**	**Надійність**
Bur	1,000	0,896	1,000	1,000	1,000	1,000	0,527	0,447	1,000	1,000	1,000	1,000	1,000	0,893
Cub	0,171	0,454	0,155	0,220	0,403	0,274	0,136	0,288	0,162	0,141	0,154	0,083	0,360	0,455
Cup	0,739	0,705	0,750	0,973	0,662	0,541	0,000	-0,002	0,260	0,847	0,482	0,541	0,420	0,705
Hav	0,222	0,414	0,219	0,282	0,321	0,204	0,254	0,417	0,108	0,253	0,163	0,081	0,120	0,416
Hod	0,192	0,408	0,182	0,225	0,330	0,211	0,258	0,449	0,114	0,231	0,158	0,073	0,200	0,410
Kol	0,000	0,038	0,000	0,000	0,039	0,019	0,918	1,000	0,000	0,000	0,000	0,000	0,840	0,038
Lob	0,161	0,493	0,146	0,136	0,369	0,244	0,141	0,313	0,078	0,119	0,094	0,021	0,000	0,495
Lot	0,122	0,369	0,125	0,153	0,362	0,237	0,295	0,390	0,165	0,111	0,145	0,083	0,620	0,369
Oli	0,155	0,412	0,143	0,142	0,308	0,193	0,420	0,596	0,107	0,162	0,128	0,025	0,060	0,414
Per	0,205	0,000	0,212	0,218	0,000	0,000	1,000	0,692	0,030	0,275	0,123	0,194	0,860	0,000
Pon	0,169	0,302	0,166	0,070	0,307	0,192	0,092	0,142	0,071	0,161	0,105	0,138	0,720	0,302
Sol	0,826	1,000	0,823	0,802	0,954	0,930	0,110	0,152	0,511	0,917	0,665	0,610	0,420	1,000
Syr	0,188	0,525	0,162	0,140	0,526	0,392	0,285	0,488	0,290	0,081	0,210	0,143	0,640	0,525

Зауважимо, що нормування означає, що всі ознаки є рівноцінними з погляду з'ясування подібності розглянутих об'єктів. Інколи, поряд з нормуванням надають кожному з показників вагу і тим самим вказують на його значущість під час встановлення подібності і відмінностей між об'єктами.

Вибір метрики для побудови матриці близькості. Якщо відстань між об'єктами природно трактувати як міру відмінності об'єктів, то обернену величину $B = \frac{1}{D}$ можна розглядати як міру подібності (близькості) об'єктів.

1. Найчастіше відстань $D(\vec{x}_1, \vec{x}_2)$ між об'єктами вимірюють в евклідовій матриці, яка найбільш узгоджена з нашими інтуїтивними представленнями про близькість об'єктів і визначається за формулою

$$D_E(\vec{x}_1, \vec{x}_2) = \sqrt{\sum_{i=1}^{N} (x_{1j} - x_{2j})^2} \quad , \qquad \text{(КП.31)}$$

де x_{1j} та x_{2j} – значення j-ї компоненти в описах 1-го та 2-го об'єктів, N – розмірність простору ознак, а в даному випадку $N = n$.

2. Широко використовується лінійна метрика (метрика міських кварталів або манхетенська метрика), яка задає відстань

$$D_M(\vec{x}_1 \vec{x}_2) = \sum_{i=1}^{N} \left| \vec{x}_{1j} - \vec{x}_{2j} \right|, \qquad \text{(КП.32)}$$

а також sup-норма, яка визначає чебишевську відстань

$$D_{ч}(\vec{x}_1, \vec{x}_2) = \max_j \left| x_{1j} - x_{2j} \right|. \qquad \text{(КП.33)}$$

Очевидно, що $D_{ч} < D_E < D_M$ з ростом розмірності N манхетенська відстань D_M підкреслює, а чебишовська згладжує відмінності між об'єктами. Відстані $D_M, D_E, D_{ч}$ визначаються для тих описів, в яких ознаки виражені кількісними шкалами і є частинними випадками узагальнення відстані Мінковського $D_p(\vec{x}_1, \vec{x}_2) = \left(\sum_{i=1}^{N} \left| x_{1j} - x_{2j} \right|^p \right)^{\frac{1}{p}}$, тобто відстані $D_M, D_E, D_{ч}$ визначаються значенням степеня $p = 1, 2, , \infty$. Існує досить велика різноманітність мір, але на практиці ці три міри є найуживанішими.

Евклідова відстань ефективна при дослідженні слабко кореляційних сукупностей об'єктів (кулеподібні класи), а манхетенська тоді, коли об'єкти утворюють плоскі витягнені класи, ортогональні до будь-яких координатних осей простору ознак. Тому опрацювання однієї і тієї ж сукупності даних одним і тим же методом або алгоритмом, але з використанням різних метрик може дати різні, інколи кардинально протилежні, результати. Отже, до вибору метрики слід підходити дуже продумано і обережно, співставляючи результати використання різних метрик між собою і з цілями опрацювання даних, яке здійснюємо. Якщо ознаки представляються в якісних шкалах, зокрема в шкалах найменувань та порядку, використовують відстань Геммінга

$$D_H(\vec{x}_1, \vec{x}_2) = \frac{1}{N} \sum_{j=1}^{N} \left| x_{1j} - x_{2j} \right|. \qquad \text{(КП.34)}$$

для якої відмінності виражаються числом неспівпадінь властивостей порівнюваних об'єктів. У випадку якісних шкал ознаки розглядаються як бінарні, тобто такі, що можуть приймати лише два значення "0" та "1". Відстань Геммінга D_H є

максимальною і рівна 1 для об'єктів з протилежними за значеннями описами, тобто елементи одного опису є протилежними до відповідних елементів опису другого об'єкту. Для об'єктів, всі ознаки яких (з числа включених в опис) співпадають, $D_H = 0$. При виборі виду міри близькості необхідно врахувати їх формальні властивості й порівняти їх зі змістовними особливостями задачі. В результаті застосування будь-якої з цих метрик до даних отримують матрицю близькості, розмірність якої $M \times M$, а за своєю специфікою вона є симетричною відносно головної діагоналі.

Побудова таблиці близькості. Для побудови матриці близькості за допомогою табличного процесора Ms Excel 2003 використовують дані табл. 5 в такий спосіб.

1. На робочому листі поміщають табл. 5 та її копію так, щоб номери стрічок в них були однакові (наприклад, якщо перший елемент табл. 5 – таблиці-оригіналу розміщений в комірці **B3**, то перший елемент таблиці-копії цієї таблиці є розміщений в комірці **R3**).

2. Далі будують матрицю близькості. При цьому виникає задача вибору міри близькості. Найчастіше відстань $D(\vec{x}_1, \vec{x}_2)$ між об'єктами вимірюють в евклідовій метриці, її ще називають *евклідовою відстанню* і обчислюють за такою формулою:

$$d_E = \sqrt{\sum_{p=1}^{q} (x_{ip} - x_{jp})^2} \,. \qquad \text{(КП.35)}$$

Евклідова метрика, є найбільш узгодженою з нашими інтуїтивними представленнями про близькість об'єктів.

3. В результаті застосування до матриці даних цієї метрики отримують матрицю близькості, розмірність якої $M \times M$, де $p = 1, 2, \ldots, q$ – номер ознаки, i, j – індекси пари об'єктів, між якими визначають відстань. Матриця відстаней за своєю специфікою є симетричною відносно головної діагоналі.

В Excel матрицю близькості будують так. Визначають комірку для розміщення першого елемента матриці близькості, наприклад, **B18**, тобто нижче на кілька стрічок під таблицею, в якій розміщують (для даного випадку) таку формулу

=**КОРЕНЬ**(СУММ((B3-$R3)^2+($C$3-$S3)^2+(D3-$T3)^2+($E$3-$U3)^2+(F3-$V3)^2+($G$3-$W3)^2+(H3-$X3)^2+($I$3-$Y3)^2+(J3-$Z3)^2+($K$3-$AA3)^2+(L3-$AB3)^2+($M$3-$AC3)^2+(N3-$AD3)^2+($O$3-$AE3)^2))

і далі, шляхом автозаповнення в M клітинок стовпчика **B** будуть записані відповідні значення першого стовпчика матриці близькості. В результаті в комірці **B19** формула буде мати вигляд

=**КОРЕНЬ**(СУММ((B3-$R4)^2+($C$3-$S4)^2+(D3-$T4)^2+($E$3-$U4)^2+(F3-$V4)^2+($G$3-$W4)^2+(H3-$X4)^2+($I$3-$Y4)^2+(J3-$Z4)^2+($K$3-$AA4)^2+(L3-$AB4)^2+($M$3-$AC4)^2+(N3-$AD4)^2+($O$3-$AE4)^2)),

тобто, номер і значення стрічок таблиці-оригіналу залишаться без змін, а зміняться лише номери стрічок таблиці-копії. В результаті, буде визначено відстань між першим об'єктом, визначеним першою стрічкою таблиці-оригіналу, і кожним об'єктом, визначеним стрічкою таблиці-копії.

4. Для визначення відстані між другим об'єктом таблиці-оригіналу і всіма іншими об'єктами таблиці-копії копіюють формулу з комірки **B18** в комірку **C18**. Після активізації формули в цій комірці, змінюють адресу першої стрічки таблиці-оригіналу

шляхом переміщення всіх кольорових рамок на одну комірку вниз. В результаті цього формула буде мати такий вигляд

=**КОРЕНЬ**(СУММ((B4-$R3)^2+($C$4-$S3)^2+(D4-$T3)^2+($E$4-$U3)^2+(F4-$V3)^2+($G$4-$W3)^2+(H4-$X3)^2+($I$4-$Y3)^2+(J4-$Z3)^2+($K$4-$AA3)^2+(L4-$AB3)^2+($M$4-$AC3)^2+(N4-$AD3)^2+($O$4-$AE3)^2)).

Реалізуємо цю формулу і шляхом автозаповнення формуємо другий стовпчик матриці близькості. Аналогічно роблять при визначенні решти стовпчиків.

Зауваження. Якщо, виходячи з розмірності власних задач, побудувати матрицю близькості для максимально можливих кількостей об'єктів і ознак, то використовуючи її як шаблон, можна знайти значення матриць близькості для будь-якої кількості об'єктів і ознак, замінюючи в таблиці-оригіналі і в таблиці-копії їхні значення і відкидаючи порожні комірки шаблону, отримуємо безпосередньо потрібну матрицю близькості. Іншими словами, маючи лист Excel з один раз зробленою процедурою побудови матриці близькості можна в таблиці ввести інші дані і, відкидаючи зайві або додаючи нові стовпчики і стрічки безпосередньо отримати матрицю близькості, але для конкретної метрики.

Після формування останнього стовпчика отримуємо матрицю близькості зображену у вигляді табл. КП.6.

Таблиця КП.6.

Матриця близькості

	Bur	Cub	Cup	Hav	Hod	Kol	Lob	Lot	Oli	Per	Pon	Sol	Syr
Bur	0,000	2,620	1,547	2,651	2,669	3,292	2,836	2,656	2,812	2,942	2,749	1,112	2,422
Cub	2,620	0,000	1,562	0,359	0,303	1,440	0,397	0,360	0,547	1,365	0,502	1,956	0,461
Cup	1,547	1,562	0,000	1,563	1,619	2,634	1,716	1,734	1,835	2,275	1,728	0,771	1,625
Hav	2,651	0,359	1,563	0,000	0,118	1,408	0,335	0,565	0,326	1,302	0,745	1,981	0,686
Hod	2,669	0,303	1,619	0,118	0,000	1,328	0,346	0,466	0,293	1,249	0,674	2,021	0,602
Kol	3,292	1,440	2,634	1,408	1,328	0,000	1,566	1,141	1,234	0,608	1,340	2,998	1,338
Lob	2,836	0,397	1,716	0,335	0,346	1,566	0,000	0,679	0,430	1,541	0,807	2,091	0,765
Lot	2,656	0,360	1,734	0,565	0,466	1,141	0,679	0,000	0,626	1,087	0,391	2,116	0,374
Oli	2,812	0,547	1,835	0,326	0,293	1,234	0,430	0,626	0,000	1,232	0,891	2,188	0,735
Per	2,942	1,365	2,275	1,302	1,249	0,608	1,541	1,087	1,232	0,000	1,226	2,695	1,307
Pon	2,749	0,502	1,728	0,745	0,674	1,340	0,807	0,391	0,891	1,226	0,000	2,174	0,649
Sol	1,112	1,956	0,771	1,981	2,021	2,998	2,091	2,116	2,188	2,695	2,174	0,000	1,890
Syr	2,422	0,461	1,625	0,686	0,602	1,338	0,765	0,374	0,735	1,307	0,649	1,890	0,000

Отримана матриця близькості, є симетричною діагональною матрицею, яка вказує на величину близькості між об'єктами. На основі такої матриці проводиться агломеративний ієрархічний кластерний аналіз.

Частина II. Проведення агломеративного ієрархічного кластерного аналізу.

Вибір стратегій об'єднання. Процедура кластерного аналізу базується на перерахунку значень матриці близькості і, в результаті, кожного такого кроку обчислень об'єднуються об'єкти, об'єкт з групою або дві групи. Після кожного такого

об'єднання розмірність матриці зменшується на одиницю, а кількість кластерів або кількість об'єктів в конкретному кластері збільшується на одиницю. Проте такі об'єднання відбуваються не будь як довільно, а в рамках конкретно вибраної стратегії, яка діє протягом усієї процедури. Зміст такої стратегії полягає в тому, що кожен новий кластер визначався значеннями ознак, отриманими в результаті перерахунку відповідних значень ознак об'єктів і кластерів, які об'єднуються в цей новий кластер. Іншими словами, процедуру об'єднання об'єктів в кластери можна подати так.

Суть стратегії групування полягає в наступному. У випадку n об'єктів обчислюються всі $\frac{n(n-1)}{2}$ мір відмінностей і пара об'єктів з найменшою мірою об'єднується в одну групу. На наступному кроці визначають відповідну міру відмінності (нове значення близькості) між цією групою і рештою $n-2$ об'єктами, а на більш пізніх стадіях треба буде визначати цю міру між об'єктом і групою будь-якого обсягу, а також між будь-якими двома групами. На кожному кроці класифікації виконується те об'єднання (між двома об'єктами, між об'єктом і групою або між двома групами), для яких міра відмінності мінімальна серед всіх існуючих на даному кроці. Міра повинна бути такою, щоб об'єкт можна було розглядати як групу з одного елемента. Стратегія об'єднання визначається саме мірою відмінності між групами.

Нехай є дві групи i та j з n_i і n_j елементами відповідно; міру відмінності між цими групами позначимо d_{ij}. Припустимо, що d_{ij} – найменша міра з усіх, що залишились, так що i та j об'єднуються і утворюють нову групу k з $n_k = n_i + n_j$ елементами. Розглянемо деяку іншу групу h з n_h елементами. Перед об'єднанням відомі значення мір d_{hi}, d_{hj}, d_{ij} та об'ємів n_h, n_i, n_j. Значення розглянутих мір обчислюють за формулою

$$d_{hk} = \alpha_i d_{hi} + \alpha_j d_{hj} + \beta d_{ij} + \gamma \left| d_{hi} - d_{hj} \right|, \qquad \text{(КП.36)}$$

де параметри α_i, α_j, β і γ визначають суть стратегії.

Найчастіше використовують наступні стратегії.

1. *Стратегія найближчого сусіда.* Відстань між двома групами визначається як відстань між двома найближчими елементами з цих груп. Ця стратегія монотонна і сильно стискає простір ознак, а її параметрами є $\alpha_i = \alpha_j = 0.5$, $\beta = 0$, $\gamma = -0.5$.
2. *Стратегія найдальшого сусіда.* Тут відстань між двома групами визначається як відстань між двома найбільш віддаленими представниками (елементами) цих груп. Вона монотонна і сильно розтягує простір. Її параметри мають значення $\alpha_i = \alpha_j = 0.5$, $\beta = 0$, $\gamma = 0.5$.
3. *Стратегія групового середнього* (середнього зв'язку). Якщо одна група складається з m_1, а друга з m_2 елементів, то відстань між цими групами в цій стратегії визначається як середнє арифметичне відстаней між елементами з цих груп. Ця

стратегія є монотонною і зберігає метрику простору. Параметри стратегії рівні: $\alpha_i = \frac{n_i}{n_k}$, $\alpha_j = \frac{n_j}{n_k}$, $\beta = \gamma = 0$.

4. *Гнучка стратегія.* Може бути застосована для будь-якої міри близькості і визначається наступними обмеженнями $\alpha_i + \alpha_j + \beta = 1$, $\alpha_i = \alpha_j$, $\beta < 1$, $\gamma = 0$. Стратегія монотонна, а її властивості повністю залежать від β. Якщо $\beta = 0$, то стратегія зберігає метрику простору. Якщо $\beta > 0$ то стратегія стискає простір, а якщо $\beta < 0$, то розтягує. Для практичного використання для параметрів приймають наступні значення $\alpha_i = \alpha_j = 0.625$, $\beta = -0.25$.

Об'єктами класифікації можуть бути практично будь-які об'єкти. Причому стратегії класифікації, тобто чисельні методи не залежать від природи об'єктів, що класифікуються, але різні стратегії, як правило, приводять до різних результатів. Тому вибір стратегії є досить складною задачею і вимагає високої кваліфікації від фахівця. Дані, які підлягають класифікації, утворюють множину елементів, кожен з яких визначається набором ознак, які відповідають узагальненому поняттю змінної. Такі дані або множину даних вважають неоднорідними, тобто множину даних розглядають як сукупність підмножин, таких, що всередині підмножини її елементи між собою є більш подібними, ніж з будь-яким іншим елементом будь-якою іншої підмножини. В цьому розумінні виділяють два підходи до аналізу. Перший є знаходженням міри впевненості, що при використанні цієї чисельної процедури можна вважати, що існують такі підмножини (тобто якщо їх немає, то і не повинно бути їх знайдено); другий – допускає, що істотних відмінностей підмножини не мають, проте для полегшення аналізу дані все таки треба розбити штучно.

В математичному плані задача класифікації даних тобто елементів формулюється як задача побудови розбиття елементів множини даних на деяке наперед задане чи знайдене в ході аналізу число непорожніх попарно неперетинних підмножин (класів) елементів.

Проведення кластерного аналізу. Процедура, яка складає суть ієрархічної класифікації полягає в тому, що в матриці близькості з двох об'єктів, між якими найменша відстань, формують перший кластер, значення якого перераховують у відповідності з вибраною стратегією. Другий об'єкт з більшим номером стовпчика і стрічки викидається, а замість першого об'єкта (з меншим номером стовпчика і стрічки) вставляється утворений з цих об'єктів кластер з перерахованими значеннями. В результаті розмірність матриці зменшується на одиницю. На наступному кроці знову знаходять найменшу відстань між її елементами і чинять аналогічно. Коли матриця близькості матиме розмірність 2×2 процедура кластеризації припиняється. На основі отриманої на кожному кроці інформації про об'єднання кластерів і знайдені значення мінімальних відстаней будується дендрограма і подається її інтерпретація.

Суть цієї процедури полягає в тому, що: в матриці близькості, в якій стовпці і стрічки є векторами пронумерованих об'єктів, шукають найменше значення, визначають, які об'єкти відповідають цьому значенню і об'єднують їх в одну групу. Далі перераховують значення векторів цих об'єктів і подають цю групу як новий об'єкт зі своїм вектором значень.

Ця процедура реалізується за таким алгоритмом.

1. Знаходять в матриці близькості найменше значення і об'єднують об'єкти, яким воно відповідає, в одну групу.

2. Вилучають стовпчики, що належать цим об'єктам, і розміщують їх, довільно і поряд, під матрицею близькості. Залишаємо порожнім місце стовпчика першого (лівого) об'єкта і зсуваємо вліво всі стовпчики, що лежали справа від стовпчика другого об'єкта, тобто ліквідуємо порожнє місце. Зсувом вгору ліквідуємо вектор стрічку другого об'єкта.

3. Вибираємо стратегію об'єднання об'єктів. Відповідно до обраної стратегії значення вилучених стовпчиків перераховуємо за відповідною формулою. В даному прикладі використано гнучку стратегію з параметрами $\alpha_i = \alpha_j = 0.625$, $\beta = -0.25$, оскільки ця стратегія за даного значення β дещо розтягує простір, а отже, і віддаляє між собою кластери, підкреслюючи їхні індивідуальності і однорідність.

4. В отриманому в результаті перерахунку стовпчику шукають два найменших значення, які знаходяться в комірках, що відповідають коміркам з нулями стовпчиків, які перераховуються, і верхнє мінімальне значення замінюємо на нуль, а нижнє ліквідуємо зсувом вгору всіх нижче розташованих комірок в цих стовпчиках.

5. Перерахований стовпчик вставляємо на місце (порожне) першого вилученого стовпчика і перевіряємо чи його нуль лежить на головній діагоналі.

6. Копіюємо всі значення цього стовпчика, транспонуємо їх у стрічку і замінюємо нею стрічку першого вилученого стовпчика.

7. Присвоюємо новому об'єкту, утвореному в результаті об'єднання вилучених об'єктів, наступний за порядком номер. В результаті цієї операції матриця близькості зменшується на один стовпчик і одну стрічку, і в ній з'являється новий об'єкт.

8. Цю процедуру повторюють до тих пір поки матриця не зменшиться до розміру 2×2. Результати кластерного аналізу наведені в табл. КП.7.

Таблиця КП.7

Результати кластерного аналізу операторського персоналу

	Процедура об'єднання кластерів		
Кроки	**Об'єднання**	**Вузол**	**Метрика**
1	4+5	**d14**	0,118
2	14+9	**d15**	0,357
3	2+8	**d16**	0,36
4	15+7	**d17**	0,427
5	16+13	**d18**	0,432
6	18+11	**d19**	0,59
7	6+10	**d20**	0,608
8	3+12	**d21**	0,771
9	19+17	**d22**	1,038
10	1+21	**d23**	1,469
11	22+20	**d24**	2,368
12	23+24	**d25**	5,059

Побудова дендрограми. Візуалізація результатів кластерного аналізу здійснюється з допомогою дендрограми, тобто графічного зображення результатів процесу послідовної кластеризації, яку проводять в термінах матриці близькості. За допомогою дендрограми можна графічно або геометрично зобразити процедуру кластеризації за умови, що ця процедура оперує тільки з елементами матриці відстаней

або подібності. Вид дендрограми залежить від вибору міри подібності або відстані між об'єктом і кластером та методу кластеризації. Найбільш важливим моментом є вибір міри подібності або міри відстані між об'єктом і кластером. Не дивлячись на те, що існує досить велика кількість різноманітних програмних засобів проведення кластерного аналізу, наприклад, ППП «Statistica» та SPSS, а також різних додатків до статистичних пакетів програм опрацювання даних, їх використання часто є дуже складним в сенсі знаходження найбільш відповідного, його придбання, узгодження з використовуваними засобами тощо.

Проте побудова дендрограми кластеризації декількох десятків об'єктів легко може бути здійснена вручну безпосередньо або за дві три ітерації в чорновому варіанті і остаточно сформована в, практично, будь якому графічному середовищі. Процедура побудови дендрограми вручну є нескладною і потребує лише уважності, використовує результати кластерного аналізу приведені в табл. КП.7, складається з побудови ескізу, на підставі якого здійснюється її графічне редагування та масштабування відстаней між між об'єктами і вузлами та включає такі кроки.

1. Побудова ескізу дендрограми. Дендрограму будують починаючи з «кореня дерева», тобто з вузла останнього об'єднання, відобразивши його точкою, відміченою його номером.

2. Оскільки кластерний аналіз є дихотомічною процедурою, то з кореневого вузла проводять дві гілки, які закінчуються точками наступних вузлів. Помічаються номери вузлів.

3. Якщо гілка закінчується на об'єкті, то її кінець помічають номером відповідного об'єкта, якщо ж вузлом, то вказують відповідний йому номер.

4. Після того, як визначені всі гілки і помічені усі об'єкти, здійснюють графічну корекцію і масштабування. Для цього в декартовій системі координат по осі абсцис відкладають значення відстаней об'єднання об'єктів, об'єктів з групами та груп, а гілки будують прямими лініями, довжини яких відповідають значенням відстаней. В результаті отримують дендрограми зображені на рис. КП.15.

Інтерпретація дендрограми. Проведення горизонтальних ліній в площині дендрограми на заданій висоті, в даному випадку, дозволяє виділити окремі кластери. Вже з першого погляду на дендрограму можна вказати, що на рівні 0.75 маємо в принципі шість досить чітко відображених кластерів, в які входять такі об'єкти.

1 кластер – об'єкти 4, 5, 9, 7;
2 кластер – об'єкти 2, 8, 13, 11;
3 кластер – об'єкти 6, 10;
4 кластер – об'єкт 3;
5 кластер – об'єкт 12;
6 кластер – об'єкт 1;

На рівні 1.0 маємо 4 кластери.

1 кластер – об'єкти 4, 5, 9, 7, 2, 8, 13, 11;
2 кластер – об'єкти 6, 10;
3 кластер – об'єкти 3, 12;
4 кластер – об'єкт 1;

На рівні 1.5 отримуємо три кластери, в які входять такі об'єкти:

1 кластер – об'єкти 4, 5, 9, 7, 2, 8, 13, 11;
2 кластер – об'єкти 6, 10;
3 кластер – об'єкти 3, 12;

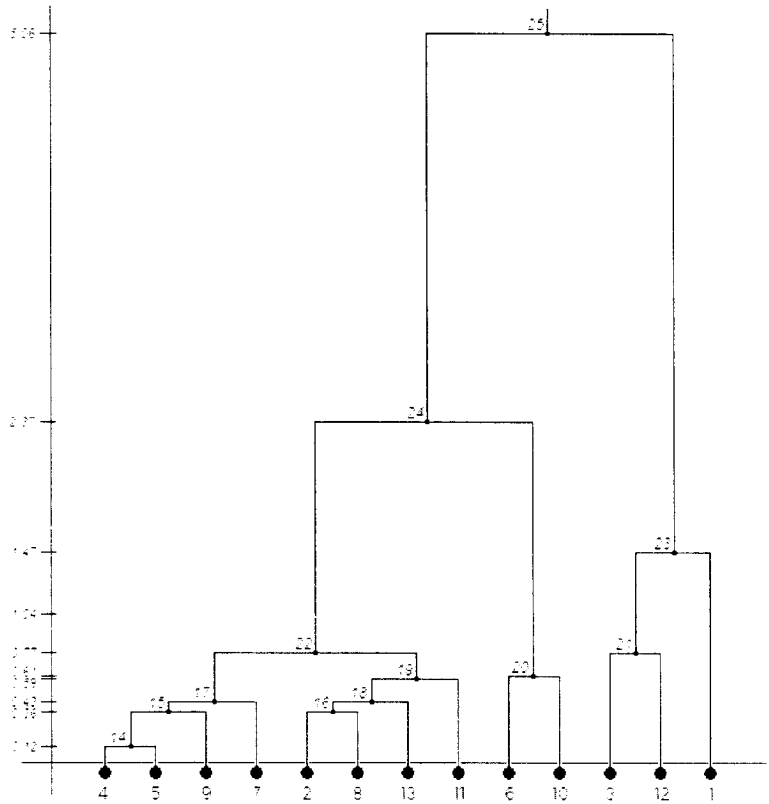


Рис. .КП.15. Дендрограма.

КП.5. Порядок роботи

Мета роботи – ознайомлення з основними методами візуалізації – графічного відображення та первинного статистичного опрацювання числових даних, які представлені вибірковою сукупністю або часовим рядом. Проведення процедури

агломеративного ієрархічного кластерного аналізу в середовищі табличного процесора Excel.

Для досягнення поставленої мети в роботі вирішені такі завдання.

1. Проведені експериментальні дослідження, а зібрані дані, сформовані у файл *.dat (або *.txt) та внесені у книгу Excel.

2. Для отриманих даних побудована звітна таблиця. (Підписи, написи, шрифт, форматування тощо).

3. Для наочного подання отримані дані візуалізовані в декартовій та полярній системах координат (підписи, написи, шрифт, форматування тощо).

4. Кількісні значення характеристик отриманих даних приведені у таблицях описової статистики.

5. Інформація про розподіл даних як випадкових величин подається характером гістограми.

6. Наочне подання графіка функції розподілу даних здійснено за допомогою емпіричної кумуляти.

7. Вихідні дані, подані у вигляді таблиці «об'єкт-властивість».

8. Повна формула для відстані в математичній символіці та в символіці Excel.

9. Таблиця близькості з виділеним найменшим значенням.

10. Назва, параметри та формула стратегії об'єднання в символіці Excel.

11. Таблиця «об'єднання – вузол – метрика».

12. Побудована дендрограма.

13. Проведена інтерпретація результатів.

14. Висновок.

Умови для завдань і проведення агломеративного ієрархічного кластерного аналізу здійснюється за вказаними: формулою нормування, метрикою та стратегією, що наведені в табл. 8.

Таблиця 8.

№	Формула нормування	Вид відстані (метрика)	Стратегія
1	$z = \frac{x - x_{\text{сер}}}{s}$	$D_E(\vec{x}_1, \vec{x}_2) = \sqrt{\sum_{i=1}^{N} (x_{1i} - x_{2i})^2}$	*Стратегія найближчого сусіда.* $\alpha_i = \alpha_j = 0.5$, $\beta = 0$, $\gamma = -0.5$.
2	$z = \frac{x}{x_{\text{етал}}}$	$D_M(\vec{x}_1, \vec{x}_2) = \sum_{i=1}^{N} \lvert \vec{x}_{1i} - \vec{x}_{2i} \rvert$	*Стратегія найдальшого сусіда.* $\alpha_i = \alpha_j = 0.5$, $\beta = 0$, $\gamma = 0.5$.
3	$z = \frac{x}{x_{\text{сер}}}$	$D_H(\vec{x}_1, \vec{x}_2) = \max_i \lvert x_{1i} - x_{2i} \rvert$	*Стратегія групового середнього* $\alpha_i = \frac{n_i}{n_k}$, $\alpha_j = \frac{n_j}{n_k}$, $\beta = \gamma = 0$.

№	Формула нормування	Вид відстані (метрика)	Стратегія
4	$z = \frac{x}{x_{\max}}$		*Гнучка стратегія.* $\alpha_i + \alpha_j + \beta = 1$, $\alpha_i = \alpha_j$, $\beta < 1$, $\gamma = 0$.
5	$z = \frac{x - x_{\text{сер}}}{x_{\max} - x_{\min}}$		
6	$z = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$		

КП.6. Хід роботи

1. Побудова звітної таблиці

1. Ввести дані з *.dat або *.txt. файлу в книгу Excel.
2. Сформувати звітну таблицю з мінімальною кількістю порожніх комірок.
3. Побудувати графіки даних в декартовій та полярній системі координат.
4. Визначити кількісні характеристики описової статистики.
5. Подати дані гістограмою.
6. Подати дані кумулятами: за даними гістограми та за інтегральним процентом.
7. Подати результати роботи у відповідній формі звітності.

2. Для досягнення поставленої мети, а саме: ознайомлення з основними методами висвітлення тенденції поведінки досліджуваного показника, яка представлена характером його тренду, за допомогою методів згладжування часових рядів та подання отриманих результатів засобами табличного процесора MS Excel необхідно провести такі дослідження.

1. Відкриваємо нову **Excel-книгу** і на новий лист вносимо свої дані так, щоб кожне завдання було виконане на одному листі.
2. Згладжування за формулами з Кендела – простого ковзного середнього.

   Особливість цих формул полягає в тому, що з їх допомогою просто розраховувати втрачені рівні на початку та вкінці згладженого ряду.

   а) згладжуємо дані, використовуючи розміри інтервалу згладжування w = 3, 5, 7, 9, 11, 13, 15. Маємо отримати підряд сім стовпчиків;

   б) згладжуємо дані, використовуючи розмір інтервалу згладжування w = 3, далі, отримані згладжені дані згладжуємо знову, але використовуємо розмір інтервалу згладжування w = 5. Продовжуємо згладжування отриманих даних інтервалом згладжування w = 7 і так до w = 15. Маємо отримати підряд сім стовпчиків.

   в) в обох випадках знайти для кожного згладжування кількість поворотних точок та коефіцієнти кореляції між оригінальними значеннями та згладженими.
3. Згладжування за формулами з Полларда. Тут, в залежності від розміру інтервалу згладжування, вага для центрального рівня змінюється. Згладжування провести аналогічно до попереднього пункту.

4. Експоненціальне згладжування. Основним параметром експоненціального згладжування є параметр α, який приймає значення в інтервалі $0.1 \le \alpha \le 0.3$. Необхідно, здійснити згладжування того самого ряду зі значеннями параметра α = 0.1, 0.15, 0.2, 0.25, 0.3 і у всіх цих випадках знайти для кожного згладжування кількість поворотних точок та коефіцієнти кореляції між оригінальними значеннями та згладженими.
5. Медіанне згладжування. В цьому випадку використати ті ж самі розміри інтервалу згладжування та операції як в пункті 2.
6. Отримані результати подати у формі зведених графіків, діаграм та таблиць.

3. Кореляція даних

1. Побудова кореляційного поля. Як вхідні дані використати звітну таблицю з файлу *.dat (або *.txt) (дані повинні бути впорядковані певним чином).
2. Визначення значення коефіцієнта кореляції.
3. Обчислити кореляційне відношення.
4. Побудувати графіки автокореляційних функцій.
5. Розбити одну з послідовностей на три рівні частини.
6. Побудувати для них кореляційну матрицю.
7. Знайти коефіцієнти множинної кореляції.
8. Автокореляція.

4. Провести кластер ний аналіз даних

1. Відкрити чистий **Лист** табличного процесора Excel.
2. Сформувати з даних таблицю «об'єкт-властивість».
3. Утворити з них близько розташовані «таблицю-оригінал» та «таблицю-копію».
4. Вибрати формулу для розрахунку близькості між об'єктами і місце для неї на **Листі**.
5. Побудувати матрицю близькості.
6. Скопіювати матрицю близькості на новий **Лист**, розірвавши зв'язки з формулами.
7. Вибрати стратегію об'єднання об'єктів у групи
8. Виконати процедуру кластерного аналізу.
9. Побудувати таблицю «об'єднання – вузол – метрика».
10. Побудувати дендрограму.
11. Виконати інтерпретацію результату кластерного аналізу.
12. Оформити звіт.

КП.7. Форма звітності

1. Титулка з назвою роботи.
2. Мета роботи.
3. Короткі теоретичні відомості
4. Форми і методи подання та попереднє статистичне орацювання числових даних часових послідовностей: вхідні дані, хід роботи, результати

4.1. Попереднє опрацювання даних та подання результатів:

4.1.1. Формування файлу даних у формі Excel та Word таблиць.

4.1.2. Графічне подання даних.

4.2. Описова статистика – кількісні характеристики даних.

4.2.1. Побудова гістограми

4.2.2. Побудова кумуляти.

5. Виявлення тенденції часового ряду методами згладжування

5.1. Методи згладжування часових рядів.

5.2. Метод ковзної середньої.

5.3. Метод зваженої ковзної середньої.

5.4. Властивості ковзного середнього.

5.4.1. Лінійне згладжування для $\mathbf{w} = \mathbf{3}$.

5.4.2. Лінійне згладжування для $\mathbf{w} = \mathbf{5}$.

5.4.3. Нелінійне згладжування $\mathbf{w} = \mathbf{7}$.

5.5. Медіанна фільтрація.

5.6. Нормування часових послідовностей.

5.7. Критерії ефективності згладжування часових рядів.

5.8. Формули для зваженого ковзного середнього.

6. Згладжування формулами з Кендела:

- подати узагальнений графік результатів згладжування для однієї реалізації даних, лише один ряд);

- побудувати кореляційну таблицю для всіх інтервалів згладжування, включаючи і ряд оригінальних значень;

- побудувати діаграму поворотних точок для всіх інтервалів згладжування.

6.1. Кореляційний аналіз часових послідовностей.

6.2. Кореляційне поле.

6.3. Коефіцієнт кореляції.

6.4. Кореляційне відношення.

6.5. Властивості кореляційного відношення.

6.6. Кореляційна матриця.

6.7. Автокореляція.

6.8. Автокореляція в часових рядах.

6.9. Розрахунок автокореляції.

7. Згладжування за формулами з Полларда:

- аналогічно до п. 6.

8. Експоненціальне згладжування:

- аналогічно до п. 6.

9. Медіанне згладжування:

- аналогічно з п. 6.

10. Ієрархічний агломеративний кластерний аналіз багатомірних даних: вхідні дані, хід роботи, результати

10.1. Постановка задачі. Виконання цієї роботи полягає у реалізації її двох частин, а саме, побудови матриці близькості на підставі таблиці об'єкт-властивість та проведення самого кластерного аналізу на основі побудованої матриці близькості.

10.2. Побудова матриці близькості.

10.2.1. Формування таблиці «об`єкт- властивість».

10.2.2. Нормування таблиці «об`єкт-властивість».

10.2.3. Вибір метрики для побудови матриці близькості.

10.2.4. Побудова таблиці близькості.

10.3. Проведення агломеративного ієрархічного кластерного аналізу.

10.3.1. Вибір стратегій об'єднання.

10.3.2. Проведення кластерного аналізу.

10.3.3. Побудова дендрограми.

11. Висновок.

Комп'ютерний проект 2. Розрахункова робота реалізації чисельних методів

- Алгоритми розв'язування нелінійних рівнянь
  - Основні поняття та визначення
  - Метод половинного ділення
    - Постановка задачі
    - Алгоритм методу
  - Метод хорд
    - Постановка задачі
    - Алгоритм методу
  - Метод Ньютона (метод дотичних)
    - Постановка задачі
    - Правила визначення рухомого кінця для методу Ньютона
  - Метод ітерацій (метод послідовних наближень)
- Система лінійних алгебраїчних рівнянь
  - Основні поняття та визначення
    - Поняття матриці
    - Приклади матриць
    - Рівність матриць
    - Додавання матриць
    - Множення матриці на число
    - Віднімання матриць
    - Множення матриць
    - Транспонування матриць
    - Обернена матриця
    - Властивості оберненої матриці
  - Класифікація методів розв'язання СЛАР на ПК
  - Особливості методів Гауса
  - Метод Гауса з послідовним виключенням невідомих
  - Метод Крамера
  - Метод оберненої матриці
  - Матричний метод
- Інтерполювання функцій
  - Математична постановка задачі інтерполювання
  - Інтерполяційний многочлен Лагранжа
  - Перша інтерполяційна формула Ньютона для рівновіддалених вузлів інтерполяції
  - Метод Гауса
  - Метод Стірлінга
  - Метод Бесселя
- Розв'язування нелінійних та трансцендентних рівнянь

- Метод половинного ділення
- Метод пропорційних частин (хорд)
- Метод Рибакова
- Метод простих ітерацій

♦ Точність чисельного інтегрування

♦ Методи розв'язку крайових задач
- Метод прогону
- Проекційні методи (на прикладі методу Гальоркіна)

♦ Методи пошуку максимуму функції двох змінних
- Метод попокоординатного спуску
- Метод спірального попокоординатного спуску

♦ Розв'язування чисельними методами диференціальних рівнянь
- Основні поняття та визначення
- Класифікація численних методів розв'язування задачі Коші
- Одноточкові методи розв'язування задачі Коші на ПК
  - Метод Ейлера
  - Метод Рунге-Кутта

♦ Градієнтні методи пошуку екстремуму
- Лінії рівня
- Найпростіший градієнтний метод оптимізації функцій
  - Властивості градієнтів
  - Алгоритм градієнтного методу
  - Вплив величини кроку на градієнтний пошук
  - Критерій закінчення пошуку
  - Недоліки градієнтного методу
  - Обчислення похідних
  - Модифікації алгоритмів градієнтного методу
- Метод найшвидшого спуску
- Метод Ньютона
- Модифікований метод Ньютона
- Методи змінної метрики

♦ Диференціальні рівняння в частинних похідних
- Різницеві методи розв'язування диференціальних рівнянь у частинних похідних
- Етапи чисельного розв'язування диференціальних рівнянь (ДР) у частинних похідних на ПК
- Гіперболічні рівняння
- Параболічні рівняння
- Приклади розв'язування диференціальних рівнянь в частинних похідних в середовищі MathCAD

♦ Контрольні питання

♦ Задачі для самостійної роботи

♦ Завдання до розрахункової роботи

Сучасний розвиток науки та обчислювальної техніки характеризується все більш зростаючим рівнем використання комп'ютерних моделей як для дослідження поведінки явищ та процесів, що оточують людину, так і для розв'язування практичних задач, що

пов'язані з управлінням та прогнозуванням [1-93, 103-109, 112-114, 125-136, 293-302]. Вивчення навчальної дисципліни «Чисельні методи в комп'ютерних науках» дозволяє студентам оволодіти знаннями в галузі практичних методів вирішення математичних проблем, які виникають у процесі інженерної діяльності та моделюванні фізичних систем; засвоїти способи обчислення на сучасних комп'ютерах із застосуванням пакетів спеціальних прикладних програм. Об'єктом вивчення навчальної дисципліни є типові математичні задачі, до яких зводиться розв'язування практичних проблем, що виникають в процесі розроблення інформаційних систем та систем моделювання. Предметом вивчення навчальної дисципліни є чисельні методи розв'язування типових математичних задач. Мета навчального посібника – ознайомлення студентів з постановками основних математичних задач і чисельних методів їх розв'язування, набуття студентами навичок реалізації на комп'ютері чисельних методів, навичок роботи з відомими комп'ютерними математичними пакетами.

Чисельні методи – методи наближеного або точного розв'язування задач прикладної математики, які ґрунтуються на побудові послідовності дій над скінченною множиною чисел [94-102, 110-111, 137-292, 303-320]. Основні вимоги – чисельні методи мають бути стійкими та збіжними [234]. Основне питання теорії чисельних методів: отримання чисельних методів, які задовольняють вимоги високої точності, стійкості та економічності. Знаходження чисельних методів, що відповідають цим вимогам, є складною задачею оптимізації чисельних методів. Статистичне опрацювання експериментальних даних зазвичай ґрунтується на граничних теоремах теорії ймовірностей та обчисленні порівняльних оцінок. Однак, для підвищення якості оцінок необхідна велика кількість даних, обсяг обчислень може виявитися дуже великим. Чисельні методи націлені на скорочення кількості обчислень при збереженні якості результатів. До найбільш ефективних чисельних методів в цій галузі відносяться методи, які застосовують швидке перетворення Фур'є. Для розв'язування задач апроксимації та обчислення функцій різних класів застосовують чисельні методи інтерполювання, найменших квадратів, ортогоналізації, врівноваження значень, умовної мінімізації тощо. Найбільш актуальними є методи кусково-многочленної та раціональної сплайнової апроксимації, а також адаптивної апроксимації та нелінійної за параметром апроксимації [91-93]. Інженеру часто необхідно вирішувати алгебраїчні та трансцендентні рівняння і системи рівнянь, що можуть бути самостійною задачею (наприклад, аналіз рівноваги сил в жорсткій системі балок або дослідження умов та параметрів рівноваги хімічної реакції, тощо) або частиною більш складних задач. В обох випадках практична цінність чисельного методу в значній мірі визначається швидкістю та ефективністю отримання розв'язку. Розглянемо найбільш відомі чисельні методи та ефективні алгоритми розв'язування систем лінійних алгебраїчних рівнянь.

Чисельне інтегрування та диференціювання починається із визначення відповідних операцій. Однак, з урахуванням необхідності економії обсягу обчислень та некоректності задачі диференціювання з'являється велика кількість чисельних методів для різних класів функцій та різного роду вихідних даних. Інженери часто зіштовхуються з диференціальними рівняннями та системами диференціальних рівнянь при розробці нових виробів чи технологічних процесів, так як більша частина законів фізики формалізується саме у вигляді диференціальних рівнянь. Будь-яка задача проектування, яка пов'язана з розрахунком потоків енергії чи руху тіл, в кінцевому результаті зводиться до розв'язку диференціальних рівнянь. Нажаль, лише дуже малу частину з них можливо розв'язати без

використання обчислювальних машин. Тому чисельні методи розв'язку диференціальних рівнянь відіграють важливу роль у практиці інженерних розрахунків.

1. Алгоритми розв'язування нелінійних рівнянь

1.1. Основні поняття та визначення

При вирішенні практичних інженерних задач часто доводиться зустрічатися з розв'язуванням рівнянь виду

$$\varphi(x) = g(x), \tag{КР.1}$$

або $f(x) = 0$ (КР.2)

де $\varphi(x)$, $g(x)$ та $f(x) = 0$ – нелінійні функції, визначені на деякій числовій множині X, яка називається *областю допустимих значень рівняння.*

Рівняння виду (КР.1) або (КР.2) називаються *нелінійними рівняннями.* Всі нелінійні рівняння можна поділити на алгебраїчні та трансцендентні (рис. КР.1)

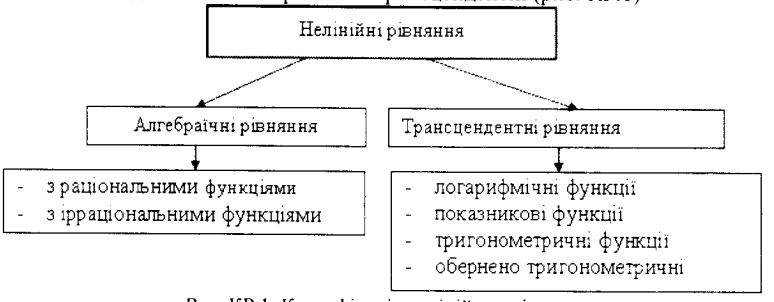


Рис. КР.1. Класифікація нелінійних рівнянь

Функція називається *алгебраїчною*, якщо для отримання значення функції на заданій множині X потрібно здійснити арифметичні операції та піднесення до степеня з раціональним або ірраціональним показником. Рівняння, які містять алгебраїчні функції, називаються *нелінійними алгебраїчними рівняннями.* До трансцендентних функцій відносять всі неалгебраїчні функції: показникові, логарифмічні, тригонометричні, обернені тригонометричні та інші. Нелінійні рівняння, які містять трансцендентні функції називаються *нелінійними трансцендентними рівняннями.*

Розв'язком нелінійного рівняння на комп'ютері називається вектор $\overline{X}$, координати якого $\{x_1, x_2, ..., x_n\}$ при підстановці в початкове рівняння перетворюють його в тотожність.

В нелінійному рівнянні виду

$$a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + ... + a_{n-1} x + a_n = 0 \tag{КР.3}$$

i-та координата вектора $\overline{X} = \{x_1, x_2, ..., x_n\}$ називається i-тим коренем рівняння. a_1, a_2, ..., a_m – коефіцієнтами рівняння (КР.3). Застосування будь-яких чисельних методів до знаходження розв'язку рівняння (КР.1) або (КР.2) передбачає два етапи: відокремлення коренів та уточнення відокремленого кореня.

Розглянемо основні чисельні алгоритми наближеного знаходження коренів нелінійних рівнянь на прикладі знаходження одного із коренів рівняння $x - \cos(x) = 0$ з точністю 0,0001. Відокремити корені рівняння означає знайти множину відрізків, що не перетинаються між собою, кожен з яких містить лише один корінь. Відокремлюють корені, використовуючи таблицю значень або графік функції $f(x)$. Для уточнення коренів нелінійного рівняння з заданою похибкою ε на деякому відрізку $[a,b]$ в інженерній практиці найчастіше використовують:

- метод половинного ділення (метод бісекції);
- метод хорд (метод пропорційних частин);
- метод дотичних (метод Ньютона);
- комбінований метод (метод хорд та дотичних);
- метод ітерацій (метод послідовних наближень).

Всі ці методи є ітераційними, тобто побудовані на алгоритмах, в яких одна з їх частин повторюється багаторазово, при чому кількість повторень залежить від початкових даних (від заданої користувачем похибки, від відрізка дослідження та інше). Розглянемо особливості цих методів та алгоритмів, на яких вони базуються.

1.2. Метод половинного ділення

1.2.1. Постановка задачі

Нехай маємо рівняння $f(x) = 0$, де $f(x)$ – неперервна, монотонна нелінійна функція, яка має на відрізку $[a.b]$ єдиний корінь ξ, тобто добуток $f(a) \cdot f(b) < 0$, причому $b - a > \varepsilon$, де ε – задана похибка обчислень. Потрібно знайти значення кореня ξ з заданою похибкою ε (рис. КР.2).

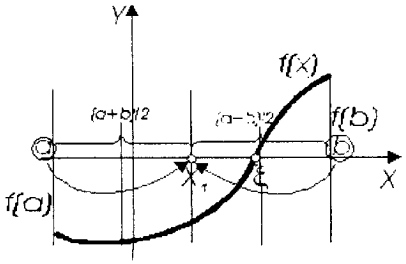


Рис. КР.2. Графічна інтерпретація методу половинного ділення

Алгоритм методу (рис. КР.2) полягає в багатократному діленні навпіл і звужуванні досліджуваного відрізка $[a.b]$, який отримали в результаті попереднього дослідження функції $f(x)$ (відокремлення коренів). Метод половинного ділення – це найпростіший метод уточнення кореня рівняння. Він сходиться для будь-яких неперервних функцій $f(x)$, в тому числі і недиференційованих. Швидкість сходження невелика $N \approx \log_2 \frac{b-a}{\varepsilon}$

1.2.2. Алгоритм методу

1. На відрізку $[a,b]$ вибираємо точку x_0, яка ділить його на два рівних відрізка $[a,x_0]$ і $[x_0,b]$, довжини яких знаходиться за формулою $x_0 = \frac{a+b}{2}$.

2. Перевіряємо чи $f(x_0)=0$, якщо так, то x_0 – точний корінь початкового рівняння і переходимо до пункту 6.

3. У випадку, коли $f(x_0)\neq 0$, то з двох отриманих відрізків $[a,x_0]$ і $[x_0,b]$ вибираємо той, на кінцях якого функція $f(x)$ приймає значення протилежних знаків. Тобто, якщо $f(x_0)\cdot f(a)<0$, то залишаємо відрізок $[a,x_0]$ і точку b переносимо в точку x ($b=x_0$); якщо $f(x_0)\cdot f(a)<0$, то залишаємо відрізок $[x_0,b]$ і переносимо точку a в точку x ($a=x_0$) і переходимо до пункту 1.

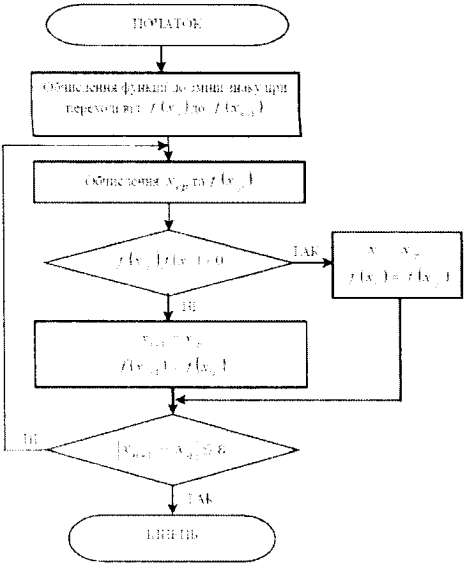
4. Процес ділення відрізка навпіл виконується доти, поки на якомусь етапі середина відрізка буде коренем або буде виконана умова закінчення ітераційного процесу: $|b-a|<\varepsilon$.

5. У цьому випадку за наближене значення кореня вибирають $x=\frac{b-a}{2}$.

6. Вивід результатів. Кінець алгоритму.

Відомо, що при цьому похибка не перевищує $\frac{b-a}{2^{k+1}}$, де k – число ітерацій.

Схема алгоритму розв'язування нелінійного рівняння методом половинного ділення представлена на рисунку КР.3.



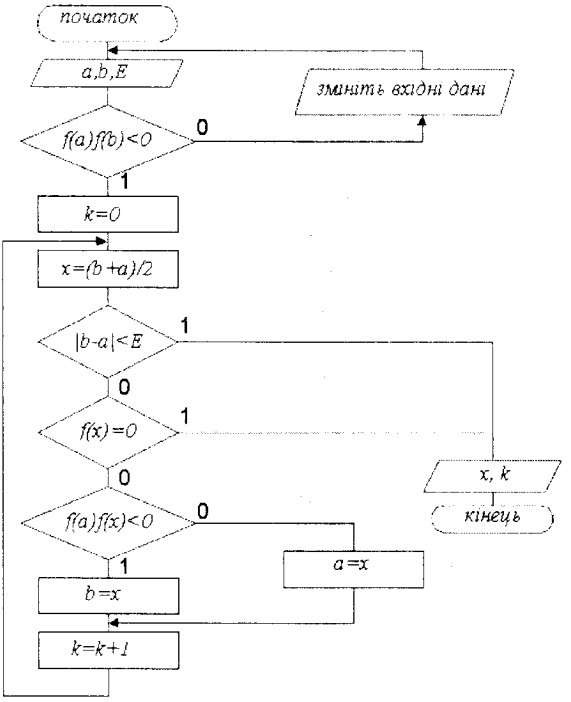


Рис. КР.3. Схема алгоритму розв'язування нелінійного рівняння методом половинного ділення

1.3. Метод хорд

Метод хорд є одним з найбільш поширених методів розв'язування алгебраїчних і трансцендентних рівнянь. В літературі він також зустрічається під назвою *метод лінійного інтерполювання* і *метод пропорційних частин*.

1.3.1. Постановка задачі

Розглянемо рівняння $f(x)=0$, де $f(x)$ – неперервна, монотонна нелінійна функція, яка на відрізку $[a,b]$ монотонна, диференційована і має єдиний корінь ξ, тобто добуток $f(a)\cdot f(b)<0$, причому $b-a>\varepsilon$, де ε – задана похибка обчислень. Потрібно знайти наближене значення кореня ξ із заданою похибкою ε.

Суть методу хорд полягає в тому, що на достатньо малому відрізку $[a,b]$ дуга функції $f(x)$ замінюється хордою ab, яка її стягує. За наближене значення кореня приймається точка x_1 перетину хорди з віссю Ox (рис.КР.4а).

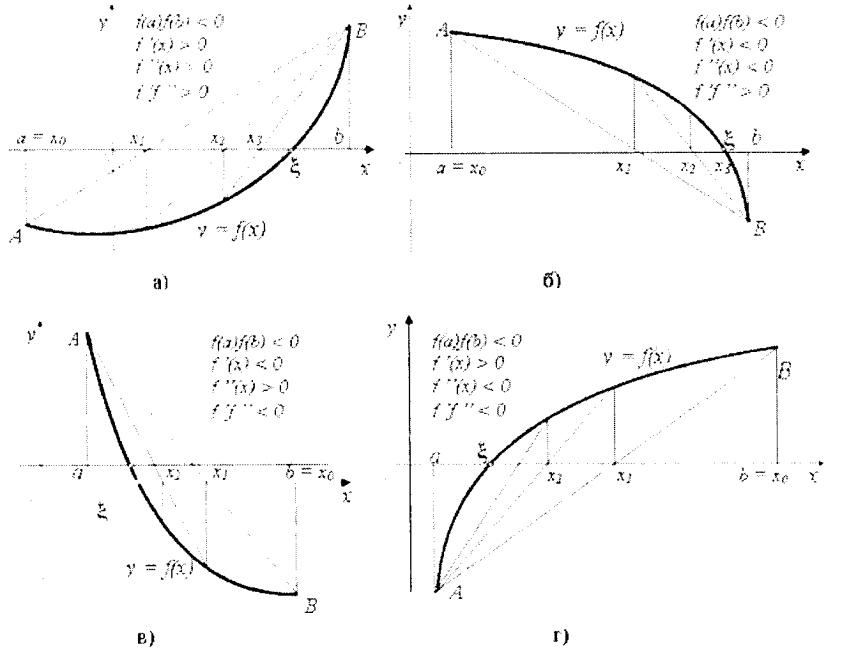


Рис.КР.4. Графічна інтерпретація методу хорд і процедури визначення рухомого кінця хорди

Рівняння хорди, яка проходить через точки має вигляд

$$\frac{y-f(x)}{f(b)-f(a)}=\frac{x-a}{b-a}. \qquad \text{(КР.4)}$$

Знайдемо значення x_1, для якого $y=0$, тобто для нерухомого кінця:

$$x_1 = a-\frac{f(a)\cdot(b-a)}{f(b)-f(a)}. \qquad \text{(КР.5)}$$

Ця формула називається формулою методу хорд. Тепер корінь ξ знаходиться всередині відрізка $[x_1, b]$. Значення кореня x_1 можна уточнити за допомогою метода хорд на відрізку $[x_1, b]$, тоді нове наближене значення кореня x_2 знаходиться за формулою $x_2 = x_1 - \frac{f(x_1)\cdot(b-x_1)}{f(b)-f(x_1)}$. Аналогічно, для довільного $(i+1)$-го наближення до точного значення кореня ξ цього рівняння використовується формула:

$$x_{i+1} = x_i - \frac{f(x_i)\cdot(b-x_i)}{f(b)-f(x_i)}. \qquad \text{(КР.6)}$$

Процес стягування хордою продовжується багаторазово доти, поки не одержано наближений корінь із заданою точністю

$$| x_{i+1} - x_i | < \varepsilon . \qquad \text{(КР.7)}$$

де x_{i+1}, x_i – наближені значення коренів рівняння $f(x) = 0$, відповідно на $(i+1)$ та i-му ітераційному кроці; ε – задана точність обчислень.

Слід відзначи, що розглянутий випадок (рис.КР.4а) перетину функції $f(x)$ на відрізку $[a,b]$ не єдиний. Існує ще три варіанти перетину функції, кожен з яких відрізняється напрямком побудови хорд і відповідно рухомими кінцями відрізку. Наприклад, на рис.КР.4а,б рухомий кінець відрізку a, а на рис.КР.4в,г рухомий кінець – b і відповідно формула (КР.5) для нього має вигляд: $x_1 = b - \dfrac{f(b)\cdot(b-a)}{f(b)-f(a)}$.

1.3.2. Алгоритм методу

Для автоматизації цього алгоритму необхідно розробити правило для автоматичного вибору рухомого кінця хорди і відповідно формули для обчислення наближеного значення кореня. Існує два правила визначення рухомого кінця хорди.

Правило КР.1. Нерухомим кінцем відрізка є той, для якого знак функції співпадає зі знаком другої похідної. Якщо $f(b)\cdot f''(x) > 0$, то нерухомим є кінець b, а всі наближення до кореня ξ лежать зі сторони кінця a. Якщо $f(a)\cdot f''(x) > 0$, то нерухомим є кінець a, а всі наближення до кореня ξ лежать зі сторони кінця b (рис.КР.4.а,б,в,г).

Правило КР.2. Якщо добуток першої на другу похідну функції $f(x)$ більший за нуль: $f' \cdot f'' > 0$, то рухомий кінець a; якщо добуток першої на другу похідну менший за нуль: $f' \cdot f'' < 0$, то рухомий кінець b. Схема алгоритму розв'язування нелінійного рівняння методом хорд представлена на рис. КР.5.

1.4. Метод Ньютона (метод дотичних)

Метод послідовних наближень, розроблений Ньютоном, дуже часто використовується при побудові ітераційних алгоритмів. Його популярність обумовлена тим, що на відміну від двох попередніх методів замість інтерполяції за двома значенням функції в методі Ньютона здійснюється екстраполяція за допомогою дотичної до кривої в одній точці.

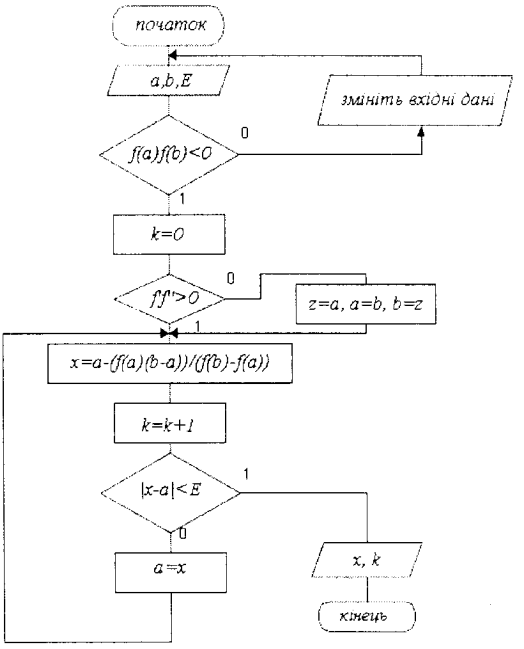


Рис. КР.5. Схема алгоритму розв'язування нелінійного рівняння методом хорд

1.4.1. Постановка задачі

Нехай корінь рівняння $f(x)=0$ є на відрізку $[a,b]$, на якому нелінійна функція $f(x)$ монотонна і має різні знаки на кінцях відрізку, причому похідні $f'(x)$ та $f''(x)$ неперервні та зберігають постійні знаки на всьому відрізку $[a,b]$. Потрібно знайти наближене значення кореня ξ з заданою похибкою ε. Геометричний зміст методу Ньютона полягає в тому, що дуга кривої $y=f(x)$ на відрізку $[a,b]$ замінюється дотичною до цієї кривої, а наближене значення кореня визначається як точка перетину дотичної, проведеної з одного з кінців досліджуваного відрізку, з віссю Ox. Рівняння дотичної має вигляд: $x_{i+1}=x_i-\dfrac{f(x_i)}{f'(x_i)}$.

Перший випадок. Нехай $f(a)<0$, $f(b)>0$, $f'(x)>0$, $f''(x)>0$ (рис. КР.6а) або $f(a)>0$, $f(b)<0$, $f'(x)<0$, $f''(x)<0$ (рис. КР.6б). Проведемо дотичну до кривої $y=f(x)$ в точці $B_0(b,f(b))$ і знайдемо абсцису точки перетину дотичної з віссю Ox. Відомо, що рівняння дотичної в точці $B_0(b,f(b))$ має вигляд: $y-f(b)=f'(b)(x-b)$. Припускаючи $y=0$, $x=x_1$, отримаємо

$$x_1=b-\frac{f(b)}{f'(b)}. \qquad \text{(КР.8)}$$

Тепер корінь рівняння є на відрізку $[a, x_1]$. Застосовуючи знову метод Ньютона, проведемо дотичну до кривої в точці $B_1(x_1, f(x_1))$ і отримаємо $x_2 = x_1 - \dfrac{f(x_1)}{f'(x_1)}$ і так далі (рис. КР.6). Цей процес ітераційний, тому формула для будь-якого n-го кроку ітерації має вигляд:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \qquad \text{(КР.9)}$$

В результаті отримана послідовність наближених значень $x_1, x_2, ..., x_n, ...$, кожен наступний член якої ближчий до кореня ξ, ніж попередній. Однак, всі x_n залишаються більшими, ніж істинний корінь ξ, тобто x_n – наближене значення кореня ξ з надлишком. Процес визначення кореня продовжується багаторазово доти, поки не одержано наближений корінь із заданою точністю

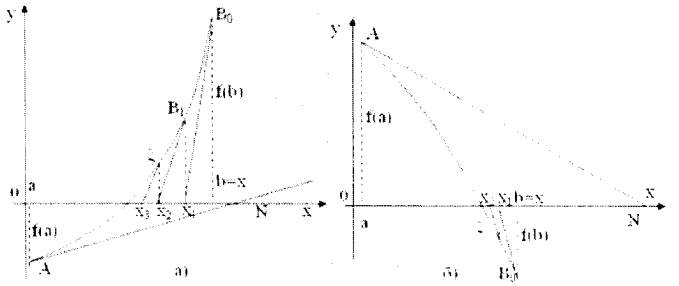


Рис. КР.6. Геометричний зміст методу Ньютона для випадків, коли:
а) функція, яка досліджується, ввігнута ($f'(x) > 0$, $f''(x) > 0$)
б) функція, яка досліджується, опукла ($f'(x) < 0$, $f''(x) < 0$)

Другий випадок. Нехай $f(a) < 0$, $f(b) > 0$, $f'(x) > 0$, $f''(x) < 0$ (рис. КР.7а) або $f(a) > 0$, $f(b) < 0$, $f'(x) < 0$, $f''(x) > 0$ (рис. КР.7б).

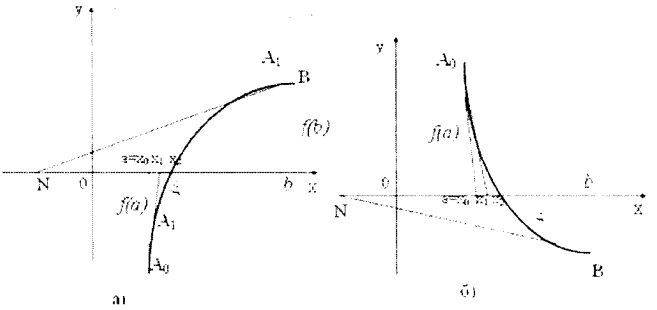


Рис. КР.7. Геометричний зміст методу Ньютона для випадків, коли:

а) функція, яка досліджується, опукла ($f'(x)>0$, $f''(x)<0$)

б) функція, яка досліджується, ввігнута ($f'(x)<0$, $f''(x)>0$)

Якщо провести дотичну до кривої $y=f(x)$ в точці B, то вона перетне вісь абсцис в точці, яка не належить відрізку $[a,b]$. Тому проведемо дотичну в точці $A_0(a,f(a))$ і запишемо рівняння дотичної в цій точці: $y-f(a)=f'(a)(x-a)$.

Врахувавши, що $y=0$, $x=x_1$, отримаємо

$$x_1=a-\frac{f(a)}{f'(a)}. \qquad \text{(КР.10)}$$

Корінь ξ знаходиться тепер на відрізку $[x_1,b]$. Застосовуючи знову метод Ньютона, проведемо дотичну до кривої в точці $A_1(x_1,f(x_1))$ і отримаємо $x_2=x_1-\frac{f(x_1)}{f'(x_1)}$, і в загальному:

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}. \qquad \text{(КР.11)}$$

В результаті отримаємо послідовність наближених значень $x_1,x_2,...,x_n,...$, кожен наступний член якої ближчий до істинного кореня ξ, ніж попередній, тобто x_n – наближене значення кореня ξ з недостачею. Порівнюючи формули (КР.10), (КР.11) з раніше виведеними, (а також враховуючи випадки, які розглядаються на рисунках КР.6а,б помічаємо, що вони відрізняються одна від одної тільки вибором початкового наближення: в першому випадку за x_0 приймався кінець b відрізка, в другому – кінець a. При виборі початкового наближення кореня необхідно використовувати наступне правило: *за початкову точку слід вибирати той кінець відрізка* $[a,b]$, *в якому знак функції співпадає зі знаком другої похідної*. В першому випадку $f(b)\cdot f''(x)>0$ і початкова точка $b=x_0$, в другому – $f(a)\cdot f''(x)>0$ і за початкове наближення беремо $a=x_0$. Для оцінки похибки можна скористатися загальною формулою

$$|\xi-x_n|\le\frac{|f(x_n)|}{m}, \qquad \text{(КР.12)}$$

де $m=\min\limits_{[a,b]}|f'(x)|$ (ця формула підходить і до методу хорд).

У випадку, коли відрізок $[a,b]$ настільки малий, що на ньому виконується умова $M_2<2m_1$, де $M_2=\max\limits_{[a,b]}|f''(x)|$, а $m_1=\min\limits_{[a,b]}|f'(x)|$, точність наближення на n-му кроці інтерполяційного процесу оцінюється так: $|x_n-x_{n-1}|<\varepsilon$, $|\xi-x_n|<\varepsilon^2$.

Якщо похідна $f'(x)$ мало змінюється на відрізку $[a,b]$, то для спрощення обчислень можна користуватися формулою

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)}, \qquad \text{(КР.13)}$$

тобто значення похідної в початковій точці достатньо обчислити тільки один раз.

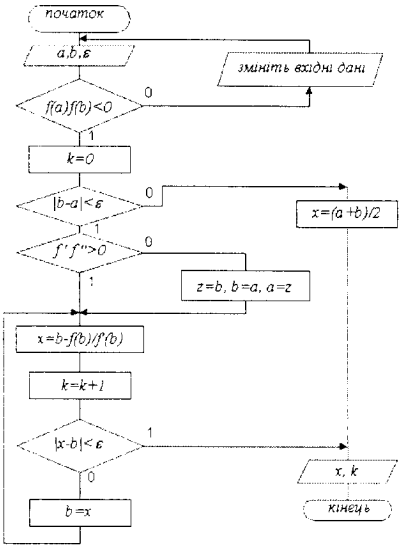


Рис. КР.8. Схема алгоритму розв'язування нелінійного рівняння методом дотичних

Процес побудови дотичної продовжується багаторазово доти, поки $|x_{i+1} - x_i| < \varepsilon$, де ε – задана точність обчислень; x_{i+1}, x_i – наближені значення кореня рівняння $f(x) = 0$, відповідно на $(i+1)$ та i-тому ітераційному кроці. На рисунку КР.8 подана схема алгоритму цього методу.

1.4.2. Правила визначення рухомого кінця для методу Ньютона

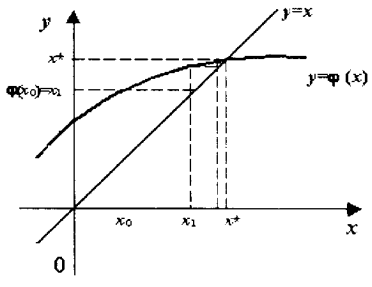


Рис .КР.9. Точка x^* – розв'язок рівняння $x = \varphi(x)$
Побудова точки x_1 по точці x_0

Правило КР.3. Якщо добуток першої на другу похідну функції $f(x)$ більший нуля: $f' \cdot f'' > 0$, то рухомий кінець – b; якщо добуток першої на другу похідну менший нуля: $f' \cdot f'' < 0$, то рухомий кінець – a, тобто дотична будується в кінці a.

Правило КР.4. Якщо знак функції на кінці відрізку співпадає зі знаком другої похідної, то цей кінець відрізку є рухомим, і в цій точці будується дотична.

1.5. Метод ітерацій (метод послідовних наближень)

Припустимо, що рівняння $f(x)=0$ за допомогою деяких тотожних перетворень зведене до вигляду $x=\varphi(x)$. Відзначимо, що таке перетворення можна робити різними способами, і при цьому матимемо різні функції $\varphi(x)$ в правій частині рівняння. Рівняння $f(x)=0$ еквівалентне рівнянню $x=x+\lambda(x)f(x)$ для будь-якої функції $\lambda(x)\neq 0$. Таким чином, можна взяти $\varphi(x)=x+\lambda(x)f(x)$ і при цьому вибрати функцію (або постійну) $\lambda\neq 0$ так, щоб функція $\varphi(x)$ задовольняла тим властивостям, які знадобляться нам для забезпечення знаходження кореня рівняння. Для знаходження кореня рівняння $x=\varphi(x)$ виберемо деяке початкове наближення x_0 (розташоване, по можливості, близько до кореня). Далі будемо обчислювати наступні наближення $x_1, x_2, .., x_i, x_{i+1}, ...$ за формулами $x_1=\varphi(x_0), x_2=\varphi(x_1)$, і так далі, тобто використовуючи кожне обчислене наближення до кореня як аргумент функції $\varphi(x)$ в черговому обчисленні. Такі обчислення за однією і тією ж формулою $x_{i+1}=\varphi(x_i)$, коли отримане на попередньому кроці значення використовується на наступному кроці, називаються ітераціями. Ітераціями називають часто і самі значення x_i, отримані в цьому процесі (тобто, в нашому випадку, послідовні наближення до кореня). Відзначимо той факт, що x^* – корінь рівняння $x=\varphi(x)$, означає, що x^* є абсциса точки перетину графіка $y=\varphi(x)$ з прямою $y=x$. Якщо ж при будь-якому x_0 обчислене значення $x_1=\varphi(x)$ і взяте за новий аргумент функції, то це означає, що через точку графіка $(x_0,\varphi(x_0))$ проводиться горизонталь до прямої $y=x$, а звідти опускається перпендикуляр на вісь. Там і знаходитиметься новий аргумент x_1.

Прослідкуємо, як змінюються послідовні наближення x_i при різних варіантах взаємного розташування графіка $y=\varphi(x)$ і прямої $y=x$.

1) Графік $y=\varphi(x)$ розташований, принаймні в деякому околі кореня , що включає початкове наближення x_0, в деякому куті зі сторонами, що мають нахил менше $\frac{\pi}{4}$ до горизонталі (тобто сторони кута –прямі $y=f(x^*)\pm k(x-x^*)$, де $0<k<1$):

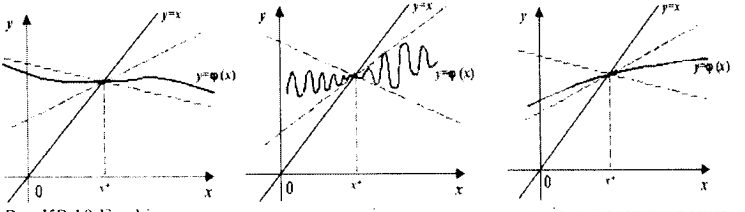


Рис.КР.10.Графік перетинає пряму $y=x$ під малим кутом: варіанти розташування.

Якщо припустити додатково, що функція $\varphi(x)$ має похідну $\varphi'(x)$, то цей випадок відповідає тому, що виконується нерівність $|\varphi'(x)|<1$, при x, близьких до

кореня x^*. Простежимо в цьому випадку за поведінкою послідовних наближень $x_0, x_1, \ldots$

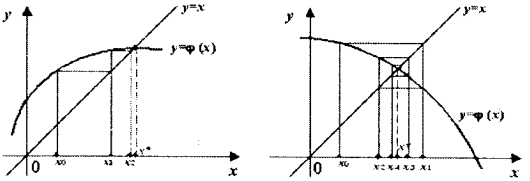


Рис. КР.11. Наближення, що збігаються до кореня у випадку $|\varphi'(x)| < 1$.

Ми бачимо, що кожне наступне наближення x_{i+1} буде в цьому випадку розташовано ближче до кореня x^*, ніж попереднє x_i. При цьому, якщо графік при $x<x^*$, лежить нижче за горизонталь $y = \varphi(x^*)$, а при $x>x^*$– вище за неї (що, у разі наявності похідної, вірно, якщо $0 < \varphi'(x) < 1$), то наближення x_i поводяться монотонно: якщо $x_0<x^*$, то послідовність $\{x_i\}$ монотонно зростає і прямує до x^*, а якщо $x_0>x^*$, то монотонно спадає і також прямує до x^*. Якщо ж графік функції $\varphi(x)$ лежить вище за горизонталь $y = \varphi(x^*)$ при $x<x^*$ і нижче за неї при $x>x^*$ (якщо $-1 < \varphi'(x) < 0$), то послідовні наближення поводяться інакше: вони "скачуть" навколо кореня, з кожним стрибком наближаючись до нього, але так само прямують до x^* при $i \to \infty$. Відзначимо, що якщо функція $\varphi(x)$ не монотонна в околі точки x^*, то послідовні наближення можуть поводитися нерегулярно (тобто не монотонно і не потрапляючи почергово то лівіше, то правіше кореня, а роблячи стрибки відносно кореня при довільних номерах.

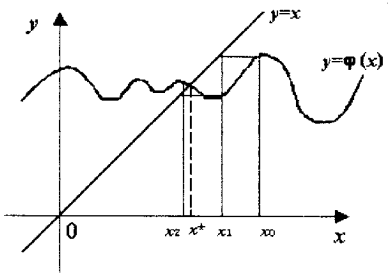


Рис. КР.12. У випадку немонотонної функції ітерації, що сходяться, можуть поводитися нерегулярно

2) Графік $y = \varphi(x)$ розташований, принаймні в деякому околі кореня, що включає початкове наближення x_0, в деякому куті зі сторонами, що мають нахил більше $\frac{\pi}{4}$ до горизонталі (тобто сторони кута – прямі $y = f(x^*) \pm k(x - x^*)$, де $k>1$):

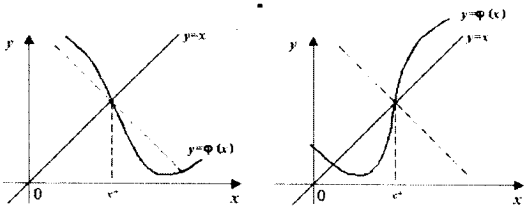


Рис. КР.13. Графік перетинає пряму $y=x$ під великим кутом: варіанти розташування

Якщо функція $\varphi(x)$ має похідну $\varphi'(x)$, то при x, близьких до кореня x^* виконується нерівність $|\varphi'(x)| > 1$.

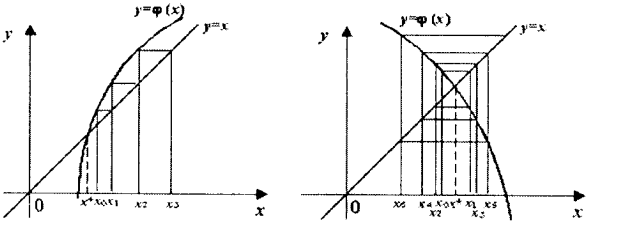
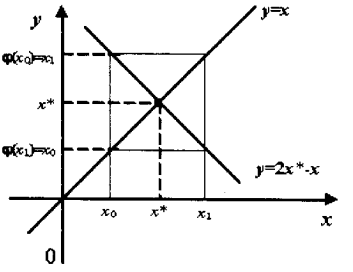


Рис. КР.14. Послідовність $x_0, x_1, x_2, \ldots$ розбіжна у випадку $|\varphi'(x)| > 1$.

Кожна наступна ітерація x_{i+1} буде в цьому випадку розташована далі від кореня x^*, ніж попередня x_i. При цьому, залежно від того, чи перетинає графік пряму $y=x$ "знизу вгору" або "згори донизу", послідовність $\{x_i\}$ монотонно віддаляється від кореня x^* або ж ітерації віддаляються від x^*, потрапляючи почергово то справа, то зліва від кореня. Ще одне зауваження: якщо не виконується ні умова $|\varphi'(x)| < 1$, ні $|\varphi'(x)| > 1$, то ітерації $x_1, x_2, \ldots, x_i, x_{i+1}, \ldots$ можуть зациклюватися. Наприклад, якщо рівняння має вигляд $x=2x^*-x$:



Суть методу полягає у заміні початкового рівняння

$$f(x) = 0 \qquad \text{(КР.14)}$$

еквівалентним рівнянням

$$x = \varphi(x). \qquad \text{(КР.15)}$$

Постановка задачі. Нехай задано рівняння $f(x) = 0$, де $f(x)$ – неперервна нелінійна функція. Потрібно визначити корінь ξ цього рівняння, який знаходиться на відрізку $[a.b]$ із заданою похибкою ε. Виберемо довільним способом $x_0 \in [a.b]$ і підставимо його в праву частину рівняння (КР.15); тоді отримаємо $x_1 = \varphi(x_0)$. Потім це значення x_1 підставимо знову в праву частину рівняння (КР.15) і отримаємо $x_2 = \varphi(x_1)$ (рис. КР.15 а,б). Повторюючи цей процес, отримаємо послідовність чисел $x_n = \varphi(x_{n-1})$. При цьому можливі два випадки:

• послідовність $x_0, x_1, ..., x_n, ...$ збігається, тобто має границю, і тоді ця границя буде коренем рівняння (КР.14);

• послідовність $x_0, x_1, ..., x_n, ...$ розбіжна, тобто не має границі.

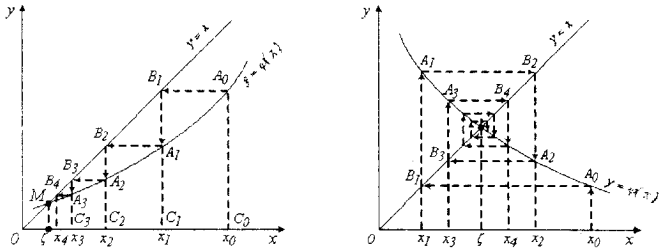


Рис. КР.15. Геометрична інтерпретація методу ітерацій

Умови, при яких ітераційний процес розв'язку нелінійного рівняння методом ітерацій на ПК збігається, сформульовано в наступній теоремі:

Теорема. Нехай на відрізку $[a,b]$ знаходиться єдиний корінь рівняння $x = \varphi(x)$ та у всіх точках цього відрізку похідна $\varphi'(x)$ задовільняє нерівності $|\varphi'(x)| \le q < 1$. Якщо при цьому виконується й умова $a \le \varphi(x) \le b$, то ітераційний процес збігається, а за нульове наближення x_0 можна взяти число з відрізку $[a,b]$.

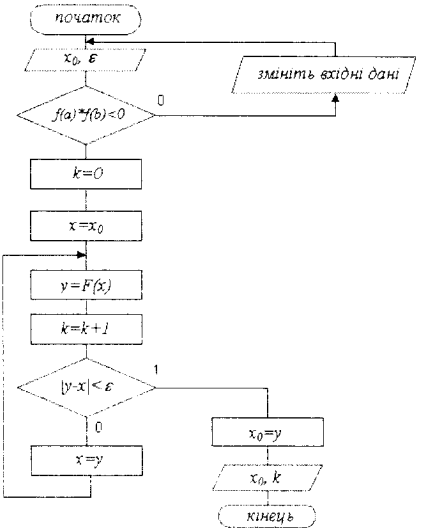


Рис. КР.16. Схема алгоритму розв'язування нелінійного рівняння методом ітерацій

Розв'яжемо один етап ітерацій. Виходячи із заданого на попередньому кроці значення x_{n-1}, обчислюємо $y = \varphi(x_{n-1})$. Якщо $|y - x_{n-1}| > \varepsilon$, візьмемо $x_n = y$ і

виконаємо наступну ітерацію. Якщо ж $|y - x_{n-1}| < \varepsilon$, то обчислення закінчують, за наближене значення кореня приймають величину $x_n = y$. При використанні методу простих ітерацій основною операцією є вибір функції $\varphi(x)$ у рівнянні $x = \varphi(x)$, яку слід підібрати так, щоб $|\varphi'(x)| \le q < 1$ і швидкість збіжності послідовності $\{x_n\}$ до кореня ξ тим вища, чим менше число q. Схема алгоритму методу ітерацій представлена на рис. КР.16.

2. Система лінійних алгебраїчних рівнянь

2.1. Основні поняття та визначення

Системою лінійних алгебраїчних рівнянь (СЛАР) називають систему вигляду:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m = b_2 \\ a_{31}x_1 + a_{32}x_2 + \ldots + a_{3m}x_m = b_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m = b_n \end{cases} \qquad \text{(КР.16)}$$

де x_i, $(i = \overline{1,n})$ – невідомі; b_i, $(i = \overline{1,n})$ – вільні члени системи; a_{ij}, $(i, j = \overline{1,n})$ – коефіцієнти системи. В матричному вигляді рівняння (КР.16) прийме вигляд: $\vec{A} \times \vec{X} = \vec{B}$, де $\vec{X} = (x_1, x_2, \ldots, x_n)$ – вектор невідомих; $\vec{B} = (b_1, b_2, \ldots, b_n)$ – вектор вільних членів; $\vec{A} = (a_{11}, a_{21}, \ldots, a_{mn})$ – матриця коефіцієнтів СЛАР.

Розв'язком системи лінійних алгебраїчних рівнянь (КР.16) називають вектор $\vec{X}$, координати якого $(x_1, x_2, \ldots, x_n)$ при підстановці у систему, яку розв'язують, перетворюють кожне рівняння системи в тотожність. Кількість невідомих m в системі називають *порядком* СЛАР. Систему лінійних алгебраїчних рівнянь називають *сумісною*, якщо вона має хоча б один ненульовий розв`язок. В протилежному випадку СЛАР називають *несумісною*. СЛАР називається *визначеною*, якщо вона має тільки один розв`язок (випадок, коли $m=n$). Систему називають *невизначеною*, якщо вона має безліч розв`язків ( $m \neq n$ ). Система називається *виродженою*, якщо головний визначник системи дорівнює нулю. Система називається *невиродженою*, якщо головний визначник системи не дорівнює нулю. Дві системи називаються *еквівалентними*, якщо ці системи сумісні, визначені і мають однаковий розв`язок. СЛАР можна розв`язати на ПК чисельними методами, якщо вона сумісна, визначена, невироджена.

2.1.1. Поняття матриці

На практиці різну інформацію досить часто представляють масивами у формі таблиць, тобто впорядкованих за рядками та стовбцями чисел.

Означення КР.1. Матрицею називають впорядковану таблицю чисел,

розташованих у вигляді m рядків та n стовпців. Матриці позначають великими літерами A, B, C, D та круглими ( ), квадратними [ ] або прямими подвійними $\| \|$ дужками. Матриця, яка містить m рядків та n стовпців, називається матрицею розміру $m \times n$. Найбільш поширеними формами представлення матриць є такі:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{12} & a_{22} & a_{23} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}, \qquad A = \left[a_{ij} \right], \left(i = \overline{1,m};\ j = \overline{1,n} \right).$$

Числа, що складають матрицю називаються її елементами. При двохіндексному позначенні елементів матриці перший індекс вказує на номер рядка, а другий індекс – номер стовпця, на перетині яких міститься елемент. Так елемент a_{23} знаходиться на перетині другого рядка і третього стовпця. Коли $m = n$, матриця називається квадратною, а число $m = n$ називається її порядком.

2.1.2. Приклади матриць

1) $A = \begin{bmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \end{bmatrix}$ – матриця є прямокутною, має розміри 2×3, причому $a_{11} = 1$, $a_{12} = 2$, $a_{13} = 3$, $a_{21} = 6$, $a_{22} = 5$, $a_{23} = 4$.

2) $B = \begin{bmatrix} -10 & 0 \\ 3 & 4 \end{bmatrix}$ – квадратна матриця другого порядку. Якщо матриця складається з одного рядка, то її називають матрицею-рядком або вектор-рядком і її розміри становлять $(1 \times n)$. Матриця, що складається з одного стовпця, називається матрицею-стовпцем або вектор-стовпцем. Її розміри становлять $(m \times 1)$.

3) $C = \begin{bmatrix} 1 & 3 & -4 & 5 \end{bmatrix}$ – матриця-рядок (вектор-рядок) розміром (1×4).

4) $D = \begin{bmatrix} 1 \\ 7 \\ 8 \\ -10 \end{bmatrix}$ – матриця-стовпець, (вектор-стовпець) розміру (4×1).

Розглянемо матрицю $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}$. Елементи матриці з однаковими індексами рядків і стовпців: a_{11}, a_{22}, a_{33}, …, a_{nn} утворюють *головну діагональ*, а елементи a_{1n}, $a_{2(n-1)}$, $a_{3(n-2)}$, …, a_{n1} утворюють *допоміжну діагональ* матриці. Квадратну матрицю, в якій усі елементи, що лежать поза головною діагоналлю дорівнюють нулю, називають діагональною матрицею.

$$C = \begin{bmatrix} c_1 & 0 & 0 & \cdots & 0 \\ 0 & c_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & c_n \end{bmatrix}.$$

Якщо усі числа c_i $(i = \overline{1,n})$ рівні між собою, матрицю називають скалярною. Діагональну матрицю, в якої всі елементи головної діагоналі дорівнюють одиниці називають *одиничною* і переважно позначають її через $E = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$.

Матриця, в якої усі елементи рівні нулю називається *нульовою*. Нульові матриці можуть мати різні розміри. Квадратна матриця називається *симетричною*, якщо для усіх i та j мають місце рівності $a_{ij} = a_{ji}$.

5) $B = \begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{Vmatrix}$ – діагональна матриця четвертого порядку.

6) $L = \begin{pmatrix} 15 & 0 & 0 \\ 0 & 15 & 0 \\ 0 & 0 & 15 \end{pmatrix}$ – скалярна матриця третього порядку.

7) $M = \begin{bmatrix} 1 & -1 & 3 \\ -1 & 2 & 2 \\ 3 & 2 & 3 \end{bmatrix}$ – симетрична матриця 3-го порядку.

8) $O = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ – нуль-матриця розміру (3×3).

2.1.3. Рівність матриць

Означення КР.2. Дві матриці $A = [a_{ij}]$ та $B = [b_{ij}]$ називаються *рівними* $(A = B)$, якщо вони мають однакові розміри – однакове число стовпців і рядків, а відповідні елементи рівні між собою: $a_{ij} = b_{ij}$.

2.1.4. Додавання матриць

Означення КР.3. *Сумою* двох матриць $A = [a_{ij}]$ та $B = [b_{ij}]$ однакових розмірів називається матриця $C = [c_{ij}]$ того ж розміру, елементи якої дорівнюють сумам відповідних елементів матриць доданків: $C = A + B$, $c_{ij} = a_{ij} + b_{ij}$ $(i = \overline{1,m};\ j = \overline{1,n})$. Операція знаходження суми матриць називається додаванням матриць. Приклад:

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} + \begin{bmatrix} d_1 & d_2 & d_3 \\ f_1 & f_2 & f_3 \end{bmatrix} = \begin{bmatrix} a_1 + d_1 & a_2 + d_2 & a_3 + d_3 \\ b_1 + f_1 & b_2 + f_2 & b_3 + f_3 \end{bmatrix}.$$

Операція додавання матриць є комутативною – має переміщувальну властивість $A + B = B + A$, а також властивість сполуки (асоціативність)

$$(A + B) + C = A + (B + C).$$

Крім того, $A + O = A$, де O – нульова матриця тих же розмірів, що й матриця A.

2.1.5. Множення матриці на число

Означення КР.4. Добутком матриці $A = [a_{ij}]$ на число λ називається матриця $C = [c_{ij}]$ тих же розмірів, що й матриця A, елементи якої отримуємо з відповідних елементів матриці A множенням їх на число λ: $C = \lambda \cdot A$, де $c_{ij} = \lambda \cdot a_{ij}$ і ($i = \overline{1,m}$; $j = \overline{1,n}$). Приклад: $\lambda \cdot \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} = \begin{bmatrix} \lambda \cdot a_1 & \lambda \cdot a_2 & \lambda \cdot a_3 \\ \lambda \cdot b_1 & \lambda \cdot b_2 & \lambda \cdot b_3 \end{bmatrix}$.

2.1.6. Віднімання матриць

Різниця $A - B$ двох матриць однакового (і тільки однакового) розміру визначається наступною рівністю: $A - B = A + (-1) \cdot B$. Приклад:

$$\begin{bmatrix} 4 & 5 \\ 6 & 7 \\ 8 & 9 \end{bmatrix} - \begin{bmatrix} 7 & 1 \\ 3 & 4 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} -3 & 4 \\ 3 & 3 \\ 6 & 9 \end{bmatrix}.$$

2.1.7. Множення матриць

Множення матриці A на матрицю B можна здійснити лише тоді, коли число стовпців матриці A відповідає числу рядків матриці B. Нехай задано дві матриці $A = [a_{ik}]_{m \times n}$ та $B = [b_{kj}]_{n \times p}$.

Означення КР.5. Добутком $A \cdot B$ матриці A розміру $(m \times n)$ та матриці B розміру $(n \times p)$ називається матриця C розміру $(m \times p)$, елементи якої c_{ij} дорівнюють сумі добутків елементів i-го рядка матриці A на відповідні елементи j-го стовпця матриці B. Отже, елементи матриці C знаходять за формулою

$$c_{ij} = a_{i1} \cdot b_{1j} + a_{i2} \cdot b_{2j} + a_{i3} \cdot b_{3j} + \cdots + a_{in} \cdot b_{nj} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}, \quad (i = \overline{1,m};\ j = \overline{1,p})$$

Приклад КР.1. Обчислити $A \cdot B$, якщо $A = \begin{bmatrix} 3 & 0 & 1 \\ 2 & 1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 2 & 3 \\ 4 & 3 & 5 \end{bmatrix}$.

$$A \cdot B = \begin{bmatrix} 3 \cdot 1 + 0 \cdot 1 + 4 \cdot 1 & 3 \cdot 0 + 0 \cdot 2 + 1 \cdot 3 & 3 \cdot 2 + 0 \cdot 3 + 1 \cdot 5 \\ 2 \cdot 1 + 1 \cdot 1 + 2 \cdot 4 & 2 \cdot 0 + 1 \cdot 2 + 2 \cdot 3 & 2 \cdot 2 + 1 \cdot 3 + 2 \cdot 5 \end{bmatrix} = \begin{bmatrix} 7 & 3 & 11 \\ 11 & 8 & 17 \end{bmatrix}.$$

Для квадратних матриць операція множення справедлива завжди, коли матриці мають однакові порядки. В загальному випадку операція множення матриць не

відповідає переміщувальній властивості – $A \cdot B \neq B \cdot A$. Приклад:

$$A \cdot B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 8 & -2 \\ 18 & -4 \end{bmatrix}. \quad B \cdot A = \begin{bmatrix} 2 & 0 \\ 3 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 0 & 2 \end{bmatrix}.$$

В частинному випадку, коли $A \cdot B = B \cdot A$, матриці називаються перестановочними, або такими, що комутують між собою. Одинична матриця комутативна з будь-якою квадратною матрицею того ж порядку.

Операції множення матриць мають такі властивості

1. Сполучний закон: $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

2. Розподільний закон: $(A + B) \cdot C = A \cdot C + B \cdot C, \quad A \cdot (B + C) = A \cdot B + A \cdot C.$

Операція множення матриць природним чином розповсюджується на випадок декількох співмножників. На цій підставі для квадратних матриць вводять поняття степеня матриці: $A^2 = A \cdot A$; $A^n = \underbrace{A \cdot A \cdot \ldots \cdot A}_{n}$.

2.1.8. Транспонування матриць

Означення КР.6. Транспонуванням матриці A називається операція заміни рядків цієї матриці її стовпцями зі збереженням їх номерів. Матриця отримана таким чином з матриці A, називається *транспонованою* відносно матриці A та позначається: A', A^T.

Приклад КР. 2. Задано матрицю A розміру $(m \times n)$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Переставимо в ній рядки зі стовпцями зі збереженням номерів

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}.$$

Таким чином отримуємо транспоновану матрицю відносно матриці A. Якщо вихідна матриця має розміри $(m \times n)$, то матриця A^T транспонована відносно матриці A буде мати розміри $(n \times m)$. В частинному випадку для вектора-рядка транспонованою матрицею буде вектор-стовпець.

Транспонована матриця має такі властивості.

1. Двічі транспонована матриця є вихідною $A^{TT} = (A^T)^T = A$.

2. Транспонована матриця суми дорівнює сумі транспонованих матриць доданків $(A + B)^T = A^T + B^T$.

3. Транспонована матриця добутку дорівнює добутку транспонованих матриць співмножників, взятих у зворотному порядку $(A \cdot B)^T = B^T \cdot A^T$.

4. Для симетричної матриці $A^T = A$.

2.1.9. Обернена матриця

Нехай A – квадратна матриця n-го порядку: $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}$, а E – одинична квадратна матриця того ж порядку.

Означення КР.7. Оберненою відносно заданої квадратної матриці A називається матриця A^{-1}, яка після множення на вихідну матрицю A як зліва, так і справа дає одиничну матрицю: $A^{-1} \cdot A = A \cdot A^{-1} = E$.

2.1.10. Властивості оберненої матриці

1. Матриця A має обернену (може бути оберненою), якщо її визначник $\det A \neq 0$.

2. Визначник оберненої матриці дорівнює величині оберненої до визначника вихідної матриці, тобто $\det A^{-1} = \dfrac{1}{\det A}$.

3. Обернена матриця добутку матриць дорівнює добутку обернених матриць, узятих в зворотному порядку $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.

4. Транспонована обернена матриця дорівнює оберненій від транспонованої даної матриці, тобто $\left(A^{-1}\right)^T = \left(A^T\right)^{-1}$.

2.2. Класифікація методів розв'язування СЛАР на ПК

В загальному вигляді система з m лінійних рівнянь з n невідомими записується так

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1j}x_j + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2j}x_j + \cdots + a_{2n}x_n = b_2, \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{ij}x_j + \cdots + a_{in}x_n = b_i, \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mj}x_j + \cdots + a_{mn}x_n = b_m. \end{cases} \qquad \text{(КР.17)}$$

Рівняння системи вважають пронумерованими: перше, друге, … , m-те. Числа $x_1, x_2, \ldots, x_n$ є *невідомими* системи: $a_{11}, a_{12}, \ldots, a_{mn}$ – *коефіцієнтами* при невідомих системи. Коефіцієнт при невідомому x_{ij} в i-му рівнянні позначається через a_{ij}, де перший індекс i вказує номер рівняння, в якому знаходиться даний коефіцієнт, а

другий індекс j – номер невідомого, при якому знаходиться даний коефіцієнт. Наприклад, коефіцієнт a_{23} знаходиться в другому рівнянні системи при невідомому x_3. Числа $b_1, b_2, \ldots, b_m$ є *вільними членами* системи. Коротко систему (КР.17) можна записати так: $\sum_{j=1}^{n} a_{ij} x_j = b_i$, $(i = 1, 2, \ldots, m)$.

Розв'язком лінійної системи (КР.17) називається будь-яка сукупність чисел $\alpha_1, \alpha_2, \ldots, \alpha_n$, яка, будучи підставлена замість невідомих $x_1, x_2, \ldots, x_n$ в рівняння цієї системи, перетворює усі рівняння системи в тотожності. Система лінійних рівнянь (КР.17) є *сумісною*, якщо вона має розв'язок. Якщо система лінійних рівнянь не має розв'язку, то вона є *несумісною* (або *суперечливою*). Сумісна система лінійних рівнянь може мати один або декілька розв'язків і називатися визначеною, якщо вона має один єдиний розв'язок, і *невизначеною*, якщо має більше одного розв'язку.

Дві системи лінійних рівнянь з одним і тим же числом невідомих називаються *еквівалентними*, якщо вони або обидві несумісні, або обидві сумісні і мають одні і ті ж розв'язки. Наведені нижче три типи перетворень називаються *елементарними перетвореннями системи лінійних рівнянь*. Це такі:

1) перестановка двох рівнянь системи;

2) множення обох частин рівняння системи на будь-яке відмінне від 0 число;

3) додавання до обох частин одного рівняння відповідних частин другого рівняння, помножених на будь-яке число.

Виконання елементарних перетворень рівносильне представленню однієї змінної через інші. Для знаходження розв'язків лінійної системи рівнянь існує багато методів. Найбільш поширеними серед них є метод Крамера, метод Гауса та матричний метод. Для розв'язування СЛАР на ПК традиційно використовують дві групи чисельних методів, що представлені на рис. КР.17:

1. **точні** (метод Гауса, метод Гауса з вибором головного елементу, метод Гауса з одиничною матрицею, метод Гауса з перетвореною матрицею, метод Гауса-Халецького, метод Гауса-Жордана, метод Крамера);
2. **наближені** (метод послідовних ітерацій, метод Гауса-Зейделя, метод векторів зміщень).

До **точних методів** належать методи, які дозволяють отримати точний розв'язок системи (КР.17) за відповідну кількість операцій перетворення без врахування похибок заокруглення.

До **наближених методів** належать методи, які дозволяють отримати розв'язок системи (КР.17) у вигляді границі послідовності векторів $\lim_{k \to \infty} \left\{ \overline{X}^0, \overline{X}^1, \overline{X}^2, \ldots \overline{X}^n \right\}$, яка збігається до точного розв'язку системи, де:

$$\overline{X}^0 = \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \dots \\ x_n^{(0)} \end{bmatrix}, \quad \overline{X}^1 = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ \dots \\ x_n^{(1)} \end{bmatrix}, \quad \dots \quad \overline{X}^n = \begin{bmatrix} x_1^{(n)} \\ x_2^{(n)} \\ \dots \\ x_n^{(n)} \end{bmatrix}$$



Рис. КР.17. Класифікація чисельних методів

2.3. Особливості методів Гауса

Найбільш відомими з точних методів розв'язування системи лінійних алгебраїчних рівнянь (КР.17) є методи Гауса, суть яких полягає в тому, що система рівнянь, яка розв'язується, зводиться до еквівалентної системи з верхньою трикутною матрицею. Невідомі знаходяться послідовними підстановками, починаючи з останнього рівняння перетвореної системи. Алгоритми Гауса складаються із виконання однотипних операцій, які легко формалізуються. Однак, точність результату й витрачений на його отримання час у більшості випадків залежить від алгоритму формування трикутної матриці системи. У загальному випадку алгоритми Гауса складаються з двох етапів:

Прямий хід, в результаті якого СЛАР (КР.17), що розв'язується, перетворюється в еквівалентну систему з верхньою трикутною матрицею коефіцієнтів вигляду:

$$\begin{cases} a_{11} \cdot x_1 & + a_{12} \cdot x_2 & + \dots + & a_{1n} \cdot x_n & = b_1 \\ 0 \cdot x_1 & + a_{22} \cdot x_2 & + \dots + & a_{2n} \cdot x_n & = b_2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 \cdot x_1 & + 0 \cdot x_2 & + \dots + & a_{nn} \cdot x_n & = b_n \end{cases} \qquad \text{(КР.18)}$$

Зворотній хід дозволяє визначити вектор розв'язку починаючи з останнього рівняння системи (КР.18) шляхом підстановки координат вектора невідомих, отриманих на попередньому кроці. Відомо декілька різних алгоритмів отримання еквівалентної системи з верхньою трикутною матрицею. Розглянемо найбільш відомі з них.

2.4. Метод Гауса з послідовним виключенням невідомих

Метод Гауса з послідовним виключенням невідомих (базовий метод) засновано на алгоритмі, в основі якого лежить послідовне виключення невідомих вектора $\bar{x}$ з усіх рівнянь, починаючи з $(i+1)$-го, шляхом елементарних перетворень: перемноження обох частин рівняння на будь-яке число, крім нуля; додавання (віднімання) до обох частин одного рівняння відповідних частин другого рівняння, помножених на будь-яке число, крім нуля. Суть алгоритму розглянемо на прикладі системи, яка складається з трьох лінійних алгебраїчних рівнянь з трьома невідомими:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases} \qquad \text{(КР.19)}$$

1. Перевіримо, щоб принаймні один із коефіцієнтів a_{11}, a_{21}, a_{31} не дорівнював нулю. Якщо, наприклад, $a_{11} = 0$, тоді необхідно переставити рівняння так, щоб коефіцієнт при x_1 у першому рівнянні не дорівнював нулю.
2. Обчислюється множник: $M_2 = \dfrac{a_{21}}{a_{11}}$.
3. Перше рівняння системи (КР.19) множиться на M_2 і віднімається від другого рівняння системи, отриманої після перестановки рівнянь, якщо вона була необхідною. Результат обчислення має вигляд:

$$(a_{21} - M_2 a_{11})x_1 + (a_{22} - M_2 a_{12})x_2 + (a_{23} - M_2 a_{13})x_3 = b_2 - M_2 b_1. \qquad \text{(КР.20)}$$

але $a_{21} - M_2 a_{11} = a_{21} - \left(\dfrac{a_{21}}{a_{11}}\right) a_{11} = 0$. Тоді x_1 виключається із другого рівняння.

Позначимо нові коефіцієнти: $a'_{22} = a_{22} - M_2 a_{12}$; $a'_{23} = a_{23} - M_2 a_{13}$; $b'_2 = b_2 - M_2 b_1$. Тоді друге рівняння системи (КР.19) набуває вигляду: $a'_{22}x_2 + a'_{23}x_3 = b_2$. Далі необхідно звільнитися від коефіцієнта a_{31} при x_1 в третьому рівнянні системи (КР.19) за аналогічним алгоритмом

4. Обчислюється множник для третього рівняння: $M_3 = \dfrac{a_{31}}{a_{11}}$.

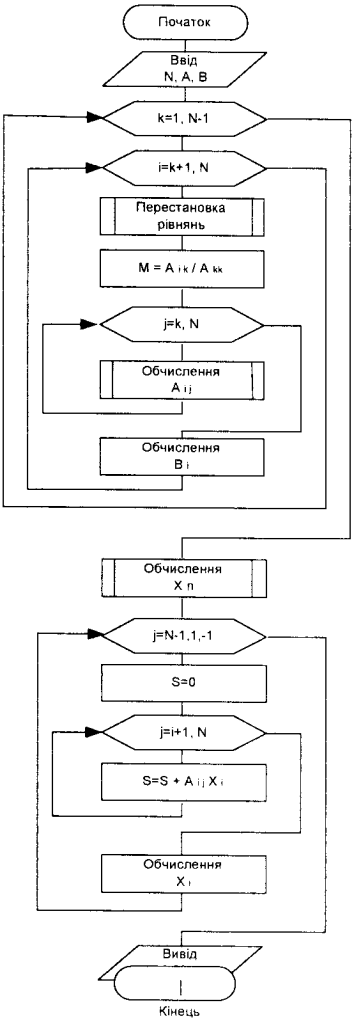


Рис. КР.18. Схема алгоритму розв'язування СЛАР методом Гауса.

5. Перше рівняння системи (КР.19) множиться на M_3 і віднімається від третього рівняння. Коефіцієнт при x_1 стає нулем, і третє рівняння набуває вигляду: $a'_{32}x_2 + a'_{33}x_3 = b_3$, де

$$a_{32} = a_{32} - M_3 a_{12}\,,\ a'_{33} = a_{33} - M_3 a_{13}\,,\ b'_3 = b_3 - M_3 b_1\,. \qquad \text{(КР.21)}$$

Перетворена таким чином система рівнянь (КР.19) набуває вигляду:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ 0 * x_1 + a'_{22}x_2 + a'_{23}x_3 = b'_2 \\ 0 * x_1 + a''_{32}x_2 + a''_{33}x_3 = b''_3 \end{cases} \qquad \text{(КР.22)}$$

Ця система рівнянь еквівалентна початковій і має певні переваги, оскільки x_1 входить тільки до першого рівняння. Спробуємо тепер виключити x_2 з останнього рівняння. Якщо $a_{22} = 0$, а $a_{32} \neq 0$, тоді переставимо друге й третє рівняння так, щоб $a_{22} \neq 0$. Інакше, система вироджена і має безліч розв'язків. Обчислюємо множник $M''_3 = \frac{a_{32}}{a_{22}}$. Друге рівняння системи (КР.23) множиться на M''_3 і віднімається від третього рівняння: $(a_{32} - M_3 a_{22})x_2 + (a_{33} - M_3 a_{23})x_3 = b_3 - b_2 M_2$. При цьому коефіцієнт біля x_2 дорівнює нулю: $a'_{32} - M_3 a'_{22} = 0$, $a''_{33} = a'_{33} - M_3 a'_{23}$, $b''_3 = b'_3 - M_3 b'_2$. Отримаємо

$$a''_{33}x_3 = b''_3. \qquad \text{(КР.23)}$$

Замінивши в системі (КР.22) третє рівняння на (КР.23), отримаємо систему рівнянь виду:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ 0 * x_1 + a'_{22}x_2 + a'_{23}x_3 = b'_2 \\ 0 * x_1 + 0 * x_2 + a''_{33}x_3 = b''_3 \end{cases} \qquad \text{(КР.24)}$$

Таку систему називають *системою з трикутною матрицею коефіцієнтів*, що еквівалентна СЛАР (КР.19). Процес знаходження такої системи називається *прямим ходом Гауса*. Знайти розв'язок такої системи просто: із 3-го рівняння знайти x_3, підставити результат у друге і знайти x_2, підставити x_2 і x_3 в 1-е рівняння системи (КР.24) і знайти x_1 за формулами: $x_3 = \frac{b''_3}{a''_{33}}$, $x_2 = \frac{b'_2 - a'_{23}x_3}{a'_{22}}$, $x_1 = \frac{b_1 - a_{12}x_2 - a_{13}x_3}{a_{11}}$.

Процес знаходження вектора розв'язку системи (КР.19) називають *зворотнім ходом методу Гауса* . На рис. КР.18 зображена схема алгоритму методу Гауса з послідовним виключенням для розв'язування системи із N рівнянь з N невідомими. Ця схема відповідає розглянутому алгоритму і може бути використана при розробці програми. Блок "Перестановка рівнянь так, щоб $a_{nn} \neq 0$ " означає деякий алгоритм, який дає змогу не допустити помилки "ділення на 0". Призначення індексів в схемі алгоритму (рис. КР.18): k – номер рівняння, яке віднімається від інших, а також номер невідомого, яке виключається із залишених k-рівнянь; i – номер рівняння, з якого виключається невідоме; j – номер стовпця.

2.5. Метод Крамера

Нехай дано систему лінійних рівнянь, в якій число рівнянь рівне числу невідомих:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n. \end{cases} \qquad \text{(КР.25)}$$

Подамо складові цієї лінійної системи в такий спосіб

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{bmatrix},$$

де $\mathbf{A}$ – матриця системи, $\mathbf{x}$ – стовпчик невідомих, $\mathbf{b}$ – стовпчик вільних членів.

Припустимо, що визначник системи не дорівнює нулю, тобто:

$$d = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \neq 0.$$

Якщо тепер замінити послідовно у визначнику d стовпчики коефіцієнтів при невідомих x_j ($j = 1, 2, \ldots, n$) стовпчиком вільних членів b_i, то отримаємо відповідно такі визначники:

$$d_1 = \begin{bmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{bmatrix}; \qquad d_2 = \begin{bmatrix} a_{11} & b_1 & \cdots & a_{1n} \\ a_{21} & b_2 & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & b_n & \cdots & a_{nn} \end{bmatrix};$$

$$d_3 = \begin{bmatrix} a_{11} & a_{12} & b_1 & \cdots & a_{1n} \\ a_{21} & a_{22} & b_2 & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & b_n & \cdots & a_{nn} \end{bmatrix}; \qquad d_n = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2,n-1} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & b_n \end{bmatrix}.$$

Теорема Крамера. Система n лінійних рівнянь з n невідомими, визначник якої відмінний від нуля, завжди сумісна і має єдиний розв'язок, який обчислюється за формулами:

$$x_1 = \frac{d_1}{d}; \; x_2 = \frac{d_2}{d}; \; x_{n-1} = \frac{d_{n-1}}{d}; \; x_n = \frac{d_n}{d}. \qquad \text{(КР.26)}$$

Формули (КР.26) називаються *формулами Крамера*.

Цей метод полягає в обчисленні визначника $\det(A)$ матриці A, а також визначників $\det(A_k)$ матриць A_k, $k=\overline{1,n}$. Матриці A_k отримуємо з матриці A шляхом заміни k-го стовпця на стовпець B. Можливі наступні варіанти:

1. Якщо $\det A \neq 0$, то система має єдиний розв'язок $\vec{X}=(x_1, x_2, ..., x_n)$, який очислюється за формулою:

$$x_k = \frac{\det(A_k)}{\det(A)}; \quad k=\overline{1,n}.$$

2. Якщо $\det(A)=0$ та всі $\det(A_k)=0$, $k=\overline{1,n}$, то система має безліч розв'язків.

3. Якщо $\det(A)=0$ і хоча б один $\det(A_k)\neq 0$, то система розв'язку не має.

Метод Крамера (визначників) не може бути застосований в більшості практичних задач через велику складність розрахунку визначників, навіть при невеликому зростанні порядку системи.

Розглянемо приклад розв'язування СЛАР методом Крамера

Приклад КР.3. Знайти розв'язок системи лінійних алгебраїчних рівнянь:

$$\begin{cases} 2x_1 + 5x_2 + 4x_3 = 30 \\ x_1 + 3x_2 + 2x_3 = 150 \\ 2x_1 + 10x_2 + 9x_3 = 110 \end{cases}$$

Визначники: $\Delta = \begin{vmatrix} 2 & 5 & 4 \\ 1 & 3 & 2 \\ 2 & 10 & 9 \end{vmatrix} = 5$; $\Delta_1 = \begin{vmatrix} 30 & 5 & 4 \\ 150 & 3 & 2 \\ 110 & 10 & 9 \end{vmatrix} = -760$; $\Delta_2 = \begin{vmatrix} 2 & 30 & 4 \\ 1 & 150 & 2 \\ 2 & 110 & 9 \end{vmatrix} = 1350$;

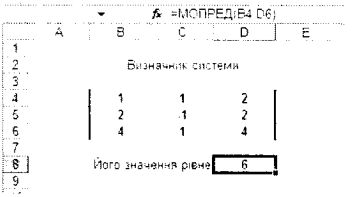
$\Delta_3 = \begin{vmatrix} 2 & 5 & 30 \\ 1 & 3 & 150 \\ 2 & 10 & 110 \end{vmatrix} = -1270$ та $x_1 = -\frac{760}{5} = -152$, $x_2 = \frac{1350}{5} = 270$, $x_3 = -\frac{1270}{5} = -254$.

Приклад КР.4. Розв'язати засобами **Ms Excel** систему лінійних алгебраїчних рівнянь методом Крамера. Система має такий вигляд:

$$\begin{cases} x_1 + x_2 + 2x_3 = -1, \\ 2x_1 - x_2 + 2x_3 = -4, \\ 4x_1 + x_2 + 4x_3 = -2. \end{cases}$$

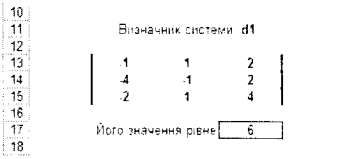
1. Відкрити чистий лист **Excel** і побудувати визначник d цієї системи. Використовуючи матричну операцію «МОПРЕД», яка є в категорії функцій «Математичні», виділяємо масив (визначника) B4:D6 і натискаємо ОК. В комірці D8 отримуємо значення $d=6$, що й зображено на скриншоті справа від визначника системи.

$$d = \begin{vmatrix} 1 & 1 & 2 \\ 2 & -1 & 2 \\ 4 & 1 & 4 \end{vmatrix}$$

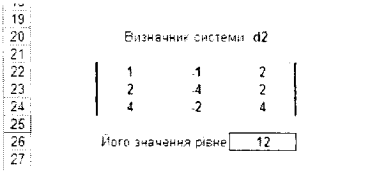


2. Аналогічно обчислюємо визначники d_1, d_2, d_3.

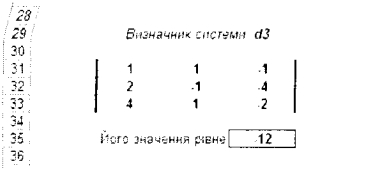
$$d_1 = \begin{vmatrix} -1 & 1 & 2 \\ -4 & -1 & 2 \\ -2 & 1 & 4 \end{vmatrix}$$



$$d_2 = \begin{vmatrix} 1 & -1 & 2 \\ 2 & -4 & 2 \\ 4 & -2 & 4 \end{vmatrix}$$



$$d_3 = \begin{vmatrix} 1 & 1 & -1 \\ 2 & -1 & -4 \\ 4 & 1 & -2 \end{vmatrix}$$



3. Остаточно, використовуючи формули Крамера (КР.26) знаходимо розв'язки цієї лінійної системи алгебраїчних рівнянь:

$$x_1 = \frac{6}{6} = 1, \qquad x_2 = \frac{12}{6} = 2, \qquad x_3 = \frac{-12}{6} = -2.$$

Таким чином, практично безпосереднє отримання значення визначника дає можливість знайти розв'язок системи n лінійних алгебраїчних рівнянь з n невідомими.

2.6. Метод оберненої матриці

Якщо задача розв'язку СЛАР вирішується у пакеті прикладних програм, в якому реалізована функція обчислення оберненої матриці, то для пошуку розв'язку можна застосовувати формулу: $X = A^{-1}B$, де A^{-1} – обернена матриця. Нагадаємо визначення оберненої матриці. *Оберненою* до квадратної матриці A називається така матриця A^{-1}.

для якої виконується співвідношення: $A \cdot A^{-1} = E = A^{-1} \cdot A$, де E – одинична матриця. Якщо матриця така, що можливо знайти обернену, то для знаходження оберненої матриці можна скористатися одним з наступних способів:

1. Метод Гауса. Візьмемо дві матриці: матрицю A та одиничну матрицю E такого самого розміру, що й A. Приведемо матрицю A до одиничної матриці методом Гауса, але використовуючи тільки операції над **рядками** (!!). Після застосування кожної операції до першої матриці, застосуємо ту саму операцію до другої. Коли зведення першої матриці A до одиничного вигляду буде завершено, друга матриця E зведеться до оберненої матриці A^{-1}. Тобто ми одержимо $(A \mid E) \Rightarrow (E \mid A^{-1})$.

2. За допомогою союзної матриці: $A^{-1} = \frac{1}{\det A} \cdot (C^*)^T$, де C^* – союзна матриця, а $(C^*)^T$ – матриця, отримана в результаті транспонування союзної матриці. *Союзна матриця* – матриця, яка складається з алгебраїчних доповнень елементів вхідної матриці.

$$C^* = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1j} \\ A_{21} & A_{22} & \dots & A_{2j} \\ \dots & \dots & \dots & \dots \\ A_{i1} & A_{i2} & \dots & A_{ij} \end{pmatrix}$$

де A_{11} – алгебраїчне доповнення. Алгебраїчним доповненням елемента a_{ij} матриці A називається число $A_{ij} = (-1)^{i+j} M_{ij}$, де M_{ij} – мінор, визначник матриці, що отримуємо із A викреслюванням i-го рядка та j-го стовпчика.

2.7. Матричний метод

1. Система n лінійних рівнянь з n невідомими.

Систему (КР.25) можна записати у вигляді матричного рівняння:

$$\mathbf{A} \times \mathbf{X} = \mathbf{B}. \qquad \text{(КР.27)}$$

В розгорнутому вигляді систему (КР.27) можна представити так:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{bmatrix}.$$

Розглянемо метод розв'язування системи (КР.25) в загальному вигляді, який часом називають ***методом оберненої матриці***. Вважаємо, що квадратна матриця $\mathbf{A}_{nn}$ є невиродженою, тобто її визначник $|\mathbf{A}| \neq 0$. В цьому випадку існує обернена матриця $\mathbf{A}^{-1}$. Помноживши зліва обидві частини матричної рівності (КР.27) на обернену матрицю $\mathbf{A}^{-1}$, отримаємо: $\mathbf{A}^{-1} \times \mathbf{A} \times \mathbf{X} = \mathbf{A}^{-1} \times \mathbf{B} \rightarrow \mathbf{E} \times \mathbf{X} = \mathbf{A}^{-1} \times \mathbf{B} \rightarrow \mathbf{E} \times \mathbf{X} = \mathbf{X}$, де $\mathbf{E}$ – одинична матриця. Звідси розв'язок системи методом оберненої матриці є таким

$$\mathbf{X} = \mathbf{A}^{-1} \times \mathbf{B}. \qquad \text{(КР.28)}$$

Таким чином, для розв'язування системи (КР.25), тобто знаходження вектора $\mathbf{X}$ необхідно знайти обернену матрицю коефіцієнтів $\mathbf{A}^{-1}$ і помножити її справа на вектор вільних членів $\mathbf{B}$.

Приклад КР.5. Розв'язати систему $\begin{cases} 3x + 2y = 7, \\ 4x - 5y = 40. \end{cases}$

Розв'язок.

1. Ввести матрицю $\mathbf{A}$ (в цьому випадку розміром 2×2) в діапазон A1:B2 $\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 4 & -5 \end{bmatrix}$. Вектор $\mathbf{B} = \begin{bmatrix} 7 & 40 \end{bmatrix}$ введіть в діапазон C1:C2

2. Знайдіть обернену матрицю $\mathbf{A}^{-1}$. Для цього:
   - виділіть блок комірок під обернену матрицю. Наприклад, виділіть блок A3:B4 (вказівником мишки при натиснутій лівій клавіші);
   - натисніть на панелі інструментів **Стандартна** кнопку **Вставка функції**;
   - в отриманому діалоговому вікні **Майстер функцій** в робочому полі **Категорія** оберіть пункт **Математичні**, а в робочому полі **Функція** — ім'я функції МОБР. Клацніть кнопкою ОК;
   - діалогове вікно МОБР, яке з'явилося мишкою відсуньте в сторону від початкової матриці і введіть діапазон вихідної матриці A1:B2 в робоче поле Масив (вказівником мишки при натиснутій лівій клавіші). Натисніть сполучення клавіш CTRL+SHIFT+ENTER;
   - якщо обернена матриця не з'явилась в діапазоні A3:B4, то треба клацнути вказівником мишки в рядку формул и повторити натискання сполучення клавіш CTRL+SHIFT+ENTER.

В результаті в діапазоні A3:B4 з'явиться обернена матриця

$$\begin{bmatrix} 0.217391 & 0.086957 \\ 0.173913 & -0.13043 \end{bmatrix}$$

3. Множенням оберненої матриці $\mathbf{A}^{-1}$ на вектор $\mathbf{B}$ знаходять вектор $\mathbf{X}$. Для цього:
   - виділяють блок комірок під результуючу матрицю (під вектор $\mathbf{X}$). Її розмірність буде $m \times p$, в цьому прикладі 2×1. Наприклад, виділіть блок комірок C3:C4 (вказівником мишки при натиснутій лівій клавіші);
   - натискають на панелі інструментів **Стандартна** кнопку **Вставка функції**;
   - в отриманому діалоговому вікні **Майстер функцій** в робочому полі **Категорія** виберіть пункт **Математичні**, а в робочому полі **Функція** — ім'я функції МуМнож. Клацнути кнопкою ОК;
   - Діалогове вікно МуМнож мишкою відсуньте в сторону від вихідних матриць і введіть діапазон оберненої матриці $\mathbf{A}^{-1}$, тобто діапазон A3:B4 в робоче поле Массив1 (вказівником мишки при натиснутій лівій клавіші), а діапазон матриці $\mathbf{B}$ тобто C1:C2 введіть в робоче поле Массив2. Після чого натиснути сполучення клавіш CTRL+SHIFT+ENTER.

- якщо вектор **X** не з'явиться в діапазоні C3:C4, треба клацнути вказівником мишки в рядку формул и повторити натискання CTRL+SHIFT+ENTER.

В результаті в діапазоні C3:C4 з'явиться вектор **X**. Причому $x = 5$ буде знаходитися в комірці C3, а $y = -4$ – в комірці C4.

Можна здійснити перевірку знайденого розв'язку. Для цього знайдений вектор **X** необхідно підставити у вихідне матричне рівняння $\mathbf{A} \times \mathbf{X} = \mathbf{B}$. Перевірка здійснюється в такий спосіб.

1. Виділити блок комірок під результуючу матрицю (під вектор **B**). Її розмірність буде $m \times p$, в цьому прикладі 2×1. Наприклад, виділіть блок комірок D1:D2 (вказівником мишки при натиснутій лівій клавіші).

2. Натисніть на панелі інструментів **Стандартна** кнопку **Вставка функції**.

3. В діалоговому вікні **Майстер функцій** на робочому полі **Категорія** виберіть **Математичні**, а на робочому полі **Функція** – ім'я функції МУМНОЖ. Клацніть кнопкою ОК.

4. Діалогове вікно МУМНОЖ мишкою відсуньте в бік від існуючих матриць і введіть діапазон вихідної матриці **A** – A1:B2 в робочому полі Масив1 (вказівником мишки при натиснутій лівій кнопці), а діапазон матриці **X** – C3:C4 – на робочому полі Масив2. Після цього натиснути сполучення клавіш CTRL+SHIFT+ENTER.

5. Якщо вектор **B** не з'явився в діапазоні D1:D2, то треба клацнути вказівником мишки в рядку формул і повторити натискання CTRL+SHIFT+ENTER.

В результате в діапазоні D1:D2 з'явиться вектор **B**, і, якщо система розв'язана правильно, то вектор, який з'явиться буде рівний вихідному $\mathbf{B} = \begin{bmatrix} 7 & 40 \end{bmatrix}$.

Приклад КР.6. Розв'язати систему $\begin{cases} 5x_1 + 3x_2 + 4x_3 = 2700, \\ 2x_1 + x_2 + x_3 = 900, \\ 3x_1 + 2x_2 + 2x_3 = 1600. \end{cases}$

Розв'язок.

1. Ввести матрицю **A** (в цьому випадку розміром 3×3) в діапазон A1:C3 $\mathbf{A} = \begin{bmatrix} 5 & 3 & 4 \\ 2 & 1 & 1 \\ 3 & 2 & 2 \end{bmatrix}$. Вектор $\mathbf{B} = \begin{bmatrix} 2700 & 900 & 1600 \end{bmatrix}$ ввести в діапазон D1:D3.

2. Знайдіть обернену матрицю $\mathbf{A}^{-1}$. Для цього:

- виділити блок комірок під обернену матрицю. Наприклад, виділити блок A4:C6 (вказівником мишки при натиснутій лівій клавіші);
- натисніть на панелі інструментів **Стандартна** кнопку **Вставка функції**;
- в отриманому діалоговому вікні **Майстер функцій** в робочому полі **Категорія** виберіть пункт **Математичні**, а в робочому полі **Функція** — ім'я функції МОБР. Клацніть кнопкою ОК;
- діалогове вікно МОБР, яке з'явилося мишкою відсуньте в бік від початкової матриці і введіть діапазон вихідної матриці A1:C3 в робоче

поле Масив (вказівником мишки при натиснутій лівій клавіші). Натисніть сполучення клавіш CTRL+SHIFT+ENTER;

- якщо обернена матриця не з'явилась в діапазоні A4:C6, то треба клацнути вкзівником мишки в рядку формул і повторити натискання сполучення клавіш CTRL+SHIFT+ENTER.

В результаті в діапазоні A4:C6 з'явиться обернена матриця.

3. Множенням оберненої матриці $\mathbf{A}^{-1}$ на вектор **B** знаходять вектор **X**. Для цього:

- виділяють блок комірок під результуючу матрицю (під вектор **X**). Її розмірність буде $m \times p$, в цьому прикладі 3×1. Наприклад, виділіть блок комірок D4:D6 (вказівником мишки при натиснутій лівій клавіші);
- натисніть на панелі інструментів **Стандартна** кнопку **Вставка функції**;
- в отриманому діалоговому вікні **Майстер функцій** в робочому полі **Категорія** виберіть пункт **Математичні**, а в робочому полі **Функція** — ім'я функції МуМнож. Клацнути кнопкою ОК;
- Діалогове вікно МуМнож мишкою відсуньте в сторону від вихідних матриць і введіть діапазон оберненої матриці $\mathbf{A}^{-1}$, тобто діапазон A4:C6 в робоче поле Масив1 (вказівником мишки при натиснутій лівій клавіші), а діапазон матриці **B** (тобто D1:D3) введіть в робоче поле Масив2. Після чого натиснути сполучення клавіш CTRL+SHIFT+ENTER;
- якщо вектор **X** не з'явиться в діапазоні D4:D6, треба клацнути вказівником мишки в рядку формул та повторити натискання CTRL+SHIFT+ENTER.

В результаті в діапазоні D4:D6 з'явиться вектор **X**. Причому $x_1 = 200$ буде в комірці D4, а $x_2 = 300$ – в комірці D5, а $x_3 = 200$ – в комірці D6.

Рекомендується зробити перевірку, підставивши знайдені значення в рівняння системи. Для цього знайдений вектор **X** треба підставити в початкове матричне рівняння $\mathbf{A} \times \mathbf{X} = \mathbf{B}$.

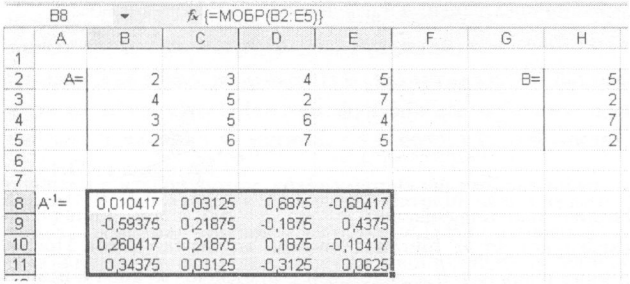
Приклад КР.7. Знайти розв'язки системи $\begin{cases} 2x_1 + 3x_2 + 4x_3 + 5x_4 = 5 \\ 4x_1 + 5x_2 + 2x_3 + 7x_4 = 2 \\ 3x_1 + 5x_2 + 6x_3 + 4x_4 = 7 \\ 2x_1 + 6x_2 + 6x_3 + 5x_4 = 2 \end{cases}$

Розв'язок.

1. Ввести значення елементів матриць **A** і **B** рівняння в комірки **Excel**.

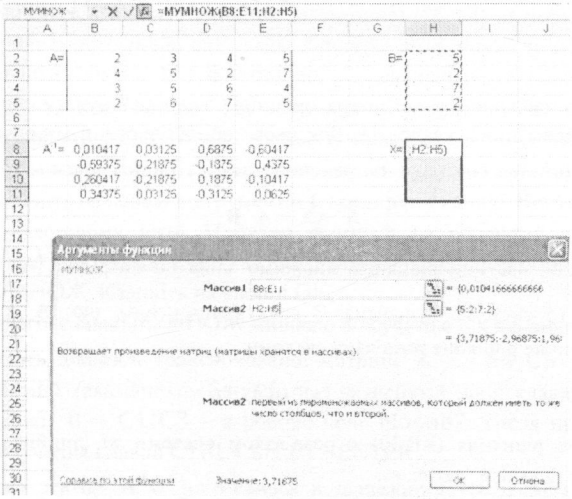
	A	B	C	D	E	F	G	H
1								
2	A=	2	3	4	5		B=	5
3		4	5	2	7			2
4		3	5	6	4			7
5		2	6	7	5			2

2. Обчислити обернену матрицю за допомогою матричної функції МОБР.

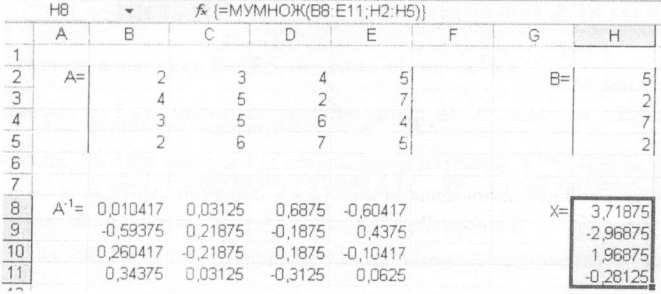


	A	B	C	D	E	F	G	H
1								
2	A=	2	3	4	5		B=	5
3		4	5	2	7			2
4		3	5	6	4			7
5		2	6	7	5			2
6								
7								
8	A^{-1}=	0,010417	0,03125	0,6875	-0,60417			
9		-0,59375	0,21875	-0,1875	0,4375			
10		0,260417	-0,21875	0,1875	-0,10417			
11		0,34375	0,03125	-0,3125	0,0625			

3. Помножити обернену матрицю $\mathbf{A}^{-1}$ на матрицю **B** за допомогою матричної функції МУМНОЖ (порядок множення є важливим: першою повинна бути матриця $\mathbf{A}^{-1}$, а другою – матриця **B**)



4. Перевірити правильність отриманої матриці розв'язків (коренів) **X**.



	A	B	C	D	E	F	G	H
1								
2	A=	2	3	4	5		B=	5
3		4	5	2	7			2
4		3	5	6	4			7
5		2	6	7	5			2
6								
7								
8	A^{-1}=	0,010417	0,03125	0,6875	-0,60417		X=	3,71875
9		-0,59375	0,21875	-0,1875	0,4375			-2,96875
10		0,260417	-0,21875	0,1875	-0,10417			1,96875
11		0,34375	0,03125	-0,3125	0,0625			-0,28125

II. Система m лінійних рівнянь з n невідомими.

Система m лінійних рівнянь з n невідомими має такий вигляд:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \quad \cdots \quad + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \quad \cdots \quad + a_{2n}x_n = b_2, \\ \quad \cdots \quad \quad \cdots \quad \quad \cdots \quad \quad \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \quad \cdots \quad + a_{mn}x_n = b_n. \end{cases}$$

Як і попередні системи цю систему також можна подати в матричному вигляді $\mathbf{A} \times \mathbf{X} = \mathbf{B}$. Можливі такі три випадки: $m < n$, $m = n$ і $m > n$. Випадок, коли $m = n$, розглянутий вище (система n лінійних рівнянь з n невідомими). При $m < n$, якщо система m лінійних рівнянь з n невідомими є сумісною, то вона є невизначена і має нескінченну множину розв'язків. У випадку, якщо $m > n$ і система є сумісною, то матриця $\mathbf{A}$ має принаймні $m - n$ лінійно залежних рядків. Тоді розв'язок може бути отриманий відбором n будь-яких лінійно незалежних рівнянь (якщо вони існують) і застосуванням формули (КР.27) (система n лінійних рівнянь з n невідомими), тобто може бути зведена до раніше розв'язуваних задач. При цьому отриманий розв'язок буде задовольняти і решту $m - n$ рівнянь.

Проте, за наявності комп'ютера зручніше використовувати більш загальний підхід – метод найменших квадратів. Для цього обидві частини матричного рівняння системи (КР.27) помножимо зліва на транспоновану матрицю системи $\mathbf{A}^{\mathrm{T}}$

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{X} = \mathbf{A}^{\mathrm{T}}\mathbf{B}. \qquad \text{(КР.29)}$$

Далі обидві частини рівняння множимо зліва на матрицю $\left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}$. Якщо ця матриця існує, то система визначена. З врахуванням того, що $\left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1} \cdot \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right) = \mathbf{E}$, отримуємо матричне рівняння розв'язку системи

$$\mathbf{X} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{B}. \qquad \text{(КР.30)}$$

Матричне рівняння (КР.30) є розв'язком системи m лінійних рівнянь з n невідомими при за умови, що $m > n$.

Приклад КР.8. Розв'язати систему $\begin{cases} 6x + 2y = 7, \\ 4x - 5y = 40, \\ 3x + 3y = 3. \end{cases}$

Розв'язок.

1. Ввести матрицю $\mathbf{A}$ (в цьому випадку розміром 3×2) в діапазон A1:B3 $\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 4 & -5 \\ 3 & 3 \end{bmatrix}$, а вектор $\mathbf{B} = \begin{pmatrix} 7 & 40 & 3 \end{pmatrix}$ в ввести в діапазон C1:C3.

2. Знайти транспоновану матрицю $\mathbf{A}^{\mathrm{T}}$. Для цього:

- виділити блок комірок під транспоновану матрицю. Його розмір в цьому прикладі буде 2×3. Наприклад, якщо виділити блок A4:C5 (вказівником мишки при натиснутій лівій клавіші мишки);
- натиснути на панелі інструментів **Стандартна** кнопку **Вставка функції;**
- в діалоговому вікні **Майстер функцій** на робочому полі **Категорія** вибрати **Посилання та масиви**, а в робочому полі **Функція** ім'я функції ТРАНСП і клацніть ОК;
- перенесіть мишкою в бік від вихідної матриці діалогове вікно ТРАНСП та введіть діапазон вихідної матриці A1:B3 в робоче поле **Масив** (вказівником мишки при натиснутій лівій клавіші). Після цього натисніть сполучення клавиш CTRL+SHIFT+ENTER;
- якщо обернена матриця не з'явилась в діапазоні A3:B4, то треба клацнути вказівником мишки в **Рядку формул** і повторити натискання клавіш CTRL+SHIFT+ENTER.

В результаті в діапазоні A4:C5 з'явиться транспонована матриця $\mathbf{A}^{\mathrm{T}}$.

3. Щоб знайти добуток $\mathbf{A}^{\mathrm{T}} \times \mathbf{B}$ потрібно:

- виділити блок комірок під результуючу матрицю (під вектор $\mathbf{A}^{\mathrm{T}}\mathbf{B}$). Її розмірність має бути $n \times 1$ (в цьому прикладі 2×1). Наприклад, виділіть блок комірок E4:E5 (вказівником мишки при натиснутій лівій клавіші);
- натисніть на панелі інструментів **Стандартна** кнопку **Вставка функції**;
- в діалоговому вікні **Майстер функцій** на робочому полі **Категорія** виберіть **Математичні**, а на робочому полі **Функція** ім`я функції МУМНОЖ. Клацніть на кнопці ОК;
- діалогове вікно МУМНОЖ мишкою відсуньте в бік від вихідних матриць і введіть діапазон транспонованої матриці $\mathbf{A}^{\mathrm{T}}$ – A4:C5 в робоче поле Масив1 (вказівником мишки при натиснутій лівій клавіші), а діапазон матриці $\mathbf{B}$ – C1:C3 – в робоче поле Масив2. Після цього натиснути сполучення клавіш CTRL+SHIFT+ENTER;
- якщо вектор $\mathbf{A}^{\mathrm{T}}\mathbf{B}$ не з'явився в діапазоні E4:E5, то треба клацнути вказівником мишки в рядку формул і повторити натискання CTRL+SHIFT+ENTER.

В результаті, в діапазоні E4:E5 з'явиться вектор $\mathbf{A}^{\mathrm{T}}\mathbf{B} = \begin{pmatrix} 190 \\ -177 \end{pmatrix}$.

Аналогічно знаходимо добуток $\mathbf{A}^{\mathrm{T}} \times \mathbf{A}$:

- виділіть блок комірок під результуючу матрицю $\mathbf{A}^{\mathrm{T}}\mathbf{A}$. Її розмірність буде $n \times n$ (в цьому прикладі 2×2). Наприклад, виділіть блок комірок A7:B8 (вказівником мишки при натиснутій лівій клавіші);
- далі виконати так, як описано в пункті 3, вказуючи відповідні діапазони.

У результаті в діапазоні A7:B8 з'явиться матриця $\mathbf{A}^T\mathbf{A}$: $\begin{pmatrix} 34 & -5 \\ -5 & 38 \end{pmatrix}$.

4. Знаходимо обернену матрицю $(\mathbf{A}^T\mathbf{A})^{-1}$. Для цього необхідно:

- виділити блок комірок під обернену матрицю. Наприклад, блок A10:B11 (вказівником миші при натиснутій лівій клавіші);
- натиснути на панелі інструментів **Стандартна** кнопку **Вставка функції**;
- в діалоговому вікні **Майстер функцій** в робочому полі **Категорія** вибрати пункт **Математичні**, а в робочому полі **Функція** – ім'я функції МОБР. Потім клацнути на кнопці OK;
- мишею відсунути убік від вихідної матриці діалогове вікно МОБР, що з'явилося і ввести діапазон вихідної матриці $\mathbf{A}^T\mathbf{A}$ – A7:B8 в робоче поле Масив (вказівником миші при натиснутій лівій клавіші). Після цього натиснути клавіші CTRL+SHIFT+ENTER;
- якщо обернена матриця не з'явилася в діапазоні A10:B11, треба клацнути вказівником миші в рядку формул і повторити натискання CTRL+SHIFT+ENTER.

В результаті, в діапазоні A10:B11 з'явиться обернена матриця $(\mathbf{A}^T\mathbf{A})^{-1}$

$$\begin{bmatrix} 0{,}029992 & 0{,}003946 \\ 0{,}\ 003946 & 0{,}\ 026835 \end{bmatrix}$$

5. Тепер множенням оберненої матриці $(\mathbf{A}^T\mathbf{A})^{-1}$ на вектор $\mathbf{A}^T\mathbf{B}$ знаходимо вектор $\mathbf{X}$. Для цього треба:

- виділити блок комірок під результуючу матрицю (під вектор X) . Її розмірність буде $n \times 1$, в даному прикладі 2×1. Наприклад, можна виділити блок комірок D1:D2 (покажчиком миші при натиснутій лівій клавіші);
- натиснути на панелі інструментів **Стандартна** кнопку **Вставка функції**;
- в діалоговому вікні **Майстер функцій** в робочому полі **Категорія** виберіть пункт **Математичні**, а в робочому полі **Функція** – ім'я функції МУМНОЖ. Клацніть на кнопці OK;
- діалогове вікно МУМНОЖ, яке з'явилося, відсуньте мишею убік від вихідних матриць і введіть діапазон оберненої матриці $(\mathbf{A}^T\mathbf{A})^{-1}$ – A10:B11 в робоче поле Масив1 (вказівником миші при натисненні лівої клавіші), а діапазон матриці $\mathbf{A}^T\mathbf{B}$ – E4:E5 – в робоче поле Масив2. Після цього натисніть поєднання клавіш CTRL+SHIFT+ENTER;

- якщо вектор **X** не з'явиться в діапазоні D1:D2, то треба клацнути вказівником миші в рядку формул і повторити натискання CTRL+SHIFT+ENTER.

В результаті в діапазоні D1:D2 з'явиться вектор **X** ($x = 5$ буде в комірці D1, а $y = -4$ – в комірці D2).

Можна здійснити перевірку знайденого розв'язку. Для цього розв'язок **X** треба підставити в початкове матричне рівняння $\mathbf{A} \times \mathbf{X} = \mathbf{B}$, причому досить підставити **X** в будь-які n рівнянь.

3. Інтерполювання функцій

Інженери в практичній діяльності постійно зіштовхуються з необхідністю визначення видів зв'язку в процесах та явищах і необхідністю їх математичного опису. Якщо деяка величина y, що характеризує процес, залежить від сукупності не зв'язаних між собою величин $(x_1,x_2,\ldots,x_n)$ таким чином, що кожному набору $(x_1,x_2,\ldots,x_n)$ відповідає значення величини y, така однозначна відповідність величини y сукупності незалежних змінних $x_1,x_2,\ldots,x_n$ називається *функціональною залежністю*, а сама змінна величина y – функцією змінних величин $x_1,x_2,\ldots,x_n$, що формально записується у вигляді $y=f(x_1,x_2,\ldots,x_n)$. Так, вираз $y = x_1^2 + 3\sqrt{x_2} + x_1 x_3^2$ є функцією трьох змінних. Наприклад, площа кола S є функцією незалежної змінної – радіуса кола R, тобто $S=f(R)$ (конкретний вигляд цієї функції $S = \pi R^2$). Об'єм фігури є вже функцією трьох вимірів: $V = f(x_1, x_2, x_3)$, і в залежності від вигляду фігури цей функціональний зв'язок відповідно конкретизується.

З курсу математичного аналізу відомі три способи представлення функціональних залежностей: аналітичний; графічний; табличний. Так, наприклад, в результаті математичного опрацювання можна отримати наступну аналітичну залежність: грошових кредитів в сільському господарстві під товарно-матеріальних цінностей і сезонних витрат від витрат на велику рогату худобу: $y = 51.0203 + 0.1059x$, де y – кредити під товарно-матеріальні цінності; x – витрати на велику рогату худобу. Інший приклад аналітичної залежності: зв'язок шляху з часом в рівноприскореному русі виражається як $s = vt + \dfrac{at^2}{2}$.

Позитивною властивістю *аналітичного способу* представлення є можливість одержувати значення y для будь-якого фіксованого аргументу x з будь-якою точністю. До недоліків цього способу слід віднести те, що потрібно повторювати всю послідовність обчислень; крім того, аналітичний спосіб не володіє наочністю. Вказані недоліки аналітичного способу усуваються у випадку *графічного* представлення функції $y = f(x)$.

Табличний спосіб представлення функцій розповсюджений у техніці, фізиці, економіці, природознавстві та найчастіше всього використовується для запису результатів експерименту. Нехай, наприклад, в результаті досліду отримана залежність омічного опору R мідного стержня від температури t^0 у вигляді табл. КР.1.

Таблиця КР.1

Результати експериментальних досліджень

R	77.80	79.75	80.80	82.35	83.90	85.10
t^0	25.0	30.1	36.0	40.0	45.1	50.0

В цьому експерименті значення омічного опору мідного стержня змінюється при коливанні температури і є залежною змінною.

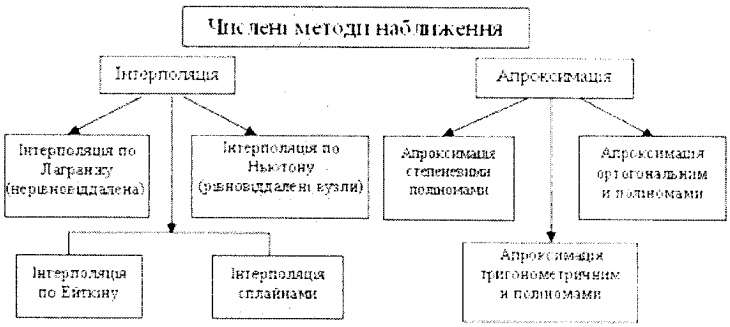


Рис. КР.19. Класифікація чисельних методів наближення табличних функцій

Перевагою табличного способу представлення експериментальної функції є те, що для кожного значення незалежної змінної з таблиці, можна відразу ж, без усяких вимірів і обчислень, знайти відповідне значення функції. *Недолік табличного способу* полягає в тому, що не можна задати всю функцію скрізь, тобто завжди знайдуться такі значення незалежної змінної, яких немає в таблиці. Тому, для аналізу результатів інженерних експериментів дуже зручно використовувати як табличний так і аналітичний способи представлення залежностей, що досліджуються. Так, якщо в результаті інженерного або наукового експерименту отримана система точок: $\{(x_0, y_0)(x_1, y_1)(x_2, y_2),...,(x_n, y_n)\}$, то дуже часто виникає задача пошуку аналітичної залежності, яка б пов'язувала експериментальні дані у вигляді аналітичної функції $Q(x)$. Для розв'язування цієї задачі за допомогою чисельних методів на комп'ютері використовуються *два підходи*:

1. ***Інтерполяція*** – підхід, за допомогою якого отримують аналітичні залежності табличних функцій за умови, що аналітична функція $Q(x)$ повинна проходити через всі задані експериментальні точки.

2. ***Апроксимація*** – підхід, за допомогою якого знаходимо аналітичну функцію $Q(x)$, що "найкращим чином" наближається до заданої табличної функції. Звичайно "найкращим чином" – це критерій, замість якого використовується критерій середньо квадратичного відхилення (СКВ), що базується на тому, що сума квадратів відхилень

аналітичної функції $Q(x_i)$ від експериментальної y_i (при $i=0,1,\ldots,k$) повинна бути мінімальною: $\sum_{i=0}^{n}(Q(x_i)-y_i)^2 \Rightarrow \min$. На рис. КР.19 дано *класифікаіцю відомих методів наближення* табличних функцій, які призначені для пошуку ***аналітичної залежності*** $Q(x)$, яка б пов'язувала експериментальні дані $\{(x_0,y_0)(x_1,y_1)(x_2,y_2),\ldots,(x_n,y_n)\}$, отримані в результаті інженерного або наукового експерименту.

3.1. Математична постановка задачі інтерполювання

В економіці та техніці постійно доводиться зіштовхуватися з необхідністю обчислення значень функції $y=f(x)$ в точках x_i, відмінних від значень аргументу, фіксованих в таблиці експериментальних досліджень. Крім того, в деяких випадках, незважаючи на те, що аналітичний вираз функції $y=f(x)$ відомий, він є занадто складним і незручним для подальших математичних перетворень. Подібні задачі формалізуються як задачі інтерполювання.

Нехай на відрізку $[a,b]$ функція $y=f(x)$ задана системою точок $\{(x_0,y_0)(x_1,y_1)(x_2,y_2),\ldots,(x_n,y_n)\}$, де значення $x_1,x_2,\ldots,x_n$ називаються *вузлами інтерполяції*. Необхідно знайти аналітичну залежність $Q(x)$, яка співпадає у вузлах інтерполяції зі значеннями заданої функції, тобто $y_0=Q(x_0)=f(x_0)$, $y_1=Q(x_1)=f(x_1)$, ..., $y_n=Q(x_n)=f(x_n)$. Процес обчислення значень функції $Q(x)$ в точках x_i, відмінних від вузлів інтерполяції, називають *інтерполюванням функції* $f(x)$ (рис. КР.20).

Якщо аргумент x знаходиться за межами відрізка інтерполювання $[x_0,x_n]$, то задача визначення значення функції $Q(x)$ в точці x називається *екстраполюванням*.

Слід відзначити, що задача інтерполювання стає однозначною, якщо для функції $Q(x)$ вибрати многочлен $Q_n(x)$ степені не вище n, такий, що $Q_n(x_0)=y_0, Q_n(x_1)=y_1,\ldots,Q_n(x_n)=y_n$. Многочлен $Q_n(x)$, що задовольняє цим умовам, називають *інтерполяційним многочленом*, а відповідні формули – *інтерполяційними формулами*.

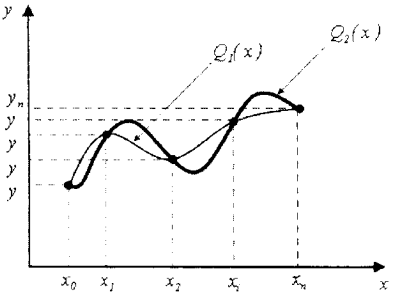


Рис.КР.20. Геометрична інтерпретація інтерполяції табличної функції

У випадку, коли $Q(x)$ береться з класу степеневих функцій, інтерполяція називається *параболічною*. Цей спосіб наближення грунтується на тому, що на невеликих відрізках експериментальна функція $f(x)$ може бути достатньо добре апроксимована параболою певного порядку. Якщо як інтерполяційну функцію використовувати многочлен виду:

$$F(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n, \qquad \text{(КР.31)}$$

то така інтерполяція називається *степеневою*.

Інколи доцільно використати інші види інтерполяції. Якщо функція, що досліджується, $f(x)$ – періодична, то в як інтерполяційну функцію $Q(x)$ обирають тригонометричну, наприклад, виду:

$$Q(x) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) \qquad \text{(КР.32)}$$

і така інтерполяція називається *тригонометричною*. Деколи як інтерполяційну функцію $Q(x)$ обирають раціональні функції.

При інтерполюванні виникає ряд задач:

1) вибір найбільш зручного способу побудови інтерполяційної функції для кожного конкретного випадку;

2) оцінка похибки при заміні $f(x)$ інтерполяційною функцією $Q(x)$ на відрізку $[a,b]$, оскільки функції $f(x)$ та $Q(x)$ співпадають тільки у вузлах інтерполяції $x_0, x_1, \ldots, x_n$.

3) оптимальний вибір вузлів інтерполяції для отримання мінімальної похибки.

Для задачі інтерполювання важливим є визначення того, як повинна вести себе інтерполяційна функція між заданими точками, так як ці точки можуть бути інтерпольовані множиною різних функцій, і необхідно мати певний критерій вибору. Звичайно критерій формується в термінах гладкості та простоти. Більшість інтерполяційних функції генеруються лінійними комбінаціями найпростіших функцій. Лінійні комбінації одночленів $\{x^k\}$ формують ***степеневі поліноми***, лінійні комбінації тригонометричних функцій $\{\cos x, \sin x\}$ формують ***тригонометричні поліноми***, використовуються також лінійні комбінації експонент $\{\exp(\beta_k \alpha)\}$. Найбільш важливим класом інтерполяційних функцій є множина алгебраїчних поліномів. Поліноми мають переваги з точки зору алгоритмізації, тому що їх значення легко обчислювати, додавати, множити, інтегрувати чи диференціювати. Важливою властивістю поліномів є те що якщо c – константа, а $p(x)$ – поліном, то поліномами будуть і $p(cx)$ і $p(x+c)$. Клас інтерполяційних функції обирають, використовуючи ***теорему Вейерштраса:*** якщо $f(x)$ – неперервна на кінцевому інтервалі $[a,b]$ функція, то для любого $\varepsilon > 0$ існує поліном $p_n(x)$ ступеня n такий, що $\max|f(x) - p_n(x)| < \varepsilon$.

3.2. Інтерполяційний многочлен Лагранжа

Найбільш загальною формулою параболічного інтерполювання є ***інтерполяційна формула Лагранжа***. Задача параболічного інтерполювання в цьому випадку формулюється наступним чином: на відрізку $[a,b]$ у вузлах інтерполяції $x_0,x_1,\ldots,x_n$ задається функція $f(x)$ своїми $(n+1)$ значеннями $y_0 = f(x_0), y_1 = f(x_1),\ldots, y_n = f(x_n)$, необхідно побудувати многочлен $L(x)$ так, щоб у вузлах інтерполяції $x_0,x_1,\ldots,x_n$ його значення співпадали зі значеннями заданої функції, тобто $L(x_0) = y_0$, $L(x_1) = y_1$, …, $L(x_n) = y_n$. Слід відзначити, що в такій постановці задачі вузли інтерполяції $x_0,x_1,\ldots,x_n$ можуть бути ***довільно розташовані*** один від одного на відрізку $[a,b]$, іншими словами, вузли інтерполяції не рівновіддалені, тобто $h = x_{i+1} - x_i \neq const \quad (i = \overline{0,n-1})$. Величина h називається *кроком інтерполяції*.

Задача інтерполювання має розв'язок, якщо степінь m многочлена $L(x) = a_0 + a_1x + a_2x^2 + \ldots + a_mx^m$, яким замінюється функція $f(x)$, не вище порядку n $(m \le n)$. Тоді задача інтерполювання зводиться до пошуку невідомих постійних коефіцієнтів многочлена $a_i \quad (i = \overline{0,m})$ з системи рівнянь, яка будується наступним чином. З початкових умов відомо, що функція $L_n(x)$ в вузлах $x_0,x_1,\ldots,x_n$ приймає значення $L_n(x_0) = y_0$, $L_n(x_1) = y_1$, …, $L_n(x_n) = y_n$. Тоді в вузлі x_0 інтерполяційний многочлен $L_n(x)$ має вигляд $L_n(x_0) = a_0 + a_1x_0 + a_2x_0^2 + \ldots + a_mx_0^m$ в вузлі інтерполяції x_1 - $L_n(x_1) = a_0 + a_1x_1 + a_2x_1^2 + \ldots + a_mx_1^m$ і так далі. Нарешті, в вузлі x_n інтерполяційний многочлен $L_n(x)$ буде мати вигляд: $L_n(x_1) = a_0 + a_1x_n + a_2x_n^2 + \ldots + a_mx_n^m$.

Запишемо це у вигляді системи $(n+1)$ рівнянь з $m+1$ невідомими $a_0,a_1,\ldots,a_m$:

$$\begin{cases} a_0 + a_1x_0 + a_2x_0^2 + \ldots + a_mx_0^m = y_0 \\ a_0 + a_1x_1 + a_2x_1^2 + \ldots + a_mx_1^m = y_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0 + a_1x_n + a_2x_n^2 + \ldots + a_mx_n^m = y_n \end{cases} \qquad \text{(КР.33)}$$

де x_i і y_i $(i = \overline{0.n})$ – табличні значення аргументу і функції, що досліджується. Невідомі коефіцієнти $a_0,a_1,\ldots,a_n$ знаходяться за *формулами Крамера*:

$$a_0 = \frac{\Delta_0}{\Delta}, a_1 = \frac{\Delta_1}{\Delta}, \ldots, a_n = \frac{\Delta_n}{\Delta}, \qquad \text{(КР.34)}$$

де Δ – визначник системи (КР.33).

Якщо $\Delta \neq 0$, (тобто, коли $x_0,x_1,\ldots,x_n$ різні), то система (КР.33) має єдиний розв'язок. Якщо знайти коефіцієнти $a_0,a_1,\ldots,a_n$, можна представити інтерполяційний многочлен у вигляді $L_n(x) = \frac{\Delta_0}{\Delta} + \frac{\Delta_1}{\Delta}x + \ldots + \frac{\Delta_n^2}{\Delta}x^n$. Перепишемо многочлен в іншій

формі: $L_n(x) = y_0 Q_0(x) + y_1 Q_1(x) + \ldots + y_n Q_n(x)$. Легко перевірити, що функція $Q_i(x)$ повинна задовольняти умовам $Q_i(x_j) = \begin{cases} 0, i \neq j, \\ 1, i = j. \end{cases}$

$$Q_i(x) = \frac{(x - x_0)(x - x_1)\ldots(x - x_{i-1})(x - x_{i+1})\ldots(x - x_n)}{(x_j - x_0)(x_j - x_1)\ldots(x_j - x_{i-1})(x_j - x_{i+1})\ldots(x_j - x_n)} \qquad \text{(КР.35)}$$

В точках $x_0, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ функція $Q_i(x)$ стає 0, а в точці x_i дорівнює 1. Остаточно отримаємо вираз

$$L_n(x) = y_0 \frac{(x - x_1)\ldots(x - x_n)}{(x_0 - x_1)\ldots(x_0 - x_n)} + y_1 \frac{(x - x_0)(x - x_2)\ldots(x - x_n)}{(x_1 - x_0)(x_1 - x_2)\ldots(x_0 - x_n)} + \ldots + \\ + y_n \frac{(x - x_0)(x - x_1)\ldots(x - x_{n-1})}{(x_n - x_0)(x_n - x_1)\ldots(x_n - x_{n-1})}. \qquad \text{(КР.36)}$$

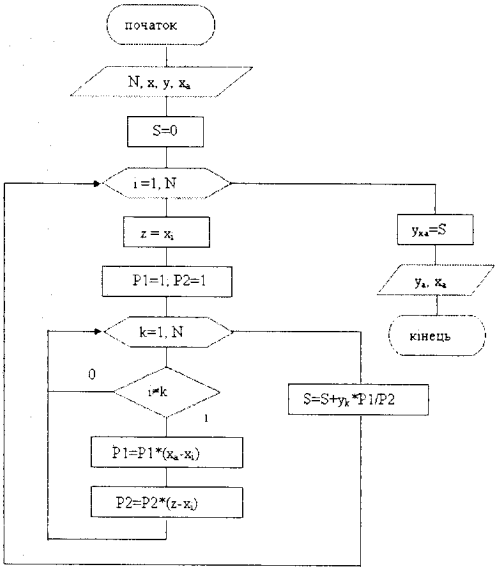


Рис. КР.21. Схема алгоритму метода Лагранжа

Цей многочлен називається *інтерполяційним многочленом Лагранжа*. В спрощеному вигляді його можна записати так:

$$L_n(x) = \sum_{i=0}^{n} y_i \frac{(x - x_0)(x - x_1)\ldots(x - x_{i-1})(x - x_{i+1})\ldots(x - x_n)}{(x_i - x_0)(x_i - x_1)\ldots(x_i - x_{i-1})(x_i - x_{i+1})\ldots(x_i - x_n)}. \qquad \text{(КР.37)}$$

Цей метод легко алгоритмізується і може використовуватися для розробки програм інтерполяції. Схема алгоритму методу подана на рис. КР.21.

Приклад КР.9. Побудувати інтерполяційний многочлен Лагранжа для функції заданої таблично, $n=4$; $m=n-1=4-1=3$. Припустимо, що $y=a_0+a_1x+a_2x^2+a_3x^3$.

n	0	1	2	3
x	1	2	3	5
y	1	5	4	81

Слід пам'ятати, що при екстраполяції функції, чим далі значення x від інтервалу спостереження, тим отримане значення функції містить більшу похибку.

$$L_m(x)=y_0\frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}+y_1\frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}+$$

$$+y_2\frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}+y_3\frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)}=$$

$$=1\cdot\frac{(x-2)(x-3)(x-5)}{(1-2)(1-3)(1-5)}+5\cdot\frac{(x-1)(x-3)(x-5)}{(2-1)(2-3)(2-5)}+$$

$$+4\cdot\frac{(x-1)(x-2)(x-5)}{(3-1)(3-2)(3-5)}+81\cdot\frac{(x-1)(x-2)(x-3)}{(5-1)(5-2)(5-3)}=$$

$$=-\frac{1}{8}((x^2-3x-2x+6)(x-5))+\frac{5}{3}((x^2-3x-x+3)(x-5))-$$

$$-((x^2-2x-x+2)(x-5))+\frac{81}{24}((x^2-2x-x+2)(x-3))=$$

$$=-\frac{1}{8}(x^3-10x^2+21x-30)+\frac{5}{3}(x^3-9x^2+23x-15)-$$

$$-(x^3-8x^2+17x-10)+\frac{27}{8}(x^3-6x^2+11x-6)=$$

$$=-\frac{1}{8}x^3+\frac{10}{8}x^2-\frac{21}{8}x+\frac{30}{8}+\frac{5}{3}x^3-15x^2+\frac{115}{3}x-25-$$

$$-x^3+8x^2-17x+10+\frac{27}{8}x^3-\frac{162}{8}x^2+\frac{297}{8}x-\frac{162}{8}=$$

$$=x^3-2x^2+3x-1$$

Таким чином: $a_0=-1, a_1=3, a_2=-2, a_3=1$. Висновки:

1. Таким чином за допомогою многочлена Лагранжа були отримані коефіцієнти інтерполяційної функції $Q(x)=a_0+a_1x+a_2x^2+a_3x^3$: $a_0=-1, a_1=3, a_2=-2, a_3=1$.

2. Використовуючи отриманий многочлен можна знайти будь-яке значення функції y для заданого x. Наприклад, для $x=4 \Rightarrow y=43$.

3. Використовуючи інтерполяційний многочлен можливо отримати значення функції y за межами спостережень. У цьому прикладі інтервал спостереження [1,5]; $x=6 \Rightarrow y=161$. Така задача називається *екстраполяція* (прогнозування функції). Для ***оцінки похибки*** інтерполяційного многочлена Лагранжа використовують формулу:

$$R_n(x)=|f(x)-L_n(x)|\le\frac{M_{n+1}}{(n+1)!}|(x-x_0)(x-x_1)...(x-x_n)| \quad \text{при} \quad M_{n+1}=\max|f^{(n+1)}(x)|,$$

$x\in[a,b]$.

3.3. Перша інтерполяційна формула Ньютона для рівновіддалених вузлів інтерполяції

Якщо функція, що досліджується, задана значеннями $y_0=f(x_0)$, $y_1=f(x_1)$, …, $y_n=f(x_n)$ в ***рівновіддалених вузлах*** інтерполяції, тобто $x_1=x_0+h$, $x_2=x_0+h$, …, $x_n=x_0+nh$, то для побудови її аналітичної залежності зручно використовувати ***першу інтерполяційну формулу Ньютона***. Для виведення інтерполяційних формул для рівновіддалених вузлів інтерполяції вводиться поняття ***кінцевої різниці.*** Поставимо наступну задачу: для функції $y=f(x)$, яка задана таблицею значень $y_0=f(x_0)$, $y_1=f(x_1)$, …, $y_n=f(x_n)$, причому x змінюється з однаковим кроком h, тобто $x_i=x_{i-1}+h$, побудувати кінцеві різниці. *Кінцевою різницею першого порядку* Δy_i називається різниця між значеннями функції в сусідніх вузлах інтерполяції: $\Delta y_0 = y_1 - y_0$, $\Delta y_1 = y_2 - y_1$, $\Delta y_2 = y_3 - y_2$, …, $\Delta y_{n-1} = y_n - y_{n-1}$. В загальному вигляді кінцеву різницю першого порядку Δy_i можна записати як $\Delta y_i = y_{i+1} - y_i$. *Кінцева різниця другого порядку* $\Delta^2 y_i$ складається з кінцевих різниць першого порядку: $\Delta^2 y_0 = \Delta y_1 - \Delta y_0$, $\Delta^2 y_1 = \Delta y_2 \quad \Delta y_1$, $\Delta^2 y_2 = \Delta y_3 - \Delta y_2$, …, $\Delta^2 y_{n-1} = \Delta y_n - \Delta y_{n-1}$. *Кінцева різниця n-го порядку* $\Delta^n y_i$ складається з кінцевих різниць (n-1)-го порядку: $\Delta^n y_0 = \Delta^{n-1} y_1 - \Delta^{n-1} y_0$, $\Delta^n y_1 = \Delta^{n-1} y_2 - \Delta^{n-1} y_1$, $\Delta^n y_2 = \Delta^{n-1} y_3 - \Delta^{n-1} y_2$, …, $\Delta^n y_{n-1} = \Delta^{n-1} y_n - \Delta^{n-1} y_{n-1}$, або в технічній літературі використовують наступну формулу кінцевої різниці n-го порядку: $\Delta^n y = \Delta(\Delta^{n-1} y)$.

Нехай необхідно побудувати інтерполяційний многочлен $P_m(x)$ степеня m такий, що $P_m(x_0)=y_1$, $P_m(x_0)=y_1$, …, $P_m(x_n)=y_n$. Будемо шукати многочлен вигляду

$$P_m(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \\ + a_3(x - x_0)(x - x_1)(x - x_2) + \ldots + a_m(x - x_0)(x - x_1)\ldots(x - x_{n-1}). \qquad \text{(КР.38)}$$

В цьому виразі невідомі коефіцієнти $a_0, a_1, \ldots, a_m$. Для того щоб знайти a_0, покладемо $x=x_0$. Тоді при підстановці $x=x_0$ у вираз (КР.38) всі складові, окрім першої, перетворяться на нуль, тобто $P_m(x_0)=a_0$, а значення функції в точці x_0 відомі з умови задачі: $P_m(x_0)=y_0$. Отже $a_0=y_0$. Щоб знайти коефіцієнт a_1 складемо першу кінцеву різницю для многочлена $P_m(x)$ в точці x: $\Delta P_m(x) = P_m(x+h) - P_m(x)$.

Зробивши всі підстановки, отримаємо:

$$\Delta P_m(x) = a_0 + a_1(x - x_0 + h) + a_2(x - x_0 + h)(x - x_1 + h) + \\ + a_3(x - x_0 + h)(x - x_1 + h)(x - x_2 + h) + \ldots + \\ + a^m(x - x_0 + h)(x - x_1 + h)\ldots(x - x_{n-1} + h) - a_0 - \\ - a_1(x - x_0) - a_2(x - x_0)(x - x_1) - a_3(x - x_0)(x - x_1)(x - x_2) - \\ - \ldots - a_m(x - x_0)(x - x_1)\ldots(x - x_{n-1}) = \\ = a_1((x - x_0 + h) - (x - x_0)) + a_2((x - x_0 + h)(x - x_1 + h) - \\ - (x - x_0)(x - x_1)) + a_3((x - x_0 + h)(x - x_1 + h)(x - x_2 + h) - \\ - -(x - x_0)(x - x_1)(x - x_2)) + \ldots + a^m((x - x_0 + h)(x - x_1 + h)\ldots \\ \ldots(x - x_{n-1} + h) - a_m(x - x_0)(x - x_1)\ldots(x - x_{n-1})) = \\ h = ha_1 + 2ha_2(x - x_0) + 3ha_3(x - x_0)(x - x_1) + \ldots + \\ + nha_m(x - x_0)(x - x_1)\ldots(x - x_{n-2}).$$

Обчислимо першу кінцеву різницю многочлена в точці x_0. Тут також всі члени, окрім першого, обернуться на нуль, і, отже, $\Delta P_m(x_0) = a_1 h$, але $\Delta P_m(x_0) = f(x_1) - f(x_0) = y_1 - y_0 = \Delta y_0$, звідки $\Delta y_0 = a_1 h$ і $a_1 = \dfrac{\Delta y_0}{h}$.

Щоб визначити коефіцієнт a_2, складаємо кінцеву різницю другого порядку: $\Delta^2 P_m(x) = \Delta P_m(x+h) - \Delta P_m(x)$. Після перетворень отримаємо

$$\Delta^2 P_m(x) = 2!h^2 a_2 + 2 \cdot 3 \cdot h^2 a_3 (x - x_0) + \ldots + (n-1) n h^2 a_m (x - x_0) \ldots (x - x_{n-3}).$$

Вважаємо $x=x_0$; тоді всі члени, окрім першого, знов обернуться на нуль і $\Delta^2 P_m(x_0) = \Delta^2 y_0 = 2!h^2 a_2$. Звідси $a_2 = \dfrac{\Delta^2 y_0}{2!h^2}$. Обчислюючи кінцеві різниці більш високих порядків і вважаючи $x=x_0$, прийдемо до загальної формули для отримання коефіцієнтів: $a_k = \dfrac{\Delta^k y_0}{k! h^k}$, $k = \overline{0,m}$, де будемо вважати, що $0!=1$ та $\Delta^0 y = y$. Підставивши знайденні значення коефіцієнтів a_k у вираз (КР.38), отримаємо ***першу інтерполяційну формулу Ньютона.***

$$P_m(x) = y_0 + \frac{\Delta y_0}{1!h}(x - x_0) + \frac{\Delta^2 y_0}{2!h^2}(x - x_0)(x - x_1) + \ldots + \frac{\Delta^m y_0}{m!h^m}(x - x_0) \ldots (x - x_{n-1}). \qquad \text{(КР.39)}$$

На практиці часто використовують формулу Ньютона в іншому вигляді. Для цього введемо заміну $q = (x - x_0)/h$, де h – крок інтерполяції, а q – число кроків. Тоді перша інтерполяційна формула Ньютона прийме наступний вигляд:

$$P_m(x) = y_0 + q\Delta y_0 + \frac{q(q-1)}{2!}\Delta^2 y_0 + \ldots + \frac{q(q-1)\ldots(q-n+1)}{m!}\Delta^m y_0. \qquad \text{(КР.40)}$$

Формулу (КР.40) зручно використати для інтерполювання на початку відрізку інтерполяції $[a,b]$, де q мале за абсолютною величиною. Якщо за число вузлів інтерполяції прийняти $n=1$, то отримаємо формулу ***лінійного інтерполювання*** $P_1(x) = y_0 + q\Delta y_0$. При $n=2$ отримаємо формулу параболічного, або ***квадратичного інтерполювання*** $P_2(x) = y_0 + q\Delta y_0 + \dfrac{q(q-1)}{2!}\Delta^2 y_0$.

На практиці часто буває необхідно зменшити крок інтерполяції будь-якої таблиці з рівновіддаленими аргументами. В таблиці можна вважати, що кількість вузлів інтерполяції необмежена. Тоді вибирають n так, щоб кінцева різниця $\Delta^n y_i$ була постійна із заданим ступенем точності. За початкове значення x_0 можна вибирати будь-яке значення аргумента. Схема алгоритму інтерполяції табличної функції многочленом Ньютона представлена на рис. КР.22.

Приклад КР.10. Припустимо, що результати експерименту подані в табл. КР.2 (перших три стовпця). З таблиці видно, що $n=6$, $m=5$, $\Delta y = y_1 - y_0$, $\Delta^2 y = \Delta y_1 - \Delta y_0$, $\Delta^k y = \Delta^{k-1} y_1 - \Delta^{k-1} y_0$, $h = x_{i+1} - x_i$.

Таблиця КР.2

Результати експерименту

N	x	y	$\Delta^1 y$	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$
0	0	-1	3	12	6	0	0
1	1	2	15	18	6	0	
2	2	17	33	24	6		
3	3	50	57	30			
4	4	107	87				
5	5	194					

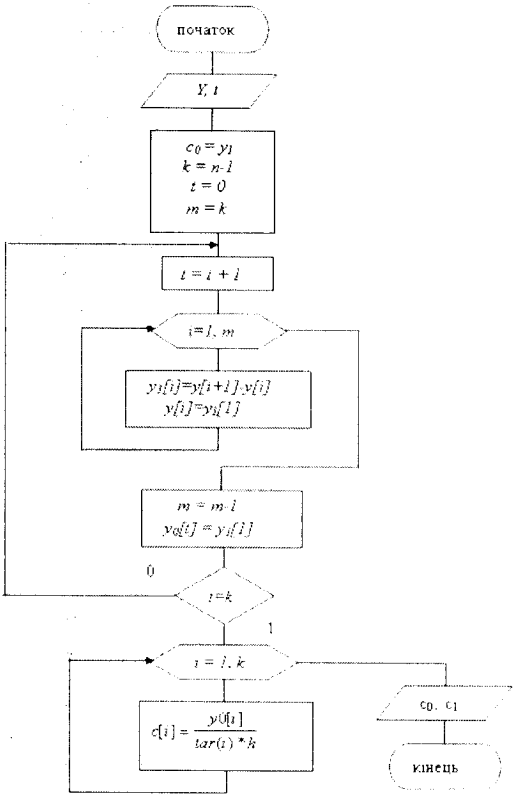


Рис. КР.22 Схема алгоритму інтерполяції табличної функції многочленом Ньютона

$$c_k = \frac{\Delta^k y_0}{(k!)h^k},\ c_0 = y_0 = -1,\ c_1 = \frac{\Delta y_0}{(1!)1} = 3,\ c_2 = \frac{\Delta^2 y_0}{(2!)1^2} = \frac{12}{2} = 6,$$

$$c_3 = \frac{\Delta^3 y_0}{(3!)1^3} = \frac{6}{6} = 1,\ c_4 = \frac{\Delta^4 y_0}{(4!)1^4} = 0.$$

Побудуємо многочлен Ньютона:

$$Q=-1+3x+6(x(x-1))+1(x(x-1)(x-2))+0=-1+3x+6x^2-6x+x^3-2x^2-x^2+2x=x^3+3x^2-x-1.$$

Висновки:

1. За допомогою аналітичної залежності x^3+3x^2-x-1 можливо отримати значення функції для x, які знаходяться між точками дослідження. Це називається задачею інтерполяції.

2. За допомогою аналітичної залежності x^3+3x^2-x-1 можна отримати значення функції за межами інтервалу дослідження. Наприклад, при $x=6 \Rightarrow y=317$. Ця задача називається екстраполяцією або прогнозуванням.

3.4. Метод Гауса

Нехай є $2n+2$ вузла інтерполяції $x_0,x_1,x_2,\ldots,x_n,x_{n+1}$, $x_k=x_0+kh$, $k\in Z$. Нехай $x=x(t)=x_0+th$, $t\in R$. Тоді $x(k)=x_k$. Інтерполяційна формула Гауса має вигляд:

$$P_{2n+2}(x_0+th)=f_0+tf_{1/2}^1+\frac{t(t-1)}{2!}f_0^2+\frac{t(t-1)(t+1)}{3!}f_{1/2}^3+\ldots+\frac{t(t-1)(t+1)\ldots(t+n)}{(2n+1)!}f_{1/2}^{2n+1}.$$

3.5. Метод Стірлінга

Нехай є $2n+1$ вузол інтерполяції $x_0,x_1,x_2,\ldots,x_n,x_{n+1}$, $x_k=x_0+kh$, $k\in Z$. Нехай $x=x(t)=x_0+th$, $t\in R$. Тоді $x(k)=x_k$. Інтерполяційна формула Стірлінга має вигляд:

$$P_{2n+1}(x_0+th)=f_0+tf_0^1+\frac{t^2}{2!}f_0^2+\ldots+\frac{t(t^2-1^2)\ldots(t^2-(n-1)^2)}{(2n-1)!}f_0^{2n-1}+$$

$$+\frac{t^2(t^2-1^2)\ldots(t^2-(n-1)^2)}{(2n)!}f_0^{2n}.$$

3.6. Метод Бесселя

Нехай є $2n+2$ вузла інтерполяції $x_0,x_1,x_2,\ldots,x_n,x_{n+1}$, $x_k=x_0+kh$, $k\in Z$. Нехай $x=x(t)=x_0+th$, $t\in R$. Тоді $x(k)=x_k$. Інтерполяційна формула Бесселя має вигляд:

$$P_{2n+2}(x_0+th)=f_{1/2}+\left(t-\frac{1}{2}\right)f_{1/2}^1+\frac{t(t-1)}{2!}f_{1/2}^2+\ldots+\frac{t(t^2-1^2)\ldots(t^2-(n-1)^2)(t-n)}{(2n)!}f_{1/2}^{2n}+$$

$$+\frac{t(t^2-1^2)\ldots(t^2-(n-1)^2)\ldots(t-n)(t-1/2)}{(2n+1)!}f_{1/2}^{2n+1}.$$

Приклад КР.11. Для функції $y=f(x)$, яка визначена на інтервалі $(0,\pi)$ і задана у вигляді таблиці значень $y_k=f(x_k)$:

1.0	2.0	3.0	4.0	5.0
6.0	7.0	8.0	9.0	10.0
11.0	12.0	13.0	14.4	15.6

скласти тригонометричний інтерполяційний многочлен.

Розв'язок. $Q(x)=2-0.7\cos 2x-0.3\sin 2x-0.3\cos 2x-0.3\cos 4x+0.2\sin 4x$.

Приклад КР.12. Для функції $y=f(x)$, яка визначена на інтервалі (0,1) і задана у вигляді таблиці значень $y_k=f(x_k)$:

16.0	17.0	18.0	19.0	20.0	21.0	22.0
23.0	24.0	25.0	26.0	27.0	28.0	29.0
30.0	31.0	32.0	33.0	34.0	35.0	36.0

скласти тригонометричний многочлен не нижче другого порядку.

Розв'язок. $Q(x) = -1 + \frac{7}{3}\cos 2\pi x - \frac{2}{\sqrt{3}}\sin 2\pi x - \cos 2\pi x$.

Приклад КР.13. Побудувати інтерполяційний многочлен Лагранжа для функції, яка задана таблицею

37.00	38.20	39.10	40.00	41.00
42.00	43.12	44.80	45.00	46.00

Розв'язок. $Q(x) = -\frac{1}{15}x^3 - \frac{3}{20}x2^3 + \frac{241}{60}x - 39$.

4. Розв'язування нелінійних та трансцендентних рівнянь

Необхідно знайти розв`язки рівняння:

$$f(x) = 0, \qquad \text{(КР.41)}$$

де $f(x)$ визначена та неперервна на деякому скінченному або нескінченному інтервалі $a<x>b$. Якщо функція є поліномом, тоді рівняння (КР.41) називається *алгебраїчним*, якщо ж у функцію $f(x)$ входять елементарні (тригонометричні, логарифмічні, показникові, тощо) функції, то таке рівняння називається *трансцендентним*.

Будь-яке значення x, що перетворює функцію на нуль, тобто таке, при якому $f(x)=0$, називається коренем рівняння (КР.41), а спосіб знаходження цього значення і є розв'язуванням рівняння (КР.41). Знайти корені рівняння вигляду (КР.41) точно вдається лише в деяких випадках. Крім того, часто рівняння містять коефіцієнти, що відомі лише приблизно і точне визначення коренів неможливе. Для цього і були розроблені методи чисельного розв'язування рівнянь виду (КР.41), які дозволяють знаходити наближені значення коренів цього рівняння.

При цьому доводиться вирішувати дві задачі:

- знаходження коренів, встановлення достатньо малих інтервалів (α,β), в яких міститься ізольований корінь рівняння (КР.41).
- уточнення коренів до заданої точності.

Для приблизного знаходження дійсних коренів або відокремлення коренів використовують різні методи. В деяких випадках межі коренів можна визначати з фізичних явищ, що описуються рівнянням (КР.41).

Іноді для приблизного визначення коренів знаходять простіше рівняння, яке має корені, що наближено дорівнюють кореням цього рівняння. Одним з найпростіших методів знаходження наближеного значення коренів є побудова графіку функції $y=f(x)$ та наближене визначення точок, в яких крива перетинає вісь X. У цих точках $y=0$ і відповідне значення x є дійсним коренем рівняння (КР.41). В деяких випадках рівняння (КР.41) краще записати у вигляді $f_1(x) = f_2(x)$, а потім побудувати графік двох функцій $y_1 = f_1(x)$ і $y_2 = f_2(x)$. Абсциси точок перетину задовольняють рівнянню $f_1(x) = f_2(x)$, а отже, і рівнянню (КР.41).Для уточнення дійсних коренів до заданої точності використовуються ітераційні методи. Використання того чи іншого методу залежить і від початкового наближення до кореня, існування та гладкість похідних функцій необхідної кількості коренів, їх кратності тощо. В ітераційних методах можна вибирати критерій закінчення процедури пошуку. Якщо функція $f(x)$ в заданій області змінюється повільно, тобто $|f(x)|<1$, то ітераційний процес слід припинити при виконанні умови $|x_{k+1}-x_k| < \varepsilon$, де x_k, x_{k+1} – наближення до кореня. Якщо ж функція змінюється швидко, тобто $|f(x)| \geq 1$, то ітераційний процес буде закінчуватися з виконанням умови: $|f(x)| < \varepsilon$.

Існує багато методів чисельного розв'язування нелінійних рівнянь. Розглянемо чотири з них: метод половинного ділення, метод пропорційних частин (хорд), метод Рибакова і метод простих ітерацій.

4.1. Метод половинного ділення

Цей метод використовується, коли відомо, що $f(x)$ неперервна на відрізку $[A,B]$ і $f(A)\,f(B)<0$. Суть методу полягає в наступному. Ділимо відрізок $[A,B]$, на якому шукаємо корінь рівняння, навпіл і, якщо $f((A+B)/2)\neq 0$, обираємо ту з половин $[A,(A+B)/2]$ чи $[(A+B)/2,B]$, на кінцях якої $f(x)$ має протилежні знаки. Новий відрізок знову ділимо навпіл і повторюємо, поки не отримаємо корінь рівняння з заданою точністю. Метод половинного ділення зручно застосовувати для грубого знаходження кореня цього рівняння, тому що зі збільшенням точності зростає об'єм роботи, яка буде виконуватися, через повільну збіжність ітераційного процесу.

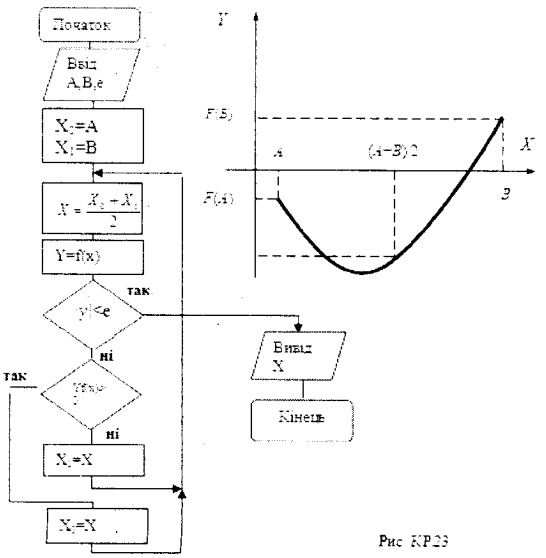


Рис. КР.23

Опишемо схему методу. До початку обчислення задаємо число ε (точність, з якою треба отримати корінь рівняння). Потім:

1) надамо значення $x_0=A$; $x_1=B$;
2) для кожного кроку процесу:
   обчислюємо $x=(x_0+x_1)/2$, $y=f(x)$; перевіряємо нерівність $|y|<\varepsilon$ і, якщо нерівність виконується, то x вважаємо коренем рівняння;
   якщо нерівність не виконується, тоді, якщо $y \cdot f(x)>0$ надамо x_0 значення x; а інакше – надамо x_1 значення x та повторюємо обчислення з пункту 2).

Дамо геометричну інтерпретацію методу половинного ділення. Нехай функція $f(x)$ має графік зображений на рис. КР.23. При обчисленні значення $(A+B)/2$ знаходимо середину відрізку $[A,B]$. Обчислення величини $f((A+B)/2)$ означає опустити перпендикуляр з точки $(A+B)/2$ до перетину з графіком функції $f(x)$. Якщо $|f(x)|>\varepsilon$, тоді для подальшого розгляду обираємо той відрізок, на кінцях якого $f(x)$ приймає значення різних знаків. У цьому випадку – це відрізок $[(A+B)/2,B]$. На цьому крок ітераційного процесу закінчується. Повторюємо, поки не буде виконана умова закінчення ітераційного процесу.

4.2. Метод пропорційних частин (хорд)

Нехай треба знайти корінь рівняння $f(x)$ на відрізку $[A,B]$ і відомо, що $f(x)$ неперервна на $[A,B]$ і $f(A)\,f(B)<0$. Крім цього, $f'(x)$ та $f''(x)$ на відрізку $[A,B]$ зберігає свій знак. Для розв'язування рівняння доцільно використати метод пропорційних частин (метод хорд), котрий дає розв'язки задачі для досить малих ε за меншу кількість

арифметичних операцій, ніж метод половинного ділення. Для наближеного кореня рівняння (КР.41) замінимо функцію $f(x)$ на відрізку $[A,B]$ лінійною функцією

$$P(x) = f(A) + (x-A)(f(B) - f(A))/(B-A),$$

що на кінцях відрізка $[A,B]$ приймає ті ж самі значення, що й функція $f(x)$. За перше наближене значення кореня $f(x)$ беремо корінь функції $P(x)$, тобто знаходимо таке значення x_1, при якому $P(x_1) = 0$. Це значення x_1 (перше наближення) визначається за формулою $x_1 = A - f(A)(B-A)/(f(B) - f(A))$. Потім розглядаємо відрізки $[A,x]$, $[x,B]$ і обираємо з цих двох відрізків той, на кінцях якого функція $f(x)$ має різні знаки. Ті ж обчислення повторюємо на обраному відрізку і отримуємо друге наближення кореня x_2 і т. д., поки не отримаємо корінь рівняння (КР.41) із заданою точністю.

Блок-схема алгоритму методу пропорційних частин зображена на рис. КР.24.

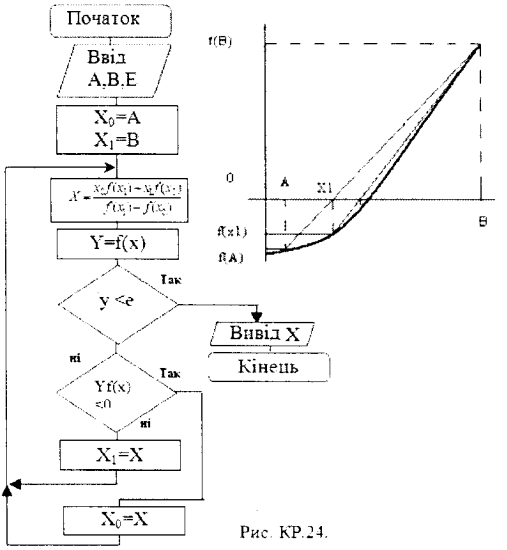


Рис. КР.24.

Схема для обчислення методом пропорційних частин реалізується так.

Спочатку задаємо точність ε, з якою треба отримати розв'язок, потім:

1) надамо значення $x_0 = A$ та $x_1 = B$;

2) обчислюємо наближення до кореня за формулою:

$$x = (x_0 \cdot f(x_1) - x_1 \cdot f(x_0))/(f(x_1) - f(x_0)). \qquad \text{(КР.42)}$$

3) перевіряємо виконання нерівності $|f(x)| < \varepsilon$, і, якщо ця нерівність виконується, тоді x вважаємо розв'язком рівняння (КР.41), а інакше – продовжуємо обчислення.

4). перевіряємо умову $f(x) \cdot f(x_0) < 0$, і при її виконанні надамо x_1 значення x, інакше – надамо x_0 значення x та повторюємо обчислення з пункту 2).

Дамо геометричну інтерпретацію методу. Нехай функція $y=f(x)$ має графік, зображений на рис. КР.24. З'єднаємо точки $(A, f(A))$ та $(B, f(B))$ прямою і точку перетину цієї прямої з віссю OX візьмемо за перше наближення кореня. Потім з'єднаємо точки $(x_1, f(x_1))$, $(B, f(B))$ прямою і точку x_2 перетину прямої з віссю OX приймаємо за друге наближення до кореня. Цей процес продовжуємо до тих пір, поки не виконається умова виходу з ітераційного процесу.

4.3. Метод Рибакова

Нехай необхідно знайти декілька дійсних коренів рівняння (КР.41) на відрізку $[A,B]$, причому похідна функції $f(x)$ на $[A,B]$ може мати розриви першого роду.

Для розв'язування цієї задачі доцільно використовувати метод Рибакова. Опишемо схему цього методу. До початку ітераційного процесу задамо точність ε, з якою необхідно отримати корені рівняння (КР.41). Обираємо деяке число M, що задовольняє нерівності (при умові $A \le x \le B$) $\max |f'(x)| \le M$.

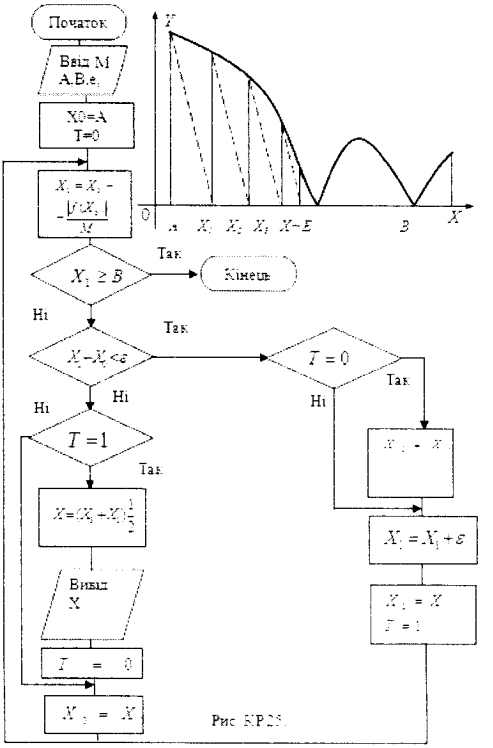


Рис. КР.23

Збільшення M не порушує збіжності методу, а тільки уповільнює. Потім:

1) надамо значення $x_0 = A$;

2) для кожного $k = 0,1,2\ldots$ (k – номер ітерації) обчислюємо (k+1)-е наближення до кореня за формулою: $x_{k+1} = x_k + |f(x_k)|/M$;
3) перевіряємо умову $x_{k+1} < B$; якщо умова не виконується, то вважаємо, що всі корені були знайдені, а інакше – продовжуємо подальші обчислення;
4) перевіряємо виконання умови $x_{k+1} - x_k > \varepsilon$,
5) при виконанні умови переходимо до п.2); а інакше – вважаємо x_{k+1} одним з коренів;
6) обчислюємо початкове наближення до наступного кореня: $x_0 = x_{k+1} + \varepsilon$.

Дамо геометричну інтерпретацію розглянутого ітераційного процесу. Нехай функція $y=|f(x)|$ має графік, зображений на рис. КР.25. Початкове наближення до кореня $x_0=A$. Через точку $(A, f(A))$ проводимо пряму (тангенс кута нахилу дорівнює M). Точку перетину цієї прямої з віссю OX вважаємо першим приближенням до першого кореня. Через точку $(x_1, f(x_1))$ проводимо пряму, що паралельна першій. Точку перетину цієї прямої з віссю OX приймаємо за друге наближення до першого кореня і т.д., поки не отримаємо корінь. Зробимо відступ від нього вправо на ε, починаємо шукати наступний корінь. Тут змінна T отримує значення 1, після першого виконання умови п.4. При цьому змінні x_2 і x_3 запам'ятовують значення x_{k+1}. Значення $x=(x_2+x_3)/2$ буде надруковано, якщо при новому значенні x_{k+1} не виконується умова 6. Інакше знову змінюється x_2 (отримує значення x_{k+1}), а x_3 зберігає своє значення.

4.4. Метод простих ітерацій

Для того, щоб розв'язати рівняння (КР.41) методом простих ітерацій його спочатку треба привести до вигляду:

$$x=\varphi(x). \qquad \text{(КР.43)}$$

При цьому повинна виконуватися умова збіжності: $|\varphi'(x)| \le q \le 1$ для x з проміжку $[A,B]$. На відрізку $[A,B]$ обирають початкове наближення x_0 (доцільно обирати в середині відрізка $[A,B]$) та знаходять наступні наближення до кореня за формулою: $x_k = \varphi(x_{k-1})$. Ітерації x_k обчислюються поки $|x_k - x_{k-1}| \le \varepsilon(1-q)/q$.

До початку ітераційного процесу задамо точність ε, з якою необхідно отримати корені рівняння (КР.41), та q – максимальне значення першої похідної. Схема методу:

1) надамо значення $x_0 = (A+B)/2$, це і буде першим наближенням до кореня;
2) приводимо рівняння типу (КР.41) до рівняння типу (КР.43), отримуємо нову функцію $\varphi(x)$:
3) знаходимо k-те (k=1,2,3,…) наближення до кореня за формулою $x_k = \varphi(x_{k-1})$;
4) перевіряємо умову $|x_k - x_{k-1}| q(1-q) > \varepsilon$. Якщо умова виконується, то шукаємо наступне наближення (повертаємося до пункту 3), а інакше – x_k вважаємо коренем рівняння. Блок-схема методу представлена на рис. КР.26.

5. Точність чисельного інтегрування

Похибку інтегрування можна оцінити. Наприклад, у формулі трапецій вона має такий вигляд: $R=-Pnh^3/12$, де P – максимальне значення похідної другого порядку на відрізку (a,b).

Для формули Сімпсона похибка визначається складніше: $R=-Pmh^5/90$, де P – максимальне значення похідної четвертого порядку на відрізку (a,b). При обчисленні інтегралу з використанням комп'ютера доцільно визначати абсолютну величину різниці двох значень визначеного інтегралу за кількість кроків n і з подвійною кількістю кроків $2n$: $|S_n - S_{2n}|=ze$, де z=3 для формул прямокутників і трапецій, z=15 для формули Сімпсона; e – похибка. Для досягнення заданої точності обчислення необхідно подвоювати значення n, поки значення w не буде менше, ніж задана точність e. Значення w знаходимо так: $w = |S_n - S_{2n}|/z$. Початкову кількість кроків n варто вибрати від 10 до 50.

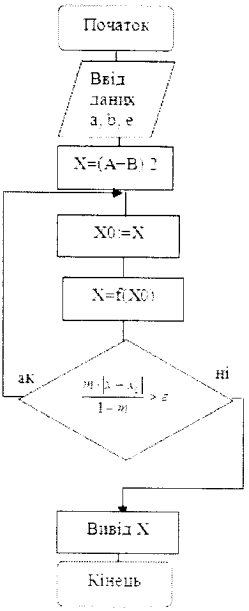


Рис. КР.26. Метод простих ітерацій

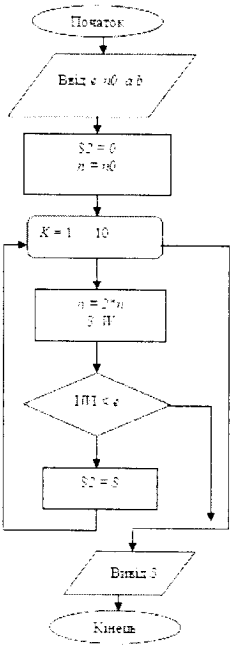


Рис. КР.27. Обчислення визначеного інтегралу

Для обчислення визначеного інтеграла можна використовувати блок-схему рис. КР.27. Тут блок 2 фіксує уведення вихідних даних. У наступному блоці задаємо початкове значення визначеного інтеграла S_2 при кількості відрізків, рівному n. Блок 4 – це початок циклу з параметром K, що обмежує кількість повторень (в цьому випадку – це 10). У циклі обчислюється подвоєне значення n, величина визначеного інтеграла S

одним із зазначених методів і різницею W між новим і попереднім значенням інтегралу: $|S_n - S_{2n}|$. Обчислюємо отриману похибку w. Блок 6 перевіряє закінчення процесу обчислень S. Якщо отримане значення W менше, ніж задана похибка e, завершуємо цикл. У протилежному випадку нове значення інтегралу запам'ятовується і процес обчислення інтегралу з подвійною кількістю кроків продовжується, але не більш 10 разів. Після виходу з циклу результати виводяться.

6. Методи розв'язування крайових задач

Методи розв'язування крайових задач розглядаються на прикладі звичайного диференціального рівняння другого порядку

$$y'' + p(x)y' + q(x)y = f(x) \qquad \text{(КР.44)}$$

при граничних умовах

$$\alpha_0 y(a) + \alpha_1 y'(a) = A, \quad \beta_0 y(b) + \beta_1 y'(b) = B; \qquad \text{(КР.45)}$$

при $|\alpha_0| + |\alpha_1| \neq 0$, $|\beta_0| + |\beta_1| \neq 0$, $a \le x \le b$. (де $\alpha_0,\alpha_1,\beta_0,\beta_1,a,b,A,B$ – *деякі числа).*

Методи розв'язування крайових задач можна розділити на три групи: різницеві, проекційні і методи, засновані на заміні розв'язку крайової задачі розв'язком декількох задач Коші (методи «стрілянини»).

6.1. Метод прогону

Розділимо область інтегрування $[a,b]$ на досить велике число рівних частин точками $x_i = x_0 + ih$, $i = 1 \ldots N$. Перетворимо систему (КР.44), (КР.45) до вигляду

$$y_{i-1} + m_i y_i + n_i y_{i-1} = h^2 f(x_i);\ i = 1 \ldots N\text{-}1, \qquad \text{(КР.46)}$$

де $r_i = 1/(1 + h/2p_i)$; $p_i = p(x_i)$; $q_i = q(x_i)$;

$$m_i = r_i(h^2 q_i - 2);\ n_i = r_i(1 - h/2p_i);\ f^*_i = h^2 r_i f(x_i). \qquad \text{(КР.47)}$$

Розв'язавши рівняння (КР.46) відносно y_i, одержуємо:

$$y_i = c_i(d_i - y_{i+1}),\ i = N, N\text{-}1, \ldots, 1; \qquad \text{(КР.48)}$$

$$c_i = 1/(m_i - n_i c_{i-1}),\ d_i = f^*_i - n_i c_{i-1} d_{i-1},\ i = 1 \ldots N\ldots \qquad \text{(КР.49)}$$

де $c_0 = \alpha_1/(\alpha_0 h - \alpha_1)$, $d_0 = h/\alpha_1$

За формулами (КР.49) обчислюються коефіцієнти c_i, d_i – прямий хід прогону. За формулою (КР.48) обчислюємо y_i, $i = N, N\text{-}1,\ldots,1$ – зворотний хід прогону. При цьому з (КР.46) і (КР.47) знаходимо $y_{N-1} = (Bh + \beta_1 c_{N-1} d)/(\beta_0 h + \beta_1(c+1))$.

Обчислювальна схема методу прогону:

1. Вибираємо крок h: $x_i = x_0 + ih$, $i = 1\ldots N$.
2. Обчислюємо

   $r_i = 1/(1 + h/2p_i)$; $m_i = r_i(h^2 q_i - 2)$; $n_i = r_i(1 - h/2p_i)$; $f^*_i = h^2 r_i f(x_i)$, $i = 1\ldots N$.
3. Обчислюємо $c_0 = \alpha_1/(\alpha_0 h - \alpha_1)$, $d_0 = hA/\alpha_1$; $h \neq \alpha_1/\alpha_0$.

$$c_i = 1/(m_i - n_i c_{i-1}),\ d_i = f_i - n_i c_{i-1} d_{i-1},\ i = 1 \ldots N.$$

4. Обчислюємо $y_{+1} = (Bh + \beta_1 c_{N-1} d)/(\beta_0 h + \beta_1 (c+1))$;

$$y_i = c_i(d_i - y_{i-1}),\ i = N, N-1, \ldots, 1...$$

```
Uses Crt; {підключення модуля Crt}
Var
      c,d,y: array [0..20] of real;{прогоночні коеф. і шукана функція}
      a,b,aa,bb,h,al0,all,b0,b1:real; {a,b - крайові умови}
      n : integer; {кількість вузлів інтегрування}
      i,j: byte;{службові}
      o:char;{відповідь користувача}
      flag : boolean;{службова}
procedure rab;forward;
procedure vyvod;forward;
procedure vvod;{процедура введення}
begin
   clrscr;
   textcolor(7);
   Write('УВЕДІТЬ  КІЛЬКІСТЬ   КРОКІВ ');ReadLn(n);
   writeln('УВЕДІТЬ КОЕФІЦІЄНТИ КРАЙОВИХ УМОВ ');
   write('a, al0, all, aa, b, b0, b1, bb');
   readln(a, al0, all, aa, b, b0, b1, bb);rab;
end;
procedure shap;{процедура виводу шапки}
begin
 repeat {основний цикл діалогу з користувачем}
   ClrScr;textcolor(15);
   writeln('СПЕЦ РОЗДІЛИ МАТЕМАТИКИ');writeln('ЛАБОРАТОРНА РОБОТА ');
   writeln('(c) AJ 7.05.2001');
   writeln('МЕТОД  ОРТОГОНАЛЬНОГО ПРОГОНУ (розв'язку КРАЙОВОЇ ЗАДАЧІ)');writeln;
   writeln('ЗРОБІТЬ ВИБІР');writeln('p   - ПРОДОВЖИТИ');writeln('Esc - ВИХІД');
   o:=readkey;   {відповідь користувача}
   o:=upcase(o);
   case o of
    'P' : vvod;{уведення даних}
     #27: halt {вихід із програми}
     else Continue;
   end;
 until flag;
end;
function p(x:real) : real;{функція p(x)}
begin p:=x*x+1; end;
function q(x:real) : real;{функція q(x)}
begin q:=x-1; end;
function f(x:real) : real;{функція f(x)}
begin f:=x+2; end;
procedure rab;{основна процедура обчислень}
var r,ff,m,mm,x : real;
begin
  h:=(b-a)/n;
  c[0]:= all/(al0*h-all);
  d[0]:= h*aa/all;
  for i:=1 to n do
   begin
     x:=a+i*h;
     r:=1/(1+h/2*p(x));
     m:=r*(h*h*q(x)-2);
     mm:=r*(1-h/2*p(x));
     ff:=h*h*r*f(x);
```

```
    c[i]:=1/(m-nn*c[i-1]);  {обчислення прогоночних коеф}
    d[i]:=ff-nn*c[i-1]*d[i-1];
   end;
  y[n+1]:=(bb*h+b1*c[n-1]*d[n])/(b0*h+b1*(c[n]+1));
  for i:=n downto 1 do
   begin                          {наповнення вектора B}
    y[i]:=c[i]*(d[i]-y[i+1]);
   end;
  vyvod;{перехід на вивід}
end;
procedure vyvod;{процедура виводу вектора B}
Var x: real;
Begin
  Clrscr; Textcolor (10);
  writeln('РЕЗУЛЬТАТ ');
  Textcolor (12);
  For i:=1 to n+1 do
   begin
    x:=a+(i-1)*h;
    writeln(x:0:2,'    ', y[i]:0:2,'  ');
   end;
  readln; shap;
end;
BEGIN
  flag := false;  shap;{виклик процедури виводу шапки}
END.
```

Результати роботи

Вхідні дані:

$p(x)=x^2+1$; $q(x)=x-1$; $f(x)=x+2$; $a=0$, $\alpha_0=2$, $\alpha_1=1$, $A=3$; $b=2$, $\beta_0=5$, $\beta_1=-2$, $B=3$; $h=5$.

Вихідні дані:

x_i	y_i
0.00	2.42
0.40	0.63
0.80	0.08
1.20	0.23
1.00	0.58
2.00	0.86

6.2. Проекційні методи (на прикладі методу Гальоркіна)

Сутність проекційних методів обчислювальної математики полягає в представленні розв'язку задачі множиною проекцій (відліків) у визначеній системі координатних функцій.

У традиційному методі, запропонованому Б. П. Гальоркіним і розвинутому у роботах М. В. Келдиша, наближений розв'язок $y(x)$ шукається у вигляді

$$y_n = a_1u_1(x) + a_2u_2(x) + \ldots + a_nu_n(x). \qquad \text{(КР.50)}$$

де $\varphi_1(x)$, $\varphi_2(x)$, ..., $\varphi_n(x)$, – система базисних функцій, що задовольняють вихідним граничним умовам; $a_1, a_2, \ldots, a_n$ – невідомі постійні коефіцієнти.

При підстановці (КР.50) у (КР.44) одержуємо функцію

$$R(x; a_1; a_2; \ldots; a_n) = L[y_n(x)] - f(x).$$

яку називають нев'язкою розв'язку. Очевидно, для точного розв'язку задачі

$$R(x; a_1; a_2; \ldots; a_n) = 0.$$

Коефіцієнти a_i визначимо з умови ортогональності нев'язки першим n функціям деякої системи функцій $\{\varphi_i, i = 0 \ldots n\}$:

$$\int_a^b L[\sum_{i=1}^{n} c_i u_i]\varphi_j(x)dx = \int_a^b f(x)\varphi_j(x)dx, \quad j = 1...n. \qquad \text{(КР.51)}$$

Такий метод розв'язку задачі називається *методом моментів*. Якщо при цьому $\varphi_i = u_i$, то отримуємо *метод Гальоркіна*. Система рівнянь (КР.51) є системою лінійних рівнянь щодо вектора $\vec{a}$ з матрицею $A_n = [\int_a^b L[u_i(x)]\varphi_j(x)dx]_1^n$. Розв'язок цієї системи є каркасом наближеного розв'язку крайової задачі. Система (КР.51) може бути представлена у вигляді

$$\sum_{k=1}^{n} c_{ki}a_i - d_i = 0, \qquad i = 1...n. \qquad \text{(КР.52)}$$

Коефіцієнти c_{ki}, d_i, обчислюються за формулами $c_{ki} = \int_a^b L(u_k)u_i(x)dx$ та

$$d_i = \int_a^b [f(x) - L(u_0)]u_i(x)dx, \qquad i = 1...n$$

Для рівняння (КР.44) $L(y) = y'' - p(x)y' - q(x)y$. Тому

$$c_{ki} = -\int_a^b [p(x)u_k'(x) + q(x)u_k(x)u_i(x)]dx \quad k = 1...n; i = 0...n-1, \; c_{ik} = c_{ki}$$

$$d_i = \int_a^b [p(x)u_0'(x)u_i'(x) + q(x)u_0(x)u_i(x) + f(x)u_i(x)]dx \quad i = 1...n$$

Функцію $u_0(x)$ можна вибирати довільно, але так щоб $u_0(a) = A$, $u_0(b) = B$. Нехай $u_0(x) = \alpha + \beta x$. Тоді з (КР.45) одержуємо: $\beta = \dfrac{A\beta_0 - B\alpha_0}{\alpha_0\beta_0(a-b) + \beta_0\alpha_1 - \alpha_0\beta_1}$ $:\alpha = A - a\beta$.

Інші функції $u_i(x)$ можна обчислити за кожним з правил:

$$u_i(x) = (x-a)^i(x-b);$$
$$u_i(x) = (x-a)(x-b)^i.$$

Обчислювальна схема проекційних методів (на прикладі методу Гальоркіна):

1. Задаємо степінь полінома n, тобто $y_n = a_1u_1(x) + a_2u_2 + \ldots + a_nu_n(x)$.
2. Обчислюємо коефіцієнти α і β функції $u_0(x)$:

$$\beta = \frac{A\beta_0 - B\alpha_0}{\alpha_0\beta_0(a-b) + \beta_0\alpha_1 - \alpha_0\beta_1} \quad :\alpha = A - a\beta.$$

3. Вибираємо функції $u_k(x)=(x-a_1)^k(x-a_2)$; $a_1=\dfrac{1}{2(d-\sqrt{d^2+4e})}$:

$$a_1=\frac{1}{2(d+\sqrt{d^2+4e})};\ e=a(2\alpha_1+a)+d(\alpha_1-a).$$

4. Обчислюємо $d=\dfrac{b^2-a^2+2(b\beta_1-a\alpha_1)}{b-a+\alpha_1-\beta_1}$

$$u_k(x)\ u_i(x)=(x\text{-}a_1)^{k+i}(x\text{-}a_2)^2;\ u'_0(x)\ u'_i(x)=i\beta(x\text{-}a_1)^{i-1}(x\text{-}a_2);$$
$$u_0(x)\ u_i(x)=(x\text{-}a_1)^i(\beta x^2+(\alpha-\beta a_1)x-\alpha a_1);\ k,i=1\ldots m;$$
$$u'_k(x)\ u'_i(x)=(x\text{-}a_1)^{k+i-2}(e_3x^2+e_2x+e_3).$$
$$e_1=(ka_2+a_1)(ia_2+a_1);\ e_2=-a_1(2+k+i);\ e_3=(k+1)(i+1).$$

5. Обчислюємо c_{ki}, d_i:

$$c_{ki}=-\int_a^b[p(x)(e_1+e_2x+e_3x^2)+q(x)\sum_{t=1}^{5}g_tx^{t-1}](x-a_1)^{k+i-2}dx,$$

$$d_i=\int_a^b[p(x)i\beta(x-a_2)+q(x)\sum_{t=1}^{4}r_tx^{t-1}+f(x)(x^2-dx-e](x-a_1)^{i-1}dx,$$

$$g_1=e^2;\ g_2=2de;\ g_3=d^2-2e;\ g_4=-2d;\ g_5=1,$$
$$r_1=\alpha a_1^2;\ r_2=a_1(\beta a_1-2\alpha);\ r_3=-a_1\beta;\ r_4=1.$$

6. Розв'язуємо систему лінійних рівнянь $\sum_{k=1}^{n}c_{ki}a_i-d_i=0,\quad i=1\ldots m$

7. Знайдені значення a_i підставляємо в ряд $y_m(x)=u_0(x)+\sum_{i=1}^{m}a_iu_i(x).$

```
Uses Crt; {підключення модуля Crt}
Var
      m,i,j,k:byte{m - степінь полінома, i,j,k - службові };
      o:char;{відповідь користувача}
      flag : boolean;{службова}
      a, al0, all, a4, b, b0, b1, bb,alf,bet:real;{крайові умови}
      d,a1,a2,e,e1,e2,e3:real;{службові}
      c: array [1..10,1..10] of real;{коефіцієнти матриці m*m}
      dd,x: array[1..10] of real;{праві частини і шуканий вектор}
      r: array [1..4] of real;{допоміжний вектор}
      g: array [1..5] of real;{допоміжний вектор}
procedure rab;forward;
procedure matrix;forward;
procedure poisk;forward;
procedure iskl;forward;  {опис процедур}
procedure vyvod;forward;
procedure matrix;{основна процедура перетворень матриці C}
var q : real;
begin
  i:=1;
  repeat
    if c[i,i]=0 then poisk;
```

```
    k:=i+1;
    repeat
     iskl;
     k:=k+1;
    until k>m;
    inc(i);
  until i>m-1;
  x[m]:=dd[m]/c[m,m]; {визначення m-ого кореня}
  i:=i-1;
  repeat {цикл обчислення всіх інших коренів}
    q:=0;
    for k:=i+1 to m do
     q:=q+c[i,k]*x[k];
    x[i]:=(dd[i]-q)/c[i,i];
    dec(i);
  until i<1;
end;
procedure poisk;{процедура пошуку ненульового ведучого елемента}
var h:byte;
    buf : real;
begin
  for j:=i+1 to m do
   if c[j,i]<>0 then
     begin
       for h:=1 to m do
         begin
          buf:=c[i,h];
          c[i,h]:=c[j,h]; {заміна рядків}
          c[j,h]:=buf;
         end;
       exit;
     end;
end;
procedure iskl;{процедура виключення елемента C<k,i>}
var r,mm: real;{допоміжні}
begin
  r:=-(c[k,i]/c[i,i]);
  for j:=1 to m do
   begin
    mm:=r*c[i,j];
     c[k,j]:=c[k,j]+mm;
   end;
  dd[k]:=dd[k]+r*dd[i];
end;
function px(x:real) : real;{функція p(x)}
begin px:=x*x-1; end;
function qx(x:real) : real;{функція q(x)}
begin qx:=x*x; end;
function fx(x:real) : real;{функція f(x)}
begin fx:=x-1; end;
procedure vvod;{процедура введення}
begin
   clrscr; textcolor(7);
   Write('ВВЕДІТЬ СТЕПІНЬ ПОЛІНОМА '); Readln(m);
   writeln('ВВЕДІТЬ КОЕФІЦІЄНТИ КРАЙОВИХ УМОВ ');
   write('a, a10, a11, aa, b, b0, b1, bb ');
   readln(a, a10, a11, aa, b, b0, b1, bb); rab;
end;
procedure shap;{процедура виводу шапки}
```

```
begin
 repeat {основний цикл діалогу з користувачем}
   ClrScr; textcolor(15);
   writeln('СПЕЦ РОЗДІЛИ МАТЕМАТИКИ');writeln('ЛАБОРАТОРНА РОБОТА ');
   writeln('(c) AJ 7.05.2001');
   writeln('МЕТОД ГАЛЬОРКІНА  (розв'язку КРАЙОВОЇ ЗАДАЧІ)'); writeln;
   writeln('ЗРОБІТЬ ВИБІР');writeln('p   - ПРОДОВЖИТИ');writeln('Esc - ВИХІД');
   o:=readkey;   {відповідь користувача}
   o:=upcase(o);
   case o of
    'P' : vvod;{уведення даних}
     #27: halt {вихід із програми}
     else Continue;
   end;
 until flag;
end;
function suml(x:real):real;{обчислення проміжної суми}
var i:byte;
    buf : real;
begin
  buf:=0;
  for i:=1 to 5 do
   buf:=buf + g[i]*exp((i-1)*ln(abs(x)));
  suml:=buf;
end;
function sum2(x:real):real;{обчислення проміжної суми}
var i:byte;
    buf : real;
begin
  buf:=0;
  for i:=1 to 4 do
   buf:=buf + r[i]*exp((i-1)*ln(abs(x)));
  sum2:=buf;
end;
FUNCTION step (x:real;n:integer):real;{піднесення до степеня}
begin
    IF n = 0 THEN step := 1
    ELSE step := step(x,n - 1) * x;
end;{процедура обчислення інтеграла}
procedure integ(a,b : real; var s : real;func : byte; o,p: byte; dd: real);
var ss,w,x,d,ch1,ch2 : real;
    n,k,i : integer;
begin
  ss := 0; n := 10;
  for k:= 1 to 20 do          {цикл - обмеження кількості повторень}
   begin
     n:=2*n;
     d := (b-a)/n;          {приріст}
     s:=0;
     x:=a;
     for i:=0 to n-1 do     {обчислення інтегральної суми}
       begin
         case func of
           1 : begin
             ch1:=(px(x)*(e1+e2*x+e3*x*x)+qx(x)*suml(x))*step((x-a),(o-p-2)) ;
             ch2:=(px(x+d)*(e1+e2*(x+d)+e3*(x+d)*(x+d))+qx(x-d)*suml(x+d))*
                step((x+d-a),(o-p-2)) ;;
              s:=s+ (ch1+ch2)*d;
               end;
```

```
          2 :begin
               ch1:=(px(x)*p*bet*(x-a2)+qx(x)*sum2(x)+fx(x)*
                    (sqr(x)-dd*x-e))*step((x-a1),(p-1));
               ch2:=(px(x+d)*p*bet*(x+d-a2)+qx(x+d)*sum2(x+d)+fx(x+d)*
                    (sqr(x+d)-dd*(x+d)-e))*step((x-a1+d),(p-1));
              s:=s+(ch1+ch2)*d;  {вибір функції}
             end;
        end;
        x:=x+d;
      end;
    s:=0.5*s; w:=s-ss;
    if abs(w)<e then break else ss:=s;   {умова кінця обчислень}
  end;
end;
procedure rab;{основна процедура обчислень}
var buf: real ;
begin
  bet:=(b0*aa-a10*bb)/(a10*b0*(a-b)+b0*a11-a10*b1); alf:=aa-a*bet;
  d:=(b*b-a*a+2*(b*b1-a*a11))/(b-a+a11-b1); e:=a*(2*a1+a)+d*(a1-a);
  a1:=1/2*(d-sqrt(d*d+4*e)); a2:=0.5*(d+sqrt(d*d+4*e));
  r[1]:=alf*a1*a1; r[2]:=a1*(bet*a1-2*alf); r[3]:=-a1*bet; r[4]:=1;
  g[1]:=e*e; g[2]:=2*d*e; g[3]:=d*d-2*e; g[4]:=-2*d; g[5]:=1;
  for k:=1 to m do
   for i:=1 to m do
    begin
      e1:=(k*a2+a1)*(i*a2+a1);
      e2:=-a1*(2+k+i)-a2*(2*k*i+k+i);
      e3:=(k+1)*(i+1);
      integ(a,b,buf,1,k,i,d);          {цикл наповнення C}
      c[k,i]:=-buf;
      integ(a,b,dd[i],2,k,i,d);    end;
  matrix;{перехід на визначення коренів a_i}
  vyvod;{перехід на вивід }
end;
procedure vyvod;{ вивід результату}
begin
  clrscr; textcolor(10);writeln('РЕЗУЛЬТАТ (вектор A коефіцієнтів)');textcolor(12);
  for i:=1 to m do writeln(x[i]:6:2);
  readln;shap;
end;
BEGIN
  flag := false;shap;{виклик процедури  виводу шапки}
END.
```

Результати роботи. Вхідні дані:

$p(x)=x^2-1$; $q(x)=x^2$; $f(x)=x+1$; $a=-1$. $\alpha_0=3$, $\alpha_1=0$. $A=1$; $b=1$, $\beta_0=2$, $\beta_1=0$. $B=4$; $m=5$.

Вихідні дані (вектор коефіцієнтів a): 36,02 –62,72 55,03 –23,98 6,24

7. Методи пошуку максимуму функції двох змінних

7.1. Метод покоординатного спуску

Викладемо цей метод на прикладі функції двох змінних $F(x,y)$. Виберемо нульове наближення x_0 y_0 Фіксуємо значення координати $y=y_0$. Тоді функція буде залежати тільки від однієї змінних x; позначимо її через $f_1(x)=F(x,y_0)$. Використовуючи один з вище розглянутих методів пошуку максимуму функції однієї змінної знайдемо

максимум функції однієї змінної $f_1(x)$ і позначимо його через x_1. Зробили крок з точки (x_0,y_0) у точку (x_1,y_0) за напрямом, рівнобіжним осі x. На цьому кроці значення функції збільшилося. Потім з нової точки зробимо спуск за напрямом, рівнобіжним осі y. Тобто розглянемо $f_2(x)=F(x_1,y)$, знайдемо її максимум і позначимо його через y_1. Другий крок приводить нас у точку (x_1,y_1). Прихід у точку (x_1,y_1) завершує цикл спусків. Будемо повторювати цикли. На кожному спуску функція не спадає, і при цьому значення функції обмежені зверху її значенням у максимумі. Отже, ітерації сходяться до деякої межі $F \le F_{max}$.

7.2. Метод спірального покоординатного спуска

Цей метод дуже подібний на попередній, але відрізняється тим, що крок h змінюється щоразу при переході від пошуку максимуму по одній змінній до пошуку максимуму по інший змінній. У тривимірному просторі це нагадує спуск у западину по спіралі. Звичайно цей метод дає деяке зменшення часу пошуку (рис. КР.28).

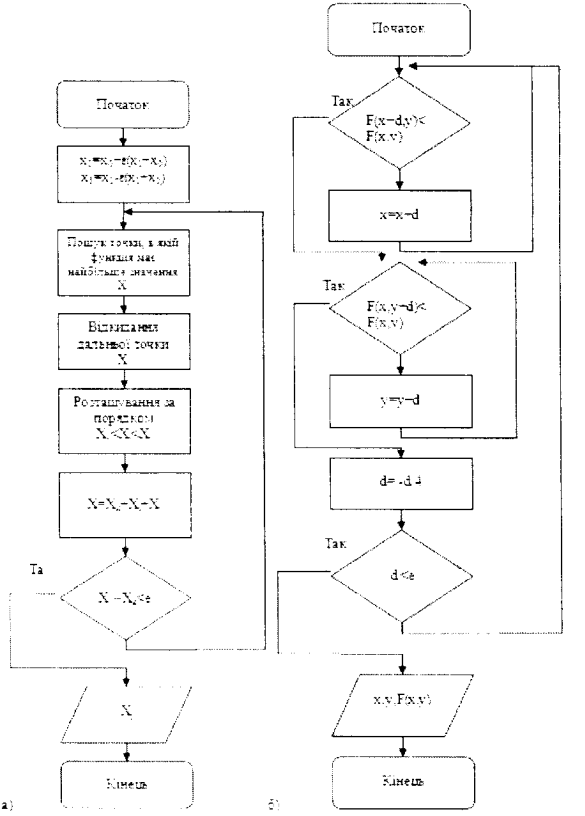


Рис. КР. 28. Блок-схема: а)методу золотої середини та б) покоординатного спуску.

Блок-схема методу спірального покоординатного спуску відрізняється від попередніх тим, що крок h змінюється щоразу при переході від пошуку максимуму по одній змінній до пошуку максимуму по інший змінній (рис. КР.29).

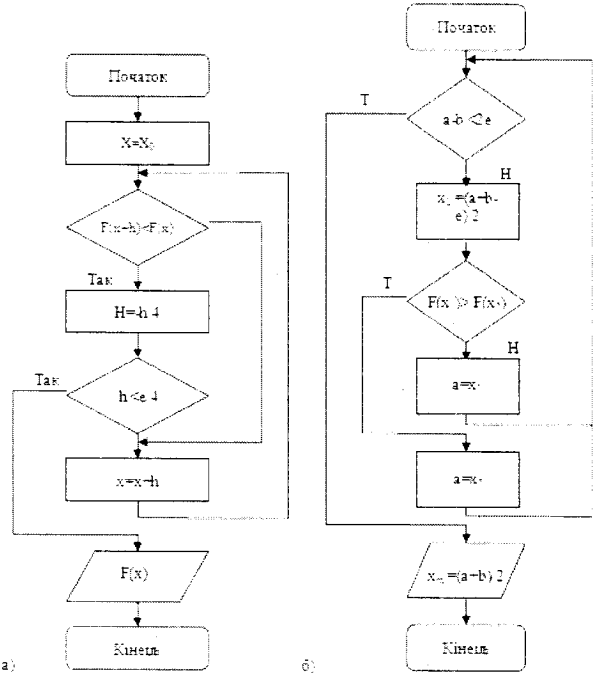


Рис. КР. 29. Блок-схема: а) порозрядного пошуку та б) методу дихотомії

8. Розв'язування чисельними методами диференціальних рівнянь

8.1. Основні поняття та визначення

Рівняння, у якому невідома функція є під знаком похідної чи диференціалу, називається *диференціальним рівнянням*. Наприклад,

$$\frac{dy}{dx} = 2(y-3); \quad \frac{d^2y}{dt^2} = t+1; \quad \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0; \quad y' = x^2; \quad xdy = y^3 dx;$$

Якщо невідома функція, що належить до диференціального рівняння, залежить тільки від однієї незалежної змінної, то диференціальне рівняння є *звичайним*. Наприклад, диференціальні рівняння $x^2 \dfrac{d^2y}{dx^2} = 2$; $2sdt = tds$ відносяться до звичайних.

Якщо ж невідома функція, що належить до диференціального рівняння, є функцією двох чи більшого числа незалежних змінних, то таке рівняння називається *диференціальним рівнянням у частинних похідних*. Наприклад, диференціальне рівняння $\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}=0;$ відноситься до рівняння в частинних похідних. *Порядком* диференціального рівняння є найвищий порядок похідної (чи диференціала), що входить у рівняння. Розглянемо звичайні диференціальні рівняння. Звичайне диференціальне рівняння n-го порядку в самому загальному випадку містить незалежну змінну, невідому функцію і її похідні чи диференціали до n-го порядку включно і має вигляд

$$F(x, y, y', y'', ..., y^{(n)}) = 0 . \qquad \text{(КР.53)}$$

У цьому рівнянні x – незалежна змінна, y – невідома функція, $y, y', y'', ..., y^{(n)}$ – похідні цієї функції. *Розв'язком* (чи *інтегралом*) рівняння (КР.53) називається будь-яка диференційована функція $y = \varphi(x)$, що задовольняє цьому рівнянню, тобто така, після підстановки якої у рівняння (КР.53), воно перетворюється в тотожність. Графік розв'язку звичайного диференціального рівняння називається *інтегральною кривою* цього рівняння. Розв'язок диференціального рівняння, що містить стільки незалежних довільних (постійних) параметрів, як його порядок, називається *загальним розв'язком* (чи *загальним інтегралом*) цього рівняння.

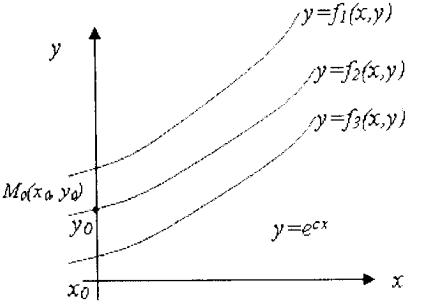


Рис. КР.30. Сімейство інтегральних кривих диференціального рівняння (КР.53)

Геометрично загальний розв'язок диференціального рівняння є сімейством інтегральних кривих цього рівняння (рис. КР.30).

Частинним розв'язком диференціального рівняння є будь-який розв'язок, що отриманий із загального при визначених числових значеннях довільних постійних (рис. КР.30). Довільні постійні, що входять до загального розв'язку, визначаються з *початкових* або *крайових умов*.

Задача з *початковими умовами* ставиться так: знайти розв'язок $y = \varphi(x)$ рівняння $y^{(n)} = f(x, y, y', y'', ..., y^{(n-1)})$, що задовольняє додатковим умовам, які полягають в тому,

що розв'язок $y=\varphi(x)$ разом зі своїми похідними до (n-1)-го порядку повинен приймати задані числові значення $y, y', y'', ..., y^{(n)}$ при заданому числовому значенні $x = x_0$ незалежної змінної x.

Такі умови є *початковими умовами*, а задача знаходження розв'язку $y=\varphi(x)$ диференціального рівняння (КР.53), що задовольняє заданим початковим умовам – *задачею з початковими умовами* або *задачею Коші*.

Задача з *крайовими умовами* ставиться так: знайти розв'язок $y=\varphi(x)$ рівняння $y^{(n)} = f(x, y, y', y'', ..., y^{(n-1)})$, що задовольняє додатковим умовам, які полягають в тому, що розв'язок $y=\varphi(x)$ разом зі своїми похідними до (n-1)-го порядку повинен приймати задані числові значення $y, y', y'', ..., y^{(n)}$ при заданому числовому значенні $x = x_0$ та $y_n, y_n', y_n'', ..., y_n^{(n-1)}$, при заданому числовому значенні $x = x_n$ незалежної змінної x. Такі умови називаються *крайовими умовами*, а задача знаходження розв'язку $y=\varphi(x)$ диференціального рівняння (КР.53), що задовольняє заданим крайовим умовам – *крайовою задачею*.

У випадку рівняння першого порядку, тобто при n=1, одержуємо задачу Коші для рівняння $y' = f(x, y)$ з початковою умовою $x = x_0$, $y = y_0$. Геометрично задача Коші для рівняння першого порядку полягає в тому, що з усіх інтегральних кривих, що представляють собою загальний розв'язок, потрібно знайти ту інтегральну криву, що проходить через точку M_0 з координатами $x = x_0$, $y = y_0$ (рис. КР.30). Часто в задачі Коші у ролі незалежної змінної виступає час t. Прикладом може бути задача про вільні коливання тіла, яке підвішене на пружині. Рухи такого тіла описуються диференційним рівнянням, в якому незалежною змінною є час t. Якщо додаткові умови задані у вигляді значень переміщень чи швидкості при t=0, то це також задача Коші. Задача Коші має єдиний розв'язок, що задовольняє умові $f(x_0) = y_0$, якщо функція $f(x, y)$ неперервна в деякій області $R_{[a,b]} = \{|x - x_0| < a, |y - y_0| < b\}$ і задовольняє в цій області *умові Ліпшица*. $|f(x, \bar{y}) - |f(x, y)| < N|\bar{y} - y|$, де N – постійна Ліпшица, що залежить від a і b (a і b – границі області).

Методи точного інтегрування диференціальних рівнянь придатні лише для порівняно невеликої частини рівнянь, що зустрічаються на практиці. Тому в задачах моделювання та дослідження складних технічних систем, наприклад, систем автоматизованого управління, велике значення мають методи наближеного розв'язування диференційних рівнянь, що у залежності від форми представлення розв'язку можна розділити на дві групи:

1) ***аналітичні методи***, що дають наближений розв'язок диференційного рівняння у вигляді аналітичного виразу;

2) ***чисельні методи***, що дають наближений розв'язок у вигляді таблиці.

Похибки. Перед тим, як перейти до розгляду методів чисельного розв'язування диференційних рівнянь, зупинимося на джерелах похибок, що пов'язані з чисельною апроксимацією. Таких джерел три:

1. *Похибка заокруглення* зумовлена обмеженнями на представлення чисел в комп'ютері, тому що число значущих цифр, що запам'ятовується і використовується в обчисленнях, обмежене.

2. *Похибка відсічення* пов'язана з тим, що для апроксимації функції замість

$$y = y_0, y' = y'_0, y'' = y''_0, \ldots, y^{(n)} = y_0^{(n)} \quad \text{при} \quad x = x_0 \qquad \text{(КР.54)}$$

нескінчених рядів часто використовується лише декілька перших їх членів.

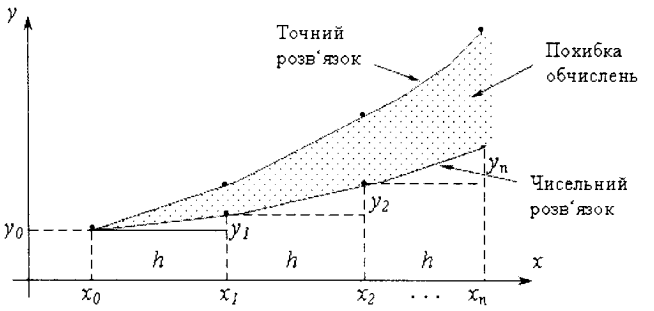


Рис.КР.31. Геометричне представлення накопичення похибки в процесі обчислень

Це звичайний для чисельних методів прийом, що є джерелом похибки, цілком зумовленої використаним методом і не залежать від характеристик ПК.

3. *Похибка поширення* є результатом накопичення похибок, що з'явились у попередніх результатах обчислення. Оскільки ні один з наближених методів не може дати зовсім точних результатів, то будь-яка похибка, що виникла в процесі обчислень, зберігається і на наступних стадіях обчислення (рис. КР.31).

Вказані три джерела похибок є причиною помилок двох типів. *Локальна помилка* – сума похибок, що вносяться у розрахунковий процес на кожному етапі обчислення. *Глобальна помилка* – різниця між обчисленим та точним значеннями величини на кожному етапі реалізації чисельного алгоритму, що визначає сумарну похибку, що накопичується з моменту початку обчислення.

8.2. Класифікація чисельних методів розв'язування задачі Коші

Протягом багатьох років чисельний розв'язок задачі Коші був об'єктом пильної уваги науковців, оскільки він часто застосовується в різних галузях науки та техніки. Тому і кількість розроблених для нього методів дуже велика.

Чисельні методи розв'язування задачі Коші поділяються на 3 групи: одноточкові; багатоточкові (методи прогнозу та корекції); методи з автоматичним вибором кроку

інтегрування. На рис. КР.32 подана класифікація найбільш відомих чисельних методів розв`язування диференційних рівнянь (ДР) на ПК.

До *одноточкових методів* відносять методи, які мають певні загальні риси, такі як:

1. В основі усіх одноточкових методів лежить розклад функції в ряд Тейлора, в якому зберігаються члени, що мають h в степені до k включно. Ціле число k називається порядком методу. Похибка на кроці має порядок $k+1$.

2. Всі одноточкові методи не потребують дійсного обчислення похідних, тому що обчислюється лише сама функція, однак можуть бути потрібними її значення в деяких проміжкових точках. Це означає додаткові затрати часу і зусиль.

3. Для отримання інформації у новій точці, потрібно мати дані лише в попередній точці. Цю властивість можна назвати „самостартуванням". Властивість „самостартування" дозволяє легко змінювати величину кроку h.

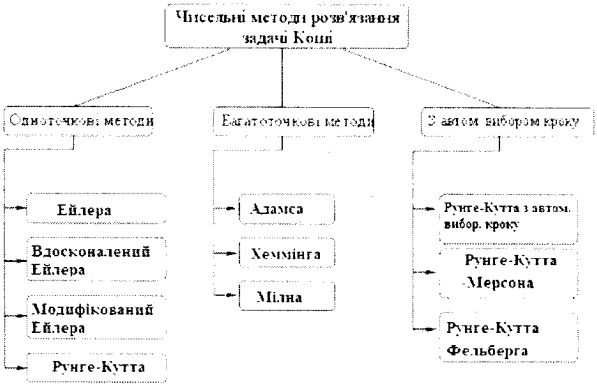


Рис.КР.32. Класифікація чисельних методів розв'язування задачі Коші.

4. В порівнянні з одноточковими методами *методи прогнозу і корекції* мають ряд особливостей:

4.1. Для реалізації методів прогнозу і корекції необхідно мати інформацію про декілька попередніх точок (вони не відносяться до „самостартуючих" методів). Тому для отримання додаткової інформації доводиться застосовувати одноточковий метод. Якщо в процесі розв`язку диференційних рівнянь методом прогнозу і корекції змінюється крок, то звичайно тимчасово доводиться переходити до одноточкового методу.

4.2. Одноточкові методи та методи прогнозу і корекції забезпечують приблизно однакову точність результатів. Однак другі на відміну від перших дозволяють лише оцінити похибку на кроці. З цієї причини, користуючись одноточковими методами, величину кроку h звичайно обирають трохи менше, ніж це необхідно, тому методи прогнозу і корекції виявляються найбільш ефективними.

4.3. Використовуючи метод Рунге-Кутта четвертого порядку точності, на кожному кроці доводиться обчислювати чотири значення функції, але для збіжності

методу прогнозу і корекції того ж порядку точності часто достатньо двох значень функції. Тому методи прогнозу і корекції вимагають майже вдвічі менше машинного часу, ніж методи Рунге-Кутта порівнюваної точності.

8.3. Одноточкові методи розв'язування задачі Коші на ПК

Розв'язати диференційне рівняння $y' = f(x, y)$ чисельним методом – це для заданої послідовності аргументів $x_0, x_1, ..., x_n$ і y_0 знайти такі значення $y_0, y_1, ..., y_n$, що $y_i = F(x_i)\,(i = 1,2,...,n)$ і $F(x_0) = y_0$. Таким чином, чисельні методи дозволяють замість отримання функції $y = F(x)$ одержати таблицю значень цієї функції для заданої послідовності аргументів. Величина $h = x_k - x_{k-1}$ називається *кроком інтегрування*. Графічно чисельний розв'язок уявляє собою послідовність коротких прямолінійних відрізків, якими апроксимується аналітичний розв'язок $y = F(x)$ рівняння (кусково-лінійна апроксимація).

Розглянемо алгоритми найбільш відомих чисельних методів.

8.3.1. Метод Ейлера

Цей метод є порівняно грубим і застосовується в основному для орієнтованих розрахунків. Однак ідеї, покладені в основу методу Ейлера, є базовими для інших методів. Нехай дано диференційне рівняння першого порядку $y' = f(x, y)$ з початковими умовами $x = x_0, y(x_0) = y_0$. Потрібно знайти розв'язок рівняння на відрізку $[x_0, x_n]$. Розіб'ємо відрізок $[x_0, x_n]$ на n рівних частин і одержимо послідовність $x_0, x_1, ..., x_n$, де $x_i = x_0 + ih\,(i = 0,1,2,...,n)$, а $h = (x_n - x_0)/n$ – крок інтегрування. Виберемо k-й відрізок $[x_k, x_{k+1}]$ і проінтегруємо рівняння $y' = f(x, y)$:

$$\int_{x_k}^{x_{k+1}} f(x,y)dx = \int_{x_k}^{x_{k+1}} y'dx = y(x)\Big|_{x_k}^{x_{k+1}} = y(x_{k+1}) - y(x_k) = y_{k+1} - y_k, \text{ або}$$

$$y_{k+1} = y_k + \int_{x_k}^{x_{k+1}} f(x,y)dx \qquad \text{(КР.55)}$$

Якщо в останньому інтегралі підінтегральну функцію на відрізку $[x_k, x_{k+1}]$ прийняти постійною і рівною початковому значенню в точці $x = x_k$, то одержимо:

$$\int_{x_k}^{x_{k+1}} f(x_k, y_k)dx = f(x_k, y_k)x\Big|_{x_k}^{x_{k+1}} = f(x_k, y_k)(x_{k+1} - x_k) = y'_k \cdot h.$$

Тоді формула (КР.55) прийме вигляд $y_{k+1} = y_k + y'_k \cdot h$. Позначивши $y_{k+1} - y_k = \Delta y_k$, тобто $y'_k \cdot h = \Delta y_k$, отримаємо $y_{k+1} = y_k + \Delta y_k$.

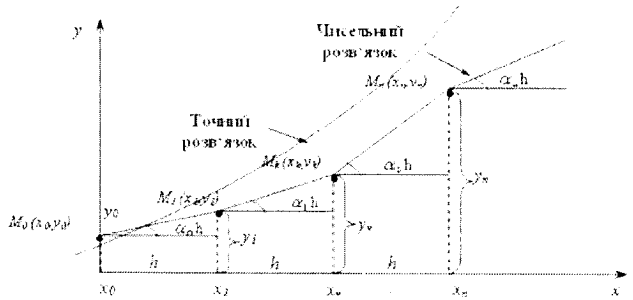


Рис. КР.33. Геометричне розв'язування методом Ейлера.

Продовжуючи цей процес, щораз приймають, що на відрізку $[x_k, x_{k+1}]$ інтегральна крива $y = F(x)$ приблизно замінюється прямолінійним відрізком, що виходить із точки $M_k(x_k, y_k)$ з кутовим коефіцієнтом $f(x_k, y_k)$. Тому за наближення шуканої інтегральної кривої отримуємо ламану лінію з вершинами в точках $M_0(x_0, y_k), M_1(x_1, y_1), \ldots, M_n(x_n, y_n)$ (рис. КР.33). Схема алгоритму метода Ейлера представлена на рисунку КР.34а.

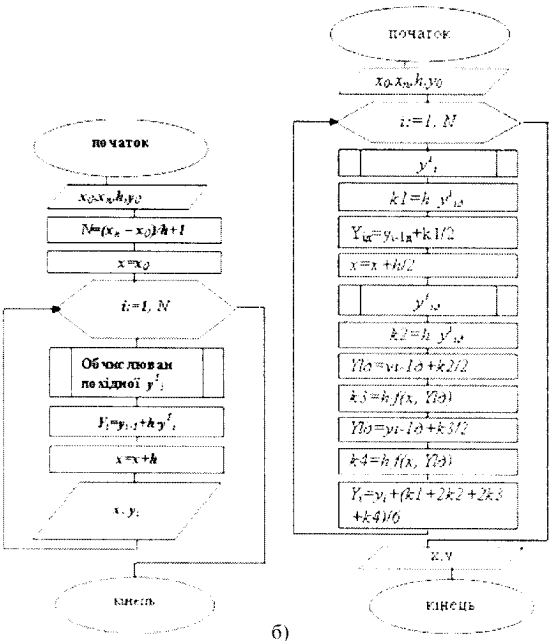


Рис. КР.34. Схема алгоритму метода: а) Ейлера та б) Рунге-Кутта.

Метод Ейлера може бути застосований до розв'язку систем диференційних рівнянь вищих порядків. Однак в останньому випадку диференційні рівняння повинні бути зведені до системи диференційних рівнянь першого порядку.

Нехай задана система двох рівнянь першого порядку $y' = f_1(x, y, z), \quad z' = f_2(x, y, z)$ з початковими умовами $y(x_0) = y_0\ z(x_0) = z_0$.

Наближені значення $y(x_i) \approx y_i\ z(x_i) \approx z_i$ знаходяться за формулами

$$y_{i+1} = y_i + \Delta y_i, \quad z_{i+1} = z_i + \Delta z_i, \tag{КР.56}$$

де $\Delta y_i = hf_1(x_i, y_i, z_i), \quad (i = 0,1,...), \quad \Delta z_i = hf_2(x_i, y_i, z_i),$

8.3.2. Метод Рунге-Кутта

Метод Рунге-Кутта є одним з методів підвищеної точності, але має багато спільного з методом Ейлера. Нехай на відрізку $[a,b]$ потрібно знайти чисельний розв`язок рівняння

$$y' = f(x, y) \tag{КР.57}$$

з початковою умовою

$$y(x_0) = y_0, \tag{КР.58}$$

$x_0, x_1, ..., x_n$, де $x_i = x_0 + ih\ (i = 0,1,2,...,n)$, а $h = (x_n - x_0)/n$ – крок інтегрування.

Розіб'ємо відрізок $[a,b]$ на n рівних частин точками $x_i = x_0 + ih\ (i = 0,1,2,....n)$, де $h = (b-a)/n$ – крок інтегрування. В методі Рунге-Кутта, так як й у методі Ейлера, послідовні значення y_i шуканої функції y визначаються за формулою $y_{i+1} = y_i + \Delta y_i$. Якщо розкласти функцію в ряд Тейлора й обмежитися членами до h_4 включно, то приріст функції Δy можна представити у вигляді

$$\Delta y = y(x+h) - y(x) = hy'(x) + \frac{h^2}{2}y''(x) + \frac{h^3}{6}y'''(x) + \frac{h^4}{24}y^{IV}(x), \tag{КР.59}$$

де похідні $y''(x), y'''(x), y^{IV}(x)$ визначаються послідовним диференціюванням.

Замість безпосередніх обчислень за формулою (КР.59) у методі Рунге-Кутта визначаються чотири числа:

$$k_1 = hf(x, y), \quad k_2 = hf(x + \frac{h}{2}, y + \frac{k_1}{2}),$$
$$k_3 = hf(x + \frac{h}{2}, y + \frac{k_2}{2}), k_4 = hf(x + h, y + k_3). \tag{КР.60}$$

Можна довести, що якщо до чисел k_1, k_2, k_3, k_4 додати відповідно вагу 1/6; 1/3; 1/3; 1/6, то середньозважене цих чисел, тобто $\frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4$ з точністю до четвертих ступенів дорівнює значенню y, яке обчислене за формулою (КР.59):

$\Delta y = \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$. Таким чином, для кожної пари поточних значень x_i та y_i за формулою (КР.60) визначаються значення:

$$k_1^{(i)} = hf(x_i, y_i), \quad k_2^{(i)} = hf(x_i + \frac{h}{2}, y_i + \frac{k_1^{(i)}}{2}),$$
$$k_3^{(i)} = hf(x_i + \frac{h}{2}, y_i + \frac{k_2^{(i)}}{2}), k_4^{(i)} = hf(x_i + h, y_i + k_3^{(i)}) \qquad \text{(КР.61)}$$

і потім $y_{i+1} = y_i + \Delta y_i$.

Числа k_1, k_2, k_3, k_4 мають простий геометричний зміст. Нехай крива M_0CM_1 (рис. КР.35) є розв'язком диференціального рівняння (КР.57) з початковою умовою (КР.58). Точка C цієї кривої лежить на прямій, яка є рівнобіжною до осі Oy і розділяючою відрізок $[x_i, x_{i+1}]$ навпіл, B и G – точки перетину дотичної, проведеної до кривої в точці M_0, з ординатами AC і N_1M_1. Тоді число k_1 з точністю до множника h (де $h = x_{i+1} - x_i$) є кутовим коефіцієнтом дотичної в точці M_0 до інтегральної кривої M_0CM_1, тобто $k_1^{(i)} = hy_i' = hf(x_i, y_i)$. Точка B має координати $x = x_i + \frac{h}{2}; y = y_i + \frac{k_1}{2}$. Отже число k_2 з точністю до множника є кутовим коефіцієнтом дотичної, що проведена до інтегральної кривої в точці B (BF – відрізок цієї дотичної).

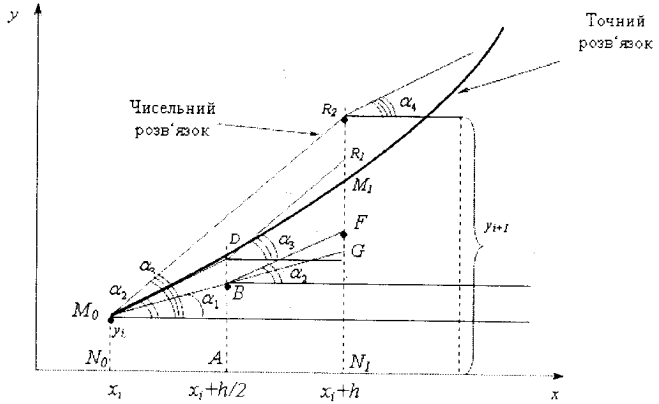


Рис.КР.35. Геометрична інтерпретація метода Рунге-Кутта.

Через точку M_0 проведемо пряму, рівнобіжну відрізку BF. Тоді k_3 з точністю до множника h – кутовий коефіцієнт дотичної, яка проведена до інтегральної кривої в точці D (DR_1 – відрізок цієї дотичної). Нарешті, через точку M_0 проведемо пряму, рівнобіжну до DR_1, що перетне продовження M_1N_1 у точці $R_2(x_i + h, y_i + k_3)$. Тоді k_4 з точністю до множника h є кутовим коефіцієнтом дотичної, проведеної до інтегральної кривої в точці R_2.

Алгоритм обчислення методом Рунге-Кутта:

1. Значення x_0 і y_0 підставляють у праву частину диференціального рівняння (КР.57), визначають $f(x, y)$.

2. Отримане значення $f(x, y)$ множать на крок інтегрування h, обчислюють $k_1 = hf(x, y)$.

3. Змінюють значення x_0: $x_{0+\frac{1}{2}} = x_0 + \frac{h}{2}$.

4. Визначають допоміжне значення y_{0d}: $y_{0d} = y_0 + \frac{k_1}{2}$.

5. Визначення похідної в точці $(x_{0+\frac{1}{2}}, y_{0d})$: $y'_{0d} = f(x_{0+\frac{1}{2}}, y_{0d})$.

6. Визначають значення k_2: $k_2 = hf(x_{0+\frac{1}{2}}, y_{0d})$.

7. Визначають нове допоміжне значення y_{0d}: $y_{0d} = y_0 + \frac{k_2}{2}$.

8. Визначення похідної в точці $(x_{0+\frac{1}{2}}, y_{0d})$: $y'_{0d} = f(x_{0+\frac{1}{2}}, y_{0d})$.

9. Визначають значення k_3: $k_3 = h \cdot y'_{0d} = hf(x_{0+\frac{1}{2}}, y_{0d})$.

10. Визначають нове значення допоміжного y_{1d}: $y_{1d} = y_0 + k_3$.

11. Змінюють значення $x_{0+\frac{1}{2}}$: $x_1 = x_{0+\frac{1}{2}} + \frac{h}{2}$.

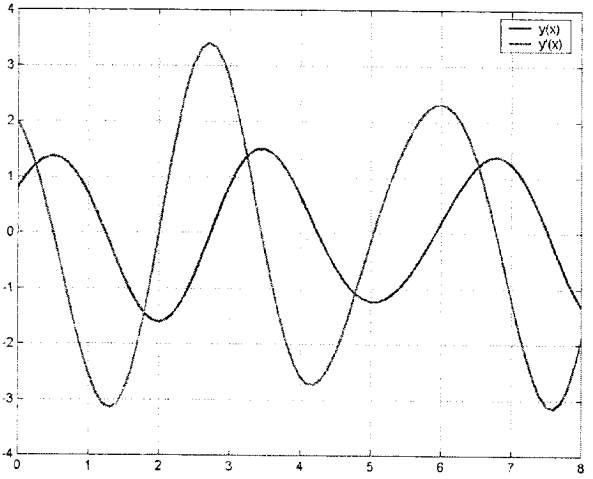
12. Визначають допоміжну похідну в точці (x_1, y_{1d}): $y'_{0d} = f(x_1, y_{1d})$.

13. Визначають значення k_4: $k_4 = h \cdot y'_{1d} = hf(x_0 + h, y_0 + k_3)$.

14. Визначають нове значення y_1 за формулою $y_1 = y_0 + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$.

Для визначення $y_2, y_3, \ldots, y_n$ повторюють ітераційний процес, починаючи з першого кроку. Схема алгоритму методу Рунге-Кутта подана на рис. КР.34б. Потім всі обчислення повторюють починаючи з 1-го кроку, доти, поки не буде пройдений весь відрізок $[a, b]$. Метод Рунге - Кутта має порядок точності h^4 на всьому відрізку $[a, b]$. Оцінка точності цього методу дуже складна. Грубу оцінку точності можна одержати за допомогою "подвійного обчислення" за формулою $|y_i^* - y(x_i)| \approx \frac{y^* - y_i}{15}$, де $y(x_i)$ – значення точного розв'язку рівняння (КР.57) у точці x_i, а y_i^* та y_i – наближені значення, отримані з кроком $h/2$ і h. Якщо ε – задана точність розв'язку, то число n (число кроків) для визначення кроку інтегрування $h = (b - a)/2$ вибирається таким чином, щоб $h^4 < \varepsilon$. Однак, крок обчислення можна змінювати при переході від однієї точки до іншої. Для оцінки правильності вибору кроку h використовують

рівність $q = \left| \frac{k_2^{(i)} - k_3^{(i)}}{k_1^{(i)} - k_2^{(i)}} \right|$, де q рівне декільком сотим, у протилежному випадку – крок h зменшують.



9. Градієнтні методи пошуку екстремуму

9.1. Лінії рівня

Функціональну залежність від декількох змінних, кількість яких більше двох, важко зобразити графічно. Функцію двох змінних на площині можна представити лініями рівня. Лінією рівня функції $y = f(\vec{u})$ називають геометричне місце точок, у яких функція $f(\vec{u})$ приймає одне й теж значення. Зображення функції за допомогою ліній рівня дає можливість наочно представити задачу оптимізації, а також траєкторії пошуку екстремуму за допомогою різних алгоритмів з метою порівняння їх ефективності. Сімейство ліній рівня одержуємо за рівнянням $f(\vec{u}) = C$, де C – постійна величина для кожної лінії рівня.

Приклад КР.14. Нехай цільова функція задачі оптимізації має вигляд:

$$F(\vec{u}) = -1.1u_1^2 - 1.5u_2^2 + 2u_1u_2 + u_1 + 5.$$

Задано обмеження: $u_1 + 1.5u_2 - 1 = 0$. Виконаємо графічну інтерпретацію цієї задачі, в координатній площині $(u_1; u_2)$ шляхом побудови ліній постійного рівня. Сімейство ліній постійного рівня функції $F(\vec{u})$ дане на рис. КР.36.

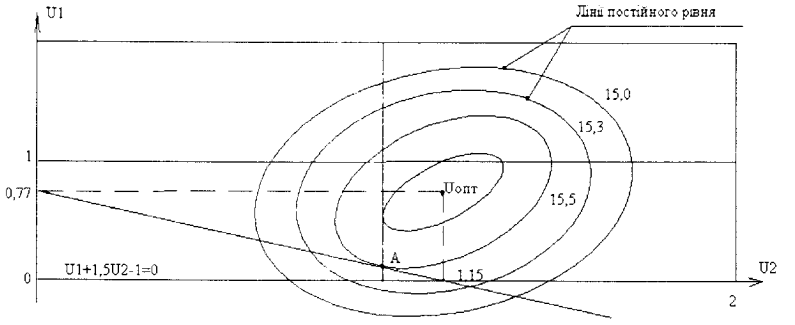


Рис.КР.36. Сімейство ліній постійного рівня функції $F(\vec{u})$.

У цьому прикладі точка безумовного екстремуму має координати: $\vec{u}_{OPT} = (1.15; 0.77)$. Умовний екстремум знаходиться в точці А.

9.2. Найпростіший градієнтний метод оптимізації функцій

Найбільш відомим ітеративним або алгоритмічним методом оптимізації є *метод градієнта*. Зміст його зводиться до організації руху системи в напрямку до екстремуму кроками, причому напрямок руху збігається із градієнтом, якщо шукається максимум, і протилежний до градієнта ("збігається з антиградієнтом"), якщо виконується пошук мінімуму.

9.2.1. Властивості градієнтів

Вектор градієнта скалярної функції $f(\vec{u})$ визначається як вектор-стовпець перших частинних похідних функції по незалежних змінних:

$$grad f(\vec{u}^{(k)}) = \nabla f(\vec{u}^{(k)}) = \begin{bmatrix} \dfrac{\partial f(\vec{u}^{(k)})}{\partial u_1} \\ ---- \\ \dfrac{\partial f(\vec{u}^{(k)})}{\partial u_n} \end{bmatrix}.$$

Частинні похідні функцій обчислюються в точці $\vec{u} = \vec{u}^{(k)}$.

Вводиться поняття одиничного вектора градієнта $\vec{S}$: $\vec{S}^{(k)} = \dfrac{\nabla f(\vec{u}^{(k)})}{\left\| \nabla f(\vec{u}^{(k)}) \right\|}$, де $\| \|$ – норма вектора градієнта. Компоненти одиничного вектора визначаються співвідношенням:

$$S_i^{(k)} = \frac{\dfrac{\partial f(\vec{u}^{(k)})}{\partial u_1}}{\sqrt{\sum_{i=1}^{n}\left(\dfrac{\partial f(\vec{u}^{(k)})}{\partial u_i}\right)^2}}.$$

Градієнт $\nabla f(\vec{u})$ у кожній точці області визначення повністю описує поведінку функції. Він спрямований по нормалі до поверхні рівня, проведеної через точку, що розглядається, а по абсолютній величині дорівнює похідній функції по напрямку нормалі. Через те, що найшвидша зміна функції відбувається при переміщенні по нормалі до поверхні рівня, напрямок градієнта збігається з напрямком найшвидшого зростання функції. Природно, що напрямок антиградієнта збігається з напрямком найшвидшого зменшення значення функції.

Приклад КР.15. Ця скалярна функція двох незалежних змінних, лінії постійного рівня якої зображені на рис. КР.37: $I(\vec{u}) = (u_1 - 1)^2 + 2(u_2 - 2)^2 + 2$. Лінії рівня являють собою еліпси з координатами центра M[1,2]. Розглянемо точку A[-3,2]. Визначимо градієнт в цій точці:

$$grad f(\vec{u}^A) = \begin{bmatrix} \dfrac{\partial f(\vec{u}^A)}{\partial u_1} \\ \dfrac{\partial f(\vec{u}^A)}{\partial u_2} \end{bmatrix} = \begin{bmatrix} 2(u_1 - 1) \\ 4(u_2 - 2) \end{bmatrix}_{\vec{u}=\vec{u}^A} = \begin{bmatrix} -8 \\ 0 \end{bmatrix} = [-8, 0]^T.$$

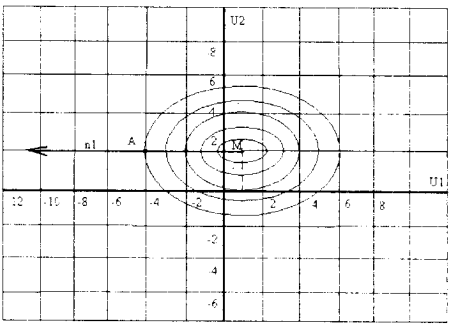


Рис. КР.37.

У просторі n незалежних змінних функції можна провести (n-1) взаємно перпендикулярних дотичних до гіперповерхні (поверхні) рівня. Вісь, перпендикулярна до дотичних, буде нормаллю до поверхні рівня. Для двомірної функції в точці A можна провести одну дотичну. Це – пряма $u_1 + 1.5u_2 - 1 = 0$. Градієнт буде збігатися з напрямком нормалі n_1. Його проекції рівні -8 і 0.

9.2.2. Алгоритм градієнтного методу

Сутність градієнтного методу оптимізації полягає в тому, що задаються довільно або виходячи з наявної апріорної інформації про положення точки екстремуму, початкові значення вектора незалежних змінних $\vec{u}^{(0)}$. Потім виконується заміна $\vec{u}^{(0)}$ на $\Delta\vec{u}^{(0)}$, тобто роблять крок $\Delta\vec{u}^{(0)}$ з метою наблизитися до точки екстремуму u_{opt}. Потім роблять новий крок $\Delta\vec{u}^{(1)}$ і т.д. Таким чином, на кожній ітерації обчислюється значення вектора для наступної ітерації:

$$\vec{u}^{(k+1)} = \vec{u}^{(k)} + \Delta\vec{u}^{(k)}.$$

Оскільки напрямок вектора градієнта вказує напрямок найшвидшого збільшення функції, то кроки Δu виконують у напрямку градієнта при пошуку максимуму й антиградієнта при пошуку мінімуму. Надалі, для визначеності, будемо розглядати задачу на мінімум. Тоді $\Delta\vec{u}^{(k)} = -\lambda\vec{S}^{(k)}$, де λ – множник, що визначає величину кроку $\Delta\vec{u}^{(k)}$; $\vec{S}^{(k)}$ – одиничний вектор градієнта; k – номер ітерації. Знак "–" вказує на напрямок антиградієнта. У такий спосіб: $\vec{u}^{(k+1)} = \vec{u}^{(k)} - \lambda\vec{S}^{(k)}$.

Алгоритм градієнтного пошуку часто застосовують у наступному вигляді:

$$\vec{u}^{(k+1)} = \vec{u}^{(k)} - h\nabla f(\vec{u}^{(k)}). \qquad \text{(КР.62)}$$

У цьому випадку величина кроку $h\nabla f(\vec{u}^{(k)})$ змінюється автоматично відповідно до зміни величини градієнта. Величина h називається *параметром кроку* й залишається постійною. Алгоритм має ту перевагу, що при наближенні до точки мінімума довжина кроку автоматично зменшується.

Ітераційна формула (КР.62) може бути записана в наступній формі:

$$\begin{bmatrix} u_1^{(k+1)} \\ --- \\ --- \\ --- \\ u_n^{(k+1)} \end{bmatrix} = \begin{bmatrix} u_1^{(k)} \\ --- \\ --- \\ --- \\ u_n^{(k)} \end{bmatrix} - h \begin{bmatrix} \dfrac{\partial f(u^{(k)})}{\partial u_1} \\ ----- \\ \dfrac{\partial f(u^{(k)})}{\partial u_n} \end{bmatrix},$$

або в скалярному вигляді: $u_i^{(k+1)} = u_i^{(k)} - h\dfrac{\partial f(\vec{u}^{(k)})}{\partial u_i} = u_i^{(k)} - h \cdot grad_{u_i} f(\vec{u}^{(k)})$.

9.2.3. Вплив величини кроку на градієнтний пошук

Питання вибору величини кроку є досить важливим і в підсумку визначає працездатність і швидкість збіжності алгоритму. Якщо розмір кроку вибраний занадто малим, то рух до оптимуму буде довгим через необхідність обчислення частинних похідних у багатьох точках. При великому кроці в районі оптимуму можуть виникнути незатухаючі коливання незалежних змінних і знижується точність знаходження екстремуму. При дуже великому кроці можливі розбіжні коливання. На рис. КР.38

зображені лінії постійного рівня функції $f(u_1,u_2)$. Процес пошуку при великому h зображений послідовністю точок A_0, A_1, A_2, A_3.

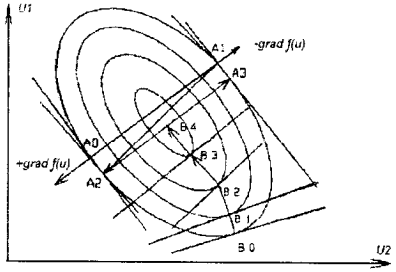


Рис. КР.38.

Як видно з рис. КР.38 при цьому можуть виникати незатухаючі коливання незалежних змінних. При меншому значенні параметра кроку h процес пошуку зображений послідовністю точок B_0, B_1, B_2, B_3, B_4.

9.2.4. Критерій закінчення пошуку

Розглянемо одномірну функцію $f(u)$ (рис. КР.39). Прямі $P_1 \div P_4$ – дотичні до функції $f(u)$. В області екстремуму $\dfrac{df}{du} = 0$, що відповідає дотичній P_4. У такий спосіб при $\left|\dfrac{df}{du}\right| \le \delta$, де δ – мале додатнє, наперед задане число, процес пошуку може бути зупинений. Для багатопараметричної функції умовою зупинки є $\sum_{j=1}^{n}\left(\dfrac{dF}{du_j}\right)^2 \le \delta$. Умова перевіряється на кожному кроці пошуку й виконується, якщо екстремум лежить у допустимій області. Інший варіант визначення моменту закінчення пошуку полягає в наступному. Після кожної серії із заданою кількістю кроків s_1 запам'ятовується значення цільової функції (ЦФ).

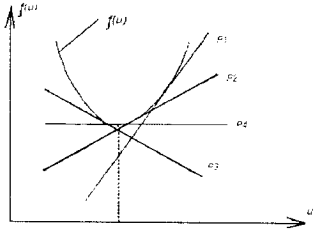


Рис. КР.39.

Якщо наступна серія дає менше значення ЦФ, то пошук триває, у іншому випадку – зупинка. Рекомендується також наступний критерій зупинки:

$f(\vec{u}^{(k+1)}) \geq f(\vec{u}^{(k)})$; $\left|\vec{u}^{(k+1)} - \vec{u}^{(k)}\right| < \varepsilon_i$, i=1.2...n, де ε_i – малі, наперед задані величини, обрані з фізичних міркувань.

9.2.5. Недоліки градієнтного методу

1. Можна виявити тільки локальний екстремум.
2. Невисока швидкість досягнення екстремуму.
3. Мала швидкість збіжності поблизу екстремальної точки.
4. Залежність методу від масштабу змінних. Якщо гіперпростір витягнутий так, що утвориться яр, то процедура градієнтного методу сходиться занадто повільно.
5. Відсутні методи обчислення оптимального значення параметра кроку *h*.

9.2.6. Обчислення похідних

Якщо аналітичний вигляд $f(\vec{u})$ відомий, то обчислення похідних, які використовуються в алгоритмі градієнтного методу, не викликає труднощів. Часто $f(\vec{u})$ у явному аналітичному вигляді записати не можна, або ж вигляд $\frac{df}{du}$ занадто складний. У таких випадках використовують різницеву апроксимацію похідних

$$\frac{df}{du_i} \cong \frac{\Delta f}{\Delta u_i} = \frac{f(u_1, u_2, ..., u_i + \delta, ..u_n) - f(u_1, ..., u_n)}{\delta},$$

де δ – величина приросту, яка однакова для всіх змінних.

9.2.7. Модифікації алгоритмів градієнтного методу

Вибір величини параметра кроку в градієнтному алгоритмі досить важливий. Відомі модифікації градієнтного алгоритму, у яких параметр кроку *h* змінюється автоматично в процесі пошуку. У випадку погано організованої функції масштабують змінні. Нижче наведені деякі алгоритми з адаптивним параметром кроку.

1) $u_i^{(k+1)} = u_i^{(k)} - h^{(k)} sign \frac{\partial f(\vec{u}^{(k)})}{\partial u_i}$; $h^{(k+1)} = \begin{cases} h^{(k)}, f(\vec{u}^{(k)}) \leq f(\vec{u}^{(k-1)}) \\ -\frac{h^{(k)}}{2}, f(\vec{u}^{(k)}) > f(\vec{u}^{(k-1)}) \end{cases}$;

2) $u_i^{(k+1)} = u_i^{(k)} - h^{(k)} \frac{\partial f(\vec{u}^{(k)})}{\partial u_i}$; $h^{(k+1)} = \begin{cases} 2h^{(k)}, a^{(k)} < a_{\min} \\ h^{(k)}, a_{\min} \leq a^{(k)} \leq a_{\max} \\ \frac{h^{(k)}}{3}, a^{(k)} > a_{\max} \end{cases}$;

$$\cos a^{(k)} = \frac{\sum_i \frac{\partial f(\vec{u}^{(k)})}{\partial u_i} \cdot \frac{\partial \vec{u}^{(k-1)}}{\partial u_i}}{\sqrt{\sum_i \left(\frac{\partial f(\vec{u}^{(k)})}{\partial u_i}\right)^2 \cdot \sum_i \left(\frac{\partial f(\vec{u}^{(k-1)})}{\partial u_i}\right)^2}},$$

де a^k – кут повороту градієнта на *k*-тому кроці пошуку, причому

$$\cos a_{\min} \approx 0.2; \cos a_{\max} \approx 0.8.$$

9.3. Метод найшвидшого спуску

Метод найшвидшого спуску (МНС) називають також *методом Коші*. У МНС на кожній ітерації параметр кроку вибирають із умови мінімуму функції, що оптимізується по напрямку градієнта. Алгоритм МНС має вигляд:

$$u_i^{(k+1)} = u_i^{(k)} - h^{(k)} \nabla f(\vec{u}^{(k)}),\ h^{(k)} = \arg\min_h f(\vec{u}_i^{(k)} - h\nabla f(\vec{u}^{(k)})).$$

Одна з головних переваг методу пов'язана з його стійкістю, а також в тому, що він забезпечує виконання нерівності $f(\vec{u}^{(k+1)}) \le f(\vec{u}^{(k)})$ в процесі пошуку. Метод Коші дозволяє, як правило, істотно зменшити значення цільової функції уже з початкової точки, яка розташована на значній відстані від екстремуму. Тому МНС часто використовується як початкова процедура при реалізації градієнтних методів. Основним недоліком його є низька швидкість збіжності поблизу екстремуму. Тому метод рідко використовується як серйозна самостійна оптимізуюча процедура.

9.4. Метод Ньютона

У МНС використовуються тільки перші похідні. Природно, що шлях подальшого вдосконалення оптимізуючої процедури – використання других похідних. Алгоритм методу Ньютона, що враховує другі похідні функції, має вигляд: $\vec{u}_i^{(k+1)} = \vec{u}_i^{(k)} - \nabla^2 f(\vec{u}^{(k)})^{-1} \nabla f(\vec{u}^{(k)})$. Він може бути одержаний на основі квадратичної апроксимації критерію оптимізації за умови досягнення нуля градієнта апроксимуючої функції. У зв'язку із цим задача мінімізації квадратичної функції вирішується за одну ітерацію.

Приклад КР.16. Розглянемо задачу мінімізації функції однієї змінної:

$$Y(u) = U^2 + 2U + 2 \to \min;$$

Початкова точка $U^{(0)} = 1$. Алгоритм Ньютона для цього завдання має вигляд: $U^{(k)-1} = U^k - \dfrac{1}{\dfrac{\partial^2 y(u^k)}{\partial u^2}} \cdot \dfrac{dy(u^{(k)})}{du}$; тоді $U^{(1)} = U^{(0)} - \dfrac{1}{2} \cdot (2U^{(0)} + 2) = 1 - \dfrac{1}{2} \cdot 4 = -1.$

Координата точки екстремуму знайдена за допомогою однієї ітерації.

Приклад КР.17. Розглянемо функцію $f(\vec{u}) = 8u_1^2 + 4u_1u_2 + 5u_2^2$. Обчислимо градієнт $\nabla f(\vec{u})$: $\nabla f(\vec{u}) = [16u_1 + 4u_2; 10u_2 + 4u_1]^T$. Визначимо матрицю других похідних $\nabla^2 f(\vec{u}) = H = \begin{bmatrix} 16 & 4 \\ 4 & 10 \end{bmatrix}$. Як початкову точку візьмемо $\vec{u}^{(0)} = [10,10]^T$. Відповідно до методу Ньютона маємо $\vec{u}^{(1)} = \vec{u}^{(0)} - \nabla^2 f(u^{(0)})^{-1} \cdot \nabla f(\vec{u}^{(0)})$. Обчислимо градієнт у точці $\vec{x}^{(0)}$: $\nabla f(\vec{u}) = [16 \cdot 10 + 4 \cdot 10; 10 \cdot 10 + 4 \cdot 10]^T = [200,140]^T$. Обчислимо матрицю $\nabla^2 f(\vec{u}^{(0)})^{-1} = H^{-1}$: $H^{-1} = \dfrac{1}{144} \cdot \begin{bmatrix} 10 & -4 \\ -4 & 16 \end{bmatrix}$. Згідно алгоритму Ньютона

одержимо $\vec{u}^{(1)} = [10,10]^T - \frac{1}{144}\cdot\begin{bmatrix} 10 & -4 \\ -4 & 16 \end{bmatrix}\cdot[200;140]^T = [0,0]^T$. Таким чином, задача мінімізації вирішується за допомогою однієї ітерації при будь-якій початковій точці.

9.5. Модифікований метод Ньютона

При оптимізації неквадратичних функцій метод Ньютона не відрізняється високою надійністю. Якщо початкова точка перебуває на значній відстані від екстремуму, то крок за методом Ньютона може бути надмірно великим, що може привести до відсутності збіжності. Модифікація методу дозволяє істотно підвищити його ефективність:

$$\vec{u}_i^{(k+1)} = \vec{u}_i^{(k)} - h^{(k)} H^{-1}\nabla f(\vec{u}^{(k)}),\ h^{(k)} = \arg\min_h f(u^{(k)} - h\nabla f(\vec{u}^{(k)})).$$

Методи Ньютона використовуються, якщо обчислення точних значень перших і других похідних не пов'язане з істотними труднощами.

9.6. Методи змінної метрики

Ідея методу полягає у використанні інформації про градієнт критерію оптимальності для наближеного обчислення оберненної матриці других похідних. Пошук ведеться за формулою: $\vec{u}^{(k+1)} = \vec{u}^{(k)} - \lambda^{(k)}\eta^{(k)}\nabla f(\vec{u}^{(k)})$, де $\eta^{(k)}$ – приблизно обернена матриця Гессе; k – параметр кроку.

Головна перевага методу змінної метрики над методом Ньютона: відмова від обчислення матриці Гессе на кожній ітерації. Методи змінної метрики відрізняються, в основному, способом обчислення матриці $\eta^{(k)}$. У методі Девідона-Флетчера-Пауелла (ДФП) матриця $\eta^{(k)}$ обчислюється за формулою:

$$\eta^{(k)} = \eta^{(k-1)} + \frac{\Delta\vec{u}^{(k)}(\Delta\vec{u}^{(k)})^T}{(\Delta\vec{u}^{(k)})^T\Delta\vec{g}^{(k)}} - \frac{\eta^{(k)}\Delta\vec{g}^{(k)}(\Delta\vec{g}^{(k)})^T(\eta^{(k)})^T}{(\Delta\vec{g}^{(k)})^T\eta^{(k)}\Delta\vec{g}^{(k)}};$$

де $\Delta\vec{u}^{(k)} = \vec{u}^{(k-1)} - \vec{u}^{(k)}; \Delta\vec{g}^{(k)} = \nabla f(\vec{u}^{(k+1)}) - \nabla f(\vec{u}^{(k)})$. Матриця $\eta^{(k)}$ зветься метрикою. Вона змінюється на кожній ітерації. Тому методи називаються методами змінної метрики. Таким чином, алгоритм ДФП має вигляд:

$$\vec{u}^{(k+1)} = \vec{u}^{(k)} - \lambda^{*(k)}\eta(\vec{u}^{(k)})\cdot\nabla f(\vec{u}^{(k)});\ \lambda^{*(k)} = \arg\min_\lambda f(\vec{u}^{(k)} - \lambda\cdot\eta(\vec{u}^{(k)})\cdot\nabla f(\vec{u}^{(k)}));$$

$$\Delta\vec{g}^{(k)} = \nabla f(\vec{u}^{(k+1)}) - \nabla f(\vec{u}^{(k)});\ \Delta\vec{u}^{(k)} = \vec{u}^{(k+1)} - \vec{u}^{(k)};\ \eta^{(k+1)} = \eta^{(k)} + A^{(k)} - B^{(k)};$$

$$A^{(k)} = \frac{\Delta\vec{u}^{(k)}\cdot(\Delta\vec{u}^{(k)})^T}{(\Delta\vec{u}^{(k)})^T\cdot\Delta\vec{g}^{(k)}};\ B^{(k)} = \frac{\eta^{(k)}\cdot\Delta\vec{g}^{(k)}\cdot(\Delta\vec{g}^{(k)})^T(\eta^{(k)})^T}{(\Delta\vec{g}^{(k)})^T\eta^{(k)}\Delta\vec{g}^{(k)}}.$$

Критерій зупинки: $\left|\frac{\Delta u_i^{(k)}}{u^{(k)}}\right| < \varepsilon_1$. *або* $\left|\Delta u_i^{(k)}\right| < \varepsilon_2$; $\left|\Delta\vec{g}^{(k)}\right| < \varepsilon_3$;

Приклад КР.18. Обчислення чотирьох кроків пошуку методом змінної метрики. Задана функція $f(\vec{u}) = 4(u_1 - 5)^2 + (u_2 - 6)^2$. Алгоритм змінної метрики має вигляд $\vec{u}^{(k+1)} = \vec{u}^{(k)} - \lambda^{*(k)} \eta^{(k)} \cdot \nabla f(\vec{u}^{(k)})$. Вибираємо початкову точку і початкове значення матриці η: $\vec{u}^{(0)} = [8,9]^T; \eta^{(0)} = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Крок 1. Обчислюємо градієнт в нульовій точці $\nabla f(\vec{u}^{(0)}) = \begin{bmatrix} 8(u_1 - 5) \\ 2(u_2 - 6) \end{bmatrix}^T = [24;6]^T$.

Крок 2. Виконуємо пошук вздовж прямої. Спочатку обчислимо координати точки $\vec{u}^{(1)}$ як функції скалярного множника $\lambda^{(0)}$.

$$\begin{bmatrix} u_1^{(1)} \\ u_2^{(1)} \end{bmatrix}^T = \begin{bmatrix} u_1^{(0)} \\ u_2^{(0)} \end{bmatrix} - \lambda^{(0)} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 24 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 \\ 9 \end{bmatrix} - \begin{bmatrix} 24 & \lambda^{(0)} \\ 6 & \lambda^{(0)} \end{bmatrix} = \begin{bmatrix} 8 - 24\lambda^{(0)} \\ 9 - 6\lambda^{(0)} \end{bmatrix};$$

Підставим в функцію, яка оптимізується, знайдені величини:

$$f(\vec{u}^{(1)}, \lambda^0) = 4(u_1^{(1)} - 5)^2 + (u_2^{(1)} - 6)^2 = [4 - (8 - 24 \cdot \lambda)^0 - 5]^2 + [(9 - 6 \cdot \lambda)^0 - 6]^2 \to \min_{\lambda^0}.$$

Виконуємо оптимізацію по $\lambda^{(0)}$ одержаного виразу. Для знаходження екстремуму можна застосувати будь-який метод одномірної оптимізації. Знайдемо першу похідну від виразу для $f\left(\vec{u}^{(1)}, \lambda^{(0)}\right)$: $\frac{\partial f(\vec{u}^{(1)})}{\partial \lambda^{(0)}} = 0 = 51 - 390 \cdot \lambda^0$. Звідси обчислимо початкове значення оптимального параметру кроку $\lambda^{(0)}$: $\lambda^{*(0)} = 0.1307$. Знайдемо координати точки $\vec{u}^{(1)}$: $\vec{u}^{(1)} = \begin{bmatrix} 8 - 24 \cdot 0.1307 \\ 9 - 6 \cdot 0.1307 \end{bmatrix} = [4.86;8.2]^T$.

Крок 3. Обчислимо градієнт в точці $\vec{u}^{(1)}$, а також значення $g^{(0)}$, $u^{(0)}$

$$\nabla f(\vec{u}^{(1)}) = [-1.108;4.43]^T;$$

$$\Delta \vec{g}^{(0)} = \nabla f(\vec{u}^{(1)}) - \nabla f(\vec{u}^{(0)}) = [-25.1;-1.57]^T;$$

$$\Delta \vec{u}^{(0)} = \vec{u}^{(1)} - \vec{u}^{(0)} = [-3.13;-0.785]^T;$$

Знайдемо матрицю $\eta^{(1)}$: $\eta^{(1)} = \eta^{(0)} + A^{(0)} - B^{(0)} = \begin{bmatrix} 0.127 & -0.03 \\ -0.03 & 1.0 \end{bmatrix}$. Обчислюємо координати точки $\vec{u}^{(2)}$: $\vec{u}^{(2)} = \vec{u}^{(1)} - \lambda^{*(1)} \eta^{(1)} \cdot \nabla f(\vec{u}^{(1)})$.

Крок 4. Пошук уздовж прямої, в результаті якого визначаємо $\lambda^{*(1)}$, тобто повертаємось на крок 2.

9.7. Недоліки і переваги методу ДФП

1. Схильність до нагромадження похибок обчислень при перерахунку матриці $\eta^{(k)}$. Тому матрицю треба під час обчислень періодично поновлювати, тобто штучно робити одиничною.

2. Висока чутливість до точності визначення перших похідних.

При тестових випробуваннях алгоритмів метод ДФП проявив себе як найбільш ефективний за кількістю необхідних кроків пошуку.

10. Диференційні рівняння в частинних похідних

Диференційні рівняння в частинних похідних класифікуються:

1. Залежно від їх математичної природи (наприклад, еліптичні, параболічні тощо);
2. Залежно від фізичного змісту задач, які розв'язуються з їх допомогою (рівняння дифузії, хвильове рівняння тощо).

Розглянемо рівняння 2-го порядку з двома невідомими незалежними змінними:

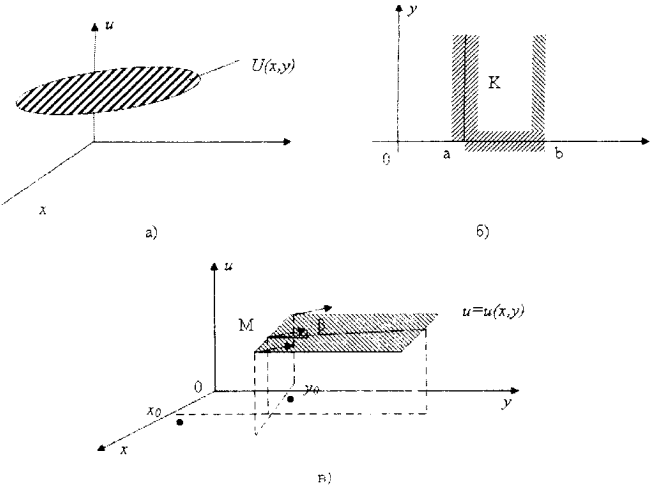
$$F(x, y, u, u_x, u_y, u_{xx}, u_{yy}, u_{xy}) = 0, \qquad \text{(КР.63)}$$

де x, y – незалежні змінні; u – шукана функція; u_x, u_y, u_{xx}, u_{yy} – її 1-і та 2-і частинні похідні за аргументами x та y.

Розв'язком рівняння (КР.63) називається функція $u = u(x, y)$, яка перетворює це рівняння в тотожність. Графік функції розв'язку є поверхнею у просторі OXY (інтегральна поверхня) (рис. КР.40).

Рівняння (КР.63) називається *лінійним*, та може мати вигляд:

$$A\frac{\partial^2 u}{\partial x^2} + B\frac{\partial^2 u}{\partial x \partial y} + C\frac{\partial^2 u}{\partial y^2} + E\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) = 0, \qquad \text{(КР.64)}$$



а) б) в)

Рис. КР.40. Приклади інтегральних поверхонь.

де коефіцієнти A, B, C, E можуть залежати від x,y і в окремих випадках можуть бути константами. Для класифікації ДР в частинних похідних вводиться термін *дискримінант рівняння:* $Д = AC - B^2$. Залежно від знака дискримінанта $Д$, лінійне диференційне рівняння (КР.64) можна віднести до одного з таких типів:

1) якщо $Д > 0$, то рівняння (КР.64) відноситься до еліптичного;
2) якщо $Д = 0$, то рівняння (КР.64) відноситься до параболічного;
3) якщо $Д < 0$, то рівняння (КР.64) відноситься до гіперболічного.

Слід відзначити, що еліптичні рівняння описують встановлені процеси (стаціонарні), а гіперболічні та параболічні рівняння описують еволюційні процеси.

10.1. Різницеві методи розв'язування диференційних рівнянь у частинних похідних

Класичне визначення похідної функції однієї змінної записується у вигляді $\frac{dy}{dx} = \lim_{h\to 0} \frac{y(x+h) - y(x)}{h}$. Природно, що в ПК не можна зробити граничного переходу. З іншого боку, можна додати деяке мале, хоча і ненульове, значення h і спробувати перевірити, що наближення виходить досить точним (проблема точності), і що помилка не зростає в ході процесу обчислень (проблема стійкості), тобто цей метод зводиться до того, що похідну замінюємо різницею.

Очевидно, на ПК неможливо провести граничний перехід. А тому із розкладу функції в ряд Тейлора в околі точки (x_0, y_0) можна наближено записати:

$$u_x(x_0, y_0) = \frac{u(x_0 + h, y_0) - u(x_0, y_0)}{h} \qquad \text{(КР.65)}$$

з помилкою обмеження $\xi_T = \frac{h}{2} u_{xx}(\xi, y_0)$; $x < \xi < x_0 + h$.

Вираз (КР.65) називають *правою різницею*, а вираз

$$u_x(x_0 + h, y_0) = \frac{u(x_0 + h, y_0) - u(x_0, y_0)}{h} \qquad \text{(КР.66)}$$

лівою різницею. Користуючись (КР.65) та (КР.66), можна отримати різницевий вираз для 2-ї похідної

$$u_{xx}(x_0, y_0) = \frac{u(x_0 + h, y_0) - 2u(x_0, y_0) + u(x_0 + h, y_0)}{h^2} \qquad \text{(КР.67)}$$

і відповідно

$$u_{yy}(x_0, y_0) = \frac{u(x_0, y_0 + k) - 2u(x_0, y_0) + u(x_0, y_0 + k)}{k^2}. \qquad \text{(КР.68)}$$

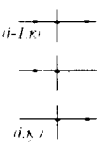
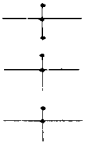
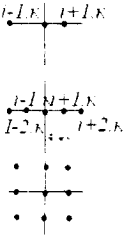
де h, k – величина кроку відповідно за координатами x та y.

10.2. Етапи чисельного розв'язування диференційних рівнянь (ДР) у частинних похідних на ПК

1. В область розв'язування рівняння, яке розглядається, вводять рівномірну сітку $N \times M$ вузлових точок відповідно до характеру задачі та граничних умов.
2. Диференційне рівняння, що розв'язується, записують у найзручнішій системі координат, замінюючи похідні кінцевими різницями (табл. КР.3), приводять його до вигляду різницевого рівняння. Отримане різницеве рівняння використовують далі для опису функціонального зв'язку між сусідніми вузлами сітки. Тобто різницеве рівняння записують для усіх вузлів сітки і отримують внаслідок цього систему $(N-1)(M-1)$ рівнянь з $(N-1)(M-1)$ невідомими.

Таблиця КР.3

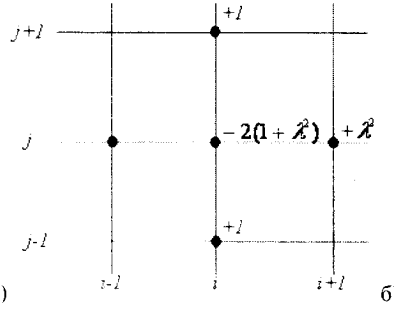
Відношення між похідними та значеннями функції у вузлах сітки

Похідна	Схема	Наближена формула	Порядок похибки
$\frac{\partial u}{\partial x}$		$\frac{\partial u_{i,k}}{\partial x} \approx \frac{u_{i+1,k} - u_{i-1,k}}{2h}$	h^2
		$\frac{\partial u_{i,k}}{\partial x} \approx \frac{u_{i,k} - u_{i-1,k}}{h}$	h
		$\frac{\partial u_{i,k}}{\partial x} \approx \frac{u_{i+1,k} - u_{i,k}}{h}$	h
$\frac{\partial u}{\partial y}$		$\frac{\partial u_{i,k}}{\partial y} \approx \frac{u_{i,k+1} - u_{i,k-1}}{2e}$	e^2
		$\frac{\partial u_{i,k}}{\partial y} \approx \frac{u_{i,k+1} - u_{i,k}}{e}$	e
		$\frac{\partial u_{i,k}}{\partial y} \approx \frac{u_{i,k} - u_{i,k-1}}{e}$	e
$\frac{\partial^2 u}{\partial x^2}$		$\frac{\partial^2 u_{i,k}}{\partial x^2} \approx \frac{u_{i+1,k} - 2u_{i,k} + u_{i-1,k}}{h^2}$	h^2
		$\frac{\partial^2 u_{i,k}}{\partial x^2} \approx \frac{1}{12h^2}(-u_{i+2,k} + 16u_{i+1,k} - 30u_{i,k} + 16u_{i-1,k} - u_{i-2,k})$	h^4
		$\frac{\partial^2 u_{i,k}}{\partial x^2} \approx \frac{1}{h^2}(u_{i,k+1} - 2u_{i,k+1} + u_{i-1,k+1} + u_{i+1,k} - 2u_{i,k} + u_{i-1,k} + u_{i+1,k-1} - 2u_{i,k-1} + u_{i-1,k-1})$	h^4

Похідна	Схема	Наближена формула	Порядок похибки
$\frac{\partial^2 u}{\partial y^2}$		$\frac{\partial^2 u_{i,k}}{\partial y^2} \approx \frac{u_{i,k+1} - 2u_{i,k} + u_{i,k-1}}{e^2}$	e^2
		$\frac{\partial^2 u_{i,k}}{\partial y^2} \approx \frac{1}{12e^2}(-u_{i,k+2} + 16u_{i,k+1} - 30u_{i,k} + 16u_{i,k-1} - u_{i,k-2})$	e^4
		$\frac{\partial^2 u_{i,k}}{\partial y^2} \approx \frac{1}{3e^2}(u_{i+1,k+1} - 2u_{i+1,k} + u_{i+1,k-1} + u_{i,k+1} - 2u_{i,k} + u_{i,k-1} + u_{i-1,k+1} - 2u_{i-1,k} + u_{i-1,k-1})$	e^4

3. Систему $(N\text{-}1)(M\text{-}1)$ рівнянь з $(N\text{-}1)(M\text{-}1)$ невідомими розв'язують одним із відомих швидких і ефективних чисельних методів, вибір яких визначається деякими властивостями системи рівнянь (наприклад, система лінійних або нелінійних рівнянь та ін.).

Рис. КР.41а називається *різницевою схемою.* Різницева схема геометрично ілюструє різницеву апроксимацію диференціального рівняння. Відомо, що при $h \to 0$ та $k \to 0$ різницеве рівняння наближується до диференційного рівняння. Однак, цікавить зовсім інше питання: чи наближається при $h \to 0$ та $k \to 0$ розв'язок різницевого рівняння до розв'язку диференційного рівняння? Для еліптичних рівнянь на це питання можна дати позитивну відповідь. Далі побачимо, що у випадку параболічних і гіперболічних рівнянь доведеться дотримуватись ряду обмежень, що забезпечать такого роду збіжність.



а) б)

Рис. КР.41. а) Різницева схема та б) побудова сітки в прямокутній області.

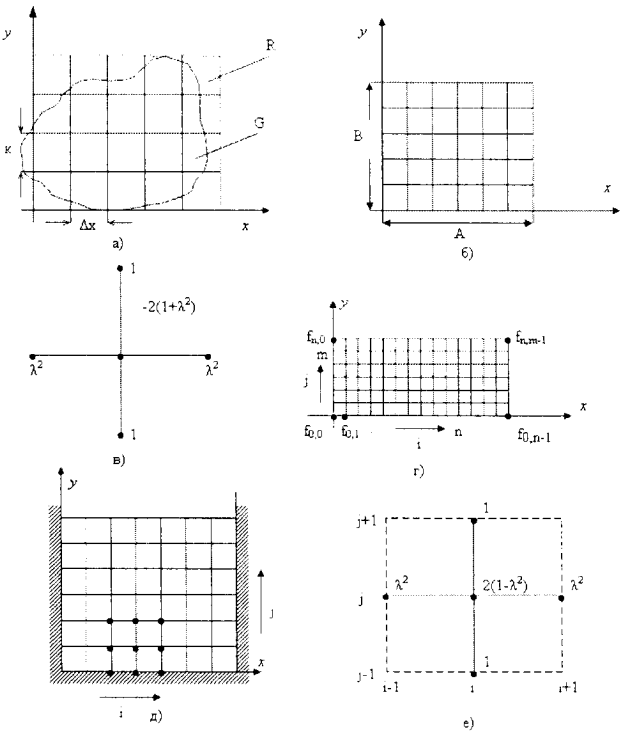


Рис. КР.42. Графічне подання переходу до кінцево-різницевих формул

До еліптичних відносять рівняння виду:

$$\text{рівняння Пуассона: } \frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=-2 \qquad \text{(КР.69)}$$

$$\text{рівняння Лапласа: } \frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0 \qquad \text{(КР.70)}$$

Розглянемо класичну задачу Діріхле $\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0$ у деякій області R і $\psi = f(x,y)$ на границі цієї області, якою є крива G. Звичайно, рівняння (КР.70) є рівнянням Лапласа, окремий випадок рівняння Пуассона. Припустимо що область R має вигляд прямокутника шириною A і висотою B (рис.КР.41б, КР.42б). Розділимо спочатку ширину прямокутника A на n інтервалів, кожен розміром $h = A/n$; а висоту B – на m частин розміром $k = B/m$. Усередині області виходять при цьому $(n-1)(m-1)$ перерізів сітки. Визначимо різницеве співвідношення. Нумерація точок

перетину сітки починається в горизонтальному напрямку зліва – направо від нуля, крайня права точка буде при цьому n-ою. Аналогічно у вертикальному напрямку нумеруємо точки знизу-вверх від нуля до m. Вузол з індексами (i, j) буде i-м ліворуч і j-м знизу. для кожної внутрішньої точки і розв'яжемо отриману систему рівнянь.

Нехай початок координат збігається з точкою, координати якої (0,0) позначимо $u(ih, jk) = u_{i,j}$, аналогічно записується $f(ih, jk) = f_{i,j}$. Використовуючи таку систему позначень граничну умову можна записати у вигляді:

$$\begin{aligned} &u_{i,0} = f_{i,0};\ \ i = 0,\ 1,\ 2,\ \ldots,\ n, \quad u_{i,m} = f_{i,m};\ \ i = 0,\ 1,\ 2,\ \ldots,\ n, \\ &u_{0,j} = f_{0,j};\ \ j = 0,\ 1,\ 2,\ \ldots,\ m, \quad u_{n,j} = f_{n,j};\ \ j = 0,\ 1,\ 2,\ \ldots,\ m. \end{aligned} \qquad \text{(КР.71)}$$

Якщо позначити $\lambda = k/h$, то диференціальне рівняння зведеться до різницевого рівняння вигляду

$$\lambda^2 u_{i+1,j} + \lambda^2 u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 2(1+\lambda^2)u_{i,j} = 0 \qquad \text{(КР.72)}$$

для i=1, 2, 3, ..., n-1 та j=1, 2, 3, ..., m-1.

При $\lambda = 1$, тобто при однакових величинах інтервалів розбиття в горизонтальному і вертикальному напрямках, це співвідношення означає, що значення $u_{i,j}$ є середнім арифметичним чотирьох сусідніх з ним:

$$u_{ij} = \frac{1}{4}\left(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}\right). \qquad \text{(КР.73)}$$

Важливо те, що одержана система – це система лінійних алгебраїчних рівнянь. Усього є $(m-1)(n-1)$ рівнянь щодо $(m+1)(n+1)$ невідомих. Після того як $2(m+n)$ невідомих будуть виключені за допомогою граничної умови (КР.71), залишається точно $(m-1)(n-1)$ рівнянь щодо $(m-1)(n-1)$ невідомих. Вираз (КР.71) є системою лінійних алгебраїчних $(n$-$1)(m$-$1)$ рівнянь відносно $(n+1)(m+1)$ невідомих. Після того, як $2(n+m)$ невідомих будуть виключені за допомогою граничних умов, залишається $(n$-$1)(m$-$1)$ рівняння відносно $(n$-$1)(m$-$1)$ невідомих.

У (КР.71) і (КР.72) кожна пара значень (i, j) визначає вузол, в якому рівняння розв'язується відносно $u_{i,j}$. Рівняння (КР.72) можна легко запам'ятати, якщо знати різницеву схему п'яти вузлів (рис. КР.41а, КР.2.д,е). Розглянемо деякі рівняння в тому вигляді, у якому з ними будуть проводитися ітерації. Позначаючи верхніми індексами порядковий номер ітерації і припускаючи, $u_{i,j}^{(0)} = 0$ для усіх i, j, одержимо наступний порядок розв'язування системи рівнянь (КР.74).

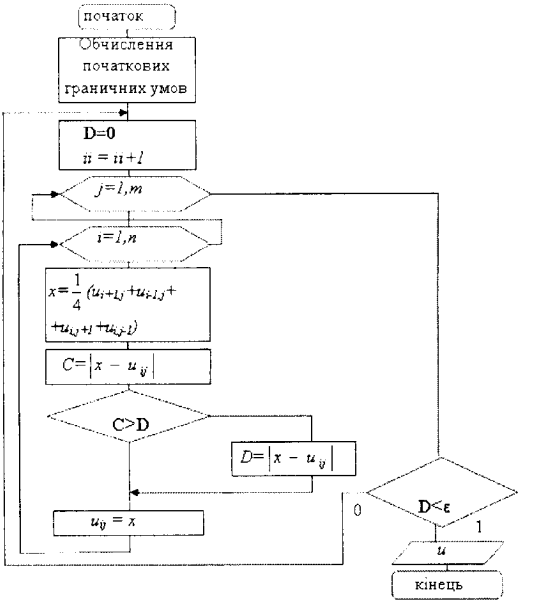


Рис. КР.43. Схема алгоритму розв'язування еліптичного рівняння на ПК

Для цієї системи рівнянь виконується умова збіжження ітераційного методу Гауса-Зейделя. Щоб розв'язати цю систему, розроблено алгоритм (рис. КР.43), який використовує метод Гауса-Зейделя і який застосовують до еліптичних рівнянь різницевих, називають методом Лібмана (або методом послідовних зсувів).

$$
\begin{aligned}
u_{1,1}^{(1)} &= \frac{1}{4}\left[f_{1,0} + f_{0,1} + u_{2,1}^{(0)} + u_{1,2}^{(0)}\right] \\
u_{2,1}^{(1)} &= \frac{1}{4}\left[f_{2,0} + u_{1,1}^{(1)} + u_{3,1}^{(0)} + u_{2,2}^{(0)}\right] \\
u_{3,1}^{(1)} &= \frac{1}{4}\left[f_{3,0} + u_{2,1}^{(1)} + u_{4,1}^{(0)} + u_{3,2}^{(0)}\right] \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
u_{n-1,1}^{(1)} &= \frac{1}{4}\left[f_{n-1,0} + f_{n,1} + u_{n-2,1}^{(1)} + u_{n-1,2}^{(0)}\right] \\
u_{1,2}^{(1)} &= \frac{1}{4}\left[f_{0,2} + u_{1,1}^{(1)} + u_{2,2}^{(0)} + u_{1,3}^{(0)}\right] \\
u_{2,1}^{(1)} &= \frac{1}{4}\left[u_{2,1}^{(1)} + u_{1,2}^{(1)} + u_{3,2}^{(0)} + u_{2,3}^{(0)}\right] \\
&\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}
\qquad \text{(КР.74)}
$$

10.3. Гіперболічні рівняння

Гіперболічними рівняннями є вирази вигляду

$$\left.\begin{aligned} \frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial y^2} \qquad & 0 \le x \le L; \\ & t > 0; \end{aligned}\right\} \tag{КР.75}$$

$$\left.\begin{aligned} u(0, y) = u(L, y) = 0; \qquad & y \ge 0; \\ u(x,0) = f(x); \qquad & 0 \le x \le L; \\ u_y(x,0) = g(x); \qquad & x \le x \le L; \end{aligned}\right\} \tag{КР.76}$$

Систему рівнянь (КР.76) називають початковими умовами. Якщо позначити $u_{i,j} = u(ih, jk)$ і $\lambda = \frac{k}{h}$, то різницеве рівняння запишеться у вигляді

$$u_{i,j+1} = 2(1-\lambda^2)u_{i,j} + \lambda^2(u_{i+1,j} + u_{i-1,j}) - u_{i,j-1} \tag{КР.77}$$

для i = 1, 2, ..., n і для j = 1, 2, ..., m. Умови збіжження та стійкість методу: розв'язок рівняння (КР.77) сходиться, якщо $\lambda < 1$, тобто $k < L$. При $\lambda > 1$ метод стає нестійким. Схема алгоритму подана на рис. КР.44.

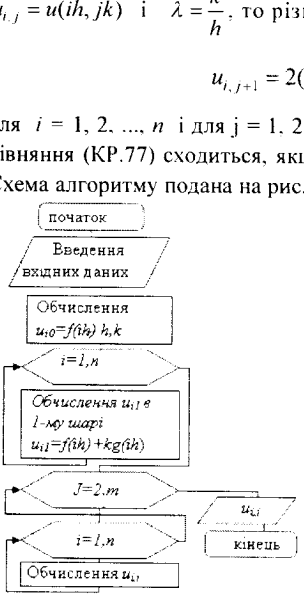


Рис. КР.44. Схема алгоритму розв'язування гіперболічного рівняння на ПК

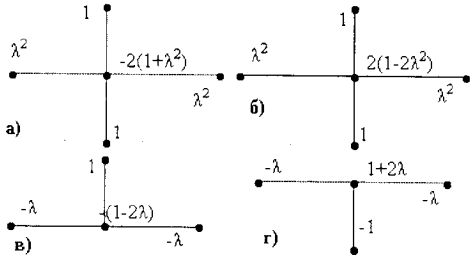


Рис. КР.45. Різницеві схеми для: еліптичного (а), гіперболічного (б) і параболічного (в, г) рівнянь

10.4. Параболічні рівняння

Параболічними рівняннями є вирази вигляду $\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial y}$. Граничними умовами для нього є: $u(0, y) = T_0$; $u(l, y) = T_l$; $u(x,0) = f(x)$. Різницеві схеми еліптичного (а), гіперболічного (б) і параболічного (в, г) рівнянь показано на рис.КР.45.

10.5. Приклади розв'язування диференціальних рівнянь в частинних похідних в середовищі MathCAD

Для розв'язування диференціальних рівнянь в частинних похідних використовується функція: ***relax (a, b, c, d, e, f, u, rjac)*** – повертає квадратну матрицю, в якій: розташування елемента в матриці відповідає його розташуванню в межах квадратної області, а також його значення наближує значення розв'язку в цій точці. Параметри:

1) ***a, b, c, d, e*** – квадратні матриці однакового розміру, що містять коефіцієнти диференціального рівняння;
2) ***f*** – квадратна матриця, що містить початковий розв'язок рівняння, що досліджується;
3) ***u*** – квадратна матриця, що містить граничні значення по межах області і початкових умов для розв'язування в області.
4) ***rjac*** – спектральний радіус Jacobi ітерації. Він керує конвергенцією алгоритму релаксації (може змінюватися від 0 до 1, але його оптимальне значення залежить від деталей вашої проблеми).

Зауваження. Ця функція використовує метод релаксації для знаходження розв'язку. Необхідно використовувати функцію ***relax***, якщо вам відоме значення, яке приймає невідома функція ***u(x, y)*** на всіх чотирьох сторонах квадратної області. Якщо функція ***u(x, y)*** приймає нульові значення на всіх чотирьох сторонах квадратної області, то замість ***relax*** використовується функція ***multigrid***.

multigrid (M, ncycle) повертає квадратну матрицю, в якій: розташування елемента в матриці відповідає його розташуванню в межах квадратної області, а також його значення наближує значення розв'язку в цій точці. *Параметри:*

1) ***M*** – 1+2^n рядок квадратної матриці, елементи якого містять початковий розв'язок рівняння, що досліджується;
2) ***ncycle*** – номер циклів на кожному рівні ітерації (значення 2 буде взагалі давати хороше наближення розв'язку).

Нижче подано фрагменти робочого документу MathCAD з прикладами використання спочатку функції ***relax()***, а потім ***multigrid()***

Приклад КР.19. Розв'язування диференціального рівняння в частинних похідних за допомогою функції ***relax()***

$$R := 32 \qquad M_{3\frac{R}{4},\frac{R}{4}} := 1$$

$$M_{R,R} := 0$$

$$M_{\frac{R}{4},3\frac{R}{4}} := -2 \qquad M_{\frac{R}{8},\frac{R}{2}} := 2$$

$$i := 0..32 \qquad k := 0..32$$

$$a_{i,k} := 1 \quad b := a \quad c := a \quad d := a \quad e := -4a \quad f := -M$$

$$v_{i,k} := 0$$

$$S := relax(a, b, c, d, e, f, v, .95)$$

Результати дослідження можна подати у вигляді таблиці та графіка

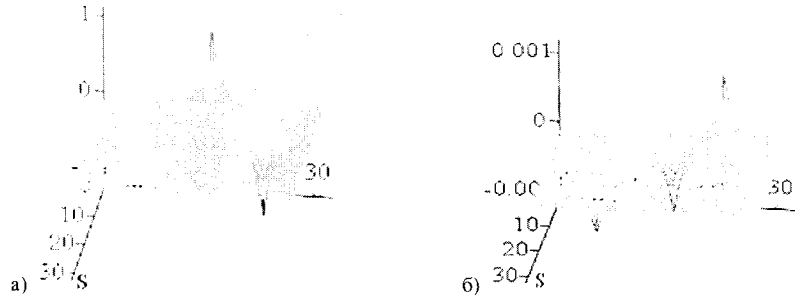


Рис. КР.46. Графіки результатів дослідження функції з прикладу а) КР.19 та б) КР.20

$v_{i,k} := 0$

$S := relax(a, b, c, d, e, f, v, .95)$

$S =$

	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	$1.681 \cdot 10^{-3}$	$3.426 \cdot 10^{-3}$	$5.303 \cdot 10^{-3}$	$7.399 \cdot 10^{-3}$	$9.829 \cdot 10^{-3}$	0.013
2	0	$3.3 \cdot 10^{-3}$	$6.718 \cdot 10^{-3}$	0.01	0.014	0.019	0.025
3	0	$4.801 \cdot 10^{-3}$	$9.76 \cdot 10^{-3}$	0.015	0.021	0.028	0.036
4	0	$6.142 \cdot 10^{-3}$	0.012	0.019	0.027	0.035	0.044
5	0	$7.301 \cdot 10^{-3}$	0.015	0.023	0.031	0.041	0.051
6	0	$8.275 \cdot 10^{-3}$	0.017	0.026	0.035	0.045	0.057
7	0	$9.083 \cdot 10^{-3}$	0.018	0.028	0.038	0.048	0.06
8	0	$9.758 \cdot 10^{-3}$	0.02	0.03	0.04	0.051	0.062
9	0	0.01	0.021	0.031	0.042	0.052	0.063
10	0	0.011	0.022	0.033	0.043	0.054	0.064
11	0	0.012	0.023	0.034	0.045	0.056	0.066
12	0	0.012	0.024	0.036	0.047	0.058	0.068
13	0	0.013	0.026	0.038	0.05	0.061	0.07
14	0	0.014	0.028	0.041	0.054	0.065	0.075
15	0	0.015	0.031	0.045	0.058	0.071	0.081

Приклад КР.20. Розв'язування диференціального рівняння в частинних похідних за допомогою функції ***multigrid()***

$$R := 32 \qquad M_{3\frac{R}{4}, \frac{R}{4}} := 1$$

$$M_{R,R} := 0$$

$$M_{\frac{R}{4}, 3\frac{R}{4}} := -2 \qquad M_{\frac{R}{8}, \frac{R}{2}} := 2$$

$$S := multigrid(M, 2)$$

Результати дослідження можна подати у вигляді таблиці та графіка

S =

	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	-1.607·10⁻⁶	-3.292·10⁻⁶	-5.117·10⁻⁶	-7.173·10⁻⁶	-9.56·10⁻⁶
2	0	-3.164·10⁻⁶	-6.444·10⁻⁶	-1.003·10⁻⁵	-1.401·10⁻⁵	-1.862·10⁻⁵
3	0	-4.604·10⁻⁶	-9.394·10⁻⁶	-1.454·10⁻⁵	-2.027·10⁻⁵	-2.683·10⁻⁵
4	0	-5.885·10⁻⁶	-1.199·10⁻⁵	-1.847·10⁻⁵	-2.568·10⁻⁵	-3.372·10⁻⁵
5	0	-6.95·10⁻⁶	-1.411·10⁻⁵	-2.166·10⁻⁵	-2.987·10⁻⁵	-3.903·10⁻⁵
6	0	-7.834·10⁻⁶	-1.582·10⁻⁵	-2.419·10⁻⁵	-3.311·10⁻⁵	-4.295·10⁻⁵
7	0	-8.562·10⁻⁶	-1.725·10⁻⁵	-2.618·10⁻⁵	-3.564·10⁻⁵	-4.575·10⁻⁵
8	0	-9.194·10⁻⁶	-1.842·10⁻⁵	-2.789·10⁻⁵	-3.753·10⁻⁵	-4.788·10⁻⁵
9	0	-9.789·10⁻⁶	-1.959·10⁻⁵	-2.942·10⁻⁵	-3.945·10⁻⁵	-4.964·10⁻⁵
10	0	-1.038·10⁻⁵	-2.073·10⁻⁵	-3.101·10⁻⁵	-4.122·10⁻⁵	-5.144·10⁻⁵
11	0	-1.101·10⁻⁵	-2.193·10⁻⁵	-3.268·10⁻⁵	-4.32·10⁻⁵	-5.339·10⁻⁵
12	0	-1.173·10⁻⁵	-2.331·10⁻⁵	-3.462·10⁻⁵	-4.55·10⁻⁵	-5.586·10⁻⁵
13	0	-1.258·10⁻⁵	-2.498·10⁻⁵	-3.698·10⁻⁵	-4.841·10⁻⁵	-5.906·10⁻⁵
14	0	-1.364·10⁻⁵	-2.704·10⁻⁵	-3.996·10⁻⁵	-5.211·10⁻⁵	-6.332·10⁻⁵
15	0	-1.493·10⁻⁵	-2.96·10⁻⁵	-4.369·10⁻⁵	-5.692·10⁻⁵	-6.896·10⁻⁵

11. Контрольні питання

1. Класифікація рівнянь. Трансцендентні та алгебраїчні рівняння.
2. Які рівняння відносяться до трансцендентних рівнянь?
3. Які рівняння відносяться до алгебраїчних рівнянь?
4. Суть відокремлення коренів нелінійних рівнянь.
5. Суть методів уточнення коренів.
6. Які способи використовуються для відокремлення коренів?
7. Які теореми використовуються для аналітичного методу відокремлення коренів?
8. В чому суть алгоритму методу половинного ділення? Дайте геометричну інтерпретацію цього методу.
9. Графічна інтерпретація методу половинного ділення та основні формули методу.
10. В чому суть алгоритму методу хорд? Дайте геометричну інтерпретацію цього методу.
11. Графічна інтерпретація методу хорд та основні формули методу.
12. В чому суть алгоритму методу січних? Дайте геометричну інтерпретацію цього методу.
13. Графічна інтерпретація методу січних та основні формули методу.
14. Покажіть особливості методу ітерацій та його обмеження.
15. Графічна інтерпретація методу ітерацій та основні формули методу.
16. Яку систему називають системою лінійних алгебраїчних рівнянь?
17. Що називається розв'язком СЛАР?
18. Яка система називається сумісною і несумісною?
19. Яка система називається визначеною і невизначеною?
20. Яка система називається виродженою і невиродженою?
21. Які системи називаються еквівалентними?
22. Які методи відносять до точних (дати означення і перелічити методи)?

23. Які методи відносять до наближених (дати означення і перелічити методи)?
24. В чому суть алгоритмів методу Гауса?
25. В чому суть прямого ходу в методах Гауса?
26. В чому суть зворотного ходу в методах Гауса?
27. В чому суть методу Крамера?
28. Що таке визначник матриці?
29. Які є способи обчислення визначників?
30. В чому суть методу оберненої матриці?
31. Яка матриця називається оберненою до матриці А?
32. Які є способи знаходження оберненої матриці?
33. Яка матриця називається союзною для матриці А?
34. Загальна постановка задачі інтерполяції.
35. Що називається інтерполяційним многочленом, вузлами інтерполяції?
36. Сформулюйте задачу Коші і крайову задачу. В чому відмінність у постановці цих задач?
37. Наведіть приклади звичайних диференціальних рівнянь, які розв'язуються лише чисельними методами.
38. Дайте порівняльну оцінку і визначте області застосування методів розв'язування звичайних диференціальних рівнянь.
39. Чи можлива крайова задача для рівнянь першого порядку?
40. Що таке похибка зрізання (обмеження)? Чим визначається порядок похибки методу?
41. Покажіть, що загальна формула для методу Рунге-Кутта, застосована до рівняння $y'=-y$ з початковою умовою $y(0)=1$, приводить до співвідношення

$$y_n = 1 - h + \frac{h^2}{2} - \ldots + e^{-x_n}$$

42. В чому подібність методів прогнозу та корекції і виправленим методом Ейлера?
43. Як оцінюються похибки методів Ейлера, Рунге-Кутта, прогнозу і корекції?
44. Дайте геометричну інтерпретацію методів Ейлера.
45. Що таке властивість "самостартування" методів розв'язування звичайних диференціальних рівнянь.?
46. Як обирається крок при використанні багатокрокових методів розв'язування?
47. Математична постановка задачі оптимізації. Приклад.
48. Графічна інтерпритація задачі оптимізації та її розв'язок. Приклади.
49. Градієнт. Його фізичний зміст.
50. Лінії рівня.
51. Вплив параметра кроку на пошуки екстремуму.
52. Формулювання задачі оптимізації. Змістовна і математична. Приклади.
53. Вплив вигляду функції на процес пошуку екстремуму.
54. Чим відрізняється метод ДФП від методу Ньютона?
55. Переваги та недоліки методів першого й другого порядку.
56. Які рівняння відносять до рівнянь у частинних похідних?
57. Що таке дискримінант?
58. Класифікація рівнянь у частинних похідних.
59. Особливості трактування задачі розв'язування рівнянь в частинних похідних на ПК.

60. Що таке дискримінант рівняння?
61. Що таке змішана задача?
62. Особливості різницевих методів. Навести наближені формули для похідних $\frac{\partial u}{\partial x}$, $\frac{\partial^2 u}{\partial x^2}$, $\frac{\partial u}{\partial y}$ (за 2, 3, 5, 9 точками).
63. Різницева схема для розв'язування еліптичного рівняння на ПК.
64. Алгоритми розв'язування еліптичного рівняння на ПК.
65. Особливість методу Лібмана.
66. Особливість прискореного методу Лібмана
67. Різницева схема для розв'язування гіперболічного рівняння на ПК.
68. Різницева схема для розв'язування еліптичного рівняння на ПК.
69. Різницева схема для розв'язування параболічного рівняння на ПК
70. Особливість алгоритму розв'язування гіперболічного рівняння на ПК.
71. Явні та неявні методи розв'язування гіперболічного рівняння на ПК.
72. Явні та неявні методи розв'язування еліптичного рівняння на ПК.
73. Різницеві схеми для розв'язування параболічного рівняння на ПК.
74. Особливість алгоритму розв'язування параболічного рівняння на ПК.
75. Що таке збліження методу?
76. Від чого залежить збліження методів розв'язування еліптичного, гіперболічного та параболічного рівнянь?
77. З допомогою яких функцій можна розв'язувати диференційні рівняння в частинних похідних в середовищі MathCAD?

12. Задачі для самостійної роботи

Задача 12.1. У задачі безумовної оптимізації виконати один крок пошуку екстремуму простим градієнтним методом і обчислити значення цільової функції в знайденій точці.

а) $I(u) = u^2 + 2u + 2$; б) $I(u) = -u^2 + 2u + 2$

Пошук починати з точки $u^{(0)} = 1$. Параметр кроку в алгоритмі пошуку взяти h=0.25.

Задача 12.2. У задачі безумовної оптимізації виконати один крок градієнтоного спуску і обчислити значення цільової функції $I(u_1^{(1)}, u_2^{(1)})$.

$$I(u) = u_1^2 + u_2^2 \to \min_{u_1, u_2};$$

Початкові умови: $u_1^{(0)} = 4$; $u_2^{(0)} = 4$. Параметр кроку: h=1.

Задача 12.3. Задача безумовної оптимізації функції має вигляд

$$I(u) = u^2 + 2u + 2 \to \min_u.$$

Обчислити оптимальний параметр кроку в алгоритмі градієнтного спуску.

Задача 12.4. Задано функцію $f(u) = 100(u_2 - u_1^2)^2 + (1 - u_1)^2$, а також дві перші точки, отримані у процесі пошуку мінімуму функції $f(u)$:

$$u^{(0)} = [-1,2;1]^T, \quad u^{(1)} = [-1,3;1,07]^T.$$

Визначте напрямок пошуку з точки $u^{(1)}$, користуючись наступними методами:

а) градієнтним методом; б) методом Ньютона.

Пошук точки екстремуму зобразити графічно.

Задача 12.5. Побудувати лінію рівня функції $I(u_1,u_2) = (u_1-2)^2 + 4(u_2-1)^2 = 10$. Обчислити та побудувати градієнт в одній із точок лінії рівня рівня.

Задача 12.6. Знайти мінімум функції $I(u_1,u_2) = u_1^2 - u_1 + u_2^2 - u_2 + 1$.

Задача 12.7. Знайти точки стаціонарності функції $I(u_1,u_2) = u_1 - u_1^2 - u_1^3 - u_1u_2$ і виявнити, які з них є точками максимуму.

Задача 12.8.Записати і обчислити матрицю Гессе функції в задачі 12.7 у точках її стаціонарності.

Задача 12.9. Виконати один цикл пошуку максимуму функції

$$I(u_1,u_2) = u_1 + u_2 - (u_1,u_2)^2 - u_1^2 - 2u_2^2,$$

з початкової точки $\vec{u}^{(0)} = (1,0)$, використовуючи градієнтинй метод зі сталим кроком h=1.

Задача 12.10. Побудувати лінію рівня функції: $I(u_1,u_2) = u_1 + u_2 - (u_1,u_2)^2 = 1$. Знайти вектор градієнта в точці $u_1 = 1;\ u_2 = 0;$

Задача 12.11. Знайти точки стаціонарності функції

$$I(u_1,u_2) = u_1 - u_1^2 - u_1^3 - u_1u_2.$$

Задача 12.12. Виконати три цикли пошуку максимуму функції

$$I(u_1;u_2) = u_1 + u_2 - u_1u_2^2 - u_1^2 - 2u_2^2$$

із початкової точки $u^{(0)} = (1,0)$, використовуючи наступні алгоритми:

а) алгоритм градієнта з постійним кроком;

б) алгоритм попокоординатного підйому.

Задача 12.13. Виконати один крок пошуку екстремуму функцій методом Коші й методом Ньютона.

а) $I = u^2 + 2u + 2$ б) $I = -u^2 + 2u + 2$

13. Завдання до розрахункової роботи

Завдання КР.13.1. Методами поділу відрізка пополам, методом хорд, дотичних та ітерацій знайти всі корені нелінійного рівняння. Точність знаходження коренів вважати рівною 0.0000001. Варіанти рівнянь подані в табл. КР.4.

Таблиця КР.4

№	Рівняння	№	Рівняння
1	$\ln^2(x-1) = 3\cos 2x + 1$	16	$\sqrt{25-x^2} = arctg\,2x$
2	$\frac{3\pi}{2}\cos x = e^{0.1x^2} arctg\,2x$	17	$10e^{-x^2} = \sqrt{2\pi x} + \sin x$

Продовження таблиці КР.4

3	$\sin x \cdot \sqrt{81 - x^2} = 5x \cdot arctgx$	18	$arctg2x - 0{,}2(x-1)^4 + \sin x = 0$
4	$\sqrt{\ln^2(x-1)} \cdot \ell^{\sin 3x} = 10\ell^{-0.1x^2}$	19	$\sin 3x \cdot \sqrt{64 - x^2} = 5x\ell^{0.1x}$
5	$\sqrt{36 - x^2}\lg x = \sin 4x$	20	$10\ell^{-0.1x^2} = \sqrt{2\pi + x} + \sin 2x$
6	$arctg2x - \frac{(x-1)^4}{5} + \sin^2 5x = 0$	21	$\frac{10}{1+x^2} = 2\sin 2x + x$
7	$\sin 4x \cdot \sqrt{81 - 25x^2} = 5x \cdot arctgx$	22	$\sin^2 3x \cdot \sqrt{16 - x^2} = 5x\ell^{0.2x}$
8	$\frac{10}{1+x^2} = 2\cos 2x + x$	23	$\frac{x^2-4}{x^2+1} = \sqrt{x}\,\ell^{x\sin x}$
9	$\arcsin x - \sin 5x \cdot \sqrt[4]{1 - x^4} = 0$	24	$4xtg(0{,}5\sqrt{9 - x^2}) = 10\sin 3x$
10	$\frac{x^2-4x}{x^2-4x+8} = \sqrt[3]{x^3+4} \cdot \ell^{\cos 3x}$	25	$\frac{x-1}{x^2-2x+2} = \sqrt[4]{x^4+4} \cdot \ell^{\sin 2x}$
11	$\frac{10x-2}{3+x^2} = 2\cos 2x + \sqrt[4]{x}$	26	$\frac{x^2-9}{x^2+4} = \sqrt{x^2+1} \cdot \ell^{x\cos x}$
12	$\sqrt{64 - x^2}\log_2 x = \sin 3x$	27	$xtg(0{,}5\sqrt{9 - x^2}) = 3\cos x$
13	$10\ell^{-0.3x^2} = \sqrt{2\pi x + x^2} + 3\sin x$	28	$5 \cdot 3^{-x^2} + 1 = \sqrt{3x} + \sin 2x$
14	$arctg2x - (x-0{,}1)^4 + \sin^2 x = 0$	29	$\sin^2 x \cdot \sqrt{81 - x^2} = 5\ell^{-x^2}$
15	$\frac{5\pi}{2}\cos 2x = 3^{0.1x^2} arctg2x$	30	$\frac{x^2-4}{x^2+1} = \sqrt{x} \cdot \ell^{x\sin x}$

Завдання КР.13.2. Скласти схему алгоритму розв'язування нелінійного рівняння (табл. КР.5) методом, вказаним викладачем. Скласти програму розв'язування нелінійного рівняння, користуючись схемою алгоритму. Вибрати рівняння відповідно до варіанту (табл. КР.5). Відокремити корені в рівнянні, що досліджується (відокремлення коренів можна виконати на ПК, розробивши для цього програму). Вибрати значення точності обчислень ε. На кожному відрізку уточнити корені, користуючись розробленою програмою. Початкові дані (відрізок, точність обчислень ε) і результати роздрукувати. Розв'язати рівняння в середовищі MathCAD. Порівняти результати, зробити висновки. До списку параметрів підпрограми слід включити вхідні параметри: [*A*, *B*] – дійсні змінні, які визначають проміжок, на якому уточнюється корінь; *EPS* – дійсне число, що визначає точність обчислень. Вихідні параметри: *X* – дійсна змінна, що визначає наближене значення кореня, який досліджується на заданому проміжку [*A*, *B*]; *K* – цілочисельна змінна, що визначає кількість ітерацій, за яке визначається наближене значення кореня *X* з заданою точністю EPS; *ERR* – ознака розв'язку (код помилки), причому: *ERR*=0 – нормальне завершення роботи програми; *RR*=1 – проміжок [*A*, *B*] заданий невірно; *RR*=2,3,... – інші типи помилок, які визначає студент.

Таблиця КР.5

Метод	Варіант	Рівняння	Примітка
Половинного ділення	1	$x^3 - x + 1 = 0$	-1.325
	2	$x^3 + 2x - 4 = 0$	1.180
	3	$x^4 + 5x - 3 = 0$	-1.876; 0.578
	4	$2.2x - 2^x = 0$	0.781; 2.401
	5	$2^x - 2x^2 - 1 = 0$	0.0;0.399;6.352
	6	$2^x - 4x = 0$	0.310; 4.0
Хорд	7	$x^3 - x - 3 = 0$	1.213
	8	$x^3 + 8x - 6 = 0$	0.703
	9	$x^3 + 10x - 9 = 0$	0.841
	10	$x^2 - \cos \pi x = 0$	-0.438; 0.438
	11	$x^2 - \sin \pi x = 0$	0.0; 0.787
	12	$\lg x - \frac{1}{x^2} = 0$	1.897
	13	$x^3 - 6x^2 + 9x - 3 = 0$	-4.071; 0.466; 0.993
	14	$x^3 - 12x - 8 = 0$	-0.695; -3.067;3.757
	15	$2\lg x - \frac{x}{2} + 1 = 0$	0.398; 4.682
	16	$x^2 - 20\sin x = 0$	0.0; 2.753
	17	$x - \cos x = 0$	0.739
Комбінаційний	18	$x^3 + 6x - 5 = 0$	0.760
	19	$x^3 - 2x + 7 = 0$	-2.258
	20	$x^3 - 2x^2 + x + 1 = 0$	-0.465
	21	$1.8x^2 - \sin 10x = 0$	-0.567;-0.335; 0.0
	22	$\lg x - \frac{7}{(2x+6)} = 0$	3.473
	23	$2x \ln x - 1 = 0$	1.422
Ітерацій	24	$\ln x + (x+1)^3 = 0$	0.187
	25	$x + \lg x = 0.5$	0.672
	26	$tg1.5x - 2.3x = 0$	0.739
	27	$5\sin 5x - x = 0$	0.760
	28	$0.83e^{-0.54x} - x = 0$	-2.258

Завдання КР.13.3. Розв'язати систему лінійних алгебраїчних рівнянь з п'ятьма невідомими $AX = B$ методами Гауса, Крамера і оберненої матриці. Матриця A і вектор B представлені в табл. КР.6.

Таблиця КР.6

№	Матриця A					B	№	Матриця A					B
1	2	4	17	0	−14	−8	18	−15	8	−17	9	−17	13
	−9	19	−14	−13	25	7		12	−1	15	4	−12	3
	9	21	25	15	−2	0		0	−9	−16	4	−10	−6
	−4	−8	−4	20	−7	13		−10	13	−19	19	8	−2
	6	−12	22	10	−4	17		14	1	16	−20	7	−4
2	15	14	13	−7	−6	8	19	25	−18	−19	22	−16	14
	10	25	9	0	9	7		−6	21	19	−3	8	6
	1	19	24	25	−5	3		−19	22	−17	−4	18	13
	−16	0	19	16	−2	13		12	−2	22	−10	8	19
	−5	12	20	5	−2	8		−10	1	19	4	−8	16
3	−4	13	5	19	6	16	20	21	14	24	16	18	−6
	−10	21	1	−13	−2	9		17	−4	−10	0	13	−2
	−15	−20	14	21	−12	2		17	21	7	−2	−7	7
	21	−16	−18	19	13	−3		−2	3	19	18	−13	19
	−13	5	19	23	9	16		9	−4	−8	−16	22	−1
4	−6	−20	10	10	−12	−8	21	−2	−2	6	−9	6	18
	8	−8	−18	−5	1	13		0	6	20	−18	20	11
	−1	22	18	22	15	−1		−8	7	−10	−7	−9	−8
	−18	−13	13	24	−8	7		−1	17	0	−6	24	20
	−7	14	17	−6	−9	8		5	5	15	−11	2	16
5	11	9	19	4	−1	16	22	2	−4	−2	8	−8	12
	−16	18	13	18	8	1		19	12	15	2	2	−5
	−1	15	−14	1	−13	9		−15	12	0	15	7	−2
	2	−8	1	3	16	−6		7	16	0	−4	−18	−1
	14	10	7	0	16	−1		−3	−13	−15	6	12	−10
6	1	19	19	−12	15	−5	23	7	−4	10	−7	−7	2
	−1	−10	−8	−16	−14	−8		−1	12	−5	−8	−12	12
	3	−3	3	1	3	−5		9	12	14	−3	3	−9
	9	1	4	−20	−18	16		1	16	25	−19	2	1
	16	−18	11	−7	15	−4		7	23	12	2	−17	−4
7	19	−17	17	10	−13	−1	24	1	3	−4	−14	−12	18
	−17	−2	15	−7	11	−2		12	4	22	15	−13	8
	−18	11	−16	−13	6	19		1	−3	24	−9	15	5
	−2	12	14	−9	22	−3		23	5	0	−4	13	19
	7	−7	−6	−8	11	9		−16	−6	3	3	−13	5
8	7	5	6	3	10	12	25	−14	11	14	−19	11	0
	13	−16	−16	−18	18	9		14	0	6	14	21	7
	14	−17	17	−10	−20	−10		13	2	−6	−12	6	−8
	20	3	19	−1	−7	17		1	19	−11	18	5	7
	17	9	21	22	10	14		9	−14	19	15	11	−3
9	−6	18	15	8	6	0	26	8	−1	−8	−8	−17	−9
	2	11	−18	17	23	−8		−16	−8	24	15	2	−10
	−13	−2	8	−6	8	1		−5	1	0	5	4	−1
	1	13	11	−10	−5	−2		5	−9	6	18	−12	0
	10	21	8	−1	11	−6		−10	19	15	10	4	3

№	Матриця A					B	№	Матриця A					B
10	-10	3	24	-9	2	9	27	-18	-10	-16	21	-12	0
	-1	15	5	20	-12	12		24	-7	-13	-5	19	2
	-10	7	-10	14	16	-9		-9	-11	-19	-10	8	15
	22	-4	-10	-6	-6	19		-3	-6	-15	-16	3	10
	1	16	7	0	-15	-1		4	-7	20	-12	20	18
11	4	0	-13	-6	23	18	28	20	-14	0	1	16	-4
	-2	23	25	-19	22	14		-5	4	14	-7	-16	17
	23	-9	-6	20	13	16		8	15	5	11	-3	15
	-2	-19	2	-9	-20	13		10	6	-1	13	-15	16
	1	24	-19	12	-13	0		-17	-9	8	18	4	0
12	-5	-11	13	21	-2	9	29	22	-6	-9	-19	22	1
	-11	-16	0	-18	-7	7		-20	2	2	-6	11	7
	-19	-1	19	17	8	7		15	-13	-2	-12	-3	14
	3	-13	25	18	-13	20		-4	2	-17	7	11	-1
	-14	12	23	6	-14	17		3	-19	-1	-18	5	19
13	-19	7	7	24	6	1	30	-3	-14	12	12	24	-1
	17	6	-2	15	8	13		-6	-6	-12	7	-7	-9
	0	16	-6	8	-16	17		15	-18	-1	19	-6	4
	-10	-2	8	16	-16	11		-17	6	4	14	15	1
	13	-18	25	-13	2	3		-8	1	-9	22	-7	18
14	-5	-19	9	15	14	19	31	-13	12	7	-19	4	11
	18	11	14	3	-6	1		-2	2	15	4	6	15
	-6	24	11	6	-7	19		16	-7	-16	8	-15	12
	-7	18	8	11	-7	14		10	2	-17	8	-2	9
	20	0	4	-13	-17	-5		0	-18	7	2	-12	6
15	-3	4	4	15	-11	0	32	2	-13	-2	25	13	19
	7	9	10	0	20	9		0	14	-14	9	24	9
	-5	-11	6	-7	-1	14		4	-9	-8	19	0	7
	16	-6	14	7	20	2		23	13	1	12	3	-2
	-6	24	-19	-2	12	-2		7	24	21	-16	-13	12
16	4	4	-5	13	9	19	33	20	19	18	17	-17	15
	24	-17	-18	-16	24	0		-3	-13	6	-6	7	-8
	-3	3	11	11	-3	1		2	0	16	9	-6	17
	-14	-19	17	11	-15	19		-11	1	-15	19	-20	-10
	19	13	12	-8	5	17		23	24	-14	7	-1	-5

Завдання КР.13.4. Інтерполювати функцію. Вихідні дані:

1. Функція $f(x, \alpha, \beta, \gamma, \delta, \varepsilon, \mu)$.

2. Метод: $1N$ – перший метод Ньютона, $2N$ – другий метод Ньютона, G – метод Гауса, St – метод Стірлінга, B – метод Бесселя.

Вхід:

1. Дійсні параметри: $-100 \le \alpha, \beta, \gamma, \delta, \varepsilon, \mu \le 100$.

2. Дійсні розміри вікна: $-100 \le A, B, C, D \le 100$.

3. Натуральне число вузлів інтерполяції $0 < n \le 100$.

4. Виділені функції зі списку: $f(x)$, $Pn(x)$, $rn(x)$, $\partial f(x)$, $\partial Pn(x)$.

Вихід:

1. Контрастні графіки виділених функцій у вікні $[A, B] \times [C, D]$, де $[A, B]$ – відрізок, на якому вирішується завдання інтерполяції зазначеним методом.

2. Значення $\delta = \max [A, B] \mid rn(x) \mid$ і $\xi \in [A, B]$ – точка, в якій реалізується цей максимум.

3. Побудувати многочлен Ньютона для функції, яка задана таблицею:

47.	48.3	49	50.	51.	52.
53.	54.15	55.7	56.	57.5	58.7

Варіанти завдань

1)	$\varepsilon \sin \frac{\beta x}{a^2 - x^2} + \delta \cos \gamma x$	18)	$\delta \cos \frac{\beta x}{a^2 - x^2} + \varepsilon \sin \gamma x$	35)	$\varepsilon \sin \frac{\beta x}{a^2 - x^2} + \cos \gamma x$
2)	$\varepsilon \sin \frac{\beta x}{a^2 - x^2} + \sin \gamma x$	19)	$\alpha \sin \frac{\beta}{(x-\gamma)^2} + \delta \cos \frac{\varepsilon}{(x-\mu)^2}$	36)	$\varepsilon \sin(\alpha + \beta x + \gamma^x)$
3)	$\alpha \sin \beta x \cdot \cos \frac{\varepsilon}{(x-\mu)^2}$	20)	$\varepsilon \sin \operatorname{tg} \frac{\alpha}{(x-\beta)(x-\gamma)(x-\delta)}$	37)	$\alpha \sin \varepsilon x \cdot \cos\left(\operatorname{tg} \frac{\beta}{x-\gamma}\right)$
4)	$\alpha + \gamma \sin \lvert x \rvert^{\beta} + \varepsilon \sin \delta x$	21)	$\alpha + \gamma \sin \lvert x \rvert^{\beta} + \varepsilon \cos \delta x$	38)	$\varepsilon \cos(\alpha + \beta x + \gamma^x)$
5)	$\varepsilon \sin \frac{\alpha}{\beta x^2 + \gamma x + \delta}$	22)	$\alpha \sin \frac{\beta}{(x-\gamma)^2} + \delta \cos \varepsilon x$	39)	$\alpha \sin \frac{\beta}{(x-\gamma)^2} \cdot \cos \delta x$
6)	$\alpha \cos(\operatorname{tg} \beta x) + \varepsilon \sin \gamma x$	23)	$(\alpha + \cos \lvert x \rvert^{\beta}) \cdot \gamma \cdot \sin \delta x$	40)	$\alpha \sin(\operatorname{tg} \beta x) + \varepsilon \cos \gamma x$
7)	$\alpha \sin \beta x + \gamma \cos \delta x$	24)	$(\alpha + \sin \lvert x \rvert^{\beta}) \cdot \gamma \cdot \cos \delta x$	41)	$(\alpha + \beta x^3) \cdot \sin^2 \gamma x$
8)	$\alpha \sin \beta x + \delta \cos \frac{\varepsilon}{(x-\mu)^2}$	25)	$\frac{\alpha}{\beta + x^2} \cdot \cos \frac{\gamma}{(x-\delta)(x-\varepsilon)}$	42)	$\varepsilon \cos\left(\operatorname{tg} \frac{\alpha x}{(x-\beta)(x-\gamma)}\right)$
9)	$\alpha \sin \beta x + \gamma \cos \lvert x \rvert^{\delta}$	26)	$\alpha \cos(\operatorname{tg}(\beta x^2 + \gamma x + \delta)) \cdot \sin \varepsilon x$	43)	$\alpha \sin \beta^{\gamma \cos \delta x}$
10)	$\varepsilon \cos \frac{\beta x}{\alpha^2 - x^3} \cdot \sin \delta x$	27)	$\alpha \cos(\operatorname{tg} \beta x) \cdot \cos \gamma x$	44)	$(\alpha + \beta x^3) \cdot \cos^2 \gamma x$
11)	$\gamma \sin \frac{\beta x}{\alpha^2 - x^3} + \varepsilon \cos \delta x$	28)	$\alpha \operatorname{tg}(\delta \sin(x-\beta) \cdot \cos(x-\gamma))$	45)	$\varepsilon \sin \frac{\beta x}{\alpha^2 - x^3} \cdot \cos \delta x$
12)	$\frac{\alpha}{\beta + x^2} \cdot \sin \frac{\gamma}{(x-\delta)(x-\varepsilon)}$	29)	$\varepsilon \sin \frac{\alpha x}{(x-\beta)(x-\gamma)(x-\delta)}$	46)	$\gamma \cos \frac{\beta x}{\alpha^2 - x^3} + \varepsilon \sin \delta x$
13)	$\alpha \cos\left(\operatorname{tg} \frac{\beta}{x-\gamma}\right) + \delta \sin \varepsilon x$	30)	$\alpha \sin \frac{\beta x^2}{(x-\gamma)(x-\delta)} + \varepsilon \cos \mu x^2$	47)	$\alpha \sin\left(\operatorname{tg} \frac{\beta}{x-\gamma}\right) + \delta \cos \varepsilon x$
14)	$\alpha \cos \varepsilon x \cdot \sin\left(\operatorname{tg} \frac{\beta}{x-\gamma}\right)$	31)	$\alpha \cos \frac{\beta x^2}{(x-\gamma)(x-\delta)} + \varepsilon \sin \mu x^2$	48)	$\varepsilon \cdot \sin \frac{\alpha}{\beta x^2 + \gamma x + \delta}$
15)	$\alpha \sin \beta x \cdot \cos \lvert x \rvert^{\delta}$	32)	$\alpha \cos(\delta \sin(x-\beta) - \varepsilon \cos(x-\gamma))$	49)	$\alpha \cos \beta^{\gamma \sin \delta x}$
16)	$\alpha \sin \lvert x \rvert^{\beta} + \gamma \cos \delta x$	33)	$\alpha \sin(\delta \sin(x-\beta) - \varepsilon \cos(x-\gamma))$	50)	$\alpha \sin(\operatorname{tg} \beta x) \cdot \sin \gamma x$
17)	$\alpha \sin \lvert x \rvert^{\beta} + \gamma \cos(\operatorname{tg} \delta x)$	34)	$\alpha \cos \beta x \cdot \sin \lvert x \rvert^{\delta}$		$\alpha \sin \beta x + \gamma \operatorname{tg} \lvert x \rvert^{\delta}$

Завдання КР.13.5. Обчислити приблизне значення коренів алгебраїчного або трансцендентного рівняння для варіанту даних, наведених в табл. КР.7, де число M – максимальне значення похідної функції на відрізку.

Таблиця КР.7

№	Рівняння	Відрізок	M	№	Рівняння	Відрізок	M
1	$2\cdot x^4 - 8\cdot x^3 + 8\cdot x^2 - 1 = 0$	[0;1]	4	16	$x\cdot\log_3(x+1) - 1 = 0$	[1,2]	2
2	$x^4 - 4\cdot x^3 + 8\cdot x^2 + 1 = 0$	[0;1]	3	17	$\operatorname{arctg}(x-1) + 2x = 0$	[0,1]	4
3	$(x-1)^2\cdot\lg(x+11) - 1 = 0$	[-10;9]	140	18	$\sqrt{(1-x)} - \operatorname{tg}(x) = 0$	[0,1]	7,5
4	$(x-4)^2\cdot\lg(x-3) - 1 = 0$	[4;6]	9	19	$3\ln^2(x) + 6\ln(x) - 5 = 0$	[1,3]	7,2
5	$\operatorname{arctg}(x) - x^3/3 = 0$	[0.5;1]	20	20	$\sin(x^2) + \cos(x^2) - 10x = 0$	[0,1]	12,5
6	$(\log_2(-x))\cdot(x+2) = -1$	[-1;-0.5]	6	21	$2x\sin(x) - \cos(x) = 0$	[0.4;1]	4,3
7	$2\cdot x^3 - 9\cdot x^2 - 60\cdot x + 1 = 0$	[0;1]	90	22	$\cos^2(x) - 5x = 0$	[0;1]	8
8	$x^2 - \cos(x) = 0$	[0.5;1]		23	$\arccos(x) - \sqrt{(1-0.3\cdot x^3)} = 0$	[0;1]	5
9	$x - 1/3 + \sin(3.6\cdot x) = 0$	[0;0.85]	1,7	24	$\sqrt{(1-x)} - 2\sin(x) = 0$	[0;2]	10
10	$x^5 - x - 0.2 = 0$	[1;1.3]	0.18	25	$x^4 - 3\cdot x^3 - 6\cdot x^2 + 2 = 0$	[0;1]	5
11	$x - \sin(x) = 0$	[0;1]	0.99	26	$3\cdot\sin(\sqrt{x}) + 0.35\cdot x - 3.8 = 0$	[2;3]	0,6
12	$e^x + \ln(x) - 10\cdot x = 0$	[3;4]	6,5	27	$\operatorname{tg}(x) - \operatorname{tg}^3(x)/3 + \operatorname{tg}^5(x)/5 = 1/3$	[0;0.8]	22
13	$\cos(x) - e^{-x} + x - 1 = 0$	[1;2]	1	28	$1 - x + \sin(x) - \ln(1+x) = 0$	[0;1.5]	1,6
14	$x^{10} - x^3 - \sqrt{x} = 0$	[1;5]	0.2	29	$3\cdot x^4 + 8\cdot x^3 + 6\cdot x^2 - 10 = 0$	[0;1]	55
15	$3\cos(x) - 4 = 0$	[0;1]	15	30	$x^{10} - x^3 - 56 = 0$	[1;2]	0.4

Завдання КР.13.6. Обчислити значення

$$Z = sin(Q_1 + Q_2)/cos(Q_1 - Q_2) + 3Q_2 - 4Q_1,$$

де $Q_1 = \int_a^b F_1(x)dx$ $\qquad Q2 = \int_c^d F_2(x)dx$

Значення параметрів свого варіанту взяти із таблиці КР.8.

У цій таблиці номер варіанту розташований у першому стовпчику. У наступних двох колонках приведені підінтегральні функції. У колонці "Метод" вказується однією буквою метод чисельного інтегрування: П – метод прямокутників; Т – трапецій; С – Сімпсона. У наступних чотирьох колонках зазначені межі інтегрування для кожної функції, а в 9-10 – точність інтегрування і початкове значення кількості розподілів. Для визначення заданого виразу рекомендується створити модуль обчислення визначеного інтеграла згідно програми та використати його двічі.

Таблиця КР.8

№	$F_1(x)$	$F_2(x)$	Метод	a	b	c	d	e	N
1	$x^3\cos(3\ln x)$	$\sin x$	П	0	1	1	2	0.001	20
2	$x^2\sin(4\ln x)/\ln x$	$\cos x$	Т	0	1	1	2	0.001	30
3	$\ln((5+3\sin x)/(5-3\sin x))$	e^x	Т	0	$\pi/2$	3	4	0.01	20
4	$x\ln 1/x/(1-x^2)$	e^x	С	0	1	2	3	0.01	20
5	$x^3(1-x)^5$	e^x	П	0	1	1	2	0.0005	20
6	$\sin^2 x/(9-12\cos x+4x)$	$\sin x$	С	0	π	0	$\pi/2$	0.001	30
7	$\sin 4x/(8-1.6\cos x+x)$	$\cos x$	П	0	π	0	$\pi/3$	0.001	20
8	$\sin^2 x/(\sin^2 x+4\cos^2 x)$	$\sqrt{x}$	Т	0	$\pi/2$	0	9	0.01	20
9	$x\sin 3x/(1+\cos 3x)$	$\sqrt{x}$	С	0	π	20	30	0.01	40

№	$F_1(x)$	$F_2(x)$	Метод	a	b	c	d	e	N
10	$x^2/((1-x)^3+2)$	e^x	Т	0	1	40	50	0.0005	20
11	$1/(4-\sin^2 x+0.6\cos^2 x)$	$\sqrt{x}$	П	0	$\pi/2$	35	55	0.001	20
12	$\ln(1-1.2\cos x)\cos 3x$	e^x	Т	0	π	3	3.9	0.001	30
13	$x^3\cos(4\ln x)$	$\sqrt{x}$	С	0	1	17	19	0.002	30
14	$\sin^2 x \ln\sin x$	e^x	Т	0	1	7	8.5	0.005	20
15	$1+x/2x^3\ln((1)/(1+x))$	$\sin x$	Т	0	1	0	$\pi/4$	0.001	30
16	$x\ln x\ln(2-x)$	$\cos x$	П	0.5	1.5	0	$\pi/4$	0.002	20
17	$\sin^{1.5}x\cos^4 x$	$\ln x$	С	0	$\pi/2$	2	3	0.01	40
18	$x\sin x/(1-0.7\cos x+2x)$	$\ln x$	С	0	$\pi/2$	4	7	0.001	30
19	$\ln(0.6\sin^2 x+\cos^2 2x)$	$\sqrt{x}$	Т	0	$\pi/4$	17	25	0.002	50
20	$1/(1+0.2\sin^2 x+2x)$	$\cos x$	С	0	$\pi/4$	0	$\pi/3$	0.0001	20
21	$\sin 3x\,(1+x^2)/(1+x^3)$	$\ln x$	П	0	$\pi/6$	1	3.5	0.002	40
22	$(1-\ln x)^3/(1+\sqrt{x})$	$\sin x$	Т	0.1	2	0	3	0.005	20
23	$(1+\ln x)^3/(2-\sin^2 3x)$	e^x	П	0.01	1.5	0	3	0.01	50
24	$\sqrt{1+x^2}\sin x/(2-\cos^3 3x)$	$\ln x$	Т	0	$\pi/6$	2	7	0.001	20
25	$\sqrt{1+x^2}/(2-\sqrt{1+\cos^2 x})$	$\ln x$	С	0	$\pi/6$	2	5	0.005	30
26	$x\sin^2 x+2\cos 3x$	e^x	Т	0.1	$\pi/4$	0.3	0.4	0.001	20
27	$(x-1)\sin 3x/(x^2+2)$	$\lg x$	П	0.3	1	0.5	1.5	0.001	40
28	$\sqrt{x}+x\sin x/(3+4x\sin x)$	e^x	Т	0.5	1	1.2	1.9	0.002	30
29	$\ln 3x+0.2\ln(5-x)^2$	$\sin x$	С	0	1	3	2	0.001	20
30	$\sqrt{1+\sin x\cos x}/(x^3+3)$	$\ln x$	П	0.1	1	2	15	0.01	40
31	$\lg 5x+\ln\sqrt{14x^3}$	$\cos x$	П	0.2	1.5	0	3	0.001	20

Завдання КР.13.7. Реалізувати методи розв'язку крайових задач: метод прогону та метод Гальоркіна.

Завдання КР.13.8. Для заданих значень X знайти значення функції, заданої дискретно. Використовувати рекомендовані чисельні методи пошуку максимуму функції двох змінних.

№	Задане значення аргументу	Табличне значення функції		Задане значення аргументу	Табличне значення функції	
		X	Y		X	Y
1	0.507	0.50	1.6487	1.207	1.20	4.60487
2	0.512	0.51	1.6653	1.212	1.21	1.13143
3	0.523	0.52	1.6820	2.823	1.22	2.28575
4	0.535	0.53	1.6989	0.535	1.23	6.69069
5	0.541	0.54	1.7160	0.941	1.24	1.25660
6	0.556	0.55	1.7333	3.256	1.25	8.44043
7	0.568	0.56	1.7507	0.555	1.26	4.75507
8	0.571	0.57	1.7683	1.221	1.27	1.71223
9	0.589	0.58	1.7860	0.639	1.28	1.53850
10	0.545	0.59	1.96356	0.845	1.29	2.66356
11	0.642	0.60	1.98545	1.521	1.30	3.28545
12	0.653	0.61	1.99749	1.328	1.31	4.95349

№	Задане значення аргументу	Табличне значення функції		Задане значення аргументу	Табличне значення функції	
		X	Y		X	Y
13	0.609	0.62	1.99957	3.109	1.32	1.57757
14	0.632	0.63	1.99166	1.732	1.33	5.99199
15	0.649	0.64	1.97385	1.849	1.34	2.72685
16	0.659	0.65	1.94630	1.929	1.35	1.25637
17	2.031	0.66	1.90930	2.031	1.36	1.90932
18	2.173	0.67	1.86321	2.173	1.37	2.26321
19	2.218	0.68	1.80850	2.218	1.38	1.83120
20	1.134	0.69	1.17520	1.134	1.39	4.77533
21	1.089	0.70	1.30254	1.089	1.40	1.30764
22	1.14	0.71	1.38631	0.144	1.41	1.48135
23	1.19	0.72	1.50946	1.192	1.42	1.20241
24	1.25	0.73	1.21730	1.258	1.43	1.11710
25	1.305	0.74	1.22361	2.304	1.44	1.32465

Завдання КР.13.9. Відокремити корені рівняння *графічно* і уточнити один із них методом ітерації з точністю до 0,001. Відокремити корені рівняння *аналітично* і уточнити один із них методом ітерації з точністю до 0,001.

№	Завдання 1	Завдання 2	№	Завдання 1	Завдання 2
1.	$\ln x + (x+1)^3$;	$x^3 + 2x^2 + 2 = 0$.	19.	$x = \sqrt{\lg(x+2)}$;	$x^3 + 3x^2 + 6x - 1 = 0$.
2.	$x \cdot 2^x = 1$;	$x^3 - 3x^2 + 9x - 10 = 0$.	20.	$x^2 = \ln(x+1)$;	$x^3 - 0{,}1x^2 + 0{,}4x - 1{,}5 = 0$.
3.	$\sqrt{x+1} = \frac{1}{x}$;	$\sin(0{,}5 + x) = 2x - 0{,}5$.	21.	$2x + \lg x = -0{,}5$;	$x^3 - 3x^2 + 6x - 2 = 0$
4.	$x - \cos x = 0$;	$0{,}5x + \lg(x-1) = 0{,}5$.	22.	$2x + \cos x = 0{,}5$;	$x^3 - 0{,}2x^2 + 0{,}3x - 1{,}2 = 0$
5.	$3x + \cos x + 1 = 0$;	$\sin 0{,}5x + 1 = x^2, x > 0$	23.	$x^3 + x - 3 = 0$;	$x^3 - 3x^2 + 12x - 9 = 0$
6.	$x + \ln x = 0{,}5$;	$x^3 + 0{,}4x^2 + 0{,}6x - 1{,}6 = 0$.	24.	$x^3 + 3x - 1 = 0$;	$x^3 + 0{,}2x^2 + 0{,}5x - 2 = 0$
7.	$2 - x = \ln x$;	$x^3 - 0{,}2x^2 + 0{,}4x - 1{,}4 = 0$.	25.	$x^3 - 2x + 2 = 0$;	$x^3 + 3x + 1 = 0$
8.	$(x-1)^2 = \frac{1}{2}e^x$;	$x^3 - 0{,}1x^2 + 0{,}4x + 2 = 0$.	26.	$\lg(2+x) + 2x = 3$;	$x^3 + 0{,}2x^2 + 0{,}5x - 1{,}2 = 0$
9.	$(2-x)e^x = 0{,}5$;	$x^3 + 3x^2 + 12x + 3 = 0$.	27.	$\lg(1+2x) = 2 - x$;	$x^3 - 3x^2 + 9x + 2 = 0$
10.	$2{,}2x - 2^x = 0$;	$x^3 - 0{,}2x^2 + 0{,}5x - 1 = 0$.	28.	$x^3 + 2x + 4 = 0$;	$x^3 - 0{,}1x^2 + 0{,}4x - 1{,}5 = 0$
11.	$x^2 + 4\sin x = 0$;	$x^3 - 0{,}1x^2 + 0{,}4x + 1{,}2 = 0$	29.	$x + \lg(1+x) = 1{,}5$;	$x^3 - 3x^2 + 6x + 3 = 0$
12.	$2x - \lg x = 7$;	$x^3 - 3x^2 + 6x - 5 = 0$.	30.	$x + \cos x = 1$;	$x^3 - 0{,}1x^2 + 0{,}3x - 0{,}6 = 0$.
13.	$5x - 8\ln x = 8$;	$x^3 - 0{,}2x^2 + 0{,}5x - 1{,}4 = 0$.	31.	$2 - x = \ln x$;	$x^3 - 0{,}2x^2 + 0{,}4x - 1{,}4 = 0$.
14.	$3x - e^x = 0$;	$2\sin(x - 0{,}6) = 1{,}5 - x$.	32.	$(x-1)^2 = \frac{1}{2}e^x$;	$x^3 - 0{,}1x^2 + 0{,}4x + 2 = 0$.
15.	$x(x+1)^2 = 1$;	$x^3 - 3x^2 + 12x - 12 = 0$.	33.	$(2-x)e^x = 0{,}5$;	$x^3 + 3x^2 + 12x + 3 = 0$
16.	$x = (x+1)^3$;	$x^3 + 0{,}2x^2 + 0{,}5x + 0{,}8 = 0$.	34.	$x^3 = \sin x$;	$x^3 + 0{,}1x^2 + 0{,}4x - 1{,}2 = 0$.
17.	$x^2 = \sin x$;	$x^3 - 4x^2 - 6 = 0$.	35.	$x = \sqrt{\lg(x+2)}$;	$x^3 + 3x^2 + 6x - 1 = 0$.
18.	$x^3 = \sin x$;	$x^3 + 0{,}1x^2 + 0{,}4x - 1{,}2 = 0$.	36.	$x^2 = \ln(x+1)$;	$x^3 - 0{,}1x^2 + 0{,}4x - 1{,}5 = 0$

Завдання КР.13.10. Розв'язати чисельними методами Ейлера та Рунге-Кутта звичайне диференційне рівняння з початковими умовами (задачу Коші) на відрізку $[a,b]$ з кроком h та побудувати графік розв'язку (інтегральну криву).

Варіанти завдань представлені в табл. КР.9.

Таблиця КР.9

№	Диференційне рівняння	Початкові умови	$[a,b]$, h	№	Диференційне рівняння	Початкові умови	$[a,b]$, h
1	$y' = xy - \cos(x^2 + 2y)$	$y(0) = -3$	[0,1], 0.001	16	$y' = \sqrt{2y + x^2}$	$y(0) = 0$	[0,1], 0.001
2	$y' = \sqrt{\sin^2(x^2 + 3y) + 1}$	$y(0) = 1$	[0,1], 0.001	17	$y' = \sqrt{y + 3x^3 y}$	$y(0) = 1$	[0,2], 0.01
3	$y' = \sin^2(2x + 3y) + 3x$	$y(0) = 3$	[0,1], 0.001	18	$y' = x^2 + 2y\sin x$	$y(0) = 0.5$	[0,1], 0.001
4	$y' = \sqrt{\ln(2x^2 + y^2) + y}$	$y(0) = 3$	[0,1], 0.001	19	$y' = y^2 + yx^2$	$y(0) = 1$	[0,2], 0.01
5	$y' = \sqrt{\cos^2(x + 3y) + y^2}$	$y(0) = 1$	[0,2], 0.01	20	$y' = \sin(y + x^2)$	$y(0) = 0$	[0,1], 0.001
6	$y' = \sqrt{\sin^2(y) + 3x^2}$	$y(0) = 0$	[0,1], 0.001	21	$y' = \sin\left(\sqrt{2x^2 + y}\right)$	$y(0) = 0$	[0,2], 0.01
7	$y' = 3xy^3 - \cos(x + 2y)$	$y(0) = -3$	[0,2], 0.01	22	$y' = \sqrt{y^2 + t^3}$	$y(0) = 0$	[0,2], 0.01
8	$y' = \sqrt{\cos^2(x + y^2) + y^2}$	$y(0) = 1$	[0,1], 0.001	23	$y' = \sin(x + y^2)$	$y(0) = 0$	[0,2], 0.01
9	$y' = \sqrt{\cos^2(2x) + \sin^2(y^2)}$	$y(0) = 1$	[0,1], 0.001	24	$y' = \sin\left(\sqrt{x^2 + y^2}\right)$	$y(0) = 0$	[0,1], 0.001
10	$y' = \sqrt{\cos^2(x^2) + \sin^2(y^3)}$	$y(0) = 1$	[0,1], 0.001	25	$y' = \sqrt{\sin^2(y^2) + 3}$	$y(0) = 0$	[0,2], 0.01
11	$y' = \sqrt{\sin^2(x + y^3) + 4}$	$y(0) = 1$	[0,2], 0.01	26	$y' = \sqrt{y + 5xy^3}$	$y(0) = 1$	[0,1], 0.001
12	$y' = \sqrt{x^2 + y^2} + \sin(x + 2y)$	$y(0) = 0$	[0,1], 0.001	27	$y' = yx^2 + 3\sin y$	$y(0) = 0.5$	[0,2], 0.01
13	$y' = \sqrt{\ln(x^2 + 5y^2) + 2y}$	$y(0) = 3$	[0,2], 0.01	28	$y' = y^2 + y^3$	$y(0) = 1$	[0,1], 0.001
14	$y' = \sqrt{x^2 + y^2} + \sin(2x + y)$	$y(0) = 0$	[0,2], 0.01	29	$y' = \ln(3x^2 + 5y)$	$y(0) = 1$	[0,2], 0.01
15	$y' = \sin^2(x - y) + \sin^2(2x + 3y)$	$y(0) = 3$	[0,2], 0.01	30	$y' = \ln(x^2 + 3y^2)$	$y(0) = 1$	[0,1], 0.001

Завдання КР.13.11. Градієнтні методи пошуку екстремуму. 1) Вивчити найпростіший градієнтний метод оптимізації, метод крутого спуску, метод найшвидшого спуску, методи Ньютона, метод ДФП. 2) Виконати вправи 11.1-11.5, відповідно до варіанту.

Вправа 11.1. Побудувати лінію рівня.

11.1.1) $I(u_1, u_2) = (u_1-2)^2 + 4(u_2-1)^2 = 10$;

11.1.2) $I(u_1, u_2) = (u_1-1)^2 + 2(u_2-4)^2 = 32$;

11.1.3) $I(u_1, u_2) = 4(u_1-2)^2 + (u_2-2)^2 = 8$;

11.1.4) $I(u_1, u_2) = 2(u_1-2)^2 + (u_2-1)^2 = 8$;

11.1.5) $I(u_1, u_2) = (u_1-2)^2 + 2(u_2-2)^2 = 12$;

11.1.6) j $I(u_1, u_2) = (u_1-4)^2 + 2(u_2-4)^2 = 32$;

11.1.7) $I(u_1,u_2)=(u_1-2)^2+2(u_2-4)^2=32$;

11.1.8) $I(u_1,u_2)=(u_1-1)^2+4(u_2-4)^2=20$;

11.1.9) $I(u_1,u_2)=u_1^2+4u_2^2=4$;

11.1.10) $I(u_1,u_2)=2(u_1-3)^2+(u_2-3)^2$.

Обчислити й побудувати градієнт в одній із точок лінії рівня.

Вправа 11.2. Побудувати сімейство ліній рівня. Графічно зобразити пошук екстремуму простим градієнтним методом.

11.2.1) $I(u_1,u_2)=(u_1-1)^2+2(u_2-2)^2$;

11.2.2) $I(u_1,u_2)=u_1^2+4(u_2-1)^2$;

11.2.3) $I(u_1,u_2)=(u_1-1)^2+2u_2^2$;

11.2.4) $I(u_1,u_2)=(u_1-1)^2+4(u_2-1)^2$;

11.2.5) $I(u_1,u_2)=2(u_1-1)^2+(u_2-2)^2$;

11.2.6) $I(u_1,u_2)=(u_1-1)^2+(u_2-4)^2$;

11.2.7) $I(u_1,u_2)=(u_1-2)^2+4(u_2-1)^2$;

11.2.8) $I(u_1,u_2)=4(u_1-2)^2+(u_2-1)^2$;

11.2.9) $I(u_1,u_2)=(u_1-2)^2+(u_2-4)^2$;

11.2.10) $I(u_1,u_2)=(u_1-4)^2+(u_2-4)^2$.

Вправа 11.3. Для заданої функції виконати декілька кроків пошуку екстремуму за методом градієнта, найшвидшого спуску, Ньютона, ДФП і зобразити траєкторію руху на площині параметрів u_1,u_2.

Вправа 11.4. Знайти точки стаціонарності функції та класифікувати їх.

11.4.1) $I=u_1^2+u_1u_2+u_2^3+u_1$;

11.4.2) $I=u_1+u_2-(u_1+u_2)^2-4u_1^2$;

11.4.3) $I=1-(u_1^2-u_2)^2+u_1u_2-10u_2^2$;

11.4.4) $I=u_1^3-u_1u_2+u_2^2-2u_1+3u_2-4$;

11.4.5) $I=2u_1^2+4u_1u_2^2-10u_1u_2+u_2^2$;

11.4.6) $I=4u_1^2+3u_2^2-4u_1u_2+u_1$;

11.4.7) $I=8u_1^2+4u_1u_2+5u_2^2$;

11.4.8) $I=u_1^3+u_2^2-3u_1-2u_2+2$;

11.4.9) $I=u_1^2+u_2+u_1u_2+u_1^3-1$.

Вправа 11.5. Тестові функції.

1) $I(\vec{u})=100(u_2-u_1^2)^2+(1-u_1)^2$; початкова точка (1.2;1): $I_{\text{OPT}}=0$; $\vec{u}_{\text{OPT}}=(1;1)$;

2) $I(\vec{u})=(u_2-u_1)^2+(1-u_1)^2$;

3) $I(\vec{u})=(u_2-u_1^2)^2+100(1-u_1)^2$;

4) $I(\vec{u})=100(u_2-u_1^3)^2+(1-u_1)^2$;

5) $I(\vec{u})=[1.5-u_1(1-u_2)]^2+[2.25-u_1(1-u_2^2)]^2+[2.625-u_1(1-u_2^3)]^2$; $I_{\text{OPT}}=0$; $\vec{u}_{\text{OPT}}:(3;\frac{1}{2})$;

6) $I(\vec{u})=(u_1^2+u_2-11)^2+(u_1+u_2^2-7)^2$;

7) $I(\vec{u})=(u_1-u_2)+\frac{(u_1+u_2-10)}{9}$; додаткова точка (0;1); $I_{\text{OPT}}=0$; $\vec{u}_{\text{OPT}}:(5;5)$

8) $I(\vec{u})=(u_1-u_2+u_3)^2+(u_2-u_1+u_3)^2+(u_1+u_2-u_3)^2$; початкова точка (0,5;1;0.5); $I_{\text{opt}}=0$; $u_{\text{opt}}:(0;0;0)$.

9) $I(\vec{u})=\left\{12+u_1^2+\frac{1+u_2^2}{u_1^2}-\frac{u_1^2u_2^2+100}{(u_1u_2)^4}\right\}\frac{1}{10}\to\min$;

10) $I(\vec{u}) = 100(u_2 - u_1)^2 + (1-u_1) + 90(u_4 - u_3^2)^2 + (1-u_3)^2 + 10.1[(u_2-1)^2 + (u_4-1)^2] + 19.8(u_4-1);$

$u_3 = u_4 = 1$

11) $I(\vec{u}) = (u_1^2 + 2u_2^2)\exp(-(u_1^2 + u_2^2)). \longrightarrow \min\limits_{u_1, u_2}$

Завдання КР.13.12. Градієнтні методи пошуку екстремуму. Сформулювати задачу оптимізації, змістовна постановка якої подана нижче. Звести задачу умовної оптимізації до задачі безумовної оптимізації, підготувати вихідні дані для розв'язку на ПК.

12.1. Виконати апроксимацію експериментальних даних лінійною залежністю за методом найменших квадратів (застосувавши пошуковий алгоритм оптимізації).

№ вар	1		2		3		4		5	
№ екс	y_i	u_i	y_i	u_i	y_i	u_i	y_i	u_i	y_i	u_i
1	2	1	4	1	2	1	2	1	2	1
2	3	2	5	2	4	2	4	2	3	2
3	4	3	9	3	5	3	6	3	4	3
4	8	4	12	4	9	4	10	4	9	4

12.2. Вибір оптимального розташування центрального вузла при прокладці трас до споживачів. Задані координати розташування споживачів сировини або електроенергії. Потрібно вибрати координати джерела енергії або проміжкової ємності для сировини, щоб сумарна довжина комунікацій була мінімальною.

Координати споживачів (u_1,u_2)

12.2.1) (2,1);(1,3);(6,2);(8,6);(4,8);
12.2.2) (3,1);(6,8);(5,8);(8,6);(4,8);
12.2.3) (4,2);(4,4);(2,3);(1,2);(6,7);
12.2.4) (1,1);(2,4);(3,1);(1,5);(2,7);
12.2.5) (2,2);(3,1);(4,2);(1,7);(2,6);
12.2.6) (5,1);(4,3);(3,2);(6.1);(4.2);
12.2.7) (1,1);(7,7);(7,1);(8,2);(4,6);
12.2.8) (2,3);(4,5);(3,1);(2,6);(2,6);
12.2.9) (2,1);(6,3);(2,3);(1,5);(2,6);
12.2.10) (4,5);(3,1);(7,1);(1,4);(2,1).

12.3. Спроектувати ємність мінімальної вартості. Матеріал і об'єм ємності задані. Запропонувати форму ємності. Штуцери для подачі і відбору рідини в ємність приварюються після її виготовлення і при розрахунку вартості до уваги не беруться.

12.4. Спроектувати ємність циліндричної форми, для якої при заданому об'ємі довжина зварювальних швів мінімальна. Наявність штуцерів для подачі і відбору матеріалу, що буде зберігатись, до уваги не приймається.

Завдання КР.13.13. Градієнтні методи пошуку екстремуму. Розробити структурну схему алгоритму оптимізації.

Варіант	1	2	3	4	5
Метод	Ньютона	Флетчера-ДФП	Ньютона. модифікований	Найшвидшого спуску	Градієнтний з адаптивним кроком

Завдання КР.13.14. Градієнтні методи пошуку екстремуму. Знайти мінімум і/або максимум функції або роз'язок системи рівнянь відповідно до номера варіанту. Починайте із точок (0.25; 2.5);(92.5; 2.5);(-2.5; -2.5).

14.1) $I = 20 + 0.3u_1 - 4u_2 + 0.3u_2^2 + 0.4u_1u_2$.

14.2) $I = (u_1^2 + u_2^2 - 1)^2 + (u_1 + u_2 - 1)^2$.

14.3) $\begin{cases} u_1^2 + (u_2 - 1)^2 - 5 = 0; \\ (u_1 - 1)^2 + u_2^2 - 1 = 0. \end{cases}$

14.4) $\begin{cases} u_1^2 + u_2 - 11 = 0; \\ u_1 + u_2^2 - 7 = 0. \end{cases}$

14.5) $I(\vec{u}) = u_1^4 + u_2^4 + 2u_1^2u_2^2 - 4u_1 + 3$.

14.6) $I = (u_1 - u_2)^2 + \frac{1}{9}(u_1 - u_2 - 10)^2$.

14.7) $I(\vec{u}) = (u_1u_2)^2(1 - u_1)^2[1 - u_1 - u_2(1 - u_1)^5]^2$.

14.8) $I(\vec{u}) = (u_1^2 + 12u_2 - 1)^2 + (49u_1^2 + 49u_2^2 + 84u_1 + 2324u_2 - 681)^2$.

14.9) $I = u_1^2 - u_1 + u_2^2 - u_2 + 1$.

14.10) $I = (u_1 + u_2)^3 + (u_1 - u_2 + 1)^2$.

14.11) $I(\vec{u}) = u_1(1 - u_1)^2[1 - u_2(1 - u_1)^2]$.

14.12) $I(u) = 7(2u_3 + 1)^2$.

14.13) $I(\vec{u}) = (u_1 - 1)^2 + (u_1^2 + u_2^2 - u_1 - u_2 + 8)$

14.14) $I = 3u_1^3 - u_1 + u_2^3 - 3u_2^2 - 1$

14.15) $I(\vec{u}) = u_1^3 - u_2^2 - 3u_1 - 2u_2 + 2$.

Завдання КР.13.15. Градієнтні методи пошуку екстремуму. Для зазначеного викладачем методу розробити структурну схему алгоритму й програму на одній з мов програмування.

15.1. Градієнтний пошук одночасно по всіх координатах.

15.2. Покоординатний градієнтний пошук.

15.3. Покоординатний пошук, що починається зі змінної з максимальною похідною.

Вибрати контрольний приклад згідно номера варіанту і виконати один ітераційний цикл. Виконати налагодження розробленої програми оптимізації.

15.4.1. Алгоритм вибору величини кроку $h^{(k+1)} = \begin{cases} h^{(k)}, Y^{(k)} < Y^{(k-1)}; \\ -\dfrac{h^{(k)}}{2}, Y^{(k)} \geq Y^{(k-1)}; \end{cases}$, де $Y^{(k)}$ – значення критерію оптимізації; k – номер кроку пошуку.

15.4.2. Алгоритм вибору величини кроку $h^{(k+1)} = \begin{cases} 2h^{(k)}, a^{(k)} < a_{\min}; \\ h^{(k)}, a_{\min} < a^{(k)} < a_{\max}; \\ \dfrac{h^{(k)}}{3}, a^{(k)} > a_{\max}. \end{cases}$

$\cos a = \dfrac{\sum_j \dfrac{dI(u^{(k)})}{du_j} \cdot \dfrac{dI(u^{(k-1)})}{du_j}}{\sqrt{\sum_j \dfrac{dI(u^{(k)})}{du_j} \cdot \dfrac{dI(u^{(k-1)})}{du_j}}}$, де $Y\left(\vec{x}^{(k)}\right)$ – критерій оптимізації обчислений в $\vec{x}^{(k)}$.

15.4.3 Алгоритм вибору величини кроку $h^{(k+1)} = \begin{cases} 1.5h^{(k)}, Y^{(k)} < Y^{(k-1)}; \\ \pm\dfrac{h^{(k)}}{2}, Y^{(k)} \geq Y^{(k-1)}. \end{cases}$ де $Y^{(k)}$ – критерій оптимізації; k –номер кроку пошуку.

15.4.4. Алгоритм вибору величини кроку $h^{(k+1)} = \begin{cases} 2h^{(k)}, Y^{(k)} < Y^{(k-1)}; \\ \dfrac{h^{(k)}}{3}, Y^{(k)} \geq Y^{(k-1)}. \end{cases}$

15.4.5. Алгоритм вибору величини кроку $h_j^{(k+1)} = \begin{cases} ga_1 \cdot h_j^{(k)}, \left(\dfrac{dI(\vec{u}^{(k)})}{dx_j} \cdot \dfrac{dI(\vec{u}^{(k-1)})}{dx_j} \right) > 0; \\ ga_2 \cdot h_j^{(k)}, \left(\dfrac{dI(\vec{u}^{(k)})}{du_j} \cdot \dfrac{dI(\vec{u}^{(k-1)})}{du_j} \right) \leq 0; \\ ga_1 = 1.2, ga_2 = 0.4. \end{cases}$

k – номер кроку пошуку, j – номер змінної.

15.4.6. $u_j^{(k+1)} = u_j^{(k)} - h^{(k)} \operatorname{sgn} \dfrac{dI(\vec{u}^{(p)})}{du_j}$, p – точка, у якій обчислюється похідна.

15.4.7. Алгоритм із адаптивним кроком: $u_j^{(k+1)} = u_j^{(k)} - h^{(k)} \dfrac{dI(\vec{u}^{(k)})}{du_j}$; $h^{(k)}$ вибирається згідно пунктів 15.4.1-15.4.5.

15.4.8. Найпростіший градієнтний алгоритм: $u_j^{(k+1)} = u_j^{(k)} - h_j \dfrac{dI(\vec{u}^{(k)})}{du_j}$, $j = \overline{1,n}$.

15.4.9. Алгоритм крутого спуску: $u_j^{(k+1)} = u_j^{(k)} - h^{(k)} \dfrac{dI(u^{(p)})}{du^{(j)}}$, де p – точка, в якій обчислений градієнт; k – номер кроку пошуку.

15.4.10. Алгоритм найшвидшого спуску.

$$u_j^{(k+1)} = u_j^{(k)} - h^{(k)} \frac{dI(\vec{u}^{(k)})}{du_j}, \qquad h^{(k)} = \arg\min_h \left[I(u^{(k)}) - h \cdot grad(I(\vec{u}^{(k)})) \right].$$

15.4.11. Спосіб обчислення градієнта. Градієнт обчислюється за допомогою різницевої апроксимації.

Завдання КР.13.16. Виконати дослідження роботи алгоритму з використанням розробленої програми.

16.1. Для дослідження взяти функцію $y = a_0 + a_1(x_1 - b_1)^2 + a_2(x_2 - b_2)^2$. Параметр a_1 дорівнює номеру варіанта. Величини a_i, $i = \overline{0.2}$ вибирати за узгодженням з викладачем. Коефіцієнти a_1, a_2 необхідно змінювати в процесі дослідження, з метою їх впливу на пошук екстремуму.

16.2. Визначити вплив параметрів алгоритму, а також величин a_1, a_2 початкових умов $x_i^{(0)}$ на процес пошуку екстремуму. Визначити найкраще значення параметрів

алгоритму.

Завдання КР.13.17. Розв'язування диференційних рівнянь у частних похідних на ПК. Скласти схему алгоритму розв'язування диференційного рівняння в частинних похідних 2-го порядку методом, вказаним викладачем. Скласти програму розв'язування диференційного рівняння в частинних похідних 2-го порядку алгоритмічною мовою за вказівкою викладача.Використовуючи дані відповідного варіанту (табл. КР.10), розв'язати диференційне рівняння в частинних похідних на ПК (початкові дані та результати надрукувати). Розв'язати диференційне рівняння у частинних похідних 2-го порядку в середовищі MathCAD. Порівняти результати, які отримані при виконані завдання.

Варіанти завдання приведені в таблиці КР.10. Для розв'язування диференційного рівняння у частинних похідних необхідно розробити:

- основну програму, яка виконує опис даних, ввід початкових даних, виклик підпрограми розв'язування диференційного рівняння, вивід початкових даних та отриманих результатів;
- підпрограму (або декілька підпрограм) розв'язування диференційного рівняння в частинних похідних.

В список параметрів підпрограми слід включити такі параметри:

– вхідні: N – кількість поділів по горизонталі; M – кількість поділів по вертикалі; $(N\text{-}1)(M\text{-}1)$ – кількість вузлів сітки; HY – крок (інтервал між поділами) по вертикалі; F_0 – масив із дійсних чисел (граничні умови);

– вихідні: F – масив із $N\times M$ дійсних чисел (розв'язування диференційного рівняння в усіх вузлах сітки); IRR – ознака розв'язання (код помилки), причому: $IRR=0$ – нормальне завершення роботи підпрограми; $IRR=1, 2, 3,\ldots$ – види помилок, які визначаються студентом.

Знайти розв'язок $u(x, y)$ задачі Діріхле в квадраті зі стороною 1 для рівняння Лапласа з граничними умовами:

$$u(0, y) = f_1(y), \quad (0 \le y \le 1); \quad u(1, y) = f_2(y), \quad (0 \le y \le 1);$$
$$u(x,0) = f_3(x), \quad (0 \le x \le 1); \quad u(x,1) = f_4(x), \quad (0 \le x \le 1).$$

Таблиця КР.10

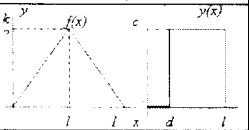
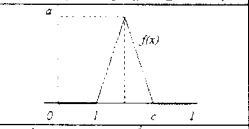
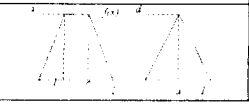
№	$f_1(y)$	$f_2(y)$	$f_3(x)$	$f_4(x)$
1	y^2	$\cos y + (2-\cos 1)y$	x^3	$x+1$
2	$e^y - ey^2$	y	$-x^3+1$	x^2
3	$-y^2+1$	y	$\sin x - (1+\sin 1)x^3 + 1$	x
4	0	y	$\sin x - x^3 \sin 1$	x
5	$e^y + (1-e)y^2 - 1$	y	0	x
6	y^2	$\cos y + (3-\cos 1)y$	x^3	$2x+1$
7	0	y	x^3	x^2
8	$2e^y - (1+2e)y^2 - 1$	$-y$	$-x^3+1$	$x-2$
9	$(-10y^2 - 8y + 6)$	$-10y^2 - 30y + 22$	$9x^2 + 7x + 6$	$9x^2 - 5x - 12$

Продовження таблиці КР.10

10	$-7y^2-5y+3$	$-7y^2-21y+13$	$6x^2+4x+3$	$6x^2-12x-9$
11	$-6y^2-4y+2$	$-6y^2-18y+10$	$5x^2+3x+2$	$5x^2-11x-8$
12	$-5y^2-3y+1$	$-5y^2-15y+7$	$4x^2+2x+1$	$4x^2-24x-21$
13	$-19y^2-17y+15$	$-19y^2-57y+49$	$18x^2-16x+15$	$18x^2-24x-21$
14	$-2y-4y^2$	$4-12y-4y^2$	$x+3x^2$	$-6-9x+3x^2$
15	1	$y+1$	1	$x+1$
16	1	$y+1$	1	x^2+1
17	1	e^y	1	e^x
18	e^{-y^2}	e^{1-y^2}	e^{x^2}	$e^{x^2}-1$
19	$-y^3$	$1-y^3$	x^2	x^2-1
20	$5y-y^2$	$4-y^2+5y$	x^2+3x	x^2+3x-5

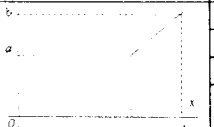
Розв'язати змішану задачу до хвильового рівняння $\frac{\partial^2 u}{\partial t^2}=\frac{\partial^2 u}{\partial x^2}$ з початковими умовами $u(x,0)=f(x)$, $\frac{\partial u}{\partial t}(x,0)=g(x)$; $0\le x\le 1$ та граничними умовами $u(0,t)=u(1,t)=0$

Таблиця КР.11

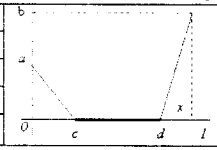
№	*f(x), g(x)*	*a*	*b*	*c*	*d*	*l*
21		1,0	0,05	1,5	0,05	0,45
22		3,0	0,15	1,7	0,15	0,55
23		7,0	0,35	-2,1	0,40	0,75
24		3,0	0,15	0,55	-	-
25		7,0	0,35	0,75	-	-
26		10,0	0,50	0,90	-	-
27		2,0	0,15	0,30	-3,0	0,20
28		5,0	0,40	0,60	-1,0	0,50
29		9,0	0,80	0,95	3,0	0,90
30		10,0	0,90	1,0	3,5	0,95

Розв'язати змішану задача для рівняння теплопровідності $\frac{\partial u}{\partial t}=\frac{\partial u}{\partial x^2}$ з початковими умовами $u(x,0)=f(x)$ та граничними $u(0,t)=a$; $u(1,t)=b$.

Таблиця КР.12

№	*f(x), g(x)*	*a*	*b*	*c*	*d*	*l*
31		1,3	4,0	0,25	-	-
32		1,5	5,0	0,45	-	-
33		1,8	6,5	0,75	0,25	-
34		10,0	5,0	22,0	0,55	-

Продовження таблиці КР.12

35		13,0	8,0	25,0	0,55	-
36		17,0	12,0	29,0	0,95	-
37		11,1	14,5	0,10	0,35	-
38		6,0	16,0	0,25	0,50	-
39		1,0	19,5	0,50	0,75	-

Завдання КР.13.18. Табулювання функцій за логічних умов. Ознайомитись з методикою та набути початкових навичок використання логічних функцій для організації розрахунків в табличному процесорі Excel.

18.1. Логічні функції використовують для організації обчислень в тому випадку, коли необхідно виконувати ті чи інші дії залежно від виконання деяких умов. **Excel** дозволяє використовувати наступні логічні функції:

Логічні оператори прийнятими мовами			Виконувані функції
українська	англійська	російська	
ЯКЩО ( )	**IF ( )**	**ЕСЛИ ( )**	Приймає те чи інше значення в залежності від виконання умови
І ( )	**AND ( )**	**И ( )**	Приймає значення **ІСТИНА**, якщо значення **ІСТИНА** приймають усі аргументи
АБО ( )	**OR ( )**	**ИЛИ ( )**	Приймає значення **ІСТИНА**, якщо хоча б один аргумент приймає значення **ІСТИНА**
НЕ ( )	**NOT ( )**	**НЕ ( )**	Змінює значення свого аргументу на протилежне

Для перевірки істинності умови використовують логічні вирази – вирази, утворені математичними операторами, функціями та аргументами, серед яких обов'язково є хоча б один логічний оператор. Результатом логічного виразу є логічні значення **ІСТИНА** ("1") або **ХИБНІСТЬ** ("0").

18.2. Реалізація логічних функцій.

18.2.1. ***Функція*** **І ( )**. Логічне множення: повертає значення **ІСТИНА**, якщо всі аргументи мають значення **ІСТИНА**: повертає значення **ХИБНІСТЬ**, якщо хоча б один аргумент має значення **ХИБНІСТЬ**. Записується в такий спосіб:

І(*логічне_значення1; логічне_значення2; …*)

Тобто, функція приймає логічне значення **ІСТИНА**, лише тоді, коли всі аргументи є істиною. Функція може мати від одного до тридцяти аргументів. Аргументи повинні бути або логічними значеннями, або масивами, або посиланнями на комірки, в яких містяться логічні значення. Приклади:

1. **І(ІСТИНА;ІСТИНА)=ІСТИНА.**
2. **І(ХИБА;ІСТИНА)=ХИБА.**
3. **І(2+2=4;2+3=5)=ІСТИНА** (обидва аргументи функції є логічними виразами, які завжди приймають значення **ІСТИНА**).
4. **І(А3;С7)** - значення яке прийме функція, залежить від значень, що містяться в комірках **А3** та **С7**. Якщо в комірках **А3** та **С7** знаходяться логічні значення **ІСТИНА**, то функція приймає значення **ІСТИНА**. В усіх інших випадках функція прийматиме значення **ХИБА.**

5. **І(В1:В3)** – значення, яке прийме функція, залежить від значень, що містяться в діапазоні комірок **В1:В3**. якщо в комірках **В1:В3** містяться значення **ІСТИНА**, **ХИБА** і **ІСТИНА** то **І(В1:В3)=ХИБА**.

18.2.2. ***Функція*** **АБО ( )**. Функція приймає логічне значення **ІСТИНА**, якщо хоча б один її аргумент має логічне значення **ІСТИНА** і приймає значення **ХИБА**, якщо усі її аргументи приймають значення **ХИБА**. Цю функцію записують так:

АБО(*логічне_значення1; логічне_значення2; ...*)

Функція може мати від 1 до 30 аргументів. Аргументи повинні бути або логічними значеннями, або масивами, або посиланнями на комірки, в яких містяться логічні значення. Приклади:

1. **АБО(ІСТИНА;ХИБА)=ІСТИНА.**
2. **АБО(ХИБА;ХИБА)=ХИБА.**
3. **АБО(2+2=4;2+3=5)=ІСТИНА** (обидва аргументи функції є логічними виразами, які завжди приймають значення **ІСТИНА**).
4. **АБО(2+2=5;2+3=1)=ХИБА** (обидва аргументи функції є логічними виразами, які завжди приймають значення **ХИБА**).
5. **АБО(А3;С7)** – значення яке прийме функція, залежить від значень, що містяться в комірках **А3** та **С7**. Якщо в комірках **А3** та **С7** знаходяться логічні значення **ХИБА**, то функція приймає значення **ХИБА**. В усіх інших випадках функція прийматиме значення **ІСТИНА.**
6. **АБО(В1:В3)** – значення, яке прийме функція, залежить від значень, що містяться в діапазоні комірок **В1:В3**. якщо в комірках **В1:В3** містяться значення **ІСТИНА, ХИБА** і **ІСТИНА** то **І(В1:В3)=ІСТИНА**.

18.2.3. ***Функція*** **ЯКЩО ( )**. Записується в такий спосіб:

ЯКЩО(*логічна умова, яку перевіряють; значення1; значення2*)

Даний вираз можна розширити, тобто він допускає розгалуження за рахунок вкладання функцій

ЯКЩО(*логічна умова, яку перевіряють; значення якщо істина;* ЯКЩО(*логічна умова, яку перевіряють; значення якщо істина;* ЯКЩО(*логічна умова, яку перевіряють; значення якщо істина; значення якщо хибність*)

Для прикладу протабулюємо функцію і побудувати її графік:

$$y=\begin{cases}1+\ln(1+x), & x<0.2;\\ \dfrac{1+x^{0.5}}{1+x}, & x\in[0.2,\ 0.8];\\ 2e^{-2x}, & x>0.8.\end{cases}$$

на проміжку $[0,\ 1]$ з кроком 0.1.

Розв'язок

1. Визначаємо діапазон зміни значень аргумента x та конкретні значення цього діапазону, використовуючи для цього процедуру автозаповнення з заданим кроком, наприклад, починаючи з комірки B6.

2. Поруч (справа) в сусідню комірку з коміркою першого значення аргументу x, тобто С6 вводимо формулу:

=ЕСЛИ(B6<0,2; 1+LN(1+B6); ЕСЛИ(И(B6>=0,2; B6<=0,8);
(1+B6^(0.5))/(1+B6); 2*EXP(-2*B6)))

3. Використовуючи автозаповнення табулюємо функцію y.

ЗАВДАННЯ для самостійного розв'язування

Варіант № 1.

1. Побудувати графіки наступних функцій для $x \in [-2; 2]$ і кроку 0.2 в різних системах координат $y = \sin(x)e^{-2x}$,

$$g = \begin{cases} \dfrac{1+x^2}{\sqrt{1+x^4}}, & x \le 0, \\ 2x + \dfrac{\sin^2(x)}{2+x}, & x > 0 \end{cases} \qquad z = \begin{cases} \dfrac{1+|x|}{\sqrt[3]{1+x+x^2}}, & x \le -1, \\ 2\ln(1+x^2) + \dfrac{1+\cos^4(x)}{2+x}, & x \in (-1, 0), \\ (1+x)^{3/5}, & x \ge 0. \end{cases}$$

2. Побудувати в одній системі координат при $x \in [-2; 2]$ з кроком 0.1 графіки функцій: $y = 2\sin(x)\cos(x)$, $z = 3\cos^2(2x)\sin(x)$.

3. Побудувати поверхню $z = x^2 - 2y^2$ при $x, y \in [-1; 1]$ (крок 0.1)

Варіант № 2.

1. Побудувати графіки наступних функцій для $x \in [-2; 2]$ і кроку 0.2:

$$y = \frac{1+x^2}{1+2x^2}, \qquad g = \begin{cases} 3\sin(x) - \cos^2(x), & x \le 0, \\ 3\sqrt{1+x^2} & x > 0. \end{cases} \qquad z = \begin{cases} \dfrac{1+x}{\sqrt[3]{1+x^2}}, & x \le 0; \\ -x + 2e^{-2x}, & x \in (0, 1); \\ |2-x|^{1/3}, & x \ge 1. \end{cases}$$

2. Побудувати в одній системі координат при $x \in [-1; 3]$ з кроком 0.2 графіки таких двох функцій: $y = 2\sin(\pi x) - 3\cos(\pi x)$, $z = \cos^2(2\pi x) - 2\sin(\pi x)$.

3. Побудувати поверхню $z = 3x^2 - 2y^2\sin(y)$ при $x, y \in [-1; 1]$ (крок 0.1).

Варіант № 3.

1. Побудувати графіки наступних функцій для $x \in [-3; 1]$ і кроку 0.2:

$$y = \frac{2+\sin^2(x)}{1+x^2}, \qquad g = \begin{cases} \dfrac{3x^2}{1+x^2}, & x \le 0; \\ \sqrt{1+\dfrac{2x}{1+x^2}}, & x > 0. \end{cases} \qquad z = \begin{cases} 3x + \sqrt{1+x^2}, & x < 0; \\ 2\cos(x)e^{-2x}, & x \in [0, 1]; \\ 2\sin(3x), & x > 1. \end{cases}$$

2. Побудувати в одній системі координат при $x \in [-1; 2]$, крок 0.1 графіки наступних двох функцій: $y = 5\sin(\pi x) - \cos(3\pi x)\sin(\pi x)$ та $z = \cos(2\pi x) - 2\sin^3(\pi x)$.

3. Побудувати поверхню $z = 5x^2\cos^2(y) - 2y^2 e^y$ при $x, y \in [-1; 1]$, крок 0.1.

Вариант № 4.

1. Побудувати графіки наступних функцій для $x \in [-5; 1]$ і кроку 0.2

$$y = \frac{1+\cos(x)}{1+e^{2x}}, \qquad g = \begin{cases} \dfrac{3+\sin^2(2x)}{1+\cos^2(x)}, & x \le 0; \\ 2\sqrt{1+2x}, & x > 0. \end{cases} \qquad z = \begin{cases} \sqrt{1+\dfrac{x^2}{1+x^2}}, & x < 0; \\ 2\cos^2(x), & x \in [0, 1]; \\ \sqrt{1+|2\sin(3x)|^{1/3}}, & x > 1. \end{cases}$$

2. Побудувати в одній системі координат при $x \in [1; 4]$, крок 0.2 графіки наступних двох функцій: $y = 3\sin(2\pi x)\cos(\pi x) - \cos^2(3\pi x)$ та $z = 2\cos^2(2\pi x) - 3\sin(3\pi x)$.

3. Побудувати поверхню

$$z = \begin{cases} 2x^2 - e^y, & |x+y| < 0.5, \\ xe^{2x} - y, & 0.5 \le |x+y| < 1, \\ 2e^x - ye^y, & 1 \le |x+y|, \end{cases}$$

при $x, y \in [-2; 1.4]$, крок 0.2.

Вариант № 5.

1. Побудувати в різних системах координат для $x \in [-1.8; 1.8]$ з кроком 0.2 графіки наступних функцій:

$$y = \sqrt[4]{1+e^{3x}}; \quad g = \begin{cases} \dfrac{3+\sin(x)}{1+x^2}, & x \le 0, \\ 2x^2\cos^2(x), & x > 0; \end{cases} \qquad z = \begin{cases} |x|^{1/3}, & x < 0, \\ -2x^2 + \dfrac{x}{1+x}, & x \in [0, 1), \\ \dfrac{3-x}{1+x}, & x \ge 1. \end{cases}$$

2. Побудувати в одній системі координат при $x \in [0; 3]$, крок 0.1 графіки наступних двох функцій: $y = 2\sin(\pi x)\cos(\pi x)$ та $z = \cos^2(\pi x)\sin(3\pi x)$.

3. Побудувати поверхню $z = 2x^2\cos^2(x) - 2y^2$ при $x, y \in [1; 4]$, крок 0.2.

Вариант № 6.

1. Побудувати в різних системах координат для $x \in [-2; 1.8]$ з кроком 0.2 графіки наступних функцій:

$$y=\frac{2+3x}{1+x+x^2},\quad g=\begin{cases}\sqrt{1+2x^2-\sin^2(x)}, & x\le 0,\\ \dfrac{2+x}{\sqrt[3]{2+e^{-0.1x}}}, & x>0;\end{cases}\qquad z=\begin{cases}\dfrac{1+x}{1+x^2}, & x<0,\\ \sqrt{1+\dfrac{x}{1+x}}, & x\in[0,1),\\ 2\left|\sin(3x)\right|, & x\ge 1\end{cases}$$

2. Побудувати в одній системі координат при $x\in[-3;0]$, крок 0.2 графіки наступних двох функцій: $y=3\sin(3\pi x)\cos(2\pi x)$ та $z=\cos^3(4\pi x)\sin(\pi x)$.

3. Побудувати поверхню $z=2e^{0.2x}x^2-2y^4$ при $x,\ y\in[-2;2]$, крок 0.4.

Вариант № 7.

1. Побудувати в різних системах координат для $x\in[-1.7;1.6]$ з кроком 0.3 графіки наступних функцій:

$$y=\frac{1+x}{1+\sqrt{2+x+x^2}},\quad g=\begin{cases}\sqrt{1+x^2}, & x\le 0,\\ \dfrac{1+x}{1+\sqrt[3]{1+e^{-0.2x}}}, & x>0;\end{cases}\quad z=\begin{cases}\dfrac{1+x+x^2}{1+x^2}, & x<0,\\ \sqrt{1+\dfrac{2x}{1+x^2}}, & x\in[0,1),\\ 2\left|0.5+\sin(x)\right|, & x\ge 1.\end{cases}$$

2. Побудувати в одній системі координат при $x\in[-3;0]$, крок 0.2 графіки наступних двох функцій: $y=2\sin(2\pi x)\cos(4\pi x)$ та $z=\cos^2(3\pi x)-\cos(\pi x)\sin(\pi x)$.

3. Побудувати поверхню $z=x^2-2e^{0.2y}y^2$ при $x,\ y\in[-1;1]$, крок 0.2.

Вариант № 8.

1. Побудувати в різних системах координат, коли $x\in[-1.5;1.8]$ з кроком 0.2 графіки таких функцій:

$$y=\frac{x^{3/5}-\sqrt[3]{\dfrac{x+3}{2}}}{1.1+\cos^2 3x},\qquad g=\begin{cases}\dfrac{5x}{x^2+1}, & x\le 0,\\ \sqrt{1-\dfrac{3x}{x^2+2^x}}, & x>0.\end{cases}\qquad z=\begin{cases}\dfrac{\ln^2 x-\sin\left(\dfrac{x}{5.7}\right)}{\sqrt{1.2x}}, & x<3,\\ tg\sqrt{x}, & x=3,\\ 1.2\sqrt{5.7+e^x} & x>3.\end{cases}$$

2. Побудувати в одній системі координат при $x\in[-1;2]$, крок 0.2, графіки двох функцій: $y=2\sin(0.5\pi x)\sin(3.2\pi x)$ та $z=\cos^2(1.5\pi x)-3.4\cos(\pi x)\sin(0.6\pi x)$.

3. Побудувати поверхню $z=2x^2+3e^{0.4y}y^{0.5}$ при $x,\ y\in[-1;1]$, крок 0.1.

Вариант № 9.

1. Побудувати в різних системах координат при $x \in [-1.7; 1.8]$ з кроком 0.3 графіки таких функцій:

$$y = \frac{1 + xe^{-x}}{2 + \sqrt{x^2 + \sin^2(x)}},\quad g = \begin{cases} \sqrt{1 + |x|}, & x \le 0 \\ \dfrac{1 + 3x}{2 + \sqrt[3]{1 + x}}, & x > 0 \end{cases} \qquad z = \begin{cases} 1 + \dfrac{3 + x}{1 + x^2}, & x < 0, \\ \sqrt{1 + (1 - x)^2}, & x \in [0, 1), \\ \dfrac{1 + x}{1 + \cos^2(x)}, & x \ge 1. \end{cases}$$

2. Побудувати в одній системі координат при $x \in [0; 2]$, крок 0.1, графіки двох функцій: $y = \sin(3\pi x) + 2\sin(2\pi x)\cos(3\pi x)$ та $z = \cos(\pi x) - \cos(3\pi x)\sin^2(\pi x)$.

3. Побудувати поверхню

$$z = \begin{cases} x - e^{2y}, & |x| + |y| < 0.5, \\ 2x^2 - e^y, & 0.5 \le |x| + |y| < 1, \\ e^{2x} - y, & 1 \le |x| + |y|. \end{cases}$$

при $x, y \in [-1; 1]$, крок 0.1.

Вариант № 10.

1. Побудувати в різних системах координат при $x \in [-1.4; 2.2]$ з кроком 0.2 графіки таких функцій:

$$y = \frac{1 + xe^{-x}}{2 + x^2}\sin^2(x),\quad g = \begin{cases} \dfrac{\sqrt{1 + |x|}}{2 + |x|}, & x \le 0, \\ \dfrac{1 + x}{2 + \cos^3(x)}, & x > 0. \end{cases} \qquad z = \begin{cases} \dfrac{1 + 2x}{1 + x^2}, & x < 0. \\ \sin^2(x)\sqrt{1 + x}, & x \in [0, 1), \\ \sin^2(x)e^{0.2x} & x \ge 1. \end{cases}$$

2. Побудувати в одній системі координат при $x \in [0; 2]$, крок 0.2, графіки двох функцій: $y = \cos(3\pi x) + 2\sin(3\pi x)\cos(2\pi x)$ та $z = \cos^2(\pi x) - \cos(3\pi x)$.

3. Побудувати поверхню

$$z = \begin{cases} x^2 - 3y^2, & x^2 + y^2 \le 1, \\ 3x^2 - y^3, & x^2 + y^2 > 1. \end{cases}$$

при $x, y \in [-1; 1]$, крок 0.2.

Вариант № 11.

1. Побудувати в різних системах координат при $x \in [-1.4; 1.4]$ з кроком 0.2 графіки наступних функцій.

$$y=\frac{1+x}{1+\sqrt{|x|e^{-x}+|\sin(x)|}},\qquad g=\begin{cases}\sqrt[3]{1+x^2}, & x\le 0,\\ \sin^2(x)+\dfrac{1+x}{1+\cos^2(x)}, & x>0.\end{cases}$$

$$z=\begin{cases}\dfrac{|x|}{1+x^2}e^{-2x}, & x<0;\\ \sqrt{1+x^2}, & x\in[0,1);\\ \dfrac{1+\sin(x)}{1+x}+3x, & x\ge 1.\end{cases}$$

Вариант № 12.

1. Побудувати в різних системах координат графіки наступних функцій на проміжку $[1;10]$ з кроком 1.0 при $a=17.3$, $b=0.36$:

$$y=\frac{e^{\sin(\mu)}+\sqrt[4]{a+\mu}}{\ln^3 b\mu},\qquad g=\begin{cases}3\sin^2(x)-\cos(x), & x\le 0,\\ \sqrt[3]{1-x^2}, & x>0.\end{cases}\qquad z=\begin{cases}\dfrac{\ln^2|12-x|}{5.7+x}, & x<4,\\ \sqrt[3]{x}, & x=4,\\ \dfrac{\sqrt{x}}{1.2+\sqrt{5.7x}}, & x>4,\end{cases}$$

Вариант № 13.

1. Побудувати в різних системах координат на проміжку $[-2;15]$ з кроком 1 при $a=1.35$, $b=8.4$ графіки таких функцій:

$$y=\frac{\sqrt[4]{1+\sqrt{a\varpi+b}}}{\sin^2(b\varpi)+\varpi},\qquad g=\begin{cases}\dfrac{1+|x|}{\sqrt[3]{1+x+x^2}}, & x\le -1,\\ (1+x)^{3/5}, & x>0.\end{cases}\qquad z=\begin{cases}\dfrac{tg^2(1.7x)}{6.2-x^3}, & x<7,\\ \sin^3\left(x^2+\sqrt{x}\right), & x=7,\\ \sqrt[3]{6.2+e^{x/6}}, & x>7.\end{cases}$$

Вариант № 14.

1. Побудувати в різних системах координат на проміжку $[-1;10]$ з кроком 1 при $a=1.8$, $b=0.56$ графіки таких функцій:

$$y=\frac{tg^2(\gamma-a)+\sqrt{\ln(\gamma)}}{e^{-b\gamma}},\qquad g=\begin{cases}\dfrac{1+x}{\sqrt{1+x^3}}, & x\le 0,\\ |2-x|^{1/3}, & x>1.\end{cases}\qquad z=\begin{cases}\dfrac{\ln^2(1.7x)}{\sqrt{6x+1}}, & x<6,\\ \sqrt[5]{e^x}, & x=6,\\ tg^2(1.7x)+\sqrt[4]{e^{5-x}}, & x>6.\end{cases}$$

Варіант № 15.

1. Побудувати в різних системах координат на проміжку $[-1;5]$ з кроком 0.5 при $a=0.3$, $b=9.5$ графіки таких функцій:

$$y=\frac{\sqrt[3]{a^3+\lambda^3}}{tg^2 b\lambda+1.6}.\qquad g=\begin{cases}2\cos(x)e^{-2x}, & x<0,\\ 2\sin(3x), & x\ge 0,\end{cases}\qquad z=\begin{cases}\dfrac{\ln^4(7.3x)}{\sqrt{9.1+x+1.23}}, & x<4,\\ \sqrt[5]{x}, & x=4,\\ \lg^2(x+4)+\sqrt{\dfrac{17x}{5.2}}\ln(x), & x>4.\end{cases}$$

2. Побудувати в одній системі координат при $x\in[0;3]$ з кроком 0.2 графіки наступних двох функцій: $y=3\cos(\pi x)-\cos(\pi x)\sin(4x)$, $z=\cos^3(\pi x)+\sin(2\pi x)$.

3. Побудувати поверхню $z=-\frac{1}{x}+\cos^3(y)$ на проміжку $x\in[10;20]$ з кроком 2 та $y\in[0;1]$ з кроком 0.1.

Завдання КР.13.19. Реалізація матричних операцій засобами MS EXCEL

1. Після отримання конкретного варіанту матриць – індивідуального завдання, на робочому полі **Excel**-таблиці заносимо їхні значення у відповідні комірки.

2. Для кожного пункту завдання, в якому є слово **"Приклад"** ***виключаємо*** це слово і вводимо в кінці теоретичної частини цього підрозділу слова **"Отриманий результат"** і проводимо відповідну операцію над своїми даними.

19.1. Виконати дії: $\lambda A\pm\mu B$ та $\lambda A\cdot\mu B$ або $\lambda A\cdot\mu B^T$.

19.1.1) $\lambda=1.1$, $\mu=0.5$, $A=\begin{bmatrix}3 & 4 & -1\\ -2 & 0 & 3/4\\ 7 & 0.1 & 8\end{bmatrix}$; $B=\begin{bmatrix}1/4 & 3 & -0.8\\ -3 & 0.6 & 3/7\\ 6 & 4 & 2\end{bmatrix}$.

19.1.2) $\lambda=0.7$, $\mu=-1/3$, $A=\begin{bmatrix}1 & 4 & -0.7\\ 1/3 & 7 & 8\end{bmatrix}$; $B=\begin{bmatrix}2/5 & 0.1 & 7\\ -3 & -1 & 4\end{bmatrix}$.

19.1.3) $\lambda=1/6$, $\mu=0.2$, $A=\begin{bmatrix}1 & 4\\ 3 & -7\\ 3/6 & 0.1\end{bmatrix}$; $B=\begin{bmatrix}-7 & 8\\ 9 & 0.1\\ -1/8 & 4\end{bmatrix}$.

19.1.4) $\lambda=-1.6$, $\mu=0.1$, $A=\begin{bmatrix}7 & -1\\ 8/11 & 0.25\end{bmatrix}$; $B=\begin{bmatrix}6/11 & 7\\ 0 & -0.8\end{bmatrix}$.

19.1.5) $\lambda=-1/9$, $\mu=0.15$, $A=\begin{bmatrix}0.25\\ -6\\ 0.54\\ 2\end{bmatrix}$; $B=\begin{bmatrix}-1\\ 3\\ -8/7\\ 0.35\end{bmatrix}$.

19.1.6) $\lambda=2/5$, $\mu=-0.7$, $A=\begin{bmatrix}-1/3 & 1/7 & -4 & 0.2\end{bmatrix}$; $B=\begin{bmatrix}-1 & 0 & 5 & -0.4\end{bmatrix}$.

19.1.7) $\lambda=3/7$, $\mu=-0.7$, $A=\begin{bmatrix}-4 & 7 & 8\\ 2/3 & 1/3 & 0.2\end{bmatrix}$; $B=\begin{bmatrix}1/3 & 4 & -0.2\\ 3 & -4 & 7\end{bmatrix}$.

19.1.8) $\lambda = 4/5$, $\mu = 0.7$, $A = \begin{bmatrix} 4 & 2/3 \\ 0.7 & -0.8 \\ 0.8 & -3 \\ 1/3 & 4 \end{bmatrix}$; $B = \begin{bmatrix} 6 & 4 \\ 2.1 & 1/7 \\ -4 & 0.8 \\ 3 & -3.4 \end{bmatrix}$.

19.1.9) $\lambda = 3/4$, $\mu = 0.1$, $A = \begin{bmatrix} 6 & 2 \\ 3/7 & -1 \end{bmatrix}$; $B = \begin{bmatrix} 4/3 & 0.7 \\ 1.3 & 2 \end{bmatrix}$.

19.1.10) $\lambda = 0.2$, $\mu = -1/3$, $A = \begin{bmatrix} 2 & -4 & -3 & 4 \\ 1/4 & 0.3 & -2 & 7 \end{bmatrix}$; $B = \begin{bmatrix} 7 & -3 & 4 & -0.8 \\ 2/3 & 4 & 8 & 1/3 \end{bmatrix}$.

19.1.11) $\lambda = 0.25$, $\mu = -2/3$, $A = \begin{bmatrix} 1 & 4 \\ -2 & 3 \\ 7 & 8 \end{bmatrix}$; $B = \begin{bmatrix} 1.1 & 4.2 \\ -3.1 & 4/3 \\ 2/7 & -1 \end{bmatrix}$.

19.1.12) $\lambda = 0.3$, $\mu = -3/7$, $A = \begin{bmatrix} 1 & 3 & -4 \\ 2 & 4 & 6 \\ -1 & 2/3 & 1/6 \end{bmatrix}$; $B = \begin{bmatrix} 6 & -2 & -3 \\ 2/3 & 1/3 & 4 \\ -2 & 1 & 4 \end{bmatrix}$.

19.1.13) $\lambda = -0.6$, $\mu = 4/7$, $A = \begin{bmatrix} 4/3 & 1 \\ 2 & 4.3 \\ 3 & -4 \end{bmatrix}$; $B = \begin{bmatrix} 6 & 4 \\ 2 & -8 \\ 4/7 & -3/11 \end{bmatrix}$.

19.1.14) $\lambda = 0.7$, $\mu = 5/7$, $A = \begin{bmatrix} 3/7 & 8/11 & -2 \\ 4 & 6 & 8 \end{bmatrix}$; $B = \begin{bmatrix} 2/5 & 6/5 & 1 \\ -3/4 & -5/7 & 1 \end{bmatrix}$.

19.1.15) $\lambda = 0.9$, $\mu = 5/9$, $A = \begin{bmatrix} 1/3 & 2/3 & -2 \\ 7 & 8 & -2/3 \\ 7.2 & 8.4 & 3 \end{bmatrix}$; $B = \begin{bmatrix} 9/2 & -1 & 3/4 \\ 9/5 & 2 & -4/5 \\ 7 & 1 & 8 \end{bmatrix}$.

19.1.16) $\lambda = 1.2$, $\mu = -2/7$, $A = \begin{bmatrix} 7/2 & 4 & -1 \\ 4 & 2 & 6 \\ -2.7 & 1 & 4 \end{bmatrix}$; $B = \begin{bmatrix} -7/2 & -7/3 & 4.2 \\ 2 & 8 & 1 \\ -2 & -14 & \end{bmatrix}$.

19.1.17) $\lambda = 2.4$, $\mu = -3/8$, $A = \begin{bmatrix} 1.5 & 4 \\ 2.7 & -8.2 \end{bmatrix}$; $B = \begin{bmatrix} 4 & -1 \\ 2 & 3/4 \end{bmatrix}$.

19.1.18) $\lambda = 2.6$, $\mu = -5/7$, $A = \begin{bmatrix} 4.2 & -3.1 & 7 & 8 \end{bmatrix}$; $B = \begin{bmatrix} 7 & -7/3 & 7/8 & -3/4 \end{bmatrix}$.

19.1.19) $\lambda = 2.8$, $\mu = -5.9$, $A = \begin{bmatrix} 2 & 1 & 7 \\ 4 & -2 & -1 \end{bmatrix}$; $B = \begin{bmatrix} 9 & -4 & -9/5 \\ -18 & 36 & -4 \end{bmatrix}$.

19.1.20) $\lambda = 3.2$, $\mu = -4/9$, $A = \begin{bmatrix} 1 \\ 0.3 \\ 2 \\ -4 \end{bmatrix}$; $B = \begin{bmatrix} 9 \\ -9/3 \\ -9/4 \\ -18 \end{bmatrix}$.

19.2. Обчислити добутки матриць $A \cdot B$, $B \cdot A$, $A \cdot B^T$ та $B \cdot A^T$ якщо вони існують.

№	Завдання	№	Завдання
1)	$A=\begin{bmatrix}1&2&0\\3&1&-4\end{bmatrix};\ B=\begin{bmatrix}-2&3\\4&5\\-1&4\end{bmatrix}$	11)	$A=\begin{bmatrix}1&3&5\\-2&0&4\\1&2&4\end{bmatrix};\ B=\begin{bmatrix}2&3&3\\0&4&0\\-1&-2&4\end{bmatrix}$
2)	$A=\begin{bmatrix}2&-1&1\\4&0&2\end{bmatrix};\ B=\begin{bmatrix}1&2\\3&4\\-1&3\end{bmatrix}$	12)	$A=\begin{bmatrix}6&8&4\\5&7&-1\\3&4&8\end{bmatrix};\ B=\begin{bmatrix}3&4&6&-1\\0&2&4&-1\\2&4&-3&0\end{bmatrix}$
3)	$A=\begin{bmatrix}1&4&3&4\\8&6&-1&8\\2&7&6&-1\end{bmatrix};\ B=\begin{bmatrix}3&4\\7&-1\\8&4\\0&-2\end{bmatrix}$	13)	$A=\begin{bmatrix}1&4\\5&0\\6&7\\0&-1\end{bmatrix};\ B=\begin{bmatrix}1&7&0&5\\2&4&8&1\end{bmatrix}$
4)	$A=\begin{bmatrix}1&2&3\\8&3&4\\5&6&-1\end{bmatrix};\ B=\begin{bmatrix}3&0\\1&-2\\3&7\end{bmatrix}$	14)	$A=\begin{bmatrix}1&2&1\\0&-1&3\\4&5&1\end{bmatrix};\ B=\begin{bmatrix}1&2&5\\0&-1&-1\\3&4&6\end{bmatrix}$
5)	$A=\begin{bmatrix}1&3&0\\4&5&-1\\3&4&8\\7&-1&3\end{bmatrix};\ B=\begin{bmatrix}3&4\\1&0\\8&-1\end{bmatrix}$	15)	$A=\begin{bmatrix}1&4&6\\5&-1&3\\2&0&7\end{bmatrix};\ B=\begin{bmatrix}4&1&0&5\\2&0&3&7\\1&-1&4&0\end{bmatrix}$
6)	$A=\begin{bmatrix}2&0&-1&4\\3&6&0&2\end{bmatrix};\ B=\begin{bmatrix}3&2\\1&7\\8&4\\-1&0\end{bmatrix}$	16)	$A=\begin{bmatrix}1&0&-2&4\\5&-1&3&4\end{bmatrix};\ B=\begin{bmatrix}1&0\\3&-2\\7&8\\9&0\end{bmatrix}$
7)	$A=\begin{bmatrix}1&3&0\\-2&3&0\\2&1&4\end{bmatrix};\ B=\begin{bmatrix}1&4\\2&3\\-1&0\end{bmatrix}$	17)	$A=\begin{bmatrix}5&6&-1&2\\4&2&3&4\\0&1&7&6\end{bmatrix};\ B=\begin{bmatrix}3&4&0\\3&-1&2\\-1&6&4\\0&2&-1\end{bmatrix}$
8)	$A=\begin{bmatrix}3&1\\2&8\\-1&3\end{bmatrix};\ B=\begin{bmatrix}2&1&0&4\\-1&3&2&8\end{bmatrix}$	18)	$A=\begin{bmatrix}1&4&3\\3&4&6\\-1&2&4\\0&-3&2\end{bmatrix};\ B=\begin{bmatrix}4&1&0&-1\\2&6&4&-1\\3&0&8&-2\end{bmatrix}$
9)	$A=\begin{bmatrix}2&1\\0&7\\0&2\end{bmatrix};\ B=\begin{bmatrix}3&1&0&-2\\6&0&-3&4\end{bmatrix}$	19)	$A=\begin{bmatrix}3&2&1\\4&5&6\\4&2&-1\end{bmatrix};\ B=\begin{bmatrix}1&6&4&-1\\5&0&3&0\\2&4&-1&0\end{bmatrix}$
10)	$A=\begin{bmatrix}1&7&9\\0&-1&2\\6&1&4\\5&0&-3\end{bmatrix};\ B=\begin{bmatrix}3&1\\4&2\\6&8\end{bmatrix}$	20)	$A=\begin{bmatrix}3&1&3\\2&0&3\end{bmatrix};\ B=\begin{bmatrix}3&-1&3&2\\0&3&-1&4\\2&0&1&3\end{bmatrix}$

19.3. Визначити матрицю C, якщо задано такі умови:

19.3.1) $C=A\cdot B-2B$; $A=\begin{bmatrix}-2&1\\2&-3\end{bmatrix}$, $B=\begin{bmatrix}1&4\\3&-4\end{bmatrix}$.

19.3.2) $C = 3A - B \cdot A$; $A = \begin{bmatrix} 1 & 4 \\ 0 & -2 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 4 & 2 \end{bmatrix}$.

19.3.3) $C = (2A - B) \cdot (3A)$; $A = \begin{bmatrix} -2 & 1 \\ 2 & -3 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 4 \\ 3 & -4 \end{bmatrix}$.

19.3.4) $C = A \cdot B + B$; $A = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 2 & 1 \\ 4 & 2 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$.

19.3.5) $C = 3B - A \cdot B$; $A = \begin{bmatrix} -1 & 1 & 0 \\ 4 & 2 & -1 \\ 2 & 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix}$.

19.3.6) $C = A^2 + B \cdot D$; $A = \begin{bmatrix} -2 & 1 \\ 0 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 4 \end{bmatrix}$.

19.3.7) $C = 2B \cdot D + A^2$; $A = \begin{bmatrix} -2 & 0 \\ 3 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $D = \begin{bmatrix} -1 & 4 \end{bmatrix}$.

19.3.8) $C = A \cdot B - 3B$; $A = \begin{bmatrix} 3 & -1 & 0 \\ 2 & 4 & 1 \\ -1 & 0 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 4 \\ 2 \\ -1 \end{bmatrix}$.

19.3.9) $C = -2B + 3A \cdot B$; $A = \begin{bmatrix} 3 & -1 & 0 \\ 2 & 4 & 1 \\ -1 & 0 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 4 \\ 2 \\ -1 \end{bmatrix}$.

19.3.10) $C = A^2 - B \cdot D$; $A = \begin{bmatrix} 4 & 2 \\ -1 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $D = \begin{bmatrix} -1 & 4 \end{bmatrix}$.

19.3.11) $C = 2B \cdot D - 4A^2$; $A = \begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix}$, $B = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$, $D = \begin{bmatrix} 2 & -1 \end{bmatrix}$.

19.3.12) $C = 4A - 3B \cdot D$; $A = \begin{bmatrix} 1 & 4 \\ 2 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$, $D = \begin{bmatrix} 2 & -1 \end{bmatrix}$.

19.3.13) $C = -\frac{1}{2}D + 3A \cdot B$; $A = \begin{bmatrix} 2 \\ -6 \\ 8 \end{bmatrix}$, $B = \begin{bmatrix} 8 & 1 & 4 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 2 & 3 \\ 2 & 0 & -2 \end{bmatrix}$.

19.3.14) $C = \frac{1}{3}D + A \cdot (4B)$; $A = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}$, $B = \begin{bmatrix} 6 & 4 & 2 \end{bmatrix}$, $D = \begin{bmatrix} 6 & 9 & 18 \\ 3 & -3 & -6 \\ 9 & -9 & 27 \end{bmatrix}$.

19.3.15) $C = A \cdot (-2B) - B$; $A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 3 & 4 \\ -2 & 3 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$.

19.3.16) $C = A^3 - 2B \cdot D$; $A = \begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}$, $B = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 \end{bmatrix}$.

19.3.17) $C = (2E + A) - A \cdot B$; $A = \begin{bmatrix} 1 & 3 & 0 \\ -2 & 4 & 6 \\ 0 & 8 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 4 & 0 & 3 \end{bmatrix}$.

19.3.18) $C = (8E) \cdot A - 3A \cdot B$; $A = \begin{bmatrix} 0 & 1 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 2 & -1 \\ 0 & 4 \end{bmatrix}$.

19.3.19) $C = 3B - 2A \cdot B$; $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 1 \\ 1 & 0 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix}$.

19.3.20) $C = 2B \cdot D - 4A^2$: $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ 3 & 0 \\ 6 & 4 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 1 & 0 \\ 2 & -1 & 2 \end{bmatrix}$.

19.4. Виконати операцію транспонування матриці A.

№	Завдання	№	Завдання	№	Завдання
1)	$A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 7 & 8 \\ 9 & 10 & -1 \end{bmatrix}$	8)	$A = \begin{bmatrix} 1 & 2 & -1 & 4 \\ 4 & 6 & 2 & -3 \\ 3 & 0 & 0 & 1 \end{bmatrix}$	15)	$A = \begin{bmatrix} 1 & 7 & 8 & 9 \\ 10 & 1 & 4 & 7 \\ 2 & 3 & -1 & 4 \end{bmatrix}$
2)	$A = \begin{bmatrix} 1 & 4 & 5 & 7 & -1 \end{bmatrix}$	9)	$A = \begin{bmatrix} 2 & 3 & -1 & 0 \\ -1 & 4 & 5 & 6 \end{bmatrix}$	16)	$A = \begin{bmatrix} 2 & 6 & 1 & 0 \\ 4 & 7 & 3 & -2 \end{bmatrix}$
3)	$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & -2 \\ 1 & -2 & 3 \end{bmatrix}$	10)	$A = \begin{bmatrix} 1 & 4 & 2 & 3 & 0 \\ -1 & -1 & 0 & 2 & -1 \end{bmatrix}$	17)	$A = \begin{bmatrix} 1 & -3 & 4 & 5 \\ 6 & 4 & 0 & 8 \end{bmatrix}$
4)	$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$	11)	$A = \begin{bmatrix} 2 & 1 & 4 \\ 6 & 8 & -1 \\ 4 & 10 & -11 \\ -9 & 4 & 5 \end{bmatrix}$	18)	$A = \begin{bmatrix} 2 & 4 \\ 0 & -1 \\ 6 & 7 \\ 8 & 9 \end{bmatrix}$
5)	$A = \begin{bmatrix} 2 & 1 & 0 & 4 \\ 6 & 7 & -1 & 3 \\ 2 & 2 & -1 & 0 \end{bmatrix}$	12)	$A = \begin{bmatrix} -1 & -2 \\ 3 & 0 \\ 4 & 8 \\ 6 & 4 \end{bmatrix}$	19)	$A = \begin{bmatrix} 2 & 3 \\ 1 & 4 \\ 6 & 0 \\ -1 & 3 \end{bmatrix}$
6)	$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$	13)	$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$	20)	$A = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \\ 4 \end{bmatrix}$
7)	$A = \begin{bmatrix} 0 & 1 & 2 & 3 \\ -1 & 4 & 2 & -7 \\ 8 & 9 & -1 & 4 \end{bmatrix}$	14)	$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$		

19.5. Обчислити визначник

№	Завдання	№	Завдання	№	Завдання
1)	$\det Q = \begin{vmatrix} 2 & -1 & 1 & 0 \\ 0 & 1 & 2 & -1 \\ 3 & -1 & 2 & 3 \\ 3 & 1 & 6 & 1 \end{vmatrix}$	8)	$\det Q = \begin{vmatrix} 2 & 3 & -3 & 4 \\ 2 & 1 & -1 & 2 \\ 6 & 2 & 1 & 0 \\ 2 & 3 & 0 & -5 \end{vmatrix}$	15)	$\det Q = \begin{vmatrix} 3 & -1 & 4 & 2 \\ 5 & 2 & 20 & 1 \\ 0 & 2 & 1 & -3 \\ 6 & -2 & 9 & 8 \end{vmatrix}$
2)	$\det Q = \begin{vmatrix} 5 & 6 & 0 & 0 & 0 \\ 1 & 5 & 6 & 0 & 0 \\ 0 & 1 & 5 & 6 & 0 \\ 0 & 0 & 1 & 5 & 6 \\ 0 & 0 & 0 & 1 & 5 \end{vmatrix}$	9)	$\det Q = \begin{vmatrix} 2 & 1 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 & 1 \\ 1 & 1 & 4 & 1 & 1 \\ 1 & 1 & 1 & 5 & 1 \\ 1 & 1 & 1 & 1 & 6 \end{vmatrix}$	16)	$\det Q = \begin{vmatrix} -1 & -1 & -2 & -2 \\ -1 & -2 & -4 & -8 \\ -1 & -1 & -9 & -8 \\ -1 & 4 & 18 & 16 \end{vmatrix}$
3)	$\det Q = \begin{vmatrix} 2 & -1 & 3 & 4 \\ 0 & -1 & 5 & -3 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 2 \end{vmatrix}$	10)	$\det Q = \begin{vmatrix} 7 & 0 & 0 & 0 \\ 4 & 2 & 0 & 0 \\ 1 & 3 & -1 & 0 \\ 1 & 5 & 3 & 5 \end{vmatrix}$	17)	$\det Q = \begin{vmatrix} 3 & 2 & 7 & 1 \\ 1 & -2 & 3 & 2 \\ 2 & 1 & -5 & 3 \\ 2 & 1 & 0 & 4 \end{vmatrix}$
4)	$\det Q = \begin{vmatrix} 2 & -1 & 2 & 2 \\ 2 & -1 & 1 & 0 \\ 0 & 1 & 2 & -1 \\ 3 & 1 & 6 & 1 \end{vmatrix}$	11)	$\det Q = \begin{vmatrix} 3 & 10 & -5 & -8 \\ 2 & 3 & 5 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 2 & -3 \end{vmatrix}$	18)	$\det Q = \begin{vmatrix} 1 & 1 & 0 & 5 \\ 1 & 4 & 2 & 1 \\ 3 & -1 & 7 & -1 \\ 5 & 2 & 6 & -1 \end{vmatrix}$
5)	$\det Q = \begin{vmatrix} 6 & 2 & 1 & 0 \\ 2 & 3 & -3 & 4 \\ 2 & 1 & -1 & 2 \\ 2 & 3 & 0 & -5 \end{vmatrix}$	12)	$\det Q = \begin{vmatrix} 0 & 8 & 2 & 7 \\ 10 & -8 & 7 & 2 \\ 5 & 4 & 4 & 4 \\ 2 & 0 & 3 & 4 \end{vmatrix}$	19)	$\det Q = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 1 & 2 \\ 3 & -1 & -1 & 0 \\ 1 & 2 & 0 & -5 \end{vmatrix}$
6)	$\det Q = \begin{vmatrix} 1 & 4 & 2 & 3 \\ 0 & 2 & 4 & 6 \\ 1 & 4 & -1 & 3 \\ 2 & 1 & 7 & 3 \end{vmatrix}$	13)	$\det Q = \begin{vmatrix} 3 & -5 & 1 & 4 \\ 2 & -2 & 1 & 7 \\ 5 & 3 & -4 & 2 \\ 1 & 2 & 3 & 4 \end{vmatrix}$	20)	$\det Q = \begin{vmatrix} 1 & 3 & 1 & -4 \\ 0 & 2 & -1 & 1 \\ 2 & 1 & 2 & 1 \\ 1 & -3 & 4 & 3 \end{vmatrix}$
7)	$\det Q = \begin{vmatrix} 2 & -1 & 1 & 0 \\ 3 & 1 & 1 & 2 \\ 0 & 1 & 2 & 3 \\ 3 & 1 & 6 & 1 \end{vmatrix}$	14)	$\det Q = \begin{vmatrix} 3 & -7 & 2 & 1 \\ 4 & 3 & 1 & 0 \\ 2 & -5 & 6 & 2 \\ 4 & 1 & 3 & 4 \end{vmatrix}$		

19.6. Знайти обернену матрицю

№	Завдання	№	Завдання	№	Завдання
1)	$Z = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$	8)	$Z = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 2 & 0 \end{bmatrix}$	15)	$Z = \begin{bmatrix} 2 & 7 & 3 \\ 3 & 9 & 4 \\ 1 & 5 & 3 \end{bmatrix}$
2)	$Z = \begin{bmatrix} 2 & -1 & 3 & 4 \\ 0 & -1 & 5 & -3 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 2 \end{bmatrix}$	9)	$Z = \begin{bmatrix} 3 & -1 & 2 & 3 \\ 2 & -1 & 1 & 0 \\ 0 & 1 & 2 & -1 \\ 3 & 1 & 6 & 1 \end{bmatrix}$	16)	$Z = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

№	Завдання	№	Завдання	№	Завдання
3)	$Z=\begin{bmatrix}2&-1&1\\3&-1&2\\5&-3&4\end{bmatrix}$	10)	$Z=\begin{bmatrix}1&2&0\\-1&1&2\\3&0&1\end{bmatrix}$	17)	$Z=\begin{bmatrix}2&1&-1\\1&0&-2\\1&1&-1\end{bmatrix}$
4)	$Z=\begin{bmatrix}1&2&3\\6&5&4\\3&1&2\end{bmatrix}$	11)	$Z=\begin{bmatrix}2&1&-1\\1&3&2\\4&7&3\end{bmatrix}$	18)	$Z=\begin{bmatrix}5&0&0\\4&5&0\\-10&4&5\end{bmatrix}$
5)	$Z=\begin{bmatrix}1&2&2\\2&1&-2\\2&-2&1\end{bmatrix}$	12)	$Z=\begin{bmatrix}-1&1&2\\3&2&1\\2&3&-1\end{bmatrix}$	19)	$Z=\begin{bmatrix}10&20&-30\\0&10&20\\0&0&10\end{bmatrix}$
6)	$Z=\begin{bmatrix}3&3&-4&-3\\0&6&1&1\\5&4&2&1\\2&3&3&2\end{bmatrix}$	13)	$Z=\begin{bmatrix}0&0&1&-1\\0&3&1&4\\2&7&6&-1\\1&2&2&-1\end{bmatrix}$	20)	$Z=\begin{bmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&0&0\\0&0&1&-1\end{bmatrix}$
7)	$Z=\begin{bmatrix}3&2&1\\4&5&2\\2&1&4\end{bmatrix}$	14)	$Z=\begin{bmatrix}2&5&7\\6&3&4\\5&-2&-3\end{bmatrix}$		

Завдання КР.13.20. Розв'яжіть системи рівнянь:

№	Завдання	№	Завдання
1)	$\begin{cases}8x_1+4x_2-6x_3+18=0,\\-2x_1-4x_3-6x_4+2=0,\\6x_1+4x_2+4x_3+4x_4+14=0,\\4x_1+6x_2+8x_3+8x_4+6=0.\end{cases}$	21)	$\begin{cases}-3x_1+x_2-11x_3+4x_4-1=0,\\-4x_1-3x_2-3x_3+4=0,\\x_1-7x_2-13x_3-2x_4+1=0,\\4x_1-5x_2+3x_3-4x_4=0.\end{cases}$
2)	$\begin{cases}-8x_1+2x_2-2x_4-34=0,\\-6x_1-4x_2-2x_3-2x_4-24=0,\\-10x_1+2x_2+4x_4-68=0,\\-2x_1-6x_2+8x_3-4x_4+36=0.\end{cases}$	22)	$\begin{cases}-5x_1-7x_2-5x_3+6x_4-1=0,\\5x_2+7x_3=0,\\-3x_1+7x_2+x_3+12x_4-9=0,\\-12x_1+5x_2-11x_3-2x_4+14=0.\end{cases}$
3)	$\begin{cases}6x_1-4x_3-4x_4+34=0,\\-10x_1+10x_3-20=0,\\-8x_1-4x_2+2x_3-44=0,\\-2x_1-10x_2+6x_3+4x_4+2=0.\end{cases}$	23)	$\begin{cases}-5x_1-3x_2+5x_3-2x_4+7=0,\\4x_1+7x_2-3x_3+8x_4-12=0,\\x_1-11_2-x_3+6x_4-7=0,\\3x_2+x_3-4x_4+4=0.\end{cases}$
4)	$\begin{cases}8x_1+2x_2-2x_3+32=0,\\2x_1+4x_3+2x_4+14=0.\\2x_1-8x_2-8x_3-6=0.\\-10x_1-4x_2+10x_3+2x_4-24=0.\end{cases}$	24)	$\begin{cases}9x_1-5x_2+5x_3-9=0.\\-2x_1+9x_2-3x_3+2x_4=0,\\x_1+3x_2+3x_3-2x_4+1=0,\\-14x_1-3x_2+3x_3-2x_4+16=0.\end{cases}$
5)	$\begin{cases}2x_1+6x_2+4x_3+16=0,\\-6x_1+8x_2+4x_3+2x_4+34=0,\\-2x_2+6x_3-10x_4+60=0,\\6x_1-10x_2+2x_3-81x_4+78=0.\end{cases}$	25)	$\begin{cases}3x_1-3x_2+x_3-2x_4+2=0,\\-9x_2-5x_3+8x_4+14=0,\\-x_1-3x_2-3x_3-6x_4+6=0,\\-3x_2-9x_3-4x_4+12=0.\end{cases}$

№	Завдання	№	Завдання
6)	$\begin{cases} 6x_1 - 2x_2 + 10x_3 + 4x_4 + 46 = 0, \\ -6x_1 - 4x_2 + 10x_3 + 10x_4 - 36 = 0, \\ x_3 - 4x_4 + 19 = 0, \\ 8x_2 - 4x_3 + 10x_4 - 60 = 0. \end{cases}$	26)	$\begin{cases} -3x_1 + 9x_2 + 7x_3 - 4x_4 - 16 = 0, \\ -6x_1 + 5x_2 - 9x_3 + 8x_4 + 4 = 0, \\ -7x_1 - 9x_2 - 7x_3 + 4x_4 + 16 = 0, \\ 7x_2 + x_3 - 8 = 0. \end{cases}$
7)	$\begin{cases} 6x_1 + 8x_3 - 6x_4 + 2 = 0, \\ 10x_1 - 10x_2 - 2x_3 - 8x_4 - 42 = 0, \\ 4x_1 - 2x_2 - 2x_3 + 10x_4 - 12 = 0, \\ -4x_1 - 2x_2 - 2x_3 - 4 = 0. \end{cases}$	27)	$\begin{cases} -9x_1 - x_2 - 5x_3 + 2x_4 + 6 = 0, \\ x_2 - 7x_3 - 2x_4 + 6 = 0, \\ 7x_1 + 3x_2 - 5x_3 + 12x_4 + 2 = 0, \\ -4x_1 - 5x_2 - 3x_3 + 10x_4 + 8 = 0. \end{cases}$
8)	$\begin{cases} -4x_1 + 6x_2 - 4x_3 - 6x_4 + 18 = 0, \\ 4x_1 + 10x_2 - 8x_3 + 2x_4 + 18 = 0, \\ 2x_2 - 6x_3 + 6x_4 = 0, \\ -2x_3 - 2x_4 - 2 = 0. \end{cases}$	28)	$\begin{cases} 3x_1 + 3x_2 - x_3 - 8x_4 - 2 = 0, \\ 2x_1 + 7x_2 - 5x_3 + 8x_4 - 2 = 0, \\ 7x_1 - 7x_2 - x_3 + 2x_4 + 8 = 0, \\ -5x_2 - 7x_3 + 8x_4 + 12 = 0. \end{cases}$
9)	$\begin{cases} 4x_1 + 4x_2 + 4x_3 + 8x_4 + 12 = 0, \\ -8x_2 - 2x_3 + 6x_4 - 26 = 0, \\ -2x_1 + 2x_2 - 8x_3 + 8x_4 = 0, \\ -8x_2 + 2x_3 - 6x_4 - 22 = 0. \end{cases}$	29)	$\begin{cases} -x_1 + 5x_2 - x_3 + 4x_4 - 4 = 0, \\ -2x_1 + 3x_2 - 11x_3 + 8 = 0, \\ -3x_1 + 5x_2 - 11x_3 + 12x_4 + 6 = 0, \\ 9x_2 - 11x_3 + 2 = 0. \end{cases}$
10)	$\begin{cases} 2x_1 + 8x_2 + 6x_3 + 28 = 0, \\ -4x_2 + 6x_3 + 8x_4 - 6 = 0, \\ -8x_1 + 4x_2 + 10x_4 + 20 = 0, \\ -6x_1 - 2x_2 - 4x_3 + 2x_4 - 4 = 0. \end{cases}$	30)	$\begin{cases} -x_1 + 9x_2 + 7x_3 - 8x_4 - 16 = 0, \\ -4x_1 + 9x_2 - x_3 - 8 = 0, \\ 5x_1 - 3x_2 + 5x_3 - 2 = 0, \\ -12x_1 - 5x_2 - 5x_3 - 4x_4 + 10 = 0. \end{cases}$
11)	$\begin{cases} -4x_1 - 8x_3 - 4x_4 - 4 = 0, \\ 6x_1 - 2x_2 - 6x_3 - 6x_4 - 18 = 0, \\ -4x_1 + 2x_2 - 8x_3 - 8x_4 + 2 = 0, \\ -8x_2 - 6x_3 - 8x_4 - 30 = 0. \end{cases}$	31)	$\begin{cases} 7x_1 + 5x_2 - 9x_3 - 6x_4 - 80 = 0, \\ -8x_1 - 5x_2 - x_3 + 6x_4 + 35 = 0, \\ 3x_1 - 5x_2 + 3x_3 = 0, \\ -6x_1 - 3x_2 - 7x_3 - 5 = 0. \end{cases}$
12)	$\begin{cases} -2x_1 - 2x_2 + 2x_4 - 4 = 0, \\ -8x_2 - 6x_3 - 8x_4 - 30 = 0, \\ -4x_1 - 10x_2 - 10x_3 + 10x_4 - 36 = 0, \\ 10x_1 + 4x_2 + 4x_3 - 4x_4 + 6 = 0. \end{cases}$	32)	$\begin{cases} 5x_1 - 9x_2 - 9x_3 - 70 = 0, \\ -6x_1 - 5x_2 - 9x_3 - 15 = 0, \\ -5x_1 - 3x_2 - 3x_3 - 8x_4 + 10 = 0, \\ -2x_1 - 3x_2 - x_3 + 2x_4 + 5 = 0. \end{cases}$
13)	$\begin{cases} -10x_1 + 10x_4 - 20 = 0, \\ 6x_1 + 2x_2 - 6x_3 + 2x_4 - 24 = 0, \\ 2x_1 + 6x_2 + 2x_3 + 10x_4 - 28 = 0, \\ 4x_1 + 6x_2 + 4x_3 - 4x_4 + 16 = 0. \end{cases}$	33)	$\begin{cases} 9x_1 + 3x_2 - 9x_3 + 2x_4 - 90 = 0, \\ 2x_1 + 3x_2 - 7x_3 - 45 = 0, \\ -7x_1 - 7x_2 - 7x_3 - 6x_4 = 0, \\ x_2 + 5x_3 + 2x_4 + 25 = 0. \end{cases}$
14)	$\begin{cases} -9x_1 - 9x_2 - 5x_3 + 10x_4 - 31 = 0, \\ -4x_1 + 7x_2 + 5x_3 + 14 = 0, \\ 9x_1 - 5x_2 + x_3 - 7 = 0, \\ -11x_2 - 13x_3 + 2x_4 - 32 = 0. \end{cases}$	34)	$\begin{cases} x_1 - 5x_2 + 5x_3 + 20 = 0, \\ -6x_1 - 9x_2 - 5x_3 + 10x_4 + 5 = 0, \\ 7x_1 + 5x_2 - x_3 - 2x_4 - 40 = 0, \\ 3x_2 - x_3 - 6x_4 - 5 = 0 \end{cases}$

№	Завдання	№	Завдання
15)	$\begin{cases} -5x_1 + x_2 - 7x_3 + 8x_4 - 33 = 0, \\ 9x_2 - 3x_3 - 4x_4 + 6 = 0, \\ -3x_1 + 7x_2 + 5x_3 + 13 = 0, \\ -7x_2 - 11x_3 - 4x_4 - 10 = 0. \end{cases}$	35)	$\begin{cases} 3x_1 - 5x_2 + 5x_3 - 6x_4 + 10 = 0, \\ 7x_2 - 5x_3 - 2x_4 - 25 = 0, \\ -9x_1 - 7x_2 - 13x_3 + 8x_4 - 20 = 0, \\ 4x_1 - 5x_2 - 3x_3 + 2x_4 - 35 = 0. \end{cases}$
16)	$\begin{cases} 3x_1 - 11x_2 + 5x_3 + 4x_4 - 5 = 0, \\ -8x_1 - 5x_2 - 3x_3 + 10x_4 - 28 = 0, \\ 3x_1 - x_2 + 5x_3 + 7 = 0, \\ -6x_1 - 11x_2 + 3x_3 + 12 = 0. \end{cases}$	36)	$\begin{cases} -5x_1 - 9x_2 + 3x_3 + 12x_4 + 40 = 0, \\ -12x_1 - 3x_2 - 11x_3 - 2x_4 + 5 = 0, \\ 5x_1 - 3x_2 - 5x_3 - 50 = 0, \\ -7x_2 + x_3 + 5 = 0. \end{cases}$
17)	$\begin{cases} 5x_1 - 7x_2 - 11x_3 + 8x_4 - 51 = 0, \\ -6x_1 - 9x_2 - 3x_3 - 2x_4 + 6 = 0, \\ 3x_1 + 5x_2 - 5x_3 - 6x_4 + 5 = 0, \\ -5x_2 - 13x_3 + 4x_4 - 38 = 0. \end{cases}$	37)	$\begin{cases} 2x + 3y + 4z = 1 \\ x + 8y - 17z = -4 \\ 3x + 32y + 41z = 7 \end{cases}$
18)	5) $\begin{cases} x_1 + 7x_2 + 3x_3 + 4x_4 - 7 = 0, \\ 5x_2 - 9x_3 - 4x_4 - 6 = 0, \\ 9x_1 - 9x_2 - 9x_3 + 8x_4 - 51 = 0, \\ -10x_1 - 7x_2 - 7x_3 - 6x_4 + 14 = 0. \end{cases}$	38)	$\begin{cases} 2x + 3y + 4z = 1 \\ 2x + 32y - 17z = 6 \\ 3x + 32y + 41z = 7 \end{cases}$
19)	$\begin{cases} 7x_1 + x_2 + 3x_3 + 4x_4 - 11 = 0, \\ 3x_2 + 5x_3 - 6x_4 + 6 = 0, \\ x_1 + 9x_2 + 5x_3 - 1 = 0, \\ 5x_2 - 3x_3 + 10x_4 - 10 = 0. \end{cases}$	39)	$\begin{cases} 2x_1 + 3x_2 + 4x_3 + 5x_4 = 1 \\ x_1 + 32x_2 + 7x_3 + 5x_4 = 6 \\ x_1 - 32x_2 + 49x_3 + 5x_4 = 2 \end{cases}$
20)	$\begin{cases} -3x_1 - 11x_2 - 13x_3 + 12x_4 - 9 = 0, \\ 7x_2 - 9x_3 + 6x_4 - 6 = 0, \\ 5x_1 - 3x_2 - x_3 + 10x_4 - 15 = 0, \\ -12x_1 - x_2 - 11x_3 - 2x_4 + 14 = 0. \end{cases}$	40)	$\begin{cases} 2x + 3y + 4z - u + t = -2 \\ x + 31y + 6z + 7u - 5t = 6 \\ 3x + 37y + 41z - 9u + 11t = 3 \\ x - y + 17z - u + 13t = 13 \\ 11x + 13y + 17z + 19u + 23t = 32 \end{cases}$

Правильність отриманих розв'язків легко перевірити. В результаті розв'язування вибраної СЛАР трьома різними способами мають бути отримані три однакових розв'язки. Приведені розв'язки систем 1)-36):

1)-6) $X = (-5,1,-3,4)^T$; 7)-12) $X = (1,-3,-1,0)^T$; 13)-18) $X = (1,0,-2,3)^T$;

19)-24) $X = (1,0,0,1)^T$; 25)-30) $X = (0,1,1,0)^T$; 30)-36) $X = (5,0,-5,0)^T$.

ЛІТЕРАТУРА

1. Andrunyk V. Electronic Content Commerce System Development / Vasyl Andrunyk, Lyubomyr Chyrun, Victoria Vysotska // Матеріали XIII Міжнародної науково–технічної конференції "Досвід розробки та застосування приладо–технологічних САПР в мікроелектроніці": CADSM 2015. – Львів: Видавництво Львівської політехніки, 2015. – 24–27 February, 2015, Polyana–Svalyava (Zakarpattya), Ukraine. – C. 368–434.
2. Andrunyk V. The Peculiarities of Electronic Digest Formation / Vasyl Andrunyk, Victoria Vysotska, Lyubomyr Chyrun // Computer Science and Information Technologies: Proc. of the IX–th Int. Conf. CSIT'2014, 18–22 November, 2014, Lviv, Ukraine. – Lviv: Publishing Lviv Polytechnic, 2014.– P.25–28.
3. Anscombe P.J. Qnick Analysis Methods for Random Balanse Screlmlng Experiments / Technometrics. – Vol. 1(2). – 1959.
4. Berko A. Content Formation Method in the Electronic Content Commerce Systems / Andriy Berko, Victoria Vysotska, Lyubomyr Chyrun // 6th International Conference of Young Scientists "Computer Science & Engineering" (CSE–2013). – Lviv, Ukraine, November 21–23, 2013. – P.174–179.
5. Berko A. Features of information resources processing in electronic content commerce [Electronic resource] / Andriy Berko, Victoria Vysotska, Lyubomyr Chyrun // Applied Computer Science. ACS journal. – Volume 10, Number 2. – Poland, 2014. – ISSN 2353–6977 (Online), ISSN 1895–3735 (Print). – PP. 5–19 [Online]. – Access mode: www.acs.pollub.pl, http://www.acs.pollub.pl/index.php/–current–issue/vol–10–no–122014.html, http://www.acs.pollub.pl/pdf/v10n2/1.pdf.
6. Berko A. Functionally logistic model of commercial content processing / Andriy Berko, Victoria Vysotska, Lyubomyr Chyrun // Комп'ютерні системи проектування. Теорія і практика. Вісник Національного університету "Львівська політехніка". – № 777. – Львів 2013. – С. 30–38.
7. Berko A. Linguistic Analysis for the Textual Commercial Content / Andriy Berko, Victoria Vysotska, Lyubomyr Chyrun // Computer Science and Information Technologies: Proc. of the IX–th Int. Conf. CSIT'2014, 18–22 November, 2014, Lviv, Ukraine..– Lviv: Publishing Lviv Polytechnic, 2014.– P.11–14.
8. Berko A. Logistic Functionally Model of Commercial Content Processing / Andriy Berko, Victoria Vysotska, Lyubomyr Chyrun // Computer Science and Information Technologies: Proc. of the VIII–th Int. Conf. CSIT'2013, 11–16 November, 2013, Lviv, Ukraine..– Lviv: Publishing Lviv Polytechnic, 2013.– P. 36–39.
9. Box G.E.P. Multi-factor experimental designs for exploring response surfaces / G.E.P. Box, J.S. Hunter // Annals of Mathematical Statistics, 28, 195. – 1957. – 315 p.
10. Box G.E.P. On the experimental attainment of optimum conditions / G.E.P. Box, K.B. Wilson // J. Royal Statistical Soc. (B) 13,1. 1951. – 200 h.
11. Box G.E.P. Experimental designs for the exploration and exploitation of response surface / G.E.P. Box // "Experimental design in Industry" Chapman and Halle Ltd. – London, 1958. – 100 p.

12.Box G.E.P. A basis for the selection of a response surface design / G.E.P. Box. N.R. Drpar // International Jornal of Abstracts of Statistical; Method in Indusstry, 6,91. – 1959. – 200 p.
13.Brandon D.B. Developing Mathematical Models far Computer Control / D.B. Brandon // ISA – Jornal, 7, 1959. – P. 13-21.
14.Box G.E.P. The exploration and exploitation of response surface / U.G.E.P. Box // Some general considerations and examples. Biometrics, 10, 16. – 1954. – 95 p.
15.Box G.E.P. The exploration and exploitation of response surface / U.G.E.P. Box, P.V. Youle // Biometrics, 11, 287. – 1955. – 214 p.
16.Brown Leek K.A. Приложение методов корреляции к обработке производственных данных / K.A. Brown Leek // Ind. Eng. Chem., 43, 9. – 1951. – P. 84-111.
17.Budne A.T. Industrial Anality Control / A.T. Budne. – 5, 11, 16. – 1959.
18.Chew (ed) V. Experimental designs in industry / V. (ed) Chew // John Williy and Sons Inc. – New York, 1958. – 95 p.
19.Cohran W.G. Experimental designs / W.G. Cohran, O.M. Box // John Williy and Sons Inc. – New York, 1957. – 154 p.
20.Chyrun L. Electronic content commerce system development [Electronic resource] / Lyubomyr Chyrun, Victoria Vysotska, Vasyl Andrunyk // MEST Journal. – Vol.4 No.2. – 2015. – PP. 120–138 [Online]. – ISSN 2334–7058 (Online). – Access mode: http://mest.meste.org/MEST_Najava/VI_Chyrun.pdf.
21.Chyrun L. Information resources analysis in electronic content commerce systems [Electronic resource] / Lyubomyr Chyrun, Victoria Vysotska, Rostik Laba // Applied Computer Science. ACS journal. – Volume 11, Number 2. – Poland, 2016. – ISSN 2353–6977 (Online), ISSN 1895–3735 (Print) – PP. 48–66 [Online]. – Access mode: www.acs.pollub.pl, http://www.acs.pollub.pl/pdf/v12n1/5.pdf, http://www.acs.pollub.pl/index.php/–current–issue/vol–12–no–1–2016.html.
22.Chyrun L. Informational resources processing intellectual systems with textual commercial content linguistic analysis usage constructional means and tools development / L. Chyrun, V. Vysotska, I. Kozak // Econtechmod : an international quarterly journal on economics in technology, new technologies and modelling processes. – Lublin ; Rzeszow. 2016. – Volum 5, nomber 2. – P. 85–94.
23.Chyrun L.V. Method for solving systems of linear algebraic equations with tree–diagonal matrix / L.V. Chyrun // Тези XVI Відкритої наук.–техн. конф. молодих науковців і спеціалістів Фізико–механічного інституту ім. Г.В.Карпенка НАН України. КМН–2001. – Львів, 2001. – С. 64.
24.Chyrun Lyubomyr. Specifics Informational Resources Processing for Textual Content Linguistic Analysis / Lyubomyr Chyrun, Victoria Vysotska, Vasyl Lytvyn // Proceeding of XIIth International Conference of Perspective Technologies and Methods in MEMS Design, MEMSTECH 2016. – 20–24 April. 2016. Lvi–Polyana, Ukraine. – Lviv Politechnic Publishing House. – P. 214–219.
25.Chyrun Lyubomyr. Using content analysis of textual information in electronic commerce: conference proceedings. October 14–16, 2010 / Lyubomyr Chyrun, Victoria Vysotska // The 5th International Scientific and Technical Conference "Computer Sciences and Information Technologies" (CSIT'2010) which will be held

October 14–16, 2010 at Lviv Polytechnic National University (Lviv, Ukraine). – Lviv 2010. – Стор.80–82.

26. Classification Methods of Text Documents Using Ontology Based Approach [Electronic resource] / [Vasyl Lytvyn, Victoria Vysotska, Oleh Veres, Ihor Rishnyak, Halya Rishnyak] // Advances in Intelligent Systems and Computing. Advances in Intelligent Systems and Computing 512. Natalya Shakhovska Editor. Selected Papers from the International Conference on Computer Science and Information Technologies, CSIT 2016, September 6–10 Lviv, Ukraine. – ISSN 2194–5357 ISSN 2194–5365 (electronic). – ISBN 978–3–319–45990–5 ISBN 978–3–319–45991–2 (eBook). – DOI 10.1007/978–3–319–45991–2. – Library of Congress Control Number: 2016950408. – Springer International Publishing AG 2017. – P. 229–240. – Access mode: http://www.springer.com/us/book/9783319459905.
27. Content Linguistic Analysis Methods for Textual Documents Classification / [Vasyl Lytvyn, Victoria Vysotska, Oleh Veres, Ihor Rishnyak, Halya Rishnyak] // Computer Science and Information Technologies: Proc. of the XI–th Int. Conf. CSIT'2016, 6–10 September, 2016, Lviv, Ukraine. – Lviv: Lviv Polytechnic Publishing House, 2016.– P. 190–192.
28. Continuous and smooth minimax spline–approximation of sensor temperature characteristic and its sensitivity / [Andrunyk Vasyl, Petro Malachivskyy, Yaropolk Pizyur, Vasyl Yatsuk] // Pomiary, automatyka, kontrola (PAK). – 2007. – Vol. 53. – P. 617–620.
29. D'Agunno B. CHPACK: A package for manipulation Chebyshev approximation / B. D'Agunno, A. Nobile // Comput. Commun. – 1983. – 29, N4. – P.361 – 374.
30. Davis O. The design and analysis of Industrial Experiments / O. Davis. – New York, 1956. – 120 p.
31. Distance Learning Method for Modern Youth Promotion and Involvement in Independent Scientific Researches [Electronic resource] / [Vasyl Lytvyn, Victoria Vysotska, Liliya Chyrun, Lyubomyr Chyrun] // DATA STREAM MINING & PROCESSING. Proceedings of the 2016 IEEE First International Conference on Data Stream Mining & Processing (DSMP). – August 23–27, 2016. – Lviv, Ukraine. – P. 269–274.
32. Everitt B. A handbook of statistical analyses using R / B. Everitt, T. Hothorn. – 2–nd ed. – Chapman and HALL/CRC, 2009. – 376 p.
33. Features of the Content–Analysis Method in Processing Online Newspaper Articles / [Iaroslav Kis, Victoria Vysotska, Liliya Chyrun, Vasyl Foltovych] // Computer Science and Information Technologies: Proc. of the IX–th Int. Conf. CSIT'2014, 18–22 November, 2014, Lviv, Ukraine.– Lviv: Publishing Lviv Polytechnic, 2014.– P. 39–42.
34. Geddes K.O. A Package for Numerical Approximation / K.O. Geddes. – Maple Tech, 10. – 1993. – P. 28–36.
35. Goodman T.P. Determination of Sistem Charactenstics from normal Operating Records / T.P. Goodman, I.B. Reswick // Trans. of the ASME. 2. – 1956. – 214 p.
36. Kabacoff Robert I. Quick–R [Electronic resource] / Robert I. Kabacoff. – Access mode: http://www.statmethods.net/index.html.
37. Lytvyn V. Designing Architecture of Electronic Content Commerce System / Vasyl Lytvyn, Victoria Vysotska // Computer Science and Information Technologies: Proc.

of the X–th Int. Conf. CSIT'2015, 14–17 September, 2015, Lviv, Ukraine..– Lviv: Publishing Lviv Polytechnic, 2015.– P.115–119.

38. Methods based on ontologies for information resources processing : Monograph [Electronic resource] / [Vasyl Lytvyn, Victoria Vysotska, Lyubomyr Chyrun, Dmytro Dosyn] // LAP Lambert Academic Publishing. Saarbrücken, Germany. – ISBN–13: 978–3–659–89905–8, ISBN–10: 3659899054, EAN: 9783659899058. – 2016. – 324 c. – Access mode: https://www.lap–publishing.com/catalog/details/store/gb/book/978–3–659–89905–8/methods–based–on–ontologies–for–information–resources–processing?locale=gb.

39. Modelling of semantics of natural language sentences using generative grammars / [Tetiana Shestakevych, Victoria Vysotska, Lyubomyr Chyrun, Liliya Chyrun] // Computer Science and Information Technologies: Proc. of the IX–th Int. Conf. CSIT'2014, 18–22 November, 2014, Lviv, Ukraine..– Lviv: Publishing Lviv Polytechnic, 2014.– P. 19–22.

40. Netlib. Class K2. Minimax (L–infinity) approximation [Electronic resource]. – Access mode: http://gams.nist.gov/serve.cgi/Class/K2.

41. R Site Search [Electronic resource]. – Access mode : http://finzi.psych.upenn.edu/nmz.html.

42. Rtips. Revival 2014! [Electronic resource]. – Access mode : http://pj.freefaculty.org/R/Rtips.html.

43. Satterthwaite F.E. Random Balance Experimentation / F.E. Satterthwaite. – 230 p.

44. Satterthwait. Discussion of the Papers of Messers / Satterthwait, Budue // Technometrics. – Vol.1 (2). – 1959. – 150 p.

45. Schrage R.W. Optimizing a catalytic cracking operation by the method of steepest ascent / R.W. Schrage // Operation Research, 6, 498. – 1958. – 98 p.

46. Shakhovska Natalya. Features of E–Learning Realization Using Virtual Research Laboratory / Natalya Shakhovska, Victoria Vysotska, Lyubomyr Chyrun // Computer Science and Information Technologies: Proc. of the XI–th Int. Conf. CSIT'2016, 6–10 September, 2016, Lviv, Ukraine.– Lviv: Lviv Polytechnic Publishing House, 2016.– P. 143–148.

47. Shakhovska Natalya. Intelligent Systems Design of Distance Learning Realization for Modern Youth Promotion and Involvement in Independent Scientific Researches [Electronic resource] / Natalya Shakhovska, Victoria Vysotska, Lyubomyr Chyrun // Advances in Intelligent Systems and Computing. Advances in Intelligent Systems and Computing 512. Natalya Shakhovska Editor. Selected Papers from the International Conference on Computer Science and Information Technologies, CSIT 2016, September 6–10 Lviv, Ukraine. – ISSN 2194–5357 ISSN 2194–5365 (electronic). – ISBN 978–3–319–45990–5 ISBN 978–3–319–45991–2 (eBook). – DOI 10.1007/978–3–319–45991–2. – Library of Congress Control Number: 2016950408. – Springer International Publishing AG 2017. – PP. 175–198. – Access mode: http://www.springer.com/us/book/9783319459905.

48. Shewell C.T. Paper studies in, catalytic cracking / C.T. Shewell // Trans. 10-h Ann. Conv. American Soc. For Quality Control, 1956, 1. – 60 p.

49. Shumway R.H. Time series analyses and its applications: With R examples / R.H. Shumway, D.S. Stoffer. – 3–rd ed. – New York : Springer, 2011. – 596 p.

50. Solution of Overdetermined Linear System in the Chebychev Norm [Electronic resource]. – Access mode: http://wwwasdoc.web.cern.ch/wwwasdoc/shortwrupsdir/e222/top.html.
51. Statistics with R [Electronic resource]. – Access mode : http://zoonek2.free.fr/UNIX/48_R/all.html.
52. The Comprehensive R Archive Network [Electronic resource]. – Access mode: http://cran.r–project.org.
53. The method of formation of the status of personality understanding based on the content analysis [Electronic resource] / [V. Lytvyn, P. Pukach, I. Bobyk, V. Vysotska] // Eastern–European Journal of Enterprise Technologies. – ISSN 1729–3774. – (№5/2(83), 2016). – P. 4–12. – Access mode: http://journals.uran.ua/eejet/issue/archive.
54. Vysotska V. Analysis Features of Information Resources Processing / Victoria Vysotska, Lyubomyr Chyrun // Computer Science and Information Technologies: Proc. of the X–th Int. Conf. CSIT'2015, 14–17 September, 2015, Lviv, Ukraine..– Lviv: Publishing Lviv Polytechnic, 2015.– P.124–128.
55. Vysotska V. Analysis of business processes in electronic content–commerce systems [Electronic resource] / V. Vysotska, L. Chyrun, P. Kozlov // ECONTECHMOD. An international quarterly journal on economics in technology, new technologies and modeling processes. – Polish Academy of Sciences, University of Engineering and Economics in Rzeszyw, Lviv Polytechnic National University, University of Life Sciences in Lublin, Faculty of Production Engineering. – Vol. 5, No 1. – LUBLIN–RZESZYW,2016. – pp. 111–125. – ISSN 2084–5715. – Access mode: http://www.pan–ol.lublin.pl/wydawnictwa/Econtechmod.html.
56. Vysotska V. Analytical methods for commercial web content processing of information resource in electronic business systems [Electronic resource] / Victoria Vysotska, Lyubomyr Chyrun // MEST Journal. – Vol.2 No.1. – PP. 57–70 [Online]. – ISSN 2334–7058. – Access mode: http://www.mest.meste.org/MEST_Najava_clanaka.html, http://mest.meste.org/MEST_Najava/IV_vysotska.pdf.
57. Vysotska V. Analytical methods for commercial web content processing of information resource in electronic business systems [Electronic resource] / Victoria Vysotska, Lyubomyr Chyrun // MEST Journal (Management Education Science & Society Technologie). – Vol.2 No.2. – PP. 285–300 [Online]. – ISSN 2334–7171, ISSN 2334–7058 (Online). DOI 10.12709/issn.2334–7058. This issue: DOI 10.12709/mest.02.02.02.0. – Access mode: http://mest.meste.org/R4.html, http://mest.meste.org/MEST_2_2014/4_29.pdf, http://www.meste.org/mest/Archive/MEST_II_2_2.pdf.
58. Vysotska V. Commercial Content Support Method in the Electronic Business Systems / Victoria Vysotska, Lyubomyr Chyrun, Liliya Chyrun // Computer Science and Information Technologies: Proc. of the VIII–th Int. Conf. CSIT'2013, 11–16 November, 2013, Lviv, Ukraine..– Lviv: Publishing Lviv Polytechnic, 2013.– P.2–5.
59. Vysotska V. Comprehensive method of commercial content support in the electronic business systems / Victoria Vysotska, Lyubomyr Chyrun, Liliya Chyrun //

Комп'ютерні системи проектування. Теорія і практика, Вісник Національного університету "Львівська політехніка". – № 777. – Львів 2013. – С. 21–30.

60. Vysotska V. Designing features of architecture for e–commerce systems [Electronic resource] / Victoria Vysotska, Lyubomyr Chyrun // MEST Journal (Management Education Science & Society Technologie). – Vol.2 No.1. – PP. 57–70 [Online]. – ISSN 2334–7171, ISSN 2334–7058 (Online), DOI 10.12709/issn.2334–7058. This issue: DOI 10.12709/mest.02.02.02.0. – Access mode: http://mest.meste.org/R3.html, http://www.meste.org/mest/Archive/MEST_II_2_1.pdf, http://mest.meste.org/MEST_1_2014/_06.pdf, http://mest.meste.org/MEST_1_2014/K1_eng.html.
61. Vysotska V. Features of Text Categorization of Commercial Content / Victoria Vysotska // Computer Science and Information Technologies: Proc. of the IX–th Int. Conf. CSIT'2014, 18–22 November, 2014, Lviv, Ukraine. – Lviv: Publishing Lviv Polytechnic, 2014.– P.5–8.
62. Vysotska V. Features of the content–analysis method for text categorization of commercial content in processing online newspaper articles [Electronic resource] / Victoria Vysotska, Lyubomyr Chyrun // Applied Computer Science. ACS journal. – Volume 11, Number 1. – Poland, 2015. – ISSN 2353–6977 (Online), ISSN 1895–3735 (Print) – PP. 5–19 [Online]. – Access mode: www.acs.pollub.pl, http://www.acs.pollub.pl/index.php/–current–issue/applied–computer–science–volume–11–number–1–2015.html, http://www.acs.pollub.pl/pdf/v11n1/2.pdf.
63. Vysotska V. Generative regular grammars application to modeling the semantics of sentences in natural language / Victoria Vysotska // Комп'ютерні системи проектування. Теорія і практика, Вісник Національного університету "Львівська політехніка" № 808, – Львів 2014 – Стор.43–56.
64. Vysotska V. Life Cycle Model of Commercial Content Processing in Electronic Commerce System [Electronic resource] / Victoria Vysotska, Lyubomyr Chyrun // Computational Problems of Electrical Engineering. Founder and Publisher Lviv Polytechnic National University. – Volume 3, Number 2. – Lviv, 2014. – ISSN 2224–0977. PP.118–122 [Online]. – Access mode: http://vlp.com.ua/files/special/24_137.pdf.
65. Vysotska V. Linguistic analysis and modelling semantics of textual content for digest formation [Electronic resource] / Victoria Vysotska, Lyubomyr Chyrun // MEST Journal (Management Education Science & Society Technologie). – Vol.3 No.1. – PP. 127–148 [Online]. – ISSN 2334–7171, ISSN 2334–7058 (Online), DOI 10.12709/issn.2334–7058. This issue: DOI 10.12709/mest.02.02.02.0. – Access mode: http://mest.meste.org/MEST_1_2015/5_15.pdf.
66. Vysotska V. Linguistic Analysis of Textual Commercial Content for Information Resources Processing / Victoria Vysotska // Proceedings of the XIIIth International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET'2016). – February 23–26, 2016. Lviv–Slavske, Ukraine. – P. 709–713.
67. Vysotska V. Methods and means of processing information resources in electronic content commerce systems [Electronic resource] / Victoria Vysotska, Lyubomyr Chyrun // Applied Computer Science. ACS journal. – Volume 11, Number 2. – Poland,

2015. – ISSN 2353–6977 (Online), ISSN 1895–3735 (Print) – PP. 68–85 [Online]. – Access mode: www.acs.pollub.pl, http://www.acs.pollub.pl/pdf/v11n2/7.pdf, http://www.acs.pollub.pl/index.php/archival–issues/vol–11–no–2–2015.html.

68. Vysotska V. Methods of Information Resources Processing in Electronic Content Commerce Systems / Victoria Vysotska, Lyubomyr Chyrun // Матеріали XIII Міжнародної науково–технічної конференції "Досвід розробки та застосування приладо–технологічних САПР в мікроелектроніці": CADSM 2015. – Львів: Видавництво Львівської політехніки, 2015. – 24–27 February, 2015, Polyana–Svalyava (Zakarpattya), Ukraine – C. 328– 332.
69. Vysotska V. Process Analysis in Electronic Content Commerce System / Victoria Vysotska, Ruslan Hasko, Volodymyr Kuchkovskiy // Computer Science and Information Technologies: Proc. of the X–th Int. Conf. CSIT'2015, 14–17 September, 2015, Lviv, Ukraine..– Lviv: Publishing Lviv Polytechnic, 2015.– P. 120–123.
70. Vysotska V. Set–theoretic models and unified methods of information resources processing in e–business systems [Electronic resource] / Victoria Vysotska, Lyubomyr Chyrun // Applied Computer Science. ACS journal. – Volume 10, Number 3. – Poland, 2014. – ISSN 2353–6977 (Online), ISSN 1895–3735 (Print). – PP. 5–22 [Online]. – Access mode: www.acs.pollub.pl, www.acs.pollub.pl, http://www.acs.pollub.pl/index.php/–current–issue/vol–10–no–3–2014.html, http://www.acs.pollub.pl/pdf/v10n3/1.pdf.
71. Vysotska V. Technology of Commercial Web–Resource Processing / Kateryna Alieksieieva, Andriy Berko, Victoria Vysotska // Матеріали XIII Міжнародної науково–технічної конференції "Досвід розробки та застосування приладо–технологічних САПР в мікроелектроніці": CADSM 2015. – Львів: Видавництво Львівської політехніки, 2015. – 24–27 February, 2015, Polyana–Svalyava (Zakarpattya), Ukraine – C. 340–344.
72. Vysotska V. The Means Structure of Information Resources Processing in Electronic Content Commerce Systems [Electronic resource] / Victoria Vysotska, Lyubomyr Chyrun // Journal of Information Sciences and Computing Technologies (JISCT). – Vol 3, No 3 (2015). – Punjab, India, 2015. – ISSN: 2394–9066. – P.241–248. – Access mode: http://scitecresearch.com/journals/index.php/jisct/article/view/134/119.
73. Vysotska V. Unified Methods of Processing Information Resources in Electronic Content Commerce Systems / Victoria Vysotska, Lyubomyr Chyrun, Liliya Chyrun // Computer Science and Information Technologies (CSIT'2012): Proc. of the VII–th Int. Conf., (Lviv, 20–24 November, 2012).– Lviv: Publishing House Vezha&Co, 2012.– P.190–194.
74. Vysotska V. Uniform Method of Operative Content Management in the Electronic Content Commerce Systems / Victoria Vysotska, Lyubomyr Chyrun, Liliya Chyrun // 6th International Conference of Young Scientists "Computer Science & Engineering" (CSE–2013). – Lviv, Ukraine, November 21–23, 2013. – P.180–185.
75. Vysotska Victoria. Analysis and evaluation of risks in electronic commerce / Victoria Vysotska, Ihor Rishnyak, Lyubomyr Chyrun // CAD Systems in Microelectronics. CADSM'07, 9th International Conference. – The Experience of Designing and Applications of CAD Systems in Microelectronics. – Lviv, 24 February 2007. – P. 332–333.

76. Vysotska Victoria. Commercial Web Content Lifecycle Model: conference proceedings, November 16–19, 2011 / Victoria Vysotska, Lubomyr Chyrun, Lilya Chyrun // The 6th International Scientific and Technical Conference "Computer Sciences and Information Technologies" (CSIT'2011) which will be held November 16–19, 2011 at Lviv Polytechnic National University (Lviv, Ukraine) – Lviv 2011. – Стор. 160–163.
77. Vysotska Victoria. Design and analysis features of generalized electronic content–commerce systems architecture [Electronic resource] / Victoria Vysotska, Lyubomyr Chyrun, Pavlo Kozlov // Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska 2016. – Redakcja IAPGOS, Poland, 2016. – 6(2). – ICV 81,34. – ICID: 1201317.– p–ISSN: 2083–0157, e–ISSN: 2391–6761. – P.48–59. – Access mode: http://www.e–iapgos.pl/.DOI: 10.5604/20830157.1201317. – http://e–iapgos.pl/stats.php?icid=1201317.
78. Vysotska Victoria. Information Technology of Processing Information Resources in Electronic Content Commerce Systems // Victoria Vysotska, Lyubomyr Chyrun, Liliya Chyrun // Computer Science and Information Technologies: Proc. of the XI–th Int. Conf. CSIT'2016, 6–10 September, 2016, Lviv, Ukraine..– Lviv: Lviv Polytechnic Publishing House, 2016.– P. 212–222.
79. Vysotska Victoria. The Commercial Content Digest Formation and Distributional Process / Victoria Vysotska, Lyubomyr Chyrun, Liliya Chyrun // Computer Science and Information Technologies: Proc. of the XI–th Int. Conf. CSIT'2016, 6–10 September, 2016, Lviv, Ukraine. – Lviv: Lviv Polytechnic Publishing House, 2016.– P. 186–189.
80. Vysotska Victoria. Unified Methods of Processing Information Resources in Electronic Content Commerce Systems / Vysotska, Lyubomyr Chyrun, Liliya Chyrun // The 7th International Scientific and Technical Conference "Computer Sciences and Information Technologies" (CSIT'2012) which will be held November 20–24, 2012 at Lviv Polytechnic National University (Lviv, Ukraine) . – Lviv. – P. 190–194.
81. Vysotska Victoria. Web Content Processing Method for Electronic Business Systems [Electronic resource] / Victoria Vysotska, Lyubomyr Chyrun // International Journal of Computers & Technology. – Vol 12, No 2. – December 2013. – PP.3211–3220. – ISSN 2277–3061. – Access mode: http://cirworld.org/journals/index.php/ijct/article/view/3299.
82. Zadachyn V. Calculation of optimal path for parallel car parking / V. Zadachyn, O. Dorokhov // Transport and Telecommunication. – Volume 13. – 2012. – P. 303–309.
83. Айвазян С.А. Статистическое исследование зависимостей / С.А. Айвазян. М.: Металлургия 1968. – 227 с.
84. Алгоритмы и программы для вычисления функций на ЭЦВМ. – Киев: Ин–т кибернетики АН УССР, вып. 1, 1972. – 184 с.; вып. 2, 1972. – 168 с.; вып. 3, 1976. – 234 с.; вып. 4, 1977. – 239с.; вып. 5(часть первая), 1981. – 241 с.; вып. 5 (часть вторая), 1980. – 211 с.
85. Алєксєєва К.А. Аналіз процесу опрацювання web–ресурсу інформаційного продукту на основі нечіткої логіки та контент–аналізу / К.А. Алєксєєва, А.Ю. Берко, В.А. Висоцька // Комп'ютерні науки та інформаційні технології, Вісник

Національного університету "Львівська політехніка". – № 843. – Львів 2016. – С. 122–134.

86. Алєксєєва К.А. Інформаційна технологія управління Web–ресурсом на основі нечіткої логіки / К.А. Алєксєєва, А.Ю. Берко, В.А. Висоцька // Інформаційні системи та мережі. Вісник Національного університету "Львівська політехніка". – № 829. – Львів, 2015. – С. 7–28.

87. Алєксєєва К.А. Особливості процесу управління web–ресурсом комерційного контенту на основі нечіткої логіки / К.А. Алєксєєва, А.Ю. Берко, В.А. Висоцька // Комп'ютерні науки та інформаційні технології. Вісник Національного університету "Львівська політехніка". – № 826. – Львів, 2015. – С. 201–211.

88. Алєксєєва К.А. Технологія управління комерційним web–ресурсом на основі нечіткої логіки [Electronic resource] / К.А. Алєксєєва, А.Ю. Берко, В.А. Висоцька // Науковий журнал «Радіоелектроніка. Інформатика. Управління.». – № 3(34). – Запоріжжя: ЗНТУ. – 2015/3. – С.71–79. – ISSN 1607–3274 (print), ISSN 2313–688X (on–line). – Access mode: http://ric.zntu.edu.ua/.

89. Алєксєєва К.А. Управління Web–ресурсами за умов невизначеності [Electronic resource] / К.А. Алєксєєва, А.Ю. Берко, В.А. Висоцька // Журнал «Технологический аудит и резервы производства». – Vol 2, No 2(22) (2015). – Харків, 2015. – ISSN (print) 2226–3780, ISSN (on–line) 2312–8372. – С. 4–7. – Access mode: http://journals.uran.ua/tarp/article/view/39139.

90. Амосов А.А. Вычислительные методы для инженеров: учебн. пособ. / А.А. Амосов, Ю.А. Дубинський, Н.В. Копченова. – М. : Высшая школа, 1994. – 544 с.

91. Андруник В.А. Неперервна мінімаксна сплайн–апроксимація температурної характеристики та чутливості термодіодного сенсора логарифмічним виразом / В.А. Андруник, П.С. Малачівський // Вісник Вінницького політехнічного інституту. – 2007. – №5. – С.108–115.

92. Андруник В. Неперервна й гладка мінімаксна сплайн–апроксимація експоненційним виразом / В. Андруник, П. Малачівський // Фізико–математичне моделювання та інформаційні технології. – 2007. – Вип. 5. – С.85–97.

93. Ахиезер Н.И. Лекции по теори апроксимации / Н.И. Ахиезер. – Москва: Наука, 1965.

94. Бабенко К.Н. Основы вычислительного анализа / К.Н. Бабенко. – М.: Наука, 1986. – 744 с.

95. Базара М. Нелинийное программирование. Теория и алгоритмы / М. Базара, К. Шетти. – М: Мир, 1982. – 584 с.

96. Барахнин В.Б. Введение в численный анализ / В.Б. Барахнин, В.П. Шапеев. – Новосибирск, 1997. – 112 с.

97. Батунер Л.М. Математические методі в химической технике / Л.М. Батунер, М.К. Позин. – Госхимштад, 1953.

98. Бахвалов Н.С. Численные методы / Н.С. Бахвалов, Н.П. Жидков, Г.М. Кобельков. – М: Бином, 2007. – 636 с.

99. Бахвалов Н.С. Численные методы. Анализ, алгебра, обычные диференциальные уравнения / Н.С. Бахвалов. – М.: Наука, 1975. – 631 с.

100. Березин Н.С. Методы вичислений / И.С. Березин, Н.П. Житков. – М.: Наука, 1966. т. 1. – 632 с.

101. Березин Н.С. Методы вычислений / Н.С. Березин, Н.П. Жидков. – М.: Наука, 1962. т. 2. – 640 с.
102. Березовская Л.И. Прогамма построения интерполяционного полинома по формуле Лагранжа для функции f(x), заданной таблично / Л.И. Березовская. – К: Ин–т кибернетики АН УССР. – 1970.
103. Берко А.Ю. Застосування маркетингових методів для аналізу життєвого циклу комерційного web–контенту / А.Ю. Берко, В.А. Висоцька // Інформаційні системи та мережі. Вісник Національного університету "Львівська політехніка". – Львів 2011. – №699. – С. 3–13.
104. Берко А.Ю. Інтелектуальна система прийняття маркетингових рішень з поширення електронних видань / А.Ю. Берко, В.А. Висоцька, О.Ю. Михайлов // Комп'ютерні науки та інформаційні технології, Вісник Національного університету "Львівська політехніка". – Львів 2010. – № 672. – С. 7–18.
105. Берко А.Ю. Системи електронної контент–комерції. Монографія / А.Ю. Берко, В.А. Висоцька, В.В. Пасічник // Видавництво національного університету "Львівська політехніка". – Львів 2009. – 612 с.
106. Берко А.Ю. Застосування методу контент–аналізу для формування інформаційних ресурсів в системах електронної контент–комерції/ А.Ю.Берко, В.А.Висоцька, М.М.Сороковський // Інформаційні системи та мережі. Вісник Національного університету "Львівська політехніка". – № 743. – Львів 2012. – С. 3–15.
107. Берко А.Ю. Методи та засоби оцінки ризиків безпеки інформації в системах електронної контент–комерції / А.Ю. Берко, В.А. Висоцька, І.В. Рішняк // Комп'ютерні системи проектування. Теорія і практика. Вісник Національного університету "Львівська політехніка". – Львів 2007. – № 591. – С. 81–86.
108. Берко, А.Ю. Методи та засоби оцінювання ризиків безпеки інформації в системах електронної комерції / А.Ю. Берко, В.А. Висоцька, І.В. Рішняк // Інформаційні системи та мережі. Вісник Національного університету "Львівська політехніка". – Львів 2008. – № 610. – С. 20–33.
109. Берко, А.Ю. Моделі оцінювання ризиків інформаційної безпеки в системах електронної комерції / А.Ю. Берко, І.В. Рішняк, В.А. Висоцька // Искусственный интеллект. Науково–технічний журнал. – Донецьк 2007. – № 2. – С. 14–21.
110. Биллингсли П. Схоимость вероятностных мер / П. Биллингсли. – М.: Наука, 1977. – 352 с.
111. Білуха М.Т. Основи наукових досліджень / М.Т. Білуха. – Київ: Вища школа, 1997. – 184 с.
112. Бісікало О.В. Виявлення ключових слів на основі методу контент–моніторингу україномовних текстів [Electronic resource] / О.В. Бісікало, В.А. Висоцька// Науковий журнал «Радіоелектроніка. Інформатика. Управління.». – № 1(36). – Запоріжжя: ЗНТУ. – 2016/1. – С. 74–83. – ISSN 1607–3274 (print), ISSN 2313–688X (on–line). – Access mode: http://ric.zntu.edu.ua/.
113. Бісікало О.В. Експериментальне дослідження пошуку значущих ключових слів україномовного контенту / О.В. Бісікало, В.А. Висоцька// Інформаційні системи та мережі. Вісник Національного університету "Львівська політехніка". – № 829. – Львів, 2015. – С. 255–272.

114. Бісікало О.В. Застосування методу синтаксичного аналізу речень для визначення ключових слів україномовного тексту [Electronic resource] / О.В. Бісікало, В.А. Висоцька // Науковий журнал «Радіоелектроніка. Інформатика. Управління.». – № 3(38). – Запоріжжя: ЗНТУ. – 2016/3. – С. 54–65. – ISSN 1607–3274 (print), ISSN 2313–688X (on–line). – Access mode: http://ric.zntu.edu.ua/.
115. Блох Л.С. Основные графические методы обработки опытных данных / Л.С. Блох. – М.: Машгиз, 1951. – 268 с.
116. Бобков В.В. Вычислительные методы / В.В. Бобков, В.И. Крылов, П.Н. Монастырный. – Т.1. – Минск: Наука и техника, 1989. – 304 с.
117. Бобков В.В. Вычислительные методы / В.В. Бобков, В.И. Крылов, П.Н. Монастырный. – Т.2. – Минск: Наука и техника, 1989. – 400 с.
118. Боглаев Ю.П. Вычислительная математика и программирование / Ю.П. Боглаев. – М: Высшая школа, 1990. – 544 с.
119. Бородин А.Н. Выдающиеся математики / /А.Н. Бородин, А.С. Бугай. – Київ: Радянська школа, 1987. – 654 с.
120. Бронтейн И.Н. Справочник по математике для инженеров и учащихся втузов / И.Н. Бронштейн, К.А. Семендяев. – М.: Нука, 1986. – 544 с.
121. Вальд А. Последовательный анализ . А. Вальд. – М.: ИЛ, 1960. – 306 с.
122. Ван-дер-Варден. Математическая статистика / Ван –дер-Варден. – М.: ИЛ, 1960. – 406 с.
123. Велодарский Е.Т. Планирование и организация сумерительного эксперимента / Велодарский Е.Т. Велодарский, Б.И. Малиновский, Ю.М. Туз. – Киев: Вища школа, 1987. – 286 с.
124. Виньерон А. Обработка результатов физико-химических наблюдений / А. Виньерон. – ОНТИ, 1936.
125. Висоцька В.А. Методи і засоби опрацювання інформаційних ресурсів в системах електронної контент–комерції : автореферат дисертації на здобуття наукового ступеня кандидата технічних наук : 05.13.06 – інформаційні технології / Вікторія Анатоліївна Висоцька; Національний університет «Львівська політехніка». – Львів, 2014. – 24 с.
126. Висоцька В.А. Аналіз та оцінювання ризиків інформаційної безпеки в інформаційних системах: тези науково–практичної конференції, 21 березня 2007 р. / В.А. Висоцька, І.В. Рішняк // Міжвузівська наукова конференція науково–педагогічних працівників. Інститут підприємництва та перспективних технологій. – Львів 2007. – С. 202–203.
127. Висоцька В.А. Аналіз та оцінювання ступеня ризиків в електронній комерції / В.А. Висоцька, І.В. Рішняк, Л.В. Чирун // Інформаційні системи та мережі. Вісник Національного університету "Львівська політехніка". – Львів 2007. – № 589. – С. 56–68.
128. Висоцька В.А. Аналітичні методи опрацювання інформаційних ресурсів в системах електронної контент–комерції / В.А. Висоцька // Інформаційні системи та мережі. Вісник Національного університету "Львівська політехніка". – № 829. – Львів, 2015. – С. 76–101.
129. Висоцька В.А. Ентропія як міра невизначеності контенту в системах електронної комерції / В.А. Висоцька // Комп'ютерні науки та інформаційні

технології. Вісник Національного університету "Львівська політехніка". – Львів 2007. – № 604. – С. 65–75.

130. Висоцька В.А. Концептуальна модель опрацювання інформаційних ресурсів в системах електронної контент–комерції [Electronic resource] / В.А. Висоцька, Л.В. Чирун // Математичні машини і системи. Науково–технічний журнал. – 2015.– № 3. – ISSN 1028–9763. – С. 179–190. – Access mode: http://www.immsp.kiev.ua/publications/articles/2015/2015_3/03_2015_Vysotska.pdf.

131. Висоцька В.А. Метод вибору оптимального алгоритму криптографічного захисту інформації / В.А. Висоцька, О.Р. Гарасим // Інформаційні системи та мережі. Вісник Національного університету "Львівська політехніка". – Львів 2010. – № 673. – С. 220–233.

132. Висоцька В.А. Моделювання системи захисту інформації в електронній комерції на основі оцінювання ризиків інформаційної безпеки: тези науково–практичної конференції, 16 травня 2007 р. / В.А. Висоцька // Науково–практична конференція "Математичне моделювання складних систем". Львівський державний інститут новітніх технологій та управління ім. В. Чорновола. – Львів 2007. – С. 17–22.

133. Висоцька В.А. Опрацювання інформаційних ресурсів у системах електронної контент–комерції / В.А. Висоцька, Л.В. Чирун // Відбір і обробка інформації. Міжвідомчий збірник наукових праць, випуск 42 (118), – Львів 2015. – С. 84–91. – ISSN 0474–8662.

134. Висоцька В.А. Особливості рубрикації текстового комерційного контенту / В.А. Висоцька // Комп'ютерні науки та інформаційні технології. Вісник Національного університету "Львівська політехніка". – № 826. – Львів, 2015. – С. 359–367.

135. Висоцька В.А. Умовна ентропія та ентропія поєднання контенту в системах електронної комерції. / В.А. Висоцька // Комп'ютерні системи проектування. Теорія і практика, Вісник Національного університету "Львівська політехніка". – Львів 2008. – № 626. – С. 116–125.

136. Висоцька В.А. Формальна модель формування інформаційних ресурсів систем електронної контент–комерції: тези науково–практичної конференції. 10–13 травня 2011 р. / В.А. Висоцька, Л.Б Чирун, Л.В. Чирун // І Міжнародна науково–технічнаа конференція Обчислювальний інтелект (результати, проблеми, перспективи)». – Черкаси, 2011. – С. 289–290.

137. Воеводин В.В. Вычислительные основы линейной алгебры / В.В. Воеводин. – М.: Наука, 1977. – 304 с.

138. Воеводин В.В. Численные методы алгебры. Теория и алгоритмы / В.В. Воеводин. – М.: Наука, 1966. – 248 с.

139. Воеводин В.В. Математические модели и методы в параллельных процессах / В.В. Воеводин. – М.: Наука, 1996. – 296 с.

140. Возняк Л.С. Чисельні методи: Методичний посібник для студентів природничих спеціальностей [Electronic resource] / Л.С. Возняк, С.В. Шарин. – Івано–Франківськ: "Плай", 2001. – 64 с. – Access mode: http://www.mif.pu.if.ua/attachments/article/24–02–11/Chys_metody_Voznjak_Sharyn.pdf.

141. Волков Е.А. Численные методы / Е.А. Волков. – М: Высшая школа, 1987. – 312 с.
142. Вольтер Б.В. Исследование процесса полимеризации этилена методом множественной корреляции / Б.В. Вольтер, Е.В. Маркова. – М.: «Химическая промышленность». - № 4. – 1961. – С. 81-92.
143. Вувченко А.Л. Основы научных исследований / А.Л. Вувченко. – Киев: Знання, 2000. – 186 с.
144. Выгодский М. Справочник по высшей математике / М. Выгодский. – М.: Наука, 1977. – 872 с.
145. Гаврилюк В.Т. О принципе равномерного приближения в применении к обработке результатов измерений, условных в особенности / В.Т. Гаврилюк, Е.Я. Ремез. – К: Ин–т кибернетики АН УССР. – 1969.
146. Гаврілюк І.П. Методи обчислень / І.П. Гаврілюк, В.Л. Макаров.– К: Вища школа, в 2–х частинах, 1995. – 367 с.
147. Гавурин М.К. Лекции по методам вычислений / М.К. Гавурин. – М.: Наука, 1971. – 148 с.
148. Гасько Р.В. Інформаційна система аналізу психологічного стану особистості / Р.В. Гасько, В.А. Висоцька, Л.Б. Чирун // Інформаційні системи та мережі. Вісник Національного університету "Львівська політехніка". – № 829. – Львів, 2015. – С. 102–128.
149. Геращенко О.А. Температурные измерения / Геращенко О.А., Гордов А.И., Еремина А.К. и др. – Киев: Наук. думка, 1989. – 704 с.
150. Гилл Ф. Численные методы условной оптимизаии / Ф. Гилл, У. Мюррэй, М. Дж. Д. Пауэлл, Р. Саджент, Д.М, Райан, В.Г. Свэн, Р. Флетчер. – М.: Мир, 1977. – 290 с.
151. Гмурман В.Е. Теория вероятностей и математическая статистика / В.Е. Гмурман. – М: Высшая кола, 1999. – 479 с.
152. Голощук Р.О. Математичне моделювання процесів дистанційного навчання / Р.О. Голощук, В.В. Литвин, Л.В. Чирун, В.А. Висоцька // Інформаційні системи та мережі. Вісник Національного університету "Львівська політехніка". – Львів 2003. – № 489. – С.100–109.
153. Грановский В.А. Методы обработки экспериментальных данных при измерениях / В.А. Грановский, Т.Н. Сирая // Ленинград: Энергоатомиздат. Ленингр. отделение, 1990. – 288с.
154. Григорів В.С. Про структурний підхід в економікоматематичному моделюванні виробничих та екологоекономічних функцій оптимальних випусків і затрат / В.С. Григорів, Л.М. Буяк // Науковий вісник КНЕТУ. Економічні науки. – В. 2. – Чернівці: КНТУ, 2002. - С. 31-45.
155. Гультяев А.К. MATLAB 5.3 Имитационное моделирование в среде WindowsЖ Практическое пособие / А.К. Гультяев. – СПб.: Корона принт, 2001. – 400 с.
156. Данилевський М. П. Елементи обчислювальної математики: Конспект лекцій з дисципліни "Вища математика II" (для студентів 1 курсу денної та 2 курсу заочної форм навчання за напрямами підготовки 6.030504 "Економіка підприємства" і 6.030509 "Облік і аудит"). з дисципліни "Вища математика" (для

студентів 2 курсу денної форми навчання за напрямом підготовки 6.050702 "Електромеханіка") / М.П. Данилевський, А.В. Якунін, Г.А. Кузнецова; Харк. нац. акад. міськ. госп–ва. – Х.: ХНАМГ, 2012. – 156 с.

157. Данилина Н.И. Вычислительная математика / Н.И. Данилина, Н.С. Дубровская, О.П. Кваша, Г.Л. Смирнов. – М.: Высшая школа, 1985. –368 с.
158. Де Бор К. Практическое руководство по сплайнам / К. Де Бор. – М: Радио и связь, 1985. – 304 с.
159. Девятов Б. Определение динамических характеристик и критериев инерционности химико-технологических аппаратов / Б. Девятов // Автоматика и телемеханика. – № 6. – 1960. – С. 779–790.
160. Демидович Б.П. Основы вычислительной математики / Б.П. Демидович. – М: Наука, 1994. – 664 с.
161. Демидович Б.П. Основы вычислительной математики / Б.П. Демидович, И.А. Марон. – М: Наука, 1966. – 664 с.
162. Демидович Б.П. Численные методы анализа / Б.П. Демидович, И.А. Марон, Е.З. Шувалова. – М: Наука, 1963. – 400 с.
163. Демьянов В.Ф. Введение в минимакс / В.Ф. Демьянов, В.Н. Малоземов. – М: Наука, 1972. – 368 с.
164. Демьянов В.Ф. Вопросы теории и элементы программного обеспечения минимаксных задач / под ред. В. Ф. Демьянова и В. Н. Малоземова.. – Л.: Изд–во Ленингр. ун–та, 1977. – 192 с.
165. Денежкина И.Е. Численные методы. Курс лекций: Учебное пособие / И.Е. Денежкина. – М: Финакадемия, 2010. – 132 с.
166. Денисюк В.П. Сплайны и их приложения в задачах моделирования и обработки измерительных сигналов / В.П. Денисюк, Б.Г. Марченко. – Киев: Национальный технический университет Украины "КПИ". 1995. – 246 с.
167. Дубров А.М. Анализ статистических связей в одномерном и многомерном анализе / А.М. Дубров, М.Е. Мазуров. – М.: МЗСИ, 1980. – 106 с.
168. Дудченко А.А. Основы научных исследований / А.А. Дудченко. – Киев: Знання, 2000. – 186 с.
169. Дэннис Дж. Численные методы безусловной оптимизации и решения нелинейных уравнений / Дж. Дэннис, Р. Шнабель. – М: Мир, 1988. – 40 с.
170. Дьяконов В.П. Математическая система MAPLE V R3/R4/R5 / В.П. Дьяконов. – М: Солон, 1998. – 399 с.
171. Емелин А. Высшая математика для заочников и не только [Electronic resource] / А. Емелин. – 2010–2016. – Access mode: http://mathprofi.ru.
172. Емец С.В. Алгоритм корекции статических характеристик измерительных преобразователей / С.В. Емец, С.А. Емец // Межвуз. науч. сб. – Уфа: УГАТУ, 1996. – С. 152–159.
173. Ермаков С.М. Статистическое моделирование / С.М. Ермаков, Г.А. Михайлов. – М.: Наука, 1982. – 296 с.
174. Ермаков С.М. Метод Монте-Карло и смежные вопросы / С.М. Ермаков. – М.: Наука, 1971. – 296 с.
175. Єлейко В.І. Економіко-статистичні методи моделювання і прогнозування / В.І. Єлейко. – К.: НИК ВО, 1988. – 88 с.

176. Єлейко В.І. Моделювання зв'язків в економіці за допомогою регресійного аналізу / В.І. Єлейко. – К.: НИК ВО, 1990– 131 с.
177. Жалдак М.І. Чисельні методи математики / М.І. Жалдак, Ю.С. Рамський. – К:Радянська школа, 1984. – 206 с.
178. Заварыкин В.М. Численные методы / В.М. Заварыкин, В.Г. Житомирский, М.П. Лапчик. – М: Просвещение, 1990. – 176 с.
179. Задачин В.М. Робоча програма навчальної дисципліни "Чисельні методи" для студентів напряму підготовки "Комп'ютерні науки" всіх форм навчання / В.М. Задачин, І.Г. Конюшенко. – Х: Вид. ХНЕУ, 2012. – 40 с.
180. Задачин В.М. Чисельні методи : навчальний посібник [Electronic resource] / В.М. Задачин, І.Г. Конюшенко. – Х: Вид. ХНЕУ ім. С. Кузнеця, 2014. – 180 с. – Access mode: http://www.repository.hneu.edu.ua/jspui/bitstream/123456789/8310/1/ЧИСЕЛЬНІ%20МЕТОДИ.pdf.
181. Иванов В.В. Методы вычислений на ЭВМ: Справочное пособие / В.В. Иванов. – Киев: Наук. думка, 1986. – 564 с.
182. Иващенко А.Н. Аппроксимация термометрических характеристик кремниевых диодных сенсоров температуры / А.Н. Иващенко, Ю.М. Шварц // Оптоэлектроника и полупроводниковая техника: Межвед. сб. науч. тр. – 2003. – Вып. 38. – С. 61–70.
183. Казаков И.Е. Приближенный вероятностный анализ точности работы существенно нелинейных систем / И.Е. Казаков // Автоматика и телемеханика. – № 5. – 1959. – С. 41-54.
184. Каленчук–Порханова А.А. Аппарат аппроксимации в составе аппаратного обеспечения суперкомпьютера с кластерной архитектурой / А.А.Каленчук–Порханова, Л.П. Вакал // Штучний інтелект. – 2009р. – №1. – С. 158–165.
185. Каленчук–Порханова А.А. Пакет программ аппроксимации функций / А.А.Каленчук–Порханова, Л.П. Вакал // Комп`ютерні засоби, мережі та системи. – 2008 р. – №7. – С. 32–38.
186. Калиткин Н.Н. Численные методы / Н.Н. Калиткин. – М: Наука, 1978. – 512 с.
187. Катренко А.В. Теорія прийняття рішень: підручник для вищих навчальних закладів / А.В. Катренко, В.В. Пасічник, ВП. Пасько. – К: Видавнича група BHV, 2009. – 448 с.
188. Кісь Я.П. Особливості застосування методу контент–аналізу для опрацювання текстової інформації / Я.П. Кісь, В.А. Висоцька, Л.Б. Чирун, В.М. Фольтович // Інформаційні системи та мережі. Вісник Національного університету "Львівська політехніка". № 814.– Львів 2015 – С. 282–292.
189. Кононюк А.Е. Дискретно–непрерывная математика. Матрицы / А.Е. Кононюк. – К: Освіта України, 2012. – 508 с.
190. Копченова Н.В. Вычислительная математика в примерах и задачах / Н.В. Копченова, И.А. Марон. – М: Наука, 1972. – 368 с.
191. Корн Г. Справочник по математике для научніх работников и инженеров. Определения, теоремы формулы / Г. Корн, Т. Корн // Издание четвертое. Перевод со второго американского переработанного издания И.Г. Арамановича, А.М.

Березмана, И.А. Вайнтейна, Л.З. Румчинского, Л.Я. Цлара под общей редакцией И.Г. И.Г. Арамановича. – М: Наука, 1977. – 832с.
192. Корнейчук А.А. Итерационный метод отыскания многочлена наилучшего приближения / А.А. Корнейчук, Н.Ю. Ширикова. – Дубна, 1967.
193. Коссак О. Методи наближених обчислень, навч. посібн. / О. Коссак, О. Тумашова, О. Коссак.– Л: БаК, 2003. – 168 с.
194. Косякин А.А. Статистическая теория квантования по уровню / А.А. Косякин // Автоматика и телемеханика. - № 6. – 1961.
195. Кравцова Л.В. Лекції з дисципліни Чисельні методи [Electronic resource]/ Л.В. Кравцова. – Херсон, 2007. – Access mode: http://dls.ksu.kherson.ua/DLS/Library/LibdocView.aspx?id=bbd94fad–a930–49da–952c–df2838670a5b.
196. Крастинь О.П. Разработка и интерпретация моделей корреляционных связей в экономике / О.П. Крастинь. – Рига: Зинатне, 1983. – 302 с.
197. Кривоносов А.И. Статические характеристики поликристалических резисторов / А.И. Кривоносов, В.Я. Кауфман. – М: Энергия, 1976. – 120 с.
198. Крыжановская М.П. Эмпирические формулы и основы номографии / М.П. Крыжановская. – 8. – Л, 1949. – 85 с.
199. Крылов В.И. Вычислительные методы высшей математики, в 2 томах / В.И. Крылов, В.В. Бобков, П.Н. Монастырный. – М: Вышэйшая школа, 1972. – т. 1. – 304 с.
200. Крылов В.И. Справочная книга по вычислительному интегрированию / В.И. Крылов, Л.Т. Шумгина. – М.: Наука, 1966. – 370 с.
201. Ланс Дж.Н. Численные методы для быстродействующих вычислительных машин. – М.: ИНЛ, 1962. – 208 с.
202. Левин Б.Р. Теория случайных процессов и ее применение в радиотехнике / Б.Р. Левин. М.: Советское радио, 1960. – 504 с.
203. Лукомский Я.И. Теория корреляции и ее применение к анализу производства / Я.И. Лукомский. М.: Госстатиздат, 1958. – 312 с.
204. Ляшенко Б.М. Методи обчислень: навчально–методичний посібник для студентів фізико–математичного факультету [Electronic resource] / Б.М. Ляшенко, О.М. Кривонос, Т.А. Вакалюк. – Ж: ЖДУ, 2014. – 228 с. – Access mode: http://eprints.zu.edu.ua/18543/1/metody_obchyslen.pdf.
205. Ляшко И.И. Методы вичислений / И.И. Ляшко, В.Л. Макаров, А.А. Скоробагатько. – Киев: Вища школа, 1977. – 406 с.
206. Ляшенко М.Я. Чисельні методи / М.Я. Ляшенко, М.С. Голованъ. – К: Либідь, 1996. – 287 с.
207. Мак–Кракен Д. Численные методы и программирование на Фортране / Д. Мак–Кракен, У. Дорн / Перевод с английского Казака Б.Н. Под редакцией Наймарка Б.М. Издание второе, стереотипное. – М.: Мир, 1977. – 584 с.
208. Малачівський П. Рівномірне сплайн–наближення / П. Малачівський, В. Андруник // "Комп'ютерні технології друкарства", № 7. – Львів: Українська академія друкарства, 2002. – С. 107–115.

209. Малачівський П. Модифікований алгоритм Валле–Пуссена / П. Малачівський // Фізико–математичне моделювання та інформаційні технології. 2005. – Вип. 2. – С. 159–166.
210. Малачівський П. Оптимальне калібрування засобів вимірювання / П. Малачівський // Наукові праці III Міжнародної науково–технічної конференції «Метрологія та вимірювальна техніка», 8–10 жовтня 2002, Харків. – Харків: ХНДІМ, 2002. – Том 1. – С. 79–82.
211. Малачівський П.С. Програмування в середовищі Visual Basic / П.С. Малачівський. – Львів: Видавництво "Бескид Біт", 2004. – 260 с.
212. Малачівський П.С. Рівномірне наближення з точним відтворенням значень функції та похідної в заданих точках / П.С. Малачівський // Доп. НАН України. – 2006. – № 9. – С. 80–85.
213. Малачівський П. С. Рівномірне наближення функцій з інтерполюванням у заданих точках / П.С. Малачівський // Фізико–математичне моделювання та інформаційні технології. – 2006. – Вип. 4. – С. 142–150.
214. Малачівський П. Розв'язування задач аналітичного представлення експериментальних даних / П. Малачівський, Б. Монцібович // Україномовне програмне забезпечення: Матеріали 4–ї та 5–ї Міжнародних науково–практичних конференцій «УкрСофт». Каталог україномовних програмних засобів. – Львів: СП «БАК», 1995. – С. 186–190.
215. Малачівський П.С. Програмне забезпечення задач рівномірної апроксимації дослідних даних / П.С. Малачівський, Б.Р. Монцібович // В сб. науч. Тр.: Математические методы в компьютерных системах / НАН Украины. Ин–т кибернетики им. В.М. Глушкова. – Киев, 1996. – С. 38–43.
216. Малоземов В.Н. Наилучшее равномерное приближение функций нескольких аргументов / В.Н. Малоземов. – Киев: Ин–т кибернетики АН УССР. – 1970.
217. Маркушевич А.И. Краткий курс теории аналитических функцій / А.И. Маркушевич. – М.: Наука, 1978. – 416 с.
218. Марчук Г.И. Методы вычислительной математики / Г.И. Марчук. – М. : Наука, 1989. – 608 с.
219. Математична лінгвістика. Книга 1. Квантитативна лінгвістика / [В.В. Пасічник, Ю.М. Щербина, В.А. Висоцька, Т.В. Шестакевич] // Навчальний посібник з грифом Міністерства освіти і науки, молоді та спорту України. – Львів: видавництво «Новий світ–2000», 2012. – 359 с.
220. Мильштейн Г.Н. Приближенное интегрирование стохастических дифференциальных уравнений / Г.Н. Милштейн // Теория вероятностей и ее применение. – Т. 20, В. 3. – 1975. – С. 583-588.
221. Митропольский А.К. Техника статистических вычислений / А.К. Митропольский. – М.: Физматгиз, 1961. – 104 с.
222. Молчанов М.И. Машинные методы решения прикладных задач. Алгебра. Приближение функций / М.И. Молчанов. – Киев: Наукова думка, 1987. – 288 с.
223. Монцибович Б.Р., Попов Б.А. Наилучшие приближения табличных функций многочленами низких степеней (алгоритмы и программы), ч. 1 и 2 / Б.Р. Монцибович, Б.А. Попов. – К: Ин–т кибернетики АН УССР , 1973. – 214, 238 с.

224. Монцибович Б.Р. Диалоговый пакет программ для аналитической обработки табличных данных (ППП РАДАН – 2) / Б.Р. Монцибович, Г.Ф. Криворучко, П.С. Малачивский и др.. – Л: 1988. – 135 с. – Деп. в Укр. РФАП 27.03.1990, № АП0278.

225. Монцибович Б.Р. Диалоговый пакет программ для равномерной аппроксимации табличных данных на языке Бейсик для СМ–4 (ППП РАДАН) / Б.Р. Монцибович, П.С. Малачивский, С.В. Тыкайло и др.. –Львов: Ин–т прикл. пробл. мех. и мат. АН УССР, 1985. – 383 с. Деп. в Укр.РФАП АН УССР 25 фев., 1986 г., №АП0091.

226. Москвіна С.М. Комп'ютерні методи дослідження та аналіз даних [Electronic resource] / Москвіна С.М. Москвіна. – ВНТУ, 2010. – Access mode: http://posibnyky.vntu.edu.ua/met/zmist.htm.

227. Мэйндоналд Дж. Вычислительные алгоритмы в прикладной статистике / Дж. Мэйндоналд // Под ред. Е.З. Демиденко. – М: Финансы и статистика, 1988. – 350 с.

228. Мэтьюз Д.Г. Численные методы. Использование MATHLAB / Д.Г. Мэтьюз, К.Д. Финк. – М.: Издательский дом «Вильямс», 2001. – 720 с.

229. Налимов В.В. Применение математической статистики при анализе вещества / В.В. Налимов. – М.: Физматгиз, 1960. – 288 с.

230. Налимов В.В. Статистические методы поиска оптимальных режимов протекания химических процессов / В.В. Налимов // «Успехи химии», т. XXIX, вып. 11. – 1960. – 125 с.

231. Никитин Н.Н. Методы цифрового моделирования стохастических дифференциональных уравнений и оценка их погрешности / Н.Н. Никитин, В.Д. Разевич // Журнал вычислительной математики и математической физики. – Т. 18, № 1. – 1978. – С. 106-117.

232. Норский В.Г. Планирование промышленных экспериментов / В.Г. Норский, Ю.П. Адлер. – М.: Металлургия, 1974. – 884 с.

233. Носач В.В. Решение задач аппроксимации с помощью персональных компьютеров / В.В. Носач // М: Микап, 1993. – С. 6–7.

234. Окуненко В.М. Чисельні методи в моделюванні систем / В.М. Окуненко, В.К. Ясинський. – Чернівці: Золоті литаври, 2006. – 592 с.

235. Окуненко В.М. Метод експертних оцінок перевірки адекватності математичних моделей / В.М. Окуненко // IV международная научно-практическая конференция «Современные информационные технологии». – Одесса: ОДУ, 2003. – С. 94-98.

236. Олександренко В.Л. Побудова чебишовського найкращого поліноміального наближення функції одної змінної за методом підвищуючої дії / В.Л. Олександренко, А.О. Порханова // Автоматика. – №4. – 1967.

237. Ортега Дж. Введение в численные методы решения дифференциальных уравнений / Дж. Ортега, У. Пул. – М: Наука, 1986. – 56 с.

238. Пізюр Я.В. Побудова многочленних Ермітово–Чебишовських сплайнів третього степеня / Я.В. Пізюр, Б.О. Попов // Волинський математичний вісник. – 1995. – Вип. 2. – С. 129–131.

239. Полиа Г. Задачи и теоремы из анализа / Г. Полиа, Г. Сеге. – Том I, II. – М.: Наука, 1978. – 392 с., 432 с.
240. Попов Б.А. Некоторые способы наилучшего приближения функцій / Б.А. Попов. – К: Наук. думка, 1973.
241. Попов Б.А. Наилучшие чебышевские приближения суммой многочлена и нелинейных функций / Б.А. Попов, П.С. Малачивский. – Львов, 1984. – 79 с.
242. Попов Б.А. Приближение функций для технических приложений / Б.А. Попов, Г.С. Теслер. – К: Наук. думка, 1980. – 352 с.
243. Попов Б.А. Равномерное приближение сплайнами / Б.А. Попов. – К: Наук. думка, 1989. – 272 с.
244. Попов Б.О. Розв'язування задач у системі комп'ютерної алгерби MAPLE V / Б.О. Попов. – К: ViP, 2001. – 312 с.
245. Попов Б.О. Чисельні методи рівномірного наближення сплайнами / Б.О. Попов. – Л: Львів. націон. ун–т, 1992. – 92 с.
246. Попов Б.А. Приближение функций для технических приложений / Б.А. Попов, Г.С. Теслер. – К: Наукова думка, 1980. – 352 с.
247. Правила устройства элелектроустановок. – М: Энергоатомиздат, 1985.
248. Пулькин С.П. Вычислительная математика / С.П. Пулькин, Л.Н. Никольская, А.С. Дьячков. – М.: Просвещение, 1980. – 176 с.
249. Ракитин В.И. Практическое руководство по методам вычислений с приложением программ для персональных компьютеров / В.И. Ракитин, В.Е. Первушин. – М: Высшая школа, 1998. – 384 с.
250. Рего К.Г. Метрологічна обробка результатів технічних вимірювань / К.Г. Рего. – Київ: Вища школа, 1987. – 128 с.
251. Ремез Е.Я. Общие вычислительные методы чебишевского приближения / Е.Я. Ремез. – К: Ин–т кибернетики АН УССР, 1957.
252. Ремез Е.Я. Основы численных методов чебышевского приближения / Е.Я. Ремез. – Киев: Наук. думка, 1969. – 623 с.
253. Рішняк І.В. Аналіз якісних та кількісних характеристик оцінювання ризиків інформаційної безпеки в системах електронної комерції / І.В.Рішняк, В.А. Висоцька // Радіоелектроніка та інформатика. – Харків, 2007. – № 2. – С. 128–132.
254. Рішняк І.В. Аналіз якісних та кількісних характеристик оцінювання ризиків інформаційної безпеки в системах електронної комерції / І.В. Рішняк, В.А. Висоцька // Комп'ютерні науки та інформаційні технології. Вісник Національного університету "Львівська політехніка". – Львів 2008. – № 616 – С. 183–189.
255. Романовский В. Математическая статистика / В. Романовский. М.: АНУз.ССР, 1961. – 288 с.
256. Рудавський Ю.К. Лінійна алгебра та аналітична геометрія: навчальний посібник / Ю.К. Рудавський, П.П. Костробій, Х.П. Луник, Д.В. Уханська. – Л.: В–во Державного університету «Львівська політехніка», 1999. – 262 с.
257. Румшинский Л.З. Элементы теории вероятностей / Л.З. Румшинский. – М.: Физматгиз, 1960. – 188 с.

258. Самарский А.А. Введение в численные методы / А.А. Самарский. – М: Наука, 1997. – 240 с. 16.
259. Самарский А.А. Численные методы / А.А. Самарский, А.В. Гулин. – М: Наука, 1989. – 432 с.
260. Саутин С.Н. Планирование эксперимента в химии и химической технологии / С.Н. Саутин. М.: Химия, 1975. – 316 с.
261. Сборник задач по методам вычислений / под ред. П. И. Монастырного. – М: Наука, 1994. – 320 с.
262. Семендяев К.А. Эмпирические формулы / К.А. Семендяев. – М.: ГТТИ, 1933.
263. Скарборо Дж. Численные методы математического анализа / Дж. Скарборо. – М.-Л.: ГТТИ. 1934.
264. Скопецький В.В. Чебишовське наближення сумою многочлена й нелінійного виразу з ермітовим інтерполюванням у крайніх точках відрізка / В.В. Скопецький, П.С. Малачівський // Доповіді НАН України. – 2010. – № 4. – С. 42–47.
265. Смирнов Н.В. Краткий курс математической статистики для технических приложений / Н.В. Смирнов. И.В. Дунин-Барковский. – М.: Физматгиз, 1959. – 332 с.
266. Солодовников В.В. Статистический анализ объектов регулирования / В.В. Солодовников. А.С. Усков. – М.: Машгиз, 1960.
267. Стечкин С.Б. Сплайны в вычислительной математике / С.Б. Стечкин, Ю.Н. Субботин. М.: Наука. 1976. – 248 с.
268. Супруненко О.О. Розв'язання задач з дисципліни «Чисельні методи в інформатиці»: навчально–методичний посбник / О.О. Супруненко, Ю.Є Гребенович. Черкаси: ЧНУ імені Богдана Хмельницького, 2001. – 80 с.
269. Тейлор Дж. Введение в теорию ошибок / Дж. Тейлор. – М: Мир, 1985. – 16 с.
270. Теслер Г.С. Аппроксимация элементарных функций с помощъю полиномов нулевой и первой степени / Г.С. Теслер. – Киев: Ин–т кибернетики АН УССР, 1969.
271. Ткачук К.В. Захисні плівкові екрани для моніторів ЕОМ / К.В. Ткачук та ін. // Охорона праці. №4. 1997.
272. Тумаркий С.А. Об оценке ошибок в методе средних / С.А. Тумаркий // Труды ЦАРИ. – Вып. 198. – 1935. – 154 с.
273. Турецкий В.Я. Математика и информатка / В.Я. Турецкий. – М: ИНФРА–М. 2000. – 560 с.
274. Турчак Л.И. Основы численных методов / Л.И. Турчак. – М: Наука, 1997. – 320 с.
275. Уиттекер Э. Математическая обработка результатов наблюдений / Э. Уиттекер, Г. Робинсон. – Л.-М.: ГТТИ. 1933. – 341 с.
276. Успенский А.К. Выбор вида и нахождение параметров эмпирической формулы / А.К. Успенский. – М.. 1960. – 216 с.
277. Федоренко Р.П. Введение в вычислительную физику / Р.П. Федоренко. – М: МФТИ. 1994. – 528 с.

278. Фельдман Л.П. Чисельні методи в інформатиці: підручник для вищих навчальних закладів / Л.П. Фельдман, А.І. Петренко, А.О. Дмитрієва – К: Видавнича група BHV, 2006. – 480 с.
279. Форсайт Дж., Малькольм М., Моулер К. Машинные методы математических вычислений. – М.: Мир, 1980. – 280 с.
280. Форсайт Дж. Численное решение систем линейных алгебраических уравнений / Дж. Форсайт, К. Молер.– М: Мир, 1969.–168 с.
281. Форсинги А. Дж. Геффеиер. Методы обработки экспериментальных данных / Форсинги А. Дж. Геффеиер. – М.: ИЛ, 1949ю – 245 с.
282. Фролов С.В. Приближенные вычисления / С.В. Фролов. – М.: МВТУ, 1948.
283. Фурунжиев Р.И. Применение математических методов и ЭВМ. Практикум: учебн. пособ. для вузов / Р.И. Фурунжиев, Ф.М. Бабушкин, В.В. Варавко. – Мн : Вышейшая школа, 1988. – 192 с.
284. Хальд А. Математическая статистика с техническими приложениями / А. Хальд. – М.: ИЛ, 1956. – 296 с.
285. Хилле Э. Функциональный анализ и полугруппы / Э. Хилле, Р. Филипс. – М.: Иностранная литература, 1962. – 688 с.
286. Холл Дж. Современные численне методы решения обыкновенных дифференциальных уравнений / Дж. Холл, Дж. Уатт. – М: Мир, 1979. – 312с.
287. Царков Є. Ф. Основи теорії випадкових процесів / Є.Ф. Царков, В.К. Ясинський. – Чернівці: Зелена Буковина, 1999. – 256 с.
288. Цегелик Г.Г. Чисельні методи: підручник / Г.Г. Цегелик. – Л: Видавничий центр Львівського національного університету імені Івана Франка, 2004. – 408 с.
289. Цукерман М.Л. Эмпирические формулы / М.Л. Цукерман. – М., 1932. – 284 с.
290. Чабан В. Чисельні методи / В.Чабан. – Львів: В–во НУ «Львівська політехніка», 2001. – 186 с.
291. Чапля Є.Я. Пакет програм для неперервного і гладкого рівномірного сплайн–наближення високоточної низькотемпературної характеристики («Апрокріо») / Є.Я. Чапля, П.С. Малачівський, М.І. Дзюбачик, Б.Р. Монцибович, А.Р. Торський, В.А. Андруник // Свідоцтво про реєстрацію авт. права на твір № 20705 від 30.05.2007 / Державний департамент інтелектуальної власності МОНУ. – 80 с.
292. Черномаз В.М. Чисельні методи в інформатиці : методичні вказівки до виконання контрольної роботи з дисципліни (студентів 4–го курсу спеціальності 7.080404 «Інтелектуальні системи прийняття рішень» денної форми навчання) [Electronic resource] / В.М. Черномаз. – Краматорськ: ДДМА, 2010. – 40 с. – Access mode: http://www.dgma.donetsk.ua/metod/pm/do/chisl_metod/mu.pdf.
293. Чирун Л. В. Розробка методів апроксимації трансферних функцій з використанням неперервних дробів в системах синтезу мови: автореферат дисертації на здобуття наукового ступеня кандидата технічних наук : 01.05.02 – математичне моделювання та обчислювальні методи / Чирун Любомир Вікторович; Національний університет «Львівська політехніка». – Львів, 2007. – 20 с.

294. Чирун Л.Б. Особливості методів контент–аналізу текстових масивів даних web–ресурсів в межах регіону / Л.Б. Чирун, В.В. Кучковський, В.А. Висоцька // Інформаційні системи та мережі. Вісник Національного університету "Львівська політехніка". – № 829. – Львів, 2015. – С. 296–320.

295. Чирун Л.В. Застосування контент–аналізу текстової інформації в системах електронної комерції / В.А. Висоцька, Л.В. Чирун // Інформаційні системи та мережі. Вісник Національного університету "Львівська політехніка". – Львів 2010. – № 689. – С. 332–347.

296. Чирун Л.В. Інтелектуальний аналіз таблиць прийняття рішень в системах електронної контент–комерції: тези науково–практичної конференції, 16 травня 2007 р. / Л.В. Чирун, В.А. Висоцька, Т.В. Шестакевич // Науково–практична конференція "Математичне моделювання складних систем". Львівський державний інститут новітніх технологій та управління ім. В. Чорновола. – Львів 2007. – С. 140–149.

297. Чирун Л.В. Модель адаптивного синтезу мови в цифрових сигнальних процесорах на основі неперервних дробів / Л.В. Чирун // Вісник Національного університету "Львівська політехніка": Інформаційні системи та мережі.– №489.– Львів, 2003.– С. 307–315.

298. Чирун Л.В. Модель розпізнавання голосових сигналів на основі неперервних дробів / Л.В. Чирун // Тези XVIII відкритої наук.–техн. конф. молодих науковців і спеціалістів Фізико–механічного інституту ім. Г.В. Карпенка НАН України КМН–2003. – Львів, 2003. – С. 239–243.

299. Чирун Л.В. Основи програмування: конспект лекцій / Л.В. Чирун, В.А. Висоцька. – Л: В–во НУ «Львівська політехніка», 2010. – 174 с.

300. Чирун Л.В. Моделювання та розв'язок алгебраїчних рівнянь в нелінійних задачах / Л.В. Чирун, Я.П. Кісь // Тези Першої міжнар. конф. з індуктивного моделювання "МКІМ–2002". – Львів, 2002. – Т.3. – С. 245–249.

301. Чирун Л.В. Адаптивні алгоритми апроксимації коефіцієнтів моделі синтезу мовних сигналів / Л.В. Чирун, Б.П. Русин // Вісник Національного університету "Львівська політехніка": Комп'ютерні системи проектування. Теорія і практика. – №522. – Львів, 2004. – С. 109–114.

302. Чирун Л.В. Проектування інформаційної моделі обробки зображень з використанням теорії ланцюгових дробів / Л.В. Чирун, Б.П. Русин, В.І. Шмойлов // Вісник Національного університету "Львівська політехніка": Інформаційні системи та мережі. – №464. – Львів, 2002. – С. 295–301.

303. Шабат Б.В. Введение в комплексный анализ. – М.: Наука, 1969. – 576 с.

304. Шаманский В.Е. Методы численного решения краевых задача на ЕВМ. – Киев: Изд-во АН УССР, 1963. – 196 с.

305. Шаповаленко В.А. Чисельне обчислення функцій, характеристик матриць і розв'язування нелінійних рівнянь та систем рівнянь: Навч. посібник [Electronic resource] / В.А. Шаповаленко, Л.М. Буката, О.Г. Трофименко. – Одеса: ВЦ ОНАЗ, 2010. – Ч.1. – 88 с. – Access mode: file:///C:/Users/vica/Downloads/Чисельні%20методи%20та%20PjPsPrPµP»CЂPIP.PDF.

306. Шварц Ю.М. Нейросетевая аппроксимация термометрической характеристики диодного сенсора / Ю.М. Шварц, П.А. Яганов, В.Г. Дзюба // Технология и конструирование в электрон. аппаратуре. – 2005. – № 5. – С. 18–22.
307. Шикин Е. В. Кривые и поверхности на экране компьютера. Руководство по сплайнам для пользователей / Е. В. Шикин, А. И. Плис. – М. : ДИАЛОГ–МИФИ, 1996. – 50 с.
308. Шмойлов В.И. Непрерывные дроби и комплексные числа: монография / В.И. Шмойлов, Л.В. Чирун. – Л: Меркатор, 2001. – 564 с.
309. Шунда Н.М. Функції та їх графіки / Н.М. Шунда. – К: Радянська школа, 1976. – 192 с.
310. Шуп Т. Решение инженерных задач на ЭВМ: практическое руководство : учебн. изд. / Т. Шуп; пер. с англ. – М: Мир, 1982. – 237 с.
311. Шураков В.В. Задачник по основам алгоритмизации, алгоязыкам и машинной обработке: учебное пособие / В.В. Шураков, В.П. Морозов. – М: Статистика, 1978. – С. 198.
312. Юрченко І.В. Методи стохастичного моделювання систем / І.В. Юрченко, Л.І. Ясинська, В.К. Ясинський. – Чернівці: Прут, 2002. – 442 с.
313. Яковлєв К.П. Математическая обработка результатов измерений / К.П. Яковлєв. – М.: ГТТИ, 1933.
314. Ясинський В.К. Курс методів обчислень / В.К. Ясинський. – Чернівці: Прут, 2003. – 354 с.
315. Ясинський В.К. Стохастичні динамічно-функціональні системи зі всією передісторією / В.К. Ясинський. – Київ: ТВіМС, 2003. – 304 с.
316. Ясинский В.К. Устойчивость и оптимальное управление стохастических динамических систем со всей предысторией / В.К. Ясинский, И.В. Ясинский. – Киев: ТВіМС, 2004. – 354 с.
317. Ясинський В.К. Вибраны лекцыъ з теорыъ ймовырностей / В.К. Ясинський. – Чернівці: Прут, 2002. – 286 с.
318. Ясинський В.К. Чисельні методи в інформатиці / В.К. Ясинський. – Чернівці: Прут, 2003. – 306 с.
319. Ястрембский А.Л. Об одном алгоритме рациональной аппроксимации таблично заданой функции / А.Л. Ястрембский, П.Г. Македонская. – М: ФТИНТ, 1971.
320. Яцук В.П. Малачівський Методи підвищення точності вимірювання / В.П. Яцук. – Л: Бескид Біт, 2008. – 356 с.

В.А. Андруник, В.А. Висоцька,
В.В. Пасічник, Л.Б. Чирун, Л.В. Чирун

ЧИСЕЛЬНІ МЕТОДИ В КОМП'ЮТЕРНИХ НАУКАХ

Том 2

Навчальний посібник

Підписано до друку 15.05.2018 р.
Формат 70x100/16. Папір друк. №2. Гарнітура Times New Roman Умовн. друк. арк. 33,63.
Тираж 300 прим.

Видавець ФОП Піча С.В.
а/с 5026, м. Львів-53, 79053, Україна
e-mail: novsv2016@ukr.net
Свідоцтво про внесення суб'єкта видавничої справи до Державного реєстру видавців, виготівників і розповсюджувачів ви-давничої продукції: серія ДК № 5069 від 22.03.2016 року, видане Державним комітетом інформаційної політики, телебачення та радіомовлення України

Надруковано в поліграфічному центрі ПП «Новий Світ - 2000»



Андруник Василь Адамович – асистент кафедри «Інформаційні системи та мережі» Інституту комп'ютерних інформаційних технологій Національного університету "Львівська політехніка». В 2002 році з відзнакою закінчив академію дрикарства за спеціальністю «Комп'ютерно-інтегровані технологічні процеси та виробництва». Працює асистента кафедри ІСМ, веде лабораторні заняття з курсів «Комп'ютерна графіка», «Системний аналіз», «М складних систем», «Комп'ютерні системи обробки текстової та графічної інформації». Публікує статті у фахови приймає активну участь у міжнародних науково-технічних конференціях, виставках та семінарах.

Область наукових інтересів – математичне моделювання, неперервна та гладка мінімаксна сплайн-а комп'ютерна графіка.



Висоцька Вікторія Анатоліївна – доцент кафедри «Інформаційні системи та мережі» НУ «Львівська політех понад 210 наукових праць, співавтор 3-х монографій та 3-х навчальних посібників. В 2014 р. захистила дисертацій тему «Методи та засоби опрацювання інформаційних ресурсів в системах електронної контент-комерції». Бере ак у роботі міжнародних конференцій (http://victana.lviv.ua). Читає лекційні курси з таких дисциплін: «Дискретна м «Математична лінгвістика», «Математична структурна та прикладна лінгвістика», «Чисельні методи», «Алгор програмування», «Методи опрацювання природної мови», «Крос-платформне програмування та хмарні сервіси», електронної комерції та Інтернет-маркетингу», «Технології створення програмних продуктів», «Системний аналіз» комп'ютерного проектування». Наукові інтереси – системи електронної контент-комерції, методи опрацювання при математична лінгвістика, NLP, SEO-технології, інтелектуальні системи прийняття рішень, технологія програ створення програмних продуктів, штучний інтелект, інтелектуальний аналіз даних.



Пасічник Володимир Володимирович — професор, доктор технічних наук, професор кафедри інформаційни мереж, Інституту комп'ютерних наук та інформаційних технологій Національного університету «Львівська політех понад 200 наукових праць, понад 10 монографій та 20 підручників та навчальних посібників. 1984 року захистив к дисертацію на здобуття наукового ступеня кандидата фізико-математичних наук, наукова спеціальність «М кібернетика». 1994 року захистив дисертацію на здобуття наукового ступеня доктора технічних наук за сп «Теоретичні основи інформатики». Відзначений державною премією України в галузі науки та техніки. Облас інтересів – створення та розроблення інформаційних систем для реальної економіки, виробничої та невиробн Керівник наукової школи з інформаційного моделювання, систем баз даних і знань. Під науковим керівництвом пр Пасічника захищено більше 10 докторських та більше 30 кандидатських дисертацій. Науковий редактор осн підручників і посібників «КОМП'ЮТИНГ», відповідальний редактор серії «Інформаційні системи та мере Національного університету «Львівська політехніка».



Чирун Лілія Борисівна – провідний спеціаліст інституту комп'ютерних наук та інформаційних технологій НУ політехніка». Працює у НУ «Львівська політехніка» з 1979 р., після закінчення Львівського державного університету за спеціальністю «Прикладна математика» факультету прикладної математики та механіки. З жовтня 2005 р. старшим викладачем кафедри інформаційних систем та мереж Інституті комп'ютерних наук та інформаційних те співавтором рацпропозиції та винаходу, понад 60-ти наукових праць та методичних розробок. Читала лек «Комп'ютерні мережі», «Сховища даних та довідково-інформаційні фонди», «Комп'ютерно-технологічні засоби ін діяльності», «Системи автоматизованого перекладу».

Наукові інтереси – сховища даних, веб-технології, прикладна лінгвістика, застосування чисельних інформаційних технологіях.



Чирун Любомир Вікторович – доцент кафедри «Системи автоматизованого проектування» НУ «Львівська п кандидат технічних наук. З 2001 року працював в НУ «Львівська політехніка» в Інституті комп'ютерних наук та інф технологій. на посаді доцента кафедри інформаційних систем і мереж. В 2007р. захистив кандидатську ди спеціальністю 01.05.02 - «Математичне моделювання та обчислювальні методи». Є співавтором монографії «Не дроби и комплексные числа». Автор понад 90 наукових публікацій. Читає лекційні курси з наступних дисциплін: програмування та операційні системи», «Основи програмування та алгоритмічні мови», «Технологія програ створення програмних продуктів», «Чисельні методи в інформатиці».

Наукові інтереси – розпізнавання образів, застосування чисельних методів в інформаційних технологія орієнтоване програмування, веб-технології.

Вища освіта в Україні