Комп'ютинг

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

ЧИСЕЛЬНІ МЕТОДИ В КОМП'ЮТЕРНИХ НАУКАХ

Том 1



omputing

Навчальний посібник

Бібліотека
Галицького коледжу
ім. В.Чорновола
38930
0
Чисельні методи в комп'ютерні

004.93(0758)
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 1

ГАЛИЦЬКИЙ КОЛЕДЖ
ІМЕНІ [illegible]
№ 38930

Навчальний посібник

Серія "Комп"ютинг"
за науковою редакцією д.т.н., професора,
Лауреата державної премії України у галузі наук та техніки
В.В. Пасічника

Видавництво «Новий Світ – 2000»
Львів

2020

УДК 519.6
ББК22.19я73
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»;

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

Чисельні методи в комп'ютерних науках: навчальний посібник за заг. ред. В.В. Пасічника – Львів: Видавництво «Новий Світ – 2000», 2020. – 470 с.

Видавництво «Новий світ – 2000» ISBN 978-617-7519-06-4

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'язання задач. Навчальний посібник призначається для студентів, що навчаються за спеціальностями 122 «Комп'ютерні науки» та 124 «Системний аналіз» і споріднених спеціальностей, які пов'язані вивченням чисельних методів в інформатиці та інформаційних технологій. 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 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».

ISBN 978-617-7519-06-4

© Андруник В.А, Висоцька В.А., Пасічник В.В., Чирун Л.Б., Чирун Л.В., 2020
ФОП С.В., Світ 2000», 2020

Зміст

Передмова наукового редактора серії підручників та навчальних посібників «КОМП'ЮТИНГ» 9
Вступне слово авторів 13
Розділ 1. Математичне моделювання 17
1.1. Чисельні методи та використання персонального комп'ютера для розв'язування прикладних задач 19
1.1.1. Наближений аналіз. Джерела та класифікація похибок 21
1.2. Обчислювальна задача 23
1.2.1. Аналіз постановки задачі 23
1.2.2. Приклади постановки задачі обчислення 24
1.3. Чисельне розв'язування коректних задач. Структура похибки розв'язку 26
1.3.1. Обчислювальна задача. Похибки 26
1.3.2. Похибка заокруглення 27
1.3.3. Похибка функції 30
1.4. Контрольні питання 32
1.5. Задачі для самостійної роботи 32

Розділ 2. Елементи теорії похибок 33
2.1. Абсолютна та відносна похибки 34
2.1.1. Постановка задачі знаходження похибок 34
2.1.2. Приклади знаходження абсолютної та відносної похибок 36
2.2. Пряма задача теорії похибок 38
2.2.1. Аналіз постановки задачі 38
2.2.2. Приклади постановки прямої задачі теорії похибок 40
2.3. Обернена задача теорії похибок 42
2.3.1. Аналіз постановки оберненої задачі теорії похибок 42
2.3.2. Приклади постановки оберненої задачі теорії похибок 43
2.4. Контрольні питання 44
2.5. Задачі для самостійної роботи 44

Розділ 3. Методи розв'язування лінійних алгебраїчних рівнянь 47
3.1. Чисельне розв'язування систем лінійних алгебраїчних рівнянь 49
3.2. Основні поняття та класифікація систем лінійних алгебраїчних рівнянь 49
3.3. Основні поняття та особливості матриць 51
3.3.1. Операції над матрицями 52
3.3.2. Квадратна матриця й суміжні визначення 53
3.3.3. Властивості матриць 54
3.3.4. Обчислення оберненої матриці 54
3.4. Метод Крамера 57
3.4.1. Аналіз постановки задачі 57
3.4.2. Приклади розв'язування системи за формулами Крамера 58
3.5. Метод оберненої матриці 64
3.6. Особливості методу Гауса 69
3.6.1. Аналіз постановки задачі для методу Гауса 70
3.6.2. Метод Гауса (метод заміни змінних) 72

3.6.3. Частковий випадок застосування методу Гауса з послідовним виключенням невідомих ... 74
3.6.4. Метод Гауса за схемою Халецького (метод LU факторизації) ... 78
3.6.5. Метод Гауса з вибором головного елемента ... 81
3.6.6. Метод Гауса з одиничними коефіцієнтами ... 82
3.6.7. Метод Гауса-Жордана ... 84
3.6.8. Приклади розв'язування системи методом Гауса ... 86
3.6.9. Несумісні системи. Системи із загальним розв'язком. Часткові розв'язки ... 98
3.6.10. Способи знаходження оберненої матриці ... 111
3.6.11. Застосування методу Гауса до обчислення визначника ... 115
3.6.12. Застосування методу Гауса до інверсії матриці ... 117
3.7. Метод простої ітерації ... 125
3.7.1. Аналіз постановки задачі ... 125
3.7.2. Приклади розв'язування системи методом простої ітерації ... 127
3.8. Метод Гауса-Зейделя (метод поліпшеної ітерації) ... 129
3.8.1. Аналіз постановки задачі ... 129
3.8.2. Приклади розв'язування системи методом Зейделя ... 130
3.9. Особливості трьохдіагональної матриці ... 131
3.9.1. Обчислення детермінанту трьохдіагональної матриці ... 131
3.9.2. Метод прогонки для розв'язування трьохдіагональних систем лінійних рівнянь ... 133
3.10. Аналіз числових методів розв'язування СЛАР ... 134
3.10.1. Загальна характеристика методів розв'язування систем лінійних алгебраїчних рівнянь ... 134
3.10.2. Збіжність метричної геометричної прогресії ... 135
3.10.3. Збіжність методу простої ітерації ... 139
3.10.4. Збіжність методу Зейделя розв'язування СЛАР ... 141
3.10.5. Збіжність метода Зейделя за елементами матриці ... 145
3.11. Методи розв'язування повної проблеми власних значень і власних векторів ... 147
3.11.1. Наближене знаходження власних значень матриці ... 147
3.11.2. Ідея методу Данилевського ... 149
3.11.3. Обчислення власних векторів методом Данилевського ... 152
3.11.4. Метод невизначених коефіцієнтів ... 153
3.11.5. Загальна постановка задачі методів розв'язування повної проблеми власних значень і власних векторів ... 155
3.11.6. Метод невизначених коефіцієнтів розв'язування повної проблеми власних значень і власних векторів ... 156
3.11.7. Метод інтерполювання розв'язування повної проблеми власних значень ... 157
3.12. Приклади розв'язування систем лінійних алгебраїчних рівнянь ... 158
3.13. Контрольні питання ... 164
3.14. Задачі для самостійної роботи ... 164
3.15. Завдання до лабораторної роботи ... 167

Розділ 4. Розв'язування нелінійних рівнянь ... 170
4.1. Загальна постановка задачі ... 172
4.2. Методи розв'язування нелінійних рівнянь ... 174
4.3. Розв'язування функціональних рівнянь з однією змінною ... 180
4.3.1. Початковий етап: відокремлення коренів ... 182

4.3.2. Відокремлення кореня трансцендентного рівняння ... 183
4.3.3. Відокремлення дійсних коренів алгебраїчного рівняння ... 186
4.3.4. Теореми про число коренів ... 189
4.3.5.Другий етап: уточнення коренів ... 190
4.4. Ітераційні методи уточнення коренів скалярних нелінійних рівнянь ... 190
4.4.1. Ідея ітераційних методів ... 190
4.4.2. Застосування принципу стислих відображень до дослідження ітераційних процесів ... 191
4.4.3. Поняття порядку ітерації ... 194
4.5. Метод ділення проміжку навпіл (метод дихотомії, Мюллера) ... 195
4.5.1. Аналіз постановки задачі ... 195
4.5.2. Геометрична інтерпретація методу ... 196
4.5.3. Обчислення похибки ... 197
4.6. Метод хорд ... 197
4.6.1. Аналіз постановки задачі ... 197
4.6.2. Геометрична інтерпретація методу ... 199
4.6.3. Обчислення похибки ... 200
4.7. Метод Ньютона (метод дотичних) ... 201
4.7.1. Аналіз постановки задачі ... 201
4.7.2. Геометрична інтерпретація методу ... 201
4.7.3. Обчислення похибки ... 203
4.7.4. Інтерпретація методу дотичних через ітераційний процес ... 204
4.7.5. Модифікований метод Ньютона ... 206
4.8. Комбінований метод хорд і дотичних ... 206
4.8.1. Аналіз постановки задачі ... 206
4.8.2. Геометрична інтерпретація методу ... 207
4.8.3. Обчислення похибки ... 208
4.9. Метод січних ... 208
4.9.1. Аналіз постановки задачі ... 208
4.9.2. Геометрична інтерпретація методу ... 208
4.9.3. Обчислення похибки ... 209
4.9.4. Інтерпретація методу січних через ітераційний процес ... 209
4.10. Метод простої ітерацій або метод послідовних наближень ... 212
4.10.1. Аналіз постановки задачі ... 212
4.10.2. Геометрична інтерпретація методу простої ітерації ... 213
4.10.3. Обчислення похибки ... 214
4.10.4. Вплив обчислювальної похибки на збіжність ітераційного процесу ... 216
4.11. Метод релаксації ... 218
4.11.1. Аналіз постановки задачі ... 218
4.11.2. Геометрична інтерпретація методу ... 220
4.11.3. Обчислення похибки ... 220
4.12. Метод парабол (метод Д. Мюллера) знаходження коренів многочлена з комплексними коефіцієнтами ... 221
4.12.1. Аналіз постановки задачі ... 221
4.12.2. Обчислення похибки ... 226
4.13. Розв'язування систем нелінійних рівнянь ... 227
4.14. Приклади розв'язування нелінійних рівнянь ... 230
4.15. Контрольні питання ... 233
4.16. Задачі для самостійної роботи ... 233

4.17. Завдання до лабораторної роботи ... 235

Розділ 5. Задача інтерполяції та наближення функцій ... 238
5.1. Основи теорії інтерполювання та числового диференціювання ... 239
5.1.1. Загальна постановка задачі інтерполювання ... 239
5.1.2. Загальна задача лінійного інтерполювання ... 240
5.1.3. Постановка задачі інтерполяції функції однієї дійсної змінної ... 242
5.1.4. Поліноміальна інтерполяція. Існування та одиничність інтерполяційного полінома ... 243
5.1.5. Типи інтерполювання ... 244
5.2. Інтерполяційний поліном Лагранжа ... 254
5.2.1. Аналіз загальної постановки задачі ... 254
5.2.2. Інтерполяційна формула Лагранжа. Частковий випадок ... 255
5.2.3. Побудова інтерполяційного многочлена Лагранжа ... 257
5.2.4. Інтерполяційний многочлен Лагранжа з рівновіддаленими вузлами ... 260
5.2.5. Інтерполяційна схема О. Ейткена ... 261
5.2.6. Похибка інтерполювання для інтерполяційного многочлена Лагранжа ... 265
5.3. Вибір вузлів інтерполювання ... 267
5.3.1. Многочлени Чебишева. Властивості ... 267
5.3.2. Вибір вузлів інтерполювання ... 268
5.3.3. Поведінка похибки на інтервалі інтерполювання ... 270
5.4. Інтерполяційний поліном Ньютона ... 271
5.4.1. Аналіз загальної постановки задачі ... 271
5.4.2. Інтерполяційний многочлен Ньютона ... 274
5.4.3. Поділені різниці та їх властивості ... 275
5.4.4. Інтерполяційна формула Ньютона за поділеними різницями та нерівновіддаленими вузлами ... 277
5.4.5. Скінченні різниці та їх властивості ... 280
5.4.6. Формула Ньютона вперед і назад за скінченними різницями ... 283
5.5. Числове диференціювання ... 285
5.5.1. Числове диференціювання для нерівновіддалених вузлів ... 285
5.5.2. Формули числового диференціювання для рівновіддалених вузлів ... 286
5.5.3. Формули числового диференціювання методом невизначених коефіцієнтів ... 287
5.6. Раціональна інтерполяція ... 288
5.7. Інтерполювання функцій комплексної змінної ... 289
5.8. Інтерполювання функцій багатьох змінних ... 290
5.8.1. Труднощі інтерполювання функцій багатьох змінних ... 290
5.8.2. Побудова інтерполяційних формул за поділеними різницями ... 291
5.9. Збіжність інтерполяційного процесу ... 292
5.10. Інтерполювання сплайнами ... 295
5.11. Приклади розв'язку задачі інтерполяції функції ... 297
5.12. Контрольні питання ... 302
5.13. Задачі для самостійної роботи ... 303
5.14. Завдання до лабораторної роботи ... 305

Розділ 6. Основні відомості про методи наближення ... 307
6.1. Основні відомості про методи наближення ... 310
6.2. Формулювання задачі наближення функцій ... 312
6.3. Апроксимація методом найменших квадратів ... 315

6.4. Поняття сплайн-апроксимації 321
6.4.1. Аналітичне подання сплайнів 321
6.4.2. Апроксимація кубічними сплайнами 324
6.5. Розв'язування задач рівномірного сплайн-наближення 331
6.5.1. Означення та властивості рівномірного сплайн-наближення 331
6.5.2. Однорідне сплайн-наближення з заданою похибкою 332
6.5.3. Побудова неоднорідне сплайн-наближення 334
6.6. Методи обчислення рівномірного (мінімаксного) наближення функцій 335
6.6.1. Загальна постановка задачі найкращого мінімаксного наближення функцій 335
6.6.2. Апроксимаційні залежності 337
6.6.3. Доведення про найкраще рівномірне наближення функцій 341
6.7. Метод Є.Я.Ремеза 343
6.7.1. Загальна постановка задачі схеми Є.Я.Ремеза 343
6.7.2.Мінімакс. Визначення параметрів мінімаксимальної моделі за алгоритмом Є.Я.Ремеза 346
6.8. Метод Валле-Пуссена 349
6.8.1. Загальна постановка задачі 349
6.8.2. Алгоритм Валле-Пуссена 350
6.9. Застосування методу найменших квадратів багатовимірного регресійного аналізу в економіці 351
6.9.1. Загальна постановка задачі 351
6.9.2. Знаходження параметрів лінійного множинного рівняння регресії методом найменших квадратів 354
6.9.3. Стандартна похибка оцінки за рівнянням 361
6.9.4. Коефіцієнт детермінації й кореляції 362
6.9.5. Вибіркова похибка коефіцієнта множинної регресії 364
6.9.6. Вибіркова похибка множинної регресії 365
6.9.7. Похибка індивідуальної оцінки множинної регресії 367
6.9.8. Вибіркова похибка коефіцієнта множинної кореляції 369
6.9.9. Часткова регресія та кореляція 370
6.9.10. Приклад розрахунку потреб ринку на основі множинних рівнянь регресії 373
6.9.11. Приклад розрахунку потреб ринку на основі економічної оцінки 378
6.9.12. Приклад дослідження на основі багатовимірного регресійного аналізу 380
6.10. Контрольні питання 384
6.11. Задачі для самостійної роботи 385
6.12. Завдання до лабораторної роботи 387

Розділ 7. Чисельне інтегрування 390
7.1. Постановка задачі чисельного інтегрування 391
7.2. Особливості чисельного інтегрування функцій 392
7.3. Формула прямокутників. Графічна інтерпретація 394
7.4. Формула трапецій 395
7.4.1. Постановка задачі 395
7.4.2. Графічна інтерптетація 396
7.5. Формула Сімпсона (парабол) 398
7.5.1. Постановка задачі 398
7.5.2. Графічна інтерптетація 399

7.6. Економічний алгоритм реалізації принципу подвійного перерахунку та автоматичний вибір кроку інтегрування ... 402
7.7. Наближене обчислення багатократних інтегралів ... 403
7.8. Метод комірок ... 404
7.9. Послідовне інтегрування ... 407
7.10. Кубатурна формула типу Сімпсона ... 409
7.11. Методи числового інтегрування ... 414
7.11.1. Задача обчислення визначеного інтеграла ... 414
7.11.2. Три підходи квадратурної формули ... 415
7.11.3. Інтерполяційні квадратурні формули Ньютона-Котеса ... 416
7.11.4. Частинні випадки інтерполяційних формул Ньютона-Котеса ... 419
7.12. Квадратурні формули Гаусса найкращого степеня точності ... 425
7.12.1. Ідея побудови квадратурних формул Гаусса ... 425
7.12.2. Метод Гаусса знаходження вузлів КФ ... 426
7.12.3. Метод Гаусса вибору коефіцієнтів ... 428
7.12.4. Застосування формули Гаусса ... 430
7.13. Квадратні формули чисельного інтегрування Чебишова ... 432
7.13.1. Постановка задачі ... 432
7.13.2. Метод Чебишова знаходження вузлів КФ ... 433
7.13.3. Побудова квадратурних формул Чебишова ... 434
7.14. Збіжність квадратних формул ... 436
7.15. Числові методи обчислення кратних інтегралів ... 440
7.15.1. Методи повторного інтегрування для двократного інтегрування ... 440
7.15.2. Метод заміни підінтегральної функції інтерполяційним многочленом ... 442
7.16. Контрольні питання ... 444
7.17. Задачі для самостійної роботи ... 445
7.18. Завдання до лабораторної роботи ... 446
7.19. Варіанти завдання (рівняння для розв'язування) ... 447

Література ... 449

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 ІТ 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 –ий класи.

При цьому ми уважно відслідковуємо базові тенденції розвитку шкільної інфоматичної 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 переваж-ної 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 Вугт 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. Декомпозиція 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 змішення 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 на-ших професійних інтересів мова йде про галузь знань 12- «Інформаційні Технології» з шістьма спеціальностями: 121- програмна інженерія; 122 – комп'ютерні науки; 123 – комп'ютерна інже-нерія; 124 – системний аналіз; 125 – кібербезпека; 126 – інфор-маційні 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. ІСТ), 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 ТІС (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 ТІС (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 ІНФОРМА-ЦІЙНОГО 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 ІТ фахівців, - освітян, бізнесменів, науковців видав-ничий 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, д.т.н., професор
Володимир Пасічник

Вступне слово авторів

Чисельні методи – методи наближеного або точного розв'язування задач прикладної математики, які грунтуються на побудові послідовності дій над скінченною множиною чисел [76-79, 200-209, 211]. Згідно основних вимог чисельні методи мають бути стійкими та збіжними [106-107, 110]. Чисельні методи називають збіжними, якщо результати прямують до точного розв'язання задачі при прямуванні параметрів чисельних методів до певних граничних значень [113, 121-124]. Основне питання теорії чисельних методів: отримання чисельних методів, які задовольняють вимоги високої точності, стійкості та економічності. Отримання чисельних методів, що задовольняють цим вимогам, є складною задачею оптимізації чисельних методів [116-134, 137]. Статистичне опрацювання експериментальних даних зазвичай грунтується на граничних теоремах теорії ймовірностей та обчисленні порівняльних оцінок [88, 92-93, 108-109, 113-114, 140-151, 184-185]. Однак, для підвищення якості оцінок необхідна велика кількість даних, об'єм обчислень може виявитися дуже великим [160-164]. Чисельні методи націлені на скорочення об'єму обчислень при збереженні якості результатів. До найбільш ефективних чисельних методів в цій галузі відносяться методи, які застосовують швидке перетворення Фур'є. Для розв'язання задач апроксимації та обчислення функцій різних класів застосовують чисельні методи інтерполювання, найменших квадратів, ортогоналізації, врівноваження значень, умовної мінімізації тощо [135-136, 168-176]. Найбільш актуальними є методи кусково-многочленної та раціональної сплайнової апроксимації, а також адаптивної апроксимації та нелінійної за параметром апроксимації [64-65, 71-80, 152-159, 177-182, 210, 212, 220-230].

Чисельне інтегрування та диференціювання починається із визначення відповідних операцій. Однак, з урахуванням необхідності економії об'єму обчислень та з урахуванням некоректності задачі диференціювання з'являється велика 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 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'язання різних класів рівнянь та багатьох інших задач зводиться до задач мінімізації функцій та функціоналів за наявності або відсутності обмежень. Чисельні методи розв'язання 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 і 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.

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 подання 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 округлення, 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:

6.1. Методи оптимізації.

6.2. 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 їхньої складності кінцевої мети досягають постановкою експерименту або методом математичного моделювання. Кожен з цих методів має переваги і 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 подання. 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 і 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 власне 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 через те, 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 [186-190]. Незважаючи на відмінності в методології, до чисельного експерименту щільно дотикають фізичний експеримент/дослідження, де необхідна оцінка достовірності здобутих результатів [191-198].

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 [1-63, 66-70]. У моделі враховують найважливіші моменти і взаємозв'язки, найхарактерніші для досліджуваного реального об'єкта [81-87, 89-91]. Разом з тим внаслідок заміни реального об'єкта відповідною йому математичною моделлю стало можливим сформулювати задачу як математичну і скористатися для її розв'язання тим чи іншим математичним апаратом [94-105, 111-112, 115, 152-159].

Алгоритм – це зрозумілий і точний припис (вказівка) виконавцеві здійснювати послідовність дій, що спрямовані на досягнення зазначеної мети або розв'язання поставленої задачі [138-139, 165-167, 183, 199]. 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 і опису її на алгоритмічній машинній мові здійснюють чисельну реалізацію задачі на комп'ютері [213-221]. 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.

Розділ 1. Математичне моделювання

- Чисельні методи та використання персонального комп'ютера для розв'язування прикладних задач
- Наближений аналіз. Джерела та класифікація похибок
- Обчислювальна задача
  - Аналіз постановки задачі
  - Приклади постановки задачі обчислення
- Чисельне розв'язування коректних задач. Структура похибки розв'язку
  - Обчислювальна задача. Похибки
  - Похибка заокруглення
  - Похибка функції
- Контрольні питання
- Задачі для самостійної роботи

В розділі розглянуто особливості та переваги використання математичного моделювання для дослідження явищ та процесів [1-64, 66-70, 81-87, 89-91, 94-105, 111, 115, 138-139, 146, 171-191, 195, 212-230]. Визначено особливості аналізу результатів моделювання та оцінювання похибки цих результатів [65, 71, 76-80, 88, 92-93, 106-110, 113-114, 116-133, 137, 140-145, 147-151, 161-164, 168-169, 192-194, 196-197, 201-209, 211]. Моделювання (англ. scientific modelling, simulation, нім. Modellieren n, Modellierung f, Simulation f, рос. моделирование) – це метод дослідження явищ і процесів, що ґрунтується на заміні конкретного об'єкта досліджень (оригіналу) іншим, що подібний до нього (моделлю). Моделювання в широкому сенсі – це особливий пізнавальний процес, метод теоретичного та практичного опосередкованого пізнання, коли суб'єкт замість безпосереднього об'єкта пізнання вибирає чи створює подібний із ним допоміжний об'єкт-замісник (модель), досліджує його, а здобуту інформацію переносить на реальний предмет вивчення. Моделювання – це процес створення та дослідження моделі, а модель – засіб, форма наукового пізнання. Під моделлю розуміють об'єкт будь-якої природи (подумки уявлена або матеріально реалізована система), який, відображаючи чи відтворюючи в певному сенсі об'єкт дослідження, здатний заміщати його так, що вивчення моделі дає нову інформацію про об'єкт. Моделювання зазвичай використовують в дослідженні систем різної природи, але особливого значення воно набуває в соціальному управлінні, в рамках методології системного підходу.

Моделювання – одна з основних категорій теорії пізнання. На ідеї моделювання базується будь-який метод наукового дослідження. Основні види моделювання – фізичне і математичне. Характерною рисою розглядуваного методу є можливість відтворення моделлю відповідно до завдань дослідження тих чи інших істотних властивостей, структур досліджуваного об'єкта, взаємозв'язків і відносин між його

ТАЛЬНОВСЬКИЙ НАЦІОНАЛЬНИЙ
ІМЕНІ ТАРАСА ШЕВЧЕНКА
БІБЛІОТЕКА

елементами. В процесі пізнання модель є наступною за об'єктом, є його копією, а у відтворенні, конструюванні, навпаки, об'єкт є наступним за моделлю, копіюючи її.

Модель фіксує існуючий рівень пізнання про досліджуваний об'єкт. Неможливо створити універсальну модель, яка могла б відповісти на всі запитання, що викликають інтерес; кожна з них дає лише наближений опис явища, причому в різних моделях знаходять відображення різні його властивості. Моделювання доречно використовувати, коли дослідження реального об'єкту з усією сукупністю його властивостей недоцільно, незручно або неможливо.

Моделювання – це метод, а модель – форма, засіб наукового пізнання. Метод моделювання володіє загальністю, оскільки змоделювати можна будь-який об'єкт: така можливість рівнозначна визнанню принципової їх пізнаваності. Перевірка достовірності моделювання – процес визначення, чи модель або виконувана імітація точно є детальним концептуальним описом, прийнятим розроблювачем. Перевірка достовірності також оцінює ступінь відповідності моделі або імітації змісту і проводиться з використанням прийнятих методів технології програмування.

Перевірка правильності – визначення, чи здається модель або імітація розумною людям, що добре інформовані щодо системи при її дослідженні, заснованому на експлуатаційних показниках моделі. Цей процес не розглядає програмний код або логіку, а скоріше розглядає входи і виходи для гарантії їх уявної реалістичності або показності. Це процес визначення ступеня точності, з яким модель або імітація відображає реальний чи створюваний світ. Фізичне моделювання, при якому модель і об'єкт, що моделюється, мають одну і ту ж фізичну природу.

Математичне моделювання – моделювання, при якому модель є системою математичних співвідношень, що описують певні технологічні, економічні чи інші процеси. Найчастіше застосовують два способи математичного моделювання:

- аналітичний, що передбачає можливість точного математичного опису строго детермінованих систем,
- ймовірнісний, що дозволяє отримати не однозначне рішення, а його імовірнісну характеристику (наприклад, параметрів технологічного процесу).

Математичне моделювання займає провідне місце в системному аналізі, а також широко застосовується для опису технологічних процесів.

Математичне моделювання (рос. моделирование математическое; англ. mathematical simulation, нім. mathematische Modellierung f) – метод дослідження процесів або явищ шляхом створення їхніх математичних моделей і дослідження цих моделей. В основу методу покладено ідентичність форми рівнянь і однозначність співвідношень між змінними в рівняннях оригіналу і моделі, тобто, їх аналогію. Математичні моделі досліджуються за допомогою комп'ютерів.

На початку 60-их років розроблено один із методів математичного моделювання – квазіаналогове моделювання. Цей метод полягає в дослідженні недосліджуваного явища, а явища або процесу іншої фізичної природи, що описується співвідношеннями, які еквівалентні відносно отримуваних результатів.

Математичне моделювання зазвичай застосовують всі природничі і суспільні науки, що використовують математичний апарат для одержання спрощеного опису реальності за допомогою математичних понять. Математичне моделювання

дозволяє замінити реальний об'єкт його моделлю і потім дослідити останню. Як і у разі будь-якого моделювання, математична модель не описує явище абсолютно адекватно, що залишає актуальним питання про застосовність отриманих таким шляхом даних. Математичне моделювання широко застосовується в інформатиці.

Формальна класифікація моделей грунтується на математичних засобах, що використовують для розв'язування поставлених задач. Розрізняють моделі:

1) лінійні або нелінійні моделі;
2) зосереджені або розподілені системи;
3) детерміновані або стохастичні;
4) статичні або динамічні.

Існує ще декілька підходів. Разом з тим, кожна побудована модель є лінійною або нелінійною, детермінованою або стохастичною. Звичайно, можливі й змішані типи: у одному відношенні зосереджені (за частиною параметрів), в іншому – розподілені моделі тощо.

1.1. Чисельні методи та використання комп'ютера для розв'язування прикладних задач

Розглядаючи *математичний аналіз явища* як приклад *теоретичного експерименту*, із загальних та достатньо природних міркувань процес *математичного моделювання* розбивається на декілька етапів:

- **Формулювання математичної моделі явища.**

Математична модель будь-якого явища, що досліджується, через його надзвичайну складність, повинна включати важливі сторони процесу для задачі, що розглядається, його істотні характеристики та формалізовані зв'язки, які потрібно враховувати. Як правило, *математична модель* досліджуваного фізичного явища формулюється у вигляді *рівнянь математичної фізики*. На цій стадії аналізу це суттєво нелінійні, багатовимірні системи рівнянь, які мають велику кількість невідомих та параметрів. Якщо *математична модель* недостатньо деталізована, то які б ми не застосували методи для подальших розрахунків, отримані результати будуть *ненадійними*, а в окремих випадках і зовсім *невірними*.

- **Проведення математичного дослідження отриманої моделі та отримання відповідного розв'язку.**

На цьому етапі моделювання, в залежності від складності моделі, що розглядається, застосовують різні підходи для її дослідження та вкладається різний зміст в поняття розв'язку задачі. Доведення теорем існування та єдності у визначеному сенсі розв'язує задачу, але, воно часто є некоректним та не дозволяє вирішити проблему дослідження якісної поведінки рішення та оцінки його кількісних характеристик. Для найбільш загальних, формальних та нескладних моделей вдається отримати їх аналітичний розв'язок. Використання засобів символьних розрахунків на ПК (MATHEMATICA, MathCAD, MatLab, інтелектуальні калькулятори тощо) суттєво покращує та пришвидшує розрахунки, ніж традиційне поле діяльності на папері з олівцем.

Для більш точних та складних моделей аналітичний розв'язок вдається отримати порівняно рідко. При теоретичному аналізі задачі, в такій ситуації, зазвичай використовують наближені математичні методи, наприклад, розклад за малим параметром, усереднення, дослідження різних асимтотик тощо. Ці прийоми дозволяють подати наближений розв'язок в аналітичній формі та за його допомогою отримати задовільні чисельні результати. Нарешті, для найбільш точних та складних моделей основними методами розв'язування є чисельні методи, які вимагають проведення великого об'єму розрахунків на ПК. Ці методи дозволяють отримати хороший кількісний і навіть якісний результат у описі моделі. Але є у них і принципові недоліки – як правило, мова йде про розгляд деякого частинного розв'язку. У зв'язку із поділом методів на точні та наближені доцільно ввести:

1. Похибка задачі, тобто похибка вхідних даних (число π, e).
2. Похибка методу (лише для наближених методів).
3. Похибка заокруглення.

Для наближених методів використовують поняття збіжності.

Метод називається *збіжним*, якщо послідовність наближень отриманих цим методом є збіжною. У граничному випадку збіжний наближений метод дає точний результат. Метод називається *стійким*, якщо невелика похибка вхідних даних викликає невелику похибку вихідних даних. Метод називається *коректним*, якщо він є збіжним та стійким, і для розв'язування задачі використані всі початкові дані. Для дослідження збіжності та стійкості використовують числові ознаки, а саме, розглядають *абсолютну* та *відносну похибки*. Перед тим, як перейти до розгляду конкретних методів розглянемо поняття математичного моделювання.

Математична модель – це опис того чи іншого явища (процесу) за допомогою математичних термінів, формул, рівнянь (алгебраїчних, диференціальних, інтегральних). Для того, щоб розв'язати довільну задачу необхідно провести так званий "обчислювальний експеримент". Цей термін передбачає три складові (рис. 1.1). Наведена схема на рис. 1.2 частково відображає взаємозв'язки етапів математичного моделювання.

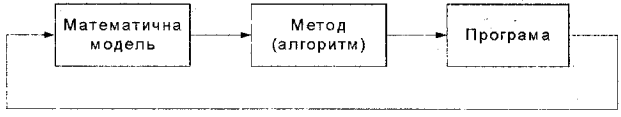


Рис. 1.1. Схема процесу математичного моделювання

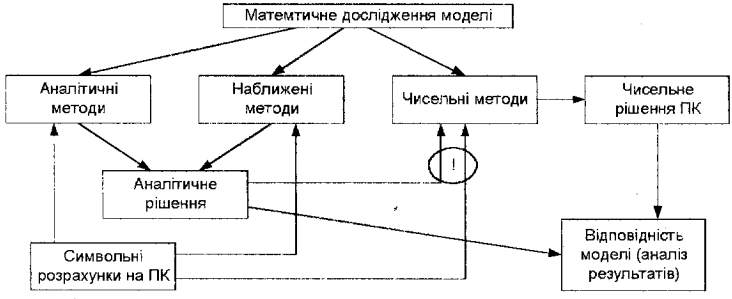


Рис. 1.2. Схема процесу математичного дослідження моделі

Кожний із етапів математичного дослідження моделі пов'язаний із використанням *чисельних методів* та отриманням *чисельного розв'язку* задачі.

- **Аналіз відповідності запропонованої моделі**, тобто осмислення результатів розв'язку, співставлення отриманного розв'язку із наявними даними фізичного експерименту. На цьому етапі розв'язується питання відповідності математичної моделі та проведеного дослідження. «Добре» узгодження з експериментом звично засвідчує правильність вибору моделі. У іншому випадку необхідні додаткові уточнення, зміни тощо, повторення попередніх етапів дослідження. Використання ПК в процесі математичного дослідження моделі вимагає специфічних, чисельних методів, тобто такої інтерпретації математичної моделі, яка може бути реалізована на комп'ютері – назвемо її *дискретною* (*обчислювальною*) моделлю. Так як комп'ютер виконує лише арифметичні та логічні операції, то для реалізації *обчислювальної моделі* необхідно розроблення відповідного *обчислювального алгоритму*. Подальша послідовність дій – це програмування, розрахунок на комп'ютері, опрацювання результатів розрахунку.

В рамках лекційного курсу зупинимося на окремих проблемах чисельних методів при аналізі порівняно простих та класичних математичних моделей.

1.2. Наближений аналіз. Джерела та класифікація похибок

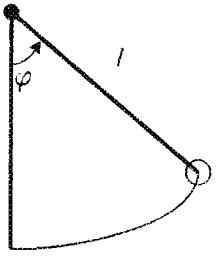


Рис. 1.3. Коливання маятника

Числові розв'язки будь-якої математичної задачі, як правило, ніколи точно не співпадають з її точними розв'язками. Одержуються результати з тією чи іншою точністю.

Означення 1.1. *Відхилення точного розв'язку математичної задачі від наближеного розв'язку називається похибкою розв'язку задачі.*

Виникнення похибки обумовлюється наступними причинами:

1) математичний опис задачі є неточним (наприклад, вихідні дані задано неточно);
2) метод, що застосовується для розв'язання математичної задачі є неточним, бо точно розв'язати, як правило, математичну задачу неможливо;
3) при вводі та виводі даних з комп'ютера відбувається заокруглення чисел.

Тому відповідно похибка задачі складається з трьох похибок:

1) неусувна похибка або похибка математичної моделі;
2) похибка наближеного методу;
3) обчислювальна похибка.

Розглянемо наступну модельну задачу.

Приклад 1.1. Математичний маятник (рис. 1.3.). Нехай маємо математичний маятник, що починає рух в момент $t=t_0$. Треба знайти кут відхилення $\varphi = \varphi(t)$ в момент часу $t>t_0$.

Розв'язок. Математичною моделлю є диференційне рівняння коливання маятника [76-79]

$$l\frac{d^2\varphi}{dt^2}+\mu\frac{d\varphi}{dt}+g\sin\varphi=0, \tag{1.1}$$

$$\varphi(t_0)=\varphi_0;\quad \varphi'(t_0)=\varphi_0', \tag{1.2}$$

де l – довжина маятника, g – прискорення сили тяжіння, μ – коефіцієнт тертя.

Неусувна похибка пояснюється тим, що:

а) $l,\ g,\ \mu,\ t_0,\ \varphi(t_0),\ \varphi'(t_0)$ визначені неточно (прилади не можуть абсолютно вимірювати ці величини);

б) модель (1.1) неточна, бо тертя залежить від швидкості, взагалі кажучи, нелінійно. У нас лінійна модель. Її можна покращити, якщо (1.1) замінити на

$$l\frac{d^2\varphi}{dt^2}+\mu(\dot\varphi)\frac{d\varphi}{dt}+g\sin\varphi=0, \tag{1.3}$$

яке еквівалентне системі

$$\dot y(t)=-\frac{g}{l}\sin\varphi-\frac{\mu(\dot\varphi)}{l}y(t). \tag{1.4}$$

Похибка методу виникає заміною (1.1), (1.2) на іншу задачу (1.3), (1.2), бо рівняння (1.3) точно не розв'язується. Рівняння (1.3) розв'язується, наприклад, методом Ейлера [76-79].

$$\Phi_i=\Phi_i+h\Psi_i+o\left(h^2\right), \tag{1.5}$$

$$\Phi(t)\equiv\left(\varphi(t),y(t)\right)^T;\Psi(t)\equiv\dot\Phi(t)\equiv\begin{pmatrix} y(t)\\ -\frac{g}{l}\sin\varphi-\frac{\mu(\dot\varphi)}{l}y(t)\end{pmatrix}.$$

Обчислювальна похибка виникає за рахунок заокруглення чисел у комп'ютері. Опишемо формально ці похибки. Нехай A – точне значення кута $\varphi(t)$ у фізичному реальному маятнику; $\widetilde{A}$ – значення цього кута в математичній моделі (в даному випадку розв'язок $\varphi(t)$ задачі (1.1), (1.2)). $\widetilde{A}_h$ – розв'язок задачі (1.1), (1.2), якщо розв'язувати наближеним методом (1.5). $\widetilde{A}_h^*$ – наближення, що одержуємо при обчисленні на комп'ютері. Тоді $\beta_1=\widetilde{A}-A$ неусувна похибка; $\beta_2=\widetilde{A}_h-\widetilde{A}$ похибка методу; $\beta_3=\widetilde{A}_h^*-\widetilde{A}_h$ обчислювальна похибка. Повна похибка $\beta_0=\widetilde{A}_h^*-A$ дорівнює

різниці між реальним та точним розв'язком задачі відхилення маятника від положення рівноваги. Тоді,

$$\beta_0 = \beta_1 + \beta_2 + \beta_3 \tag{1.6}$$

У багатьох випадках зручно трактувати не різниці $\beta_1, \beta_2, \beta_3$, а деякі їх міри. Наприклад: $\beta_0 = \left|\widetilde{A}_h^* - A\right|; \beta_1 = \left|\widetilde{A} - A\right|; \beta_2 = \left|\widetilde{A}_h - \widetilde{A}\right|; \beta_3 = \widetilde{A}_h^* - \widetilde{A}_h$. При таких позначеннях $\beta_0 < \beta_1 + \beta_2 + \beta_3$. Тобто повна похибка оцінюється справа.

1.2. Обчислювальна задача

1.2.1. Аналіз постановки задачі

Звичайно задачу обчислення величини y за відомою величиною x записують із врахуванням досліджуваних причинно-наслідкових зв'язків у вигляді

$$y = A(x), \tag{1.7}$$

де $y \in Y$, $x \in X$ – елементи відповідних функціональних просторів (якщо не обумовлено, то X, Y, як правило, *лінійні*, *нормовані*, *повні*, тобто *банахові* простори); A – оператор (правило), що реалізує обчислення.

Першочергово досліджують коректно поставлені задачі розрахунку.

Означення 1.2. *Обчислювальна задача $y = A(x)$ є коректно поставленою, якщо для будь-яких вхідних даних із деякого класу розв'язок задачі існує, єдиний та стійкий за вхідними даними (тобто неперервно залежить від вхідних даних задачі).*

В сформульовановану понятті *коректності* поставленої задачі (за Адамаром) враховані достатньо природні вимоги, дійсно: щоб чисельно розв'язувати задачу, необхідно бути впевненим, що її розв'язок *існує*. Так як дослідження носить принципово наближений характер, то для конкретних умов і вимоги *єдності* розв'язку необхідна вимога *стійкості* розв'язку. Зробимо декілька зауважень щодо *стійкісті*. Тобто, чи розв'язок y задачі (1.7) відповідає вхідним даним x. Маємо неточні вхідні дані із похибкою δx, тобто $x + \delta x$ і знаходимо розв'язок

$$y + \delta y = A(x + \delta x). \tag{1.8}$$

Ця похибка вхідних даних породжує неусувну похибку розв'язку

$$\delta y = A(x + \delta x) - A(x). \tag{1.9}$$

Якщо розв'язок неперервно залежить від вхідних даних, то

$$\|\delta y\| \to 0 \text{ завжди при } \|\delta x\| \to 0 \tag{1.10}$$

і задача (1.7) є стійкою за вхідними даними.

Відсутність стійкості означає, що навіть невеликим похибкам δx можуть відповідати великі похибки δy, тобто побудований розв'язок при виконанні обчислень сильно відрізняється від істинного. Застосовувати безпосередньо до такої нестійкої задачі чисельні методи немає сенсу. Але і не всяку формально стійку задачу зручно вирішувати практично. Нехай має місце оцінка

$$\|\delta y\| \le C \cdot \|\delta x\|, \text{ але } C - \text{ велике.} \tag{1.11}$$

Задача формально стійка, але *неусувна похибка* розв'язку може бути великою. Це випадок *поганої обумовленості* або *слабкої стійкості* задачі обчислення. Наведемо декілька прикладів постановки *задачі обчислення* (1.7).

1.2.2. Приклади постановки задачі обчислення

Приклад 1.2. Задача знаходження коренів поліному. Розглянемо деякий поліном степеня n в такому вигляді (старший коефіцієнт рівний одиниці):

$$p_n(x) \equiv x^n + a_1 x^{n-1} + a_2 x^{n-3} + \ldots + a_{n-1} x + a_n \tag{1.12}$$

з комплексними коефіцієнтами $(a_k, x \in C)$. Необхідно знайти його корені.

Розв'язок. Нехай E^n – n-мірний, комплексний евклідовий простір. Припустимо, що компоненти деякого вектора $\vec{z} = \{z_1, z_2, \ldots, z_n\}$ цього простору є коренями поліному $p_n(x)$, тобто

$$p_v(z_i) = 0, i = \overline{1,n}. \tag{1.13}$$

Тоді, згідно теореми Безу, можемо $p_n(x)$ записати у вигляді:

$$\begin{aligned} p_n(x) &\equiv x^n + a_1 x^{n-1} + a_2 x^{n-3} + \ldots + a_{n-1} x + a_n = \\ &= (x - z_1)(x - z_2)\ldots(x - z_n) = \prod_{i=1}^{n} (x - z_i) \end{aligned} \tag{1.14}$$

Звідси отримаємо відомі формули Вієтта:

$$a_k = (-1)^k \sigma_k, \; k = \overline{1,n}, \tag{1.15}$$

де σ_k – елементарні, симетричні відносно $z_1, z_2, \ldots, z_n$ однорідні функції k-го порядку (кожне σ_k має C_n^k доданків):

$$\begin{cases} \sigma_1 = z_1 + z_2 + \ldots + z_n \\ \sigma_1 = z_1 z_2 + z_1 z_3 + \ldots + z_1 z_n + z_2 z_3 + \ldots + z_{n-1} z_n \\ \quad \ldots \\ \sigma_1 = z_1 z_2 \ldots z_n. \end{cases} \tag{1.16}$$

Таким чином формули Вієтта (1.15-1.16) співставляють кожному вектору $z \in E^n$ вектор $\vec{a} = \{a_1, a_2, ..., a_n\} \in E^n$ того ж простору, тобто визначають відображення $V : E^n \Rightarrow E^n$ простору E^n на себе. За допомогою цього відображення V задача визначення коренів поліному $p_n(x)$ формулюється наступним чином:

Для заданого вектору $\vec{a}$ знайти вектор $\vec{z} \in E^n$ такий, що

$$V(\vec{z}) = \vec{a}\,. \tag{1.17}$$

Відображення V *взаємно однозначне та взаємно неперервне*, тобто задача (1.17) є *коректною*.

Приклад 1.3. Основна задача лінійної алгебри. Нехай задано матрицю $A_{(p\times q)} = \left\| a_j^i \right\|_q^p$ та два евклідових простори E^p та E^q. Визначено відображення $A : E^q \Rightarrow E^p; \vec{y} \in E^p; \vec{x} \in E^q$, ($\vec{x}$, $\vec{y}$ – стовпці відповідних розмірностей).

Розв'язок. *Основна задача лінійної алгебри* полягає в тому, щоб за заданим вектором $\vec{f} \in E^p$ знайти вектор $\vec{x} \in E^q$ такий, що

$$A\vec{x} = \vec{f} \;\; A\vec{x} = \vec{f}\,. \tag{1.18}$$

Задача (1.18) полягає у знаходженні розв'язку *системи лінійних алгебраїчних рівнянь* (СЛАР). Задача знаходження розв'язку СЛАР детально досліджувалась в курсі *лінійної алгебри*:

1) якщо $p = q$ та $\det A \neq 0$, то задача (1.18) поставлена *коректно* (її розв'язок знаходиться за формулами Крамера);

2) якщо система (1.18) сумісна ($rangA = rangA'$), то розв'язок не єдиний. В іншому випадку рішення зовсім відсутнє, тобто задача (1.18) в цих випадках *некоректно поставлена.*

Приклад 1.4. Задача Коші для звичайного диференційного рівняння. Нехай необхідно знайти рішення звичайного диференційного рівняння (ЗДР), яке відповідає початковій умові $y(a)=b$:

$$\begin{cases} \dfrac{dy}{dx} = f(x,y), \quad a < x \le c, \\ y(a) = b, \end{cases} \tag{1.19}$$

де a,b – задані числа; $f(x,y)$ – визначена на відрізку $C = \{(x,y); a \le x \le c; y \in (-\infty;\infty)\}$ та C задовольняє умові теореми про неперервність розв'язку (1.19) на відрізку $[a;c]$.

Розв'язок. Позначимо через R_0 множину всіх можливих розв'язків задачі Коші (1.19), що відповідають різним значенням початкової умови b. Визначимо відображення $K : R_0 \Rightarrow R^1$, враховуючи, що

$$K(y(x)) = y(a)\,,\;\; \forall y \in R_0\,. \tag{1.20}$$

Тоді розв'язок задачі Коші для ЗДР (1.20) можна сформулювати так:
За заданим числом b знайти функцію y(x) таку, що

$$K(y(x)) = b\,. \tag{1.21}$$

В курсі *диференційних рівнянь* було *доведено коректність* задачі (1.20).

Кількість розглянутих прикладів *задачі обчислення* можна було б збільшити, але обмежимося розглянутими прикладами постановки *задачі обчислення*.

1.3. Чисельне розв'язування коректних задач. Структура похибки розв'язку

1.3.1. Обчислювальна задача. Похибки

Звернемося знову до *задачі обчислення* (1.22)

$$y = A(x)\,. \tag{1.22}$$

В розглянутих прикладах (1.12)-(1.21) відповідне правило A, що реалізує обчислення, задане явно неконструктивно. Мова йде про знаходження обернених операторів V^{-1}, A^{-1}, K^{-1}, точніше про чисельну реалізацію оберненого відображення для (1.12)-(1.21). Така ситуація типова та зайвий раз показує, що, як правило, знаходження оберненого оператора не може бути реалізоване просто. Щоб усунути ці складності задачу (1.22) замінюють іншою, близькою до неї задачею, але яка легко розв'язується чисельно. При цьому в першу чергу аналізують питання про внесення похибки в розв'язок поставленої задачі.

Є чотири основних джерела похибки результату обчислень: *математична модель*; *початкові дані* задачі; *наближений метод* та *похибка при реалізації обчислення* (зокрема *похибка заокруглення*):

1) $\delta_1 y$ – *похибка математичної моделі*, що пов'язана із фізичними припущеннями при виборі математичної моделі; на аналізі цієї похибки детально зупинятися не будемо;

2) $\delta_2 y$ – *похибка вхідних даних*, що породжує *неусувну похибку розв'язку*

$$\delta_2 y = A(x + \delta x) - A(x)\,; \tag{1.23}$$

3) $\delta_3 y$ – *похибка методу*. Вираз $A(x)$ не може бути просто чисельно реалізований. Задачу $y = A(x)$ замінюють близькою задачею

$$\overline{y} = \overline{A}(\overline{x})\,. \tag{1.24}$$

Переходимо до інших функціональних просторів $X, Y \Rightarrow \overline{X}, \overline{Y}$, елементи яких дозволяють порівняно просту чисельну реалізацію. Відповідним чином змінюється

й відображення $A \Rightarrow \overline{A}$. При цьому природньо вимагати, щоб задача (1.24) була *коректна* та щоб розв'язок $\overline{y}$ був близьким до розв'язку y. Величина

$$\delta_3 y = y - \overline{y} = A(x) - A(\overline{x}), \tag{1.25}$$

є похибкою методу.

4) $\delta_4 y$ – *похибка обчислень*. При чисельній реалізації $\overline{y}$, можливо отримають елемент $\widetilde{y}$, так як проміжні результати заокруглювались. Тоді *похибка обчислень* може бути записана у вигляді

$$\delta_4 y = \overline{y} - \widetilde{y} = \overline{A}(\overline{x}) - \widetilde{y}. \tag{1.26}$$

Корисно відразу ж сформулювати деякі емпіричні правила, котрих дотримуються при реалізації задачі розрахунку:

$$\|\delta_2 y\| \approx (2 \div 5)\|\delta_3 y\| >> \|\delta_4 y\|. \tag{1.27}$$

1) При проведенні обчислень потрібно досягнути того, щоб похибка методу $\delta_3 y$ була б в декілька разів меншою за *неусувну похибку* розв'язку $\delta_2 y$.

2) *Похибка обчислень* $\delta_4 y$ повинна бути істотно меншою ніж всі інші похибки розв'язку, тобто обчислення потрібно вести з такою кількістю значущих цифр, щоб похибка заокруглення була істотно меншою ніж всі інші похибки.

Тепер можемо ще раз окреслити коло питань, що розглядаються в рамках лекційного курсу «Чисельні методи в інформатиці» – це

1) побудова *дискретної* (або обчислювальної) моделі $\{\overline{X}, \overline{x}, \overline{A}\}$;

2) розроблення на її основі відповідних алгоритмів розв'язку наближеної задачі обчислень $\overline{y} = \overline{A}(\overline{x})$;

3) аналіз похибки методу $\delta_3 y$ та частково обчислювальної похибки $\delta_4 y$ алгоритму, що реалізує обчислення $\overline{A}$.

Предмет лекційного курсу може бути і більш змістовним та великим, але є обмеження, що визначаються специфікою навчального плану дисципліни.

1.3.2. Похибка заокруглення

Зупинимось детально на аналізі *обчислювальної похибки* $\delta_4 y$, яка існує через похибки заокруглення при реалізації чисельного алгоритму.

1. Похибка одиничного заокруглення. В сучасних комп'ютерах дійсні числа подають у формі *з плаваючою комою*, тобто число a в позиційній системі числення з основою r записувють у вигляді r-ого дробу

$$\begin{aligned} a &= sign\, a(a_n a_{n-1} \ldots a_1 a_0, a_{-1} a_{-2} \ldots)_r = \\ &= sign\, a(a_n r^n + \ldots + a_1 r + a_0 + \frac{a_{-1}}{r} + \frac{a_{-2}}{r^2} + \ldots), \end{aligned} \tag{1.28}$$

тоді таку форму запису числа a називають *поданням з фіксованою комою*. Тут $a_k \in \{0;1;...;(r-1)\}$, тобто r-і числа (r=2 в двійковій системі чисення, r=8 – у вісімковій, r=10 – у десятковій, r=16 – у шістнадцятковій тощо).

Подання числа a у формі *з плаваючою комою* або *нормалізоване подання* означає його запис у вигляді

$$a = sign\, aMr^p = sign\, a \cdot r^p \cdot \left(\frac{b_1}{r} + \frac{b_2}{r^2} + ...\right), \tag{1.29}$$

де p – порядок числа (ціле); M – мантиса числа a, причому $\frac{1}{r} \le M < 1$, тобто перша r-а цифра у записі мантиси b_1 нерівна нулю.

В сучасних комп'ютерах в якості основи системи числення r вибирається двійка (r=2). Тоді, якщо для запису мантиси відводиться тільки t двійкових розрядів, то це дозволяє із діапазону $[M_0; M_\infty = M_0^{-1}]$ (для додатніх чисел) записати лише скінчену кількість раціональних чисел, а всі інші дійсні числа підлягають заокругленню при їх подані в комп'ютері.

Точність подання числа a за допомогою заокругленого числа $\tilde{a}$ характеризується відносною похибкою заокруглення

$$\delta_a = \frac{|a-\tilde{a}|}{|a|} = \frac{\Delta_a}{|a|}. \tag{1.30}$$

При простому способі заокруглення *усіканням*, коли всі зайві розряди мантиси просто відкидаються, можна легко отримати оцінку величини відносної похибки δ_a одиничного заокруглення. Дійсно

$$|a-\tilde{a}| = 2^p \left|\frac{b_{t+1}}{2^{t+1}} + ...\right| \le 2^p \cdot \frac{1}{2^{t+1}}\left(1 + \frac{1}{2} + ...\right) = 2^{p-t}. \tag{1.31}$$

Тут при оцінці всі двійкові цифри в залишку замінені одиницею

$$b_k \le 1,\ k \ge t+1. \tag{1.32}$$

З іншого боку $|a| \ge 2^p \cdot (1/2)$. Ми припускаємо a_i=0, при $i \ge 2$; a_1=1 завжди. Таким чином для похибки одиничного заокруглення отримаємо

$$\delta_a = \frac{|a-\tilde{a}|}{|a|} = \frac{\Delta_a}{|a|} \le \frac{2^{p-t}}{2^{p-1}} = 2^{-(t-1)}. \tag{1.33}$$

Більш точний спосіб заокруглення дає для похибки одиничного заокруглення вдвічі меншу оцінку через *машинний епсилон*

$$\delta_a \le 2^{-t} = \varepsilon_M. \tag{1.34}$$

Означення 1.3. *Відносна похибка подання числа з плаваючою комою на комп'ютері визначається числом розрядів мантиси, що не перевищує* $\varepsilon_M = 2^{-t}$.

Спираючись на оцінку (1.34) *можемо* вважати, що саме число a та його заокруглене значення $\widetilde{a}$ пов'язані відношенням

$$\widetilde{a} = fl(a) = a(1+\varepsilon_a), \tag{1.35}$$

де $|\varepsilon_a| \le \varepsilon_M = 2^{-t}$. Але зауважимо, що для чисел $|a| < M_0$ в результаті заокруглення отримаємо $\widetilde{a} = 0$,та для цих чисел $\varepsilon_a = -1$.

Означення 1.4. *Арифметичний пристрій (АП) сучасних комп'ютерів сконструйований таким чином, що будь-яка арифметична операція при наступному заокругленні дає відносну похибку не більше* ε_M.

Для оцінки впливу похибки заокруглень на результат того чи іншого розрахункового алгоритму користуються припущенням, що *результат обчислень, спотворений похибкою заокруглення співпадає із результатом точного розрахунку за таким же алгоритмом, але з іншими –* $\widetilde{x}$*, вхідними даними.* Таким чином,

$$\widetilde{y} = \overline{A}(\widetilde{x}), \tag{1.36}$$

$$\delta_4 y = \overline{y} - \widetilde{y} = \overline{A}(\overline{x}) - \overline{A}(\widetilde{x}). \tag{1.37}$$

Це припущення дозволяє пов'язати аналіз *розрахункової похибки* $\delta_4 y$ з аналізом *стійкості* алгоритму $\overline{A}$ за *вхідними даними.*

Приклад 1.5. Розглянемо задачу про знаходження добутку n співмножників

$$z_n = \prod_{k=1}^{n} y_k . \tag{1.38}$$

Нехай обчислення реалізовані за алгоритмом $\overline{A}$ наступним чином:

$$\begin{cases} z_k = y_k \cdot z_{k-1}, \quad k = \overline{1,n} \\ z_0 = 1. \end{cases} \tag{1.39}$$

Розв'язок. Припустимо, що в результаті заокруглення замість точного значення z_{k-1} отримано значення $\widetilde{z}_{k-1}$. Тоді замість величини z_k отримаємо величину

$$\widetilde{z}_k = fl(y_k \cdot \widetilde{z}_{k-1}) = y_k \cdot \widetilde{z}_{k-1}(1+\varepsilon_k);\ |\varepsilon_k| \le \varepsilon_M . \tag{1.40}$$

Таким чином отримаємо алгоритм $\overline{A}$

$$\begin{cases} \widetilde{z}_k = \widetilde{y}_k \cdot \widetilde{z}_{k-1}, \quad k = \overline{1,n} \\ \widetilde{z}_0 = 1. \end{cases} \tag{1.41}$$

Структура отриманого алгоритму $\overline{A}$ підтверджує сформульоване припущення. Оцінимо результуючу відносну похибку

$$\delta_{z_n} = \left|\frac{z_n - \tilde{z}_n}{z_n}\right| = \frac{\left|\prod_{k=1}^{n} y_k - \prod_{k=1}^{n}(1-\varepsilon_k)y_k\right|}{\left|\prod_{k=1}^{n} y_k\right|} \le (1-\varepsilon_k)^n - 1 = n\varepsilon_M + O(\varepsilon_M^2), \qquad (1.42)$$

або, не враховуючи доданки другого та вищих порядків за ε_M, отримаємо

$$\delta_{z_n} \le n\varepsilon_M . \qquad (1.43)$$

1.3.3. Похибка функції

У практичних задачах виникає така проблема. Шукана величина $y=A(x)$ є функцією аргументів $x=(x_1, x_2, \ldots x_n)$, тобто

$$y = A(x_1, x_2, \ldots, x_n), \qquad (1.44)$$

де $x \in G \subset E^n$. Вимагається одержати наближення до y та оцінити похибку.

Означення 1.5. *Нехай $\bar{y}$ – наближене значення величини y, тоді граничною абсолютною похибкою $\overline{A}(\bar{y})$ називають найкращу оцінку похибки величини $\bar{y}$ у такому розумінні*

$$\overline{A}(\bar{y}) \equiv \sup_{x \in G} |y - \bar{y}| . \qquad (1.45)$$

Означення 1.6. *Граничною відносною похибкою називають величину*

$$\delta(\bar{y}) \equiv \frac{\overline{A}(\bar{y})}{\bar{y}} . \qquad (1.45)$$

Теорема 1.1. Нехай задано

1) n-вимірний паралелепіпед $G = \{x \in E^n \mid |x_j - \bar{x}_j| \le \Delta(\bar{x}_j), j = \overline{1,n}\}$;

2) $y \in C^1(G)$;

3) задано наближені величини $\bar{y} = A(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$, $\Delta(\bar{x}_j)$, $\bar{x}_j$, $j = \overline{1,n}$.

Тоді

$$|y(x) - \bar{y}| \le \overline{A}_0(\bar{y}) \equiv \sum_{j=1}^{n} B_j \Delta(\bar{x}_j), \qquad (1.46)$$

$$B_j \equiv \sup_{x \in G} \left| \frac{\partial y(x)}{\partial (x_j)} \right|, \tag{1.47}$$

$$\left| y(x) - \bar{y} \right| \le \bar{A}^0(\bar{y}) \equiv \sum_{j=1}^{n} \left| \frac{\partial y(x)}{\partial (x_j)} \right| \Delta(\bar{x}_j). \tag{1.48}$$

Доведення. За формулою Лагранжа для функції багатьох змінних маємо

$$y(x) - \bar{y} = \sum_{j=1}^{n} b_j(\theta)(x_j - \bar{x}_j), \tag{1.49}$$

$$b_j(\theta) = \frac{\partial y\left(\bar{x}_1 + \theta(x_1 - \bar{x}_1), \ldots, \bar{x}_n + \theta(x_n - \bar{x}_n)\right)}{\partial x_j}, 0 < \theta < 1. \tag{1.50}$$

З (1.50) випливає (1.46). Покладемо $\rho \equiv \sqrt{\sum_{j=1}^{n} \Delta^2(\bar{x}_j)}$. Похідні $\frac{\partial y(x)}{\partial (x_j)}$ неперервні за умовою 2. Тоді $B_j = \left| b_j(\theta) \right| + o(\rho)$, де вираз $x = y + o(\rho)$ розуміємо так, що $(x - y) \to 0$ при $\rho \to 0$. Таким чином, $\bar{A}_0(\bar{y}) = \bar{A}^0(\bar{y}) + o(\rho)$, тобто маємо (1.48).

Наслідок 1.1. Гранична абсолютна похибка суми або різниці дорівнює сумі абсолютних похибок доданків

$$\bar{A}(\bar{y}) = \bar{A}_0(\bar{y}) = \bar{A}^0(\bar{y}) = \Delta(\bar{x}_1) + \ldots + \Delta(\bar{x}_n), \tag{1.51}$$

$$\Delta(\bar{y}) = \Delta(\bar{x}_1) + \ldots + \Delta(\bar{x}_n), \tag{1.52}$$

$$y = \gamma_1 x_1 + \gamma_2 x_2 + \ldots + \gamma_n x_n, \gamma_j \in \{-1, +1\}, \ j = \overline{1,n}. \tag{1.53}$$

Доведення. Нехай знаємо похибки аргументів $\Delta(\bar{x}_j)$. У даному випадку $b_j(\theta) \equiv \gamma_j$, $\left| b_j(\theta) \right| \equiv 1$. Тому з (1.50) отримаємо (1.52). А оскільки абсолютною похибкою називають будь-яку оцінку для $y - \bar{y}$, то (1.46) дає одразу (1.51) або (1.52).

Наслідок 1.2. Гранична відносна похибка добутку або частки наближеного дорівнює сумі граничних відносних похибок.

Доведення. Нехай $y = x_1^{p_1} \cdot \ldots \cdot x_n^{p_n}$, де $p_i \in \{-1, +1\}$. Тоді для будь-якого $j = \overline{1,n}$: $b_j = p_j \left[(\bar{x}_j)^{-1} \cdot \bar{y} \right]$ та

$$\bar{A}(\bar{y}) \approx \bar{A}^0(\bar{y}) = \sum \left| \bar{p}_j \right\| \bar{x}_j \left\| \bar{y}_j \right| \Delta(\bar{x}_j). \tag{1.54}$$

Після ділення на $|\bar{y}|$ одержимо

$$\frac{\bar{A}(\bar{y})}{|\bar{y}|} \approx \frac{\bar{A}^0(\bar{y})}{|\bar{y}|} = \sum_{j=1}^{n} |\bar{p}_j| \frac{\Delta(\bar{x}_j)}{|\bar{x}_j|} = \sum_{j=1}^{n} |\bar{p}_j| \delta(\bar{x}_j). \qquad (1.55)$$

Звідки матимемо доведення наслідку 2.

Приклад 1.6. Нехай $y = x^{10}; \bar{x} = \bar{y} = 1; \Delta(\bar{x}_j) = 10^{-3}$. Знайти $\bar{A}(\bar{y}); \bar{A}_0(\bar{y}); \bar{A}^0(\bar{y})$.

Розв'язок. $y'_x = 10x^9; b(1) = 10$. $B = \sup\limits_{|x-1| \le 10^{-3}} |10x^9| = 10(0{,}999)^9 = 10{,}09$.

$$\bar{A}(\bar{y}) \equiv \sup_{|x-1| \le 10^{-3}} |x^{10} - 1| \le 0{,}001^{10} - 1 = 0{,}010045,...$$

Тоді $\bar{A}_0(\bar{y}) \equiv B \cdot \Delta(\bar{x}) = 10{,}09 \cdot 0{,}001 = 0{,}01009,...$; $\bar{A}^0(\bar{y}) \equiv |b(0)| \Delta(\bar{x}) = 0{,}01$.

1.4. Контрольні питання

1. Дайте визначення математичної моделі.
2. Коли обчислювальна задача $y=A(x)$ називається коректно поставленою?
3. Що означає відсутність стійкості?
4. Наведіть приклади постановки задачі обчислення.
5. Назвіть основні джерела похибки результату обчислень.
6. В чому полягає основна задача лінійної алгебри?

1.5. Задачі для самостійної роботи

Задача 1.5.1. Довести, що гранична абсолютна похибка $\bar{A}(\bar{y})$ мінімальна при $y = \frac{Y_1 + Y_2}{2}$, $Y_1 = \inf\limits_{G} y(x_1, x_2, ..., x_n)$ Довести $\bar{A}_0(\bar{y}) - \bar{A}(\bar{y}) = o(\rho)$, $\bar{A}^0(\bar{y}) - \bar{A}(\bar{y}) = o(\rho)$.

Задача 1.5.2. Знайти $\bar{A}(\bar{y})$; $\bar{A}_0(\bar{y})$; $\bar{A}^0(\bar{y})$ для функції y, яка задана неявно $F(y, a_1, a_2, ..., a_n) = 0$.

Задача 1.5.3. Задано наближення $\bar{y}$ до кореня рівняння $f(y) = a$. Знайти його похибки.

Задача 1.5.4. Оцінити похибку кореня рівняння $F(y, a_1, a_2) = y^2 + a_1 y + a_2 = 0$ при заданих $\bar{a}_1, \bar{a}_2, \Delta(\bar{a}_1), \Delta(\bar{a}_2)$.

Задача 1.5.5. З якою точністю треба задати значення аргументів $\bar{a}_1, \bar{a}_2, ..., \bar{a}_n$ функції $y = A(\bar{a}_1, \bar{a}_2, ..., \bar{a}_n)$, щоб похибка $A(\bar{a}_1, \bar{a}_2, ..., \bar{a}_n)$ не перевищувала заданої величини ε?

Задача 1.5.6. Знайти граничну відносну похибку для $y = \gamma_1 x_1 + \gamma_2 x_2 + ... + \gamma_n x_n$ (за означенням).

Розділ 2.
Елементи теорії похибок

- Абсолютна та відносна похибки
  - Постановка задачі знаходження похибок
  - Приклади знаходження абсолютної та відносної похибок
- Пряма задача теорії похибок
  - Аналіз постановки задачі
  - Приклади постановки прямої задачі теорії похибок
- Обернена задача теорії похибок
  - Аналіз постановки оберненої задачі теорії похибок
  - Приклади постановки оберненої задачі теорії похибок
- Контрольні питання
- Задачі для самостійної роботи

В розділі розглянуто основні елементи теорії похибок [65, 71, 76-80, 88, 92-93, 106-110, 113-114, 116-133, 137, 140-145, 147-151, 161-164, 168-169, 192-194, 196-197, 201-209, 211]. Визначено поняття абсолютної та відносної похибок, прямої та оберненої задачі теорії похибок. Розглянуто приклади знаходження абсолютної та відносної похибок, прямої та оберненої задач теорії похибок.

Під похибкою будемо розуміти величину, що характеризує точність результату. Похибки, що виникають при розв'язуванні задачі, можна поділити на три групи:

1) неусувна похибка;
2) похибка методу;
3) похибка обчислень.

Неусувна похибка є наслідком:

а) неточності вхідних даних, що входять до математичного опису задачі;

б) невідповідності математичної моделі реальній задачі (інколи цю похибку називають похибкою математичної моделі).

Похибка методу пояснюється тим, що для розв'язування математичної задачі доводиться використовувати наближені методи, оскільки отримання точного розв'язку необмеженої або неприйнятно великої кількості арифметичних операцій є складним, а в багатьох випадках просто неможливим.

Похибка обчислень виникає при вводі-виводі даних до ПК та при виконанні математичних операцій.

Основна задача теорії похибок – знаходження області невизначеності результату.

Розглянемо процес заокруглення чисел. Якщо число x=4,167493 і його потрібно заокруглити до п'яти десяткових знаків після коми, то будемо мати x^*=4,16749. Тобто, якщо старший розряд, що відкидається менше 5, то попередня цифра не змінюється. Якщо x=4,167493 потрібно заокруглити до чотирьох знаків після коми,

то x^*=4,1675. Тобто, якщо старший розряд, що відкидається дорівнює або більше 5, то попередня цифра в числі збільшується на 1.

Зауваження. Інколи вважають, якщо старший розряд, що відкидається дорівнює 5, а попередня цифра парна, то вона не змінюється, якщо ж попередня цифра непарна, то вона збільшується на одиницю.

Розглянемо приклади заокруглення чисел:

x=2,8497621	x=345,453275
x^*=2,849762	x^*=345,45328
x^*=2,84976	x^*=345,4533
x^*=2,8498	x^*=345,453
x^*=2,850	x^*=345,45
x^*=2,85	x^*=345,5
x^*=2,8	x^*=345
x^*=3	x^*=3,5·10^2
	x^*=3·10^2

Визначимо, що при заокруглені цілого числа відкинуті знаки не можна замінняти нулями, а потрібно застосовувати множення на відповідний степінь 10.

2.1. Абсолютна та відносна похибки

2.1.1. Постановка задачі знаходження похибок

Нехай x – точне значення деякої величини, а x^* – її відоме наближене значення. Якщо $x^* < x$, то x^* наближає x з недостачею. Якщо $x^* > x$, то x^* наближає x з надлишком.

Абсолютною похибкою *числа* x^* називається *деяка величина* $\Delta(x^*)$, *що задовольняє умові*

$$|x^* - x| \le \Delta(x^*). \qquad (2.1)$$

Або під абсолютною похибкою $\Delta(x^*)$ наближеного числа x^* розуміють абсолютне значення різниці точного і наближеного значень.

Відносною похибкою *числа* x^* *називається деяка величина* $\delta(x^*)$, *що задовольняє умові*

$$\left|\frac{x^* - x}{x^*}\right| \le \delta(x^*). \qquad (2.2)$$

$$\left|\frac{x^* - x}{x^*}\right| \cdot 100\% \le \delta(x^*).$$

Або під відносною похибкою $\delta(x^*)$ наближеного числа x^* розуміють відношення абсолютної похибки до модуля точного значення числа. Часто відносна похибка вимірюється відсотками точного значення. Відомо, що будь-яке число

може бути подане у вигляді десяткового дробу, скінченого або нескінченого. Наближені десяткові числа виражаються лише кінцевими десятковими дробами.

З (2.1), (2.2) випливає, що $\Delta(x^*) \geq 0$ та $\delta(x^*) \geq 0$. Відзначимо, що точність результату краще характеризує відносна похибка. Інформацію про абсолютну та відносну похибки можна використати для наступного подання числа x:

$$\begin{aligned} x &= x^* \pm \Delta(x^*), \\ x &= x(1 \pm \delta(x^*)), \end{aligned} \tag{2.3}$$

де числа x, $\Delta(x^*)$ та $\delta(x^*)$ задано з однаковою кількістю знаків після коми.

Приклад 2.1. Нехай x = 0,20819940. Якщо $\Delta(x^*) = 0{,}00003$, знайти діапазон зміни x^*.

Розв'язок. Тоді $0{,}20816940 < x^* < 0{,}20822940$.

Означення 2.1. *Значущими цифрами числа називаються всі цифри в його записі, починаючи з першої ненульової зліва.*

Приклад 2.2. Знайти значущі цифри.

Розв'язок. Значущими цифрами відповідних чисел є:

1. x=4,570345 – всі цифри в запису цього числа значущі;
2. x=0,007614 – значущі цифри тільки 7,6,1,4;
3. x=0,03105600 – значущі цифри 3,1,0,5,6,0,0 (два останні нулі в запису числа є значущими);
4. а) x=3750000 – всі цифри значущі;
   б) $x = 3{,}75 \cdot 10^6$ – значущі цифри тільки 3,7,5.

Означення 2.2. *Значущу цифру називають вірною, якщо абсолютна похибка числа не перевищує ½ одиниці розряду, що відповідає цій цифрі. Інші цифри називають сумнівними.*

Означення 2.3. *Якщо всі значущі цифри вірні, то говорять, що число записано з усіма вірними цифрами.*

Говорять, що n перших значущих цифр десяткового числа є *вірними*, якщо абсолютна похибка цього числа не перевищує половину одиниці розряду, вираженого n-ю значущою цифрою, рахуючи зліва направо, тобто якщо

$$\Delta(x^*) = |x^* - x| \leq \frac{1}{2} \cdot 10^{m-n+1}, \tag{2.4}$$

то n цифр вірні.

Приклад 2.3. Нехай $x^* = 0{,}0208$; $\Delta(x^*) = 0{,}000003$. Знайти вірні цифри.

Розв'язок. Тоді x^* записано з усіма вірними цифрами.

Приклад 2.4. Пояснити запис $x = 0{,}02081994 \pm 4 \cdot 10^{-6}$.

Розв'язок. Запис слід розуміти $0{,}02081994 - 4 \cdot 10^{-6} \leq x \leq 0{,}02081994 + 4 \cdot 10^{-6}$.

Приклад 2.5. Пояснити записи $x = 1{,}0208(1 \pm 0{,}003)$ або $x = 1{,}0208(1 \pm 3 \cdot 10^{-3})$, або $x = 1{,}0208(1 \pm 0{,}3\%)$.

Розв'язок. Записи означають, що $(1-0{,}003)1{,}0208 \le x \le (1+0{,}003)1{,}0208$.

При проведенні розрахунків з наближеними числами слід пам'ятати та керуватися наступними *правилами*:

1. Кількість вірних знаків числа відраховується від першої значущої цифри числа до першої значущої цифри його абсолютної похибки (за означенням вірної цифри числа).
2. У математичних таблицях всі числа заокруглені до вірних знаків, причому абсолютна похибка не перевищує половини одиниці останнього розряду в зображенні числа.

Приклад 2.6. Нехай $\varepsilon = 2{,}718$. Знайти абсолютну похибку ε.

Розв'язок. Тоді $\Delta(\varepsilon) = 0{,}001/2 = 0{,}005 = 0{,}5 \cdot 10^{-3}$.

3. У проміжкових результатах обчислень, як правило, зберігають 2-3 сумнівні знаки, щоб не накопичувалися похибки від заокруглень.
4. У кінцевих результатах обчислень залишають, крім вірних знаків, один сумнівний знак.
5. Правило доповнення при заокругленні: якщо перша цифра, що відкидається, більша або дорівнює 5, то остання цифра, що зберігається, збільшується на одиницю.
6. При додаванні та відніманні наближених чисел у результаті повинні зберігатися стільки вірних розрядів, скільки їх у найменш точного із доданків.
7. Правило 6 застосовують у випадку операції множення і ділення.
8. У випадку операції піднесення до степеня слід зберігати стільки значущих цифр, скільки їх має основа степеня.
9. Правило 7 застосовують для операції видобування кореня.
10. В усіх проміжкових результатах розрахунків слід брати на одну-дві цифри більше. В остаточних результатах вони відкидаються.
11. Якщо дані можна брати з довільною точністю, то для отримання результату з k вірними цифрами вихідні дані слід брати з таким числом вірних цифр, щоб одержати k+1 вірну цифру в результаті.

2.1.2. Приклади знаходження абсолютної та відносної похибок

Приклад 2.7. Нехай x^*=14,537 та відомо, що $\Delta(x^*)$=0,04. Скільки вірних значущих цифр має число x^*?

Розв'язок. Маємо $\Delta(x^*) > 0{,}5 \cdot 10^{-2}$ і $\Delta(x^*) < 0{,}5 \cdot 10^{-1}$. Отже у числа x^* вірними будуть значущі цифри 1,4,5, а цифри 3 і 7 – сумнівні.

Приклад 2.8. Нехай x^*=8,677142 та $\Delta(x^*)=3 \cdot 10^{-4}$. Скільки вірних значущих цифр має число x^*?

Розв'язок. Оскільки $\Delta(x^*)=0{,}3 \cdot 10^{-3} < 0{,}5 \cdot 10^{-3}$, то x^* має вірні три значущі цифри після коми, тобто вірними будуть значущі цифри 8,6,7,7.

Приклад 2.9. Нехай x^*=0,046725 та $\Delta(x^*)$=0,008. Скільки вірних значущих цифр має число x^*?

Розв'язок. Маємо $\Delta(x^*)=0{,}8\cdot10^{-2}>0{,}5\cdot10^{-2}$. Отже у числа x^* всі значущі цифри сумнівні.

Приклад 2.10. Знайти для числа 402,35 з абсолютною похибкою $\Delta=0{,}004$ кількість вірних цифр.

Розв'язок. Очевидно, $\Delta<\frac{1}{2}\cdot10^{-2}$. Усі цифри цього числа вірні.

Приклад 2.11. Заокруглити наступні числа до чотирьох значущих цифр, визначити абсолютну Δ і відносну δ похибки отриманих наближених чисел A_1=625,51; A_2=0,0039227.

Розв'язок. Отримаємо a_1=625,5; a_2=0,003923;

$$\Delta_1=|A_1-a_1|=|625{,}51-625{,}5|=0{,}01;$$

$$\Delta_2=|A_2-a_2|=|0{,}0039227-0{,}003923|=0{,}0000003;$$

$$\delta_1=\frac{\Delta_1}{|A_1|}\cdot100\%=\frac{0{,}01}{625{,}51}\cdot100\%=0{,}0016\%;$$

$$\delta_2=\frac{\Delta_2}{|A_2|}\cdot100\%=\frac{0{,}0000003}{0{,}0039227}\cdot100\%=0{,}0076\%.$$

Приклад 2.12. Визначити абсолютну похибку наближеного числа за його відносною похибкою: a=20,725, $\delta=1\%$.

Розв'язок. Маємо $\delta=\frac{\Delta}{|A|}\cdot100\%$; $\Delta=\frac{\delta|a|}{100\%}$; $\Delta=\frac{1\%\cdot20{,}725}{100\%}\cong0{,}21$.

Приклад 2.13. Визначити кількість вірних знаків чисел a_1 і a_2, якщо відомі їхні абсолютні похибки Δ_1 і Δ_2: $a_1=293{,}452$; $\Delta_1=0{,}1$; $a_2=-0{,}00576$; $\Delta_1=0{,}1\cdot10^{-4}$.

Розв'язок. Число a_1 має три вірні значущі цифри $a_1=293\pm0{,}1$. Число a_2 має дві вірні значущі цифри, оскільки $\Delta_2=0{,}00001$, $a_2=0{,}0058\pm0{,}00001$.

Приклад 2.14. Визначити кількість вірних знаків числа a, якщо відома його відносна похибка δ: $a=9{,}6589$; $\delta=0{,}1\%$.

Розв'язок. Знаходимо $\Delta=\frac{0{,}1\%\cdot9{,}6589}{100\%}=0{,}0096589\approx0{,}01$. Тому a має дві вірних цифри $a=9{,}7\pm0{,}01$.

Приклад 2.15. Дані числа $a_1=1{,}137$ і $a_2=1{,}073$ з абсолютними похибками $\Delta_1=\Delta_2=0{,}011$. Оцінити відносну похибку їх різниці $a=a_1-a_2$.

Розв'язок. Маємо a=1,137-1,073=0,064; $\Delta a=\Delta_1+\Delta_2=0{,}022$;

$$\delta_a=\frac{0{,}022}{0{,}064}\cdot100\%=35\%;\ \delta_{a_1}=\frac{0{,}011}{1{,}137}\cdot100\%\approx1\%;\ \delta_{a_2}=\frac{0{,}011}{1{,}073}\cdot100\%\approx1\%\ .$$

Таким чином, відносна похибка різниці в 35 разів більша відносних похибок вихідних даних. Результат не має жодної вірної цифри.

Приклад 2.16. Обчислити $a = \dfrac{a_1 \cdot a_2 \cdot a_3}{a_4 \cdot a_5}$, вважаючи, що всі числа дані з вірними знаками: a_1=3,2; a_2=356,7; a_3=0,04811; a_4=7,1948; a_5=34,56.

Розв'язок. Найбільшу відносну похибку має число a_1=3,2:

$$\delta_{a_1} = \frac{0{,}05}{3{,}2} \cdot 100\% = 1{,}6\%.$$

Тому відносна похибка результату складе також приблизно 1,6%, тобто результат буде містити дві вірні цифри. Збережемо у вихідних даних по три вірні цифри (один запасний знак), одержимо $a = \dfrac{3{,}2 \cdot 357 \cdot 0{,}0481}{7{,}19 \cdot 34{,}6} = 0{,}221$. Абсолютну похибку результату обчислюємо за його відносною похибкою

$$\Delta_a = \frac{\delta_a \cdot a}{100\%} = \frac{1{,}6 \cdot 0{,}221}{100} = 0{,}0036.$$

Заокруглюємо результат, відкидаючи запасний знак, одержуємо a=0,22 з абсолютною похибкою $\Delta_a < 0{,}005$.

Приклад 2.17. Знайти суму наближених чисел $a = a_1 + a_2 + a_3 + a_4$: a_1=0,146; a_2=321,5; a_3=78,27; a_4=39,1. Усі дані знаки є вірними.

Розв'язок. Найбільшу абсолютну похибку мають числа a_2=321,5 і a_4=39,1. Тому можна вважати, що абсолютна похибка суми складає $\Delta = 2 \cdot 0{,}05 = 0{,}10$. Заокруглюємо доданки за зразком a_2 і a_4, збережемо один запасний розряд

$$a$=0,15+321,5+78,27+39,1=439,02.

В остаточному результаті запасний знак відкидаємо, одержимо a=439,0. До абсолютної похибки $\Delta = 0{,}10$ додаємо похибку округлення $\Delta_{окр} \le 0{,}05$. Тоді $\Delta_{пр\,a} = 0{,}10 + 0{,}05 = 0{,}15 < 0{,}2$. Таким чином, $a = 439 \pm 0{,}2$.

2.2. Пряма задача теорії похибок

2.2.1. Аналіз постановки задачі

В деякій області G n-вимірного простору розглядається неперервно-диференційована функція $y=f(x_1, x_2,\ldots, x_n)$. Необхідно обчислити значення цієї функції в точці $(x_1, x_2,\ldots, x_n)\in G$. Відомі тільки наближені значення $x_1^*, x_2^*,\ldots, x_n^*$ такі, що точка $(x_1^*, x_2^*,\ldots, x_n^*) \in G$, та їх похибки. Обчислимо наближене значення $y^* = f(x_1^*, x_2^*,\ldots, x_n^*)$ та оцінимо його абсолютну похибку.

Використовуючи формулу Лагранжа, отримуємо

$$\Delta(y^*) = \left| f(x_1^*, x_2^*,\ldots, x_n^*) - f(x_1, x_2,\ldots, x_n) \right| \le \sum_{j=1}^{n} B_j \Delta(x_j^*), \qquad (2.5)$$

$$B_j = \sup_G \left| \frac{\partial f(x_1, x_2, ..., x_n)}{\partial x_j} \right|. \tag{2.6}$$

В практичних розрахунках окрім оцінки (2.5) використовують оцінку

$$\Delta(y^*) \le \sum_{j=1}^{n} \left| \frac{\partial f(x_1^*, x_2^*, ..., x_n^*)}{\partial x_j} \right| \Delta(x_j^*), \tag{2.7}$$

яку називають *лінійною оцінкою похибки.*

Виходячи з оцінки (2.7), знайдемо відносну похибку:

$$\delta(y^*) \le \sum_{j=1}^{n} \left| \frac{\Delta(y^*)}{f(x_1^*, x_2^*, ..., x_n^*)} \right| = \sum_{j=1}^{n} \left| \frac{\dfrac{\partial f(x_1^*, x_2^*, ..., x_n^*)}{\partial x_j} \Delta(x_j^*)}{f(x_1^*, x_2^*, ..., x_n^*)} \right|. \tag{2.8}$$

Використовуючи формули (2.7), (2.8), визначимо похибки результатів математичних операцій.

Похибка суми.

$$y = f(x_1, x_2) = x_1 + x_2, \quad x_1, x_2 > 0. \tag{2.9}$$

Оскільки $f'_{x_j}(x^*) = 1$, то з (2.7) отримуємо

$$\Delta(y^*) = \Delta(x_1^*) + \Delta(x_2^*), \tag{2.10}$$

а з (2.8) відповідно

$$\delta(y^*) = \left| \frac{x_1^*}{x_1^* + x_2^*} \right| \delta(x_1^*) + \left| \frac{x_2^*}{x_1^* + x_2^*} \right| \delta(x_2^*). \tag{2.11}$$

Аналогічно знаходимо похибки для інших математичних операцій.

Похибка різниці.

$$y = f(x_1, x_2) = x_1 - x_2, \quad x_1 > x_2 > 0. \tag{2.12}$$

$$\Delta(y^*) = \Delta(x_1^*) + \Delta(x_2^*), \tag{2.13}$$

$$\delta(y^*) = \frac{x_1^* \delta(x_1^*) + x_2^* \delta(x_2^*)}{x_1^* - x_2^*}. \tag{2.14}$$

Похибка множення.

$$y = f(x_1, x_2) = x_1 \cdot x_2, \quad x_1, x_2 > 0. \tag{2.15}$$

$$\Delta(y^*) = \left|x_2^*\right|\Delta(x_1^*) + \left|x_1^*\right|\Delta(x_2^*), \tag{2.16}$$

$$\delta(y^*) = \delta(x_1^*) + \delta(x_2^*). \tag{2.17}$$

Похибка ділення.

$$y = f(x_1, x_2) = x_1 / x_2, \quad x_1, x_2 > 0. \tag{2.18}$$

$$\Delta(y^*) = \frac{\left|x_2^*\right|\Delta(x_1^*) + \left|x_1^*\right|\Delta(x_2^*)}{(x_2^*)^2}, \tag{2.19}$$

$$\delta(y^*) = \delta(x_1^*) + \delta(x_2^*). \tag{2.20}$$

Відзначимо, що для суми та різниці абсолютні похибки додаються, а для операцій множення та ділення складаються відносні похибки. З формули (2.13) видно, що якщо віднімаються два близьких числа, то відносна похибка результату може значно зрости. А при діленні на досить мале число може значно зрости абсолютна похибка. Розглянемо деякі приклади.

2.2.2. Приклади постановки прямої задачі теорії похибок

Приклад 2.18. Заокруглюючи наступні числа до трьох значущих цифр, визначити абсолютну та відносну похибки отриманих наближених чисел:

1) 0,1545; 2) 1,343; 3) –372,75.

Розв'язок.

1) x=0,1545. Заокруглення до трьох значущих цифр дає x^*=0,155, тоді $\Delta(x^*)$=0,0005=$5\cdot10^{-4}$, а відносна похибка

$$\delta(x^*)=5\cdot10^{-4}/0{,}155\approx0{,}32\cdot10^{-4}.$$

2) x=1,343. Тоді x^*=1,34, $\Delta(x^*)=|x^*-x|$=0,003. Відповідно відносна похибка

$$\delta(x^*)=3\cdot10^{-3}/1{,}34=2{,}2\cdot10^{-3}.$$

3) x=–372,75. Тоді x^*=–373, $\Delta(x^*)$=0,25, а

$$\delta(x^*)=0{,}25/373=6{,}7\cdot10^{-4}.$$

Приклад 2.19. Визначити кількість вірних цифр в числі x^*, якщо відома його відносна похибка:

1) x^*=22,351, $\delta(x^*)$=0,1;

2) x^*=9,4698, $\delta(x^*)=0{,}1\cdot10^{-2}$;

3) x^*=47361, $\delta(x^*)$=0,01;

Розв'язок.

1) Обчислимо абсолютну похибку $\Delta(x^*)=x^*\delta(x^*)$=2,2351. Тоді маємо, що в числі x^* вірною є тільки цифра 2, тобто одна вірна цифра.

2) Обчислимо абсолютну похибку

$$\Delta(x^*)=x^*\delta(x^*)=9{,}4698\cdot0{,}1\cdot10^{-2}=0{,}0094698.$$

Тоді в числі x^* будуть вірними дві цифри 9 та 4.

3) Абсолютна похибка буде дорівнювати

$$\Delta(x^*)=47361\cdot0{,}01=473{,}61.$$

Отже в числі x^* будуть вірними дві цифри 4 та 7.

Визначимо, що поведінка обчислювальної похибки залежить від правил заокруглення та алгоритму чисельного розв'язування задачі.

Приклад 2.20. На гіпотетичному ПК з мантисою довжини чотири знайти суму

$$S=0{,}2764+0{,}3944+1{,}475+26{,}46+1364$$

а) сумуючи від меншого доданку до більшого;

б) сумуючи від більшого доданку до меншого.

Розв'язок.

а) Маємо $S_2=0{,}2764+0{,}3944=0{,}6708$, $S_3=S_2+1{,}475$. Вирівнюючи порядки у цих двох доданків будемо мати $S_3=1{,}475+0{,}671=2{,}146$. Аналогічно далі

$$S_4=S_3+26{,}46=2{,}15+226{,}46=28{,}61,$$
$$S=S_5=S_4+1364=29+1393.$$

б) Маємо $S_2=1364+26{,}46=1364+26=1390$,

$$S_3=S_2+1{,}475=1390+1=1391,$$
$$S_4=S_3+0{,}3944=1391,$$
$$S=S_5=S_4+0{,}2764=1391.$$

Враховуючи, що точне значення $S=1392{,}6058$, бачимо, що сумування потрібно проводити починаючи з менших доданків. В протилежному випадку може мати місце значна втрата значущих цифр.

Приклад 2.21. Оцінити похибку обчислення функції

$$f(x,y,z)=\frac{x^2z}{y^3},$$

якщо $x=0{,}15\pm0{,}005$, $y=2{,}13\pm0{,}01$, $z=1{,}14\pm0{,}007$.

Розв'язок. Згідно з формулою (2.6), для абсолютної похибки результату отримаємо

$$\Delta(f^*)=\left|\frac{2x^*z^*}{(y^*)^3}\right|\Delta(x^*)+\left|\frac{3(x^*)^2z}{(y^*)^4}\right|\Delta(y^*)+\left|\frac{(x^*)^2}{(y^*)^3}\right|\Delta(z^*)=\frac{2\cdot0{,}15\cdot1{,}14}{(2{,}13)^3}\cdot0{,}005+$$

$$+\frac{3\cdot(0{,}15)^2\cdot1{,}14}{(2{,}13)^4}\cdot0{,}01+\frac{(0{,}15)^2}{(2{,}13)^3}\cdot0{,}007=0{,}00017695+0{,}00003738+$$

$$+0{,}000016298\approx0{,}00023=2{,}3\cdot10^{-4}.$$

Знайдемо $f(x^*,y^*,z^*)=\dfrac{(0{,}15)^2\cdot1{,}14}{(2{,}13)^3}=0{,}0022265429$.

Тоді $\delta(f^*)=\dfrac{2{,}3\cdot10^{-4}}{0{,}00265429}=0{,}08665$.

Приклад 2.22. Висота h та радіус основи циліндра виміряні з точністю до 0,5%. Яка відносна похибка при обчисленні об'єму циліндра, якщо $\pi^*=3{,}14$?

Розв'язок. $V=\pi R^2h$. Більш точне значення $\pi=3{,}14159265$, отже $\Delta(\pi^*)=0{,}16\cdot10^{-2}$, а $\delta(\pi^*)=0{,}16\cdot10^{-2}/3{,}14=0{,}0005=0{,}05\%$. Тоді, згідно формули відносної похибки добутку отримуємо:

$$\delta(V^*)=\delta(\pi^*)+2\delta(R^*)+\delta(h)=1{,}55\%.$$

Приклад 2.23. Ребро куба виміряне з точністю до 0,02 см. дорівнює 8 см. Знайти абсолютну та відносну похибки при обчисленні об'єму куба.

Розв'язок. Позначимо сторону куба через a. Тоді $V = a^3$, $V^* = (a^*)^3 = 512$ см. Застосовуючи формулу (2.7), отримуємо

$$\Delta(V^*) = 3(a^*)^2 \cdot \Delta(a^*) = (3 \cdot 8^2 \cdot 0{,}02)\text{см}^3 = 3{,}84\text{см}^3, \text{ а } \delta(V^*) = (3{,}84/512) = 0{,}0075.$$

Приклад 2.24. Визначити відносну похибку числа, що записане в ПК в системі числення β та довжиною мантиси t.

Розв'язок. Число x^* можна записати в ПК у вигляді

$$x^* = \pm(d_1\beta^{-1} + d_2\beta^{-2} + ... + d_t\beta^{-t})\beta^{\ell},$$

де ℓ визначає порядок числа, d_i – цілі, причому $0 \le d_i \le \beta - 1$, $d_1 \ne 0$. Нехай точне значення числа дорівнює $x^* = \pm(d_1\beta^{-1} + d_2\beta^{-2} + ... + d_t\beta^{-t} + d_{t+1}\beta^{-t-1})\beta^{\ell}$. Тоді

$$\frac{|x_* - x|}{|x_*|} = \frac{d_{t+1}\beta^{\ell-t-1}}{|x_*|} = \left|\frac{d_{t+1}}{d_1\beta^t + d_2\beta^{t-1} + ... + d_t\beta}\right| \le \frac{d_{t+1}}{d_1\beta^t} \le \frac{d_{t+1}}{\beta^t} = \beta^{1-t}\left(1 - \frac{1}{\beta}\right) \le \beta^{1-t}.$$

Отже, $\delta(x^*) \le \beta^{1-t}$.

Якщо ж числа вводяться за правилами заокруглення, то $d_{t+1} \le 0{,}5\beta$ і тоді отримуємо, що $\delta(x^*) \le \frac{1}{2}\beta^{1-t}$.

2.3. Обернена задача теорії похибок

2.3.1. Аналіз постановки оберненої задачі теорії похибок

Обернена задача теорії похибок полягає в наступному: з якою точністю потрібно задати значення аргументів $x_1^*, x_2^*, ..., x_n^*$ функції $f(x_1, x_2, ..., x_n)$, щоб похибка значення функції $f(x_1^*, x_2^*, ..., x_n^*)$ не перевищувала заданої величини ε.

Для функції однієї змінної $y=f(x)$ абсолютну похибку можна наближено обчислити за формулою:

$$\Delta(x^*) = \frac{\Delta(y^*)}{|f'(x^*)|}, \quad f'(x^*) \ne 0. \tag{2.21}$$

Для функції декількох змінних $y = f(x_1, x_2, ..., x_n)$ задача розв'язується за допомогою наступних рекомендацій:

а) *принцип рівних впливів*, тобто вважаємо, що всі доданки $c_i = |\partial f / \partial x_i|\Delta(x_i^*)$, $i = \overline{1,n}$ рівні між собою. Тоді абсолютні похибки всіх аргументів визначаються формулою:

$$\Delta(x_i^*) = \frac{\Delta(y^*)}{n\left|\frac{\partial f^*}{\partial x_i}\right|}, \quad i = \overline{1,n}; \tag{2.22}$$

б) *вважаємо всі похибки рівними*, причому максимально можливими, тобто покладемо

$$\Delta(x_1^*) = \Delta(x_2^*) = ... = \Delta(x_n^*) = \delta, \tag{2.23}$$

де $\delta = \varepsilon/(c_1 + c_2 + ... + c_n)$.

2.3.2. Приклади постановки оберненої задачі теорії похибок

Приклад 2.25. Сторона квадрату дорівнює 2м. З якою точністю її потрібно виміряти, щоб похибка знаходження площі не перевищувала 1см^2?

Розв'язок. Позначимо сторону квадрату через x; $S=x^2$, $S'=2x$. Тоді за формулою (2.20) отримаємо

$$\Delta(x^*) = \frac{1}{2\cdot 200} = \frac{1}{4}10^{-2} \text{ см}.$$

Приклад 2.26. З якою кількістю вірних значущих цифр потрібно взяти вільний член квадратного рівняння

$$x^2 - 2x + lg2 = 0,$$

щоб отримати корені рівняння з чотирма вірними значущими цифрами?

Розв'язок. Для коренів рівняння маємо $x_1 = 1 + \sqrt{1 - \lg 2}$, $x_2 = 1 - \sqrt{1 - \lg 2}$. Оскільки $\lg 2 \approx 0{,}3...$, тоді $x_1 \approx 1{,}8...$, $x_2 \approx 0{,}1...$. Отже, за змістом задачі x_1^* потрібно визначити так, щоб $\Delta(x_1^*) \le 0{,}5\cdot 10^{-3}$, а для x_2^*, щоб $\Delta(x_2^*) \le 0{,}5\cdot 10^{-4}$. Позначимо $z=ln2$ та розглянемо функцію $f(z) = 1 + \sqrt{1 - z}$. З'ясуємо, з якою точністю потрібно обчислити z^* в околі точки 0,3, щоб $\Delta f'(z) = \frac{1}{2\sqrt{1-z}}$. Використовуючи формулу (2.20), отримуємо

$$\Delta(z^*) \le 0{,}5\cdot 10^{-3} \cdot \frac{1}{2\sqrt{0{,}7}} = 0{,}000299.$$

Звідси робимо висновок, що для знаходження кореня x_1 потрібно обчислити $lg2$ з трьома вірними значущими цифрами після коми, тобто $lg2=0{,}301$.

Аналогічно, розглядаючи функцію $f(z) = 1 - \sqrt{1 - z}$, отримаємо, що для знаходження кореня x_2 з точністю $0{,}5\cdot 10^{-4}$ потрібно обчислити $lg2$ з чотирма вірними значущими цифрами після коми, тобто $lg2=0{,}3010$.

Приклад 2.27. В п'ятизначних логарифмічних таблицях дано значення десяткових логарифмів з точністю до $\varepsilon=0{,}5\cdot 10^{-6}$. Оцінити величину можливої похибки при знаходженні числа за його логарифмом, якщо саме число знаходиться між 300 та 400.

Розв'язок. Позначимо $y = \lg x,\ x \in [300;400]$. За умовою задачі $\Delta(y^*) \le 0{,}5 \cdot 10^{-6}$ та потрібно знайти $\Delta(x^*)$. Маємо $y' = \dfrac{1}{x \ln 10}$. Тоді за формулою (2.20) отримуємо

$$\Delta(x^*) = x(\ln 10) \cdot \Delta(y^*) \le 400 \cdot 2{,}30 \cdot 0{,}5 \cdot 10^{-6} = 0{,}00046.$$

Отже, x можна знайти принаймні з трьома вірними значущими цифрами після коми.

2.4. Контрольні питання

1. Назвіть групи похибок, що виникають при розв'язуванні задачі.
2. Коли виникає *похибка обчислень*?
3. Дайте визначення *абсолютній похибці*.
4. Дайте визначення *відносній похибці*.
5. При якій умові значуща цифра називається вірною?
6. В чому полягає *обернена задача теорії похибок*?

2.5. Задачі для самостійної роботи

Задача 2.5.1. Заокруглюючи наступні числа до трьох значущих цифр, визначити абсолютну та відносну похибки наближених чисел:

1) 3,2523
2) 0,17153
3) 0,02103
4) 1,445
5) –0,0035392
6) –583,71
7) 0,004966
8) 315,55
9) 71,534

Задача 2.5.2. Визначити кількість вірних цифр в числі x, якщо його відносна похибка:

1) x=2,7981, $\delta(x)=0{,}1 \cdot 10^{-2}$;
2) x=12,8370, $\delta(x)$=1%;
3) x=0,3328, $\delta(x)=0{,}2 \cdot 10^{-1}$;
4) x=372,8, $\delta(x)$=2%;
5) x=23,652, $\delta(x)$=0,1;
6) x=17261, $\delta(x)$=1%;
7) x=0,03575, $\delta(x)=0{,}5 \cdot 10^{-2}$;
8) x=0,22453, $\delta(x)$=10%;
9) x=0,000335, $\delta(x)$=0,15;
10) x=6,3495, $\delta(x)$=0,1%.

Задача 2.5.3. Визначити, яка рівність точніша:

1) $6/7=0{,}857$, $\sqrt{4{,}8}=2{,}19$;
2) $2/21=0{,}095$, $\sqrt{22}=4{,}69$;
3) $7/19=0{,}895$, $\sqrt{52}=7{,}21$;
4) $49/13=3{,}77$, $\sqrt{14}=3{,}74$.

Задача 2.5.4. Якою буде відносна похибка, якщо число π наблизити числом 3,14?

Задача 2.5.5. Записати число π з п'ятьма вірними значущими цифрами та визначити відносну похибку отриманого наближення.

Задача 2.5.6. Знайти $\sqrt{3{,}02}-\sqrt{3}$ з трьома вірними значущими цифрами.

Задача 2.5.7. При вимірі радіуса кола з точністю до 0,5 см, отримали число 14 см. Знайти абсолютну та відносну похибки при обчисленні площі кола.

Задача 2.5.8. Кожне ребро куба, виміряне з точністю 0,02 см виявилося рівним 15 см. Знайти абсолютну та відносну похибки при обчисленні площі куба.

Задача 2.5.9. Визначити відносну похибку обчислення повної поверхні зрізаного конуса, якщо радіуси його основ R і r та твірна ℓ виміряні з точністю до 0,01 см, рівні: $R=23{,}64$ см, $r=17{,}31$ см, $\ell=10{,}21$ см.

Задача 2.5.10. Обчислити значення функції f. Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних вірними: $f(x_1,x_2)=x_1 * x_2$, де

1) $x_1=5{,}49$, $x_2=7{,}6$;
2) $x_1=15{,}1$, $x_2=2{,}543$;
3) $x_1=0{,}03$, $x_2=12{,}5$.

Задача 2.5.11. Обчислити значення функції f. Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних вірними: $f(x_1,x_2,x_3)=x_1*x_2*x_3$, де

1) $x_1=381{,}56$, $x_2=6157$, $x_3=0{,}0053$;
2) $x_1=0{,}147$; $x_2=653$, $x_3=76{,}3$;
3) $x_1=1{,}28$, $x_2=6{,}3$, $x_3=2{,}173$.

Задача 2.5.12. Обчислити значення функції f. Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних вірними: $f=x_1 x_2+x_2x_3$, де $x_1=2{,}104$, $x_2=1{,}935$, $x_3=0{,}845$.

Задача 2.5.13. Обчислити значення функції f. Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних вірними: $f=x_1/x_2$

1) $x_1=526{,}677$, $x_2=829$;
2) $x_1=745{,}8371$, $x_2=336{,}2$;
3) $x_1=6{,}3$, $x_2=449$;
4) $x_1=5{,}684$, $x_2=5{,}032$.

Задача 2.5.14. Обчислити значення функції f. Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних вірними: $f=\ln(x_1+x_2^2)$, де $x_1=0{,}93$, $x_2=1{,}123$.

Задача 2.5.15. Обчислити значення функції f. Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних вірними: $f=\dfrac{x_1+x_2^2}{x_3}$, де x_1=3,15, x_2=0,831, x_3=1,123.

Задача 2.5.16. Оцінити абсолютну та відносну похибки обчислення функції:

1) $f(x,y,z)=\ln\dfrac{xy}{z}$, при x=2,34±0,01, y=1,25±0,02, z=3,05±0,02;

2) $f(x,y,z)=\sqrt{\dfrac{xy}{z}}$, при x=0,757±0,001, y=21,7±0,05, z=1,84±0,05;

3) $f(x,y,z)=\dfrac{\sqrt{x}+y}{\sqrt[3]{z}}$, при x=4±0,1, y=3±0,05, z=1±0,08;

4) $f(x,y,z)=\ln\left(xy+\dfrac{z}{x}\right)$, при x=1,02±0,01, y=2,35±0,02, z=3,04±0,01;

5) $f(x,y,z)=\dfrac{(x+y)(2z-1)^2}{x-y}$, при x=5,8±0,01, y=0,65±0,02, z=1,1753±0,0002;

6) $f(x,y,z)=\dfrac{(x^2+4xy+y^2)(2z-1)^2}{(x+y)^2}\cdot\dfrac{z^2}{18}$, при x =27,51±0,001, y=21,78±0,003, z=32,5±0,06;

7) $f(x,y,)=\dfrac{1}{64}\pi\sqrt{x^4-y^4}$, при x=36,5±0.01, y=26,35±0,005, π=3,14.

Задача 2.5.17. Знайти межі абсолютної та відносної похибки аргументів, які дозволяють обчислити з чотирма вірними знаками функцію $f=\dfrac{x_1+x_2^2}{x_3}$, де x_1=2,10415, x_2=1,93521, x_3=0,84542.

Задача 2.5.18. Оцінити похибку в визначені кута x=60° за п'ятизначною таблицею синусів.

Задача 2.5.19. З якою кількістю вірних значущих цифр потрібно взяти значення аргументу x, щоб обчислити значення функції $f(x)=x^3\sin x$ з точністю до $0{,}1\cdot10^{-5}$?

Задача 2.5.20. З якою точністю потрібно обчислити $\sin\dfrac{\pi}{8}$, щоб відносна похибка обчислення коренів рівняння $x^2-2x+\sin\dfrac{\pi}{8}=0$ не перевищувала 10^{-3}?

Задача 2.5.21. З якою відносною похибкою треба виміряти висоту h =0.5 м та радіус основи r=10 для того, щоб відносна похибка обчислення об'єма конуса не перевищувала 0,1%?

Розділ 3.
Методи розв'язування лінійних алгебраїчних рівнянь

- Чисельне розв'язування систем лінійних алгебраїчних рівнянь
- Основні поняття та класифікація систем лінійних алгебраїчних рівнянь
- Основні поняття та особливості матриць
  - Операції над матрицями
  - Квадратна матриця й суміжні визначення
  - Властивості матриць
  - Обчислення оберненої матриці
- Метод Крамера
  - Аналіз постановки задачі
  - Приклади розв'язування системи за формулами Крамера
- Метод оберненої матриці
- Особливості методу Гауса
  - Аналіз постановки задачі для методу Гауса
  - Метод Гауса (метод заміни змінних)
  - Частковий випадок застосування методу Гауса з послідовним виключенням невідомих
  - Метод Гауса за схемою Халецького (метод LU факторизації)
  - Метод Гауса з вибором головного елемента
  - Метод Гауса з одиничними коефіцієнтами
  - Метод Гауса-Жордана
  - Приклади розв'язування системи методом Гауса
  - Несумісні системи. Системи із загальним розв'язком. Часткові розв'язки
  - Способи знаходження оберненої матриці
  - Застосування методу Гауса до обчислення визначника
  - Застосування методу Гауса до інверсії матриці
- Метод простої ітерації
  - Аналіз постановки задачі
  - Приклади розв'язування системи методом простої ітерації
- Метод Гауса-Зейделя (метод поліпшеної ітерації)
  - Аналіз постановки задачі
  - Приклади розв'язування системи методом Зейделя
- Особливості трьохдіагональної матриці

- Обчислення детермінанту трьохдіагональної матриці
- Метод прогонки для розв'язування трьохдіагональних систем лінійних рівнянь

- Аналіз числових методів розв'язування СЛАР
  - Загальна характеристика методів розв'язування систем лінійних алгебраїчних рівнянь
  - Збіжність метричної геометричної прогресії
  - Збіжність методу простої ітерації
  - Збіжність методу Зейделя розв'язування СЛАР
  - Збіжність метода Зейделя за елементами матриці
- Методи розв'язування повної проблеми власних значень і власних векторів
  - Наближене знаходження власних значень матриці
  - Ідея методу Данилевського
  - Обчислення власних векторів методом Данилевського
  - Метод невизначених коефіцієнтів
  - Загальна постановка задачі методів розв'язування повної проблеми власних значень і власних векторів
  - Метод невизначених коефіцієнтів розв'язування повної проблеми власних значень і власних векторів
  - Метод інтерполювання розв'язування повної проблеми власних значень
- Приклади розв'язування систем лінійних алгебраїчних рівнянь
- Контрольні питання
- Задачі для самостійної роботи
- Завдання до лабораторної роботи

Інженеру часто доводиться розв'язувати алгебраїчні та трансцендентні рівняння і системи рівнянь, що можуть являти собою самостійну задачу (наприклад, аналіз рівноваги сил в жорсткій системі балок або дослідження умов та параметрів рівноваги хімічної реакції, тощо), або частину 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 [65, 71, 76-80, 88, 92-93, 106-110, 113-114, 116-133, 137, 140-145, 147-151, 161-164, 168-169, 192-194, 196-197, 201-209, 211].

3.1. Чисельне розв'язування систем лінійних алгебраїчних рівнянь

Нехай дано систему n лінійних рівнянь з n невідомими. Системою лінійних алгебраїчних рівнянь (СЛАР) називають систему вигляду:

$$\begin{cases} a_{11} \cdot x_1 + a_{12} \cdot x_2 + ... + a_{1n} \cdot x_n = b_1; \\ a_{21} \cdot x_1 + a_{22} \cdot x_2 + ... + a_{2n} \cdot x_n = b_2; \\ ..... \\ a_{n1} \cdot x_1 + a_{n2} \cdot x_2 + ... + a_{nn} \cdot x_n = b_n. \end{cases} \quad (3.1)$$

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. Необхідно знайти *розв'язок* системи, тобто таку сукупність значень невідомих x_1, x_2, …, x_n, яка при підстановці в цю систему перетворює всі рівняння у тотожності. Запишемо систему (3.1) у матричній формі

$$\mathbf{A} \cdot \mathbf{X} = \mathbf{B}, \quad (3.2)$$

де $A = \{a_{11}, a_{21}, ..., a_{nn}\}$ – матриця коефіцієнтів СЛАР при невідомих; $B = \{b_1, b_2, ..., b_n\}$ – вектор-стовпець вільних членів; $X = \{x_1, x_2, ..., x_n\}$ – вектор-стовпець невідомих. Вони мають вигляд:

$$A = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nn} \end{bmatrix}; B = \begin{bmatrix} b_1 \\ b_2 \\ ... \\ b_n \end{bmatrix}; X = \begin{bmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{bmatrix}. \quad (3.3)$$

Розв'язком системи лінійних алгебраїчних рівнянь (3.1) називають вектор X, координати якого $X = \{x_1, x_2, ..., x_n\}$ при підстановці у систему, яку розв'язують, перетворюють кожне рівняння системи в тотожність. Якщо детермінант матриці A відрізняється від нуля, тобто $\det \mathbf{A} \neq 0$, то матриця A є неособливою і система лінійних рівнянь (3.1) або (3.2) має єдиний розв'язок.

3.2. Основні поняття та класифікація систем лінійних алгебраїчних рівнянь

Для розв'язування СЛАР на комп'ютері традиційно використовують дві групи чисельних методів, що представлені на рисунку 3.1.

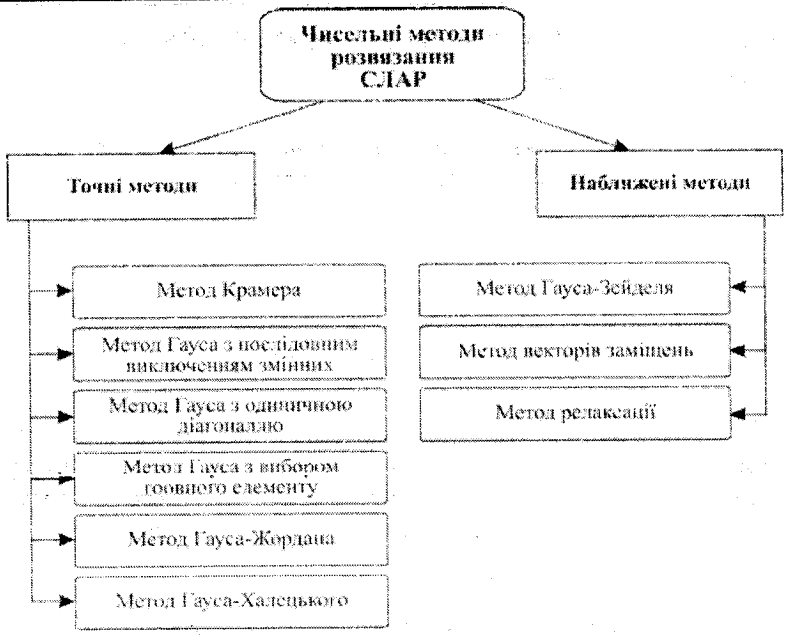


Рис. 3.1. Класифікація чисельних методів

Методи чисельного розв'язування систем лінійних алгебраїчних рівнянь поділяються на дві групи:

1. **Точні** (прямі) методи, які дозволяють одержати розв'язок, якщо він існує, як скінченну кількість математичних операцій (наприклад: 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 тощо). До точних методів відносять методи, які дозволяють отримати точний розв'язок системи (3.1) через відповідну кількість операцій перетворення без урахування похибок заокруглення.
2. **Наближені** (ітераційні) методи, які дозволяють одержати лише наближені до коренів значення із певною похибкою (наприклад, метод послідовних ітерацій, метод Гауса-Зейделя, метод векторів зміщень). Точний розв'язок можна отримати як результат нескінченно збіжного процесу дослідження або експерименту. До таких методів відносять метод простої ітерації, метод Зейделя та ін. "Точні" методи використовують при розв'язку на ПК систем невисокого порядку ($n<10^3$, де n – число лінійних алгебраїчних рівнянь системи). Наближені методи використовують для систем високого порядку $n=10^3...10^6$. Переваги наближеного методу (простих ітерацій) над точним методом (Гауса) наступні:
   - Час обчислень пропорційний n^2 на ітерацію, тоді як для методу Гауса – n^3. Якщо для розв'язку потрібно менше ніж n ітерацій, то втрати машинного часу будуть менші.

- Як правило похибки округлення при ітераційному методі впливають на остаточні результати значно менше, ніж при методі Гауса, оскільки при його використанні похибки не нагромаджуються.
- Ітераційний метод стає особливо зручним при розв'язуванні систем, переважна кількість коефіцієнтів яких дорівнює 0. Рівняння, які отримують, наприклад, при розв'язуванні крайових задач, відносяться саме до цього класу.

Недоліком ітераційного методу є те, що він не завжди збігається. Ітераційні методи дозволяють одержати приблизний розв'язок системи за скінченне число наближень (ітерацій) із заданою похибкою результату. До наближених методів відносять методи, які дозволяють отримати розв'язок системи (рис. 3.1) у вигляді границі послідовності векторів $\lim_{k\to\infty}\{X^0, X^1, X^2, ..., X^n\}$, яка збігається до точного розв'язку системи, де:

$$X^0 = \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \dots \\ x_n^{(0)} \end{bmatrix}; X^1 = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ \dots \\ x_n^{(1)} \end{bmatrix}; \dots; X^n = \begin{bmatrix} x_1^{(n)} \\ x_2^{(n)} \\ \dots \\ x_n^{(n)} \end{bmatrix}. \qquad (3.4)$$

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 комп'ютері чисельними методами, якщо вона сумісна, визначена, невироджена.

3.3. Основні поняття та особливості матриць

Матриця – математичний об'єкт, що записується у вигляді прямокутної таблиці чисел (або елементів кільця), і допускає алгебраїчні операції (додавання, віднімання, множення тощо) між ним та іншими подібними об'єктами. Правила виконання операцій над матрицями визначені для зручності запису системи лінійних алгебраїчних рівнянь. Звичайно матрицю позначають великою буквою латинського алфавіту й виділяють круглими дужками (...) (зустрічається також виділення квадратними дужками – [...], подвійними прямими лініями – ||...||). А числа (елементи матриці), позначають тією же буквою, що й саму матрицю, але маленькою. У кожного елемента матриці є 2 нижніх індекси (a_{ij}) – перший «i» позначає номер рядка, у якому перебуває елемент, а другий «j» – номер стовпця.

3.3.1. Операції над матрицями

Нехай a_{ij} – елементи матриці A, а b_{ij} – елементи матриці B.

1. Множення матриці A на число λ (позначення: λA) полягає в побудові матриці B, елементи якої отримані шляхом множення кожного елемента матриці A на це число, тобто кожен елемент матриці B дорівнює:

$$b_{ij} = \lambda a_{ij} . \tag{3.5}$$

2. Додавання матриць $A+B$ – операція знаходження матриці C, всі елементи якої дорівнюють попарній сумі всіх відповідних елементів матриць A і B, тобто кожен елемент матриці C дорівнює:

$$c_{ij} = a_{ij} + b_{ij} . \tag{3.6}$$

Наприклад:

$$A + B = \begin{bmatrix} 2 & 0 & -1 & 4 \\ 1 & 3 & 0 & 5 \end{bmatrix} + \begin{bmatrix} 3 & 1 & 0 & 10 \\ 8 & 2 & 3 & -5 \end{bmatrix} = \begin{bmatrix} 2+3 & 0+1 & -1+0 & 4+10 \\ 1+8 & 3+2 & 0+3 & 5+(-5) \end{bmatrix} = \begin{bmatrix} 5 & 1 & -1 & 14 \\ 9 & 5 & 3 & 0 \end{bmatrix}$$

3. Різниця матриць $A-B$ визначається аналогічно додаванню, це операція знаходження матриці C, елементи якої визначаються як:

$$c_{ij} = a_{ij} - b_{ij} . \tag{3.7}$$

Додавання й віднімання допускається тільки для матриць однакового розміру.

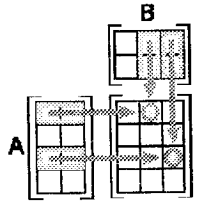
4. Існує нульова матриця Θ така, що її додавання до іншої матриці A не змінює A, тобто

$$A + \Theta = A . \tag{3.8}$$

Всі елементи нульової матриці рівні нулю.

5. Добуток матриць (позначення: AB, рідше зі знаком множення $A \times B$) – операція обчислення матриці C, елементи якої дорівнюють сумі добутків елементів у відповідному рядку першого множника й стовпці другого.

$$c_{ij} = \sum_k a_{ik} b_{kj} . \tag{3.9}$$



У першому множнику повинно бути стільки ж стовпців, скільки рядків у другому. Якщо матриця A має розмірність $m \times n$, B – $n \times k$, то розмірність їх добутку $AB = C$ є $m \times k$. Множення матриць не є комутативною операцією. Це видно хоча б з того, що якщо матриці не квадратні, то можна множити тільки одну

на іншу, але не навпаки. Для квадратних матриць результат множення залежить від *порядку співмножників. Наприклад:*

$$AB = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix} \times \begin{bmatrix} 3 & 1 \\ 8 & 2 \end{bmatrix} = \begin{bmatrix} 2\cdot 3 + 0\cdot 8 & 2\cdot 1 + 0\cdot 2 \\ 1\cdot 3 + 3\cdot 8 & 1\cdot 1 + 3\cdot 2 \end{bmatrix} = \begin{bmatrix} 6 & 2 \\ 27 & 7 \end{bmatrix}$$

$$BA = \begin{bmatrix} 3 & 1 \\ 8 & 2 \end{bmatrix} \times \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3\cdot 2 + 1\cdot 1 & 3\cdot 0 + 1\cdot 3 \\ 8\cdot 2 + 2\cdot 1 & 8\cdot 0 + 2\cdot 3 \end{bmatrix} = \begin{bmatrix} 7 & 3 \\ 18 & 6 \end{bmatrix}$$

Підносити до степеня можна тільки квадратні матриці.

6. Транспонування матриці (позначення: A^T) — операція, при якій матриця відображається відносно головної діагоналі, тобто

$$a_{ij}^T = a_{ji}. \tag{3.10}$$

Якщо A – матриця розміру $m \times n$, то A^T – матриця розміру $n \times m$.

3.3.2. Квадратна матриця й суміжні визначення

Якщо кількість рядків матриці дорівнює кількості стовпців, то така матриця називається *квадратною*. Для квадратних матриць існує одинична матриця E така, що множення будь-якої матриці на неї не впливає на результат, а саме:

$$EA = AE = A. \tag{3.11}$$

В одиничній матриці одиниці розташовані тільки по діагоналі, інші елементи дорівнюють нулю

$$E = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{pmatrix} \tag{3.12}$$

Для деяких квадратних матриць можна знайти так звану обернену матрицю. *Обернена матриця* A^{-1} є такою, що якщо помножити матрицю на неї, то отримаємо одиничну матрицю

$$AA^{-1} = E. \tag{3.13}$$

Обернена матриця існує не завжди. Матриці, для яких обернена існує, називаються *невиродженими*, а для яких не існує – *виродженими*. Матриця невироджена, якщо всі її рядки (стовпці) лінійно незалежні як вектори. Максимальне число лінійно незалежних рядків (стовпців) називається *рангом матриці*. *Визначником* (детермінантом) *матриці* називається нормований кососиметричний лінійний функціонал на рядках матриці. Квадратна матриця над числовим полем вироджена тоді й тільки тоді, коли її визначник дорівнює нулю.

3.3.3. Властивості матриць

Розглянемо основні властивості матриць в (3.14)-(3.21).

$$A+(B+C)=(A+B)+C, \tag{3.14}$$

$$A(BC)=(AB)C, \tag{3.15}$$

$$A(B+C)=AB+AC, \tag{3.16}$$

$$(B+C)A=BA+CA, \tag{3.17}$$

$$A\cdot 0=0, \tag{3.18}$$

$$1\cdot A=A, \tag{3.19}$$

$$A_{k\times l}\cdot B_{l\times n}=C \Rightarrow c_{ij}=\sum_{k=1}^{l} a_{ik}b_{kj}, \tag{3.20}$$

$$\left(A^T\right)^T=A. \tag{3.21}$$

Квадратна матриця має обернену тоді й тільки тоді, коли вона невироджена, тобто її визначник не дорівнює нулю. Для неквадратних матриць і сингулярних матриць обернених матриць не існує.

$$\det A^{-1}=\frac{1}{\det A}, \tag{3.22}$$

$$(AB)^{-1}=B^{-1}\cdot A^{-1}. \tag{3.23}$$

Якщо необхідно розв'язати систему лінійних рівнянь $Ax=b$, де x – шуканий вектор, і якщо A^{-1} існує, то $x=A^{-1}b$. В протилежному випадку розмірність простору розв'язків більше нуля.

3.3.4. Обчислення оберненої матриці

Нехай задано матрицю

$$A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}. \tag{3.24}$$

Матриця A^{-1} називається оберненою до заданої матриці A, якщо її множення як зліва, так і справа на задану матрицю дає одиничну матрицю, тобто

$$A^{-1}\cdot A=A\cdot A^{-1}=E. \tag{3.25}$$

Одинична матриця має вигляд

$$E = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{bmatrix}. \tag{3.26}$$

Необхідною і достатньою умовою існування оберненої матриці є неособливість вихідної матриці. Матриця називається *неособливою*, якщо детермінант матриці відрізняється від нуля, тобто $\det \mathbf{A} \neq 0$.

Обернена матриця A^{-1} обчислюється так:

1. Знаходиться визначник (детермінант) заданої матриці A

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}. \tag{3.27}$$

2. Обчислюються алгебраїчні доповнення до елементів a_{ij} матриці (3.24) за формулою

$$A_{ij} = (-1)^{i+j} \cdot M_{ij}, \tag{3.28}$$

де M_{ij} – мінор елемента a_{ij}, тобто визначник матриці, що залишилася після викреслювання з даної матриці A i-го рядка та j-го стовпця.

3. Складається матриця $\widetilde{\mathbf{A}}$, елементами якої є алгебраїчні доповнення

$$\widetilde{\mathbf{A}} = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix}. \tag{3.29}$$

4. Транспонується матриця $\widetilde{\mathbf{A}}$ (матриця, у якої рядки поміняні місцями зі стовпцями, називається *транспонованою*)

$$C = \begin{bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{bmatrix}. \tag{3.30}$$

5. Обчислюються елементи оберненої матриці A^{-1} діленням елементів матриці C на детермінант початкової матриці

$$A^{-1}=\frac{C}{\det A}=\begin{bmatrix}\frac{A_{11}}{\det A} & \frac{A_{21}}{\det A} & \frac{A_{n1}}{\det A}\\ \frac{A_{12}}{\det A} & \frac{A_{22}}{\det A} & \frac{A_{n2}}{\det A}\\ \frac{A_{1n}}{\det A} & \frac{A_{2n}}{\det A} & \frac{A_{nn}}{\det A}\end{bmatrix}. \qquad (3.31)$$

У правильності обчислення оберненої матриці переконуються, виконуючи перевірку за формулою (3.25).

Приклад 3.1. Знайти матрицю A^{-1}, обернену до заданої

$$A=\begin{bmatrix}1 & 2 & 3\\ 3 & 2 & 1\\ 1 & 1 & 2\end{bmatrix}.$$

Розв'язок.

1. Обчислимо визначник матриці A

$$\det A=\begin{vmatrix}1 & 2 & 3\\ 3 & 2 & 1\\ 1 & 1 & 2\end{vmatrix}=1\cdot2\cdot2+3\cdot1\cdot3+2\cdot1\cdot1-3\cdot2\cdot1-3\cdot2\cdot2-1\cdot1\cdot1=-4.$$

2. Знайдемо алгебраїчні доповнення A_{ij}.

$$A_{11}=(-1)^{1+1}\cdot\begin{vmatrix}2 & 1\\ 1 & 2\end{vmatrix}=3 \quad A_{12}=(-1)^{1+2}\cdot\begin{vmatrix}3 & 1\\ 1 & 2\end{vmatrix}=-5 \quad A_{13}=(-1)^{1+3}\cdot\begin{vmatrix}3 & 2\\ 1 & 1\end{vmatrix}=1;$$

$$A_{21}=(-1)^{2+1}\cdot\begin{vmatrix}2 & 3\\ 1 & 2\end{vmatrix}=-1 \quad A_{22}=(-1)^{2+2}\cdot\begin{vmatrix}1 & 3\\ 1 & 2\end{vmatrix}=-1 \quad A_{23}=(-1)^{2+3}\cdot\begin{vmatrix}1 & 2\\ 1 & 1\end{vmatrix}=1;$$

$$A_{31}=(-1)^{3+1}\cdot\begin{vmatrix}2 & 3\\ 2 & 1\end{vmatrix}=-4 \quad A_{32}=(-1)^{3+2}\cdot\begin{vmatrix}1 & 3\\ 3 & 1\end{vmatrix}=8 \quad A_{33}=(-1)^{3+3}\cdot\begin{vmatrix}1 & 2\\ 3 & 2\end{vmatrix}=-4.$$

3. Складемо матрицю з алгебраїчних доповнень $\widetilde{\mathbf{A}}$

$$\widetilde{A}=\begin{bmatrix}3 & -5 & 1\\ -1 & -1 & 1\\ -4 & 8 & -4\end{bmatrix}.$$

4. Транспонуємо матрицю $\widetilde{\mathbf{A}}$

$$C=\begin{bmatrix}3 & -1 & -4\\ -5 & -1 & 8\\ 1 & 1 & -4\end{bmatrix}.$$

5. Запишемо обернену матрицю

$$A^{-1}=\begin{bmatrix}-0.75 & 0.25 & 1\\ 1.25 & 0.25 & -2\\ -0.25 & -0.25 & 1\end{bmatrix}.$$

Виконаємо перевірку обчислень

$$A\cdot A^{-1}=\begin{bmatrix}1 & 2 & 3\\ 3 & 2 & 1\\ 1 & 1 & 2\end{bmatrix}\cdot\begin{bmatrix}-0.75 & 0.25 & 1\\ 1.25 & 0.25 & -2\\ -0.25 & 0.25 & 1\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

3.4. Метод Крамера

3.4.1. Аналіз постановки задачі

Відомий з курсу математики метод Крамера є одним з найпростіших в теоретичному плані, проте одним з найбільш об'ємних при практичній реалізації, тобто вимагає здійснення дуже великої кількості арифметичних операцій, хоча по запису та операціях дуже простий. Метод використовується для визначення сумісності чи несумісності алгебраїчних рівнянь. Нехай задана система лінійних рівнянь:

$$A\bar{x}=\bar{b}, \tag{3.32}$$

$$A=\begin{pmatrix}a_{11}\dots a_{1n}\\ \text{-------}\\ a_{n1}\dots a_{nn}\end{pmatrix}\quad \bar{x}=\begin{pmatrix}x_1\\ \vdots\\ .x_n\end{pmatrix}\quad \bar{b}=\begin{pmatrix}b_1\\ \vdots\\ \mathrm{b_n}\end{pmatrix}. \tag{3.33}$$

Система (3.32) може бути записана в розгорнутому вигляді наступним чином:

$$\begin{cases}a_{11}x_1+a_{12}x_2+\dots+a_{1n}x_n=b_1\\ \text{----------}\\ a_{n1}x_1+a_{n2}x_2+\dots+a_{nn}x_n=b_n\end{cases} \tag{3.34}$$

Існування та єдиність розв'язку, відсутність розв'язку або ж наявність нескінченої кількості розв'язків можна з'ясувати за допомогою методу Крамера.

$$x_i=\frac{\Delta_i}{\Delta}, \tag{3.35}$$

де $i=\overline{1,n}$ $\Delta=\det A$, $\Delta_i=\det A_i$. Величина $\det A_i$ утворена з матриці A шляхом заміни i-го стовпчика стовпчиком вільних членів.

1) Якщо $\Delta\neq 0$, $\Delta_i\neq 0$, то існує єдиний розв'язок; система (3.32) сумісна;

Якщо $\Delta \neq 0$, $\Delta_i = 0$, то розв'язок нульовий;

2) Якщо $\Delta = 0$, $\Delta_i \neq 0$, то розв'язків немає; система (3.32) несумісна;

3) Якщо $\Delta = 0$, $\Delta_i = 0$, то існує безліч розв'язків; система (3.32) невизначена.

Значення невідомих x_i (i=1, 2, ..., n) одержують за формулами Крамера:

$$x_i = \frac{\det \mathbf{A}_i}{\det \mathbf{A}}, \tag{3.36}$$

де $\det \mathbf{A}$ – детермінант матриці A; $\det \mathbf{A}_i$ – детермінант матриці A_i, яку одержано з матриці A шляхом заміни i-го стовпця на стовпець вільних членів.

3.4.2. Приклади розв'язку системи за формулами Крамера

Приклад 3.2. Розв'язати систему лінійних рівнянь методом Крамера

$$\begin{cases} 9x_1 + 3x_2 + x_3 = 2 \\ 4x_1 + 2x_2 + x_3 = 1 \\ 18x_1 + 6x_2 + 2x_3 = 3 \end{cases}.$$

Розв'язок. Знайдемо $\det A$

$$\det A = \begin{vmatrix} 9 & 3 & 1 \\ 4 & 2 & 1 \\ 18 & 6 & 2 \end{vmatrix} = 9\cdot 2\cdot 2 + 3\cdot 1\cdot 18 + 4\cdot 6\cdot 1 - 18\cdot 2\cdot 1 - 4\cdot 3\cdot 2 - 9\cdot 6\cdot 1 = 0.$$

Визначник завжди буде рівний 0, якщо присутні однакові або кратні рядки чи стовпчики. В прикладі 3.2 перший рядок кратний третьому. Навіть при транспонуванні матриці або переставлянні рядків місцями визначник буде рівний 0. Кратні рядки або стовпчики можна вилучати з матриці. Система лінійних рівнянь розв'язків немає та є несумісною.

$$\det B = \begin{vmatrix} 9 & 4 & 18 \\ 3 & 2 & 6 \\ 1 & 1 & 2 \end{vmatrix} = 9\cdot 2\cdot 2 + 3\cdot 1\cdot 18 + 4\cdot 6\cdot 1 - 18\cdot 2\cdot 1 - 4\cdot 3\cdot 2 - 9\cdot 6\cdot 1 = 0.$$

$$\det C = \begin{vmatrix} 4 & 2 & 1 \\ 9 & 3 & 1 \\ 18 & 6 & 2 \end{vmatrix} = 4\cdot 3\cdot 2 + 2\cdot 1\cdot 18 + 9\cdot 6\cdot 1 - 18\cdot 3\cdot 1 - 6\cdot 1\cdot 4 - 9\cdot 2\cdot 2 = 0.$$

Наприклад, для

$$\begin{cases} 3x_1 + x_2 + 9x_3 = 2 \\ 4x_1 + x_2 + 12x_3 = 1 \\ 5x_1 + x_2 + 15x_3 = 3 \end{cases}.$$

також визначник буде рівний 0:

$$\det A = \begin{vmatrix} 3 & 1 & 9 \\ 4 & 1 & 12 \\ 5 & 1 & 15 \end{vmatrix} = 3\cdot 1\cdot 15 + 4\cdot 1\cdot 9 + 12\cdot 5\cdot 1 - 5\cdot 1\cdot 9 - 12\cdot 1\cdot 3 - 4\cdot 15\cdot 1 = 0.$$

Приклад 3.3. Розв'язати систему лінійних рівнянь методом Крамера

$$\begin{cases} x_1 + 2x_2 - 4x_3 = 2 \\ 3x_1 - x_2 + x_3 = 1 \\ 2x_1 + 2x_2 - x_3 = 3 \end{cases}.$$

Розв'язок. Знайдемо $\det A$, $\det A_1$, $\det A_2$, $\det A_3$:

$$\det A = \begin{vmatrix} 1 & 2 & -4 \\ 3 & -1 & 1 \\ 2 & 2 & -1 \end{vmatrix} = -23; \qquad \det A_1 = \begin{vmatrix} 2 & 2 & -4 \\ 1 & -1 & 1 \\ 3 & 2 & -1 \end{vmatrix} = -14;$$

$$\det A_2 = \begin{vmatrix} 1 & 2 & -4 \\ 3 & 1 & 1 \\ 2 & 3 & -1 \end{vmatrix} = -22; \qquad \det A_3 = \begin{vmatrix} 1 & 2 & 2 \\ 3 & -1 & 1 \\ 2 & 2 & 3 \end{vmatrix} = -3.$$

Тоді

$$x_1 = \frac{\det A_1}{\det A} = \frac{-14}{-23} = 0{,}6087; \qquad x_2 = \frac{\det A_2}{\det A} = \frac{-22}{-23} = 0{,}9565;$$

$$x_3 = \frac{\det A_3}{\det A} = \frac{-3}{-23} = 0{,}1304.$$

Приклад 3.4. Розв'язати систему лінійних рівнянь методом Крамера

$$\begin{cases} x_1 + 2x_2 + x_3 = 5 \\ 3x_1 + 4x_2 + x_3 = 11 \\ 7x_1 + 2x_2 + 5x_3 = 21 \end{cases}.$$

Розв'язок. Знайдемо $\det A$, $\det A_1$, $\det A_2$, $\det A_3$:

$$\det A = \begin{vmatrix} 1 & 2 & 1 \\ 3 & 4 & 1 \\ 7 & 2 & 5 \end{vmatrix} = 20 + 14 + 6 - 28 - 30 - 2 = 40 - 60 = -20 \neq 0.$$

$$x_1 = \frac{\Delta_1}{\Delta} = -\frac{1}{20}\begin{vmatrix} 5 & 2 & 1 \\ 11 & 4 & 1 \\ 21 & 2 & 5 \end{vmatrix} = -\frac{1}{20}(100 + 22 + 42 - 84 - 110 - 10) = -\frac{1}{20}(-40) = 2.$$

$$x_2 = \frac{\Delta_2}{\Delta} = -\frac{1}{20}\begin{vmatrix} 1 & 5 & 1 \\ 3 & 11 & 1 \\ 7 & 21 & 5 \end{vmatrix} = -\frac{1}{20}(55 + 35 + 63 - 77 - 75 - 21) = -\frac{1}{20}(-20) = 1.$$

$$x_3 = \frac{\Delta_3}{\Delta} = -\frac{1}{20}\begin{vmatrix} 1 & 2 & 5 \\ 3 & 4 & 11 \\ 7 & 2 & 21 \end{vmatrix} = -\frac{1}{20}(84 + 154 + 30 - 140 - 126 - 22) = -\frac{1}{20}(-20) = 1.$$

Для того, щоб засвоїти цю частину необхідно вміти розкривати визначники «два на два» і «три на три». Спочатку докладно розглянемо правило Крамера для системи двох лінійних рівнянь з двома невідомими.

Нехай іноді, але зустрічається завдання – розв'язати систему двох лінійних рівнянь з двома невідомими за формулами Крамера. Крім того, простий приклад допоможе зрозуміти, як використовувати правило Крамера для більш складного випадку – системи трьох рівнянь з трьома невідомими. Крім того, існують системи лінійних рівнянь з двома змінними, які доцільно вирішувати саме за правилом Крамера!

Розглянемо систему рівнянь

$$\begin{cases} a_1 x + b_1 y = s_1, \\ a_2 x + b_2 y = s_2. \end{cases} \tag{3.37}$$

Алгоритм Крамера.

Крок 1. Обчислити *головний визначник системи* визначник $\Delta = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$.

Крок 2. Якщо $\Delta = 0$, то система має нескінченно багато розв'язків або несумісна (не має розв'язків). У цьому випадку правило Крамера не допоможе, потрібно використовувати метод Гауса. Якщо $\Delta \neq 0$, то система має єдиний розв'язок, і для знаходження коренів необхідно обчислити ще два визначника: $\Delta_x = \begin{vmatrix} s_1 & b_1 \\ s_2 & b_2 \end{vmatrix}$ та $\Delta_y = \begin{vmatrix} a_1 & s_1 \\ a_2 & s_2 \end{vmatrix}$. На практиці вищевказані визначники також можуть позначатися латинською літерою D.

Крок 3. Корені рівняння знаходимо за формулами: $x = \frac{\Delta_x}{\Delta}$ та $y = \frac{\Delta_y}{\Delta}$.

Приклад 3.5. Розв'язати систему лінійних рівнянь $\begin{cases} 506a + 66b = 2315{,}1; \\ 66x + 11y = 392{,}3. \end{cases}$

Розв'язок. Коефіцієнти рівняння досить великі, в правій частині присутні десяткові дроби з комою, яка рідко зустрічається у практичних завданнях з математики. Щоб розв'язати таку систему, можна спробувати виразити одну змінну через іншу, але в цьому випадку вийдуть дроби, з якими вкрай незручно працювати. Можна помножити друге рівняння на 6 і провести по елементне віднімання, але й тут виникнуть ті ж дроби.

У подібних випадках і приходять на допомогу формули Крамера.

$$\Delta = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \begin{vmatrix} 506 & 66 \\ 66 & 11 \end{vmatrix} = 506 \cdot 11 - 66 \cdot 66 = 5566 - 4356 = 1210 \neq 0,$$

отже, система має єдиний розв'язок.

$$\Delta_a = \begin{vmatrix} s_1 & b_1 \\ s_2 & b_2 \end{vmatrix} = \begin{vmatrix} 2315{,}1 & 66 \\ 392{,}3 & 11 \end{vmatrix} = 2315{,}1 \cdot 11 - 392{,}3 \cdot 66 = 25466{,}1 - 25891{,}8 = -425{,}7;$$

$$\Delta_b = \begin{vmatrix} a_1 & s_1 \\ a_2 & s_2 \end{vmatrix} = \begin{vmatrix} 506 & 2315{,}1 \\ 66 & 392{,}3 \end{vmatrix} = 66 \cdot 392{,}3 - 66 \cdot 2315{,}1 = 198503{,}8 - 152796{,}6 = 45707{,}2;$$

$$a = \frac{\Delta_a}{\Delta} = \frac{-425{,}7}{1210} \approx -0{,}35; \; b = \frac{\Delta_b}{\Delta} = \frac{45707{,}2}{1210} \approx 37{,}77.$$

Тоді $a \approx -0{,}35$ та $b \approx 37{,}77$. Обидва корені знайдені наближено, що цілком прийнятно для завдань економетрики. Коментарі тут не потрібні, оскільки завдання розв'язується за відомими формулами, однак, є один нюанс. Коли використовуєте даний метод, обов'язковим фрагментом оформлення завдання є наступне: "$\neq 0$, значить, система має єдиний розв'язок ". Зовсім не зайвою буде перевірка, яку зручно провести на калькуляторі: підставляємо наближені значення $a \approx -0{,}35$ та $b \approx 37{,}77$ в ліву частину кожного з рівнянь системи. В результаті з невеликою похибкою повинні вийти числа, які знаходяться в правих частинах.

Переходимо до розгляду правила Крамера для системи трьох рівнянь з трьома невідомими:

$$\begin{cases} a_1 x_1 + b_1 x_2 + c_1 x_3 = s_1, \\ a_2 x_1 + b_2 x_2 + c_2 x_3 = s_2, \\ a_3 x_1 + b_3 x_2 + c_3 x_3 = s_3. \end{cases} \tag{3.38}$$

Знаходимо головний визначник системи: $\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$. Якщо $\Delta = 0$, то система має нескінченно багато розв'язків або несумісна (не має розв'язків). У цьому випадку правило Крамера не допоможе, потрібно використовувати метод Гауса. Якщо $\Delta \neq 0$, то система має єдиний розв'язок і для знаходження коренів ми повинні обчислити ще три визначника:

$$\Delta_1 = \begin{vmatrix} s_1 & b_1 & c_1 \\ s_2 & b_2 & c_2 \\ s_3 & b_3 & c_3 \end{vmatrix}, \; \Delta_2 = \begin{vmatrix} a_1 & s_1 & c_1 \\ a_2 & s_2 & c_2 \\ a_3 & s_3 & c_3 \end{vmatrix}, \; \Delta_3 = \begin{vmatrix} a_1 & b_1 & s_1 \\ a_2 & b_2 & s_2 \\ a_3 & b_3 & s_3 \end{vmatrix}.$$

І, нарешті, відповідь обчислюється за формулами: $x_1 = \frac{\Delta_1}{\Delta}$, $x_2 = \frac{\Delta_2}{\Delta}$ та $x_3 = \frac{\Delta_3}{\Delta}$.

Як видно, випадок «три на три» принципово нічим не відрізняється від випадку «два на два», стовпець вільних членів s_1, s_2, s_3 послідовно зчитується зліва направо по стовпчиках головного визначника.

Приклад 3.6. Розв'язати систему за формулами Крамера:

$$\begin{cases} 3x_1 - 2x_2 + 4x_3 = 21, \\ 3x_1 + 4x_2 - 2x_3 = 9, \\ 2x_1 - x_2 - x_3 = 10. \end{cases}$$

Розв'язок. Розв'яжемо систему за формулами Крамера.

$$\Delta = \begin{vmatrix} 3 & -2 & 4 \\ 3 & 4 & -2 \\ 2 & -1 & -1 \end{vmatrix} = 3 \cdot \begin{vmatrix} 4 & -2 \\ -1 & -1 \end{vmatrix} + 2 \cdot \begin{vmatrix} 3 & -2 \\ 2 & -1 \end{vmatrix} + 4 \cdot \begin{vmatrix} 3 & 4 \\ 2 & -1 \end{vmatrix} =$$

$$= 3 \cdot (-4-2) + 2 \cdot (-3+4) + 4 \cdot (-3-8) = -18 + 2 - 44 = -60 \neq 0,$$

отже, система має єдиний розв'язок.

$$\Delta_1 = \begin{vmatrix} 21 & -2 & 4 \\ 9 & 4 & -2 \\ 10 & -1 & -1 \end{vmatrix} = 21 \cdot \begin{vmatrix} 4 & -2 \\ -1 & -1 \end{vmatrix} + 2 \cdot \begin{vmatrix} 9 & -2 \\ 10 & -1 \end{vmatrix} + 4 \cdot \begin{vmatrix} 9 & 4 \\ 10 & -1 \end{vmatrix} =$$

$$= 21 \cdot (-4-2) + 2 \cdot (-9+20) + 4 \cdot (-9-40) = -126 + 22 - 196 = -300,$$

$$\Delta_2 = \begin{vmatrix} 3 & 21 & 4 \\ 3 & 9 & -2 \\ 2 & 10 & -1 \end{vmatrix} = 3 \cdot \begin{vmatrix} 9 & -2 \\ 10 & -1 \end{vmatrix} + 21 \cdot \begin{vmatrix} 3 & -2 \\ 2 & -1 \end{vmatrix} + 4 \cdot \begin{vmatrix} 3 & 9 \\ 2 & 10 \end{vmatrix} =$$

$$= 3 \cdot (-9+20) - 21 \cdot (-3+4) + 4 \cdot (30-18) = -33 - 21 + 48 = 60,$$

$$\Delta_3 = \begin{vmatrix} 3 & -2 & 21 \\ 3 & 4 & 9 \\ 2 & -1 & 10 \end{vmatrix} = 3 \cdot \begin{vmatrix} 4 & 9 \\ -1 & 10 \end{vmatrix} - 3 \cdot \begin{vmatrix} -2 & 21 \\ -1 & 10 \end{vmatrix} + 2 \cdot \begin{vmatrix} -2 & 21 \\ 4 & 9 \end{vmatrix} =$$

$$= 3 \cdot (40+9) - 3 \cdot (-20+21) + 2 \cdot (-18-88) = 147 - 3 - 284 = -60,$$

$$x_1 = \frac{\Delta_1}{\Delta} = \frac{-300}{-60} = 5; \; x_2 = \frac{\Delta_2}{\Delta} = \frac{60}{-60} = -1; \; x_3 = \frac{\Delta_3}{\Delta} = \frac{-60}{-60} = 1;$$

Тоді $x_1 = 5; \; x_2 = -1; \; x_3 = 1.$

Розв'язок знаходиться за готовими формулами. Але є декілька зауважень.

Зауваження 3.1. Буває так, що в результаті обчислень виходять «погані» дроби, які не скорочуються, наприклад: $x_1 = \frac{53}{17}$. Рекомендують застосувати наступний алгоритм. Якщо під рукою немає комп'ютера, чинимо так:

1) Можливо, допущена помилка в обчисленнях. Як тільки зіткнулися з «поганим» дробом, відразу необхідно перевірити, чи правильно переписано умова. Якщо умова переписана без помилок, то потрібно перерахувати визначники, використовуючи розклад по іншому рядку (стовпцю).

2) Якщо в результаті перевірки помилок не виявлено, то найімовірніше, допущена помилка в умові завдання. В цьому випадку розв'язуємо завдання до кінця.

Приклад 3.7. Розв'язати систему за формулами Крамера.

$$\begin{cases} 7x + y = 23, \\ -5x + 3y = 1. \end{cases}$$

Розв'язок. Розв'яжемо систему за формулами Крамера.

$\Delta = \begin{vmatrix} 7 & 1 \\ -5 & 3 \end{vmatrix} = 21 + 5 = 26 \neq 0$, отже, система має єдиний розв'язок.

$$\Delta_x = \begin{vmatrix} 23 & 1 \\ 1 & 3 \end{vmatrix} = 69 - 1 = 68,\ x = \frac{\Delta_x}{\Delta} = \frac{68}{26} = \frac{34}{13};$$

$$\Delta_y = \begin{vmatrix} 7 & 23 \\ -5 & 1 \end{vmatrix} = 7 + 115 = 122,\ y = \frac{\Delta_y}{\Delta} = \frac{122}{26} = \frac{61}{13};$$

Тоді $x = \frac{34}{13};\ y = \frac{61}{13}$.

Перевірка. Підставимо знайдений розв'язок в ліву частину кожного рівняння системи.

1) $7 \cdot \frac{34}{13} + \frac{61}{13} = \frac{238}{13} + \frac{61}{13} = \frac{299}{13} = 23.$

2) $-5 \cdot \frac{34}{13} + 3 \cdot \frac{61}{13} = -\frac{170}{13} + \frac{183}{13} = \frac{13}{13} = 1.$

Отримано відповідні праві частини, таким чином розв'язок знайдено правильно.

Зауваження 3.2. Деколи зустрічаються системи, в рівняннях яких відсутні деякі змінні, наприклад:

$$\begin{cases} 3x_2 + 4x_3 = 21, \\ 3x_1 + 4x_2 = 9, \\ 2x_1 - x_2 - x_3 = 10. \end{cases}$$

Тут в першому рівнянні відсутня змінна x_1, у другому – змінна x_3. В таких випадках дуже важливо правильно і уважно записати головний визначник (на місці відсутніх змінних ставляться нулі):

$$\Delta = \begin{vmatrix} 0 & 3 & 4 \\ 3 & 4 & 0 \\ 2 & -1 & -1 \end{vmatrix}.$$

Визначники з нулями раціонально розкривати по тому рядку (стовпцю), в якому є нуль, так як обчислень виходить значно менше.

Приклад 3.8. Розв'язати систему за формулами Крамера.

$$\begin{cases} 2x - 3y + z = 2, \\ 2x + y - 4z = 9, \\ 6x - 5y + 2z = 17. \end{cases}$$

Розв'язок. Розв'яжемо систему за формулами Крамера.

$$\Delta = \begin{vmatrix} 2 & -3 & 1 \\ 2 & 1 & -4 \\ 6 & -5 & 2 \end{vmatrix} = 2 \cdot \begin{vmatrix} 1 & -4 \\ -5 & 2 \end{vmatrix} + 3 \cdot \begin{vmatrix} 2 & -4 \\ 6 & 2 \end{vmatrix} + 1 \cdot \begin{vmatrix} 2 & 1 \\ 6 & -5 \end{vmatrix} =$$

$$= 2 \cdot (2 - 20) + 3 \cdot (4 + 24) + 1 \cdot (-10 - 6) = -36 + 84 - 16 = 32 \neq 0,$$

отже, система має єдиний розв'язок.

$$\Delta_x = \begin{vmatrix} 2 & -3 & 1 \\ 9 & 1 & -4 \\ 17 & -5 & 2 \end{vmatrix} = 2 \cdot \begin{vmatrix} 1 & -4 \\ -5 & 2 \end{vmatrix} + 3 \cdot \begin{vmatrix} 9 & -4 \\ 17 & 2 \end{vmatrix} + 1 \cdot \begin{vmatrix} 9 & 1 \\ 17 & -5 \end{vmatrix} =$$

$$= 2 \cdot (2 - 20) + 3 \cdot (18 + 68) + 1 \cdot (-45 - 17) = -36 + 258 - 62 = 160,$$

$$\Delta_y = \begin{vmatrix} 2 & 2 & 1 \\ 2 & 9 & -4 \\ 6 & 17 & 2 \end{vmatrix} = 2 \cdot \begin{vmatrix} 9 & -4 \\ 17 & 2 \end{vmatrix} - 2 \cdot \begin{vmatrix} 2 & -4 \\ 6 & 2 \end{vmatrix} + 1 \cdot \begin{vmatrix} 2 & 9 \\ 6 & 17 \end{vmatrix} =$$

$$= 2 \cdot (18 + 68) - 2 \cdot (4 + 24) + 1 \cdot (34 - 54) = 172 - 56 - 20 = 96,$$

$$\Delta_z = \begin{vmatrix} 2 & -3 & 2 \\ 2 & 1 & 9 \\ 6 & -5 & 17 \end{vmatrix} = 2 \cdot \begin{vmatrix} 1 & 9 \\ -5 & 17 \end{vmatrix} + 3 \cdot \begin{vmatrix} 2 & 9 \\ 6 & 17 \end{vmatrix} + 2 \cdot \begin{vmatrix} 2 & 1 \\ 6 & -5 \end{vmatrix} =$$

$$= 2 \cdot (17 + 45) + 3 \cdot (34 - 54) + 2 \cdot (-10 - 6) = 124 - 60 - 32 = 32,$$

$$x = \frac{\Delta_x}{\Delta} = \frac{160}{32} = 5;\ y = \frac{\Delta_y}{\Delta} = \frac{96}{32} = 3;\ z = \frac{\Delta_z}{\Delta} = \frac{32}{32} = 1;$$

Тоді $x = 5;\ y = 3;\ z = 1.$

Для випадку системи 4-х рівнянь з 4-ма невідомими формули Крамера записуються за аналогічним принципом. Зниження порядку визначника -- п'ять визначників 4-го порядку цілком придатні для розв'язування.

3.5. Метод оберненої матриці

Метод оберненої матриці - це, по суті, окремий випадок матричного рівняння. Для вивчення цього параграфа необхідно вміти розкривати визначники, знаходити обернену матрицю та виконувати матричне множення. Відповідні посилання будуть надані в процесі пояснень. Система лінійних рівнянь записана у матричному вигляді в (3.2). Розв'язування матричного рівняння зводиться до знаходження оберненої матриці $\mathbf{A}^{-1}$ і наступного множення її на матрицю вільних членів $\boldsymbol{B}$, тобто

$$\mathbf{X} = \mathbf{A}^{-1} \cdot \mathbf{B}. \tag{3.39}$$

Приклад 3.9. Розв'язати систему матричним методом

$$\begin{cases} 3x_1 - 2x_2 + 4x_3 = 21, \\ 3x_1 + 4x_2 - 2x_3 = 9, \\ 2x_1 - x_2 - x_3 = 10. \end{cases}$$

Розв'язок. Запишемо систему в матричної формі: $AX = B$, де

$$A = \begin{pmatrix} 3 & -2 & 4 \\ 3 & 4 & -2 \\ 2 & -1 & -1 \end{pmatrix}, X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, B = \begin{pmatrix} 21 \\ 9 \\ 10 \end{pmatrix}.$$

Якби в рівняннях були відсутні деякі змінні, то на відповідних місцях в матриці потрібно було поставити нулі. Розв'язок системи знайдемо за формулою $\mathbf{X} = \mathbf{A}^{-1} \cdot \mathbf{B}$. Відповідно до формули потрібно знайти зворотну матрицю $\mathbf{A}^{-1}$ і виконати матричне множення $\mathbf{A}^{-1} \cdot \mathbf{B}$. Обернену матрицю знайдемо за формулою: $A^{-1} = \frac{1}{|A|} \widetilde{A}^T$, де $\widetilde{A}^T$ - транспонована матриця алгебраїчних доповнень відповідних елементів матриці. Спочатку розглянемо визначник:

$$|A| = \Delta = \begin{vmatrix} 3 & -2 & 4 \\ 3 & 4 & -2 \\ 2 & -1 & -1 \end{vmatrix} = 3 \cdot \begin{vmatrix} 4 & -2 \\ -1 & -1 \end{vmatrix} + 2 \cdot \begin{vmatrix} 3 & -2 \\ 2 & -1 \end{vmatrix} + 4 \cdot \begin{vmatrix} 3 & 4 \\ 2 & -1 \end{vmatrix} =$$

$$= 3 \cdot (-4 - 2) + 2 \cdot (-3 + 4) + 4 \cdot (-3 - 8) = -18 + 2 - 44 = -60 \neq 0.$$

Тут визначник розкритий по першому рядку. Якщо $|A| = 0$, то обернена матриця не існує, і розв'язати систему матричним методом неможливо. В цьому випадку система розв'язується методом виключення невідомих (методом Гауса). Тепер потрібно обчислити 9 мінорів і записати їх в матрицю мінорів

$$M = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix}.$$

В процесі розв'язку мінори краще розписати докладно, хоча, при певному досвіді їх можна пристосуватися рахувати усно.

$$M_{11} = \begin{vmatrix} 4 & -2 \\ -1 & -1 \end{vmatrix} = -4 - 2 = -6;\ M_{12} = \begin{vmatrix} 3 & -2 \\ 2 & -1 \end{vmatrix} = -3 + 4 = 1;\ M_{13} = \begin{vmatrix} 3 & 4 \\ 2 & -1 \end{vmatrix} = -3 - 8 = -11;$$

$$M_{21} = \begin{vmatrix} -2 & 4 \\ -1 & -1 \end{vmatrix} = 2 + 4 = 6;\ M_{22} = \begin{vmatrix} 3 & 4 \\ 2 & -1 \end{vmatrix} = -3 - 8 = -11;\ M_{23} = \begin{vmatrix} 3 & -2 \\ 2 & -1 \end{vmatrix} = -3 + 4 = 1;$$

$$M_{31} = \begin{vmatrix} -2 & 4 \\ 4 & -2 \end{vmatrix} = 4 - 16 = -12;\ M_{32} = \begin{vmatrix} 3 & 4 \\ 3 & -2 \end{vmatrix} = -6 - 12 = -18;$$

$$M_{33} = \begin{vmatrix} 3 & -2 \\ 3 & 4 \end{vmatrix} = 12 + 6 = 18$$

Порядок розрахунку мінорів абсолютно не важливий, тут обчислено зліва направо по рядках. Можна розрахувати мінори по стовпчиках (це навіть зручніше). Таким чином:

$M = \begin{pmatrix} -6 & 1 & -11 \\ 6 & -11 & 1 \\ -12 & -18 & 18 \end{pmatrix}$ – матриця мінорів відповідних елементів матриці A.

$\widetilde{A} = \begin{pmatrix} -6 & -1 & -11 \\ -6 & -11 & -1 \\ -12 & 18 & 18 \end{pmatrix}$ – матриця алгебраїчних доповнень.

$\widetilde{A}^T = \begin{pmatrix} -6 & -6 & -12 \\ -1 & -11 & 18 \\ -11 & -1 & 18 \end{pmatrix}$ – транспонована матриця алгебраїчних доповнень.

Тепер записуємо обернену матрицю:

$$A^{-1} = \frac{1}{|A|}\widetilde{A}^T = \frac{1}{-60}\begin{pmatrix} -6 & -6 & -12 \\ -1 & -11 & 18 \\ -11 & -1 & 18 \end{pmatrix} = \frac{1}{60}\begin{pmatrix} 6 & 6 & 12 \\ 1 & 11 & -18 \\ 11 & 1 & -18 \end{pmatrix}.$$

Не вносимо $\frac{1}{60}$ в матрицю, це серйозно ускладнить подальші обчислення. Ділення потрібно було б виконати, якби всі числа матриці ділилися на 60 без залишку. А внести мінус в матрицю в даному випадку дуже потрібно, це, навпаки – спростить подальші обчислення. Залишилося провести матричне множення.

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = A^{-1}B = \frac{1}{60}\begin{pmatrix} 6 & 6 & 12 \\ 1 & 11 & -18 \\ 11 & 1 & -18 \end{pmatrix} \cdot \begin{pmatrix} 21 \\ 9 \\ 10 \end{pmatrix} = \frac{1}{60}\begin{pmatrix} 6\cdot 21+6\cdot 9+12\cdot 10 \\ 1\cdot 21+11\cdot 9-18\cdot 10 \\ 11\cdot 21+1\cdot 9-18\cdot 10 \end{pmatrix} = \frac{1}{60}\begin{pmatrix} 300 \\ -60 \\ 60 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \\ 1 \end{pmatrix}.$$

Зверніть увагу, що ділення на 60 виконується в останню чергу. Іноді може і не розділитися без залишку, тобто можуть вийти «погані» дроби. Що в таких випадках робити подано в розділі про правило Крамера. Отримали $x_1 = 5$; $x_2 = -1$; $x_3 = 1$.

Приклад 3.10. Розв'язати систему лінійних рівнянь матричним методом.

$$\begin{cases} x_1 + x_2 + 2x_3 = -1 \\ 2x_1 - x_2 + 2x_3 = -4 \\ 4x_1 + x_2 + 4x_3 = -2 \end{cases}.$$

Розв'язок. Початкова матриця коефіцієнтів при невідомих має вигляд:

$$A=\begin{bmatrix}1 & 1 & 2\\ 2 & -1 & 2\\ 4 & 1 & 4\end{bmatrix}.$$

Знайдемо обернену матрицю A^{-1}.

$$A^{-1}=\begin{bmatrix}-1 & -0{,}3333 & 0{,}6667\\ 0 & -0{,}6667 & 0{,}3333\\ 1 & 0{,}5 & -0{,}5\end{bmatrix}.$$

Тоді при розв'язуванні системи одержимо:

$$X=\begin{bmatrix}-1 & -0{,}3333 & 0{,}6667\\ 0 & -0{,}6667 & 0{,}3333\\ 1 & 0{,}5 & -0{,}5\end{bmatrix}\cdot\begin{bmatrix}-1\\ -4\\ -2\end{bmatrix}=\begin{bmatrix}1\\ 2\\ -2\end{bmatrix}.$$

Таким чином, $x_1=1;\quad x_2=2;\quad x_3=-2.$

Приклад 3.11. Розв'язати систему за допомогою оберненої матриці.

$$\begin{cases}2x-3y+z=2,\\ 2x+y-4z=9,\\ 6x-5y+2z=17.\end{cases}$$

Розв'язок. Запишемо систему в матричні формі:

$$AX=B,\ \text{де}\ A=\begin{pmatrix}2 & -3 & 1\\ 2 & 1 & -4\\ 6 & -5 & 2\end{pmatrix},\ X=\begin{pmatrix}x\\ y\\ z\end{pmatrix},\ B=\begin{pmatrix}2\\ 9\\ 17\end{pmatrix}.$$

Розв'язок рівняння шукаємо у вигляді $X=A^{-1}B$. Обернену матрицю знайдемо за формулою $A^{-1}=\dfrac{1}{|A|}\widetilde{A}^T$, де $\widetilde{A}^T$ – транспонована матриця алгебраїчних доповнень відповідних елементів матриці A.

$$|A|=\Delta=\begin{vmatrix}2 & -3 & 1\\ 2 & 1 & -4\\ 6 & -5 & 2\end{vmatrix}=2\cdot\begin{vmatrix}1 & -4\\ -5 & 2\end{vmatrix}+3\cdot\begin{vmatrix}2 & -4\\ 6 & 2\end{vmatrix}+1\cdot\begin{vmatrix}2 & 1\\ 6 & -5\end{vmatrix}=$$

$$=2\cdot(2-20)+3\cdot(4+24)+1\cdot(-10-6)=-36+84-16=32.$$

$M=\begin{pmatrix}-18 & 28 & -16\\ -1 & -2 & 8\\ 11 & -10 & 8\end{pmatrix}$ – матриця мінорів відповідних елементів матриці A.

$\widetilde{A} = \begin{pmatrix} -18 & -28 & -16 \\ 1 & -2 & -8 \\ 11 & 10 & 8 \end{pmatrix}$ – матриця алгебраїчних доповнень.

$\widetilde{A}^T = \begin{pmatrix} -18 & 1 & 11 \\ -28 & -2 & 10 \\ -16 & -8 & 8 \end{pmatrix}$ – транспонована матриця алгебраїчних доповнень.

Тепер записуємо обернену матрицю:

$$A^{-1} = \frac{1}{|A|}\widetilde{A}^T = \frac{1}{32}\begin{pmatrix} -18 & 1 & 11 \\ -28 & -2 & 10 \\ -16 & -8 & 8 \end{pmatrix}.$$

Таким чином:

$$X = A^{-1}B = \frac{1}{32}\begin{pmatrix} -18 & 1 & 11 \\ -28 & -2 & 10 \\ -16 & -8 & 8 \end{pmatrix}\cdot\begin{pmatrix} 2 \\ 9 \\ 17 \end{pmatrix} = \frac{1}{32}\begin{pmatrix} -18\cdot 2+1\cdot 9+11\cdot 17 \\ -28\cdot 2-2\cdot 9+10\cdot 17 \\ -16\cdot 2-8\cdot 9+8\cdot 17 \end{pmatrix} = \frac{1}{32}\begin{pmatrix} 160 \\ 96 \\ 32 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 1 \end{pmatrix}.$$

Звідси, $x=5; y=3; z=1$.

Приклад 3.12. Розв'язати систему за допомогою оберненої матриці.

$$\begin{cases} x+2y+z=5, \\ 3x+4y+z=11, \\ 7x+2y+5z=21. \end{cases}$$

Розв'язок. Запишемо систему в матричні формі:

$$AX = B, \text{ де } A = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 4 & 1 \\ 7 & 2 & 5 \end{pmatrix}, X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, B = \begin{pmatrix} 5 \\ 11 \\ 21 \end{pmatrix}.$$

Розв'язок рівняння шукаємо у вигляді $X = A^{-1}B$. Обернену матрицю знайдемо за формулою $A^{-1} = \frac{1}{|A|}\widetilde{A}^T$, де $\widetilde{A}^T$ – транспонована матриця алгебраїчних доповнень відповідних елементів матриці A.

$$|A| = \Delta = \begin{vmatrix} 1 & 2 & 1 \\ 3 & 4 & 1 \\ 7 & 2 & 5 \end{vmatrix} = 1\cdot\begin{vmatrix} 4 & 1 \\ 2 & 5 \end{vmatrix} - 2\cdot\begin{vmatrix} 3 & 1 \\ 7 & 5 \end{vmatrix} + 1\cdot\begin{vmatrix} 3 & 4 \\ 7 & 2 \end{vmatrix} =$$

$$= 1\cdot(20-2) - 2\cdot(15-7) + 1\cdot(6-28) = 18-16-22 = -20.$$

$$M_{11} = \begin{vmatrix} 4 & 1 \\ 2 & 5 \end{vmatrix} = 18;\ M_{12} = \begin{vmatrix} 3 & 1 \\ 7 & 5 \end{vmatrix} = 8;\ M_{13} = \begin{vmatrix} 3 & 4 \\ 7 & 2 \end{vmatrix} = -22;$$

$$M_{21} = \begin{vmatrix} 2 & 1 \\ 2 & 5 \end{vmatrix} = 8;\ M_{22} = \begin{vmatrix} 1 & 1 \\ 7 & 5 \end{vmatrix} = -2;\ M_{23} = \begin{vmatrix} 1 & 2 \\ 7 & 2 \end{vmatrix} = -12;$$

$$M_{31}=\begin{vmatrix}2 & 1\\ 4 & 1\end{vmatrix}=-2;M_{32}=\begin{vmatrix}1 & 1\\ 3 & 1\end{vmatrix}=-2;M_{33}=\begin{vmatrix}1 & 2\\ 3 & 4\end{vmatrix}=-2$$

$M=\begin{pmatrix}18 & 8 & -22\\ 8 & -2 & -12\\ -2 & -2 & -2\end{pmatrix}$ – матриця мінорів відповідних елементів матриці A.

$\widetilde{A}=\begin{pmatrix}18 & -8 & -22\\ -8 & -2 & 12\\ -2 & 2 & -2\end{pmatrix}$ – матриця алгебраїчних доповнень.

$\widetilde{A}^T=\begin{pmatrix}-18 & -8 & -2\\ -8 & -2 & 2\\ -22 & 12 & -2\end{pmatrix}$ – транспонована матриця алгебраїчних доповнень.

Тепер записуємо обернену матрицю:

$$A^{-1}=\frac{1}{|A|}\widetilde{A}^T=-\frac{1}{20}\begin{pmatrix}-18 & -8 & -2\\ -8 & -2 & 2\\ -22 & 12 & -2\end{pmatrix}.$$

$$X=A^{-1}B=-\frac{1}{20}\begin{pmatrix}-18 & -8 & -2\\ -8 & -2 & 2\\ -22 & 12 & -2\end{pmatrix}\cdot\begin{pmatrix}5\\ 11\\ 21\end{pmatrix}=\begin{pmatrix}-9/10 & 2/5 & 1/10\\ 2/5 & 1/10 & -1/10\\ 11/10 & -3/5 & 1/10\end{pmatrix}\cdot\begin{pmatrix}5\\ 11\\ 21\end{pmatrix}=$$

$$=\begin{pmatrix}-9/10\cdot 5+4/10\cdot 11+1/10\cdot 21\\ -9/10\cdot 5+4/10\cdot 11+1/10\cdot 21\\ -9/10\cdot 5+4/10\cdot 11+1/10\cdot 21\end{pmatrix}=\begin{pmatrix}(-45+44+21)/10\\ (20+11-21)/10\\ (55-66+21)/10\end{pmatrix}=\begin{pmatrix}20/10\\ 10/10\\ 10/10\end{pmatrix}=\begin{pmatrix}2\\ 1\\ 1\end{pmatrix}.$$

Звідси, $x=2;y=1;z=1$.

Найбільш універсальним способом розв'язування системи є метод виключення невідомих (метод Гауса).

3.6. Особливості методу Гауса

Найбільш відомим з точних методів розв'язання СЛАР (3.1) є метод Гауса, суть якого 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 час у більшості випадків залежить від алгоритму формування трикутної матриці системи.

У загальному випадку алгоритми Гауса складаються з двох етапів:

1. **Прямий хід**, в результаті якого СЛАР (3.1), що розв'язується, перетворюється в еквівалентну систему з верхньою трикутною матрицею коефіцієнтів виду:

$$\begin{cases} a_{11} \cdot x_1 + a_{12} \cdot x_2 + \ldots + a_{1n} \cdot x_n = b_1; \\ 0 \cdot x_1 \quad + a_{22} \cdot x_2 + \ldots + a_{2n} \cdot x_n = b_2; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ 0 \cdot x_1 \quad + 0 \cdot x_2 \quad + \ldots + a_{nn} \cdot x_n = b_n. \end{cases} \qquad (3.40)$$

2. **Зворотній хід** дозволяє визначити вектор розв'язку починаючи з останнього рівняння системи (3.1) шляхом підстановки координат вектора невідомих, отриманих на попередньому кроці.

Метод Гауса - це просто. Відомий німецький математик Йоганн Карл Фрідріх Гаус ще за життя отримав визнання видатного математика всіх часів, генія і навіть прізвисько «короля математики». Метод Гауса простий тим, що для його освоєння достатньо вміти додавати і множити. Не випадково метод послідовного виключення невідомих викладачі часто розглядають. Але у студентів метод Гауса викликає найбільші труднощі. Вся справа в методиці і в доступній формі розповімо про алгоритм методу.

Спочатку систематизуємо знання про системи лінійних рівнянь. Система лінійних рівнянь може:

1) мати єдиний розв'язок;

2) мати нескінченно багато розв'язків;

3) не мати розв'язків (бути несумісною).

Метод Гауса - найбільш потужний і універсальний інструмент для знаходження розв'язку будь-якої системи лінійних рівнянь. Правило Крамера та матричний метод непридатні в тих випадках, коли система має нескінченно багато розв'язків або несумісна. А метод послідовного виключення невідомих в будь-якому випадку приведе до відповіді. В цьому розділі розглянемо метод Гауса для випадку 1 (єдиний розв'язок системи), для випадкам 2, 3 присвячений наступний розділ: несумісні системи і системи із загальним розв'язком. Зауважимо, що сам алгоритм методу у всіх трьох випадках працює однаково.

Відомо декілька різних алгоритмів отримання еквівалентної системи з верхньою трикутною матрицею. Розглянемо найбільш відомі з них.

3.6.1. Аналіз постановки задачі для методу Гауса

Метод Гауса застосовується для розв'язування системи лінійних алгебраїчних рівнянь (3.1). У матричній формі ця система має вигляд (3.2)-(3.3). Використаємо той факт, що розв'язок системи не змінюється при виконанні кожної з наступних операцій:

а) перестановка двох рівнянь місцями;

б) множення одного з рівнянь на число, яке не дорівнює нулю;

в) віднімання одного рівняння, помноженого на деяке число, від іншого.

Якщо $a_{11}=0$, поміняємо місцями перше рівняння з j-м рівнянням, таким, що $a_{j1}\neq 0$. Тепер коефіцієнт у першому рівнянні при першому невідомому, відмінний від нуля. Позначимо його через a'_{11} і будемо називати *ведучим елементом* першого кроку. Поділимо перше рівняння на ведучий елемент. Потім віднімемо його від k-го рівняння ($k=2, 3, ..., n$) отриманої системи, попередньо помноживши на a'_{k1}. Після таких перетворень перший стовпчик коефіцієнтів рівнянь буде складатися з одиниці на першому місці і нулів на інших місцях.

Розглянемо отримані рівняння з номерами 2, ..., n. Вони утворюють систему з (n-1) рівняннями з (n-1) невідомими. Виконаємо з цією системою ті ж операції, що і з попередньою (другий крок методу Гауса).

Наступний крок виконуємо для останніх (n-2) рівнянь і так далі.

Якщо на кожному кроці вдається вибрати ведучий елемент, то після ряду перетворень система рівнянь набуває трикутного вигляду:

$$\begin{aligned} x_1 + a''_{12}x_2 + a''_{13}x_3 + ... + a''_{1n}x_n &= b''_1; \\ x_2 + a''_{23}x_3 + ... + a''_{2n}x_n &= b''_2; \\ \cdots\cdots\cdots\

3.6.2. Метод Гауса (метод заміни змінних)

Цей метод полягає у послідовному виключенні змінних з кожних наступних рівнянь і приведенні заданої системи до так званого "трикутного вигляду".

Якщо задана система $A\bar{x} = \bar{b}$ (3.32), то внаслідок того, що над вектором вільних членів здійснюються такі ж алгоритмічні перетворення, як і над коефіцієнтами матриці A, її зводять до допоміжної матриці:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ \text{----------} \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases} \tag{3.42}$$

$$\tilde{A} = \begin{pmatrix} a_{11} \ldots a_{1n} b_1 \\ \cdots \\ a_{n1} \ldots a_{nn} b_n \end{pmatrix} \tag{3.43}$$

Метод Гауса поділяється на два етапи – *прямий хід і зворотний хід.* В результаті *прямого ходу* система (3.32) зводиться до наступної системи $\hat{A}\bar{x} = \hat{\bar{b}}$. Прямий хід складається з n кроків. На першому кроці, припускаючи $a_{11}^{(0)} \neq 0$, перше рівняння системи (3.32) ділять на елемент a_{11}. В результаті отримаємо зведене рівняння:

$$x_1 + a_{12}^{(1)} \cdot x_2 + \ldots + a_{1n}^{(1)} \cdot x_n = a_{1,n+1}^{(1)} \tag{3.44}$$

Зауважимо, що верхній індекс визначатиме номер кроку. Стовпчик вільних членів позначається як a_{n+1} стовпчик. Якщо через $A^{(0)} = \tilde{A}$, то коефіцієнт a_{1j} на першому кроці:

$$a_{1j}^{(1)} = \frac{a_{1j}^{(0)}}{a_{11}^{(0)}} \qquad \mathrm{j} = \overline{2, n+1} \tag{3.45}$$

Для вилучення змінної x_1 у всіх інших рівняннях, зведене рівняння (3.44) множать на коефіцієнт a_{i1}, після чого його віднімають від i–го рівняння. В результаті отримаємо матрицю:

$$A^{(1)} = \begin{pmatrix} 1\, a_{12}^{(3)} \ldots a_{1,n+1} \\ 0\, a_{22}^{(1)} \ldots a_{2,n+1}^{(2)} \\ \cdots \\ 0\, a_{n2}^{(2)} \ldots a_{n,n+1} \end{pmatrix} \tag{3.46}$$

$$a_{ij}^{(1)} = a_{ij}^{(0)} - a_{i1}^{(0)} \cdot \frac{a_{1j}^{(0)}}{a_{11}^{(0)}} = a_{ij}^{(0)} - a_{i1}^{(0)} \cdot a_{1j}^{(1)}, \quad \begin{matrix} i = \overline{2, n} \\ j = \overline{2, n+1} \end{matrix} \tag{3.47}$$

На другому кроці залишають незмінними перший стовпчик і першу стрічку $A^{(1)}$. Аналогічні перетворення проводять зі стрічками починаючи з 2-ї і закінчуючи n-ю при умові, що $a_{22} \neq 0$. На k-ому кроці розрахункові формули (3.45) і (3.47) будуть мати вигляд:

$$a_{kk}^{(k-1)} \neq 0, \quad a_{kj}^{(k)} = \frac{a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}}, \quad \begin{matrix} j = \overline{k+1, n+1} \\ k = \overline{1, n} \end{matrix} \tag{3.48}$$

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - a_{ik}^{(k-1)} \frac{a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}} = a_{ij}^{(k-1)} - a_{ik}^{(k-1)} \cdot a_{kj}^{(k)}, \quad \begin{matrix} i = \overline{k+1, n} \\ j = \overline{k+1, n+1} \\ k = \overline{1, n-1} \end{matrix} \tag{3.49}$$

Зауважимо, що формули (3.48) і (3.49) є загальними для будь-якого кроку. Проте, якщо формула (3.48) працює для всіх кроків, то формула (3.49) працює лише для випадків $k = \overline{1, n-1}$.

В результаті *прямого ходу методу Гауса* отримали систему лінійних рівнянь верхньої трикутної форми, в якій елементи головної діагоналі рівні одиниці, а піддіагональні елементи – нулі.

$$A^{(n)} = \begin{pmatrix} 1 & a_{12} & a_{13} & \dots & a_{1n} & a_{1,n+1} \\ 0 & 1 & a_{23} & \dots & a_{2n} & a_{2,n+1} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & a_{n,n+1} \end{pmatrix}. \tag{3.50}$$

Зворотний хід методу Гауса полягає в тому, що $x_n = a_{n,n+1}^{(n)}$ одразу ж прирівнюється останній координаті вільних членів. Це означає, що на останньому кроці *прямого ходу* методу Гауса нами була отримана остання координата вектора невідомих. Останній крок прямого ходу методу Гауса є першим кроком зворотного ходу цього методу.

$$x_n = a_{n,n+1}^{(n)} \tag{3.51}$$

$$x_{n-1} + a_{n-1,n}^{(n-1)} \cdot x_n = a_{n-1,n+1}^{(n-1)} \tag{3.52}$$

$$x_{n-1} = a_{n-1,n+1}^{(n-1)} - a_{n-1,n}^{(n-1)} \cdot x_n \tag{3. 53}$$

$$x_k = a_{k,n+1}^{(k)} - \sum_{j=k+1}^{n} a_{kj}^{(k)} \cdot x_j, \quad k = \overline{n,1} \tag{3.54}$$

Формула (3.54) є загальним виглядом формул (3.51)-(3.53). Цю формулу у програмуванні можна приймати як загально прийняту умову: якщо нижній індекс суми перевищує верхній, то значення суми вважається рівним нулю. Для реалізації методу Гауса розрахунковими є формули (3.48), (3.49) і (3.54). При реалізації

прямого ходу зроблено припущення, що діагональний елемент a_{kk} був відмінний від нуля. Якщо ж цей елемент дорівнює 0, то в цьому ж k-му стовпчику, починаючи з $k+1$-го елемента, шукається відмінний від нуля елемент. Якщо такий елемент є і має номер l ($a_{lk} \neq 0$), то здійснюється перестановка коефіцієнтів k-ї і l-ї стрічки. Як відомо, перестановка рівнянь системи місцями розв'язку не змінює.

Якщо ж всі елементи k-го стовпчика вийдуть рівними 0, то це свідчить про несумісність системи. В цьому випадку процес зупиняють і генерують відповідне повідомлення. Досліджено, що для реалізації методу Гауса потрібно виконати приблизно $\frac{2}{3}n^3$ арифметичних операцій. *Висновок*: для систем розмірності більше 1000 метод Гауса застосовувати не слід, тому що похибка заокруглень (у випадку її невдалого накопичення) може бути занадто великою.

3.6.3. Частковий випадок застосування методу 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: множення 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:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases} \tag{3.55}$$

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:

$$M_2 = \frac{a_{21}}{a_{11}}. \tag{3.56}$$

3) Перше рівняння системи (3.55) множиться на M_2 і віднімається від другого рівняння системи, яка отримана після перестановки рівнянь (якщо перестановка була необхідна). Результат обчислення має вигляд:

$$(a_{21} - M_2 a_{11})x_1 + (a_{22} - M_2 a_{12})x_2 + (a_{23} - M_2 a_{13})x_3 = b_2 - M_2 b_1, \tag{3.57}$$

$$a_{21} - M_2 a_{11} = a_{21} - \left(\frac{a_{21}}{a_{11}}\right) a_{11} = 0. \tag{3.58}$$

Тоді x_1 виключається із другого рівняння. Позначимо нові коефіцієнти:

$$a'_{22} = a_{22} - M_2 a_{12};\ a'_{23} = a_{23} - M_2 a_{13};\ b'_2 = b_2 - M_2 b_1. \tag{3.59}$$

Тоді друге рівняння системи (3.55) набуває вигляду:

$$a'_{22} x_2 + a'_{23} x_3 = b'_2. \tag{3.60}$$

Далі необхідно звільнитися від коефіцієнта a_{31} при x_1 в третьому рівнянні системи (3.55) за аналогічним алгоритмом.

4) Обчислюється множник для третього рівняння:

$$M_3 = \frac{a_{31}}{a_{11}}. \tag{3.61}$$

5) Перше рівняння системи (3.55) множиться на M_3 і віднімається від третього рівняння. Коефіцієнт при x_1 стає нулем, і третє рівняння набуває вигляду:

$$a'_{32} x_2 + a'_{33} x_3 = b'_3, \tag{3.62}$$

$$a'_{32} = a_{32} - M_3 a_{12}, \tag{3.63}$$

$$a'_{33} = a_{33} - M_3 a_{13}, \tag{3.64}$$

$$b'_3 = b_3 - M_3 b_1. \tag{3.65}$$

Перетворена таким чином система рівнянь (3.55) набуває вигляду:

$$\begin{cases} a_{11} x_1 + a_{12} x_2 + a_{13} x_3 = b_1 \\ 0 \cdot x_1 + a'_{22} x_2 + a'_{23} x_3 = b'_2 \\ 0 \cdot x_1 + a'_{32} x_2 + a'_{33} x_3 = b'_3 \end{cases} \tag{3.66}$$

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.

7) Обчислюємо множник

$$M'_3 = \frac{a_{32}}{a_{22}}. \tag{3.67}$$

8) Друге рівняння системи (3.63) множиться на M'_3 і віднімається від 3-го рівняння:

$$\left(a'_{32} - M'_{33}a'_{22}\right)x_2 + \left(a'_{33} - M'_3 a'_{23}\right)x_3 = b'_3 - M'_3 b'_2. \tag{3.68}$$

При цьому коефіцієнт біля x_2 дорівнює нулю:

$$a'_{32} - M'_{33}a'_{22} = 0, \tag{3.69}$$

$$a''_{33} = a'_{33} - M'_3 a'_{23}, \tag{3.70}$$

$$b''_3 = b'_3 - M'_{32} b'_2. \tag{3.71}$$

$$a''_{33}x_3 = b''_3. \tag{3.72}$$

Замінивши в системі (3.66) третє рівняння на (3.72), отримаємо систему рівнянь виду:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ 0 \cdot x_1 + a'_{22}x_2 + a'_{23}x_3 = b'_2 \\ 0 \cdot x_1 + 0 \cdot x_2 + a''_{33}x_3 = b''_3 \end{cases} \tag{3.73}$$

Таку систему називають ***системою з трикутною матрицею коефіцієнтів,*** що еквівалентна СЛАР (3.55). 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-ше рівняння системи (3.73) і знайти x_1 за формулами:

$$x_3 = \frac{b''_3}{a''_{33}}, \tag{3.74}$$

$$x_2 = \frac{b'_2 - a'_{23}x_3}{a'_{22}}, \tag{3.75}$$

$$x_1 = \frac{b_1 - a_{12}x_2 - a_{13}x_3}{a_{11}}. \tag{3.76}$$

Процес знаходження вектора розв'язку системи (3.55) називають ***зворотним ходом метода Гауса***. На рис. 3.2а показана схема алгоритму метода Гауса з послідовним виключенням для розв'язання системи із N рівнянь з N невідомими.

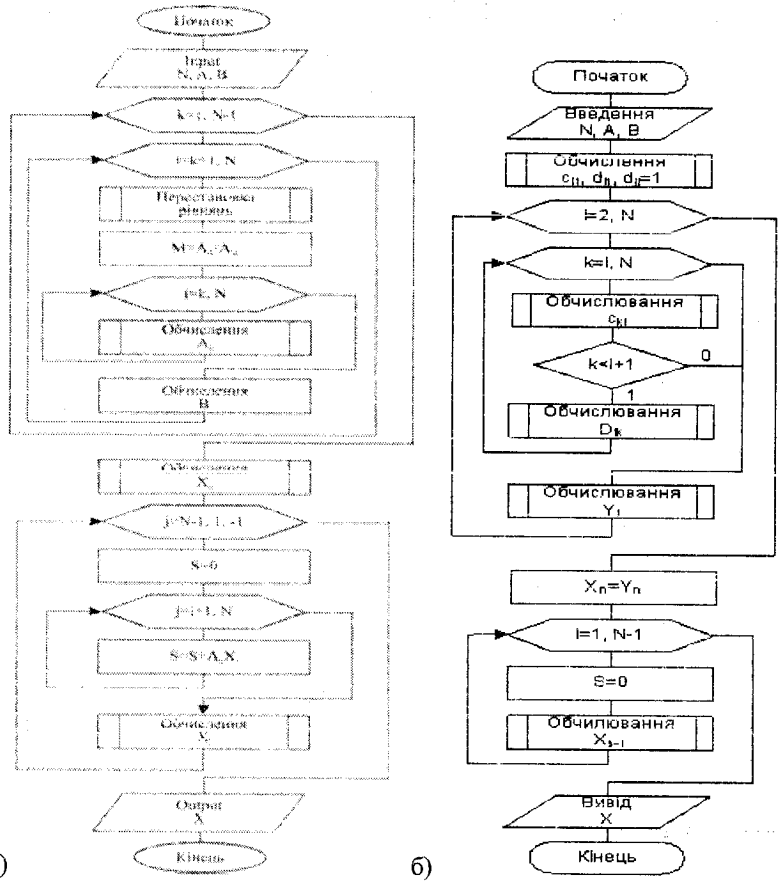


Рис. 3.2. Схема алгоритму розв'язування системи лінійних алгебраїчних рівнянь методом Гауса: а) з послідовним виключенням; б) за схемою Халецького.

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". Якщо прямувати до можливого зменшення помилок округлення, то можна використати алгоритм з вибором головного елемента. Призначення індексів в схемі алгоритму (рис. 3.2а): k – номер рівняння, яке віднімається від інших, а також номер невідомого, яке виключається із k рівнянь, які залишились; i – номер рівняння, з якого в даний момент виключається невідоме; j – номер стовпця.

Приклад 3.13. Розв'язати систему за допомогою методу Гауса (приклад з підрозділу про метод Крамера).

$$\begin{cases} x + 2y + z = 5, \\ 3x + 4y + z = 11, \\ 7x + 2y + 5z = 21. \end{cases}$$

Розв'язок. Запишемо систему в матричні формі:

$$AX = B, \text{ де } A = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 4 & 1 \\ 7 & 2 & 5 \end{pmatrix}, X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, B = \begin{pmatrix} 5 \\ 11 \\ 21 \end{pmatrix}.$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 5 \\ 3 & 4 & 1 & 11 \\ 7 & 2 & 5 & 21 \end{array}\right) \xrightarrow{(1)} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 5 \\ 0 & -2 & -2 & -4 \\ 0 & -12 & -2 & -14 \end{array}\right) \xrightarrow{(2)} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 5 \\ 0 & -2 & -2 & -4 \\ 0 & 0 & 10 & 10 \end{array}\right) \xrightarrow{(3)} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 5 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

(1) Перший рядок множимо на -3 та додаємо до другого рядка. Початковий перший рядок множимо на -7 та додаємо до третього рядка. Перший рядок переписуємо без змін, а наступні два рядка – зі змінами.

(2) Другий рядок множимо на -6 та додаємо до третього отриманий результат. Перший і другий рядки переписуємо без змін, а третій – зі змінами..

(3) Третій рядок ділимо на 10, другий рядок ділимо на -2, а перший переписуємо без змін. Тоді $z = 1;\ y + z = 2 \Rightarrow y = 1;\ x + 2y + z = 5 \Rightarrow x = 2$.

3.6.4. Метод Гауса за схемою Халецького (метод LU факторизації)

В літературі цей метод ще називається *схема Халецького* або *метод LU факторизації*. Алгоритм методу включає також прямий і зворотній хід. Кінцевою метою прямого ходу є отримання СЛАР, яка еквівалентна заданій, з верхньою трикутною матрицею коефіцієнтів. Суть цього методу ґрунтується на наступній теоремі алгебри. Для цього матрицю коефіцієнтів початкової системи рівнянь A або інакше довільну квадратну дійсну матрицю розбивають на дві трикутні:

$$A = L \cdot U, \tag{3.77}$$

де матриця L – нижня трикутна матриця; U – верхня трикутна матриця з одиничною головною діагоналлю при $l_{ij} = 0\ (i < j)$, $u_{ij} = 0\ (i > j)$ та $u_{ii} = 0$:

$$L = \begin{pmatrix} l_{11} & 0 & 0 & \dots & 0 \\ l_{21} & l_{22} & 0 & \dots & 0 \\ l_{31} & l_{32} & l_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ l_{n1} & l_{n2} & l_{n3} & \dots & l_{nn} \end{pmatrix}, U = \begin{pmatrix} 1 & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & 1 & u_{23} & \dots & u_{2n} \\ 0 & 0 & 1 & \dots & u_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}. \tag{3.78}$$

Проте такий розклад передбачає безліч розв'язків. Тому, щоб виділити з цих розв'язків один, необхідно в одній із матриць L та U зафіксувати будь-яку стрічку, стовпчик чи діагональ. Таким чином, необхідно визначити ненульові елементи цих матриць. Для цього скористаємось правилом добутку двох матриць:

$$a_{ij} = \sum_{k=1}^{n} l_{ik} u_{kj} \qquad i, j = \overline{1, n} \tag{3.79}$$

при $l_{ij}=0,\ \forall i<j;\ u_{ij}=0,\ \forall i>j,\ u_{ii}=1$.

Розпишемо суму у формулі (3.79) на три складові частини наступним чином:

$$a_{ij}=\sum_{k=1}^{j-1}l_{ik}u_{kj}+l_{ij}u_{jj}+\sum_{k=j+1}^{n}l_{ik}u_{kj} \tag{3.80}$$

$$l_{ij}=a_{ij}-\sum_{k=1}^{j-1}l_{ik}u_{kj},\ i\geq j,\ l_{i1}=a_{i1},\ \ i=\overline{1,n} \tag{3.81}$$

Запишемо формулу (3.79) в іншому вигляді:

$$a_{ij}=\sum_{k=1}^{i-1}l_{ik}u_{kj}+l_{ii}u_{ij}+\sum_{k=i+1}^{n}l_{ik}u_{kj},\quad \begin{matrix} i=\overline{2,n} \\ j=\overline{2,n+1} \end{matrix} \tag{3.82}$$

$$u_{ij}=\left(a_{ij}-\sum_{k=1}^{i-1}l_{ik}u_{kj}\right)/l_{ii},\ \ u_{1j}=a_{1j}/a_{11},\ \ j=\overline{2,n} \tag{3.83}$$

Отже, за допомогою формул (3.81) і (3.83), в залежності від співвідношення між індексами i та j, можна почергово обчислювати елементи матриці U по стовпчиках.

Запишемо систему лінійних рівнянь (3.77) у вигляді

$$L\underbrace{U\bar{x}}_{\bar{y}}=\bar{b} \tag{3.84}$$

Позначимо $U\bar{x}=\bar{y}$. Тоді, враховуючи, що матриця L – нижня трикутна,

$$L\bar{y}=\bar{b} \tag{3.85}$$

$$y_i=(b_i-\sum_{k=1}^{i-1}l_{ij}y_k)/l_{ij}\quad i=\overline{1,n}. \tag{3.86}$$

$$\begin{pmatrix} l_{11} & 0 & 0 & \dots & 0 \\ l_{21} & l_{22} & 0 & \dots & 0 \\ l_{31} & l_{32} & l_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ l_{n1} & l_{n2} & l_{n3} & \dots & l_{nn} \end{pmatrix}\cdot\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_n \end{pmatrix}=\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \dots \\ b_n \end{pmatrix}. \tag{3.87}$$

Враховуючи введене позначення $U\bar{x}=\bar{y}$, і те, що матриця U є верхньою трикутною, то можна отримати координати шуканого вектор-розв'язку, скориставшись методом Гауса у зворотному порядку

$$\begin{pmatrix} 1 & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & 1 & u_{23} & \dots & u_{2n} \\ 0 & 0 & 1 & \dots & u_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}\cdot\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{pmatrix}=\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \dots \\ y_n \end{pmatrix}. \tag{3.88}$$

$$x_i = y_i - \sum_{k=i+1}^{n} u_{ik} x_k \tag{3.89}$$

при $x_n = y_n$ та $i = \overline{n-1,1}$

Існує певний зв'язок методу LU-факторизації з методом Гауса, так як в цьому методі прямий хід реалізується через обчислення значень l_{ij}, u_{ij}, y_i. При цьому слід зауважити, що y_i можна обчислювати одночасно з l_{ij} та u_{ij}.

Обчислення x_i представляє собою зворотній хід методу LU-факторизації.

Цей метод зручно використовувати для обчислення детермінант

$$\det A = \det LU = \det L \cdot \det U = \det L = \prod_{i=1}^{n} l_{ii} \tag{3.90}$$

$$\det A \neq 0 \Leftrightarrow \forall i : l_{ii} \neq 0 \tag{3.91}$$

Якщо компоненти l_{ij} є відмінні від нуля, то можна зробити висновок про сумісність чи несумісність системи, що не вимагає доповнюючої перестановки стрічок.

Алгоритм визначення коефіцієнтів матриць L і U.

1) Обчислюється перший стовпець матриці L, перший рядок матриці U та y_1 за формулами:

$$l_{i1} = a_{i1}, \quad i = \overline{1,n}; \; u_{11} = 1, \quad u_{1i} = \frac{a_{1i}}{a_{11}}, \quad i = \overline{2,n}. \tag{3.92}$$

$$y_1 = \frac{b_1}{a_{11}}. \tag{3.93}$$

2) Обчислюються елементи другого стовпця матриці L:

$$l_{i2} = a_{i2} - a_{i1} u_{12}, \quad i = \overline{2,n}, \tag{3.94}$$

елементи другого рядка матриці U та елемент y_2:

$$u_{21} = \frac{a_{2i} - l_{2i} u_{1i}}{a_{22}}, \tag{3.95}$$

$$y_2 = \frac{b_2 - l_{21} y_1}{a_{22}}. \tag{3.96}$$

3) Обчислюють елементи третього стовпця матриці L:

$$l_{i3} = a_{i3} - \left(l_{i1} u_{31} - l_{i2} u_{23}\right), \quad i = \overline{2,n}, \tag{3.97}$$

елементи третього рядка матриці D та елемент y_3:

$$u_{31} = \frac{a_{31} - \left(l_{31} u_{1i} + l_{32} u_{2i}\right)}{l_{33}}, \tag{3.98}$$

$$y_3 = \frac{b_3 - \left(l_{31} y_1 + l_{32} y_2\right)}{l_{33}} \tag{3.99}$$

і т.д.

Схема алгоритму метода Гауса за схемою Халецького показана на рисунку 3.2б. Загальний вигляд формул для обчислення l_{ki}, u_{ki}, y_i елементів матриць L, U і Y:

$$u_{ii} = 1;\ u_{1i} = \frac{a_{1i}}{a_{11}};\ l_{i1} = a_{i1};\ i = \overline{1,n};\ y_1 = \frac{b_1}{a_{11}}. \tag{3.100}$$

$$l_{ki} = a_{ki} - \sum_{j-1}^{i} l_{kj} u_{ji}, \quad k = \overline{i,n}, \quad i = \overline{2,n}. \tag{3.101}$$

$$u_{ik} = \frac{a_{ik} - \sum_{j=1}^{i} l_{ij} u_{jk}}{l_{ii}}, \quad k = \overline{i+1,n}. \tag{3.102}$$

$$y_i = \frac{b_i - \sum_{j=1}^{i} l_{ij} y_j}{l_{ii}}. \tag{3.103}$$

Метод LU-факторизації володіє кращою стійкістю порівняно з методом Гауса по відношенню до нагромадження похибок обчислень, а саме, виконує $n^3/3$ арифметичних операцій. Виходячи з цього, метод Гауса доцільно використовувати лише для систем розмірності не більше 100, так як в іншому випадку похибка заокруглення може бути непомірно великою, що взагалі призводить до неправильного результату.

3.6.5. Метод Гауса з вибором головного елемента

Ідея цього методу виникла у зв'язку з тим, що коефіцієнти СЛАР є параметрами реальних інженерних систем та в більшості є наближеними значеннями, тому що отримані в результаті вимірювання або як статистичні дані. Для таких систем рівнянь при обчисленні масштабного множника

$$M = \frac{a_{ik}}{a_{kk}} \tag{3.104}$$

можлива ситуація при визначені a_{kk}, що ділення наближеного числа a_{ik} на достатньо мале число a_{kk} веде до різкого збільшення похибки методу. Тому для того, щоб не збільшувати похибку результату необхідно виконувати такі дії:

1) в системі (3.34) необхідно знайти в k-му стовпці найбільший за абсолютним значенням коефіцієнт a_{kj};

2) переставити k-те рівняння з рівнянням, у якому знаходиться цей максимальний коефіцієнт;

3) масштабний множник буде обчислюватись за формулою (3.104), де a_{kk} – максимальний коефіцієнт, а тому похибка розв'язування СЛАР у результаті арифметичних операцій не збільшується.

Схема алгоритму метода Гауса з вибором головного елемента (прямий та обернений хід) показана на рис. 3.3а.

3.6.6. Метод Гауса з одиничними коефіцієнтами

В цьому методі зроблена спроба зменшити недоліки перших двох методів, які пов'язані з багаторазовим діленням одного наближеного числа на інше. Для цього перед введенням масштабного множника k-те рівняння системи ділиться один раз на діагональний елемент a_{kk} так, щоб коефіцієнт при $x_k = 1$, а масштабний множник M_i буде дорівнювати a_{kj}. Результатом прямого ходу є система, яка еквівалентна СЛАР (3.34) з одиничними коефіцієнтами на головній діагоналі виду:

$$\begin{cases} 1 \cdot x_1 + a_{12} \cdot x_2 + \ldots + a_{1n} \cdot x_n = b_1; \\ 0 \cdot x_1 + 1 \cdot x_2 + \ldots + a_{2n} \cdot x_n = b_2; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ 0 \cdot x_1 + 0 \cdot x_2 + \ldots + 1 \cdot x_n = b_n. \end{cases} \qquad (3.105)$$

Ця система схожа на систему (3.40), яка отримується в результаті прямого ходу базового методу Гауса з послідовним вилученням невідомих і відрізняється від неї тільки діагональними коефіцієнтами. Для отримання такої системи необхідно використовувати алгоритм, який включає в себе наступні етапи:

1. Організація циклу по всіх рівняннях від 1 до N-1 (k = 1, 2, …, N-1).

2. В кожному k-му стовпці визначається номер l-го рівняння з головним елементом (тобто номер l-го рівняння, в якому знаходиться коефіцієнт при x_n зі всіх рівнянь, починаючи з k-го до N-го).

3. Якщо номер цього рівняння l не дорівнює k ($l<>k$), тоді необхідно переставити місцями l-е рівняння з k-м.

4. Нормування k-го рівняння, тобто ділення всіх коефіцієнтів k-го рівняння на a_{kk} (головний елемент при x_n), включаючи b_n.

5. Перетворення всіх i-х рівнянь, починаючи з (k+1) до N, у відповідності з базовим алгоритмом Гауса з метою отримання еквівалентної системи з верхньою трикутною матрицею коефіцієнтів.

6. Кінець циклу по k.

Формула зворотного ходу для систем виду (3.105) спрощується і має вигляд:

$$\begin{aligned} x_n &= b_n, \\ &\ldots\ldots\ldots\ldots \\ x_2 &= b_2 - a_{2n}x_n, \\ x_1 &= b_1 - a_{1n}x_n - a_{12}x_2. \end{aligned} \qquad (3.106)$$

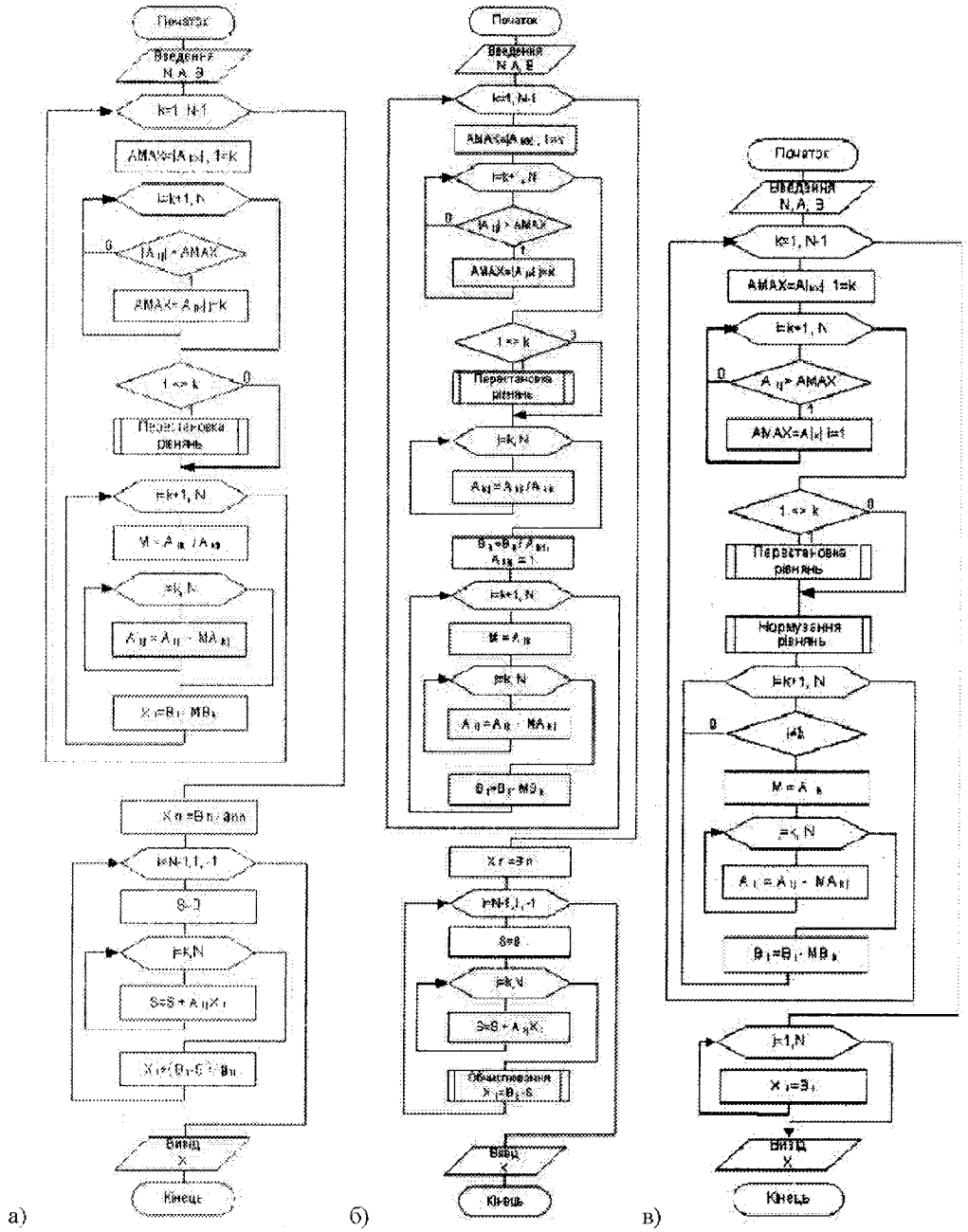


Рис. 3.3. Схема алгоритму метода Гауса: а)з вибором головного елемента, б) з одиничними коефіцієнтами; в) Гауса-Жордана.

Схема алгоритму методу Гауса з одиничними діагональними коефіцієнтами наведена на рисунку 3.3б.

Приклад 3.14. Розв'язати систему за допомогою методу Гауса з одиничними коефіцієнтами.

$$\begin{cases} 9x+3y+z=13, \\ 4x+2y+z=7, \\ x+y+z=3. \end{cases}$$

Розв'язок. Запишемо систему в матричні формі:

$$AX=B, \text{ де } A=\begin{pmatrix} 9 & 3 & 1 \\ 4 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}, X=\begin{pmatrix} x \\ y \\ z \end{pmatrix}, B=\begin{pmatrix} 13 \\ 7 \\ 3 \end{pmatrix}.$$

$$\left(\begin{array}{ccc|c} 9 & 3 & 1 & 13 \\ 4 & 2 & 1 & 7 \\ 1 & 1 & 1 & 3 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{ccc|c} -36 & -12 & -4 & -52 \\ 36 & 18 & 9 & 63 \\ 1 & 1 & 1 & 3 \end{array}\right) \overset{(2)}{\to} \left(\begin{array}{ccc|c} -36 & -12 & -4 & -52 \\ 0 & 6 & 5 & 11 \\ 1 & 1 & 1 & 3 \end{array}\right) \overset{(3)}{\to}$$

$$\overset{(3)}{\to} \left(\begin{array}{ccc|c} -36 & -12 & -4 & -52 \\ 0 & 6 & 5 & 11 \\ 0 & 24 & 32 & 56 \end{array}\right) \overset{(4)}{\to} \left(\begin{array}{ccc|c} -36 & -12 & -4 & -52 \\ 0 & -24 & -20 & -44 \\ 0 & 0 & 12 & 12 \end{array}\right).$$

(1) Перший рядок множимо на -4, а другий на 9.

(2) До другого рядка додаємо перший рядок.

(3) Перший рядок множимо на 1, а третій на 36. До третього рядка додаємо перший рядок.

(4) Другий рядок множимо на -4, а третій на 1. До третього рядка додаємо другий рядок.

Тоді $12z=12 \Rightarrow z=1$;

$-24y+20z=-44 \Rightarrow -24y+20=-44 \Rightarrow -24y=-24 \Rightarrow y=1$;

$-36x-12y-4z=-52 \Rightarrow -36x-12-4=-52 \Rightarrow -36x=-36 \Rightarrow x=1$.

3.6.7. Метод Гауса-Жордана

Припустимо, що дано систему з трьома рівняннями, трьома невідомими і записана її розширена матриця

$$\left(\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{array}\right).$$

У найбільш поширеному випадку виходять стандартні ступені вигляду:

$$\left(\begin{array}{ccc|c} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \end{array}\right) \text{ або } \left(\begin{array}{ccc|c} * & 0 & 0 & * \\ * & * & 0 & * \\ * & * & * & * \end{array}\right),$$

причому останній абсолютно рівноцінний і може бути незручним тільки через суб'єктивне сприйняття. Жордан запропонував розв'язувати задачу такого типу без зі зворотного ходу гаусівського алгоритму:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{array}\right).$$

Метод Гауса-Жордана є модифікацією методу Гауса.

Особливістю метода Гауса-Жордана є перетворення системи (3.34) (прямий хід) до еквівалентної з одиничною матрицею коефіцієнтів виду:

$$\begin{cases} 1 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = b_1; \\ 0 \cdot x_1 + 1 \cdot x_2 + \ldots + 0 \cdot x_n = b_2; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ 0 \cdot x_1 + 0 \cdot x_2 + \ldots + 1 \cdot x_n = b_n. \end{cases} \qquad (3.107)$$

тобто системи, яка містить тільки одиничну діагональ.

Для отримання такої системи до прямого ходу алгоритму базового методу Гауса (з послідовним виключенням невідомих) додатково вводяться наступні дії:

1. Організація циклу по k по всіх рівняннях від 1 до N-1 (k = 1, 2, …, N-1).

2. Процедура вибору головного елементу в кожному k-му стовпці при x_k.

3. Процедура нормування k-го рівняння системи, тобто в k-му рівнянні кожен коефіцієнт a_{kj} розділити на a_{kk}, включаючи b_k, так, щоб коефіцієнт a_{kk}=1.

4. Перетворення всіх рівнянь системи, починаючи з 1-го до N у відповідності з базовим алгоритмом Гауса з метою отримати еквівалентну систему з одиничною діагоналлю. В цьому випадку для розрахунку коефіцієнтів a_{ij} використовуються ті самі формули, що і в базовому алгоритмі Гауса:

$$M = a_{ik}; \quad a_{ij} = a_{ij} - M \cdot a_{kj}; \quad b_i = b_i - M \cdot b_k, \qquad (3.108)$$

але використовуються вони для всіх рівнянь з 1-го до N крім k-го, в якому залишається коефіцієнт a_{kk} рівний одиниці.

5. Кінець циклу по k.

Зворотний хід методу Гауса-Жордана дуже простий і використовує наступні формули: $x_k = b_k$, при $k = \overline{1,n}$. Схема алгоритму подана на рисунку 3.3в.

Приклад 3.15. Розв'язати систему за допомогою методу Гауса-Жордана.

$$\begin{cases} 9x + 3y + z = 13, \\ 4x + 2y + z = 7, \\ x + y + z = 3. \end{cases}$$

Розв'язок. Запишемо систему в матричні формі:

$$AX = B, \text{ де } A = \begin{pmatrix} 9 & 3 & 1 \\ 4 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}, X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, B = \begin{pmatrix} 13 \\ 7 \\ 3 \end{pmatrix}.$$

$$\left(\begin{array}{ccc|c} 9 & 3 & 1 & 13 \\ 4 & 2 & 1 & 7 \\ 1 & 1 & 1 & 3 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{ccc|c} -36 & -12 & -4 & -52 \\ 36 & 18 & 9 & 63 \\ 1 & 1 & 1 & 3 \end{array}\right) \overset{(2)}{\to} \left(\begin{array}{ccc|c} -36 & -12 & -4 & -52 \\ 0 & 6 & 5 & 11 \\ 1 & 1 & 1 & 3 \end{array}\right) \overset{(3)}{\to}$$

$$\overset{(3)}{\to} \left(\begin{array}{ccc|c} -36 & -12 & -4 & -52 \\ 0 & 6 & 5 & 11 \\ 0 & 24 & 32 & 56 \end{array}\right) \overset{(4)}{\to} \left(\begin{array}{ccc|c} -36 & -12 & -4 & -52 \\ 0 & -24 & -20 & -44 \\ 0 & 0 & 12 & 12 \end{array}\right) \overset{(5)}{\to}$$

$$\overset{(5)}{\to} \left(\begin{array}{ccc|c} -36 & -12 & -4 & -52 \\ 0 & -24 & -20 & -44 \\ 0 & 0 & 1 & 1 \end{array}\right) \overset{(6)}{\to} \left(\begin{array}{ccc|c} 1 & 12/36 & 4/36 & 52/36 \\ 0 & 1 & 20/24 & 44/24 \\ 0 & 0 & 1 & 1 \end{array}\right) \overset{(7)}{\to}$$

$$\overset{(7)}{\to} \left(\begin{array}{ccc|c} 1 & 12/36 & 4/36 & 52/36 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right) \overset{(8)}{\to} \left(\begin{array}{ccc|c} 1 & 0 & 4/36 & 40/36 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right) \overset{(9)}{\to} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

(1) Перший рядок множимо на -4, а другий – на 9.

(2) До другого рядка додаємо перший рядок.

(3) Перший рядок множимо на 1, а третій на 36. До третього рядка додаємо перший рядок.

(4) Другий рядок множимо на -4, а третій – на 1. До третього рядка додаємо другий рядок.

(5) Третій рядок ділимо на 12.

(6)Перший рядок ділимо на 36, а другий – на 24 (залишаємо дробами)

(7) Згідно формули (3.95) $M = a_{23} = 20/24$; $a_{2j} = a_{2j} - M \cdot a_{3j}$; $b_2 = b_2 - M \cdot b_3$.

(8) Згідно формули (3.95) $M = a_{12} = 12/36$; $a_{1j} = a_{1j} - M \cdot a_{2j}$; $b_1 = b_1 - M \cdot b_2$.

(9) Згідно формули (3.95) $M = a_{13} = 4/36$; $a_{1j} = a_{1j} - M \cdot a_{3j}$; $b_1 = b_1 - M \cdot b_3$.

Тоді $z = 1$; $y = 1$; $x = 1$.

3.6.8. Приклади розв'язування системи методом Гауса

Матриця системи – це матриця, яка складається тільки з коефіцієнтів при невідомих, в даному прикладі СЛАР:

$$\begin{cases} x - y = -5, \\ 2x + y = -7, \end{cases}$$

матриця системи – це

$$\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix}.$$

Розширена матриця системи – це та ж матриця системи та стовпець вільних членів, в даному випадку:

$$\left(\begin{array}{cc|c} 1 & -1 & -5 \\ 2 & 1 & -7 \end{array}\right).$$

Будь-яку з матриць для стислості називають просто матрицею. Після запису розширеної матриці системи необхідно виконати деякі дії, які називають елементарними перетвореннями. Існують наступні елементарні перетворення:

1) Рядки матриці можна переставляти місцями. Наприклад, в цій матриці можна переставити перший і другий рядки:

$$\left(\begin{array}{cc|c} 1 & -1 & -5 \\ 2 & 1 & -7 \end{array}\right) \to \left(\begin{array}{cc|c} 2 & 1 & -7 \\ 1 & -1 & -5 \end{array}\right).$$

2) Якщо в матриці є (або з'явилися) пропорційні (як окремий випадок – однакові) рядки, то слід видалити з матриці всі ці рядки крім одного. Розглянемо, наприклад матрицю

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & -1 & 2 & 3 \\ 0 & 1 & -1 & 2 & 3 \\ 0 & 2 & -2 & 4 & 6 \end{array}\right).$$

У цій матриці останні три рядки пропорційні, тому досить залишити тільки один із них:

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & -1 & 2 & 3 \\ 0 & 1 & -1 & 2 & 3 \\ 0 & 2 & -2 & 4 & 6 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & -1 & 2 & 3 \end{array}\right).$$

3) Якщо в матриці в ході перетворень з'явився нульова рядок, то його також необхідно знищити. Нульовий рядок – це рядок, в якому тільки нулі.

4) Рядок матриці можна помножити (поділити) на будь-яке число, відмінне від нуля. Розглянемо, наприклад, матрицю

$$\left(\begin{array}{cc|c} -3 & 9 & 15 \\ 0{,}5 & 0 & 2{,}5 \end{array}\right).$$

Тут доцільно перший рядок розділити на -3, а другий рядок помножити на 2:

$$\left(\begin{array}{cc|c} -3 & 9 & 15 \\ 0{,}5 & 0 & 2{,}5 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & -3 & -5 \\ 1 & 0 & 5 \end{array}\right).$$

Ця дія дуже корисна, оскільки спрощує наступні перетворення матриці.

5) Це перетворення викликає найбільші труднощі, але насправді – нічого складного немає. До рядка матриці можна додати інший рядок, помножений на число, відмінне від нуля. Розглянемо матрицю з практичного прикладу:

$$\left(\begin{array}{cc|c} 1 & -1 & -5 \\ 2 & 1 & -7 \end{array}\right).$$

Спочатку розпишемо перетворення докладно. Множимо перший рядок на –2:

$$\left(\begin{array}{cc|c} 1 & -1 & -5 \\ 2 & 1 & -7 \end{array}\right) \to \left(\begin{array}{cc|c} -2 & 2 & 10 \\ 2 & 1 & -7 \end{array}\right),$$

і до другого рядка додаємо перший рядок помножений на –2:

$$\left(\begin{array}{cc|c} 1 & -1 & -5 \\ 2 & 1 & -7 \end{array}\right) \to \left(\begin{array}{cc|c} -2 & 2 & 10 \\ 2 & 1 & -7 \end{array}\right) \to \left(\begin{array}{cc|c} -2 & 2 & 10 \\ 0 & 3 & 3 \end{array}\right).$$

Тепер перший рядок можна розділити «знову» на –2:

$$\left(\begin{array}{cc|c} 1 & -1 & -5 \\ 2 & 1 & -7 \end{array}\right) \to \left(\begin{array}{cc|c} -2 & 2 & 10 \\ 2 & 1 & -7 \end{array}\right) \to \left(\begin{array}{cc|c} -2 & 2 & 10 \\ 0 & 3 & 3 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & -1 & -5 \\ 0 & 3 & 3 \end{array}\right).$$

Як бачите, рядок, який додавали – не змінився. Завжди змінюється рядок, до якого додають. На практиці так докладно, звичайно, не розписують, а пишуть коротше:

$$\left(\begin{array}{cc|c} 1 & -1 & -5 \\ 2 & 1 & -7 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & -1 & -5 \\ 0 & 3 & 3 \end{array}\right).$$

Ще раз: до другого рядка додали перший рядок, помножений на -2. Множать рядок зазвичай усно або на чернетці, при цьому уявний хід розрахунків приблизно такий:

1. Переписують матрицю і переписують перший рядок:

$$\left(\begin{array}{cc|c} 1 & -1 & -5 \\ 2 & 1 & -7 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & -1 & -5 \\ * & * & * \end{array}\right).$$

2. Спочатку перший стовпець. Знизу потрібно отримати нуль. Тому одиницю вгорі множать на -2: $1 \cdot (-2) = -2$, до другого рядка додають першу: 2 + (-2) = 0. Записують результат у другому рядку:

$$\left(\begin{array}{cc|c} 1 & -1 & -5 \\ 2 & 1 & -7 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & -1 & -5 \\ 0 & * & * \end{array}\right).$$

3. Тепер другий стовпець. Зверху -1 помножити на -2: $-1 \cdot (-2) = 2$. До другої рядку додаю першу: 1 + 2 = 3. Записую результат у другому рядку:

$$\left(\begin{array}{cc|c} 1 & -1 & -5 \\ 2 & 1 & -7 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & -1 & -5 \\ 0 & 3 & * \end{array}\right).$$

4. І третій стовпець. Зверху -5 множать на –2: $-5 \cdot (-2) = 10$. До другого рядка додають перший: -7 + 10 = 3. Записують результат у другому рядку:

$$\left(\begin{array}{cc|c} 1 & -1 & -5 \\ 2 & 1 & -7 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & -1 & -5 \\ 0 & 3 & 3 \end{array}\right).$$

Необхідно ретельно осмислити цей приклад і розібратись в послідовному алгоритмі обчислень. Якщо це зрозуміло, то метод Гауса практично «в кишені». Але, над цим перетворенням потрібно попрацювати.

Елементарні перетворення не змінюють розв'язку системи рівнянь.

Розглянуті дії не можна використовувати, якщо Вам запропоновано завдання, де матриці дано «самі по собі». Наприклад, при «класичних» діях з матрицями щось переставляти всередині матриць ні в якому разі не можна.

Приклад 3.16. Розв'язати систему за допомогою метода Гауса (заміни змінних).

$$\begin{cases} x - y = -5, \\ 2x + y = -7. \end{cases}$$

Розв'язок. На першому етапі (прямий хід методу Гауса) потрібно записати розширену матрицю системи:

$$\left(\begin{array}{cc|c} 1 & -1 & -5 \\ 2 & 1 & -7 \end{array}\right).$$

Запишемо розширену матрицю системи і за допомогою елементарних перетворень приведемо її до східчастого вигляду:

$$\left(\begin{array}{cc|c} 1 & -1 & -5 \\ 2 & 1 & -7 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{cc|c} 1 & -1 & -5 \\ 0 & 3 & 3 \end{array}\right) \overset{(2)}{\to} \left(\begin{array}{cc|c} 1 & -1 & -5 \\ 0 & 1 & 1 \end{array}\right).$$

(1) До другого рядка додали перший рядок, помножений на -2. І знову: чому перший рядок множимо саме на -2? Для того щоб внизу отримати нуль, а значить, позбутися однієї змінної у другому рядку.

(2) Ділимо другий рядок на 3. Мета елементарних перетворень – привести матрицю до ступінчастого вигляду (під головною діагоналлю всі 0, а діагональ складена з 1):

$$\left(\begin{array}{cc|c} 1 & -1 & -5 \\ 0 & 1 & 1 \end{array}\right).$$

В результаті елементарних перетворень отримана система рівнянь, яка еквівалентна до вихідної (початок зворотного ходу):

$$\begin{cases} x - y = -5, \\ y = 1. \end{cases}$$

Тепер систему потрібно «розкрутити» у зворотному напрямку – знизу до верху, цей процес називається зворотним ходом методу Гауса. У нижньому рівнянні у нас вже готовий результат: $y = 1$. Розглянемо перше рівняння системи $x - y = -5$ і підставимо в нього вже відоме значення $y = 1$: $x - 1 = -5$ та $x = -4$.

Тоді розв'язок СЛАР $x = -4$ та $y = 1$.

Розглянемо найбільш поширену ситуацію, коли методом Гауса потрібно розв'язати систему трьох лінійних рівнянь з трьома невідомими.

Приклад 3.17. Розв'язати систему за допомогою метода Гауса (заміни змінних).

$$\begin{cases} 3x+2y-5z=-1, \\ 2x-y+3z=13, \\ x+2y-z=9. \end{cases}$$

Розв'язок. Запишемо розширену матрицю системи:

$$\left(\begin{array}{ccc|c} 3 & 2 & -5 & -1 \\ 2 & -1 & 3 & 13 \\ 1 & 2 & -1 & 9 \end{array}\right).$$

Мета – за допомогою елементарних перетворень привести матрицю до ступінчастого вигляду. Спочатку дивимося на ліве верхнє $a_{11}=3$. У більшості випадків тут повинна знаходитися одиниця. Взагалі кажучи, влаштує і -1 (а іноді і інші числа), але якось так традиційно склалося, що туди зазвичай поміщають одиницю. Як організувати одиницю? Дивимося на перший стовпець – готова одиниця у нас є! Перетворення перше: міняємо місцями перший і третій рядки:

$$\left(\begin{array}{ccc|c} 3 & 2 & -5 & -1 \\ 2 & -1 & 3 & 13 \\ 1 & 2 & -1 & 9 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 2 & -1 & 3 & 13 \\ 3 & 2 & -5 & -1 \end{array}\right).$$

Тепер перший рядок у нас залишиться незмінним до кінця розв'язування. Одиниця в лівому верхньому кутку отримана. Тепер потрібно отримати нулі на a_{21} та a_{31} (тобто в першому стовпці під головною діагоналлю). Нулі отримуємо саме за допомогою «важкого» перетворення. Спочатку розбираємося з другим рядком (2, -1, 3, 13). Що потрібно зробити, щоб на першій позиції отримати нуль? Потрібно до другого рядка додати перший рядок, помножений на -2. Подумки або на чернетці множимо перший рядок на -2: (-2, -4, 2, -18). І послідовно проводимо (знову ж подумки або на чернетці) додавання: до другого рядка додаємо перший рядок, вже помножений на -2:

$$\begin{array}{cccc} 0 & -5 & 5 & -5 \\ \| & \| & \| & \| \\ -2 & -4 & 2 & -18 \\ + & + & + & + \\ 2 & -1 & 3 & 13 \end{array}$$

Результат записуємо у другому рядку:

$$\left(\begin{array}{ccc|c} 3 & 2 & -5 & -1 \\ 2 & -1 & 3 & 13 \\ 1 & 2 & -1 & 9 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 2 & -1 & 3 & 13 \\ 3 & 2 & -5 & -1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & -5 & 5 & -5 \\ 3 & 2 & -5 & -1 \end{array}\right).$$

Аналогічно розбираємося з третім рядком (3, 2, -5, -1). Щоб отримати на першій позиції нуль, потрібно до третього рядка додати перший рядок, помножений на -3. Подумки або на чернетці множимо перший рядок на -3: (-3, -6, 3, -27). І до третього рядка додаємо перший рядок, помножений на -3:

$$\begin{array}{cccc} 0 & -4 & -2 & -28 \\ \| & \| & \| & \| \\ -3 & -6 & 3 & -27 \\ + & + & + & + \\ 3 & 2 & -5 & -1 \end{array}$$

Результат потрібно записати в третій рядок:

$$\left(\begin{array}{ccc|c} 3 & 2 & -5 & -1 \\ 2 & -1 & 3 & 13 \\ 1 & 2 & -1 & 9 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 2 & -1 & 3 & 13 \\ 3 & 2 & -5 & -1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & -5 & 5 & -5 \\ 3 & 2 & -5 & -1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & -5 & 5 & -5 \\ 0 & -4 & -2 & -28 \end{array}\right).$$

На практиці ці дії зазвичай виконуються усно і записуються в один крок:

$$\left(\begin{array}{ccc|c} 3 & 2 & -5 & -1 \\ 2 & -1 & 3 & 13 \\ 1 & 2 & -1 & 9 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 2 & -1 & 3 & 13 \\ 3 & 2 & -5 & -1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & -5 & 5 & -5 \\ 0 & -4 & -2 & -28 \end{array}\right).$$

Не потрібно рахувати все відразу й одночасно. Порядок обчислень і «вписування» результатів послідовний і зазвичай такий: спочатку переписуємо перший рядок і рухаємось ПОСЛІДОВНО та УВАЖНО:

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ * & * & * & * \\ * & * & * & * \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & * & * & * \\ * & * & * & * \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & -5 & * & * \\ * & * & * & * \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & -5 & 5 & * \\ * & * & * & * \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & -5 & 5 & -5 \\ * & * & * & * \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & -5 & 5 & -5 \\ 0 & * & * & * \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & -5 & 5 & -5 \\ 0 & -4 & * & * \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & -5 & 5 & -5 \\ 0 & -4 & -2 & * \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & -5 & 5 & -5 \\ 0 & -4 & -2 & -28 \end{array}\right).$$

Уявний хід розрахунків вже розглянуто вище. Далі потрібно отримати одиницю наступної «сходинки» a_{22} (одиниці на головній діагоналі). В цьому прикладі це зробити легко, другий рядок ділимо на -5 (оскільки там всі числа діляться на 5 без залишку). Ділимо третій рядок на -2, адже чим менше числа, тим простіший розв'язок:

$$\left(\begin{array}{ccc|c} 3 & 2 & -5 & -1 \\ 2 & -1 & 3 & 13 \\ 1 & 2 & -1 & 9 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 2 & -1 & 3 & 13 \\ 3 & 2 & -5 & -1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & -5 & 5 & -5 \\ 0 & -4 & -2 & -28 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & 1 & -1 & 1 \\ 0 & 2 & 1 & 14 \end{array}\right)$$

На заключному етапі елементарних перетворень потрібно отримати ще один нуль на a_{32}. Для цього до третього рядка додаємо другий рядок, помножений на -2:

$$\left(\begin{array}{ccc|c} 3 & 2 & -5 & -1 \\ 2 & -1 & 3 & 13 \\ 1 & 2 & -1 & 9 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 2 & -1 & 3 & 13 \\ 3 & 2 & -5 & -1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & -5 & 5 & -5 \\ 0 & -4 & -2 & -28 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & 1 & -1 & 1 \\ 0 & 2 & 1 & 14 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 3 & 12 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 4 \end{array}\right).$$

Спробуйте розібрати цю дію самостійно – подумки помножте другий рядок на -2 і проведіть додавання. Остання виконана дія – приведення результату: ділимо третій рядок на 3. В результаті елементарних перетворень отримана еквівалентна до вихідної система лінійних рівнянь:

$$\begin{cases} x + 2y - z = 9, \\ y - z = 1, \\ z = 4. \end{cases}$$

Застосовуємо зворотний хід методу Гауса. Рівняння «розкручуються» знизу до верху. У третьому рівнянні у нас вже готовий результат: $z = 4$. Дивимося на друге рівняння: $y - z = 1$. Значення z вже відомо, таким чином: $y - 4 = 1$ та $y = 5$. І, нарешті, перше рівняння: $x + 2y - z = 9$. Дві змінні відомі, далі: $x + 2 \cdot 5 - 4 = 9$; $x + 6 = 9$ та $x = 3$. Тоді: $x = 3$, $y = 5$, $z = 4$.

Як вже зазначалося, для будь-якої системи рівнянь можна і потрібно зробити перевірку знайденого розв'язку (це нескладно і швидко).

Приклад 3.18. Розв'язати систему лінійних рівнянь методом Гауса:

$$\begin{cases} x + 2y + 3z = 1, \\ 2x - y + 2z = 6, \\ x + y + 5z = -1. \end{cases}$$

Розв'язок. Запишемо розширену матрицю системи і за допомогою елементарних перетворень зведемо її до східчастого вигляду:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 2 & -1 & 2 & 6 \\ 1 & 1 & 5 & -1 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -5 & -4 & 4 \\ 0 & -1 & 2 & -2 \end{array}\right) \overset{(2)}{\to} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -1 & 2 & -2 \\ 0 & 5 & 4 & -4 \end{array}\right) \overset{(3)}{\to}$$

$$\overset{(3)}{\to} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -1 & 2 & -2 \\ 0 & 0 & 14 & -14 \end{array}\right) \overset{(4)}{\to} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & -1 \end{array}\right).$$

Виконано елементарні перетворення:

(1) до другого рядка додали перший рядок, помножений на -2. До третього рядка додали перший рядок, помножений на -1. Може виникнути бажання від третього рядка відняти перший, але це не рекомендовано – сильно підвищується ризик помилки.

(2) У другому рядку змінили знак (помножили на -1). Другий і третій рядки поміняли місцями. Зверніть увагу, що на «сходинках» нас влаштовує не тільки одиниця, але ще й -1, що навіть зручніше.

(3) До третього рядка додали другий рядок, помножений на 5.

(4) У другому рядку змінили знак (помножили на -1). Третій рядок розділили на 14.

$z=-1$, та $y-2z=2 \Rightarrow y+2=2 \Rightarrow y=0$, і $x+2y+3z=1 \Rightarrow x+0-3=1 \Rightarrow x=4$.

Тоді $x=4$, $y=0$, $z=-1$.

Слід зазначити, що один хід розв'язування може не збігтися з іншим ходом розв'язування, і це – особливість методу Гауса. Але відповіді обов'язково повинні вийти однаковими.

Приклад 3.19. Розв'язати систему лінійних рівнянь методом Гауса:

$$\begin{cases} 4x+2y-z=1, \\ 5x+3y-2z=2, \\ 3x+2y-3z=0. \end{cases}$$

Розв'язок. Запишемо розширену матрицю системи і за допомогою елементарних перетворень зведемо її до східчастого вигляду:

$$\left(\begin{array}{ccc|c} 4 & 2 & -1 & 1 \\ 5 & 3 & -2 & 2 \\ 3 & 2 & -3 & 0 \end{array}\right).$$

Дивимося на ліву верхню «сходинку». Там у нас повинна бути одиниця. Проблема полягає в тому, що в першому стовпці одиниць немає взагалі, тому перестановкою рядків нічого не вирішити. У таких випадках одиницю потрібно організувати за допомогою елементарного перетворення. Зазвичай це можна зробити декількома способами.

(1) Наприклад, до першої рядка додаємо другий рядок, помножений на -1. Тобто, подумки помножили другий рядок на -1 і виконали додавання першого та другого рядка, при цьому другий рядок у нас не змінився.

$$\left(\begin{array}{ccc|c} 4 & 2 & -1 & 1 \\ 5 & 3 & -2 & 2 \\ 3 & 2 & -3 & 0 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{ccc|c} -1 & -1 & 1 & -1 \\ 5 & 3 & -2 & 2 \\ 3 & 2 & -3 & 0 \end{array}\right).$$

Тепер зліва вгорі «мінус один», що нас цілком влаштує. Щоб отримати +1, можна виконати додатково: помножити перший рядок на -1 (змінити знак). Далі алгоритм працює вже як раніше:

$$\left(\begin{array}{ccc|c} 4 & 2 & -1 & 1 \\ 5 & 3 & -2 & 2 \\ 3 & 2 & -3 & 0 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{ccc|c} -1 & -1 & 1 & -1 \\ 5 & 3 & -2 & 2 \\ 3 & 2 & -3 & 0 \end{array}\right) \overset{(2)}{\to} \left(\begin{array}{ccc|c} -1 & -1 & 1 & -1 \\ 0 & -2 & 3 & -3 \\ 0 & -1 & 0 & -3 \end{array}\right) \overset{(3)}{\to}$$

$$\overset{(3)}{\to} \left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & -2 & 3 & -3 \end{array}\right) \overset{(4)}{\to} \left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 3 & 3 \end{array}\right) \overset{(5)}{\to} \left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

(2) До другого рядка додали перший рядок, помножений на 5. До третього рядка додали перший рядок, помножений на 3.

(3) Перший рядок помножили на -1. У третього рядка також змінили знак і переставили його на друге місце. Таким чином, на другій «сходинці» у нас з'явилася потрібна одиниця.

(4) До третього рядка додали другий рядок, помножений на 2.

(5) Третій рядок розділили на 3.

Поганою ознакою, що свідчить про помилку в обчисленнях (рідко про друкарську помилку), є «поганий» нижній рядок. Тобто, якщо у нас внизу вийшло щось на кшталт $\left(\begin{array}{ccc|c} 0 & 0 & 11 & 23 \end{array}\right)$, і відповідно, $11z = 23 \Rightarrow z = \frac{23}{11}$, то з великою ймовірністю можна стверджувати, що допущена помилка в ході елементарних перетворень. Проводимо зворотний хід, в оформленні прикладів часто вже не переписують саму систему, а рівняння «беруть прямо з наведеної матриці» (зворотний хід працює знизу до верху). Отримуємо:

$$z = 1;\ y = 3 \text{ та } x + y - z = 1 \Rightarrow x + 3 - 1 = 1 \Rightarrow x = -1.$$

Тоді: $x = -1;\ y = 3;\ z = 1$.

Приклад 3.20. Розв'язати систему лінійних рівнянь методом Гауса:

$$\begin{cases} 8x + 7y + 3z = 18, \\ -7x - 4y - 4z = -11, \\ -6x + 5y - 4z = -15. \end{cases}$$

Розв'язок. Цей приклад для самостійного розв'язування, він трохи складніший. Запишемо розширену матрицю системи і за допомогою елементарних перетворень зведемо її до східчастого вигляду:

$$\left(\begin{array}{ccc|c} 8 & 7 & 3 & 18 \\ -7 & -4 & -4 & -11 \\ -6 & 5 & -4 & -15 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{ccc|c} 1 & 1 & -1 & 7 \\ -7 & -4 & -4 & -11 \\ -6 & 5 & -4 & -15 \end{array}\right) \overset{(2)}{\to} \left(\begin{array}{ccc|c} 1 & 3 & -1 & 7 \\ 0 & 17 & -11 & 38 \\ 0 & 23 & -10 & 27 \end{array}\right) \overset{(3)}{\to}$$

$$\overset{(3)}{\to} \left(\begin{array}{ccc|c} 1 & 3 & -1 & 7 \\ 0 & 17 & -11 & 38 \\ 0 & 6 & 1 & -11 \end{array}\right) \overset{(4)}{\to} \left(\begin{array}{ccc|c} 1 & 3 & -1 & 7 \\ 0 & -1 & -14 & 71 \\ 0 & 6 & 1 & -11 \end{array}\right) \overset{(5)}{\to} \left(\begin{array}{ccc|c} 1 & 3 & -1 & 7 \\ 0 & -1 & -14 & 71 \\ 0 & 0 & -83 & 415 \end{array}\right) \overset{(6)}{\to}$$

$$\overset{(6)}{\to} \left(\begin{array}{ccc|c} 1 & 3 & -1 & 7 \\ 0 & 1 & 14 & -71 \\ 0 & 0 & 1 & -5 \end{array}\right).$$

Виконані перетворення:

(1) До першого рядка додали другий. Таким чином, організована потрібна одиниця на лівій верхній «сходинці».

(2) До другого рядка додали перший рядок, помножений на 7. До третього рядка додали перший рядок, помножений на 6. З другої «сходинкою» все гірше, «кандидати» на неї – числа 17 і 23, а нам потрібна або +1, або -1. Перетворення (3) і (4) будуть спрямовані на отримання потрібної одиниці.

(3) До третього рядка додали другий, помножений на -1.

(4) До другого рядка додали третій, помножений на -3. Потрібне на другій сходинці отримано.

(5) До третього рядка додали другий, помножений на 6.

(6) Другий рядок помножили на -1, третій рядок розділили на -83.

Зворотній хід: $z = -5$;

$y + 14z = -71 \Rightarrow y - 70 = -71 \Rightarrow y = 1$;

$x + 3y - z = 7 \Rightarrow x - 3 + 5 = 7 \Rightarrow x = 5$.

Тоді: $x = 5$, $y = 1$, $z = -5$.

Розглянемо деякі *особливості алгоритму Гауса*. Перша особливість полягає в тому, що іноді в рівняннях системи відсутні деякі змінні, наприклад:

$$\begin{cases} 2x + y - z = 0, \\ 3y + 4z + 6 = 0 \\ x + z = 1. \end{cases}$$

У розширеній матриці системи на місці відсутніх змінних ставимо нулі:

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 0 \\ 0 & 3 & 4 & -6 \\ 1 & 0 & 1 & 1 \end{array}\right).$$

До речі, це досить легкий приклад, оскільки в першому стовпці вже є один нуль, і необхідно виконати менше елементарних перетворень. Друга особливість полягає ось у чому. У всіх розглянутих прикладах на «сходинки» ми розміщали або -1, або

+1. Чи можуть там бути інші числа? У деяких випадках можуть. Розглянемо систему:

$$\begin{cases} 2x-3y+z=2, \\ 2x+y-4z=9, \\ 6x-5y+2z=17. \end{cases}$$

Тут на лівій верхній «сходинці» у нас двійка. Але візначимо, що всі числа в першому стовпці діляться на 2 без залишку: і інша 2, і 6. І двійка зліва вгорі нас влаштує. На першому етапі необхідно виконати наступні перетворення: до другого рядка додати перший рядок, помножений на -1; до третього рядка додати перший рядок, помножений на -3. Таким чином, ми отримаємо потрібні нулі в першому стовпчику. Або ще такий умовний приклад:

$$\left(\begin{array}{ccc|c} 1 & 2 & -7 & 4 \\ 0 & 3 & -1 & 17 \\ 0 & 12 & 2 & 1 \end{array}\right).$$

Тут трійка на другий «сходинці» теж нас влаштовує, оскільки 12 (місце, де нам потрібно отримати нуль) ділиться на 3 без залишку. Необхідно провести наступне перетворення: до третього рядка додати другий рядок, помножений на -4, в результаті чого і буде отримано потрібний нам нуль.

Метод Гауса універсальний, але є одне особливість. Впевнено навчитися розв'язувати системи іншими методами (методом Крамера, матричним методом) можна з першого разу – там дуже жорсткий алгоритм. Але от щоб впевнено себе почувати в методі Гауса, слід розв'язати хоча б 5-10 систем. Тому спочатку можливі помилки в обчисленнях, і в цьому немає нічого незвичайного.

Приклад 3.21. Розв'язати методом Гауса систему лінійних рівнянь з чотирма невідомими:

$$\begin{cases} 2x+5y+4z+t=20, \\ x+3y+2z+t=11, \\ 2x+10y+9z+7t=40, \\ 3x+8y+9z+2t=37. \end{cases}$$

Розв'язок. Таке завдання на практиці зустрічається не так вже й рідко. Алгоритм розв'язування такої системи інтуїтивно зрозумілий. Принципово все таке ж, просто дій більше. Випадки, коли система не має розв'язків (несумісна) або має нескінченно багато розв'язків, розглянуті в пункті про несумісні системи і системи з частковим розв'язком.

Запишемо матрицю системи і за допомогою елементарних перетворень зведемо її до східчастого вигляду:

$$\left(\begin{array}{cccc|c} 2 & 5 & 4 & 1 & 20 \\ 1 & 3 & 2 & 1 & 11 \\ 2 & 10 & 9 & 7 & 40 \\ 3 & 8 & 9 & 2 & 37 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{cccc|c} 1 & 3 & 2 & 1 & 11 \\ 2 & 5 & 4 & 1 & 20 \\ 2 & 10 & 9 & 7 & 40 \\ 3 & 8 & 9 & 2 & 37 \end{array}\right) \overset{(2)}{\to} \left(\begin{array}{cccc|c} 1 & 3 & 2 & 1 & 11 \\ 0 & -1 & 0 & -1 & -2 \\ 0 & 4 & 5 & 5 & 18 \\ 0 & -1 & 3 & -1 & 4 \end{array}\right) \overset{(3)}{\to}$$

$$\overset{(3)}{\to} \left(\begin{array}{cccc|c} 1 & 3 & 2 & 1 & 11 \\ 0 & -1 & 0 & -1 & -2 \\ 0 & 0 & 5 & 1 & 10 \\ 0 & 0 & 3 & 0 & 6 \end{array}\right) \overset{(4)}{\to} \left(\begin{array}{cccc|c} 1 & 3 & 2 & 1 & 11 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 5 & 1 & 10 \end{array}\right) \overset{(5)}{\to} \left(\begin{array}{cccc|c} 1 & 3 & 2 & 1 & 11 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right).$$

Виконані перетворення:

(1) Перший і другий рядки поміняли місцями.

(2) До другого рядка додали перший рядок, помножений на -2. До третього рядка додали перший рядок, помножений на -2. До четвертого рядка додали перший рядок, помножений на -3.

(3) До третього рядка додали другий, помножений на 4. До четвертого рядка додали другий, помножений на -1.

(4) У другого рядка змінили знак. Четвертий рядок розділили на 3 та помістили замість третього рядка.

(5) До четвертого рядка додали третій рядок, помножений на -5.

Зворотній хід: $t = 0$; $z = 2$; $y + t = 2 \Rightarrow y = 2$;

$x + 3y + 2z + t = 11 \Rightarrow x + 6 + 4 + 0 = 11 \Rightarrow x = 1$.

Тоді: $x = 1$; $y = 2$; $z = 2$ та $t = 0$.

Приклад 3.22. Розв'язати систему за допомогою методу Гауса-Жордана.

$$\begin{cases} 3x + 2y - 5z = -1, \\ 2x - y + 3z - 13, \\ x + 2y - z = 9. \end{cases}$$

Розв'язок. Запишемо систему в матричні формі та зведемо до східчастого вигляду:

$$\left(\begin{array}{ccc|c} 3 & 2 & -5 & -1 \\ 2 & -1 & 3 & 13 \\ 1 & 2 & -1 & 9 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 2 & -1 & 3 & 13 \\ 3 & 2 & -5 & -1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & -5 & 5 & -5 \\ 0 & -4 & -2 & -28 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & 1 & -1 & 1 \\ 0 & 2 & 1 & 14 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 3 & 12 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 4 \end{array}\right).$$

Тепер замість зворотного ходу використовуємо додаткові елементарні перетворення. Спочатку необхідно отримати нулі на місцях, a_{13} та a_{23}, а потім ще один нуль на a_{12}. Ідеальний з точки зору простоти випадок:

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 9 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 4 \end{array}\right) \xrightarrow{(1)} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 13 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 4 \end{array}\right) \xrightarrow{(2)} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 4 \end{array}\right).$$

(6) До другого рядка додали третій рядок. До першого рядка додали третій рядок.

(7) До першого рядка додали другий рядок, помножений на -2.

Підсумкова система:

$$\begin{cases} x = 3, \\ y = 5, \\ z = 4. \end{cases}$$

3.6.9. Несумісні системи. Системи із загальним розв'язком. Часткові розв'язки

Продовжуємо розглядати системами лінійних рівнянь. До цих пір це були системи, які сумісні і мають єдиний розв'язок. Такі системи можна розв'язати будь-яким способом: методом підстановки, за формулами Крамера, матричним методом, методом Гауса. Однак на практиці широко поширені ще два випадки:

- Система *несумісна* (не має розв'язків);
- Система *сумісна* і має нескінченно багато розв'язків.

Зауваження. Термін «сумісність» означає, що у системи існує хоч один розв'язок. У ряді завдань потрібно попередньо дослідити систему на сумісність. Для цих систем застосовують найбільш універсальний з усіх способів розв'язування – метод Гауса. У вищій математиці прийнято використовувати гаусівський метод послідовного виключення невідомих. Спочатку розглянемо пару прикладів, коли система не має розв'язків (несумісна).

Приклад 3.23. Розв'язати систему лінійних рівнянь

$$\begin{cases} 4x_1 - 3x_2 + 2x_3 - x_4 = 8, \\ 3x_1 - 2x_2 + x_3 - 3x_4 = 7, \\ 5x_1 - 3x_2 + x_3 - 8x_4 = 1. \end{cases}$$

Розв'язок. В цій системі кількість рівнянь менша, ніж кількість змінних. Якщо кількість рівнянь менша, ніж кількість змінних, то відразу можна сказати, що система або несумісна, або має нескінченно багато розв'язків. І це залишилося тільки з'ясувати. Початок розв'язування абсолютно звичайний: запишемо розширену матрицю системи і за допомогою елементарних перетворень зведемо її до східчастого вигляду:

$$\left(\begin{array}{cccc|c} 4 & -3 & 2 & -1 & 8 \\ 3 & -2 & 1 & -3 & 7 \\ 5 & -3 & 1 & -8 & 1 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{cccc|c} -1 & 0 & 1 & 7 & 7 \\ 3 & -2 & 1 & -3 & 7 \\ 5 & -3 & 1 & -8 & 1 \end{array}\right) \overset{(2)}{\to} \left(\begin{array}{cccc|c} -1 & 0 & 1 & 7 & 7 \\ 0 & -2 & 4 & 18 & 28 \\ 0 & -3 & 6 & 27 & 36 \end{array}\right) \overset{(3)}{\to}$$

$$\overset{(3)}{\to} \left(\begin{array}{cccc|c} -1 & 0 & 1 & 7 & 7 \\ 0 & -1 & 2 & 9 & 14 \\ 0 & 1 & -2 & -9 & 12 \end{array}\right) \overset{(4)}{\to} \left(\begin{array}{cccc|c} -1 & 0 & 1 & 7 & 7 \\ 0 & -1 & 2 & 9 & 14 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right).$$

(1) На лівій верхній сходинці нам потрібно отримати +1 або -1. Таких чисел в першому стовпці немає, тому перестановка рядків нічого не дасть. Одиницю доведеться організувати самостійно, і зробити це можна кількома способами. Робимо так: До першого рядка додаємо третій рядок, помножений на -1.

(2) Тепер отримуємо два нулі в першому стовпці. До другого рядка додаємо перший рядок, помножений на 3. До третього рядка додаємо перший рядок, помножений на 5.

(3) Після виконаного перетворення завжди доцільно подивитися, чи не можна спростити отримані рядки. Можна, можливо. Другий рядок ділимо на 2, заодно отримуючи потрібну -1 на другій сходинці. Третій рядок ділимо на -3.

(4) До третього рядка додаємо другий рядок.

Напевно, звернули увагу на рядок, який вийшов в результаті елементарних перетворень: $(0\ \ 0\ \ 0\ \ 0|2)$. Зрозуміло, що так бути не може. Дійсно, перепишемо отриману матрицю назад в систему лінійних рівнянь:

$$\left(\begin{array}{cccc|c} -1 & 0 & 1 & 7 & 7 \\ 0 & -1 & 2 & 9 & 14 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right).$$

Якщо в результаті елементарних перетворень отримано рядок виду $(0\ \ 0\ \ \ldots\ \ 0|\lambda)$, *де* λ *- число, відмінне від нуля, то система несумісна (не має розв'язків).*

Як записати кінець завдання? В результаті елементарних перетворень отримано рядок виду $(0\ \ 0\ \ \ldots\ \ 0|\lambda)$, де $\lambda \neq 0$ і відповідь: система не має розв'язків (несумісна). Якщо ж за умовою потрібно досліджувати систему на сумісність, тоді необхідно оформити розв'язок в іншому стилі із залученням поняття рангу матриці і теореми Кронекера-Капеллі. Зверніть увагу, що тут немає ніякого зворотного ходу алгоритму Гауса – розв'язків немає і знаходити просто нічого.

Приклад 3.24. Розв'язати систему лінійних рівнянь

$$\begin{cases} 7x - 2y - z = 2, \\ 6x - 4y - 5z = 3, \\ x + 2y + 4z = 5. \end{cases}$$

Розв'язок. Запишемо розширену матрицю системи і за допомогою елементарних перетворень зведемо її до східчастого вигляду.

$$\left(\begin{array}{ccc|c} 7 & -2 & -1 & 2 \\ 6 & -4 & -5 & 3 \\ 1 & 2 & 4 & 5 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{ccc|c} 1 & 2 & 4 & 5 \\ 6 & -4 & -5 & 3 \\ 7 & -2 & -1 & 2 \end{array}\right) \overset{(2)}{\to} \left(\begin{array}{ccc|c} 1 & 2 & 4 & 5 \\ 0 & -16 & -29 & -27 \\ 0 & -16 & -29 & -33 \end{array}\right) \overset{(3)}{\to} \left(\begin{array}{ccc|c} 1 & 2 & 4 & 5 \\ 0 & -16 & -29 & -27 \\ 0 & 0 & 0 & -6 \end{array}\right).$$

Виконано елементарні перетворення:

(1) Перший і третій рядок поміняли місцями.

(2) До другого рядка додали перший рядок, помножений на -6. До третього рядка додали перший рядок, помножений на -7.

(3) До третього рядка додали другий рядок, помножений на -1. В результаті елементарних перетворень отримано рядок виду $(0 \quad 0 \quad \ldots \quad 0|\lambda)$, де $\lambda \neq 0$, значить, система несумісна.

Відповідь: розв'язків немає.

Знову нагадуємо, що ваш хід розв'язування може відрізнятися від цього ходу розв'язування: у алгоритму Гауса немає сильної «жорсткості». Ще одна технічна особливість: елементарні перетворення можна припиняти відразу ж, як тільки з'явився рядок виду $(0 \quad 0 \quad \ldots \quad 0|\lambda)$, де $\lambda \neq 0$.

Розглянемо умовний приклад: припустимо, що після першого ж перетворення вийшла матриця

$$\left(\begin{array}{cccc|c} 1 & 6 & 3 & -2 & 3 \\ 0 & 12 & -9 & 8 & 3 \\ 0 & 3 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 7 \end{array}\right).$$

Матриця ще не приведена до ступінчастого вигляду, але в подальших елементарних перетвореннях немає ніякої необхідності, тому що з'явився рядок виду $(0 \quad 0 \quad \ldots \quad 0|\lambda)$, де $\lambda \neq 0$. Слід відразу дати відповідь, що система несумісна.

Коли система лінійних рівнянь не має розв'язків - це майже подарунок, з огляду на те, що виходить короткий розв'язок, іноді буквально в 2-3 дії. Завдання, в якому система має нескінченно багато розв'язків – довше.

Приклад 3.25. Розв'язати систему лінійних рівнянь

$$\begin{cases} 2x+3y-z+t=1, \\ 8x+12y-9z+8t=3, \\ 4x+4y+4z-2t=3, \\ 2x+3y+9z-7t=3. \end{cases}$$

Розв'язок. Тут 4 рівняння і 4 невідомі. Таким чином, система може мати або єдиний розв'язок, або не мати розв'язків, або мати нескінченно багато розв'язків. Метод Гауса в будь-якому випадку приведе до відповіді. У цьому його і універсальність. Початок знову стандартний. Запишемо розширену матрицю системи і за допомогою елементарних перетворень зведемо її до східчастого вигляду:

$$\left(\begin{array}{cccc|c} 2 & 3 & -1 & 1 & 1 \\ 8 & 12 & -9 & 8 & 3 \\ 4 & 6 & 3 & -2 & 3 \\ 2 & 3 & 9 & -7 & 3 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{cccc|c} 2 & 3 & -1 & 1 & 1 \\ 0 & 0 & -5 & 4 & -1 \\ 0 & 0 & 5 & -4 & 1 \\ 0 & 0 & 10 & -8 & 2 \end{array}\right) \overset{(2)}{\to} \left(\begin{array}{cccc|c} 2 & 3 & -1 & 1 & 1 \\ 0 & 0 & 5 & -4 & 1 \end{array}\right).$$

(1) Всі числа в першому стовпці діляться на 2, тому на лівій верхній сходинці нас влаштовує і двійка. До другого рядка додаємо перший рядок, помножений на -4. До третього рядка додаємо перший рядок, помножений на -2. До четвертого рядка додаємо перший рядок, помножений на -1.

Може виникнути спокуса з четвертого рядка відняти перший рядок. Так робити можна, але не потрібно, досвід показує, що ймовірність помилки в обчисленнях збільшується в декілька разів. Тільки додаємо: до четвертого рядка додаємо перший рядок, помножений на -1 .

(2) Останні три рядки пропорційні, два з них можна видалити. Потрібно з'ясувати, чи дійсно рядки пропорційні? Для перевірки не зайвим буде другий рядок помножити на -1, а четвертий рядок розділити на 2, отримавши в результаті три однакові рядки. І тільки після цього видалити два з них. В результаті елементарних перетворень розширена матриця системи приведена до ступінчастого вигляду:

$$\left(\begin{array}{cccc|c} 2 & 3 & -1 & 1 & 1 \\ 0 & 0 & 5 & -4 & 1 \end{array}\right).$$

Перепишемо відповідну систему рівнянь:

$$\begin{cases} 2x+3y-z+t=1, \\ 5z-4t=1. \end{cases}$$

«Звичайного» єдиного розв'язку системи тут не видно. «Поганого» рядка $(0 \quad 0 \quad \ldots \quad 0|\lambda)$ теж немає. Значить, це залишився третій випадок – система має нескінченно багато розв'язків. Іноді за умовою потрібно досліджувати сумісність системи (тобто довести, що розв'язок взагалі існує). Але поки розглянемо таке:

Нескінченну множину розв'язків системи коротко записують у вигляді так званого загального розв'язку системи.

Загальний розв'язок системи знайдемо за допомогою зворотного ходу методу Гауса. Спочатку потрібно визначити, які змінні у нас є *базисними*, а які змінні *вільними*. Існують так звані *базисні змінні і вільні змінні*.

Базисні змінні завжди розміщені строго на сходинках матриці.

В даному прикладі базисними змінними є x і z.

Вільні змінні – це всі інші змінні, для яких не знайшлося сходинки. У нашому випадку їх дві: y,t – вільні змінні. Тепер потрібно всі *базисні змінні* виразити тільки через *вільні змінні*.

Зворотний хід алгоритму Гауса традиційно працює знизу до верху. З другого рівняння системи записуємо базисну змінну z:

$$5z-4t=1 \Rightarrow 5z=4t+1 \Rightarrow z=\frac{4}{5}t+\frac{1}{5}.$$

Тепер дивимося на перше рівняння: $2x+3y-z+t=1$. Спочатку в нього підставляємо знайдений вираз: $z=\frac{4}{5}t+\frac{1}{5}$:

$$2x+3y-(\frac{4}{5}t+\frac{1}{5})+t=1.$$

Залишилося записати базисну змінну x через вільні змінні y,t:

$$2x+3y-\frac{4}{5}t-\frac{1}{5}+t=1; \qquad 2x=1-3y+\frac{4}{5}t+\frac{1}{5}-t;$$

$$2x=-3y-\frac{1}{5}t+\frac{6}{5}; \qquad x=-\frac{3}{2}y-\frac{1}{10}t+\frac{3}{5}.$$

У результаті вийшло те, що потрібно – всі базисні змінні (x і z) виражені лише через вільні змінні y,t: $x=-\frac{3}{2}y-\frac{1}{10}t+\frac{3}{5}$ та $z=\frac{4}{5}t+\frac{1}{5}$. Загальний розв'язок готовий:

$$\left(-\frac{3}{2}y-\frac{1}{10}t+\frac{3}{5};y;\frac{4}{5}t+\frac{1}{5};t\right).$$

Як правильно записати загальний розв'язок? Вільні змінні записуються в загальний розв'язок строго на своїх місцях. В даному випадку вільні змінні y,t слід записати на другій і четвертій позиції. Отримані ж вирази для базисних змінних (x і z), очевидно, потрібно записати на першій і третій позиції. Надаючи вільним змінним y,t довільні значення, можна знайти нескінченно багато часткових розв'язків. Найпопулярнішими значеннями є нулі, оскільки часткові розв'язки виходять найпростіше. Підставимо $y=0,t=0$ в загальний розв'язок:

$$\left(-0-0+\frac{3}{5};0;0+\frac{1}{5};0\right).$$

$\left(\frac{3}{5};0;\frac{1}{5};0\right)$ – частковий розв'язок.

Іншими популярними значеннями є одиниці, підставимо $y=1,t=1$ в загальний розв'язок:

$$\left(-\frac{3}{2}-\frac{1}{10}+\frac{3}{5};1;\frac{4}{5}+\frac{1}{5};1\right).$$

$(-1;1;1;1)$ – ще один частковий розв'язок.

Легко помітити, що система рівнянь має *нескінченно багато розв'язків* (так як вільним змінним ми можемо надати будь-які значення).

Кожен частковий розв'язок має задовольняти *кожне* рівняння системи. На цьому заснована «швидка» перевірка правильності розв'язку. Візьміть, наприклад, частковий розв'язок $(-1;1;1;1)$ і підставте його в ліву частину кожного рівняння вихідної системи:

$$\begin{cases} 2x+3y-z+t=1, \\ 8x+12y-9z+8t=3, \\ 4x+4y+4z-2t=3, \\ 2x+3y+9z-7t=3. \end{cases}$$

Все повинно зійтися. І з будь-яким отриманим вами частковими розв'язками теж все має зійтися. Але, строго кажучи, перевірка часткового розв'язку є оманливою, тобто якийсь частковий розв'язок може задовольняти кожне рівняння системи, а загальний розв'язок насправді знайдений невірно. Тому більш ґрунтовна і надійна перевірка загального розв'язку. Як перевірити отримане загальний розв'язок:

$$\left(-\frac{3}{2}y-\frac{1}{10}t+\frac{3}{5};y;\frac{4}{5}t+\frac{1}{5};t\right)?$$

Це нескладно, але досить марудно. Потрібно взяти вирази *базисних* змінних, в даному випадку $x=-\frac{3}{2}y-\frac{1}{10}t+\frac{3}{5}$ і $z=\frac{4}{5}t+\frac{1}{5}$, і підставити їх в ліву частину кожного рівняння системи. У ліву частину першого рівняння системи:

$$2x+3y-z+t=2\left(-\frac{3}{2}y-\frac{1}{10}t+\frac{3}{5}\right)+3y-\left(\frac{4}{5}t+\frac{1}{5}\right)+t=$$

$$=-3y-\frac{1}{5}t+\frac{6}{5}+3y-\frac{4}{5}t-\frac{1}{5}+t=\frac{6}{5}-\frac{1}{5}=1.$$

Отримано права частина вихідного рівняння. У ліву частину рівняння 2 системи:

$$8x+12y-9z+8t=8\left(-\frac{3}{2}y-\frac{1}{10}t+\frac{3}{5}\right)+12y-9\left(\frac{4}{5}t+\frac{1}{5}\right)+8t=$$

$$=-12y-\frac{4}{5}t+\frac{24}{5}+12y-\frac{36}{5}t-\frac{9}{5}+8t=\frac{24}{5}-\frac{9}{5}=3.$$

Отримано права частина вихідного рівняння. І далі – в ліві частини третього і четвертого рівняння системи. Це довше, але зате гарантує стовідсоткову правильність загального розв'язку. Крім того, в деяких завданнях вимагають перевірку загального розв'язку.

Приклад 3.26. Розв'язати систему лінійних рівнянь методом Гауса. Знайти загальний розв'язок і два часткових. Зробити перевірку загального розв'язку. Тут знову кількість рівнянь менша, ніж кількість невідомих, а значить, відразу зрозуміло, що система буде або несумісною, або з нескінченною множиною розв'язків.

$$\begin{cases} -x_1+2x_2+x_3-4x_4+3x_5=1, \\ 2x_1-x_2-3x_3+x_4-x_5=-4, \\ 3x_1+2x_2-x_3-2x_4-5x_5=1. \end{cases}$$

Розв'язок. Запишемо розширену матрицю системи і за допомогою елементарних перетворень зведемо її до східчастого вигляду:

$$\left(\begin{array}{ccccc|c} -1 & 2 & 1 & -4 & 3 & 1 \\ 2 & -1 & -3 & 1 & -1 & -4 \\ 3 & 2 & -1 & -2 & -5 & 1 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{ccccc|c} -1 & 2 & 1 & -4 & 3 & 1 \\ 0 & 3 & -1 & -7 & 5 & -2 \\ 0 & 8 & 2 & -14 & 4 & 4 \end{array}\right) \overset{(2)}{\to}$$

$$\overset{(2)}{\to} \left(\begin{array}{ccccc|c} -1 & 2 & 1 & -4 & 3 & 1 \\ 0 & 3 & -1 & -7 & 5 & -2 \\ 0 & -1 & 5 & 7 & -11 & 10 \end{array}\right) \overset{(3)}{\to} \left(\begin{array}{ccccc|c} -1 & 2 & 1 & -4 & 3 & 1 \\ 0 & -1 & 5 & 7 & -11 & 10 \\ 0 & 3 & -1 & -7 & 5 & -2 \end{array}\right) \overset{(4)}{\to}$$

$$\overset{(4)}{\to} \left(\begin{array}{ccccc|c} -1 & 2 & 1 & -4 & 3 & 1 \\ 0 & -1 & 5 & 7 & -11 & 10 \\ 0 & 0 & 14 & 14 & -28 & 28 \end{array}\right) \overset{(5)}{\to} \left(\begin{array}{ccccc|c} 1 & -2 & -1 & 4 & -3 & -1 \\ 0 & 1 & -5 & -7 & 11 & -10 \\ 0 & 0 & 1 & 1 & -2 & 2 \end{array}\right).$$

Виконані перетворення:

(1) До другого рядка додали перший рядок, помножений на 2. До третього рядка додали перший рядок, помножений на 3. На другій сходинці немає одиниці, і перетворення (2) скероване на її отримання.

(2) До третього рядка додали другий рядок, помножений на -3.

(3) Другий рядок з третім поміняли місцями (переставили отриману -1 на другу сходинку).

(4) До третього рядка додали другий рядок, помножений на 3.

(5) У перших двох рядках змінили знак (помножили на -1), третій рядок розділили на 14.

Зворотній хід. x_1, x_2, x_3 – базисні змінні (ті, які на сходинках), x_4, x_5 – вільні змінні (ті, яким не дісталися сходинки). Запишемо базисні змінні через вільні змінні. З третього рівняння: $x_3 + x_4 - 2x_5 = 2 \Rightarrow x_3 = -x_4 + 2x_5 + 2$. Розглянемо друге рівняння: $x_2 - 5x_3 - 7x_4 + 11x_5 = -10$. Підставимо в нього знайдений вираз $x_3 = -x_4 + 2x_5 + 2$:

$$x_2 - 5(-x_4 + 2x_5 + 2) - 7x_4 + 11x_5 = -10$$

$$x_2 + 5x_4 - 10x_5 - 10 - 7x_4 + 11x_5 = -10$$

$$x_2 - 2x_4 - 10 + x_5 = -10$$

$$x_2 = 2x_4 - x_5$$

Розглянемо перше рівняння: $x_1 - 2x_2 - x_3 + 4x_4 - 3x_5 = -1$. Підставимо в нього знайдені вирази: $x_2 = 2x_4 - x_5$ та $x_3 = -x_4 + 2x_5 + 2$:

$$x_1 - 2(2x_4 - x_5) - (-x_4 + 2x_5 + 2) + 4x_4 - 3x_5 = -1$$

$$x_1 - 4x_4 + 2x_5 + x_4 - 2x_5 - 2 + 4x_4 - 3x_5 = -1$$

$$x_1 + x_4 - 3x_5 - 2 = -1$$

$$x_1 = -x_4 + 3x_5 + 1$$

Загальний розв'язок: $(-x_4 + 3x_5 + 1; 2x_4 - x_5; -x_4 + 2x_5 + 2; x_4; x_5)$.

Знайдемо два часткових розв'язки. Якщо $x_4 = x_5 = 0$ то $(1;0;2;0;0)$. Якщо $x_4 = x_5 = 1$ то $(3;1;3;1;1)$. Тоді відповідь: часткові розв'язки: $(1;0;2;0;0)$, $(3;1;3;1;1)$, а

загальний розв'язок: $(-x_4+3x_5+1;2x_4-x_5;-x_4+2x_5+2;x_4;x_5)$. Перевірка: підставимо знайдений розв'язок (базисні змінні: $x_1=-x_4+3x_5+1$, $x_2=2x_4-x_5$ і $x_3=-x_4+2x_5+2$) в ліву частину кожного рівняння системи:

$$-x_1+2x_2+x_3-4x_4+3x_5=-(-x_4+3x_5+1)+2(2x_4-x_5)+(-x_4+2x_5+2)-4x_4+3x_5=$$
$$=x_4-3x_5-1+4x_4-2x_5-x_4+2x_5+2-4x_4+3x_5=-1+2=1,$$
$$2x_1-x_2-3x_3+x_4-x_5=2(-x_4+3x_5+1)-(2x_4-x_5)-3(-x_4+2x_5+2)+x_4-x_5=$$
$$=-2x_4+6x_5+2-2x_4+x_5+3x_4-6x_5-6+x_4-x_5=2-6=-4,$$
$$3x_1+2x_2-x_3-2x_4-5x_5=3(-x_4+3x_5+1)+2(2x_4-x_5)-(-x_4+2x_5+2)-2x_4-5x_5=$$
$$=-3x_4+9x_5+3+4x_4-2x_5+x_4-2x_5-2-2x_4-5x_5=3-2=1.$$

Отримано відповідні праві частини, тому загальний розв'язок знайдено вірно.

Розглянемо приклади для закріплення матеріалу.

Приклад 3.27. Розв'язати систему лінійних рівнянь. Якщо система має нескінченно багато розв'язків, знайти два часткових розв'язки і зробити перевірку загального розв'язку.

$$\begin{cases} x+3y-2z-2t=-3, \\ -x-2y+z+2t=2, \\ -2x-y+3z+t=-2, \\ -3x-2y+3z+3t=-1. \end{cases}$$

Розв'язок. Запишемо розширену матрицю системи і за допомогою елементарних перетворень зведемо її до східчастого вигляду:

$$\left(\begin{array}{cccc|c} 1 & 3 & -2 & -2 & -3 \\ -1 & -2 & 1 & 2 & 2 \\ -2 & -1 & 3 & 1 & -2 \\ -3 & -2 & 3 & 3 & -1 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{cccc|c} 1 & 3 & -2 & -2 & -3 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 5 & -1 & -3 & -8 \\ 0 & 7 & -3 & -3 & -10 \end{array}\right) \overset{(2)}{\to}$$

$$\overset{(2)}{\to} \left(\begin{array}{cccc|c} 1 & 3 & -2 & -2 & -3 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 4 & -3 & -3 \\ 0 & 0 & 4 & -3 & -3 \end{array}\right) \overset{(3)}{\to} \left(\begin{array}{cccc|c} 1 & 3 & -2 & -2 & -3 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 4 & -3 & -3 \end{array}\right).$$

(1) До рядка 2 додаємо рядок 1. До третього рядка додаємо перший рядок, помножений на 2. До четвертого рядка додаємо перший рядок, помножений на 3.

(2) До третього рядка додаємо другий рядок, помножений на -5. До четвертого рядка додаємо другий рядок, помножений на -7.

(3) Третій і четвертий рядки однакові, один з них видаляємо.

Базисні змінні знаходяться на сходинках, тому x, y, z – базисні змінні. Вільна змінна тут всього одна: t.

Зворотній хід. Запишемо базисні змінні через вільну змінну. З третього рівняння: $4z-3t=-3\Rightarrow z=\frac{3}{4}t-\frac{3}{4}$. Розглянемо друге рівняння $y-z=-1$ і підставимо в нього знайдений вираз $z=\frac{3}{4}t-\frac{3}{4}$:

$$y-\left(\frac{3}{4}t-\frac{3}{4}\right)=-1;\qquad y-\frac{3}{4}t+\frac{3}{4}=-1;\qquad y=\frac{3}{4}t-\frac{7}{4}.$$

Розглянемо перше рівняння $x+3y-2z-2t=-3$ і підставимо в нього знайдені вирази $y=\frac{3}{4}t-\frac{7}{4}$ і $z=\frac{3}{4}t-\frac{3}{4}$:

$$x+3\left(\frac{3}{4}t-\frac{7}{4}\right)-2\left(\frac{3}{4}t-\frac{3}{4}\right)-2t=-3;\qquad x+\frac{9}{4}t-\frac{21}{4}-\frac{3}{2}t+\frac{3}{2}-2t=-3;$$

$$x-\frac{5}{4}t-\frac{21}{4}+\frac{3}{2}=-3;\qquad x=\frac{5}{4}t+\frac{3}{4}.$$

Таким чином, загальний розв'язок: $\left(\frac{5}{4}t+\frac{3}{4};\frac{3}{4}t-\frac{7}{4};\frac{3}{4}t-\frac{3}{4};t\right)$.

Вільна змінна t знаходиться на своєму законному четвертому місці. Отримані вирази для базисних змінних $\frac{5}{4}t+\frac{3}{4};\frac{3}{4}t-\frac{7}{4};\frac{3}{4}t-\frac{3}{4}$, теж зайняли свої порядкові місця. Відразу виконаємо перевірку загального розв'язку. Підставляємо $\frac{5}{4}t+\frac{3}{4};\frac{3}{4}t-\frac{7}{4};\frac{3}{4}t-\frac{3}{4}$ в ліву частину кожного рівняння системи:

$$x+3y-2z-2t=\frac{5}{4}t+\frac{3}{4}+3\left(\frac{3}{4}t-\frac{7}{4}\right)-2\left(\frac{3}{4}t-\frac{3}{4}\right)-2t=$$

$$=\frac{5}{4}t+\frac{3}{4}+\frac{9}{4}t-\frac{21}{4}-\frac{6}{4}t+\frac{6}{4}-2t=-3,$$

$$-x-2y+z+2t=-\left(\frac{5}{4}t+\frac{3}{4}\right)-2\left(\frac{3}{4}t-\frac{7}{4}\right)+\frac{3}{4}t-\frac{3}{4}+2t=$$

$$=-\frac{5}{4}t-\frac{3}{4}-\frac{6}{4}t+\frac{14}{4}+\frac{3}{4}t-\frac{3}{4}+2t=2,$$

$$-2x-y+3z+t=-2\left(\frac{5}{4}t+\frac{3}{4}\right)-\left(\frac{3}{4}t-\frac{7}{4}\right)+3\left(\frac{3}{4}t-\frac{3}{4}\right)+t=$$

$$=-\frac{10}{4}t-\frac{6}{4}-\frac{3}{4}t+\frac{7}{4}+\frac{9}{4}t-\frac{9}{4}+t=-2,$$

$$-3x-2y+3z+3t=-3\left(\frac{5}{4}t+\frac{3}{4}\right)-2\left(\frac{3}{4}t-\frac{7}{4}\right)+3\left(\frac{3}{4}t-\frac{3}{4}\right)+3t=$$

$$=-\frac{15}{4}t-\frac{9}{4}-\frac{6}{4}t+\frac{14}{4}+\frac{9}{4}t-\frac{9}{4}+3t=-1.$$

Отримано відповідні праві частини рівнянь, таким чином, загальний розв'язок знайдено вірно. Тепер із знайденого загального розв'язку $\left(\frac{5}{4}t+\frac{3}{4};\frac{3}{4}t-\frac{7}{4};\frac{3}{4}t-\frac{3}{4};t\right)$ отримаємо два часткових через єдину вільну змінну t.

Нехай $t=0$, тоді $\left(\frac{3}{4};-\frac{7}{4};-\frac{3}{4};0\right)$ – частковий розв'язок. Нехай $t=1$, тоді $\left(\frac{5}{4}+\frac{3}{4};\frac{3}{4}-\frac{7}{4};\frac{3}{4}-\frac{3}{4};1\right)=(2;-1;0;1)$ – ще один частковий розв'язок.

Відповідь: Загальний розв'язок: $\left(\frac{5}{4}t+\frac{3}{4};\frac{3}{4}t-\frac{7}{4};\frac{3}{4}t-\frac{3}{4};t\right)$, часткові розв'язки: $\left(\frac{3}{4};-\frac{7}{4};-\frac{3}{4};0\right)$, $(2;-1;0;1)$.

Приклад 3.28. Знайти загальний розв'язок системи лінійних рівнянь.

$$\begin{cases} x-2y-z+3t=2, \\ -2x+y+3z-2t=-3, \\ 2x-y-2z+t=2, \\ 3x-3y-2z+3t=3. \end{cases}$$

Розв'язок. Запишемо розширену матрицю системи і за допомогою елементарних перетворень зведемо її до східчастого вигляду:

$$\left(\begin{array}{cccc|c} 1 & -2 & -1 & 3 & 2 \\ -2 & 1 & 3 & -2 & -3 \\ 2 & -1 & -2 & 1 & 2 \\ 3 & -3 & -2 & 3 & 3 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{cccc|c} 1 & -2 & -1 & 3 & 2 \\ 0 & -3 & 1 & 4 & 1 \\ 0 & 3 & 0 & -5 & -2 \\ 0 & 3 & 1 & -6 & -3 \end{array}\right) \overset{(2)}{\to}$$

$$\overset{(2)}{\to} \left(\begin{array}{cccc|c} 1 & -2 & -1 & 3 & 2 \\ 0 & -3 & 1 & 4 & 1 \\ 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 2 & -2 & -2 \end{array}\right) \overset{(3)}{\to} \left(\begin{array}{cccc|c} 1 & -2 & -1 & 3 & 2 \\ 0 & -3 & 1 & 4 & 1 \\ 0 & 0 & 1 & -1 & -1 \end{array}\right).$$

(1) До другого рядка додаємо перший рядок, помножений на 2. До третього рядка додаємо перший рядок, помножений на -2. До четвертого рядка додаємо перший рядок, помножений на -3.

(2) До третього рядка додаємо другий рядок. До четвертого рядка додаємо другий рядок.

(3) Третій і четвертий рядки пропорційні, один із них видаляємо. x, y, z – базисні змінні, t – вільна змінна. Запишемо базисні змінні через вільну змінну:

$z-t=-1 \Rightarrow z=t-1$;

$-3y+z+4t=1 \Rightarrow -3y+t-1+4t=1 \Rightarrow y=\frac{5}{3}t-\frac{2}{3}$;

$$x-2y-z+3t=2 \Rightarrow x-2\left(\frac{5}{3}t-\frac{2}{3}\right)-(t-1)+3t=2 \Rightarrow x=\frac{4}{3}t-\frac{1}{3}.$$

Загальний розв'язок: $\left(\frac{4}{3}t-\frac{1}{3};\frac{5}{3}t-\frac{2}{3};t-1;t\right)$.

Перевірка загального розв'язку вже зроблена, відповіді можна довіряти. Напевно, можна помітити неприємний момент в розв'язуванні: дуже часто при зворотному ході методу Гауса довелося мати справу зі звичайними дробами. На практиці це дійсно так, випадки, коли дробів немає – зустрічаються значно рідше. Зупинимося на деяких особливостях розв'язування, які не зустрілися в прикладах, що розв'язувались. У загальний розв'яз системи іноді може входити константа (або константи), наприклад: $(5;t-2;7t+3;t)$. Тут одна з базисних змінних дорівнює постійному числу: $x=5$. Очевидно, що в даному випадку будь-який частковий розв'язок буде містити п'ятірку на першій позиції. Рідко, але зустрічаються системи, в яких *кількість рівнянь більша кількості змінних*. Метод Гауса працює в різних умовах, слід привести розширену матрицю системи до східчастого вигляду за стандартним алгоритмом. Така система може бути несумісною, може мати нескінченно багато розв'язків, і може мати єдиний розв'язок.

Для методу Гауса-Жордана є характерними специфічні прийоми та незручні обчислення.

Приклад 3.29. Розв'язати систему за допомогою методу Гауса-Жордана.

$$\begin{cases} x-8y+z-9t=6, \\ x-4y-z-5t=2, \\ -3x+2y+8z+5t=4, \\ 5x+2y+2z+3t=12. \end{cases}$$

Розв'язок. Запишемо систему в матричній формі та зведемо до східчастого вигляду:

$$\left(\begin{array}{cccc|c} 1 & -8 & 1 & -9 & 6 \\ 1 & -4 & -1 & -5 & 2 \\ -3 & 2 & 8 & 5 & 4 \\ 5 & 2 & 2 & 3 & 12 \end{array}\right) \overset{(1)}{\rightarrow} \left(\begin{array}{cccc|c} 1 & -8 & 1 & -9 & 6 \\ 0 & 4 & -2 & 4 & -4 \\ 0 & -22 & 11 & -22 & 22 \\ 0 & 42 & -3 & 48 & -18 \end{array}\right) \overset{(2)}{\rightarrow} \left(\begin{array}{cccc|c} 1 & -8 & 1 & -9 & 6 \\ 0 & 2 & -1 & 2 & -2 \\ 0 & -2 & 1 & -2 & 2 \\ 0 & 14 & -1 & 16 & -6 \end{array}\right) \overset{(3)}{\rightarrow}$$

$$\overset{(3)}{\rightarrow} \left(\begin{array}{cccc|c} 1 & -8 & 1 & -9 & 6 \\ 0 & 2 & -1 & 2 & -2 \\ 0 & 0 & 6 & 2 & 8 \end{array}\right) \overset{(4)}{\rightarrow} \left(\begin{array}{cccc|c} 1 & -8 & 1 & -9 & 6 \\ 0 & 2 & -1 & 2 & -2 \\ 0 & 0 & 3 & 1 & 4 \end{array}\right).$$

(1) До другого рядка додали перший рядок, помножений на -1. До третього рядка додали перший рядок, помножений на 3. До четвертого рядка додали перший рядок, помножений на -5.

(2) Другий рядок розділили на 2, третій рядок розділили на 11, четвертий рядок розділили на 3.

(3) Другий і третій рядки пропорційні, 3-й рядок видалили. До четвертого рядка додали другий рядок, помножений на -7

(4) Третій рядок розділили на 2.

Очевидно, що система має нескінченно багато розв'язків, і завдання – привести її розширену матрицю до вигляду

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{array}\right).$$

Знаходимо найменше спільне кратне чисел третього стовпчика (1, -1 і 3), тобто - найменше число, яке б ділилося без залишку і на 1, і на -1 і на 3. В даному випадку, це, звичайно ж, «трійка». Тепер в третьому стовпці нам потрібно отримати однакові по модулю числа, і саме такі міркування обумовили 5-е перетворення матриці:

(5) Перший рядок множимо на -3, другий рядок множимо на 3. Перший рядок можна було помножити теж на 3, але це було б менш зручно для наступної дії:

$$\left(\begin{array}{cccc|c} 1 & -8 & 1 & -9 & 6 \\ 0 & 2 & -1 & 2 & -2 \\ 0 & 0 & 3 & 1 & 4 \end{array}\right) \overset{(5)}{\to} \left(\begin{array}{cccc|c} -3 & 24 & -3 & 27 & -18 \\ 0 & 6 & -3 & 6 & -6 \\ 0 & 0 & 3 & 1 & 4 \end{array}\right) \overset{(6)}{\to} \left(\begin{array}{cccc|c} -3 & 24 & 0 & 28 & -14 \\ 0 & 6 & 0 & 7 & -2 \\ 0 & 0 & 3 & 1 & 4 \end{array}\right) \overset{(7)}{\to}$$

$$\overset{(7)}{\to} \left(\begin{array}{cccc|c} -3 & 24 & 0 & 28 & -14 \\ 0 & -24 & 0 & -28 & 8 \\ 0 & 0 & 3 & 1 & 4 \end{array}\right) \overset{(8)}{\to} \left(\begin{array}{cccc|c} -3 & 0 & 0 & 0 & -6 \\ 0 & -24 & 0 & -28 & 8 \\ 0 & 0 & 3 & 1 & 4 \end{array}\right) \overset{(9)}{\to} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 7/6 & -1/3 \\ 0 & 0 & 1 & 1/3 & 4/3 \end{array}\right)$$

(6) До другого рядка додали третій рядок. До першого рядка додали третій рядок.

(7) У другому стовпці два ненульових значення (24 і 6) і знову потрібно отримати однакові по модулю числа. В даному випадку все склалося досить вдало: найменше кратне 24, і найефективніше помножити другий рядок на -4.

(8) До першого рядка додали другий.

(9) Заключний штрих: перший рядок розділили на -3, другий рядок розділили на -24 і третій рядок розділили на 3. Ця дія виконується В ОСТАННЮ ЧЕРГУ. Ніяких дробів дочасно!

В результаті елементарних перетворень отримана еквівалентна до вихідної система:

$$\begin{cases} x = 2, \\ y + \dfrac{7}{6}t = -\dfrac{1}{3}, \\ z + \dfrac{1}{3}t = \dfrac{4}{3}. \end{cases}$$

Елементарно записуємо базисні змінні через вільну:

$$\begin{cases} x = 2, \\ y = -\frac{7}{6}t - \frac{1}{3}, \\ z = -\frac{1}{3}t + \frac{4}{3}. \end{cases}$$

і записуємо **загальний розв'язок:** $\left(2; -\frac{7}{6}t - \frac{1}{3}; -\frac{1}{3}t + \frac{4}{3}; t\right)$.

У подібних прикладах застосування розглянутого алгоритму найчастіше виправдано, оскільки зворотний хід методу Гауса зазвичай вимагає трудомістких і неприємних обчислень з дробами. І, зрозуміло, необхідна перевірка, яка виконується за звичайною схемою, яка розглядалась попередньо. Так, для знаходження базисного розв'язку за допомогою елементарних перетворень

$$\begin{cases} 2x - 3y - z - 7t + 5v = 4, \\ x - 2y - 3z + t - 4v = -5, \\ -5x + 4y + 6z + 6t + 5v = 4. \end{cases}$$

таке формулювання завдання передбачає використання методу Гауса-Жордана, і матриця приводиться до стандартного вигляду з базисними змінними:

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & * & * & * \\ 0 & 1 & 0 & * & * & * \\ 0 & 0 & 1 & * & * & * \end{array}\right).$$

Однак, завжди тримайте на увазі, що в якості базисних можна вибрати й інші змінні. Так, наприклад, якщо в першому стовпці громіздкі числа, то цілком допустимо привести матрицю до вигляду

$$\left(\begin{array}{ccccc|c} * & 1 & 0 & 0 & * & * \\ * & 0 & 1 & 0 & * & * \\ * & 0 & 0 & 1 & * & * \end{array}\right).$$

(базисні змінні y, z, t), або до виду

$$\left(\begin{array}{ccccc|c} * & * & 1 & 0 & 0 & * \\ * & * & 0 & 1 & 0 & * \\ * & * & 0 & 0 & 1 & * \end{array}\right).$$

(базисні змінні z, t, v), або навіть до виду

$$\left(\begin{array}{ccccc|c} * & 1 & * & 0 & 0 & * \\ * & 0 & * & 1 & 0 & * \\ * & 0 & * & 0 & 1 & * \end{array}\right).$$

з базисними змінними y, t, v.

Існують й інші варіанти, наприклад,

$$\left(\begin{array}{ccccc|c} 0 & * & 1 & * & 0 & * \\ 0 & * & 0 & * & 1 & * \\ 1 & * & 0 & * & 0 & * \end{array}\right).$$

Втім, буває важко втриматися від нетипового базису, коли у вихідній матриці, скажімо, в 4-м стовпці є два готових нуля. Якщо в розширеній матриці даних розмірів раптом виявляється пара лінійно залежних рядків, то її слід спробувати привести до звичного вигляду з базисними змінними x, y

$$\left(\begin{array}{cccc|c} 1 & 0 & * & * & * \\ 0 & 1 & * & * & * \end{array}\right).$$

3.6.10. Способи знаходження оберненої матриці

Квадратна матриця має обернену тоді й тільки тоді, коли вона *невироджена*, тобто її визначник не дорівнює нулю. Для неквадратних матриць і сингулярних матриць обернених матриць не існує.

$$\det A^{-1} = \frac{1}{\det A} \text{ та } (AB)^{-1} = B^{-1} \cdot A^{-1}.$$

Якщо необхідно розв'язати систему лінійних рівнянь $Ax = b$, де x — шуканий вектор, і якщо A^{-1} існує, то $x = A^{-1}b$. В протилежному випадку розмірність простору розв'язків більше нуля.

Якщо матриця така, що можливо знайти обернену, то для знаходження оберненої матриці можна скористатися одним із наступних способів:

1. Метод Гауса

I. Допишемо до матриці A одиничну матрицю E такого ж розміру що й A. Одержимо блокову матрицю $(A|E)$. Приведемо блок A до одиничної матриці методом Гауса, але використовуючи тільки операції над рядками. Після зведення блоку A до E, E зведеться до оберненої до A. Ми одержимо $(A|E) \Rightarrow (E|A^{-1})$.

II. Візьмемо дві матриці: саму A і одиничну E. Зведемо матрицю A до одиничної матриці методом Гауса. Після застосування кожної операції до першої матриці застосуємо ту ж операцію до другої. Коли зведення першої матриці до одиничного вигляду буде завершено, друга матриця виявиться зведена до вигляду A^{-1}.

При використанні методу Гауса перша матриця буде множитися ліворуч на одну з елементарних матриць Λ_i (трансвекцію або діагональну матрицю з одиницями на головній діагоналі, крім однієї позиції): $\Lambda_1 \cdot \ldots \cdot \Lambda_n \cdot A = \Lambda A = E \Rightarrow \Lambda = A^{-1}$. Друга матриця після застосування всіх операцій стане рівною Λ, тобто буде шуканою.

Алгоритм

1. Вибирається перший стовпчик ліворуч, у якому є хоч одне відмінне від нуля значення.

2. Якщо саме верхнє число в цьому стовпчику є нуль, то міняється весь перший рядок матриці на інший рядок матриці, де в цьому стовпчику немає нуля.
3. Всі елементи першого рядка діляться на верхній елемент обраного стовпчика.
4. З рядків, що залишилися, віднімається перший рядок, помножений на перший елемент відповідного рядка, щоб одержати першим елементом кожного рядка (крім першого) нуля.
5. Далі проводимо таку ж процедуру з матрицею, що отримуємо із вихідної матриці після викреслювання першого рядка й першого стовпця.
6. Після повторення цієї процедури n-1 раз одержуємо верхню трикутну матрицю.
7. Віднімаємо з передостаннього рядка останній рядок, помножений на відповідний коефіцієнт, для того, щоб у передостанньому рядку залишилася тільки 1 на головній діагоналі.
8. Повторюємо попередній крок для наступних рядків. В результаті одержуємо одиничну матрицю і розв'язок на місці вільного вектора (з ним необхідно проводити всі ті ж перетворення).

Приклад 3.30. Розв'язати наступну систему рівнянь:

$$\begin{cases} a+b+c=0, \\ 4a+2b+c=1, \\ 9a+3b+c=3. \end{cases}$$

Розв'язок. Запишемо її у вигляді матриці 3×4, де останній стовпець є вільним членом:

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 4 & 2 & 1 & 1 \\ 9 & 3 & 1 & 3 \end{pmatrix} \overset{(1)}{\to} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & -2 & -3 & 1 \\ 0 & -6 & -8 & 3 \end{pmatrix} \overset{(2)}{\to} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 3/2 & -1/2 \\ 0 & 0 & 1 & 0 \end{pmatrix} \overset{(3)}{\to}$$

$$\overset{(3)}{\to} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1/2 \\ 0 & 0 & 1 & 0 \end{pmatrix} \overset{(4)}{\to} \begin{pmatrix} 1 & 0 & 0 & 1/2 \\ 0 & 1 & 0 & -1/2 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Проведемо наступні дії:
(1) до Рядка_2 додамо: -4*Рядок_1; до Рядка_3 додамо: -9*Рядок_1;
(2) до Рядка_3 додамо: -3*Рядок_2; Рядок_2 ділимо на -2;
(3) до Рядка_1 додамо: -1*Рядок_3; до Рядка_2 додамо: -3/2*Рядок_ 3;
(4) до Рядка_1 додамо: -1*Рядок_2. У правому стовпці одержуємо розв'язок:

$$a=\frac{1}{2}; b=-\frac{1}{2}; c=0.$$

2. За допомогою союзної матриці

Формула для знаходження оберненої матриці:

$$A^{-1}=\frac{1}{\det A}\cdot\left(C^{*}\right)^{T}, \tag{3.110}$$

де C^* – союзна матриця; $(C^*)^T$ – матриця, отримана в результаті транспонування союзної матриці; отримана матриця A^{-1} і буде оберненою.

Союзна матриця – матриця, складена з алгебраїчних доповнень до елементів вихідної матриці.

$$C^* = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1j} \\ A_{21} & A_{22} & \dots & A_{2j} \\ \dots & \dots & \dots & \dots \\ A_{i1} & A_{i2} & \dots & A_{ij} \end{pmatrix}, \tag{3.111}$$

де A_{ij} – алгебраїчне доповнення.

Алгебраїчним доповненням елемента a_{ij} матриці A називається число

$$A_{ij} = (-1)^{i+j} A_{ij}, \tag{3.112}$$

де M_{ij} – мінор, визначник матриці, що отримується із A викреслюванням i-го рядка і j-го стовпця.

Властивості методу знаходження оберненої матриці через союзну.

1. Назва «алгебраїчне доповнення» пов'язана з формулами розкладу визначника матриці по рядку (по стовпцю):

$$\det A = \sum_{j=1}^{n} a_{ij} A_{ij} = \sum_{i=1}^{n} a_{ij} A_{ij}. \tag{3.113}$$

2. Лема про фалышивий розклад визначника стверджує, що

$$\sum_{j=1}^{n} a_{i_1 j} A_{i_2 j} = \sum_{i=1}^{n} a_{ij_1} A_{ij_2} = 0. \tag{3.114}$$

3. З цих тверджень випливає правильність алгоритму знаходження оберненої матриці:

1) замінити кожний елемент на його алгебраїчне доповнення (у результаті буде отримано союзну матрицю);
2) транспонувати матрицю;
3) розділити кожний елемент на визначник вихідної матриці.

Визначник або **детермінант** – одна з найважливіших характеристик квадратних матриць. Визначник матриці розмірності $n \times n$ дорівнює орієнтованому n-мірному об'єму паралелепіпеда, який натягнуто на її вектори рядків (або стовпців). Для матриці $n \times n$ визначник виражається у вигляді багаточлена ступеня n від елементів матриці, що являє собою суму добутків елементів матриці з різними комбінаціями номерів, що відрізняються, рядків і стовпців, причому в кожному з добутків елемент із будь-якого рядка й будь-якого стовпця рівно один. Кожному добутку приписується знак плюс або мінус залежно від парності перестановки номерів.

Якщо елементами матриці є числа, то визначник – це теж число. У загальному випадку визначник може бути функціональним, векторним тощо, тобто поданим іншими виразами, що складені з елементів. Визначник матриці $n \times n$ задається формулою:

$$\det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_{i=1}^{n!} (-1)^{p(i)} \cdot a_{1k_{i1}} \cdot a_{1k_{i2}} \dots \cdot a_{1k_{in}}, \qquad (3.115)$$

де $\det(A)$ і $|A|$ – позначення визначника; k_{ij} – i-та перестановка послідовності $k_1 = 1,\dots,n$, тобто, $k_{1j} = j$; $p(i)$ – кількість перестановок пар номерів у послідовності до k_{1j}, яка необхідна для того, щоб вона перетворилася в послідовність k_{ij},

або формула для обчислення визначника по заданому рядку матриці:

$$\det(A) = |A| = \sum_{i=1}^{n} (-1)^{k+i} \cdot a_{1k} \cdot M_{ik}, \qquad (3.116)$$

де i – номер рядка матриці; M_{ik} – додатковий мінор; k – номер стовпця матриці.

Зауваження. Можна виділити наступні особливості побудови виразів для визначника матриці $n \times n$:

- вираз є сума членів, кожний з яких складається з n співмножників;
- кількість доданків у сумі дорівнює кількості перестановок n номерів, тобто, $n!$;
- номери рядків і стовпців елементів, що входять в один доданок, не повторюються;
- доданки входять у суму або із плюсом, або з мінусом, залежно від парності перестановки;
- доданок з елементів головної діагоналі матриці, тобто, $a_{11}a_{22}\dots a_{nn}$ є із плюсом.

$$\det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}, \qquad (3.117)$$

$$\det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - \qquad (3.118)$$

$$- a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}.$$

3.6.11. Застосування методу Гауса до обчислення визначника матриці

Для обчислення визначників матриць застосовують два підходи:

- рекурсивний розрахунок за допомогою розкладення по рядку чи стовпцю;
- обчислення на основі прямого ходу метода Гауса.

Перший спосіб грунтується на використанні тієї властивості визначників, що визначник матриці дорівнює сумі добутків елементів будь-якого рядка чи стовпця на їх алгебраїчне доповнення, тобто:

$$\det(A) = \sum_{j=1}^{n} a_{ij} \cdot A_{ij} \ , \ \forall i = \overline{1,n} \ . \tag{3.119}$$

Обчислення одного визначника n-го порядку зводиться до розрахунку n визначників порядку n-1. Реалізується даний спосіб за допомогою рекурсії. Рекурсивний спосіб зручно застосовувати по відношенню до рядків чи стовпців, що мають велику кількість нульових елементів. Якщо ж нульових елементів у матриці немає або дуже мало, то застосування цього способу є вкрай неефективним. Для визначника n порядку доведеться розрахувати $n!/2$ визначників другого порядку.

Другий спосіб грунтується на методі, що базується на алгоритмі прямого ходу метода Гауса, використовує властивість визначника трикутної матриці. Для такої матриці визначник дорівнює добутку елементів головної діагоналі.

Для обчислення визначника використовується алгоритм побудови послідовності матриць $A^{(0)}, A^{(1)}, A^{(2)}, ...A^{(n)}$ прямого ходу метода Гауса з тією відмінністю, що при перестановці рядків чи стовпців знак визначника змінюється на протилежний. Значення визначника розраховується за формулою:

$$\det(A) = (-1)^m \cdot a_{22}^{(0)} \cdot a_{22}^{(1)} \cdot \ldots \cdot a_{nn}^{(n-1)}, \tag{3.120}$$

де m – кількість перестановок.

Цей метод дозволяє обчислювати визначники матриць великих порядків.

Для обчислення визначника методом Гауса використовують формули:

$$\begin{aligned} A^{(0)} &= a_{11}^{(0)} \cdot A^{(1)} \\ A^{(1)} &= a_{22}^{(1)} A^{(2)} = a_{22}^{(1)} a_{11}^{(0)} A^{(1)} \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ A &= a_{nn}^{(n-1)} \cdot \ldots \cdot a_{22}^{(1)} \cdot a_{11}^{(0)} A^{(n)} \\ \det A &= \prod_{i=1}^{n} a_{i,i}^{(i-1)} \end{aligned} \tag{3.121}$$

Принцип вибору головного (ведучого) елемента. Часто буває, що на деякому k-ому кроці методу Гауса елемент $a_{kk}^{(k-1)}$ дорівнює або наближено дорівнює нулю. В такому випадку доцільно в цьому стовпчику знайти найбільший по модулю

елемент, який, наприклад, буде знаходитись в p-ій стрічці, де $p>k$, і замінити елементи k-ої стрічки елементами p-ої. Це приводить до того, що:

1) уникаємо загрози ділення на нуль;
2) похибка в результаті ділення стає мінімально можливою.

Зауваження: взагалі таку процедуру (знаходження найбільшого по модулю елемента) варто було б проводити не лише в k-ому стовпчику, а серед всіх елементів матриці

Приклад 3.31. Обчислити визначник для матриці

$$\begin{pmatrix} 9 & 3 & 1 & 1 \\ 4 & 2 & 1 & 1 \\ 4 & 4 & 3 & 3 \\ 2 & 1 & 2 & 2 \end{pmatrix}.$$

Розв'язок. Запишемо розв'язок через алгебраїчне доповнення

$$\det(A)=\begin{vmatrix} 9 & 3 & 1 & 1 \\ 4 & 2 & 1 & 1 \\ 4 & 4 & 3 & 3 \\ 2 & 1 & 2 & 2 \end{vmatrix} = 9\cdot A_{11}+3\cdot A_{12}+1\cdot A_{13}+1\cdot A_{14}, \text{ де } A_{ij}=(-1)^{i+j}\cdot M_{ij}.$$

$$\det(A)=9\cdot\begin{vmatrix} 2 & 1 & 1 \\ 4 & 3 & 3 \\ 1 & 2 & 2 \end{vmatrix}-3\cdot\begin{vmatrix} 4 & 1 & 1 \\ 4 & 3 & 3 \\ 2 & 2 & 2 \end{vmatrix}+1\cdot\begin{vmatrix} 4 & 2 & 1 \\ 4 & 4 & 3 \\ 2 & 1 & 2 \end{vmatrix}-1\cdot\begin{vmatrix} 4 & 2 & 1 \\ 4 & 4 & 3 \\ 2 & 1 & 2 \end{vmatrix}=$$

$$=9\cdot(2\cdot3\cdot2+4\cdot2\cdot1+1\cdot3\cdot1-1\cdot3\cdot1-4\cdot1\cdot2-2\cdot3\cdot2)-$$
$$-3\cdot(4\cdot3\cdot2+4\cdot2\cdot1+1\cdot3\cdot2-1\cdot2\cdot3-2\cdot3\cdot1-4\cdot3\cdot2)+$$
$$+1\cdot(4\cdot4\cdot2+2\cdot3\cdot2+4\cdot1\cdot-2\cdot4\cdot1-4\cdot2\cdot2-1\cdot3\cdot4)-$$
$$-1\cdot(4\cdot4\cdot2+4\cdot1\cdot1+2\cdot3\cdot2-2\cdot4\cdot1-4\cdot2\cdot2-1\cdot2\cdot4)=$$
$$=9\cdot0-3\cdot2+8-8=6.$$

3.6.12. Застосування методу Гауса до інверсії матриці

Позначимо обернену матрицю через $X=A^{-1}$ причому $|A|\neq0$. Нагадаємо форму оберненої матриці $AX=E$. Візьмемо з E лише перший стовпчик і запишемо першу підсистему:

$$\begin{cases} a_{11}x_{11}+a_{12}x_{21}+\ldots+a_{1n}x_{n1}=1 \\ \vdots \\ a_{n1}x_{11}+a_{n2}x_{21}+\ldots+a_{nn}x_{n1}=0 \end{cases} \qquad (3.122)$$

другу підсистему

$$\begin{cases} a_{11}x_{12} + a_{12}x_{22} + \ldots + a_{1n}x_{n2} = 0 \\ \vdots \\ a_{n1}x_{12} + a_{n2}x_{22} + \ldots + a_{nn}x_{n2} = 0 \end{cases} \tag{3.123}$$

Як видно, застосовуючи цей підхід n разів, можна за один прямий прохід методу Гауса обчислити розв'язки всіх n підсистем по n рівнянь, які мають однакову ліву частину і відрізняються лише правою.

$$\sum_{k=1}^{n} a_{ik}x_{kj} = \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases} \qquad i, j = \overline{1,n} \tag{3.124}$$

Записуємо розрахункові формули, приведені до однієї системи з n рівнянь, для випадку системи (3.69), і враховуємо, що для розширеної матриці буде матриця розмірності $2n$, де права половина - одинична матриця:

$$\hat{A} = \left(\underbrace{\begin{matrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{matrix}}_{2n} \right). \tag{3.125}$$

$$a_{kj}^{(k)} = \frac{a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}}, \qquad \begin{matrix} j = \overline{k+1, 2n} \\ k = \overline{1,n} \end{matrix} \tag{3.126}$$

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - a_{ik}^{(k-1)}a_{kj}^{(k)} = a_{ij}^{(k-1)} - a_{ik}^{(k-1)}\frac{a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}}, \qquad \begin{matrix} j = \overline{k+1, 2n} \\ i = \overline{k+1, n} \\ k = \overline{1, n-1} \end{matrix} \tag{3.127}$$

$$x_{nj} = \frac{a_{n,n+j}^{(n-1)}}{a_{n,n}^{(n-1)}}, \qquad j = \overline{1,n} \tag{3.128}$$

$$x_i = a_{i,n+1}^{(i)} - \sum_{k=i+1}^{n} a_{i,k}^{(i)}x_k, \qquad \begin{matrix} j = \overline{1,n} \\ i = \overline{n-1,1} \end{matrix} \tag{3.129}$$

$$x_{ij} = a_{i,n+j}^{(i)} - \sum_{k=i+1}^{n} a_{ik}^{(i)}x_{kj}, \qquad \begin{matrix} j = \overline{1,n} \\ i = \overline{n-1,1} \end{matrix} \tag{3.130}$$

$$\det A \neq 0 \Leftrightarrow \forall i : \ell_{ii} \neq 0. \tag{3.131}$$

Тому, перевіряючи ℓ_{11} на рівність нулю при умові, що $\ell_{11} \neq 0$, на початку можна відразу зробити висновок про сумісність чи несумісність системи (що не

вимагає доповнюючої перестановки стрічок). Відповідні теореми доводять кращу стійкість методу LU-факторизації (який детально описаний далі в цьому розділі) порівняно з методом Гауса по відношенню до нагромадження похибки обчислень. Інші прямі методи розв'язування систем лінійних рівнянь розраховані на їх спеціальний вигляд, а саме: метод квадратного кореня для так званих розріджених матриць.

Приклад 3.32. Знайти обернену матрицю методом Гауса

$$\begin{pmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{pmatrix}.$$

Розв'язок. Запишемо її у вигляді матриці 3×4, де останній стовпець є вільним членом:

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 4 & 2 & 1 & 0 & 1 & 0 \\ 9 & 3 & 1 & 0 & 0 & 1 \end{array}\right) \overset{(1)}{\to} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -2 & -3 & -4 & 1 & 0 \\ 0 & -6 & -8 & -9 & 0 & 1 \end{array}\right) \overset{(2)}{\to} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -2 & -3 & -4 & 1 & 0 \\ 0 & 0 & 1 & 3 & -3 & 1 \end{array}\right) \overset{(3)}{\to}$$

$$\overset{(3)}{\to} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & -2 & 3 & -1 \\ 0 & -2 & 0 & 5 & -8 & 3 \\ 0 & 0 & 1 & 3 & -3 & 1 \end{array}\right) \overset{(4)}{\to} \left(\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & -2 & 1 \\ 0 & -2 & 0 & 5 & -8 & 3 \\ 0 & 0 & 1 & 3 & -3 & 1 \end{array}\right) \overset{(5)}{\to} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0{,}5 & -1 & 0{,}5 \\ 0 & 1 & 0 & 2{,}5 & 4 & -1{,}5 \\ 0 & 0 & 1 & 3 & -3 & 1 \end{array}\right).$$

(1) Перший рядок множимо на -4 і додаємо результат до другого рядка. Початкові значення першого рядка множимо на -3 і додаємо результат до третього рядка. Перший рядок переписуємо без змін, а наступні два зі змінами.

(2) Другий рядок множимо на -3 та додаємо результат до третього рядка. Перший та другий рядки переписуємо без змін, а третій зі змінами. Отримали нулі під головною діагоналлю. Тепер маємо рухатися зворотно таким же чином.

(3) Третій рядок множимо на 3 та результат додаємо до другого. Третій рядок (початкові результати, отримані перед кроком 3) множимо на -1 та додаємо до першого рядка. Третій рядок переписуємо без змін, а інші два – із отриманими результатами.

(4) Перший рядок множимо на 2 та додаємо до результату другий рядок. Третій та другий рядки переписуємо без змін, а перший – із змінами.

(5) Перший рядок ділимо на 2, а другий рядок ділимо на -2. Отримаємо зліва одиничну матрицю, а справа – обернену матрицю до початкових даних. Перевірка:

$$\begin{pmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{pmatrix} \times \begin{pmatrix} 0{,}5 & -1 & 0{,}5 \\ 2{,}5 & 4 & -1{,}5 \\ 3 & -3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Для знаходження оберненої матриці методом Гауса використовують, по суті, метод Гауса-Жордана. Короткий зміст майбутніх дій: спочатку слід записати квадратну матрицю A в тандемі з одиничною матрицею: $(A|E)$. Потім за допомогою елементарних перетворень необхідно отримати одиничну матрицю зліва, при цьому (не вдаючись у теоретичні подробиці) праворуч вийде обернена матриця. Схематично рішення виглядає наступним чином:

$$(A|E)\to\ldots\to(E|A^{-1}).$$

(Зрозуміло, що зворотна матриця повинна існувати)

Приклад 3.33. Знайти обернену матрицю методом Гауса за допомогою елементарних перетворень для матриці.

$$A=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}.$$

Розв'язок. Для цього запишемо її з одиничною матрицею:

$$(A|E)=\left(\begin{array}{cc|cc}1 & 2 & 1 & 0\\ 3 & 4 & 0 & 1\end{array}\right)\overset{(1)}{\to}\left(\begin{array}{cc|cc}1 & 2 & 1 & 0\\ 0 & -2 & -3 & 1\end{array}\right)\overset{(2)}{\to}\left(\begin{array}{cc|cc}1 & 0 & -2 & 1\\ 0 & -2 & -3 & 1\end{array}\right)\overset{(3)}{\to}$$

$$\overset{(3)}{\to}\left(\begin{array}{cc|cc}1 & 0 & -2 & 1\\ 0 & 1 & 3/2 & -1/2\end{array}\right)=(E|A^{-1}).$$

(1) До другого рядка додали перший рядок, помножений на -3.

(2) До першого рядка додали другий рядок.

(3) Другий рядок розділили на -2.

Відповідь: $A^{-1}=\begin{pmatrix}-2 & 1\\ 3/2 & -1/2\end{pmatrix}=\frac{1}{2}\begin{pmatrix}-4 & 2\\ 3 & -1\end{pmatrix}$.

Приклад 3.34. Знайти обернену матрицю за допомогою елементарних перетворень

$$\begin{pmatrix}3 & 4 & 2\\ 2 & -1 & -3\\ 1 & 5 & 1\end{pmatrix}.$$

Розв'язок. Приєднали одиничну матрицю і починаємо виконувати перетворення, дотримуючись алгоритму «звичайного» методу Гауса:

$$(A|E)=\left(\begin{array}{ccc|ccc}3 & 4 & 2 & 1 & 0 & 0\\ 2 & -1 & -3 & 0 & 1 & 0\\ 1 & 5 & 1 & 0 & 0 & 1\end{array}\right)\overset{(1)}{\to}\left(\begin{array}{ccc|ccc}1 & 5 & 1 & 0 & 0 & 1\\ 2 & -1 & -3 & 0 & 1 & 0\\ 3 & 4 & 2 & 1 & 0 & 0\end{array}\right)\overset{(2)}{\to}$$

$$\overset{(2)}{\to}\left(\begin{array}{ccc|ccc}1 & 5 & 1 & 0 & 0 & 1\\ 0 & -11 & -5 & 0 & 1 & -2\\ 0 & -11 & -1 & 1 & 0 & -3\end{array}\right)\overset{(3)}{\to}\left(\begin{array}{ccc|ccc}1 & 5 & 1 & 0 & 0 & 1\\ 0 & -11 & -5 & 0 & 1 & -2\\ 0 & 0 & 4 & 1 & -1 & -1\end{array}\right).$$

(1) Перший і третій рядки поміняли місцями. На перший погляд, перестановка рядків здається нелегальною, але насправді переставляти їх можна, адже в результаті $(E|A^{-1})$ зліва нам потрібно отримати одиничну матрицю, а праворуч «примусово» вийде саме матриця A^{-1} (незалежно від того чи будемо переставляти рядки в ході розв'язування чи ні). Зверніть увагу, що тут замість перестановки можна організувати «шістки» в 1-му стовпці (найменше спільне кратне (НСК) чисел 3, 2 і 1). Розв'язувати через НСК особливо зручно, коли в першому стовпці відсутні «одиниці».

(2) До 2-го і 3-го рядків додали 1-й рядок, помножений на -2 і -3 відповідно.

(3) До 3-го рядка додали 2-й рядок, помножений на -1.

Друга частина розв'язування проводиться за вже відомою з попереднього параграфа схемою: перестановки рядків стають безглуздими, і ми знаходимо найменше спільне кратне чисел третього стовпчика (1, -5, 4): 20. Існує строгий алгоритм знаходження НСК, але тут зазвичай вистачає підбору. Нічого страшного, якщо взяти більше число, яке ділиться і на 1, і на -5, і на 4, наприклад, число 40. Відмінність буде в більш громіздких обчисленнях. Отже:

(4) Третій рядок множимо на 5, другий рядок на 4, перший рядок на «мінус двадцять»:

$$\xrightarrow{(3)}\left(\begin{array}{ccc|ccc}1&5&1&0&0&1\\0&-11&-5&0&1&-2\\0&0&4&1&-1&-1\end{array}\right)\xrightarrow{(4)}\left(\begin{array}{ccc|ccc}-20&-100&-20&0&0&-20\\0&-44&-20&0&4&-8\\0&0&20&5&-5&-5\end{array}\right)\xrightarrow{(5)}\left(\begin{array}{ccc|ccc}-20&-100&0&5&-5&-25\\0&-44&0&5&-1&-13\\0&0&20&5&-5&-5\end{array}\right)\xrightarrow{(6)}$$

$$\xrightarrow{(6)}\left(\begin{array}{ccc|ccc}-4&-20&0&1&-1&-5\\0&44&0&-5&1&13\\0&0&4&1&-1&-1\end{array}\right)\xrightarrow{(7)}\left(\begin{array}{ccc|ccc}-44&-220&0&11&-11&-55\\0&220&0&-25&5&65\\0&0&4&1&-1&-1\end{array}\right)\xrightarrow{(8)}\left(\begin{array}{ccc|ccc}-44&0&0&-14&-6&10\\0&220&0&-25&5&65\\0&0&4&1&-1&-1\end{array}\right)\xrightarrow{(9)}$$

$$\xrightarrow{(9)}\left(\begin{array}{ccc|ccc}44&0&0&14&6&-10\\0&44&0&-5&1&13\\0&0&4&1&-1&-1\end{array}\right)\xrightarrow{(10)}\left(\begin{array}{ccc|ccc}44&0&0&14&6&-10\\0&44&0&-5&1&13\\0&0&44&11&-11&-11\end{array}\right)\xrightarrow{(11)}$$

$$\xrightarrow{(11)}\left(\begin{array}{ccc|ccc}1&0&0&14/44&6/44&-10/44\\0&1&0&-5/44&1/44&13/44\\0&0&1&11/44&-11/44&-11/44\end{array}\right)=(E\,|\,A^{-1}).$$

(5) До 1-го та 2-го рядків додали третій рядок.

(6) Перший і третій рядки розділили на 5, другий рядок помножили на -1.

(7) Найменше спільне кратне ненульових чисел другого стовпчика (-20 і 44) одне – це 220. Перший рядок множимо на 11, другий рядок на 5.

(8) До першого рядка додали другий рядок.

(9) Перший рядок помножили на -1, другий рядок розділили «назад» на 5.

(10) Тепер на головній діагоналі лівої матриці доцільно отримати найменше спільне кратне чисел діагоналі (44, 44 і 4). Цілком зрозуміло, що це число 44. Третій рядок множимо на 11.

(11) Кожен рядок ділимо на 44. Ця дія виконується в останню чергу.

Таким чином, обернена матриця:

$$A^{-1}=\begin{pmatrix}14/44&6/44&-10/44\\-5/44&1/44&13/44\\11/44&-11/44&-11/44\end{pmatrix}=\frac{1}{44}\begin{pmatrix}14&6&-10\\-5&1&13\\11&-11&-11\end{pmatrix}.$$

Внесення і винесення $1/44$, в принципі, зайві дії, але того вимагає протокол оформлення завдання.

Приклад 3.35. Знайти обернену матрицю методом Гауса-Жордана.

$$A=\begin{pmatrix}1 & 2 & 4\\ 5 & 1 & 2\\ 3 & -1 & 1\end{pmatrix}.$$

Розв'язок. Обернену матрицю знайдемо за допомогою елементарних перетворень:

$$(A\,|\,E)=\left(\begin{array}{ccc|ccc}1 & 2 & 4 & 1 & 0 & 0\\ 5 & 1 & 2 & 0 & 1 & 0\\ 3 & -1 & 1 & 0 & 0 & 1\end{array}\right)\overset{(1)}{\to}\left(\begin{array}{ccc|ccc}-15 & -30 & -60 & -15 & 0 & 0\\ 15 & 3 & 6 & 0 & 3 & 0\\ 15 & -5 & 5 & 0 & 0 & 5\end{array}\right)\overset{(2)}{\to}$$

$$\overset{(2)}{\to}\left(\begin{array}{ccc|ccc}-15 & -30 & -60 & -15 & 0 & 0\\ 0 & -27 & -54 & -15 & 3 & 0\\ 0 & -35 & -55 & -15 & 0 & 5\end{array}\right)\overset{(3)}{\to}\left(\begin{array}{ccc|ccc}1 & 2 & 4 & 1 & 0 & 0\\ 0 & 9 & 18 & 5 & -1 & 0\\ 0 & 7 & 11 & 3 & 0 & -1\end{array}\right)\overset{(4)}{\to}$$

$$\overset{(4)}{\to}\left(\begin{array}{ccc|ccc}1 & 2 & 4 & 1 & 0 & 0\\ 0 & 63 & 126 & 35 & -7 & 0\\ 0 & -63 & -99 & -27 & 0 & 9\end{array}\right)\overset{(5)}{\to}\left(\begin{array}{ccc|ccc}1 & 2 & 4 & 1 & 0 & 0\\ 0 & 63 & 126 & 35 & -7 & 0\\ 0 & 0 & 27 & 8 & -7 & 9\end{array}\right)\overset{(6)}{\to}\ldots$$

(1) Перший рядок помножили на -15, другий рядок помножили на 3, третій рядок помножили на 5.

(2) До 2-го і 3-го рядків додали перший рядок.

(3) Перший рядок розділили на -15, другий рядок розділили на -3, третій рядок розділили на -5.

(4) Другий рядок помножили на 7, третій рядок помножили на -9.

(5) До третього рядка додали другий рядок.

$$\overset{(6)}{\ldots}\to\left(\begin{array}{ccc|ccc}1 & 2 & 4 & 1 & 0 & 0\\ 0 & 9 & 18 & 5 & -1 & 0\\ 0 & 0 & 27 & 8 & -7 & 9\end{array}\right)\overset{(7)}{\to}\left(\begin{array}{ccc|ccc}27 & 54 & 108 & 27 & 0 & 0\\ 0 & 54 & 108 & 30 & -6 & 0\\ 0 & 0 & -108 & -32 & 28 & -36\end{array}\right)\overset{(8)}{\to}$$

$$\overset{(8)}{\to}\left(\begin{array}{ccc|ccc}27 & 54 & 0 & -5 & 28 & -36\\ 0 & 54 & 0 & -2 & 22 & -36\\ 0 & 0 & -108 & -32 & 28 & -36\end{array}\right)\overset{(9)}{\to}\left(\begin{array}{ccc|ccc}27 & 0 & 0 & -3 & 6 & 0\\ 0 & 54 & 0 & -2 & 22 & -36\\ 0 & 0 & 27 & 8 & -7 & 9\end{array}\right)\overset{(10)}{\to}$$

$$\overset{(10)}{\to}\left(\begin{array}{ccc|ccc}27 & 0 & 0 & -3 & 6 & 0\\ 0 & 27 & 0 & -1 & 11 & -18\\ 0 & 0 & 27 & 8 & -7 & 9\end{array}\right)\overset{(11)}{\to}\left(\begin{array}{ccc|ccc}1 & 0 & 0 & -\frac{3}{27} & \frac{6}{27} & 0\\ 0 & 1 & 0 & -\frac{1}{27} & \frac{11}{27} & -\frac{18}{27}\\ 0 & 0 & 1 & \frac{8}{27} & -\frac{7}{27} & \frac{9}{27}\end{array}\right)=(E\,|\,A^{-1}).$$

(6) Другий рядок розділили на 7.

(7) Перший рядок помножили на 27, другий рядок помножили на 6, третій рядок помножили на -4.

(8) До першого і другого рядків додали третій рядок.

(9) Третій рядок розділили на -4. До першого рядка додали другий рядок, помножений на -1.

(10) Другий рядок розділили на 2.

(11) Кожен рядок розділили на 27. В результаті отримали:

$$A^{-1}=\begin{pmatrix} -3/27 & 6/27 & 0 \\ -1/27 & 11/27 & -18/27 \\ 8/27 & -7/27 & 9/27 \end{pmatrix}=\frac{1}{27}\begin{pmatrix} -3 & 6 & 0 \\ -1 & 11 & -18 \\ 8 & -7 & 9 \end{pmatrix}.$$

Ця схема, якщо і не сама, то одна з найнадійніших. Іноді буває зручно використовувати коротший «модерністський» розв'язок, який полягає в наступному: на першому кроці все як завжди

$$\left(\begin{array}{ccc|ccc} * & * & * & * & * & * \\ 0 & * & * & * & * & * \\ 0 & * & * & * & * & * \end{array}\right).$$

На другому кроці відомим прийомом (через НСК чисел 2-го стовпця) організовуються відразу два нулі в другому стовпці:

$$\left(\begin{array}{ccc|ccc} * & 0 & * & * & * & * \\ 0 & * & * & * & * & * \\ 0 & 0 & * & * & * & * \end{array}\right).$$

І, нарешті, на третьому етапі точно так отримуємо потрібні нулі в третьому стовпці:

$$\left(\begin{array}{ccc|ccc} * & 0 & 0 & * & * & * \\ 0 & * & 0 & * & * & * \\ 0 & 0 & * & * & * & * \end{array}\right).$$

Приклад 3.36. Знайти загальний розв'язок за допомогою елементарних перетворень

$$\begin{cases} 2x-3y-z-7t+5v=4, \\ x-2y-3z+t-4v=-5, \\ -5x+4y+6z+6t+5v=4. \end{cases}$$

Розв'язок. Запишемо розширену матрицю системи і за допомогою елементарних перетворень отримаємо загальний розв'язок:

$$\left(\begin{array}{ccccc|c} 2 & -3 & -1 & -7 & 5 & 4 \\ 1 & -2 & -3 & 1 & -4 & -5 \\ -5 & 4 & 6 & 7 & 5 & 4 \end{array}\right)\overset{(1)}{\to}\left(\begin{array}{ccccc|c} 1 & -2 & -3 & 1 & -4 & -5 \\ 2 & -3 & -1 & -7 & 5 & 4 \\ -5 & 4 & 6 & 7 & 5 & 4 \end{array}\right)\overset{(2)}{\to}$$

$$\overset{(2)}{\to}\left(\begin{array}{ccccc|c} 1 & -2 & -3 & 1 & -4 & -5 \\ 0 & 1 & 5 & -9 & 13 & 14 \\ 0 & -6 & -9 & -2 & -15 & -21 \end{array}\right)\overset{(3)}{\to}\left(\begin{array}{ccccc|c} 1 & -2 & -3 & 1 & -4 & -5 \\ 0 & 1 & 5 & -9 & 13 & 14 \\ 0 & -2 & -3 & 4 & -5 & -7 \end{array}\right)\overset{(4)}{\to}$$

$$\overset{(4)}{\to}\left(\begin{array}{ccccc|c} 1 & -2 & -3 & 1 & -4 & -5 \\ 0 & 1 & 5 & -9 & 13 & 14 \\ 0 & 0 & 7 & -14 & 21 & 21 \end{array}\right)\overset{(5)}{\to}\left(\begin{array}{ccccc|c} 1 & -2 & -3 & 1 & -4 & -5 \\ 0 & 1 & 5 & -9 & 13 & 14 \\ 0 & 0 & 1 & -2 & 3 & 3 \end{array}\right)\overset{(6)}{\to}$$

$$\xrightarrow{(6)}\left(\begin{array}{ccccc|c}5&-10&-15&5&-20&-25\\0&-3&-15&27&-39&-42\\0&0&15&-30&45&45\end{array}\right)\xrightarrow{(7)}\left(\begin{array}{ccccc|c}5&-10&0&-25&25&20\\0&-3&0&-3&6&3\\0&0&15&-30&45&45\end{array}\right)\xrightarrow{(8)}$$

$$\xrightarrow{(8)}\left(\begin{array}{ccccc|c}1&-2&0&-5&5&4\\0&1&0&1&-2&-1\\0&0&1&-2&3&3\end{array}\right)\xrightarrow{(9)}\left(\begin{array}{ccccc|c}1&-2&0&-5&5&4\\0&2&0&2&-4&-2\\0&0&1&-2&3&3\end{array}\right)\xrightarrow{(10)}$$

$$\xrightarrow{(10)}\left(\begin{array}{ccccc|c}1&0&0&-3&1&2\\0&2&0&2&-4&-2\\0&0&1&-2&3&3\end{array}\right)\xrightarrow{(11)}\left(\begin{array}{ccccc|c}1&0&0&-3&1&2\\0&1&0&1&-2&-1\\0&0&1&-2&3&3\end{array}\right).$$

(1) Перший і другий рядки поміняли місцями.

(2) До другого рядка додали перший рядок, помножений на -2. До третього рядка додали перший рядок, помножений на 5.

(3) Третій рядок розділили на 3.

(4) До третього рядка додали другий рядок, помножений на 2.

(5) Третій рядок розділили на 7.

(6) Найменша кратне чисел 3-го стовпця (-3, 5, 1) дорівнює 15. Перший рядок помножили на 5, другий рядок помножили на -3, третій рядок помножили на 15.

(7) До першого рядка додали 3-ій рядок. До другого рядка додали 3-ій рядок.

(8) Перший рядок розділили на 5, другий рядок розділили на -3, третій рядок розділили на 15.

(9) Найменша кратне ненульових чисел 2-го стовпця (-2 і 1) одне: 2. Другий рядок помножили на 2.

(10) До першого рядка додали другий рядок.

(11) Другий рядок розділили на 2.

Запишемо базисні змінні x, y, z через вільні змінні t, v:

$$\begin{cases}x-3t+v=2\\y+t-2v=-1\\z-2t+3v=3\end{cases}\Rightarrow\begin{cases}x=3t-v+2\\y=-t+2v-1\\z=2t-3v+3\end{cases}$$

Відповідь: загальний розв'язок $(3t-v+2;-t+2v-1;2t-3v+3;t;v)$.

Приклад 3.37. Знайти обернену матрицю за допомогою елементарних перетворень

$$A=\begin{pmatrix}2&3&2&2\\-1&-1&0&-1\\-2&-2&-2&-1\\3&2&2&2\end{pmatrix}.$$

Розв'язок. Знайдемо обернену матрицю методом Гауса-Жордана:

(1) До 1-го та 4-го рядків додали 3-ій рядок.

(2) Перший і четвертий рядки поміняли місцями.

(3) До 2-го рядка додали 1-й рядок. До 3-го рядка додали 1-й рядок, помножений на 2.

(4) До 3-го рядка додали 2-й рядок, помножений на -2. До 4-го рядка додали 2-й рядок.

(5) До 1-го і 3-го рядка додали 4-й рядок, помножений на -1.

(6) Другий рядок помножили на -1, третій рядок розділили на -2.

$$(A|E)=\left(\begin{array}{cccc|cccc} 2 & 3 & 2 & 2 & 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & -1 & 0 & 1 & 0 & 0 \\ -2 & -2 & -2 & -1 & 0 & 0 & 1 & 0 \\ 3 & 2 & 2 & 2 & 0 & 0 & 0 & 1 \end{array}\right) \overset{(1)}{\rightarrow} \left(\begin{array}{cccc|cccc} 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ -1 & -1 & 0 & -1 & 0 & 1 & 0 & 0 \\ -2 & -2 & -2 & -1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{array}\right) \overset{(2)}{\rightarrow}$$

$$\overset{(2)}{\rightarrow}\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ -1 & -1 & 0 & -1 & 0 & 1 & 0 & 0 \\ -2 & -2 & -2 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \end{array}\right) \overset{(3)}{\rightarrow} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & -1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & -2 & -2 & 1 & 0 & 0 & 3 & 2 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \end{array}\right) \overset{(4)}{\rightarrow}$$

$$\overset{(4)}{\rightarrow}\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & -1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & -2 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 2 & 1 \end{array}\right) \overset{(5)}{\rightarrow} \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -1 & -1 & -1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & -2 & 1 & -1 & -3 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 2 & 1 \end{array}\right) \overset{(6)}{\rightarrow} \cdot(-1)$$

$$\overset{(5)}{\rightarrow}\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & -1 & -1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 1 & \frac{1}{2} & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 1 & 1 & 2 & 1 \end{array}\right) = (E|A)$$

3.7. Метод простої ітерації

3.7.1. Аналіз постановки задачі

Для обчислення розв'язку системи лінійних рівнянь великої розмірності прямі методи використовувати недоцільно. Тому застосовуємо наближені методи, які в разі їх збіжності при нескінченній кількості операцій (кроків) дають точний результат. Розглянемо систему

$$A\bar{x} = \bar{b}. \tag{3.132}$$

Метод простої ітерації застосовується для систем типу

$$\bar{x} = \bar{\beta} + \alpha\bar{x}.\ \bar{x} = \bar{\beta} + \alpha\bar{x}\,, \tag{3.133}$$

де $\bar{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $\bar{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$, $\alpha = \begin{pmatrix} \alpha_{11} \dots \alpha_{1n} \\ \alpha_{21} \dots \alpha_{2n} \\ \dots \\ \alpha_{n1} \dots \alpha_{nn} \end{pmatrix}$.

Система (3.132) зводиться до (3.133) так:

1-й спосіб передбачає в кожному i-му рівнянні залишити i-ий елемент в лівій частині та перенесення в праву частину всіх невідомих разом з їх перетвореними коефіцієнтами. Таким чином, співвідношення між коефіцієнтами системи виражаються через такі відношення:

$$\beta_i = \frac{b_i}{a_{ii}},\ a_{ii} \neq 0,\ \alpha_{ij} = -\frac{a_{ij}}{a_{ii}} \tag{3.134}$$

2-й спосіб передбачає розбиття a_{ii} на дві частини $a_{ii} = a_{ii}^{(1)} + a_{ii}^{(2)}$, в якому одну з частин доцільно вибирати у вигляді цілого числа $a_{ii}^{(1)} \in Z$. Тоді поділивши i-те рівняння на a_{ii}, потрібно перенести в праву частину всі невідомі включаючи x_i разом з перетвореними коефіцієнтами $a_{ii}^{(2)}$.

$$\beta_i = \frac{b_i}{a_{ii}^{(1)}},\quad \alpha_{ii} = -\frac{a_{ii}^{(2)}}{a_{ii}^{(1)}},\quad \alpha_{ij} = -\frac{a_{ij}}{a_{ii}^{(1)}} \tag{3.135}$$

Другий спосіб значно ефективніший, так як уникає ділення на нуль і дає змогу зменшити похибку заокруглень при діленні у зв'язку з вибором знаменника у вигляді цілого числа. Беремо $\bar{x}^{(0)} = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_n^{(0)} \end{pmatrix}$ – початкове наближення. Тоді будуємо наступні наближення у такому вигляді:

$$\begin{cases} \bar{x}^{(1)} = \beta + \alpha\bar{x}^{(0)} \\ \bar{x}^{(2)} = \beta + \alpha\bar{x}^{(1)} \\ \quad \dots \\ \bar{x}^{(k+1)} = \beta + \alpha\bar{x}^{(k)} \end{cases}. \tag{3.136}$$

Легко довести, що в разі збіжності це приводить до точного розв'язку $\exists \lim\limits_{k\to\infty} \bar{x}^{(k)} = \bar{x}^*$. Переходячи до граничного випадку будемо мати:

$$\lim_{k\to\infty} \bar{x}^{(k+1)} = \lim_{k\to\infty} \left(\beta + \alpha\bar{x}^{(k)}\right) = \beta + \alpha \lim_{k\to\infty} \bar{x}^{(k)},\quad \bar{x}^* \equiv \beta + \alpha\bar{x}^* \tag{3.137}$$

Запишемо формулу (3.122) у розгорнутій формі:

$$x_i^{(k+1)} = \beta + \sum_{j=1}^{n} \alpha_{ij} x_j^{(k)} \qquad (3.138)$$

Візьмемо $x_j^{(k)}$ рівне відповідному стовпчику вільних членів. Умовою зупинки при виконанні (3.138) можна використати $\forall i : \left| x_i^{(k+1)} - x_i^{(k)} \right| < E$. Щоб вивести достатні умови збіжності методу простої ітерації, введемо певні позначення, що стосуються евклідових норм.

$$\|\alpha\|_m = \max_i \sum_{j=1}^{n} |\alpha_{ij}|; \quad \|\alpha\|_l = \max_l \sum_{i=1}^{n} |\alpha_{ij}|; \quad \|\alpha\|_k = \frac{1}{n} \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij}^2}. \qquad (3.139)$$

Теорема. Якщо одна з евклідових норм матриць $\|\alpha\| \le q < 1$ і довільно вибрати початковий вектор $\bar{x}^{(0)}$, то метод простої ітерації є збіжним $\lim_{n\to\infty} \bar{x}^{(n)} = \bar{x}^*$, причому $\bar{x}^* \equiv \beta + \alpha \bar{x}^*$ і цей розв'язок єдиний.

$\bar{x}^{(k+1)} = \bar{\beta} + \alpha \bar{x}^{(k)}$, $\bar{x}^{(1)} = \bar{\beta} + \alpha \bar{x}^{(0)}$,

$\bar{x}^{(2)} = \bar{\beta} + \alpha \bar{x}^{(1)} = \bar{\beta} + \alpha\left(\bar{\beta} + \alpha \bar{x}^{(0)}\right) = \left(\bar{E} + \bar{\alpha}\right)\bar{\beta} + \alpha^2 \bar{x}^{(0)}$, ...

$$\bar{x}^{(k+1)} = (\bar{E} + \bar{\alpha} + \bar{a}^2 + \ldots + \bar{a}^{(k)})\bar{\beta} + \alpha^{(k+1)} \bar{x}^{(0)} \qquad (3.140)$$

Відомо, що якщо $\|\alpha\| < 1$, то $\lim_{k\to\infty} \alpha^k = 0$. При прямуванні $k \to \infty$ вираз в круглих дужках формули (3.140) являє собою аналог нескінченно спадної геометричної прогресії, яка, як відомо, збігається, так як її знаменник. Тобто $\|\alpha\| < 1$. Тоді

$$E + \alpha + \alpha^2 + \ldots = \frac{E}{E - \alpha} = (E - \alpha)^{-1} \qquad (3.141)$$

Перше твердження є справедливим, тому що границя нескінченно спадної геометричної прогресії існує. Друге твердження про єдність та існування розв'язку випливає при прямуванні відношення (3.141) до нескінченності.

$$\lim_{k\to\infty} \bar{x}^{(k+1)} = (E - \alpha)^{-1} \bar{\beta} + 0 \quad \Rightarrow \bar{x}^* = (E - \alpha)^{-1} \bar{\beta} \quad \left| \cdot E^{-\alpha} \text{ зліва}, \right. \qquad (3.142)$$

$$(E - \alpha)\bar{x}^* = \bar{\beta}, \ \bar{x}^* - \alpha \bar{x}^* = \bar{\beta}, \ \bar{x}^* = \bar{\beta} + \alpha \bar{x}^*.$$

Виходячи з того, що умови теоретично є достатніми, говорити, яким буде метод $\|\alpha\| < 1$ розбіжним чи збіжним є безпідставним. Тому на практиці, крім загальноприйнятої зупинки $\left\| x^{(k+1)} - x^{(k)} \right\| < E$, треба використовувати ще додаткову умову $k > k_{\max}$.

3.7.2. Приклади розв'язку системи методом простої ітерації

Систему лінійних рівнянь, що записана у формі (3.2), необхідно привести до вигляду

$$\mathbf{X} = \mathbf{C} \cdot \mathbf{X} + \mathbf{F}. \tag{3.143}$$

так, щоб виконувались умови збіжності ітераційного процесу

$$\max_i \sum_{j-1}^{n} \left| C_{ij} \right| < 1; \quad i = 1, 2, ..., n \tag{3.144}$$

$$\max_j \sum_{i=1}^{n} \left| C_{ij} \right| < 1; \quad j = 1, 2, ..., n. \tag{3.145}$$

Умови закінчення ітераційного процесу

$$\max \left| x_i^{(k+1)} - x_i^{(k)} \right| \le \varepsilon, \tag{3.146}$$

де ε - задана точність наближення до кореня.

Приклад 3.37. Розв'язати систему лінійних рівнянь методом простої ітерації, прийнявши ε =0,005.

$$\begin{cases} 3x_1 + x_2 - 5x_3 = 10 \\ 4x_1 + x_2 - 2x_3 = 8 \\ x_1 - 5x_2 + x_3 = -10 \end{cases}$$

Розв'язок. Легко помітити, якщо з першого рівняння системи виділити невідоме x_3, з другого – x_1, а з третього – x_2, то умови збіжності (3.135) будуть виконуватись. Спочатку перше рівняння поділимо на 5, друге на 4, а третє на 5.Дійсно,

$$\begin{cases} x_3 = 0{,}6x_1 + 0{,}2x_2 - 2 \\ x_1 = -0{,}25x_2 + 0{,}5x_3 + 2 \\ x_2 = 0{,}2x_1 + 0{,}2x_3 + 2 \end{cases}, \text{ при } F = \begin{pmatrix} -2 \\ 2 \\ 2 \end{pmatrix}.$$

Матриця C системи (3.143) та умови збіжності мають вигляд:

$$C = \begin{bmatrix} 0{,}6 & 0{,}2 & 0 \\ 0 & -0{,}25 & 0{,}5 \\ 0{,}2 & 0 & 0{,}2 \end{bmatrix}; \qquad \max_i \sum_{j-1}^{n} \left| C_{ij} \right| = 0{,}8 < 1.$$

За нульове наближення коренів системи приймаємо вектор $\overline{F}$ (табл. 3.1 при K=0), тобто $x_1^{(0)} = 2; x_2^{(0)} = 2; x_3^{(0)} = -2$. Підставляємо отримані значення $\overline{F}$ у систему рівнянь та отримаємо нові значення $x_1^{(1)} = 0{,}5; x_2^{(2)} = 2; x_3^{(3)} = -0{,}4$:

$$\begin{cases} x_1^{(1)} = -0{,}25 \cdot 2 + 0{,}5 \cdot (-2) + 2 = 0{,}5 \\ x_2^{(1)} = 0{,}2 \cdot 2 + 0{,}2 \cdot (-2) + 2 = 2 \\ x_3^{(1)} = 0{,}6 \cdot 2 + 0{,}2 \cdot 2 - 2 = -0{,}4. \end{cases}$$

Підставляючи отримані дані в початкову умову (систему рівнянь) знаходять друге наближення коренів і так далі (табл. 3.1), доки не буде досягнуто задану точність наближення до кореня ε. Результати розв'язування наведено у табл. 3.1.

Таблиця 3.1

K	x1	x2	x3	$\varepsilon = \max \left\| x_i^{(k+1)} - x_i^{(k)} \right\|$	
0	2	2	-2		
1	0,5	2	-0,4	1,6	> ε
2	1,3	2,02	-1,3	0,9	> ε
3	0,845	2	-0,816	0,484	> ε
4	1,092	2,0058	-1,093	0,277	> ε
5	0,9521	1,9998	-0,9436	0,1494	> ε
6	1,0282	2,0017	-1,0288	0,0852	> ε
7	0,9852	1,9999	-0,9827	0,0461	> ε
8	1,0087	2,0005	-1,0089	0,0262	> ε
9	0,9954	1,9999	-0,9947	0,0142	> ε
10	1,0027	2,0001	-1,0028	0,0081	> ε
11	0,9986	2,0000	-0,9984	0,0044	< ε

Відповідь: $x_1^{(11)} = 0{,}9986 \quad x_2^{(11)} = 2{,}0000 \quad x_3^{(11)} = -0{,}9984$.

3.8. Метод Гауса-Зейделя (метод поліпшеної ітерації)

3.8.1. Аналіз постановки задачі

Метод Зейделя є удосконаленим методом простої ітерації, який полягає в знаходженні розв'язків системи рівнянь та застосовується для систем типу

$$\bar{x} = \overline{\beta} + \alpha\bar{x},$$

$$x_1^{(k+1)} = \beta_1 + \alpha_{11}x_1^{(k)} + \ldots + \alpha_{1n}x_n^{(k)}$$

$$x_2^{(k+1)} = \beta_2 + \alpha_{21}x_1^{(k+1)} + \alpha_{22}x_2^{(k)} + \ldots + \alpha_{2n}x_n^{(k)}$$

$$\ldots$$

$$x_n^{(k+1)} = \beta_n + \alpha_{n1}x_1^{(k+1)} + \alpha_{n2}x_2^{(k)} + \ldots + \alpha_{nn}x_n^{(k+1)}$$

Звідси

$$x_i^{(k+1)} = \beta_i + \sum_{j=1}^{i-1} \alpha_{ij}x_j^{(k+1)} \qquad (3.147)$$

Незважаючи на більш громіздкий запис формули (3.147) у порівнянні з методом простої ітерації, програмна реалізація цієї формули є більш простою. Так як для методу простої ітерації в ПК потрібно формувати два вектори (масиви), які відповідатимуть попередньому і наступному наближенню, а в методі Зейделя достатньо використовувати лише один вектор (матрицю) для наступного наближення. Проте доповнюючи $x_j^{(k+1)}$-тими змінними та аналізуючи перевірки, які пов'язані з умовою закінчення розрахунку методу, вимагатиметься введення нових змінних, які будуть служити як ознака умов закінчення алгоритму. Самі умови зупинки аналогічні до тих, що є в методі простої ітерації. Аналогічне твердження має місце при формулюванні достатніх умов збіжності методу Зейделя.

Виходячи з того, що умови достатні, можливі такі випадки: метод простої ітерації збігається, а метод Зейделя не збігається і навпаки, тобто

$$\left\| x^{(k+1)} - x^{(k)} \right\| < E \quad \Leftrightarrow \quad \forall i : \left| x_i^{(k+1)} - x_i^{(k)} \right| < E_i \tag{3.148}$$

а) $E_i = E_j$

б) $E_i \neq E_j$

Для методу Зейделя умова збіжності може і не виконуватись, якщо матриця α має так звану нормальну форму. То, яка б не була норма, метод Зейделя буде збігатись. Це стверджує одна з теорем алгебри. Матриця A має нормальну форму якщо:

1) вона є симетричною $A = A^T \quad \left(a_{ij} = a_{ji}\right)$;

2) її квадратична форма є достатньо визначеною

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j = a_{11}x_1^2 + \ldots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \ldots > 0\,. \tag{3.149}$$

Твердження. Якщо система $\widetilde{A}\bar{x} = \bar{\widetilde{b}}$ має матрицю $\widetilde{A}$ нормальної форми, то після зведення її до вигляду $\bar{x} = \bar{\beta} + \alpha\bar{x}$ одним з вищеописаних способів, метод Зейделя буде збігатись при довільному виборі початкового наближення.

Залишається показати спосіб, як систему $\widetilde{A}\bar{x} = \bar{\widetilde{b}}$ привести до такого вигляду, щоб $\widetilde{A}$ мала нормальну форму при $\det A \neq 0$, $A^T A\bar{x} = A^T\bar{b}$, $\widetilde{A} = A^T A$, $\bar{\widetilde{b}} = A^T\bar{b}$:

$$\widetilde{a}_{ij} = \sum_{k=1}^{n} a_{ik}^T a_{kj} = \sum_{k=1}^{n} a_{ik}a_{kj}\,, \tag{3.150}$$

$$\widetilde{b}_i = \sum_{k=1}^{n} a_{ki}b_k\,. \tag{3.151}$$

Доведемо, що матриця A, яка обчислена за формулою (3.150), має нормальну форму.

1) $\widetilde{A} = \widetilde{A}^T$ при

$$\widetilde{A}^T = (A^T A) = A^T (A^T)^T = A^T A = \widetilde{A}, \qquad (3.152)$$

2) $\sum_{\substack{i=1 \\ j=1}}^{n} \widetilde{a}_{ij} x_i x_j > 0$.

3.8.2. Приклади розв'язування системи методом Зейделя

Цей метод є модифікацією методу простої ітерації. Іноді він має більшу швидкість збіжності результатів у порівнянні з методом простої ітерації. Суть його полягає в тому, що для розрахунку $(K+1)$-го наближення кореня x_{i+1} одразу використовуються щойно знайдені $(K+1)$ наближення попередніх коренів x_i. Умови збіжності ті ж самі, що і в методі простої ітерації.

Приклад 3.38. Розв'язати систему лінійних рівнянь методом Зейделя, прийнявши ε =0,005.

$$\begin{cases} 3x_1 + x_2 - 5x_3 = 10 \\ 4x_1 + x_2 - 2x_3 = 8 \\ x_1 - 5x_2 + x_3 = -10 \end{cases}$$

Розв'язок. Перетворимо систему до вигляду (3.143). Перевіримо умови збіжності (3.144). За нульове наближення приймемо вектор $\overline{F}$. Тоді отримуємо перші наближення коренів:

$$x_1^{(1)} = -0{,}25 \cdot 2 + 0{,}5 \cdot (-2) + 2 = 0{,}5;$$
$$x_2^{(1)} = 0{,}2 \cdot 0{,}5 + 0{,}2 \cdot (-2) + 2 = 1{,}7;$$
$$x_3^{(i)} = 0{,}6 \cdot 0{,}5 + 0{,}2 \cdot 1{,}7 - 2 = -1{,}36.$$

Отримане значення $x_1^{(1)}$ з першого рівняння підставляємо в друге рівняння та отримаємо значення $x_2^{(1)}$. Так само отримане значення $x_2^{(1)}$ з першого рівняння підставляємо в друге рівняння та отримаємо значення $x_3^{(1)}$. Далі виконаємо друге наближення, третє та ін. Результати розв'язування занесемо у табл. 3.2.

Таблиця 3.2

K	x1	x2	x3	$\varepsilon = \max\left\lvert x_i^{(k+1)} - x_i^{(k)} \right\rvert$	
0	2	2	-2		
1	0,5	1,7	-1,36	1,5	$> \varepsilon$
2	0,895	1,907	-1,0816	0,395	$> \varepsilon$
3	0,9825	1,9802	-1,0145	0,0875	$> \varepsilon$
4	0,9977	1,9966	-1,0020	0,0165	$> \varepsilon$
5	0,9998	1,9996	-1,0002	0,0029	$< \varepsilon$

Відповідь: $x_1^{(5)} = 0{,}9998$; $x_2^{(5)} = 1{,}9996$; $x_3^{(5)} = -1{,}0002$.

3.9. Особливості трьохдіагональної матриці

3.9.1. Обчислення детермінанту трьохдіагональної матриці

Трьохдіагональною матрицею або матрицею Якобі є матриця такого вигляду:

$$A=\begin{pmatrix} a_{11} & a_{12} & 0 & \dots & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & \dots & 0 & 0 & 0 \\ 0 & a_{32} & a_{33} & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-2,n-2} & a_{n-2,n-1} & 0 \\ 0 & 0 & 0 & \dots & a_{n-1,n-2} & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 & a_{n,n-1} & a_{nn} \end{pmatrix}. \qquad (3.153)$$

де у всіх місцях, окрім головної діагоналі та двох сусідніх з нею, стоять нулі. Трьохдіагональні матриці розглядаються принаймні для $n \geq 3$.

Для обчислення детермінанта матриці A використовується формула:

$$\Delta_k = a_{kk}\Delta_{k-1} - a_{k,k-1}a_{k-1,k}\Delta_{k-2}, k=\overline{2,n}, \qquad (3.154)$$

$$\Delta_1 = a_{11}, \quad \Delta_0 = 1, \qquad (3.155)$$

$$\det A = \Delta_n. \qquad (3.156)$$

Виведення цієї формули здійснюється поетапно з розкладом детермінанта матриці A по останній стрічці чи останньому розкладу. Формули вимагають введення певних позначень через $\Delta_{k-1}, \Delta_{k-2}$.

Доведення здійснюється методом математичної індукції:

1) припустимо справедливість формули (3.156) при $n = 2$:
$$\Delta_2 = a_{22}a_{11} - a_{12}a_{21}\cdot 1 = \det A_{2\times 2},$$
2) припустимо справедливість формули (3.156) для $n = m$;
3) доведемо справедливість формули (3.156) для $n = m + 1$.

Доведення формули (3.156) для $n=m+1$ здійснюється за розкладом першої стрічки або першого стовпчика. Вважаємо, що Δ_m, Δ_{m-1} вже обчислені.

Як видно, рекурентна формула (3.154) дає значний ефект при певній кількості затрачених операцій типу множення, віднімання (приблизно $4n$ операцій $<<\frac{2}{3}n^3$). Проте резервувати оперативну пам'ять для матриці (3.154) в традиційних системах програмування все одно необхідно. Покажемо, як змінити формулу (3.154), щоб в оперативній пам'яті резервувати не n^2 елементів, а $3n$ елементів: $3n \approx n+(n-1)+(n-1)$. Для цього матрицю A подамо у вигляді (3.157):

$$A=\begin{pmatrix} s_1 & u_1 & 0 & \dots & 0 & 0 & 0 \\ l_2 & s_2 & u_2 & \dots & 0 & 0 & 0 \\ 0 & l_3 & s_3 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & s_{n-2} & u_{n-2} & 0 \\ 0 & 0 & 0 & \dots & l_{n-1} & s_{n-1} & u_{n-1} \\ 0 & 0 & 0 & \dots & 0 & l_n & s_n \end{pmatrix}. \qquad (3.157)$$

Як бачимо в (3.157) ми використали $3n$-вимірні простори $\bar{l},\bar{s},\bar{u}$. Залишається встановити відповідність між $a[i,j]$ і цими векторами.

$$s_i=a_{ii},\ i=\overline{1,n},\ l_k=a_{k,k-1},\ k=\overline{2,n},\ u_j=a_{j,j-1},\ j=\overline{1,n-1}$$ тобто

$$\left.\begin{array}{l} l_2=a_{21} \\ l_3=a_{31} \\ \dots \\ l_n=a_{n,n-1} \end{array}\right\} \text{ та } \left\{\begin{array}{l} u_1=a_{12} \\ u_2=a_{23} \\ \dots \\ u_{n-1}=a_{n-1,n} \end{array}\right.$$

$$\Delta_k=s_k\Delta_{k-1}-u_{k-1}l_k\Delta_{k-2} \text{ та } \Delta_1=s_1,\ \Delta_0=1,\ k=\overline{2,n}.$$

3.9.2. Метод прогонки для розв'язування трьохдіагональних систем лінійних рівнянь

Якщо матриця системи є розрідженою, тобто містить велику кількість нульових елементів, то в такому випадку застосовують ще одну модифікацію методу Гауса - **метод прогонки**. Розглянемо систему рівнянь такого вигляду

$$A\bar{x}=\bar{b}, \qquad (3.158)$$

$$A=\begin{pmatrix} s_0 & u_0 & 0 & \dots & 0 & 0 & 0 \\ l_1 & s_1 & u_1 & \dots & 0 & 0 & 0 \\ 0 & l_2 & s_2 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & s_{n-2} & u_{n-2} & 0 \\ 0 & 0 & 0 & \dots & l_{n-1} & s_{n-1} & u_{n-1} \\ 0 & 0 & 0 & \dots & 0 & l_n & s_n \end{pmatrix}. \qquad (3.159)$$

Запишемо i-те рівняння системи (3.158) враховуючи подання (3.159)

$$l_ix_{i-1}+s_ix_i+u_ix_{i+1}=b_i,\ i=\overline{0,n},\ l_0=u_n=0. \qquad (3.160)$$

Ідея методу прогонки базується на тому, що в результаті прямого ходу методу Гауса від системи виду (3.158) переходимо до системи, де матриця A буде мати

вигляд верхньої трикутної матриці з одиничною діагоналлю. Це означає, що в i-му рівнянні отримана система буде містити невідомі $x_i, x_{i+1}, \ldots, x_n$. Очевидно, що шляхом лінійних комбінацій ці рівняння можна перетворити так, щоб в кожному i-му з них залишились x_i та x_{i+1} невідомі. Якщо при x_i невідомому стоїть коефіцієнт 1, то при x_{i+1} невідомому стоїть невизначений коефіцієнт.

Припустимо, що обчислення i-ої координати шуканого розв'язку здійснюється через $i+1$ координату за допомогою спеціально введеної формули:

$$x_i = \alpha_{i+1} x_{i+1} + \beta_{i+1}, \quad i = \overline{0, n-1}. \tag{3.161}$$

У формулі (3.161) α_i і β_i є невизначеними коефіцієнтами, які потрібно знайти.

Перепишемо (3.161) замінюючи i на i-1:

$$x_{i-1} = \alpha_i x_i + \beta_i \tag{3.162}$$

та підставимо (3.162) в (3.160) при цьому об'єднаємо коефіцієнти при x з однаковими індексами

$$l_i(\alpha_i x_i + \beta_i) + s_i x_i + u_i x_{i+1} = b_i, \tag{3.163}$$

$$(l_i \alpha_i + s_i) x_i + u_i x_{i+1} = b_i - l_i \beta_i. \tag{3.164}$$

Зводимо останнє співвідношення до вигляду (3.161), в результаті отримаємо:

$$x_i = -\frac{u_i}{l_i \alpha_i + s_i} x_{i+1} + \frac{b_i - l_i \beta_i}{l_i \alpha_i + s_i}. \tag{3.165}$$

Порівнюючи коефіцієнти в останній нерівності з α_i і β_i

$$\alpha_{i+1} = -\frac{u_i}{l_i \alpha_i + s_i}, \quad \beta_{i+1} = \frac{b_i - l_i \beta_i}{l_i \alpha_i + s_i}, \quad i = \overline{0, n-1}. \tag{3.166}$$

Для організації виведення формули (3.165) необхідні значення α_0 і β_0, які можуть бути знайдені виходячи з запису першого рівняння

$$l_0 x_0 + u_0 x_1 = b_0, \tag{3.167}$$

$$x_0 = -\frac{u_0}{l_0} x_1 + \frac{b_0}{l_0} \Rightarrow \alpha_0 = -\frac{u_0}{l_0}, \beta_0 = \frac{b_0}{l_0}. \tag{3.167}$$

Таким чином здійснили процес аналогічний прямому ходу методу Гауса. Для реалізації зворотної прогонки використовується формула (3.161), для якої треба знайти останній елемент x_n. Для цього записуємо останнє рівняння системи

$$l_n x_{n-1} + u_n x_n = b_n, \tag{3.168}$$

$$x_{n-1} = \alpha_n x_n + \beta_n. \tag{3.169}$$

Підставляємо останнє співвідношення (3.169) в (3.168) та отримаємо

$$x_n = \frac{b_n - l_n\beta_n}{l_n\alpha_n + u_n}. \tag{3.170}$$

Не будемо проводити окремо аналіз можливості виникнення ділення на нуль, проте перевірку такої ситуації в програмі доцільно проводити.

3.10. Аналіз числових методів розв'язання СЛАР

3.10.1. Загальна характеристика методів розв'язування систем лінійних алгебраїчних рівнянь

Нехай задана система лінійних алгебраїчних рівнянь(СЛАР)

$$Ax = b, \tag{3.171}$$

де $x \in R^n$; $A \equiv \{a_{ij}\}$; $i, j = \overline{1,n}$; $a_{ij} \in R^1$; $b \equiv (b_1, b_2, ..., b_n)^T \in R^n$.

Якщо $\Delta \equiv \det A$, тоді

$$x_i = \frac{\Delta_{x_i}}{\Delta}. \tag{3.172}$$

Однак, при великих n формули (3.172) практично не використовуються. Труднощі:

1) Відомо що

$$\Delta \equiv \sum (-1)^{I(a_1, a_2, ..., a_n)} a_{1a_1}, a_{2a_2}, ..., a_{na_n}, \tag{3.173}$$

де $I(a_1, a_2, ..., a_n)$ – кількість інверсій множини $(1,2,...,n)$, $(a_1, a_2, ..., a_n)$ – перестановки. Тоді при n=30 кількість операцій на комп'ютері для обчислення одного визначника Δ дорівнює N=30!, а для обчислення 31 визначника на комп'ютері з 1 млрд.опер/сек необхідно $\approx 10^{16}$ років (час існування всієї Галактики).

2) Дуже впливає обчислювальна похибка при заокругленні на кінцевий результат. Так, для n=30 у 32 бітовому поданні числа при обчисленні одного доданку у (3.173) можуть бути аварійні зупинки ($|a_{ia_i}| > 1$).

В залежності від порядку системи (3.171) методи розв'язування поділяються на:

1) Точні $(n \le 10^4)$: метод Гауса, метод квадратного кореня, метод ортогоналізації тощо;
2) Ітераційні $(n \le 10^7)$: метод простої ітерації (МПІ), метод Зейделя (МЗ), метод градієнтного спуску тощо;
3) Ймовірнісні методи $(n > 10^7)$.

Точні методи розв'язання СЛАР розглядалися в курсі вищої алгебри.

3.10.2. Збіжність метричної геометричної прогресії

Нехай задано матричний ряд

$$E + A + A^2 + \ldots + A^n + \ldots, \tag{3.174}$$

де $E = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 1 \end{pmatrix}$; $A \equiv \{a_{ij}\}$; $i, j = \overline{1,n}$; $a_{ij} \in R$.

Працюємо у лінійному нормованому просторі векторів $x \in R^n$ з нормою

$$\|x\|_p = \left(\sum_{j=1}^{n} |x_j|^p \right)^{1/p}, 1 \le p \le \infty. \tag{3.175}$$

При різних $p \ge 1$ в силу скінченності простору всі ці норми еквівалентні.

Розглянемо 3 норми:

1) Чебишовська для $p = \infty$

$$\|x_\infty\| = \max_{1 \le j \le n} |x_j|; \tag{3.176}$$

2) Сферична для $p = 2$

$$\|x\|_2 = \left(\sum_{j=1}^{n} |x_j|^2 \right)^{1/2} = \sqrt{(x,x)}; \tag{3.177}$$

3) Октаедрична для $p = 1$

$$\|x\|_1 = \sum_{j=1}^{n} |x_j|. \tag{3.178}$$

Визначимо норму матриці A, як норму лінійного оператора:

$$\|A\|_p = \sup_{\|x\|=1} \frac{\|Ax\|_p}{\|x\|_p}. \tag{3.179}$$

Якщо А – симетрична, тоді $\|A\|_2 = \max_{1 \le i \le n} |\lambda_i(A)|$.

Говорять, що норми $\|X\|$ та $\|A\|$ узгоджені, якщо $|\lambda| \le \|A\|$.

Зауваження **3.10.1.** Якщо $Ax = \lambda x$, то є узгодженість норм, бо $\|A\|\|X\| \ge \|\lambda x\| = |\lambda|\|x\|$, тобто $|\lambda| \le \|A\|$.

***Зауваження* 3.10.2.** Нижче всі твердження мають місце для всіх трьох норм. Будемо писати $\|\circ\|$, бо норми еквівалентні.

Говорять, що послідовність векторів $\{x^k \equiv (x_1^k, x_2^k, ..., x_n^k)\}$ збігається до вектора $x^k \equiv (x_1, x_2, ..., x_n)$ за нормою, якщо $\|x^k - x\| \to 0 \ (k \to \infty)$. Ця збіжність не залежить від розмірності простору (скінченного або зліченного). Але якщо простір R^n має розмірність n, то збіжність за нормою та збіжність покоординатна рівносильні $([8], c.31)$.

Лема 3.10.1. $\lim\limits_{m\to\infty} A^m = 0$ тоді і тільки тоді, коли

$$|\lambda_i(A)| < 1, i = 1, 2, ..., n,. \tag{3.180}$$

Доведення. Відомо, що $\forall A$ з $a_{ij} \in R$ існує $C(\det C \neq 0)$ така, що вихідна матриця A має представлення [12]

$$A = CIC^{-1}. \tag{3.181}$$

$$I = \left[I_{t_1}^{\lambda_1}, I_{t_2}^{\lambda_2}, ..., I_{t_r}^{\lambda_r} \right]. \tag{3.182}$$

діагональна матриця з клітинами Жордана порядку $t = 1, 2, ..., r, a\lambda_1, \lambda_2, ..., \lambda_r$ – власні значення матриці A; r – кількість лінійно-незалежних власних векторів матриці A. Клітина Жордана має вигляд:

$$I_{t_i}(\lambda_i) \equiv \begin{pmatrix} \lambda_i & 0 & 0 & ... & 0 & 0 \\ 1 & \lambda_i & 0 & ... & 0 & 0 \\ 0 & 1 & \lambda_i & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 1 & \lambda_i \end{pmatrix}, i = \overline{1, r}. \tag{3.183}$$

Якщо $\dim A = n$, тоді $n = \sum\limits_{i=1}^{r} t_i(\lambda_i)$. Отримаємо A^m з (3.181):

$$A^m = CIC^{-1}CIC^{-1}...CIC^{-1} = CI^mC^{-1}. \tag{3.184}$$

Тоді зрозумілий ланцюжок еквівалентних тверджень

$$\left[\lim_{m\to\infty} A^m = 0 \right] \Leftrightarrow \left[\lim_{m\to\infty} I^m = 0 \right] \Leftrightarrow \left[\lim_{m\to\infty} I_{t_i}^m(\lambda_i) = 0, t = t_{1,...,t_2} \right]. \tag{3.185}$$

Обчислимо степінь клітини Жордана

$$I_{i_1}^m(\lambda_i)=\begin{Vmatrix} \lambda_i^m & 0 & 0 & \dots & 0 & 0 \\ \frac{1}{1!}\left(\lambda_i^m\right)^{(1)} & \lambda_i^m & 0 & \dots & 0 & 0 \\ \frac{1}{2!}\left(\lambda_i^m\right)^{(2)} & \frac{1}{1!}\left(\lambda_i^m\right)^{(1)} & \lambda_i^m & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{1}{(t-1)!}\left(\lambda_i^m\right)^{(t-1)} & \frac{1}{(t-2)!}\left(\lambda_i^m\right)^{(t-2)} & \frac{1}{(t-3)!}\left(\lambda_i^m\right)^{(t-3)} & \dots & \frac{1}{2!}\left(\lambda_i^m\right)^{(2)} & \frac{1}{1!}\left(\lambda_i^m\right)^{(1)} \end{Vmatrix}. \quad (3.186)$$

А це подання разом з твердженням дають ланцюжок

$$\left[A^m \to 0\right] \Leftrightarrow \left[\lim_{m\to\infty} I_{i_1}^m(\lambda_i)=0\right] \Leftrightarrow \left|\lambda_i(A)\right|<1,$$ що і доводить лему 1.

Лема 3.10.2. Якщо $\|A\|<1$, то

$$\lim_{m\to\infty} A^m = 0. \quad (3.187)$$

Доведення. Це випливає з очевидних рівностей $\left\|A^m-0\right\| = \left\|A^m\right\| = \|A\|^m$.

Лема 3.10.3. Для будь-якої дійсної матриці A маємо співвідношення між власними значеннями і нормою:

$$\left|\lambda_i(A)\right| \le \|A\|, i=1,2,\dots,n. \quad (3.188)$$

Доведення. $\forall \varepsilon > 0$ побудуємо матрицю

$$B=\frac{A}{\|A\|+\varepsilon}. \quad (3.189)$$

Звідси обчислюємо будь-яку з трьох норм

$$\|B\| \equiv \frac{\|A\|}{\|A\|+\varepsilon} < 1. \quad (3.190)$$

Власне значення $\lambda(A)$ визначається з характеристичного рівняння за означенням

$$\det(\lambda \cdot E - A) = 0. \quad (3.191)$$

Доведемо, що

$$\lambda_i(B) = \frac{\lambda_i(A)}{\|A\|+\varepsilon}, i=1,2,\dots,n. \quad (3.192)$$

Дійсно, $\lambda(B)$ визначається за означенням (3.191) з характеристичного рівняння для матриці B: $\det\left(\frac{\lambda(A)}{\|A\|+\varepsilon}\cdot E-\frac{A}{\|A\|+\varepsilon}\right)=\left(\frac{1}{\|A\|+\varepsilon}\right)^n\det\left(E\lambda(A)-A\right)=0$. Тобто $\lambda(B)$, що обчислене за формулою (3.192), є за означенням (3.191) власним значенням матриці B. Тоді за лемою 3.10.2 з $\|B\|<1$ матимемо $\lim\limits_{m\to\infty}B^m=0$. Лема 3.10.1 дає можливість стверджувати, що $|\lambda_i(B)|<1$, а за означенням матриці B (3.189) це означає $\frac{|\lambda(A)|}{\|A\|+\varepsilon}<1$, тобто $|\lambda_i(A)|<\|A\|+\varepsilon$. Але з довільності вибору $\varepsilon>0$ матимемо $|\lambda_i(A)|\le\|A\|$.

Теорема 3.10.1. Для збіжності матричної геометричної прогресії (3.174) необхідно і досить, щоб

$$|\lambda_i(A)|<1,\ i=1,2,...,n, \tag{3.193}$$

причому $\exists!(E-A)^{-1}$ для $(E-A)$ та

$$\sum_{m=0}^{\infty}A^m=(E-A)^{-1}, \tag{3.194}$$

де $A^0=E$.

Доведення. Необхідність. Нехай матрична геометрична прогресія збігається, тобто загальний член $\lim\limits_{m\to\infty}A^m=0$. Тоді за лемою 1 маємо $|\lambda_i(A)|<1,\ i=1,2,...,n$.

Достатність. Нехай $|\lambda_i(A)|<1$. Доведемо (3.194). Спочатку покажемо, що $\exists(E-A)^{-1}$, і якщо $\det(E-A)\neq0$. Доведемо спочатку, що $\lambda(E-A)=1-\lambda(A)$. Нехай $\lambda(A)$ – власне значення A, тобто $\lambda(A)$ задовольняє $\det\left(E\cdot\lambda(A)-A\right)=0$.

Тоді $\det\left[\left(1-\lambda(A)\right)\cdot E-(E-A)\right]=\det\left(E-E\cdot\lambda(A)-E+A\right)=$

$=\det\left(A-E\cdot\lambda(A)\right)=(-1)^n\det\left(\lambda(A)E-A\right)=0$. Тобто $1-\lambda(A)$ – власне значення для матриці $E-A$. Але $|\lambda(A)|<1$, тоді $|\lambda(A)|\neq1$, що визначає $\det\left(E\cdot1-A\right)\neq0$, так як тільки при $|\lambda(A)|<1$ маємо $\det\left[\lambda(A)\cdot E-A\right]=0$. Значить $\exists!(E-A)^{-1}$ за означенням.

Доведемо (3.194). Очевидна матрична рівність: $(E+A+A^2+...+A^m)(E-A)=E-A^{m+1}$. Помножимо справа цю рівність та

$(E-A)^{-1}$. Тоді одержимо $E+A+...+A^m=(E-A^{m+1})(E-A)^{-1}=E(E-A)^{-1}-A^{m+1}(E-A)^{-1}$. За лемою 1 загальний член матричної геометричної прогресії $\lim\limits_{m\to\infty}A^m=0$, бо $|\lambda(A)|<1$ за припущенням достатності. Таким чином $\sum\limits_{m=0}^{\infty}A^m=(E-A)^{-1}$.

Наслідок 3.10.1. Нехай $\|A\|<1$, тоді очевидно, що $\left\|(E-A)^{-1}-(E+A+...+A^m)\right\|\le\frac{\|A\|^{m+1}}{1-\|A\|}$.

Далі розглянемо ітераційний процес знаходження розв'язку СЛАР (3.171).

3.10.3. Збіжність методу простої ітерації

Нехай задана СЛАР

$$Ax=b, \tag{3.195}$$

де $x\in R$, $A\equiv\left\{a_{ij}\in R, i,j=\overline{1,n}\right\}$. Одержимо еквівалентну СЛАР

$$x=\alpha x+\beta. \tag{3.196}$$

Лема 3.10.4. Якщо $A=C+D$, $\det C\neq 0$, то (3.195) еквівалентна (3.196) з

$$\alpha=D^{-1}(D-A);\beta=D^{-1}\cdot b. \tag{3.197}$$

Доведення. Доведемо, що з (3.196) випливає (3.195). Дійсно, підставимо (3.197) у (3.196), причому врахуємо $D=A-C$ і $x=-C^{-1}Dx+C^{-1}b=-C^{-1}(A-C)x+C^{-1}b$. Звідки $x=-C^{-1}Ax+C^{-1}Cx+C^{-1}b$, що еквівалентне $-C^{-1}(Ax-b)=0$, а значить еквівалентне (3.195).

Метод простої ітерації (МПІ) можна побудувати наступним чином

$$x^{(k+1)}=\alpha x^{(k)}+\beta;\ k=0,1,2,...;\ x^0=\beta. \tag{3.198}$$

Нехай x^* розв'язок СЛАР (3.195) (або(3.196)), тобто $Ax^*=b\ (x^*=\alpha x+\beta)$.

Теорема 3.10.2. Для збіжності МПІ (3.198) для довільної початкової ітерації $x^{(0)}\in R$:

$$\lim_{k\to\infty}x^{(k)}=x^*. \tag{3.199}$$

Необхідно і досить, щоб

$$|\lambda(\alpha)|<1. \tag{3.200}$$

Доведення. Необхідність з $\lim\limits_{k\to\infty} x^{(k)} = x^*$ треба довести (3.200). Очевидні рівності: $x^* - x^{(k)} = \left(\alpha x^* + \beta\right) - \left(\alpha x^{(k-1)} + \beta\right) = \alpha\left(x^* - x^{(k-1)}\right) =$

$= \alpha^2\left(x^* - x^{(k-2)}\right) = \ldots = a^k\left(x^* - x^{(0)}\right)$, тобто $x^* - x^{(k)} = \alpha^k\left(x^* - x^{(0)}\right)$. Якщо перейти до границі зліва і справа при $k \to \infty$, то одержимо: $0 = \lim\limits_{k\to\infty}\alpha^k\left(x^* - x^{(0)}\right)$. Але $x^* - x^{(0)} \neq 0$. Значить $\lim\limits_{k\to\infty}\alpha^k = 0$. Тоді за лемою 1 маємо $\left|\lambda(\alpha)\right| < 1$.

<u>Достатність.</u> Нехай $\left|\lambda(\alpha)\right| < 1$, доведемо (3.199). Одержимо k-ту ітерацію через $x^{(0)}$:

$$x^{(k)} = \alpha x^{(k-1)} + \beta = \alpha\left(\alpha x^{(k-2)} + \beta\right) + \beta = \alpha^2 x^{(k-2)} + \alpha\beta + \beta =$$

$$= \alpha^2 x^{(k-2)} + \left(E + \alpha\right)\beta = \alpha^2\left(\alpha x^{(k-3)} + \beta\right) + \left(E + \alpha\right)\beta =$$

$$= \alpha^3 x^{(k-3)} + \left(E + \alpha + \alpha^2\right)\beta = \ldots = \alpha^k x^{(0)} + \left(E + \alpha + \alpha^2 + \ldots + \alpha^{k-1}\right)\beta.$$

За лемою 1 з $\left|\lambda(\alpha)\right| < 1$ випливає $\lim\limits_{k\to\infty}\alpha^k = 0$.

Тоді $\lim\limits_{k\to\infty} k^{(k)} = \lim\limits_{k\to\infty}\alpha^k x^{(0)} + \lim\limits_{k\to\infty}\left[E + \alpha + \alpha^2 + \ldots + \alpha^{k-1}\right]\beta = \left(E - \alpha\right)^{-1}\beta \equiv x^*$, для $\forall x^{(0)} \in R^n$. Знайшли формулу для обчислення точного розв'язку $x^k = (E - \alpha)^{-1}\beta$.

Зауважимо, що проблема знаходження СЛАР звелася до проблеми знаходження оберненої матриці.

<u>Теорема 3.10.3.</u> Якщо $\|\alpha\| < 1$, то $\lim\limits_{k\to\infty} x^{(k)} = x^*$.

Доведення. За лемою 3 маємо $\left|\lambda(\alpha)\right| \le \|\alpha\| < 1$. Тоді теорема 2 стверджує, що $\lim\limits_{k\to\infty} k^{(k)} = x^*, \forall x^{(0)} \in R^n$.

<u>Наслідок 3.10.2.</u> Достатніми умовами збіжності МПІ є умови:

$$\sum_{j=1}^{n}\left|\alpha_{ij}\right| < 1; i = \overline{1,n} \text{ або } \sum_{i=1}^{n}\left|\alpha_{ij}\right| < 1; j = \overline{1,n}. \qquad (3.201)$$

Це випливає з означення норм матриць.

<u>Наслідок 3.10.3.</u> Якщо норма a узгоджена з нормою x, тобто $\|a \cdot x\| \le \|a\| \cdot \|x\|$ та $\|a\| \le 1$, тоді МПІ для $\forall x^{(0)} \in R^n$ задається нерівністю

$$\left\|r^{(k)}\right\| \equiv \left\|x^{(k)} - x^*\right\| \le \left\|\alpha^k\right\| \cdot \left\|x^{(0)}\right\| + \frac{\|\alpha\|^k}{1 - \|\alpha\|}.$$

Доведення. Дійсно, маємо $x^* = \lim\limits_{k\to\infty} x^k = \left(E - \alpha\right)^{-1}\beta = \left(E + \alpha + \ldots + \alpha^k + \ldots\right)\beta$

та з достатньої умови теореми 2 k-та ітерація має вигляд: $x^{(k)} = \alpha^k x^{(0)} + (E + \alpha + \ldots + \alpha^{k-1})\beta$. Тоді $\left\|x^* - x^{(k)}\right\| = \left\|\left(\alpha^k + \alpha^{k+1} + \ldots\right)\beta - \alpha^k x^{(0)}\right\| \le$

$$\le \left(\|\alpha\|^k + \|\alpha\|^{k+1} + \ldots\right)\|\beta\| + \|\alpha\|^k \left\|x^{(0)}\right\| \le \|\alpha\|^k \left\|x^{(0)}\right\| + + \|\beta\| \cdot \|\alpha\|^k / \left(1 - \|\alpha\|\right).$$

Наслідок 3.10.4. Якщо $x^{(0)} = \beta$, тоді $\left\|x^* - x^{(k)}\right\| \le \dfrac{\|\alpha\|^{k+1}\|\beta\|}{1 - \|\alpha\|}$. Це випливає з доведення теореми 3.10.2.

Наслідок 3.10.5. Похибка іноді контролюється очевидною нерівністю $\left\|r^{(k)}\right\| \equiv \left\|x^* - x^{(k)}\right\| \le \|\alpha\| \cdot \left\|x^* - x^{(k-1)}\right\|$. МПІ можна модифікувати.

3.10.4. Збіжність методу Зейделя розв'язування СЛАР

Метод простої ітерації (3.198) запишемо покоординатно

$$x_i^{(k+1)} = \sum_{j=1}^{n} a_i x_j^{(k)} + \beta_i = \sum_{j=1}^{i-1} a_i x_j^{(k)} + \sum_{j=i}^{n} a_i x_j^{(k)} + \beta_i. \tag{3.202}$$

Тоді методом Зейделя (МЗ) називається наступний ітераційний процес

$$x_i^{(k+1)} = \sum_{j=1}^{i-1} a_i x_j^{(k)} + \sum_{j=i}^{n} a_i x_j^{(k)} + \beta_i, i = \overline{1,n}. \tag{3.203}$$

Методом Зайделя отримаємо наближення

$$(x_1^{(k+1)}, x_2^{(k+1)}, \ldots, x_i^{(k+1)}, x_{i+1}^{(k)}, \ldots, x_n^{(k)}), \tag{3.204}$$

якщо відомо

$$x_1^{(k+1)}, x_2^{(k+1)}, \ldots, x_i^{(k+1)}, x_{i-1}^{(k+1)}, x_i^{(k)}, x_{i+1}^{(k)}, \ldots, x_n^{(k)}. \tag{3.205}$$

Лема 3.10.5. Метод Зейделя (3.202) еквівалентний МЛІ (3.203) з матрицею

$$a^* = [E - H]^{-1} F, \tag{3.206}$$

$$\begin{aligned} H &\equiv \{\alpha_{ij} \mid h_{ij} = 0, i \le j \text{для } i, j = \overline{1,n}\}, \\ F &\equiv \{\alpha_{ij} \mid h_{ij} = 0, i > j \text{для } i, j = \overline{1,n}\}. \end{aligned} \tag{3.207}$$

Доведення. Зрозуміло, що з врахуванням виразу H і F матриця α має представлення $\alpha = H + F$. Тоді МЗ (3.203) еквівалентний $x^{(k+1)} = Hx^{(k+1)} + Fx^{(k)} + \beta$, тобто $[E - H]x^{(k+1)} = Fx^{(k)} + \beta$. Але $\det[E - H] = 1 \ne 0$, значить існує $(E - H)^{-1}$. Тоді домножимо цю систему на $[E - H]^{-1}$, в результаті одержимо

$$[E-H]^{-1}[E-H]x^{(k+1)} = x^{(k+1)} = [E-H]^{-1}Fx^{(k)} + (E-H)^{-1}\beta, \qquad (3.208)$$

що і доводить (3.206).

***Зауваження* 3.10.3.** *Геометричний зміст МЗ.* Метод простої ітерації покоординатно збігається $x_i^{(k+1)} = \sum_{j=1}^{n} \alpha_{ij} x_j^{(k)} + \beta_i$.

Позначимо через $h = \sum_{j=1}^{n} \alpha_{ij} x_j^{(k)} - b_i = 0, \quad i = 1, 2, ..., n,$ – гіперплощину.

При отриманні наближення (3.204) з наближення (3.205) проходить переміщення наближення (3.205) паралельно до вісі Ox_i до перетину з площиною L_{i+1}. На рис. 3.4-3.6 проілюстровано випадки для n=2, коли МЗ збігається, розбігається, зациклюється. Порівняння рис.3.4 та рис.3.5 показує, що збіжність МЗ може змінюватися при перестановці рівнянь.

***Теорема* 3.10.4.** Метод Зейделя збігається $\lim_{n\to\infty} x^{(n)} = x^*$ тоді і тільки тоді, коли $|\tilde{\lambda}| < 1$, *де* $\tilde{\lambda}$ – характеристичні корені рівняння

$$\det[F + \lambda(H - E)] = 0. \qquad (3.209)$$

Доведення. За теоремою 2 нерівність $|\lambda(\alpha)| < 1$ еквівалентна збіжності $\lim_{n\to\infty} x^{(n)} = x^*$. Якщо застосувати теорему 2 до МЗ (3.203) з $a^* = (E-H)^{-1}F$, то власне значення $|\lambda((E-H)^{-1}F)| < 1$.

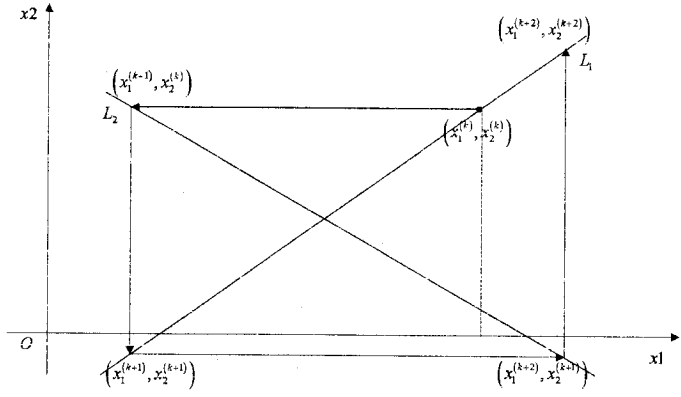


Рис. 3.4. Випадок, коли МЗ збігається

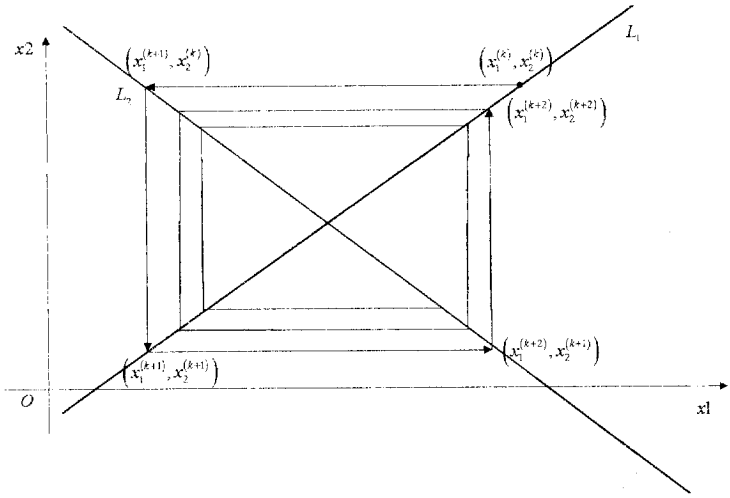


Рис. 3.5. Випадок, коли МЗ розбігається

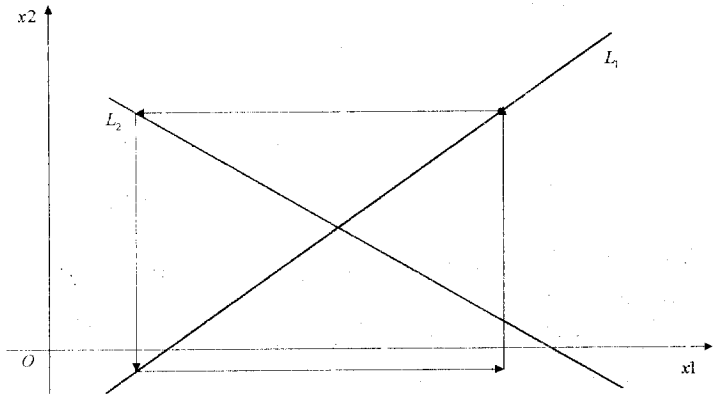


Рис. 3.6. Випадок, коли МЗ зациклюється

Власні значення матриці $(E-H)^{-1}F$ визначаються з характеристичного рівняння $\det[\lambda E-(E-H)^{-1}F]=0$, яке еквівалентне $(-1)^n\det((E-H)^{-1}F-\lambda E)=0$. Помножимо цю рівність на $\det(E-H)\neq 0$. За властивостями визначників маємо:

$$\det(E-h)\det((E-H)^{-1}F-\lambda E)=$$
$$=\det\left[(E-H)(E-H)^{-1}F-\lambda(E-H)E\right]=0,$$

що еквівалентне $\det(F+\lambda(H-E))=0$.

***Зауваження* 3.10.4.** МПІ не є еквівалентний МЗ.

Дійсно, з одного боку:

$$F+\lambda(H-E)=\begin{pmatrix}\alpha_{11}-\lambda & \lambda\alpha_{12} & \dots & \lambda\alpha_{1,n-1} & \lambda\alpha_{1n}\\ \lambda\alpha_{21} & \alpha_{22}-\lambda & \dots & \lambda\alpha_{2,n-1} & \lambda\alpha_{2n}\\ \dots & \dots & \dots & \dots & \dots\\ \lambda\alpha_{n1} & \lambda\alpha_{n2} & \dots & \lambda\alpha_{n,n-1} & \alpha_{nn}-\lambda\end{pmatrix},$$

а з іншого боку

$$\alpha-\lambda E=\begin{pmatrix}\alpha_{11}-\lambda & \alpha_{12} & \dots & \alpha_{1n}\\ \alpha_{21} & \alpha_{22}-\lambda & \dots & \alpha_{2n}\\ \dots & \dots & \dots & \dots\\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn}-\lambda\end{pmatrix}.$$

Таким чином, $\det(\alpha-\lambda E)\neq\det(F+\lambda(H-E))$, а це і доводить твердження, бо корені характеристичних рівнянь не співпадають (характеристичні рівняння різні).

Наслідок 3.10.6. Якщо $\left\|(E-H)^{-1}F\right\|<1$, то МЗ збігається $\lim\limits_{k\to\infty}x^{(k+1)}=x^*$.

Доведення. За лемою 3 випливає: $\left|\lambda((E-H)^{-1}F\right|\le\left\|(E-H)^{-1}F\right\|<1$, тоді $\left|\lambda((E-H)^{-1}F)\right|<1$. Значить, за теоремою 3 матимемо збіжність $\lim\limits_{k\to\infty}x^{(k+1)}=x^*$.

Але на практиці умови незручні, бо треба шукати $(E-H)^{-1}$.

3.10.5. Збіжність метода Зейделя за елементами матриці

Доведемо, що збіжність МЗ можна характеризувати елементами основної матриці СЛАР $Ax=B$, якщо матриця A є матрицею з домінуючими елементами.

Означення 3.1. *Діагональні елементи матриці* $A=(a_{ij})_{i,j=1}^{n}$ *домінують за рядками (стовпчиками), якщо*

$$\sum_{\substack{j=1\\ j\neq i}}^{n}\left|a_{ij}\right|<\left|a_{ii}\right|,\quad i=1,2,\dots,n; \tag{3.210}$$

$$\sum_{\substack{j=1\\ j\neq i}}^{n}\left|a_{ij}\right|<\left|a_{jj}\right|,\quad i=1,2,\dots,n. \tag{3.211}$$

Лема 3.10.6. Якщо діагональні елементи $A=(a_{ij})$ домінують за рядками (стовпчиками), тоді $\det A\neq 0$.

Доведення еквівалентне твердженню про СЛАР Ax=0 і має тривіальний розв'язок $x\equiv(0,0,\dots,0,)^T$.

Від супротивного: Нехай $\det A\neq 0$, але система (3.210)-(3.211) має нетривіальний розв'язок

$$y = (y_1, y_2, ..., y_n)^T \neq 0. \tag{3.212}$$

Серед координат $y_1, ..., y_n$ візьмемо максимальну за модулем:

$$|y_i| \geq |y_j|, \; j = 1, 2, ..., i-1, i+1, ..., n. \tag{3.213}$$

Підставимо (3.212) у систему (3.210)-(3.211) та запишемо i-те рівняння:

$$a_{i1}y_1 + ... + a_{ii}y_i + ... + a_{in}y_n = 0. \tag{3.214}$$

З іншого боку

$$|a_{i1}y_1 + ... + a_{ii}y_i + ... + a_{in}y_n| = \left| a_{ii}y_i - \left(-\sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}y_j \right) \right| \geq \tag{3.215}$$

$$\geq |a_{ii}y_i| - \left| \sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}y_j \right| \geq |a_{ii}y_i| - \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}||y_j| \geq |y_j| \left(|a_{ii}| - \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \right) > 0.$$

що є протиріччям (3.214).

Теорема 3.10.5. *Для* збіжності МЗ (3.203) достатньо, щоб

$$\max_{1\leq i\leq n} \left(\sum_{j=1}^{n} |\alpha_{ij}| \right) \equiv \|\alpha\| < 1, \tag{3.216}$$

$$\max_{1\leq j\leq n} \left(\sum_{i=1}^{n} |\alpha_{ij}| \right) \equiv \|\alpha\| < 1. \tag{3.217}$$

Доведення. З (3.217) випливає нерівність

$$\sum_{j=1}^{n} |\alpha_{ij}| < 1, \quad i = 1, 2, ..., n.. \tag{3.218}$$

Доведемо, що $\tilde{\lambda}$, для якого $|\tilde{\lambda}| \geq 1$, не є коренем рівняння

$$\det(F + \lambda(H - E)) = 0. \tag{3.219}$$

Доведемо від супротивного. Нехай $\tilde{\lambda}$ таке, що $|\tilde{\lambda}| \geq 1$ є коренем рівняння (3.219): $\det(F + \tilde{\lambda}(H - E)) = 0$. Розглянемо суму елементів i-го рядка без i-го елемента :

$$|\alpha_{11}\tilde{\lambda} + \alpha_{12}\tilde{\lambda} + ... + \alpha_{i,i-1}\tilde{\lambda} + \alpha_{i,i+1} + ... + \alpha_{in}| \leq$$

$$\le |\alpha_{11}||\tilde{\lambda}| + |\alpha_{12}||\tilde{\lambda}| + \ldots + |\alpha_{i,i-1}||\tilde{\lambda}| + |\alpha_{i,i+1}| + \ldots + |\alpha_{in}| \le |\tilde{\lambda}| \sum_{\substack{j=1 \\ j\neq i}}^{n} |\alpha_{ij}| =$$

$$= |\tilde{\lambda}| \left(\sum_{\substack{j=1 \\ j\neq i}}^{n} |\alpha_{ij}| - |\alpha_{ii}| \right) < |\tilde{\lambda}| \left(1 - |\alpha_{ii}|\right) = |\tilde{\lambda}| - |\tilde{\lambda}||\alpha_{ii}| \le |\tilde{\lambda}| - |\alpha_{ii}| \le$$

$$\le |\tilde{\lambda} - \alpha_{ii}| = |\alpha_{ii} - \tilde{\lambda}|, \quad i = 1, 2, \ldots, n.$$

Тоді елементи матриці $F + \lambda(E - H)$ мають домінуючі елементи за рядками. За лемою 5 $(F + \lambda(H - E)) \neq 0$ при $|\tilde{\lambda}| < 1$. А тоді за теоремою 5 МЗ збігається. Доведення для (3.217) аналогічне. Встановимо похибку МЗ.

Теорема 3.10.6. Якщо для вихідної матриці A СЛАР (3.203)

$$\sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| < |\alpha_{ii}|, \quad i = 1, 2, \ldots, n, \tag{3.220}$$

то МЗ збігається швидше ніж МПІ.

Доведення. Розглядаємо СЛАР $Ax=b$, яка еквівалентна $x = \alpha x + \beta$, де $\alpha_{ij} = \dfrac{-a_{ij}}{a_{ii}}$; $\dfrac{b_i}{a_{ii}} = \beta_i$; $\alpha_{ii} = 0$, $i \neq j$. Значить, $\|\alpha\| < 1$. Покоординатно розглянемо МЗ:

$$x_i^{(k)} = \sum_{j=1}^{i-1} \alpha_{ij} x_j^{(k)} + \sum_{j=i}^{n} \alpha_{ij} x_j^{(k-1)} + \beta_i, \quad i = \overline{1,n}, \tag{3.221}$$

а точний розв'язок x^* задовольняє СЛАР, (3.171) покоординатно

$$x^* = \sum_{j=1}^{i-1} \alpha_{ij} x_j^* + \beta_i. \tag{3.222}$$

Тоді очевидні покоординатні нерівності:

$$|x_i^* - x_i^{(k)}| \le \sum_{j=1}^{i-1} |a_{ij}| \cdot |x_i^* - x_i^{(k)}| + \sum_{j=i-1}^{n} |a_{ij}| \cdot |x_i^* - x_i^{(k-1)}|, \tag{3.223}$$

що еквівалентне нерівності

$$\|x^* - x^{(k)}\| \le \max_i \frac{\omega_i}{1 - \gamma_i} \|x^* - x^{(k-1)}\| \le \ldots \le (M^*) \|x^* - x^{(0)}\|, \tag{3.224}$$

де $M^* \equiv \max\limits_{1 \le i \le n} \dfrac{\omega_i}{1 - \gamma_i}$; $\omega_i \equiv \sum\limits_{j=i}^{n} |\alpha_{ij}|$; $\gamma_i \equiv \sum\limits_{j=1}^{i-1} |\alpha_{ij}|$.

Для МПІ легко одержати

$$\left\|x^* - x^{(k)}\right\| \le M\left\|x^* - x^{(k-1)}\right\|, \tag{3.225}$$

що дає $\left\|x^* - x^{(k)}\right\| \le M^k \left\|x^* - x^{(0)}\right\|$, $\forall x^{(0)} \in \mathbf{R}^n$, де $M \equiv \max\limits_{1\le i\le n} \sum\limits_{j=i}^{n} \left|\alpha_{ij}\right| < 1$. Зрозуміло, що $\sum\limits_{j=i}^{n} \left|\alpha_{ij}\right| = \gamma_i + \omega_i \le M < 1$. Знайдемо різницю $\gamma_i + \omega_i - \dfrac{\omega_i}{1-\gamma_i} = \dfrac{\gamma_i(1-\gamma_i-\omega_i)}{1-\gamma_i} > 0$.

Значить $\max\limits_i (\gamma_i + \omega_i) > \max \dfrac{\omega_i}{1-\gamma_i}$, тобто, $M > M^*$. Це означає, що МЗ збігається швидше за МПІ у випадках домінування елементів матриці A.

Існують інші ітераційні методи розв'язання СЛАР [1]-[8].

3.11. Методи розв'язання повної проблеми власних значень і власних векторів

3.11.1. Наближене знаходження власних значень матриці

Сукупність власних значень дійсної квадратної матриці називають її *спектром*. Тому відповідна задача називається *спектральною задачею*.

$$A\bar{x} = \bar{b}\,, \tag{3.226}$$

$$A\bar{x} = 0\,. \tag{3.227}$$

Існування та єдиність розв'язку (3.226) еквівалентні існуванню та єдиності тривіального або нульового розв'язку (3.226). Проте відомо, що (3.227) може мати при $\det A = 0$ безліч ненульових розв'язків. Тому в алгебрі є актуальною задача знаходження власних значень матриці. Власним значенням є $\lambda \ne 0$, $\bar{x} \ne 0$ у

$$A\bar{x} = \lambda\bar{x}\,, \tag{3.228}$$

$$(A - \lambda E)\bar{x} = \bar{0}\,, \tag{3.229}$$

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots \\ a_{21} & a_{22}-\lambda & \dots \\ & \dots\dots\dots & \\ a_{n1} & \dots\dots\dots & a_{nn}-\lambda \end{vmatrix} = 0\,. \tag{3.230}$$

Формула (3.229) зводиться до розв'язування (3.230), в якій ліва частина називається **характеристичним визначником**. З нього отримуємо рівняння, розв'язавши яке, маємо n власних значень. Для $n = 2$:

$$D(\lambda) = \begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = (a_{11}-\lambda)(a_{22}-\lambda) - a_{12}a_{21} = 0\,,$$

$$\lambda^2 - \lambda(a_{11} + a_{22}) - a_{12}a_{21} + a_{11}a_{22} = 0.$$

Позначимо $\sigma_1 = a_{11} + a_{22}$, $\sigma_2 = a_{11}a_{22} - a_{12}a_{21}$, $\sigma_3 = \det A$. Тоді характеристичне рівняння набуде вигляду $\lambda^2 - \lambda\sigma_1 + \sigma_2 = 0$. Проводячи аналогію для випадків $n = 3,\ldots$ можна побачити, що розкриваючи характеристичний визначник, отримаємо поліном n-го степеня, в якому при λ^n буде стояти $(-1)^n$.

$$D(\lambda) \equiv (-1)^n \left[\lambda^n - \sigma_1\lambda^{n-1} + \sigma_2\lambda^{n-2} - \ldots + (-1)^n\sigma_n\right] = 0 \qquad (3.231)$$

Досліджено, що σ_1 являє собою суму всіх діагональних елементів

$$\sigma_1 = \sum_{i}^{n} a_{ii}, \qquad (3.232)$$

$$\sigma_2 = \sum_{\substack{i,j=1 \\ i<j}}^{n} \begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}, \qquad (3.233)$$

$$\sigma_3 = \sum_{i<j<k}^{n} \begin{vmatrix} a_{ii} & a_{ij} & a_{ik} \\ a_{ji} & a_{jj} & a_{jk} \\ a_{ki} & a_{kj} & a_{kk} \end{vmatrix}, \qquad (3.234)$$

$$\sigma_n = \det A. \qquad (3.235)$$

Можна підрахувати кількість всіх діагональних мінорів

$$M = C_n^1 + C_n^2 + C_n^3 + \ldots C_n^n = 2^n - 1, \qquad (3.236)$$

$$\sum_{i=0}^{n} C_n^i = 2^n. \qquad (3.237)$$

При фіксованому n програмна реалізація формул (3.232)-(3.235) не є складною. Якщо ж n змінна величина, то в програмній реалізації стикаємося з організацією сум вигляду

$$\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_n=1}^{n}, \qquad (3.238)$$

де кількість цих сум наперед невідома. Це означає, що не будемо знати скільки операторів циклу треба записати при довільному значенні n. Розглянутий метод називається **прямим методом розгортання характеристичного визначника** за допомогою діагональних мінорів. **Дві матриці називаються подібними**, якщо існує така матриця S, яка є не виродженою, тобто її $\det S \neq 0$, тоді $B = S^{-1}AS$.

Відомо, що дві подібні матриці мають один і той же спектр, тобто однакові власні значення.

$$\det\left(B-\lambda E\right)=\det\left(S^{-1}AS-\lambda E\right), \tag{3.239}$$

$$\det B=\det\left(S^{-1}AS\right)=\det S^{-1}\cdot\det A\cdot\det S=\det A, \tag{3.240}$$

$$c=\det S \quad\Rightarrow\quad \det S^{-1}=\frac{1}{c}=\det A. \tag{3.241}$$

В деяких методах з використанням перетворення подібності матриці зводять до такого вигляду, що розкривати її характеристичний визначник не є складним. Інші ж методи тим же перетворенням подібності зводять до діагонального або трикутного вигляду, при цьому використавши відомий факт: спектром або власними значеннями матриці є самі діагональні елементи.

$$D=\begin{pmatrix} d_{11}\ldots 0 \\ \ddots \\ 0\ldots\ d_{nn}\end{pmatrix}, \tag{3.242}$$

$$\det\left(D-\lambda E\right)=\left(d_{11}-\lambda\right)\left(d_{22}-\lambda\right)\ldots\left(d_{nn}-\lambda\right). \tag{3.243}$$

3.11.2. Ідея методу Данилевського

Ідея методу Данилевського передбачає зведення заданої матриці A до так званого нормального вигляду *Фробеніуса.* Деяка матриця P називається *матрицею Фробеніуса,* якщо вона має вигляд

$$P=\begin{pmatrix} p_1 & p_2 & p_3 & \ldots & p_{n-1} & p_n \\ 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 & 0 \end{pmatrix}. \tag{3.244}$$

Перетворення подібності використовується тому, що подібні матриці мають одинакові спектри, а зведення до нормального вигляду проводять, бо характерний многочлен легко розкривається.

$$\det(P-\lambda E)=\begin{vmatrix} p_1-\lambda & p_2 & p_3 & \dots & p_{n-1} & p_n \\ 1 & -\lambda & 0 & \dots & 0 & 0 \\ 0 & 1 & -\lambda & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & -\lambda \end{vmatrix}. \tag{3.245}$$

Розгортаючи визначник по першій стрічці, отримаємо

$$P(\lambda)=(p_1-\lambda)(-1)^{n-1}(-\lambda)^{n-1}-p_2(-1)^{n-2}(-\lambda)^{n-2}+ \\ +p_3(-1)^{n-3}(-\lambda)^{n-3}+\dots+p_n=(-1)^n\left[\lambda^n-p_1\lambda^{n-1}-p_2\lambda^{n-2}-\dots-p_n\right]=0. \tag{3.246}$$

Коефіцієнтами характеристичного визначника для *матриці Фробеніуса* є самі коефіцієнти її ж першої стрічки зі знаками плюс чи мінус, в залежності від їх парного чи непарного індексу. Розв'яжемо рівняння

$$D(\lambda)=0. \tag{3.247}$$

Матрицю A потрібно (n-1) перетвореннями подібності звести до вигляду *матриці Фробеніуса*. Для цього на першому кроці лінійними перетвореннями приходять до такої матриці C, в якої остання стрічка має такий же вигляд як і в матриці P. Це означає, що нам потрібно підібрати деяку матрицю M_{n-1}, таку що $\det M_{n-1}\neq 0$. Тоді буде мати місце формула

$$C=M_{n-1}^{-1}AM_{n-1}. \tag{3.248}$$

Запишемо проходження першого кроку:

1) якщо $a_{n,n-1}\neq 0$, то (n-1)-й стовпчик матриці A ділимо на цей елемент;

2) для того, щоб в цій матриці A в останній стрічці за винятком (n-1)-ї позиції всюди були нулі необхідно від j-го стовпчика ($j=\overline{1,n-2}$) відняти (n-1)-й стовпчик помножений на елемент a_{nj}. Тоді отримаємо матрицю $\widetilde{A}$, в якій остання стрічка прийме бажаний вигляд. Як видно в процесі перетворень виконували лише операції ділення на константу та віднімання стрічок, що являють собою лінійні перетворення матриці. В теорії матриць доводиться, що будь-які лінійні перетворення еквівалентні таким же перетворенням, які проводяться над одиничною матрицею і наступним її застосуванням у формулі (3.248). Маємо одиничну матрицю:

$$E=\begin{pmatrix}1&0&0&\dots&0\\0&1&0&\dots&0\\0&0&1&\dots&0\\\dots&\dots&\dots&\dots&\dots\\0&0&0&\dots&1\end{pmatrix},M_{n-1}=\begin{pmatrix}1&\dots&0&0\\\dots&\dots&\dots&\dots\\m_{n-1,1}&\dots&m_{n-1,n-1}&m_{n-1,n}\\0&\dots&0&1\end{pmatrix}, \tag{3.249}$$

$$m_{n-1,j}=-\frac{a_{nj}}{a_{n,n-1}}, \tag{3.250}$$

$$m_{n-1,n-1}=\frac{1}{a_{n,n-1}}. \tag{3.251}$$

За формулами (3.250) і (3.251) визначають елементи матриці (3.249):

$$\det M_{n-1}\neq 0\Rightarrow\exists M_{n-1}^{-1}, \tag{3.252}$$

$$M_{n-1}^{-1}=\begin{pmatrix}1&\dots&0&0\\\dots&\dots&\dots&\dots\\a_{n,1}&\dots&a_{n,n-1}&a_{n,n}\\0&\dots&0&1\end{pmatrix}, \tag{3.253}$$

$$M_{n-1}^{-1}M_{n-1}=E. \tag{3.254}$$

Зробивши перше перетворення подібності, друге перетворення подібності вже здійснюється над матрицею C, множенням її зліва на M_{n-2}^{-1} і справа – на M_{n-2}.

$$\widetilde{C}=M_{n-2}^{-1}CM_{n-2}, \tag{3.255}$$

$$\widetilde{C}=M_{n-2}^{-1}M_{n-1}^{-1}AM_{n-1}M_{n-2}, \tag{3.256}$$

$$C^{(n-1)}=M_1^{-1}M_2^{-1}\dots M_{n-2}^{-1}M_{n-1}^{-1}AM_{n-1}M_{n-2}\dots M_2M_1, \tag{3.257}$$

$$C^{(n-1)}\approx A,\ A^{(i+1)}=M_{n-i}^{-1}A^{(i)}M_{n-i}. \tag{3.258}$$

Повторюючи цей процес (n-1) разів, отримаємо матрицю, яка буде подібною до матриці A, тобто їх спектри будуть співпадати. Залишилося розкрити характеристичний визначник матриці $C^{(n-1)}$, яка має вигляд Фробеніуса та знайти корені відповідного характеристичного рівняння. Отже, на першому кроці за формулами (3.250) та (3.251) обчислюються елементи матриці M_{n-1}.

3.11.3. Обчислення власних векторів методом Данилевського

Розглянемо систему рівнянь такого вигляду

$$P\bar{y} = \lambda\bar{y} \tag{3.259}$$

при $\lambda \neq 0$, $\bar{y} \neq \bar{0}$. Відомо, що одному власному значенню може відповідати скінченна або нескінченна сукупність власних векторів. З них можна виділити *нормований* вектор – це вектор, довжина якого дорівнює 1. Нехай одне з λ знайдене методом Данилевського, тоді рівняння (3.259) можна записати у такому вигляді:

$$P = \begin{pmatrix} p_1 & p_2 & p_3 & \dots & p_{n-1} & p_n \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}.$$

$$(P - \lambda E)\bar{y} = \bar{0}. \tag{3.260}$$

$$\begin{cases} (p_1 - \lambda)y_1 + p_2 y_2 + \dots + p_n y_n = 0 \\ y_1 - \lambda y_2 + \dots = 0 \\ \dots \\ y_{n-2} - \lambda y_{n-1} = 0 \\ y_{n-1} - \lambda y_n = 0 \end{cases} \tag{3.261}$$

Система (3.261) є однорідною, і тому може мати лише нульовий розв`язок або безліч ненульових розв'язків. До того, щоб серед ненульових розв'язків знайти один необхідний, потрібно одне з невідомих оголосити вільним, тобто зафіксувати і перенести в праву частину. Тоді отримаємо (n-1) рівняння з (n-1)-им невідомим. При відомих умовах така система буде мати лише єдиний ненульовий розв`язок. Як видно, в останній рівності є лише два невідомі, а тому вибираючи $y_n = 1$ отримуємо, що $y_{n-1} = \lambda$.

$$\begin{cases} y_n = 1 \\ y_{n-1} = \lambda \\ y_{n-2} = \lambda^2 \\ \dots \\ y_2 = \lambda^{n-2} \\ y_1 = \lambda^{n-1} \end{cases} \tag{3.262}$$

3.11.4. Метод невизначених коефіцієнтів

Розглянемо систему рівнянь такого вигляду

$$A\overline{x} = \lambda\overline{x}\,, \tag{3.263}$$

де $\overline{x} \neq \overline{0}$ при

$$\det(A - E\lambda) = 0\,. \tag{3.264}$$

Подамо характеристичний визначник матриці A у вигляді:

$$D(\lambda) = \lambda^n + p_1\lambda^{n-1} + p_2\lambda^{n-2} + \ldots + p_n = \det(\lambda E - A)\,. \tag{3.265}$$

У формулі (3.265) коефіцієнти $p_1, p_2, \ldots, p_n$ є невизначені. Щоб їх знайти, знаходять значення полінома (3.265) при n різних значеннях λ.

Нехай $\lambda = \overline{0, n-1}$. Підставивши послідовно λ в (3.265) отримаємо

$$p_n = \det(-A) = D(0)\,. \tag{3.266}$$

$$\begin{cases} 1^n + p_1 1^{n-1} + \ldots + p_n 1^0 = D(1) \\ 2^n + p_1 2^{n-1} + \ldots + p_n 2^0 = D(2) \\ \ldots \\ (n-1)^n + p_1(n-1)^{n-1} + \ldots + p_n(n-1)^0 = D(n-1) \end{cases} \tag{3.267}$$

Система (3.267) по відношенню до елементів $p_1, p_2, \ldots, p_n$ є лінійною. Представимо цю систему у вигляді зручному для подальшого застосування в раніше вивченому методі (Крамера, Гауса, Зейделя).

$$\begin{cases} p_1 1^{n-1} + .p_2 1^{n-2} + .. + p_{n-1} 1^1 = D(1) - D(0) - 1^n \\ p_1 2^{n-1} + p_2 2^{n-2} + \ldots + p_{n-1} 2^1 = D(2) - D(0) - 2^n \\ \ldots \\ p_1(n-1)^{n-1} + p_2(n-1)^{n-2} + \ldots + p_{n-1}(n-1)^1 = D(n-1) - D(0) - (n-1)^n \end{cases} \tag{3.268}$$

Подамо систему (3.268) у матричному вигляді:

$$C\overline{p} = \overline{d}\,, \tag{3.269}$$

$$\overline{p} = \begin{pmatrix} p_1 \\ p_2 \\ \ldots \\ p_{n-1} \end{pmatrix}, \tag{3.270}$$

$$\overline{d} = \begin{pmatrix} D(1) - D(0) - 1^n \\ D(2) - D(0) - 2^n \\ \dots \\ D(n-1) - D(0) - (n-1)^n \end{pmatrix}. \tag{3.271}$$

$$C = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 2^{n-1} & 2^{n-2} & \dots & 2 \\ \dots & \dots & \dots & \dots \\ (n-1)^{n-1} & (n-1)^{n-2} & \dots & n-1 \end{pmatrix}. \tag{3.272}$$

Відомо, що ця матриця, яка носить назву *матриці Вандермонда*, має ненульовий визначник, а це означає, що система лінійних рівнянь (3.269) має розв'язок і цей розв'язок буде єдиний. Як бачимо, в матриці C коефіцієнти матриці A не беруть участі, ці коефіцієнти неявно входять у вектор вільних членів. А тому цей метод досить зручно використовувати для великої сукупності різних матриць однієї і тієї ж розмірності. Якщо $\det C \neq 0 \Rightarrow \exists C^{-1}$, тоді

$$\overline{p} = C^{-1}\overline{d}. \tag{3.273}$$

Залишається лише один раз обчислити елементи матриці C, знайти до неї обернену і порахувати n визначників вигляду:

$$D_k = \det(kE - A), \; k = \overline{0, n-1}. \tag{3.274}$$

3.11.5. Загальна постановка задачі методів розв'язання повної проблеми власних значень і власних векторів

Нехай задано дійсну матрицю $A = (a_{ij})_{i,j=1}^{n}; \quad a_{ij} \in \mathbf{R}, \quad \dim A = n.$

Означення 3.2. *Власними значеннями* $\lambda \in \mathbf{C}$ *(комплексної площини) матриці A називаються такі* λ*, при яких система*

$$Ax = \lambda x. \tag{3.275}$$

має тривіальний розв'язок.

Як на практиці знайти власне значення λ?

$$Ax = \lambda x \Leftrightarrow (A - \lambda E)x = 0. \tag{3.276}$$

Система (3.275) має нетривіальний розв'язок $x \neq 0$ тоді і тільки тоді, коли

$$p(\lambda) \equiv \det(A - \lambda E) = 0. \tag{3.277}$$

Означення 3.3. $p(\lambda)$ *називається характеристичним многочленом для матриці A, а рівняння (3.277), яке еквівалентне* $(A - \lambda E) = (-1)^n \det(\lambda E - A) = 0$, *назвемо характеристичним рівнянням для матриці A.*

Якщо розкрити визначник (3.277), то отримаємо алгебраїчне рівняння з точністю до знака:

$$\lambda^n + p_1\lambda^{n-1} + p_2\lambda^{n-2} + \ldots + p_{n-1}\lambda + p_n = 0, \qquad (3.278)$$

де $p_1 = \sum_{i=1}^{n} a_{ii}$, $p_2 = \sum_{i<j} \begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}$, ..., $p_n \det A$. Обчислення коефіцієнтів $p_1, p_2, \ldots, p_n$ еквівалентне обчисленню $C_n^1 + C_n^2 + \ldots + C_n^n = 2^n - 1$ визначників різних порядків. Це технічно важка задача навіть при наявності сучасних інформаційних технологій. Тому існують спеціальні методи знаходження $p_1, p_2, \ldots, p_n$ без обчислення мінорів матриці A, а саме, метод невизначених коефіцієнтів, метод інтерполювання, метод Левер'є*-Фадєєва, метод О.М.Крилова та багато інших [2] – [14], [17], [23]. Розв'язувати алгебраїчні рівняння можемо методом половинного ділення відрізка, де знаходиться корінь; методом січних, комбінованим методом дотичних, методом парабол тощо. Якщо $\{p_1, p_2, \ldots, p_n\} \subset \mathbf{C}$ то за основною теоремою алгебри існує n коренів $\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \subset \mathbf{C}$.

Означення 3.4. *Власними векторами* $\{x^{(i)}, i = \overline{1,n}\} \subset \mathbf{C}^n$, *які відповідають власним значення* $\left\{\lambda_i, i = \overline{1,n}\right\}$, *назвемо розв'язки СЛАР*

$$Ax = \lambda_i x, \quad i = \overline{1,n}. \qquad (3.279)$$

Означення 3.5. *Знаходження всіх власних значень* $\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \subset \mathbf{C}$ *та всіх власних векторів* $\{x^{(i)}, i = \overline{1,n}\} \subset \mathbf{C}^n$ *назвемо повною проблемою власних значень і векторів (ППВЗ). Задача знаходження декількох власних значень матриці назвемо частковою проблемою власних значень (ЧПВЗ).*

Розв'язати ППВЗ – це означає знайти n коренів алгебраїчного рівняння (3.278). ППВЗ еквівалентна розкриттю визначника $\det(\lambda E - A) = 0$, тобто знаходження $p_0 = 1;\ p_1; \ldots;\ p_n$.

3.11.6. Метод невизначених коефіцієнтів розв'язання повної проблеми власних значень і власних векторів

Потрібно від

$$p(\lambda) \equiv \det(\lambda E - A). \qquad (3.280)$$

перейти до

$$p(\lambda) = \lambda^n + p_1\lambda^{n-1} + p_2\lambda^{n-2} + \ldots + p_{n-1}\lambda + p_n, \qquad (3.281)$$

тобто знайти $p_1, p_2, \ldots, p_n$?

Зауважимо, що $p(\lambda)$ визначається для $\forall \lambda \in \mathbf{C}$, бо область визначення функції $p(\lambda)$ є вся площина $\mathbf{C}$. Тоді вона визначена і для $\lambda = 0,1,...,n-1$. Запишемо її значення :

$$p_n = p(0),. \tag{3.282}$$

$$1^n + p_1 1^{n-1} + ... + p_{n-1} 1 + p_n = p(1), \tag{3.283}$$

$$2^n + p_1 2^{n-1} + ... + p_{n-1} 2 + p_n = p(2), \tag{3.284}$$

$$(n-1)^n + p_1 (n-1)^{n-1} + ... + p_{n-1}(n-1) + p_n = p(n-1). \tag{3.285}$$

Тоді СЛАР (3.285) еквівалентна розв'язанню матричного рівняння

$$D_{n-1} p = d, \tag{3.286}$$

де $p = (p_1, p_2, ..., p_{n-1})^T$;

$$d \equiv (p(1) - p(0) - 1; p(2) - p(0) - 2^n; ...; \tag{3.287}$$

$$p(n-1) - p(0) - (n-1)^n)^T ; p_n = p(0) = \det A, \tag{3.288}$$

а $p(t) = \det(iE - A), i = 1, 2, ..., n-1;$

$$D_{n-1} \equiv \begin{pmatrix} 1 & 1 & ... & 1 \\ 2^{n-1} & 2^{n-2} & ... & 2 \\ ... & ... & ... & .. \\ (n-1)^{n-1} & (n-1)^{n-2} & ... & n-1 \end{pmatrix}. \tag{3.289}$$

$D_{n-1} \neq 0$, як *визначник Вандермонда**. Значить існує єдиний розв'язок СЛАР (3.287). Сформулюємо результат у вигляді теореми.

Теорема 3.11.1. Якщо $A \equiv (a_{ij})_{i,j=1}^n \subset \mathbf{C}$ то $p = D_{n-1}^{-1} d$, де p, d, D_{n-1} – визначено(3.288).

Зауваження 5. D_{n-1}^{-1} залежить тільки від порядку $\det D_{n-1}^{-1}$. Отже, можна обчислити D_{n-1}^{-1} на комп'ютері при різних n, якщо маємо справу з масовим розкриттям визначників для матриці D_{n-1} однакового порядку.

3.11.7. Метод інтерполювання розв'язання повної проблеми власних значень

Власні значення $\{\lambda_j, 1 \le j \le n\}$ для A знаходяться як корені рівняння $p(\lambda) = \lambda^n + p_1\lambda^{n-1} + p_2\lambda^{n-2} + ... + p_{n-1}\lambda + p_n = 0$, де $p_1, p_2, ..., p_n$ подаються через елементи матриці A. Теорія інтерполювання має місце і для комплексних функцій. Значить функція $p(\lambda)$ визначена і для $\lambda_0 = 0$; $\lambda_1 = 1$; $\lambda_n = n$. Запишемо $(n+1)$ визначник $p(i) = \det(iE - A), i = 0, 1, ..., n$. Складемо таблицю скінченних різниць та запишемо формулу Ньютона вперед:

$$p \le(\lambda) = p(0) + \frac{p^1_{1/2}}{1!}\lambda + \frac{p^2_1}{2!}\lambda(\lambda-1) + ... + \frac{p^n_{n/2}}{n!}\lambda(\lambda-1) = \qquad (3.290)$$

$$= p(0) + \sum_{i=1}^{n} \frac{p^i_{\frac{i}{2}}}{i!}\lambda(\lambda-1)...(\lambda-i+1),$$

де $R_n(\lambda) = 0$.

Позначимо коефіцієнти $C_{mi}; m = \overline{1,i}; i = \overline{1,n}$, при розкритті добутку у формулі (3.290):

$$\frac{\lambda(\lambda-1)...(\lambda-i+1)}{i!} = \sum_{m=0}^{i} C_{mi}\lambda^m. \qquad (3.291)$$

Тоді має місце твердження:

Теорема 3.11.2. Для $A \equiv (a_{ij})^n_{i,j=1}$, $a_{ij} \subset \mathbf{C}^1$ характеристичний многочлен має вигляд

$$p(\lambda) = p(0) + \sum_{m=1}^{n} p_m \cdot \lambda^m. \qquad (3.292)$$

де $p_m = \sum_{i=m}^{n} C_{mi} p^i_{\frac{i}{2}}, \quad m = 1, 2, ..., n.$

Зауважимо, що метод інтерполювання вимагає $\approx \frac{n^4}{3}$ множень та ділень.

3.12. Приклади знаходження розв'язку систем лінійних алгебраїчних рівнянь

Приклад 3.12.1. Здійснюючи трикутний розклад методом Гауса з вибором ведучого елемента в стовпчику, знайти розв'язок системи $Ax = b$,

$$\text{де } A=\begin{pmatrix}2 & -4 & -1\\ -2 & 3 & -2\\ 4 & -11 & -13\end{pmatrix},\ b=\begin{pmatrix}1\\ 1\\ 10\end{pmatrix}.$$

Розв'язок. Введемо розширену матрицю системи

$$\overline{A}=\left(\begin{array}{ccc|c}2 & -4 & -1 & 1\\ -2 & 3 & -2 & 1\\ 4 & -11 & -13 & 10\end{array}\right).$$

Розв'язування системи запишемо у вигляді перетворення матриці $\overline{A}$ множенням на матриці перестановок рядків P_k та на матриці виключення невідомих M_k. Маємо

$$P_1=\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix};\ P_1\overline{A}=\left(\begin{array}{ccc|c}4 & -11 & -13 & 10\\ -2 & 3 & -2 & 1\\ 2 & -4 & -1 & 1\end{array}\right);$$

$$M_1=\begin{pmatrix}1 & 0 & 0\\ 0{,}5 & 1 & 0\\ -0{,}5 & 0 & 1\end{pmatrix};\ \overline{A_1}=M_1P_1\overline{A}=\left(\begin{array}{ccc|c}4 & -11 & -13 & 10\\ 0 & -2{,}5 & -8{,}5 & 6\\ 0 & 1{,}5 & 5{,}5 & -4\end{array}\right);$$

$$M_2=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0{,}6 & 1\end{pmatrix};\ \overline{A_2}=M_2P_2\overline{A_1}=\left(\begin{array}{ccc|c}4 & -11 & -13 & 10\\ 0 & -2{,}5 & -8{,}5 & 6\\ 0 & 0 & 0{,}4 & -0{,}4\end{array}\right).$$

З останньої трикутної системи отримуємо розв'язок системи:

$$x_3=\frac{-0{,}4}{0{,}4}=-1,\ x_2=\frac{6-8{,}5}{-2{,}5}=1,\ x_1=\frac{10-13+11}{4}=2.$$

Приклад 3.12.2. Обчислити визначник матриці A попередньої задачі.

Розв'язок. Виходячи з трикутного розкладу матриці A, маємо:

$$\det A=(-1)^l\cdot u_{11}\cdot u_{22}\cdot u_{33}=-4\cdot(-2{,}5)\cdot 0{,}4=4,$$

де l – кількість перестановок рядків при перетворенні матриці A.

Приклад 3.12.3. Знайти матрицю обернену до матриці A задачі 3.12.1.

Розв'язок. Шукана матриця $C=A^{-1}$ задовольняє матричне рівняння

$$AC=B=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

Із правими частинами цієї системи виконаємо ті самі перетворення, що і з матрицею A в задачі 3.12.1:

$$P_1B=\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix},\ M_1P_1B=\begin{pmatrix}1&0&1\\0&1&0{,}5\\0&0&-0{,}5\end{pmatrix},\ B_2=M_2P_2M_1P_1B=\begin{pmatrix}0&0&1\\0&1&0{,}5\\0&0{,}6&-0{,}2\end{pmatrix}.$$

Розв'язуючи матричне рівняння $A_2C=B_2$ (тобто три системи з трикутною матрицею A_2) знаходимо

$$C=\begin{pmatrix}-15{,}25&-10{,}25&-2{,}75\\-8{,}5&-5{,}5&-1{,}5\\2{,}5&1{,}5&-0{,}5\end{pmatrix}.$$

Приклад 3.12.4. Методом квадратних коренів знайти розв'язок системи $Ax=b$,

$$\text{де } A=A^*=\begin{pmatrix}4&2&2&2\\2&-3&-1&-1\\2&-1&-4&-2\\2&-1&-2&3\end{pmatrix},\ b=\begin{pmatrix}8\\-6\\-9\\3\end{pmatrix}.$$

Розв'язок. Згідно з формулами методу знаходимо розклад матриці $A=S^*DS$,

$$\text{де } D=\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&1\end{pmatrix},\ S=\begin{pmatrix}2&1&1&1\\0&2&1&1\\0&0&2&1\\0&0&0&2\end{pmatrix},$$

S^* – матриця спряжена (транспонована) до S.

Розв'язуючи систему з нижньою трикутною матрицею $S^*Dy=b$, знаходимо $y_1=4$, $y_2=5$, $y_3=4$, $y_4=4$. Розв'язуючи систему з верхньою трикутною матрицею $Sx=y$, знаходимо $x_4=2$, $x_3=1$, $x_2=1$, $x_1=0$.

Приклад 3.12.5. Обчислити визначник матриці A попередньої задачі.
Розв'язок. Згідно з розкладом матриці A задачі 3.12.4 маємо

$$\det A=\prod_{i=1}^{n}d_{ii}\bullet\prod_{i=1}^{n}s_{ii}^2=256.$$

Приклад 3.12.6. Знайти матрицю обернену до матриці A задачі 3.12.4.
Розв'язок.

$$A^{-1}=\begin{pmatrix}\frac{45}{256}&\frac{13}{128}&\frac{5}{64}&-\frac{1}{32}\\\frac{13}{128}&-\frac{19}{64}&\frac{5}{32}&-\frac{1}{16}\\\frac{5}{64}&\frac{5}{32}&-\frac{3}{15}&-\frac{1}{8}\\-\frac{1}{32}&-\frac{1}{16}&-\frac{1}{8}&\frac{1}{4}\end{pmatrix}.$$

Приклад 3.12.7. Знайти розв'язок системи $Ax = b$,

$$\text{де } A = \begin{pmatrix} 3{,}00 & 1{,}00 & -1{,}00 \\ 6{,}00 & 2{,}01 & 0{,}00 \\ 3{,}00 & 1{,}02 & 3{,}01 \end{pmatrix};\ b = \begin{pmatrix} -2{,}00 \\ -2{,}01 \\ 1{,}99 \end{pmatrix}$$

методом Гауса, виконуючи обчислення з чотирма значущими цифрами.

Розв'язок. Запишемо послідовність перетворень розширеної матриці A системи методом Гауса з вибором ведучого елемента в стовпчику:

$$P_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix};\ P_1\overline{A} = \left(\begin{array}{ccc|c} 6{,}00 & 2{,}01 & 0{,}00 & -2{,}01 \\ 3{,}00 & 1{,}00 & -1{,}00 & -2{,}00 \\ 3{,}00 & 1{,}02 & 3{,}01 & 1{,}99 \end{array}\right);$$

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ -0{,}5 & 1 & 0 \\ -0{,}5 & 0 & 1 \end{pmatrix};\ \overline{A}_1 = M_1\overline{A}P_1 = \left(\begin{array}{ccc|c} 6 & 2{,}01 & 0 & -2{,}01 \\ 0 & -5\cdot10^{-3} & -1 & -0{,}995 \\ 0 & 1{,}5\cdot10^{-2} & 3{,}01 & 2{,}995 \end{array}\right);$$

$$P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix};\ P_2\overline{A} = \left(\begin{array}{ccc|c} 6 & 2{,}01 & 0 & -2{,}01 \\ 0 & 1{,}5\cdot10^{-2} & 3{,}01 & 2{,}995 \\ 0 & -5\cdot10^{-3} & -1 & -0{,}995 \end{array}\right);$$

$$M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0{,}3333 & 1 \end{pmatrix};\ \overline{A}_2 = M_2\overline{A}_1P_2 = \left(\begin{array}{ccc|c} 6 & 2{,}01 & 0 & -2{,}01 \\ 0 & 1{,}5\cdot10^{-2} & 3{,}01 & 2{,}995 \\ 0 & 0 & 3\cdot10^{-3} & 3{,}2\cdot10^{-3} \end{array}\right).$$

Звідси $x_3 = 1{,}067$, $x_2 = -14{,}47$, $x_1 = 4{,}511$, що дуже відрізняється від точного розв'язку $x = (0, -1.1)$.

Приклад 3.12.8. Розв'язати задачу 3.12.7, виконуючи обчислення з п'ятьма значущими цифрами.

Розв'язок. На відміну від попередньої задачі обчислення на останньому кроці дають

$$M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0{,}33333 & 1 \end{pmatrix};\ \overline{A}_2 = \left(\begin{array}{ccc|c} 6 & 2{,}01 & 0 & -2{,}01 \\ 0 & 1{,}5\cdot10^{-2} & 3{,}01 & 2{,}995 \\ 0 & 0 & 3{,}3\cdot10^{-3} & 3{,}3\cdot10^{-3} \end{array}\right).$$

Звідси $x_3 = 1$, $x_2 = -1$, $x_1 = 0$, що збігається з точним розв'язком системи.

Приклад 3.12.9. Визначити число обумовленості матриці A із задачі 3.12.7.

Розв'язок. Використовуючи знайдені в задачі 3.12.7 матриці P_k, M_k маємо

$$A^{-1} = \begin{pmatrix} 2{,}240\cdot10^{3} & -1{,}492\cdot10^{3} & 7{,}447\cdot10^{3} \\ -6{,}687\cdot10^{3} & 4{,}455\cdot10^{3} & -2{,}223\cdot10^{3} \\ 3{,}33\cdot10^{2} & -2{,}2\cdot10^{2} & 1{,}111\cdot10^{2} \end{pmatrix}.$$

Звідси $\|A^{-1}\|_\infty = 1{,}121\cdot 10^4$, $\|A\|_\infty = 8{,}01$, $\mathrm{cond}_\infty(A) = 8{,}979\cdot 10^4$. Таким великим числом обумовленості пояснюється відмінність розв'язку задачі 3.12.7 від точного:

$$c_k = \frac{(k-1)!!}{k!!}\delta_k,\ \delta_k = \begin{cases} \dfrac{\pi}{2}, & k - \text{парне}, \\ 1, & k - \text{непарне} \end{cases}.$$

Приклад 3.12.10. Знайти розв'язок системи $Ax = b$,

$$\text{де } A = \begin{pmatrix} 3 & 1 & -1 \\ 6 & 2{,}1 & 0 \\ 3 & 1{,}15 & 1{,}9999 \end{pmatrix},\ b = \begin{pmatrix} +22{,}01 \\ +2{,}15\cdot 10^{-2} \\ -0{,}8499 \end{pmatrix}.$$

виконуючи обчислення з чотирма значущими цифрами. Пояснити результат.

Розв'язок. В силу припущень, замість системи, що задана, розв'язується система з наближеною матрицею

$$\widetilde{A} = \begin{pmatrix} 3 & 1 & -1 \\ 6 & 2{,}1 & 0 \\ 3 & 1{,}15 & 2{,}000 \end{pmatrix}.$$

Але вона вироджена, оскільки $\det\widetilde{A} = 0$. В той час як $\det A = -3\cdot 10^{-5}$.

Виродженість матриці $\widetilde{A}$ обумовлена похибкою заокруглення елемента a_{33}=1,9999≈2,000. Для того щоб одержати задовільний результат, потрібно збільшити кількість значущих цифр при обчисленнях.

Розв'яжіть цей приклад з п'ятьма значущими цифрами. Точний розв'язок: x=(0.1,–1).

Приклад 3.12.11. Яку похибку в розв'язку системи лінійних алгебраїчних рівнянь $Ax = b$, де

$$A = \begin{pmatrix} 1 & a & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & a \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

спричиняє похибка ε в компоненті b_n вектора $b(b_n = b_n + \varepsilon)$? Які властивості матриці A обумовлюють цю похибку?

Розв'язок. Оскільки розв'язок системи можна знайти за рекурентними формулами $x_n = b_n$, $\widetilde{x}_{n-1} = b_{n-1} - ab_n, \dots$, то не важко знайти, що похибки в x_i є:

$$\widetilde{x}_{n-1} - x_{n-1} = b_{n-1} - ab_n - b_{n-1} + ab_n = -a\varepsilon,$$

$$\widetilde{x}_{n-2} - x_{n-2} = a^2\varepsilon, \dots,\ \widetilde{x}_1 - x_1 = (-1)^{n-1}a^{n-1}\varepsilon.$$

Якщо a>1, то ця похибка швидко зростає з ростом n. Матриця A при цьому є погано обумовленою, що і спричиняє велику похибку розв'язку.

Приклад 3.12.12. Знайти оцінку швидкості збіжності методу простої ітерації

$$y_{k+1} = By_k + b$$

для системи лінійних алгебраїчних рівнянь

$$x = Bx + b,$$

де B – матриця розмірності n, причому відомо, що для деякої норми $\|B\| = \rho < 1$.

Розв'язок. Якщо x^* – розв'язок системи, то $x^* = Bx^* + b$. Тому маємо $x^* - y_k = B(x^* - y_{k-1})$. Звідси $\|x^* - y_k\| \le \rho\|x^* - y_{k-1}\| \le \ldots \le \rho^k\|x^* - y_0\|$. Оскільки $x^* - y_{k-1} = y_k - y_{k-1} + B(x^* - y_{k-1})$, то $\|x^* - y_{k-1}\| \le \|y_k - y_{k-1}\| + \|B\|\|x^* - y_{k-1}\|$ та $\|x^* - y_{k-1}\| \le \frac{1}{1-\rho}\|y_k - y_{k-1}\|$. Остаточно маємо: $\|x^* - y_k\| \le \frac{\rho^k}{1-\rho} \cdot \|y_1 - y_0\|$. Оскільки $\rho < 1$, то для $k \to \infty$ $\|x^* - y_k\| \to 0$, що означає збіжність методу.

Приклад 3.12.13. Дослідити збіжність ітераційного процесу $y_{k+1} = By_k + b$ з матрицями:

$$\text{а) } B = \begin{pmatrix} -3/5 & 3/5 \\ 2/5 & 1/5 \end{pmatrix}; \quad \text{б) } B = \begin{pmatrix} 1/10 & 4/5 \\ 3/5 & -1/5 \end{pmatrix}.$$

Розв'язок. а) $\|B\|_\infty = 6/5 > 1$, $\|B\|_E = \left(\sum_{i,j} b_{ij}^2\right)^{1/2} = \sqrt{23}/5 < 1$, що означає збіжність ітераційного процесу в нормі $\|x\| = \left(\sum_{i=1}^{n} x^2\right)^{1/2}$, яка узгоджена з матричною нормою Шура $\|B\|_E$.

б) $\|B\|_\infty$ =9/11<1, що означає збіжність ітераційного процесу в нормі $\|x\|_\infty$. Одночасно для цієї матриці $\|B\|_E = \sqrt{21/20} > 1$.

Матрична норма Шура називається також нормою Ерхарда– Шмідта або нормою Фробеніуса.

Приклад 3.12.14. Нехай матриця системи лінійних алгебраїчних рівнянь $Ax = b$ має діагональну перевагу, тобто

$$\sum_{\substack{j=1 \\ j \ne i}}^{n} |a_{ij}| \le q|a_{ii}|, i = \overline{1,n}, 0 < q < 1,$$

а також є симетричною і додатньо визначеною. Дати оцінку швидкості збіжності методу Зейделя.

Розв'язок. Для похибки $z_k = y_k - x = \left(z_k^j\right)_{j-1}^n$ маємо

$$a_{ii}z_{k+1}^{(i)} = -\sum_{j<i} a_{ij}z_{k+1}^{(j)} - \sum_{j>i} a_{ij}z_k^{(j)} \text{ та } |a_{ii}|\left|z_{k+1}^{(i)}\right| = \sum_{j<i}|a_{ij}|\left|z_{k+1}^{(j)}\right| + \sum_{j>i}|a_{ij}|\left|z_k^{(j)}\right|.$$

Нехай $\max_i \left|z_k^{(i)}\right| = \left|z_k^{(i_0)}\right| = \|z_k\|_C$. Тоді $|a_{i_0i_0}|\|z_{k+1}\|_C \le \sum_{j<i_0}|a_{i_0j}|\|z_{k+1}\|_C + \sum_{j>i_0}|a_{i_0j}|\|z_k\|_C$,

$$\|z_{k+1}\|_C \le \left[\sum_{j<i_0} |a_{i_0 j}| / |a_{i_0 i_0}| - \sum_{j>i_0} |a_{i_0 j}| \right] \|z_k\|_C .$$

Користуючись умовою діагональної переваги, маємо

$$\sum_{j>i_0} |a_{i_0 j}| \le q|a_{i_0 i_0}| - \sum_{j<i_0} |a_{i_0 j}| < q\left(|a_{i_0 i_0}| - \sum_{j<i_0} |a_{i_0 j}| \right)$$

тобто $\|z_{k+1}\|_C \le q\|z_k\|_C \le \ldots \le q^{k+1}\|z_0\|_C$.

Приклад 3.12.15. Дослідити збіжність методу простої ітерації та методу Зейделя для системи $Ax = b$, де

$$A = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, |\rho| < 1 .$$

Розв'язок. Щоб застосувати метод простої ітерації зведемо систему до вигляду $x=Bx+b$, де $B = I - A = \begin{pmatrix} 0 & \rho \\ \rho & 0 \end{pmatrix}$.

Власні значення матриці B визначаються з рівняння

$$|B - \lambda I| = \begin{pmatrix} -\lambda & \rho \\ \rho & -\lambda \end{pmatrix} = \lambda^2 - \rho^2 = 0 .$$

Звідси $\lambda = \pm\rho, |\lambda| = \rho < 1$. Це означає збіжність методу. Із критерію Сільвестра маємо $A^* = A > 0$, бо $\det(a_{ii}) = 1 > 0$, $\det A = 1 - \rho^2 > 0$. Тому метод Зейделя також збіжний.

3.13. Контрольні питання

1. Дайте визначення поняття «матриця».
2. Які операції визначені над матрицями?
3. Коли матриця є невиродженою?
4. Дайте визначення рангу матриці.
5. В чому суть зворотного ходу методу Гауса?
6. Як визначається алгебраїчне доповнення?
7. Назвіть умову подібності двох матриць.
8. Яка матриця називається матрицею Фробеніуса?
9. Яку систему називають системою лінійних алгебраїчних рівнянь?
10. Що називається розв'язком СЛАР?
11. Яка система називається сумісною і несумісною?
12. Яка система називається визначеною і невизначеною?
13. Яка система називається виродженою і невиродженою?
14. Які системи називаються еквівалентними?
15. Яку СЛАР можна розв'язати на ЕОМ?
16. Які методи відносять до точних (дати означення та перелічити методи)?

17. Які методи відносять до наближених (дати означення та перелічити методи)?
18. В чому суть алгоритмів методу Гауса?
19. В чому суть прямого ходу в методах Гауса?
20. В чому суть зворотного ходу в методах Гауса?
21. Для чого в методі Гауса з послідовним виключенням елементів вводиться множник М і переставляються рівняння?
22. Чим відрізняються алгоритми методів Гауса з послідовним виключенням елементів і вибором головного елементу?
23. Чим відрізняються алгоритми методів Гауса з вибором головного елементу і з одиничною діагоналлю?
24. Чим відрізняються алгоритми методів Гауса з одиничною діагоналлю і Гауса-Жордана?
25. В чому суть методу Гауса за схемою Халецького?
26. Яку систему отримано в результаті прямого ходу методу Гауса-Жордана?

3.14. Задачі для самостійної роботи

Задача 3.14.1. Розв`язати систему за формулами Крамера. Відповідь представити в звичайних неправильних дробах. Зробити перевірку.

$$\begin{cases} 7x + y = 23 \\ -5x + 3y = 1 \end{cases}$$

Задача 3.14.2. Нехай векторна норма $\|x\|_p$ визначається формулою

$$\|x\|_p = \left\{ \sum_{k=1}^{n} |x_k|^p \right\}^{1/p}, \; p \geq 1, \; \|x\|_\infty = \max_{1 \leq k \leq n} |x_k| .$$

Довести, що для узгоджених норм матриці $A = (a_{ik})_{i,k=1}^{n}$

$$\|A\|_p = \sup_{\|x\|_p \neq 0} \frac{\|Ax\|_p}{\|x\|_p}$$

мають місце такі формули:

а) $\|A\|_1 = \max\limits_k \sum\limits_{i=1}^{n} |a_{ik}|$; б) $\|A\|_\infty = \max\limits_i \sum\limits_{k=1}^{n} |a_{ik}|$; в) $\|A\|_2 = \sqrt{\max \lambda\left(A^* A\right)}$,

де $\lambda\left(A^* A\right)$ – власне значення матриці $A^* A$.

Задача 3.14.3. Довести такі нерівності:

а) $\frac{1}{\sqrt{n}} \|A\|_2 \leq \|A\|_1 \leq \sqrt{n} \|A\|_2$; в) $n^{-1} \leq \frac{\mathrm{cond}_\infty A}{\mathrm{cond}_2 A} \leq n$;

б) $\frac{1}{\sqrt{n}} \|A\|_2 \leq \|A\|_\infty \leq \sqrt{n} \|A\|_2$; г) $n^{-1} \leq \frac{\mathrm{cond}_1 A}{\mathrm{cond}_2 A} \leq n$,

де $\mathrm{cond}_k A = \|A\|_k \|A^{-1}\|_k$ – число обумовленості матриці A.

Задача 3.14.4. Довести, що матриця обернена до

$$A = \begin{pmatrix} 1 & a & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & a \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

має вигляд

$$A^{-1} = \begin{pmatrix} 1 & -a & a^2 & -a^3 & \dots & (-a)^{n-1} \\ 0 & 0 & -a & a^2 & \dots & (-a)^{n-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & -a \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

і матриця A має число обумовленості

$$\mathrm{cond}_\infty A = (1+a)\frac{a^{n-1}-1}{a-1}.$$

Підрахувати число обумовленості для n=20 , a=5.

Задача 3.14.5. Дати відповіді на запитання:

а) Як зв'язані між собою число обумовленості матриці та її детермінант?

б) Чи обов'язково погано обумовлена матриця має малі власні значення?

Задача 3.14.6. Довести, що детермінант матриці Коші

$$K_n = \left(\frac{1}{a_i + b_j}\right)_{i,j=1}^{n}$$

дорівнює

$$\det K_n = D_n = \prod_{p=1}^{n-1}\left[\prod_{q=p+1}^{n}(a_p - a_q)\right]\cdot\prod_{r=1}^{n-1}\left[\prod_{s=r+1}^{n}(b_r - b_s)\right]\Bigg/\prod_{i=1}^{n}\left[\prod_{j=1}^{n}(a_i + b_j)\right] =$$

$$= \prod_{1\le i<j\le n}(a_i - a_j)(b_i - b_j)\left(\prod_{1\le i<j\le n}(a_i + b_j)\right)^{-1}.$$

Задача 3.14.7. Знайти елементи матриці $K_n^{-1} = (b_{ij})_{i,j=1}^{n}$, яка є оберненою до матриці Коші

$$K_n = \left(\frac{1}{a_i + b_j}\right)_{i,j=1}^{n}.$$

Задача 3.14.8. Знайти елементи матриці $H_n^{-1} == (a_{ij})_{i,j=1}^{n}$, яка є оберненою до матриці Гільберта

$$H_n = \left(\frac{1}{j+i-1} \right)_{i,j=1}^{n}$$

і показати, що a_{ij} – цілі числа.

Задача 3.14.9. Довести, що (див. задачу 3.14.7)

$$\max|a_{ij}| = \frac{1}{4\sqrt{2}\pi^2 n}\left(\sqrt{2}+1\right)^{4n}\left(1+O\left(n^{-1}\right)\right),$$

$$\left\|H_n^{-1}\right\|_\infty = \frac{1}{(2\pi)^{3/2}2^{7/4}n^{1/2}}\left(\sqrt{2}+1\right)^{4n}\left(1+O\left(n^{-1}\right)\right)$$

і дати оцінку числа обумовленості матриці Гільберта H_n.

Задача 3.14.10. Привести приклади систем $Ax=b$, для яких метод простої ітерації є розбіжним, але збігається метод Зейделя та навпаки.

Задача 3.14.11. Обчислити норми матриці A:

1) $\|A\|_\infty = \max\limits_{1\le i\le n}\left(\sum\limits_{j-1}^{n}|a_{ij}|\right)$;

2) $\|A\|_2 = \sqrt{\max\limits_{1\le i\le n}\lambda_i\left(A^*\cdot A\right)}$, де $\lambda_i(A)$ власні значення матриці A^i, де A^* – матриця спряжена до A;

3) $\|A\|_1 = \max\limits_{1\le i\le n}\left(\sum\limits_{i=1}^{n}|a_{ij}|\right)$.

3.15. Завдання до лабораторної роботи

Перший рівень.

Реалізувати основні операції над матрицями. Матриці брати розмірності $n>3$. Перевірити основні властивості матриць.

Другий рівень.

Реалізувати знаходження визначника матриці для матриць розмірності $n>3$. Знайти обернену матрицю для невироджених матриць розмірності $n>3$.

Третій рівень.

1. Розв'язати СЛАР, використовуючи обернену матрицю. Матриці брати розмірності $n>3$. Здійснити перевірку отриманого розв'язку.
2. Реалізувати метод Гауса розв'язування СЛАР. Матриці брати розмірності $n>3$. Знайти аналітичний розв'язок СЛАР. Здійснити перевірку отриманого розв'язку.

Зміст звіту до лабораторної роботи:

1. Мета роботи
2. Алгоритм розв`язування задачі.
3. Текст програми з коментарями.

4. Результати виконання програми.
5. Висновки.

Варіанти завдань

Таблиця 3.3

№	Матриця A	Вектор b	№	Матриця A	Вектор b
1	$\begin{pmatrix} 2 & 4 & 17 & 0 & -14 \\ -9 & 19 & -14 & -13 & 25 \\ 9 & 21 & 25 & 15 & -2 \\ -4 & -8 & -4 & 20 & -7 \\ 6 & -12 & 22 & 10 & -4 \end{pmatrix}$	$\begin{pmatrix} -8 \\ 7 \\ 0 \\ 13 \\ 17 \end{pmatrix}$	18	$\begin{pmatrix} 25 & -18 & -19 & 22 & -16 \\ -6 & 21 & 19 & -3 & 8 \\ -19 & 22 & -17 & -4 & 18 \\ 12 & -2 & 22 & -10 & 8 \\ -10 & 1 & 19 & 4 & -8 \end{pmatrix}$	$\begin{pmatrix} 14 \\ 6 \\ 13 \\ 19 \\ 16 \end{pmatrix}$
2	$\begin{pmatrix} 15 & 14 & 13 & -7 & -6 \\ 10 & 25 & 9 & 0 & 9 \\ 1 & 19 & 24 & 25 & -5 \\ -16 & 0 & 19 & 16 & -2 \\ -5 & 12 & 20 & 5 & -2 \end{pmatrix}$	$\begin{pmatrix} 8 \\ 7 \\ 3 \\ 13 \\ 8 \end{pmatrix}$	19	$\begin{pmatrix} -2 & -2 & 6 & -9 & 6 \\ 0 & 6 & 20 & -18 & 20 \\ -8 & 7 & -10 & -7 & -9 \\ -1 & -17 & 0 & -6 & 24 \\ 5 & 5 & 15 & -11 & 2 \end{pmatrix}$	$\begin{pmatrix} 18 \\ 11 \\ -8 \\ 20 \\ 16 \end{pmatrix}$
3	$\begin{pmatrix} -4 & 13 & 5 & 19 & 6 \\ -10 & 21 & 1 & -13 & -2 \\ -15 & -20 & 14 & 21 & -12 \\ 21 & -16 & -18 & 19 & 13 \\ -13 & 5 & 19 & 23 & 9 \end{pmatrix}$	$\begin{pmatrix} 16 \\ 9 \\ 2 \\ -3 \\ 16 \end{pmatrix}$	20	$\begin{pmatrix} 2 & -4 & -2 & 8 & -8 \\ 19 & 12 & 15 & 2 & 2 \\ -15 & 12 & 0 & 15 & 7 \\ 7 & 16 & 0 & -4 & -18 \\ -3 & -13 & -15 & 6 & 12 \end{pmatrix}$	$\begin{pmatrix} 12 \\ -5 \\ -2 \\ -1 \\ -10 \end{pmatrix}$
4	$\begin{pmatrix} -6 & -20 & 10 & 10 & -12 \\ 8 & -8 & -18 & -5 & 1 \\ -1 & 22 & 18 & 22 & 15 \\ -18 & -13 & 13 & 24 & -8 \\ -7 & 14 & 17 & -6 & -9 \end{pmatrix}$	$\begin{pmatrix} -8 \\ 13 \\ -1 \\ 7 \\ 8 \end{pmatrix}$	21	$\begin{pmatrix} 7 & 4 & 10 & -7 & -7 \\ -1 & -20 & -5 & -8 & -12 \\ 9 & -1 & 14 & -3 & 3 \\ 1 & 16 & 25 & -19 & 2 \\ 7 & 23 & 12 & 2 & -17 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 12 \\ -9 \\ 1 \\ -4 \end{pmatrix}$
5	$\begin{pmatrix} 11 & 9 & 19 & 4 & -1 \\ -16 & 18 & 13 & 18 & 8 \\ -1 & 15 & -14 & 1 & -13 \\ 2 & -8 & 1 & 3 & 16 \\ 14 & 10 & 7 & 0 & 16 \end{pmatrix}$	$\begin{pmatrix} 16 \\ 1 \\ 9 \\ -6 \\ -1 \end{pmatrix}$	22	$\begin{pmatrix} 1 & 3 & -4 & -14 & -12 \\ 12 & 4 & 22 & 15 & -13 \\ 1 & -3 & 24 & -9 & 15 \\ 23 & 5 & 0 & -4 & 13 \\ -16 & -6 & 3 & 3 & -13 \end{pmatrix}$	$\begin{pmatrix} 18 \\ 8 \\ 5 \\ 19 \\ 5 \end{pmatrix}$
6	$\begin{pmatrix} 1 & 19 & 19 & -12 & 15 \\ -1 & -10 & -8 & -16 & -14 \\ 3 & -3 & 3 & 1 & 3 \\ 9 & 1 & 4 & -20 & -18 \\ 16 & -18 & 11 & -7 & 15 \end{pmatrix}$	$\begin{pmatrix} -5 \\ -8 \\ -5 \\ 16 \\ -4 \end{pmatrix}$	23	$\begin{pmatrix} -14 & 11 & 14 & -19 & 11 \\ 14 & 0 & 6 & 14 & 21 \\ 13 & 2 & -6 & -12 & 6 \\ 1 & 19 & -11 & 18 & 5 \\ 9 & -14 & 19 & 15 & 11 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 7 \\ -8 \\ 7 \\ -3 \end{pmatrix}$
7	$\begin{pmatrix} 19 & -17 & 17 & 10 & -13 \\ -17 & -2 & 15 & -7 & 11 \\ -18 & 11 & -16 & -13 & 6 \\ -2 & 12 & 14 & -9 & 22 \\ 7 & -7 & -6 & -8 & 11 \end{pmatrix}$	$\begin{pmatrix} -1 \\ -2 \\ 19 \\ -3 \\ 9 \end{pmatrix}$	24	$\begin{pmatrix} 8 & -1 & -8 & -8 & -17 \\ -16 & -8 & 24 & 15 & 2 \\ -5 & 1 & 0 & 5 & 4 \\ 5 & -9 & 6 & -18 & -12 \\ -10 & 19 & 15 & 10 & 4 \end{pmatrix}$	$\begin{pmatrix} -9 \\ -10 \\ -1 \\ 0 \\ 3 \end{pmatrix}$

8	$\begin{pmatrix} 7 & 5 & 6 & 3 & 10 \\ 13 & -16 & -16 & -18 & 18 \\ 14 & -17 & 17 & -10 & -20 \\ 20 & 3 & 19 & -1 & -7 \\ 17 & -9 & 21 & 22 & 10 \end{pmatrix}$	$\begin{pmatrix} 12 \\ 9 \\ -10 \\ 17 \\ 14 \end{pmatrix}$	25	$\begin{pmatrix} -18 & -10 & -16 & 21 & -12 \\ 24 & -7 & -13 & -5 & 19 \\ -9 & -11 & -19 & -10 & 8 \\ -3 & -6 & -15 & -16 & 3 \\ 4 & -7 & 20 & -12 & 20 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 2 \\ 15 \\ 10 \\ 18 \end{pmatrix}$
9	$\begin{pmatrix} -6 & 18 & 15 & 8 & 6 \\ 2 & 11 & -18 & 17 & 23 \\ -13 & -2 & 8 & -10 & -5 \\ 1 & 13 & 11 & -10 & -5 \\ 10 & 21 & 8 & -1 & 11 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -8 \\ 1 \\ -2 \\ -6 \end{pmatrix}$	26	$\begin{pmatrix} 20 & -14 & 0 & 1 & -16 \\ -5 & 4 & 14 & -7 & -16 \\ 8 & 15 & 5 & 11 & -3 \\ 10 & 6 & -1 & 13 & -15 \\ -17 & -9 & 8 & 18 & 4 \end{pmatrix}$	$\begin{pmatrix} -4 \\ 17 \\ 15 \\ 16 \\ 0 \end{pmatrix}$
10	$\begin{pmatrix} -10 & 3 & 24 & -9 & 2 \\ -1 & 15 & 5 & 20 & -12 \\ -10 & 7 & -10 & 14 & 16 \\ 22 & -4 & -10 & -6 & -6 \\ 1 & 16 & 7 & 0 & -15 \end{pmatrix}$	$\begin{pmatrix} 9 \\ 12 \\ -9 \\ 19 \\ -1 \end{pmatrix}$	27	$\begin{pmatrix} 22 & -6 & -9 & -19 & 22 \\ -20 & 2 & 2 & -6 & 11 \\ 15 & -13 & -2 & -12 & -3 \\ -4 & 2 & -17 & 7 & 11 \\ 3 & -19 & -1 & -18 & 5 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 7 \\ 14 \\ -1 \\ 19 \end{pmatrix}$
11	$\begin{pmatrix} 4 & 0 & -13 & -6 & 23 \\ -2 & 23 & 25 & -19 & 22 \\ 23 & -9 & -6 & 20 & 13 \\ -2 & -19 & 2 & -9 & -20 \\ 1 & 24 & -19 & 12 & -13 \end{pmatrix}$	$\begin{pmatrix} 18 \\ 14 \\ 16 \\ 13 \\ 0 \end{pmatrix}$	28	$\begin{pmatrix} -3 & -14 & 12 & 12 & 24 \\ -6 & -6 & -12 & 7 & -7 \\ 15 & -18 & -1 & 19 & -6 \\ -17 & 6 & 4 & 14 & 15 \\ -8 & 1 & -9 & 22 & -7 \end{pmatrix}$	$\begin{pmatrix} -1 \\ -9 \\ 4 \\ 1 \\ 18 \end{pmatrix}$
12	$\begin{pmatrix} -5 & -11 & 13 & 21 & -2 \\ -11 & -16 & 0 & -18 & -7 \\ -19 & -1 & 19 & 17 & 8 \\ 3 & -13 & 25 & 18 & -13 \\ -14 & 12 & 23 & 6 & -14 \end{pmatrix}$	$\begin{pmatrix} 9 \\ 7 \\ 7 \\ 20 \\ 17 \end{pmatrix}$	29	$\begin{pmatrix} -13 & 12 & 7 & -19 & -4 \\ -2 & 2 & 15 & -4 & 6 \\ 16 & -7 & -16 & 8 & -15 \\ 10 & 2 & -17 & -8 & -2 \\ 0 & -18 & 7 & 2 & -12 \end{pmatrix}$	$\begin{pmatrix} 11 \\ 15 \\ 12 \\ 9 \\ 6 \end{pmatrix}$
13	$\begin{pmatrix} -19 & 7 & 7 & 24 & 6 \\ 17 & 6 & -2 & 15 & 8 \\ 0 & 16 & -6 & 8 & -16 \\ -10 & -2 & 8 & 16 & -16 \\ 13 & -18 & 25 & -13 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 13 \\ 17 \\ 11 \\ 3 \end{pmatrix}$	30	$\begin{pmatrix} 2 & -13 & -2 & 25 & 13 \\ 0 & 14 & -14 & 9 & 24 \\ 4 & -9 & -8 & 19 & 0 \\ 23 & 13 & 1 & 12 & 3 \\ 7 & 24 & 21 & -16 & -13 \end{pmatrix}$	$\begin{pmatrix} 19 \\ 9 \\ 7 \\ -2 \\ 12 \end{pmatrix}$
14	$\begin{pmatrix} -5 & -19 & 9 & 15 & 14 \\ 18 & 11 & 14 & 3 & -6 \\ -6 & 24 & 11 & 6 & -7 \\ -7 & 18 & 8 & 11 & -7 \\ 20 & 0 & 4 & -13 & -17 \end{pmatrix}$	$\begin{pmatrix} 19 \\ 1 \\ 19 \\ 14 \\ -5 \end{pmatrix}$	31	$\begin{pmatrix} 20 & 19 & 18 & 17 & -17 \\ -3 & -13 & 6 & -6 & 7 \\ 2 & 0 & 16 & 9 & -6 \\ -11 & 1 & -15 & 19 & -20 \\ 23 & 24 & 14 & 7 & -1 \end{pmatrix}$	$\begin{pmatrix} 15 \\ -8 \\ 17 \\ -10 \\ -5 \end{pmatrix}$
15	$\begin{pmatrix} -3 & 4 & 4 & 15 & -11 \\ 7 & 9 & 10 & 0 & 20 \\ -5 & -11 & 6 & -7 & -1 \\ 16 & -6 & 14 & 7 & 20 \\ -6 & 24 & -19 & -2 & 12 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 9 \\ 14 \\ 2 \\ -2 \end{pmatrix}$	32	$\begin{pmatrix} 22 & 17 & -16 & 18 & -7 \\ -8 & 21 & 1 & 24 & 16 \\ 2 & 23 & -18 & 4 & -8 \\ 22 & -4 & 10 & 7 & -18 \\ 18 & -2 & -15 & -20 & 8 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 5 \\ 5 \\ -4 \\ -5 \end{pmatrix}$

16	$\begin{pmatrix} 4 & 4 & -5 & 13 & 9 \\ 24 & -17 & -18 & -16 & 24 \\ -3 & 3 & 11 & 11 & -3 \\ -14 & -19 & 17 & 11 & -15 \\ 19 & 13 & 12 & -8 & 5 \end{pmatrix}$	$\begin{pmatrix} 19 \\ 0 \\ 1 \\ 19 \\ 17 \end{pmatrix}$	33	$\begin{pmatrix} 21 & 14 & 24 & 16 & 18 \\ 17 & -4 & -10 & 0 & 13 \\ 17 & 21 & 7 & -2 & -7 \\ -2 & 3 & 19 & 18 & -13 \\ 9 & -4 & -8 & -16 & 22 \end{pmatrix}$	$\begin{pmatrix} -6 \\ -2 \\ 7 \\ 19 \\ -1 \end{pmatrix}$
17	$\begin{pmatrix} -15 & 8 & -17 & 9 & -17 \\ 12 & -1 & 15 & 4 & -12 \\ 0 & -9 & -16 & 4 & -10 \\ -10 & 13 & -19 & 19 & 8 \\ 14 & 1 & 16 & -20 & 7 \end{pmatrix}$	$\begin{pmatrix} 13 \\ 3 \\ -6 \\ -2 \\ -4 \end{pmatrix}$		$\begin{pmatrix} 3 & 0 & 0 & 7 & 1 \\ 9 & 7 & 9 & -1 & 4 \\ 0 & 2 & 2 & 0 & 0 \\ 5 & -2 & 5 & 0 & 6 \\ 1 & 9 & 7 & 8 & -3 \end{pmatrix}$	$\begin{pmatrix} 7 \\ -9 \\ 6 \\ 15 \\ -7 \end{pmatrix}$

Розділ 4. Розв'язування нелінійних рівнянь

- Загальна постановка задачі
- Методи розв'язування нелінійних рівнянь
- Розв'язування функціональних рівнянь з однією змінною
  - Початковий етап: відокремлення коренів
  - Відокремлення кореня трансцендентного рівняння
  - Відокремлення дійсних коренів алгебраїчного рівняння
  - Теореми про число коренів
  - Другий етап: уточнення коренів
- Ітераційні методи уточнення коренів скалярних нелінійних рівнянь
  - Ідея ітераційних методів
  - Застосування принципу стислих відображень до дослідження ітераційних процесів
  - Поняття порядку ітерації
- Метод ділення проміжку навпіл (метод дихотомії, Мюллера)
  - Аналіз постановки задачі
  - Геометрична інтерпретація методу
  - Обчислення похибки
- Метод хорд
  - Аналіз постановки задачі
  - Геометрична інтерпретація методу
  - Обчислення похибки
- Метод Ньютона (метод дотичних)
  - Аналіз постановки задачі
  - Геометрична інтерпретація методу
  - Обчислення похибки
  - Інтерпретація методу дотичних через ітераційний процес
  - Модифікований метод Ньютона
- Комбінований метод хорд і дотичних
  - Аналіз постановки задачі
  - Геометрична інтерпретація методу
  - Обчислення похибки
- Метод січних
  - Аналіз постановки задачі

- Геометрична інтерпретація методу
- Обчислення похибки
- Інтерпретація методу січних через ітераційний процес

- Метод простої ітерацій або метод послідовних наближень
  - Аналіз постановки задачі
  - Геометрична інтерпретація методу простої ітерації
  - Обчислення похибки
  - Вплив обчислювальної похибки на збіжність ітераційного процесу
- Метод релаксації
  - Аналіз постановки задачі
  - Геометрична інтерпретація методу
  - Обчислення похибки
- Метод парабол (метод Д. Мюллера) знаходження коренів многочлена з комплексними коефіцієнтами
  - Аналіз постановки задачі
  - Обчислення похибки
- Розв'язування систем нелінійних рівнянь
- Приклади розв'язування нелінійних рівнянь
- Контрольні питання
- Задачі для самостійної роботи
- Завдання до лабораторної роботи

Розв'язування багатьох інженерних задач зводиться до знаходження коренів одного нелінійного рівняння або до розв'язування систем нелінійних рівнянь [65, 71, 76-80, 88, 92-93, 106-110, 113-114, 116-133, 137, 140-145, 147-151, 161-164, 168-169, 192-194, 196-197, 201-209, 211]. В обох випадках нелінійні рівняння, що утворюються, можна поділити на два типи – *алгебраїчні* та *трансцендентні*. *Алгебраїчними* називають рівняння, що містять лише алгебраїчні функції (цілі, раціональні, ірраціональні). Нелінійні рівняння, що містять тригонометричні, логарифмічні, показникові, степеневі чи інші спеціальні функції, називають *трансцендентними*.

Розглянемо задачу знаходження коренів рівняння $f(x)=0$, де $f(x)$ – задана функція дійсної змінної. Розв'язування цієї задачі можна розкласти на декілька етапів:

а) дослідження розташування коренів (у загальному випадку – на комплексній площині) та їх кратність;

б) відокремлення коренів, тобто виділення областей, що містять тільки один корінь;

в) обчислення кореня з заданою точністю за допомогою одного з ітераційних алгоритмів.

Далі розглядаються ітераційні процеси, що дають можливість побудувати числову послідовність x_n, яка збігається до шуканого кореня x_* рівняння $f(x)=0$.

4.1. Загальна постановка задачі

Розглянемо задачу знаходження коренів рівняння. Нелінійне рівняння можна подати у таких формах запису;

$$f(x) = 0, \tag{4.1}$$

або

$$\varphi(x) = g(x), \tag{4.2}$$

де $f(x)$ – задана функція дійсної змінної, тобто $f : \mathbf{R}^1 \to \mathbf{R}^1$.

Функції, що входять до (4.1), (4.2), визначені та неперервні на множині X, що називається областю визначення рівняння. Сукупність значень змінної x, при яких рівняння (4.1) чи (4.2) перетворюється на тотожність, називають розв'язком цього рівняння, а кожне значення x із цієї сукупності – коренем рівняння.

Корені можуть бути дійсними або комплексними. Крім того, деякі з них можуть бути кратними, тобто кілька коренів можуть збігатися. Якщо функції $f(x)$, $\varphi(x)$, $g(x)$, що входять до рівнянь (4.1) і (4.2), трансцендентні, то ці рівняння можуть не мати коренів, мати скінченну кількість їх або нескінченну множину.

У багатьох задачах, як правило, розв'язують алгебраїчні або трансцендентні рівняння зі скінченною кількістю дійсних коренів, тому далі будемо розглядати тільки такі нелінійні рівняння.

Серед алгебраїчних рівнянь особливе місце належить рівнянням, що містять поліноми виду $\sum_{i=0}^{n} a_i x^{n-1}$. Ці рівняння записують так:

$$a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_{n-1} x + a_n = 0, \tag{4.3}$$

де n – натуральне число чи нуль; a_0, a_1, a_2, ..., a_n – коефіцієнти полінома (будь-які дійсні числа).

Порівняно з трансцендентними алгебраїчні нелінійні рівняння мають ту перевагу, що наперед відома точна кількість їх коренів, а отже, відомо, коли слід закінчити пошук при дослідженні алгебраїчного нелінійного рівняння.

1. Алгебраїчне рівняння n-го порядку має n коренів, які можуть бути дійсними або комплексними.

2. Кількість додатних дійсних коренів дорівнює (або менша на ціле число) кількості змін знаків у послідовності коефіцієнтів a_i.

3. Кількість від'ємних дійсних коренів дорівнює (або менша на ціле число) числу змін знаків у послідовності коефіцієнтів a_i при заміні x на $-x$.

Сукупність кількох рівнянь з кількома невідомими називають системою рівнянь. Систему n рівнянь з n невідомими можна зобразити у вигляді

$$\begin{aligned} f_1(x_1, x_2, ..., x_n) &= 0, \\ f_2(x_1, x_2, ..., x_n) &= 0, \\ &\cdots\cdots\cdots\cdots \\ f_n(x_1, x_2, ..., x_n) &= 0. \end{aligned} \tag{4.4}$$

Розв'язком системи нелінійних рівнянь називають сукупність значень невідомих, яка перетворює кожне з рівнянь на тотожність. Наприклад, система

$$x_1 + x_2^2 = 5,$$
$$x_1^2 + x_2 = 3.$$

має розв'язок x_1=1, x_2=2, оскільки при цих значеннях невідомих рівняння системи перетворюються на тотожності.

Методи розв'язування нелінійних рівнянь поділяють на *прямі* та *ітераційні*. Прямі дають змогу отримати розв'язок безпосередньо за допомогою формул і тому забезпечують точні значення коренів. Як приклад можна навести формули для визначення коренів квадратного та кубічного рівнянь. Існує також спосіб обчислення коренів алгебраїчного рівняння четвертого порядку (n = 4), проте він настільки складний, що практично його не застосовують.

Для трансцендентних рівнянь та систем нелінійних рівнянь прямих методів обчислення коренів не існує. Тому на практиці найчастіше застосовують наближені методи розв'язування цих рівнянь, які дають змогу за допомогою скінченного набору арифметичних операцій обчислити корені будь-якого нелінійного рівняння чи системи нелінійних рівнянь з достатньою точністю. Особливо ефективні наближені методи при реалізації на ПК, оскільки використовувані для цього алгоритми є простими, зручними, легко програмуються.

Наближеним значенням кореня x нелінійного рівняння з точністю до e вважають будь-яке число між a і b, при якому виконується умова $b - a \le \varepsilon$. Числа a і b – це наближені значення кореня x відповідно з недостачею і з надлишком з точністю ε. Наприклад, якщо корінь лежить між числами 1,133 і 1,134, то за наближене значення кореня з точністю до 0,001 можна взяти будь-яке число у межах між цими числами, наприклад число 1,1335.

Універсальні алгоритми обчислення коренів нелінійних рівнянь грунтуються на тому, що виходять з будь-якого вже відомого наближеного значення $x^{(0)}$ одного з коренів, якщо розв'язується одне рівняння, або значень $x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}$, якщо розв'язується система рівнянь (n – порядок системи). Ці значення далі уточнюються до заданого ступеня точності. При цьому несуттєво, як знайдено початкове значення обчислюваного кореня (коренів). Часто початкове наближення отримують за допомогою грубого попереднього підрахунку, графіка або аналізу фізичної суті задачі. Алгоритм наближеного обчислення коренів нелінійного рівняння складається з двох етапів:

1. Знаходження достатньо малих відрізків (інтервалів), у кожному з яких міститься один і тільки один корінь. Цей етап називають відокремленням коренів (або визначенням відрізків ізоляції кореня).

2. Обчислення кореня з наперед заданою точністю e, якщо відоме його деяке початкове наближення в інтервалі, що не містить інших коренів. Цей етап називають уточненням наближених значень коренів.

Відокремити корені можна кількома способами – графічним, аналітичним або методом послідовного перебирання. Останній метод зручний при використанні ПК.

4.2. Методи розв'язування нелінійних рівнянь

Знаходження кореня x для рівняння (4.1) можна розбити на 4 задачі.

1. Розв'язання питання існування коренів рівняння (4.1) в $\mathbf{R}^1$.
2. Відокремлення кореня. В $\mathbf{R}^1$ виділяємо: $X_i \subset \mathbf{R}^1$, $i=\overline{1,k}$, $\bigcup_{i=1}^{k} X_i \subset \mathbf{R}^1$, де $\alpha_i \in X_i$ – єдиний корінь рівняння (4.1) в $X_i (f(\alpha_i)=0)$.
3. Алгоритм знаходження кореня α_i, $i=\overline{1,k}$, з довільною точністю ε здійснюється за допомогою ітераційного методу. Від (4.1) перейдемо до еквівалентного рівняння

$$x = \varphi(x). \tag{4.5}$$

Будуємо

$$x_n = \varphi(x_{n-1}), n = 1,2,..k. \tag{4.6}$$

при заданому $x_0 \in \mathbf{R}^1$. Причому $\lim\limits_{n\to\infty} x_n = \alpha$.

4. Вивчення якості алгоритму знаходження кореня α з точністю $\varepsilon > 0$.

В математиці нелінійні рівняння поділяються на алгебраїчні та трансцендентні.

І. Постановка задачі для розв'язку алгебраїчних рівнянь. Під алгебраїчними рівняннями будемо розуміти степеневі рівняння вигляду:

$$a_0 x^n + a_1 x^{n-1} + ... + a_n = 0. \tag{4.7}$$

До трансцендентних рівнянь відносять логарифмічні, ірраціональні, тригонометричні та інші. Зрозуміло, що тільки окремі типи таких рівнянь піддаються точному або аналітичному розв'язуванню. Серед тригонометричних рівнянь ще менше рівнянь піддаються точному розв'язуванню. Тому широкого використання набули методи наближеного розв'язування.

Розглянемо рівняння $f(x)=0$, де $x \in [a;b]$. Необхідно визначити проміжок, на якому розв'язок існував би і був єдиним, так як можливі такі випадки:

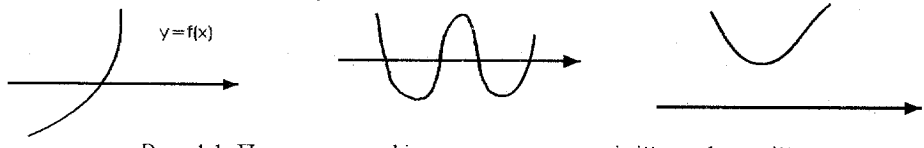


Рис. 4.1. Приклади графічного подання нелінійних функцій

Тому перед тим, як застосувати один з наближених методів, необхідно локалізувати такий інтервал $[a,b]$, на якому справедливі умови існування та єдиності розв'язків рівняння:

1) функція $f(x)$ неперервна на $[a,b]$;

2) $f(x)$ приймає на кінцях проміжку різні знаки $f(a)\cdot f(b)<0$;

3) $f'(x)<0$ або $f'(x)>0$;

3′) $f''(x)<0$ або $f''(x)>0$.

Якщо виконуються умови 1)-3), то можна говорити про існування і єдиність розв'язку на проміжку $[a,b]$. Якщо інтервал $[a,b]$ наперед не заданий, то довільним чином на дійсній осі вибирають досить великий, але скінчений проміжок, який ділять на n частин і табулюють значення функції у вузлових точках. При цьому як тільки зустрічається зміна знаку функції, то цей проміжок або їх сукупність, виділяють для наступного дослідження на єдиність. Такий підхід отримав назву **методу перебору**. Але повної гарантії на локалізацію потрібного інтервалу він не дає, тому що залежить від ширини підінтервалу. Ще простішим способом є **графічний спосіб виділення кореня** для чого використовують графічні редактори або системи комп'ютерної алгебри. Після того, як інтервал існування кореня локалізований, його необхідно далі дослідити або тим же методом перебору, або тими ж системами комп'ютерної алгебри на предмет єдиності.

При виконанні вказаних умов розв'язування рівняння допускає використання наближених обчислювальних методів до яких належать: метод хорд (січних), метод дотичних, метод простої ітерації, градієнтні методи (методи найшвидшого спуску), інформаційно-статичні методи тощо.

II. Постановка задачі для розв'язку трансцендентних рівнянь. *Трансцендентні рівняння* – це такі, що містять тригонометричні, показникові, степеневі чи інші спеціальні функції. Для їх розв'язання необхідно наступні дії.

Задано функцію (4.1), де $f(x)$ деяка функція аргументу x, що визначена на інтервалі $[a;b]$. Коренем рівняння (4.1) називається всяке число $\xi \in [a;b]$, що перетворює функцію $f(x)$ в нуль, тобто $f(\xi) = 0$. Задача пошуку кореня рівняння поділяється на два етапи.

1. Відокремлення кореня, тобто виділення відрізку, на якому розміщено тільки один корінь. При цьому один з кінців відрізку або його середину вибирають за початкове наближення.
2. Уточнення кореня – це звуження границь виділеного відрізку ізоляції кореня за допомогою одного з методів дотi, поки довжина відрізку не стане меншою, ніж насамперед задана точність ε.

В багатьох випадках відділення кореня можна провести графічно. Беручи до уваги, що дійсні корені рівняння (4.1) – це точки перетину графіка функції $f(x)$ з віссю абсцис, достатньо побудувати графік $f(x)$ і відмітити на осі $0x$ відрізки, які містять один корінь. Побудову графіків часто вдається сильно спростити, замінивши рівняння (4.1) рівносильним йому рівнянням $\varphi(x) = \psi(x)$. В цьому випадку будуються графіки функцій $\varphi(x)$ і $\psi(x)$, а потім на осі $0x$ виділяють відрізки, які локалізують абсциси точок перетину цих графіків.

Умова існування кореня

1. Якщо неперервна на відрізку $[a;b]$ функція $f(x)$ приймає на його кінцях значення різних знаків, тобто $f(a)f(b) < 0$, то рівняння (4.1) має на цьому відрізку по меншій мірі один корінь.
2. Якщо функція $f(x)$ строго монотонна, то на $[a;b]$ корінь єдиний.

Приклад 4.1. Виділити корені для рівняння $e^x + x - 2 = 0$ (рис. 4.2).

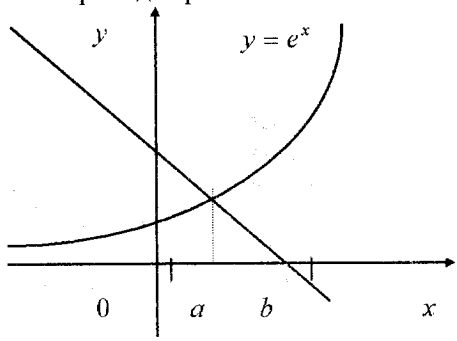


Рис. 4.2. Графічне подання функції для прикладу 4.1

Розв'язок. Для функції $\varphi(x) = e^x$, тобто $\psi(x) = x - 2$ визначимо:

$$a = 0.3;\ f(a) = -0{,}35014;\ b = 0.8;\ f(b) = 1.0255.$$

Отже, корінь існує на проміжку $[a;b]$. Достатня умова – постійність знаку похідної на $[a;b]$.

Визначення та основні властивості опуклих функцій

Визначення. Нехай $f: S \to E_1$, де S – непорожня *опукла* множина в E_n. Тоді функція f опукла на S, якщо

$$f[\lambda x_1 + (1-\lambda)x_2] \le \lambda f(x_1) + (1-\lambda) f(x_2), \tag{4.8}$$

для будь-яких $x_1, x_2 \in S_\alpha$ і $\lambda \in (0,1)$.

Функція f *строго опукла* на S, якщо

$$f[\lambda x_1 + (1-\lambda)x_2] < \lambda f(x_1) + (1-\lambda) f(x_2), \tag{4.9}$$

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'єднує x_1 та x_2, не перевищує середньозваженого (з тим же λ) значення величин $f(x_1)$ та $f(x_2)$. Наступні функції є опуклими:

$$
\begin{aligned}
&f_1(x) = 3x + 4,\\
&f_2(x) = |x|,\\
&f_3(x) = x^2 - 2x,\\
&f_4(x) = -x^{\frac{1}{2}}, x \geq 0,\\
&f_5(x_1, x_2) = 2x_1^2 + x_2^2 - 2x_1x_2,\\
&f_6(x_1, x_2, x_3) = x_1^4 + 2x_2^2 + 3x_3^2 - 4x_1 - 4x_2x_3.
\end{aligned}
\tag{4.10}
$$

Поставимо у відповідність функції f множину $S_\alpha = \{x \in S : f(x) \leq \alpha\}$, що зазвичай має назву *множина рівня*. Наступна лема показує, що S_α опукла при будь-якому дійсному α.

Лема 4.1. Нехай S – непорожня опукла множина в E_n, $f : S \to E_1$ – опукла функція. Тоді множина $S_\alpha = \{x \in S : f(x) \leq \alpha\}$ опукла при будь-якому дійсному значенні α.

Доведення. Нехай $x_1, x_2 \in S_\alpha$, тобто $x_1, x_2 \in S_\alpha$ та $f(x_1) \leq \alpha$, $f(x_2) \leq \alpha$. Візьмемо будь-яке $\lambda \in (0,1)$ та розглянемо точку $x = \lambda x_1 + (1 - \lambda)x_2$. Із-за опуклості S точка x належить S. Так як функція f опукла, то

$$
f(x) \leq \lambda f(x_1) + (1 - \lambda) f(x_2) \leq \lambda\alpha + (1 - \lambda)\alpha = \alpha. \tag{4.11}
$$

Відповідно, $x \in S_\alpha$, тобто S_α – опукла множина.

Неперервність опуклих функцій. Важливою властивістю опуклих функцій є їх неперервність всередині області ефективності.

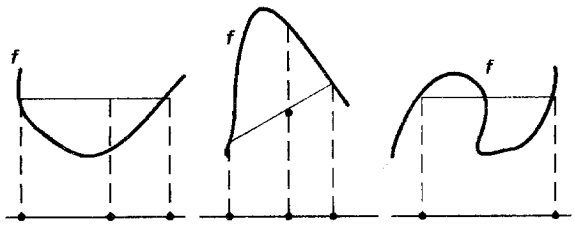
Теорема 4.1. Нехай S – непорожня опукла множина в E_n, $f : S \to E_1$ – опукла функція. Тоді f неперервна у внутрішніх точках множини S.

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. Кажуть, що функція $f : S \to E_1$ квазіопукла, якщо для всіх $x_1, x_2 \in S$ виконується нерівність

$$
f[\lambda x_1 + (1 - \lambda)x_2] \leq \max\{f(x_1), f(x_2)\}, \tag{4.12}
$$

при будь-якому $\lambda \in (0,1)$.

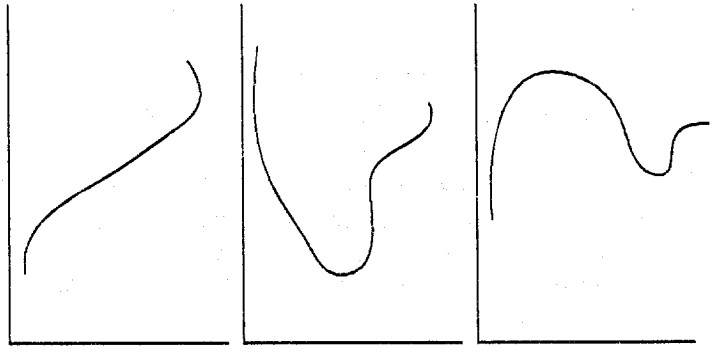


Опуклі та ввігнуті функції. a - опукла функція;b - ввігнута функція;c - функція, що не є ні опуклою, ні ввігнутою.

Рис. 4.3. Графічне подання опуклих функцій

Функція f строго квазіопукла, 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 квазіопукла функція.

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.



Квазіопуклість та псевдоопуклість. a - функція квазіопукла та псевдоопукла; b - функція квазіопукла, але не псевдоопукла; c - функція не квазіопукла і не псевдоопукла

Рис. 4.4. Графічне подання квазіопуклості та псевдоопуклості функцій

Наведемо короткі відомості важливих властивостей різних типів опуклих функцій. Всюди тут $f: S \to E_1$ та S – непорожня опукла множина в E_n.

Строго опуклі функції

1. Функція f неперервна в $\operatorname{int} S$.
2. Множина $\{(x, y): x \in S, y \geq f(x)\}$ опукла.
3. Множина $\{x \in S: f(x) \leq \alpha\}$ опукла для кожного дійсного α.
4. Диференційована функція f строго опукла тоді і тільки тоді, коли $f(x) > f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x})$ при всіх різних $x, \bar{x} \in S$.

5. Нехай f подвійно 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 $x \in S$.
6. Кожний локальний мінімум функції f на опуклій множині $X \subset S$ є єдиним глобальним мінімумом.
7. Якщо $\nabla f(\bar{x}) = 0$, то $\bar{x}$ – єдина точка глобального мінімуму функції f на множині S.
8. Функція f досягає максимуму на компактній багатогранній множині $X \subset S$ в екстремальній точці цієї множини.



Рис. 4.5. Схема зв'язків типів опуклих функції

Опуклі функції

1. Функція f неперервна в $\operatorname{int} S$.
2. Функція f опукла тоді і лише тоді, коли множина $\{(x,y): x \in S, y \geq f(x)\}$ опукла.
3. Множина $\{x \in S: f(x) \leq \alpha\}$ опукла для кожного дійсного α.
4. Диференційована функція f опукла тоді і тільки тоді, коли $f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x})$ при всіх різних $x, \bar{x} \in S$.
5. Подвійно диференційована функція f опукла тоді і лише тоді, коли матриця Гессе $H(x)$ позитивно визначена для всіх $x \in S$.
6. Кожний локальний мінімум функції f на опуклій множині $X \subset S$ є глобальним мінімумом.
7. Якщо $\nabla f(\bar{x}) = 0$, то $\bar{x}$ – точка глобального мінімуму функції f на множині S.
8. Функція f досягає максимуму на компактній багатогранній множині $X \subset S$ в екстремальній точці цієї множини.

Псевдоопуклі функції

1. Множина $\{x \in S: f(x) \leq \alpha\}$ опукла для кожного дійсного α.

2. Кожний локальний мінімум функції f на опуклій множині $X \subset S$ є глобальним мінімумом.
3. Якщо $\nabla f(\bar{x}) = 0$, то $\bar{x}$ – точка глобального мінімуму функції f на множині S.
4. Функція f досягає максимуму на компактній багатогранній множині $X \subset S$ в екстремальній точці цієї множини.

Квазіопуклі функції

1. Функція f квазіопукла тоді і тільки тоді, коли множина $\{x \in S : f(x) \le \alpha\}$ опукла для будь-якого дійсного α.
2. Функція f досягає максимуму на компактній багатогранній множині $X \subset S$ в екстремальній точці цієї множини.
3. Диференційована функція f на множині S квазіопукла тоді і тільки тоді, коли із $x_1, x_2 \in S$ та $f(x_1) \le f(x_2)$ випливає, що $\nabla f(x_2)^T (x_1 - x_2) \le 0$.

Локальний мінімум 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.

4.3. Розв'язування функціональних рівнянь з однією змінною

У загальному вигляді функціональне рівняння записується у вигляді:

$$F(x) = 0, \tag{4.13}$$

де $F(x)$ визначена і неперервна на скінченому інтервалі $[a,b]$. Число ζ (ксі) називається коренем r кратності, якщо воно задовольняє рівняння:

$$F(\zeta) = 0, \quad F'(\zeta) = 0, \quad \ldots, \quad F^{r-1}(\zeta) = 0, \tag{4.14}$$

Однократний корінь називається простим.

Два рівняння $F(x)$ та $G(x)$ називаються рівносильними, якщо будь-який розв'язок кожного з них є розв'язком іншого.

Функціональні рівняння поділяються на алгебраїчні (наприклад: $x^3 - 3x + 10 = 0$) та трансцендентні (наприклад: $\sin(2x) - \lg(x/3) = 0$).

Алгебраїчні рівняння є лінійні та нелінійні.

Шляхом алгебраїчних перетворень з будь-якого алгебраїчного рівняння можна отримати рівняння у канонічній формі:

$$P_n(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_n = 0, \tag{4.15}$$

де n - ступінь алгебраїчного рівняння. Кожне алгебраїчне рівняння має хоча б один дійсний корінь чи пару комплексних.

При зведенні алгебраїчного рівняння до канонічної форми отримаємо ті ж корені, що і для початкового рівняння, але при цьому можуть з'явитися сторонні корені. Наприклад,

$$\sqrt{2x^2-1}+x=\sqrt{2x^2+1}-1 \Rightarrow 7x^4+12x^3+2x^2-4x-5. \qquad (4.16)$$

Якщо функція $f(x)$ не є алгебраїчною, то рівняння (4.13) називається трансцендентним. Наприклад,

$$\lg(2x)=\cos(0{,}5\cdot x). \qquad (4.17)$$

У деяких випадках розв'язування трансцендентних рівнянь зводиться до розв'язування алгебраїчних.

Означення 4.1. *У теорії чисельних методів розв'язати рівняння – значить встановити, чи має воно корені на вказаному проміжку, з'ясувати кількість коренів, відшукати значення всіх виявлених коренів із заданою точністю (корені мають знаходитися у ε -околі точного розв'язку).*

Задача чисельного знаходження дійсних і комплексних коренів рівняння зазвичай складається з двох основних етапів:

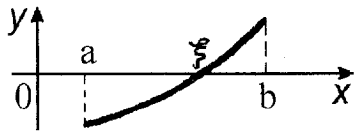
- ***відокремлення коренів*** – знаходження достатньо малих околів у заданій області, в кожному з яких знаходиться тільки один корінь,
- ***уточнення коренів*** – обчислення коренів із заданим ступенем точності ε в деякому околі (ε -околі).

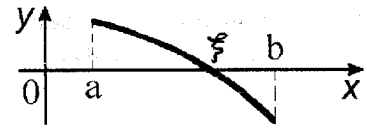
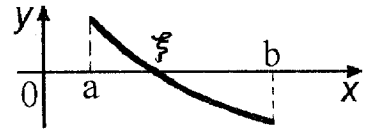
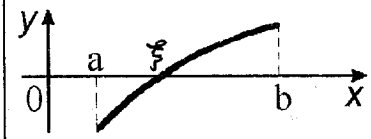
Для знаходження дійсних коренів рівняння (4.1) застосовується ряд методів, серед яких найчастіше використовують такі чисельні методи:

- графічного розв'язку;
- метод половинного ділення (бісекції або поділ навпіл);
- метод хорд;
- метод січних;
- метод дотичних (Ньютона);
- метод простої ітерації (ітерацій);
- комбінований (хорд і дотичних).

Всі можливі випадки при уточненні коренів можна класифікувати за допомогою таблиці 4.1.

Таблиця 4.1

№	СХЕМА	$f(a)$	$f(b)$	$f'(x)$	$f''(x)$	Нульове наближення x_0	
						метод хорд	*метод дотичних*
I		-	+	+	+	*a*	*b*

II		+	-	-	-	a	b
III		+	-	-	+	b	a
IV		-	+	+	-	b	a

Постановка задачі. Дано рівняння $f(x)=0$, де $f(x)$ – неперервна функція, що має в $[a,b]$ інтервалі неперервні і знакосталі похідні першого і другого порядків. Корінь ξ ізольований і відділений на $[a,b]$, тобто виконується умова $f(a)\cdot f(b)<0$. Необхідно уточнити корінь ξ до заданого ступеня точності ε.

Розглянемо перший основний етап знаходження коренів - ***відокремлення коренів.***.

4.3.1. Початковий етап: відокремлення коренів

Відокремлення коренів полягає у встановленні тісних проміжків, кожен з яких містить тільки один корінь. Найпростіший алгоритм відокремлення коренів полягає у наступному:

Проміжок, на якому шукаються розв'язки, ділять рівномірно на невеликі відрізки, на кінцях кожного відрізку розраховують значення функції. Якщо функція на кінцях відрізку змінює знак – відрізок містить корінь (значення початку і кінця відрізку записується для подальшого опрацювання), інакше:

- коренів кілька і відрізок ділять на менші відрізки,
- кореня на відрізку немає і відрізок не розглядається.

Величина відрізків поділу залежить від коефіцієнта гладкості функції. У розрахунково-графічній роботі можемо прийняти $h=|b-a|/10$. Для визначення відрізку, що містить один корінь, можемо скористиатися наступною теоремою.

Теорема 4.2. Якщо функція $y=F(x)$ неперервна на інтервалі $[a,b]$ і якщо $F(a)$ та $F(b)$ мають протилежні знаки, тобто $F(a)\cdot F(b)<0$, то $F(x)$ має хоча б один дійсний корінь на інтервалі $[a,b]$. Якщо при цьому $F(x)$ має першу і другу похідні, що не змінюють знак на інтервалі $[a,b]$ (кількох точках інтервалу), то *корінь єдиний.*

Варіантів розташування функції на відрізку може бути кілька (рис. 4.6):

1) $f(x_i)\cdot f(x_{i+1})<0$ та $f'(x_i)\cdot f'(x_{i+1})>0$ (рис. 4.6а);
2) $f(x_i)\cdot f(x_{i+1})<0$ та $f'(x_i)\cdot f'(x_{i+1})<0$ (рис. 4.6б);

3) $f(x_i)\cdot f(x_{i+1})>0$ та $f'(x_i)\cdot f'(x_{i+1})<0$ (рис. 4.6в);

4) $f(x_i)\cdot f(x_{i+1})<0$ та $f'(x_i)\cdot f'(x_{i+1})<0$ (рис. 4.6г).

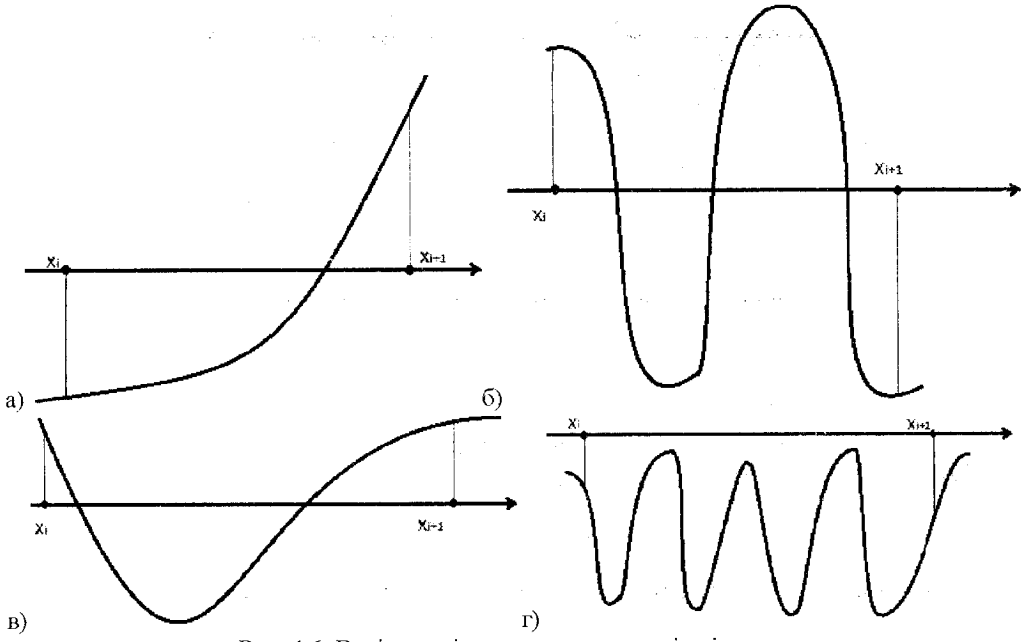


Рис. 4.6. Варіанти відокремлення коренів рівняння

В залежності від реалізації алгоритм відокремлення коренів може бути модифікований. Так у більшості випадків, коли немає однозначного рішення про відкидання поточного проміжку (випадки б-г на рис. 4.6), його ділять на кілька менших проміжків.

4.3.2. Відокремлення кореня трансцендентного рівняння

Розглянемо рівняння (4.1) $f(x)=0$, де $f\in C^1\left((a,b)\right)$. *Задача відокремлення* полягає у відшуканні інтервалів (a_i,b_i), $a_i\in(a_i,b_i)$, $(a_i,b_i)\in(a,b)$, в кожному з яких лежить один корінь α_i, причому $(a_i,b_i)\cap(a_j,b_j)=\varnothing$, $i\neq j$.

Найпростішим шляхом відокремлення коренів є графічний метод (рис. 4.7). Від заданого рівняння $f(x)=0$ переходимо до еквівалентного рівняння

$$\varphi(x)=\psi(x), \tag{4.18}$$

тоді абсциси точок перетину цих двох функцій $\varphi(x)$ та $\psi(x)$ будуть коренями вихідного рівняння (4.1) (рис. 4.8).

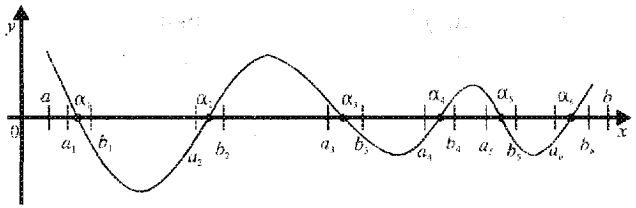


Рис. 4.7. Варіанти відокремлення коренів трансцендентного рівняння

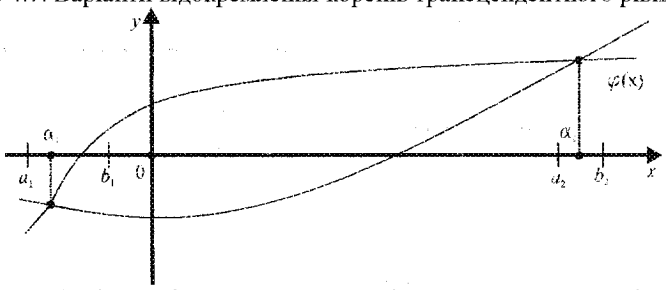


Рис. 4.8. Варіанти відокремлення коренів трансцендентного рівняння

Приклад 4.2. Знайти корені рівняння

$$x \cdot \sin(x) - 1 = 0. \tag{4.19}$$

Розв'язок. Від (4.19) переходимо до еквівалентного рівняння.

$$\sin x = \frac{1}{x}, \tag{4.20}$$

де $\varphi(x) = \sin x$, $\psi(x) = \frac{1}{x}$. Окремо будуємо графіки (рис. 4.9), абсциси точок перетину й будуть дійсними коренями вихідного рівняння (4.19).

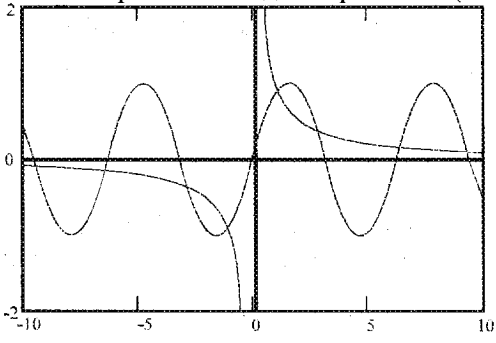


Рис. 4.9. Варіанти відокремлення коренів рівняння

Для великих $k \geq 10$ корені на графіку не зображено, але $\alpha_k \approx \pi k$. Уточнимо їх значення:

$$\alpha_k = \pi k + \varepsilon_k, \tag{4.21}$$

де $\varepsilon_k > 0$ – нескінченно малі. Тепер (4.20) матиме вигляд:

$$\frac{1}{\pi k + \varepsilon_k} - \sin\left(\pi k + \varepsilon_k\right) = \frac{1}{\pi k + \varepsilon_k} - \left(-1\right)^k \sin \varepsilon_k = 0.$$

Але для малих $\varepsilon_k > 0$ відомо, що $\sin \varepsilon_k \approx \varepsilon_k$, тоді $\frac{1}{\pi k + \varepsilon_k} \approx \frac{1}{\pi k}$. Значить $\frac{1}{\pi k} - \varepsilon_k \left(-1\right)^k \approx 0$, що дає $\varepsilon_k \approx \frac{\left(-1\right)^k}{\pi k}$. Таким чином, $\alpha_k \approx \pi k + \frac{\left(-1\right)^k}{\pi k}$. Можна продовжити уточнення коренів.

Зауваження 4.1. *(по відокремленню коренів трансцендентного рівняння* $f(x) = 0$*).* Нехай $f \in C\left(\left[a_i, b_i\right]\right)$. Якщо $f\left(\alpha_i\right) \cdot f\left(b_i\right) < 0$, то на (a_i, b_i) існує хоча б один корінь.

Зауваження 4.2. Якщо $f\left(\alpha_i\right) \cdot f\left(b_i\right) < 0$ і $f'(x) > 0 (< 0)$ для $\forall x \in \left(a_i, b_i\right)$, то a, єдиний корінь.

Зауваження 4.3. Нехай $f(x)$ аналітична функція на $[a,b]$, тоді

1) якщо $f\left(\alpha\right) f\left(b\right) < 0$, то на $[a,b]$ знаходиться непарна кількість коренів (4.1);
2) якщо $f\left(\alpha\right) f\left(b\right) > 0$, то на $[a,b]$ або немає коренів, або парна кількість коренів.

Зауваження 4.4. Якщо

$$f\left(z\right) = 0,\ Z \in \mathbf{R}, \tag{4.22}$$

де $z = x + iy$, то $f\left(z\right) = 0$ еквівалентне

$$\varphi\left(x, y\right) + i \cdot \psi\left(x, y\right) = 0, \tag{4.23}$$

де φ, ψ – функції дійсних змінних $(x, y) \in \mathbf{R}^2$.

Значить рівняння (4.23) еквівалентне системі рівнянь, але з дійсними функціями дійсних аргументів

$$\begin{cases} \varphi\left(x, y\right) = 0, \\ \psi\left(x, y\right) = 0. \end{cases}$$

Отримаємо графіки кривих $\varphi(x,y), \psi(x,y)$ в площині XOY. Їх точки перетину дають дійсні та уявні частини коренів $f\left(z\right) = 0$.

4.3.3. Відокремлення дійсних коренів алгебраїчного рівняння

Розглянемо рівняння

$$f(z) = a_0 z^n + a_1 z^{n-1} + \ldots + a_{n-1} z + a_n = 0, \tag{4.24}$$

для якого існує n коренів, якщо $a_i \in \mathbf{R}^1$, $i = 0,1,..,n$.

Теорема 4.3 (метод знаходження меж розташування коренів). Розглянемо рівняння (4.24). Якщо позначимо

$$\begin{aligned} a &= \max\left\{|a_1|,|a_2|,..,|a_n|\right\}, \\ a^* &= \max\left\{|a_0|,|a_1|,..,|a_{n-1}|\right\}, \end{aligned} \tag{4.25}$$

то всі корені рівняння(4.24) знаходяться в кільці

$$\frac{|a_n|}{a^* + |a_n|} = z_0 \le |z| \le z_1 \equiv 1 + \frac{a}{|a_0|}. \tag{4.26}$$

Доведення. 1. Доведемо, що $|z| \le z_1$. Якщо $|f(z)| > 0$ при $|z| > z_1$, то $|f(z)| = 0$ при $|z| \le z_1$. Якщо позначимо через $A \equiv a_0 z^n$; $B \equiv a_1 z^{n-1} + \ldots + a_{n-1} z + a_n$, то матимемо $(|A+B| \ge |A| - |B|)$

$$|f(z)| \ge |a_0 z^n| - |a_1 z^{n-1} + \ldots + a_{n-1} z + a_n|. \tag{4.27}$$

Для $(z) > 1$ зрозумілий ланцюжок нерівностей

$$|a_1 z^{n-1} + a_2 z^{n-2} + \ldots + a_{n-1} z + a_n| \le a\left(|z|^{n-1} + |z|^{n-2} + \ldots + |z| + 1\right) = a\frac{|z|^n - 1}{|z| - 1} < \frac{a|z|^n}{|z| - 1}. \tag{4.28}$$

Враховуючи (4.28), нерівність (4.27) тільки посилиться

$$|f(z)| > |a_0||z|^n - \frac{a|z|^n}{|z| - 1} > 0.$$

Значить $|f(z)| > 0$, як тільки $|a_0||z|^n - \frac{a|z|^n}{|z| - 1} > 0$ при $|z| > 1$, тоді

$$|z|^n \left(|a_0||z| - |a_0| - a\right) > 0.$$

Якщо $|a_0||z| - |a_0| - a > 0$, то $|f(z)| > 0$. Таким чином $f(z) = 0$ при

$$|z| \le 1 + \frac{a}{|a_0|} \equiv z_1.$$

2. Доведемо другу нерівність (4.26): $z_0 \le |z|$. Зауважимо, що рівняння

$$a_0 + a_1 y + \ldots + a_n y^n = 0 \tag{4.29}$$

має корінь $y = \frac{1}{z}$, де z - розв'язок рівняння (4.24).

За теоремою 4.3 пункту 1 виконується нерівність $|y| \le 1 + \frac{a^*}{|a_n|}$, тоді $\left|\frac{1}{z}\right| \le 1 + \frac{a^*}{|a_n|}$,

а значить $|z| \ge \frac{|a_n|}{a^* + |a_n|} \equiv z_0$. Далі припустимо, що $a_i \in R, i = 1,2,..,n,$ та $a_0 > 0$.

Знайдемо межі дійсних коренів рівняння (4.24). Очевидно, що треба мати методи знаходження меж додатних коренів, бо роблячи зміну x на $-x$, матимемо рівняння, корені якого відрізняються від коренів вихідного рівняння знаком.

Теорема 4.4. (метод знаходження додатних коренів). Якщо

1) $a_0 > 0$;

2) $\tilde{a} \equiv \max\left\{|a_i, i \in \{i_1,\ldots,i_k, k \le n\} | a_i < 0\right\}$;

3) Позначимо через m індекс першого від'ємного коефіцієнта $\{a_0 > 0, a_1 > 0, \ldots, a_{m-1} > 0, a_{m+1} < 0, \ldots, a_n < 0\}$, то

$$x = z^+ < 1 + \sqrt{\frac{\tilde{a}}{a_0}}; x > 0. \tag{4.30}$$

Зауваження 4.5. Якщо від'ємних коефіцієнтів немає, то немає і додатних коренів.

Доведення. В рівнянні (4.24) замінимо додатні коефіцієнти, починаючи з другого, нулями, а всі інші на a:

$$a_1 = 0, a_2 = 0, \ldots, a_{m-1} = 0; a_m = -\tilde{a}; a_{m+1} = -\tilde{a}; \ldots; a_n = -\tilde{a}.$$

Тоді при $x > 1$ виконується нерівність

$$\begin{aligned} f(x) &= a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n > a_0 x^n - \tilde{a}(x^{n-m} + x^{n-m+1} + \ldots + x + 1) = \\ &= a_0 x^n - \tilde{a}\frac{x^{n-m+1} - 1}{x-1} > a_0 x^n - \tilde{a}\frac{x^{n-m+1}}{x-1} = \frac{x^{n-m+1} - 1}{x-1}(a_0 x^{m-1}(x-1) - \tilde{a}). \end{aligned} \tag{4.31}$$

З (4.31) випливає $f(x) > 0$ для $x > \sqrt{\tilde{a}/a_0 + 1}$, бо

$$a_0 x^{m-1}(x-1) - \tilde{a} \ge a_0 \sqrt[m]{\frac{\tilde{a}}{a_0}}\left(1 + \sqrt[m]{\frac{\tilde{a}}{a_0}}\right)^{m-1} - \tilde{a} > 0.. \tag{4.32}$$

Нерівність (4.32) означає, що $f(x) \le 0$, тобто має місце (4.30).

Теорема 4.5. (метод Ньютона знаходження верхньої межі додатних коренів). Якщо

1) $a_0 > 0$;

2) при $x = c > 0$ мають місце нерівності $f(c) > 0;\ f^{(1)}(c);\ldots;f^{(n)}(c) > 0$, то

$$c = \sup\left\{x_i > 0, i = i_1,\ldots,i_k, k \le n \middle| f(x_i) = 0\right\}. \tag{4.33}$$

Нагадаємо: Ісаак Ньютон (04.01.1643 – 31.03.1727) – геніальний англійський фізик, механік, астроном і математик. Автор диференціального та інтегрального числення. Ввів метод дослідження функцій за допомогою рядів.

Доведення. За формулою Тейлора для $\forall x > c$ маємо

$$f(x) = f(c) + (x-c)f^{(1)}(c) + \ldots + (x-c)^n \frac{f^{(n)}(c)}{n!} > 0.$$

Тоді не існує жодного кореня, більшого ніж c, тобто $\forall x: f(x) = 0$ випливає, що $x \le c$, тобто має місце (4.33).

Приклад 4.3. Знайти нижню межу додатних коренів рівняння (4.24).

Розв'язок. Нижня межа додатних коренів може бути знайдена з рівняння (4.29) за теоремою 4.5. Нехай β - верхня межа рівняння (4.29), тоді $\frac{1}{\beta}$ буде нижньою межею додатних коренів вихідного рівняння (4.24).

Теорема 4.6. (метод знаходження додатних та від'ємних коренів). Нехай

1) $f(z) \equiv a_0 z^n + a_1 z^{n-1} + \ldots + a_{n-1} z + a_n$;

2) $\begin{cases} \varphi_1(z) = z^n f(\frac{1}{z}), \\ \varphi_2(z) = f(-z), \\ \varphi_3(z) = z^n f(-\frac{1}{z}); \end{cases}$

3) N_0, N_1, N_2, N_3 – верхні межі додатних коренів многочленів $\{f(z)\};\{\varphi_1(z)\};\{\varphi_2(z)\};\{\varphi_3(z)\}$.

Тоді $z^+ \in \left[\frac{1}{N_1}, N_0\right]; z^- \in \left[-N_2, -\frac{1}{N_3}\right]$, де z^+ позначено додатні корені, z^- – від'ємні. Доведення можна знайти у будь якому підручнику з вищої алгебри.

4.3.4. Теореми про число коренів

Теорема 4.7. (теорема Декарта про число додатних коренів). Якщо:

$$f(z)=0; \tag{4.34}$$

та $\{a_k\} \equiv \{a_0, a_1, a_2, ..., a_{n-1}, a_n,$ без врахування $a_i = 0, i = i_1, ..., i_k, k \le n\}$, то число додатних коренів N^+ дорівнює числу S зміни знаків у послідовності $\{a_n\}$, тобто $N^+ = S(\{a_k\})$.

Нагадаємо: Декарт Рене (31.03.1596 – 11.02.1650) – французький філософ, математик, фізик, фізіолог.

Приклад 4.4. Знайти число від'ємних коренів многочлена (4.34).

Розв'язок. Якщо застосувати теорему Декарта до многочлена $f(-z)=0$, то одержимо число додатних коренів, яке рівне числу від'ємних коренів вихідного многочлена (4.34).

Теорема 4.8. (теорема Штурма про число дійсних коренів на відрізку $[a,b]$**).** Нехай рівняння $f(x)=0$:

1) не має кратних коренів;

2) послідовність

$$f(x), f_1(x), ..., f_n(x) \tag{4.35}$$

побудована таким чином:

$$f_1(x) = f'(x); \qquad \frac{f(x)}{f_1(x)} = q_1(x) + r_1(x); \qquad f_2(x) = -r_1(x);$$

$$f_1(x) / f_2(x) = q_2(x) + r_2(x); \qquad f_3(x) = -r_2(x);$$ і т. д. $f_n(x) \equiv const.$

Тоді число дійсних коренів рівняння $f(x)=0$ на $[a,b]$ дорівнює числу зміни знаків у послідовності (4.35): $N = S(a) - S(b)$.

Нагадаємо: Штурм Жан Шарль Франсуа (22.09.1803 – 18.12.1855) – французький математик.

Приклад 4.5. Відокремити корені многочлена $f(x)=0$ за допомогою теореми 4.6.

Розв'язок. Нехай кратних коренів немає.

1. Визначимо межі, де знаходяться дійсні корені, для чого розіб'ємо відрізок $[a,b]$ точками $a = \alpha_0 < \alpha_1 < \alpha_2 < ... < \alpha_{m-1} < \alpha_m = b$.
2. За теоремою Штурма на $[\alpha_i, \alpha_{i+1}]$ визначаємо число коренів. Якщо $N > 1$, тоді $\frac{\alpha_i + \alpha_{i+1}}{2}$ і знову на $\left[\alpha_i, \alpha_i + \frac{\alpha_i + \alpha_{i+1}}{2}\right], \left[\frac{\alpha_i + \alpha_{i+1}}{2} + \alpha_i, \alpha_{i+1}\right]$ застосуємо теорему Штурма.
3. Розбиття відрізків продовжується, поки на $[\alpha_i, \alpha_{i+1}]$ не отримаємо по одному кореню.

Приклад 4.6. Відокремити дійсні корені у рівнянні $4x^3 - 2x^2 - 4x - 3 = 0$.

Розв'язок. Побудуємо ряд Штурма

$$f(x) = 4x^3 - 2x^2 - 4x - 3$$

$$f_1(x) = 3x^2 - x - 1$$

$$f_2(x) = 26x + 29$$

$$f_3(x) = -1.$$

Таблиця 4.2

$sign\ \ x$	$-\infty$	0	1	2	$+\infty$
$sign\ \ f(x)$	-	-	-	+	+
$sign\ \ f_1(x)$	+	-	+	+	+
$sign\ \ f_2(x)$	-	+	+	+	+
$sign\ \ f_3(x)$	-	-	-	-	-
Кількість коренів	2	2	2	11	

4.3.5. Другий етап: уточнення коренів

Уточнення коренів полягає в обчисленні коренів у заданих границях $[a,b]$ із заданим ступенем точності ε на кожному відокремленому проміжку, який містить один корінь (такі проміжки отримаємо на попередньому етапі). Методи уточнення коренів мають алгоритм наближення кореня із заданою точністю та умову виходу з ітераційного (покрокового) процесу. Розглянемо деякі з перелічених вище методів.

4.4. Ітераційні методи уточнення коренів скалярних нелінійних рівнянь

4.4.1 Ідея ітераційних методів

Розглянемо рівняння

$$f(x) = 0, \tag{4.36}$$

яке зведемо до еквівалентного рівняння

$$x = \varphi(x). \tag{4.37}$$

Нехай для $x = \beta$ маємо $f(\beta) = 0$, тобто $\beta \in (a,b)$ – єдиний корінь рівняння (4.36) (методами з розділу 4.3 відокремлюємо корінь β). Беремо довільне $x_0 \in (a,b)$ і будуємо послідовність $\{x_k\} \subset (a,b)$

$$x_k = \varphi_k(x_0, x_1, \ldots, x_{k-1}), \tag{4.38}$$

так, щоб $\lim\limits_{k\to\infty} x_k = \beta$. Збіжність забезпечується вибором функцій φ_k і початкового наближення x_0.

Означення 4.1. *Ітераційний процес* (4.38) *назвемо нестаціонарним, якщо він залежить від f(x) і номера k.*

Ітераційний процес

$$x_k = \varphi(x_{k-1}, x_{k-2}, \ldots, x_{k-l}) \tag{4.39}$$

назвемо стаціонарним, якщо він не залежить від k, а залежить від початкових умов $x_0, x_1, \ldots, x_{k-1}, k = 1, 2, \ldots, l$.

Означення 4.2. *Ітераційний процес* (4.39) *назвемо l-кроковим.*

Зауважимо, що будемо вивчати ітераційні процеси

$$x_k = \varphi(x_{k-1}), \tag{4.40}$$

$$x_k = \varphi(x_{k-1}, x_{k-2}) \tag{4.41}$$

які назвемо відповідно однокроковим і двокроковим ітераційним процесом.

4.4.2. Застосування принципу стислих відображень до дослідження ітераційних процесів

Нехай задано метричний простір $\mathbf{R} \equiv \mathbf{R}^1$ з елементом $x \in \mathrm{R}$ і оператор $A: \mathbf{R} \to \mathbf{R}$.

Означення 4.3. *Оператор A здійснює стисле відображення* **R** *у себе, якщо існує дійсне число* $\alpha \in (0,1)$ *таке, що для довільної пари дійсних чисел x та y виконується нерівність:*

$$\rho(Ax, Ay) \le \alpha \cdot \rho(x, y). \tag{4.42}$$

Теорема 4.9. (принцип стислих відображень). Якщо

1) **R** — повний метричний простір;

2) для оператора A виконується нерівність (4.42), то

a) рівняння $x = Ax$ має єдиний розв'язок β;

b) $\lim\limits_{k\to\infty} x_k = \beta$, де $x_k = Ax_{k-1}$, $x_0 \in \mathbf{R}$.

Доведення у будь-якому підручнику з функціонального аналізу.

Приведемо твердження, яке уточнює принцип стислих відображень.

Теорема 4.10. Нехай

1) $S \equiv \{x \in \mathbf{R}, y_0 \in \mathbf{R} \mid \rho(x, y_0) \le r\}$;

2) визначено оператор списку A в S так, що $\forall x, y \in D$ існує число $\alpha \in (0,1)$ і виконуються такі нерівності:

$$\rho(Ax, Ay) \le \alpha \cdot \rho(x, y), \tag{4.43}$$

$$\rho(Ax_0, Ay_0) \le (1-\alpha)r. \tag{4.44}$$

Тоді в S існує єдиний розв'язок $Ax=x$, і $\lim\limits_{k\to\infty} x_k = \beta, (\beta = A\beta) \quad \forall x_0 \in S.$

Доведення теореми 4.10 базується на основі теореми 4.9. Потрібно довести, що окіл $S \equiv \{x \in \mathbf{R}, y_0 \in \mathbf{R} \mid \rho(x, y_0) \le r\}$ є повним метричним простором.

1. Спочатку доведемо, що оператор A відображає S у себе; тобто $A: S \to S$. Дійсно для $\forall x \in S$ виконується

$$\rho(Ax, y_0) \le \rho(Ax, Ay_0) + \rho(Ay_0, y_0) \le$$
$$\le \alpha\rho(x, y_0) + (1-\alpha)r \le ar + (1-\alpha)r = r,$$

Значить $Ax \in S$, що означає $A: S \to S$.

2. Нерівність (4.43) означає, що A – оператор стиску. Доведемо, що S є повним, тобто метричним простором з тим же поняттям віддалі $\rho(x, y)$. Беремо для фундаментальної послідовності $\{y_n, n \ge 1\} \subset S \subset \mathbf{R}$ границю $\lim\limits_{n\to\infty} y_n = y \in \mathbf{R}$. Доведемо, що $y \in S$. Спираючись на пункт 1, можемо записати $\rho(y_n, y_0) \le r$, тоді $\lim\limits_{n\to\infty} \rho(y_n, y_0) = \rho(\lim\limits_{n\to\infty}(y_n, y_0)) = \rho(y, y_0) \le r$, тобто $y \in S$.

Отже, для S виконано умови теореми 4.9. Застосуємо принцип стислих відображень для наближеного розв'язування нелінійного рівняння (4.36), яке еквівалентно (4.37). Побудуємо ітераційний процес

$$x_k = \varphi(x_{k-1}), x_0 \in \mathbf{R} \tag{4.45}$$

такий, що $\lim\limits_{k\to\infty} x_k = \beta, f(\beta) = 0.$

Теорема 4.11. (про збіжність ітераційного процесу). Нехай:

1) для рівнянь (4.36) (або (4.37)) існує корінь в околі $S \equiv \left\{x \in R \middle| x-\beta\right| \le r\right\}$;

2) функція $\varphi(x)$ задовольняє умову Ліпшица $\forall x, y \in S$:

$$\left|\varphi(x) - \varphi(y)\right| \le K\left|x - y\right|, 9 < K < 1. \tag{4.46}$$

Тоді

1) $\lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} \varphi(x_{n-1}) = \beta; \forall x_0 \in S;$

2) $\left|x_n - \beta\right| \le K^n \left|x_0 - \beta\right|, \forall x_0 \in S.$

Нагадаємо: Ліпшиць Рудольф (14.05.1832 – 7.10.1902) – німецький математик і фізик. Учень К.Ф. Гауса.

Доведення. Перевіримо умови теореми 4.10.

1) Сукупність точок S утворюють повний метричний простір, якщо визначити

$$\rho(x, y) = \left|x - y\right|, \forall \{x, y\} \subset S.$$

a) $\left|x - y\right| = 0$, якщо $\left|x = y\right|$;

b) $\left|x - y\right| = \left|y - x\right|$;

c) $|x-y| \le |x-z| + |y-z|$.

Множина $\{x_n, n \ge 1\} \subset S$ є фундаментальною послідовністю, бо її границя $x \in S$. Дійсно, $A \equiv \varphi(x): S \to S, \forall x \in S,$ бо $\rho(\varphi(x), \beta) = |\varphi(x) - \beta| = |\varphi(x) - \varphi(\beta)| \le K|x - \beta| \le r,$ звідки $\varphi(x) \in S,$ якщо $x \in S.$ Тому $|x_n - \beta| \le r,$ значить $\lim\limits_{n\to\infty} |x_n - \beta| = \left|\lim\limits_{n\to\infty} x_n - \beta\right| = |x - \beta| < r,$ а це дає $\lim\limits_{n\to\infty} x_n = x \in S.$ Таким чином, довели, що S – повний метричний простір.

2) $\varphi(x)$ – стисле відображення: $\forall x, y \in S$ очевидна нерівність $\rho(\varphi(x), \varphi(y)) = |\varphi(x) - \varphi(y)| \le K|x - y| = K\rho(x, y), 0 < K < 1.$ Довели, що умови теореми 4.10 виконано, а значить $\lim\limits_{n\to\infty} x_n = x = \beta; \forall x_0 \in S; x = \beta \in S.$

Пункт 1) теореми 4.11 доведено. Доведення пункту 2) полягає у виконанні ланцюжка нерівностей

$$|x_n - \beta| = |\varphi(x_{n-1}) - \varphi(\beta)| \le K|x_{n-1} - \beta| \le K|\varphi(x_{n-2}) - \varphi(\beta)| \le K^2|x_{n2} - \beta| \le \ldots \le K^n|x_0 - \beta|.$$

Звідки $x_n \underset{n\to\infty}{\to} \beta$ зі швидкістю геометричної прогресії зі знаменником $K \in (0,1).$

Теорема 4.12. (існування кореня рівняння (4.36) (або 4.37)). Нехай:

1) функція $\varphi(x)$ визначена в околі $S \equiv \{x \in R \mid |x - x_0| \le r\};$

2) значення

$$\forall x, y \in S: |\varphi(x) - \varphi(y)| \le K|x - y|, K \in (0,1); \tag{4.47}$$

3) для $x_0 \in S$ виконується нерівність

$$|x_0 - \varphi(x_0)| \le (1 - K)r. \tag{4.48}$$

Тоді в околі S рівняння (4.36) (або (4.37)) має єдиний корінь $x = \beta,$ причому для ітераційного процесу $y_k = \varphi(y_{k-1}); k = 1, 2, \ldots, \forall y_0 \in S, \lim\limits_{k\to\infty} = \beta.$

Доведення полягає у перевірці умов теореми 4.10.

1) Значення

$$\forall (x) \in S: \rho(\varphi(x), x_0) = |\varphi(x) - x_0| = |\varphi(x) - \varphi(x_0) + \varphi(x_0) - x_0| \le |\varphi(x) - \varphi(x_0)| + |\varphi(x_0) - x_0| \le$$
$$\le K|x - x_0| + (1 - K)r \le r,$$

тобто $\varphi: S \to S.$

2) Відображення $\varphi(x)$ є стислим. Дійсно

$$\forall x, y \in S: \rho(\varphi(x), \varphi(y)) = |\varphi(x) - \varphi(y)| \le K|y - x| = K\rho(x, y).$$

Значить виконано умови теореми 4.10, які гарантують існування єдиного розв'язку β рівняння $x = \varphi(x),$ причому $\beta = \lim\limits_{k\to\infty} y_k, y_0 \in S.$

Зауваження 4.6. Умова Ліпшиця випливає з припущення існування неперервної похідної $\varphi'(x)$ в околі S кореня $x=\beta$: $|\varphi'(x)| \le K < 1$. Тоді за теоремою Лангранжа: $|\varphi(x)-\varphi(y)| = |\varphi'(c)||x-y| \le K|x-y|$, а значить і теореми 4.11 та 4.12 мають місце.

Означення 4.4. Однокроковий стаціонарний ітераційний процес $x_n = \varphi(x_{n-1})$, $n = 1,2,\ldots$, $\forall x_0 \in \mathrm{R}$ називається *методом простої ітерації.*

Зауваження 4.7. Чим менше K, тим швидкість збіжності вище.

4.4.3. Поняття порядку ітерації

Розглянемо рівняння (4.36), яке еквівалентне (4.37). Побудуємо ітераційний процес

$$x_n = \varphi(x_{n-1}), \tag{4.49}$$

$\forall x_0 \in S \equiv \left\{x \in R \middle| |x-\beta| \le r\right\}$, такий, що $\lim\limits_{n\to\infty} x_n = \beta$, тобто $\varphi(\beta) = \beta$ (або $f(\beta) = 0$).

Означення 4.5. Ітерації (4.49) мають порядок m, якщо

$$\varphi'(\beta) = \varphi''(\beta) = .. = \varphi^{(m-1)}(\beta) = 0; \varphi^{(m)}(\beta) \ne 0.. \tag{4.50}$$

Нехай $\varphi(x)$ в околі точки β має неперервні похідні до m-го порядку. За формулою Тейлора в околі β можна записати

$$\begin{aligned}\varphi(x) - \beta = (x-\beta)\varphi'(\beta) + \frac{(x-\beta)^2}{2!}\varphi''(\beta) + \ldots + \\ + \frac{(x-\beta)^{m-1}}{(m-1)!}\varphi^{(m-1)}(\beta) + \frac{(x-\beta)^m}{m!}\varphi^{(m)}(\xi),\end{aligned} \tag{4.51}$$

де ξ з околу точки β.

Якщо ітерації (4.49) мають порядок m, тобто виконується (4.50), то (4.51) набуде вигляду

$$\varphi(x) - \beta = \frac{(x-\beta)^m}{m!}\varphi^{(m)}(\xi). \tag{4.52}$$

Якщо здійснити ітераційний процес (4.48), то з (4.52) матимемо:

$$\varphi(x_{k-1}) - \beta = \frac{(x_{k-1}-\beta)^m}{m!}\varphi^{(m)}(\xi),$$

ξ з околу точки β. Таким чином, $x_k - \beta = \frac{(x_{k-1}-\beta)^m}{m!}\varphi^{(m)}(\xi)$.

Позначимо $M_m \equiv \max\limits_{x\in S}\left|\varphi^{(m)}(x)\right|$, тоді

$$x_k-\beta \le \frac{M_m}{m!}\left|x_{k-1}-\beta\right|^m, k=1,2,\dots \quad . \tag{4.53}$$

Запишемо (4.53), для $k-1, k-2, \dots, 1$:

$$\left|x_{k-1}-\beta\right| \le \frac{M_m}{m!}\left|x_{k-2}-\beta\right|^m, \dots, \; x_1-\beta \le \frac{M_m}{m!}\left|x_0-\beta\right|^m. \tag{4.54}$$

Зворотна підстановка дасть

$$x_k-\beta \le \frac{M_m}{m!}\left|\frac{M_m}{m!}\left|x_{k-2}-\beta\right|^m\right| = \frac{M_m M_m^m}{m!(m!)^m}\left|x_{k-2}-\beta\right|^{m^2}, \dots,$$

тобто

$$\begin{aligned}\left|x_k-\beta\right| &\le \left(\frac{M_m}{m!}\right)^{1+m+m^2+\dots+m^{k-1}}\left|x_0-\beta\right|^{m^k} = \left(\frac{M_m}{m!}\right)^{\frac{m^k-1}{m-1}}\left|x_0-\beta\right|^{m^k} = \\ &= \left(\frac{M_m}{m!}\right)\left|x_0-\beta\right|^{\frac{m^k-1}{m-1}}\left|x_0-\beta\right|^{\frac{m^{k+1}-2m^k+1}{m-1}}\end{aligned} \tag{4.55}$$

Таким чином, довели наступне твердження.

Теорема 4.13. Якщо $\varphi(x), \varphi'(x), \dots, \varphi^{(m)}(x)$ неперервні в S, $\left|x_0-\beta\right|<1$ і $\omega \equiv \frac{M_m}{m!}\left|x_0-\beta\right|<1$, *то* $\left|x_k-\beta\right| \le \omega^{\frac{m^k-1}{m-1}}$, що дає швидку збіжність x_k до β при $k\to\infty$.

4.5. Метод ділення проміжку навпіл (метод дихотомії, Мюллера)

4.5.1. Аналіз постановки задачі

Припустимо, що функція $F(x)$ на відрізку $[a,b]$ неперервна і рівняння $F(x)=0$ має на цьому відрізку єдиний корінь. Нехай $f\in C[a,b]$, $f(a)f(b)<0$ і відомо, що рівняння (4.1) має єдиний корінь $x_*\in[a,b]$. Покладемо $a_0=a$, $b_0=b$, $x_0=(a_0+b_0)/2$.

Якщо $f(x_0)=0$, то $x_*=x_0$. Якщо $f(x_0)=0$, то візьмемо

$$a_{n+1}=\begin{cases} x_n, & \text{якщо } \operatorname{sign} f(a_n)=\operatorname{sign} f(x_n), \\ a_n, & \text{якщо } \operatorname{sign} f(a_n)\ne\operatorname{sign} f(x_n), \end{cases} \tag{4.56}$$

$$b_{n+1} = \begin{cases} x_n, & \text{якщо } \operatorname{sign} f(b_n) = \operatorname{sign} f(x_n), \\ b_n, & \text{якщо } \operatorname{sign} f(b_n) \neq \operatorname{sign} f(x_n), \end{cases} \tag{4.57}$$

$$x_{n+1} = \frac{a_{k+1} + b_{k+1}}{2}, \quad n = 0,1,2,\ldots,, \tag{4.58}$$

і обчислимо $f(x_{n+1})$. Якщо $f(x_{n+1}) = 0$, то ітераційний процес зупинимо і будемо вважати, що $x_* \approx x_{n+1}$. Якщо $f(x_{n+1}) \neq 0$, то повторюємо розрахунки за формулами (4.56)-(4.58).

З формул (4.56), (4.57) видно, що $\operatorname{sign} f(a_{n+1}) = \operatorname{sign} f(a_n)$ і $\operatorname{sign} f(b_{n+1}) = \operatorname{sign} f(b_n)$. Тому $f(a_{n+1}) f(b_{n+1}) < 0$, а отже шуканий корінь x_* знаходиться на проміжку $[a_{n+1}, b_{n+1}]$. При цьому має місце оцінка збіжності

$$|x_n - x_*| \leq \frac{b-a}{2^{n+1}}. \tag{4.59}$$

Звідси випливає, що кількість ітерацій. які необхідно провести для знаходження наближеного кореня рівняння (4.1) із заданою точністю ε задовольняє співвідношенню

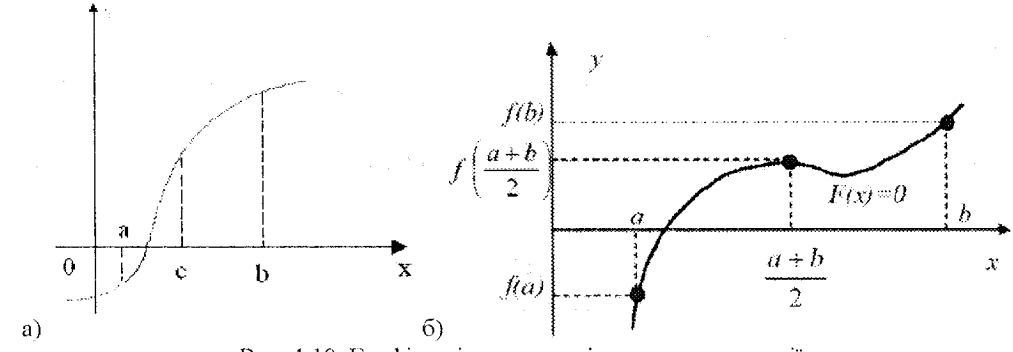
$$n \geq \left\lfloor \log_2 \frac{b-a}{\varepsilon} \right\rfloor. \tag{4.60}$$

де $\lfloor c \rfloor$ – ціла частина числа c.

Серед переваг цього методу слід відзначити простоту реалізації та надійність. Послідовність $\{x_n\}$ збігається до кореня x_* для довільних неперервних функцій $f(x)$. До недоліків можна віднести невисоку швидкість збіжності методу та неможливість безпосереднього узагальнення систем нелінійних рівнянь.

4.5.2. Геометрична інтерпретація методу

Нехай рівняння (4.1) на відрізку $[a,b]$ має єдиний корінь і функція на ньому неперервна. Поділимо відрізок $[a,b]$ пополам точкою $c=(a+b)/2$ (рис. 4.10).



а) б)

Рис. 4.10. Графічна інтерпретація методу дихотомії

Якщо $f(c) \neq 0$, то можливі два випадки (рис. 4.10б):

- або $f(x)$ змінює знак на $[a,c]$,
- або $f(x)$ змінює знак на $[c,b]$.

Тобто $f \in C[a,c]$, $f(a)f(c)<0$ або $f \in C[c,b]$, $f(c)f(b)<0$ і далі ділимо обраний відрізок, вибираючи в кожному випадку той з відрізків, на якому функція змінює знак. Продовжуючи процес поділу далі, можна дійти до скільки завгодно малого проміжку, що містить корінь. На кожному кроці процесу половинного ділення відрізок зменшується вдвічі.

4.5.3. Обчислення похибки

Ітераційний процес продовжуємо доти, поки довжина відрізка не стане менше заданої точності ε :

$$|a-b|<\varepsilon. \tag{4.61}$$

Як розв'язок приймаємо ліву (a) чи праву (b) границю інтервалу. Якщо як відповідь обрати точку в середині проміжку, то точний результат не може відхилятися від неї більш ніж на $\varepsilon/2$, тому для досягнення заданої точності ε потрібно вийти з ітераційного процесу, коли виконується умова $|a-b|<2\cdot\varepsilon$.

4.6. Метод хорд

Метод хорд – один з поширених ітераційних методів. Його ще називають *методом лінійного інтерполювання*, *методом пропорційних частин* або *методом хибного положення*. При виконанні попереднього методу більш раціональним було б ділити $[a,b]$ не навпіл, а пропорційно значенню функції в точках a та b. В цьому випадку точка поділу відрізка буде знаходитися на перетині хорди AB з віссю $0x$. Ідея методу хорд в тому, що на досить малому відрізку дуга кривої $y=f(x)$ замінюється хордою, і абсциса точки перетину хорди з віссю Ox є наближеним значенням кореня.

4.6.1. Аналіз постановки задачі

Ідея методу хорд полягає в тому, що на достатньо малому проміжку $[a,b]$ дуга кривої $y=f(x)$ замінюється стягуючою її хордою. Шуканий корінь рівняння $f(x)=0$ є абсциса точки перетину графіка функції $y=f(x)$ з віссю Ox. Ця точка нам невідома, але замість неї ми візьмемо точку x_1 перетину хорди AB із віссю Ox.

Розглянемо випадок, коли перша і друга похідні мають однакові знаки, тобто $f'(x)\cdot f''(x)>0$ (рис. 4.11а). У якості нульового наближення кореня в даному випадку приймається ліва межа інтервалу ізольованого кореня, тобто $x_0=a$.

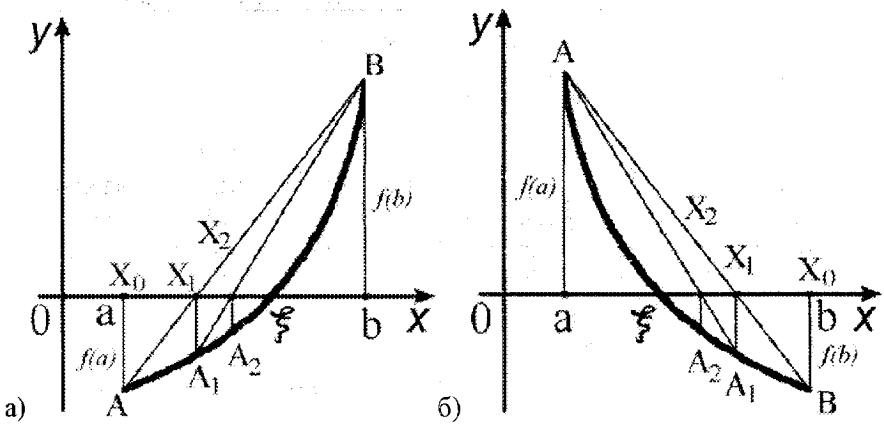


Рис. 4.11. Графічна інтерпретація методу хорд

Перше, друге та інші наближення кореня знаходимо з формули, яку отримуємо з рівняння хорди. Рівняння хорди AB записується як рівняння прямої, що проходить через дві точки з відомими координатами

$$\frac{x-a}{b-a}=\frac{y-f(a)}{f(b)-f(a)}. \tag{4.62}$$

Значення $x=x_1$, для якого $y=0$, тобто точка перетину хорди з віссю абсцис, розташовується ближче до точного значення кореня, ніж $x_0=a$, і визначається з виразу

$$x_1=a-f(a)\frac{b-a}{f(b)-f(a)}. \tag{4.63}$$

Обчислимо значення $f(x_1)$. Геометрично $f(x_1)$ – довжина перпендикуляра до осі Ox, проведеного з точки до кривої $f(x)$. Якщо $f(x_1)<0$, то ми знайшли більш вузький інтервал існування кореня $[x_1,b]$, оскільки знаки $f(x_1)$ і $f(a)$ збігаються. Тепер корінь знаходиться у середині відрізка $[x_1,b]$. Якщо значення кореня x_1 нас не влаштовує, то його можна уточнити, застосовуючи метод хорд до відрізка $[x_1,b]$. Тобто будуємо хорду A_1B, записуючи її рівняння та визначаючи точку перетину x_2 хорди A_1B із віссю абсцис

$$x_2=x_1-f(x_1)\frac{b-x_1}{f(b)-f(x_1)}, \tag{4.64}$$

та ін.

1. Якщо мають місце варіанти I і II (табл. 4.1), тоді $f'(b)\cdot f''(x)>0$ на відрізку $[a,b]$, то наближені значення коренів $x_1,x_2,\ldots,x_n$ будуть знаходитися всередині відрізків $[x_1,b]$, $[x_2,b]$, …, тобто нерухомим кінцем відрізка $[a,b]$ буде точка b, а наближені значення коренів будуть знаходитися за формулою

$$x_{n+1} = x_n - f(x_n)\frac{b - x_n}{f(b) - f(x_n)}, \qquad (4.65)$$

при цьому $x_0 = a$ (рис. 4.11а).

2. Якщо мають місце варіанти III і IV (табл. 4.1), тоді $f'(a) \cdot f''(x) > 0$ на відрізку $[a, b]$, то наближені значення $x_1, x_2, \ldots, x_n$ коренів будуть знаходитися усередині відрізків $[x_1, a]$, $[x_2, a]$, ..., тобто нерухомим кінцем відрізка $[a, b]$ буде кінець a, а наближені значення коренів будуть знаходитися за формулою

$$x_{n+1} = x_n - f(x_n)\frac{x_n - a}{f(x_n) - f(a)}, \qquad (4.66)$$

при цьому $x_0 = b$ (рис. 4.11б).

Вибір тих або інших формул можна здійснити, користуючись простим правилом: нерухомим кінцем відрізка є той, для якого знак функції збігається зі знаком другої *похідної, а нульове наближення* x_0 *вибирається відповідно до умови*

$$f'(x_0) \cdot f''(x_0) > 0. \qquad (4.67)$$

Процес послідовного наближення до кореня слід продовжувати доти, поки не буде виконана умова $|x_n - x_{n-1}| \le \varepsilon$, де ε - задана точність; x_n і x_{n-1} - наближення, отримані на n-му та $(n-1)$-му кроках. При цьому уточнене значення кореня приймається $\xi = x_n \pm \varepsilon$.

4.6.2. Геометрична інтерпретація методу

Даний метод грунтується на лінійній інтерполяції функції (4.1) по двох значеннях, що мають протилежні знаки у значення $f(a)$ та $f(b)$ (рис. 4.12).

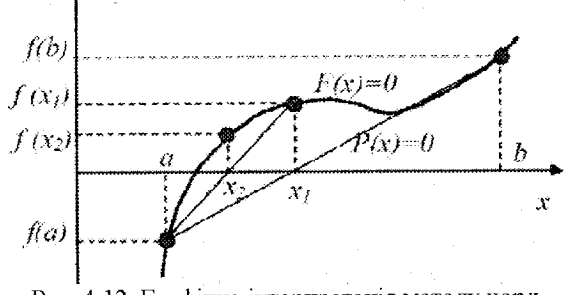


Рис. 4.12. Графічна інтерпретація методу хорд

Нехай потрібно знайти корінь рівняння (4.1) на проміжку $[a,b]$, і відомо, що $F(x)$ неперервна на $[a,b]$ та $f(a)f(b) < 0$. Крім того, перша $F'(x)$ і друга $F''(x)$

похідні функції $F(x)$ зберігають на проміжку $[a,b]$ свій знак. Замінимо функцію $F(x)$ лінійною функцією, яка проходить через вузлові точки $(a, f(a))$ та $(b, f(b))$:

$$\frac{y-f(a)}{f(b)-f(a)}=\frac{x-a}{b-a}\Rightarrow P(x)=y=f(a)+\frac{x-a}{b-a}(f(b)-f(a)). \qquad (4.68)$$

Лінійна функція $P(x)$ на кінцях відрізку $[a,b]$ приймає такі ж значення, як і функція $F(x)=0$. В якості першого наближення при знаходженні кореня функції $F(x)=0$ візьмемо точне значення кореня функції $P(x)=0$, тобто x_1, яке визначимо з рівняння:

$$f(a)+\frac{x_1-a}{b-a}(f(b)-f(a))=0\Rightarrow x_1=a-f(a)\frac{b-a}{f(b)-f(a)}. \qquad (4.69)$$

4.6.3. Обчислення похибки

При подальшому дослідженні відрізків $[a,x_1]$ та $[x_1,b]$, виберемо той, на якому функція змінює знак. На рис 4.12 таким відрізком є $[a,x_1]$. Для обраного відрізка побудуємо лінійне наближення функції та виконаємо розрахунки кореня для лінійного наближення за формулою (4.69). В результаті розрахунків отримаємо координати точки x_2. Ітераційни процес закінчуємо, коли $|F(x_i)|<\varepsilon$.

Приклад 4.7. Знайти корінь рівняння $x^3-12.2x^2+7.45x+42=0$ на відрізку [10, 12].

Розв'язок. Обчислюємо значення функції на кінцях відрізка: $f(10)=-103.5; f(12)=102.6$. Оскільки $f''(x)>0$, то за нульове наближення приймаємо x_0=10 та обчислення будемо проводити за формулою (4.25).

$$x_1=x_0-f(x_0)\frac{b-x_0}{f(b)-f(x_0)}=10+103.5\frac{(12-10)}{102.6+103.5}\cong 11.$$

Скористаємося схемою Горнера для обчислення значення поліному $f(x)$ у точці x_1=11.

	1	-12.2	7.45	42
11		11.0	-13.2	-63.25
	1	-1.2	-5.75	*-21.25*

Отже, $f(11)=-21.25$. Це говорить про те, що істинний корінь розташований в інтервалі [11,12]. Повторюючи процес для визначення другого наближення кореня, одержимо $x_2=11.17$, для якого значення функції $f(11.17)=-3.55$. Тепер корінь знаходиться в інтервалі [11.17, 12]. Нарешті, третє наближення дає $x_3=11.2$, для якого $f(11.2)=0$. Таким чином, $x_3=\xi=11.2$. Тобто в цьому прикладі на третьому кроці ми отримали точне значення кореня.

4.7. Метод Ньютона (метод дотичних)

4.7.1. Аналіз постановки задачі

Метод Ньютона застосовується до розв'язування задачі (4.1), де $f(x)$ є неперервно-диференційованою функцією. На початку обчислень вибирається початкове наближення x_0. Наступні наближення обчислюються за формулою

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0,1,2,..., \quad f'(x_n) \neq 0. \tag{4.70}$$

З геометричної точки зору x_{n+1} є значенням абсциси точки перетину дотичної до кривої $y=f(x)$ в точці $(x_n, f(x_n))$ з віссю абсцис. Тому метод Ньютона називають також методом дотичних.

Теорема 4.14. Якщо $f(x) \in C^2[a,b]$, $f(a)f(b) < 0$, а $f''(x)$ не змінює знак на $[a,b]$, то виходячи з початкового наближення $x_0 \in [a,b]$, що задовольняє умові $f(x_0)f''(x_0) > 0$, можна обчислити методом Ньютона єдиний корінь x_* рівняння (4.1) з будь-якою степінню точності.

Теорема 4.15. Нехай x_* – простий дійсний корінь рівняння (4.1) і $f(x) \in C^2(S)$, де $S = \{x : |x - x_*| \leq \delta\}$,

$$0 < m_1 = \min_{x \in S} |f'(x)|, \quad M_2 = \max_{x \in S} |f''(x)|, \tag{4.71}$$

$$q = \frac{M_2 |x_0 - x_*|}{2m_1} < 1. \tag{4.72}$$

Тоді для $x_0 \in S$ метод Ньютона збігається, причому для похибки справедлива оцінка

$$|x_n - x_*| \leq q^{2^n - 1} |x_0 - x_*|. \tag{4.73}$$

З оцінки (4.73) видно, що метод Ньютона має квадратичну збіжність, тобто похибка на $(n+1)$-й ітерації пропорційна квадрату похибки на n-й ітерації.

4.7.2. Геометрична інтерпретація методу

Ідея методу полягає в послідовній заміні ділянки кривої $f(x)$ дотичною в точці c, що належить відрізку $[a,b]$ і перетинає вісь $0x$ в точці x_k. Точка c вибирається з умови: $f(c)*f''(c)>0$, яка гарантує збіжність процесу. При цьому необхідно, щоб

- $f(x)=0$ мало єдиний корінь на $[a,b]$;
- $f(x)$ була неперервна на $[a,b]$.
- $f'(x)$ і $f''(x)$ не змінювали на ньому знак.

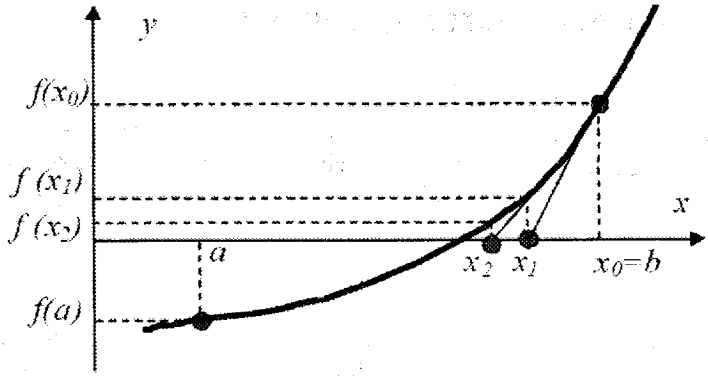


Рис. 4.13. Графічна інтерпретація методу дотичних

Геометрично метод Ньютона еквівалентний заміні невеликої дуги кривої $y = f(x)$ дотичною, яка проведена до деякої точки кривої, наприклад, точки B (рис. 4.14). Точка перетину цієї дотичної з віссю абсцис дає нам перше наближення x_1 кореня ξ. Рівняння дотичної до кривої у точці B має наступний вигляд

$$y - f(x_0) = f'(x_0)(x - x_0). \tag{4.74}$$

Вважаючи $y = 0$, знаходимо абсцису x_1 точки перетину цієї дотичної з віссю Ox:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}. \tag{4.75}$$

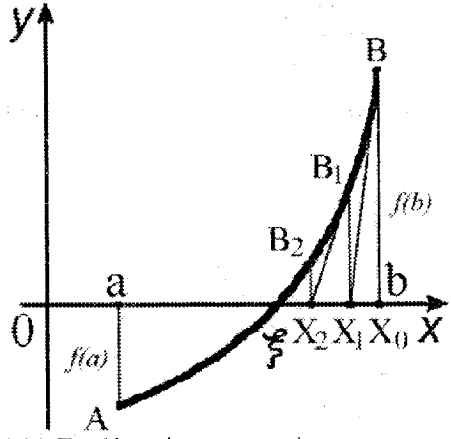


Рис. 4.14. Графічна інтерпретація методу дотичних

Значення x_1 приймаємо в якості першого наближення кореня ξ. В точці x_1 визначаємо значення функції $f(x)$. Геометрично – це довжина перпендикуляра до осі Ox, відновленого в x_1 до його перетину з кривою $y = f(x)$ в точці B_1.

Процес продовжується таким чином: проведемо дотичну до точки B_1 і знайдемо точку перетину цієї дотичної з віссю абсцис x_2. Точку x_2 приймаємо в якості другого наближення кореня ξ. Процес продовжується доти, поки ми не одержимо значення

кореня з заданим ступенем точності. Будь-яке $(n+1)$-ше наближення кореня визначається рівністю

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{4.76}$$

Вибір нульового наближення кореня x_0 здійснюється наступним чином:

- якщо $f(a)\cdot f''(x) > 0$ на $[a,b]$, то $x_0 = a$;
- якщо $f(b)\cdot f''(x) > 0$ на $[a,b]$, то $x_0 = b$.

4.7.3. Обчислення похибки

Чим більше чисельне значення похідної $f'(x)$ в околі цього кореня, тим менша похибка, яку необхідно враховувати в n-му наближенні. Тому метод Ньютона особливо зручно застосовувати тоді, коли в околі цього кореня графік функції має велику крутизну. Якщо чисельне значення похідної біля кореня мале, то похибки будуть великими і процес уточнення кореня може виявитися тривалим. Якщо крива поблизу точки перетину з віссю Ox майже горизонтальна, то застосовувати метод Ньютона не рекомендується. Точність наближення на n-му кроці оцінюється таким чином: якщо $\left|x_n - x_{n-1}\right| \le \varepsilon$, то $\left|\xi - x_n\right| < \varepsilon^2$.

Приклад 4.8. Методом дотичних уточнити до $\varepsilon = 0.001$ корінь ξ рівняння $x^3 - 3x^2 - 3 = 0$, який розташований на відрізку [-2. 75, -2. 5].

Розв'язок. За умовою $f(x) = x^3 - 3x^2 - 3$. Визначаємо другу похідну $f''(x)$:

$$f'(x) = 3x^2 + 6x; \quad f''(x) = 6x + 6.$$

Таким чином, $f(-2.75)\cdot f''(x) > 0$, тому $x_0 = -2.75$.

Визначаємо значення першої похідної у точці x_0:

$$f'(x_0) = f'(-2.75) = 6.1875.$$

Для зручності подальші обчислення зводимо в табл. 4.3.

Таблиця 4.3

n	x_n	x_n^2	x_n^3	$3x_n^2$	$f(x_n)$	$-\frac{f(x_n)}{6.1875}$
0	-2.75	-20.797	7.5625	22.6875	-1.111	0.179
1	-2.571	-16.994	6.6100	19.8300	-0.164	0.026
2	-2.545	-16.484	6.4770	19.431	-0.053	0.008
3	-2.537	-16.329	6.4364	19.309	0.020	0.003
4	-2.534	-16.271	6.4212	19.2636	0.007	0.001
5	-2.533					

Остаточно одержимо $\xi = -2.533 \pm 0.001$.

4.7.4. Інтерпретація методу дотичних через ітераційний процес

Теорема 4.16. (Метод Ньютона). Нехай

1) $x = \beta$ корінь рівняння $f(x) = 0$.

2) $f \in C^2(S)$ для $x \in S \equiv \{x \in \mathbf{R} \,||\, x - \beta| \le 2\}$;

3) похідні першого та другого порядку функції *f(x)* не змінюють знаку $f'(x) > 0 (< 0), f''(x) > 0 (< 0), \forall x \in S$;

4) $f(x_0) f''(x_0 > 0)$

5) Для $\varepsilon > 0, |x_0 - \beta| < \varepsilon, x_0 \in S$.

Тоді

1) при $\psi(x) \equiv \dfrac{1}{f'(x)}$ для рівняння $x \equiv x - \psi(x) f(x)$ метод ітерацій прийме вигляд

$$\left\{ x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}, n \ge 1 \right\} \subset S, \qquad \lim_{x \to \infty} x_n = \beta, \forall x_0 \in S; \tag{4.77}$$

2) та похибка обчислень

$$|x_n - \beta| \le \frac{M_2}{2m_1} |x_{n-1} - \beta|^2, \tag{4.78}$$

де $M_2 \equiv \max_{x \in S} |f''(x)|, m_1 \equiv \min_{x \in S} |f'(x)|$.

Доведення. 1) Зауважимо, що $x = \beta$ є коренем рівняння (4.2). Дійсно, $\beta = \beta - \dfrac{f(\beta)}{f'(\beta)}$, де $f(\beta) = 0$. Оскільки $\varphi'(x) = \dfrac{f'(x) f'(x) - f''(x) f(x)}{[f'(x)]^2}$, тоді $\varphi'(\beta) = 0$. Виконано умови теореми 4.11, $|\varphi'(x)| \le K < 1$. Значить за теоремою 4.10 $\lim_{n \to \infty} x_n = \beta$.

2) За умови 2) теореми 4.8 за формулою Тейлора в околі x_n маємо $f(x) = f(x_n) + f'(x_n)(x - x_n) + \dfrac{1}{2} f''(\xi)(x - x_n)^2$, де ξ з околу точки x_n. Підставимо в цю рівність $x = \beta$: $0 = f(x_n) + f'(x_n)(\beta - x_n) + \dfrac{1}{2} f''(\xi)(\beta - x_n)^2$. Звідси випливає

$$\frac{\frac{1}{2}f''(\xi)(\beta-x_n)^2}{f'(x_n)} = x_n - \beta - \frac{f(x_n)}{f'(x_n)}. \qquad (4.79)$$

Ітераційний процес (4.77) при $x_n=x_{n+1}$ набуде вигляду:

$$x_{n+1} - \beta = x_n - \beta - \frac{f(x_n)}{f'(x_n)},$$

звідки з (4.79) випливає $x_{n+1} - \beta = \frac{1}{2}\frac{f''(\xi)}{f'(\xi)}(x_n - \beta)^2$, що доводить (4.78).

Наслідок. 4.1 *(геометричний зміст методу дотичних).* Геометричний зміст методу дотичних полягає в тому, що значення $\{x_n, n \ge 1\}$ (4.77) є абсцисою точки перетину з віссю Ox дотичної до кривої $y = f(x)$ в точці $(x_{n-1}, f(x_{n-1}))$ (рис. 4.15).

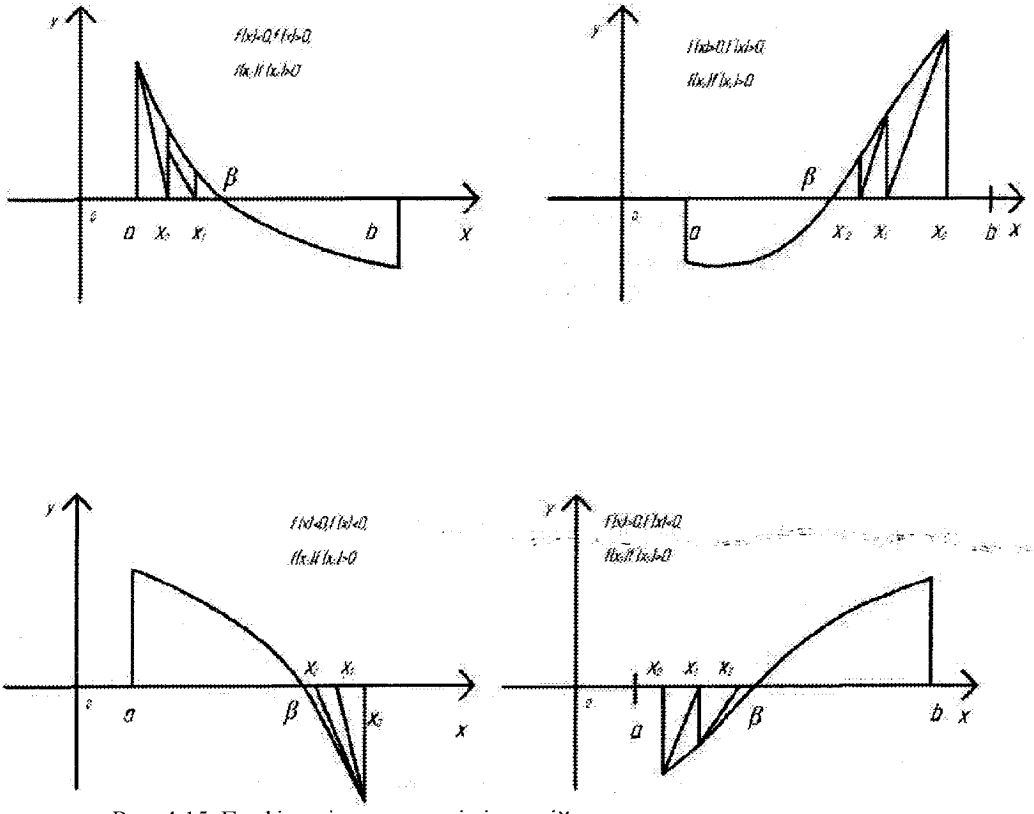


Рис. 4.15. Графічна інтерпретація ітераційного процесу методу дотичних

Доведення. Запишемо рівняння дотичної в точці $(x_{n-1}, f(x_{n-1}))$:

$$f(x) - f(x_{n-1}) = f'(x_{n-1})(x - x_{n-1}),$$

де покладемо $x = x_n, f(x_n) = 0$ (x_n – точка перетину дотичної з віссю Ox)

$$-f(x_{n-1}) = f'(x_{n-1})(x_n - x_{n-1}).$$

Звідки $x_n = x_{n-1} - \dfrac{f(x_{n-1})}{f'(x_{n-1})}$.

Зауваження 4.8 Метод Ньютона ітераційний метод другого порядку з $\varphi'(\beta) = 0; \varphi''(\beta) \neq 0$.

4.7.4. Модифікований метод Ньютона

Цей метод аналогічний методу дотичних за однією відмінністю – кожна дотична будується з початкової точки. Якщо похідна $f'(x)$ мало змінюється на відрізку $[a, b]$, то для спрощення обчислень можна використовувати формулу

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)}, \quad n = 0,1,2,\ldots, \tag{4.80}$$

тобто значення похідної в початковій точці достатньо обчислити один раз.

Геометрично це означає, що дотичні в точках $B_n(x_n; f(x_n))$ заміняємо прямими, які паралельні до дотичної, що проведена до кривої $y = f(x)$ у точці $B(x_0; f(x_0))$ (рис. 4.14). Метод дозволяє не обчислювати похідну $f'(x)$ на кожній ітерації, а отже і позбутися можливого ділення на нуль. Однак, цей алгоритм має тільки лінійну збіжність. Кількість ітерацій, які потрібно виконати для знаходження розв'язку задачі (4.1) з точністю ε задовольняє нерівності:

$$n \geq \left[\log_2 \left(\frac{\ln(|x_0 - x_*|/\varepsilon)}{\ln(1/q)} \right) + 1 \right] + 1. \tag{4.81}$$

4.8. Комбінований метод хорд і дотичних

4.8.1. Аналіз постановки задачі

Методи хорд і дотичних дають наближення кореня з різних сторін (більше або менше істинного значення кореня), тому їх часто застосовують один з одним, і уточнення кореня відбувається швидше. Якщо $f'(x) \cdot f''(x) > 0$, то метод хорд дає наближення кореня з недостачею, а метод дотичних – з надлишком. Якщо ж $f'(x) \cdot f''(x) < 0$, то методом хорд одержуємо значення кореня з надлишком, а методом дотичних – із недостачею. Проте в усіх випадках істинне значення кореня знаходиться між наближеними значеннями коренів, що отримуються методом хорд і методом дотичних.

Нехай x_{n+1} і $\bar{x}_{n+1}$ – наближені значення кореня з недостачею і з надлишком.

1. Якщо $f(a) \cdot f''(x) > 0$ на $[a.b]$ (при цьому $x_0 = a, \quad \bar{x}_0 = b$), то

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)};\ \bar{x}_{n+1} = \bar{x}_n - f(\bar{x}_n)\frac{\bar{x}_n - x_n}{f(\bar{x}_n) - f(x_n)}, \tag{4.82}$$

2. Якщо $f(b)\cdot f''(x) > 0$ на $[a,b]$ (при цьому $x_0 = a, \quad \bar{x}_0 = b$), то

$$x_{n+1} = x_n - f(x_n)\frac{\bar{x}_n - x_n}{f(\bar{x}_n) - f(x_n)};\ \bar{x}_{n+1} = \bar{x}_n - \frac{f(\bar{x}_n)}{f'(\bar{x}_n)}, \tag{4.83}$$

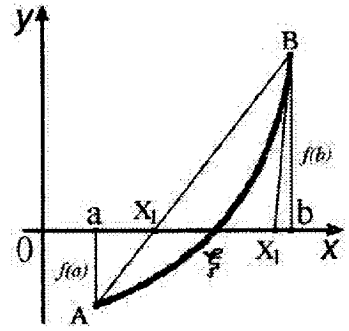


Рис. 4.16. Графічна інтерпретація методу дотичних

4.8.2. Геометрична інтерпретація методу

Нехай $x = \beta \in (a, b)$ єдиний корінь рівняння (4.1) та виконуються умови теорем 4.1 та 4.2. Якщо $f(a)f'(a) > 0$, то $x_0 = a - \frac{f(a)}{f'(a)}; x_1 = \frac{af(b) - bf(a)}{f(b) - f(a)}$. Якщо $f(b)f'(b) > 0$, то $x_0 = b - \frac{f(b)}{f'(v)}; x_1 = \frac{af(b) - bf(a)}{f(b) - f(a)}$. Значить, можна наближатися до кореня, наприклад, зліва методом Ньютона , а справа методом січних (рис.4.17):

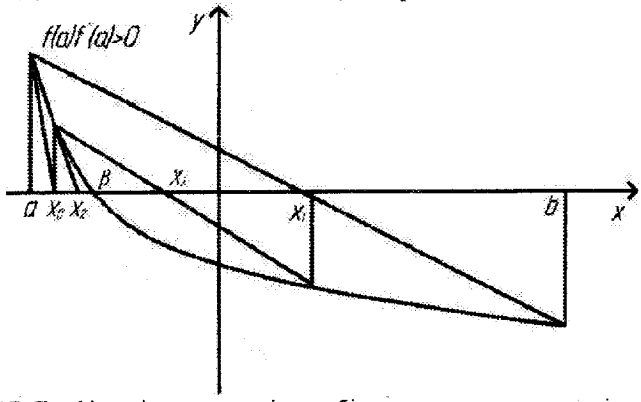


Рис. 4.17. Графічна інтерпретація комбінованого методу хорд і дотичних

$$x_{2n} = x_{2n-2} - f\frac{f(x_{2n-2})}{f'(x_{2n-2})}, \tag{4.84}$$

$$x_{2n+1} = \frac{x_{2n-2} f(x_{2n-1}) - x_{2n-1} f(x_{2n-2})}{f(x_{2n-1}) - f(x_{2n-2})}, n = 1, 2... \tag{4.85}$$

Очевидно, що $\beta \in [x_n, x_{n-1}]$ і $\lim\limits_{x\to\infty} x_{2n} = \lim\limits_{x\to\infty} x_{2n+1} = \beta$.

4.8.3. Обчислення похибки

Процес обчислень припиняється, як тільки буде виконуватися нерівність $|x_n - \bar{x}_n| \le \varepsilon$. Значення кореня, який є уточненим, буде $\xi = \frac{1}{2}(x_n + \bar{x}_n)$.

4.9. Метод січних

4.9.1. Аналіз постановки задачі

Метод січних подібний до методу хорд, тільки точки з координатами $(x_0, F(x_0))$ та $(x_1, F(x_1))$ взяті з одного боку від кореня рівняння $F(x) = 0$. Метод січних отримуємо з методу Ньютона заміною похідної на розподілені різниці:

$$f(x_{n-1}, x_n) = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}. \tag{4.86}$$

Отримуємо алгоритм

$$x_{n+1} = x_n - \frac{f(x_n)}{f(x_n, x_{n-1})} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \cdot f(x_n), \quad n = 1, 2, \ldots, \tag{4.87}$$

де x_0, x_1 – задані. В результаті отримано ітераційний метод на два кроки.

4.9.2. Геометрична інтерпретація методу

Геометрична інтерпретація методу подана на рис. 4.18.

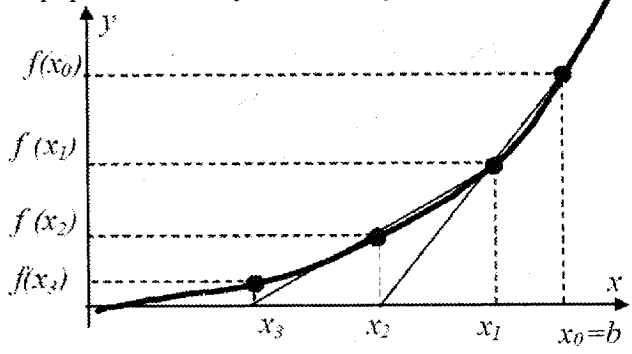


Рис. 4.18. Графічна інтерпретація методу січних

Як початкове наближення обираємо точки $(x_0, F(x_0))$ та $(x_1, F(x_1))$, що розташовані від краю відрізка, наприклад, на $1/10$ його довжини. Через точки $(x_0, F(x_0))$ та $(x_1, F(x_1))$ проводимо січну до графіку функції, яка перетинає вісь x в точці $(x_2, 0)$. Перевіряємо виконання умови $|f(x_2)| < \varepsilon$, якщо вона не виконується, проводимо січну через точки $(x_1, F(x_1))$ та $(x_2, F(x_2))$, знаходимо точку перетину січної з віссю x (точка $(x_3, 0)$) і перевіряємо виконання чергової умови $|f(x_3)| < \varepsilon$.

4.9.3. Обчислення похибки

Ітерації продовжуємо до виконання умови виходу з ітераційного процесу,тобто $|f(x_n)| < \varepsilon$. Загальна формула для знаходження точки перетину січної з віссю x:

$$x_{n+1} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}, \quad n = 1,2,\ldots. \tag{4.88}$$

4.9.4. Інтерпретація методу січних через ітераційний процес

Нехай $x = \beta$ корінь рівняння (4.1), тобто $f(\beta) = \beta$ функція $\psi \in C(S)$, де $S \equiv \{x \in \mathbf{R} \,||\, x - \beta| < \varepsilon\}$. Тоді еквівалентне рівняння (4.2) прийме вигляд

$$x = \varphi(x) = x - \psi(x) f(x). \tag{4.89}$$

і також має корінь $x = \beta$. Дійсно, підставимо в (4.89) $x = \beta$:

$$\beta = \beta - \psi(\beta) f(\beta) = \beta.$$

Теорема 4.17. Нехай:

1) $x = \beta$ корінь рівняння $f(x)=0$;

2) для $\forall x \in S : f \in C^2(S)$;

3) похідні першого та другого порядку функції $f(x)$ не змінюють знаку: $f'(x) > 0 (< 0); f''(x) > 0 (< 0)$;

4) $f(x_0) f''(x_0) > 0$ в області $|x_0 - \beta| < \varepsilon, x_0 \in S, \varepsilon > 0$.

Тоді:

1) для

$$\psi(x) \equiv \frac{x - x_0}{f(x) - f(x_0)}. \tag{4.90}$$

Для рівняння (4.89) метод ітерацій матиме вигляд

$$x_n = \frac{x_0 f(x_{n-1}) - x_{n-1} f(x_0)}{f(x_{n-1}) - f(x_0)}, n \geq 2. \tag{4.91}$$

(надалі назвемо (4.90) методом січних), причому

$$\lim_{x\to\infty} x_n = \beta; \forall x_0, x_1 \in S, \text{ де } f(x_0)f(x_1) < 0; \tag{4.92}$$

2) вірна оцінка

$$|x_n - \beta| \le \frac{|f(x_n)|}{m}. \tag{4.93}$$

$$m \equiv \min_{x\in S} |f'(x)|. \tag{4.94}$$

Доведення. Нехай $x = \beta$ корінь рівняння (4.1). Підставимо (4.90) у (4.89).

$$x = \varphi(x) \equiv x - \psi(x)f(x) = x - \frac{x - x_0}{f(x) - f(x_0)} f(x) = \frac{x_0 f(x) - xf(x_0)}{f(x) - f(x_0)}. \tag{4.95}$$

При підстановці β в праву частину (4.95) маємо

$$\frac{x_0 f(\beta) - \beta f(x_0)}{f(\beta) - f(x_0)} = \beta, \text{ бо } f(\beta) = 0.$$

Доведемо, що визначена за (4.91) послідовність $\{x_n, n \ge 1\}$ збігається. З одного боку згідно (4.95):

$$\varphi'(x) = \frac{[x_0 f'(x) - f(x_0)][f(x) - f(x_0)] - f'(x)[x_0 f(x) - xf(x_0)]}{[f(x) - f(x_0)]^2},$$

Звідки

$$\frac{f(x_0) + (\beta - x_0)}{f(x_0)} = \varphi'(\beta). \tag{4.96}$$

З іншого боку, за формулою Тейлора

$$f(x) = f(\beta) + (x - \beta)f'(\beta) + \frac{(x-\beta)^2}{2!} f''(\xi). \tag{4.97}$$

Тоді для $x=x_0$ одержуємо

$$f(x_0) = (\beta - x_0)f'(\beta) = \frac{(x_0 - \beta)^2}{2!} f''(\xi). \tag{4.98}$$

Підставимо (4.98) у (4.96):

$$\varphi'(\beta) = \frac{(x_0 - \beta)^2 f''(\xi)}{2!f(x_0)}. \tag{4.99}$$

Якщо $|x_0-\beta|<\varepsilon$, $|\varphi'(\beta)|\le K\le 1$ за неперервністю $\varphi'(\mathrm{x})$ в S. За теоремою 4.10 отримаємо $\lim\limits_{x\to\infty} x_n=\beta$. Далі, за теоремою Лагранжа $f(\mathrm{x})-\mathrm{f}(\beta)=\mathrm{f}'(\xi)(\mathrm{x}-\beta)$, що для $x=x_n$, дає $f(\mathrm{x}_n)=\mathrm{f}'(\beta)(\mathrm{x}_n-\beta)$. Якщо $m\equiv\min\limits_{x\in S}|f'(\mathrm{x})|$, то $|f(\mathrm{x}_n)|\ge m|x_n-\beta|$, тобто

$$|x_n-\beta|\le\frac{|f(\mathrm{x}_n)|}{m}. \tag{4.100}$$

Наслідок 4.2 (геометричний зміст методу січних). Геометричний зміст методу січних полягає в тому, що значення x_{n+1} є абсциса перетину прямої, що проходить через точки $(\mathrm{x}_0,\mathrm{f}(\mathrm{x}_0))$ і $(\mathrm{x}_n,\mathrm{f}(\mathrm{x}_n))$, з віссю Ox. Тому цей метод (4.91) називають методом січних (рис. 4.19).

Приклад 4.9. Одержати формулу (4.91) із геометричних міркувань.

Розв'язок. Проведемо пряму через точки $(\mathrm{x}_0,\mathrm{f}(\mathrm{x}_0))$ і $(\mathrm{x}_n,\mathrm{f}(\mathrm{x}_n))$

$$\frac{x-x_0}{x_n-x_0}=\frac{f(\mathrm{x})-\mathrm{f}(\mathrm{x}_0)}{f(\mathrm{x}_n)-\mathrm{f}(\mathrm{x}_0)}. \tag{4.101}$$

Пряма (4.101) перетне Ox в точці x_{n+1}, де $f(\mathrm{x}_{n+1})=0$, а (4.101) набуде вигляду

$$\frac{x_{n+1}-x_0}{x_n-x_0}=\frac{\mathrm{f}(\mathrm{x}_0)}{f(\mathrm{x}_n)-\mathrm{f}(\mathrm{x}_0)}.$$

звідки матимемо (4.91).

Приклад 4.10. Показати, що метод (4.91) є методом лінійної інтерполяції.

Розв'язок. Побудуємо многочлен Лагранжа за точками $(\mathrm{x}_0,\mathrm{f}(\mathrm{x}_0)),(\mathrm{x}_n,\mathrm{f}(\mathrm{x}_n))$:

$$L_1(\mathrm{x})=\mathrm{f}(\mathrm{x}_0)\frac{x-x_n}{x_0-x_n}+f(\mathrm{x}_n)\frac{x-x_n}{x_n-x_0}.$$

Легко бачити, що $x_{n+1}=\dfrac{x_0 f(\mathrm{x}_n)-\mathrm{x}_n f(\mathrm{x}_0)}{f(\mathrm{x}_n)-\mathrm{f}(\mathrm{x}_0)}$ є коренем многочлена Лагранжа $L_1(\mathrm{x})$, тобто $L(\mathrm{x}_{n+1})=0$.

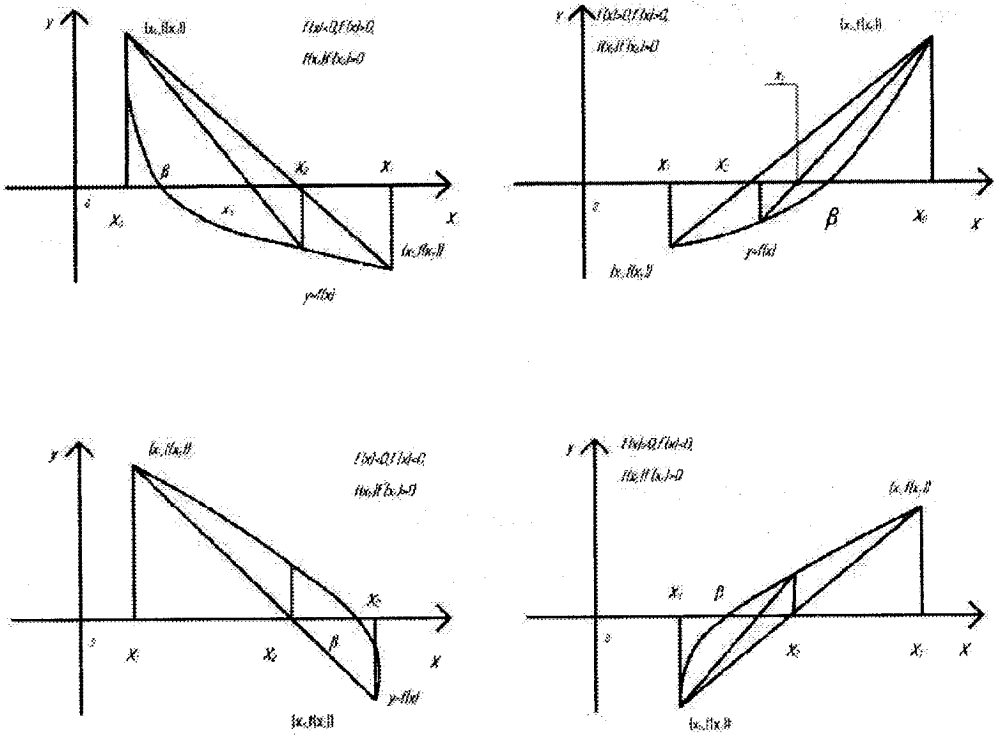


Рис. 4.19. Графічна інтерпретація методу січних

4.10. Метод простої ітерації або метод послідовних наближень

4.10.1. Аналіз постановки задачі

Метод простої ітерації застосовується до розв'язування нелінійного рівняння виду

$$x = \varphi(x)\,. \tag{4.102}$$

Перейти від рівняння (4.1) до рівняння (4.102) можна багатьма способами, наприклад, вибравши

$$\varphi(x) = x + \psi(x)\,f(x)\,, \tag{4.103}$$

де $\psi(x)$ – довільна знакостала неперервна функція.

Вибравши нульове наближення x_0, наступні наближення знаходимо за формулою

$$x_{n+1} = \varphi(x_n), \quad n = 0,1,2,\ldots. \tag{4.104}$$

Наведемо достатні умови збіжності методу простої ітерації.

Теорема 4.18. Нехай для вибраного початкового наближення x_0 на проміжку

$$S = \{x : |x - x_0| \le \delta\}. \tag{4.105}$$

функція $\varphi(x)$ задовольняє умову Ліпшиця

$$|\varphi(x') - (x'')| \le q|x' - x''|, \quad x', x'' \in S, \tag{4.106}$$

де $0<q<1$, і виконується нерівність

$$|\varphi(x_0) - x_0| \le (1-q)\delta. \tag{4.107}$$

Тоді рівняння (4.102) має на проміжку S єдиний корінь x_*, до якого збігається послідовність (4.104), причому швидкість збіжності визначається нерівністю

$$|x_n - x_*| \le \frac{q^n}{1-q}|\varphi(x_0) - x_0|. \tag{4.108}$$

Зауваження 4.9. якщо функція $\varphi(x)$ має на проміжку S неперервну похідну $\varphi'(x)$, яка задовольняє умові

$$|\varphi'(x)| \le q < 1, \tag{4.109}$$

то функція $\varphi(x)$ буде задовольняти умові (4.106) теореми 4.18.

З (4.108) можна отримати оцінку кількості ітерацій, які потрібно провести для знаходження розв'язку задачі (4.102) з наперед заданою точністю ε:

$$n \ge \left[\frac{\ln \dfrac{|\varphi(x_0) - x_0|}{(1-q)\cdot\varepsilon}}{\ln(1/q)} \right] + 1. \tag{4.110}$$

Наведемо ще одну оцінку, яка характеризує збіжність методу простої ітерації:

$$|x_n - x_*| \le \frac{q}{1-q}|x_n - x_{n-1}|. \tag{4.111}$$

4.10.2. Геометрична інтерпретація методу простої ітерації

Розглянемо рівняння (4.1), яке еквівалентне (4.2). Нехай побудовано ітераційний процес (4.49). Тоді в залежності від знаку похідної $\{\varphi'(x)\}$ маємо на рис. 4.20 та на рис. 4.21 геометричну картину збіжності простої ітерації.

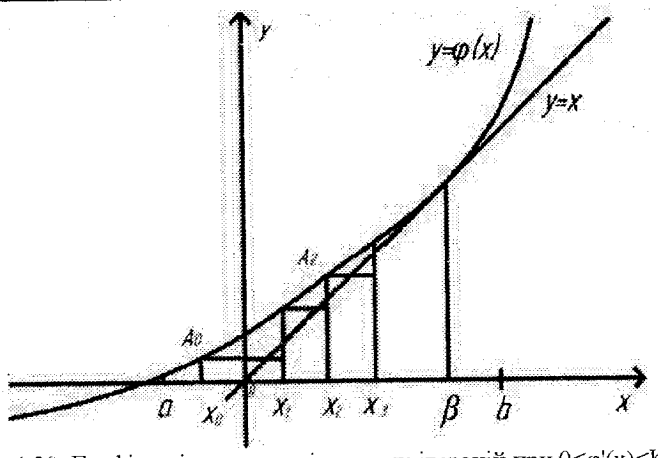


Рис. 4.20. Графічна інтерпретація методу ітерацій при $0<\varphi'(x)<K<1$

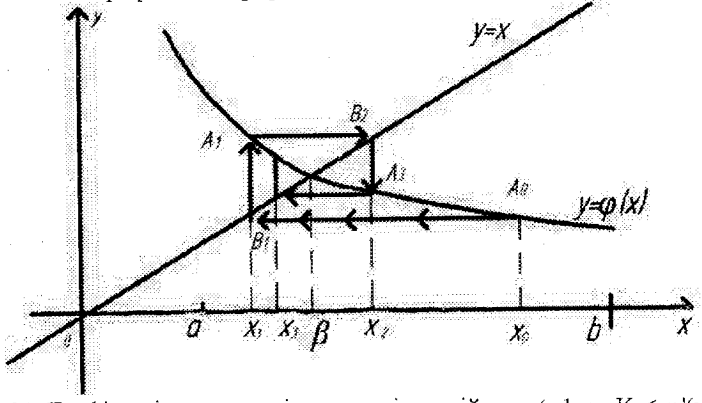


Рис. 4.21. Графічна інтерпретація методу ітерацій при $(-1<-K\le\varphi'(x)<0)$

Геометрично можна побачити, що якщо в околі кореня $x=\beta$ нерівність $0<\varphi'(x)\le K<1$, тоді послідовність $\{x_n, n\ge 1\}$ збігається до кореня, причому з тієї сторони, з якої розміщено початкове наближення x_0. У випадку $-1<-K\le\varphi'(x)<0$ послідовні наближення розміщені почергово з різних сторін. Точність наближення для цього методу $|x_n-\beta|\le|x_n-x_{n-1}|<\varepsilon, \forall\varepsilon>0$.

4.10.3. Обчислення похибки

Для застосування методу ітерацій (латинське "*iteratio*" – повторення) вихідне рівняння $f(x)=0$ ($f(x)$ – неперервна функція) необхідно, по-перше, записати у вигляді $x=\varphi(x)$, по-друге, виділити інтервал $[a,b]$ ізоляції кореня цього рівняння і по-третє, обрати нульове наближення кореня x_0. Для одержання першого наближення x_1 в праву частину рівняння $x=\varphi(x)$ замість x підставляємо x_0, так що $x_1=\varphi(x_0)$.

Наступні наближення утворюються за схемою

$$x_2 = \varphi(x_1); x_3 = \varphi(x_2); ...; x_n = \varphi(x_{n-1}); .... \qquad (4.112)$$

Таким чином, у результаті застосування деякого однакового процесу будуються послідовні наближення $x_0, x_1, x_2, ..., x_n, ...$

При цьому можуть бути два випадки:

1) процес може збігатись, тобто послідовні наближення прямують до деякої кінцевої межі ξ, що є коренем рівняння;
2) процес може бути розбіжним, тобто кінцева межа побудованих наближень не буде існувати; із цього не випливає, що розв'язок вихідного рівняння не існує, тобто виявилось, що процес послідовних наближень обраний невдало.

Збіжність процесу ітерації визначається наступною теоремою.

Теорема 4.19. Нехай інтервал $[a,b]$ є інтервалом кореня рівняння $x = \varphi(x)$, а функція $\varphi(x)$ визначена і диференційована на всьому інтервалі, причому усі її значення $\varphi(x) \in [a,b]$.

Тоді, якщо існує правильний дріб q, такий, що $|\varphi'(x)| \le q < 1$, то:

1) процес ітерації $x_n = \varphi(x_{n-1}), (n = 1, 2, ...)$ є збіжним незалежно від початкового значення $x_0 \in [a,b]$;
2) граничне значення $\xi = \lim\limits_{n\to\infty} x_n$ є єдиним коренем рівняння $x = \varphi(x)$ на відрізку $[a,b]$.

Наближення $x_0, x_1, x_2, ..., x_n, ...$ слід обчислювати доти, поки не буде виконана нерівність

$$|x_n - x_{n-1}| \le \frac{\varepsilon(1-q)}{q}, \qquad (4.113)$$

де ε – задана гранична абсолютна похибка кореня ξ.

Якщо $|\varphi'(x) \le q < 1|$ і $\varphi'(x)$ додатня в околі кореня, то послідовні наближення $x_n = \varphi(x_{n-1}), (n = 1, 2, ...)$ і $x_0 \in [a,b]$ збігаються до кореня монотонно. Якщо ж похідна $\varphi'(x)$ від'ємна, то послідовні наближення коливаються біля кореня ξ.

Приклад 4.11. Методом ітерацій уточнити з $\varepsilon = 10^{-4}$ корінь рівняння $5x^3 - 20x + 3 = 0$, ізольований на відрізку [0, 1].

Розв'язок. Зведемо рівняння до вигляду $x = \varphi(x)$. Це можна зробити таким чином:

1) $x = x + \left(5x^3 - 20x + 3\right)$, тоді $\varphi_1(x) = 5x^3 - 19x + 3$;
2) $x = \dfrac{5x^3 + 3}{20}$, тоді $\varphi_2(x) = \dfrac{5x^3 + 3}{20}$.

Визначимо, яку з отриманих функцій слід використовувати. Знаходимо:

$$|\varphi_1'(x)| = |15x^2 - 19| > 1 \quad \text{на } [0,1];$$

$$|\varphi_2'(x)| = \left|\frac{15x^2}{20}\right| = \frac{3}{4}x^2 < 1 \quad \text{на } [0,1].$$

Отже, можна використовувати функцію $\varphi_2(x)$ і шукати послідовні наближення за формулою

$$x_n = \frac{5x_{n-1}{}^3 + 3}{20}.$$

Визначимо, якою повинна бути різниця між двома послідовними наближеннями

$$|x_n - x_{n-1}| \le \frac{0.0001 \cdot (1 - 0.75)}{0.75} = 0.00003.$$

Обчислення зручно вести за допомогою табл. 4.4

Таблиця 4.4

№ ітерації	x_n	$x_n{}^3$	$x_{n+1} = \varphi(x_n)$
0	0.75	0.42188	0.25547
1	0.2555	0.016777	0.154144
2	0.1541	0.000565	0.151413
3	0.1514	0.005443	0.151361
4	0.15136	0.005442	0.151361

Можна вважати, що $\xi \approx 0.1514$.

4.10.4. Вплив обчислювальної похибки на збіжність ітераційного процесу

Нехай рівняння $f(x) = 0$ еквівалентне $x = \varphi(\mathrm{x})$. Побудуємо ітераційний процес

$$\{x_n, n \ge 1\} \subset S \equiv \{x \in R \mid |x - \beta| \le r\} : x_n = \varphi(\mathrm{x}_{n-1}), \forall\, \mathrm{x}_0 \in S, \lim_{n\to\infty} x_n = \beta, f(\beta) = 0.$$

Але при заокругленні на комп'ютері маємо $\tilde{x}_n$ і може бути, що $\lim_{x\to\infty} \tilde{x}_n \ne \beta$.

Теорема 4.20. Нехай $\forall x, y \in S \equiv \{x \in R \mid |x - \beta| \le r\}$ виконується

$$|\varphi(\mathrm{x}) - \varphi(\mathrm{y})| \le K|x - y|, 0 < K < 1; \tag{4.114}$$

$$|x_0 - \varphi(\mathrm{x}_0)| \le (1 - \mathrm{K})\,\mathrm{r} - \rho_0; x_0 \in S; 0 \le \rho_0 \le (1 - \mathrm{K})\,\mathrm{r};, \tag{4.115}$$

$$\tilde{x}_n = \varphi(\tilde{x}_{n-1}) + \rho_n, \text{ причому } \tilde{x}_0 = x_0 \in S, |x - x_0| \le r; n \ge 1, \tag{4.116}$$

де ρ_n – похибка обчислення $\varphi(\tilde{x}_{n-1})$

$$|\rho_n| \le \rho_0 q^{n-1} (0 < \mathrm{q} \le 1). \tag{4.117}$$

Тоді

$$\left|\tilde{x}_n - x_n\right| \le \rho_0 \left(K^{n-1} + K^{n-2}q + \ldots + Kq^{n-2} + q^{n-1}\right), \tag{4.118}$$

$$\lim_{n\to\infty} \tilde{x}_n = \beta. \tag{4.119}$$

Доведення. 1) Твердження (4.118) доведемо за допомогою методу математичної індукції. Якщо $\tilde{x}_n \in S; x_0 \in S$, то

$$\left|\tilde{x}_1 - x_0\right| = \left|\varphi(\mathrm{x}_0) + \rho_1 - x_0\right| \le \left|\varphi(\mathrm{x}_0 - \mathrm{x}_0)\right| + \left|\rho_1\right| \le (\mathrm{1} - \mathrm{K})\,\mathrm{r} - \rho_0 + \rho_0 < r.$$

Для n=1 виконується нерівність (4.118):

$$\left|\tilde{x}_1 - x_1\right| = \left|\varphi(\mathrm{x}_0) + \rho_1 - \varphi(\mathrm{x}_0)\right| = \left|\rho_1\right| \le \rho_0.$$

Нехай (4.118) виконується для $m, x_m \in S, \tilde{x}_m \in S$:

$$\left|\tilde{x}_m - x_m\right| \le \rho_0 \left(K^{m-1} + K^{m-2}q + \ldots + Kq^{m-2} + q^{m-1}\right), \tag{4.120}$$

Покажемо, що $x_{m+1} \in S$ для $n = m+1$. Дійсно

$$\left|\tilde{\mathrm{x}}_{m+1} - x_{m+1}\right| = \left|\varphi(\tilde{\mathrm{x}}_m) + \rho_{m+1} - \varphi(\mathrm{x}_0) - \mathrm{x}_0\right| \le \left|\varphi(\tilde{\mathrm{x}}_m) - \varphi(\mathrm{x}_0)\right| +$$

$$+ \left|\rho_{m+1}\right| + \left|\varphi(\mathrm{x}_0) - x_0\right| \le K\left|\tilde{\mathrm{x}}_m - x_0\right| + \rho_0 q^m + (1 - \mathrm{K})\,\mathrm{r} - \rho_0 \le$$

$$\le r + \rho_0\left(q^m - 1\right) \le r.$$

Звідки $\tilde{\mathrm{x}}_{m+1} \in S$, а значить

$$\left|\tilde{\mathrm{x}}_{m+1} - x_{m-1}\right| = \left|\varphi(\tilde{\mathrm{x}}_m) - \varphi(x_m) + \rho_{m+1}\right| \le K\left|\tilde{\mathrm{x}}_m - x_m\right| + \rho_0 q^m \le$$

$$\le \rho_0\left(K^m + K^{m-1}q + \ldots + Kq^{m-1} + q^m\right)$$

Звідси має місце (4.118) для $n = m+1$, тобто нерівність (4.118) доведена.

2) Очевидно, що

$$\left|\tilde{\mathrm{x}}_n - \beta\right| \le \left|\tilde{\mathrm{x}}_n - x_n\right| + \left|x_n - \beta\right| \le \left|\tilde{\mathrm{x}}_n - x_n\right| + K^n\left|x_0 - \beta\right| \le \left|\tilde{\mathrm{x}}_n - x_n\right| + K^n r \le$$

$$\le \rho_0\left(K^{n-1} + K^{n-2}q + \ldots + Kq^{n-2} + q^{n-1}\right).$$

Наслідок 4.3. Якщо 0<q<1, то $\left|\tilde{\mathrm{x}}_n - \beta\right| \le K^n r + n\rho_0(\max(\mathrm{K}, \mathrm{q}))^{n-1}$, звідки маємо (4.119).

Наслідок 4.4. Якщо q=1, то $\left|\tilde{\mathrm{x}}_n - \beta\right| \le K^n r + \rho_0 \dfrac{1 - K^n}{1 - K} \le K^n r + \dfrac{\rho_0}{1 - K}$.

Таким чином при q=1 не можемо гарантувати обчислення кореня з похибкою меншою за $\dfrac{\rho_0}{1-K}$.

4.11. Метод релаксації

4.11.1. Аналіз постановки задачі

Для збіжності ітераційного процесу (4.104) суттєве значення має вибір функції $\varphi(x)$. Зокрема, якщо в (4.103) вибрати $\psi(x)=\tau=\text{const}$, то отримаємо *метод релаксації*

$$x_{n+1}=x_n+\tau f(x_n),\quad n=0,1,2,\ldots, \tag{4.121}$$

який збігається при

$$-2<\tau f'(x)<0. \tag{4.122}$$

Якщо в деякому околі кореня виконуються умови

$$f'(x)<0,\quad 0<m_1<\left|f'(x)\right|<M_1, \tag{4.123}$$

то метод релаксації збігається при $\tau\in(0,2/M_1)$. Збіжність буде найкращою при

$$\tau=\tau_{\text{опт}}=2/(m_1+M_1). \tag{4.124}$$

При такому виборі τ для похибки $z_n=x_n-x_*$ буде мати місце оцінка

$$\left|z_n\right|\le q^n\left|z_0\right|,\quad n=0,1,2,\ldots, \tag{4.125}$$

де $q=(M_1-m_1)/(M_1+m_1)$.

Кількість ітерацій, які потрібно провести для знаходження розв'язку з точністю ε визначається нерівністю

$$n\ge\left[\frac{\ln\left(\left|z_0\right|/\varepsilon\right)}{\ln(1/q)}\right]+1. \tag{4.126}$$

Зауваження 4.10. Якщо виконується умова $f'(x)>0$, то ітераційний метод (4.121) потрібно записати у вигляді $x_{n+1}=x_n-\tau f(x_n)$.

Детально розглянемо метод простої ітерації для розв'язування нелінійного рівняння. Розглянемо нелінійне рівняння

$$x=\varphi(x). \tag{4.127}$$

Якщо в рівнянні (4.127) можна виділити x, то його залишаємо в лівій частині, а решта переносимо в праву

$$f(x)=0. \tag{4.128}$$

Якщо ні, то рівняння (4.127) множать на число α і додають до обох частин рівняння x:

$$\alpha f(x)=0,\ \alpha\neq 0,\ \underbrace{x+\alpha f(x)}_{\varphi(x)}=x. \tag{4.129}$$

Вибір α (великим чи малим, додатнім чи від'ємним) залежить від того, які умови будуть накладені на функцію $\varphi(x)$ для забезпечення збіжності методу. Вибираємо деяке довільне початкове наближення $x_0\in[a,b]$.

$$x_1=\varphi(x_0),\ x_2=\varphi(x_1),\ \ldots,\ x_{k+1}=\varphi(x_k),\ k=0,1,2,\ldots \tag{4.130}$$

Ітераційний процес (4.130) отримав назву *методу простої ітерації*. Якщо він збігається, то границею є точний розв'язок. Якщо функція $\varphi(x)$ є неперервною, то в процесі (4.130) можна прийти до границі

$$x^*=\lim_{k\to\infty}x_{k+1}=\lim_{k\to\infty}\varphi(x_k)=\varphi(\lim_{k\to\infty}x_k)=\varphi(x^*). \tag{4.131}$$

Теорема 4.21. Якщо $\varphi(x)$ є неперервною і диференційованою на деякому проміжку $[a,b]$, а x_0, $\varphi(x_k)\in[a,b]$ і $|\varphi'(x_k)|\le q<1$, то границя $\lim_{k\to\infty}x_k=x^*$ існує і x^* єдиний розв'язок рівняння на $[a,b]$.

Доведення.

$$x_k=\varphi(x_{k-1}), \tag{4.132}$$

Згідно (4.130) – (4.132)

$$|x_{k+1}-x_k|=|\varphi(x_k)-\varphi(x_{k-1})|=$$

використовуючи теорему Лагранжа, $c\in[x_{k-1},x_k]$

$$=|\varphi'(c)|\cdot|x_k-x_{k-1}|\le q|x_k-x_{k-1}|.$$

Проводячи процес k разів отримаємо

$$|x_{k+1}-x_k|<q^k|x_1-x_0|. \tag{4.133}$$

Розглянемо ряд

$$x_0+(x_1-x_0)+(x_2-x_1)+\ldots+(x_k-x_{k-1})+\ldots \tag{4.134}$$

Збіжність ряду означає збіжність послідовних частинних сум.

Отже, збіжність послідовності x_k зводиться до доведення збіжності ряду (4.134). Збіжність останнього доводиться завдяки збіжності деякого мажорантного ряду. Замінимо ряд (4.134) мажорантним рядом, використовуючи співвідношення (4.133):

$$\begin{aligned}&|x_0|+|x_1-x_0|+q|x_1-x_0|+q^2|x_1-x_0|+\ldots+q^k|x_1-x_0|+\ldots=\\&=|x_0|+|x_1-x_0|\cdot(1+q+q^2+\ldots.).\end{aligned} \tag{4.135}$$

Вираз в останніх дужках є спадною геометричною прогресією, бо $q<1$, тому доведена збіжність ряду.

4.11.2. Геометрична інтерпретація методу

Нехай $0<\varphi'(x)<1$, так як $x=\varphi(x)$. Розв'язком рівняння є точка перетину двох плоских кривих (рис. 4.22а). Нехай $-1<\varphi'(x)<0$, тоді розглядають варіант, зображений на рис. 4.22б. Якщо $\varphi'(x)>1$, тоді розглядають варіант на рис. 4.22в.

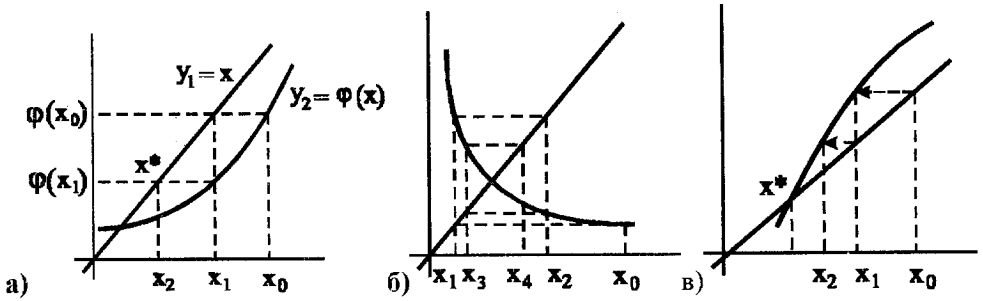


Рис. 4.22. Графічна інтерпретація методу релаксації

4.11.3. Обчислення похибки

Введемо позначення: $\psi(x)=x-\varphi(x)$, причому $\psi'(x)=1-\varphi'(x)$, так як $|\varphi'(x)|<1$, то $-|\varphi'(x)|>1$. Якщо $\psi'(x)\geq 0$, то нехай $\psi'(x)\geq 1-q$. Розглянемо оцінку:

$$|x_k-x_{k-1}|=|x_k-\varphi(x_k)|=|\varphi(x_k)|=$$

Відомо, що

$$\left.\begin{array}{l} x=\varphi(x) \\ x^*-\varphi(x^*)=0 \end{array}\right\} \psi(x^*)=0. \qquad (4.136)$$

Тоді оцінка

$$=|\psi(x_k)-\psi(x^*)|=|\psi'(c)|\cdot|x_k-x^*|=|1-\varphi'(c)|\cdot|x_k-x^*|,\ c\in[x_k,x^*]$$

$$\begin{aligned} |x_k-x^*| &= \frac{1}{1-\varphi'(c)}|x_{k+1}-x_k|\leq\frac{1}{1-q}|x_{k+1}-x_k|\leq \\ &\leq\frac{q}{1-q}|x_k-x_{k-1}|\leq\ldots\leq\frac{q^k}{1-q}|x_1-x_0|. \end{aligned} \qquad (4.137)$$

Зазначимо, що знаючи величину q, яка приблизно вибирається як $q\approx\max\limits_{x\in[a,b]}|q'(x)|$ можна оцінити число n, тобто кількість кроків, які потрібно провести методом простої ітерації, щоб досягнути наперед заданої точності ε.

З нерівності

$$\left|x_k - x^*\right| \le \frac{q}{1-q}\left|x_k - x_{k-1}\right| \tag{4.138}$$

при $q = \frac{1}{2}$ отримаємо, що

$$\left|x_k - x^*\right| \le \left|x_k - x_{k-1}\right| < \varepsilon . \tag{4.139}$$

Як тільки x_k відрізняється від x_{k-1} на ε, то це наближення і від точного розв'язку відрізняється на величину $<\varepsilon$.

Якщо q вибране довільне, але $0 < q < 1$, то як тільки

$$\left|x_k - x_{k-1}\right| < \frac{1-q}{q}\varepsilon \Rightarrow \left|x_k - x^*\right| < \varepsilon . \tag{4.140}$$

Умова (4.140) без особливих змін може бути записана як умова зупинки у вигляді $\left|x_k - x_{k-1}\right| < \delta$, де δ – достатньо мала абсолютна величина.

В силу того, що початкові наближення в певній мірі можна вибирати довільні (наприклад δ), то метод набув властивості самовиправлення, тобто якщо по якійсь причині здійснюється обчислювальна похибка, то при спотворенні значення x_k воно може бути вибране у ролі нового початкового наближення, а процес ітерації все одно приведе то точного розв'язку. Можна дещо полегшити умови теореми, щоб початкове наближення хоча і вибирати довільним, але з дещо вужчого проміжку, а саме якщо $[a,b]$ розбити на три частини $\overbrace{a \quad \alpha \quad \beta \quad b}$, тоді $\alpha = a + \frac{1}{3}(b-a)$ та $\beta = b + \frac{2}{3}(b-a)$. Якщо $x_0 \in [\alpha, \beta]$, то всі $\varphi(x_k) \in [a,b]$.

4.12. Метод парабол (метод Д. Мюллера) знаходження коренів многочлена з комплексними коефіцієнтами

4.12.1. Аналіз постановки задачі

Метод парабол призначений для знаходження всіх коренів алгебраїчного многочлена (метод Д. Мюллера)

$$P_n(z) = a_n + a_{n-1}z + \ldots + a_0 z^n, a_0 \neq 0, \tag{4.141}$$

з комплексними коефіцієнтами без знання їх наближених значень. Цей метод легко реалізувати на комп'ютері: він забезпечує хорошу швидкість обчислень та дозволяє визначати прості корені многочлена з високого точністю. Кратні і близькі корені

знаходяться дещо гірше, ніж прості, але точніше ніж іншими методами. Всі корені многочлена визначаються послідовно один за одним.

Збіжність методу парабол у загальному випадку не доведена, але в практиці обчислень невідомо жодного прикладу, коли б цей метод не збігався або збігався повільно.

Основна ідея методу полягає в наступному. Нехай задані три довільні попарно різні комплексні числа z_0, z_1, z_2. За цими числами як вузлами інтерполювання будується інтерполяційний многочлен Лагранжа для $P_n(z)$. Це буде деякий многочлен другого порядку з комплексними коефіцієнтами. Знаходяться обидва його корені і за z_3 береться той з його коренів, який розташований ближче до z_3. Наступна трійка чисел буде z_1, z_2, z_3. За цими числами будується новий інтерполяційний многочлен, і за z_4 береться його корінь ближчий до z_3. Отримаємо нову трійку чисел z_2, z_3, z_4 тощо.

Послідовність точок, яка побудована таким чином, збігається завжди до деякого кореня многочлена (цей факт встановлений емпіричним шляхом). Після того, як обчислено один з коренів, він виділяється і далі знову застосовуємо цей метод до многочлена меншого степеня. Приведемо алгоритм методу. Припустимо, що відомо значення многочлена $P_n(z)$ у трьох попарно різних точках z_{i-2}, z_{i-1}, і z_i. За цими точками будується інтерполяційний многочлен Лагранжа для $P_n(z)$.

$$L^{(1-2)}(z)=P_n(z_{i-2})\frac{(z-z_i)(z-z_{i-1})}{(z_{i-2}-z_i)(z_{i-2}-z_{i-1})}+P_n(z_{i-1})\frac{(z-z_i)(z-z_{i-2})}{(z_{i-1}-z_i)(z_{i-1}-z_{i-2})}+P_n(z_i)\frac{(z-z_{i-1})(z-z_{i-2})}{(z_i-z_{i-1})(z_i-z_{i-2})}. \quad (4.142)$$

Введемо нові позначення

$$\left.\begin{aligned} h \equiv z-z,\ \ h_i \equiv z_i - z_{i-1},\ \ h_{i-1} \equiv z_{i-1}-z_{i-2}, \\ \lambda \equiv \frac{h}{h_i},\ \ \lambda_i \equiv \frac{h_i}{h_{i-1}},\ \ \delta_i \equiv 1+\lambda_i,\ \ i=2,3... \end{aligned}\right\}. \quad (4.143)$$

Якщо підставити (4.143) в (4.142), то одержимо

$$L^{(1-2)}(\lambda) = \lambda^2\delta_i^{-1}[P_n(z_{i-2})\lambda_i^2 - P_n(z_{i-1})\lambda_i\delta_i + P_n(z_i)\lambda_i] +$$
$$\lambda\delta_i^{-1}[P_n(z_{i-2})\lambda_i^2 - P_n(z_{i-1})\delta_i^2 + P_n(z_i)(\lambda_i+\delta_i)] + P_n(z_i)$$

Якщо прирівняти $L^{(1-2)}(\lambda)$ до нуля і розв'язати це рівняння відносно λ, то одержимо формулу

$$\lambda = \frac{-2P_n(z_i)\delta_i}{g_i \pm \sqrt{g_i^2 - 4P_n(z_i)\delta_i\lambda_i[P_n(z_{i-2})\lambda_i - P_n(z_{i-1})\delta_i + P_n(z_i)]}}. \quad (4.144)$$

Тут

$$g_i = P_n(z_{i-2})\lambda_i^2 - P_n(z_{i-1})\delta_i^2 + P_n(z_i)(\lambda_i+\delta_i)$$

З двох можливих значень λ береться найменше за модулем і далі обчислюється

$$\lambda_{i+1} = \lambda,\ \ h_{i+1} = \lambda_{i+1}h_i,\ \ z_{i+1} = z_i + h_{i+1}. \quad (4.145)$$

На цьому можна було б закінчити процес обчислення наступного наближення кореня, але для більшої гарантії збіжності і для подолання труднощів, що пов'язані з появою великих чисел, пропонується додатково зробити наступне.

Перевіряється виконання нерівності

$$|h_{i+1}| \geq R_i \equiv \sqrt[n]{|P_n(z_i)|\frac{1}{|a_0|}}. \qquad (4.146)$$

Якщо це виконується, то λ береться рівним одиниці, і проводимо обчислення за формулами (4.145). Якщо нерівність (4.146) не виконується, то і λ не змінюється.

Після цього перевіряється виконання іншої нерівності:

$$|P_n(z_{i+1})| \leq 10|P_n(z_i)|. \qquad (4.147)$$

Якщо ця нерівність не виконується, то λ замінюємо на $\lambda/2$, перераховуються величини (4.145) і знову перевіряється (4.147). Заміна λ на $\lambda/2$ і перерахунок (4.145) здійснюється доти, поки нерівність (4.147) вперше не виконується. Останнє отримане z_{i+1} береться за наступне наближення до кореня. Після цього замість точок z_{i-2}, z_{i-1}, z_i беруться точки z_{i-1}, z_i, z_{i+1} і процес повторюється.

Таким чином, отримується послідовність точок $\{z_p\}$, яка завжди збігається до деякого кореня многочлена $P_n(z)$.

Метод дозволяє визначити лише один, взагалі кажучи, випадковий корінь, тому для знаходження інших коренів потрібно кожен знайдений корінь, за схемою Горнера [13] виділити і далі застосовувати метод до многочлена меншого степеня.

Глобальна збіжність методу парабол не доведена, але на практиці обчислень невідомо жодного випадку, коли б цей метод не збігався, або збігався повільно.

Дослідимо локальну збіжність методу. Будемо вважати, що інтерполяційний многочлен $L(z)$ має степінь $n \geq 1$ і початкові точки $z_0, z_1, \ldots, z_m$ взяті досить близько до кореня z^* многочлена $P_n(z)$. Нехай p-кратність кореня z^*, причому $p \leq m$. Маємо

$$L(z) = P_n(z) + \frac{P_n^{(m+1)}(\xi)}{(m+1)!}\prod_{i=0}^{m}(z - z_i). \qquad (4.148)$$

де точка ξ належить найменшій опуклій області, що містить $z_0, z_1, \ldots, z_m$, z. Так як $L(z_{m+1}) = 0$ то з (4.148) випливає

$$P_n(z_{m+1}) + \frac{P_n^{(m+1)}(\xi_{m+1})}{(m+1)!}\prod_{i=0}^{m}(z_{m+1} - z_i) = 0. \qquad (4.149)$$

Без обмежень загальності, можна вважати, що потрібний корінь z_{m+1} цього рівняння розташований ближче інших коренів до точки z^*. Насправді, зі збіжності методу парабол випливає, що на деякому кроці z_m буде ближче до z^*, ніж $z_0, z_1, \ldots, z_{m-1}$. Якщо z^* – простий корінь, то корінь z_{m+1}, найбільш близький до z^* буде ізольованим. Тому цей корінь буде найбільш близьким до z_m. В загальному випадку

для виконання цього твердження, можливо, потрібно вибрати z_{m+1} не найбільш близьким до z_m, а з найменшим за модулем значенням многочлена $P_n(z)$.

Позначимо $\varepsilon_i = z_i - z^*$ і запишемо в околі z^* рівняння (4.149) у такому вигляді

$$\frac{P_n^{(p)}(z^*)}{p!}\left((z_{m+1}-z^*)^p + o\left(\left|z_{m+1}-z^*\right|^{p+1}\right)\right) +$$

$$+\left(\frac{P_n^{(m+1)}(z^*)}{(m+1)!} + o\left(\left|\xi_{m+1}-z^*\right|\right)\right)\prod_{i=0}^{m}[(z_{m+1}+z^*)-(z_i+z^*)] = 0$$

Звідси отримуємо

$$\frac{P_n^{(p)}(z^*)}{p!}\varepsilon_{m+1}^p + \frac{P_n^{(m+1)}(z^*)}{(m+1)!}\prod_{i=0}^{m}[(\varepsilon_{m+1}+\varepsilon_i) +$$

$$+o\left(\left|\varepsilon_{m+1}\right|^{p+1}\right)] + o\left(\left|\xi_{m+1}-z^*\right|\right)\prod_{i=0}^{m}(\varepsilon_{m+1}-\varepsilon_i) = 0 \qquad (4.150)$$

Оцінимо величину мінімального за модулем кореня ε_{m+1} цього рівняння. Очевидно, що ε_{m+1} тільки зменшиться, якщо $P_n^{(m+1)}(z^*) = 0$. Нехай $P_n^{(m+1)}(z^*) \neq 0$. При малих ε_i мінімальні корені рівняння (4.150) з точністю до нескінченно малої вищого порядку будуть співпадати з відповідними коренями рівняння

$$\frac{P_n^{(p)}(z^*)}{p!}\varepsilon_{m+1}^p + \frac{P_n^{m+1}(z^*)}{(m+1)!}\prod_{i=0}^{m}[(\varepsilon_{m+1}-\varepsilon_i) = 0.$$

Введемо нову змінну

$$\eta \equiv \left(\frac{p!P_n^{(m-1)}(z^*)}{(m+1)!P_n^{(p)}(z^*)}\right)^{\frac{1}{m-p+1}}\varepsilon,$$

Тоді $\eta_{m+1}^p + \prod_{i=0}^{m}(\eta_{m+1}-\eta_i) = 0$.

При $\eta_i = 0$, $i = 0,1,\ldots, m$, це рівняння має p-кратний нульовий корінь, а добуток інших коренів рівний – 1. Вільний член многочлена з точністю до знаку рівний добутку всіх його коренів. Тому, виділяючи корені за модулем близькі до одиниці, отримаємо, що з точністю до нескінченно малої вищого порядку добутку p найменших коренів рівний $\prod_{i=0}^{m}(-\eta_i)$. Найменшим за модулем корінь буде не більшим за середнє геометричне , тобто $\left|\eta_{m+1}\right|^p \le \prod_{i=0}^{m}(\eta_i)$. Будемо знаходити $|\eta_{m+1}|$ з рівності $\left|\eta_{m+1}\right|^p = \prod_{i=0}^{m}(\eta_i)$. При цьому $|\eta_{m+1}|$ може тільки збільшуватися. Позначимо

$y_i \equiv \ln|\eta_i|$, $i = 0,1,...,m+1$. Легко помітити, що величини y_i задовольняють лінійне різницеве рівняння $p \cdot y_{m+1} = y_m + y_{m-1} + ... + y_0$.

Згідно загальної теорії розв'язування цих рівнянь будемо мати

$$y_{k+1} = \alpha_{mp}^{k+1}\left[Q_{p-1}(k+1) + o(\alpha^{k+1})\right], \tag{4.151}$$

де $|\alpha| < 1$, α_{mp} – максимальний за модулем корінь характеристичного многочлена

$$\varphi_{mp}(t) = p \cdot t^{m+1} - t^m - t^{m-1} - ... - t - 1, \tag{4.152}$$

$Q_{p-1}(k)$ – многочлен степеня p-1.

З (4.151) знаходимо $y_{k+1} = \alpha_{mp} y_k \left[1 + o(\alpha^k)\right]$ і далі

$$\left|z_{k+1} - z^*\right| \le \left|\frac{p!P_n^{(m+1)}(z^*)}{(m+1)!P_n^{(p)}(z^*)}\right|^{\frac{\alpha_{mp}-1}{m-p+1}} \left|z_k - z^*\right|^{\alpha_{mp}} + o\left(\left|z_k - z^*\right|^{\alpha_{mp}}\right).. \tag{4.153}$$

Теорема 4.22. Справедливе таке твердження [76-80].Якщо в рівнянні

$$x^n - b_1 x^{n-1} - ... - b_n = 0 \tag{4.154}$$

всі $b_i \ge 0$, i = 1,2,...,n, а індекси додатніх b_i, мають найбільший спільний дільник 1, то рівняння (4.154) має єдиний простий додатний корінь r і модулі всіх інших коренів менші r.

В силу цієї теореми α_{mp} буде дійсним додатним числом.

Многочлен φ_{mp} *(t)* не має кореня рівного 1 для $p \le m$, тому сумуючи його останні члени як геометричну прогресію, отримаємо:

$$\varphi_{mp}(t) = \frac{t^{m+1}(pt - 1 - p) + 1}{t - 1}.$$

Звідси та з (4.152) встановлюємо, що $1 < \alpha_{mp} < \frac{1+p}{p}$, так як φ_{mp} *(1)*<0 для $p \le m$, а $\varphi_{mp}((1+p)/p)$>0. Крім того, співвідношення

$$\varphi_{m+1,p}(t) = t\varphi_{mp}(t) - 1;\ \varphi_{m,p+1}(t) = \varphi_{mp}(t) - t^{m+1}$$

дозволяє зробити висновок, що $\alpha_{m,p+1} < \alpha_{mp} < \alpha_{m+1,p}$.

Отже, всі числа $\alpha_{mp} > 1$, тому з формули (4.153) випливає, що метод парабол забезпечує локальну збіжність, якщо тільки всі різниці $z_i - z^*$ на початку процесу були достатньо малі. Однак, швидкість збіжності не перевищить квадратичної в силу того, що $\alpha_{mp} < 2$ для всіх m і p.

Зокрема, при використанні інтерполяційного многочлена другого степеня, отримаємо відповідно при однократному і двократному корені z^*

$$\left|z_{k+1}-z^*\right| \le \left|\frac{P_n^m(z^*)}{6P_n'(z^*)}\right|^{0.42} \left|z_k-z^*\right|^{1.84} + o\left(\left|z_k-z^*\right|^{1.84}\right),$$

$$\left|z_{k+1}-z^*\right| \le \left|\frac{P_n^m(z^*)}{6P_n^n(z^*)}\right|^{0.23} \left|z_k-z^*\right|^{1.23} + o\left(\left|z_k-z^*\right|^{1.23}\right).$$

Розглянемо деякі основні характеристики методу Д. Мюллера.

4.12.2. Обчислення похибки

Метод парабол є одним із найбільш простих і зручних методів визначення коренів многочлена. Його зручність полягає не тільки у простоті обчислювальної схеми та великої швидкості роботи, але і в тому що на кожному кроці він вимагає обчислення тільки значення многочлена у деякій точці. Ця властивість методу стає особливо важливою у тому випадку, коли многочлен не заданий явно, а заданий, наприклад, у вигляді характеристичного визначника. Тоді інші методи практично майже не застосовні, а метод парабол можна успішно застосувати.

Цікаво відзначити, що хоч і можна побудувати багато аналогічних методів, використовуючи інтерполяційні многочлени більш високого порядку, робити це недоцільно. По-перше, для визначення параметра λ доведеться знаходити найменший за модулем корінь многочлена більш високого порядку, для чого потрібно буде застосувати деякі спеціальні методи. По-друге, як було показано вище, швидкість збіжності все одно не перевищить квадратичної. Числа a_{mp} із ростом m дуже швидко збігаються до своєї границі, тому достатньо використовувати многочлен другого степеня.

Цей метод дозволяє визначити тільки один корінь многочлена, тому для знаходження інших коренів кожний знайдений корінь потрібно виділяти і далі застосовувати цей метод до многочлена меншого степеня. При цьому для зменшення впливу помилок як і в інших методах корінь потрібно уточнити перед виділенням за вихідним многочленом. Початковими вузлами для уточнення будуть останні три наближення. Починати процес можна з будь-яких значень z_0, z_1, z_2, наприклад, $z_0 = -1$, $z_1 = +1$, $z_2 = 0$.

Порівнюючи інші методи з методом парабол за швидкістю збіжності можна зробити висновок, що останній збігається повільніше (за числом ітерацій) для простих коренів, але швидше для кратних. Якщо для кореня z^* кратності p вибрати параметр з напівінтервалу (0,p], то в інших методах можна отримати квадратичну збіжність. Показник швидкості методу парабол не перевищить число $(1+p)/p$, що менше двох для $p > 1$. Значення показника збіжності a_{mp} для різних m і p приведені у таблиці 4.5.

Беручи до уваги те, що многочлени є швидкозростаючими функціями, значення $P_n(z_i)$ навіть для не дуже великих n та z_i (порядку 10-20) стають настільки великими, що перевищать максимально допустиме у комп'ютері число. Саме ж λ, яке обчислюється за формулою (4.144), майже завжди є малим числом.

Таблиця 4.5

Значення показників збіжності

p \ m	1	2	3	4	5	6
1	1,618	1,839	1,928	1,966	1,984	1,992
2	—	1,234	1,349	1,410	1,445	1,466
3	—	—	1,126	1,199	1,243	1,271
4	—	—	—	1,079	1.130	1,163
5	—	—	—	—	1,055	1,092
6	—	—	—	—	—	1,040

Щоб позбутися цього недоліку, представимо $P_n(z_i)$ наступним чином: $P_n(z_i) = z_i^{t_i} R(z_i)$, де $t_i = 0$, $R(z_i) = a_0 z_i^n + a_1 z_i^{n-1} + \ldots + a_n$, якщо $|z_i| \le 1$, $t_i = n$, $R(z_i) = a_0 + a_1 z_i^{-1} + \ldots + a_n z_i^{-n}$, якщо $|z_i| > 1$. Знаючи t_i, $R(z_i)$ і відношення чисел z_i z_{i-1}, z_{i-2}, легко обчислити λ, без появи великих чисел. Аналогічно слід діяти і при обчисленні R_i згідно (4.146).

4.13. Розв'язування систем нелінійних рівнянь

Розглянемо систему нелінійних рівнянь

$$f_i(x_1, x_2, \ldots, x_n) = 0,\ i = \overline{1,n}. \tag{4.155}$$

1. ***Метод ітерацій розв'язку систем спеціального вигляду.*** Нехай система (4.155) еквівалентна

$$x_i = \varphi_i(x_1, x_2, \ldots, x_n),\ i = \overline{1,n}, \tag{4.156}$$

а система (4.156) в області $G \in \mathbf{R}^n$ має єдиний розв'язок $x_i = \alpha_i$, $i = \overline{1,n}$.

Для $\forall x_i^{(0)} \in R^1$, $i = \overline{1,n}$, будуємо ітераційний процес

$$\left.\begin{aligned} x_i^{(1)} &= \varphi_i(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}), \\ x_i^{(2)} &= \varphi_i(x_1^{(1)}, x_2^{(1)}, \ldots, x_n^{(1)}), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_i^{(k+1)} &= \varphi_i(x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}), \end{aligned}\right\} \tag{4.157}$$

де $\lim\limits_{x\to\infty} x_i^{(k)} = a_i$, $i = 1, 2, \ldots, n$.

Визначимо оператор

$$y = Ax, \tag{4.158}$$

де $y_1 = \varphi_1(x_1, x_2, \ldots, x_n) \equiv Ax$, $x, y \in \mathbf{R}^n$, $x \equiv (x_1, x_2, \ldots, x_n)$, $y \equiv (y_1, y_2, \ldots, y_n)$.

Теорема 4.23. Нехай:

1) $Q \equiv \{x \in R^n | \rho(x, \alpha) \le r\}$, $\alpha \equiv (\alpha_1, \alpha_2, .., \alpha_n)^T$;

2) $|\varphi_i(x') - \varphi_i(x'')| \le K \rho(x', x'')$;

3) $\rho(\varphi(x') - \varphi(x'')) \le K \rho(x', x'')$, $0 < K < 1$.

Тоді

I) $\forall x^{(0)} \in Q : x^{(k+1)} = Ax^{(k)}$; $\lim\limits_{x \to \infty} x_i^{(k)} = \alpha$;

II) $\rho(x^{(K)}, \alpha) < K^i \ \rho(x^{(0)}, \alpha)$.

Доведення випливає з принципу стислих відображень (теорема 4.7).

Теорема 4.24. (про єдиність розв'язку).

Нехай:

1) $Q \equiv \{x \in \mathbf{R}^n : \rho(x, y^{(0)}) \le r\}$, $y^{(0)} \in \mathbf{R}^n$ – фіксований вектор;

2) $\varphi_i(x)$ задовольняє умову Ліпшиця;

3) $\forall y(0) \in Q$ має місце $\rho(Ax^{(0)}, x^{(0)}) < (1 - k)z$.

Тоді система (4.156) має єдиний розв'язок $\lim\limits_{x \to \infty} x_i^{(k)} = \alpha$, при $\forall x^0 \in \mathbf{R}^n$.

Доведення аналогічно до скалярного випадку.

Теорема 4.25. (достатні умови збіжності).

Нехай:

1) $\varphi_i(x) \in C(G)$, $1 \le i \le n$, $a \in G$;

2) $\dfrac{\partial \varphi_i(x)}{\partial x_i} \in C(G)$, $1 \le i, \ j \le n$.

Тоді для збіжності методу ітерації достатньо, щоб власні значення матриці

$$M = \left\{ \max_G \left| \frac{\partial \varphi_i}{\partial x_j} \right| \right\}, \ i, \ j = \overline{1, n}, \tag{4.158}$$

задовольняли умову $|\lambda(M)| < 1$.

Доведення. Якщо $x^{(k)} = \varphi(x^{k-1}) \in G$, то

$$x_i^{(k)} - \alpha_i = \varphi_i(x^{k-1}) - \varphi_i(\alpha) = \sum_{j=1}^{n} \frac{\partial \varphi_i(p_i^{k-1})}{\partial x_i} (x_j^{k-1} - \alpha_j), \tag{4.159}$$

$i = \overline{1, n}$, $p_j^{(k-1)}$ – деяка точка відрізка прямої, що з'єднує точки $x^{(k-1)}$ і α. Позначимо

$$M_k = \begin{pmatrix} \dfrac{\partial \varphi_1(p_1^{(k)})}{\partial x_1} & \cdots & \dfrac{\partial \varphi_1(p_1^{(k)})}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial \varphi_n(p_1^{(k)})}{\partial x_1} & \cdots & \dfrac{\partial \varphi_n(p_1^{(k)})}{\partial x_n} \end{pmatrix}. \tag{4.160}$$

Тоді за (4.159) маємо $x^{(k)} - \alpha = M_{k-1}(x^{(k-1)} - \alpha)$, звідки, ітеруючи процес, одержимо

$$x^{(k)} - \alpha = M_{k-1}M_{k-2}\ldots M_1 M_0 (x^{(0)} - \alpha).$$

Значить для $x^0 \neq \alpha$ одержимо $\lim\limits_{x\to\infty} x^{(k)} = \alpha$, якщо $M_{k-1}\ldots M_0 \underset{k\to\infty}{\longrightarrow} 0$.

Позначимо $M_{ij} = \max\limits_{G}\left|\dfrac{\partial \varphi_i}{\partial x_j}\right|$, тоді $\|M_{k-1}M_{k-2}\ldots M_1 M_0\| \le \|M\| \underset{k\to\infty}{\longrightarrow} 0$, бо $|\lambda(M)|<1$.

2. *Метод Ньютона.* Нехай

$$f_i(x_1, x_2, \ldots, x_n) = 0, \tag{4.161}$$

$x \equiv (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbf{R}^n$.

Позначимо

$$f(x) \equiv (f_1(x), f_2(x), \ldots, f_n(x)), \tag{4.162}$$

тоді (4.155) прийме вигляд

$$f(x) = 0,\ f:\ \mathbf{R}^n \to \mathbf{R}^n. \tag{4.163}$$

Теорема 4.26. Нехай:

1) $f \in C^1(G)$, де $a \in G \equiv \{x \in \mathbf{R}^n \mid \rho(x, a) \le r\}$;

2) матриця

$$f_x(\alpha) = \begin{pmatrix} \dfrac{\partial f_1(\alpha)}{\partial x_1} & \cdots & \dfrac{\partial f_1(\alpha)}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial f_n(\alpha)}{\partial x_1} & \cdots & \dfrac{\partial f_n(\alpha)}{\partial x_n} \end{pmatrix}. \tag{4.164}$$

не вироджена, тобто $\det f_x(x) \neq 0$.

Тоді існує окіл $Q = \{x \in G:\ \rho(x, a) \le \delta < r\} \subset G$ такий, що $\forall x^0 \in \mathbf{R}^n$

$$x^{(m+1)} = \Phi(x^{(m)}) = x^{(m)} - f_x^{-1}(x^{(m)}) \cdot f(x^{(m)}) \underset{m\to\infty}{\longrightarrow} \alpha. \tag{4.165}$$

Доведення. Нехай

$$x = \varphi(x) = x - f_x^{-1}(x) f(x), \tag{4.166}$$

тоді корінь α системи (4.163) є коренем (4.166). Зауважимо, що (4.165) еквівалентне

$$f_x(x^{(m)})(x^{(m+1)} - x^{(m)}) = -f(x^{(m)}). \tag{4.167}$$

За теоремою 4.17 виконується умова $\|M\| < 1$, що еквівалентне $A(M) < 1$, тобто $\lim\limits_{x\to\infty} x^{(m)} = \alpha$. Отже, $\lim\limits_{x\to\infty} x^{(m)} = \alpha$ є коренем (4.166), а, отже і коренем (4.163).

4.14. Приклади розв'язування нелінійних рівнянь

Приклад 4.12. Розв'язати рівняння

$$x + \sin x - 1 = 0. \tag{4.168}$$

методом ділення проміжку навпіл з точністю $\varepsilon=10^{-4}$.

Розв'язок. Спочатку знайдемо проміжок, на якому рівняння має єдиний корінь. Оскільки похідна функції $f(x) = x + \sin x - 1$ не змінює знак, то корінь у рівнянні (4.168) буде один. Легко побачити, що $f(0)=-1<0$, а $f\left(\frac{\pi}{2}\right) = \frac{\pi}{2} > 0$. Отже, корінь належить проміжку $\left[0, \frac{\pi}{2}\right]$. Виберемо $a_0 = 0$, $b_0 = \frac{\pi}{2}$. Згідно формули (4.60), отримаємо, що для знаходження кореня з точністю 10^{-4} необхідно провести 13 інтеграцій. Відповідні значення x_n наведені в таблиці 4.6.

Таблиця 4.6

n	x_n	$f(x_n)$
0	0785398E+00	0492505E+00
1	0392699E+00	0224617E+00
2	0589049E+00	0144619E+00
3	0490874E+00	0377294E- 01
4	0539961E+00	0540639E- 01
5	0515418E+00	0831580E- 02
6	0503146E+00	0146705E- 01
7	0509282E+00	0316819E- 02
8	0512350E+00	0257611E- 02
9	0510816E+00	0295467E- 03
10	0511583E+00	0114046E- 02
11	0511199E+00	0422535E- 03
12	0511007E+00	0635430E- 04
13	0510911E+00	0116016E- 03

Приклад 4.13. Знайти додатні корені рівняння

$$x^3 - x - 1 = 0 \tag{4.169}$$

методом простої ітерації з точністю $\varepsilon=10^{-4}$.

Розв'язок. Графічне дослідження рівняння (4.169) показує, що існує єдиний дійсний додатній корінь цього рівняння і він належить проміжку [1,2]. Оскільки на цьому проміжку $x \neq 0$, то рівняння (4.169) можна подати у вигляді

$$x = \sqrt{\frac{1}{x} + 1}. \tag{4.170}$$

Позначимо $\varphi(x) = \left(\frac{1}{x} + 1\right)^{1/2}$. Перевіримо виконання умов теореми про збіжність методу простої ітерації. Виберемо $x_0=1,5$, тоді $\delta=0,5$. Розглянемо

$$\varphi'(x) = -\frac{1}{2\sqrt{x^3 + x^4}}; \quad \max_{1\le x\le 2}|\varphi'(x)| = \frac{1}{2\sqrt{2}},$$

тобто $q = \frac{1}{2\sqrt{2}}$.

Тоді $|\varphi(x_0) - x_0| = \left|\sqrt{\frac{2}{3}+1} - 1{,}5\right| = 0{,}205, \quad (1-q)\delta = 0{,}5\cdot\left(1 - \frac{1}{2\sqrt{2}}\right) \approx 0{,}3232,$

а отже, умова (4.107) виконується. З формули (4.110) маємо, що кількість ітерацій, які необхідно провести для знаходження кореня з точністю $\varepsilon=10^{-4}$, повинна задовольняти умові $n \ge 8$. Відповідні значення x_n та $x_n - \varphi(x_n)$ наведені в табл. 4.7.

Таблиця 4.7

n	x_n	$x_n - \varphi(x_n)$
0	0150000E+01	0209006E+00
1	0129099E+01	0411454E– 01
2	0133214E+01	0901020E– 02
3	0132313E+01	0193024E– 02
4	0132506E+01	0415444E– 03
5	0132464E+01	0892878E– 04
6	0132473E+01	0191927E– 04
7	0132471E+01	0417233E– 05
8	0132472E+01	0953674E– 06

З нерівності (4.111) і отриманих результатів видно, що для досягнення заданої точності достатньо було провести 5 ітерацій (n=5). Взагалі слід зазначити, що апостеріорна оцінка (4.111) є більш точною і її використання може заощадити деяку кількість обчислень.

Приклад 4.14. Методом релаксації знайти найменший за модулем від'ємний корінь рівняння з точністю $\varepsilon=10^{-4}$

$$x^3 - 3x^2 - 1 = 0. \qquad (4.171)$$

Розв'язок. Спочатку виділимо корені рівняння (4.171) користуючись наступною таблицею:

Таблиця 4.8

x	–4	–3	–2	–1	0	1	2	3
sign$f(x)$	–	–	+	+	–	+	+	+

З таблиці видно, що рівняння має три корені, що розташовані на проміжках [–3;–2], [–1;0], [0;1]. Будемо знаходити корінь на проміжку [–1;0]. Обчисливши значення $f(-0{,}5) = -0{,}375$ можна уточнити проміжок існування кореня [–1;–0,5].

Позначимо $f(x) = x^3 - 3x^2 - 1$. Тоді $f'(x) = 3x^2 + 6x < 0, \quad x \in [-1;-0{,}5]$ і є монотонно зростаючою функцією на [–1;–0,5] (оскільки $f''(x) = 6x + 6 \ge 0$).

Тому
$$m_1 = \min_{x\in[-1;-0{,}5]}|f'(x)| = |f'(-0{,}5)| = 2{,}25,$$
$$M_1 = \max_{x\in[-1;-0{,}5]}|f'(x)| = |f'(-1)| = 3.$$

Тоді, згідно формул (4.124) і (4.125), рівняння матиме вигляд

$$x_{n+1} = x_n + \tau_{\text{опт}}(x_n^3 + 3x_n^2 - 1). \qquad (4.172)$$

Вибравши за початкове наближення точку x_0=–0,5 отримаємо оцінку $|z_0| \le 0{,}5$, а кількість ітерацій, які потрібно провести для знаходження розв'язку з точністю $\varepsilon=10^{-4}$ буде дорівнювати 5 (4.126). В табл. 4.9 наведені відповідні дані ітераційної послідовності:

Таблиця 4.9

n	x_n	$f(x_n)$
0	0500000E+00	0142857E+00
1	0642857E+00	0985700E– 02
2	0652714E+00	0105500E– 04
3	0652704E+00	0596046E– 07
4	0652704E+00	0000000E+00
5	0652704E+00	0000000E+00

Із наведених даних видно, що необхідна точність досягається раніше 5-ї ітерації. Це досить характерно для апріорних оцінок типу (4.126).

Приклад 4.15. Методом Ньютона знайти найменший додатній корінь рівняння

$$x^3+3x^2-1=0. \qquad (4.173)$$

з точністю $\varepsilon=10^{-4}$.

Розв'язок. З табл. 3 видно, що рівняння (4.173) має єдиний додатній корінь, що належить проміжку [0;1]. обчислимо $f(0{,}5)$=–0,125. Тепер знаходимо корінь на проміжку [0,5;1]. Нехай $f(x)=x^3+3x^2-1$. Тоді

$$f'(x) = 3x^2 + 6x > 0, \quad f''(x) = 6x + 6 > 0, \quad x \in [0{,}5;1].$$

$$m_1 = \min_{x\in[0{,}5,1]} |f'(x)| = |f'(0{,}5)| = 3{,}75,$$

$$M_2 = \max_{x\in[-1;-0{,}5]} |f''(x)| = |f''(1)| = 12.$$

Виберемо x_0=1, тоді $|x_0 - x_*| \le 0{,}5$. З формули (4.72) маємо

$$q = \frac{12 \cdot 0{,}5}{2 \cdot 3{,}75} = 0{,}8 < 1.$$

Тобто всі умови теореми про збіжність методу Ньютона виконані. З формули (4.81) маємо, що для досягнення заданої точності достатньо провести 7 ітерацій. Відповідні обчислення наведені в табл. 4.10.

Таблиця 4.10

n	x_n	$f(x_n)$
0	01000000E+01	03000000E+01
1	06666667E+00	06296297E+00
2	05486111E+00	06804019E– 01
3	05323902E+00	01218202E– 02
4	05320890E+00	04395228E– 06
5	05320889E+00	04230802E– 07
6	05320889E+00	04230802E– 07
7	05320889E+00	04230802E 07

4.15. Контрольні питання

1. Які функціональні рівняння називають алгебраїчними?
2. Які функціональні рівняння називають трансцендентними?
3. Що означає розв'язати задачу чисельним методом?
4. Які етапи є обов'язковими для знаходження всіх коренів функціонального рівняння на заданому проміжку?
5. В чому полягає суть алгоритму реалізації етапу відокремлення коренів функціонального рівняння на заданому проміжку?
6. В чому суть методу половинного ділення?
7. Як обирається остаточне значення розв'язку методом дихотомії?
8. В чому суть методу хорд?
9. Яка умова виходу з ітераційного процесу розв'язку методом хорд?
10. В чому суть методу січних?
11. Чим метод січних відрізняється від методу хорд?
12. В чому подібність методів січних та хорд?
13. Від яких факторів залежить швидкість збіжності методу січних?
14. В чому полягає суть методу дотичних?
15. В чому полягає суть методу Ньютона?
16. В чому полягає суть модифікованого методу Ньютона?
17. До якого вигляду потрібно привести функціональне рівняння, щоб провести розрахунок кореня за методом простої ітерації?
18. Які умови збіжності методу простої ітерації і як їх врахувати при розв'язуванні функціонального рівняння?
19. В чому полягає суть комбінованого методу?
20. В чому полягає суть методу релаксації?

4.16. Задачі для самостійної роботи

Задача 4.1. Визначте, скільки коренів має функціональне рівняння $\sin(x) - x^3 = 0$ на проміжку $[-3,3]$.

Задача 4.2. Знайдіть розв'язок функціонального рівняння $x^2 - 2^x = 0$ на проміжку $[-1,1]$, який отриманий в результаті відокремлення коренів.

Задача 4.3. Знайдіть розв'язки функціонального рівняння $x^2 - 0{,}5^x = 0$ на проміжку $[-5,5]$. Перевірте отримані розв'язки графічно.

Задача 4.4. Розв'яжіть функціональне рівняння $\sin(x)^{x^1-3x} = 0$ методом простої ітерації на проміжку $[0.5,3.5]$. Перевірте отримані розв'язки, скориставшись методом половинного ділення.

Задача 4.5. Вкажіть проміжок, на якому розв'язки рівняння $0{,}5^x - 0{,}1 = 0$ методом хорд буде неефективним, якщо $\varepsilon = 10^{-3}$.

Задача 4.6. Довести, що нулі многочлена

$$P_n(x) \equiv x^n + a_1 x^{n-1} + \ldots + a_{n-1}x + a_n,$$

для $a_n \neq 0$; $\{a_i, i = \overline{1,n}\} \subset \mathbf{R}^1$ за модулем не менше єдиного додатного нуля многочлена $P_n(x) \equiv x^n + |a_1|x^{n-1} + \ldots + |a_{n-1}|x - a_n$.

Задача 4.7. Нехай $\{d_0, d_1, \ldots, d_n\} \subset (0, +\infty)$ при $d_n \geq |a_1| d_{n-1} + |a_2| d_{n-2} \ldots + |a_n| d_0$. Довести, що нулі многочлена $P_n(x) \equiv x^n + a_1 x^{n-1} + \ldots + a_{n-1}x + a_n$ не перевищують за модулем

$$d_n^* \equiv \max\left\{\frac{d_n}{d_{n-1}}; \sqrt{\frac{d_n}{d_{n-2}}}; \sqrt[3]{\frac{d_n}{d_{n-3}}}; \ldots; \sqrt[n]{\frac{d_n}{d_0}}\right\}.$$

Задача 4.8. Нехай $p_0 > p_1 > p_2 > \ldots > p_n$. Довести, що у одиничному колі $|x| < 1$ не містить ні одного нуля многочлена $P_n(x) \equiv p_0 + p_1 x + \ldots + p_{n-1}x^{n-1} + p_n x^n$.

Задача 4.9. Використовуючи теорему Ролля, знайти умови того, що корені рівняння $x^m + px^n + 1 = 0$ є дійсними числами.

Задача 4.10. Знайти додатні корені рівняння $x = tg\, x$ з точністю $\varepsilon = 10^{-4}$.

Задача 4.11. Використати метод Ньютона для розв'язку рівняння

$$x^2 - N = 0 \text{ для обчислення } \sqrt{N}.$$

Задача 4.12. Обчислити з п'ятьма знаками два найменших додатних корені рівняння

$$\cos x \, ch\, x = 1.$$

Задача 4.13. Знайти з п'ятьма десятковими знаками розв'язок системи рівнянь

$$\sin x = y + 1.32; \quad \cos y = x - 0.85.$$

Задача 4.14. Знайти з п'ятьма десятковими розв'язок систем рівнянь

$$x^7 - 5x^2y^4 + 1510 = 0; \qquad y^5 - 3x^4y - 105 = 0,$$

якщо $x_0 = 2$, $y_0 = 3$.

Задача 4.15. Знайти одним з ітераційних методів дійсні корені рівнянь з точністю ε (наприклад $\varepsilon = 10^{-4}$).

№	Рівняння	№	Рівняння
1	$x^3 - 5x^2 + 4x + 0{,}092 = 0$	25	$x + \sin x - 12x = 0{,}25$
2	$x^3 - 4x^2 - 7x + 13 = 0$	26	$3x + \cos x + 1 = 0$
3	$x^4 + x^3 - 6x^2 + 20x - 16 = 0$	27	$x^2 + 4\sin x - 1 = 0$
4	$x^3 + \sin x - 12x + 1 = 0$	28	$x^3 - 3x^2 - 17x + 22 = 0$
5	$x^3 - 10x^2 + 44x + 29 = 0$	29	$3x - \cos x - 1 = 0$
6	$x^4 - 2x^3 - 3{,}74x^3 + 8{,}18x - 3{,}48 = 0$	30	$(x-1)^3 + 0{,}5e^x = 0$
7	$x^4 - 10x^3 + 48{,}16x^2 + 108{,}08x + 70{,}76 = 0$	31	$x^3 + 4\sin x = 0$
8	$x^4 - 3x^3 + 20x^2 + 44x + 54 = 0$	32	$x^3 - 3x^2 - 14x - 8 = 0$
9	$x^4 - 2x^3 + x^2 - 2x + 1 = 0$	33	$x^2 + 4\sin x = 0$
10	$x^3 - 2x^2 + x + 1 = 0$	34	$3x^2 - \cos^2 \pi x = 0$
11	$x^3 + 6x^2 + 9x + 2 = 0$	35	$x^3 - x - 1 = 0$
12	$(x-1)^2 - \sin 2x = 0$	36	$x^3 + 4x - 6 = 0$

№	Рівняння	№	Рівняння
13	$x^4+2x^2-6x+2=0$	37	$x^5+x-3=0$
14	$x^3-x^2+x-3=0$	38	$x^2\lg x-1=0$
15	$x^5-5x+2=0$	39	$e^x-2(x-1)^2=0$
16	$x^7+6x-5=0$	40	$x^4+4x-2=0$
17	$x^4+2x-2=0$	41	$3^{-x}-x^2+1=0$
18	$x^5-3x^2+1=0$	42	$x^7+x+4=0$
19	$x^6-3x^2+x-1=0$	43	$x^4+2x-1=0$
20	$(x-1)^2-0{,}5e^x=0$	44	$2^x+x^2-1{,}15=0$
21	$x^2\cos 2x=1$	45	$e^{-x}+x^2-2=0$
22	$x^2-3+0{,}5^x=0$	46	$x^2-10\sin x=0$
23	$5x^3+2x^2-15x-6=0$	47	$\mathrm{sh}x-12\mathrm{th}x-0{,}311=0$
24	$3x^4+4x^3-12x^2-5=0$	48	$x^3+5\cos x=0$

4.17. Завдання до лабораторної роботи

Варіант завдання 1 та 2 слід обрати згідно порядкового номера в списку групи (*відповідальним є староста групи*). Метод для реалізації обирається як: залишок від ділення порядкового номера на 6. Перед початком тестування програми необхідно провести графічне дослідження функції з метою виділення проміжків, на яких є дійсні корені рівняння. Програмна реалізація повинна забезпечувати знаходження *лише одного з коренів рівняння*, в залежності від початкового наближення.

В результаті тестування програми побудувати наступну таблицю:

Точність	Значення кореня	Кількість ітерацій
$\varepsilon=10^{-1}$		
$\varepsilon=10^{-2}$		
$\varepsilon=10^{-3}$		
$\varepsilon=10^{-4}$		

Зміст звіту до лабораторної роботи

1. Мета роботи.
2. Індивідуальне завдання.
3. Алгоритм розв'язування задачі (блок-схема, тобто ***алгоритм, а не теоретичні відомості***).
4. Текст програми з коментарями.
5. Результати виконання програми.
6. Висновки.

Варіанти завдання 1 (рівняння для розв'язування)

№	Рівняння	№	Рівняння
1	$x^4 - 3x^3 + 20x^2 + 44x + 54 = 0$	16	$x^3 - x - 1 = 0$
2	$x^4 + x^3 - 6x^2 + 20x - 16 = 0$	17	$x^3 + 4x - 6 = 0$
3	$x^3 - 10x^2 + 44x + 29 = 0$	18	$x^5 - 3x^2 + 1 = 0$
4	$3x^4 + 4x^3 - 12x^2 - 5 = 0$	19	$e^x - 2(x-1)^2 = 0$
5	$x^3 - 3x^2 - 17x + 22 = 0$	20	$e^{-x} + x^2 - 2 = 0$
6	$x^3 - 3x^2 - 14x - 8 = 0$	21	$x^4 + 4x - 2 = 0$
7	$5x^3 + 2x^2 - 15x - 6 = 0$	22	$x^4 + 2x - 1 = 0$
8	$x^3 + \sin x - 12x + 1 = 0$	23	$x^3 - x^2 + x - 3 = 0$
9	$x^3 - 4x^2 - 7x + 13 = 0$	24	$x^5 + x - 3 = 0$
10	$x^4 + 2x^2 - 6x + 2 = 0$	25	$x^7 + x + 4 = 0$
11	$x^3 + 6x^2 + 9x + 2 = 0$	26	$2^x + x^2 - 1{,}15 = 0$
12	$x^6 - 3x^2 + x - 1 = 0$	27	$3^{-x} - x^2 + 1 = 0$
13	$(x-1)^3 + 0{,}5e^x = 0$	28	$x^5 - 5x + 2 = 0$
14	$x^3 - 2x^2 + x + 1 = 0$	29	$x^7 + 6x - 5 = 0$
15	$x^4 - 2x^3 + x^2 - 2x + 1 = 0$	30	$x^4 + 2x - 2 = 0$

Варіанти завдання 2 (рівняння для розв'язування)

№	Рівняння	№	Рівняння
1	$\ln^2(x-1) = 3\cos 2x + 1$	2	$\sqrt{25 - x^2} = \operatorname{arctg} 2x$
3	$\frac{3\pi}{2}\cos x = e^{0{,}1x^2}\operatorname{arcctg} 2x$	4	$\sin x\sqrt{81 - x^2} = 5x\operatorname{arctg} x$
5	$10e^{-x^2} = \sqrt{2\pi x} + \sin x$	6	$\operatorname{arctg} 2x - 0{,}2(x-1)^4 + \sin x = 0$
7	$\sqrt{\ln^2(x-1)e^{\sin 3x}} = 10e^{-0{,}1x^2}$	8	$\sin 3x\sqrt{64 - x^2} = 5xe^{0{,}1x}$
9	$\sqrt{36 - x^2}\lg x = \sin 4x$	10	$\operatorname{arctg} 2x - \frac{(x-1)^4}{5} + \sin^2 5x = 0$
11	$\frac{10}{1+x^2} = 2\sin 2x + x$	12	$10e^{-0{,}1x^2} = \sqrt{2\pi + x} + \sin 2x$
13	$\sin 4x\sqrt{81 - 25x^2} = 5x\operatorname{arctg} x$	14	$\sin^2 3x\sqrt{16 - x^2} = 5ex^{0{,}2x}$
15	$\frac{10x}{1+x^2} = 2\sin 2x + x$	16	$\frac{x^2 - 4}{x^2 + 1} = \sqrt{x}e^{x\sin x}$
17	$\arcsin x - \sin 5x\sqrt[4]{1 - x^4} = 0$	18	$4x\operatorname{tg}(0{,}5\sqrt{9 - x^2} = 10\sin 3x$
19	$\frac{x^2 - 4x}{x^2 - 4x + 8} = \sqrt[3]{x^3 + 4}e^{\cos 3x}$	20	$\frac{x-1}{x^2 - 2x + 2} = \sqrt[4]{x^4 + 4}e^{\sin 2x}$

21	$\frac{10x-2}{3+x^2} = 2\cos 2x + \sqrt[4]{x}$	22	$\frac{x^2-9}{x^2-4} = \sqrt{x^2+1}e^{x\cos x}$
23	$\sqrt{64-x^2}\log_2 x = \sin 3x$	24	$\frac{x^2-4}{x^2+1} = \sqrt{x}e^{x\sin x}$
25	$10e^{-0,3x^2} = \sqrt{2\pi x + x^2} + x^2 + 3\sin x$	26	$4x\mathrm{tg}(0,5\sqrt{9-x^2}) = 10\sin 3x$
27	$5(3^{-x^2}) + 1 = \sqrt{3x} + \sin 2x$	28	$\mathrm{arctg}2x - (x-0,1)^4 + \sin^2 x = 0$
29	$\frac{5\pi}{2}\cos 2x = 3^{0,1x^2}\mathrm{arctg}2x$	30	$\sin^2 x\sqrt{81-x^2} = 5e^{-x^2}$
31	$\sqrt{64-x^2}\log_2 x = \sin 3x$	32	$\mathrm{arctg}^2 x - \frac{(x-1)^3}{5} + 2\sin^2 5x = 0$
33	$\frac{3x-4}{2+x^2} = 2\cos^2 3x + \sqrt[4]{x^2+2}$	34	$\frac{x^2-16}{x^2+4} = \sqrt{x^4+x}e^{-x\cos x}$

Розділ 5.
Задача інтерполяції та наближення функцій

- Основи теорії інтерполювання та числового диференціювання
  - Загальна постановка задачі інтерполювання
  - Загальна задача лінійного інтерполювання
  - Постановка задачі інтерполяції функції однієї дійсної змінної
  - Поліноміальна інтерполяція. Існування та одиничність інтерполяційного полінома
  - Типи інтерполювання
- Інтерполяційний поліном Лагранжа
  - Аналіз загальної постановки задачі
  - Інтерполяційна формула Лагранжа. Частковий випадок
  - Побудова інтерполяційного многочлена Лагранжа
  - Інтерполяційний многочлен Лагранжа з рівновіддаленими вузлами
  - Інтерполяційна схема О. Ейткена
  - Похибка інтерполювання для інтерполяційного многочлена Лагранжа
- Вибір вузлів інтерполювання
  - Многочлени Чебишева. Властивості
  - Вибір вузлів інтерполювання
  - Поведінка похибки на інтервалі інтерполювання
- Інтерполяційний поліном Ньютона
  - Аналіз загальної постановки задачі
  - Інтерполяційний многочлен Ньютона
  - Поділені різниці та їх властивості
  - Інтерполяційна формула Ньютона за поділеними різницями та нерівновіддаленими вузлами
  - Скінченні різниці та їх властивості
  - Формула Ньютона вперед і назад за скінченними різницями
- Числове диференціювання
  - Числове диференціювання для нерівновіддалених вузлів
  - Формули числового диференціювання для рівновіддалених вузлів
  - Формули числового диференціювання методом невизначених коефіцієнтів

- Раціональна інтерполяція
- Інтерполювання функцій комплексної змінної
- Інтерполювання функцій багатьох змінних
    - Труднощі інтерполювання функцій багатьох змінних
    - Побудова інтерполяційних формул за поділеними різницями
- Збіжність інтерполяційного процесу
- Інтерполювання сплайнами
- Приклади розв'язку задачі інтерполяції функції
- Контрольні питання
- Задачі для самостійної роботи
- Завдання до лабораторної роботи

Інтерполяція – в обчислювальній математиці спосіб знаходження проміжних значень величини за наявним дискретним набором відомих значень [65, 71, 76-80, 88, 92-93, 106-110, 113-114, 116-133, 137, 140-145, 147-151, 161-164, 168-169, 192-194, 196-197, 201-209, 211].

Багатьом із тих, хто стикається з науковими та інженерними розрахунками часто доводиться оперувати наборами значень, отриманих експериментальним шляхом чи методом випадкової вибірки. Як правило, на підставі цих наборів потрібно побудувати функцію, зі значеннями якої могли б з високою точністю збігатися інші отримувані значення. Така задача називається апроксимацією кривої. Інтерполяцією називають такий різновид апроксимації, при якій крива побудованої функції проходить точно через наявні точки даних. Існує також близька до інтерполяції задача, що полягає в апроксимації якої-небудь складної функції іншою, простішою функцією. Якщо деяка функція занадто складна для продуктивних обчислень, можна спробувати обчислити її значення в декількох точках, а за ними побудувати, тобто інтерполювати, простішу функцію. Зрозуміло, використання спрощеної функції не дозволяє одержати такі ж точні результати, які давала б початкова функція. Але, для деяких класів задач, досягнутий виграш у простоті і швидкості обчислень може переважити отриманий огріх у результатах.

Варто також згадати і зовсім інший різновид математичної інтерполяції, відому за назвою «інтерполяція операторів». До класичних робіт з інтерполяції операторів відносяться теорема Рисса-Торина (Riesz-Thorin theorem) і теорема Марцинкевича (Marcinkiewicz theorem), що є основою для багатьох інших робіт.

5.1. Основи теорії інтерполювання та числового диференціювання

5.1.1. Загальна постановка задачі інтерполювання

При розв'язуванні багатьох задач фізики та техніки в деякі моменти часу одержуємо значення функції, які зводять у табл. 5.1.

Таблиця 5.1

x_0	x_1	x_2	...	x_{n-1}	x_n
$f(x_0)$	$f(x_1)$	$f(x_2)$	...	$f(x_{n-1})$	$f(x_n)$

Аналітичного вигляду функції $f(x)$, як правило, не знаємо, але треба знайти значення цієї функції $f(x)$ для $x \neq x_j$, $j=\overline{0,n}$, $x_j \in [a,b] \subset \mathbf{R}^1$. Як поступити?

Наприклад, слід побудувати якимось чином нову функцію $\varphi(x)$, яка в тому чи іншому розумінні близьку до шуканої функції $f(x)$, але так щоб

$$\varphi(x_j) = f(x_j),\ j=\overline{0,n}, \tag{5.1}$$

тобто за умови, що $|\varphi(x) - f(x)| < \varepsilon$, для $x \neq x_j$, $j=\overline{0,n}$, де $\varepsilon > 0$ мале.

5.1.2. Загальна задача лінійного інтерполювання

Нехай задана функція $f(x)$. З деякого класу $\gamma([a,b])$, $x \in [a,b]$, та відомі її значення

$$f_j \equiv f(x_j),\ x \in [a,b],\ j=\overline{0,n}. \tag{5.2}$$

Візьмемо систему $(n+1)$ - лінійно-незалежних на $[a,b]$ функцій

$$\{\varphi_i(x), i=\overline{0,n}\}, \tag{5.3}$$

Наприклад: $\{\varphi_i(x)\} \equiv \{1, x, x^2, ..., x^n\}$; $\{\varphi_i(x)\} \equiv \{1, e^{\alpha_1 x}, e^{\alpha_2 x}, ..., e^{\alpha_n x}, \alpha_i \neq \alpha_j\}$;

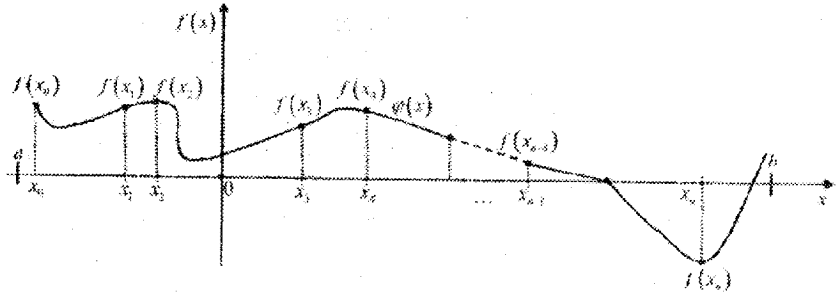


Рис. 5.1. Приклад задачі лінійного інтерполювання

$\{\varphi_i(x)\} \equiv \left\{1, \sin n, \cos n, ..., \cos\frac{\pi}{2}n\right\}$.

Утворюємо нову функцію $\varphi(x)$ з класу $\gamma([a,b])$ за законом

$$\varphi(x) = \sum_{i=0}^{n} a_i \varphi_i(x), \tag{5.4}$$

де $a_i \in \mathbf{R}^1$ поки що невідомі параметри, $\varphi_i \in \gamma([a,b])$.

Означення 5.1. *Функція $\varphi(x)$, що задана формулою (5.4), назвемо узагальненим многочленом за системою функцій $\{\varphi_i(x), i=\overline{0,n}\}$.*

Виникла задача побудувати $\varphi(x)$, підбираючи $a_i \in \mathbf{R}^1$, $i=\overline{0,n}$ так щоб $\varphi(x) \approx f(x)$ у деякому розумінні, але за умови (5.2).

Етапи її розв'язання наступні :

1) Вибираємо довільні $(n+1)$ точок з $[a,b]$: $\{x_0, x_1, ..., x_n, x_i \neq x_j, i \neq j\} \subset [a,b]$, які назвемо вузлами інтерполяції;
2) Задовільняємо умову (5.2) яка еквівалентна

$$\sum_{i=0}^{n} a_i \varphi_i(x_j) = f(x_j), \; j=\overline{0,n}, \tag{5.5}$$

яка представляє собою систему лінійних алгебраїчних рівнянь(СЛАР).

3) Розв'язуючи (5.5) матимемо $a_i \in \mathbf{R}^1$, $i=\overline{0,n}$.

Означення 2. *Узагальнений многочлен $\varphi(x)$, який володіє властивістю (5.5) називається узагальненим інтерполяційним многочленом для функції $f(x)$ за системою вузлів $\{x_0, x_1, ..., x_n\} \subset [a,b]$, а сама умова (5.2) (або 5.5) називається умовою інтерполювання функції $f(x)$.*

Щоб система лінійних алгебраїчних рівнянь (5.5) мала єдиний розв'язок достатньо, щоб головний визначник (5.5) не дорівнював нулеві:

$$\Delta = \begin{vmatrix} \varphi_0(x_0) & \varphi_1(x_0) & \dots & \varphi_n(x_0) \\ \varphi_0(x_1) & \varphi_1(x_1) & \dots & \varphi_n(x_1) \\ \dots & \dots & \dots & \dots \\ \varphi_0(x_n) & \varphi_1(x_n) & \dots & \varphi_n(x_n) \end{vmatrix} \neq 0. \tag{5.6}$$

Означення 5.3. *Система лінійно-незалежних функцій $\{\varphi_0(x), \varphi_1(x), ..., \varphi_n(x)\} \subset \gamma([a,b])$, яка задовольняє (5.6), називається системою функцій Чебишова.*

Чебишов Пафнутій Львович (16.05.1821 – 08.12.1894) – російський математик, механік, засновник Петербурзької математичної школи. Академік Петербурзької АН. Автор понад 70 наукових праць з теорії наближень функцій, теорії ймовірностей.

Одержимо вигляд узагальненого многочлена (5.4), якщо в якості $(n+1)$ - лінійно-незалежних функцій $\{\varphi_i(x), i=\overline{0,n}\}$ взяти $(n+1)$-лінійно-незалежні функції $\{x^i, i=\overline{0,n}\}$ $\{x^i, i=0,1,...,n\}$.

5.1.3. Постановка задачі інтерполяції функції однієї дійсної змінної

Обмежимося розглядом задачі інтерполяції для функції. Нехай на проміжку $a \le x \le b$ задана невироджена сітка $\overline{\omega}$

$$\overline{\omega} = \{x_0 = a < x_1 < x_2 < ... < x_n = b\}$$

та в її вузлах відомі значення функції $y = f(x)$:

$$f(x_0) \equiv y_0; f(x_1) \equiv y_1; ...; f(x_{n-1}) \equiv y_{n-1}; f(x_n) \equiv y_n.$$

В задачі інтерполяції необхідно побудувати інтерполяційну функцію – інтерполянту $g(x)$ так, щоб її значення співпали б зі значеннями функції $f(x)$, що інтерполюється, у вузлах сітки $\overline{\omega}$, тобто

$$g(x_i) = y_i, \; i = \overline{0,n}. \tag{5.7}$$

При цьому основна мета інтерполяції – отримати швидкий та економічний алгоритм розрахунку значень функції $g(x)$ для значень x, які не містяться в початковій таблиці, тобто $x \in [a,b]$ та $x_i \ne x_j$. Як обрати інтерполянту $g(x)$ та як оцінити похибку інтерполяції $\|f(x) - g(x)\|$ (в деякій формі)?

Відповідаючи на ці питання, ми обмежимось розглядом задачі лінійної інтерполяції (відносно базисних функцій $\Phi_k(x)$), тобто розглянемо ситуацію, коли інтерполяційна функція $g(x)$ будується у вигляді лінійної комбінації деяких базисних функцій $\{\Phi_k(x)\}$ в свою чергу достатньо елементарних $g(x) = \sum_{k=0}^{n} c_k \Phi_k(x)$, де $\{\Phi_k(x)\}$ – фіксовані, лінійно незалежні функції: $c_0, c_1, ..., c_n$ – невизначені поки коефіцієнти. Із умов інтерполяції (5.7) отримуємо систему $(n+1)$ рівняння відносно коефіцієнтів $\{c_k\}$: $\sum_{k=0}^{n} c_k \Phi_k(x_i) = y_i, \; i = \overline{0,n}$.

Припустимо, що система функцій $\{\Phi_k(x)\}$ така, що при будь-якому невиродженому виборі вузлів сітки $\overline{\omega}$ відмінний від нуля визначник.

$$\Delta(\Phi_0, \Phi_1, ..., \Phi_n) = \begin{vmatrix} \Phi_0(x_0) & \Phi_1(x_0) & ... & \Phi_n(x_0) \\ \Phi_0(x_1) & \Phi_1(x_1) & ... & \Phi_n(x_1) \\ ... & ... & ... & ... \\ \Phi_0(x_n) & \Phi_1(x_n) & ... & \Phi_n(x_n) \end{vmatrix} \ne 0.$$

Таку систему інтерполяційних функцій називають чебишевською системою інтерполяційних функцій. В цьому випадку на даній сітці $\overline{\omega}$ за відомими значеннями y_i, $i = \overline{0,n}$ однозначно визначаються коефіцієнти інтерполяційного багаточлена $\{c_k\}_{k=\overline{0,n}}$.

Теорема. Для розвязання задачі лінійної інтерполяції необхідно і достатньо, щоб система функцій $\{\Phi_k(x)\}$ утворювала на $[a,b]$ чебишевську систему інтерполяційних функцій.

В якості систем лінійно незалежних інтерполяційних функцій $\{\Phi_k(x)\}$ частіше всього обирають степеневі або поліноміальні функції $\Phi_k(x) = x^k$; тригонометричні функції $\Phi_k(x) = \begin{Bmatrix}\cos(kx)\\ \sin(kx)\end{Bmatrix}$; показникові функції $\Phi_k(x) = r^{xk}$ тощо.

Обмежимося розглядом випадку поліноміальної інтерполяції.

Зауваження. Табличний спосіб подання функції $y = f(x)$ та сітці $x \in \overline{\omega}$ та пов'язана із цим необхідність інтерполяції функції найбільш характерні для представлення результатів фізичного експерименту та для опису дискретної або обчислювальної моделі на ПК.

Задача інтерполяції, як задача перебудови таблиць значень функції із одної сітки на іншу, є характерною задачею опрацювання даних фізичного експерименту.

В подальшому ми використовуємо розв'язок задачі інтерполяції при побудові наближених методів розрахунку інтегралів; при різницевій апроксимації диференціальних рівнянь на основі інтегральних тотожностей; в задачах мінімізації та при розгляді інших питань. Постановка задачі інтерполяції в формі (5.7) не єдино можлива. Можливі й інші постановки задачі інтерполяції, наприклад, задача інтерполяції Ерміта та інші види задачі інтерполяції. Задача інтерполяції Ерміта повязана із інтерполяцією функції за відомими значеннями функції та її похідних до n-го порядку включно, заданими в деякій точці $x = x_0$.

5.1.4. Поліноміальна інтерполяція. Існування та одиничність інтерполяційного полінома

Нехай вхідною інформацією для нас є множина точок $\{(x_k, y_k)\}_{k=\overline{0,n}}$. Інтерполянту $g(x)$ шукають у вигляді інтерполяційного полінома

$$g(x) \equiv P_n(x) = \sum_{k=0}^{n} c_k x^k . \qquad (5.8)$$

Умови інтерполяції (5.7) $g(x_i) = y_i$ приводять до системи лінійних алгебраїчних рівнянь для коефіцієнтів $\{c_k\}$:

$$\begin{cases} c_0 + c_1 x_0 + \ldots + c_n x_0^n = y_0 \\ c_0 + c_1 x_1 + \ldots + c_n x_1^n = y_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ c_0 + c_1 x_n + \ldots + c_n x_n^n = y_n \end{cases}$$

Визначник цієї системи (визначник Вандермонда) відмінний від нуля на довільній невирожденій сітці $\overline{\omega}$

$$\Delta = \begin{vmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n-1} & x_{n-1}^2 & \dots & x_{n-1}^n \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{vmatrix} =$$

$$= (x_n - x_{n-1})\dots(x_n - x_0)(x_{n-1} - x_{n-2})\dots(x_{n-2} - x_0)\dots(x_1 - x_0) = \prod_{0 \le m \le k \le n} (x_k - x_m) \neq 0$$

Тим самим, система функцій $\{x^k\}$ – чебишевська система інтерполяційних функцій на $[a,b]$ та справедлива наступна теорема.

Нагадаємо Вандермонд Олександр Теофіл (Шарль Огюст) (28.02.1735 - 01.01.1796) – французький математик, член Паризько АН (1771), учасник Великої французької революції.

Теорема. Інтерполяційний поліном (5.8) існує та єдиний. Але інтерполяційний поліном може бути записаний в різній формі.

5.1.5. Типи інтерполювання

Нехай функція $y=f(x)$ задана таблично, тобто у вигляді множини пар чисел $(x_i; y_i), i = 0,1,\dots,n$, що отримані в результаті експерименту, але невідома аналітично, тобто:

x	x_0	x_1	x_2	…	x_n
y	y_0	y_1	y_2	…	y_n

Апроксимуючу функцію $\varphi(x)$ будемо будувати таким чином, щоб її значення в точках $\{x_i, i = \overline{0,n}\}$ співпадали с табличними значеннями заданої функції $f(x)$:

$$\varphi(x_0) = y_0,\ \varphi(x_1) = y_1, \dots, \varphi(x_n) = y_n.$$

Такий спосіб введення апроксимуючої функції називають ***інтерполяцією за Лагранжем***. Апроксимуюча функція $\varphi(x)$, що задовільняє умовам Лагранжа, називається *інтерполяційною функцією*. Значення $\{x_i, i = \overline{0,n}\}$ в даному контексті називають вузлами інтерполяції.

Якщо значення аргументу разташовано всередині інтервалу $[x_0, x_n]$, то знаходження наближеного значення функції $f(x)$ називають ***інтерполяцією***, якщо апроксимуючу функцію знаходять за інтервалом $[x_0, x_n]$, то такий процес називають ***екстраполяцією***. Походження цих термінів пов'язано з латинськими словами: *inter* – між, вередині, *extra* – поза, *pole* – вузол.

Для знахождення $\varphi(x)$ необхідно знайти коефіцієнти полінома $a_0, a_1, \dots, a_n$. При цьому інтерполяція буде називатися ***глобальною***, якщо поліном $\varphi(x)$ застосовується на всьому інтервалі зміни x, при цьому степінь полінома співпадає з кількістю точок в таблиці і коефіцієнти $a_0, a_1, \dots, a_n$ визначається з розв'язку системи лыныйних рывнянь:

$$\begin{cases} \varphi(x_0) = a_0 + a_1 x_0 + a_2 x_0^2 + \ldots + a_n x_0^n = y_0 \\ \varphi(x_1) = a_0 + a_1 x_1 + a_2 x_1^2 + \ldots + a_n x_1^n = y_1 \\ \varphi(x_2) = a_0 + a_1 x_2 + a_2 x_2^2 + \ldots + a_n x_2^n = y_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \varphi(x_n) = a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_n x_n^n = y_n \end{cases}$$

Розв'язуючи систему $n+1$ рівнянь з $n+1$ невідомими відносно a_0, a_1, ..., a_n, отримаємо формулу для інтерполяційного поліному $\varphi(x)$. Розв'язування таким чином сформульованої задачі має недолік: якщо наявна велика кількість точок, то доводиться розв'язувати систему дуже великої розмірності. Тому при розв'язуванні задач глобальної інтерполяції використовують більш простіші допоміжні задачі. Інтерполяція називається ***локальною***, якщо інтерполяційні поліноми будуються окремо для різних частин наявного інтервалу змін x. При цьому степінь поліному є меншою від кількості розглянутих точок. У зв'язку з цим разрізняють різні подходи до побудови інтерполяційних поліномів. Для розв'язку задач глобальної інтерполяції потрібно розглянути ряд допоміжних задач, наприклад: нехай у вузлах інтерполяції функція $f(x_i) = y_i = 0$ у всіх точках, крім однієї $i = k$, а в цій точці приймають значення, рівне 1 (рис. 4), тобто виконується умова $y(x_k) = \begin{cases} 1, & i = k; \\ 0, & i \neq k. \end{cases}$

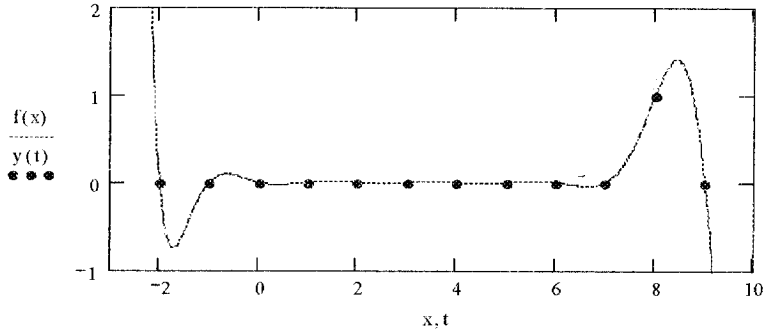


Рис. 5.2а Допоміжна функція

Необхідно побудувати такий поліном $y_k(x)$, щоб $y_k(x_i) = f(x_i)$. Якщо поліном $y_k(x)$ прямує до нуля в точці x_k, то в ньому присутній множник $(x - x_k)$. Таким чином: $y_k(x) = c_k(x - x_0)(x - x_1)(x - x_2)\ldots(x - x_{k-1})(x - x_{k+1})\ldots(x - x_n)$, де c_k –сталі коефіцієнти. Знайдемо значення полінома в точці x_k, враховуючи, що $1 = y_k(x_k)$, отримаємо:

$$1 = y_k(x_k) = c_k(x_k - x_0)(x_k - x_1)(x_k - x_2)\ldots(x_k - x_{k-1})(x_k - x_{k+1})\ldots(x_k - x_n)$$

$$c_k = \frac{1}{(x_k - x_0)(x_k - x_1)(x_k - x_2)\ldots(x_k - x_n)}.$$

Підставимо знайдений вираз c_k у формулу для $y_k(x)$:

$$y_k(x) = \frac{(x - x_0)(x - x_1)(x - x_2)\ldots(x - x_{k-1})(x - x_{k+1})\ldots(x - x_n)}{(x_k - x_0)(x_k - x_1)(x_k - x_2)\ldots(x_k - x_{k-1})(x_k - x_{k+1})\ldots(x_k - x_n)} = \frac{\prod_{j=0, j\neq k}^{n}(x - x_j)}{\prod_{j=0, j\neq k}^{n}(x_k - x_j)}.$$

Нехай тепер в точці x_k поліном приймає значення y_k, а в решту вузлах інтерполяції він, як і раніше, буде рівний 0. Необхідно побудувати поліном $\hat{y}_k(x)$ такий, що $\hat{y}_k(x_i) = f(x_i)$. Виходячи із розв'язку попередньої задачі, отримаємо:

$$c_k = \frac{y_k}{(x_k - x_0)(x_k - x_1)\ldots(x_k - x_{k-1})(x_k - x_{k+1})\ldots(x_k - x_n)},$$

тоді:

$$\hat{y}_k(x) = y_k y_k(x) = y_k\left(\frac{\prod_{j=0, j=k}^{n}(x - x_k)}{\prod_{j=0, j\neq k}^{n}(x_j - x_k)}\right).$$

Для розв'язку загальної задачі інтерполяції необхідно знайти поліном $L_n(x_i) = y_i$ тобто розв'язати $n+1$ допоміжних задач, допускаючи, що $y_k = 0$ окрім $k = 0, 1, \ldots, n$, тоді

$$L_n(x) = y_0(x)y_0 + y_1(x)y_1 + y_2(x)y_2 + \ldots + y_n(x)y_n = \sum_{k=0}^{n} y_n(x)y_k.$$

Підставляючи шукані значення $y_k(x)$ у формулу для $L_n(x)$, будемо мати:

$$L_n(x) = \sum_{k=0}^{n} y_k \frac{\prod_{j=0, j\neq k}^{n}(x - x_j)}{\prod_{j=0, j\neq k}^{n}(x_k - x_j)}. \qquad (5.9)$$

Для заданої функції $f(x)$ побудований таким чином поліном буде єдиним. Поліном, поданий формулою (5.9) називається ***інтерполяційним поліномом Лагранжа.*** Зокрема, при $n = 1$ отримаємо ***лінійну інтерполяцію***

$$L_1(x) = y_0\frac{(x - x_1)}{(x_0 - x_1)} + y_1\frac{(x - x_0)}{(x_1 - x_0)}. \qquad (5.10)$$

При $n = 2$ отримаємо ***квадратичну інтерполяцію***. В такому випадку інтерполюючою функцією буде парабола. Розв'язок:

$$L_2(x) = y_0 \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} + y_1 \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} + y_2 \frac{(x-x_1)(x-x_0)}{(x_2-x_1)(x_2-x_0)}. \qquad (5.11)$$

З допомогою інтерполяційного поліному можно знайти значення невідомої функції $f(x)$ в довільних точках, що не являються вузлами інтерполювання. Поліном Лагранжа є зручний тим, що вузли інтерполяції можуть бути разташовані довільним чином.

Приклад 5.1. Нехай функція $y = f(x)$ задана таблично:

X	0.5	0.6	0.7	0.8
Y	1.6	1.8	2.1	2.7

Завдання:

1) побудувати лінійний поліном Лагранжа на відрізку [0.5; 0.6];

2) побудувати квадратичний поліном Лагранжа на відрізку [0.6; 0.8];

3) побудувати поліном Лагранжа третього степеня на відрізку [0.5; 0,8];

4) з допомогою лінійної інтерполяції знайти $y(0.53)$;

5) з допомогою квадратичної інтерполяції знайти $y(0.75)$;

6) всі результати подати у графічному вигляді.

Розв'язок: 1. Для побудови лінійного інтерполяційного полінома Лагранжа на відрізку [0.5; 0.6] визначимо: $x_0 = 0.5$, $x_1 = 0.6$ – вузли інтерполяції; $y_0 = 1.6$, $y_1 = 1.8$ – значення функції в цих вузлах; $h = x_1 - x_0 = 0.6 - 0.5 = 0.1$ – крок, підставивши вказані значення у формулу (5.10): $L_1(x) = 1.6\frac{x-0.6}{0.5-0.6} + 1.8\frac{x-0.5}{0.6-0.5}$, а потім спростивши вказані вирази:

$$L_1(x) = \frac{1.6(x-0.6)}{-0.1} + \frac{1.8(x-0.5)}{0.1} = \frac{1}{0.1}(-1.6(x-0.6) + 1.8(x-0.5)) =$$

$$= \frac{1}{0.1}(-1.6x + 0.96 + 1.8x - 0.9) = 10 \cdot (0.2x + 0.06) = 2x + 0.6$$

– лінійний інтерполяційний поліном Лагранжа на відрізку [0.5; 0.6].

2. Для побудови квадратичного інтерполяційного поліному Лагранжа на відрізку [0.6; 0.8] визначимо: $x_0 = 0.6$, $x_1 = 0.7$, $x_2 = 0.8$ – вузли інтерполяції; $y_0 = 1.8$, $y_1 = 2.1$, $y_2 = 2.7$ – значення функції у вузлах інтерполяції; $h = x_2 - x_1 = = x_1 - x_0 = 0.1$ – крок, підставивши значення у формулу (5.11):

$$L_2(x)=1.8\frac{(x-0.7)(x-0.8)}{(0.6-0.7)(0.6-0.8)}+2.1\frac{(x-0.6)(x-0.8)}{(0.7-0.6)(0.7-0.8)}+2.7\frac{(x-0.6)(x-0.7)}{(0.8-0.6)(0.8-0.7)}=$$

$$=1.8\frac{x^2-0.7x-0.8x+0.56}{0.02}+2.1\frac{x^2-0.6x-0.8x+0.48}{-0.01}+2.7\frac{x^2-0.6x-0.7x+0.42}{0.02}=$$

$$=\frac{1}{0.01}(\frac{1.8}{2}(x^2-1.5x+0.56)-2.1(x^2-1.4x+0.48)+\frac{2.7}{2}(x^2-1.8x+0.42))=$$

$$=15x^2-16.5x+6.3$$

– квадратичний інтерполяційний поліном Лагранжа на відрізку [0.5; 0.7].

3. Для побудови $L_3(x)$ слід використати всі точки з таблиці. При цьому, якщо $n = 3$, то формула (5.9) набуде вигляду:

$$L_3(x)=y_0\frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}+y_1\frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}+$$

$$+y_2\frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}+y_3\frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)}$$

Підставивши табличні дані у формулу для $L_3(x)$, отримаємо:

$$L_3(x)=1.6\frac{(x-0.6)(x-0.7)(x-0.8)}{(0.5-0.6)(0.5-0.7)(0.5-0.8)}+1.8\frac{(x-0.5)(x-0.7)(x-0.8)}{(0.6-0.5)(0.6-0.7)(0.6-0.8)}+$$

$$+2.1\frac{(x-0.5)(x-0.6)(x-0.8)}{(0.7-0.5)(0.7-0.6)(0.7-0.8)}+2.7\frac{(x-0.5)(x-0.6)(x-0.7)}{(0.8-0.5)(0.8-0.6)(0.8-0.7)}=$$

$$=1.6\frac{x^3-2.1x^2+1.46x-0.336}{-0.1(-0.2)(-0.3)}+1.8\frac{x^3-2x^2+1.31x-0.28}{0.1(-0.1)(-0.2)}+$$

$$+2.1\frac{x^3-1.9x^2+1.18x-0.24}{0.2\cdot0.1(-0.1)}+2.7\frac{x^3-1.8x^2+1.03x-0.21}{0.3\cdot0.2\cdot0.1}=$$

$$=\frac{1}{0.1\cdot0.2\cdot0.1}(-\frac{1.6}{3}(x^3-2.1x^2+1.46x-0.336)+1.8(x^3-2x^2+1.31x-0.28)-$$

$$-2.1(x^3-1.9x^2+1.18x-0.24)+\frac{2.7}{3}(x^3-1.8x^2+1.03x-0.21)\approx$$

$33.5x^3-5.55x^2+792.8x-5$ – поліном Лагранжа третьої степені.

4. Так як точка $x=0.53\in[0.5;0.6]$, для обчислення $y(0.53)$ розглянемо лінейний інтерполяційний поліном Лагранжа $y=2x+0.6$, замість x підставимо значення 0.53: $y(0.53)=2\cdot0.53+0.6=1.06+0.6=1.66$.

5. Так як точка $x_* = 0.75 \in [0.6;0.8]$, для знаходження $y(0.75)$ обсремо квадратичний інтерполяційний поліном, замість x підставимо значення 0.75:

$$y(0.75) = 15 \cdot (0.75)^2 - 16.5 \cdot 0.75 + 6.3 = 8.44 - 12.38 + 6.3 = 2.36$$

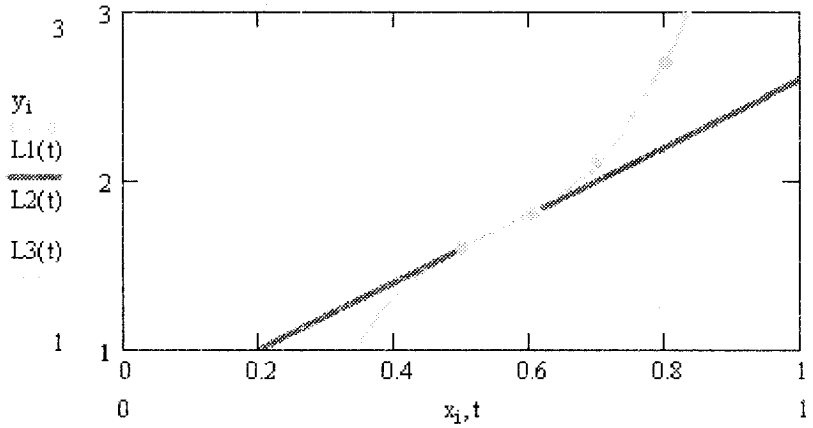


Рис. 5.2б. Інтерполяційні поліноми

Побудувавши графіки (рис.5) можна побачити, що:

1) поліном $L_1(x)$ проходить через дві задані точки $x = 0.5$, $x = 0.6$;

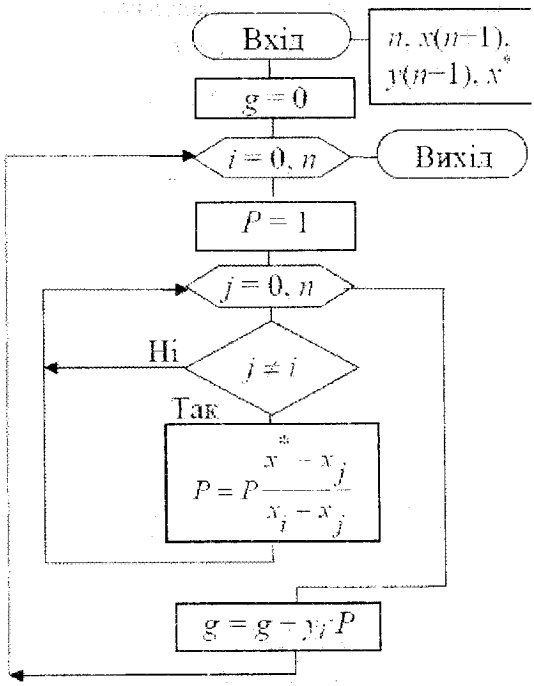
2) поліном $L_2(x)$ проходить через три занні точки $x = 0.6$, $x = 0.7$, $x = 0.8$;

3) поліном $L_3(x)$ проходить через усі, задані в таблиці точки.

Приклад 5.2. Приклад реалізації задачі інтерполяції поліномом Лагранжа

Завдання: Ввести з клавіатури парні набори значень *xi* та *yi* для функції $y = f(x)$ й обчислити для цієї функції інтерполяційний многочлен Лагранжа. Обчислити наближені значення на цьому проміжку у заданій кількості точок.

Схема алгоритму інтерполяції цієї задачі поліномом Лагранжа.

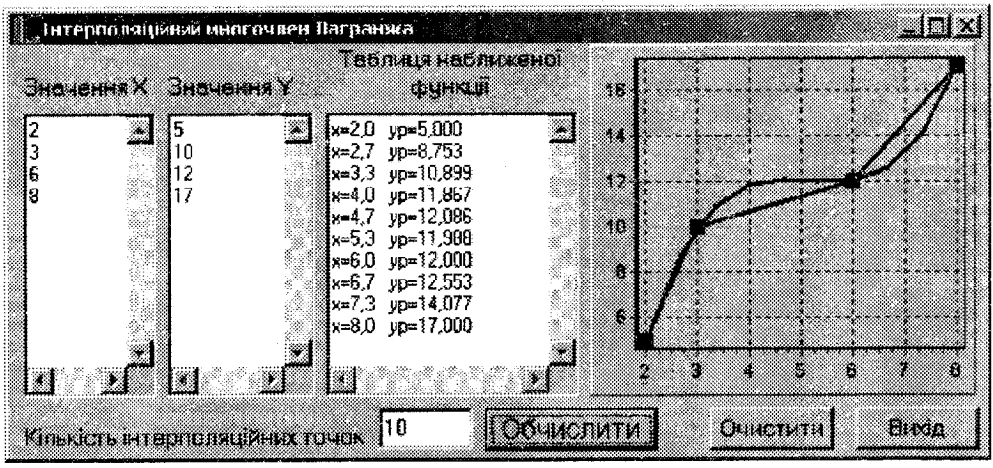


Текст програми

```
//--------------- Функція полінома Лагранжа----------------------
double lagrang(double *x, double *y, int n, double xp)
{ int i, j; double p, yr;
yr = 0;
for(i=0; i<n; i++)
{
p = 1;
for (j=0; j<n; j++)
if(i != j) p *= (xp - x[j]) / (x[i] - x[j]) ;
yr += y[i] * p;
}
return yr;
}
//--------------- Кнопка "Обчислити"----------------------
void __fastcall TForm1::Button1Click(TObject *Sender)
{ double x[20], y[20]; int i, n, n1;
Series1->Clear(); Series2->Clear(); Memo3->Clear();
n=Memo1->Lines->Count;
n1=Memo2->Lines->Count;
if(n != n1)
{ ShowMessage("Перевірте кількість введених даних!"); exit; }
for(i=0; i<n; i++)
{ x[i] = StrToFloat(Memo1->Lines->Strings[i]); // Введення масиву X
y[i] = StrToFloat(Memo2->Lines->Strings[i]); // Введення масиву Y
Series1->AddXY(x[i], y[i], "", clRed);
}
double xbeg, xend, h, xt, yt, yp;
if(Edit1->Text == "" )
{ ShowMessage("Введіть кількість контрольних точок!"); exit; }
int kt = StrToInt(Edit1->Text);
```

```
xbeg = x[0];
xend = x[n-1];
h = (double) (xend - xbeg) / (Pt - 1);
if(xbeg>xend || h<=0) { ShowMessage("Неправильно введені дані"); exit; }
xt = xbeg;
do
{ yp = lagrang(x, y, n, xt);
Memo3->Lines->Add("x=" + FloatToStrF(xt, ffFixed, 5, 1) +
+" yp=" + FloatToStrF(yp, ffFixed, 6, 3));
Series2->AddXY(xt, yp, "", clGreen);
xt += h;
} while (xt <= xend);
}
//-------------------- Кнопка "Очистити" --------------------------
void __fastcall TForm1::Button2Click(TObject *Sender)
{
Edit1->Clear(); Series1->Clear(); Series2->Clear();
Memo1->Clear(); Memo2->Clear(); Memo3->Clear();
}
//--------------------- Кнопка "Вихід"-----------------------------------
void __fastcall TForm1::Button3Click(TObject *Sender)
{
Close();
}
```

Форма проекту з результатами:



Інтерполяційний поліном Н'ютона. Розглянемо рівновіддалені вузли, тобто. $x_{i+1} - x_i = h$. $i = 0, 1, \ldots, n-1$. Поліном Лагранжа подамо у вигляді:

$$L_n(x) = L_0(x) + [L_1(x) - L_0(x)] + [L_2(x) - L_1(x)] + \ldots [L_n(x) - L_{n-1}(x)]. \qquad (5.12)$$

Введемо позначення:

$$\begin{cases} \Delta y_0 = y_1 - y_0; \\ \Delta y_1 = y_2 - y_1; \\ \ldots\ldots\ldots\ldots \\ \Delta y_{n-1} = y_n - y_{n-1}. \end{cases} \qquad (5.13)$$

Величини, що визначаються формулами (5.12) будмо називати ***кінцевими різницями першого порядку.*** Аналогічно введемо:

$$\begin{cases} \Delta^2 y_0 = \Delta y_1 - \Delta y_0; \\ \Delta^2 y_1 = \Delta y_2 - \Delta y_1; \\ \ldots\ldots\ldots\ldots\ldots \\ \Delta^2 y_{n-1} = \Delta y_n - \Delta y_{n-1}. \end{cases} \tag{5.14}$$

$$\Delta^k y_i = \Delta^{k-1} y_{i+1} - \Delta^{k-1} y_i. \tag{5.15}$$

Величини, що визначаються формулами (5.14) і (5.15) відповідно будуть називатися ***кінцевими різницями другого і k-го порядків.*** Кінцеві різниці характеризують швидкість зміни функції на різних інтервалах однакової довжини. Враховуючи формули (5.10)–(5.11) і (5.13)–(5.15), отримаємо:

$$L_1(x) - L_0(x) = \frac{x - x_1}{x_0 - x_1} y_0 + \frac{x - x_0}{x_1 - x_0} y_1 - y_0 = \frac{y_0 x - y_0 x_1}{-h} +$$

$$+ \frac{y_1 x - y_1 x_0}{h} - \frac{y_0 (x_1 - x_0)}{h} = \frac{y_0 x_1 - y_0 x + y_1 x - y_1 x_0 - y_0 x_1 + y_0 x_0}{h} =$$

$$= \frac{y_1 (x - x_0) - y_0 (x - x_0)}{h} = \frac{(y_1 - y_0)(x - x_0)}{h} = \frac{\Delta y_0}{h} (x - x_0);$$

$$L_2(x) - L_1(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} y_0 + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} y_1 +$$

$$+ \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} y_2 - \frac{x - x_1}{x_0 - x_1} y_0 - \frac{x - x_0}{x_1 - x_0} y_1 =$$

$$= \frac{y_0 x^2 - y_0 x x_1 - y_0 x x_2 + y_0 x_1 x_2}{(-h) \cdot (-2h)} + \frac{y_1 x^2 - y_1 x x_0 - y_1 x x_2 + y_1 x_0 x_2}{h \cdot (-h)} +$$

$$+ \frac{y_2 x^2 - y_2 x x_0 - y_2 x x_1 - y_2 x_0 x_1}{2h \cdot h} - \frac{y_0 x + y_0 x_1}{-h} - \frac{y_1 x + y_1 x_0}{h} =$$

$$= \frac{1}{2h^2} (y_0 x^2 - y_0 x x_1 - y_0 x x_2 + y_0 x_1 x_2 - 2 y_1 x^2 + 2 y_1 x x_0 +$$

$$+ 2 y_1 x x_2 - 2 y_1 x_0 x_2 + y_2 x^2 - y_2 x x_0 - y_2 x x_1 - y_2 x_0 x_1 +$$

$$+ (x_2 - x_0)(y_0 x + y_0 x_1) - (x_2 - x_0)(y_1 x + y_1 x_0)) =$$

$$= \frac{1}{2h^2} (y_2 - 2 y_1 + y_0)(x^2 - x x_0 - x x_1 + x_1 x_0) = \frac{\Delta^2 y_0}{2! h^2} (x - x_0)(x - x_1).$$

Аналогічно можна довести, що $L_n(x) - L_{n-1}(x) = \frac{\Delta^n y_0}{n! h^n} (x - x_0)(x - x_1) \ldots (x - x_{n-1})$.

Підставляючи знайдені вирази у формулу (5.12) в підсумку отримаємо:

$$L_n(x) = N_n(x) = y_0 + \frac{\Delta y_0}{h}(x - x_0) + \frac{\Delta^2 y_0}{2!h^2}(x - x_0)(x - x_1) + \\ + \ldots + \frac{\Delta^n y_0}{n!h^n}(x - x_0)(x - x_1)\ldots(x - x_n). \qquad (5.16)$$

Поліном, наведений формулою (5.16) називається **інтерполяційним поліномом Н'ютона.** Зазвичай цю формулу записують в іншому вигляді, ввівши нову змінну $t = \frac{x - x_0}{h}$. Тоді деякі величини в (5.16) набудуть вигляду:

$$\frac{x - x_1}{h} = \frac{x - x_0 + x_0 - x_1}{h} = \frac{x - x_0}{h} - \frac{x_1 - x_0}{h} = t - 1;$$

$$\frac{x - x_2}{h} = t - 2, \ldots, \frac{x - x_{n-1}}{h} = t - n + 1.$$

Використовуючи наведені позначення, формула (5.16) набуде вигляду:

$$N_n(x_0 + th) = y_0 + t\Delta y_0 + \frac{t(t-1)}{2!}\Delta^2 y_0 + \ldots + \frac{t(t-1)\ldots(t-n+1)}{n!}\Delta^n y_0.$$

Використовуючи формулу (5.16) в частковому вигляді, можно отримати формули для лінійної або квадратичної інтерполяції, у випадку рівновіддалених вузлів. При $n = 1$ отримаємо формулу для лінійної інтерполяції (***лінейний поліном Н'ютона)***:

$$N_1(x) = y_0 + \frac{y_1 - y_0}{h}(x - x_0). \qquad (5.17)$$

При $n = 1$ отримаємо формулу квадратичної інтерполяції (***квадратичний поліном Н'ютона***):

$$N_2(x) = y_0 + \frac{\Delta y_0}{h}(x - x_0) + \frac{\Delta^2 y_0}{2!h^2}(x - x_0)(x - x_1) = \\ = y_0 + \frac{y_1 - y_0}{h}(x - x_0) + \frac{y_2 - 2y_1 + y_0}{2!h^2}(x - x_0)(x - x_1). \qquad (5.18)$$

***Приклад* 5.3.** Використовуючи умови з попереднього прикладу 5.1 виконати наступне завдання:

1) побудувати лінійний поліном Н'ютона на відрізку [0.5; 0.6];

2) побудувати квадратичний поліном Н'ютона на відрізку [0.6; 0.8].

Розв'язок: 1. Для побудови лінійного інтерполяційного поліному Н'ютона на відрізку [0.5; 0.6] визначимо: $x_0 = 0.5$, $x_1 = 0.6$ – вузли интерполяції; $y_0 = 1.6$, $y_1 = 1.8$ – значення функції в цих вузлах; $h = x_1 - x_0 = 0.6 - 0.5 = 0.1$ – крок, підставивши вказані значення у формулу (5.17), отримаємо:

$$N_1(x) = 1.6 + \frac{1.8-1.6}{0.1}(x-0.5) = 1.6 + 2(x-0.5) = 2x + 0.6$$ – лінійний

інтерполяційний поліном Н'ютона на відрізку [0.5; 0. 6].

2. Для побудови квадратичного інтерполяційного поліному Н'ютона на відрізку [0.6; 0.8] визначимо аналогічно попередньому прикладу 5.1 вузли і крок інтерполяції, а також значення функції в цих вузлах і підставивши ці значення у формулу (5.18), отримаємо:

$$N_2(x) = 1.8 + \frac{2.1-1.8}{0.1}(x-0.6) + \frac{2.7-2\cdot 2.1+1.8}{2\cdot 0.01}(x-0.6)(x-0.7) =$$

$$= 1.8 + 3(x-0.6) + 15(x^2 - 1.3x + 0.42) =$$

$$= 15x^2 - 16.5x + 6.3.$$

Розрахунки показують, що при рівновіддалених вузлах інтерполяційні поліноми Лагранжа і Н'ютона полністю співпадають. Але, в даному випадку використання поліномів Н'ютона буде переважати, оскільки є простіші в обчисленнях.

5.2. Інтерполяційний поліном Лагранжа

5.2.1. Аналіз загальної постановки задачі

Лагранж Жозеф Луї (25.011736-10.04.1813) – французький математик, механік і астроном. Член Берлінської АН (1787) та її президент (1766 - 1787), член Паризької АН (1787). Твори Лагранжа з математики, механіки і астрономії складають 14 томів.

При побудові інтерполяційного поліному $P_n(x)$ (5.8) взяли в якості системи функцій $\Phi_k(x) = x^k$. З визначеної точки зору більш зручною є система поліномів ступеня n, що називається *базисом Лагранжа*, $\left\{l_k^{(n)}(x)\right\}_{k=\overline{0,n}}$, визначена із наступних міркувань: кожен $l_k^{(n)}(x)$ – поліном n-ої степені, рівний нулю у всіх вузлах сітки $\overline{\omega}$ крім k-го, де він рівний 1

$$l_k^{(n)}(x) = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases} \quad i,k = \overline{0,n}.$$

Не складно побудувати ці поліноми. Дійсно, знаючи корені поліному, можемо стверджувати, що поліном розв'язує поставлену задачу.

$$l_k(x) = l_k^{(n)}(x) = \frac{(x-x_0)(x-x_1)...(x-x_{k-1})(x-x_{k+1})...(x-x_n)}{(x_k-x_0)(x_k-x_1)...(x_k-x_{k-1})(x_k-x_{k+1})...(x_k-x_n)} = \prod_{\substack{i=0 \\ i\neq k}}^{n} \frac{(x-x_i)}{(x_k-x_i)}.$$

Перетворимо базис $\left\{l_k^{(n)}(x)\right\}$. Введемо в розгляд поліном $(n+1)$-го степеня

$$w(x) \equiv w_{\overline{0,n}}(x) = (x-x_0)...(x-x_n) = \prod_{i=0}^{n}(x-x_i).$$

Знайдемо його похідну в точці $x = x_k$, отримаємо

$$w'(x)|_{x=x_k} = \left| \begin{matrix} (n+1), \\ (x-x_k) \Rightarrow x = x_k \Rightarrow (x-x_k) = 0 \end{matrix} \right| = (x_k - x_0)(x_k - x_1)...(x_k - x_n).$$

Тоді

$$l_k(x) = \frac{w(x)}{(x - x_k)w'(x_k)}$$

і є базис поліномів Лагранжа.

Зауважимо, що побудований базис єдиний. Дійсно, якщо існує поліном $\bar{l}_k(x)$ при тих же умовах, то поліном $q^{(n)}(x) = l_k(x) - \bar{l}_k(x)$ є поліном n-го ступеня, що перетворюється в нуль в $(n+1)$-ій точці $x_0,...,x_n$. Він тотожно рівний нулю і, відповідно, $\bar{l}_k(x) = l_k(x)$. Тепер легко записати розв'язок задачі поліноміальної інтерполяції. Поліном $y_k \cdot l_k(x)$ приймає у вузлі x_k значення y_k та рівний нулю у всіх інших вузлах сітки $\overline{\omega}$ (при $x_i \neq x_k$). Тоді

$$L_n(x) \equiv \sum_{k=0}^{n} l_k(x) y_k = \sum_{k=0}^{n} f(x_k) \frac{w(x)}{(x - x_k)w'(x_k)} \tag{5.19}$$

представляє собою поліном степеня не вищого за n та $L_n(x_i) = y_i$, тобто є інтерполяційним поліномом.

Формулу (5.19) називають інтерполяційною формулою Лагранжа, а відповідний поліном $L_n(x)$ – інтерполяційния поліномом Лагранжа.

Зауваження. Нескладно оцінити кількість арифметичних дій при розрахунку за формулою (5.19). В головному порядку за n це є величина $O(n^2)$.

5.2.2. Інтерполяційна формула Лагранжа. Частковий випадок

Найбільш загальною формулою параболічної інтерполяції є інтерполяційна формула Лагранжа. Розглянемо спочатку окремий *випадок задачі параболічної інтерполяції*: побудуємо многочлен $p_i(x)$ такий, що $p_i(x_j)=0$ при $j \neq i$ та при $p_i(x_i)=1$ $(i, j=0, 1, 2, ..., n)$. Оскільки шуканий многочлен $p_i(x)$ повинен перетворитися на нуль у n точках $x_0, x_1, ..., x_{i-1}, x_{i+1}, ..., x_n$, то його можна представити у вигляді

$$p_i(x)=C_i (x-x_0) (x-x_1) ... (x-x_{i-1}) (x-x_{i+1}) ... (x-x_n), \tag{5.20}$$

де C_i – постійний коефіцієнт.

Покладемо $x=x_i$ у формулі (5.20) і з огляду на те, що $p_i(x_i)=1$, одержимо

$$C_i (x_i-x_0) (x_i-x_1) ... (x_i-x_{i-1}) (x_i-x_{i+1}) ... (x_i-x_n)=1. \tag{5.21}$$

Звідси

$$C_i = \frac{1}{(x_i - x_0)(x_i - x_1)...(x_i - x_{i-1})(x_i - x_{i+1})...(x_i - x_n)}. \tag{5.22}$$

Після підстановки виразу (5.22) у формулу (5.20) многочлен $p_i(x)$ матиме наступний вигляд

$$p_i(x)=\frac{(x-x_0)(x-x_1)...(x-x_{i-1})(x-x_{i+1})...(x-x_n)}{(x_i-x_0)(x_i-x_1)...(x_i-x_{i-1})(x_i-x_{i+1})...(x_i-x_n)}. \qquad (5.23)$$

Тепер перейдемо до знаходження многочлена $p_n(x)$, що задовольняє умовам

$$P_n(x)=f(x_i)=y_i \ (i=0,\ 1,\ 2,\ ...,\ n), \qquad (5.24)$$

тобто многочлен набуває в заданих точках $x=x_i\ (i=0,\ 1,\ 2,\ ...,\ n)$ задані значення y_i.

Легко перевірити, що такий многочлен буде мати наступний вигляд

$$P_n(x)=\sum_{i=0}^{n} p_i(x)y_i. \qquad (5.25)$$

Дійсно, при фіксованому $j\ (j=0,\ 1,\ 2,\ ...,\ n)$ маємо

$$P_n(x_j)=\sum_{i=0}^{n} p_i(x_j)y_i = p_j(x_j)y_j = y_j, \qquad (5.26)$$

отже виконуються умови (5.24). Степінь же многочлена $P_n(x)$ не вище n.

Підставивши у формулу (5.25) вираз (5.23), одержимо

$$P_n(x_j)=\sum_{i=0}^{n}\frac{(x-x_0)(x-x_1)...(x-x_{i-1})(x-x_{i+1})...(x-x_n)}{(x_i-x_0)(x_i-x_1)...(x_i-x_{i-1})(x_i-x_{i+1})...(x_i-x_n)}y_i, \qquad (5.27)$$

Вираз (5.27) є інтерполяційною формулою Лагранжа.

Запишемо її в розгорнутому вигляді:

$$P_n(x)=y_0\frac{(x-x_1)(x-x_2)(x-x_3)...(x-x_n)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)...(x_0-x_n)}+y_1\frac{(x-x_0)(x-x_2)(x-x_3)...(x-x_n)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)...(x_1-x_n)}+...$$
$$...+y_i\frac{(x-x_0)(x-x_1)...(x-x_{i-1})(x-x_{i+1})...(x-x_n)}{(x_i-x_0)(x_i-x_1)...(x_i-x_{i-1})(x_i-x_{i+1})...(x_i-x_n)}+...$$
$$...+y_n\frac{(x-x_0)(x-x_1)(x-x_2)...(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)(x_n-x_2)...(x_n-x_{n-1})}.$$

При $n=1$ формула Лагранжа приймає вигляд $P_1(x)=\frac{x-x_1}{x_0-x_1}y_0+\frac{x-x_0}{x_1-x_0}y_1$ і називається формулою *лінійної інтерполяції.*

При $n=2$ одержимо формулу *квадратичної інтерполяції:*

$$P_2(x)=\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}y_0+\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}y_1+\frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}y_2. \qquad (5.28)$$

Інтерполяційний многочлен Лагранжа можна побудувати при будь-якому розташуванні вузлів інтерполяції.

Перевірка результатів. Для того щоб перевірити результат обчислення, необхідно знайти похибку обчислень, тобто залишковий член.

Оскільки нам не відомий аналітичний вигляд функції, є тільки таблиця значень, то оцінити похибку для інтерполяційного многочлена Лагранжа буде важко. Для цього необхідно знайти похідні вищих порядків за допомогою інтерполяційних формул Ньютона. Тому, обмежимося знаходженням залишкового члена для інтерполяційного многочлена Ньютона за формулою

$$R_n(x) \approx \frac{q(q-1)...(q-n)}{(n+1)!}\Delta^{n+1}y_0. \tag{5.29}$$

Формула (5.29) отримана з формули $R_n(x) = h^{n+1}\frac{q(q-1)(q-2)...(q-n)}{(n+1)!}f^{(n+1)}(\xi)$, де ξ - деяке значення між вузлами інтерполяції $x_0, x_1, \ldots, x_n$ і розглянутою точкою x, з огляду на те, що $f^{(n+1)}(x) = \lim\limits_{h\to 0}\frac{\Delta^{n+1}y}{h^{n+1}}$, приблизно можна припустити $f^{(n+1)}(\xi) \approx \frac{\Delta^{n+1}y_0}{h^{n+1}}$.

5.2.3. Побудова інтерполяційного многочлена Лагранжа

Теорема 5.1. *Нехай* $\{\varphi_j(x) \equiv x^j, j = 0,1,...,n\}$ *тоді існує єдиний інтерполяційний многочлен* $\varphi(x)$ *для функції* $f \in \Im([a,b])$ *за вузлами інтерполяції* $\{x_0, x_1, ..., x_n\} \subset [a,b]$ *і значенням функції* $\{f_j \equiv f(x_j), x_i \neq x_j, i \neq j, j = 0,1,2,...,n\}$, *який записується у вигляді*

$$\varphi(x) = \sum_{i=0}^{n} a_i x^i \; ; \; a_i = \frac{\Delta_i}{\Delta}. \tag{5.30}$$

Доведення. Шукатимемо інтерполяційний узагальнений многочлен (5.4), що задовольнятиме умову інтерполювання (5.5) :

$$\sum_{i=0}^{n} a_i x_j^i = f_j, j = 0,1,...,n. \tag{5.31}$$

Визначник системи (5.31) для обчислення $\{a_i, i = 0,1,...,n\}$ є визначником Вандермонда:

$$\Delta = \begin{vmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{vmatrix} \neq 0. \qquad (5.32)$$

якщо $x_i \neq x_j$ з $[a,b]$; $i \neq j$; $i, j = 0,1,2,..,n$. .Тоді за формулами Крамера СЛАР (5.31) має єдиний розв'язок $a_i = \frac{\Delta_i}{\Delta}, i = 0,1,...,n$, що з урахуванням вигляду узагальненого многочлена (5.4) дає (5.30).

Теорема 5.2. *Нехай:*

1) $\{f_j \equiv f(x_j)\}; i = 0,1,...,n;$

2) $\{\varphi_j(x) = x^j, j = 0,1,...,n\}$ - *система функцій Чебишова за вузлами інтерполювання* $\{x_0, x_1, ..., x_n\} \subset [a,b]$.

Тоді інтерполяційний многочлен набуває так званого, вигляду інтерполяційного многочлена Лагранжа

$$\varphi(x) = \sum_{i=0}^{n} a_i x^i \equiv L_n(x) = \sum_{i=0}^{n} f(x_i)\Phi_i(x). \qquad (5.33)$$

$$\Phi_i(x) = \prod_{j \neq i}(x - x_j) / \prod_{j \neq i}(x_i - x_j). \qquad (5.34)$$

Доведення. Інтерполювання – це побудова многочлена *n*-го степеня $\varphi(x)$ за умови $\{\varphi(x_j) = f_j, j = 0,1,...,n\}$.Позначимо через

$$\delta_i^j \equiv \begin{cases} 1, \text{ якщо } i = j \\ 0, \text{ якщо } i \neq j \end{cases}. \qquad (5.35)$$

символ Кронекера. Кронекер Леопольд (07.17.1823 – 29.12.1891) – німецький математик. Член Берлінскої АН, член – кореспондент Петербурзької АН (1872).

Твердження доведемо, якщо побудуємо $\Phi_i(x)$ у формулі (5.33). Для цього візьмемо многочлени $\Phi_i(x)$ степеня не вище *n* за умови

$$\Phi_i(x_j) = \delta_i^j; i, j = 0,1,...,n. \qquad (5.36)$$

Многочлени

$$L_n(x) = \sum_{i=0}^{n} f(x_i) * \Phi_i(x). \qquad (5.37)$$

є шуканими. Дійсно він має степінь не вище n, оскільки він є лінійною комбінацією многочленів степеня не вище n. Залишається перевірити умову інтерполювання $L_n(x_j) = f(x_j)$. Дійсно,

$$L_n(x_j) = \sum_{i=0}^{n} f(x_i) * \Phi_i(x_j) = \sum_{i=0}^{n} f(x_i)\delta_i^j = f(x_0)\delta_0^j + f(x_1)\delta_1^j + \ldots + \\ + f(x_j)\delta_j^j + f(x_{j+1})\delta_{j+1}^j + f(x_n)\delta_n^j = f(x_j) \tag{5.38}$$

Отже, $L_n(x) = \sum_{i=0}^{n} f(x_i) * \Phi_i(x)$ є інтерполяційним многочленом Лагранжа за означенням. Тобто $\varphi(x) = \sum_{i=0}^{n} a_i \varphi_i(x)$ є многочленом n-го степеня та $\varphi(x_j) \equiv L_n(x_j) = f_j, j = 0,1,\ldots,n$. Побудуємо $\Phi_i(x)$, а значить і $L_n(x)$. $\Phi_i(x_j) = 0, i \neq j$ тоді $\{x_j\}$ - корені $\{\Phi_i(x)\}$. За теоремою Вієта

$$\Phi_i(x) = K * \prod_{j \neq i} (x - x_j). \tag{5.39}$$

Вієт Франсуа (1540 – 13.12.1603) – французький математик і астроном, батько алгебри.

Для визначення K обчислимо (5.39) у вузлі x_i: $\Phi_i(x_i) = K \times \prod_{j \neq i} (x_i - x_j) = 1$, бо $\Phi_i(x_i) = \delta_i^j = 1$. Тоді $K = \dfrac{1}{\prod_{j \neq i} (x_i - x_j)}$, а значить формула (5.39) прийме вигляд

(5.34) $$\Phi_i(x) = \frac{\prod_{i \neq j} (x - x_j)}{\prod_{i \neq j} (x_i - x_j)}.$$

Наслідок 5.1. *Інтерполяційний многочлен Лагранжа можна записати у вигляді*

$$L_n(n) = \sum_{i=0}^{n} f(x_i) \frac{\omega_n(x)}{(x - x_i)\omega_n'(x_i)}, \tag{5.40}$$

де $\omega_n(x) \equiv \prod_{j=0}^{n} (x - x_j) \equiv (x - x_0)(x - x_1) \cdots (x - x_n)$.

Це випливає з вигляду $\{\Phi_i(x)\}$ у формулі (5.33).

5.2.4. Інтерполяційний многочлен Лагранжа з рівновіддаленими вузлами

Теорема 5.2 стверджує, що інтерполяційний многочлен Лагранжа для $f \in \Im([a,b])$ з вузлами $\{x_i, i=0,1,...,n\} \subset [a,b]$, який побудовано за системою функцій $\{x^i, i=0,1,...,n\}$ має вигляд

$$\varphi(x) \equiv L_n(x) = \sum_{i=0}^{n} f(x_i) * \Phi_i(x). \tag{5.41}$$

$$\Phi_i(x) = \frac{(x-x_0)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n)}{(x_i-x_0)\cdots(x_i-x_{i-1})(x_i-x_{i+1})\cdots(x_i-x_n)}. \tag{5.42}$$

Теорема 5.3. *Нехай :*

1) Значення $f \in \Im([a,b])$ *задано у рівновіддалених вузлах*

$$x_i = x_0 + ih,\ i=0,1,...,n;\ x_i \in [a,b],\ h>0;\ f_i \equiv f(x_i); \tag{5.43}$$

2) $\{\varphi_i(x) \equiv x^i, i=0,1,...,n\}$ *- система функцій Чебишова.*

Тоді

$$L_n(x) = L_n(x_0+th) = \sum_{i=0}^{n} \Lambda(t) \cdot f(x_i), \tag{5.44}$$

$$\Lambda(t) \equiv (-1)^{n+i} C_n^i \frac{\omega_n(t)}{(t-i)n!};\ \omega_n(t) = t(t-1)\cdots(t-n). \tag{5.45}$$

Доведення. Для $\Phi_i(x)$, $i=0,1,...,n$ у (5.42) зробимо заміну

$$\frac{x-x_0}{h} = t,\ x \in [a,b]. \tag{5.46}$$

Одержимо кожний із співмножників у чисельнику для $\{\Phi_i(x)\}$ згідно заміни (5.46):

$$\left.\begin{array}{l} x-x_0 = th; \\ x-x_1 = (x-x_0)+(x_0-x_1) = th-h = h(t-1); \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x-x_i = (x-x_0)+(x_0-x_i) = th-ih = h(t-i); \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x-x_n = h(t-n). \end{array}\right\}. \tag{5.47}$$

Далі з (5.46) одержимо кожний із співмножників знаменника для $\Phi_i(x)$

$$\left.\begin{array}{l} x_i = x_0 = ih; \\ x_i = x_1 = (x_i - x_0) + (x_0 - x_1) = ih - h = (i-1)h; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_i - x_{i-1} = h; \ldots\ldots; x_i - x_{i+1} = -h; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_i - x_n = (i-n)h. \end{array}\right\}. \qquad (5.48)$$

Отже, з урахуванням (5.47), (5.48) кожен з $\Phi_i(x)$ матиме вигляд

$$\Phi_i(x) = \frac{th(t-1)h\cdots(t-i+1)h(t-i-1)h\cdots(t-n)h(t-i)}{ih(i-1)h\cdots h(-1)h\cdots(i-n)h(t-i)} = \\ = \frac{t(t-1)\cdots(t-i)\cdots(t-n)n!}{t!(-1)^{n-i}(n-i)!(t-i)n!} = (-1)^{n-i}C_n^i\frac{t(t-1)\cdots(t-n)}{(t-i)n!} = \Lambda_i(t) \qquad (5.49)$$

оскільки

$$\frac{1}{(-1)^{n-i}} = (-1)^{n-i},\ C_n^i = \frac{n!}{(n-i)!i!}, \qquad (5.50)$$

що і доводить теорему 5.3.

Зауваження 5.1. Для різних t та n можна одержати за допомогою комп'ютера коефіцієнти $\Lambda_i(t)$, які будуть обслуговувати у випадку рівновіддалених вузлів обчислення інтерполяційних многочленів $L_n(x) = L_n(x_0 + th)$, для $\forall f \in \Im([a,b])$. Ці таблиці наводяться у довідковій літературі методів обчислень [76-80].

5.2.5. Інтерполяційна схема О. Ейткена

Ейткен Олександр Крег (01.04.1895 – 03.11.1967) – англійський математик і фізик. Недоліком ІМЛ $L_n(x)$:

- Велика обчислювальна робота;
- Для побудови $L_{n+1}(x)$ не використовується інформація , що була отримана для побудови $L_n(x)$.

Тоді виникала задача знаходження функції $f(x)$, $x \neq x_i$, не виписуючи ІМЛ $L_n(x)$.

Теорема 5.4. *Якщо* $f \in \Im([a,b])$, *то ІМП* $L_n(x)$ *дорівнює многочлену Ейткена n-го порядку:*

$$L_n(x) = L_{0,1,2,\ldots,n}(x) \equiv \frac{1}{x_n - x_0}\begin{vmatrix} L_{0,1,\ldots,n-1} & x_0 - x \\ L_{1,2,\ldots,n} & x_n - x \end{vmatrix}, \tag{5.51}$$

де інтерполяційний многочлен Ейткена першого порядку за двома вузлами $\{x_i, x_{i+1}\} \subset [a,b]$ *визначається*

$$L_{i,i+1}(x) \equiv \frac{1}{x_{i+1} - x_i}\begin{vmatrix} f_i & x_i - x \\ f_{i+1} & x_{i+1} - x \end{vmatrix}, \tag{5.52}$$

інтерполяційний многочлен Ейткена другого порядку за трьома вузлами $x_i, x_{i+1}, x_{i+2} \in [a,b]$ *визначається*

$$L_{i,i+1,i+2}(x) \equiv \frac{1}{x_{i+2} - x_i}\begin{vmatrix} L_{i,i+1} & x_i - x \\ L_{i+1,i+2} & x_{i+2} - x \end{vmatrix}, \tag{5.53}$$

інтерполяційний многочлен Ейткена третього порядку за чотирма вузлами $x_i, x_{i+1}, x_{i+2}, x_{i+3} \in [a,b]$ *визначається*

$$L_{i,i+1,i+2,i+3}(x) \equiv \frac{1}{x_{i+3} - x_i}\begin{vmatrix} L_{i,i+1,i+2} & x_i - x \\ L_{i+1,i+2,i+3} & x_{i+3} - x \end{vmatrix}. \tag{5.54}$$

Доведення. Для доведення треба перевірити, що многочлени Ейткена (5.52), (5.53), (5.54) є інтерполяційними многочленами. Використаємо метод повної математичної індукції. Дійсно,

1) $L_i(x) = f_i$ - многочлен Ейткена нульового порядку за одним вузлом,

2) $L_{i,i+1}(x)$ є многочлен Ейткена першого порядку за двома вузлами x_i, x_{i+1} (див.(5.42)).

Нехай $L_{i,i+1,\ldots,i+k,i+k+1}(x)$ многочлен Ейткена k- го порядку за $k+1$ вузлами $x_i, x_{i+1}, \ldots, x_{i+k}$, тоді

$$L_{i,i+1,\ldots,i+k,i+k+1}(x) \equiv \frac{1}{x_{i+k+1} - x_i}\begin{vmatrix} L_{i,\ldots i+k}(x) & x_i - x \\ L_{i+1,\ldots,i+k+1}(x) & x_{i+k+1} - x \end{vmatrix}. \tag{5.55}$$

Тобто маємо отримати многочлен Ейткена $k+1$ порядку за $k+2$ вузлами $\{x_i, x_{i+1}, \ldots, x_{i+k+1}\} \subset [a,b]$, а значить за індукцією

$$L_n(x_j) = L_{0,1,2\ldots,n}(x_j) = f(x_j), j = 1,2,\ldots,n. \quad (5.56)$$

Дійсно для двох вузлів маємо $L_{i,i+1}(x_j) = f(x_j); j = i, i+1$:

$$L_{i,i+1}(x_i) = \frac{1}{x_{i+1} - x_i} \begin{vmatrix} f(x_i) & x_i - x_i \\ f(x_{i+1}) & x_{i+1} - x_i \end{vmatrix} = f(x_i). \quad (5.57)$$

$$L_{i,i+1}(x_i) = \frac{1}{x_{i+1} - x_i} \begin{vmatrix} f(x_i) & x_i - x_{i+1} \\ f(x_{i+1}) & x_{i+1} - x_{i+1} \end{vmatrix} = f(x_{i+1}). \quad (5.58)$$

Нехай для многочлена k- го порядку за $k+1$ вузлом виконується

$$L_{i,i+1,i+k}(x_j) = f(x_j), j = i, i+1,\ldots,i+k, \quad (5.59)$$

тоді для многочлена $k+1$ -го порядку за $k+2$ вузлами матимемо

$$L_{i,i+1,\ldots,i+k,i+k+1}(x_j) \equiv \frac{1}{x_{i+k+1}} \begin{vmatrix} L_{i,\ldots i+k} & x_i - x_j \\ L_{i+1,\ldots,i+k+1} & x_{i+k+1} - x_j \end{vmatrix} = f(x_j), \quad (5.60)$$
$$j = i, i+1,\ldots,i+k+1$$

Значить довели що многочлен Ейткена $L_{0,1,2,\ldots,n}(x)$ є інтерполяційним многочленом n-го степеня для функції $f(x) \in \Im([a,b])$ за $n+1$ вузлом , оскільки за теоремою 5.1 існує єдиний інтерполяційний многочлен Лагранжа $L_n(x)$, тобто $L_n(x) = L_{0,1,2,\ldots,n}(x)$.

Зауваження 5.2. Інтерполяційна схема Ейткена зводиться в табл. 5.2. Наприклад для $f_j \equiv f(x_j), j = 0,1,2,3,4$.

Таблиця 5.2

x_i	f_i	$x_i - x$	$L_{i,i+1}$	$L_{i,i+1.i+2}$	$L_{i,i+1,i+2,i+3}$	$L_{i,i+1,i+2,i+3,i+4}$
x_0	f_0	$x_0 - x$	–	–	–	–
x_1	f_1	$x_1 - x$	$L_{01}(x)$	–	–	–
x_3	f_2	$x_2 - x$	$L_{12}(x)$	$L_{012}(x)$	–	–
x_3	f_3	$x_3 - x$	$L_{23}(x)$	$L_{123}(x)$	$L_{0123}(x)$	–
x_4	f_4	$x_4 - x$	$L_{34}(x)$	$L_{234}(x)$	$L_{1234}(x)$	$L_{01234}(x)$

Зауваження 5.3. Обчислення за схемою Ейткена здійснюється до виконання умови $|L_{012\ldots k}(x) - L_{012\ldots k+1}(x)| < \varepsilon$, тоді $L_n(x) \approx L_{0,1,2,\ldots,k+1}(x)$, $k+1 \le n$, з точністю $\varepsilon > 0$. Схема Ейткена легко реалізується на комп'ютері.

Приклад 5.1. Записати інтерполяційний многочлен Лагранжа для функції $f(x)$, значення якої задано в табл. 5.3.

Таблиця 5.3

i	0	1	2	3
x_i	0	0,1	0,3	0,5
$f(x_i)$	–0,5	0	0,2	1

Розв'язок.

$$\Phi_0(x)=\frac{(x-0,1)(x-0,3)(x-0,5)}{-0,1*(0,3)*(0,5)}=\frac{x^3-0,9x^2+0,23x-0,015}{0,015}.$$

$\Phi_1(x)$ не обчислюємо , бо $f(x_1)=0$ і $f(x_1)*\Phi_1(x)$ не впливає на вигляд $L_3(x)$..Далі

$$\Phi_2(x)=\frac{x(x-0,1)(x-0,5)}{0,3*0,2*(-0,2)}=\frac{x^3-0,8x^2+0,05x}{0,012},$$

$$\Phi_2(x)=\frac{x(x-0,1)(x-0,5)}{0,5*0,4*0,2}=\frac{x^3-0,4x^2+0,03x}{0,04}.$$

$$L_3(x)=\Phi_0(x)f_0+\Phi_1(x)f_2+\Phi_3(x)f_3=\frac{125}{3}x^3-30x^2+\frac{73}{12}x-\frac{1}{2}.$$

Приклад 5.2. Функція $y=\sqrt[3]{x}$ задана табл. 5.4:

Таблиця 5.4

x_i	1,0	1,1	1,3	1,5	1,6
$f(x_i)$	1,000	1,032	1,091	1,145	1,17

Застосувавши схему Ейткена, знайти $\sqrt[3]{1,15}$ з точністю $\varepsilon=10^{-3}$.

Розвязок. Покладемо $x=1,15$ і побудуємо таблицю 5 для інтерполяційних многочленів Ейткена:

Таблиця 5.5

i	x_i	f_i	x_i-x	$L_{i,i+1}$	$L_{i,i+1,i+2}$
0	1,0	1,000	–0,15		
1	1,1	1,032	–0,05	1,048	
2	1,3	1,091	0,15	1,047	1,048
3	1.5	1,145	0,35	1.05	1,05
4	1,6	1,170	0,45	1,057	1,057

Бачимо, що $L_{01}(1,15)=L_{012}(1,15)=1,048$, тоді побудову інтерполяційних многочленів Ейткена слід припинити і покласти $\sqrt[3]{1,15}=1,048$ з точністю $\varepsilon=10^{-3}$.

5.2.6. Похибка інтерполювання для інтерполяційного многочлена Лагранжа

За теоремою 5.2, якщо

$f \in \Im([a,b]); \{\varphi_i(x) = \mathrm{x}^i, i = 0,1,..,n\} \subset \in \Im([a,b]); \{\mathrm{x}_i, \mathrm{x}_i \neq \mathrm{x}_j, i \neq j\} \subset [a,b]; f_i \equiv f(x_i) \triangleright$

то

$$\varphi(x) \equiv L_n(x) = \sum_{i=0}^{n} f(x)\Phi_i(x), \tag{5.61}$$

$$f(x) \neq L_n(x), x \neq x_i, \tag{5.62}$$

$$f(x_j) = L_n(x_j). \tag{5.63}$$

Якщо $f(x) = P_n(x)$, то

$$L_n(x) = P_n(x). \tag{5.64}$$

За теоремою 5.1 існує єдиний многочлен Лагранжа $L_n(x)$ степеня n.

Означення 5.4. *Похибкою методу Лагранжа назвемо різницю*

$$R_n(x) \equiv f(x) - L_n(x). \tag{5.65}$$

Зауваження 5.4. Табличні значення $\{x_i, f_i\}$ несуть неусувну похибку. Заокруглення при обчисленні несуть похибку заокруглення.

Теорема 5.5.(про похибку інтерполяційного многочлена Лагранжа). *Нехай* :

1) $f \in C^n([a,b])$;

2) $\exists f^{(n+1)}(x)$ на (a,b);

3) *задано значення функцій* $\{f_i \equiv f(x_i), i = \overline{0,n}\} \subset \mathbf{R}^1$;

4) $x_i \in [a,b]$.

Тоді похибка інтерполяційного многочлена Лагранжа має таку оцінку:

$$|R_n(x)| \leq \frac{M_{n+1}}{(n+1)!}|\omega_n(x)|, \tag{5.66}$$

$$\omega_n(x) \equiv (x - x_0)(x - x_1)\cdots(x - x_n), M_{n+1} \equiv \sup_{x \in [a,b]} \left|f^{(n+1)}(x)\right|. \tag{5.67}$$

Доведення. За вузлами $\{x_j, j = 0,1,..,n\}$ будуємо інтерполяційний многочлен Лагранжа за теоремою 5.2. Розглянемо функцію дійсної змінної $z \in [a,b]$:

$$\varphi(z) = f(z) - L_n(z) - K\omega_n(z). \tag{5.68}$$

Вузли є дійсними нулями ,бо

$$\varphi(x_0)=\varphi(x_1)=...=\varphi(x_n)=0. \tag{5.69}$$

Це випливає з того ,що

1) $f(x_j)=L_n(x_j)$, $j=\overline{0,n}$,

2) $\omega(x_j)=0$, $j=\overline{0,n}$.

Підбираємо K за умови

$$\varphi(x)=0, \tag{5.70}$$

де $x\in(a,b)$. З (5.8) випливає, що

$$\varphi(x)=f(x)-L_n(x)-K(x-x_0)...(x-x_n)=0. \tag{5.71}$$

Таким чином, $\varphi(z)$ перетворюється у нуль на (a,b) у $(n+2)$ точках $\{x_0,x_1,...,x_i,x_{i+1},...,x_n\}\subset(a,b)$,тобто маємо $(n+1)$ елементарних відрізків на осі Ox.

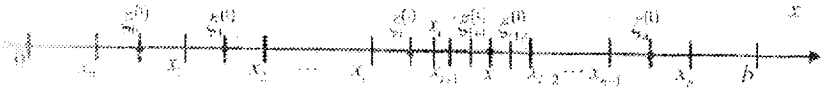


Рис. 5.2. Приклад лінійного інтерполювання для інтерполяційного многочлена Лагранжа

До кожного з $(n+1)$ відрізка застосуємо теорему Ролля , тобто якщо неперервна функція $\varphi(x)$ визначена на замкненому відрізку $[a,b]$, $\varphi\in C([a,b])$, існує $\varphi'(x)$ на (a,b) ; $\varphi(a)=\varphi(b)=0$, тоді знайдеться єдина точка $x=\xi\in[a,b]$ така, що $\varphi'(\xi)=0$. Значить, якщо застосувати це твердження на кожному з елементарних відрізків , матимемо

$$\varphi'(\xi_0^{(1)})=\varphi'(\xi_1^{(1)})=...=\varphi'(\xi_n^{(1)})=0. \tag{5.72}$$

Умова (5.72) та існування похідних другого порядку на інтервалах $(\xi_0^{(1)}),(\xi_1^{(1)})$, $i=\overline{0,n+1}$, дає можливість застосувати теорему Ролля для перших похідних. Тоді

$$\varphi^n(\xi_0^{(2)})=\varphi^n(\xi_1^{(2)})=...=\varphi^n(\xi_n^{(2)})=0. \tag{5.73}$$

Знову слід застосувати для n інтервалів теорему Ролля. Процес повторюємо $(n+1)$ раз. Тоді існує єдина точка $\xi\in[a,b]$, що $L_n(\xi)=0$. Далі обчислимо $(n+1)$ похідну від $\varphi(z)$

$$\varphi^{(n+1)}(z)=f^{(n+1)}(z)-0-K(n+1)!. \tag{5.74}$$

Підставимо замість z точку ξ в рівність (5.74),тоді

$$0=\varphi^{(n+1)}(\xi)=f^{(n+1)}(\xi)-K(n+1)!. \tag{5.75}$$

Звідки

$$K = \frac{f^{(n+1)}(\xi)}{(n+1)!}, \xi \in (a,b). \tag{5.76}$$

Відмітимо , що з (5.76) випливає оцінка (5.67), бо за умови (5.70) і (5.71) маємо

$$0 = \varphi(x) = f(x) - L_n(x) - \frac{f^{(n+1)}(\xi)}{(n+1)!}\omega_n(x), \tag{5.77}$$

а значить одержуємо формулу (5.66) для $R_n(x)$ та

$$|f(x) - L_n(x)| \le \frac{M_{n+1}}{(n+1)!}|\omega_n(x)|. \tag{5.78}$$

5.3. Вибір вузлів інтерполювання

Вище отримали, що

$$f(x) = L_n(x) + R_n(x). \tag{5.79}$$

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\omega_n(x). \tag{5.80}$$

$$\omega_n(x) = (x - x_0)(x - x_1)...(x - x_n). \tag{5.81}$$

Виникає природна задача вибрати вузли $\{x_j, j = \overline{0,n}\} \subset [a,b]$ так, щоби

$$\sup_{x\in[a,b]} |\omega_n(x)| \to \min. \tag{5.82}$$

Доведемо, що

$$\inf_{P} \sup_{x\in[a,b]} |\omega_n(x)|. \tag{5.83}$$

досягається на коренях многочлена Чебишева.

5.3.1. Многочлени Чебишева. Властивості

Означення 5.5. *Вираз*

$$T_n(x) \equiv \cos[n \arccos x], |x| \le 1. \tag{5.84}$$

запишемо многочленом Чебишова n-го степеня .

Лема 5.1. *Многочлен Чебишева* (5.84) *володіє трьома властивостями:*

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \tag{5.85}$$

$$T_n(x) = 2^{n-1}x^n + A_1x^{n-1} + \ldots + A_n. \tag{5.86}$$

$$x_m = \cos\frac{(2m+1)\pi}{2n},\ m = \overline{1,n-1}. \tag{5.87}$$

є коренями $T_n(x)$.

Доведення. 1) $\forall n \in \mathbf{R}$, $\forall \theta \in \mathbf{R}^1$ має місце тригонометрична рівність

$$\cos(n+1)\theta + \cos(n-1)\theta = 2\cos\theta\cos n\theta. \tag{5.88}$$

Покладемо $\theta = \arccos x$ та підставимо у (5.88):

$$\cos[(n+1)\arccos x] + \cos[(n-1)\arccos x] = 2\cos[\arccos x]\cos[n\arccos x],$$

звідки за означенням 5.5 матимемо (5.85).

2) Методом повної математичної індукції доводимо (5.86).

Очевидні співвідношення: для $n=0$ з (5.74) маємо $T_0(x) = 1$; для $n=1$ з (5.74) випливає $T_1(x) = x$. Якщо $n=2$,то (5.85) з врахуванням того, що $T_0(x) = 1$ і $T_1(x) = x$, дає $T_1(x) = 2x \cdot x - 1 = 2x^2 - 1$, що співпадає з (5.86) при $n=2$.

Нехай має місце для $n=k$ формула (5.86): $T_k(x) = 2^{k-1}x^k + A_1x^{k-1} + \ldots + A_k$. Тоді запишемо (5.85) для $n=k+1$: $T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x) = Ax^{k+1} + A_1x^k + \ldots + A_{k+1}$, тобто для $\forall k \le n+1$ многочлен Чебишова (5.84) є многочленом k-го степеня.

3) $T_n(x) = 2^{n-1}x^n + A_1x^{n-1} + \ldots + A_n$ є многочленом n-го степеня. Тоді для рівняння $\cos[n\arccos x] = 0$ маємо n коренів. Дійсно,

$$n\arccos x = \frac{(2m+1)\pi}{2} \in [-1;+1], m = \overline{0,n-1}. \tag{5.89}$$

а значить

$$-1 \le x_m = \cos\frac{(2m+1)\pi}{2n}, m = \overline{0,n-1}. \tag{5.90}$$

5.3.2. Вибір вузлів інтерполювання

Теорема 6 (про вибір вузлів). *Нехай:*

1) $f \in C^n([-1,1])$; $f^{n+1} \in C([-1,1])$;

2) $f_i \equiv f(x_i)$, $\{x_i, i = \overline{0,n}\} \subset [-1;+1]$;

3) $\{\varphi_i(x) \equiv x^i, i = \overline{0,n}\}$ - *система функцій Чебишева.*

Тоді

1) $$|f(x) - L_n(x)| \equiv |R_n(x)| \le \frac{M_{n+1}}{(n+1)!}\frac{1}{2^n}, \tag{5.91}$$

де M_{n+1} *визначене як співвідношення* (5.67);

2) $\inf\limits_{P}\sup\limits_{x\in[-1,1]}|\omega_n(x)|$ *досягається, якщо за вузли* $\{x_i\}\subset[-1,+1]$ *вибрати корені многочлена Чебишова* (5.66):

$$x_i=\cos\frac{(2i+1)\pi}{2n},\ i=\overline{0,n-1}. \tag{5.92}$$

Доведення.1) Нехай для побудови $L_n(x)$ вибрали (5.92). Тоді за теоремою Вієта $T_{n+1}(x)=2^nx^{n+1}+B_1x^n+\ldots+B_{n+1}=2^n\omega_n(x)$. Звідки

$$\omega_n(x)=\frac{1}{2^n}T_{n+1}(x). \tag{5.93}$$

Але $\sup\limits_{x\in[-1,1]}|T_{n+1}(x)|=1$. Значить

$$\sup_{x\in[-1,1]}|\omega_n(x)|=\frac{1}{2^n}. \tag{5.94}$$

Мали $|R_n(x)|\le\frac{M_{n+1}}{(n+1)!}\omega_n(x)$, тоді з (5.94) випливає (5.91).

Доведення пункту 2) базується на Лемі 5.2.

Лема 5.2. *Для многочлена степеня n з старшим коефіцієнтом 1 має місце оцінка*

$$\sup_{x\in[-1,1]}|P_n(x)|\ge\frac{1}{2^{n-1}}. \tag{5.95}$$

Доведення від супротивного. Нехай

$$\sup_{x\in[-1,1]}|P_n(x)|=\sup_{x\in[-1,1]}|x^n+A_1x^{n-1}+\ldots+A_n|<\frac{1}{2^{n-1}}. \tag{5.96}$$

Вибираємо за вузли інтерполювання для $P_n(x)$ корені $\sup\limits_{x\in[-1,1]}|T_{n+1}(x)|=\sup\limits_{x\in[-1,1]}\cos[n\arccos x]=\pm1$,тобто

$$x_m=\cos\frac{m\pi}{n},\ m=\overline{0,n}. \tag{5.97}$$

Розглянемо різницю

$$\frac{1}{2^{n-1}}T_n(x)-P_n(x). \tag{5.98}$$

Це многочлен степеня не більше (n-1), тобто він має рівно (n-1) коренів. Але многочлен $P_n(x)$ приймає в (n+1) точках (5.97) значення +1,...,–1 ,а старший коефіцієнт для $P_n(x)$ дорівнює 1.

Значить, з другого боку, для (5.98) маємо рівно n коренів. А маємо в дійсності (n-1) корінь. Припущення не вірне.

З леми 5.2 випливає той факт, що $\inf\limits_{P(x)=T_{n+1}(x)} \sup\limits_{x\in[a,b]} |\omega_n(x)|$ набуває значення 0 на коренях $x_m = \cos\dfrac{(2m+1)\pi}{2n}$, $m = \overline{0,n-1}$.

Зауваження 5.5. Оцінку для многочлена Лагранжа можна одержати для відрізка [-1;+1] заміною $x = \dfrac{1}{2}[(b-a)z+(b+a)]$. Дійсно, якщо $x\in[a,b]$, тоді $z = \dfrac{1}{b-a}[2x-b-a]$, для якого $z\in$ [-1;+1]. Для $T_{n+1}(z)$ має місце теорема 5.6. Корені $z_m = \cos\dfrac{(2m+1)\pi}{2n}$, $m = \overline{0,n}$ перейдуть у

$$x_m = \frac{1}{2}[(b-a)\cos\frac{(2m+1)\pi}{2n} + b + a], \tag{5.99}$$

а для похибки методу інтерполювання виконується оцінка

$$|R_n(x)| \le \frac{M_{n+1}}{(n+1)!}\frac{(b-a)^{n+1}}{2^{n-1}}. \tag{5.100}$$

Величина похибки $R_n(x)$ для інтерполяційного многочлена змінюється на відрізку інтерполювання $[a,b]$.

5.3.3. Поведінка похибки на інтервалі інтерполювання

Вивчимо поведінку похибки $R_n(x)$ для ІМЛ $L_n(x)$ на випадок рівновіддалених вузлів.

Теорема 5.7. *Якщо виконуються умови побудови інтерполяційного многочлену Лагранжа $L_n(x)$ за рівновіддаленими вузлами $x_i = x_0 + ih, h>0$, то похибка $R_n(x)$ найменша у середині відрізка інтерполяції , збільшується біля кінців відрізка $[a,b]$, найбільша поза $[a,b]$ $(x\ne x_j)$.*

Доведення. Використаємо заміну $x - x_0 = th$, тоді многочлен $\omega_n(x)$ набуває вигляду

$$\omega_n(x) = \omega_n(x_0+th) = h^{n+1}t(t-1)(t-2)...(t-n) = h^{n+1}\varphi(t). \tag{5.101}$$

$$\varphi(t) \equiv t(t-1)...(t-n). \tag{5.102}$$

Лема 5.3. *Для многочлена* $\varphi(t)$ *n-го степеня має місце рівність*

$$\varphi(t) = (-1)^{n+1}\varphi(n-t), \tag{5.103}$$

тобто $\varphi(t)$ *-парна або непарна відносно точки* $(n/2, 0)$ *у залежності від парності* n. Доведення випливає з означення (5.102) многочлена $\varphi(t)$:

$$\varphi(n-t) = (n-t)(n-t-1)...(n-t-n) = (-1)^{n+1}\varphi(t).$$

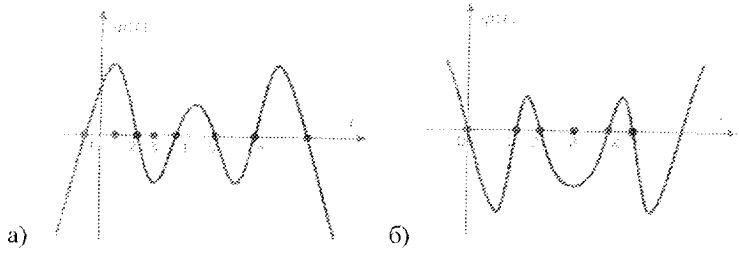
Лема 5.4. *Якщо* $[0,n] = \bigcup_{i=0}^{n-1}[i, i+1]$, *тоді*

$$\varphi(t+1) = \frac{t+1}{t-1}\varphi(t). \tag{5.104}$$

Доведення. Запишемо многочлен $\varphi(t)$ за означенням (5.102)

$$\varphi(t+1) = (t+1)t(t-1)...(t+1-n)\frac{t-n}{t-n} = \frac{t+1}{t-n}\varphi(t).$$

Доведення теореми випливає із леми 5.4 . Екстремальні значення $\varphi(t)$ зменшуються при переході від попереднього до наступного інтервалу до середини відрізку $[0, n]$ за рахунок того ,що $\left|\frac{t+1}{t-n}\right| < 1$ бо $i = \underline{1,n}$ є коренями $\varphi(t)(\varphi(i) = 0, i = \overline{0,n})$. Далі за лемою 5.3 $\varphi_n(t)$ симетрично зростає, бо $\left|\frac{t+1}{t-n}\right| > 1$, $\forall t = n+1,...$. Поведінку $\varphi(t)$ зображено на рис. 5.3. Теорема 5.7 доведена.



а) б)

Рис. 5.3. Поведінка $\varphi(t)$ для а) *n=6 та* б) *n=5*

5.4. Інтерполяційний поліном Ньютона

5.4.1. Аналіз загальної постановки задачі

Зручним представленням інтерполяційного поліному для практичних розрахунків (особливо вручну) є запис інтерполяційного поліному (того ж самого) у вигляді інтерполяційного полінома Ньютона. Для цього введемо в розгляд так звані

роздільні різниці сіткової функції $\{f_i\}$. Визначимо їх рекурентно. Розділені різниці першого порядку, побудовані на вузлах x_i, x_j визначають наступним чином

$$f(x_i, x_j) = \frac{f(x_i) - f(x_j)}{x_i - x_j} = \left| \begin{matrix} вони \\ симетричні \end{matrix} \right| = \frac{f(x_j) - f(x_i)}{x_j - x_i} = f(x_j, x_i).$$

Розділена різниця другого порядку на вузлах x_i, x_j, x_k визначається як перша різниця від попередніх розділених різниць

$$f(x_i, x_j, x_k) = \frac{f(x_i, x_j) - f(x_j, x_k)}{x_i - x_k} = \frac{f(x_j, x_k) - f(x_i, x_j)}{x_k - x_i}; \text{ і т.д.}$$

Якщо відомі розділені різниці k-го порядку, то розділені різниці $(k+1)$-го порядку на вузлах $x_j, x_{j+1}, \ldots, x_{j+k+1}$ визначаються як

$$f(x_j, x_{j+1}, \ldots, x_{j+k+1}) = \frac{f(x_j, \ldots, x_{j+k}) - f(x_{j+1}, \ldots, x_{j+k+1})}{x_j - x_{j+k+1}} =$$

$$= \frac{f(x_{j+1}, \ldots, x_{j+k-1}) - f(x_j, \ldots, x_{j+k})}{x_{j+k+1} - x_j}.$$

Зауважимо, що при розгляді введених понять сусідство вузлів сітки не обов'язкове, важлива їх кількість: для розділеної різниці k-го порядку – $(k+1)$ вузол. Далі будемо розглядати розділені різниці, які складені лише за сусідніми вузлами. Вихідна таблиця значень функції $\{f(x_i)\}$ дозволяє побудувати наступні розділені різниці вказаного типу:

Таблиця 5.4

різниці	0-го	1-го	2-го	…	$(n-1)$-го	n-го
x_0	$f(x_0)$			…		
		$f(x_0, x_1)$		…		
x_1	$f(x_1)$		$f(x_0, x_1, x_2)$	…		
		$f(x_1, x_2)$		…		
x_2	$f(x_2)$		$f(x_1, x_2, x_3)$	…		
…	…	…	…	…		
				…	$f(x_0, \ldots, x_{n-1})$	
				…		$f(x_0, \ldots, x_n)$
				…	$f(x_1, \ldots, x_n)$	
…	…	…	…	…	…	…
x_{n-1}	$f(x_{n-1})$		$f(x_{n-2}, x_{n-1}, x_n)$	…		
		$f(x_{n-1}, x_n)$		…		
x_n	$f(x_n)$			…		
Кіл-ть	$(n+1)$	n	$(n-1)$	…	2	1

Зауважимо особливості цих розділених різниць:

1. Якщо таблична функція сама собою є поліномом n-ого степеня, тобто $f(x) \equiv P_n(x) \equiv p(x)$, то її перша розділена різниця

$$p(x, x_0) = \frac{p(x) - p(x_0)}{x - x_0}$$

є поліномом $(n-1)$-го степеня, так як в чисельнику дробу знаходиться поліном, рівний нулю при $x = x_0$, тобто ділиться націло на $(x - x_0)$.

2. Друга роділена різниця

$$p(x, x_0, x_1) = \frac{p(x, x_0) - p(x_0, x_1)}{x - x_1}$$

є поліномом $(n-2)$-го степеня, так як в чисельнику поліном $(n-1)$-го степеня, рівний нулю в точці $x = x_1$.

3. Таким чином $(n+1)$ розділена різниця для поліному n-го степеня тотожньо рівна нулю. Із визначення розділених різниць отримаємо:

$$\begin{aligned}
p(x) &= p(x_0) + (x - x_0)p(x, x_0), \\
p(x, x_0) &= p(x_0, x_1) + (x - x_1)p(x, x_0, x_1), \\
p(x, x_0, x_1) &= p(x_0, x_1, x_2) + (x - x_2)p(x, x_0, x_1, x_2), \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots, \\
p(x, x_0, x_1, x_{n-2}) &= p(x_0, x_1, x_2, \ldots, x_{n-1}) + (x - x_{n-1})p(x, x_0, x_1, x_2, \ldots, x_{n-1}), \\
p(x, x_0, x_1, x_{n-1}) &= p(x_0, x_1, x_2, \ldots, x_n) + (x - x_n)\underbrace{p(x, x_0, x_1, x_2, \ldots, x_n)}_{\equiv 0}.
\end{aligned}$$

Здійснюючи зворотну підстановку (рекурентно) знайдемо

$$p(x) = p(x_0) + (x - x_0)\{p(x_0, x_1) + (x - x_1)\{p(x_0, x_1, x_2) + (x - x_2)\{p(x_0, x_1, x_2, x_3) +$$
$$+ \ldots + (x - x_{n-2})[p(x_0, x_1, x_2, \ldots, x_{n-1}) + (x - x_{n-1})(p(x_0, x_1, x_2, \ldots, x_n) + 0)]\}\}\}.$$

Тоді

$$\begin{aligned}
p(x) = {} & p(x_0) + (x - x_0)p(x_0, x_1) + (x - x_0)(x - x_1)p(x_0, x_1, x_2) + \\
& + (x - x_0)(x - x_1)(x - x_2)p(x_0, x_1, x_2, x_3) + \ldots + \\
& + (x - x_0)(x - x_1)(x - x_2)\ldots(x - x_{n-2})p(x_0, x_1, x_2, \ldots, x_{n-1}) + \\
& + (x - x_0)(x - x_1)(x - x_2)\ldots(x - x_{n-2})(x - x_{n-1})p(x_0, x_1, x_2, \ldots, x_n).
\end{aligned}$$

Залишилось врахувати наступне: якщо $p(x)$ – інтерполяційний поліном для функції $f(x)$, то $p(x_i) = f(x_i)$ і отримаємо наочну форму запису інтерполяційного багаточлена:

$$\begin{aligned}
N_n(x) \equiv {} & f(x_0) + (x - x_0)f(x_0, x_1) + (x - x_0)(x - x_1)f(x_0, x_1, x_2) + \\
& + (x - x_0)(x - x_1)(x - x_2)f(x_0, x_1, x_2, x_3) + \ldots + \\
& + (x - x_0)(x - x_1)(x - x_2)\ldots(x - x_{n-2})f(x_0, x_1, x_2, \ldots, x_{n-1}) + \\
& + (x - x_0)(x - x_1)(x - x_2)\ldots(x - x_{n-2})(x - x_{n-1})f(x_0, x_1, x_2, \ldots, x_n).
\end{aligned} \qquad (5.105)$$

5.4.2. Інтерполяційний многочлен Ньютона

Будемо шукати многочлен $P_n(x)$ степеня n, що задовольняє умовам у вигляді

$$P_n(x)=a_0+a_1(x-x_0)+a_2(x-x_0)(x-x_1)+\ldots+a_n(x-x_0)(x-x_1)\ldots(x-x_{n-1}), \qquad (5.106)$$

де $x_0, x_1, \ldots, x_n$ - задані значення аргументу x; $x_i-x_{i-1}=h=const$, $(\overline{i=0,n})$; коефіцієнти $a_0, a_1, \ldots, a_n$ невідомі.

Визначимо невідомі коефіцієнти $a_0, a_1, \ldots, a_n$, виходячи з умов (5.24). Покладемо у формулі (5.106) $x=x_0$. Тоді $P_n(x_0)=a_0$. Однак, у силу умов (5.24), $P_n(x_0)=y_0$. Отже, $a_0=y_0$. Для визначення a_1 покладемо у формулі (5.106) $x=x_1$, після чого одержимо

$$P_n(x_1) = a_0 + a_1(x_1 - x_0). \qquad (5.107)$$

З огляду на те, що $P_n(x_1)=y_1$, $a_0=x_0$, $x_1-x_0=h$, можемо записати $y_1=y_0+a_1h$, з цього виразу отримаємо $a_1=\dfrac{y_1-y_0}{h}$. Однак, $y_1-y_0=\Delta y_0$ – *кінцева різниця* 1-го порядку, отже, $a_1=\dfrac{\Delta y_0}{h}$. Далі, покладемо $x=x_2$, одержимо

$$P_n(x_2)=a_0+a_1(x_2-x_0)+a_2(x_2-x_0)(x_2-x_1). \qquad (5.108)$$

Оскільки $P_n(x_2)=y_2$, $a_0=y_0$, $a_1=\dfrac{\Delta y_0}{h}$, $x_2-x_0=2h$, $x_1-x_0=h$, можемо записати

$$y_2 = y_0 + \frac{\Delta y_0}{h}2h + a_2 2hh; \qquad (5.109)$$

звідси $a_2=\dfrac{y_2-y_0-2\Delta y_0}{2h^2}$.

Але $\Delta y_0=y_1-y_0$, тому $y_2-y_0-2\Delta y_0=y_2-y_0-2(y_1-y_0)=y_2-2y_1+y_0=\Delta^2 y_0$. Отже, $a_2=\dfrac{\Delta^2 y_0}{2h^2}$.

Аналогічні подальші обчислення (з урахуванням формули, що виражає різниці різних порядків через значення функції) дозволяють записати інші коефіцієнти:

$$a_3=\frac{\Delta^3 y_0}{3!h^3},\ a_4=\frac{\Delta^4 y_0}{4!h^4},\ \ldots,\ a_k=\frac{\Delta^k y_0}{k!h^k},\ \ldots,\ a_n=\frac{\Delta^n y_0}{n!h^n}.$$

Підставивши знайдені вираження коефіцієнтів у формулу (5.106), одержимо

$$\begin{aligned} P_n(x) &= y_0 + \frac{\Delta y_0}{h}(x-x_0) + \frac{\Delta^2 y_0}{2!h^2}(x-x_0)(x-x_1) + \\ &+ \frac{\Delta^3 y_0}{3!h^3}(x-x_0)(x-x_1)(x-x_2) + \ldots + \frac{\Delta^n y_0}{n!h^n}(x-x_0)(x-x_1)\ldots(x-x_{n-1}). \end{aligned} \qquad (5.110)$$

Це і є *інтерполяційна формула Ньютона*. Її можна подати в трохи іншому вигляді, більш зручному для практичного використання. Позначимо $\dfrac{x-x_0}{h}=q$. Тоді

$$\frac{(x-x_1)}{h}=\frac{x-(x_0+h)}{h}=q-1;\ \frac{x-x_2}{h}=q-2;\ ...;\ \frac{x-x_{n-1}}{h}=q-n+1$$

і формула (5.110) приймає вигляд

$$P_n(x)=y_0+q\Delta y_0+\frac{q(q-1)}{2!}\Delta^2 y_0+\frac{q(q-1)(q-2)}{3!}\Delta^3 y_0+...+ \\ +\frac{q(q-1)(q-2)...(q-n+1)}{n!}\Delta^n y_0 \qquad (5.111)$$

Формулу (5.111) доцільно використовувати для інтерполяції (екстраполювання) функції $y=f(x)$ в околиці початкового значення x_0, де q мале за абсолютною величиною. Якщо у формулі (5.111) прийняти $n=1$, одержимо формулу *лінійної інтерполяції*: $P_1(x)=y_0+q\Delta y_0$. При $n=2$ будемо мати формулу *параболічної (квадратичної) інтерполяції*: $P_2(x)=y_0+q\Delta y_0+\frac{q(q-1)}{2!}\Delta^2 y_0$.

При застосуванні першої інтерполяційної формули Ньютона зручно користатися горизонтальною таблицею кінцевих різниць, оскільки тоді необхідні значення різниць функції знаходяться у відповідному горизонтальному рядку таблиці.

Степінь n многочлена $P_n(x)$ на практиці бажано вибирати так, щоб кінцеві різниці $\Delta^n y_i$ були практично постійними. За початкове значення x_0 можна приймати будь-як табличне значення аргументу x.

5.4.3. Поділені різниці та їх властивості

Для $f\in C([a,b])$ існує ІМЛ $L_n(x)$, причому єдиний. Одержимо інший вигляд інтерполяційного многочлена Лагранжа $L_n(x)$ за допомогою розділених різниць.

Нехай функція $f(x)$ задана таблично

$$f_j\equiv f(x_j),\ j=\overline{0,n}. \qquad (5.112)$$

Означення 5.6. *Поділена різниця нульового порядку $f^0(x_i)$ співпадає з значеннями функції $f(x_i)$. Поділена різниця першого порядку за двома вузлами $x_i,x_j\in[a,b]$ визначимо рівністю*

$$f(x_i,x_j)\equiv\frac{f(x_j)-f(x_i)}{x_j-x_i}. \qquad (5.113)$$

Поділена різниця 2-го порядку у за трьома вузлами $x_i,x_j,x_k\in[a,b]$

$$f(x_i,x_j,x_k)\equiv\frac{f(x_j,x_k)-f(x_i,x_j)}{x_k-x_i}. \qquad (5.114)$$

Різницю k- го порядку визначимо за різницею (k - 1)- го порядку

$$f(x_i,\dots,x_{i+k}) \equiv \frac{f(x_{i+1},\dots,x_{i+k}) - f(x_i,\dots,x_{i+k-1})}{x_{i+k} - x_i}. \qquad (5.115)$$

Зауваження 5.6. Поділені різниці зручно записувати у вигляді табл. 5.6.

Таблиця 5.6

x_0	$f(x_0)$	-	-	-
x_1	$f(x_1)$	$f(x_0,x_1)$	-	-
x_2	$f(x_2)$	$f(x_1,x_2)$	$f(x_0,x_1,x_2)$	-
x_3	$f(x_3)$	$f(x_2,x_3)$	$f(x_1,x_2,x_3)$	$f(x_0,x_1,x_2,x_3)$

Теорема 5.8 (подання поділених різниць *k*-го порядку за значеннями функції). *Якщо маємо таблицю значень функції, тоді кожну поділену різницю k-го порядку можна подати через табличні значення функцій за такою формулою:*

$$f(x_i,x_{i+1},\dots,x_{i+k}) = \sum_{j=i}^{i+k} \frac{f(x_j)}{\prod\limits_{j\neq i}(x_j - x_i)}. \qquad (5.116)$$

Доведення. Доведення проведемо методом математичної індукції.

Для k=0 маємо $f^0(x_i) \equiv f_i$, $i = \overline{0,n}$. Для k=1 за означенням 5.6 $f(x_i,x_{i+1}) = \dfrac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} = \dfrac{f(x_i)}{x_i - x_{i+1}} + \dfrac{f(x_{i+1})}{x_{i+1} - x_i}$. Нехай формула (5.116) має місце для $k \le l-1$. Доведемо для k=l. Запишемо l-ту різницю за означенням (5.115), а потім для (l-1)-їх різниць застосуємо формулу (5.116):

$$f(x_i;x_{i+1};\dots;x_{i+l}) \equiv \frac{f(x_{i+1};x_{i+2};\dots;x_{i+l}) - f(x_i;x_{i+1};\dots;x_{i+l-1})}{x_{i+l} - x_i} = \frac{1}{x_{i+l} - x_i} \times$$

$$\times\{f(x_{i+1};x_{i+2};\dots;x_{i+l}) - f(x_i;x_{i+1};\dots;x_{i+l-1})\} = \frac{1}{x_{i+l} - x_i} \times$$

$$\times\left\{\frac{f(x_{i+1})}{(x_{i+1}-x_{i+2})(x_{i+1}-x_{i+3})(x_{i+1}-x_{i+l})} + \dots + \frac{f(x_{i+l-1})}{(x_{i+l-1}-x_{i+l})\dots(x_{i+l-1}-x_{i+1})} + \right.$$

$$+ \frac{f(x_{i+1})}{(x_{i+1}-x_{i+2})(x_{i+1}-x_{i+3})\dots(x_{i+1}-x_{i+l})} - \left[\frac{f(x_i)}{(x_i-x_{i+1})(x_i-x_{i+2})\dots(x_{i+1}-x_{i+l-1})} + \right.$$

$$\left.\left. + \frac{f(x_{i+1})}{(x_{i+1}-x_i)\dots(x_{i+1}-x_{i+l-1})} + \dots + \frac{f(x_{i+l-1})}{(x_{i+l-1}-x_i)(x_{i+l-1}-x_{i+1})\dots(x_{i+l-1}-x_{i+l-2})}\right]\right\}$$

Зауважимо, що $f(x_i)$ та $f(x_{i+1})$ входять один раз, а $f(x_{i+1}),...,f(x_{i+l-1})$ зустрічаються по два рази. Кожну з цих пар перетворимо наступним чином,якщо $j=i+1,...,i+l-1$,то

$$\frac{1}{x_{i+l}-x_i}\left[\frac{f(x_j)}{(x_j-x_{i+1})...(x_j-x_{i-1})(x_j-x_{i+1})...(x_j-x_{i+l})}-\right.$$

$$\left.-\frac{f(x_j)}{(x_j-x_i)...(x_j-x_{j-1})(x_j-x_{j+1})...(x_j-x_{i+l-1})}\right]=\frac{f(x_j)}{(x_j-x_{i+1})...(x_j-x_{j-1})}\times$$

$$\times\frac{1}{(x_j-x_{j+1})...(x_j-x_{i+l-1})(x_{i+l}-x_i)}\left[\frac{1}{x_j-x_{i+l}}-\frac{1}{x_j-x_i}\right]=$$

$$=\frac{f(x_j)}{(x_j-x_i)\otimes(x_j-x_{i+1})...(x_j-x_{j-1})(x_j-x_{j+1})...(x_j-x_{i+l-1})\otimes(x_j-x_{i+l})}.$$

Таким чином, одержали (l-2) доданків з $f(x_{i+1}),...,f(x_{i+l})$. Значить

$$f(x_i;x_{i+1};...;x_{i+k})=\frac{f(x_i)}{(x_i-x_{i+1})...(x_i-x_{i+k})}+\frac{f(x_{i-1})}{(x_{i+1}-x_i)(x_{i+1}-x_{i+2})...(x_{i+1}-x_{i+k})}+$$

$$+...+\frac{f(x_{i+k})}{(x_{i+k}-x_i)...(x_{i+k}-x_{i+k-1})}. \qquad (5.117)$$

Маємо (5.116) для кожного $k=0,1,...,n$.

Наслідок 5.2. *Поділена різниця від лінійної комбінації функцій є лінійною комбінацією поділених різниць цих функцій , тобто*

$$(a_1f_1+a_2f_2)(x_i,x_{i+1},...,x_{i+k})=a_1f_1(x_i,x_{i+1},...,x_{i+k}). \qquad (5.118)$$

Наслідок 5.3. *Поділені різниці є симетричні функції своїх аргументів, тобто*

$$f(x_i,x_{i+1},...,x_{i+k})=f(x_{i+1},x_i,...,x_{i+k})=f(x_{i+k},x_{i+k-1},...,x_i). \qquad (5.119)$$

Доведення наслідків випливає з формули (5.116).

5.4.4. Інтерполяційна формула Ньютона за поділеними різницями та нерівновіддаленими вузлами

Для інтерполяційної формули Лагранжа

$$L_n(x)=\sum\Phi_i(x)f(x_i). \qquad (5.120)$$

отримаємо іншу форму запису.

Теорема 5.9 (інтерполяційний многочлен Ньютона за поділеними різницями). *Нехай:*

1) $f\in C^n([a,b]);\exists f^{(n+1)}(x)\forall x\in(a,b)$;

2) *задана система функцій Чебишова* $\{\varphi_i(x) = x^i, i = 0,\dots,n\} \subset C^n([a,b])$;

3) $\{x_i, i = 0,1,\dots,n, x_i \neq x_j, i \neq j\} \subset [a,b]$;

4) $f_j \equiv f(x_i), j = 0,1,\dots,n.$

Тоді стверджується:

А) многочлен Лагранжа записується за поділеними різницями

$$L_n(x) = f(x_0) + (x - x_0) f(x_0; x_1) + (x - x_0)(x - x_1) f(x_0; x_1; x_2) + \dots + \\ + (x - x_0)(x - x_1)\dots(x - x_{n-1}) f(x_0; x_1; \dots; x_n) \quad (5.121)$$

В) похибка методу інтерполювання прийме вигляд

$$R_n(x) = f(x; x_0; x_1; \dots; x_n)\omega_n(x);$$

С) значення функції $f(x)$ *у будь-якій точці* $x \in [a,b]$ *дорівнює*

$$f(x) = L_n(x) + R_n(x). \quad (5.122)$$

Доведення. А) Побудуємо $(n+1)$ інтерполяційний многочлен Лагранжа $\{L_n(x)\}, k = 0,1,\dots,n$. Дійсно, згідно теореми 5.2 існують $f_i(x)$, який побудовано за вузлами $\{x_0, x_1, \dots, x_k\}$. Тоді інтерполяційний многочлен Лагранжа $L_n(x)$ можна подати у вигляді

$$L_n(x) = L_0(x) + [L_1(x) - L_0(x)] + [L_2(x) - L_1(x)] + \dots + [L_n(x) - L_{n-1}(x)]. \quad (5.123)$$

Для $k = 1,\dots,n$ запишемо різниці $L_k(x) - L_{k-1}(x)$, які є многочленами k-го степеня з коренями $x_0, x_1, \dots, x_{k-1}$, оскільки $f_i(x_i) = L_{k-1}(x_i) = f(x_i), i = 0,1,\dots,k-1$ за умови інтерполювання.

Тоді за Вієтом кожну з різниць подаємо у вигляді добутку двочленів

$$L_k(x) = L_{k-1}(x) = A(x - x_0)(x - x_1)\dots(x - x_{k-1}), k = 1,2,\dots,n. \quad (5.124)$$

У (5.124) підставимо $x = x_k$, тоді

$$L_k(x_k) - L_{k-1}(x_k) = A(x_k - x_0)(x_k - x_1)\dots(x_k - x_{k-1}). \quad (5.125)$$

Але $L_k(x_k) = f(x_k)$, значить

$$A = \frac{f(x_k)}{(x_k - x_0)\dots(x_k - x_{k-1})} -$$

$$\frac{\sum_{j=0}^{k-1} f(x_j)\frac{(x_j-x_0)\ldots(x_j-x_{j-1})(x_j-x_{j+1})\ldots(x_j-x_{k-1})}{(x_k-x_0)\ldots(x_k-x_{j-1})(x_k-x_{j+1})\ldots(x_k-x_{k-1})}}{(x_k-x_0)\ldots(x_k-x_{j-1})(x_k-x_j)(x_k-x_j)\ldots(x_k-x_{k-1})} = \cdot \tag{5.126}$$

$$= \sum_{j=0}^{k} \frac{f(x_j)}{(x_j-x_0)\ldots(x_j-x_{j-1})(x_j-x_{j+1})\ldots(x_j-x_k)} \equiv f(x_0;x_1;\ldots;x_k)\cdot$$

Якщо підставити A з (5.126) у (5.124), то матимемо:

$$L_k(x_k) - L_{k-1}(x_k) = f(x_0;x_1;\ldots;x_k)\omega_{k-1}(x), k = 1,2,\ldots,n. \tag{5.127}$$

Розпишемо (5.127):

$$\left.\begin{aligned} &L_0(x) = f(x_0);\\ &L_1(x) - L_0(x_k) = f(x_0;x_1)\omega_0(x);\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ &L_k(x) - L_{k-1}(x) = f(x_0;x_1;\ldots;x_k)\omega_{k-1}(x)\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ &L_n(x) - L_{n-1}(x) = f(x_0;x_1;\ldots;x_n)\omega_{n-1}(x) \end{aligned}\right\}. \tag{5.128}$$

Підставимо (5.128) у (5.123), що дає, так звану, інтерполяційну формулу Ньютона (5.121) за поділеними різницями.

В) Запишемо поділену різницю $f(x;x_0;x_1;\ldots;x_n)$ за формулою 5.116 за значеннями функції $f(x), f(x_0), f(x_1), \ldots, f(x_n)$:

$$f(x;x_0;x_1;\ldots;x_n) = \frac{f(x)}{(x-x_0)(x-x_1)\ldots(x-x_n)} + \frac{f(x_0)}{(x_0-x)(x_0-x_1)(x_0-x_2)\ldots(x_0-x_n)} + \tag{5.129}$$

$$+\ldots+\frac{f(x_n)}{(x_n-x)(x_n-x_0)(x_n-x_1)\ldots(x_n-x_{n-1})}.$$

Домножимо різність (5.129) зліва і справа на $\omega_n(x) \equiv (x-x_0)(x-x_1)\ldots(x-x_n)$ та враховуючи представлення $L_n(x)$ за формулою Лагранжа одержимо

$$f(x) - L_n(x) = f(x;x_0;\ldots;x_n)\omega_n(x) \equiv R_n(x). \tag{5.130}$$

що доводить пункт В) і С).

Наслідок 5.4. *За умов теореми 5.9 маємо формулу подання поділеної різниці* (n+1) *порядку для вузлів* $x_0, x_1, \ldots, x_n \in [a,b]$ *через* (n+1)*-похідну* $\forall x \in (a,b)$

$$f\left(x;x_0;x_1;...;x_n\right)=\frac{f^{(n+1)}\left(\zeta\right)}{\left(n+1\right)!}. \tag{5.131}$$

Доведення. За теоремою 5.5 маємо

$$f\left(x\right)-L_n\left(x\right)=\frac{f^{(n+1)}\left(\zeta\right)}{\left(n+1\right)!}\omega_n\left(x\right)\equiv R_n\left(x\right). \tag{5.132}$$

Але за пунктом В) теореми 5.9 маємо інший вигляд похибки через поділену різницю (n+1) порядку, де поряд з усіма вузлами $\left\{x_0;x_1;...;x_n\right\}$ має місце $\forall x\in\left[a,b\right]$

$$f\left(x;x_0;x_1;...;x_n\right)\equiv R_n\left(x\right). \tag{5.133}$$

Співставляючи (5.132) і (5.133) одержимо вираз (n+1)-ої поділеної різниці через похідну порядку (n+1) (5.131).

Наслідок 5.5. *На практиці результатом вимірювань є неусувні похибки* $\left\{\eta\right\}$, *тобто* $f_i^*=f_i+\eta_i$. *Тоді за формулою Лагранжа*

$$f^*\left(x\right)=\sum_{i=0}^{n}f_i\Phi_i\left(x\right)+\sum_{i=0}^{n}\eta_i\Phi_i\left(x\right). \tag{5.134}$$

Нехай відомо межі $\left\{\eta_i\right\}$, тобто $\left|\eta_i\right|\le\eta$. Тоді

$$\left|\varepsilon\right|\equiv\left|\sum_{i=0}^{n}\eta_i\Phi_i\left(x\right)\right|\le\eta\sum_{i=0}^{n}\left|\Phi_i\left(x\right)\right|. \tag{5.135}$$

Ця похибка може бути велика.

5.4.5. Скінченні різниці та їх властивості

Нехай функція $f\in\Im\left(\left[a,b\right]\right)$ задана табличною у вузлах $x_i=x_0+ih,i=0,1,...,n$ зі значеннями $f_i\equiv f\left(x_i\right)$.

Означення 5.7. *Скінченною різницею нульового порядку назвемо*

$$f_i^0\equiv f_i\equiv f\left(x_i\right). \tag{5.136}$$

Означення 5.8. *Скінченною різницею 1-го порядку за вузлами* $\left\{x_i,\mathrm{x}_{i+1}\right\}\subset\left[a,b\right]$ *назвемо різницю скінченних різниць нульового порядку і позначимо*

$$f^1_{i+\frac{1}{2}}\equiv f_{i+1}^0-f_i^0. \tag{5.137}$$

Непарні скінченні різниці (2k+1)-го порядку визначимо через парні, та позначимо ($k+\nu$)

$$f^{2k+1}_{i-\frac{1}{2}} \equiv f^{2k}_{i+1} - f^{2k}_{i}. \tag{5.138}$$

Таблиця 5.7

x_i	f_i^0	$f^1_{i+\frac{1}{2}}$	f_i^2	$f^3_{i+\frac{3}{2}}$	...	f
x_0	f_0					
		$f^1_{\frac{1}{2}}$				
x_1	f_1		f_1^2			
		$f^1_{\frac{3}{2}}$		$f^3_{\frac{3}{2}}$		
x_2	f_2		f_2^2			
		$f^1_{\frac{5}{2}}$				
x_3	f_3					$f^n_{\frac{n}{2}}$
...	...	...	...	...	...	
x_{n-2}	f_{n-2}					
		$f^1_{n-\frac{3}{2}}$		$f^3_{n-\frac{3}{2}}$		
x_{n-1}	f_{n-1}		f_n^2			
		$f^1_{n-\frac{1}{2}}$				
x_n	f_n					
	$f_n - f_0$	$f^1_{n-\frac{1}{2}} - f^1_{\frac{1}{2}}$	$f^2_{n-1} - f_1^2$			
		$\sum f^1$	$\sum f^2$	$\sum f^3$	...	$f^n_{\frac{n}{2}}$

Парні скінченні різниці $2k$-го порядку визначимо через непарні ($k+\nu$)

$$f_i^{2k} = f^{2k-1}_{i+\frac{1}{2}} - f^{2k-1}_{i-\frac{1}{2}}. \tag{5.139}$$

Скінченні різниці природно зводити в табл. 5.7 скінченних різниць.

Лема 5.5. *У таблиці скінченних різниць сума чисел кожного стовпчика дорівнює різниці крайніх чисел попереднього стовпчика.*

Доведення проведемо тривіальною перевіркою:

$$f^1_{\frac{1}{2}} + f^1_{\frac{3}{2}} + \ldots + f^1_{n-\frac{1}{2}} = (f_1 - f_0) + (f_2 - f_1) + \ldots + (f_{n-1} - f_{n-2}) + (f_n - f_{n-1}) = f_n - f_0. \tag{5.140}$$

Аналогічно перевіряється лема 5.5 для інших скінченних різниць.

Теорема 5.10. *Поділена різниця k-го порядку подасться за допомогою скінчених різниць k-го порядку у випадку рівновіддалених вузлів* $x_i = x_0 + ih, h > 0, i = 0,1,...,n,$ *наступною формулою*

$$f\left(x_i; x_{i+1}; ...; x_{i+k}\right) = \frac{f^k_{i+\frac{k}{2}}}{k!h^k}. \tag{5.141}$$

Доведення проводимо методом математичної індукції.

- Якщо k=0, то $f\left(x_i\right) = \frac{f_i^0}{0!h^0}$;

- Якщо k=1, то $f\left(x_i; x_{i+1}\right) = \frac{f_{i+1}^0 - f_i^0}{x_{i+1} - x_i} = \frac{f^1_{i+\frac{1}{2}}}{1!h^1}$;

- Якщо k=2, то $f\left(x_i; x_{i+1}; x_{i+2}\right) = \frac{f\left(x_{i+1}; x_{i+2}\right) - f\left(x_i; x_{i+1}\right)}{x_{i+2} - x_i} = \frac{f^1_{i+\frac{1}{2}} - f^1_{i+\frac{1}{2}}}{2 \cdot h \cdot h} = \frac{f^2_{i+1}}{2!h^2}$.

Нехай (5.141) справджується для $k \le l$, доведемо справедливість (5.141) для $k = l+1$. Дійсно,

$$f\left(x_i; x_{i+1}; ...; x_{i+l+1}\right) = \frac{f\left(x_{i+1}; ...; x_{i+l+1}\right) - f\left(x_i; ...; x_{i+l}\right)}{x_{i+l+1} - x_i} = \frac{f^l_{i+1+\frac{l}{2}} - f^l_{i+\frac{l}{2}}}{(l+1) \cdot h \cdot h^l \cdot l!} = \frac{f^{l+1}_{i+\frac{l+1}{2}}}{(l+1)h^{l+1}}. \tag{5.142}$$

Лема 5.6. *Скінчена різниця k-го порядку подається через значення функції за формулою*

$$f_i^k = \sum_{v=0}^{k} (-1)^v C_k^v f_{i+\frac{k}{2}-v}. \tag{5.143}$$

Доведення методом математичної індукції.

- Якщо k=0, то $f_i^0 = f_i$;
- якщо k=1,то $f^1_{i+\frac{1}{2}} = f_{i+1} - f_i = \sum_{v=0}^{1} (-1)^v C_1^v f_{i+1-v}$;
- якщо k=2, то

$$f_i^2 = f^1_{i+\frac{1}{2}} - f^1_{i-\frac{1}{2}} = f_{i+1} - f_i - \left(f_i - f_{i-1}\right) = f_{i+1} - 2f_i + f_{i-1} = \sum_{v=0}^{2} (-1)^v C_2^v f_{i+1-v}.$$

Нехай (5.143) справджується для $l \ge 1$: $f_i^l = \sum_{v=0}^{l} (-1)^v C_2^v f_{i+\frac{l}{2}-v}$. Доведемо, що (5.143) справджується для $k=l+1$

$$f_i^{l+1} = f^l_{i+\frac{1}{2}} - f^l_{i-\frac{1}{2}} = \sum_{v=0}^{l}(-1)^v C_l^v f_{i+\frac{1}{2}+\frac{l}{2}-v} - \sum_{v=0}^{l}(-1)^v C_l^v f_{i-\frac{1}{2}+\frac{l}{2}-v} = \left\{ f_{i+\frac{1}{2}+\frac{l}{2}} - \right.$$

$$\left. -C_l^1 f_{i+\frac{1}{2}+\frac{l}{2}-1} + C_l^2 f_{i+\frac{1}{2}+\frac{l}{2}-2} - \ldots + (-1)^m C_l^m f_{i+\frac{1}{2}+\frac{l}{2}-m} + \ldots + (-1)^l f_{i+\frac{1}{2}+\frac{l}{2}-l} \right\} -$$

$$- \left\{ f^1_{i+\frac{l}{2}-v} - C_l^1 f_{i-\frac{1}{2}+\frac{l}{2}-1} + C_l^2 f_{i-\frac{1}{2}+\frac{l}{2}-2} - \ldots + (-1)^m C_l^m f_{i-\frac{1}{2}+\frac{l}{2}-m} + \ldots + (-1)^l f_{i-\frac{1}{2}+\frac{l}{2}-l} \right\} =$$

$$= f^1_{i+\frac{l+1}{2}} - C_{l+1}^1 f_{i+\frac{l+2}{2}-1} + \ldots + (-1)^m C_{l+1}^m f_{i+\frac{l+1}{2}-m} + \ldots + (-1)^{l+1} f_{i-\frac{l+1}{2}} = \sum_{v=0}^{l+1}(-1)^v C_{l+1}^v f_{i+\frac{l+1}{2}-v}$$

де на попередньому кроці рівностей використали формулу комбінаторики $C_l^m + C_l^{m+1} = C_{l+1}^{m+1}$.

Наслідок 5.6. *Якщо* $f = \varphi \pm g$, *то* $f_i^k = \varphi_i^k \pm g_i^k$.

Наслідок 5.7. $(cf)_i^k = cf_i^k$.

Наслідок 5.8. *Якщо* $P_l(x) = \sum_{j=0}^{l} b_j x^i$, *то* $P_i^l = b_i; P_i^j = b_l; P_i^j = 0$, *при* j>l.

Доведення випливають з формули (5.143).

5.4.6. Формула Ньютона вперед і назад за скінченними різницями

Теорема 5.11. *Нехай виконуються умови теореми 5.2. Тоді має місце формула Ньютона вперед через скінченні різниці*

$$L_n(x) = L_n(x_0 - th) = f_0 + t \cdot f^1_{\frac{1}{2}} + \frac{t(t-1)}{2!} f_1^2 + \ldots + \frac{t(t-1)\ldots-(t-(n-1))}{n!} f^n_{\frac{n}{2}}. \quad (5.144)$$

Доведення. Побудуємо інтерполяційний многочлен Лагранжа $L_n(x)$ за поділеними різницями

$$L_n(x) = f(x_0) + (x-x_0) f(x_0; x_1) + (x-x_0)(x-x_1) f(x_0; x_1; x_2) + \ldots + (x-x_0)(x-x_1)\ldots(x-x_{n-1}) f(x_0; x_1; \ldots; x_n) \quad (5.145)$$

Зробимо заміну $t = \frac{x - x_0}{h}$, тоді

$$x - x_i = h(t-i), i = 0,1,\ldots,n. \quad (5.146)$$

А за (5.141) для i=0 можна записати

$$f(x_0; x_1; \ldots; x_n) = \frac{f_{\frac{k}{2}}^k}{k! h^k}. \quad (5.147)$$

Підставивши (5.146) і (5.147) у формулу Лагранжа (5.145) за поділеними різницями, одержимо інтерполяційну формулу Ньютона вперед (5.144) через скінченні різниці.

Наслідок 5.9. *Нехай справджуються умови теореми 5.11, а вузли* $x_i = x_0 - ih, \mathrm{h} > 0, i = 0,1,...,n$. *Тоді має місце формула Ньютона назад за скінченними різницями*

$$L_n(x) = L_n(x_0 - th) = f_0 + t \cdot f_{-\frac{1}{2}}^1 + \frac{t(t-1)}{2!} f_{-1}^2 + ... + \frac{t(t+1)...-(t+(n-1))}{\mathrm{n}!} f_{-\frac{n}{2}}^n. \qquad (5.148)$$

Доведення. Побудуємо інтерполяційний многочлен Лагранжа $L_n(x)$ через поділені різниці

$$L_n(x) = f(x_0) + (x - x_0) f(x_0; x_0 - h) + (x - x_0)(x - x_1) f(x_0, x_0 - h; x_0 - 2h) + ... + \\ + (x - x_0)(x - x_1)...(x - x_{n-1}) f(x_0; x_0 - h; x_0 - nh) \qquad (5.149)$$

Врахуємо вираз поділених різниць через скінченні різниці

$$\left.\begin{aligned} f(x_0 - h; x_0) &= f(x_0; x_0 - h) = \frac{f_{-\frac{1}{2}}^1}{1!(-h)}; \\ f(x_0 - 2h; x_0 - h; x_0) &= f(x_0; x_0 - h; x_0 - 2h) = \frac{f_{-1}^2}{2!(-h)^2}; \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f(x_0 - nh; ...; x_0 - h; x_0) &= f(x_0; x_0 - h; ...; x_0 - nh) = \frac{f_{-\frac{n}{2}}^n}{\mathrm{n}!(-h)^n}; \end{aligned}\right\} \qquad (5.150)$$

$$x_i - x_0 = t \cdot h, x - x_{-i} = (t + i)h. \qquad (5.151)$$

Підставляючи (5.150), (5.151) в (5.149) одержимо (5.148).

Наслідок 5.10. (про залишковий член інтерполяційної формули Ньютона вперед і назад). *Якщо неперервна* $f^{(n+1)}(x)$ *змінюється повільно на інтервалі (a,b), тоді для формули Ньютона вперед (5.144) залишковий член маємо структуру загального члена формули (5.144)*

$$R_n(x) = R_n(x_0 + th) \approx f_{\frac{n+1}{2}}^{n+1} \frac{t(t-1)...(t-n)}{(n+1)!}. \qquad (5.152)$$

а для формули Ньютона назад (5.148).

$$R_n(x) \approx f_{\frac{n+1}{2}}^{n+1} \frac{t(t+1)...(t+n)}{(n+1)!}. \qquad (5.153)$$

Доведення. Якщо неперервна $f^{n+1}(x)$ змінюється повільно на інтервалі (a,b), тоді повільно змінюються поділені різниці (n+1)-го порядку і, враховуючи (5.148), матимемо

$$f(x;x_0;...;x_n) \approx f(x_0;x_1;...;x_{n+1}) = \frac{f^{n+1}_{\frac{n+1}{2}}}{(n+1)!h^{n+1}}. \qquad (5.154)$$

Отже, $R_n(x) \approx f^{n+1}_{\frac{n+1}{2}} \frac{t(t-1)...(t-n)}{(n+1)!}$. Аналогічно для формули Ньютона назад має місце (5.153).

5.5. Числове диференціювання

5.5.1. Числове диференціювання для нерівновіддалених вузлів

Ідея числового диференціювання полягає у побудові $L_n(x)$ для $f \in C^{n+k+1}([a,b])$,тобто

$$f(x) = L_n(x) + R_n(x), \qquad (5.155)$$

де $L_n(x) = \sum_{i=0}^{n} f(x_j)\Phi_j(x)$ - інтерполяційна формула Лагранжа для вузлів $\{x_i, x_i \neq x_j, i = 0,1,...,n\}$ і $\{\varphi_i(x) \equiv x^i\}$ -- система Чебишева, i=0,1,…,n.

Якщо $f \in C^{n-k+1}([a,b])$, тоді зрозуміла рівність для похідних

$$f^{(j)}(x) = L^{(j)}(x) + R_n^{(j)}(x). \qquad (5.156)$$

j=1,2,…,k, випливають відповідні наближені формули для обчислення похідної k-го порядку.

Теорема 5.12 (про значення k-тої похідної). *Нехай* $\alpha_j \equiv x - x_j, j = 0,1,...,n$. *Тоді*

$$f^{(k)}(x) = k![f(x_0;x_1;...;x_k) + (\alpha_0 + \alpha_1 + ... + \alpha_k) f(x_0;x_1;...;x_{k+1}) + \\ + (\alpha_0\alpha_1 + ... + \alpha_k\alpha_{k+1}) f(x_0;x_1;...;x_{k+2}) + ... + (\alpha_0\alpha_1...\alpha_{n-k} + ... + \alpha_{k+1}\alpha_{k+2}...\alpha_{n-1}) \times \\ \times f(x_0;x_1;...;x_n)] + R_n^{(k)}(x) \qquad (5.157)$$

$$R_n^{(k)}(x)=\sum_{i=0}^{k}C_k^i\frac{d^i}{dx^i}f(x;x_0;...;x_n)\frac{d^{k-i}}{dx^{k-i}}\omega_n(x)=\sum_{i=0}^{i}\frac{k!}{(k-i)!}f(x;x;...;x;$$

$$x_0;...;x_n)\frac{d^{k-i}}{dx^{k-i}}\omega_n(x)=\sum_{i-0}^{i}\frac{k!}{(k-i)!(n+i+1)}f^{(n+i+1)}(\xi_l)\omega_n^{k-i}(x) \qquad (5.158)$$

$$\xi_l\in(x',x^n),x'=\inf\{x_0,x_1,...,x_n\}. \qquad (5.159)$$

Доведення. За теоремою 5.10 у рівності (5.155) ІМЛ $L_n(x)$ і $R_n(x)$ набувають вигляду:

$$L_n(x)=f(x_0)+\alpha_0 f(x_0,x_1)+\alpha_0\alpha_1 f(x_0,x_1,x_2)+...+\alpha_0\alpha_1...\alpha_n f(x_0,x_1,...,x_n)+$$

$$+\alpha_0\alpha_1...\alpha_n f(x,x_0,x_1,...,x_n);R_n(x)=\alpha_0\alpha_1...\alpha_n f(x,x_0,x_1,...,x_n)$$

Формулу (5.157) отримаємо за індукцією: якщо k=0, то маємо (5.145); якщо k=1:

$$f'(x)=f(x_0,x_1)+(\alpha_0+\alpha_1)f(x_0,x_1,x_2)+...+(\alpha_0\alpha_1...\alpha_{n-2}+\alpha_1\alpha_2...\alpha_{n-1})\times$$

$$\times f(x_0,x_1,...,x_n)+\frac{d\omega_n(x)}{dx}f(x,x_0,x_1,...,x_n)+\omega_n(x)\frac{df(x,x_0,x_1,....x_n)}{dx} \qquad (5.160)$$

Одержимо

$$\frac{df(x,x_0,...,x_n)}{dx}=\lim_{x'\to x}\frac{f(x',x_0,...,x_n)-f(x,x_0,...,x_n)}{x'-x}=\lim_{x'\to x}f(x',x,x_0,...,x_n)=f(x,x,x_0,...,x_n)$$

Далі легко закінчити доведення теореми 5.12.

5.5.2. Формули числового диференціювання для рівновіддалених вузлів

Лема 5.7. *Перші три похідні для формули Н'ютона впереді скінченними різницями мають вигляд (5.162).*

Доведення. Нехай

$$f(x)=f(x_0+th)=f_0+tf_{\frac{1}{2}}^1+\frac{t(t-1)}{2!}f_1^2+...+\frac{t(t-1)...\left[t-(n-1)\right]}{n!}f_{\frac{n}{2}}^n+R_n(x). \qquad (5.161)$$

Тоді маємо в результаті заміни $x=x_0+t\cdot h,h>0$, відповідні похідні

$$\left.\begin{aligned}
f'(x)&=\frac{df}{dt}\frac{dt}{dx}=\frac{1}{h}\left[f_{\frac{1}{2}}^1+\frac{2t-1}{2!}f_1^2+\frac{3t^2-6t+2}{3!}f_{\frac{n}{2}}^3+\frac{4t^3-18t^2+22t-6}{4!}f_2^4+...\right],\\
f''(x)&=\frac{1}{h^2}\left[f_1^2+\frac{6t-6}{3!}f_{\frac{3}{2}}^3+\frac{12t^2-36t+22}{4!}f_2^4+...\right],\\
f'''(x)&=\frac{1}{h^3}\left[f_{\frac{3}{2}}^3+\frac{2t-36}{4!}f_2^4+...\right],
\end{aligned}\right\} \qquad (5.162)$$

Для $x = x_0$

$$f'(x_0) = \frac{1}{h}\left[f^1_{\frac{1}{2}} - \frac{1}{2} f_1^2 + \frac{1}{3} f^3_{\frac{n}{2}} - \frac{1}{4} f_2^4 + \ldots \right],$$

$$f''(x_0) = \frac{1}{h^2}\left[f_1^2 - f^3_{\frac{3}{2}} + \frac{11}{12} f_2^4 - \ldots \right].$$

Якщо використати позначення $f_{i+1} - f_i = f^1_{i+\frac{1}{2}} = \Delta^1_i$ тощо, то

$$\left(h\frac{d}{dx} \right)^n f(x_0) = \left\{ \ln[1+\Delta] \right\}^n f(x_0). \tag{5.163}$$

де $\ln(1+\Delta) = \Delta - \frac{\Delta^2}{2} + \frac{\Delta^3}{2} - \ldots$.

5.5.3. Формули числового диференціювання методом невизначених коефіцієнтів

Нехай $f \in \gamma([a,b])$ задана таблицею $\{f_i = f(x_i), i = \overline{0,n}\}$, для якої існує похідна k-го порядку. Знаходимо її для $x^* \in (a,b)$ *у вигляді*:

$$f^{(k)}(x^*) \approx \sum_{i=0}^{n} c_i f(x_i). \tag{5.164}$$

Коефіцієнти $\{c_i, i = 0,1\ldots,n\}$ вибираємо таким чином, щоб формула (5.164) була точною для многочленів найвищого степеня:

$$\sum_{j=0}^{m} a_j x^j. \tag{5.165}$$

Будемо вимагати, щоб формула (5.164) була точною для (5.165):

$$\sum_{j=0}^{m} a_j (x^j)^{(k)}\big|_{x^*} = \sum_{i=0}^{n} c_i \left(\sum_{j=0}^{m} a_j x_i^j \right). \tag{5.166}$$

Для того, щоб рівність виконувалась для кожного многочлена m-го степеня необхідно і досить, щоб коефіцієнти при a_j в правій та лівій частинах були рівними. Враховуючи, що похідна k-го порядку має вигляд:

$$(x^j)^{(k)} = j(j-1)\ldots(j-k-1)x^{j-k},$$

рівність відповідних коефіцієнтів дасть

$$j(j-1)\ldots(j-k+1)(x^*)^{j-k} = \sum_{i=0}^{n} c_i x_i^j, \quad j = \overline{0,m}. \tag{5.167}$$

Якщо $m = n-1$, то кількість рівнянь дорівнює кількості невідомих. Оскільки визначник Вандермонда відмінний від нуля, тоді однозначно одержимо $\{c_i\}, i = 0,1,...n$.

Висновок. Завжди можна побудувати формулу числового диференціювання за n вузлами, яка точна для многочленів степеня $n-1$.

5.6. Раціональна інтерполяція

Нехай $f \in \Im([a,b])$ та $\{f_i \equiv f(x_i)\}$, $i = 0,n$. Тоді наближення до $f(x)$ шукаємо у вигляді раціональної функції, тобто

$$f(x) \approx R(x) = \frac{a_0 + a_1 x + ... + a_p x^p}{b_0 + b_1 x + ... + b_q x^q} \quad p + q + 1 = n.$$

Із співвідношення $f(x_j) = R(x_j), j = 0,...,n$ випливає

$$\sum_{j=0}^{p} a_i x_i^j - f(x_i) \sum_{j=0}^{q} b_j x_i^j = 0, \; i = 0,n. \tag{5.168}$$

Система (5.168) має $(n+1)$ незалежних рівнянь відносно $(n+2)$ невідомих. Така система завжди має нетривіальний розв'язок, тому можна знайти коефіцієнти $\{a_i, b_i\}$. Функція $R(x)$ може бути записана у явному вигляді для випадку непарного n для $p = q$ та парного n для $p - q = 1$.

Для цього треба обчислити обернені поділені різниці, що визначаються умовами

$$f^{-1}(x_l; x_k) \equiv \frac{x_l - x_k}{f(x_l) - f(x_k)}. \tag{5.169}$$

та рекурентними співвідношеннями

$$f^{-1}(x_k; ...; x_l) \equiv \frac{x_l - x_k}{f^{-1}(x_{k+1}; ...; x_l) - f^{-1}(x_k; ...; x_{l-1})}.. \tag{5.170}$$

Інтерполяційна раціональна функція записується у вигляді ланцюгового дробу

$$f(x) = f(x_0) + \cfrac{x - x_0}{f^{-1}(x_0; x_1) + \cfrac{x - x_1}{f^{-1}(x_0; x_1; x_2) + ... + \cfrac{x - x_{n-1}}{f^{-1}(x_0; ...; x_n)}}}$$

Використання раціональної інтерполяції за вдало вибраними вузлами часто більш ефективніше інтерполяції многочленами [76-80].

5.7. Інтерполювання функцій комплексної змінної

Задача інтерполювання для функцій комплексної змінної можна розв'язати таким чином.

Теорема 5.13. *Нехай:*

1) *задана аналітична функція* $f(z)$ *комплексної змінної на простому замкнутому контурі* $\Gamma \in C$ *і всередині цього контуру* $z \in \bigcup_\Gamma$:

2) вузли $\{z_j\} \subset \bigcup_\Gamma, j = 0,1,...,n;$

3) *задано значення функції* $f_j \equiv f(z_j), j = 0,1,...,n.$

Тоді

$$f(z) = L_n(z) + R_n(z); \tag{5.171}$$

$$L_n(z) = \sum_{j-0}^{n} f_j \frac{\omega_n(z)}{\omega_n'(z_j)(z - z_j)}. \tag{5.172}$$

$$R_n(z) = \frac{\omega_n(z)}{2\pi i} \int_\Gamma \frac{f(\xi)}{\omega_n(\xi)(\xi - z)} d\xi, \tag{5.173}$$

де $\omega_n(z) \equiv (z - z_0)(z - z_1)...(z - z_n), i = \sqrt{-1}$

Доведення. Розглянемо інтеграл

$$P(z) = \frac{1}{2\pi i} \int_\Gamma \frac{\omega_n(\xi) - \omega_n(z)}{\omega_n(\xi)(\xi - z)} f(\xi) d\xi. \tag{5.174}$$

З одного боку

$$P(z) = \frac{1}{2\pi i} \int_\Gamma \frac{f(\xi)}{\xi - z} d\xi - \frac{\omega_n(z)}{2\pi i} \int_\Gamma \frac{f(\xi)}{\omega_n(\xi)(\xi - z)} d\xi \equiv f(z) - R_n(z). \tag{5.175}$$

де перший інтервал за формулою Коші рівний $f(z)$.

З другого боку розглянемо підінтегральну функцію в (5.174) $\varphi(\xi) = \frac{\omega_n(\xi) - \omega_n(z)}{\omega_n(\xi)(\xi - z)} f(\xi)$. Ця функція має полюси z_j, вона аналітична на Γ та всередині контуру $\bigcup_\Gamma$.

За формулою про знаходження лишків [76-80] одержимо:

$$2\pi i \cdot P(z) = \int_1 \varphi(\xi) d\xi = 2\pi i \sum_{j=0}^{n} \operatorname{Re} s[\varphi(\xi), z_j] = 2\pi i \sum_{j=0}^{n} \lim_{\xi \to z_j} \varphi(\xi)(\xi - z) =$$

$$2\pi i\sum_{j=0}^{n}\frac{1}{(\alpha_j-1)!}\frac{d^{\alpha_j-1}}{dz^{\alpha_j-1}}\left[\varphi(\xi)(\xi-z_j)^{\alpha_j}\right]=2\pi i\sum_{j=0}^{n}f(z_j)\frac{\omega_n(z)}{\omega_n'(z_j)\cdot(z-z_j)},$$

де α_j -кратність z_j -го полюса. Тоді, спрямувавши $\xi\to z_j$, маємо

$$\sum_{j=0}^{n}f(z_j)\frac{\omega_n(z)}{\omega_n'(z_j)(z-z_j)}\equiv L_n(z). \tag{5.176}$$

Якщо (5.176) підставити в (5.175), одержимо формули (5.171), (5.172), (5.173).

5.8. Інтерполювання функцій багатьох змінних

5.8.1. Труднощі інтерполювання функцій багатьох змінних

Труднощі інтерполювання функцій багатьох змінних розглянемо на прикладі функції двох змінних

$$u=f(x,y). \tag{5.177}$$

$\{x,y,z\}\subset\mathbf{R}^1$, $(x,y)\subset G$, де G - замкнена область.

Задано значення цієї функції $u_j\equiv f(x_j,y_j)$, $j=\overline{0,n}$.

Ставиться задача: знайти аналітичний вигляд функції $f(x,y)$. Будемо шукати функцію $f(x,y)$ у вигляді многочлена m -го степеня:

$$\begin{aligned}P_m(x,y)=a_{00}+a_{10}x^1y^0+a_{01}x^0y^1+a_{11}x^1y^1+a_{12}x^1y^2+\\+\ldots+a_{m0}x^my^0+\ldots+a_{0m}x^0y^m\end{aligned}. \tag{5.178}$$

Назвемо сукупність (x_j,y_j) вузлами інтерполювання. Для побудови многочлена $P_m(x,y)$ методом невизначених коефіцієнтів, треба вимагати, щоби

$$P_m(x_j,y_j)=f(x_j,y_j). \tag{5.179}$$

Умову (5.179) назвемо умовою інтерполювання. Система (5.179) рівносильна системі:

$$a_{00}+a_{10}x_j+a_{01}y_j+a_{11}x_jy_j+\ldots+a_{0m}y_j^m=f_j,\ j=\overline{0,n}. \tag{5.180}$$

Сукупність рівностей (5.180) є системою лінійних алгебраїчних рівнянь $(n+1)$ -го порядку з $\frac{(m+1)(m+2)}{2}$ невідомими. Система лінійних алгебраїчних рівнянь має єдиний розв'язок, коли кількість рівнянь співпадає з кількістю невідомих за умови відмінності від нуля основного визначника [76-80]. Тому труднощі інтерполювання полягають в тому, що мають виконуватись такі умови:

I. Система лінійних алгебраїчних рівнянь (5.180) замкнена, якщо $n+1=\frac{(m+1)(m+2)}{2}$, а коефіцієнти тоді знаходимо за формулами Крамера $a_{ij}=\frac{\Delta_{ij}}{\Delta}, \Delta\neq 0$.

II. Вузли інтерполювання (x_j, y_j) не можуть розташовуватись довільно. Доведемо:

1) якщо кількість вузлів дорівнює трьом і вузли розташовані на одній прямій, то $\Delta_3=0$;
2) якщо вузлів шість і вузли лежать на довільній кривій другого порядку, то $\Delta_6=0$;
3) якщо вузлів одинадцять і вузли лежать на поверхні другого порядку, то $\Delta_{11}=0$.

III. Трудність полягає в оцінці залишкового члена. Теорема Ролля тут не діє, тому для функцій багатьох змінних оцінку не проводимо. У зв'язку з вказаними труднощами інтерполювання функції двох змінних розглянемо лише частинні випадки.

5.8.2. Побудова інтерполяційних формул за поділеними різницями

Нехай функція $u=f(x,y)$ задана у вузлах інтерполювання:

$$\left.\begin{array}{l}(x_0,y_0),(x_1,y_0),\ldots,(x_n,y_0)\\(x_0,y_1),(x_1,y_1),\ldots,(x_n,y_1)\\\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\(x_0,y_m),(x_1,y_m),\ldots,(x_n,y_m)\end{array}\right\}. \qquad (5.181)$$

тобто задано значення функції $f(x_j,y_k), j=0,1,\ldots,n; k=0,1,\ldots,m.$

Теорема 5.14 (про побудову інтерполяційного многочлена двох змінних). *Нехай* $f\in C^{n,m}([a,b],[c,d])$. *існує* $f\in C^{n+1,m+1}((a,b),(a,b))$. *Тоді:*

$$f(x,y)=L_{n,m}(x,y)+R_{n,m}(x,y). \qquad (5.182)$$

$$L_{n,m}(x,y)=\sum_{i=1}^{n}\sum_{j=1}^{m}(x-x_0)\ldots(x-x_{i-1})(y-y_0)\ldots(y-y_{j-1})\times f(x_0,\ldots,x_i,y_0,\ldots,y_j). \qquad (5.183)$$

$$R_{n,m}(x,y)=\frac{\omega_n(x)}{(n+1)!}\frac{\partial^{n+1}}{\partial x^{n+1}}f(\xi,y)+\frac{\omega_m(y)}{(m+1)!}\frac{\partial^{m+1}}{\partial y^{m+1}}f(x,\eta)-\frac{\omega_n(x)}{(n+1)!}\frac{\omega_m(y)}{(m+1)!}\frac{\partial^{m+n+2}}{\partial x^{n+1}\partial y^{m+1}}f(\xi_1,\eta_1), \qquad (5.184)$$

$\xi,\xi_1\in[a,b], \eta,\eta_1\in[c,d].$

Доведення. Застосуємо інтерполяційну формулу Ньютона з поділеними різницями за вузлами $\{y_0, y_1, ..., y_m\} \subset [c,d]$, при фіксованому $x \in [a,b]$.

$$L_m(x,y) = f(x,y_0) + (y-y_0)f(x,y_0,y_1) + ... + (y-y_0)...(y-y_{m-1}) \times \\ \times f(x,y_0,y_1,...,y_m) + (y-y_0)...(y-y_m)f(x,y_0,y_1,...,y_m,y). \quad (5.185)$$

Далі застосовуємо формулу Ньютона до $f(x,y_0,y_1...,y_n)$ по x за вузлами $\{x_0,x_1,...,x_n\} \subset [a,b]$:

$$f(x,y_0,y_1...,y_n) = f(x_0,y_0,...,y_n) + (x-x_0)f(x_0,x_1,y_0,y_1...,y_n) + \\ +...+(x-x_0)(x-x_1)...(x-x_{n-1})f(x_0,x_1,...,x_n y_0,y_1...,y_n) + \quad (5.186) \\ +(x-x_0)(x-x_1)...(x-x_n)f(x_0,x_1,...,x_n;x;y_0,y_1...,y_n), n=0,1,...m.$$

Підставимо (5.186) в (5.185) одержимо формули (5.182), (5.183). Зауважимо, що аналогічно можна знайти вигляд похибки $R_{n,m}(x,y)$.

5.9. Збіжність інтерполяційного процесу

Нехай задана трикутна зліченна матриця

$$\left.\begin{array}{l} x_0^{(0)} \\ x_0^{(1)}, x_1^{(1)} \\ x_0^{(2)}, x_1^{(2)}, x_2^{(2)} \\ \ldots\ldots\ldots\ldots\ldots \\ x_0^{(n)}, x_1^{(n)}, ..., x_n^{(n)} \\ \ldots\ldots\ldots\ldots\ldots \end{array}\right\}. \quad (5.187)$$

для якої $\{x_i^{(j)}, i \geq 1, j \geq 1\} \subset [a,b]$.

Для $f \in \Im([a,b])$ будується послідовність інтерполяційних многочленів $\{L_n(x), n=0,1,2...\}$, які назвемо інтерполяційним процесом. Тоді

$$f(x) = L_n(x) + R_n(x). \quad (5.188)$$

$$L_n(x) = \sum_{i=0}^{n} f(x_i) \frac{\prod_{j \neq i}(x-x_j)}{\prod_{j \neq i}(x_i-x_j)}. \quad (5.189)$$

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)(x-x_1)...(x-x_n), \quad (5.190)$$

$\xi \in (a,b)$. Для побудови $L_n(x)$ використовуємо всі елементи матриці (5.187).

Означення 5.9. *Інтерполяційний процес назвемо збіжним в точці* $x \in [a,b]$, *якщо*

$$\lim_{n\to\infty} L_n(x) = f(x). \tag{5.191}$$

Означення 5.10. *Інтерполяційний процес називається рівномірно збіжним, якщо він збігається для* $\forall x \in [a,b]$, *тобто виконується (5.181)* $\forall x \in [a,b]$, *і це будемо позначати*

$$L_n(x) \rightrightarrows f(x). \tag{5.192}$$

Теорема 5.15 (Теорема Фабера). *Для довільної заданої матриці вузлів (5.187)* $\{x_i^j, i \ge 1, j \ge 1\} \subset [a,b]$ *існує* $f \in C^1([a,b])$ *така, послідовність інтерполяційних многочленів* $\{L_n(x), n \ge 1\}$ *,що збігається рівномірно*

$$L_n(x) \neq f(x). \tag{5.193}$$

Теорема 5.16 (Теорема Бернштейна). *Нехай задана функція*

$$f(x) = |x|, \forall x \in [-1,+1]. \tag{5.194}$$

Якщо побудувати для $f(x)$ *інтерполяційний многочлен Лагранжа за рівновіддаленими вузлами* $x_i = -1 + ih, h > 0$, *такими що* $x_0^{(n)} = -1, x_n^{(n)} = 1$, *то не спостерігається збіжності до* $f(x) = |x|$ *тобто*

$$\lim_{n\to\infty} L_n(x) \neq |x|, \forall x \in (-1,0) \cup (0,1).. \tag{5.195}$$

Зрозуміло, що питанню збіжності інтерполяційного процесу присвячено багато літератури з застосуванням тонких методів математичного аналізу. Тому розглянемо найпростіший випадок цілих функцій f.

Означення 5.11. *Функцію* $f: \mathbf{R}^1 \to \mathbf{R}^1$ *назвемо цілою, якщо її можна представити у вигляді суми степеневого ряду:*

$$f(x) = \sum_{i=0}^{\infty} a_i (x - x_0)^i, \tag{5.196}$$

який збігається для $\forall x \in \mathbf{R}^1$, $\{a_i\} \subset \mathbf{R}^1$

Теорема 5.17. *Якщо функція* $f: \mathbf{R}^1 \to \mathbf{R}^1$ *ціла* $\forall x \in [a,b]$, тобто *представлена у вигляді* (5.196), *тоді* $L_n(x) \rightrightarrows f(x)$, *якщо ІМЛ побудовано за довільною інтерполяційною матрицею* (5.187).

Доведення. Зауважимо, що $f \in C^{\infty}([a,b])$, тоді вірна оцінка

$$|f(x)-L_n(x)|<\frac{M_{n+1}}{(n+1)!}(b-a)^{(n+1)};\ M_{n+1}\equiv\sup_{x\in[a,b]}|f^{(n+1)}(x)|,\quad бо\quad |\omega_n(x)|<(b-a)^{n+1},$$

де $\{x_i^{(j)}\}\subset[a,b], i,j=0,1,\ldots,n.$

Якщо при $n\to\infty$ права частина прямує до 0, будемо мати твердження теореми 5.17. Очевидна нерівність

$$|f^{n+1}(x)|\le(n+1)!|a_{n+1}|+(n+2)(n+1)\ldots2|a_{n+2}||x-x_0|+$$
$$+\ldots+(n+k)(n+k-1)\ldots k|a_{n+k}||x-x_0|^{k-1}+\ldots<(n+1)^{n+1}|a_{n-1}|+(n+2)^{n+1}|a_{n+2}||x-x_0|+$$
$$+\ldots+(n+k)^{n+1}|a_{n-k}||x-x_0|^{k-1}+\ldots,$$

де використали нерівність $(n+k)(n+k-1)\ldots k<(n+k)^{n+1}$

Якщо $(\frac{n+k}{n+1})^{n+1}=(1+\frac{k-1}{n+1})^n<e^{k-1}$, то маємо

$$\frac{|f^{(n+1)}(x)|}{(n+1)^{n+1}}\le|a_{n+1}|+|a_{n+2}|(e|x-x_0|)+\ldots+|a_{n+k}|(e|x-x_0|)^{k-1}+\ldots. \qquad (5.197)$$

Множимо (5.197) на $S^{n+1}\equiv(e(b-a))^{n+1}$, звідки

$$\frac{M_{n+1}}{(n+1)!}(b-a)^{n+1}<\frac{M_{n+1}}{(n+1)^{n+1}}(e(b-a))^{n+1}\le\sum_{k=n+1}^{\infty}|a_k|K^k\underset{n\to\infty}{\to}0. \qquad (5.198)$$

$$K\equiv\sup\{S,\max_{x\in[a,b]}(e|x-x_0|)\};e^{n+1}>\frac{(n+1)^{n+1}}{(n+1)!}. \qquad (5.199)$$

З (5.198) випливає $\lim_{n\to\infty}|f(x)-L_n(x)|=0$ для $\forall x\in[a,b]$, тобто $L_n(x)\Rightarrow f(x)$, $\forall x\in[a,b]$.

Зауваження 5.7. Вимога того, що функція f є цілою, дуже важлива.

Нехай для $x\in[-1,+1]$

$$f(x)=\begin{cases}e^{\frac{1}{x^2}},\forall x\ge0;\\0,\forall x<0;\end{cases}\ \text{та}\ f\in C^{\infty}(R\setminus\{0\}). \qquad (5.200)$$

Якщо вибрати $x_j\in[-1,0)$, тоді $L_n(x)\equiv0$. Звідки випливає $L_n(x)\ne f(x)$ для $\forall_x>0$.

5.10. Інтерполювання сплайнами

Нехай $f \in \Im([a,b])$ задана таблицею $\{f_i = f(x_i)\}$, $x_i \in [a,b], i = 0,1,...,N$. Заміною $x = (b-a)t + a$, завжди можна звести $[a,b]$ до $[0,1]$. Тому будемо розглядати випадок $f \in \Im([0,1])$. Розбиття $0 = x_0 < x_1 < x_2 < ... < x_{N-1} < x_N = 1$ позначимо через Δ.

Означення 5.12. Назвемо сплайном $S_\Delta^m(f,x)$ порядку m функцію, яка є многочленом степеня m на кожному з відрізків $[x_{n-1}, x_n]$, тобто

$$S_\Delta^m(f,x) \equiv P_{nm}(x) = a_{n0} + a_{n1}x + ... + a_{nm}x^m,\ x_{n-1} \le x \le x_n,\ n = \overline{1,N}. \qquad (5.201)$$

$n = 1,2,...,N$ і який задовольняє умовам неперервності похідних порядку $(m-1)$ в точках $x_1,...,x_{N-1}$:

$$P_{nm}^{(k)}(x_n) = P_{n+1}^{(k)}(x_n),\ k = \overline{0,m-1}. \qquad (5.202)$$

Всього маємо $Q = N(m+1)$ невідомих коефіцієнтів a_{nm}. a співвідношення (5.202) утворюють СЛАР з $(N-1)m$ -рівняннями. Щоб мати замкнену систему треба додати рівняння коефіцієнтів з умови наближення сплайнів до функцій:

$$\left.\begin{aligned} P_{nm}(x_{n-1}) &= f(x_{n-1}), n = \overline{1,N-1} \\ P_{nm}(x_n) &= f(x_n), n = \overline{1,N-1} \end{aligned}\right\}. \qquad (5.203)$$

Одержимо два випадки.

1) Цих умов (5.203) достатньо, якщо $(N-1)m + 2N = N(m+1)$, тобто $m = N$.

2) $m \ne N$. Тоді з умов перетворення в нуль на кінцях x_0 та x_N похідних $(m-1)$ - го порядку матимемо

$$P_{nm}^{(m-1)}(x_0) = 0 \ldots P_{nm}^{(m-1)}(x_N) = 0. \qquad (5.204)$$

Приклад 5.1. Побудувати сплайн 1-го порядку $S_\Delta^1(f,x)$ за умови співпадіння їх значень із значеннями функцій $f_0, f_1, ..., f_N$.

Розв'язок. Нехай $m=1$, тоді $S^1(f,x) \equiv f(x) \equiv P_{n1}(x) = a_{n0} + a_{n1}x, n = 0,1,2,...,N$, а значить

$$\begin{cases} P_{n1}(x_n) = f(x_n) = a_{n0} + a_{n1}x_n \\ P_{n1}(x_{n-1}) = f(x_{n-1}) = a_{n0} + a_{n1}x_{n-1} \end{cases} \qquad (5.205)$$

Матимемо $2N$ невідомих $\{a_{ij}\}, j=0,1$:

$$a_{n1}=\frac{f(x_n)-f(x_{n-1})}{x_n-x_{n-1}}=f(x_n,x_{n-1}). \tag{5.206}$$

$$a_{n0}=f(x_n)-f(x_n,x_{n-1})x_n. \tag{5.207}$$

Висновок. Многочлен $P_{n1}(x)=S_{\Delta}^{1}(f,x)$ є інтерполяційним многочленом першого ступеня з вузлами інтерполювання $\{x_{n-1};x_n\}$.

Приклад 5.2. Побудувати сплайн $S_{\Delta}^{2}(f,x)$, який співпадає з $f(x)$ у вузлах $x_0,x_1,\ldots,x_N(m=2), x_i=x_0+ih, n=0,1,2,\ldots,N$.

Розв'язок. Нехай $m=2$, тоді шукаємо

$$S_{\Delta}^{2}(f,x)\equiv P_{n2}(x)=a_{n0}+a_{n1}(x-x_{n-1})+a_{n2}(x-x_{n-1})^2, \tag{5.208}$$

$$\begin{cases} P_{n2}(x_{n-1})=f(x_{n-1})=a_{n0}, n=1,\ldots,N-1; \\ P_{n2}(x_n)=f(x_n)=a_{n0}+a_{n1}h+a_{n2}h^2. \end{cases} \tag{5.209}$$

Звідки матимемо $3N$ невідомих $\{a_{ij}\}, j=0,1,2$.

Система (5.209) є системою лінійних алгебраїчних рівнянь з $N-1+N-1=2N-2$ рівняннями. В точці $x=x_n$ маємо

$$P_{n2}^{(1)}(x_n)=P_{n+1,2}^{(1)}(x_n), \quad n=\overline{1,N}, \tag{5.210}$$

$$P_{12}^{(1)}(x_0)=0, P_{N2}^{(1)}(x_N)=0, \quad n=\overline{1,N-1}. \tag{5.211}$$

Тобто всього $3N-1$ рівнянь (5.209), (5.210), (5.211).

Приклад 5.3. Побудувати сплайн $S_{\Delta}^{3}(f,x)$, який співпадає з $f(x)$ у вузлах $x_0,x_1,\ldots,x_N, x_i=x_0+ih$.

Ров'язок. Шукаємо

$$S_{\Delta}^{3}(f,x)\equiv P_{n3}(x)=a_{n0}+a_{n1}(x-x_{n-1})+a_{n2}(x-x_{n-1})^2+a_{n3}(x-x_{n-1})^3 \tag{5.212}$$

з $4N$ невідомими $\{a_{ij}\}, x_{n-1}<x<x_n$. Таким чином,

$$\left.\begin{aligned} &P_{n3}(x_n)=f_n=a_{n0}+a_{n1}h+a_{n1}h^2+a_{n2}h^3 \\ &P'_{n3}(x_n)=P'_{n+1,3}(x_n), n=\overline{1,N-1} \\ &P''_{n3}(x_n)=P''_{n+1,3}(x_n), n=\overline{1,N-1} \end{aligned}\right\}. \tag{5.213}$$

$a_{n0}=P_{n3}(x_{n-1})=f_{n-1}; n=1,2,\ldots,N-1(2N-2невідомих)$.

Треба задати ще умови на кінцях, тобто в точках x_0 та x_{N-1} ($4N$-4 невідомих коефіцієнтів $\{a_{ij}\}$):

$$\left.\begin{aligned} P'_{13}(x_0) &= 0 \\ P'_{N3}(x_N) &= 0 \end{aligned}\right\}. \tag{5.214}$$

$$\left.\begin{aligned} P''_{13}(x_0) &= 0, \\ P''_{N3}(x_N) &= 0. \end{aligned}\right\}. \tag{5.215}$$

Всього $4N$ невідомих коефіцієнтів $\{a_{ij}\}, j = 0,1,2,3$ та $4N$ рівнянь (5.213), (5.214), (5.215).

5.11. Приклади розв'язку задачі інтерполяції функції

Приклад 5.11.1. За допомогою формули Лагранжа побудувати інтерполяційний многочлен для функції, що задана табл. 5.8. Обчислити значення в точці x=0,5.

Таблиця 5.8

x_i	–2	–1	0	1	2
f_i	3	–1	–1	3	–13

Розв'язок. За формулою Лагранжа

$$L_4(x) = 3\frac{(x+1)x(x-1)(x-2)}{(-1)(-2)(-3)(-4)} - 1\frac{(x+2)x(x-1)(x-2)}{1(-1)(-2)(-3)} - 1\frac{(x+2)(x+1)(x-1)(x-2)}{2\cdot 1(-1)(-2)} +$$

$$+ 3\frac{(x+2)(x+1)x(x-2)}{3\cdot 2\cdot 1(-1)} - 13\frac{(x+2)(x+1)x(x-1)}{4\cdot 3\cdot 2\cdot 1}.$$

Далі

$$L_4(0,5) = 3\frac{1,5\cdot 0,5\cdot(-0,5)(-1,5)}{(-1)(-2)(-3)(-4)} - 1\frac{2,5\cdot 0,5(-0,5)(-1,5)}{1(-1)(-2)(-3)} - 1\frac{2,5\cdot 0,5\cdot(-0,5)(-1,5)}{2\cdot 1(-1)(-2)} +$$

$$+ 3\frac{2,5\cdot 1,5\cdot 0,5\cdot(-0,5)}{3\cdot 2\cdot 1(-1)} - 13\frac{2,5\cdot 1,5\cdot 0,5(-0,5)}{4\cdot 3\cdot 2\cdot 1} = \frac{23}{16} = 1,438.$$

Приклад 5.11.2. Побудувати інтерполяційний многочлен за формулою Ньютона для функції, що задана таблицею (x), та обчислити значення в точці x=0,5.

Розв'язок. Таблиця розділених різниць для цих даних має вигляд

Таблиця 5.9

x_i	f_i				
–2	3				
		–4			
–1	–1		2		
		0		0	
0	–1		2		–1
		4		–4	
1	3		–10		
		–16			
2	–13				

Тоді за формулою Ньютона

$$L_4(x) = 3 - 4(x+2) + 2(x+2) + 0(x+2)(x+1)x - 1(x+2)(x+1)x(x-1).$$

Далі

$$L_4(0{,}5) = 3 - 4 \cdot 2{,}5 + 2 \cdot 2{,}5 \cdot 1{,}5 - 2{,}5 \cdot 1{,}5 \cdot 0{,}5(-0{,}5) = \frac{23}{16} = 1{,}438.$$

Приклад 5.11.3. За допомогою інтерполяції обчислити $e^{0{,}15}$, якщо

Таблиця 5.10

x_i	0.	0.1	0.2
e^{x_i}	1	1,10517	1,22140

Оцінити похибку.

Розв'язок. Позначимо $f(x) = e^x$. Тоді таблиця розділених різниць для даних в таблиці має вигляд

Таблиця 5.11

x_i	f_i		
0	1,0		
		1,0517	
0,1	1,105187		0,553
		1,1623	
0,2	1,22140		

Тому

$$L_2(x) = 1 + 1{,}0517x + 0{,}553x(x - 0{,}1),$$

$$L_2(0{,}15) = 1 + 0{,}10517 + 0{,}00415 = 1{,}16190.$$

Оцінка для похибки буде мати вигляд

$$|R_2(x)| \le \frac{M_3}{3!}|x(x-0{,}1)(x-0{,}2)|,$$

де $M_3 = \max\limits_{0 \le x \le 0{,}2} |f^{(3)}(x)| = e^{0{,}2}$. Тому

$$|R_2(0{,}15)| \le \frac{1{,}2214}{6} 0{,}15 \cdot 0{,}05 \cdot 0{,}05 = 7{,}63 \cdot 10^{-5}.$$

Приклад 5.11.4. За значеннями $\cos x$ в точках $x = 0,\ \pi/4,\ \pi/2$ знайти $\cos(\pi/5)$. Оцінити похибку.

Розв'язок. Таблиця розділених рівнянь має вигляд

Таблиця 5.12

x_i	f_i		
0	1,0		
		$\left(\sqrt{2}/2 - 1\right)\pi/4$	
$\pi/4$	$\sqrt{2}/2$		$\left(1 - \sqrt{2}\right)/(\pi/4)/(\pi/2)$
		$-\sqrt{2}/2/(\pi/4)$	
$\pi/2$	0		

Приклад 5.11.5. Скільки чебишевських вузлів інтерполяції необхідно вибрати, щоб похибка інтерполяції для функції $f(x)=\ln(1+x),\ x\in[0;1]$ не перевищувала $\varepsilon=10^{-4}$.

Розв'язок. З оцінки залишкового члена інтерполяційної формули маємо

$$|R_n(x)|\le\frac{M_{n+1}}{(n+1)!}|\varpi_n(x)|=\frac{M_{n+1}}{(n+1)!}\left|T_{n+1}^{[0,1]}(x)\right|\le\frac{M_{n+1}}{(n+1)!}\frac{1}{2^{2n+1}}.$$

Далі

$$f^{(n+1)}(x)=(-1)^{n+1}\frac{n!}{(1+x)^{n+1}}.$$

Тому n шукаємо з умови

$$|R_n(x)|\le\frac{1}{(n+1)2^{2n+1}}<10^{-4}.$$

Підстановкою знаходимо, що

$$|R_5(x)|\le\frac{1}{6\cdot2^{11}}\approx8\cdot10^{-5}<10^{-4}.$$

Таким чином треба вибрати 6 чебишевських вузлів.

Приклад 5.11.6. Дані по тиску водню на лінії насичення наведені в таблиці. Побудувати інтерполяційні багаточлени Лагранжа і Ньютона і оцінити їх похибку. Визначити тиск водню при температурі насичення T=12,5; T=15,4; T=17,7.

Таблиця 5.13

T,K	10	11	12	13	14	15	16	17	18
P,мм рт ст	1.93	5,62	13,9	30,2	58,8	100,3	161,1	246,0	360,3

Розв'язок. Значення тиску при температурі, яка дорівнює T=12.5K, отримане за допомогою інтерполяційного многочлена Лагранжа за формулою, складає

$$\mathbf{P}=\mathbf{P_8}(12.5)=\sum_{i=0}^{8}\frac{(12.5-x_0)(12.5-x_1)\ldots(12.5-x_{i-1})(12.5-x_{i+1})\ldots(12.5-x_n)}{(x_i-x_0)(x_i-x_1)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)}y_i=20.524$$

і за допомогою інтерполяційного многочлена Ньютона за формулою, де $\mathbf{q}=\frac{x-x_0}{h}=\frac{12.5-10}{1}=2.5$, а кінцеві різниці представлені в табл.7.1, також складає

$$P_8(12.5)=1.93+2.5\Delta y_0+\frac{2.5(2.5-1)}{2!}\Delta^2y_0+\frac{2.5(2.5-1)(2.5-2)}{3!}\Delta^3y_0+\ldots+$$

$$+\frac{2.5(2.5-1)(2.5-2)\ldots(2.5-8)}{8!}\Delta^8y_0=20.524\,мм.рт.ст.$$

Значення тиску при температурі, яка дорівнює T=15.4K, отримане за допомогою інтерполяційного многочлена Лагранжа за формулою, складає

$$\mathbf{P}=\mathbf{P_8}(15.4)=$$

$$=\sum_{i=0}^{8}\frac{(15.4-x_0)(15.4-x_1)\ldots(15.4-x_{i-1})(15.4-x_{i+1})\ldots(15.4-x_n)}{(x_i-x_0)(x_i-x_1)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)}y_i=121.847$$

і за допомогою інтерполяційного многочлена Ньютона за формулою, де $\mathbf{q} = \frac{x - x_0}{h} = \frac{15.4 - 10}{1} = 5.4$, а кінцеві різниці представлені в табл. 5.14, також складає

$$P_8(15.4) = 1.93 + 5.4\Delta y_0 + \frac{5.4(5.4-1)}{2!}\Delta^2 y_0 + \frac{5.4(5.4-1)(5.4-2)}{3!}\Delta^3 y_0 + \ldots +$$
$$+ \frac{5.4(5.4-1)(5.4-2)\ldots(5.4-8)}{8!}\Delta^8 y_0 = 121.847 мм.рт.ст.$$

Таблиця 5.14

Таблиця кінцевих різниць

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$	$\Delta^6 y$	$\Delta^7 y$	$\Delta^8 y$
10	1,93	3,69	4,59	3,43	0,85	-4,53	14,01	-30,89	57,27
11	5,62	8,28	8,02	4,28	-3,68	9,48	-16,88	26,38	
12	13,9	16,3	12,3	0,6	5,8	-7,4	9,5		
13	30,2	28,6	12,9	6,4	-1,6	2,1			
14	58,8	41,5	19,3	4,8	0,5				
15	100,3	60,8	24,1	5,3					
16	161,1	84,9	29,4						
17	246	114,3							
18	360,3								

Значення тиску при температурі рівної Т=17.7К, отримане за допомогою інтерполяційного многочлена Лагранжа за формулою, складає

$$\mathbf{P} = \mathbf{P_8}(17.7) =$$
$$= \sum_{i=0}^{8} \frac{(17.7 - x_0)(17.7 - x_1)\ldots(17.7 - x_{i-1})(17.7 - x_{i+1})\ldots(17.7 - x_n)}{(x_i - x_0)(x_i - x_1)\ldots(x_i - x_{i-1})(x_i - x_{i+1})\ldots(x_i - x_n)} y_i = 320.895$$

і за допомогою інтерполяційного многочлена Ньютона за формулою, де $\mathbf{q} = \frac{x - x_0}{h} = \frac{17.7 - 10}{1} = 7.7$, а кінцеві різниці представлені в табл. 5.14, також складає

$$P_8(17.7) = 1.93 + 7.7\Delta y_0 + \frac{7.7(7.7-1)}{2!}\Delta^2 y_0 + \frac{7.7(7.7-1)(7.7-2)}{3!}\Delta^3 y_0 + \ldots +$$
$$+ \frac{7.7(7.7-1)(7.7-2)\ldots(7.7-8)}{8!}\Delta^8 y_0 = 320.895 мм.рт.ст.$$

Для визначення залишкового члена скористаємося таблицею кінцевих різниць (табл. 5.14). Знайдемо залишковий член, який отримали при обчисленні першої заданої проміжної точки, тобто Т=12.5К. Для цього знайдемо q: $\mathbf{q} = \frac{x - x_0}{h} = \frac{12.5 - 10}{1} = 2.5$, а також з таблиці кінцевих різниць візьмемо Δ^8y=57,27 і підставимо у формулу (5.29):

$$R_n(12.5) \approx$$

$$\approx \frac{2.5(2.5-1)(2.5-2)(2.5-3)(2.5-4)(2.5-5)(2.5-6)(2.5-7)(2.5-8)}{9!} 57.27 \approx 0.048.$$

Таким засобом знайдемо залишковий член для двох інших заданих проміжних точок. Для точки Т=15.4К знаходимо $\mathbf{q} = \frac{15.4-10}{1} = \frac{15.4-10}{1} = 5.4$ та підставляємо у формулу (5.29), одержуємо:

$$R_n(15.4) \approx$$

$$\approx \frac{5.4(5.4-1)(5.4-2)(5.4-3)(5.4-4)(5.4-5)(5.4-6)(5.4-7)(5.4-8)}{9!} 57.27 \approx -0.043.$$

Для точки Т=17.7К знаходимо $\mathbf{q} = \frac{17.7-10}{1} = \frac{17.7-10}{1} = 7.7$ та підставляємо у формулу (5.19), одержуємо:

$$R_n(17.7) \approx$$

$$\approx \frac{7.7(7.7-1)(7.7-2)(7.7-3)(7.7-4)(7.7-5)(7.7-6)(7.7-7)(7.7-8)}{9!} 57.27 \approx -0.78.$$

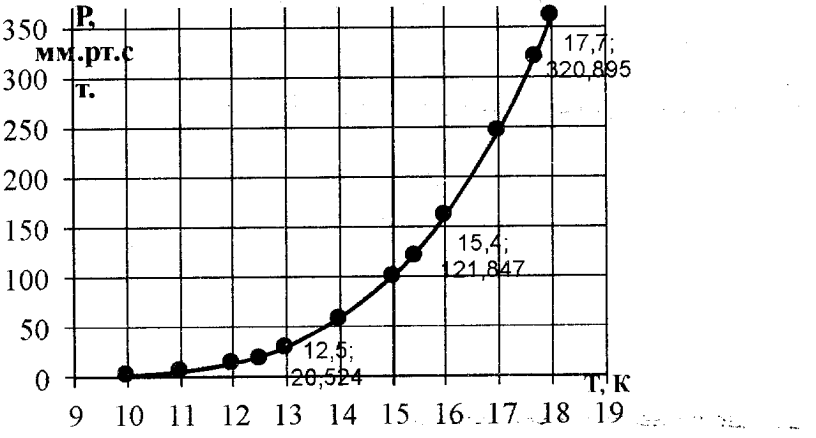


Рис. 5.7. Графічне зображення результатів приклада 5.6

На поданому графіку можна побачити вузли інтерполяції, а також проміжні точки, що мають підписи. Цей графік відображає ту залежність, що була подана у таблиці. Дивлячись на графік, видно, що знайдені значення функції входять у цю залежність, тобто розв'язок було знайдено вірно.

5.12. Контрольні питання

1. Що називається інтерполюванням та числовим диференціюванням?
2. Розкрити загальну постановку задачі інтерполювання.
3. Що називається задачею лінійного інтерполювання?
4. Розкрити постановку задачі інтерполяції функції однієї дійсної змінної
5. Що називається поліноміальною інтерполяцією? Розкрити існування та одиничність інтерполяційного полінома.
6. Що називається інтерполяційним поліномом Лагранжа?
7. Провести аналіз загальної постановки задачі для інтерполяційного полінома Лагранжа.
8. Розказати про Інтерполяційну формулу Лагранжа. Навести частковий випадок.
9. В чому полягають особливості побудови інтерполяційного многочлена Лагранжа?
10. Що називається інтерполяційний многочленом Лагранжа з рівновіддаленими вузлами?
11. Що називається інтерполяційною схемою О. Ейткена?
12. Як обчислюється похибка інтерполювання для інтерполяційного многочлена Лагранжа?
13. Як здійснюється вибір вузлів інтерполювання?
14. Що називаються многочленами Чебишева? Якими вони володіють властивостями?
15. Як змінюється поведінка похибки на інтервалі інтерполювання?
16. Що називається інтерполяційним поліномом Ньютона?
17. Провести аналіз загальної постановки задачі.
18. Як особливості інтерполяційного многочлена Ньютона?
19. Які є поділені різниці та їх властивості?
20. Розказати про інтерполяційну формулу Ньютона за поділеними різницями та нерівновіддаленими вузлами.
21. Що називаються скінченні різниці та які їх властивості?
22. Розказати про формулу Ньютона вперед і назад за скінченними різницями.
23. Що називається числовим диференціюванням?
24. Що називається числовим диференціюванням для нерівновіддалених вузлів?
25. Які є формули числового диференціювання для рівновіддалених вузлів?
26. Які є формули числового диференціювання методом невизначених коефіцієнтів?
27. Що називається раціональною інтерполяцією?
28. Розкрити особливості інтерполювання функцій комплексної змінної.
29. Розкрити особливості інтерполювання функцій багатьох змінних.
30. Які є труднощі інтерполювання функцій багатьох змінних?
31. Як будуються інтерполяційні формули за поділеними різницями?

32. Розкрити особливості збіжності інтерполяційного процесу.
33. Що називається інтерполюванням сплайнами?

5.13. Задачі для самостійної роботи

Задача 13.1. Нехай:

1) $\{\varphi_i(x)\} \subset C^{n+1}([a,b])$, $i = \overline{0,n}$.

2) Вронскіан

$$W[\varphi_0,\varphi_1,\ldots,\varphi_k] = \begin{vmatrix} \varphi_0(x) & \varphi_1(x) & \ldots & \varphi_k(x) \\ \varphi_0'(x) & \varphi_1'(x) & \ldots & \varphi_k'(x) \\ \ldots & \ldots & \ldots & \ldots \\ \varphi_0^{(k)}(x) & \varphi_1^{(k)}(x) & \ldots & \varphi_k^{(k)}(x) \end{vmatrix} \neq 0,\ k = \overline{0,n}$$

Довести, що $\{\varphi_i(x)\}$ - система функцій Чебишева.

Задача 13.2. (узагальнена теорема Ролля). Нехай виконуються умови 1) та 2) задачі 13.1, $f \in C^{n+1}([a,b])$. Довести, що тоді існує точка $\xi \in [a,b]$ така, що

$$L_{n+1}[f]|_{x=\xi} \equiv \left.\frac{W[\varphi_0,\ldots,\varphi_n,f]}{W[\varphi_0,\ldots,\varphi_n]}\right|_{x=\xi} = 0.$$

Задача 13.3. Одержання формули диференціювання Ньютона назад за скінченними різницями.

Задача 13.4. Якщо $P_n(x) = x^n + \ldots \neq T_{n(x)} \cdot 2^{1-n}$, то довести нерівність

$$\max_{[-1,1]} |P_n(x)| > 2^{1-n}.$$

Задача 13.5. У процесі обчислення $T_n(x) \equiv \cos(n \cdot \arccos x)$ отримано $T_0^* = 1$,

$$T_{n+1}^* = 2xT_N^* - T_{n-1}^* + \delta_{n+1},$$

де δ_k - похибка заокруглень. Довести

$$T_N^*(x) - T_N(x) = \sum_{k=1}^{n} \delta_k \frac{\sin((N+1-k)\arccos x}{\sqrt{1-x^2}}.$$

Задача 13.6. Отримати за умовою вправи 13.5 оцінку похибки

$$|T_N^*(x) - T_N(x)| \leq \max_{1\leq k\leq N} |\delta_k|\, N \min(N, \frac{1}{\sqrt{1-x^2}}).$$

Задача 13.7. Перевірити, що в нулях многочлена Чебишева

$$x_m = \cos\frac{(2m+1)\pi}{2N}$$

справедлива рівність $|T_N'(x_m)| = \dfrac{N}{\sqrt{1-x_m^2}}$.

Задача 13.8. Довести, що вектори, які утворені многочленами $\widetilde{T}_n = T_n(x)\cdot 2^{1-n}$, $n<N$, де N – фіксоване, в нулях $T_n(x)$ утворюють деяку ортогональну систему, а саме:

$$\frac{2}{N}\sum_{j=1}^{N}\widetilde{T}_n\left(\frac{\pi(2j+1)}{2N}\right)\cdot\widetilde{T}_m\left(\frac{\pi(2j+1)}{2N}\right)=\delta_n^m$$

для $0\le n,\ m<N-1$.

Задача 13.9. Використовуючи формулу задачі 13.8 показати, що многочлен з вузлами в нулях многочленів Чебишева записується у вигляді

$$L_n(x)=\sum_{j=0}^{n-1}a_jT_j(x)2^{1-n},\ a_j=\frac{2}{n}\sum_{k=1}^{n}f(x_j)T_j\left(\frac{\pi(2k+1)}{2n}\right)2^{1-n}.$$

Задача 13.10. Довести, що на [-1, 1] два многочлена Чебишова взаємно ортогональні з вагою $\rho(x)=\dfrac{1}{\sqrt{1-x^2}}$ (підказка $x=\cos\theta$).

Задача 13.11. Нехай $x_0<x_1<\ldots<x_n$, $m!f(x_i,\ldots,x_{i+m})=f^{(m)}(\xi_i)$, $x_i\le\xi_i$. Позначимо $M_m\equiv\sup\limits_{1\le k\le n-m}m!|f(x_k,\ldots,x_{k+m}|$, $M^{(m)}\equiv\sup\limits_{[x_0,x_n]}|f^{(m)}(x)|$. Довести, що $|M_m-M^{(m)}|\le M^{m+1}h_m\equiv\max\limits_{1\le k\le n-m}(x_{k+m}-x_k)$.

Задача 13.12. Нехай $f(x)$ інтерполюється за вузлами $x_j=(j-1)h$, $h=\dfrac{1}{(n-1)}$, $j=\overline{1,n-1}$. Довести, що $\max\limits_{0\le x\le1}\lambda(x)\ge const\dfrac{2^n}{\sqrt{n}}$, де $\lambda(x)=\sum\limits_{j=0}^{n}\Phi_j(x)$.

Задача 13.13. Довести, що якщо вузли інтерполювання співпадають з нулями Чебишева, то $\max\limits_{[-1,1]}|\lambda(x)|\le const\cdot\ln n$.

Задача 13.14. Довести, що для лінійної інтерполяції похибка. Яка є наслідком заокруглення значень функцій, не перевищує похибки цих значень.

Задача 13.15. Якщо x і t пов'язані співвідношеннями $x=\alpha t+\beta\quad(\alpha\ne0)$, то $f(x_i,\ldots,x_{i+k})=\dfrac{1}{\alpha^k}g(t_i,\ldots,t_{i+k})$, де $g(t)=f(\alpha t+\beta)$, $t_j=\dfrac{x_j-\beta}{\alpha}$.

Задача 13.16. Довести, що поділені різниці порядку n від многочлена степеня n є константами, а різниці більш високого степеня дорівнюють нулю, тобто якщо

$$P_l(x)=\sum_{j=0}^{l}b_jx^j,\ l\le n,\ P_l(x_0,x_1,\ldots,x_n)=\begin{cases}b_n, & l=n\\ 0, & l<n\end{cases}\ \text{для}\ \forall x_0,x_1,\ldots,x_n\in[a,b].$$

5.14. Завдання до лабораторної роботи

Методом Лагранжа та методом Ньютона апроксимувати функцію, задану таблично. Кількість точок апроксимації дорівнює шість. Абсциси точок для всіх варіантів прийняти рівними: 0; 0.5; 1; 2; 3.5; 4; 6. Масив ординат представлений в табл. 5.15. Побудувати графік отриманої функції Лагранжа, на графіку відзначити задані точки апроксимації. Знайти значення функції Лагранжа та функції Ньютона при x=5. Порівняти результати. Провести перевірку результати, знайшовши похибки.

Таблиця 5.15

№	1 точка	2 точка	3 точка	4 точка	5 точка	6 точка	7 точка
1.	1.213	2.125	-1.345	0.001	5.756	9.019	11.345
2.	12.234	9.239	8.567	-1.098	5.756	7.345	5.678
3.	1.947	0.345	-2.456	-4.569	-1.941	3.938	6.709
4.	23.391	12.562	9.268	12.981	16.528	20.284	19.318
5.	-2.371	-4.592	-6.937	-3.297	2.895	7.981	11.625
6.	7.382	9.271	23.744	34.561	22.128	14.795	15.345
7.	34.589	45.871	67.081	53.901	22.456	12.459	20.017
8.	-23.016	-34.802	-27.391	-9.701	0.249	15.23	13.906
9.	23.691	12.459	21.345	34.671	61.091	72.321	113.935
10.	-34.469	-56.207	-63.208	-50.109	-10.349	0.381	12.391
11.	12.619	7.380	3.018	0.682	3.947	15.378	10.026
12.	-30.492	-19.014	-9.012	1.309	-2.305	-6.826	-15.271
13.	173.601	121.039	97.391	56.729	19.823	56.501	43.467
14.	46.701	67.204	84.291	56.390	47.713	35.716	43.503
15.	-372.09	-281.17	-101.28	-54.801	-67.91	-123.41	-93.173
16.	-17.801	-34.812	-45.839	-28.917	-1.572	15.571	9.348
17.	81.351	56.728	14.824	29.381	45.727	33.628	19.429
18.	-15.351	1.396	34.206	82.917	57.901	98.128	123.178
19.	7.238	-11.817	-27.924	-23.581	-12.482	2.381	27.906
20.	71.082	83.915	88.801	101.451	74.120	56.297	60.168
21.	5.781	-12.836	-45.817	-98.104	-56.927	-23.752	0.129
22.	678.237	561.908	482.931	394.781	501.937	643.183	754.928
23.	-56.392	-44.906	-28.504	-0.581	-12.948	-31.736	-67.938
24.	77.031	48.406	32.731	16.072	-0.837	6.935	28.052
25.	141.367	192.549	145.671	98.238	56.719	101.361	120.385
26.	-141.27	-83.148	-34.301	-23.906	-31.391	-12.471	-3.518
27.	2.417	7.517	12.471	6.329	5.715	9.271	15.371
28.	37.257	72.167	92.152	67.403	24.610	38.208	34.928
29.	-3.719	-7.108	-10.281	-4.201	5.921	14.072	11.492
30.	91.302	139.391	103.491	87.201	78.037	98.983	102.491

№	1 точка	2 точка	3 точка	4 точка	5 точка	6 точка	7 точка
31.	-7.303	-34.801	-57.928	-89.003	-76.938	-67.940	-45.929
32.	82.914	102.491	172.301	281.908	202.791	159.072	187.951
33.	-3.037	9.281	23.972	23.027	8.238	12.731	18.271

Розділ 6.
Основні відомості про методи апроксимації

- Формулювання проблеми апроксимації даних
- Формулювання задачі наближення функцій
- Апроксимація методом найменших квадратів
- Поняття сплайн-апроксимації
  - Аналітичне подання сплайнів
  - Апроксимація кубічними сплайнами
- Розв'язування задач рівномірного сплайн-наближення
  - Означення та властивості рівномірного сплайн-наближення
  - Однорідне сплайн-наближення з заданою похибкою
  - Побудова неоднорідне сплайн-наближення
- Методи обчислення рівномірного (мінімаксного) наближення функцій й
  - Загальна постановка задачі найкращого мінімаксного наближення функцій
  - Апроксимаційні залежності
  - Доведення про найкраще рівномірне наближення функцій
- Метод Є.Я.Ремеза
  - Загальна постановка задачі схеми Є.Я.Ремеза
  - Мінімакс. Визначення параметрів мінімаксимальної моделі за алгоритмом Є.Я.Ремеза
- Метод Валле-Пуссена
  - Загальна постановка задачі
  - Алгоритм Валле-Пуссена
- Застосування методу найменших квадратів багатовимірного регресійного аналізу в економіці
  - Загальна постановка задачі
  - Знаходження параметрів лінійного множинного рівняння регресії методом найменших квадратів
  - Стандартна похибка оцінки за рівнянням
  - Коефіцієнт детермінації й кореляції
  - Вибіркова похибка коефіцієнта множинної регресії
  - Вибіркова похибка множинної регресії
  - Похибка індивідуальної оцінки множинної регресії
  - Вибіркова похибка коефіцієнта множинної кореляції

- Часткова регресія та кореляція
- Приклад розрахунку потреб ринку на основі множинних рівнянь регресії
- Приклад розрахунку потреб ринку на основі економічної оцінки
- Приклад дослідження на основі багатовимірного регресійного аналізу

- Контрольні питання
- Задачі для самостійної роботи
- Завдання до лабораторної роботи

Математики минулого поєднували в своїх роботах вивчення явищ природи, отримували їх математичний опис (математичні моделі) та досліджували їх. На сьогоднішній день характерним є розширення додатків математики, що пов'язане з бурхливим розвитком обчислювальної техніки, адже швидкість операцій розрахунків стала просто фантастичною. Проте, думка про всемогутність сучасних обчислювальних машин породжує міф, що математики позбулися всіх проблем, пов'язаних як з чисельним розв'язком задач, так і з розробкою нових методів їх розв'язування. Але це не так. Можемо спостерігати математизацію інших розділів науки: хімії, економіки, біології, психології, медицини та конкретних розділів техніки. Процес математизації полягає в побудові математичних моделей процесів і явищ, а також в розробці нових методів їх дослідження та моделювання [1-64, 66-70, 81-87, 89-91, 94-105, 111, 115, 138-139, 146, 171-191, 195, 212-230].

В фізиці чи механіці, побудова математичних моделей для опису різних явищ природи і вивчення цих моделей з метою пояснення старих чи передбачення нових ефектів явищ є традиційною. Прикладом успішного вивчення математичних моделей в поєднанні з опрацюванням результатів спостережень, задовго до появи обчислювальної техніки, було відкриття невідомої колись планети – Нептун. Але робота в цьому напрямку просувалась повільно, бо не вдавалось отримати розв'язок певних математичних задач і доводилось обмежуватись розглядом більш простіших моделей. Застосування обчислювальної техніки значно збільшило можливості в напрямку побудови і дослідження математичних моделей. Все частіше результати розрахунків дозволяють знаходити і передбачувати нові явища, що дає можливість говорити про математичний експеримент. В деяких напрямках довіра до результатів чисельних розрахунків така велика, що при розбіжності між результатами розрахунків та експериментами в першу чергу шукають помилку в результатах експериментів. Сучасні успіхи у вирішенні таких важливих для суспільства задач як атомні, космічні, економічні, не були б можливі без застосування сучасної обчислювальної техніки, класичних та новітніх методів обчислень [65, 71, 76-80, 88, 92-93, 106-110, 113-114, 116-133, 137, 140-145, 147-151, 161-164, 168-169, 192-194, 196-197, 201-209, 211].

Потреба чисельного розв'зку нових задач призвела до появи великої кількості нових методів. Теоретичні дослідження в галузі чисельних методів групуються

навколо чисельних методів розв'зку так званих типових математичних задач, до яких відносять і апроксимацію (наближення) даних [72-74, 159, 212, 220].

Необхідність в апроксимації (наближенні) функцій виникає при проведенні теоретичних досліджень і практичних розрахунків в багатьох областях техніки. Для функції, яка задана у вигляді таблиці, це перш за все пов'язано з бажанням отримати аналітичний вираз, так як його простіше дослідити та інтерпритувати. У випадку аналітично заданої функції часто корисно замінити дуже складний вираз більш простим так, щоб основні властивості функції зберігалися.

Окремий інтерес представляють наближення функцій, які є придатними для розрахунків із застосуванням обчислювальної техніки. В цьому випадку апроксимуюча функція може бути отримана, наприклад, у вигляді полінома, раціонального полінома нелінійного виду, рекурентної (ітераційної) формули, популярною стала сплайн-апроксимація, тощо.

Задача знаходження найкращого (чебишовського) рівномірного наближення виникає, наприклад, для:

- проектування контрольно-вимірювальних приладів, зокрема пірометрів або термометрів (градуювання, лінеаризація статичних характеристик);
- проектування різноманітних функціональних перетворювачів (опис відповідних функцій);
- реалізація алгоритмів обчислення спеціальних функцій на високопродуктивних процесорах опрацювання даних (апроксимація складних функціональних залежностей простими аналітичними виразами);
- побудова математичних моделей різноманітних неперервних процесів на основі дискретних вимірювань їх значень;
- в машинобудуванні для представлення контурів проекцій деталей з складними геометричними формами, а також аналітичного подання кривих, які задають траєкторію різальних інструментів в верстатах із програмним управлінням, тощо.

Крім того, апроксимацію дослідних даних, доцільно застосовувати у різноманітних задачах ущільнення масивів результатів вимірювань, що містять методичну похибку у відомих межах (це може бути телеметрична інформація або масиви геофізичних чи сейсмічних вимірювань).

Поряд з класичними поліномами й раціональними найкращими рівномірними наближеннями, все частіше розглядається наближення нелінійними виразами. Це пояснюється тим, що фізичні процеси здебільшого мають нелінійний характер і для більш адекватного їх представлення використовуються відповідні отримані теоретично або інтуїтивно побудовані нелінійні залежності. Вивчаються також можливості використання рівномірного наближення функцій для представлення розв'язків диференційних та інтегральних рівнянь, розглядаються й інші можливості застосування такого наближення.

В багатьох технічних задачах застовується сплайн-апроксимація функцій. При цьому необхідно, щоб абсолютна і відносна похибки ні в одній точці не перевищували наперед заданої величини. Це призводить до задачі знаходження

параметрів чебишовського наближення. Якщо на кожному інтервалі (ланці сплайну) максимальна похибка наближення однакова, то таке наближення називається рівномірним сплайн-наближенням. У відомих випадках рівномірне наближення сплайном є оптимальним в тому сенсі, що при виборі границь ланок (вузлів) із умови рівномірності наближення при заданій кількості ланок отримуємо мінімальну похибку, а при заданій похибці – мінімальну кількість ланок [76-80, 88, 92-93].

Таким чином, поруч із наявним програмним забезпеченням (Maple, MathLab, тощо), актуальною є розробка прикладного забезпечення для персональних комп'ютерів, яке є орієнтоване на розв'язування задач, що зводяться до знаходження наближення функцій. Таке програмне забезпечення для персональних комп'ютерів повинно грунтуватися на аналізі задач з наближення функцій і включати нові теоретичні й практичні розробки [76-80, 88, 92-93].

6.1. Формулювання проблеми апроксимації даних

Проблема задачі апроксимації функцій виникає в найрізноманітніших технічних задачах. Різноманіття методів настільки велике, що іноді виникає наступне питання. Може наявність великої кількості різних методів наближення пояснюється простою відсутністю наукового підходу до постановки і розв'язування проблеми? Якби такий підхід існував, то може вдалось би запропонувати один оптимальний спосіб наближення? Слід визнати, що вибір методів наближення викликано різноманіттям постановок проблем. Різні розділи теорії наближень, наприклад інтерполювання, апроксимації, можна розглядати як вивчення поведінки абстрактних моделей.

Функції для опису математичних моделей можуть бути задані як аналітичним способом (формулою), так і табличним, при якому функція відома тільки при визначених дискретних значеннях агрументу. Якщо функціональні залежності отримані в результаті розрахунків, які проведені за допомогою обчислювальної техніки або в процесі експериментів, то вони задаються саме табличним способом. Імовірно, що на практиці виникне потреба в інформації про значення функції в інших точках, відмінних від тих, що наведені в таблиці. І отримати такі значення можна шляхом складних розрахунків або проведенням ще ряду експериментів, які заберуть багато часу і ресурсів. Таким чином, для економії часу і коштів, доцільним є застосування різних задач апроксимації функцій при довільному значенні аргумента на основі наявних табличних даних.

Якщо відтворення значення функції відбувається за точними значеннями то застосовують ***інтерполяцію***.

Така задача розв'язується, наприклад, заміною функції $f(x)$ більш простішою функцією $\varphi(x)$, яку нескладно знайти при довільному значенні аргументу x на заданому інтервалі. Наближення функції $f(x)$ більш простішою функцією називають *апроксимацією* (від латинської ***approximo*** – наближатися). Апроксимуючу функцію $\varphi(x)$ будують таким чином, щоб відхилення від $f(x)$ на заданому проміжку було мінімальним. При цьому функцию $\varphi(x)$ будуть називати ***апроксимуючою*** (рис. 6.1).

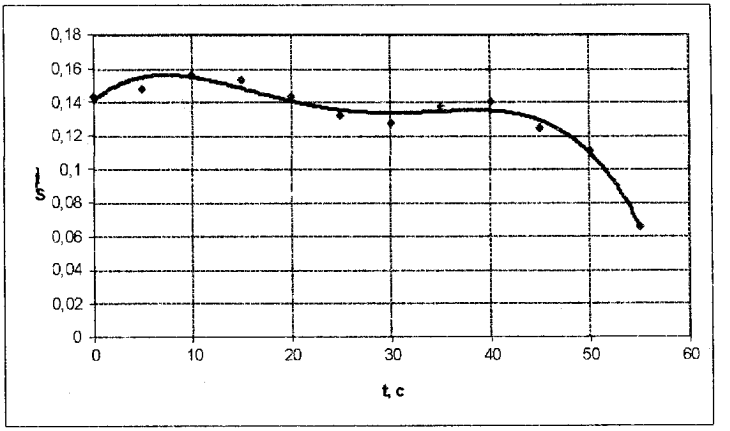


Рис. 6.1. $\varphi(x)$ – апроксимуюча функція

Розв'язуючи задачі апроксимації, потрібно враховувати:

- вибір вузлів апроксимації, формування початкового вузла x_0, вибір постійного кроку h. розрахунок усіх наступних вузлів (наприклад, якщо $x_0 = -0{,}3$, $h = 0{,}5$, то $x = (-0{,}3;\ 0{,}2;\ 0{,}7;\ldots)$; при цьому з метою зменшення похибки апроксимації вузли обчислюються за спеціальними формулами;
- підбір апроксимуючої функції $\varphi(x)$, наприклад поліномом

$$\varphi(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n . \qquad (6.1)$$

В такому випадку апроксимація буде називатися ***поліноміальною***. Коефіцієнти поліному $a_0, a_1, \ldots, a_n$ і його степінь n вибираються таким чином, щоб забезпечити найменше відхилення апроксимуючої функції від заданої. Крім того, наближення функцій може бути тригонометричним, експоненційними, логарифмічними та іншими виразами.

Якщо функція $f(x)$ задається аналітичним виразом, то при побудові апроксимуючої функції потрібно дотримуватися мінімального відхилення однієї функції від іншої на деякій неперервній множині точок $[a,b]$. Така апроксимація буде називатися ***неперервною*** або ***інтегральною***. Теоретично для найвдалішого наближення потрібно, щоб похибка на усіх точках відрізку відхилення апроксимуючої функції $\varphi(x)$ від функції $f(x)$ була меншою від заданої величини $\varepsilon>0$

$$|f(x) - \varphi(x)| < \varepsilon\ ,\ a \le x \le b\ . \qquad (6.2)$$

В такому випадку кажуть, що функція $\varphi(x)$ ***рівномірно наближує*** функцію $f(x)$ з точністю ε на відрізку $[a,b]$.

Також, дуже часто застосовують метод ***середньоквадратичного наближення***, для якого найменше значення має величина

$$M = \int_a^b [f(x) - \varphi(x)]^2 dx . \qquad (6.3)$$

Для отримання точкового середньоквадратичного наближення функції $f(x)$, яка задається таблично, апроксимуючу функцію $\varphi(x)$ будують виходячи з умови мінімуму величини:

$$S = \sum_{i=0}^{n} [y_i - \varphi(x_i)]^2 , \qquad (6.4)$$

де y_i – значення функции $f(x)$ в точках x_i.

Величина S являє собой суму квадратів відхилень апроксимируючої функції $\varphi(x)$ від $f(x)$ на заданій множині точок $x_0, x_1, \ldots, x_n$. У випадку наближення поліномом, коефіцієнти $a_0, a_1, \ldots, a_m$ підбираються, виходячи з умови, що величина S є мінімальною. Такий підхід до апроксимації називається ***методом найменших квадратів.*** В цьому випадку не обов'язково, щоб апроксимуюча функція проходила через вузли. Недоліком середньоквадратичного наближення є те, що функції $\varphi(x)$ та $f(x)$ в окремих точках можуть суттєво відрізнятися. Найчастіше средньоквадратичне наближення застосовується для опрацювання експериментальних даних.

Ще одним способом точкової апроксимації є ***інтерполювання***, за яким апроксимуюча функція приймає в заданих точках x_i ті ж значення y_i, що і функція $f(x)$, тобто: $\varphi(x_i) = y_i$, $i=0,1,\ldots, n$; при цьому точки x_i називаються ***вузлами інтерполяції***. На рис. 6.2а і рис. 6.2б показано графік інтерполяційної функції і результати середньоквадратичного наближення. Точками на графіку марковані табличні значення функції $f(x)$

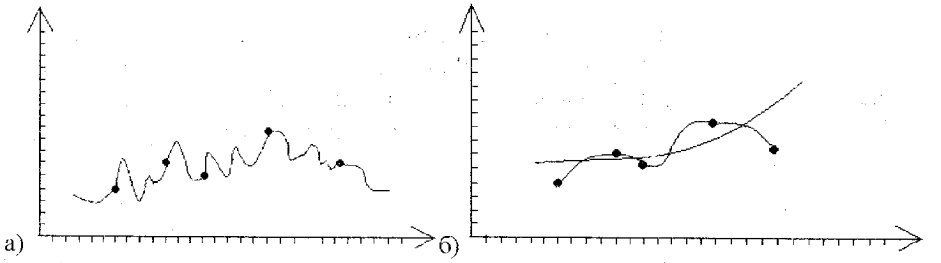


Рис. 6.2. а) Приклад інтерполювання та б) середньоквадратичне наближення

6.2. Формулювання задачі наближення функцій

Постановка задачі наближення функцій виникає в багатьох ситуаціях. Різні розділи теорії наближень, а саме інтерполяції, можна розглядати як вивчення абстрактних моделей деяких реальних класів проблем.

а) Найпростіша задача, що приводить до наближення функцій полягає в наступному. В дискретні моменти часу $x_1, \ldots x_n$ спостерігається значення функції $f(x)$. Потрібно відновити її значення при інших x. Аналогічна задача виникає в

наступному: в ході обчислень в комп'ютерних розрахунках доводиться багаторазово обчислювати одну і ту ж складну функцію в різних точках. Замість її безпосереднього розрахунку інколи доцільно обчислювати її значення в окремих точках, а значення в інших точках обчислювати, використовуючи інформацію про відомі точки.

Деколи з певних додаткових міркувань відомо, що функцію, яка наближує, доцільно шукати в вигляді:

$$F(x) = g(x; a_1, \ldots, a_n). \tag{6.5}$$

Якщо параметри $a_1, \ldots a_n$ визначаються із умови співпадіння $f(x)$ і наближаючої функції в точках $x_1, \ldots x_n$

$$G(x_j; a_1, \ldots, a_n) = f(x_j), \quad j = \overline{1,n}. \tag{6.6}$$

то такий спосіб наближення називають ***інтерполяцією***.

Якщо a лежить за межами проміжку інтерполювання, то таке наближення називається **екстраполяцією**.

б) Часто буває відомо, що функція добре наближається функціями певного вигляду, наприклад многочленами, але не зрозуміло, як краще всього вибрати степінь многочлена, що наближує. Проблема ускладнюється, коли задані значення функції містять великі похибки.

в) В задачах планування експериментів виникає наступна проблема. Відомий вигляд функції, що добре наближує, наприклад, функція добре наближується многочленом другого степеня. В той же час значення функції, що вимірюються, містять великі похибки. Потрібно отримати найкраще наближення при мінімальному числі вимірюваних значень функції. Така задача виникає при плануванні експериментів в біології, хімії, фізиці, географії тощо.

Торкнемося докладніше планування експериментів на хімічних підприємствах. Наприклад, сорт капрону визначається цілим рядом параметрів, частина яких залежить від властивостей сировини, яка поступає, а частина піддається управлінню; потрібно максимально підвищити випуск капрону першого сорту. Для цього потрібно мати в розпорядженні функцію, що відображає залежність виходу продукції від всіх цих параметрів. Розглянемо випадок, коли ця функція є многочленом другого степеня від 10 параметрів. Такий многочлен містить 66 коефіцієнтів, тому для їх знаходження потрібно щонайменше 66 вимірювань. Кожне вимірювання містить суттєву випадкову похибку через наявність ряду неконтрольованих факторів і факторів, що не вимірюються; тому для пониження впливу випадковості необхідно збільшити число вимірювань. В той же час кожне вимірювання обходиться відносно дорого, бо потребує зміни режиму роботи реального підприємства, до того ж такі вимірювання часто потребують тривалого проміжку часу, тому що підприємство повинно вийти на стаціонарний режим роботи. В результаті аналізу всіх цих протирічь і виникла проблема наближення функцій, про яку йшла мов вище.

г) Задача наближення з'являється при написанні стандартних програм розрахунку елементарних і спеціальних функцій. Звичайно такі функції володіють спеціальними властивостями, які дозволяють значно зменшити обсяг розрахунків.

Проблема, що виникає, може бути сформульована так. Розглядаються всі функції $g(x)$, результати розрахунків яких записуються в k комірках комп'ютора, такі, що деяка норма похибки $\|f-g\|$ не перевищить ε. Зі всіх таких функцій слід вибрати таку, розрахунок якої потребує мінімальних затрат часу.

В залежності від обставин норма може вибиратися по різному. В переважній більшості випадків береться $\|f\| = \sup|f|$ на проміжку $[a,b]$, на якому наближається функція. Досить часто потрібно підвищити точність в окремих точках. Наприклад, одна зі стандартних програм $\sin(x)$ забезпечує похибку в нормі:

$$\|f\| = \sup_{[0,\pi/2]} |p(x)f(x)|, \quad p(x) = \min(10^{19}, x^{-1}). \tag{6.7}$$

Введення множника $p(x)$ викликане вимогою мінімуму відносно похибки $\sin(x)$ при малих x. Таким чином можуть по-різному тлумачитися вимоги мінімальності затрат часу ПК. Затрати можуть залежати від точки, в якій вираховується значення функції. Позначимо їх через $t(x)$. Якщо відсутня інформація про те, з якою частотою вираховуються значення функції в тих чи інших частинах відрізка, то доцільно в якості спільної міри затрат прийняти:

$$T = \sup_x t(x). \tag{6.8}$$

Якщо така інформація є, то можна прийняти:

$$T = Mt(x). \tag{6.9}$$

д) Якщо функція задається графіком або складним аналітичним виразом, то виникають варіаційні задачі інших типів. Наприклад, нехай вирішено розбити відрізок $[a,b]$ на l частин:

$$[a_{i-1}, a_i], \ i = 1,\ldots,l, \ a_0 = a, \ a_n = b$$

і на кожному відрізку $[a_{i-1}, a_i]$ наближають функцію многочленом

$$f(x) = g(x) = \sum a_k^i x^k$$

або дробово–раціональною функцію

$$f(x) = g(x) = \sum a_k^i x^k \Big/ \sum b_k^i x^k. \tag{6.10}$$

Серед таких способів наближення шукають оптимальне в тому чи іншому сенсі. Найчастіше наперед накладаються умови $n_1 = \ldots = n_l$, $m_1 = \ldots = m_l$, фіксується число відрізків розбиття l і проводиться оптимізація методу по $a_1,\ldots,a_{l-1}$, n і m.

В окремому випадку, коли l=1, виникає задача найкращого рівномірного наближення многочленами чи дробово–раціональними функціями на фіксованому відрізку.

е) Вигляд наближаючої функції суттєво залежить від мети, з якою здійснюється наближення. Припустимо, що з деякою точністю функція може бути наближена многочленом 10-го степеня або виразом $a_1 \sin w_1 x + a_2 \sin w_2 x$. Якщо отримане наближення використовують в теоретичних дослідженнях для розв'язку задачі на моделюючому пристрої чи в технологічному процесі, то друга форма запису є більш зручною. Але якщо значення функції вираховується на ПК, то друга форма запису потребує при своїй реалізації більшого числа арифметичних операцій [76-80, 88, 92-93].

6.3. Апроксимація методом найменших квадратів

Нехай у результаті експерименту одержано залежність, зображену у табл. 6.1.

Таблиця 6.1

Вихідні дані

t	t_0	t_1	t_2	...	t_k	...	t_n
y^e	$y_0{}^e$	$y_1{}^e$	$y_2{}^e$	...	$y_k{}^e$	...	$y_n{}^e$

Необхідно знайти аналітичну формулу $y = y(t)$, яка апроксимує експериментальну (табличну) залежність. Виберемо залежність $y(t)$ у вигляді поліному другого ступеня, тобто

$$y(t) = a_0 + a_1 \cdot t + a_2 \cdot t^2 . \qquad (6.11)$$

У виразі (6.11) коефіцієнти a_0, a_1, a_2 підлягають визначенню, причому ці коефіцієнти повинні бути підібрані таким чином, щоб залежність $y(t)$ найкраще наближала експериментальну залежність. Назвемо ***відхиленням*** σ відмінність між табличним значенням $y(t)$ у точці t_i і значенням у тій самій точці, тобто

$$\sigma_i = y^e(t_i) - y(t_i) = [\, y^e(t_i) - a_0 - a_1 \cdot t_i - a_2 \cdot t_i{}^2 \,] . \qquad (6.12)$$

Згідно з методом найменших квадратів (МНК) "найкращими" коефіцієнтами залежності (6.11) будуть ті, для яких сума квадратів відхилень буде мінімальна, тобто

$$\Phi(a_0, a_1, a_2) = \sum \sigma_i^2 \to \min . \qquad (6.13)$$

Використовуючи необхідні умови існування екстремуму для функції декількох змінних $\Phi(a_0, a_1, a_2)$, знайдемо рівняння для визначення коефіцієнтів залежності (6.11):

$$\frac{\partial \Phi}{\partial a_0} = 0; \quad \frac{\partial \Phi}{\partial a_1} = 0; \quad \frac{\partial \Phi}{\partial a_2} = 0. \tag{6.14}$$

З умов (6.14) одержимо нормальну систему лінійних алгебраїчних рівнянь

$$\begin{cases} a_0 \cdot \sum\limits_k t_k^0 + a_1 \cdot \sum\limits_k t_k^1 + a_2 \cdot \sum\limits_k t_k^2 = \sum\limits_k t_k^0 \cdot y^e{}_k; \\ a_0 \cdot \sum\limits_k t_k^1 + a_1 \cdot \sum\limits_k t_k^2 + a_2 \cdot \sum\limits_k t_k^3 = \sum\limits_k t_k^1 \cdot y^e{}_k; \\ a_0 \cdot \sum\limits_k t_k^2 + a_1 \cdot \sum\limits_k t_k^3 + a_2 \cdot \sum\limits_k t_k^4 = \sum\limits_k t_k^2 \cdot y^e{}_k. \end{cases} \tag{6.15}$$

Розв'язавши систему (6.15), знайдемо коефіцієнти a_0, a_1, a_2 апроксимуючої залежності (6.11).

Приклад 6.1. Залежність температури досяжного перегріву рідин T_s для різних значень тиску P подана у табл. 6.2.

Таблиця 6.2

Залежність температури досяжного перегріву рідин T_s для різних значень тиску P

T_s,^{0}C	36,5	78,3	100,1	114,0	122,8	131,5	138,2
P, МПа	0,2	0,98	2,00	3,04	3,92	5,00	6,00

Необхідно узагальнити експериментальні дані у вигляді аналітичної залежності $P=f(T)$.

Розв'язок. Для проведення аналізу вихідних даних з метою вибору вигляду апроксимуючого многочлена подамо у вигляді графіку експериментальні дані з табл. 6.2. Графік наведений на рис. 6.3. У результаті аналізу даних виберемо за апроксимуючий многочлен параболу, яка задана залежністю

$$P_2(x) = a_0 + a_1 x + a_2 x^2. \tag{6.16}$$

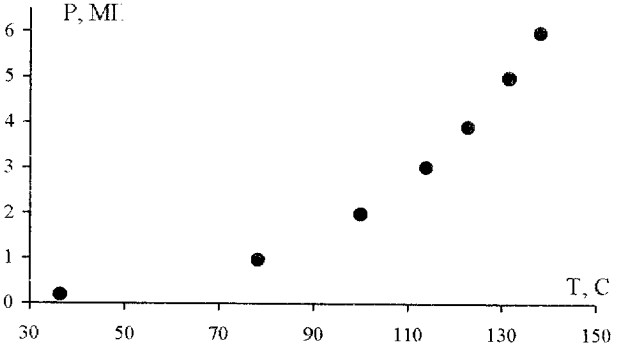


Рис. 6.3. Експериментальна залежність $P=f(T)$ прикладу 6.1.

Для спрощення обчислень зробимо наступну заміну

$$y_i = P_2(x_i); \quad x_i = \frac{T_i}{100}. \tag{6.17}$$

Для визначення коефіцієнтів a_0, a_1, a_2 необхідно записати систему рівнянь (6.15). Для складання системи зручно скористатися даними, наведеними в табл. 6.3.

Таблиця 6.3

Допоміжні дані до складання системи лінійних алгебраїчних рівнянь для визначення коефіцієнтів параболічної апроксімації

i	T_i	X_i	Y_i	X_iY_i	X_i^2	$X_i^2Y_i$	X_i^3	X_i^4
0	36,5	0,365	0,20	0,0730	0,13323	0,02665	0,04863	0,01775
1	78,3	0,783	0,98	0,7673	0,61309	0,60083	0,48004	0,37587
2	100,1	1,001	2,00	2,0020	1,00200	2,00400	1,00300	1,00401
3	114,0	1,140	3,04	3,4656	1,29960	3,95078	1,48154	1,68896
4	122,8	1,228	3,92	4,8138	1,50798	5,91128	1,85179	2,27401
5	131,5	1,315	5,00	6,5750	1,72920	8,64613	2,27393	2,99020
6	138,2	1,382	6,00	8,2920	1,90992	11,45954	2,63951	3,64781
N=7	Σ	7,214	21,14	25,9887	8,19804	32,59921	9,77846	11,99862

Використовуючи дані, які наведені у останньому рядку табл. 6.3, систему рівнянь (6.15) запишемо у вигляді

$$\begin{cases} 7a_0 + 7{,}214 \cdot a_1 + 8{,}19504 \cdot a_2 = 21{,}14; \\ 7{,}214 \cdot a_0 + 8{,}19504 \cdot a_1 + 9{,}77846 \cdot a_2 = 25{,}9887; \\ 8{,}19504 \cdot a_0 + 9{,}77846 \cdot a_1 + 11{,}99862 \cdot a_2 = 32{,}59921. \end{cases} \tag{6.18}$$

У результаті розв'язування системи (6.18) одержимо наступні значення коефіцієнтів a_0=1,82743; a_1=-6,89062; a_2=7,08440. Отже, шуканий апроксимуючий многочлен має вигляд $y = 1{,}82743\text{-}6{,}89062 \cdot x + 7{,}08440 \cdot x^2$. За допомогою формул (6.16) перейдемо до вихідних позначень та одержимо

$$P = 1{,}827\text{-}6{,}891 \frac{T}{100} + 7{,}084 \frac{T^2}{10000}.$$

Після спрощення виразу

$$P = 1{,}827\text{-}0{,}6891 \cdot T + 0{,}0007084 \cdot T^2 \tag{6.19}$$

Отримана аналітична залежність (6.19) узагальнює експериментальні дані табл. 6.2. Для оцінки похибки апроксимуючої залежності складемо таблицю значень P. Для цього визначимо тиск P за формулою (6.19). Результати занесемо до табл.6.4.

Таблиця 6.4

Розрахункові значення P

T	36,5	78,3	100,1	114	122,8	131,5	138,2
P	0,256	0,775	1,953	3,179	4,049	5,072	5,835

Для оцінки точності параболічної апроксимації необхідно порівняти значення P з табл.6.1 і табл.6.4. Модуль різниці відповідних значень дає ΔP-похибку апроксимації, значення якої подані в табл. 6.5. У таблиці наведена також відносна похибка δP, яка дорівнює відношенню ΔP до P.

Таблиця 6.5

Абсолютна та відносна похибки апроксімації

T	36,5	78,3	100,1	114	122,8	131,5	138,2
ΔP	0,056	0,205	0,047	0,139	0,129	0,072	0,165
$\delta P, \%$	28	21	2,4	4,6	3,3	1,4	2,8

Порівняльний аналіз похибок показує, що отримана аналітична залежність задовільно узагальнює вихідні експериментальні дані.

Для інтегральної оцінки апроксимації можна скористати формулу

$$\delta = \sqrt{\frac{\sum_{i=0}^{n} \sigma_i^2}{n+1}}$$

На рис. 6.4 наведені два графіки, один із яких побудований за даними апроксимації (табл. 6.4), а другий – за вихідними даними (табл. 6.1).

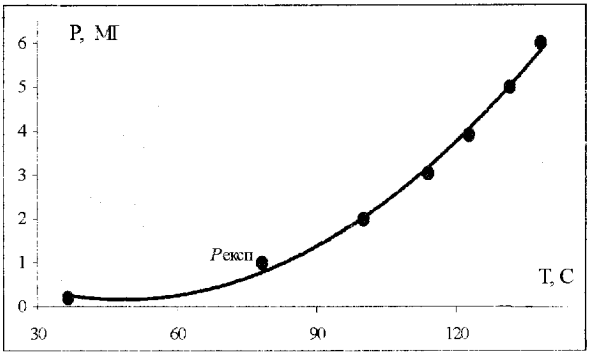


Рис. 6.4. Крива апроксимації і вихідні дані прикладу 6.1.

Порівнюючи ці графіки, можна також відзначити задовільну збіжність теоретичних та експериментальних даних.

Приклад 6.2. Залежність теплоємності Cp фториду магнію від температури T подано в табл. 6.6. Необхідно апроксимувати ці дані многочленом і оцінити похибку апроксимації.

Таблиця 6.6

Вихідні дані

T	300	400	500	600	700	800	900	1000
Cp	70,35	75,38	80,53	85,81	91,26	96,83	102,53	108,27

Розв'язок. Для вибору апроксимуючого многочлена доцільно побудувати графік залежності $Cp=f(T)$ і проаналізувати його (рис. 6.5).

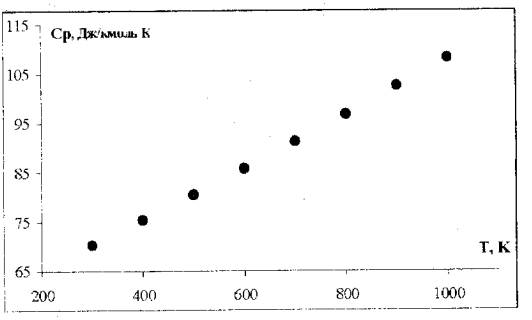


Рис. 6.5. Вихідні дані прикладу 6.2

Апроксимуємо дану табличну залежність многочленом першого ступеню (лінійна апроксімація)

$$P_1(x)=a_0+a_1x. \tag{6.20}$$

Щоб не оперувати з великими числами, зробимо наступну заміну

$$x=\frac{T-300}{100}. \tag{6.21}$$

Функцію Cp позначимо y. Для визначення коефіцієнтів a_0, a_1 необхідно скласти систему рівнянь

$$\begin{cases} a_0 n + a_1 \sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} y_i, \\ a_0 \sum\limits_{i=1}^{n} x_i + a_1 \sum\limits_{i=1}^{n} x_i^2 = \sum\limits_{i=1}^{n} x_i y_i. \end{cases} \tag{6.22}$$

Побудуємо табл. 6.7, що містить допоміжні дані для складання системи рівнянь.

Таблиця 6.7

Допоміжні дані до складання системи лінійних алгебраїчних рівнянь для визначення коефіцієнтів лінійної апроксімації

i	T_i	X_i	y_i	x_i^2	x_iy_i
1	300	0	70,35	0	0
2	400	1	75,38	1	75,38
3	500	2	80,53	4	161,06
4	600	3	85,81	9	257,43
5	700	4	91,26	16	365,04
6	800	5	96,83	25	484,15
7	900	6	102,53	36	615,18
8	1000	7	108,27	49	757,89
	Σ	28	710,96	140	2716,13

Підставивши дані з останнього рядка табл. 6.7 в систему рівнянь (6.22), одержимо:

$$\begin{cases} 8 \cdot a_0 + 28 \cdot a_1 = 710{,}96; \\ 28 \cdot a_0 + 140 \cdot a_1 = 2716{,}13. \end{cases} \tag{6.23}$$

Знайдемо a_0 і a_1 за формулами Крамера.

$$a_0 = \frac{\begin{vmatrix} 710{,}96 & 28 \\ 2716{,}13 & 140 \end{vmatrix}}{\begin{vmatrix} 8 & 28 \\ 28 & 140 \end{vmatrix}} = 69{,}89; \qquad a_1 = \frac{\begin{vmatrix} 8 & 710{,}96 \\ 28 & 2716{,}13 \end{vmatrix}}{\begin{vmatrix} 8 & 28 \\ 28 & 140 \end{vmatrix}} = 5{,}42.$$

Отже, шуканий апроксимуючий многочлен має вигляд

$$y = 69{,}89 + 5{,}42 \cdot x. \tag{6.24}$$

За допомогою формули (6.21) перейдемо до вихідних позначень. Одержимо

$$C_p = 69{,}89 + 5{,}42 \frac{T - 300}{100},$$

після перетворень

$$C_p = 53{,}62 + 5{,}42 \cdot 10^{-2} T. \tag{6.25}$$

Формула (6.25) є аналітичною залежністю, що узагальнює експериментальні дані табл. 6.6. Для оцінки лінійної апроксимації необхідно порівняти значення y_i з табл. 6.6 зі значеннями, які отримані за формулою (6.25) для всіх точок (i=1, 2, ..., 8). Результати порівняння подані в табл. 6.8.

Таблиця 6.8

Абсолютна та відносна похибки апроксімації

i	1	2	3	4	5	6	7	8
x_i	0	1	2	3	4	5	6	7
y_i	70,35	75,38	80,53	85,81	91,26	96,83	102,53	108,27
аналіт. залежність	69,89	75,31	80,74	86,16	91,58	97,00	102,43	107,85
абс. похибка	0,46	0,07	0,21	0,35	0,32	0,17	0,10	0,42
δy, %	0,65	0,1	0,3	0,4	0,4	0,2	0,1	0,4

Проаналізувавши δy з табл. 6.7, можна зробити висновок, що формула (6.25) є аналітичною залежністю, яка узагальнює експериментальні дані табл. 6.6. На рис. 6.6 наведені графік функції (6.25) і вихідні дані. Порівняльний аналіз показує задовільну збіжність лінійної апроксимації.

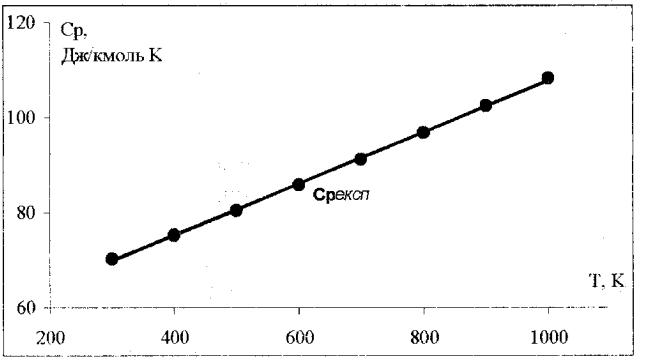


Рис. 6.6. Графік лінійного апроксимуючого многочлена і вихідні дані

6.4. Поняття сплайн-апроксимації

6.4.1. Аналітичне подання сплайнів

Якщо функція $f(x)$ задана своїми значеннями в $(n+1)$ – точці, то її можна апроксимувати інтерполяційним поліномом Лагранжа або Ньютона $P_n(x)$ степені n. Але при великих n виникають пульсації полінома $P_n(x)$ між вузлами, оскільки він має точки максимуму і мінімуму у дійсних нулях похідної. Усунути це можна зменшенням степені полінома, інтерполюючи тільки частину точок – так звана "кускова інтерполяція". Тоді для всієї множини точок одержується результуюча "зшита" або "склеєна" крива, яка неперервна на всьому інтервалі, але у вузлах "склеювання" окремих кривих її похідні матимуть розрив. Похибка такого наближення може виявитися значною. Крім того, таке наближення неможливе у тих випадках, коли важливе значення має гладкість апроксимуючої функції.

Побудова полінома невисокої степені для великого числа точок методом найменших квадратів не забезпечує задовільну відповідність кривої заданим точкам (сума квадратів різниць між значеннями многочлена і функції в даних точках хоча і мінімальна, але значна). Підвищення ж степені многочлена приводить до значного ускладнення в обчисленнях. Тому останнім часом в теорії апроксимації широкого використання набули сплайни (spline), які дозволяють в деякій мірі усунути недоліки класичних методів наближення.

У загальному випадку сплайн являє собою функцію, яку "склеєно" з узагальнених поліномів таким чином, що в точках "склейки" *їхні значення та значення похідних до деякого порядку збігаються*. Найбільш простими є кусково-поліноміальні сплайни, для яких за базисні функції беруться *алгебраїчні поліноми*.

Сплайни широко застосовуються в різних комп'ютерних технологіях: в автоматизації проектування з допомогою обчислювальної техніки, у конструюванні кривих і поверхонь, в опрацюванні й подальшому поданні геометричної інформації, а також у прикладній математиці – у теорії апроксимації, чисельному інтегруванні й диференціюванні, розв'язуванні диференціальних і інтегральних рівнянь.

Сплайни володіють рядом переваг у порівнянні із класичною інтерполяцією:

- ✓ *простота реалізації на ПК*;
- ✓ *забезпечення високої точності* апроксимації одночасно функції та її похідних;
- ✓ сплайни дозволяють отримати наближення з *мінімально можливою похибкою* заданого класу функцій у порівнянні з іншими методами;
- ✓ сплайни можна застосовувати для наближення функцій, у яких *ступінь гладкості неоднакова* на різних ділянках відрізка апроксимації.
- ✓ сплайни порівняно просто *узагальнюються* на багатовимірні задачі.

Назва сплайн (англ. Spline - гнучка креслярська лінійка) походить від технічного креслення. Для проведення гладких кривих через систему точок у креслярській практиці використають гнучкі лінійки-лекала. Їх згинають так, щоб вони наближалися до заданих точок і після цього проводять лінію.

Нехай на відрізку $[a;b]$ задано набір вузлів Δ:

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b. \tag{6.26}$$

Тоді позначимо:

- $C_k[a;b]$ – множина функцій, що неперервна на відрізку $[a;b]$ і має неперервні похідні k-го порядку:

- P_n – множина поліномів степеня не вище n.

Означення 6.1. Функція називається ***поліноміальним сплайном*** степеня n з множини k $(1 < k < n)$ з вузлами на сітці Δ, якщо:

- $S_n^k(x) \in P_n(x)\ \forall x \in [x_i;x_{i+1}]$ $(i = 0,1,2,\ldots,n\text{-}1)$;
- $S_n^k(x) \in C^{n-k}[a;b]$.

Набір k задає гладкість сплайну у вузлах і означає, що:

- функція $S_n^k(x)$ та її похідні до $(n\text{-}k)$-го порядку є *неперервними* на $[a;b]$;
- похідна порядку $(n\text{-}k\text{+}1)$ може бути *розривною* в точках x_i $(i = 1, 2,..., n\text{-}1)$.

Якщо k=1, то для сплайнів множини k можна задати позначення $S_n(x)$.

Будемо розглядати сплайни як апарат наближення функції $f(x)$, що задана на дискретно-заданій множині точок. Для сплайнів *непарного степеня* вузли інтерполяції вибирають, як правило, *у вузлах сплайна*. Нехай на відрізку $[a;b]$ у вузлах сітки Δ задані значення функції $f(x_i)=y_i$ $(i = 0, 1, .., n)$. Сплайн $S_{2n+1}^k(x)$ називається *інтерполяційним*, якщо він задовольняє умовам інтерполяції:

$$S_{2n+1}^k(x_i) = f(x_i)\ (i = 0,1,2,\ldots,n). \tag{6.27}$$

Для визначення параметрів сплайна необхідно задати умови інтерполювання иа додаткові крайові умови. Для сплайнів *парного степеня* вузли інтерполювання вибирають *відмінними від вузлів сплайну*, тому що інакше може виявитися, що при деяких умовах сплайн не існує або ж процес обчислення параметрів сплайну буде нестійким. Розглянемо способи ***аналітичного представлення сплайнів*** залежно від вибору базисних функцій.

Кусково-поліноміальне подання. Сплайн на кожному відрізку $[x_i;x_{i+1}]$ $(i = 0,1,...,n\text{-}1)$ можна подати у вигляді:

а) полінома відносно *початкової* точки відрізка:

$$S_n^k(x) = \sum_{j=0}^{n} a_{ij}(x - x_i)^j \,, x \in [x_i; x_{i+1}], \tag{6.28}$$

б) полінома відносно кінцевої точки відрізка:

$$S_n^k(x) = \sum_{j=0}^{n} b_{ij}(x - x_{i+1})^j \,, x \in [x_i; x_{i+1}]. \tag{6.29}$$

Побудова сплайну через усічену степеневу функцію: усічена степенева функція для змінної x і $N \in n$ визначається як:

$$x_+^N = \begin{cases} x^N; & x \geq 0; \\ 0; & x < 0; \end{cases} \tag{6.30}$$

і є сплайном *степеня n* множини k з вузлом в точці x=0.

Аналогічно *усічена ступенева функція* $(x - x_i)_+^N$, що зв'язана з вузлами сітки Δ (рис. 6.7), є неперервною в точці x=x_i разом із своїми похідними порядку до N-1. Отже вона є сплайном степеня n множини k з єдиним вузлом x_i.

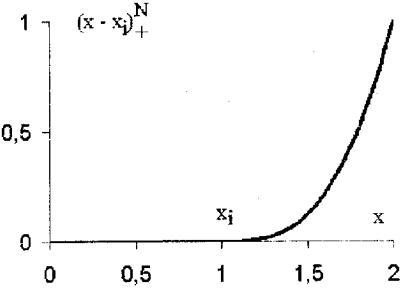


Рис. 6.7. Усічена степенева функція $(x - x_i)_+^m$ при N=3, $x_i = 1$

Функції x_j при j =0,1,...,n і $(x - x_i)_+^m$ при $m = N - k + 1,\ldots,N$; $i = 1,2,\ldots,n$-1 є лінійно незалежними на відрізку $[a;b]$, тому їх можна прийняти за базис. Звідси отримаємо сплайн на відрізку $[a;b]$ у вигляді усічених степеневих функцій.

$$S(x) \equiv S_n^k(x) = \sum_{r=0}^{N} c_r(x - x_0)^r + \sum_{i=1}^{n-1} \sum_{j=n=k+1}^{N} a_{ij}(x - x_i)_+^j \,; x \in [a;b], \tag{6.31}$$

де $c_r = \dfrac{S^{(r)}(x_0)}{r!}$; $a_{ij} = \dfrac{[S^{(j)}(x_i + 0) - S^{(j)}(x^i - 0)]}{j!}$ - величина розкиду похідних, розривних у вузлах; $r = 0,1,2,\ldots,N$; $i = 1,2,\ldots,n$-1; $j = N$-k+1,...,N.

Доведено, що така побудова сплайну є *єдина*.

6.4.2. Апроксимація кубічними сплайнами

На практиці найбільш широке застосування отримали інтерполяційні *кубічні сплайни*. Це обумовлено тим, що:

- поліном третього степеня - це найпростіша крива, що має точку згину, яка забезпечує його необхідні інтерполяційні можливості;
- кубічні сплайни мають на всьому відрізку апроксимації *неперервні похідні до другого порядку*. Така гладкість звичайно виявляється достатньою для більшості завдань;
- невисокий степінь полінома в кубічному сплайні спрощує обчислення та зменшує обчислювальну похибку.

Нехай функція $f(x)$ задана на відрізку $[a;b]$ у вузлах сітки Δ: $a = x_0 < x_1 < x_2 < ... < x_n = b$ значеннями $f(x_i) = y_i$ (i =0, 1, ..., n). Розглянемо сплайни третього степеня.

Функція $S_3(x)$ називається ***кубічним сплайном***, що інтерполює функцію $f(x)$ у вузлах сітки Δ, якщо:

- на кожному з відрізків $[x_i; x_{i+1}]$ (i = 0, 1, ...,n-1) вона є поліномом не вище третього степеня:

$$S_3(x) \in P_3; \tag{6.32}$$

- $S(x)$ є неперервною на $[a;b]$ разом зі своїми похідними *другого порядку* включно, тобто:

$$S_3(x) \in \mathbf{C}^2[a,b]; \tag{6.33}$$

- у вузлах сітки функція $S_3(x)$ задовольняє умовам інтерполяції:

$$S_3(x_i) = f(x_i)\ (i = 0,1,2,\ldots,n). \tag{6.34}$$

Скористаємося кускочно-поліноміальним поданням (6.28) кубічного сплайна на відрізку $[x_i; x_{i+1}]$:

$$S_3(x) = a_{i0} + a_{i1}(x-x_i) + a_{i2}(x-x_i)^2 + a_{i3}(x-x_i)^3,\ x \in [x_i;x_{i+1}]. \tag{6.35}$$

Завдання побудови сплайна полягає у знаходженні усіх коефіцієнтів (параметрів) $a_{i0},a_{i1},a_{i2},a_{i3}$. **При цьому на кожному** частковому відрізку $[x_i; x_{i+1}]$ кубічний сплайн визначається за *чотирма параметрами* $(a_{i0},a_{i1},a_{i2},a_{i3})$; на всьому відрізку $[a;b]$ кількість параметрів буде рівним $4 \cdot n$.

Для визначення параметрів сплайна потрібно скористатися:

- (n+1)-м умовою інтерполяції (6.32),
- умовами неперервності сплайна у внутрішніх вузлах:

$$S_3(x_i\text{-}0) = S_3(x_i\text{+}0)\ (i = 1,2,\ldots,n\text{-}1). \tag{6.46}$$

- умовами неперервності першої й другої похідних сплайна у внутрішніх вузлах:

$$S_3'(x_i\text{-}0) = S_3'(x_i\text{+}0)\ (i = 1,2,\ldots,n\text{-}1). \tag{6.47}$$

$$S_3''(x_i\text{-}0) = S_3''(x_i\text{+}0)\ (i = 1,2,\ldots,n\text{-}1). \tag{6.48}$$

Загальна кількість зв'язків (умов) має вигляд: $(n+1) + 3(n-1) = 4n-2$. Для розв'язку задачі не вистачає двох рівнянь зв'язку. Відсутні для однозначного визначення сплайна дві умови задаються додатково у вигляді обмеження на значення сплайна або його похідних *на* кінцях відрізка $[a,b]$, залежно від фізичного змісту конкретного завдання. Найбільше часто застосовують наступні *крайові умови:*

а) $S_3'(a) = f'(a), S_3'(b) = f'(b)$;

б) $S_3''(a) = f''(a), S_3''(b) = f''(b)$;

в) $S_3^{(r)}(a) = S_3^{(r)}(b)$, $r = 1,2$ (умова періодичності);

г) $S_3'''(x_i\text{-}0) = S_3'''(x_i\text{+}0)$ $(i = 1,2,\ldots,n\text{-}1)$ – додаткові умови неперервності, де $S_3'''(x_i\text{-}0)$ і $S_3'''(x_i\text{+}0)$ - відповідно, праві й ліва однобічні похідні третього порядку в точці x_i.

Для **визначення параметрів** кубічного сплайну задамо наступні позначення:

$$h_i = x_{i+1} - x_i \ (i = 0,1,\ldots,n\text{-}1); \tag{6.49}$$

$$M_i = S_3''(x_i), \tag{6.50}$$

$$m_i = S_3'(x_i) \ (i = 0,1,2,\ldots,n). \tag{6.51}$$

Величина M_i називається *моментом сплайна* в точці x_i. Величина m_i називається *нахилом сплайна* в точці x_i. Оскільки $S_3(x)$ - кубічна функція, то на кожному частковому відрізку друга похідна $S_3''(x)$ буде лінійною функцією від x. Тому для $x \in [x_i, x_{i+1}]$ за формулою лінійного інтерполювання запишемо:

$$\begin{aligned} S_3''(x) &= S_3''(x_i) + \frac{S_3''(x_{i+1}) - S_3''(x_i)}{x_{i+1} - x_i} \cdot (x - x_i) = \\ &= M_i + \frac{M_{i+1} - M_i}{h_i} \cdot (x - x_i). \end{aligned} \tag{6.52}$$

Проінтегруємо двічі (6.52):

$$S_3'(x) = S_3'(x_i) + M_i(x - x_i) + \frac{M_{i+1} - M_i}{2h_i}(x - x_i)^2; \tag{6.53}$$

$$S_3(x) = y_i + S_3'(x_i)(x - x_i) + \frac{M_i}{2}(x - x_i)^2 + \frac{M_{i+1} - M_i}{6h_i}(x - x_i)^3. \tag{6.54}$$

Підставимо у (6.54) $x = x_{i+1}$:

$$y_{i+1} = y_i + S_3'(x_i)h_i + \frac{M_i}{2}h_i^2 + \frac{M_{i+1} - M_i}{6h_i} \cdot h_i^3, \tag{6.55}$$

$$S_3'(x_i) = \frac{y_{i+1} - y_i}{h_i} - \frac{h_i}{6}(2M_i + M_{i+1}). \tag{6.56}$$

Підставимо вираз (6.56) у формули (6.53), (6.54):

$$S_3(x) = y_i + \left[\frac{y_{i+1}-y_i}{h_i} - \frac{h_i}{6}(2M_i + M_{i+1})\right](x-x_i) + \frac{M_i}{2}(x-x_i)^2 + \frac{M_{i+1}-M_i}{6h_i}(x-x_i)^3, \tag{6.57}$$

$$S_3'(x) = \frac{y_{i+1}-y_i}{h_i} - \frac{h_i}{6}(2M_i + M_{i+1}) + M_i(x-x_i) + \frac{M_{i+1}-M_i}{2h_i}(x-x_i)^2, \tag{6.58}$$

де $x \in [x_i;x_{i+1}]$ (i=0,1,2,…,n-1). Формули (6.57) і (6.58) виражають кубічний сплайн і його першу похідну на відрізку $[x_i;x_{i+1}]$ через значення функції й моменти у вузлових точках. З (6.57), (6.58) можна визначити параметри сплайна на кожному інтервалі в кусково-поліноміальному вигляді:

$$\begin{cases} a_{i0} = y_i; \\ a_{i1} = \dfrac{y_{i+1}-y_i}{h_i}; \\ a_{i2} = \dfrac{M_i}{2}; \\ a_{i3} = \dfrac{M_{i+1}-M_i}{6h_i}. \end{cases} \tag{6.59}$$

Для визначення моментів M_i, M_{i+1} використаємо властивість неперервності першої похідної. Для цього дорівняємо вираз першої похідної праворуч і ліворуч від кожної граничної крапки. Після перетворень одержимо n-1 рівняння вигляду:

$$\alpha_i M_{i-1} + 2M_i + \beta M_{i+1} = d_i, \tag{6.60}$$

$$\alpha_i = \frac{h_{i+1}}{h_i + h_{i-1}}, \tag{6.61}$$

$$\beta_i = \frac{h_{i-1}}{h_i + h_{i-1}} = 1 - \alpha_i, \tag{6.62}$$

$$d_i = \frac{6}{h_i + h_{i-1}}\left[\frac{y_{i+1}-y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}}\right], (i=1,2,\ldots,n\text{-}1). \tag{6.63}$$

Для визначення моментів M_i у всіх вузлах необхідно до системи (6.60) додати ще *два рівняння*, що випливають із крайових умов. Можуть бути задані наступні крайові умови:

а) У крайових точках збігаються перші похідні сплайна й функції:

$$S_3'(a) = f'(a),\ S_3'(b) = f'(b). \tag{6.64}$$

Додаткові рівняння мають вигляд:

$$2M_0 + M_1 = d_0, \tag{6.65}$$

$$M_{n-1} + 2M_n = d_n, \tag{6.66}$$

$$d_0 = \frac{6}{h_0}\left(\frac{y_1 - y_0}{h_0} - f'(a)\right); \tag{6.67}$$

$$d_n = \frac{6}{h_{n-1}}\left(f'(b) - \frac{y_n - y_{n-1}}{h_{n-1}}\right). \tag{6.68}$$

Ці крайові умови можна використати, коли відбувається сплайн-апроксимація заданої аналітичної функції. Спосіб непридатний для апроксимації табличнозаданих функцій.

б) У крайніх точках збігаються другі похідні функції й сплайна:

$$S_3''(a) = f''(a),\ S_3''(b) = f''(b). \tag{6.69}$$

Оскільки друга похідна – це шуканий момент, це означає, що:

$$M_0 = f''(a),\ M_n = f''(b). \tag{6.70}$$

Особливості:

- ✓ умова також використається для апроксимації функцій, заданих аналітично;
- ✓ якщо є підстави вважати, що $f''(a)$, $= f''(b) = 0$ (тобто на краях є точка перегину), те цей спосіб придатний для апроксимації табличних функцій;
- ✓ у цьому способі для знаходження моментів досить розв'язати n-2 рівняння.

г) Сплайн - періодичний з періодом (b-a). Значення сплайна і його першої й другої похідних у крайніх точках попарно збігаюся:

$$S_3^{(r)}(a) = S_3^{(r)}(b),\ r = 1,2. \tag{6.71}$$

Крайові умови:

$$2M_1 + \beta_1 M_2 + \alpha_1 M_n = d_1; \tag{6.72}$$

$$\beta_n M_1 + \alpha_n M_{n-1} + 2M_n = d_n; \tag{6.73}$$

$$M_0 = M_n. \tag{6.74}$$

Цей спосіб використається:

- для апроксимації періодичних функцій
- для апроксимації функцій, що утворюють замкнені криві.

д) Уводяться додаткові умови неперервності - рівність похідних 3-го порядку праворуч і ліворуч від всіх внутрішніх точок:

$$S_3'''(x_i - 0) = S_3'''(x_i + 0)\ (i = 1,2,\ldots,n-1). \tag{6.75}$$

Крайові умови:

$$(1 + \beta_1)\cdot M_1 + (\beta_1 - \alpha_1)\cdot M_2 = \beta_1 \cdot \alpha_1; \tag{6.76}$$

$$(\alpha_{n-1} - \beta_{n-1})M_{n-2} + (1 + \alpha_{n-1})M_{n-1} = \alpha_{n-1} d_{n-1}; \tag{6.77}$$

$$M_0 = \frac{M_1 - \alpha_1 M_2}{\beta_1}; \tag{6.78}$$

$$M_n = \frac{M_{n-1} - \beta_{n-1} \cdot M_{N-2}}{\alpha_{n-1}}. \tag{6.79}$$

У результаті виходить система з (n-1) лінійних алгебраїчних рівнянь із трьохдіагональною матрицею, з перевагою головної діагоналі. Внаслідок цього матриця системи буде невиродженою, що свідчить про єдине рішення системи. Два моменти, що залишилися розраховують по рівняннях (6.78) і (6.79). Цей метод може бути використаний для побудови сплайн-апроксимації табличних функцій.

Для розв'язання системи (6.60), (6.76), (6.77) використають метод прогону, який у цьому випадку є особливо ефективним. Оскільки в матриці коефіцієнтів є діагональна перевага, процес прогону є стійким щодо похибок заокруглення.

При кусково-поліноміальному представленні сплайна потрібно в пам'яті комп'ютера фіксувати: вузли сітки x_i, значення функції y_i, моменти M_i ($i = 0,1,...,n$).

Приклад 6.3. Побудувати сплайн і графік сплайн-апроксимації для табличної функції (табл. 6.9)

Таблиця 6.9

Вхідні дані для прикладу 6.3

x_i	0	1	3	3,5	5	6
y_i	1	2	6	6,5	6	3,5

Розв'язок.

а) Визначити моменти сплайна. За крайові використаємо умови (6.78). Робочий аркуш проекту представлений на скріншоті (рис. 6.8);

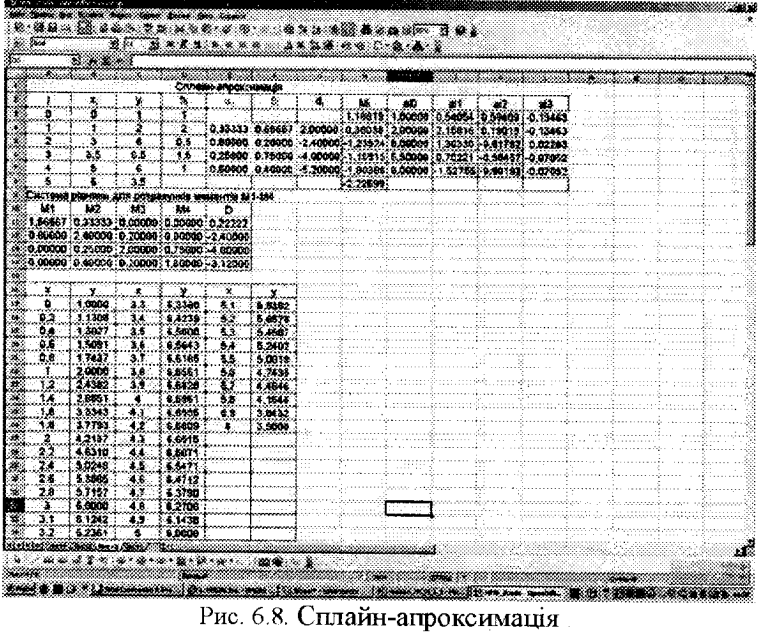


Рис. 6.8. Сплайн-апроксимація

б) задати вихідні дані - значення x_i, y_i;

в) у комірках D3-D7 визначити значення h_i; формула комірки D3: =B4-B3. Далі - копіювання на весь діапазон;

г) в комірках E4 - E7 визначити значення параметрів α по формулі (6.61). Формула комірки E4: =D3/(D3+D4). Далі - копіювання на весь діапазон;

д) в комірках F4 - F7 визначити значення параметрів β по формулі (6.62). Формула комірки F4: =1-E4. Далі - копіювання на весь діапазон;

е) в комірках G4:G7 визначити значення параметрів d за формулою (6.63) Формула комірки G4: =6/(D4+D3)*((C5-C4)/D4-(C4-C3)/D3). Далі - копіювання на весь діапазон;

ж) в комірках A11:D14 створюємо систему рівнянь для визначення моментів M_1 - M_4 за відповідними формулами. Формули ненульових комірок наведені в табл. 6.10.

з) у масиві комірок H4:H7 будуємо розв'язування системи рівнянь для M_1 - M_4. Формула масиву: =МУМНОЖ(МОБР(A11:D14);E11:E14);

Таблиця 6.10

Формули ненульових комірок для системи рівнянь сплайну (рис. 6.8)

Комірка	Рівняння ненульового елементу	Формула комірки
A11	$a_{11} = 1 + \beta_1$·	=1+F4
B11	$a_{12} = \beta_1 - \alpha_1$	=F4-E4
E11	$b_1 = \beta_1 \cdot \alpha_1$	=E4*F4
A12	$a_{21} = \alpha_2$	=E5
B12	$a_{22} = 2$	=2
C12	$a_{23} = \beta_2$	=F5
E12	$b_2 = d_2$	=G5
B13	$a_{32} = \alpha_3$	=E6
C13	$a_{33} = 2$	=2
D13	$a_{34} = \beta_3$	=F6
E13	$b_3 = d_3$	=G6
C14	$a_{43} = \alpha_4 - \beta_4$	=E7-F7
D14	$a_{44} = 1 + \alpha_4$	=1+E7
E14	$b_4 = \beta_4 \cdot d_4$	=E7*G7

і) в комірці H3 визначити значення моменту M_0 по формулі (6.67). Формула комірки: =(H4-E4*H5)/F4;

к) в комірці H8 визначити момент M_5 по формулі (6.68). Формула комірки: =(H7-F7*H6)/E7;

л) в комірці I3:I7 визначити коефіцієнти a_0 для кожного з діапазонів. Розрахунок виробляється за формулою (6.59). Формула комірки I3: = C3. Далі копіюємо на весь діапазон;

м) в комірках J3:J7 визначати коефіцієнти a_1 для кожного з діапазонів. Розрахунок виробляється по формулі (6.59). Формула комірки J3:
=(C4-C3)/D3-D3/6*(2*H3+H4). Далі копіюємо на весь діапазон;

н) в комірках K3:K7 визначити коефіцієнти a_2 для кожного з діапазонів. Розрахунок виробляється по формулі (6.49). Формула комірки K3: =H3/2. Скопіювати на весь діапазон;

о) в комірках L3:L7 визначити коефіцієнти a_3 для кожного з діапазонів. Розрахунок виконується за формулою (6.49). Формула комірки L3: =(H4-H3)/6/D3. Скопіюєвати на весь діапазон;

п) далі визначити значення сплайна в кожному з діапазонів. У діапазоні значень x від 0 до 1 розрахунок- по формулі комірки B17:
=$I$3+$J$3*(A17-$A$17)+$K$3*(A17-$A$17)^2+$L$3*(A17-$A$17)^3

Скопіювати на діапазон до B22. Аналогічно зробити розрахунок для інших діапазонів;

р) по отриманій таблиці побудувати графік сплайна (рис. 6.9).

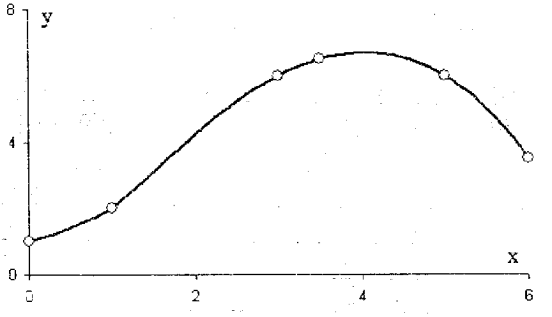


Рис. 6.9. Графік сплайн-апроксимації

Метод наближення сплайн має значне число модифікацій, у тому числі - метод згладжених сплайнів, симбіоз методу сплайнів і найменших квадратів, що дозволяє будувати гладкі апроксимації складних функцій за даними з похибками.

6.5. Розв'язування задач рівномірного сплайн-наближення

Задача рівномірного сплайн-наближення неперервної функції $f(x)$ на відрізку $[\alpha,\beta]$ полягає в знаходженні рівномірного наближення $f(x)$ декількома окремими виразами на часткових підінтервалах $[t_i, t_{i+1}]$, $i=\overline{1,q}$, $t_1=\alpha$, $t_{q+1}=\beta$, відрізку $[\alpha, \beta]$. Ці наближення підбираються так, щоб похибка ні в одній точці на всьому відрізку $[\alpha,\beta]$ не перевищувала наперед заданої величини. Доцільність такої постановки задачі з роздільним найкращим рівномірним наближенням на окремих підінтервалах відрізку $[\alpha,\beta]$ пояснюється можливістю отримання потрібної точності наближення при невеликих значеннях кількості параметрів наближаючих виразів. Кусково неперервні апроксимації використовуються, зокрема, у випадку, коли наближувана функція має особливості в деяких точках, які спричиняють погану апроксимацію на всьому відрізку.

6.5.1. Означення та властивості рівномірного сплайн-наближення

Розглянемо сплайн

$$S(x)=F_{k,m}\left(x;a^{(i)}\right), t_i \le x \le t_{i+1},\ i=1,...,q, \qquad (6.80)$$

в якому кожен вираз $F_{k,m}\left(x;a^{(i)}\right), i=1,...,q$ є найкращим рівномірним наближенням функції $f(x)$ на відрізку $[t_i,t_{i+1}]$ з вагою $w(x)$. Точки $t_i, i=1,...,q+1$ називають вузлами сплайна, а вираз $F_{k,m}\left(x;a^{(i)}\right)$, що задає значення сплайна на відрізку $[t_i,t_{i+1}]$, i-тою ланкою сплайна.

Якщо G_i – значення похибки наближення на i-тій ланці, то похибкою наближення сплайном називається величина G, яка рівна

$$G=\max_{1\le i\le q} G_i. \qquad (6.81)$$

Сплайни (6.80) поєднують позитивні властивості найкращого рівномірного наближення одним виразом (невелика кількість параметрів) і наближень класичними сплайнами (стійкість при обчисленні) [179].

Рівномірне сплайн-наближення характеризується функціональними властивостями й способом побудови. За функціональними властивостями розрізняють кусково-неперервні (кускові) й неперервні (склеєні) сплайни. В випадку неперервних сплайн-наближень на кожній ланці будується найкраще рівномірне наближення з певними додатковими умовами, які переважно визначають значення похибки у вузлах. Ця похибка може бути рівною нулю або визначатись допустимою похибкою сплайн-наближення. Необхідність розгляду кусково-неперервних сплайн-наближень визначається їх застосуванням, наприклад, при описі статичних характеристик термоперетворювачів або апаратній реалізації обчислення функцій [179]. При розв'язуванні таких задач допускається розривність апроксимуючої функції, лише б похибка наближення була достатньо малою.

За способом побудови розрізняють ***однорідні*** й ***неоднорідні*** сплайни [179]. *Однорідні* сплайни – це сплайни, в яких ланки представляються наближеними виразами одного й того ж виду, а *неоднорідні* – це сплайни, в яких ланки представляються різними наближеними виразами. Можливість вибору різних наближених виразів на окремих ланках сплайну дозволяє досягати більш точного представлення функцій на великих інтервалах. Неоднорідні сплайни використовуються при обчисленні спеціальних функцій і моделюванні деяких процесів [179]. Сплайни розрізняють ще й за способом вибору вузлів. Бувають сплайни з однаковою довжиною ланки (фіксовані вузли), так і з різною довжиною (довільні вузли). Побудова сплайнів з однаковою довжиною ланки по суті зводиться до послідовного знаходження рівномірного наближення для відомих підінтервалів і тому далі розглядаються тільки сплайни з довільними вузлами. В сплайнах з довільними вузлами довжина ланки визначається в залежності від умов оптимальності [179], тобто вибір вузлів при заданій похибці наближення повинен забезпечити мінімальну кількість ланок, а при заданій кількості ланок – мінімальну

похибку. Зустрічаються також задачі, в яких довжини ланок визначаються певними залежностями, наприклад, щоб довжина ланки була кратною степені двійки [179] тощо.

Серед задач на знаходження рівномірного сплайн-наближення розрізняють задачі, що зводяться до знаходження:

– однорідного сплайн-наближення, що забезпечує задану похибку при найменшій кількості ланок;

– однорідного сплайн-наближення з заданою кількістю ланок, що забезпечує мінімальну похибку наближення;

– неоднорідного сплайн-наближення, що забезпечує задану похибку при найменшій кількості ланок;

– неоднорідного сплайн-наближення з заданою кількістю ланок, що забезпечує мінімальну похибку наближення.

Для зручності приведені вище задачі скорочено позначатимемо S_1, S_2, S_3, S_4. Ці задачі характерні як для кусково-неперервних, так і неперервних сплайн-наближень.

6.5.2. Однорідне сплайн-наближення з заданою похибкою

Побудова однорідного сплайн-наближення з заданою похибкою (задача S_1) полягає в знаходженні найкращого рівномірного сплайн-наближення виразом $F_{k,m}(a;x)$ функції $f(x)$ на відрізку $C[\alpha,\beta]$ з вагою $w(x)$, яке забезпечує задану похибку G_0 при найменшій кількості ланок. Розв'язування цієї задачі полягає в такому виборі вузлів $t_i, i=1,...,q+1$ сплайн-наближення, при якому довжини всіх ланок, можливо, крім останньої, є максимально допустимими для заданої похибки наближення G_0.

Нехай $G(f,F_{k,m};\alpha_1,\alpha_2)$ – похибка найкращого рівномірного наближення функції $f(x)$ виразом $F_{k,m}(x;a)$ на відрізку $C[\alpha_1,\alpha_2]$ з вагою $w(x)$

$$G(f,F_{k,m};\alpha_1,\alpha_2)=\min_{A}\max_{\alpha_1\le x\le\alpha_2}\left|\frac{f(x)-F_{k,m}(x;a)}{w(x)}\right|. \qquad (6.82)$$

Далі, нехай точки $x_{i(j)}$ співпадають з вузлами сплайна t_j ($x_{i(j)}=t_j$, $j=1,2,...,q+1$). Тоді вузли $t_j\left(j=2,...,q;t_1=x_1=\alpha,t_{q+1}=x_N=\beta\right)$, які є розв'язком задачі однорідного сплайн-наближення з заданою похибкою, задовольняють умовам:

$$G(f,F_{k,m};x_{i(j)},x_{i(j+1)})\le G_0, j=1,...,q; \qquad (6.83)$$

$$G(f,F_{k,m};x_{i(j)},x_{i(j+1)+1})> G_0, j=1,...,q-1.$$

Сплайн-наближення, яке є розв'язком задачі S_1, будемо називати ***рівномірним сплайн-наближенням***.

Алгоритм побудови однорідного рівномірного сплайн-наближення з заданою похибкою полягає в послідовному визначенні проміжних вузлів сплайну t_j,

$j = 2,...,q$, $t_1 = \alpha$, $t_{q+1} = \beta$. Спочатку визначається права границя t_2 першої ланки $[t_1, t_2]$, потім - другої і т.д. Ці вузли визначаються ітераційним способом, при якому уточнення пробного значеня вузла t_j проводиться в залежності від похибки наближення на відрізку $[t_{j-1}, t_j]$ відповідної поточної ланки за формулою:

$$t_{j\,r+1} = x_{i_{j,\,r+1}},\quad j = 2,...,q,\ r = 1, 2,..., \tag{6.84}$$

$$i_{j,\,r+1} = i_{j-1} + \left[\left(i_{j_r} - i_{j-1}\right)\sqrt[m]{G_0/G_{j_r}}\right], \tag{6.85}$$

t_{j_r} - r-те наближення до j-го вузла сплайну; x_{i_j} - точка множини x, яка співпадає з j-им вузлом сплайна t_j $(t_j = x_{i_j})$; [ ]- квадратні дужки означають цілу частину від вмістимого; G_0- похибка сплайн-наближення; G_{j_r} - похибка наближення при r-ій ітерації на j-ій ланці $[t_{j-1}, t_j]$; m - кількість невідомих параметрів в наближаючому виразі; $i_{j,\,r}$ - номер точки x в підмножині X, яка відповідає r-му наближенню до j-го вузла сплайну.

Формула (6.85), для визначення номера точки чергового наближення до шуканого вузла, отримується з формули для наближеного значення похибки наближення рівномірним сплайном [179] в припущенні, що точки x_i, $i = 1,..., n$ рівновіддалені.

Початкове пробне значення вузла t_{2_o} може визначатись, виходячи з мінімально допустимої довжини ланки, або задаватись користувачем. Початкові пробні значення для решти вузлів t_{j_0} $(j = 3,... ,q)$- приймаються рівними довжинам попередніх ланок, тобто: $t_{j_0} = x_{i_{j,o}}$, де $i_{j,\,o} = 2i_{j-1} - i_{j-2}$.

У випадку, коли при деякому t_{j_r} з якихось причин не вдалось знайти найкраще рівномірне наближення на відрізку $[t_{j-1}, t_{j_r}]$, то за наступне значення $t_{j_{r+1}}$ вибирається сусідня зліва до t_{j_r} точка множини X - $(x_{i_{j,\,r+1}} = x_{i_{j,\,r}-1})$, при умові, що $i_{j,\,r} - 1 - i_{j-1} > m + 1$. При виборі наступного значення $t_{j_{r+1}}$ можна використовувати й інші стратегії, наприклад покладати його рівним середньому значенню між гранично допустимими. Задача S_1 при наближенні таблично заданих функцій не завжди має розв'язок. Вона не має розв'язку, якщо на одному з підінтервалів $[x_{i(j)}, x_{i(j+1)}]$, $j = 1,2,...,q$ мінімально допустимими для наближаючого виразу з m $(i(j+1) - i(j) = m)$ параметрами довжини ланки, отримується похибка G, більша від заданої $G_0(G > G_0)$. Можливий ще також випадок, коли не вдається знайти наближення на одному з підінтервалів $[x_{i(j)}, x_{i(j+1)}]$ через невиконання достатніх умов його існування.

Якщо при побудові рівномірного сплайн-наближення з заданою похибкою все таки на вдалося знайти наближення через невиконання необхідних умов його

існування або отримувалась більша від заданої похибка, то в залежності від конкретної задачі можна вибрати наближаючий вираз з більшою кількістю параметрів або збільшити похибку наближення. В такому випадку можливі й інші рішення, наприклад, в пропонують на підінтервалах, на яких найкраще рівномірне наближення не вдалося знайти, використовувати інтерполяцію тим самим виразом або наближають поліномом [166-167]. При програмній реалізації цього алгоритму корисно передбачити можливість задання за побажанням користувача початкової довжини першої ланки, а також оперативне коригування пробних значень вузлів сплайна. В процесі уточнення вузлів сплайну бажано виводити повідомлення про шуканий вузол, отриману на даній ланці похибку й наступне пробне значення для вузла. Така інформація дозволить користувачеві визначити стан розв'язування задачі й при потребі оперативно скоригувати подальший його хід.

6.5.3. Побудова неоднорідне сплайн-наближення

При побудові неоднорідного сплайн-наближення наближаючі вирази підбираються окремо для кожної ланки з деякої підмножини $V_m(V_m \le V)$. Тому при побудові неоднорідного сплайн-наближення з заданою похибкою необхідно попередньо визначити допустиму підмножину наближаючих виразів, які можуть використовуватись для апроксимації на окремих ланках. Якщо допустима підмножина наближаючих виразів не задається користувачем, то розглядаються всі можливі вирази з заданою кількістю параметрів. При цьому можна передбачити ще й задання пріоритетності застосування наближаючих виразів на окремих ланках. Початкове значення довжини першої ланки може визначатись так, як і при побудові однорідного сплайн-наближення з заданою похибкою.

Побудова неоднорідного сплайн-наближення з заданою похибкою відрізняється від побудови однорідного сплайн-наближення тим, що після знаходження оптимальної довжини ланки при наближенні одним з допустимих виразів перевіряється можливість збільшення її довжини при використанні інших наближаючих виразів. При цьому мінімально допустима довжина поточної ланки приймається на одну точку довшою від отриманої оптимальної. Такий вибір мінімально допустимої довжини ланки при апроксимації іншими допустимими виразами дозволяє з першої спроби визначити, чи можливе збільшення довжини поточної ланки.

Якщо пріоритетність наближаючих виразів не встановлена, то при переході до пошуку наступної ланки за початковий вид наближаючого виразу приймається вираз, який забезпечив максимальну довжину на попередній ланці.

При програмній реалізації цього алгоритму доцільно передбачити формування шляхом вибору бажаної підмножини наближаючих виразів, а також можливість оперативного коректування цієї підмножини в процесі розв'язування задачі [72-73, 152]. В праці [152] наведено приклад апроксимації номінальної статичної характеристики терп омари для діапазону кріогенних температур з необхідною точністю. В результаті апроксимації для опису залежності температури від перепаду напруги було отримано рівномірне сплайн-наближення з 12-ма параболічними ланками.

6.6. Методи обчислення рівномірного (мінімаксного) наближення функцій

6.6.1. Загальна постановка задачі найкращого мінімаксного наближення функцій

Розглянемо дискретний випадок задачі рівномірного (мінімаксного) наближення функції. Нехай C[α, β] – простір неперервних на відрізку [α,β] дійсних функцій з нормою:

$$\|f\| = \max_{\alpha \le x \le \beta} |f(x)|, \quad f(x) \in C[\alpha, \beta]. \tag{6.86}$$

і $V \le C[\alpha, \beta]$ – непустий підпростір виразів

$$F(x;a) = F(x;a_1, a_2, ..., a_m). \tag{6.87}$$

від дійсних параметрів a ($a \in A, A \le R^m$, де R^m – евклідовий m- мірний простір).

Нехай таблично задана функція $f(x)$ визначена значеннями $f_k = f(x_k)$, $k = \overline{1,N}$ на множині точок:

$$X = \{\alpha \le x_1 < x_2 < ... < x_N \le \beta\}. \tag{6.88}$$

і вагова функція $w(x)$ – відмінна від нуля на [α, β].

Означення 6.1. Вираз $F(x;a^*)$, для якого найбільше значення абсолютної величини зваженої похибки

$$G(f,F) = \max_{x_1 \le x \le x_N} \left| \frac{f(x) - F(x;a^*)}{w(x)} \right|, \tag{6.89}$$

досягає на множині X найменшого значення

$$G(f,F) = \min_{A} \max_{x_1 \le x \le x_N} \left| \frac{f(x) - F(x;a)}{w(x)} \right|, \tag{6.90}$$

називається найкращим рівномірним наближенням функції $f(x)$ виразом $F(x;a)$ на множині точок X з вагою $w(x)$. При цьому величину $G(f,F)$ (6.89) називають похибкою найкращого рівномірного наближення функції $f(x)$ виразом $F(x,a)$ на множині точок X з вагою $w(x)$.

Отже, метод рівномірного (мінміаксного) наближення функцій полягає в такому виборі невідомих коефіцієнтів апроксимуючого виразу, при якому найбільша абсолютна похибка наближення набуває найменшого значення:

$$\max_{x_1 \le x \le x_N} |(f(x) - F(x;a))/w(x)| \to \min. \tag{6.91}$$

Апроксимацію з таким критерієм визначення невідомих параметрів ще називають *найкращим чебишовським наближенням або* ***мінімаксним***.

Найбільш поширені рівномірні наближення з вагою $w(x)=1$ і $w(x)=f(x)$. У випадку $w(x)=1$ вираз $F(x;a^*)$ називають найкращим рівномірним наближенням з абсолютною похибкою, а у випадку $w(x)=f(x)$ – з відносною похибкою. На практиці зустрічаються й інші вагові функції, наприклад:

$$w(x)=\sqrt{1-f'(x)^2}\,. \tag{6.92}$$

Задача знаходження мінімаксного наближення виникає, наприклад, при таких розробках:

- проектуванні контрольно-вимірювальних приладів, зокрема термометрів або пірометрів (градуювання, лінеаризація статичних характеристик);
- проектуванні різноманітних функціональних перетворювачів (опис передаточних функцій);
- реалізації алгоритмів обчислення спеціальних функцій на високопродуктивних процесорах опрацювання даних (апроксимація складних функціональних залежностей простими аналітичними виразами);
- побудові математичних моделей різноманітних неперервних процесів на основі дискретних вимірювань їх значень;
- в машинобудуванні для представлення контурів проекцій деталей зі складними геометричними формами, а також аналітичного представлення кривих, які задають траєкторію різальних інструментів в верстатах з програмним управлінням.

Крім того, рівномірну апроксимацію даних доцільно застосовувати у різноманітних задачах ущільнення масивів результатів вимірювань, що містять методичну похибку у відомих межах (це може бути телеметрична інформація або масиви геофізичних чи сейсмічних вимірювань тощо).

Останнім часом поряд з класичними многочленними й раціональними найкращими рівномірними наближеннями все частіше розглядаються наближення нелінійними виразами. Це пояснюється тим, що фізичні процеси здебільшого мають нелінійний характер і для більш адекватного їх представлення використовуються відповідні теоретично отримані або інтуїтивно побудовані нелінійні залежності. Вивчаються також можливості використання рівномірного наближення функцій для представлення розв'язків диференціальних та інтегральних рівнянь, розглядаються й інші можливі застосування рівномірного наближення.

Таким чином, актуальною є розробка прикладного програмного забезпечення для персональних комп'ютерів, яке орієнтоване на розв'язування задач, що зводяться до знаходження найкращого рівномірного наближення функцій. В свій час для машин серії МИР і СМ-4 були розроблені пакети програм для розв'язування деяких задач з рівномірної апроксимації функцій. Таке програмне забезпечення було також розроблене й за кордоном. Розробка такого програмного забезпечення для персональних комп'ютерів повинна ґрунтуватись на аналізі задач з рівномірного наближення функцій і включати нові теоретичні й практичні розробки.

6.6.2. Апроксимаційні залежності

Підпростір наближаючих виразів V в загальному випадку складається з сукупності виразів:

$$F_i(x;a) = F_i(x;a_1, a_2, ..., a_m), \quad m = \overline{m_0, m_1}, \quad i = \overline{i, k}, \tag{6.93}$$

кожен з яких залежить від дійсних параметрів a, які в свою чергу є елементами підмножини A ($a \in A, A \le R^m$, де R^m – евклідовий m–вимірний простір). Відображення $F_i(x;a)$ співставляє кожному елементу $a \in A$ одну функцію $F_i(x;a) \in C[\alpha, \beta]$. Далі конкретне $F_k(x;a)$ називатимемо видом наближаючого виразу. Звичайно розрізняють многочленні, раціональні й нелінійні наближаючі вирази. Вирази одного виду відрізняються один від одного кількістю параметрів (наприклад, степенем полінома). Наближаючий вираз вважається заданим, якщо задано його вид і кількість параметрів. Для позначення наближаючого виразу k–го виду з m параметрами використовуватимемо запис $F_{k,m}(x;a)$.

А що ж до практичного вибору виразу $F(x;a)$, то він повинен бути придатним до якомога точнішого відтворення поведінки наближуваної функції *f(x)* на відрізку [α,β]. Конкретизуючи вигляд наближаючого виразу $F(x;a)$, потрібно в певних випадках враховувати функціональні властивості *f(x)*, а саме: парність, монотонність, асимптотику. Деколи вибір наближаючого виразу зумовлюється чисто практичними вимогами. Наприклад, зустрічаються задачі, в яких щодо апроксимуючого виразу ставляться вимоги по забезпеченню зручності при аналітичних перетвореннях над ним. В такому випадку доцільно використовувати алгебраїчні многочлени:

$$P_m(x;a) = a_0 + a_1 x + a_2 x^2 + ... + a_m x^m. \tag{6.94}$$

Многочленні наближення зручні також при обчисленні значень для конкретних x, їх нескладно технічно реалізувати в електронних приладах.

При розв'язуванні реальних задач як апроксимуючі залежності слід вибирати вирази $F(x;a)$, які відповідають очікуваній поведінці відповідних фізичних процесів. Наприклад, для опису затухаючих процесів доцільно використовувати експоненціальні многочлени виду:

$$E_m(x;a) = \sum_{i=1}^{m} a_1 \exp(-\lambda_i x), \tag{6.95}$$

де λ_i – фіксовані константи, а при періодичних процесах – тригонометричні поліноми:

$$T_m(x;a) = a_0 + \sum_{i=1}^{m} (a_i \cos(i \cdot x) + b_i \sin(i \cdot x)), \tag{6.96}$$

тощо. На практиці в задачах наближення найчастіше використовують вирази, в яких невідомі параметри входять лінійно, тобто квазіполіноми:

$$F(x;a) = a_1\varphi_1(x) + a_2\varphi_2(x) + \ldots + a_m\varphi_m(x), \tag{6.97}$$

де $\varphi_i(x)$ – лінійно-незалежні, неперервні функції на C[α,β], які задовільняють умові Хаара. Система функцій $\varphi_i(x)$ буде задовільняти умові Хаара, якщо визначник m-го порядку

$$\left|\varphi_s(x_i)\right|, \qquad i \le i, s \le m \tag{6.98}$$

не перетворюється в нуль ні для якого набору різних між собою m точок x на відрізку [α,β]. Системи функцій, що задовільняють умові Хаара називають ще Т-системами або чебишовськими системами функцій. Наприклад, умові Хаара задовільняють такі системи функцій:

$$\varphi_i(x) = x^i; \varphi_i = \cos(i \cdot x); \varphi_i = \exp(-\lambda_i x), \tag{6.99}$$

де λ_i – задані константи та інші.

Але, нажаль, лінійні щодо наближаючих параметрів апроксимації, зокрема й многочлени, не завжди відповідають характеру наближуваних даних. Многочлени, наприклад, мають ще й таку погану властивість, що їх значення швидко зростає при $|x| \to \infty$, що навіть нескладну функцію

$$\varphi(x) = 1/(1 + x^2) \tag{6.100}$$

многочлен погано апроксимує на великих відрізках. Іншим широко вживаним видом залежності від коефіцієнтів є раціональні вирази:

$$R_{m,l}(x;a) = \sum_{i=0}^{m} a_i x^i \Big/ \left(x^l + \sum_{i=1}^{l-1} b_i x^i \right), \tag{6.101}$$

які гнучкіше пристосовуються до *f(x)* у випадку, коли крайня точка відрізку [α,β] співпадає з точкою вертикальної дотичної.

Виявляється, що в багатьох випадках використання замість поліномів раціональних виразів при тій же кількості параметрів дозволяє досягнути кращої точності, тобто понижується значення максимальної похибки. Аналогічно до квазіполіноміальної залежності (6.101), на практиці використовують і квазіраціональні вирази, в яких на місці функцій x^i використовують $\varphi_i(x)$.

Деколи доцільно наближати не функцію $y = f(x)$, а спеціально підібрану $v(y)$. При цьому намагаються так підібрати функцію $v(y)$, щоб залежність від аргумента x описувалась простішою функцією, наприклад, многочленом. Але тоді слід пам'ятати, що отримане наближення не буде для функції $y=f(x)$ найкращим рівномірним, воно буде найкращим рівномірним наближенням тільки для функції $v(y)$. Відомо, що у випадку використання систем функцій, які задовільняють умові

Хаара найкраще рівномірне наближення існує завжди. Раціональні вирази умові Хаара задовільняють.

Більш складні задачі виникають при використанні нелінійних наближаючих виразів. Найкраще рівномірне наближення нелінійними виразами не завжди існує. Нелінійні вирази в загальному випадку не задовільняють умові Хаара, і тому їх використання вимагає дослідження існування для кожного конкретного випадку.

Досліджено властивості найкращого рівномірного наближення й розроблено відповідні алгоритми для обчислення їх параметрів такими нелінійними виразами:

$$A \cdot x^{\sum_{i=0}^{m} a_i x^i} \cdot e^{\sum_{i=0}^{m} a_i x^i}; \tag{6.102}$$

$$\sum_{i=0}^{m} a_i x^i + A \cdot e^{px}, p \neq 0; \tag{6.103}$$

$$\sum_{i=0}^{m} a_i x^i + A \cdot x^p, x \geq 0, p \neq k\left(k = \overline{0,m}\right); \tag{6.104}$$

$$\sum_{i=0}^{m} a_i x^i + A \cdot \ln(x+p), p > -x; \tag{6.105}$$

$$\sum_{i-0}^{m} a_i x^i + A \cdot sh(px), x > 0, p > 0; \tag{6.106}$$

$$\sum_{i=0}^{m} a_i x^i + A \cdot ch(px), x > 0, p > 0; \tag{6.107}$$

$$\sum_{i=0}^{m} a_i x^i + A \cdot (x+q)^p, q \geq x, p \neq k\left(k = \overline{0,m}\right); \tag{6.108}$$

$$a_0 + a_1 \cdot e^{qx} + a_2 \cdot e^{px}. \tag{6.109}$$

Звичайно, вирази (6.102-6.109) не охоплюють всіх можливих нелінійних представлень, які використовуються для наближення експериментальних даних, тому що фізичні чи хімічні процеси можуть описуватись будь-якими аналітичними залежностями. Здебільшого вид наближення вибирається, виходячи з фізичних представлень або емпіричних допущень, і, відповідно, невідомі параметри цих наближень мають певний фізичний зміст.

Серед задач з наближення функцій, крім знаходження наближення функції одним виразом на всьому відрізку $[\alpha,\beta]$, зустрічаються ще задачі знаходження наближення декількома окремими виразами на часткових підінтервалах $[t_i,\ t_{i+1}]$, $i = \overline{1,q}$, $t_1 = \alpha$, $t_{q+1} = \beta$, відрізку $[\alpha,\beta]$. Ці наближення підбираються так, щоб похибка ні в одній точці на всьому відрізку не перевищувала наперід заданої величини. Доцільність такої постановки задачі з роздільним найкращим

рівномірним наближенням на окремих підінтервалах відрізку $[\alpha,\beta]$ пояснюється можливістю отримання потрібної точності наближення при невеликих значеннях кількості параметрів m. Кусково-неперервні апроксимації використовуються, зокрема у випадку, коли наближувана функція має особливості в деяких точках, які спричиняють погану апроксимацію на всьому відрізку. В залежності від накладених на наближаючий вираз умов розрізняють такі задачі знаходження найкращого рівномірного наближення одним виразом на всьому відрізку:

- наближення заданим виразом;
- наближення певного виду із заданою похибкою й найменшою кількістю параметрів;
- наближення із заданою кількістю параметрів, що зебезпечує найменшу похибку серед всіх доступних видів наближення;
- наближення із заданою похибкою й найменшою кількістю параметрів, що зебезпечує найменшу похибку серед всіх доступних видів наближення.

Для зручності ці задачі позначатимемо відповідно через P_1, P_2, P_3, P_4.

Задача P_1 – це класична задача найкращого рівномірного наближення. Нехай на множині точок X (6.88) задано функції $f(x)$ ($f(x) \in C[\alpha, \beta]$) і $w(x)$ ($w(x) \in C[\alpha, \beta]$, $w(x) \neq 0$ на $[\alpha,\beta]$) і задано наближаючий вираз $F_{k,m}(x;a)$ з m ($m < N$) параметрами. Тоді задача P_1 полягає у визначенні таких значень невідомих параметрів $a_i^*(i = 1,2,...,m)$ наближаючого виразу $F_{k,m}(x;a)$, для яких

$$\left\| \frac{f(x)-F_{k,m}(x;a^*)}{w(x)} \right\| = \min_{A} \max_{x_1 \le x \le x_N} \left| \frac{f(x)-F_{k,m}(x;a)}{w(x)} \right|, \qquad (6.110)$$

де $\|\cdot\|$ – норма в просторі C.

Розв'язок задачі P_1 при наближенні неперервних функцій виразами, які лінійно залежать від параметрів, а також раціональними виразами, завжди існує. У випадку нелінійних наближаючих виразів існування найкращого рівномірного наближення визначається конкретними умовами.

6.6.3. Доведення про найкраще рівномірне наближення функцій

Розглянемо наступну задачу. Є елемент f лінійно нормованого простору R. Потрібно знайти його найкраще наближення лінійною комбінацією $\sum_{i=1}^{n} c_i g_i$ даних лінійно-незалежних елементів $g_1,...,g_n \in R$. Це означає: знайти елемент $\sum_{i=1}^{n} c_i^0 g_i$ такий, що:

$$\left\| f - \sum_{i=1}^{n} c_i^0 g_i \right\| = \Delta = \inf_{c_1,\ldots c_n} \left\| f - \sum_{i=1}^{n} c_i g_i \right\|. \tag{6.111}$$

Якщо такий елемент існує, то він називається елементом найкращого наближення.

Теорема 6.1. Елемент найкращого наближення існує.

Доведення. Внаслідок співвідношень

$$\left\| f - \sum_{i=1}^{n} c_i^1 g_i \right\| - \left\| f - \sum_{i=1}^{n} c_i^2 g_i \right\| \le \left\| \sum_{i=1}^{n} \left(c_i^1 - c_i^2 \right) g_i \right\| \le \sum_{i=1}^{n} \left| c_i^1 - c_i^2 \right| \left\| g_i \right\| \tag{6.112}$$

функція

$$F_f(c_1,\ldots c_n) = \left\| f - \sum_{i=1}^{n} c_i g_i \right\| \tag{6.113}$$

є неперервною функцією аргументів c_j при будь-якому $f \in R$. Нехай $\|c\|$ – евклідова норма вектора $c = (c_1,\ldots c_n)$. Функція

$$F_0(c_1,\ldots c_n) = \left\| c_1 g_1 + \ldots + c_n g_n \right\| \tag{6.114}$$

неперервна на одиничній сфері $\|c\| = 1$ і, відповідно, в деякій її точці $(\widetilde{c}_1,\ldots \widetilde{c}_n)$ досягає своєї нижньої границі $\widetilde{F}$ по сфері. $\widetilde{F} \neq 0$, так як рівність $\widetilde{F}\left\| \widetilde{c}_1 g_1 + \ldots + \widetilde{c}_n g_n \right\| = 0$ суперечить лінійній незалежності елементів $g_1,\ldots,g_n$. Для будь-якого $c = (c_1,\ldots c_n) \neq (0,\ldots,0)$ справедлива оцінка

$$\left\| c_1 g_1 + \ldots + c_n g_n \right\| = F_0(c_1,\ldots c_n) = \|c\| F_0\left(\frac{c_1}{\|c\|},\ldots,\frac{c_n}{\|c\|} \right) \ge \|c\| \widetilde{F}. \tag{6.115}$$

Нехай $\gamma \succ \dfrac{2\|f\|}{\widetilde{F}}$. Функція $F_f(c_1,\ldots c_n)$ неперервна в кулі $\|c\| \le \gamma$ і, відповідно, в деякій точці $(c_1^0,\ldots c_n^0)$ досягає своєї нижньої границі по кулі F_0. Маємо $F^0 \le F_f(0,\ldots,0) = \|f\|$. Поза цією кулею

$$F_f(c_1,\ldots c_n) \ge \left\| c_1 g_1 + \ldots + c_n g_n \right\| - \|f\| \succ \widetilde{F}\frac{2\|f\|}{\widetilde{F}} - \|f\| = \|f\| \ge F^0. \tag{6.116}$$

Таким чином $F_f(c_1,\ldots c_n) \ge F^0 = F_f(c_1^0,\ldots c_n^0)$ при будь-яких $c_1,\ldots c_n$. Теорема доведена. Елементів найкращого наближення може бути кілька.

Теорема 6.2. Якщо простір R строго нормований, то елемент найкращого наближення єдиний.

Доведення. Припустимо протилежне, тобто, що існують два елементи $f_1 \neq f_2$, $f_j = \sum_{i=1}^{n} a_{ij} g_i$ такі, що $\|f - f_1\| = \|f - f_2\| = \Delta$. Очевидно, $\Delta \neq 0$, тому що інакше $f_1 = f = f_2$, крім того,

$$\left\| f - \frac{f_1 + f_2}{2} \right\| = \left\| \frac{f - f_1}{2} + \frac{f - f_2}{2} \right\| \leq \frac{\|f - f_1\|}{2} + \frac{\|f - f_2\|}{2} = \Delta. \tag{6.117}$$

Оскільки $\frac{f_1 + f_2}{2}$ є лінійною комбінацією елементів $g_1, \ldots, g_n$ то $\left\| f - \frac{f_1 + f_2}{2} \right\| \geq \Delta$ внаслідок визначення Δ. Разом з попередніми співвідношеннями це означає, що $\left\| \frac{f - f_1}{2} \right\| + \left\| \frac{f - f_2}{2} \right\| = \left\| \frac{f - f_1}{2} + \frac{f - f_2}{2} \right\|$. Внаслідок припущення строгого нормованого простору $\frac{f - f_1}{2} = \alpha \frac{f - f_2}{2}$. Якщо $\alpha \neq 1$, то $f = \frac{f_1 - \alpha f_2}{1 - \alpha}$, f є лінійною комбінацією $g_1, \ldots, g_n$ і тому $\Delta = 0$. При $\alpha = 1$ маємо $f_1 = f_2$ і приходимо до протиріччя з допущенням $f_1 \neq f_2$.

Якщо норма в лінійному нормованому просторі визначається не через скалярний добуток, знаходження елемента найкращого наближення суттєво ускладнюється [76-80, 88, 92-93]. Нехай R – простір обмежених речових функцій, визначених на відрізку $[a,b]$ дійсної осі, з нормою:

$$\|f\| = \sup_{[a,b]} |f(x)|. \tag{6.118}$$

Шукаємо найкраще наближення вигляду

$$Q_n(x) = \sum_{j=0}^{n} a_j x^j. \tag{6.119}$$

Згідно теореми 6.1 існує елемент найкращого наближення, тобто многочлен $Q_n^0(x)$ такий, що

$$E_n(f) = \|f - Q_n^0\| \leq \|f - Q_n\| \tag{6.120}$$

при будь-якому многочлені $Q_n(x)$ степені n Такий многочлен $Q_n^0(x)$ називається поліномом рівномірного (мінмаксного) наближення [76-80, 88, 92-93].

6.7. Метод Є.Я.Ремеза

6.7.1. Загальна постановка задачі схеми Є.Я.Ремеза

При знаходженні мінмаксного наближення найчастіше використовуються алгоритми, що побудовані за схемою Ремеза, або апарат лінійного й випуклого програмування. Використовують й інші підходи, наприклад, знаходження найкращого рівномірного наближення реалізується шляхом послідовних найкращих наближень в метриці простору L_p. Цей підхід грунтується на тому, що в граничному випадку при $p \to \infty$ найкраще рівномірне наближення – це наближення в нормі простору L. При цьому найкраще наближення в метриці L_p зводять до знаходження зважених середньо-квадратичних апроксимацій.

Алгоритмам, які побудовані за схемою Ремеза, в порівнянні з іншими методами, властива більш швидка збіжність. Вони збігаються зі швидкістю геометричної прогресії. Практична ефективність методів типу Ремеза перевірена великою кількістю чисельних реалізацій. Теоретичною основою цих методів для випадку многочленних наближень є характеристична теорема Чебишова. Розглянемо умови існування і властивості найкращого рівномірного наближення в загальному випадку. Нехай підпростір наближаючих виразів V (6.115) задовільняє таким властивостям:

- A – відкрита підмножина m-вимірного евклідового простору параметрів a_i, $i = 1,..., m$;
- компактна підмножина X (2.1.3) складається не менш ніж з $m+1$ точки і функції $F(x;a)$ – неперервно-диференційовні по $a(a \in A)$ та $x(x \in X)$.

Говорять, що вираз $F(x;a)$ на підмножині A задовільняє локальній умові Хаара, якщо для кожного $a(a \in A)$ лінійна комбінація компонентів градієнта

$$\left(\frac{dF(x;a)}{da_1}, \ldots, \frac{dF(x;a)}{da_m} \right) \tag{6.121}$$

задовільняє класичній умові Хаара при $x(x \in X)$, тобто лінійна комбінація компонентів вектора (6.121) має на X не більше (m-1)-го кореня.

Встановлено, що коли для виразу $F_{k,m}(x;a)$ виконується локальна умова Хаара і для будь-яких параметрів a і b з A функція $F_{k,m}(x;a) - F_{k,m}(x;b)$ має на X не більше (m-1)-го кореня, то для будь-якої функції $f(x)$ ($f(x)$ є $C[\alpha,\beta]$) на множині точок X (6.88) існує єдине найкраще рівномірне наближення виразом $F_{k,m}(x;a)$ з вагою $w(x)$ ($w(x) \in C[\alpha,\beta]$, $w(x) \neq 0$ на $[\alpha,\beta]$).

Відомо також [76-80, 88, 92-93], що коли для функції $f(x)$ ($f(x) \in C[\alpha,\beta]$) на множині точок X (6.88) існує єдине найкраще рівномірне наближення виразом $F_{k,m}(x;a)$ з вагою $w(x)$ ($w(x) \in C[\alpha,\beta]$, $w(x) \neq 0$ на $[\alpha,\beta]$), то його параметри задовільняють систему рівнянь:

$$\frac{f(z_i)-F(z_i;a)}{w(z_i)}=(-1)^i\mu,\ i=1,...,m+1 \qquad (6.122)$$

де z_i – точки чебишовського альтернансу, а m – похибка наближення.

Згідно з цією властивістю існує $(m+1)$ точка на множині X (6.88), в яких похибка наближення набуває найбільшого за абсолютною величиною значення, і при цьому її знак почергово змінюється в сусідніх точках, тобто:

$$\mu=\max_{x_1\le x\le x_N}\left|G(f,F:x)\right|,\ G(f,F:z_i)=-G(f,F:z_{i+1}),\ i=1,...,m,$$

де $G(f,F:x)=\dfrac{f(x)-F(a;x)}{w(x)}$.

Властивість (6.122) називають характеристичною властивістю найкращого рівномірного наближення функцій або критерієм чебишовського альтернансу.

Якщо відомі точки чебишовського альтернансу $z_i(i=1,...,m+1)$, то значення параметрів a_i (i=1,...,m) найкращого рівномірного наближення можна визначити з системи рівнянь (6.122). Визначення наближаючих параметрів найкращого рівномірного наближення з системи рівнянь (6.122) називають розв'язком задачі чебишовської інтерполяції.

Отже, для побудови найкращого рівномірного наближення необхідно визначити точки чебишовського альтернансу. Точне їх значення може бути визначене лише в конкретних випадках, які відомі ще з часів Чебишева. Для визначення точок чебишовського альтернансу використовують ітераційний метод Ремеза, який полягає в послідовному розв'язуванні задач чебишовської інтерполяції для наборів точок $z_i^{(j)}$ $(i=1,...,m+1;\ j=0,1,..)$ – наближень до точок альтернансу. За методом Ремеза визначення параметрів наближення проводиться за такою схемою:

1) Візьмемо $j=0$. З множини точок X (6.88) вибирається певним чином підмножина точок:

$$z^{(j)}=\left\{z_i^{(j)},i=1,...,m+1\right\}.$$

Підмножина точок $z^{(0)}$ називається початковим наближенням до точок чебишовського альтернансу. Для початкового наближення до точок альтернансу рекомендується брати екстремуми многочлена Чебишова:

$$z_i=\frac{\alpha+\beta}{2}+\frac{\beta-\alpha}{2}\cos\left(\frac{(m+1-i)\pi}{m}\right). \qquad (6.123)$$

Такий вибір початкового наближення до точок альтернансу обгрунтовується орієнтацією на отримання не дуже малої початкової похибки наближення m.

При наближенні таблично заданих функцій на множині точок X (6.88) за початкове значення точок альтернансу приймаються величини з множини X по можливості близькі до значень (6.123), наприклад:

$$z_i^{(0)}=x_r,\qquad i=1,...,m+1,$$

де $r=\begin{cases}p_i, & if\ p_i>i;\\ p_{i+1}+1, & if\ p_i\le i\end{cases}$ та $p_i=\left[\dfrac{N-1}{2}\cos\dfrac{(m+1-i)\pi}{m}+\dfrac{N+1}{2}\right]$,

квадратні дужки [ ] означають цілу частину від вмістимого.

2) Розв'язується задача чебишовської інтерполяції для функції $f(x)$ наближаючим виразом $F_{k,m}(x;a)$ на множині точок $z^{(j)}$, тобто визначаються коефіцієнти наближення $a_i^{(j)}$, i=1,...,m та величина похибки μ_j із системи рівнянь:

$$\frac{f(z_i^{(j)}) - F(z_i^{(j)};a)}{w(z_i^{(j)})} = (-1)^i \cdot \mu_j,\ i=1,...,m+1. \qquad (6.124)$$

Ця система рівнянь залежить від вигляду наближаючого виразу $F_{k,m}(x;a)$. Якщо наближаючий вираз лінійно залежить від невідомих параметрів a, то система (6.124) буде лінійною. В загальному випадку – це система нелінійних рівнянь. Спроби розв'язати систему рівнянь (6.124) для загального випадку не були успішними. Як стверджують самі автори цих спроб, їм вдалось отримати розв'язок лише в часткових випадках. Методи розв'язування задачі чебишовської інтерполяції в випадку многочленних, раціональних, а також деяких нелінійних виразів, що зводяться до лінійних, описані в [76-80, 88, 92-93].

3) Визначається величина

$$\rho_j = \max_{x_1 \le x \le x_N} \left| \frac{f(x) - F(x;a)}{w(x)} \right|, \qquad (6.125)$$

тобто величина максимальної по модулю похибки наближення функції $f(x)$ наближаючим виразом $F_{k,m}(x;a)$ з вагою $w(x)$ на множині точок X.

4) Перевіряється виконання умови:

$$\rho_j = |\mu_j|. \qquad (6.126)$$

Якщо ця умова виконується, то шукані коефіцієнти a^* найкращого рівномірного наближення функції $f(x)$ наближаючим виразом $F_{k,m}(x;a)$ з вагою $w(x)$ на множині точок X знайдені, вони рівні відповідно $a_i^*=a_i^{(j)}$, i=1,...,m.

Через наявність похибок заокруглення в програмних реалізаціях алгоритму Ремеза цю умову замінюють перевіркою нерівності виду:

$$\rho_j - \mu_j \le \varepsilon \cdot |\mu_j|,$$

де ε – допустима відносна похибка.

На практиці достатньо досягти співпадіння двох–трьох значущих цифр в величинах ρ_j і μ_j. Тобто значення ε = 0,003 здебільшого задовільне.

5) Якщо умова (6.126) не виконується, то певним чином здійснюється перехід до (j+1)-го наближення до точок чебишовського альтернансу. Покладаємо j=j+1 та повторюємо виконання цієї схеми, починаючи з пункту 2.

Отже, знаходження найкращого рівномірного наближення за методом Ремеза полягає в ітераційному виконанні пунктів 2-5 наведеної схеми, поки не виконається умова пункту 4. Для проведення заміни точок альтернансу в пункті 5, відомо декілька способів. В деяких з них проводиться заміна однієї точки, в інших –

декількох або й всіх. Одним з найпростіших і поширених способів заміни точок альтернансу є алгоритм Валле-Пуссена. Його ще називають алгоритмом Ремеза з одноточковою заміною.

6.7.2. Мінімакс. Визначення параметрів мінімаксимальної моделі за алгоритмом Є. Я. Ремеза

Мінімакс – це один з принципів оптимального вибору параметрів. Суть цього принципу легко пояснити на прикладі опрацювання емпіричних даних [76-80, 88, 92-93]. Нехай задані результати спостереження, що являють собою таблицю значень деякої функції:

$$y_k = y(t_k),\ \ k \in \|0 : N\|. \tag{6.127}$$

Потрібно вибрати серед всіх алгебраїчних поліномів $P_n(A,t) = \sum_{i=0}^{n} a_i t^i$, де $A = (a_0,...,a_n)$ такий алгебраїчний поліном $P_n(A^*,t)$, який найкраще апроксимує таблицю 6.1 [76-80, 88, 92-93]. Поліном $P_n^*(a,x)$ називатимемо мінімаксимальною моделлю з найменшою абсолютною похибкою результатів спостереження y_k на множині точок x, якщо виконуватиметься рівність:

$$P_n^*(a,x) = \sum_{i-0}^{n} a_k^* x^k y_k,\ \text{де}\ k = \overline{0,N},\ \ x = \{x_0, x_1,..., x_n\};$$

$$\max_k \left| y_k - P_n^*(a,x) \right| = \min_a \max_k \left| y_k - P_n(a,x) \right|, \tag{6.128}$$

$$\max_k \left| y_k - P_n^*(a,x_k) \right| = \max_k \left| y_k - P_n(a,x_k) \right|. \tag{6.129}$$

$P_n^*(a,x)$ буде мінімаксимальною моделлю з найменшою абсолютною похибкою, коли виконуватиметься нерівність (6.129), а з найменщою відносною, коли виконуватиметься нерівність (6.128). Розглянемо задачу пошуку мінімаксимальної моделі з ваговою функцією $W(x) = 1$.

Доведено, що для результатів спостереження, які представляють неперервні процеси, існує єдина мінімаксимальна поліноміальна модель степеня не вище n, яка забезпечує найменше значення зваженої похибки на множині точок спостереження.

Нехай результати спостереження y_k описують заданий неперервний процес та задана вагова функція $W(x) \neq 0$. Тоді для того, щоб деякий поліном $P_n(a,x)$ був поліномом, який найменше відхиляється від результатів спостереження y_k на множині точок з ваговою функцією $W(x)$ необхідно і достатньо, щоб серед точок x знайшлось щонайменше $n+z$ точок, в яких зважена похибка почергово приймала значення різного знаку і досягала найбільшого за модулем значення.

$$E_n(z_k)=\frac{y(z_k)-P_n(a,z_k)}{W(z_k)}, \tag{6.130}$$

$$E_n(z_1)=-E_n(z_2)=\ldots=(-1)^{n+2-1}E_n(z_{n+2})=\max_{0\le k\le m}\left|\frac{y_k-P_n(a,x_k)}{W(x_k)}\right|.$$

Точки z_k в яких зважена похибка E_n приймає найбільше за модулем значення називаються **точками чебишевського альтернансу**.

В загальному випадку для визначення точок альтернансу використовують **алгоритм Є.Я.Ремеза**, ідея якого грунтується на властивості мінімаксної поліноміальної моделі на множині точок x. Ремез запропонував:

а) з множини точок x вибрати $n+2$ точки, які наближені до точок альтернансу z_i з найменшим відхиленням: $z_k^l=x_{r_k}$ де $k=\overline{1,n+2}$, $l=\overline{k,n+2}$ – номер ітерації,

$$r_k=\begin{cases}r_{k-1}+1, \textit{якщо } P_k\le k,\\ P_k, \textit{якщо } P_k\succ k.\end{cases}$$

$$P_k=\left[\frac{N-1}{2}\cos\frac{(n+1-k)\pi}{n}+\frac{m+1}{2}\right]. \tag{6.131}$$

б) на цій множині точок, які наближені до точок альтернансу, шукаємо мінімаксну поліноміальну модель:

$$\begin{gathered}y(z_1^{(l)})-\sum_{k=0}^{n}a_k^{(l)}z_1^k=\mu^{(l)},\\ y(z_2^{(l)})-\sum_{k=0}^{n}a_k^{(l)}z_2^k=-\mu^{(l)},\\ \vdots\\ y(z_{n+2}^{(l)})-\sum_{k=0}^{n}a_k^{(l)}z_{n+2}^k=(-1)^{n+2-1}\mu^{(l)},\end{gathered} \tag{6.132}$$

де $\mu^{(l)}$ – похибка в точках альтернансу (похибки моделі, вона менша максимальної для I-ї точки альтернансу z_1); $-\mu^{(l)}$ – похибка моделі, вона менша максимальної для II-ї точки альтернансу z_2

Розв'язавши систему, знайдемо параметри $a_0,a_1,\ldots,a_n$.

в) визначаємо максимальну похибку $\rho^{(l)}$ мінімаксимальної поліноміальної моделі:

$$\rho^{(l)}=\max_{0\le k\le m}\left|y_k-P_n\left(a^{(l)},x_k\right)\right|. \tag{6.133}$$

г) перевіряємо умову досягнення максимальної похибки в точках альтернансу мінімаксимальної поліноміальної моделі. Якщо умова виконується, то вона є

мінімаксною, якщо ж ні, то певним чином змінюємо наближення до точок альтернансу.

д) попередню умову заміняємо умовою:

$$\rho^{(l)} - \left|\mu^{(l)}\right| \le E\left|\mu^{(l)}\right|, \tag{6.134}$$

де E – допустима відносна похибка відхилення максимального за модулем значення похибки моделі відносно похибки в точках альтернансу.

Якщо ця умова виконується, то знайдене наближення є результуючим: $P_n(a^*, x) = P_n(a^{(l)}, x)$, де $a_k^* = a_k^{(l)}$, $k = \overline{0,n}$ Якщо ця умова не виконується, то переходимо до пункту б) і змінюємо крок ітерації на одиницю, тобто $l = l + 1$.

Алгоритм Ремеза при визначенні точок альтернансу полягає у виконанні пунктів б) і г) до тих пір, поки максимальне значення моделі не буде досягатися в точках альтернансу. Алгоритмам, які побудовані за схемою Ремеза, в порівнянні з іншими методами, властива більш швидка збіжність, вони збігаються з швидкістю геометричної прогресії. Практична ефективність методів типу Ремеза перевірена великою кількістю чисельних реалізацій як в нашій країні так і за кордоном. Теоретичною основою цих методів для випадку многочленних наближень є характеристична теорема Чебишова [76-80, 88, 92-93].

6.8. Метод Валле-Пуссена

6.8.1. Загальна постановка задачі

Алгоритм Валле-Пуссена переходу від j-го до (j+1)-го наближення до точок альтернансу полягає у наступному. Нехай уже виконані п.1-4 загальної схеми Ремеза, і виявилось, що найбільша по модулю похибка ρ_j (6.70) досягається у точці x_r, тобто:

$$\rho_j = \left|\rho(x_r)\right|, \quad \rho(x_r) = \frac{f(x_r) - F_{k,m}(x;a)}{W(x)}, \tag{6.135}$$

де $\rho_j > \left|\mu_j\right|$. Можливі три випадки взаємного розміщення точки x_r відносно точок альтернансу $z_i^{(j)}$:

1) $z_0^{(j)} < x_r < z_{m+1}^{(j)}$,

2) $x_r < z_0^{(j)}$,

3) $x_r > z_{m+1}^{(j)}$.

Розглянемо кожен із цих випадків.

1) Знаходимо ціле число l таке, що

$$z_l^{(j)} < x_r < z_{l+1}^{(j)}.$$

Якщо $sign\ \rho(x_r)=sign\ \rho(z_{l+1}^{(j)})$, то вибираємо $z_{l+1}^{(j)}=x_r$, у протилежному випадку змінюємо значення точки $z_{l-1}^{(j+1)}=x_r$.

Решта точок альтернансу в обох випадках залишається без зміни.

2) Якщо $\rho(x_r)=sign\ \rho(z_0^{(j)})$, то беремо $z_0^{(j+1)}=x_r$, а решту точок альтернансу залишаємо без зміни. Якщо ця умова не виконується, то здійснюємо зсув точок альтернансу:

$$z_{m+1}^{(j+1)}=z_m^{(j)};z_m^{(j+1)}=z_{m-1}^{(j)};\cdots;z_1^{(j+1)}=z_0^{(j)};z_0^{(j+1)}=x_r,$$

тобто із j-го альтернансу викидається точка $z_{m+1}^{(j)}$, а вводиться x_r як точка $z_0^{(j+1)}$.

3) Якщо $\rho(x_r)=sign\ \rho(z_{m+1}^{(j)})$, то беремо $z_{m+1}^{(j+1)}=x_r$, а решту точок альтернансу залишаємо без зміни. У випадку, коли остання умова не виконується, здійснюємо такий зсув точок альтернансу:

$$z_0^{(j+1)}=z_1^{(j)};z_1^{(j+1)}=z_2^{(j)};\cdots;z_m^{(j+1)}=z_{m+1}^{(j)};z_{m+1}^{(j+1)}=x_r\ ,$$

тобто з j-го альтернансу викидається точка z_0, а вводиться x_r як точка $z_{m+1}^{(j+1)}$.

Наступне наближення до точок альтернансу відрізняється від попереднього тим, що точка x_r вводиться замість однієї із його попередніх точок. При цьому потрібно слідкувати за зміною знаку похибки наближення в сусідніх точках.

6.8.2. Алгоритм Валле-Пуссена

Для зміни точок альтернансу (вибору наступного наближення до них) застосовують алгоритм Валле-Пуссена. За цим алгоритмом наступне наближення до точок альтернансу отримується зміною однієї точки, тобто в альтернанс вводиться точка, в якій досягалось найбільше значення похибки моделі $z_k^{(k)}$, де $k=\overline{1,n+2}$, $\rho^{(l)}\succ\left|\mu^{(l)}\right|$. Припустимо, що максимальна похибка за модулем досягається у точці:

$$\rho^{(l)}=\max_{0\le k\le m}\left|y_k-P_n\left(a^{(l)},x_k\right)\right|=\left|y_r-P_n\left(a^{(l)},x_r\right)\right|.$$

Можливі три випадки взаємного розміщення точки x_r відносно точки z_i:

1. Точка x_r розташована поміж точок альтернансу: $z_1^{(l)}\prec z_r\prec z_{n+2}^{(l)}$.
2. Точка x_r розташована ліворуч від точок альтернансу: $x_r\prec z_1^{(l)}$.
3. Точка x_r розташована праворуч від точок альтернансу: $x_r\succ z_{n+2}^{(l)}$.

У **першому** випадку знайдемо такий номер k, для якого виконується співвідношення $z_k^{(l)}\prec x_r\prec z_{k+1}^{(l)}$. Точку x_r вводимо замість точок $z_k^{(l)}$ і $z_{k+1}^{(l)}$, в яких будуть чергуватися знаки "+" і "–". Валле-Пуссен запропонував вводити точку x_r так, щоб знаки "+" і "–" чергувалися: $sign(\rho(z_k^{(l)}))=sign(\rho(x_r))$, де $z_k^{(l+1)}=x_r$, $z_i^{(l+1)}=z_i^{(l)}$, $i=\overline{1,n+1}$, $i\ne k$. Якщо знаки похибки $z_k^{(l)}$ і точка x_r співпадають, то змінюємо точку $z_k^{(l+1)}$ якщо не співпадають – точку $x_{k+1}^{(l+1)}$: $z_{k+1}^{(l+1)}x_r$, де $i\ne k$.

У **другому** випадку, при $x_r \prec z_1^{(l)}$, коли знаки похибки у точках $z_1^{(l)}$ і x_r: $sign(\rho(z_1^{(l)})) = sign(\rho(x_r))$, то точку x_r вводимо замість точок альтернансу, а решту точок залишаємо без змін. Якщо ця умова не виконується, то Валле-Пуссен запропонував зсунути точку x_r вліво: $z_k^{(l+1)} = z_{k-1}^{(l)}$, $z_1^{(l+1)} = x_r$, щоб досягнути чергування знаків "+" і "–".

У **третьому** випадку, коли $x_r \succ z_{n+2}^{(l)}$, при виконані умови, вводиться $x_r = z_{n+2}^{(l+1)}$ і чергування знаків при цьому забезпечується. Найближчу точку до x_r ми зсуваємо, а решта залишаються без змін. При невиконанні умови $sign(\rho(z_{n+2}^{(l)})) = sign(\rho(x_r))$, здійснюємо зсув вправо, відкидаючи першу точку: $z_k^{(l+1)} = z_{k+1}^{(l)}$, де $k = \overline{1, n+1}$, $z_{n+2}^{(l+1)} = x_r$.

Встановлено [76-80, 88, 92-93], що при виконанні умов існування найкращого рівномірного наближення, алгоритм Валле-Пуссена для заміни точок альтернансу збігається за скінченне число кроків незалежно від початкового наближення до точок альтернансу. Він досить ефективний при наближенні виразами з невеликою кількістю параметрів m. Як правило, він збігається не більше ніж за k ітерацій.

Алгоритми з одночасною заміною декількох точок альтернансу характеризуються дещо швидшою збіжністю в порівнянні з алгоритмом Валле-Пуссена, що дається взнаки при наближенні виразами з великою кількістю невідомих параметрів. При обробці експериментальних даних використовуються в основному вирази з невеликою кількістю параметрів, тому й обмежились описом алгоритму Валле-Пуссена.

Зазначимо, що при оптимальному способі заміни точок альтернансу можна досягнути й квадратичної збіжності алгоритму Ремеза (за 1-2 ітерації) [76-80, 88, 92-93]. Оптимальна стратегія заміни точок альтернансу пропонує в якості нового наближення до точок альтернансу приймати k+1 точку з максимальним по модулю відхиленням і дотриманням чергування зміни знаку, тобто вибираються екстремальні точки з кожного інтервалу зміни знаку похибки наближення. При такій заміні точок альтернансу відхилення від оптимальної стратегії теоретично можливе лише тоді, коли інтервалів зміни знаку похибки наближення більше ніж k+1 і при виборі нових точок альтернансу помилково не враховується (відкидається) інтервал з ненайменшим значенням похибки.

6.9. Метод найменших квадратів багатовимірного регресійного аналізу в економіці

При аналізі діяльності промислових та сільськогосподарських підприємств, особливо тих, що працюють в умовах оренди, постійно виникає потреба в розв'язуванні різних економічних задач. Наприклад, необхідність правильно оцінити можливості наявних у господарстві виробничих, трудових та земельних ресурсів, порівняти й оцінити результати господарської діяльності підприємств і

районів з точки зору їх реальних можливостей. При плануванні подальшого розвитку окремих господарств важливо встановити правильну й, по можливості, оптимальну їх спеціалізацію, визначити оптимальні заходи в окремих галузях господарства, оптимального використання земельних площ тощо [76-80, 88, 92-93]. Для звітного періоду й періоду, що планується, необхідно вивчити економічну ефективність різних видів капіталовкладень та інших господарських заходів (раціонального використання якості землі, затрат виробництва, підвищення матеріальної зацікавленості в результатах роботи тощо.).

6.9.1. Загальна постановка задачі

При статистичному аналізі економічних показників господарської діяльності підприємств зміна, яка досліджується, як правило, залежить від впливу не одного, а багатьох факторів [76-80, 88, 92-93]. Наприклад, врожайність зернових культур залежить не лише від економічної оцінки землі, але й від затрат виробництва, норм внесення мінеральних та органічних добрив, наявних фондів, агротехніки тощо. Послідовне вивчення впливу кожного фактора на результуючу ознаку має певні недоліки. Зокрема, господарства з вищою оцінкою землі мають, як правило, більші затрати виробництва, основні фонди тощо. Таким чином, вивчаючи вплив економічної оцінки землі на врожайність зернових у господарствах з різною родючістю ґрунту, одержуємо різні значення урожайності. Отже, пояснити різницю у врожайності лише впливом родючості землі можна тільки умовно. Ця різниця обумовлена не тільки неоднорідністю оцінки землі, але й впливом деяких інших факторів, що кореляційно пов'язані з економічною оцінкою землі. В той же час показники впливу окремих факторів на результуючу змінну є досить значними.

Якщо досліджувати вплив іншого фактора (наприклад, затрат (виробництва) на врожайність зернових культур, одержимо, також умовні, досить високі показники. Якщо ж будемо розглядати спільний вплив двох факторів, то прийдемо до повторного розрахунку.

Продемонструємо останню особливість на прикладі послідовного використання парної регресії. При дослідженні впливу різних факторів на врожайність зернових культур (на основі даних господарств Львівської області) було одержано, що затрати виробництва на культуру пояснюють близько 35% дисперсії урожайності (коефіцієнт детермінації дорівнює 0,35), якість ріллі – 53%, норми внесення добрив – 40%, забезпеченість сільськогосподарськими машинами – 17% дисперсії урожайності зернових культур. Підсумовуючи дисперсії лише розглянутих факторів, одержуємо 35%+53%+40%+17%=145%. Таким чином, загальна дисперсія виходить більше ніж 100%. Причини подібних парадоксів полягають у постановці задачі. Результати послідовного аналізу не можна так об'єднувати, оскільки це призводить до подвійного розрахунку. В сукупності господарств, на основі даних яких були проведені розрахунки, коефіцієнт кореляції між оцінкою землі та забезпеченістю основними фондами дорівнює 0,19; між оцінкою землі та затратами виробництва – 0,41 тощо.

Відомо, що навіть кваліфікований дослідник, аналізуючи фактичні дані, завжди дещо перебільшує важливість досліджуваного фактора [76-80, 88, 92-93]. Таким

чином, виникає потреба в дослідженні впливу всіх факторів, що входять у модель, без повторного розрахунку. При цьому спільний вплив кількох факторів повинен бути розподілений між окремими факторними змінними з відповідними статистичними вагами. На практиці повністю реалізувати таку постановку неможливо, оскільки факторів, що впливають на результуючу ознаку, дуже багато і врахувати їх одночасний вплив неможливо. Необхідно на основі якісного аналізу й експертних оцінок відібрати лише такі факторні змінні, які найбільш суттєво впливають на результуючу змінну. Після такого обмеження числа факторів можна дослідити чистий вплив кожного з них на результуючу ознаку без повторного розрахунку взаємно-корельованих факторів. Один із найпростіших способів розв'язування поставленої задачі – аналітичне (комбіноване) групування, згідно з яким спочатку всі одиниці сукупності розподіляються за однією ознакою, далі в межах виділених таким чином груп – за іншою і т.д. При аналітичному групуванні кількість очікуваних груп визначається добутком числа інтервалів кожної групуючої ознаки. Наприклад, якщо провести групування за чотирма ознаками (виділяючи в кожній по 13 інтервалів), то в результаті одержимо 13*13*13*13=28561 групу. Внаслідок кореляції між групуючими ознаками більшість груп виявиться пустими. Відзначимо, що для вибірок малих обсягів у кожній із груп, які виявляться непустими, буде лише по одній-дві одиниці сукупності. Середні значення результуючої змінної, які обчислені на основі таких малочисельних груп, не будуть об'єктивними характеристиками господарської діяльності, яка очікувана від впливу факторних змінних.

Кореляційно-регресійні моделі досить об'єктивно описують соціально-економічні процеси за умови добре перевіреної та надійної інформації. При цьому, як модель у цілому, так і окремі її параметри, завжди можна оцінити з певною ймовірністю, що визначає надійність моделі на строго науковому рівні.

Кореляційні методи дають можливість кількісно вивчати і характеризувати діючі реально економічні закономірності та зв'язки. Ці закономірності менеджери мають змогу використовувати у різних галузях аналізу та прогнозування господарської діяльності.

Порівняно з простим групуванням кореляційно-регресійні методи мають суттєві переваги.

1. При використанні такого аналізу кілька спостережень замінюють таблиці даних, які необхідні для відображення результатів групування, що важливо при дослідженні взаємозв'язку між багатьма ознаками в різних комбінаціях.

2. Регресійні методи менш чутливі до різних випадкових причин, дію яких відображає вихідна інформація. Вони виявляють важливі закономірності у більш загальному, чистому вигляді.

3. Методи множинної та часткової кореляції дають можливість вивчати взаємозв'язки між кількома ознаками одночасно, що при використанні простого групування, зробити неможливо.

4. Кореляційні методи дають можливість визначити тісноту зв'язку у між факторами, що при використанні простого групування зробити неможливо.

5. Показники кореляції, які одержані на основі обмеженої вибірки, можуть бути оцінені в якості відповідних характеристик усієї генеральної сукупності, що неможливо зробити на основі простого групування.

Остання властивість кореляційно-регресійних методів особливо важлива при їх використанні для побудови економіко-статистичних моделей. Такі моделі, що одержано на основі статистичної обробки даних звітного періоду, дуже часто використовуються при плануванні та прогнозуванні. Звітний рік (роки) у цьому випадку вважається вибіркою з деякої множини років (включаючи рік, на який планується), що утворює генеральну сукупність.

Наведемо причини недостатнього використання кореляційно-регресійних методів при аналізі економічних процесів:

- розрахунки кореляційно-регресійних показників вимагають великої кількості обчислень і, як правило, здійснюються за допомогою сучасних комп'ютерів із використанням пакетів прикладних програм статистичного аналізу;
- складні математичні методи часто недоцільно використовувати внаслідок значних їх недоліків при кількісному визначенні різних економічних категорій, наприклад ціни, собівартості, чистого прибутку тощо, а також через наявність грубих похибок вихідної інформації;
- недостатня підготовленість менеджерів до широкого використання на практиці економіко-математичних методів та сучасних комп'ютерів.

6.9.2. Знаходження параметрів лінійного множинного рівняння регресії методом найменших квадратів

Багатофакторні зв'язки соціальних, економічних, біологічних явищ найчастіше моделюються за допомогою лінійного рівняння регресії. Рівняння, в яке входять одна результуюча змінна x_0 та довільне число факторних змінних $x_1,x_2,\ldots,x_k$, досить добре відображає дійсність тоді, коли фактори змінюються пропорційно до зміни результуючої ознаки. Причому, коефіцієнт пропорціональності для кожного фактора може бути довільним. На практиці найчастіше використовуються лінійні рівняння регресії з обмеженим числом від 3-х до 8-ми факторних змінних. Хоча лінійне множинне рівняння регресії досить точно описує більшість об'єктивно існуючих зв'язків економіки, зокрема, в сільському господарстві такими рівняннями не можна моделювати всі зв'язки. В загальному випадку можуть використовуватися складніші рівняння регресії, наприклад параболічні. Характер зв'язку і вибір типу рівняння регресії необхідно обгрунтовувати перед початком розрахунків.

Для k факторів лінійне рівняння множинної регресії має такий вигляд:

$$\tilde{x}_0 = b_0 + b_1x_1 + b_2x_2 + \ldots + b_kx_k, \qquad (6.136)$$

де x_0 – результуюча змінна; b_0 – вільний член рівняння регресії; $b_1,b_2,\ldots,b_k$ – коефіцієнти регресії, $x_1,x_2,\ldots,x_k$ – факторні змінні.

У випадку дії трьох факторів на результуючу змінну одержимо:

$$\tilde{x}_0 = b_0 + b_1x_1 + b_2x_2 + b_3x_3,$$

Як правило, рівняння регресії має лише одну результуючу змінну та довільне (скінченне) число факторних змінних. Збільшення числа результуючих ознак лише на одиницю значно ускладнює задачу аналізу й практично не розглядається. При дослідженні економічних показників майже всю інформацію несуть, як правило, три – вісім найсуттєвіших факторів. Для їх подальшого збільшення необхідно значно збільшити обсяг вихідних даних. У першому наближенні вважають, що число одиниць сукупності повинно бути принаймні в 10 разів більше від числа факторних змінних. Якщо ознаки тісно корелюють між собою, то й десятикратне перевищення числа одиниць сукупності над числом факторів може виявитися недостатнім.

Область існування множинної регресії визначається найменшим і найбільшим значеннями факторних ознак, що зустрічаються у вихідних даних.

Нехай обґрунтовано, що всі фактори лінійно пов'язані з результуючою змінною. Основна задача регресійного аналізу полягає в знаходженні параметрів рівняння

$$\widetilde{x}_0 = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k,$$

Для визначення параметрів $b_0, b_1, b_2, \ldots, b_k$ будемо використовувати метод найменших квадратів, суть якого полягає в знаходженні мінімуму функції

$$Q = \sum_{i=1}^{n} (x_{i0} - \widetilde{x}_{i0})^2 \to \min, \tag{6.137}$$

де n – число одиниць сукупності в кожній ознаці; x_{i0} – теоретичні (нормативні) значення результуючої змінної, знайдені на основі рівняння множинної регресії з урахуванням значень факторів $x_{i1}, x_{i2}, \ldots, x_{ik}$ в i-й одиниці сукупності.

Підставляючи у (6.137) замість x_{i0} відповідні значення (6.136), диференціюючи одержаний функціонал за параметрами $b_0, b_1, b_2, \ldots, b_k$ й прирівнюючи частинні похідні до нуля, одержимо систему нормальних рівнянь для множинної лінійної регресії:

$$\begin{aligned}
&b_0 n + b_1 \sum x_1 + b_2 \sum x_2 + \ldots + b_k \sum x_k = \sum x_0;\\
&b_0 \sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_1 x_2 + \ldots + b_k \sum x_1 x_k = \sum x_0 x_1;\\
&b_0 \sum x_2 + b_1 \sum x_1 x_2 + b_2 \sum x_2^2 + \ldots + b_k \sum x_2 x_k = \sum x_0 x_2;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&b_0 \sum x_k + b_1 \sum x_1 x_k + b_2 \sum x_2 x_k + \ldots + b_k \sum x_k^2 = \sum x_0 x_k;
\end{aligned} \tag{6.138}$$

Розв'язуючи систему лінійних алгебраїчних рівнянь (6.138) відносно $b_0, b_1, b_2, \ldots, b_k$, одержимо невідомі параметри рівняння багатовимірної регресії. Надалі межі підсумовування від i до n будемо опускати. Як вихідну інформацію для побудови системи нормальних рівнянь на основі вибірки необхідно обчислити такі суми:

$$n;\sum x_1;\sum x_2;...,\sum x_k;$$
$$\sum x_1^2;\sum x_2^2;\ ....,\sum x_k^2;$$
$$\sum x_0x_1;\sum x_0x_2;...,\sum x_0x_k;$$
$$.....$$
$$\sum x_{k-1}x_k.$$

Відзначимо, що оскільки $\sum x_i x_j = \sum x_j x_i$, то симетричні величини визначаються один раз.

Теорема 6.3. Параметри $\{b_0,b_1,...,\ b_n\} \subset \Re^1$ основної задачі регресійного аналізу знаходяться методом найменших квадратів як розв'язок СЛАР (6.83).

Систему нормальних рівнянь (6.138) будемо розв'язувати одним із відомих методів розв'язування систем лінійних алгебраїчних рівнянь (точні методи Гауса, метод простої ітерації, метод Зейделя тощо). Точність розрахунків перевіряють підстановкою одержаних параметрів у систему нормальних рівнянь.

Якщо система має більше чотирьох параметрів, то для знаходження невідомих параметрів $\{b_0,b_1,...,b_n\} \subset \Re^1$ множинних рівнянь регресії використовують сучасні комп'ютери, а обчислення здійснюють за допомогою пакету STATISTICA або оболонки MATLAB. Для знаходження параметрів регресії систему нормальних рівнянь можна записати не тільки на основі сум, сум квадратів і сум парних добутків вихідних даних, але й в іншому вигляді, наприклад сум, сум квадратів й сум парних добутків відхилень від середніх арифметичних; дисперсій і коваріацій; коефіцієнтів парної кореляції. Якщо вихідні спостереження (вибірка) задаються у вигляді відхилень від середніх арифметичних, то система нормальних рівнянь (6.138), враховуючи, що $\sum(x_i - \widetilde{x}_j) = 0$, набуде вигляду:

$$\begin{aligned}
&b_1\sum(x_1-\widetilde{x}_1)^2 + b_2\sum(x_1-\widetilde{x}_1)(x_2-\widetilde{x}_2)+...+\\
&+b_k\sum(x_1-\widetilde{x}_1)(x_k-\widetilde{x}_k) = \sum(x_0-\widetilde{x}_0)(x_1-\widetilde{x}_1);\\
&b_1\sum(x_1-\widetilde{x}_1)(x_2-\widetilde{x}_2)+b_2\sum(x_2-\widetilde{x}_2)+...+\\
&+b_k\sum(x_2-\widetilde{x}_2)(x_k-\widetilde{x}_k) = \sum(x_0-\widetilde{x}_0)(x_2-\widetilde{x}_2);\\
&.....\\
&b_1\sum(x_1-\widetilde{x}_1)(x_k-\widetilde{x}_k)+b_2\sum(x_2-\widetilde{x}_2)(x_k-\widetilde{x}_k)+...+\\
&+b_k\sum\ (x_k-\widetilde{x}_k)^2 = \sum(x_0-\widetilde{x}_0)(x_1-\widetilde{x}_1);
\end{aligned} \qquad (6.139)$$

Розв'язуючи систему (6.139), одержуємо коефіцієнти множинної регресії. Вільний член $b_0 = \widetilde{x}_0 - b_1x_1 - ... - b_k\widetilde{x}_k$, де $\widetilde{x}_j = \sum\frac{x_j}{n}, j = 0,1,2,...,k.$

Означення 6.2. *Знаходження параметрів рівняння регресії на основі системи (6.84) називають також методом відхилень.*

Розділивши всі рівняння системи (6.139) на обсяг вибірки n, одержимо систему нормальних рівнянь:

$$\begin{aligned} & b_1 s_1^2 + b_2 \operatorname{cov}_{12} + \ldots + b_k \operatorname{cov}_{1k} = \operatorname{cov}_{01}; \\ & b_1 \operatorname{cov}_{12} + b_2 s_2^2 + \ldots + b_k \operatorname{cov}_{2k} = \operatorname{cov}_{02}; \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ & b_1 \operatorname{cov}_{1k} + b_2 \operatorname{cov}_{2k} + \ldots + b_k s_k^2 = \operatorname{cov}_{0k}; \end{aligned} \qquad (6.140)$$

$$s_j^2 \equiv \frac{\sum (x_j - \widetilde{x}_j)^2}{n} = \frac{\sum x_j^2}{n} - (\widetilde{x}_j)^2;$$

$$\operatorname{cov}_{ij} \equiv \frac{\sum (x_i - \widetilde{x}_i)(x_j - \widetilde{x}_j)}{n} = \frac{\sum x_i x_j}{n} - \widetilde{x}_i.$$

Визначення параметрів регресії за допомогою системи (6.140) будемо називати *способом коваріацій*, оскільки величину $\mathbf{cov}_{ij}$ у статистичній літературі називають *змішаним центральним моментом другого порядку* або *коваріацією*. Вільний член рівняння регресії знаходимо так само, як і в попередньому випадку.

Розділимо, нарешті, всі члени першого рівняння системи (6.140) на добуток дисперсій $s_0 s_1$, а другого – на $s_0 s_2$ і т.д. У результаті одержимо

$$\begin{aligned} & \beta_1 + \beta_2 r_{12} + \ldots + \beta_k r_{1k} = r_{01}; \\ & \beta_1 r_{1k} + \beta_2 + \ldots + \beta_k r_{2k} = r_{0k}; \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ & \beta_1 r_{1k} + \beta_2 r_{2k} + \ldots + \beta_k = r_{0k}; \end{aligned} \qquad (6.141)$$

Означення 6.3. *Метод побудови систем нормальних рівнянь (6.141) назвемо методом кореляційної матриці.*

Розв'язуючи систему (6.141), одержуємо коефіцієнти множинної регресії $\beta_1, \beta_2, \ldots, \beta_k$ у стандартизованому вигляді. Вони характеризують вихідні спостереження, які подані в стандартних відхиленнях від відповідних середніх та мають вигляд $z_j = \dfrac{(x_j - \widetilde{x}_j)}{s_j}, j = 0,1,\ldots,k.$

Якщо рівняння регресії в початкових одиницях вже знайдено, то стандартизовані коефіцієнти регресії визначаємо, очевидно, за формулами $\beta_i = b_j \dfrac{s_j}{s_0}; b_j = \beta \dfrac{s_0}{s_j}.$

Приклад 6.4. Розглянемо задачу залежності врожайності зернових x_0 від впливу основних факторів (економічної оцінки землі x_1 і затрат виробництва x_2), що наведені у табл. 6.11.

Таблиця 6.11

Вихідні дані для визначення впливу економічної оцінки землі та затрат виробництва на врожайність зернових на основі річних звітів господарств Львівської області

Номер господарства	Назва	Врожайність зернових, ц/га	Економічна оцінка землі, бал	Затрати виробництва на 1 га, грн.
		x_0	x_1	x_2
1	ім. В. Чорновола	38	27	280
2	«Дружба»	37	39	238
3	ім. Ю.Федьковича	38	27	362
4	«Правда»	53	77	352
5	ім. Д. Галицького	40	46	293
6	«Чумацький шлях»	38	66	219
7	«28 червня»	43	68	315
8	ім. Л. Українки	37	64	277
9	«Україна»	38	58	331
10	ім. Івана Франка	35	58	250
11	«Незалежності»	35	72	262
12	ім. В. Стефаника	43	73	265
13	ім. В. Стуса	53	74	426
14	ім. Т. Шевченка	56	63	484
14	Всього	584	812	4354

Розв'язок. Врожайність зернових x_0 в ц/га слід шукати у вигляді $x_0=b_0+b_1x_1+b_2x_2$, де x_1 – економічна оцінка землі, бал; x_2 – затрати виробництва на 1 га, грн.

Система нормальних рівнянь має вигляд:

$$b_0 n + b_1 \sum x_1 + b_2 \sum x_2 = \sum x_0;$$

$$b_0 \sum x_1 + b_1 \sum x_1^2 + b_2 \sum x_1 x_2 = \sum x_0 x_1;$$

$$b_0 \sum x_2 + b_1 \sum x_1 x_2 + b_2 \sum x_2^2 = \sum x_0 x_2;$$

Таблиця 6.12

n	x_0	x_1	x_2	x_0^2	x_1^2	x_1x_2	x_0x_1	x_2^2	x_0x_2
1	38	27	280	1444	729	7560	1026	78400	10640
2	37	39	238	1369	1521	9282	1443	56644	8806
3	38	27	362	1444	729	9774	1026	131044	13756
4	53	77	352	2809	5929	27104	4081	123904	18656
5	40	46	293	1600	2116	13478	1840	85849	11720
6	38	66	219	1444	4356	14454	2508	47961	8322
7	43	68	315	1849	4624	21420	2924	99225	13545
8	37	64	277	1369	4096	17728	2368	76729	10249
9	38	58	331	1444	3364	19198	2204	109561	12578
10	35	58	250	1225	3364	14500	2030	62500	8750

11	35	72	262	1225	5184	18864	2520	68644	9170
12	43	73	265	1849	5329	19345	3139	70225	11395
13	53	74	426	2809	5476	31524	3922	181476	22578
14	56	63	484	3136	3969	30492	3528	234256	27104
14	584	812	4354	25016	50786	254723	34559	1426418	187269

$$14b_0 + 812b_1 + 4354b_2 = 584$$
$$812b_0 + 50786b_1, + 254723b_2 = 34559;$$
$$4354b_0 + 254723b_1 + 1426418b_2 = 187269.$$

Розв'язуючи це рівняння (наприклад, методом Гаусса), одержуємо коефіцієнти регресії: $b_0 = 10.5344$; $b_1 = 0.1424$; $b_2 = 0.0737$.

Як правило, обчислення виконують на персональних комп'ютерах з використанням пакету «СТАТИСТИКА» або MATLAB. Вихідні дані бажано мати з 3-4 значущими цифрами. Таким чином, шукане рівняння регресії набуває вигляду

$$\widetilde{x}_0 = 10.534 + 0.142x_1 + 0.074x_2.$$

З практики відомо, що коефіцієнти парної та множинної регресій, визначені на основі однакових даних для однієї факторної змінної, майже завжди різняться. Зокрема, для даного прикладу парне рівняння регресії залежності врожайності зернових x_0 від економічної оцінки землі x_1 має вигляд $\widetilde{x}_0 = 30.9 + 0.19x_1$.

Коефіцієнти при x_1, показують середній вплив якості землі на врожайність зернових. У парному рівнянні коефіцієнт регресії на 34% більший за відповідний коефіцієнт множинної регресії. Неоднаковість коефіцієнтів рівнянь простої і множинної регресії пояснюється наявністю кореляційного зв'язку між факторними змінними. Зокрема, для цього прикладу коефіцієнт кореляції між економічною оцінкою землі та затратами виробництва $r_{12} = 0.13$.

Висновок 1. Коефіцієнт парного (простого) рівняння регресії відображає умовний вплив фактора, що аналізується, на результуючу змінну. Причому умовність впливу факторної змінної полягає в тому, що вона досить точно відображає дійсність, коли її вплив інтерпретується як результат дії даної ознаки разом з іншими факторами, кореляційно пов'язаними з факторною змінною, що досліджується. Іншими словами, коефіцієнт парного рівняння регресії показує вплив деякої фіксованої факторної змінної на результуючу ознаку при наявності певного «фону» решти факторів, причому цей «фон» може змінюватися при зміні аналізуючої ознаки. Наприклад, підвищення якості землі на 1 бал при відповідному покращенні затрат виробництва й інших факторів повинно забезпечити середній приріст врожайності зернових на 0,19 ц/га (тобто на коефіцієнт простого рівняння регресії).

Висновок 2. Коефіцієнт множинного рівняння регресії відображає чистий вплив факторної змінної, якщо регресійна модель охоплює всі фактори, які впливають на результуючу змінну. Тоді такий вплив факторів фактично розподіляється між ними з від повідними вагами, а «фон» дії кожного фактора вважається фіксованим на певному постійному рівні. Реальне число факторів, які впливають на результуючу змінну, настільки велике, що всі їх врахувати в регресійній моделі неможливо. Звичайно приймають, що серед усіх факторних змінних є кілька (від 3 до 5) найважливіших, а вплив решти факторів несуттєвий. Відносно до основних

факторів «фон» є фіксованим, а відносно до решти факторів, які не включені в регресійну модель, він залишається змінним (за аналогією з «фоном» парного рівняння регресії). Таким чином, необхідно пам'ятати, що коефіцієнт множинного рівняння регресії описує умовно чистий вплив факторної змінної. Наприклад, збільшення якості землі на 1 бал без додаткових затрат виробництва приводить до середнього приросту врожайності зернових на 0,14 ц/га (на коефіцієнт множинної регресії).

Якщо в багатовимірну регресійну модель включити ще один (третій) фактор, то степінь умовності коефіцієнтів регресії зменшиться. Якщо фактори взаємно зв'язані додатньою кореляцією, то в міру зниження ступеня умовності зменшуються коефіцієнти регресії

Зауваження 1. Інтерпретація коефіцієнтів регресії тісно пов'язана з одиницями результуючої та факторної змінних. Якщо змінити одиницю однієї чи обох змінних, то відповідно зміняться коефіцієнти регресії. Зокрема, якщо одиницю результуючої змінної зменшити (збільшити) в k разів, то праву частину рівняння регресії необхідно збільшити (зменшити) в k разів. Якщо ж збільшуємо (зменшуємо) одиницю лише однієї факторної змінної, то на k слід помножити (поділити) тільки відповідний коефіцієнт регресії.

Зауваження 2. При побудові множинного рівняння регресії вихідну інформацію необхідно перетворити так, щоб усі дані були числами одного порядку. Останнє дає можливість краще інтерпретувати економічний зміст окремих коефіцієнтів регресії.

Оскільки, в загальному випадку коефіцієнти регресій не можна співставляти, то для порівняльної оцінки впливу факторів на результуючу змінну їх необхідно подати в стандартизованому вигляді. Якщо рівняння регресії в початкових одиницях вже знайдено, то стандартизовані коефіцієнти регресії знайти порівняно легко. Зокрема, для даного прикладу на основі вихідних даних (b_1 =0.1424; b_2 =0.0737; s_0 =6.8131; s_2 =71.8749) одержуємо, що β_1=0.3393; β_2 =0.7775. У стандартизованому вигляді рівняння регресії записується $\tilde{z}_0 = 0.3393 z_1 + 0{,}7775 z_2$,

де $z_j = \dfrac{x_j - \tilde{x}_j}{s_j}, j = 1,2..$

Вільний член рівняння регресії дорівнює нулю, оскільки всі змінні виражені у відхиленнях від середніх значень.

Висновок 3. Стандартизовані коефіцієнти регресії показують, на скільки стандартних одиниць зміниться в середньому результуюча змінна при зміні факторної ознаки на одну стандартну одиницю. Зокрема, зміна оцінки землі на одну одиницю стандартного відхилення в даному прикладі забезпечує приріст врожайності в середньому на 0,3393 стандартних одиниць, а зміна на одне стандартне відхилення затрат виробництва забезпечує приріст врожайності на 0,7775 стандартних відхилень.

Зауваження 3. Стандартні коефіцієнти множинного рівняння регресії можна співставляти. Для цього прикладу таке співставлення показує, що на врожайність зернових найбільше впливають затрати виробництва, а найменше – якість землі, оскільки при z_2 коефіцієнт є більшим, ніж при Z_1.

Нарешті, розглянемо інтерпретацію теоретичної величини результуючої змінної. Найчастіше в статистичному аналізі багатофакторні рівняння регресії використовуються для оцінки результатів роботи окремих одиниць сукупності (господарств, підприємств тощо). Основна ідея такої оцінки полягає в тому, що в статистичному аналізі можна врахувати вплив на результуючу змінну відразу кількох факторів, наприклад, вплив якості землі та затрат виробництва на урожайність дернових. Із сукупності господарств, наведених у табл. 1, розглянемо колгосп ім. В.Стефаника. Якість землі цього колгоспу оцінюється 73 балами, а витрати виробництва складають у середньому 265 грн. з розрахунку на 1га. Підставляючи факторні рівняння регресії, одержуємо $x_0 = 40.51$ ц/га.

Одержана врожайність є нормативною (теоретичною) при заданих факторних змінних та середньому рівні їх використання. Фактична урожайність становить 43 ц/га, тобто є більшою від нормативної на 43–40,51 = 1,49 ц/га. Таким чином, розглянуті два фактори, що включені в регресійну модель, використовуються в цьому господарстві вище середнього рівня.

В той самий час на основі використання парного рівняння регресії, в якому враховується лише один фактор (економічна оцінка землі), теоретична врожайність x_0 =44,77 ц/га і є більшою від фактичної на 44,77–43 = 1,77 ц/га. Отже, використання парного рівняння регресії показує, що ресурси господарства використовувалися нижче середнього рівня (від'ємні відхилення фактичної урожайності від теоретичної становлять 1,77 ц/га). У той самий час багатофакторна модель показала, що ресурси господарства, навпаки, використовувалися вище середнього рівня (додатні відхилення фактичних значень від *теоретичних* дорівнюють 1,49 ц/га). Оскільки більш точною є багатофакторна регресійна модель, то побудова моделі врожайності на основі простого (парного) рівняння регресії за даними прикладу, що розглядається, може привести до неправильної оцінки роботи господарства.

Висновок 4. Якщо рівняння множинної регресії містить усі основні фактори виробництва, то теоретичне (нормативне) значення результуючої змінної характеризує потенціал цього господарства, враховуючи встановлену організацію виробництва. Високий теоретичний рівень результуючої ознаки свідчить про значний виробничий потенціал господарства і, навпаки, невеликі значення результуючої змінної свідчать про наявність у господарстві невикористаних резервів. Відхилення фактичного значення результуючої змінної від теоретичного відображає результат дії факторів, що не входять у рівняння множинної регресії. Якщо регресійна модель містить усі фактори, що впливають на результуючу ознаку (тобто об'єктивно обумовлюють результати виробництва), то відхилення фактичних значень результуючої змінної від теоретичних вважаються результатом дії суб'єктивних (організаційних) факторів. Таким чином, відхилення фактичного значення результуючої ознаки від відповідного теоретичного значення, знайденого на основі рівняння множинної регресії, свідчить про рівень використання наявних виробничих факторів та умов.

6.9.3. Стандартна похибка оцінки за рівнянням

Стандартна похибка оцінки за рівнянням множинної регресії визначається так само, як і для парного рівняння регресії. Спочатку знаходять наближену дисперсію:

$$s^2_{0,1,2,\ldots,k} \equiv \frac{\sum (x_0 - \widetilde{x}_0)^2}{n}, \tag{6.142}$$

де x_0 – теоретичне значення результуючої змінної, яке обчислене на основі множинного рівняння регресії.

Незміщена оцінка цієї дисперсії дорівнює добутку зміщеної оцінки на відповідний коефіцієнт [76-80, 88, 92-93, 138, 142, 163]:

$$\bar{s}^2_{0,1,2,\ldots,k} \equiv s^2_{0,1,2,\ldots,k} \frac{n}{n-(k+1)}, \tag{6.143}$$

де k – число факторів у рівнянні множинної регресії.

Стандартна похибка оцінки за рівнянням визначається як квадратний корінь із наближеної дисперсії:

$$s_{0,1,2,\ldots,k} \equiv \sqrt{\frac{\sum (x_0 - \widetilde{x}_0)^2}{n}}, \quad \bar{s}_{0,1,2,\ldots,k} \equiv \sqrt{\bar{s}^2_{0,1,2,\ldots,k}}. \tag{6.144}$$

На практиці найчастіше використовується інша формула для обчислення стандартної похибки оцінки за рівнянням регресії [76-80, 88, 92-93, 138, 142, 163]:

$$s^2_{0,1,2,\ldots,k} \equiv \frac{\sum x_0{}^2 - b_0 \sum x_0 - b_1 \sum x_0 x_1 - \ldots - b_k \sum x_0 x_k}{n} \tag{6.145}$$

Якщо відомі показники коваріації та дисперсії результуючої змінної, то формула (6.142) набуває вигляду

$$s^2_{0,1,2,\ldots,k} \equiv s^2_0 - b_1 \operatorname{cov}_{01} - b_2 \operatorname{cov}_{02} - \ldots - b_k \operatorname{cov}_{0k}. \tag{6.146}$$

Економічна інтерпретація стандартної похибки оцінки для рівняння регресії така сама, як і для парного рівняння регресії. Величина $s_{0,1,2,\ldots,k}$ характеризує варіацію фактичних даних навколо теоретичних (нормативних), одержаних на основі множинного рівняння регресії.

Ця величина виражається в тих же одиницях, що й результуюча змінна. Якщо зв'язок між результуючою та факторними змінними є функціональним, то $s_{0,1,2,\ldots,k} = 0$. Якщо ж такий зв'язок відсутній, то $s_{0,1,2,\ldots,k}$ приймає найбільші значення.

Гранична похибка $\Delta_{0,1,2,\ldots,k}$ дорівнює добутку стандартної похибки на ймовірнісний коефіцієнт t_p:

$$\Delta_{0,1,2,\ldots,k} = t_p s_{0,1,2,\ldots,k} \tag{6.147}$$

На основі граничної похибки визначають довірчі інтервали: нижню межу – як різницю між теоретичним значенням результуючої ознаки і граничною похибкою, верхню – як суму даних величин: $\widetilde{x}_0 - \Delta_{0,1,2,\ldots,k} \le x_{0i} \le \widetilde{x}_0 + \Delta_{0,1,2,\ldots,k}$.

Приклад 6.5. Обчислити стандартну похибку для множинного рівняння регресії, побудованого на основі даних прикладу 6.4.

Роз'язок. Підставляючи дані $\sum x_0 = 584$, $\sum x_0^2 = 25016$, $\sum x_0 x_1 = 34559$, $\sum x_0 x_2 = 187269$ у формулу стандартної похибки оцінки за рівнянням регресії, записану через суми, одержуємо

$$s^2_{0,1,2} = \frac{\sum x_0{}^2 - b_0 \sum x_0 - b_1 \sum x_0 x_1 - b_2 \sum x_0 x_2}{n} = 11{,}8081,$$

$$\bar{s}^2_{0,1,2} = 15{,}0285,\ \bar{s}_{0,1,2} = 3{,}8767 \text{ ц/га}.$$

Співставляючи незміщену стандартну похибку множинного рівняння регресії (3,9 ц/га) із стандартним відхиленням результуючої змінної (S_0=6.8 ц/га) одержуємо, що в результаті виключення із впливу на врожайність зернових двох факторів (економічної оцінки землі та затрат виробництва) наближена варіація (стандартне відхилення) зменшується на 6,8 - 3,9 = 2,9 ц/га.

Таким чином, на врожайність зернових, крім економічної оцінки землі та витрат виробництва, впливають також інші факторні змінні, не враховані в цій моделі, серед яких важливе значення мають суб'єктивні (організаційні) факторні ознаки.

6.9.4. Коефіцієнт детермінації й кореляції

Як і коефіцієнт парної детермінації, коефіцієнт множинної детермінації, дорівнює відношенню поясненої дисперсії $\delta^2_{0,1,2,\ldots,k}$ до загальної дисперсії s_0^2:

$$R^2_{0,1,2,\ldots,k} = \frac{\delta^2_{0,1,2,\ldots,k}}{s_0^2} = \frac{s_0^2 - s^2_{0,1,2,\ldots,k}}{s_0^2} = 1 - \frac{s^2_{0,1,2,\ldots,k}}{s_0^2}. \qquad (6.148)$$

Коефіцієнт множинної кореляції визначається як квадратний корінь з коефіцієнта детермінації:

$$R_{0,1,2,\ldots,k} = \sqrt{1 - \frac{s^2_{0,1,2,\ldots,k}}{s_0^2}}. \qquad (6.149)$$

Коефіцієнт множинної кореляції має такі ж властивості, як і відповідний парний коефіцієнт. Зокрема, якщо $s^2_{0,1,2,\ldots,k}$=0, то всі фактичні дані збігаються з теоретичними і взаємозв'язок є функціональним, а коефіцієнт детермінації дорівнює одиниці. Якщо $s^2_{0,1,2,\ldots,k} = s_0^2$, то взаємозв'язок між результуючою та факторними змінними відсутній. Крім того, на відміну від коефіцієнта простої кореляції, коефіцієнт множинної кореляції не має знаку, оскільки в рівнянні регресії можуть бути одночасно як додатні, так і від'ємні коефіцієнти.

При вивченні залежності врожайності зернових культур від економічної оцінки землі та затрат виробництва наближена дисперсія $s^2_{0,1,2}$ =11.808, загальна дисперсія s^2_0 =46.418. Тоді коефіцієнт детермінації $R^2_{0,1,2}$ =0.7456. Таким чином, близько 75% всієї дисперсії врожайності зернових пояснюється варіацією розглянутих факторів: економічною оцінкою землі та затратами виробництва. Решта 25% викликана дією інших факторів, не врахованих у даній регресивній моделі.

Коефіцієнт множинної кореляції $R^2_{0,1,2}$ =0.863 і, як правило, є більшим від коефіцієнтів парних кореляцій, які характеризують зв'язок між результуючою змінною та окремими факторами. Якщо показники коваріації відомі, то коефіцієнт множинної детермінації визначаємо за формулою:

$$R^2_{0,1,2,\ldots,k} = \frac{b_1 \operatorname{cov}_{01} + b_2 \operatorname{cov}_{02} + \ldots + b_k \operatorname{cov}_{0k}}{s_0^2}.$$

Враховуючи, що $b_j = \beta_j \dfrac{s_0}{s_j}$, одержимо

$$R^2_{0,1,2,\ldots,k} = b_1 \frac{\operatorname{cov}_{01}}{s_0^2} + b_2 \frac{\operatorname{cov}_{02}}{s_0^2} + \ldots + b_k \frac{\operatorname{cov}_{0k}}{s_0^2} =$$

$$= \beta_1 \frac{\operatorname{cov}_{01}}{s_0 s_1} + \beta_2 \frac{\operatorname{cov}_{02}}{s_0 s_2} + \ldots + \beta_k \frac{\operatorname{cov}_{0k}}{s_0 s_k}.$$

Нарешті, оскільки $r_{0j} = \dfrac{\operatorname{cov}_{0j}}{s_0 s_j}$, то $R^2_{0,1,2,\ldots,k} = \beta_1 r_{01} + \beta_2 r_{02} + \ldots + \beta_k r_{0k}$.

6.9.5. Вибіркова похибка коефіцієнта множинної регресії

Квадрат стандартної похибки коефіцієнта парної лінійної регресії визначаємо за формулою

$$s^2_{b_{y,x}} = \frac{\bar{s}^2_{y,x}}{s_x^2 n}. \qquad (6.150)$$

Аналогічно визначається квадрат похибки коефіцієнта множинної регресії, який також залежить від стандартної похибки за рівнянням, дисперсії факторної змінної та числа одиниць сукупності. Крім того, на похибку коефіцієнта множинної регресії впливає також кореляційний зв'язок між факторами. Щоб визначити зв'язок між факторними змінними, в формулу (6.150) необхідно ввести додатковий множник

$$\frac{1}{1 - R^2_{j,1,2,\ldots,(j-1),(j+1),\ldots,k}},$$

який збільшує стандартну похибку при наявності кореляційного зв'язку між факторами. Отже, формула для знаходження квадрату стандартної похибки коефіцієнта множинної регресії має вигляд

$$s^2_{b_{0j},1,2,\ldots,(j-1),(j+1),\ldots,k} = \frac{s^2_{y,x}}{s^2_j n} \frac{1}{1-R^2_{j,1,2,\ldots,(j-1),(j+1),\ldots,k}} \qquad (6.151)$$

де $s^2_{b_{0j},1,2,\ldots,(j-1),(j+1),\ldots,k}$ – квадрат стандартної похибки коефіцієнта множинної регресії j-го фактора для цієї результуючої змінної та при наявності в моделі інших факторів з індексами $1,2,\ldots,(j-1),(j+1),\ldots,k$; $s^2_{0,1,2,\ldots,k}$ – квадрат стандартної похибки множинного рівняння регресії; s^2_j – дисперсія j-го фактора; n – число одиниць сукупності; $R^2_{j,1,2,\ldots,(j-1),(j+1),\ldots,k}$ – квадрат множинного коефіцієнта кореляції, причому j-й фактор вважається умовною результуючою ознакою (функцією), а інші факторні змінні з індексами $1,2,\ldots,(j-1),(j+1),\ldots,k$ – аргументами.

Якщо j-й фактор, для якого визначається коефіцієнт множинної регресії, кореляційно не пов'язаний з іншими (з індексами $1,2,\ldots,(j-1),(j+1),\ldots,k$) факторами, то $R^2_{j,1,2,\ldots,(j-1),(j+1),\ldots,k}=0$ і похибка коефіцієнта множинної регресії буде найменшою.

Якщо фактор з індексом j майже повністю визначається іншими факторними змінними, то $R^2_{j,1,2,\ldots,(j-1),(j+1),\ldots,k}$ прямує до 1 і стандартна похибка коефіцієнта регресії є як завгодно великою, а сам коефіцієнт втрачає пізнавальну цінність. У граничному випадку при $R^2_{j,1,2,\ldots,(j-1),(j+1),\ldots,k}=0$ система нормальних рівнянь стає виродженою і не має розв'язку (тобто має місце повна мультиколінеарність). Отже, чим більше значення $R^2_{j,1,2,\ldots,(j-1),(j+1),\ldots,k}=0$, тим більша стандартна похибка коефіцієнта множинної регресії й менша його пізнавальна цінність і, навпаки, чим менше значення $R^2_{j,1,2,\ldots,(j-1),(j+1),\ldots,k}=0$, тим більша пізнавальна цінність коефіцієнта множинної регресії. Відзначимо також, що при збільшенні числа факторних змінних (більше п'яти) збільшується обсяг обчислень.

Приклад 6.6. Знайти стандартні похибки коефіцієнтів регресії множинного рівняння, одержаного на основі даних прикладу 6.4 (b_1=0,142 і b_2 =0,074). Для цього обчислимо $s^2_{0,1,2}=15{,}028$; $s^2_1=16.235$; $s^2_2=71.875$; n=14; $r_{12}=0.134$.

Розв'язок. Підставляючисдані значення в (6.151) і враховуючи, що $R^2_{j,1,2,\ldots,(j-1),(j+1),\ldots,k}=r_{12}$ одержуємо $s^2_{b_1}=0.0673$; $s_{b_1}=0.2595$; $s^2_{b_2}=0.0152$; $s_{b_2}=0.1233$.

Так, як і для парної регресії, стандартні похибки коефіцієнтів множинної регресії використовують для знаходження граничних похибок і довірчих границь коефіцієнтів регресії. Гранична похибка визначається як добуток стандартної похибки на ймовірнісний коефіцієнт. Довірчі границі обчислюються шляхом віднімання граничної похибки від коефіцієнта регресії (нижня границя) і додавання до коефіцієнта регресії (верхня границя).

6.9.6. Вибіркова похибка множинної регресії

Якщо відомі вибіркові похибки окремих параметрів рівняння множинної регресії, то так саме, як і для парного рівняння регресії, можна побудувати вибіркову похибку цієї регресії. З точки зору геометричної інтерпретації це означає побудову в $(k+1)$-мірному просторі довірчої області, в якій з певною (заданою) ймовірністю буде знаходитися гіперплощина регресії генеральної сукупності. При цьому враховуються вибіркові похибки як коефіцієнтів регресії $b_1, b_2, ..., b_k$, так і вільного члена рівняння b_0. Ця задача рівнозначна задачі знаходження вибіркових похибок теоретичних значень результуючої ознаки, яка обчислена за рівнянням регресії.

Теоретичне значення результуючої змінної

$$\tilde{x}_0 = x_0 + b_1(x_1 - x_1) + b_2(x_2 - x_2) + ... + b_k(x_k - x_k) \qquad (6.152)$$

Таким чином, з рівняння (6.152) випливає, що вибіркова похибка x_0 залежить від похибки середнього арифметичного Δ_{x_0} й похибок коефіцієнтів регресії $b_1, b_2, ..., b_k$.

Нехай коефіцієнти $b_1, b_2, ..., b_k$ – попарно некорельовані й вибіркові похибки обчислюються на основі дисперсій. Тоді, враховуючи теорему про суму дисперсій, загальну дисперсію вважаємо рівною сумі дисперсій доданків. Причому, оскільки в (6.152) коефіцієнти регресії $b_1, b_2, ..., b_k$ перемножуються відповідно на постійні величини $(x_1 - \Delta_{x_1}), (x_2 - \Delta_{x_2}), ..., (x_k - \Delta_{x_k})$, які не мають вибіркової похибки, то при визначенні загальної вибіркової похибки дисперсії окремих доданків необхідно домножити відповідно на квадрат постійної величини $(x_i - \Delta_{x_i})$, $i = \overline{1,k}$.

З врахуванням цього квадрат стандартної похибки теоретичного значення $s^2_{\Delta_{x_0},1,2,...,k}$, обчислюємо за формулою

$$s^2_{\tilde{x}_0,1,2,...,k} = s^2_{x_0,1,2,...,k} + s_1^2(x_1 - x_1)^2 + s_2^2(x_2 - x_2)^2 + ... + s_k^2(x_k - x_k)^2. \qquad (6.153)$$

Підставляючи в (6.153) замість $s^2_{\Delta_{x_0},1,2,...,k}$, $s_1^2, ...,$ їх відповідні значення, одержуємо

$$s^2_{\tilde{x}_0,1,2,...,k} = s^2_{0,1,2,...,k}\left[\frac{1}{n} + \frac{(x_1 - x_1)^2}{s_1^2 n} - \frac{1}{1 - R^2_{1,2,3,...,k}} + \right.$$
$$\left. + \frac{(x_2 - x_2)^2}{s_2^2 n} - \frac{1}{1 - R^2_{2,1,3,...,k}} + ... + \frac{(x_k - x_k)^2}{s_k^2 n} - \frac{1}{1 - R^2_{k,1,2,3,...,(k-1)}}\right], \qquad (6.154)$$

де $s^2_{0,1,2,...,k} = s^2_{0,1,2,...,k}\frac{n}{n-(k+1)}$ (див. 6.8.3).

Звідси, стандартна похибка теоретичного значення визначається таким чином

$$s_{x_0,1,2,\ldots,k} = s_{0,1,2,\ldots,k}\left[\frac{1}{n} + \frac{(x_1 - x_1)^2}{s_1^2 n} - \frac{1}{1 - R^2_{1,2,3,\ldots,k}} + \right.$$
$$\left. + \frac{(x_2 - x_2)^2}{s_2^2 n} - \frac{1}{1 - R^2_{2,1,3,\ldots,k}} + \ldots + \frac{(x_k - x_k)^2}{s_k^2 n} - \frac{1}{1 - R^2_{k,1,2,3,\ldots,(k-1)}}\right]^{\frac{1}{2}}, \quad (6.155)$$

Гранична похибка $\Delta_{x_0,1,2,\ldots,k}$ обчислюється як добуток стандартної похибки на ймовірнісний коефіцієнт t_p. Необхідно відрізняти стандартну похибку теоретичного значення $s_{\Delta_{x_0,1,2,\ldots,k}}$ (стандартну похибку множинної регресії) та стандартні похибки оцінки за рівнянням $s_{0,1,2,\ldots,k}$. Величина $s_{\Delta_{x_0,1,2,\ldots,k}}$ відображає лише похибку вибірки, тобто сукупність спостережень, в якій досліджується зв'язок між результуючою змінною Δ_{x_0} та факторними змінними $x_1, x_2, \ldots, x_k$. Величина $s_{0,1,2,\ldots,k}$ характеризує непояснену варіацію, яка присутня завжди, коли зв'язок між змінними Δ_{x_0} і $x_1, x_2, \ldots, x_k$ є кореляційним незалежно від того, чи є сукупність вибіркою, чи генеральною сукупністю.

Приклад 6.7. Обчислити теоретичне значення Δ_{x_0} для значень аргументів x_1=60 та x_2=300 для прикладу 6.4, а також довірчі границі з імовірністю p=0.95. Теоретичне значення результуючої змінної x_0=41.25.

Розв'язок. Стандартна похибка даної оцінки визначається за формулою (6.155). Раніше були отримані значення n=14; x_1=58; x_2=311; $s_{0,1,2}$ =3.877. Дисперсії та коефіцієнти кореляцій факторних змінних відповідно s_1^2=263.571; s_2^2=5166; $R^2_{1,2} = r^2_{1,2} = R^2_{2,1} = r^2_{2,1}$ =0.134. Звідки

$$s_{x_0,1,2} = s_{0,1,2}\left[\frac{1}{n} + \frac{(x_1 - x_1)^2}{s_1^2 n} - \frac{1}{1 - R^2_{1,2}} + \frac{(x_2 - x_2)^2}{s_2^2 n} - \frac{1}{1 - R^2_{2,1}}\right]^{\frac{1}{2}} = 1{,}059.$$

Для знаходження граничної похибки при заданих імовірності p=0.95 і числі ступенів вільності v=14-3=11 за таблицями Ст'юдента обчислимо ймовірнісний коефіцієнт (p=0.95; v=11): t_p=2.201. Гранична похибка $\Delta_{x_0,1,2} = t_p s_{\Delta_{x_0,1,2}} = 2{,}331$. Таким чином, в господарствах, економічна оцінка землі яких становить 60 балів, а витрати виробництва на 1 га складають 300 гри., середнє значення урожайності зернових з ймовірністю 0,95 очікується не нижче від 41,25 - 2,33 = 30,92 ц/га і не вище від 41,25 + 2,33 = 43,58 ц/га. Такий висновок має місце лише для сукупності господарств, з яких взято вибірку, тобто сукупності значень, наведених у табл. 6.1 причому при певних погодних умовах, рівні агротехніки, території тощо.

6.9.7. Похибка індивідуальної оцінки множинної регресії

Вибіркова похибка множинної регресії визначає вибіркові похибки середніх значень результуючих змінних, які пов'язані з конкретними рівнями факторних ознак. Проте для всіх одиниць сукупності, які відповідають однаковим значенням факторних ознак, результуюча змінна є неоднаковою, оскільки рівняння множинної регресії не визначає точного значення результуючої ознаки при довільній одиниці сукупності за заданими рівнями факторних ознак. А тому, виникає необхідність визначення довірчих інтервалів не для середніх значень результуючої змінної, а для фіксованих індивідуальних оцінок. Розглянемо теоретичне значення (оцінку) результуючої змінної i-го господарства з тієї генеральної сукупності, з якої взято вибірку. Відомо, що i-му господарству відповідають факторні змінні $x_{1i}, x_{2i}, ..., x_{ki}$. Середню оцінку визначаємо на основі множинного рівняння регресії:

$$x_0 = b_0 + b_1 x_{1i} + b_2 x_{2i} + ... + b_k x_{ki}.$$

Така оцінка містить не тільки похибки вибірки (параметри $b_0, b_1, b_2, ..., b_k$), але й похибку, викликану непоясненою множинним рівнянням регресії варіацією, тобто похибку оцінки ε за рівнянням. У даному випадку необхідно об'єднати стандартну похибку самої регресії зі стандартною похибкою оцінки ε за рівнянням регресії. Згідно з теоремою про додавання дисперсій одержуємо

$$\begin{aligned} s^2_{x_0,1,2,...,k} &= s^2_{0,1,2,...,k} + s^2_{x_0,1,2,..,k} + s^2_{b_{01}}(x_{1i} - \bar{x}_1)^2 + \\ &+ s^2_{b_{02}}(x_{2i} - \bar{x}_2)^2 + ... + s^2_{b_{0k}}(x_{ki} - \bar{x}_k)^2. \end{aligned} \tag{6.156}$$

Звідси після відповідних перетворень маємо

$$\begin{aligned} s^2_{\bar{x}_0,1,2,...,k} = s^2_{0,1,2,..,k}\Bigg[&\frac{1}{n} + \frac{(x_{1i} - \bar{x}_1)^2}{s_1^2 n} - \frac{1}{1 - R^2_{1,2,3,...,k}} + \\ &+ \frac{(x_{2i} - \bar{x}_2)^2}{s_2^2 n} - \frac{1}{1 - R^2_{2,1,3,...,k}} + ... + \frac{(x_{ki} - \bar{x}_k)^2}{s_k^2 n} - \frac{1}{1 - R^2_{k,1,2,3,...,(k-1)}} \Bigg], \end{aligned} \tag{6.157}$$

$$\begin{aligned} s_{\bar{x}_0,1,2,...,k} = s_{0,1,2,..,k}\Bigg[&\frac{1}{n} + \frac{(x_{1i} - \bar{x}_1)^2}{s_1^2 n} - \frac{1}{1 - R^2_{1,2,3,...,k}} + \\ &+ \frac{(x_{2i} - \bar{x}_2)^2}{s_2^2 n} - \frac{1}{1 - R^2_{2,1,3,...,k}} + ... + \frac{(x_{ki} - \bar{x}_k)^2}{s_k^2 n} - \frac{1}{1 - R^2_{k,1,2,3,...,(k-1)}} \Bigg]^{\frac{1}{2}}, \end{aligned} \tag{6.158}$$

Розглянемо більш детально різницю між $s_{\Delta_{x_0,1,2,...,k}}$ та $s_{\Delta_{x_0,1,1,2,...k}}$, тобто між формулами (6.155) і (6.158). Збільшуючи обсяг вибірки, можна домогтися як завгодно близького збігання характеристик вибірки та генеральної сукупності, в тому числі умовних середніх арифметичних результуючої ознаки, які відповідають фіксованим факторним змінним. Звідси при $n \to \infty$ і фіксованих $x_{1i}, x_{2i}, ..., x_{ki}$;

$\Delta_{x_1}, \Delta_{x_2}, ..., \Delta_{x_k}$; $s_1^2, ...$, величина $s_{\Delta_{x_0,1,1,2,...,k}}$ прямує до нуля. Таким чином, не можна зменшувати варіацію індивідуальних даних і тим самим покращити індивідуальну оцінку, оскільки точність останньої завжди обмежена залишковою (непоясненою) варіацією. Отже, при $n \to \infty$ і фіксованих $x_{1i}, x_{2i}, ..., x_{ki}$; $\Delta_{x_1}, \Delta_{x_2}, ..., \Delta_{x_k}$; $s_1^2, ...$, величина $s^2_{\Delta_{x_0,1,1,2,...,k}}$ прямує до $s^2_{\Delta_{0,1,2,...,k}}$. Для достатньо великих обсягів вибірки та невеликих значень $(x_{1i} - \Delta_{x_1}), (x_{2i} - \Delta_{x_2}), ..., (x_{ki} - \Delta_{x_e})$ дисперсія індивідуальної оцінки наближається до дисперсії оцінки за рівнянням регресії. Останнє твердження випливає зі співвідношення

$$\frac{s^2_{\bar{x}_0,1,2,...,k}}{s^2_{0,1,2,...,k}} = \frac{1}{n} + \frac{(x_{1i} - \bar{x}_1)^2}{s_1^2 n} - \frac{1}{1 - R^2_{1,2,3,...,k}} + \\ + \frac{(x_{2i} - \bar{x}_2)^2}{s_2^2 n} - \frac{1}{1 - R^2_{2,1,3,...,k}} + ... + \frac{(x_{ki} - \bar{x}_k)^2}{s_k^2 n} - \frac{1}{1 - R^2_{k,1,2,3,...,(k-1)}}. \quad (6.159)$$

Якщо обсяг вибірки дуже великий ($n \to \infty$), то права частина рівняння (6.159) прямує до одиниці й точність індивідуального передбачення (прогнозу) може бути оцінена за допомогою стандартної похибки оцінки за рівнянням регресії. Стандартну похибку індивідуальної оцінки доцільно обчислювати за формулою (6.158), коли вибірка досить мала та якщо рівняння регресії використовують для прогнозу. Повернемося до прикладу 6.3. Якщо індивідуальна оцінка $\Delta_{x_{0i}}$ =41.25 ц/га (при x_1=60; x_2 =300), то похибка $s^2_{\Delta_{x_0,1,1,2}}$ =4.019, а інтервальна оцінка

$$41{,}25 - 4{,}02 \le \bar{x}_{0i} \le 41{,}25 + 4{,}02 \text{ або } 37{,}23 \le \bar{x}_{0i} \le 45{,}27.$$

6.9.8. Вибіркова похибка коефіцієнта множинної кореляції

Якщо обсяг вибірки досить великий, а коефіцієнт кореляції малий, то стандартну похибку коефіцієнта множинної кореляції визначаємо за формулою

$$s_R = \frac{1 - R^2}{\sqrt{n - k - 1}}, \quad (6.160)$$

де n число одиниць у вибірці; k – число факторних змінних у рівнянні множинної регресії. Гранична похибка й довірчі границі визначаються звичайним числом на основі нормального закону розподілу. Якщо обсяг вибірки невеликий, то для оцінки значущості коефіцієнта множинної кореляції використовують відношення

$$F \equiv \frac{s_A^2}{s_Z^2}, \quad (6.161)$$

де s_A^2 – пояснена дисперсія, розрахована на один ступінь вільності:

$$s_A^2 \equiv \frac{\sum(\tilde{x}_0 - x_0)^2}{k}, \tag{6.162}$$

s_Z^2 – непояснена дисперсія, розрахована на один ступінь вільності:

$$s_Z^2 \equiv \frac{\sum(x_0 - \tilde{x}_0)^2}{n-k-1}, \tag{6.163}$$

Після нескладних перетворень формула (6.161) набуває вигляду

$$F \equiv \frac{R^2(n-k-1)}{(1-R^2)k}, \tag{6.164}$$

де R – оцінюючий коефіцієнт множинної кореляції.

Критичні значення F_a зі ступенями вільності $\nu_1=k$ і $\nu_2=n-k-1$ знаходимо за математичними таблицями F-розподілу. Якщо $F>F_a$, то нульову гіпотезу відхиляють і статистичний зв'язок вважається значним і, навпаки, якщо $F< F_a$, то кореляційний зв'язок є незначним.

Приклад 6.8. Оцінити значущість коефіцієнта множинної кореляції для прикладу 6.4 при $R_{0,1,2} = 0{,}863$ ($R^2_{0,1,2} = 0{,}0746$); n=14; k=2.

Розв'язок. За формулою (6.161) знаходимо, що F=16.1. Нехай для перевірки нульової гіпотези задано ймовірність p=0.95. Тоді за таблицями F-розподілу (при α=0.05; ν_1=2; ν_2=11) знаходимо F_a=3.96. Оскільки $F>F_a$, то нульова гіпотеза відхиляється. Такий самий висновок можна зробити і з ймовірністю p= 0.99. При α=0.01; ν_1=2; ν_2=11 F_a =7.2 і $F>F_a$. Таким чином, досліджуваний зв'язок є статистично значущим.

6.9.9. Часткова регресія та кореляція

Раніше було показано, що парна регресія та кореляція описують зв'язок між двома результуючими змінними (результуючою та факторною), а множинне рівняння регресії – це вплив факторів на результуючу ознаку. На практиці часто виникає необхідність знаходження зв'язку між результуючою змінною та деякою фіксованою факторною змінною x_j за умови, що решта факторів постійні. В загальному випадку цю задачу розв'язати неможливо, оскільки факторних змінних може бути дуже багато. У випадку регресійної моделі з k факторами зв'язок між результуючою ознакою x_0 і факторною змінною x_j при деяких фіксованих (наприклад, середніх арифметичних) значеннях решти факторів з індексами $1,2,...,(j-1),(j+1),...,k$ описується рівнянням часткової регресії

$$\begin{aligned} &x_{0,j,1,2,...,(j-1)(j+1),..,k} = b_0 + b_j x_j, \\ &b_0^* = b_0 + b_1 x_1 + ... + b_{j-1} x_{j-1} + b_{j+1} x_{j+1} + ... + b_k x_k, \end{aligned} \tag{6.165}$$

де $\Delta_{x_1}, \Delta_{x_2}, ..., \Delta_{x_{j-1}}, \Delta_{x_{j+1}}, ..., \Delta_{x_k}$ – середні (фіксовані) значення факторних змінних.

Наприклад, для знаходження рівняння часткової регресії, яке характеризує зв'язок між врожайністю зернових x_0 й економічною оцінкою землі x_0 при деякому постійному (середньому) рівні затрат виробництва x_2, у множинне рівняння регресії $x_0 = 10{,}534 + 0{,}142x_1 + 0{,}074x_2$ замість фактора x_2 необхідно підставити його середнє значення $\Delta_{x_2} = 311$. У результаті одержимо $\tilde{x}_{0,1,2} = 33{,}548 + 0{,}142x_2$.

Значення $x_{0,1,2}$ дорівнює теоретичному значенню результуючої ознаки з врахуванням заданого значення першої факторної змінної та при фіксованій на середньому рівні факторній змінній x_2. Коефіцієнт часткової регресії збігається з відповідним коефіцієнтом множинної регресії і має аналогічний економічний зміст. Зокрема, в прикладі 6.3 частковий коефіцієнт регресії показує, що збільшення економічної оцінки землі на 1 бал при середніх (постійних) затратах виробництва дає приріст урожайності зернових в середньому на 0,14 ц/га.

Зауважимо, що при переході від множинного рівняння регресії до часткового фактори можна фіксувати не лише на середніх, але й на інших довільних рівнях. Крім того, на відміну від множинних рівнянь рівняння часткових регресій можна зображати у вигляді графіків на площині.

Поряд із частковою регресією розглядають також часткову кореляцію, яка характеризує зв'язок між результуючою ознакою x_0 і деякою факторною змінною x_j при фіксованих (середніх) рівнях інших факторів множинного рівняння регресії. Основою для виведення формули коефіцієнта часткової кореляції є коефіцієнт простої кореляції:

$$r_{0j}^2 \equiv \frac{s_0^2 - s_{0j}^2}{s_0^2} = 1 - \frac{s_{0j}^2}{s_0^2},$$
$$r_{0j} \equiv \sqrt{\frac{s_0^2 - s_{0j}^2}{s_0^2}} = \sqrt{1 - \frac{s_{0j}^2}{s_0^2}}, \qquad (6.166)$$

де s_0^2 – дисперсія результуючої змінної; s_{0j}^2 – дисперсія, яка не пояснюється j-м фактором.

Подібно до коефіцієнта простої кореляції запишемо вираз для квадрата коефіцієнта часткової кореляції для регресійної

$$r_{0,j,1,2,...,(j-1),(j+1),...,k}^2 \equiv \frac{s_{0,1,2,...,(j-1),(j+1),...,k}^2 - s_{0,1,2,...,(j-1),j,(j+1),...,k}^2}{s_{0,1,2,...,(j-1),(j+1),...,k}^2}. \qquad (6.167)$$

Коефіцієнт часткової кореляції дорівнює квадратному кореню з квадрата коефіцієнта часткової кореляції:

$$r_{0,j,1,2,...,(j-1),(j+1),...,k} = \sqrt{1 - \frac{s_{0,1,2,...,(j-1),j,(j+1),...,k}^2}{s_{0,1,2,...,(j-1),(j+1),...,k}^2}}. \qquad (6.168)$$

У випадку лінійної регресійної моделі з однією результуючою і трьома факторними змінними формула (6.168) набуває вигляду:

$$r_{0,1,2,3} = \sqrt{1 - \frac{s^2_{0,1,2,3}}{s^2_{0,2,3}}}, \qquad (6.169)$$

де $s^2_{0,2,3}$ – залишкова дисперсія результуючої змінної x_0 після виключення впливу факторів x_2 і x_3; $s^2_{0,1,2,3}$ – залишкова дисперсія змінної x_0 після виключення впливу факторів x_2 і x_3.

Значення $s^2_{0,1,2,3}$ відрізняється від $s^2_{0,2,3}$ виключно наявністю кореляції між факторами x_0 і x_1, а різниця $s^2_{0,2,3} - s^2_{0,1,2,3}$ відображає зменшення залишкової дисперсії внаслідок того, що крім факторів x_2 і x_3 враховується також фактор x_1.

Якщо $s^2_{0,2,3} = s^2_{0,1,2,3}$ $r_{0,1,2,3} = 0$ і включення в модель факторної ознаки x_1 додатково не пояснює дисперсію результуючої змінної після виключення впливу факторів x_2 і x_3. Якщо $s^2_{0,1,2,3} = 0$, то включення x_1 вичерпує всю залишкову дисперсію результуючої змінної та коефіцієнт часткової кореляції дорівнює одиниці.

Таким чином, коефіцієнт часткової кореляції так само, як і коефіцієнти парної та множинної кореляцій, змінюється в межах від нуля до одиниці, його значення характеризує додатковий вплив факторної ознаки x_1 у пояснення дисперсії результуючої змінної x_0 після попереднього врахування впливу факторів x_2 і x_3.

$$r_{0,1,2} = \frac{r_{0,1} - r_{0,2}r_{1,2}}{\sqrt{1 - r^2_{0,2}}\sqrt{1 - r^2_{1,2}}}. \qquad (6.170)$$

Коефіцієнти часткової кореляції можуть бути знайдені також на основі коефіцієнтів часткової кореляції найближчих нижчих порядків. Так, частковий коефіцієнт кореляції парного порядку (один фактор за крапкою в індексі):

$$r_{0,1,2,3} = \frac{r_{0,1,2} - r_{0,3,2}r_{1,3,2}}{\sqrt{1 - r^2_{0,3,2}}\sqrt{1 - r^2_{1,3,2}}}. \qquad (6.171)$$

Приклад 6.9. Для прикладу 6.4 знайти часткові коефіцієнт кореляції на основі відповідних коефіцієнтів парної кореляції: $r_{0,1}$=0.4436; $r_{0,2}$=0.8234; $r_{1,2}$=0.1341.

Розв'язок. Підставляючи ці значення в (6.170), одержуємо:

$$r_{0,1,2} = \frac{r_{0,1} - r_{0,2}r_{1,2}}{\sqrt{1 - r^2_{0,2}}\sqrt{1 - r^2_{1,2}}} = 0{,}5924 \text{ та } r_{0,2,1} = \frac{r_{0,2} - r_{0,1}r_{1,2}}{\sqrt{1 - r^2_{0,1}}\sqrt{1 - r^2_{1,2}}} = 0{,}8601.$$

З аналізу часткових коефіцієнтів кореляції випливає, що другий фактор (витрати виробництва) має більш сильний чистий зв'язок між результуючою змінною, ніж перший (економічна оцінка землі), оскільки $r_{0,2,1} > r_{0,1,2}$. Цей висновок збігається з тим, який було одержано на основі стандартизованих коефіцієнтів регресії. Часткові

коефіцієнти кореляції менші від коефіцієнта множинної кореляції (для цього прикладу $R_{0,1,2} = 0{,}0863$). Якщо між факторними змінними існує додатній кореляційний зв'язок, то часткові коефіцієнти кореляції, як правило, є меншими від відповідних коефіцієнтів простої кореляції, що й підтверджують результати дослідження цієї задачі. Нарешті, розглянемо коефіцієнти часткової детермінації. Відомо, що квадрат коефіцієнта простої кореляції дорівнює коефіцієнту детермінації. Останній характеризує долю дисперсії, що пояснена зв'язком між ознаками x_0 і x_2 Квадрат коефіцієнта часткової кореляції не є коефіцієнтом часткової детермінації. Загальну детермінацію на часткові детермінації, що пов'язані із впливом кожної незалежної змінної, можна розкласти за формулою множинної кореляції).

$$R^2_{0,1,2,\ldots,k} = \frac{b_1 \operatorname{cov}_{01} + b_2 \operatorname{cov}_{02} + \ldots + b_k \operatorname{cov}_{0k}}{s_0^2}. \qquad (6.172)$$

Зокрема якщо множинна регресія містить дві факторні ознаки,то

$$R^2_{0,1,2} = \frac{b_1 \operatorname{cov}_{01} + b_2 \operatorname{cov}_{02}}{s_0^2}, \qquad (6.173)$$

де $R^2_{0,1,2}$- коефіцієнт множинної детермінації (квадрат коефіцієнта множинної кореляції). Відповідно до числа факторів права частина рівняння (6.172) розкладається на дві:

$$R^2_{0,1,2} = b_1 \frac{\operatorname{cov}_{01}}{s_0^2} + b_2 \frac{\operatorname{cov}_{02}}{s_0^2}.$$

Кожний з одержаних доданків виражає кількісний вклад відповідної факторної змінної в пояснення варіації результуючої змінної.

Таким чином, для j-ої факторної ознаки коефіцієнт часткової детермінації:

$$d_{0,j,1,2,\ldots,k} = b_{0,j,1,2,\ldots,k} \frac{\operatorname{cov}_{0j}}{s_0^2}.$$

Сума часткових коефіцієнтів детермінації дорівнює коефіцієнту множинної детермінації (квадрату множинного коефіцієнта кореляції).

6.9.10. Приклад розрахунку потреб ринку на основі множинних рівнянь регресії

У нових умовах ринкового господарювання на перший план виступають завдання по вдосконаленню економічних важелів управління, розширенню самостійності, підвищенню відповідальності на всіх рівнях. Вдосконалення економічного механізму охоплює в першу чергу вдосконалення планування, фінансування й впровадження господарського розрахунку. Для цього планується здійснити перехід до економічної оцінки ресурсних потенціалів господарств, зокрема до планування економічних показників на основі наявних у сільськогосподарських підприємствах ресурсів (родючості землі, забезпеченням

основними виробничими фондами та трудовими ресурсами тощо). Це дасть можливість на нормативній теоретично обгрунтованій основі встановлювати рівнонапружені державні замовлення для господарств, які знаходяться в однакових умовах. Розглянемо деякі методики розрахунку замовлень на сільськогосподарську продукцію в розрізі адміністративних районів й окремих господарств Львівської області, в основі яких лежать відповідні агро-ресурсні потенціали. Причому для розрахунку останніх використовуються як множинні рівняння регресії залежності валової продукції від основних факторів (основних і оборотних фондів, економічної оцінки землі тощо), так і дані оцінки сільськогосподарських угідь, розрахованих за валовою продукцією, окупністю затрат та диференціального доходу. Згідно з цією методикою спочатку в розрізі адміністративних районів (можна також у розрізі природно-сільськогосподарських районів або природноекономічних зон) будують множинні рівняння регресії залежності валової продукції $\tilde{x}_0$ від впливу основних факторних змінних $x_1, x_2, ..., x_k$. Практика показує, що такі зв'язки досить добре моделюються лінійними багатовимірними рівняннями регресії вигляду:

$$\tilde{x}_0 = b_0 + b_1 x_1 + b_2 x_2 + ... + b_k x_k , \tag{6.174}$$

де $\tilde{x}_0$ - валова продукція (результуюча змінна), тис. грн.; b_0 – вільний член; $b_1, b_2, ..., b_k$ – коефіцієнти регресії; $x_1, x_2, ..., x_k$ – факторні змінні, від яких залежить результуюча змінна – валова продукція (у відповідних одиницях).

Далі на основі рівняння регресії (6.174) для кожного адміністративного району визначають відповідні індекси агроресурсних потенціалів:

$$P_i = \frac{b_0 + \sum_{i=1}^{k} b_i x_{ij}}{b_0 + \sum_{i=1}^{k} b_i \bar{x}_i} = \frac{x_{0j}}{\bar{x}_0} \tag{6.175}$$

де x_{ij} – фактичне значення i-го ресурсу на одиницю земельних угідь для j-го району; $\bar{x}_i$ – середнє значення i-го ресурсу по області; $\tilde{x}_{0j}$ – теоретичне (нормативне) значення результуючої змінної для j-го району; $\bar{x}_0$ – середній по області рівень виробництва валової продукції з сільськогосподарських угідь.

Для теоретичних значень результуючої змінної $\tilde{x}_0$ й індексів алгроресурсних потенціалів можна обчислювати різні наведені раніше статистичні оцінки. Найбільшу зацікавленість викликає стандартна похибка індивідуальної оцінки рівняння регресії $s_{x_0 j,1,2,...,k}$ й стандартна похибка індивідуальної оцінки коефіцієнтів агроресурсного потенціалу S_p. Враховуючи (6.158) маємо такі оцінки:

$$s_{\bar{x}_0,1,2,\ldots,k} = s_{0,1,2,\ldots,k}\left[\frac{1}{n} + \frac{(x_{1j}-\bar{x}_1)^2}{s_1^2 n} - \frac{1}{1-R^2_{1,2,3,\ldots,k}} + \right.$$

$$\left. + \frac{(x_{2j}-\bar{x}_2)^2}{s_2^2 n} - \frac{1}{1-R^2_{2,1,3,\ldots,k}} + \ldots + \frac{(x_{kj}-\bar{x}_k)^2}{s_k^2 n} - \frac{1}{1-R^2_{k,1,2,3,\ldots,(k-1)}}\right]^{\frac{1}{2}}, \qquad (6.176)$$

$$s_{p_j} = \frac{s_{\bar{x}_0,1,2,\ldots,k}}{\bar{x}_0}.$$

Нормативний вихід валової продукції є об'єктивним показником, оскільки відображає можливість виробництва валової продукції з 1 га сільськогосподарських угідь при наявних у районі виробничих ресурсах та інших рівних умов виробництва.

Таким чином, диференціація районів за показником нормативного виходу валової продукції відображає різну ступінь забезпеченості їх виробничими ресурсами. Отже, індекси агроресурсних потенціалів у акумульованому вигляді відображають об'єктивну відмінність районів в умовах господарювання, а тому повинні безпосередньо враховуватися при встановленні рівнонапружених планів державних замовлень адміністративним районам області. Для адміністративних районів Львівської області на основі даних річних звітів за рік кореляційна модель залежності валової продукції від основних факторів має вигляд:

Таблиця 6.13

Розрахунок індексів агроресурсних потенціалів для районів Львівської області

Район	Нормативні рівні валової продукції з 1 га сільськогосподарських угідь $\bar{x}_{0j}$, крб.	Середній по області рівень виробництва валової продукції з 1 га сільськогосподарських угідь x_0, крб.	Значення індексу агроресурсних потенціалів P_j
Бродівський	0,5601	0,4488	1,2480
Бузький	0,5994	0,4488	1,3356
Городицький	0,5306	0,4488	1,1823
Дрогобицький	0,3904	0,4488	0,8699
Жидачівський	0,4628	0,4488	1,0312
Золочівський	0,6889	0,4488	1,5350
Кам'янсько-Бузький	0,5847	0,4488	1,3028
Миколаївський	0,4118	0,4488	0,9176
Мостиський	0,3732	0,4488	0,8316
Нестеровський	0,4699	0,4488	1,0470
Перемишлянський	0,4198	0,4488	0,9335
Радехівський	0,5714	0,4488	1,2732
Самбірський	0,4367	0,4488	0,9730
Сколівський	0,2305	0,4488	0,5136
Сокальський	0,5931	0,4488	1,3215
Старо-Самбірський	0,3788	0,4488	0,8440
Стрийський	0,4981	0,4488	1,1098
Турківський	0,1288	0,4488	0,2870
Яворівський	0,2883	0,4488	0,6424

$$\tilde{x}_0 = -0.3065 + 0.0001x_1 + 0.0003x_2 - 0.0033x_3 + 0.0129x_4,$$

де x_0 – вихід валової продукції на 1 га сільськогосподарських угідь, грн.; x_1, x_2 – основні та оборотні фонди на 1 га сільськогосподарських угідь, грн.; x_3 – трудові ресурси (умовних працездатних на 100 га ріллі); x_4 – економічна оцінка землі, розрахована за окупністю затрат, бал.

З рівняння регресії випливає, що середній вихід валової продукції на 1 га сільськогосподарських угідь підвищують (незначно) такі фактори, як забезпеченість виробництва основними й оборотними фондами, а також економічна оцінка землі. У той самий час вихід валової продукції зменшується в середньому на 0,33 коп. із розрахунку на 1 га угідь при збільшенні кількості працездатних на 100 га ріллі на одну людину. Значення індексів агроресурсних потенціалів у розрізі районів Львівської області наведені в табл. 6.13. Середній по області рівень виробництва валової продукції з 1 га сільськогосподарських угідь (x_0=0.4488) було одержано при таких середніх значеннях факторних змінних: x_1=1229.6 грн.; x_2=577.3 грн.; x_3=31.31 = 31 чол.; x_4=43.6 балів. З аналізу табл. 6.3 випливає, що найвищий агроресурсний потенціал має Золочівський (p=1.5350), Бузький (p=1.3356) і Кам'янсько-Бузький (p=1.3028) райони, найнижчий – Турківський (p=0.2870), Яворівський (p=0.5136) і Сколівський (p=0.6424).

Для обгрунтування рівнонапруженого плану замовлень на сільськогосподарську продукцію індекси агроресурсних потенціалів використовують так. Перш за все доведений для області план державних замовлень окремих видів продукції в натуральному виразі необхідно перевести в порівняльну (сумарну) продукцію за допомогою нормативних собівартостей C_j, обчислених відповідно до їх досягнутого рівня (середнього) за останні 3-5 років і встановленої тенденції зміни за часом:

$$N_b = \sum_{j=1}^{k} Q_j C_j , \qquad (6.177)$$

де N_b – обсяг замовлень для області у вартісному вираженні; k - загальна кількість видів продукції, яка включена у замовлення; Q – замовлення на j-й вид продукції в натуральному вираженні; C_j – нормативна собівартість продукції j-го виду.

Перевага цього способу порівняння продукції перед методами порівняння її за допомогою закупівельних цін або коефіцієнтів кормових переваг полягав в тому, що в одержаних на основі собівартості обсягах продукції відображені загальні затрати, які повинні закласти райони в суспільне виробництво для виконання зобов'язань перед замовником (табл. 6.14).

Таблиця 6.14

Середні нормативні собівартості та замовлення на основні види продукції для Львівської області

Продукція	Нормативна собівартість 1 ц C_j, крб.	Нормативні замовлення у натуральному виді Q_j
Зерно	8,78	130000 т.
Цукрові буряки	4,66	1440 тис. т.
Льон - насіння	82,3	3400 т.

Картопля	13,95	150000 т.
Овочі відкритого грунту	80,66	124000 т.
М'ясо (жива вага), в тому числі ВРХ	238,53	158000 т.
Птиця (яйця)	225,85	336 тис. шт.
Молоко	27,29	580000 т.
Вовна	1651,80	700 ц.
Льон тріста	48,34	17310т.
Плоди зерняткові, кісткові, ягоди	64,38	9000 тис. т.
Насіння багаторічних грав	203,90	830 т.

Перемножуючи планові обсяги замовлень на сільськогосподарську продукцію в натуральному вигляді на нормативи собівартості, за формулою (6.177) одержуємо порівняльний обсяг товарної продукції в цілому по області. Зокрема, для перших десяти видів продукції табл. 6.14 – 615655 тис. грн. Далі назначаються середньорайонна норма (з розрахунку на 1 га) замовлення N_i, як результат ділення обчисленого по області обсягу замовлення у вартісному виразі на площу сільськогосподарських угідь S. Зокрема, на 1 га сільськогосподарських угідь N_i=6639,88 тис. крб.м. Для кожного адміністративного району нормативний обсяг замовлень:

$$Q_{jp} = S_j N_i P_j, \tag{6.178}$$

де S_j - площа сільськогосподарських угідь j-го району, м^2. Середньорайонна норма замовлень на 1 га сільськогосподарських угідь N_r =6639.88 тис. крб. Із формули (6.178) випливає, що загальний обсяг замовлень по районах диференціювалася з відповідним індексом їх агроресурсного потенціалу. Чим більше значення P_j тим більше продукції при інших однакових параметрах j-й район повинен продати державі, і навпаки. Завдяки цьому всі райони області ставляться в однакові умови за ступенем напруженості доведених обсягів замовлень (табл. 6.15).

Таблиця 6.15

Розрахунок нормативних обсягів замовлень на продукцію для Львівської області

Район	Площа сільськогосподарських угідь S_j, га	Індекс агроресурсного потенціалу P_j	Нормативний обсяг замовлень Q_{jp}, тис. трн.
Бродівський	55830	1,25	463380620
Бузький	48173	1,34	428616320
Городицький	42664	1,18	334274930
Дрогобицький	40586	0,87	234452950
Жидачівський	52918	1,03	361910230
Золочівський	63882	1,54	653219960
Кам'янсько-	54657	1,30	471790690
Миколаївський	45904	0,83	252981550
Мостиський	70149	1,05	489069980
Нестерове ький	33924	0,92	207231170
Перемишлянський	52326	0,93	323117670
Радехівський	55613	1,27	468964820

Район	Площа сільськогосподарських угідь S_j, га	Індекс агроресурсного потенціалу P_j	Нормативний обсяг замовлень Q_{jp}, тис. трн.
Самбірський	63029	0,97	405949840
Сколівський	5234	0,51	17724096
Сокальський	88900	1,32	779176630
Старо-	38737	0,84	216055580
Стрийський	39120	*1,11*	288324830
Турківський	19728	0,29	37984549
Львівський	55936	0,64	237701320

Ступінь напруженості планів державних замовлень визначається відношенням фактично доведеного плану у вартісному виразі до нормативного плану, визначеного за індексом агроресурсного потенціалу. Важливо підкреслити, що для більшості районів плани були доведені достатньо обґрунтованими. В той самий час відсутність інформації про агроресурсний потенціал призводить до того, що окремим районам доводяться завищені плани, а деяким – занижені. В результаті виявилося, що ті райони, яким було доведено завищені плани державних замовлень, як правило, їх виконали і, навпаки, райони із ненапруженими планами перевиконали їх.

На основі обчисленого замовлення для розглядуваного району аналогічно можна визначити цей показник для кожного окремого господарства цього району. Для цього на основі даних господарств досліджуваного району будують рівняння регресії вигляду (6.175) залежності валової продукції x_0 від впливу основних x_1, і оборотних x_2 фондів, трудових ресурсів x_3 і економічної оцінки землі x_4. Далі, за формулою (6.126) розраховують відповідні індекси агроресурсних потенціалів для j-го господарства цього району. І, нарешті, на основі одержаних індексів агроресурсних потенціалів за формулою (6.178) розраховувалися плани замовлень для кожного господарства окремо.

6.9.11. Приклад розрахунку потреб ринку на основі економічної оцінки

Розглянемо інший підхід знаходження індексів агроресурсних потенціалів адміністративних районів (господарств), який базується на основі даних про економічну оцінку землі. Для одержання найбільш достовірної інформації про виробничу діяльність сільськогосподарських угідь проведемо економічну оцінку земель Львівської області. Для оціночних показників було прийнято вихідні дані, розраховані за багаторічними даними господарств для кожного з одинадцяти природно-сільськогосподарських районів: врожайність і валовий збір основних сільськогосподарських культур, багаторічних насаджень і природних кормових угідь; вартість валової продукції землеробства; виробничі затрати. Для забезпечення порівняльності в межах країни кількість валової продукції з одиниці площі подано в єдиних кадастрових цінах. Економічну оцінку земель проведено в двох аспектах: загальна оцінка ріллі, багаторічних насаджень, сіножатостей, пасовищ і всього сільськогосподарських угідь; часткова оцінка ріллі за ефективністю вирощування основних сільськогосподарських культур: всього зернових без кукурудзи, озимої

пшениці, ярого ячменю, вівса, кукурудзи на зерно, цукрових буряків, картоплі, всього кормових. В основу загальної та часткової оцінок земель закладено бонітет ґрунту, який є спеціалізованою класифікацією ґрунтів за їх продуктивністю, побудованою на підставі провідних природних властивостей ґрунту, які корелюють із багаторічною урожайністю сільськогосподарських культур.

При розробці бонітету взято такі головні ознаки ґрунтів: вміст гумусу в орному шарі (у процентах), глибина гумусового горизонту (у сантиметрах), агрофізичні властивості ґрунту (механічний склад, агрегативність і вологоємкість). Крім того, при визначенні шкали бонітету враховані поправочні коефіцієнти на такі властивості ґрунтів: кислотність, оглеєння, еродованість (змитість або дефльованість) та щебенюватість, кам'янистість, карбонатність, щільність.

На підставі математичної обробки даних показників для основних сільськогосподарських культур у розрізі природно-сільськогосподарських районів складено шкалу бонітету ґрунтів у балах. За 100 балів по кожній сільськогосподарській культурі окремо взято показники найбільш родючого та найбільш поширеного ґрунту в оціночному районі.

Загальна і часткова оцінки земель представлені трьома оціночними показниками:

- продуктивністю земель по окремих угіддях, розрахованою за виходом валової продукції рослинництва з 1 га в кадастрових цінах при загальній оцінці земель або за урожайністю сільськогосподдарських культур в частковій оцінці (крб. за 1 га чи в балах);
- окупністю затрат (вартістю валової продукції в кадастрових цінах) або кількістю її в натуральному виразі (з розрахунку на 100 грн. витрат);
- диференціальним доходом (частиною чистого доходу), який виражав вартість додаткового продукту на відносно кращих і середніх землях.

На основі описаних груп оціночних показників розраховували індекси ресурсних потенціалів адміністративних районів (господарств) області:

$$P_j^{VP} = \frac{x_j^{VP}}{x^{VP}}, \quad P_j^{OZ} = \frac{x_j^{OZ}}{x^{OZ}}, \quad P_j^{DD} = \frac{x_j^{DD}}{x^{DD}}, \tag{6.179}$$

де P_j^{VP}, P_j^{OZ}, P_j^{DD} – індекс ресурсних потенціалів для j-го району (господарства) області, розрахований на основі економічних оцінок землі відповідно за валовою продукцією, окупністю затрат і диференціальним доходом, x_j^{VP}, x_j^{OZ}, x_j^{DD} – економічні оцінки землі, які обчислені відповідно за валовою продукцією, окупністю затрат і диференціальним доходом (бал); x^{VP}, x^{OZ}, x^{DD} – середньообласні значення відповідних економічних оцінок землі.

У табл. 6.16 наведені значення ресурсних потенціалів, що обчислені за формулою (6.179) та розраховані при таких середньообласних значеннях економічних оцінок землі: $x^{VP} = 31$, $x^{OZ} = 64$, $x^{DD} = 21$.

Таблиця 6.16

Розрахункові значення індексів ресурсних потенціалів, визначених на основі економічної оцінки землі для Львівської області

Район	Індекс ресурсних потенціалів, обчислений за валовою продукцією	Індекс ресурсних потенціалів, обчислений за окупністю затрат	Індекс ресурсних потенціалів, обчислений за диференціальним доходом
Броді вський	1,3226	1,0469	1,4762
Бузький	1,2258	1,0312	1,3333
Городи цький	1,1935	1,0625	1,2857
Дрогобицький	0,7742	0,9375	0,6190
Жидачівський	1,0968	0,9531	**1,0000**
Золочівський	1,3871	1,0937	1,5714
Кам'янсько-	1,1935	1,0312	1,2381
Миколаївський	0,8387	1,0937	0,8571
Мостиський	1,1290	0,9687	1,1429
Нестеровський	1,0323	0,9062	0,9048
Перемишлянський	**1,0000**	0,8437	0,6667
Нустомитівський	1,2581	1,1250	1,4286
Радехівський	1,1290	1,0625	1,1905
Симбірський	1,0323	1,0781	1,0476
Сколі вський	0,3226	0,9687	0,2381
Сокальський	1,2903	1,0469	1,3809
Старо-Самбірський	0,7742	0,9062	0,7143
Стрийський	0,9355	1,0625	0,9048
Турківський	0,5161	0,8906	0,3810
Яворівський	0,7742	0,8750	0,5714

За формулою (6.178) знаходимо відповідні значення державних замовлень на сільськогосподарську продукцію для кожного району області (господарства в районі). Зіставлення замовлень, розрахованих за описаною методикою, дало можливість найбільш об'єктивно, з точки зору наявних ресурсів, визначити обсяги замовлень для районів або господарств області. За допомогою наведених методик розрахунку індексів агроресурсних потенціалів знайдено нормативні (теоретичні) значення замовлень (стосовно відповідних середньообласних показників), які і будуть, як правило, більшими від замовлень, заданих планом. Якщо необхідно розподілити замовлення по районах (господарствах) області згідно з наявними виробничими та земельними ресурсами, то для знаходження індексів ресурсних потенціалів можна використати такі формули [76-80, 88, 92-93, 138, 142, 163]:

$$P_j^{VP} = \frac{x_j^{VP} S_j}{\sum_{i=1}^{n} x_i^{VP} S_i}, \quad P_j^{OZ} = \frac{x_j^{OZ} S_j}{\sum_{i=1}^{n} x_i^{OZ} S_i}, \quad P_j^{DD} = \frac{x_j^{DD} S_j}{\sum_{i=1}^{n} x_i^{DD} S_i}, \qquad (6.180)$$

де S_i, S_j – площа сільськогосподарських угідь відповідно для i-го та j-го району (господарства) області; n – кількість районів області (господарств у досліджуваному районі).

При використанні множинних рівнянь регресії, для знаходження індексів агроресурсних потенціалів формули (6.178) для обчислення обсягів замовлень на сільськогосподарську продукцію можна також нормувати, домножуючи результат на нормуючий множник. При переході будь-якого господарства на ринкову економіку, орієнтовану на світові (міжнародні) ціни на відповідну продукцію, замовлення необхідно розраховувати за описаними раніше методиками не у вартісному, а в натуральному виді, що дасть можливість власникам об'єктивніше враховувати наявні у них виробничі, трудові та інші ресурси при плануванні як замовлень, так і замовлень на окремі види такої продукції.

6.9.12. Приклад дослідження на основі багатовимірного регресійного аналізу

Одержання високих та стабільних економічних показників завжди було і залишається основною метою виробництва. Однак на різних етапах розвитку виробництва в поняття "високий" вкладався різний вміст. Комплекс факторів, від яких залежить рівень виробництва, різноманітний. Він містить, наприклад, для сільського господарства, особливості наявної системи землеробства, способи обробки ґрунту, ступені використання добрив, способи боротьби із різними хворобами та шкідниками, можливості сорту і відповідності вирощуваних сортів агрокліматичним ресурсам території, енергоозброєність сільськогосподарського виробництва, меліорація мікроклімату (ґрунту, повітря) тощо. Проведені в США дослідження показали, що в середньому підвищення врожайності зернових культур в цій країні в значній мірі (41%) зумовлене використанням добрив; вклад сорту складає 27%; за рахунок змін у розміщенні сільськогосподарських культур урожайність зросла на 2% і близько 30% припадає на агротехніку та інші фактори. Отже, врожайність довільної сільськогосподарської культури формується під впливом сукупності природнокліматичних, економічних і технологічних факторів. Ступінь їх впливу на урожайність є неоднаковим, а тому дуже важливо своєчасно виявити цей вплив та поєднати фактори так, щоб забезпечити оптимальне розміщення культур і в результаті одержати максимальний врожай. Процес виробництва в землеробстві здійснюється так, що на рівень врожайності впливають багато факторів як у досліджуваному році, так і в попередні роки. Тому врожайність необхідно розглядати як результат дії сукупності різних факторів за часом і простором. Такий характер зв'язку між врожайністю сільськогосподарських культур і різними факторами, що впливають на неї, дає можливість використовувати для побудови моделі інформацію, зібрану за принципом «об'єкт–роки». Вихідні дані розглядаються як багаторічні дані масового експерименту. У дослідженні врожайності на широко і досить успішно використовуються багатофакторні регресійні моделі, які досить точно відображають існуючі закономірності. Побудову кореляційно-регресійних моделей залежності врожайності від впливу різних факторів на регіональному рівні, як правило, здійснюють за такою схемою:

1. Формулюється гіпотеза стосовно характеру впливу факторів на врожайність.

2. Сукупність показників (вхідних даних) досліджується на однорідність.
3. Згідно з певними критеріями відбираються найбільш значущі фактори.
4. Визначається форма зв'язку між результуючою та факторним змінними і будується відповідна регресійна модель (як правило, лінійна чи параболічна).
5. Аналізуються та інтерпретуються параметри одержаних моделей.

Всю множину факторів, що впливають на врожайність сільськогосподарських культур, можна розділити на два класи. До першого слід віднести фактори, які характеризують (зумовлюють) рівень культури землеробства. Сюди входять способи обробки ґрунту, боротьби с хворобами та шкідниками, озброєність сільського господарства виробничими та трудовими ресурсами, відповідність сортів агрокліматичним умовам, раціональне використання добрив тощо. Другий клас містить природні властивості ґрунту та метеорологічні умови. Зокрема, вплив на врожайність вмісту в орному шарі гумусу; вмісту в ґрунті фосфору, калію, азоту; кислотності ґрунту; температури повітря, кількість опадів тощо.

Розглянемо залежність врожайності сільськогосподарських культур від сукупності факторів, які характеризують природну родючість землі: вмісту в орному шарі гумусу, в ґрунті – поживних речовин (у міліграмах на 100 г ґрунту), зокрема фосфору (P_2O_5), калію (K_2O) і азоту (N), а також кислотності (РН) ґрунту. Найпростіша лінійна багатофакторна регресійна модель такої залежності має вигляд

$$x_0 = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5, \qquad (6.181)$$

де x_0 – врожайність сільськогосподарських культур, ц/га; x_1 – кислотність ґрунту, %; x_2, x_3, x_4 – вміст відповідно фосфору, калію, азоту (у міліграмах на 100 г ґрунту) %; x_5 – вміст в орному шарі гумусу, %. Такі множинні рівняння регресії необхідно будувати для сукупності господарств, розміщених на землях з однаковими чи досить близькими природно-кліматичними умовами. Залежно від особливостей геологічної будови ґрунтів, їх рельєфу, кліматичного та водного режимів, характеру ґрунтів і рослинності та господарської діяльності людини територія Львівської області розподілена на 11 природно-сільськогосподарських районів, які охоплюють чотири ґрунтово-кліматичні зони. Останнім відповідають такі природноекономічні зони: Поліська, Лісостепова, Передкарпатська та Гірська.

Приклад 6.10. Побудувати множинні рівняння регресії залежності врожайності основних сільськогосподарських культур від сукупності факторних ознак (вмісту в орному шарі гумусу, фосфору, калію, азоту і кислотності ґрунту) для господарств Перемишлянського району Львівської області. Цей район входить у Перемишлянський природно-сільськогосподарський район (шифр 04, площа 186,3 тис. га сільськогосподарських угідь) і є найбільшим у Лісостеповій природноекономічній зоні. У табл. 6.17 наведено структуру земель та оціночні показники по природно-сільськогосподарських районах.

Таблиця 6.17

Структура земель і оціночні показники природно-сільськогосподарських районів Львівської області

Шифр	Назва	Площа	У тому числі ріллі		Фонди	Добрив на 1 га ріллі		Всього затрат на 1га посіву, крб.	Валова продукція на 1 га посіву в кадастрових
			Всього, тис. га	Осушених		Мінеральних, ц	Органічних, т		
01	Сокальський	69,39	57,89	20,66	510,43	1,66	12,05	442,40	1438,95
02	Городоцький	126,13	90,28	38,45	768,97	2,05	11,60	386,44	1062,11
03	Борщовицький	86,22	59,99	24,75	1111,31	2,04	11,74	500,97	1372,24
04	Перемишлянський	186,31	135,11	25,10	590,58	2,20	10,96	383,47	853,39
05	Яворівський	93,89	62,25	42,41	667,75	1,97	9,23	371,80	868,19
06	Кам'яно-Бузький	146,86	83,64	54,97	928,87	2,32	11,90	398,14	1103,12
07	Радехівський	66,38	44,12	6,19	519,92	2,28	17,04	380,33	960,76
08	Золочівський	69,97	50,54	17,88	675,78	2,70	13,79	411,95	1267,63
09	Дрогобицький	96,46	74,22	38,47	783,51	1,75	9,65	300,64	789,89
10	Самбірсько-Жидачівський	82,91	51,95	30,20	874,35	1,82	11,29	365,45	977,01
11	Турківський	106,80	42,68	1,12	537,45	1,80	9,40	322,01	549,21

Розв'язок. Перемишлянський природно-сільськогосподарський район об'єднує землі господарств, що входять у наступні адміністративні райони: Бродівський, Золочівський, Жидачівський, Миколаївський, Нестерівський, Перемишлянський, Пустомитівський і Яворівський. Причому він містить усі господарства Перемишлянського району і лише по кілька господарств з інших районів. Перемишлянський природно-сільськогосподарський район охоплює найбільш підвищену (високий висотний пояс Західного Лісостепу), густо та глибоко розчленовану частину Подільської височини, так зване Західно-Подільське плоскогір'я. Абсолютні висоти досягають 350...410 м. Тому цей район більш вологий, одержує менше тепла. Як і в усіх природно-сільськогосподарських районах Лісостепової зони, тут переважають опідзолені ґрунти (84% ріллі). Ясно-сірі та сірі опідзолені ґрунти не дуже поширені (12%), Темно-сірі та чорноземи опідзолені займають 32%, тобто третину району. Ці ґрунти родючі, добре забезпечені гумусом і поживними речовинами. В той самий час опідзолені ґрунти тут більш кислі та холодні, 12% займають ґрунти, оглеєні поверхнево або по всьому профілю. У зв'язку зі складністю рельєфу та інтенсивним сільськогосподарським використанням майже 40% ріллі на схилах охоплено змивом. Продуктивність земель цього природно-сільськогосподарського району серед лісостепових районів найнижча.

На основі даних річних звітів колгоспів Перемишлянського району про результати господарської діяльності в землеробстві побудовані регресійні моделі залежності врожайності основних сільськогосподарських культур (зернових без

кукурудзи, цукрових буряків, картоплі та ін.) від впливу кислотності ґрунту, вмісту в орному шарі гумусу, фосфору, калію й азоту. Причому всебічна оцінка характеру взаємозв'язку між результуючою та факторними змінними показала, що математично цей зв'язок найбільш наближено можна виразити у вигляді лінійного множинного рівняння регресії. Після обробки даних річних звітів одержано такі рівняння регресії залежності урожайності $\tilde{x}_{0j}$ основних сільськогосподарських культур від впливу кислотності ґрунту x_1, вмісту в ґрунті фосфору x_2, калію x_3, азоту x_4 та вмісту в орному шарі гумусу x_5:

- зернових без кукурудзи

$\tilde{x}_{01} = 35.5337 + 3.2926x_1 - 1.0618\ x_2 - 0.0976\ x_3 - 1.7921\ x_4 - 1.9418\ x_5$;

- кукурудзи на зерно

$\tilde{x}_{02} = -43.3924 + 8.1117x_1 - 0.8310\ x_2 + 0.0468\ x_3 - 0.1430\ x_4 + 2.7186\ x_5$;

- цукрових буряків

$\tilde{x}_{03} = 109.6387 + 67.4334x_1 - 7.6618\ x_2 + 13.8880\ x_3 - 17.7088\ x_4 - 66.2776\ x_5$;

- картоплі

$\tilde{x}_{04} = 102.1082 + 26.0755\ x_1 - 5.4443\ x_2 - 2.7112\ x_3 - 4.1510\ x_4 - 27.4166\ x_5$;

- овочів

$\tilde{x}_{05} = -296.2336 + 56.1439\ x_1 - 12.0160\ x_2 + 6.8366\ x_3 + 4.6091\ x_4 + 26.1503\ x_5$.

Одержанні значення коефіцієнтів регресії зручно подати у вигляді табл. 6.18.

Таблиця 6.18

Залежність врожайності сільськогосподарських культур від впливу факторних змінних: кислотності фунту b_1, вмісту в ґрунті фосфору b_2, калію b_3, азоту b_4, гумусу b_5 господарств Перемишлянського району Львівської області

Назва культури	Вільний член	Коефіцієнти регресії				
	b_0	b_1	b_2	b_3	b_4	b_5
Зернові без кукурудзи	35,534	3,293	-1,062	-0,098	-1,792	-1,942
Кукурудза на зерно	-43,392	8,121	-0,831	0,047	-0,143	2,719
Цукрові буряки	109,639	67,433	-7,662	13,888	-17,709	-66,278
Картопля	102,108	26,075	-5,444	-2,711	-4,151	-27,417
Овочі	-296,234	56,144	-12,016	6,837	4,609	26,150

Висновок 1. З аналізу табл. 6.8 випливає, що кожний з наведених факторів кількісно впливав на відповідну результуючу змінну. Зокрема, кислотність грунту b_1 найбільше впливала на врожайність цукрових буряків і овочів. Наприклад, при збільшенні кислотності ґрунту на 1% очікується збільшення середньої врожайності цукрових буряків на 67,43 ц/га, картоплі – на 26,07, овочів – на 56,14 ц/га. вплив фосфору є негативним на врожайність усіх досліджуваних культур, причому найбільш негативним – на врожайність овочів і цукрових буряків.

Висновок 2. Вміст калію в ґрунті найсильніше впливає на врожайність цукрових буряків і овочів, а найслабше – на врожайність картоплі. Так, збільшення в ґрунті вмісту калію на 1% приводить до збільшення врожайності на 13,9 ц/га.

Висновок 3. Збільшення вмісту в ґрунті азоту негативно впливає на врожайність майже всіх досліджуваних культур, за винятком овочів. Зокрема, збільшення вмісту азоту в ґрунті на 1% сприяє зменшенню врожайності цукрових буряків на 17,7 ц/га, а картоплі на 4,2 ц/га.

Висновок 4. Вміст в орному шарі гумусу найбільше впливав на врожайність овочів. Так, збільшення в ґрунті гумусу на 1% сприятиме збільшенню урожайності овочів на 26,1 ц/га і в той самий час приведе до зменшення середньої врожайності цукрових буряків на 66,3 ц/га.

6.10. Контрольні питання

1. Способи задання функцій для опису математичних моделей.
2. Дайте визначення поняття *апроксимація*.
3. Яка апроксимація називається *поліноміальною*?
4. Яка апроксимація називається *інтегральною*?
5. В чому суть методу *середньоквадратичного наближення*?
6. Який метод наближення називається *інтерполяцією*?
7. Який метод наближення називається *екстраполяцією*?
8. Опишіть метод найменших квадратів.
9. В чому полягає задача рівномірного сплайн-наближення?
10. Чим характеризується рівномірне сплайн-наближення?
11. Чим відрізняються однорідні й неоднорідні сплайни?
12. Яку апроксимацію називають *найкращим чебишовським наближенням або мінімаксним*?
13. При таких розробках виникає задача знаходження найкращого рівномірного наближення?
14. Чим характеризуються алгоритми, які побудовані за схемою Ремеза?
15. Опишіть схему визначення параметрів наближення за методом Ремеза.
16. Що таке *мінімакс*?
17. Які точки називають *точками чебишевського альтернансу*?
18. Який алгоритм використовують для визначення точок альтернансу?
19. В чому полягає Алгоритм Валле-Пуссена?
20. Як визначається коефіцієнт множинної кореляції?

6.11. Задачі для самостійної роботи

Задача 6.9.1. Користуючись даними таблиці, знайдіть методом найменших квадратів множинне рівняння регресії, коефіцієнт детермінації і відповідні оцінки

врожайності зернових x_0 (ц/га) від оцінки землі x_1 (бал) і витрат виробництва x_2 (крб. на 1 га).

Таблиця 6.19

Вхідні дані для задачі 6.9.1

№ п/п	Господарство	x_0	x_1	x_2
1	«Незалежність»	47,2	91	308,0
2	ім. Т. Шевченка	45,8	69	316,6
3	«Україна»	33,8	66	315,9
4	ім. Лесі Українки	52,3	79	365,7
5	«Відродження»	46,4	69	343,1
6	ім. Івана Франка	37,6	74	278,8
7	«Буковина»	48,5	72	261,8
8	ім.10-річчя Незалежності України	35,1	72	350,8
9	ім. В. Чорновола	45,9	53	334,0
10	ім.В.Стуса	36,3	74	306,9
11	«Прогрес»	38,2	69	314,5

Задача 6.9.2. Користуючись даними таблиці, знайдіть параметри множинного рівняння регресії, коефіцієнт детермінації і їх оцінки залежності урожайності зернових x_0 (ц/га); оцінки землі x_1 (бал) і витрат виробництва x_2 (крб. на 1 га).

Таблиця 6.20

Вхідні дані для задачі 6.9.2

№ п/п	Господарство	x_0	x_1	x_2
1	ім. Т. Шевченка	41,0	71	260,0
2	«Дружба»	46,0	67	369,5
3	ім. Івана Франка	42,3	69	255,8
4	ім. Ю. Федьковича	47,1	68	298,7
5	«Незалежність»	42,2	68	263,2
6	ім. Української Армії	40,6	69	338,1
7	«Світанок»	42,2	69	280,3
8	«Перемога»	38,2	66	216,7
9	«Правда»	42,5	66	235,7
10	ім. В. Стуса	39,7	55	240,4
11	«У країна»	40,3	70	237,2
12	«Маяк»	42,7	63	176,5
13	«28 червня»	39,3	60	231,4
14	«Буковина»	42,3	60	316,9
15	ім. 5-річчя Конституції	39,8	72	269,1
16	ім. Л. Українки	37,3	64	257,4
17	ім. М. Грушевського	54,0	76	237,0
18	«Нове життя»	50,0	80	237,0
19	ім. 29 березня	37,6	72	238,6

Задача 6.9.3. Користуючись даними таблиці, обчисліть параметри множинного рівняння регресії, коефіцієнт детермінації, а також всілякі оцінки рівняння регресії, яке характеризує зв'язок між урожайністю зернових x_0 (ц/га) і факторами: оцінкою землі x_1 (бал) і витратами виробництва x_2 (крб. на 1 га). Зробіть економічний аналіз одержаних результатів.

Таблиця 6.21

Вхідні дані для задачі 6.9.3

№ п/п	Господарство	x_0	x_1	x_2
1	ім. Ю. Федьковича	36,9	51	480,0
2	«Заповіт Т. Шевченка»	26,6	53	269,7
3	ім. В. Чорновола	34,2	53	290,1
4	«Чумацький шлях»	30,3	42	420,0
5	ім. Лесі Українки	32,4	43	318,1
6	«Україна»	37,0	51	299,9
7	«Буковина»	33,8	49	302,1
8	ім. Т. Шевченка	32,6	50	324,8
9	ім. М. Вовчка	21,6	49	251,8
10	«Карпати»	31,6	41	308,1
11	ім. Стуса	24,2	41	262,7
12	ім. Г. Сковороди	24,1	40	269,7
13	ім. Щорса	42,2	63	534,1
14	ім. О. Кобилянської	35,2	44	264,3
15	«Дружба»	26,7	45	283,1
16	ім. 1-го Вересня	34,9	40	416,2
17	ім. Івана Франка	30,6	46	345,2
18	«Нове життя»	27,5	49	209,0
19	ім. Я. Галана	28,1	41	270,6

Задача 6.9.4. Запишіть на основі даних таблиці лінійне рівняння залежності врожайності зернових x_0 (ц/га) від оцінки землі x_1 (бал) і витрат виробництва x_2 (крб. на 1 га).

Таблиця 6.22

Вхідні дані для задачі 6.9.4

№ п/п	Господарство	x_0	x_1	x_2
1	«Перемога»	25,2	73	263,3
2	«Прогрес»	35,8	79	273,5
3	ім. М. Грушевського	34,1	70	305,6
4	ім. Лесі Українки	29,8	76	295,3
5	ім. Т. Шевченка	32,0	73	316,7
6	ім. Я. Галана	32,2	73	224,3
7	«Правда»	23,6	59	277,7
8	«Карпати»	34,6	77	277,0
9	ім. В. Стуса	33,9	74	194,8
10	ім. ІванаФранка	27,0	82	230,9
11	ім. М. Щорса	37,6	80	276,8
12	«Заповіт Т. Шевченка»	28,5	87	267,1

Задача 6.9.5. Одержати методом статистичного моделювання (моделювання рівномірно розподілених випадкових величин) x_0, x_1, x_2 в межах $14.5<x_0<65$; $33.8<x_1<52.3$; $261.8< x_2<365.7$ по 11 різних господарствах для вправи 6.9.1 і розв'язати поставлену задачу.

Задача 6.9.6. Одержати методом статистичного моделювання (моделювання рівномірно розподілених випадкових величин) x_0, x_1, x_2 в межах $21.6 < x_0 < 42.2$; $40<x_1<63$; $209.0 < x_2< 534.5$ по 15 і 19 різних господарствах для вправ 6.9.2 і 6.9.3 і розв'язати поставлені задачі.

Задача 6.9.7. Одержати методом статистичного моделювання (моделювання рівномірно розподілених випадкових величин) x_0, x_1, x_2 в межах $25.2 < x_0 < 37.6$; $70<x_1<90$;$150.5 <x_2< 350.1$ по 12 господарствах для вправи 6.9.4 і розв'язати задачу.

6.12. Завдання до лабораторної роботи

Мета: Опрацювання експериментальних дослідних даних МНК. Опрацювання експериментальних дослідних даних. Набуття навичок з визначення параметрів моделі за мінімаксним методом.

Завдання 6.10.1. Визначити за методом найменших квадратів параметри моделей

$$\text{а)}\ Y_1 = A + Be^{x/k};$$

$$\text{б)}\ Y_2 = A + B\,x^k + Cxe^{kx^2};$$

де k – номер студента за списком, A, B і C – невідомі параметри.

В якості результатів спостереження використати значення функції

$$f(x) = 3x^k + 2e^{x/k}$$

в точках $x_i = 2 + i*0.2$ для i = 1, 2,..., 10+k.

За кожною з отриманих моделей:

– визначити її значення в точках x=3,975 і x=9+k;

– для кожної з отриманих моделей визначити оцінку середньоквадратичного відхилення за формулою

$$\Delta = \sqrt{E/(n-1)},$$

де n – кількість точок спостереження, а E – сума квадратів відхилень моделі від результатів спостереження в точках спостереження

$$E = \sum_{i-1}^{n} \left(y_i - Y(x_i)\right)^2;$$

– побудувати графіки моделей і точками на них зобразити результати спостереження.

Завдання 6.10.2. Виконати апроксимацію табличних даних поліномом другого степеня за методом найменших квадратів

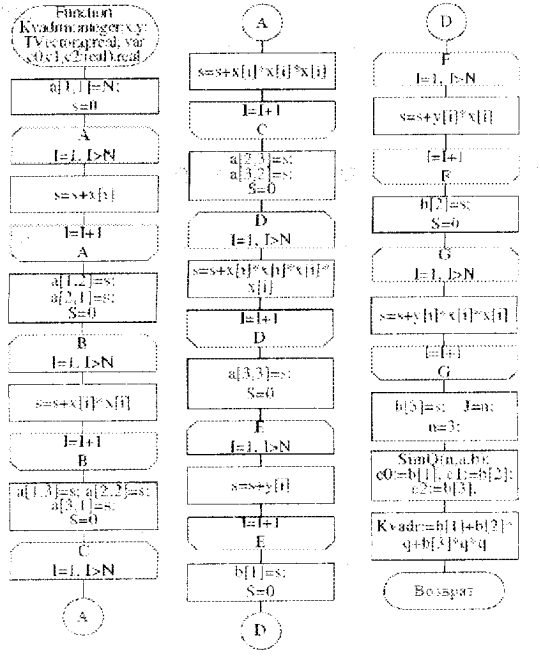


Рис. 6.10. Апроксимація табличних даних поліномом другого степеня за методом найменших квадратів

Виконати апроксимацію функції, що задана таблично, лінейною і квадратичною функціями.

Таблиця 6.23

Варіанти індивідуальних завдань

№0	***x***	**0,1**	**0,2**	**0,3**	**0,4**	**0,5**	**0,6**	**0,7**	**0,8**
	y	**3,030**	**3,142**	**3,358**	**3,463**	**3,772**	**3,251**	**3,170**	**3,665**
№1	*x*	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8
	y	3,314	3,278	3,262	3,292	3,332	3,397	3,487	3,563
№2	*x*	2,0	2,1	2,2	2,3	2,4	2,5	26	2,7
	y	1,045	1,162	1,264	1,172	1,070	0,898	0,656	0,344
№3	*x*	0,3	0,6	0,9	1,2	1,5	1,8	2,1	2,4
	y	6,715	6,735	6,750	6,741	6,645	6,639	6,647	6,612
№4	*x*	1,2	1,4	1,6	1,8	2,0	2,2	2,4	2,6
	y	2,325	2,515	2,638	2,700	2,696	2,626	2,491	2,291
№5	*x*	2,1	2,3	2,5	2,7	2,9	3,1	3,3	3,5
	y	1.752	1,762	1,777	1,797	1,821	1,850	1,884	1,944
№6	*x*	1,5	1,6	1,7	1,8	1,9	2,0	2,1	2,2
	y	1,924	1,710	1,525	1,370	1,264	1,190	1,148	1,127
№7	*x*	0,5	0,6	0,7	0,8	0,9	1,0	1,1	1,2
	y	1,025	1,144	1,336	1,419	1,479	1,530	1,568	1,248
№8	*x*	2,2	2,3	2,4	2,5	2,6	2,7	2,8	2,9
	y	5,785	5,685	5,605	5,545	5,505	5,480	5,495	5,510
№9	*x*	0,5	0,7	0,9	1,1	1,3	1,5	1,7	1,9
	y	4,052	4,092	4,152	4,234	4,338	4,468	4,599	4,771

Завдання 6.10.3. Для значень функції

$$y = 23{,}5ke^{0{,}34kx} + \sqrt{x^{1{,}4k}}$$

на відрізку $[0{,}2;\ 5{,}4 + k]$ отримати мінімаксну модель вигляду

$$F(x) = \frac{A + Bx}{1 + Cx}$$

з найменшою абсолютною похибкою. Побудувати графік функції, моделі та її похибки.

Завдання 6.10.4. Для значень функції (завдання 6.10.3) на відрізку $[0.7,\ 8.5]$ отримати мінімаксну модель вигляду

$$F(x) = A + Bx + Cx^2 + Dx^3$$

з найменшою відносною похибкою. Побудувати графік функції, моделі та її похибки.

Розділ 7.
Чисельне інтегрування

- Постановка задачі чисельного інтегрування
- Особливості чисельного інтегрування функцій
- Формула прямокутників. Графічна інтерпретація
- Формула трапецій
  - Постановка задачі
  - Графічна інтерптетація
- Формула Сімпсона (парабол)
  - Постановка задачі
  - Графічна інтерптетація
- Алгоритм реалізації принципу подвійного перерахунку та автоматичний вибір кроку інтегрування
- Наближене обчислення багатократних інтегралів
- Метод комірок
- Послідовне інтегрування
- Кубатурна формула типу Сімпсона
- Методи чисельного інтегрування
  - Задача обчислення визначеного інтеграла
  - Три підходи квадратурної формули
  - Інтерполяційні квадратурні формули Ньютона-Котеса
  - Частинні випадки інтерполяційних формул Ньютона-Котеса
- Квадратурні формули Гауса найкращого степеня точності
  - Ідея побудови квадратурних формул Гауса
  - Метод Гауса знаходження вузлів квадратурних формул
  - Метод Гаус вибору коефіцієнтів
  - Застосування формули Гауса
- Квадратні формули чисельного інтегрування Чебишова
  - Постановка задачі
  - Метод Чебишова знаходження вузлів квадратурних формул
  - Побудова квадратурних формул Чебишова
- Збіжність квадратурних формул
- Чисельні методи обчислення кратних інтегралів
  - Методи повторного інтегрування для двократного інтегрування
  - Метод заміни підінтегральної функції інтерполяційним многочленом

- Контрольні питання
- Задачі для самостійної роботи
- Завдання до лабораторної роботи
- Варіанти завдання (рівняння для розв'язування)

Формули для чисельного інтегрування називають квадратурними формулами (КФ) [65, 71, 76-80, 88, 92-93, 106-110, 113-114, 116-133, 137, 140-145, 147-151, 161-164, 168-169, 192-194, 196-197, 201-209, 211]. Вони досить часто використовуються на практиці. Це зумовлено декількома факторами: для багатьох часто вживаних функцій первісні не виражаються через елементарні функції; первісна функція має надто складний аналітичний вираз; функція є задана таблично. У загальному випадку формули для наближеного обчислення інтегралів $\int_c^d f(x)dx$ отримують заміною підінтегрального виразу інтерполяційним многочленом Лагранжа, або частіше його окремим випадком для рівновіддалених вузлів. Розглянемо окремі випадки, що випливають із узагальнених результатів, які називають *методами прямокутників, трапецій, парабол* [65, 71, 76-80, 88, 92-93, 106-110, 113-114, 116-133, 137, 140-145, 147-151, 161-164, 168-169, 192-194, 196-197, 201-209, 211].

7.1. Постановка задачі чисельного інтегрування

Необхідно знайти означений інтеграл

$$I = \int_a^b f(x)dx. \tag{7.1}$$

Відомо, що якщо функція $f(x)$ неперервна на відрізку $[a,b]$, то інтеграл (7.1) може бути знайдений за формулою Ньютона-Лейбніца:

$$\int_a^b f(x)dx = F(b) - F(a), \tag{7.2}$$

де $F(x)$ – первісна для функції $f(x)$.

Однак, на практиці формула (7.2) не завжди може бути використана: первісна функція $F(x)$ або занадто складна, або її знаходження викликає труднощі; крім того, функція $f(x)$ дуже часто задається таблицею. Тому великого значення набуває наближене та, в першу чергу, чисельне інтегрування.

Чисельне інтегрування – це знаходження означеного інтегралу на відрізку $[a,b]$, якщо підінтегральну функцію задано таблицею. Розіб'ємо відрізок інтегрування

$[a,b]$ на n рівних частин системою точок $x_i = x_0 + i \cdot h$ (i=0,1,2, ..., n), x_0=a, x_n=b, $h = \frac{b-a}{n}$ та обчислимо підінтегральну функцію у знайдених вузлах:

$$y_i = f(x_i),\ i = 0,\ 1,\ 2,\ ...,\ n. \tag{7.3}$$

Тоді

$$\int_a^b f(x)dx = \sum_{i=0}^{n} A_i \cdot f(x_i) + R, \tag{7.4}$$

де x_i – вузли інтегрування; A_i – коефіцієнти, залежні від вибору вузлів; R – залишковий член або похибка квадратурної формули.

Чисельне визначення певного інтеграла називається механічною *квадратурою*, а відповідні формули чисельного інтегрування – *квадратурними*.

7.2. Особливості чисельного інтегрування функцій

Розглянемо формулу Ньютона-Лейбніца

$$I = \int_a^b f(x)dx = F(a) - F(b). \tag{7.5}$$

Якщо $f(x)$ має первісну, то значення інтегралу знаходиться за формулою Ньютона–Лейбніца. Але визначення первісної в більшості випадків інженерних задач є неможливим, тому для визначення інтегралу використовують чисельні методи. Всі ці методи базуються на геометричній інтерпритації визначеного інтегралу, значення якого чисельно дорівнює площі фігури, що обмежена зверху – графіком функції $f(x)$, знизу – віссю $0X$, зліва та права – межами інтегрування a та b. Для знаходження площі відрізок AB розбивають на рівні частини, довжиною h, де h=(b– a)/n – крок інтегральної функції.

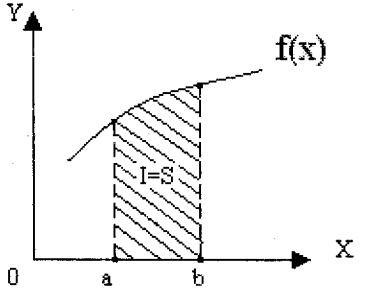


Рис. 7.1. Приклад знаходження загальної площі

1. Формула прямокутників. У випадку n=1 функцію $f(x)$ на відрізку $[c,d]$ замінюють інтерполяційним многочленом нульового порядку, побудованим за значенням функції в середній точці $x=(c+d)/2$, тобто функцією

$$F_0(x)=f\left(\frac{c+d}{2}\right) \text{ на } [c,d]. \tag{7.6}$$

Таким чином,

$$\int_c^d f(x)dx=(d-c)f\left(\frac{c+d}{2}\right)+R_1(f). \tag{7.7}$$

Геометрично це означає, що площа криволінійної трапеції замінюється площею прямокутника з основою $d-c$ і висотою $f\left(\frac{c+d}{2}\right)$. з метою покращення точності інтегрування відрізок $[c,d]$ розбивають на m рівних частин з кроком $h=\frac{d-c}{m}$ і застосовують формулу прямокутників до кожної частини окремо. Тоді

$$\int_c^d f(x)dx=\frac{(d-c)}{m}\left[f\left(c+\frac{h}{2}\right)+f\left(c+\frac{3h}{2}\right)+\ldots+f\left(c+\frac{2m-1}{2}h\right)\right]+R_1(f). \tag{7.8}$$

Ця формула називається *узагальненою формулою прямокутників.*

2. Формула трапецій. У випадку n=2 функцію $f(x)$ на відрізку $[c,d]$ замінюють інтерполяційним многочленом другого порядку, побудованим за значеннями функції в кінцевих точках проміжку, тобто крива замінюється хордою, що з'єднує кінцеві точки кривої. Інтеграл від інтерполяційного многочлена дасть площу трапеції, тому відповідна формула отримала назву *формули трапецій.*

Площа трапеції дорівнює

$$\frac{d-c}{2}[f(c)+f(d)]. \tag{7.9}$$

Таким чином,

$$\int_c^d f(x)dx=\frac{(d-c)}{2}[f(c)+f(d)]+R_2(f). \tag{7.10}$$

Узагальнена формула трапецій має вигляд:

$$\int_c^d f(x)dx=\frac{d-c}{2m}[y_0+2y_1+2y_2+\ldots+2y_{m-1}+y_m]+R_2(f), \tag{7.11}$$

де $y_0, y_1, y_2, \ldots, y_m$ – послідовні ординати у m точках розбиття.

3. Формула Сімпсона (парабол). Вузлами *інтерполяції* при n=3 обираються точки $c, \frac{c+d}{2}, d$. Через ці точки проводиться поліном другого степеня – парабола. Інтегрування рівняння параболи дає

$$\int_c^d F_2(x)dx = \frac{(d-c)}{6}\left[f(c)+4f\left(\frac{c+d}{2}\right)+f(d)\right]+R_3(f). \qquad (7.12)$$

Таким чином,

$$\int_c^d f(x)dx = \frac{(d-c)}{6}\left[f(c)+4f\left(\frac{c+d}{2}\right)+f(d)\right]+R_3(f). \qquad (7.13)$$

Для підвищення точності інтегрування проміжок доцільно розбити на $2m$ рівних частин. *Узагальнена формула парабол* буде мати вигляд:

$$\int_c^d f(x)dx = \frac{d-c}{6m}\left[y_0+4y_1+2y_2+4y_3+\ldots+4y_{2m-1}+y_{2m}\right]+R_3(f). \qquad (7.14)$$

7.3. Формула прямокутників. Графічна інтерпретація

Замінимо елементарні (криволінійні) трапеції в діапазоні $[a;b]$ прямокутниками, і обчислимо загальну площу фігури, як суму площ окремих прямокутників. Для випадку а) знайдену площу назвемо площею лівих, а для б) – правих прямокутників.

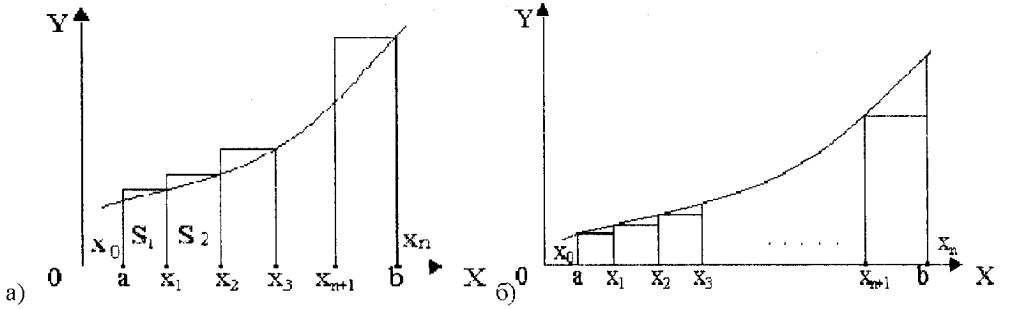


Рис. 7.2. Приклади знаходження загальної площі методом прямокутників.

Для а) маємо:

$$S^{\wedge} = f(x_1)(x_1-x_0)+f(x_2)(x_2-x_0)+\ldots..+$$
$$+f(x_n)(x_n-x_n-1)=\frac{b-a}{N}\sum_{k=1}^{N}f(x_k). \qquad (7.15)$$

Для б) маємо:

$$S^n = f(x_0)(x_1 - x_0) + f(x_2)(x_2 - x_0) + \ldots\ldots + \\ + f(x_{n-1})(x_n - x_{n-1}) = \frac{b-a}{N}\sum_{k=0}^{N-1} f(x_k). \qquad (7.16)$$

Істинне значення інтегралу обчислимо, як середнє арифметичне значення площ лівих і правих прямокутників: $I = (S^{\wedge} + S^n)/2$. Отримана формула називається формулою лівих і правих прямокутників, або формулою трапеції.

$$I = \frac{b-a}{2N}\left[\sum_{k=0}^{N} f(x_k) + \sum_{k=1}^{N} f(x_k)\right] = \\ = \frac{b-a}{2N}\left[f(a) + f(b) + 2\sum_{k=1}^{N-1} f(x_k)\right]. \qquad (7.17)$$

7.4. Формула трапецій

7.4.1. Постановка задачі

Обчислення означеного інтеграла (7.1) геометрично зводиться до обчислення площі криволінійної трапеції, яка обмежена графіком підінтегральної функції $f(x)$, відрізком $[a,b]$, віссю абсцис і прямими $x=a$ та $x=b$ (рис. 7.3). Замінимо дугу AB стягуючою її хордою. Одержимо прямолінійну трапецію $aABb$, площу якої візьмемо як наближене значення інтеграла.

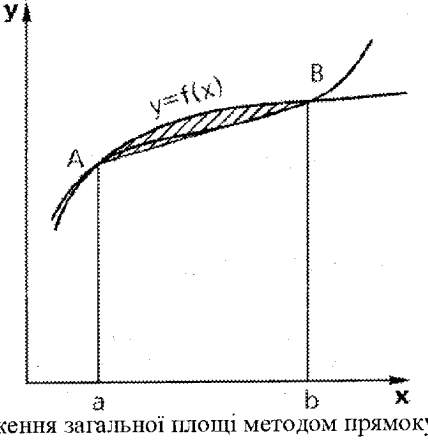


Рис. 7.3. Знаходження загальної площі методом прямокутників

$$\int_a^b f(x)dx \approx (b-a)\frac{f(a)-f(b)}{2}. \qquad (7.18)$$

Це і є формула трапецій. Для підвищення точності обчислень доцільно відрізок $[a,b]$ розділити на декілька частин точками $a = x_0 < x_1 < x_2 < \ldots < x_n = b$ та застосувати формулу (7.18) до кожного часткового відрізка $[x_i, x_{i+1}]$ (i=0, 1, …, n-1).

Тоді чисельне значення інтеграла на відрізку $[x_i, x_{i+1}]$ дорівнює:

$$\int_{x_i}^{x_{i+1}} f(x)dx \approx \frac{f(x_i)+f(x_{i+1})}{2}(x_{i+1}-x_i). \qquad (7.19)$$

На відрізку $[a,b]$

$$\int_a^b f(x)dx \approx \frac{1}{2}\sum_{i=0}^{n-1}[f(x_i)+f(x_{i+1})](x_{i+1}-x_i). \qquad (7.20)$$

Формула (7.20) називається *узагальненою формулою трапецій*.

Для простоти обчислень зручно ділити відрізок $[a,b]$ на рівні частини. У цьому випадку довжина кожного з n відрізків розбивки постійна, дорівнює $\frac{b-a}{n}$ і називається *кроком інтегрування*. Позначивши $h = \frac{b-a}{n}$, $y_0 = f(x_0)$, $y_1 = f(x_1)$, … , $y_n = f(x_n)$, де $x_0 = a$, $x_1 = x_0 + h$, $x_2 = x_0 + 2h$, …, $x_n = x_0 + nh = b$, одержимо узагальнену формулу трапецій у вигляді

$$\int_a^b f(x)dx \approx h\left(\frac{y_0+y_n}{2} + y_1 + y_2 + \ldots + y_{n-1}\right). \qquad (7.21)$$

Залишковий член або похибка квадратурної формули має вигляд:

$$R = -\frac{b-a}{12}\cdot\overline{\Delta^2 y}. \qquad (7.22)$$

де $\overline{\Delta^2 y}$ – середнє арифметичне різниць другого порядку.

7.4.2. Графічна інтерптетація

Розглянемо інший спосіб побудови квадратурних формул, що пов'язаний з апроксимацією підінтегральної функції інтерполяційним многочленом. Розглянемо найпростіший випадок. Метод трапецій використовує лінійну інтерполяцію, тобто графік функції $y=f(x)$ представляться у вигляді ламаної, що з'єднує точки (x_i,y_i). В цьому випадку площа всієї фігури (криволінійної трапеції) складається з площ елементарних прямокутних трапецій (рис. 7.4).

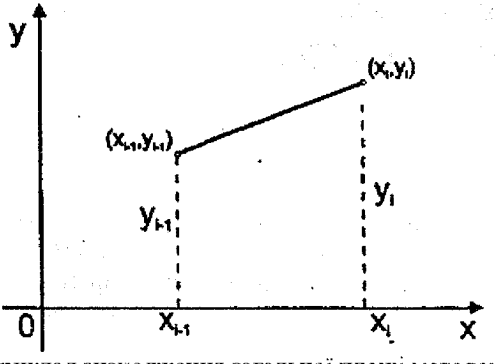


Рис. 7.4. Приклад знаходження загальної площі методом трапецій

Площа кожної такої трапеції дорівнює добутку напівсуми основи на висоту:

$$S_i = \frac{y_{i-1} + y_i}{2} h_i, \quad h=1,2,3\ldots,n. \tag{7.23}$$

Просумувавши ці рівності, отримаємо *формулу трапецій* для чисельного інтегрування:

$$I \equiv \int_a^b f(x)d(x) \approx \frac{1}{2} \sum_{i=1}^{n} h_i (y_{i-1} + y_i). \tag{7.24}$$

Для рівномірної сітки (h_i=h) ця формула має такий вигляд:

$$I \equiv \int_a^b f(x)d(x) \approx h\left(\frac{y_0 + y_n}{2} + \sum_{i=1}^{n} y_i\right) \equiv I_2. \tag{7.25}$$

Залишковий член має вигляд:

$$R_2[f] = I - I_2 = -\frac{(b-a)^3}{12} f''(\eta);\ \eta\ \epsilon\ (a,b). \tag{7.26}$$

Використовуючи вираз (7.26) для залишкового члена, оцінку похибки квадратурної формули (7.25) можна подати у вигляді:

$$\Delta_1 = \left| \int f(x)d(x) - \frac{(b-a)}{2}(f(a) - f(b)) \right| \le \frac{(b-a)^3}{12} M_2. \tag{7.27}$$

де $M_2 = \max\limits_{[a,b]} |f''(x)|$.

Оцінка обчислювальної похибки при розрахунках за формулою (7.25) для випадку, коли значення функції обчислені з однаковою точністю є, має вигляд:

$$\Delta_2 \le \frac{b-a}{2}(\varepsilon + \varepsilon) = (b-a)\varepsilon\ . \tag{7.28}$$

7.5. Формула Сімпсона (парабол)

7.5.1. Постановка задачі

Геометрична ілюстрація формули Сімпсона полягає в тому, що на кожному зі здвоєних часткових відрізків замінюємо криву $y=f(x)$ квадратичною параболою (рис. 7.5). Площа криволінійної трапеції приблизно дорівнює сумі так званих параболічних трапецій. Тому формулу Сімпсона називають також *формулою парабол*.

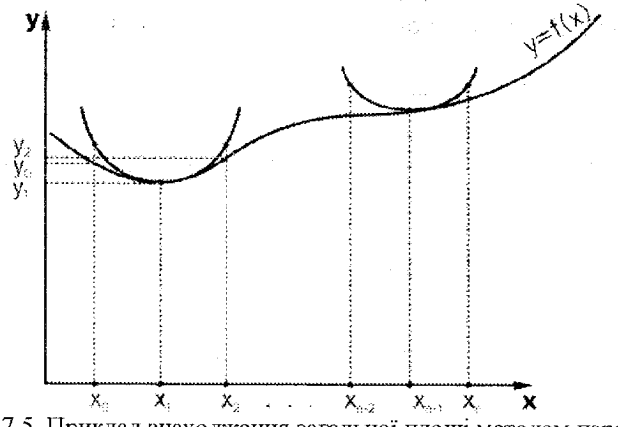


Рис. 7.5. Приклад знаходження загальної площі методом парабол

Розіб'ємо відрізок $[a,b]$ на парне число n рівних відрізків довжини $h=\dfrac{b-a}{n}$.

На кожних двох часткових відрізках замінимо цю функцію $f(x)$ інтерполяційним многочленом Ньютона другого степеня:

- на відрізку $[x_0, x_2]$: $P_2(x)=y_0+\dfrac{\Delta y_0}{h}(x-x_0)+\dfrac{\Delta^2 y_0}{h^2}(x-x_0)(x-x_1)$;
- на відрізку $[x_2, x_4]$: $P_2(x)=y_2+\dfrac{\Delta y_2}{h}(x-x_2)+\dfrac{\Delta^2 y_2}{h^2}(x-x_2)(x-x_3)$;
- на відрізку $[x_{n-2}, x_n]$: $P_2(x)=y_{n-2}+\dfrac{\Delta y_{n-2}}{h}(x-x_{n-2})+\dfrac{\Delta^2 y_{n-2}}{h^2}(x-x_{n-2})(x-x_{n-1})$.

Тоді

$$\int_a^b f(x)dx \approx \int_a^b P_2(x)dx = \int_{x_0}^{x_2} P_2(x)dx + \int_{x_2}^{x_4} P_2(x)dx + \ldots + \int_{x_{n-2}}^{x_n} P_2(x)dx. \qquad (7.29)$$

Обчислимо інтеграл на відрізку $[x_0,x_2]$:

$$\int_{x_0}^{x_2} P_2(x)dx = \int_{x_0}^{x_2}\left[y_0 + \frac{\Delta y_0}{h}(x-x_0) + \frac{\Delta^2 y_0}{h^2}(x-x_0)(x-x_1)\right]dx. \qquad (7.30)$$

Остаточно, формула Сімпсона має вигляд:

$$\int_a^b f(x)dx \approx \frac{h}{3}\left[y_0 + y_n + 4(y_1 + y_3 + y_5 + \ldots + y_{n-1}) + 2(y_2 + y_4 + \ldots + y_{n-2})\right]. \qquad (7.31)$$

Залишковий член або похибка квадратурної формули має вигляд:

$$R = -\frac{b-a}{180}\cdot\overline{\Delta^4 y}, \qquad (7.32)$$

де $\overline{\Delta^4 y}$ – середнє арифметичне різниць четвертого порядку.

7.5.2. Графічна інтерптетація

Більш точним методом визначення інтегралу є метод парабол (Сімпсона).

За цим методом відрізок AB ділять на $2n$ рівних частин, тобто кожен з проміжків ділять пополам. Розглянемо n-послідовність точок проміжку AB, $x_1, x_2, \ldots, x_n$, де $x_1 < x_2 < x_n$, $x_2 = x_1 + h$, $x_3 = x_1 + 2h$, $x_n = x_1 + (n-1)h$ (рис. 7.6). Вважаємо, що в цих точках задано значення деякої функції $f(x) = y$ й необхідно знайти інтеграл $\int_{x_1}^{x_n} f(x)\,dx$.

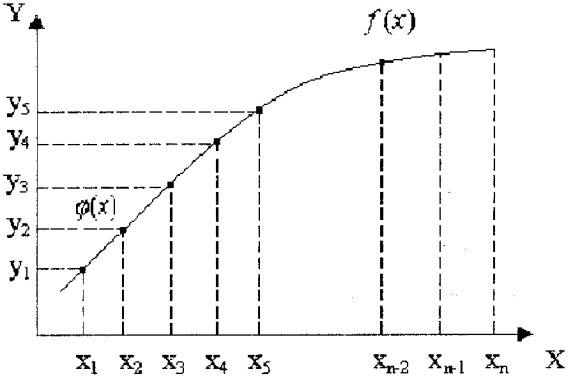


Рис. 7.6. Приклад знаходження загальної площі методом парабол.

Через кожні послідовні три точки проводимо параболу $y = a + bx + cx^2$ і обчислюємо інтеграл від функції, яка виражена у вигляді цієї параболи. Цей інтеграл і вважається наближеним значенням шуканого інтегралу.

Розглянемо три перші точки $(x_1;y_1)$, $(x_2;y_2)$, $(x_3;y_3)$. Проведемо через них параболу: $y = a + bx + cx^2$ і обчислимо інтеграл:

$$S_1 = \int_{x_1}^{x_3} (a_1 + b_1 x + c_1 x^2)\,dx = a_1(x_3 - x_1) + b_1\left(\frac{x_3^2 - x_1^2}{2}\right) + c_1\left(\frac{x_3^3 - x_1^3}{3}\right) =$$

$$= \frac{x_3 - x_1}{6}\,(6a_1 + 3b_1(x_3 + x_1) + 2c_1(x_3^2 + x_3 x_1 + x_1^2)) = \frac{x_3 - x_1}{6}\,(c_1 x_1^2 + b_1 x_1 + a_1 +$$

$$+ 4(c_1\left(\frac{x_3 + x_1}{2}\right)^2 + b_1\left(\frac{x_3 + x_1}{2}\right) + a_1) + c_1 x_3^2 + b_1 x_3 + a_1) = \frac{x_3 - x_1}{6}(y_1 + 4y_2 + y_3) =$$

$$= \frac{h}{3}(y_1 + 4y_2 + y_3).$$

Проводячи параболу через наступні три точки $(x_3;y_3)$,$(x_4;y_4)$, $(x_5;y_5)$ і обчислюючи інтеграл, отримаємо: $S_2 = \frac{h}{3}(y_3 + 4y_4 + y_5)$.

Для отримання наближеного значення інтегралу по всій області від x_1 до x_n необхідно знайти суму отриманих значень.

$$S = S1 + S2 + S\frac{n-1}{2} = \frac{h}{3}(y_1 + 4y_2 + y_3 + 4y_4 + y_5 + \ldots + y_{n-2} + 4y_{n-1} + y_n) =$$

$$= \frac{h}{3}[y_1 + y_n + 4(y_2 + y_4 + \ldots + y_{n-1}) + 2(y_3 + y_5 + \ldots + y_n)] = \frac{h}{3}[f(x_1) + f(x_n) +$$

$$+ 4\sum_{i-n} f(x_i) + 2\sum_{неп} f(x_i).$$

Якщо інтеграл необхідно обчислити із заданою точністю, то необхідно знайти n (число проміжків) для її забезпечення.

Приклад 7.1. Обчислити означений інтеграл

$$\int_{1.5}^{19.5} \frac{\ln(x^2 + 0.5)}{x + 2.8}\,dx, \quad n = 6$$

за формулами трапецій та Сімпсона з оцінкою похибок обчислень.

Розв'язок. Маємо підінтегральну функцію $f(x) = \frac{\ln(x^2 + 0.5)}{x + 2.8}$. Складемо таблицю значень підінтегральної функції та кінцевих різниць до 4-го порядку включно. Обчислимо значення $x_i = x_0 + i \cdot h$ $(i = 0,\ 1,\ 2,\ \ldots,\ n)$, $x_0 = a$, $x_n = b$ та обчислимо значення підінтегральної функції у знайдених вузлах $y_i = f(x_i)$. Кінцеві різниці першого порядку розраховують за формулою $\Delta y_i = y_{i+1} - y_i$ $(i = 0,\ 1,\ 2,\ \ldots,\ n-1)$, кінцеві різниці другого порядку: $\Delta^2 y_i = \Delta y_{i+1} - \Delta y_i$, $i = 0,\ 1,\ 2,\ \ldots,\ n-2$ та ін.

Таблиця 7.1

i	x_i	y_i	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
0	1,5	0,23526	0,18014	-0,2034	0,1885	-0,1698
1	4,5	0,4154	-0,0233	-0,0149	0,0187	-0,01706
2	7,5	0,392097	-0,0382	0,00382	0,00164	0,0391
3	10,5	0,3539	-0,03438	0,00382	0,04074	
4	13,5	0,31952	-0,02892	0,00546		
5	16,5	0,2906	0,01728	0,0462		
6	19,5	0,30788				
				Δ^2y=-0,03256		Δ^4y=-0,04925

Обчислимо інтеграл за формулою трапецій:

$$\int_{1.5}^{19.5} \frac{\ln(x^2+0,5)}{x+2,8}\,dx =$$

$$= 3\cdot\left(\frac{0,23516+0,30788}{2}+0,4154+0,392097+0,3539+0,31952+0,2906\right);$$

$$\int_{1.5}^{19.5} \frac{\ln(x^2+0,5)}{x+2,8}\,dx = 6,12926\cdot$$

Залишковий член

$$R_{TP} = -\frac{b-a}{12}\cdot\overline{\Delta^2 y} = -\frac{19,5-1,5}{12}\cdot(-0,03256) = 0,04884\,.$$

Обчислимо інтеграл за формулою Сімпсона:

$$\int_{1.5}^{19.5} \frac{\ln(x^2+0,5)}{x+2,8}\,dx =$$

$$= 1\cdot[0,23526+0,30788+2\cdot(0,392097+0,31952)+4\cdot(0,4154+0,3539+0,2906)];$$

$$\int_{1.5}^{19.5} \frac{\ln(x^2+0,5)}{x+2,8}\,dx = 6,20597\,.$$

Залишковий член

$$c = -\frac{b-a}{180}\cdot\overline{\Delta^4 y} = -\frac{19,5-1,5}{180}\cdot(-0,04925) = 0,004925\,.$$

7.6. Алгоритм реалізації принципу подвійного перерахунку та автоматичний вибір кроку інтегрування

Зауважимо, що на рис.7.7 значення в точках старої нумерації відповідають парним номерам в *новій нумерації*, тому ці значення можна не обчислювати, а зберегти в оперативній пам'яті для подальшого використання. Проте виявилось, що для наступного подвійного перерахунку зберігати достатньо лише інформацію про їх суму. Це одразу зменшує кількість необхідної оперативної пам'яті та розрахунків для перерахунку квадратурної формули.

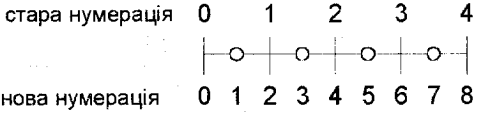


Рис. 7.7. Приклад реалізації подвійного перерахунку

Запишемо квадратурну формулу Сімпсона, яка застосована до інтервалу $[a;b]$ з кроком h:

$$I_0 = \frac{h}{3}\left[\sigma_0^{(0)} + 2\sigma_1^{(0)} + 2\sigma_1^{(0)} + 2\sigma_2^{(0)}\right]. \tag{7.33}$$

Позначивши $\sigma_3^{(0)} = \sigma_1^{(0)} + \sigma_2^{(0)}$, отримаємо:

$$I_0 = \frac{h}{3}\left[\sigma_0^{(0)} + 2\sigma_1^{(0)} + 2\sigma_3^{(0)}\right]. \tag{7.34}$$

Позначимо $\sigma_0^{(0)} = y_0 + y_n$, тоді

$$\sigma_1^{(0)} = \sum_{i=1}^{m} y_{2i-1}; \quad \sigma_2^{(0)} = \sum_{i=1}^{m-1} y_{2i}. \tag{7.35}$$

$$I_1 = \frac{h}{2\cdot 3}\left[\sigma_0^{(0)} + 2\sigma_1^{(1)} + 2\sigma_3^{(1)}\right] \tag{7.36}$$

$$\sigma_1^{(1)} = \sum_{i=1}^{2m} y_{2i-1}; \quad \sigma_3^{(1)} = \sigma_1^{(1)} + \sigma_2^{(1)}; \quad \sigma_2^{(1)} = \sigma_3^{(0)}. \tag{7.37}$$

Загальні формули для довільного кроку матимуть вигляд:

$$I_k = \frac{h}{2^k\cdot 3}\left[\sigma_0^{(0)} + 2\sigma_1^{(k)} + 2\sigma_3^{(k)}\right]. \tag{7.38}$$

$$\sigma_1^{(k)} = \sum_{i-1}^{2^k\cdot m} y(x_{2i-1}). \tag{7.39}$$

$$\sigma_2^{(k)} = \begin{cases} \sum_{i=1}^{m-1} y(x_{2i}), & k = 0 \\ \sigma_3^{(k-1)}, & k > 0 \end{cases} \tag{7.40}$$

$$\sigma_3^{(k)} = \sigma_1^{(k)} + \sigma_2^{(k)}, \quad k = 0,1,\ldots . \tag{7.41}$$

$$\begin{aligned} x_i &= a + ih, \quad k = 0 \\ x_i &= a + i\frac{h}{2}, \quad k = 1 \\ &\ldots\ldots\ldots\ldots \\ x_i &= a + i\frac{h}{2^n}, \quad k = крок \end{aligned} \tag{7.42}$$

Умова зупинки циклу така: $|I_{k+1} - I_k| < E$.

7.7. Наближене обчислення багатократних інтегралів

Розглянемо k-мірний інтеграл виду:

$$I = \underbrace{\int \ldots \int}_{k} f(x)dx_1 \ldots dx_k , \tag{7.43}$$

де $x = (x_1, x_2, \ldots, x_k)$ – деяка k-мірна точка. Далі для простоти всі рисунки будуть зроблені для випадку k=2 9(рис. 7.8).

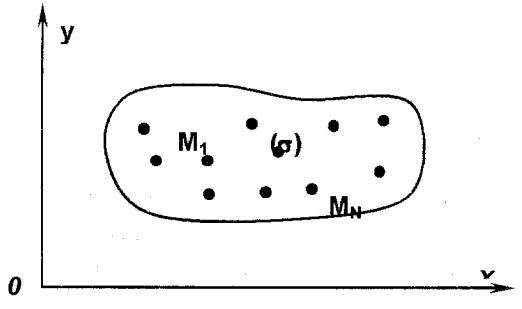


Рис. 7.8. Приклад деякої *2*-мірної точки

Кубатурні формули чи інакше **формули чисельних кубатур** призначені для чисельного обчислення кратних інтегралів. Нехай функція $z = f(x_1, x_2, \ldots, x_k)$

визначена і неперервна в деякій обмеженій області σ. У цій області σ вибирається система точок (вузлів) $M_i(x_{1_i}, x_{2_i}, ..., x_{k_i})$ при $i = \overline{1,N}$. Для обчислення інтеграла $I = \underbrace{\int ... \int}_{k} f(x_1, x_2, ..., x_k) dx_1 ... dx_k$ приблизно вважають:

$$\underbrace{\int ... \int}_{k} f(x_1, x_2, ..., x_k) dx_1 ... dx_k = \sum_{i=1}^{N} A_i f(x_{1_i}, x_{2_i}, ..., x_{k_i}). \qquad (7.44)$$

Щоб знайти коефіцієнти A_i, поставимо вимогу точного виконання кубатурної формули (7.44) для всіх поліномів

$$P_n(x_1, x_2, ..., x_k) = \sum_{l_1 + l_2 + ... + l_k \le n} c_{l_1, l_2, ..., l_k} x_1^{l_1} \cdot x_2^{l_2} \cdot ... \cdot x_k^{l_k}, \qquad (7.45)$$

степінь яких не перевищує заданого числа n. Для цього необхідно й достатньо, щоб формула (7.44) була точною для добутку степенів $x_1^{l_1} \cdot x_2^{l_2} \cdot ... \cdot x_k^{l_k}$ $(l_1, l_2, ..., l_k = \overline{0,n}$; $l_1 + l_2 + ... + l_k \le n)$. Візьмемо в (7.43) $f(x_1, x_2, ..., x_k) = x_1^{l_1} \cdot x_2^{l_2} \cdot ... \cdot x_k^{l_k}$, отримаємо:

$$I_{l_1, l_2, ..., l_k} = \underbrace{\int ... \int}_{k} x_1^{l_1} \cdot x_2^{l_2} \cdot ... \cdot x_k^{l_k} dx_1 ... dx_k = \sum_{i=1}^{N} A_i x_1^{l_1} \cdot x_2^{l_2} \cdot ... \cdot x_k^{l_k}$$
$$(l_1, l_2, ..., l_k = \overline{0,n}; \quad l_1 + l_2 + ... + l_k \le n). \qquad (7.46)$$

Таким чином, коефіцієнти A_i формули (7.44), взагалі кажучи, можуть бути визначені з системи лінійних рівнянь (7.46).

Для того щоб система (7.46) була визначеною, необхідно, щоб число невідомих N було рівне числу рівнянь. У випадку k=2 одержуємо:

$$N = (n+1) + n + ... + 1 = \frac{(n+1)(n+2)}{2}. \qquad (7.47)$$

7.8. Метод комірок

Розглянемо k-мірний інтеграл по просторовому паралелепіпеду $G\ (a_1 \le x_1 \le b_1,\ a_2 \le x_2 \le b_2, ..., a_k \le x_k \le b_k)$. Аналогічно до формули середніх можна приблизно замінити функцію на її значення в центральній точці паралелепіпеда. Тоді інтеграл легко обчислюється:

$$\underbrace{\int\ldots\int}_{k} f(x_1,x_2,\ldots,x_k)dx_1\ldots dx_k \approx Sf(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_k),$$
$$S=\prod_{i=1}^{k}(b_i-a_i),\quad x_i=\frac{a_i+b_i}{2},\quad (i=\overline{1,k}). \tag{7.48}$$

Для підвищення точності можна розбити область на прямокутні комірки (рис. 7.9). Приблизно обчислюючи інтеграл у кожній комірці за формулою середніх і позначаючи через $S_i,\bar{x}_{1_i},\bar{x}_{2_i},\ldots,\bar{x}_{k_i}$ відповідно площу комірки й координати її центра, одержимо:

$$I=\underbrace{\int\ldots\int}_{k} f(x_1,x_2,\ldots,x_k)dx_1\ldots dx_k \approx \sum_i S_i f(x_{1_i},x_{2_i},\ldots,x_{k_i}). \tag{7.49}$$

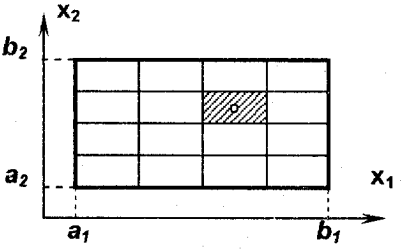


Рис. 7.9. Приклад заміни функції на її значення в центральній точці паралелепіпеда

Праворуч стоїть інтегральна сума. Отже, для будь-якої неперервної функції $f(x_1,x_2,\ldots,x_k)$ вона збігається зі значенням інтеграла, коли периметри всіх комірок прямують до нуля. Оцінимо похибку інтегрування. Формула (7.48) є точною для $f(x_1,x_k,\ldots,x_k)=const$. Але безпосередньо підстановкою легко переконатися, що формула є точною і для будь-якої лінійної функції. Справді, розкладемо функцію за формулою Тейлора:

$$f(x_1,x_2,\ldots,x_k)=f(\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_k)+\sum_{i=1}^{k}\xi_i f'_{x_i}+\frac{1}{2}\sum_{i=1}^{k}\xi_i^2 f''_{x_ix_i}+\ldots, \tag{7.50}$$

де $\xi_i=x_i-\bar{x}_i$, $(i=\overline{1,k})$, а всі похідні беруться в центрі комірки. Підставляючи цей розклад в праву і ліву частини квадратурної формули (7.48) і порівнюючи їх, аналогічно одновимірному випадку легко отримаємо вираз похибки цієї формули:

$$R\equiv\underbrace{\int\ldots\int}_{k\text{ раз}} f(x_1,x_2,\ldots,x_k)dx_1\ldots dx_k \approx \frac{1}{24}S\sum_{i=1}^{k}(b_i-a_i)^2 f''_{x_ix_i}, \tag{7.51}$$

тому що всі члени розкладу, які непарні щодо центра симетрії комірки, взаємно знищуються.

Нехай в узагальненій квадратурній формулі (7.49) сторони просторового паралелепіпеда розбиті відповідно на N_1, N_2, …, N_k рівних частин. Тоді похибка інтегрування (7.51) для одиничної комірки дорівнює:

$$R_i \approx \frac{1}{24} S_i \left[\sum_{j=1}^{k} \left(\frac{b_j - a_j}{N_j} \right)^2 f''_{x_j x_j} \right]_i . \tag{7.52}$$

Підсумовуючи цей вираз по всіх комірках, отримаємо похибку узагальненої формули:

$$R \approx \frac{1}{24} \sum_{i=1}^{k} \left(\frac{b_i - a_i}{N_i} \right)^2 \underbrace{\int \ldots \int}_{k} f''_{x_i x_i}(x_1, x_2, \ldots, x_k) dx_1 \ldots dx_k = \tag{7.53}$$

$$= O(N_1^{-2} + N_2^{-2} + \ldots + N_k^{-2})$$

тобто формула має другий порядок точності. При цьому, як і для одного виміру, можна застосовувати метод Рунге-Ромберга, але при одному додатковому обмеженні: збільшується густина сітки по кожній змінній в однакову кількість разів.

Узагальнимо формулу комірок на більш складні області. Розглянемо випадок k=2. Для лінійної функції $f(x_1, x_2)$ формула типу (7.48) буде точною в області довільної форми, якщо під S розуміти площу області, а під x_1, x_2 – координати центру ваги, що обчислюються за звичайними формулами:

$$S = \iint dx_1 dx_2, \quad x_1 = \frac{1}{S} \iint x_1 dx dy, \quad x_2 = \frac{1}{S} \iint x_2 dx_1 dx_2. \tag{7.54}$$

Зрозуміло, що практичну цінність це має тільки для областей простої форми, де площа і центр ваги легко визначається; наприклад, для трикутника, правильного багатокутника, трапеції. Але це значить, що узагальнену формулу (7.49) можна застосовувати до областей, які обмежені ламаною лінією, тому що таку область завжди можна розбити на прямокутники й трикутники.

Для області з довільною границею формулу (7.49) застосовують іншим способом. Накладемо на область G сітку з K-мірних паралелепіпедів (рис. 7.10). Ті комірки сітки, усі точки яких належать області, назвемо **внутрішніми**; якщо частина точок комірки належить області, а частина – ні, то назвемо комірку **граничною**. Об'єм внутрішньої комірки дорівнює добутку її сторін. Об'ємом граничної комірки будемо вважати об'єм тієї її частини, що попадає в середину G; цей об'єм обчислимо приблизно. Ці площі підставимо в (7.49) і обчислимо інтеграл.

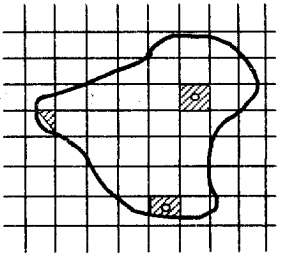


Рис. 7.10. Накладена на область G сітка з K-мірних паралелепіпедів

Оцінимо похибку формули (7.49). У кожній внутрішній комірці похибка складає $O(N^{-2})$ відносно значення інтеграла по цій комірці. У кожній граничній комірці відносна похибка є $O(N^{-1})$, тому що центр комірки не співпадає з центром ваги частини, що входить до інтегралу. Але самих граничних комірок приблизно в N разів менше, ніж внутрішніх. Тому при сумуванні по комірках загальна похибка буде $O(N^{-2})$, якщо функція двічі неперервно-диференційована; це означає другий порядок точності. Обчислення об'єму граничної комірки досить складне, тому що вимагає визначення положення границі усередині комірки. Можна обчислювати інтеграли по граничних комірках більш грубо або взагалі не включати їх у суму (7.49). Похибка при цьому буде $O(N^{-1})$, і для більшої точності буде потрібна більш щільна сітка. До області довільної форми метод комірок важко застосовувати; тому завжди бажано заміною змінних перетворити область інтегрування в прямокутний паралелепіпед (це відноситься практично до всіх методів обчислення кратних інтегралів).

7.9. Послідовне інтегрування

Знову розглянемо інтеграл по k-мірній області, яка розбита сіткою на комірки (рис. 7.9). Його можна обчислити послідовним інтегруванням:

$$I = \underbrace{\int_{a_1}^{b_1} \dots \int_{a_k}^{b_k}}_{k} f(x_1, x_2, \dots, x_k) dx_1 \dots dx_k = \int_{a_k}^{b_k} F_k(x_k) dx_k \,, \tag{7.55}$$

$$F_k(x_k) = \underbrace{\int_{a_1}^{b_1} \dots \int_{a_{k-1}}^{b_{k-1}}}_{k-1} f(x_1, x_2, \dots, x_k) dx_1 \dots dx_{k-1} = \int_{a_{k-1}}^{b_{k-1}} F_{k-1}(x_{k-1}) dx_{k-1} \,, \tag{7.56}$$

$$F_{k-1}(x_{k-1}) = \underbrace{\int_{a_1}^{b_1} \dots \int_{a_{k-2}}^{b_{k-2}}}_{k-2} f(x_1, x_2, \dots, x_k) dx_1 \dots dx_{k-2} \tag{7.57}$$

Кожен однократний інтеграл легко обчислюється на цій сітці по квадратурних формулах типу:

$$F = \sum_{i=0}^{n} c_i f(x_i) + R, \quad c_i = \int_a^b \varphi_i(x)\rho(x)dx, \quad R = \int_a^b r(x)\rho(x)dx, \tag{7.58}$$

Послідовне інтегрування в усіх напрямках приводить до кубатурних формул, що є прямим добутком одновимірних квадратурних формул:

$$I \approx \sum_{l_1, l_2, \dots, l_k} c_{l_1, l_2, \dots, l_k} f(x_1, x_2, \dots, x_k), \quad c_{l_1, l_2, \dots, l_k} = \prod_{i=1}^{l} c_{l_i}. \tag{7.59}$$

Наприклад, при k=2, якщо по кожному напрямку обрана узагальнена формула трапецій, а сітка рівномірна, то вага кубатурної формули дорівнює $\frac{c_{l_1 l_2}}{h_{x_1} h_{x_2}} = 1, \frac{1}{2}$ і $\frac{1}{4}$ відповідно для внутрішніх, граничних та кутових вузлів сітки. Легко показати, що для двічі неперервно-диференційованих функцій ця формула має другий порядок точності і застосуємо до неї метод Рунге-Ромберга. Взагалі кажучи, для різних напрямків можна використовувати квадратурні формули різних порядків точності $p_i \ (i = \overline{1,k})$. Тоді головний член похибки має вигляд:

$$R = O\left(h_{x_1}^{p_1} + h_{x_2}^{p_2} + \dots + h_{x_k}^{p_k}\right). \tag{7.60}$$

Бажано для всіх напрямків використовувати квадратурні формули однакового порядку точності. Можна підібрати вагу і положення ліній сітки так, щоб одновимірна квадратурна формула була б точною для багаточлена максимального степеня, тобто була би формулою Гауса. Тоді, для випадку k=2:

$$\begin{aligned} c_{ij} &= \frac{1}{4}(b_1 - a_1)(b_2 - a_2)\gamma_i\gamma_j, \quad x_{1_i} = \frac{1}{2}(a_1 + b_1) + \frac{1}{2}(b_1 - a_1)\xi_i, \\ x_{2_j} &= \frac{1}{2}(a_2 + b_2) + \frac{1}{2}(b_2 - a_2)\xi_i, \quad 1 \le i, j \le n, \end{aligned} \tag{7.61}$$

де ξ, γ – нулі багаточленів Лежандра і відповідні ваги. Ці формули розраховані на функції високої гладкості і дають для них велику економію в кількості вузлів у порівнянні з більш простими формулами.

Метод послідовного інтегрування можна застосовувати до області довільної форми, наприклад, із криволінійною границею. Розглянемо цей випадок при k=2. Для цього проведемо через область хорди, паралельні осі x_1, і на них введемо вузли, розташовані на кожній хорді так, як нам потрібно (рис. 7.11).

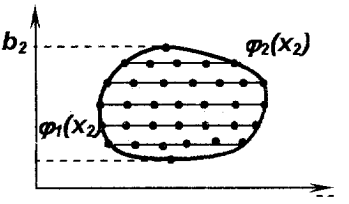


Рис. 7.11

Подамо інтеграл у вигляді:

$$I=\iint\limits_{G} f(x_1,x_2)dx_1dx_2=\int\limits_{a_2}^{b_2} F(x_2)dx_2,\quad F(x_2)=\int\limits_{\varphi_1(x_2)}^{\varphi_2(x_2)} f(x_1,x_2)dx_1\,. \tag{7.62}$$

Спочатку обчислимо інтеграл по x_1 вздовж кожної хорди по якій-небудь одновимірній квадратурній формулі, використовуючи введені вузли. Потім обчислимо інтеграл по x_2; тут вузлами будуть служити проекції хорд на вісь ординат. При обчисленні інтеграла по x_2 є одна особливість. Якщо область обмежена гладкою кривою, то при $x_2\to a_2$ довжина хорди прямує до нуля не лінійно, а як $\sqrt{x_2-a_2}$. Отже, поблизу цієї точки $F(x_2)\sim\sqrt{x_2-a_2}$. Те ж буде при $x_2\to b_2$. Тому інтегрувати безпосередньо $F(x_2)$ по формулах високого порядку точності безглуздо. Доцільно виділити з $F(x_2)$ основну особливість у вигляді ваги $\rho(x_2)=\sqrt{(b_2-x_2)(x_2-a_2)}$, якій відповідають ортогональні багаточлени Чебишева другого роду. Тоді друге інтегрування виконується по формулах Гауса-Кристофеля:

$$I=\int\limits_{a_1}^{b_2} F(x_2)dx_2\approx\sum_{j=1}^{n}\frac{b_2-a_2}{2}\gamma_j\psi\left(\frac{a_2+b_2}{2}+\frac{a_2-b_2}{2}\xi_j\right), \tag{7.63}$$

де $\psi(x_2)=\dfrac{F(x_2)}{\rho(x_2)}$, а $\xi_j=\cos\left[\dfrac{\pi j}{n+1}\right]$ і γ_j – нулі і ваги багаточленів Чебишева другого роду. Щоб можна було застосовувати цю формулу, треба ординати хорд на рис. 7.11 заздалегідь вибрати відповідно до вузлів (7.63). Якщо це не було зроблено, то доведеться обмежитися інтегруванням $F(x_2)$ по узагальненій формулі трапецій, причому її ефективний порядок точності в цьому випадку буде нижче другого.

7.10. Кубатурна формула типу Сімпсона

Нехай спочатку область інтегрування – K-мірний просторовий паралелепіпед $R=\{a_1\le x_1\le A_1;a_2\le x_2\le A_2;...;a_k\le x_k\le A_k\}$ (рис. 7.12), сторони якого паралельні до осей координат.

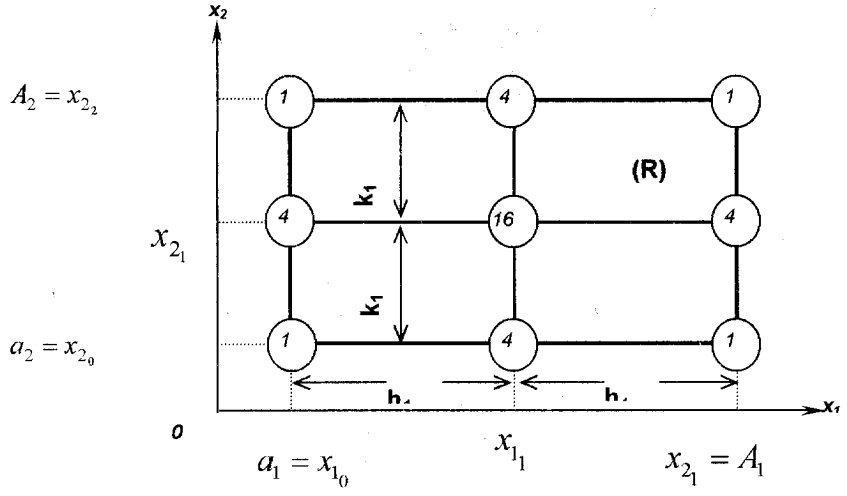


Рис. 7.12.

Кожен з проміжків $[a_i, A_i]$ $(i = \overline{1,k})$ розбиємо навпіл точками:

$$x_{i_0} = a_i,\ x_{i_1} = a_i + h_i,\ x_{i_2} = a_i + 2h_i = A_i,\ i = \overline{1,k}\ , \tag{7.64}$$

де $h_i = \dfrac{A_i - a_i}{2}$. Усього таким чином, одержимо 3^k точок сітки. Маємо:

$$\underbrace{\int\ldots\int}_{k} f(x_1, x_2, \ldots, x_k)dx_1\ldots dx_k = \underbrace{\int_{a_1}^{A_1} dx_1 \int_{a_2}^{A_2} dx_2 \ldots \int_{a_k}^{A_k} f(x_1, x_2, \ldots, x_k)dx_k}_{k} \tag{7.65}$$

Знаходимо k-мірний інтеграл, обчислюючи кожен внутрішній інтеграл за квадратурною формулою Сімпсона на відповідному відрізку. Проведемо все обчислення для випадку K=2:

$$\begin{aligned}\iint f(x_1, x_2)dx_1dx_2 &= \int_{a_1}^{A_1} dx_1 \cdot \frac{h_2}{3}\left[f(x_1, x_0) + 4f(x_1, x_{2_1}) + f(x_1, x_{2_2})\right] = \\ &= \frac{h_2}{3}\left[\int_{a_1}^{A_1} f(x_1, x_0)dx_1 + 4\int_{a_1}^{A_1} f(x_1, x_{2_1})dx_1 + \int_{a_1}^{A_1} f(x_1, x_{2_2})dx_1\right].\end{aligned} \tag{7.66}$$

Застосовуючи до кожного інтеграла знову формулу Сімпсона, одержимо:

$$\iint f(x_1,x_2)dx_1dx_2 = \frac{h_1h_2}{9}\{[f(x_{1_0},x_{2_0})+4f(x_{1_1},x_{2_0})+f(x_{1_2},x_{2_0})]+$$
$$+4[f(x_{1_0},x_{2_1})+4f(x_{1_1},x_{2_1})+f(x_{1_2},x_{2_1})]+ \qquad (7.67)$$
$$+[f(x_{1_0},x_{2_2})+4f(x_{1_1},x_{2_2})+f(x_{1_2},x_{2_2})]\}.$$

$$\iint f(x_1,x_2)dx_1dx_2 = \frac{h_1h_2}{9}\{[f(x_{1_0},x_{2_0})+4f(x_{1_2},x_{2_0})+f(x_{1_2},x_{2_2})]+ \qquad (7.68)$$
$$+4[f(x_{1_1},x_{2_0})+f(x_{1_0},x_{2_1})+f(x_{1_2},x_{2_1})+f(x_{1_1},x_{2_2})]+16f(x_{1_1},x_{2_1})\}.$$

Формулу (7.68) називають **кубатурною формулою Сімпсона**. Отже,

$$\iint\limits_{(R)} f(x_1,x_2)dx_1dx_2 = \frac{h_1h_2}{9}(\sigma_0+4\sigma_1+16\sigma_2), \qquad (7.69)$$

де σ_0 – сума значень підінтегральної функції $f(x_1,x_2)$ у вершинах прямокутника R, σ_1 – сума значень $f(x_1,x_2)$ у серединах сторін прямокутника R, $\sigma_2 = f(x_{1_1},x_{2_1})$ – значення функції $f(x_1, x_2)$ в центрі прямокутника R. Кратності цих значень позначені на рис. 7.12.

Якщо розміри просторового паралелепіпеда $R=\{a_1 \le x_1 \le A_1; a_2 \le x_2 \le A_2;...; a_k \le x_k \le A_k\}$ великі, то для збільшення точності кубатурної формули область R розбивають на систему паралелепіпедів, до кожного з яких застосовують кубатурну формулу Сімпсона.

Знову розглянемо випадок K=2. Покладемо, що сторони прямокутника R ми розділили відповідно на n_1 і n_2 рівних частин; у результаті вийшла відносно велика сітка n_1n_2 прямокутників (на рис. 7.13 вершини цих прямокутників позначені більшими кружечками). Кожен з цих прямокутників у свою чергу розділимо на чотири рівні частини. Вершини цієї останньої дрібної сітки прямокутників приймемо за вузли M_{ij} кубатурної формули.

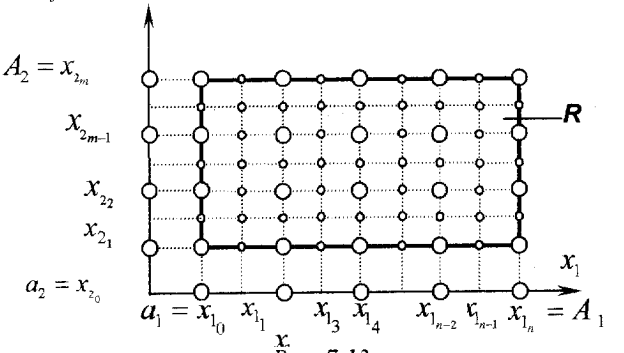


Рис. 7.13.

Нехай $h_1 = \frac{A_1 - a_1}{2n_1}$ і $h_2 = \frac{A_2 - a_2}{2n_2}$. Тоді сітка вузлів буде мати наступні координати:

$$x_{1_i} = x_{1_0} + ih_1 \quad (x_{1_0} = a_1;\ i = \overline{0,2n_1}), \tag{7.70}$$

$$x_{2_j} = x_{2_0} + jh_2 \quad (x_{2_0} = a_2;\ j = \overline{0,2n_2}). \tag{7.71}$$

Для скорочення введемо позначення $f(x_{1_i}, x_{2_j}) = f_{ij}$.

Застосовуючи формулу (7.68) до кожного з прямокутників великої сітки, будемо мати (рис. 7.14):

$$\iint\limits_{(R)} f(x_1, x_2)dx_1dx_2 = \frac{h_1h_2}{9}\sum_{i=0}^{n}\sum_{j=0}^{m}[(f_{2i,2j} + f_{2i+2,2j} + f_{2i+2,2j+2} + f_{2i,2j+2}) + \\ + 4(f_{2i+1,2j} + f_{2i+2,2j+1} + f_{2i+1,2j+2} + f_{2i,2j+1}) + 16f_{2i+1,2j+1}]. \tag{7.72}$$

Звідси, приведенням подібних членів, остаточно знаходимо:

$$\iint\limits_{(R)} f(x_1, x_2)dx_1dx_2 = \frac{h_1h_2}{9}\sum_{i=0}^{2n}\sum_{j=0}^{2m}\lambda_{ij}f_{ij}, \tag{7.73}$$

де коефіцієнти λ_{ij} є відповідними елементами матриці

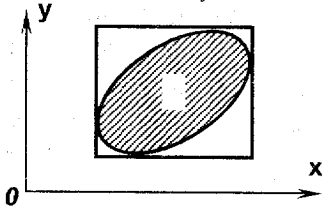


Рис. 7.14.

$$A = \begin{bmatrix} 1 & 4 & 2 & 4 & 2 & \dots & 4 & 2 & 4 & 1 \\ 4 & 16 & 8 & 16 & 8 & \dots & 16 & 8 & 16 & 4 \\ 2 & 8 & 4 & 8 & 4 & \dots & 8 & 4 & 8 & 2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 2 & 8 & 4 & 8 & 4 & \dots & 8 & 4 & 8 & 2 \\ 4 & 16 & 8 & 16 & 8 & \dots & 16 & 8 & 16 & 4 \\ 1 & 4 & 2 & 4 & 2 & \dots & 4 & 2 & 4 & 1 \end{bmatrix}$$

Якщо область інтегрування σ – довільна, то будуємо паралелепіпед $R \supset \sigma$, сторони якого паралельні до осей координат. Розглянемо допоміжну функцію

$$f^*(x_1, x_2, \dots, x_k) = \begin{cases} f(x_1, x_2, \dots, x_k), & якщо\ (x_1, x_2, \dots, x_k) \in \sigma, \\ 0, & якщо\ (x_1, x_2, \dots, x_k) \in R - \sigma. \end{cases} \tag{7.74}$$

У такому випадку маємо:

$$\underbrace{\int_{(\sigma)}\dots\int}_{k} f(x_1, x_2, \dots, x_k)dx_1\dots dx_k = \underbrace{\int_{(R)}\dots\int}_{k} f^*(x_1, x_2, \dots, x_k)dx_1\dots dx_k \tag{7.75}$$

Останній інтеграл наближено може бути обчислений за загальною кубатурною формулою (7.73). Очевидно, що застосовуючи квадратурні формули Ньютона-Котеса

спочатку до внутрішнього інтеграла, а потім до зовнішнього, отримаємо новий тип формул – кубатурний. Обмежимось розглядом кубатурної формули Сімпсона, яку побудуємо для прямокутника, який розбитий довільним чином на чотири частини так, щоб по кожній стороні крок інтегрування був однаковий (рис. 7.15).

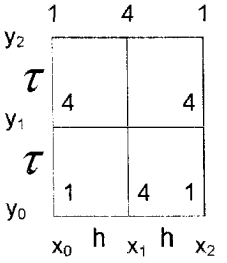


Рис. 7.15. Прямокутник, розбитий довільним чином на чотири частини так, щоб по кожній стороні крок інтегрування був однаковий при $h \neq \tau$

$$\int_{x_0}^{x_2} dx \int_{y_0}^{y_2} f(x,y)dy = \frac{\tau}{3}\int_{x_0}^{x_2}\left[f(x,y_0)+4f(x,y_1)+f(x,y_2)\right]dx =$$

$$= \frac{\tau}{3}\left[\int_{x_0}^{x_2} f(x,y_0)dx + 4\int_{x_0}^{x_2} f(x,y_1)dx + \int_{x_0}^{x_2} f(x,y_2)dx\right] = \tag{7.76}$$

$$= \frac{h\tau}{9}[f(x_0,y_0)+4f(x_1,y_0)+f(x_2,y_0)+4f(x_0,y_1)+$$

$$+16f(x_1,y_1)+4f(x_2,y_1)+f(x_0,y_2)+4f(x_1,y_2)+4f(x_2,y_2)].$$

Зводимо всі подібні члени та отримаємо:

$$\int_{x_0}^{x_2} dx \int_{y_0}^{y_2} f(x,y)dy = \frac{h\tau}{9}\left[\sigma_0+4\sigma_1+16\sigma_2\right] \tag{7.77}$$

$$f_{ij} \equiv f(x_i, y_j). \tag{7.78}$$

$$\sigma_0 = f_{00}+f_{02}+f_{20}+f_{22},\ \sigma_1 = f_{10}+f_{12}+f_{01}+f_{21},\ \sigma_2 = f_{11}. \tag{7.79}$$

Механізм побудови кубатурної формули Сімпсона полягає у розбитті на парну кількість прямокутників як по вісі абсцис, так і по вісі ординат, та подальшому групуванні всіх отриманих елементів (рис. 7.16).

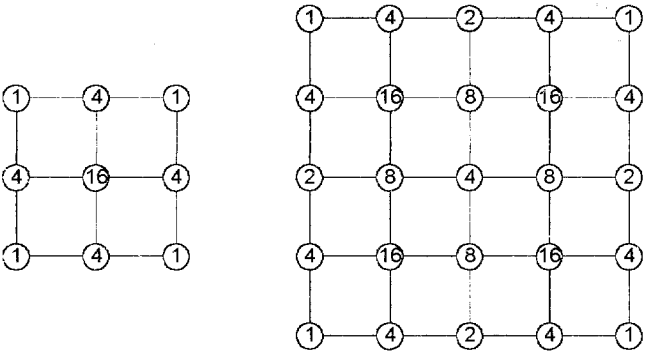


Рис. 7.16. Приклади побудови кубатурної формули Сімпсона

Загальна кубатурна формула Сімпсона матиме вигляд

$$\iint\limits_D f(x,y)dxdy = \frac{h\tau}{9}\sum_{i-1}^{2n}\sum_{j-1}^{2m}\lambda_{ij}f_{ij}\,, \tag{7.80}$$

причому $h = \dfrac{b-a}{2n}$ і $\tau = \dfrac{d-c}{2m}$. Як видно з симетричного вигляду матриці Δ, в памяті комп'ютера можна зберігати або її половину, або одну стрічку, яка відповідатиме першій чи останній стрічці матриці Δ, або стовпчику цієї матриці. Цей підхід можна застосовувати і при обчисленні потрійних інтегралів, проте роль відповідної матриці коефіцієнтів буде відігравати гіперкуб, що має просторову структуру (аналог кришталевої решітки).

7.11. Методи чисельного інтегрування

7.11.1. Задача обчислення визначеного інтеграла

Відомо, що існує первісна F до функції f, то визначений інтеграл

$$I = \int\limits_b^a f(x)dx \tag{7.81}$$

обчислюють за формулою

$$I = \int\limits_b^a f(x)dx = F(b) - F(a)\,. \tag{7.82}$$

Якщо існує первісна F, але в елементарних функціях не існує, то інтеграл Рімана можна наближено обчислити за означенням

$$I = \int_b^a f(x)dx \approx \sum_{i=1}^{n} f(x_i)\Delta x_i \,. \tag{7.83}$$

Але ця сума збігається дуже повільно при n, що прямує до нескінченності. Це означає, що для обчислення формула (7.83) непридатна.

Наближені методи обчислення визначеного інтегралу I об'єднують у формулу:

$$I \approx \sum_{i=1}^{n} K_i^{(n)} f(x_i^{(n)}) \equiv I_n(\mathbf{f}) \,. \tag{7.84}$$

Означення 1. Формулу(7.84) назвемо квадратурною формулою (КФ) обчислення визначеного інтегралу; $\{x_i^{(n)}\} \subset [a,b] \subset \mathbf{R}^1$ – вузлами КФ, $\{K_i^{(n)}, i = \overline{1,n}\} \subset \mathbf{R}^1$ – коефіцієнтами квадратурної формули.

Різницю

$$R_n(\mathbf{f}) \equiv I - \sum_{i=1}^{n} K_i^{(n)} f(x_i^{(n)}) \equiv I - I_n(\mathbf{f}) \tag{7.85}$$

назвемо *похибкою квадратурної формули* (7.84).

7.11.2. Три підходи квадратурної формули

1. Квадратурні формули з найкращою оцінкою функцій

Нехай $f \in \zeta$, тоді $\{x_i^{(n)}\}$ і $\{K_i^{(n)}\}$ знаходимо за умови, щоб супреміум був мінімальним з усіх залишкових членів:

$$\inf \sup_{f \in \zeta} |R_n(\mathbf{f})| \,. \tag{7.86}$$

Формули (7.84) називаються *формулами з найкращою оцінкою в класі функцій* $\mathbf{f}$, що побудовані за умови (7.86).

2. Квадратурні формули найкращого степеня точності

Вузли $\{x_i^{(n)}\}$, $\{K_i^{(n)}\}$, $i = \overline{1,n}$ у формулі (7.84) вибираємо так, щоб співвідношення (7.84) було точним $I = \sum_{i=1}^{n} K_i^{(n)} f(x_i^{(n)})$ для многочлена максимільного степеня, тобто

$$f(x) = P_m(x) \,,\ m \le n \,. \tag{7.87}$$

Тоді, квадратурна формула (7.84), яка побудована за (7.87) називається *квадратурною формулою найкращого степеня точності.*

3. Інтерполяційні квадратурні фомули

При цьому підході фіксуємо вузли рівновіддаленими $\{x_i^{(n)} \equiv x_0 + i \cdot h,\ i = \overline{1,n}\} \subset [a,b]$ і знаходимо $\{K_i^{(n)}\}$, будуючи інтерполяційний многочлен за вузлами КФ

$$f(x) \approx L_{n-1}(x) = \sum_{i=1}^{n} f(x_i, \mathbf{f}(x_i^{(n)})) \frac{\omega_n(x)}{\omega_n'(x_i^{(n)})(x - x_i^{(n)})}. \tag{7.88}$$

$$\omega_n(x) \equiv (x - x_1^{(n)})(x - x_2^{(n)})...(x - x_n^{(n)}). \tag{7.89}$$

$\{K_i^{(n)}\}$ визначаємо при інтегруванні $\{L_{n-1}(x)\}$. Формули (7.84), як будуються при інтегруванні $\mathbf{f}(x) = L_n(x)$ назвемо *інтерполяційною квадратурною формулою* (ІКФ).

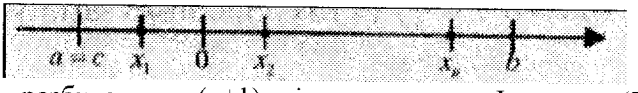
7.11.3. Інтерполяційні квадратурні формули Ньютона-Котеса

Отримаємо інтерполяційну квадратурну формулу для обчислення інтегралу $I(1,1)$. Фіксуємо вузли(КФ): $\{x_i^{(n)},\ i = \overline{1,n}\} \subset [a,b], i \neq j$, причому

$$x_i^{(n)} = c + i \cdot h,\ i = \overline{1,n},\ h > 0. \tag{7.90}$$

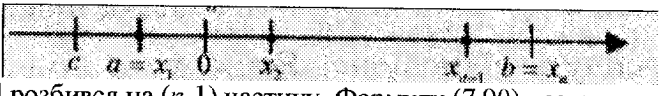
Розглянемо два випадки:

$$i, a = c \neq x_1, b = c + (n+1)h \neq x_n. \tag{7.91}$$



Відрізок $[a,b]$ розбився на $(n+1)$ рівних частин. Формулу (7.90) назвемо відкритого типу, якщо виконується (7.91).

$$c + h = a = x_1,\ b = c + nh = x_n. \tag{7.92}$$



Відрізок $[a,b]$ розбився на $(n-1)$ частину. Формулу (7.90) назвемо закритого типу, якщо виконується (7.92).

Теорема 1. Нехай

1) $f \in C^{n-1}([a,b]); \exists \mathbf{f}^n \in C([a,b])$;

2) $\{x_i^{(n)} = c + ih,\ i = \overline{1,n}\} \subset [a,b]$;

3) $\{x^{(j)}, j = \overline{1,n}\}$ – Чебишевська система функцій. Тоді

$$I = I_n(\mathbf{f}) + R_n(\mathbf{f}). \tag{7.93}$$

$$I(\mathbf{f}) = \sum_{i=1}^{n} K_{i,m}^{(n)} f(x_i^{(n)}). \tag{7.94}$$

$$K_{i,m}^{(n)} \equiv \frac{b-a}{n-1+2m} \cdot \frac{(-1)^{n-1}}{(i-1)!(n-1)!} \int_{1-m}^{n-m} \frac{(y-1)...(y-n)}{(y-i)} dy. \tag{7.95}$$

$$R_n(\mathbf{f}) \equiv \frac{b-a}{n-1+2m} \int_{1-m}^{n+m} (y-1)...(y-n) F(y,1,...,n) dy. \tag{7.96}$$

$$F(y) = f(c+hy). \tag{7.97}$$

$$\begin{cases} a = c + (1-m)h, \\ b = c + (n+m)h. \end{cases} \tag{7.98}$$

При m=1 одержимо ІКФ відкритого типу, при m=0 – ІКФ закритого типу.

Доведення. Маємо $I = \int_b^a f(x)dx$. Зробимо заміну $x=c+hy$, $dx=hdy$. Тоді вузли $\{x_i^{(n)}, i=\overline{1,n}\}$ перейдуть у вузли нової змінної $y_i = i$, $i=\overline{1,n}$. Дійсно, $x_i^{(n)} \equiv c+ih = c+yh$, звідки нижня і верхня межі інтегрування перейдуть відповідно: якщо $x=a=c+hy=c+(1-m)$, то $y=1-m$; якщо $x=b=c+hy=c+(n+m)h$, то $y=n+m$. Тобто маємо таблицю старих і нових вузлів:

X	x_1^n	x_2^n	...	x_n^n	a	b
Y	1	2	...	n	1-m	$n+m$

Значить

$$I = h \int_{1-m}^{n+m} F(y)dy. \tag{7.99}$$

За умови теореми 1: $F \in C^{(n-1)}([1-m;n+m])$ і $\exists F^{(n)}$, $y \in (1-m,n+m)$. Тоді можна застосувати теорему Лагранжа з рівновіддаленими вузлами

$$F(y) = \sum_{i=1}^{n} \frac{(y-1)(y-2)...(y-i+1)(y-i-1)...(y-n)}{(y-1)(y-2)...1(-1)(-2)...(i-n)} F(i) +$$

$$+ (y-1)(y-2)...(y-n)F(y;1;2;...;n) = \sum_{i=1}^{n} \frac{(-1)^{n-1} C_n^i F(i)\omega_n(y)}{(y-i)n!} + \tag{7.100}$$

$$+ \omega_n(y)F(y;1;2;...;n).$$

де $\omega_n(y) = (y-1)(y-2)...(y-n)$.

Знайдемо h, враховуючи (7.98): $b-a=c+(n+m)h-c(1-m)h$, звідки:

$$h=\frac{b-a}{n-1+2m}. \tag{7.101}$$

Зауважимо, що $f(i)=\mathbf{f}(x_i^{(n)})$. Якщо підставити (7.100) та (7.101) у (7.99), то маємо формулу (7.93).

Означення 2. Формула (7.93) називається ІКФ Ньютона-Котеса.

Наслідок 1. Коефіцієнти, що рівновіддалені від кінців на проміжку інтегрування, рівні між собою: $K_{i,m}^{(n)}=K_{n-i+1,m}^{(n)}$.

Доведення. За означенням (7.95) матимемо:

$$K_{n-i+1,m}^{(n)}=\frac{(-1)^{i-1}(b-a)}{(n-1+2m)(n-i)!(i-1)!}\int_{1-m}^{m=n}\frac{\omega_n(y)}{y-(n+i-1)}dy. \tag{7.102}$$

Зробимо заміну $y=n-z+1$. Тоді $y-i=n-z+1-i$, звідки знайдемо:

$$y-1=n-z+1-1=n-z;\quad y-2=n-z-1;\quad \ldots;\quad y-n=-z+1.$$

Отже,

$$\begin{aligned}K_{n-i+1,m}^{(n)}&=\frac{(-1)^{i}(b-a)}{(n-1+2m)(n-i)!(i-1)!}\int_{1-m}^{m=n}\frac{(n-z)\ldots(1-z)}{i-z}dz=\\&=\frac{(-1)^{n-i}(b-a)}{(n-1+2m)(n-i)!(i-1)!}\int_{1-m}^{m=n}\frac{(z-1)\ldots(z-n)}{z-i}dz=K_{i,m}^{(n)}.\end{aligned} \tag{7.103}$$

Наслідок 2(Теорема Кузьміна Р.О.). Границя для модулів непарних вузлів квадратурної формули закритого типу дорівнює нескінченності:

$$\lim_{m\to\infty}\left|K_{i,0}^{(2m+1)}\right|=\infty. \tag{7.104}$$

Наслідок 3.Сума коефіцієнтів квадратурної формули дорівнює проміжку інтегрування: $\sum_{i=1}^{n}K_{i,m}^{(n)}=b-a$.

Доведення. Якщо взяти $f(x)=1$, тоді $\sum_{i=1}^{n}K_{i,m}^{(n)}=\int_a^b f(x)dx=b-a$.

Зауваження 1. Для $I_{i,m}^{(n)}=\frac{K_{i,m}^{(n)}}{b-a}$ складено таблиці, що обслуговують функції $\{f(x)\}\subset\mathbf{R}^1, x\in[a,b]$, для яких існує визначений інтервал (7.95).

Слід зауважити, що $I_{i,m}^{(n)}$ не залежить від відрізка інтегрування $[a,b]$ та підінтегральної функції $f:[a,b]\to\mathbf{R}^1$. Для них наведено інші таблиці.

7.11.4. Частинні випадки інтерполяційних формул Ньютона-Котеса

1. Формула прямокутників (n=1; m=1) – ІКФ відкритого типу

Обчислимо $K_{1,1}^{(1)} = (b-1)I_{1,1}^{(1)} = b-a$ Вузол $x_1^{(1)}$ обчислюємо з врахуванням $h = \frac{b-1}{n-1+2m} = \frac{b-a}{2}$. Тод $x_1^{(1)} = a + \frac{b-a}{2} = \frac{b+a}{2}$. Таким чином маємо так звану формулу прямокутників:

$$I = \int_a^b f(x)dx = f(\frac{b+a}{2})(b-a) + R_1(\mathbf{f}). \qquad (7.105)$$

Геометрично це означає, що площа криволінійної трапеції, яка дорівнює визначеному інтегралу I, замінюється площею прямокутника з основою (b-a) і висотою $f(\frac{b+a}{2})$.

Зауваження 2. Використовуються ще формули правих і лівих прямокутників:

$$\int_a^b f(x)dx = (b-1)\mathbf{f}(b) + \overline{R}_1(\mathbf{f}). \qquad (7.106)$$

$$\int_a^b f(x)dx = (b-a)\mathbf{f}(a) + \overline{R}_2(\mathbf{f}). \qquad (7.107)$$

Ці формули не є КФ Ньютона-Котеса і мають гірший порядок точності.

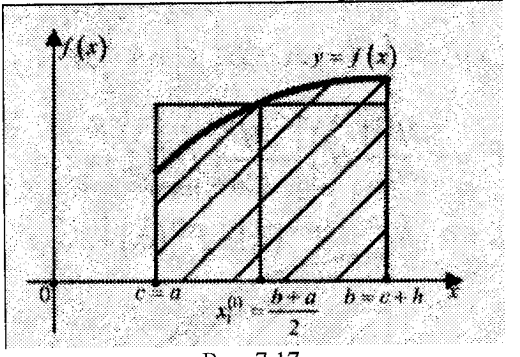


Рис. 7.17.

Зауваження 3. Нехай $f^{(1)} \in C([a,b])$, $f^{(2)} \in ((a,b))$. Тоді за формулою Тейлора з залишковим членом у формі Лагранжа маємо:

$$f(x) = \mathbf{f}(\frac{a+b}{2}) + (x - \frac{a+b}{2})\mathbf{f}^{(1)}(\frac{a+b}{2}) + (\frac{x - \frac{a+b}{2}}{2})\mathbf{f}^{(2)}(\xi). \qquad (7.108)$$

де $\xi \in (a,b)$. Проінтегруємо (7.108) у межах від a до b:

$$\int_a^b f(x)dx = (b-a)\mathbf{f}(\frac{a+b}{2}) + \frac{1}{2}\int_a^b (x-\frac{a+b}{2})^2 \mathbf{f}^{(2)}(\xi)dx. \qquad (7.109)$$

За теоремою про середнє отримаємо:

$$R_1(\mathbf{f}) = \frac{M}{2}\int_a^b (x-\frac{a+b}{2})^2 dx = \frac{(b-1)^3}{24}M, \qquad (7.110)$$

де $M \equiv \sup_{x\in[a,b]} \left|\mathbf{f}^{(2)}(x)\right|$.

Зауваження 4(використання на практиці формул прямокутників). Поділимо відрізок $[a,b]$ на k частин з кроком $H \equiv \frac{b-a}{k}$. Відрізок $[a,b]$ подамо у вигляді:

$$[a,b] = [a,a+H] \cup \ldots \cup [a+(k-1)H, b] = \bigcup_{i=1}^{k} [a+(i-1)H, a+iH]. \qquad (7.111)$$

На кожному проміжку застосовуємо формулу (7.105):

$$\sum_{i=0}^{k} \int_{a+(i-1)H}^{a+iH} f(x)dx = \int_a^b f(x)dx = H(\mathbf{f}(a+\frac{1}{2}H] + \mathbf{f}(a+\frac{3}{2}H) + \ldots + \\ + f(a+\frac{2i-1}{2}H)] + \frac{H^3}{24}[\mathbf{f}^{(2)}(\xi_1) + \mathbf{f}^{(2)}(\xi_2) + \ldots + \mathbf{f}^{(2)}(\xi_k)] = k\cdot\mathbf{f}^{(2)}(\eta). \qquad (7.112)$$

де $a < \eta < b$. Звідси випливає

$$\int_a^b f(x)dx = \frac{b-a}{k}(\mathbf{f}(a+\frac{H}{2}] + \mathbf{f}(a+\frac{3H}{2}) + \ldots + \\ + f(a+\frac{2k-1}{2}H)] + \frac{(b-1)^3}{24k^2} f^{(2)}(\eta), \quad \eta \in (a,b). \qquad (7.113)$$

загальна формула прямокутників.

2. Формула трапецій (n=2 , m=0) – ІКФ закритого типу

Обчислюємо за формулами (7.95) коефіцієнти ІКФ:

$$K_{1,0}^{(2)} = (b-a)I_{1,0}^{(2)} = \frac{b-a}{2};\ K_{2,0}^{(2)} = (b-a)I_{2,0}^{(2)} = \frac{b-a}{2}. \qquad (7.114)$$

Звідки випливає:

$$\int_a^b f(x)dx = \sum_{i=1}^{2} K_{i,0}^{(2)} f(x_i^{(2)}) + R_2(\mathbf{f}) = \frac{b-a}{2}[f(a)+\mathbf{f}(b)] + R_2(\mathbf{f}). \qquad (7.115)$$

$$R_2(\mathbf{f}) = \frac{b-a}{1}\int_1^2 (y-1)(y-2)F(y;1;2)dy = \int_a^b (x-a)(x-b)\mathbf{f}(x,a,b)dx = \\ = \frac{f^{(2)}(\eta)^{(b)}}{2!}\int_a^b (x-a)(x-b)dx = -\frac{(b-a)^3}{12} f^{(2)}(\eta). \qquad (7.116)$$

Отже, *формула трапецій* має вигляд:

$$\int_a^b f(x)dx = \frac{b-a}{2}[f(a)+\mathbf{f}(b)] + R_2(\mathbf{f}), \qquad (7.117)$$

$$R_2(\mathbf{f}) = -\frac{(b-a)^3}{12} f^{(2)}(\eta). \qquad (7.118)$$

Геометричний зміст поданий на рис. 7.18.

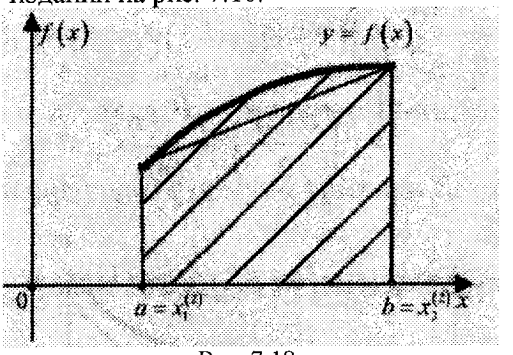


Рис. 7.18.

Зауваження 5. Похибка може бути дуже великою. Тому $[a.b]$ розіб'ємо на k частин: $H = \frac{b-a}{k}$ і $y_0, y_1, ..., y_k$ такі, що $y_0 = f(a)$ та

$$y_1 = f(a+H); ...; y_n = f(b), \qquad (7.119)$$

де $b=a+kH$. Тоді

$$\int_a^b f(x)dx = \frac{b-a}{2k}[y_0 + 2y_1 + 2y_2 + ... + 2y_{k-1} + y_k] - \\ - \frac{(b-a)^3}{12k^2}[f^{(2)}(\xi_1) + f^{(2)}(\xi_2) + ... + f^{(2)}(\xi_k)], \qquad (7.120)$$

де $k \cdot f^{(2)}(\xi) = [f^{(2)}(\xi_1) + f^{(2)}(\xi_2) + ... + f^{(2)}(\xi_k)]$ і $x_{i-1} \le \xi_i \le x_i$, $\xi \in (a,b)$.

Звідки маємо, так звану, *узагальнену формулу трапецій*

$$\int_a^b f(x)dx = \frac{b-a}{2k}\left[y_0 + y_k + 2\sum_{i=1}^{k-1} y_i \right] + R_2(\mathbf{f}). \qquad (7.121)$$

де $R_2(\mathbf{f}) = -\frac{(b-a)^3}{12k} f^{(2)}(\xi),\ \xi \in (a,b)$.

3. Формула Сімпсона (формула парабол, n=3; m=0) – ІКФ закритого типу

Сімпсон Томас (1710-1761) – англійський математик. Праці з елементарної математики, тригонометрії, аналізу й теорії ймовірності.

За формулами (7.93)-(7.98) маємо:

$$\int_a^b f(x)dx = \frac{b-a}{6}\left[f(a) + 4\mathbf{f}(\frac{a+b}{2}) + \mathbf{f}(b) \right] + R_3(\mathbf{f}), \qquad (7.122)$$

$$R_3(\mathbf{f}) = \frac{b-a}{3-1+2\cdot 0}\int_1^2 (y-1)(y-2)(y-3)F(y;1;2;3)dy = -\left(\frac{b-a}{2}\right)^5 \frac{f^{(4)}(\eta)}{90}, \qquad (7.123)$$

де $\eta \in (a,b)$.

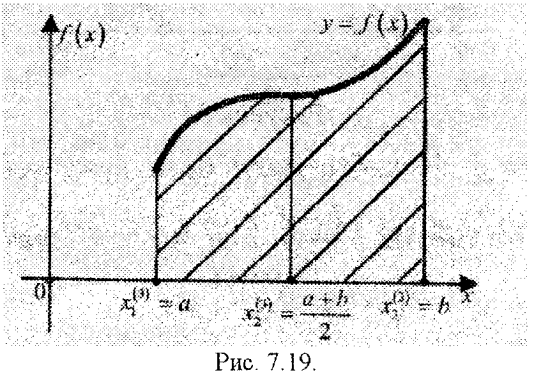


Рис. 7.19.

Зауваження 6 ***(геометричний зміст).*** Так як $x_1^{(3)} = 0$, $x_2^{(3)} = \frac{a+b}{2}$, $x_3^{(3)} = b$, то інтегральний многочлен Лагранжа має другу степінь. Проводимо $L_2(x)$ через три точки кривої.

Зауваження 7 ***(узагальнена формула трапеції).*** Розіб'ємо відрізок $[a,b]$ на $2m$ частин , тобто $H \equiv \frac{d-c}{2m}$

$$\int_a^b f(x)dx = \sum_{i=1}^{2m} \int_{a+(i-1)}^{a+iH} f(x)dx = \frac{b-a}{12m}[y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \ldots + \\ + 2y_{2m-2} + 4y_{2m-1} + 2y_{2m}] - \left(\frac{b-a}{2}\right)^5 \frac{f^{(2)}(\eta)}{90m}. \qquad (7.124)$$

Формулу (7.124) називають узагальненою формулою Сімпсона.

Приклад 7.2. Записати таблицю для $I_{i,m}^{(n)}$, $m = 0$, $2 \le n \le 5$ (ІКФ закритого типу).

Таблиця 1. Умова задачі 7.2 (m=0, ІКФ закритого типу)

№	m	n	I	№	m	n	I
1	0	1	$I_{1,0}^{(1)}$	3	0	3	$I_{1,0}^{(3)}$, $I_{2,0}^{(3)}$, $I_{3,0}^{(3)}$
2	0	2	$I_{1,0}^{(2)}$, $I_{2,0}^{(2)}$	4	0	4	$I_{1,0}^{(4)}$, $I_{2,0}^{(4)}$, $I_{3,0}^{(4)}$, $I_{4,0}^{(4)}$

Аналогічно у випадку, коли n=k, буде $I_{1,0}^{(k)}$, $I_{2,0}^{(k)}$, $I_{3,0}^{(k)}$, ..., $I_{k,0}^{(k)}$ при $k \le n$.

Розв'язок. Якщо $m = 0$, $n = 0$, то існує формула закритого типу. Якщо $n = 1$, то не існує формули Ньютона-Котеса. Якщо $n = 2$, $m = 0$, i=1, то

$$I_{1,0}^{(2)} = \frac{(-1)^{2-1}}{0!1!(2-1)} \int_1^2 \frac{(y-1)(y-2)}{y-1} dy = (-1)^1 \frac{(y-2)^2}{2}\Big|_1^2 = \frac{1}{2}.$$

Якщо $n = 2$, $m = 0$, i=2, то

$$I_{2,0}^{(2)} = \frac{(-1)^{2-2}}{1!0!(2-1)} \int_1^2 \frac{(y-1)(y-2)}{y-2} dy = \frac{(y-1)^2}{2}\Big|_1^2 = \frac{1}{2}.$$

Можна було б використати симетричність коефіцієнтів: $I_{i,k}^{(n)} = I_{n-i+1,k}^{(n)}$. Отже, $I_{2,0}^{(2)} = I_{2-2+1,0}^{(2)} = \frac{1}{2}$. Якщо $n = 3$, $m = 0$, i=1, то

$$I_{1,0}^{(3)} = \frac{(-1)^{3-1}}{0!2!(2-1+2\cdot 0)} \int_1^3 \frac{(y-1)(y-2)(y-3)}{y-1} dy = \frac{1}{4}\int_1^3 (y^2 - 5y + 6)dy =$$

$$= \frac{1}{4}\left[\frac{y^3}{3}\Big|_1^3 - 5\frac{y^2}{2}\Big|_1^3 + 6y\Big|_1^3\right] = \frac{1}{4}\left[9 - \frac{1}{3} - 5\left(\frac{9}{2} - \frac{1}{2}\right) + 12\right] = \frac{1}{4}\left[1 - \frac{1}{3}\right] = \frac{1}{6}.$$

Для i = 2 маємо:

$$I_{2,0}^{(3)} = \frac{(-1)^{3-2}}{1!1!(3-1+2\cdot 0)} \int_1^3 (y-1)(y-3)dy = \frac{1}{2}\int_1^3 (y^2 - 4y + 3)dy =$$

$$= \frac{1}{2}\left[\frac{y^3}{3}\Big|_1^3 - 2y^2\Big|_1^3 + 3y\Big|_1^3\right] = \frac{1}{2}\left[9 - \frac{1}{3} - 4\left(\frac{9}{2} - \frac{1}{2}\right) + 3(3-1)\right] = \frac{4}{6}.$$

Для $i = 3$ маємо: $I_{3,0}^{(3)} = I_{1,0}^{(3)} = \frac{1}{6}$. Якщо $n = 4$, $m = 0$, $i = \overline{1,4}$, то $I_{1,0}^{(4)} = \frac{1}{8}$, $I_{2,0}^{(4)} = \frac{3}{8}$, $I_{3,0}^{(4)} = \frac{3}{8}$, $I_{4,0}^{(4)} = \frac{1}{8}$. Самостійно обчислити коефіцієнти для $I_{i,0}^{(5)}$, $I_{i,0}^{(6)}$, n=5,6.

Результат оформити у вигляді таблиці 2.

Таблиця 2 (m=0, ІКФ закритого типу)

n\I	1	2	3	4	5	6
1	-	-	-	-	-	-
2	1/2	1/2	-	-	-	-
3	1/6	4/6	1/6	-	-	-
4	1/8	3/8	3/8	1/8	-	-
4	7/90	32/90	12/90	32/90	7/90	-
6	19/288	75/288	50/288	50/288	75/277	19/288

Приклад 7.3. Скласти таблицю: $I_{i,m}^{(n)}$, $m = 1$, $2 \le n \le 5$ (ІКФ відкритого типу).

Таблиця 3 Умова задачі 7.3 (m=0, ІКФ закритого типу)

№	m	n	I	№	m	n	I
1	1	1	$I_{1,1}^{(1)}$	3	1	3	$I_{1,1}^{(3)}$, $I_{2,1}^{(3)}$, $I_{3,1}^{(3)}$
2	1	2	$I_{1,1}^{(2)}$, $I_{2,1}^{(2)}$	4	1	4	$I_{1,1}^{(4)}$, $I_{2,1}^{(4)}$, $I_{3,1}^{(4)}$, $I_{4,1}^{(4)}$

Аналогічно у випадку, коли n=k, буде $I_{1,0}^{(k)}$, $I_{2,0}^{(k)}$, $I_{3,0}^{(k)}$, …, $I_{k,0}^{(k)}$ при $k \le n$.

Таблиця 3 (m=1, ІКФ відкритого типу)

n\I	1	2	3	4	5	6
1	1	-	-	-	-	-
2	1/2	1/2	-	-	-	-
3	2/3	-1/3	2/3	-	-	-
4	11/24	1/24	1/24	11/24	-	-
4	11/20	-14/20	26/20	-14/20	11/20	-
6	611/1440	-453/1440	562/1440	562/1440	-453/1440	611/1440

Розв'язок. Коефіцієнти формули Ньютона-Котеса відкритого типу m=1:

$n = 1$, $i = 1$: $I_{1,1}^{(1)} = \frac{(-1)^0}{(1-1+2)} \int_0^2 dy = \frac{1}{2} y \Big|_0^2 = 1$.

$n = 2$, $i = 1$:

$$I_{1,1}^{(2)} = \frac{(-1)^1}{(2-1+2)} \int_0^3 \frac{(y-1)(y-2)}{y-1} dy = -\frac{1}{3} \frac{(y-2)^2}{2} \Big|_0^3 = -\frac{1}{3}\left[\frac{1}{2} - \frac{4}{3}\right] = -\frac{1}{3}\left(-\frac{3}{2}\right) = \frac{1}{2}.$$

$$n = 2,\ i = 2\text{: } I_{2,1}^{(2)} = \frac{(-1)^0}{0!1!3!} \int_0^3 (y-1) dy = \frac{1}{3} \frac{(y-1)^2}{2} \Big|_0^3 = \frac{1}{3}\left(2 - \frac{1}{2}\right) = \frac{1}{2}.$$

$n = 3$, $i = \overline{1,3}$: $I_{1,1}^{(3)} = \frac{2}{3}$, $I_{2,1}^{(3)} = -\frac{1}{3}$, $I_{3,1}^{(3)} = I_{3-3+1,1}^{(3)} = I_{1,1}^{(3)} = \frac{2}{3}$.

$n = 4$, $i = \overline{1,4}$: $I_{1,1}^{(4)} = \frac{11}{24}$, $I_{2,1}^{(4)} = -\frac{1}{24}$, $I_{3,1}^{(4)} = I_{4-3+1,1}^{(4)} = I_{2,1}^{(4)} = \frac{1}{24}$, $I_{4,1}^{(4)} = I_{1,1}^{(4)} = \frac{11}{24}$.

7.12. Квадратурні формули Гауса найкращого степеня точності

Карл Фрідіх Гаус (30.04.1777 – 23.02.1855) – німецький математик. Дисертацію захистив у 22 роки, яка містила перше доведення, так званої, основної теореми алгебри. Наукові праці видано в 11 томах.

Квадратурні формули Ньютона-Котеса є точками для многочленів степеня n-1, якщо n – парне число; є точками для многочленів степеня n, якщо n – непарне (випадок рівновіддалених пузлів). Розглянемо квадратурні формули найвищого степеня точності.

7.12.1. Ідея побудови квадратурних формул Гауса

В формулах Ньютона-Котеса вузли були рівновіддаленими: $x_i^{(n)} = c + ih$, $i = \overline{1,n}$, $h > 0$. Виявляється, що є можливість певним чином вибирати вузли, для яких квадратурні формули будуть точними для многочленів максимального степеня. Побудуємо

$$\int_a^b f(x)dx = \sum_{i=1}^{n} K_i^{(n)} f(x_i^{(n)}) + R_n(f). \tag{7.125}$$

Будемо вимагати, щоб $R_n(f) = 0$ для будь-якого многочлена m-го степеня:

$$P_m(x) = \sum_{j=0}^{m} d_j x^j, \quad d_j \in \mathbf{R}^1, \quad j = \overline{0,m}. \tag{7.126}$$

Якщо підставити (7.126) у (7.125) і скористатися властивостями інтегралу та суми, то маємо:

$$\sum_{j=0}^{m} d_j \int_a^b x^j dx = \sum_{i=1}^{n} K_i^{(n)} \sum_{j=0}^{m} d_j x_i^j = \sum_{j=0}^{m} d_j \sum_{i=1}^{n} K_i^{(n)} \left(x_i^{(n)}\right)^j. \tag{7.127}$$

Позначимо:

$$\mu_k \equiv \int_a^b x^k dx, \; k = \overline{0,m}. \tag{7.128}$$

Завдяки тому, що $d_j \in \mathbf{R}^1$ є довільним, то з рівняння (7.127) маємо нелінійну систему (прирівнюючи коефіцієнти при однакових d_j зліва і справа):

$$\begin{cases} K_1^{(n)} + K_2^{(n)} + \ldots + K_n^{(n)} = \mu_0; \\ K_1^{(n)} x_1^{(n)} + K_2^{(n)} x_2^{(n)} + \ldots + K_n^{(n)} x_n^{(n)} = \mu_1; \\ \ldots \\ K_1^{(n)} (x_1^{(n)})^m + K_2^{(n)} (x_2^{(n)})^m + \ldots + K_n^{(n)} (x_n^{(n)})^m = \mu_m. \end{cases} \tag{7.129}$$

Отримаємо систему (m+1) нелінійних рівнянь для знаходження $2n$ невідомих $\{K_i^{(n)}\}$ і $\{x_i^{(n)}\}$. Для єдиності розвязку необхідно, щоб m+1=2n, тобто m=2n-1. Система (7.129) є нелінійною системою, отже її дослідження досить складне. Проблема єдиності й існування розв'язків, знаходження дійсних розв'язків (7.129) надзвичайно трудомістка. Тому постає питання чи немає іншого, простішого шляху побудуви КФ.

7.12.2. Метод Гауса знаходження вузлів квадратурних формул

Надалі будемо шукати вузли $\{x_i^{(n)}, i = \overline{1,n}\} \subset [a,b]$, як нулі многочлена :

$$\omega_n(x) \equiv (x - x_1^{(n)})(x - x_2^{(n)})\ldots(x - x_n^{(n)}), \tag{7.130}$$

де $x_i^{(n)}$ — шукані вузли.

Виявляється, що $\omega_n(x)$ задовільняє необхідну й достатню умову, яка дозволяє $\omega_n(x)$ визначити явно.

Теорема 2 (теорема Гауса). Для того, щоб квадратурна формула (7.125) була точною для будь-якого многочлена степеня m-2n-1 необхідно й достатньо, щоб її вузли $\{x_1^{(n)}, x_2^{(n)}, \ldots, x_n^{(n)}\} \subseteq [a,b]$ були коренями многочлена (7.130), ортогонального на $[a,b]$ до будь-якого многочлена $g_{n-1}(x)$ степеня не більше n-1, тобто

$$\int_a^b \omega_n(x) g_{n-1}(x) dx = 0. \tag{7.131}$$

Символічно записується так:

$$\int_a^b P_m(x) dx = \sum_{i=1}^{n} K_i^{(n)} P_m(x_i^{(n)}) \tag{7.132}$$

тоді і тільки тоді, коли $\{x_i^{(n)}, i = \overline{1,n}\} \subset [a,b]$ є коренями многочлена $\omega_n(x)$ і виконується (7.131).

Доведення. Необхідність. Нехай виконується (7.132) для довільного многочлена степеня m=2n-1, тобто $R_n(P_m(x)) = 0$. Покажемо, що виконується (7.131).

За умови теореми $\omega_n(x)g_{n-1}(x)$ є многочлен степеня не більше $2n$-1, тому формула (7.125) є точною (має місце (7.132)) і для цього многочлена $\int_a^b \omega_n(x)g_{n-1}(x)dx = \sum_{i=1}^{n} K_i^{(n)}\omega_n(x_i^{(n)})g(x_i^{(n)}) = 0$, бо $\omega_n(x_i^{(n)}) = 0$.

Достатність. Нехай маємо (7.131). Доведемо (7.132) для довільного многочлена $P(x)$ степеня не більше $2n$-1. Для цього розділимо $P(x)$ на $\omega_n(x)$:

$$P(x) = \omega_n(x)g_{n-1}(x) + r(x), \tag{7.133}$$

де $g_{n-1}(x)$, $r(x)$ – степеня не більше n-1. Тепер проінтегруємо (7.133):

$$\int_a^b P(x)dx = \int_a^b \omega_n(x)g_{n-1}(x)dx + \int_a^b r(x)dx = \int_a^b r(x)dx. \tag{7.134}$$

Оскільки $r(x)$ – многочлен степеня не більше $2n$-1, то $r(x)$ дорівнює інтерполяційному многочлену, що побудований за вузлами $\{x_1^{(n)}, x_2^{(n)}, ..., x_n^{(n)}\}$, квадратурної формули Ньютона-Котеса (теорема 1), для якої $R_n(r) = \frac{r^{(n)}(\xi)}{n!}\omega_n(x) = 0$, де $\xi \in (a,b)$. Таким чином, маємо точну КФ для $r(x)$:

$$\int_a^b r(x)dx = \sum_{i=1}^{n} K_i^{(n)} r(x_i^{(n)}). \tag{7.135}$$

Інакше:

$$P(x_i^{(n)}) = \omega_n(x_i^{(n)})g_{n-1}(x_i^{(n)}) + r(x_i^{(n)}). \tag{7.136}$$

Звідки випливає $P(x_i^{(n)}) = r(x_i^{(n)})$, бо $\omega_n(x_i^{(n)}) = 0$. Таким чином

$$\int_a^b r(x)dx = \sum_{i=1}^{n} K_i^{(n)} P_m(x_i^{(n)}). \tag{7.137}$$

З (7.137) і (7.134) маємо $\int_a^b P_m(x)dx = \sum_{i=1}^{n} K_i^{(n)} P_m(x_i^{(n)}) + 0$, що означає $R_n(P_m) = 0$.

Теорема 3. Нехай виконано умови теореми 2. Тоді вузли квадратурної формули $\{x_1^{(n)}, x_2^{(n)}, ..., x_n^{(n)}\} \subseteq [a,b]$ є нулями многочлена

$$\omega_n(x) = \frac{n!}{(2n!)!}\frac{d!}{dx^n}\left[(x-a)^n(x-b)^n\right]. \tag{7.138}$$

Доведення. Позначимо

$$\int_a^b \omega_n(x)dx \equiv \varphi_1(x); \int_a^x \varphi_1(x)dx \equiv \varphi_2(x); \ldots; \int_a^x \varphi_{n-1}(x)dx \equiv \varphi_n(x). \qquad (7.139)$$

За теоремою 2 маємо: $\int_a^b \omega_n(x)g(x)dx = 0$. Інакше, інтегруванням частинами одержимо:

$$\int_a^b \omega_n(x)g(x)dx = \varphi_1(b)g(b) - \varphi_1(a)g(a) - \int_a^b \varphi_1(x)g^{(1)}(x)dx = \ldots =$$

$$= \left[\varphi_1(x)g_{n-1}^{(0)}(x) - \varphi_2(x)g_{n-1}^{(1)}(x) + \ldots + (-1)^{n-1}\varphi_n(x)g_{n-1}^{(n-1)}(x)\right]_a^b = 0,$$

де використали (7.132).

Звідки за умови (7.139) випливає, що $\varphi_1(a) = \varphi_2(a) = \ldots = \varphi_n(a) = 0$. Тоді

$$\varphi_1(b)g_{n-1}^{(0)}(b) - \varphi_2(b)g_{n-1}^{(1)}(b) + \ldots + (-1)^{n-1}\varphi_n(b)g_{n-1}^{(n-1)}(b) = 0. \qquad (7.140)$$

Якщо $\{g_{n-1}(x)\}$ є будь-яким многочленом степеня не більше n-1 для $\forall x \in [a,b]$, то (7.140) виконується тоді і тільки тоді, коли

$$\varphi_1(b) = \varphi_2(b) = \ldots = \varphi_n(b) = 0. \qquad (7.141)$$

Звідки випливає, що $\varphi_n(x)$ має n-кратні корені a і b. Тоді за формулою Вієта:

$$\varphi_n(x) = A(x-a)^n(x-b)^n. \qquad (7.142)$$

З (7.139) випливає рівність $\omega_n(x) = \frac{d^n}{dx^n}\varphi_n(x)$, а з (7.142) маємо: $\omega_n(x) = A\frac{d^n}{dx^n}\left[(x-a)^n(x-b)^n\right]$. Сталу слід підбирати за умови, що коефіцієнт при x^n рівний $\frac{(2n)!}{n!}$, тобто A=1, звідки $A = \frac{(2n)!}{n!}$.

7.12.3. Метод Гауса вибору коефіцієнтів

Теорема 4 (про вибір коефіцієнтів квадратурної формули Гауса). Нехай виконуються умови теореми 2. Тоді коефіцієнти КФ обчислюються за формулою

$$K_k^{(n)} \equiv \frac{(n!)^4(b-a)^{2n+1}}{\left[(2n)!\right]^2(x_k^{(n)}-a)(b-x_k^{(n)})\left[\omega_n'(x_k^{(n)})\right]^2}. \qquad (7.143)$$

Доведення. Розглянемо

$$\psi_{n,k}(x) \equiv \frac{\omega_n(x)}{x - x_k^{(n)}}, \tag{7.144}$$

де $\psi_{n,k}(x)$ є многочленом степеня n-1. Отже, $\psi_{n,k}^2(x)$ є многочленом степеня $2n$-2. Тому формула (7.125) за теоремою 2 є справедливою і для $\psi_{n,k}^2(x)$.

Так як степінь $\psi_{n,k}^2(x)$ рівний $2n$-2 ((7.125) справедливий для многочленів степеня менше ніж $2n$), тоді

$$\int_a^b \psi_{n,k}^2(x)dx = \sum_{i=1}^{n} K_l^{(n)} \psi_{n,k}^2(x_l^{(n)}) = \sum_{i=1}^{n} K_l^{(n)} \frac{\omega_n^2(x_l^{(n)})}{(x_l^{(n)} - x_k^{(n)})^2} = K_k^{(n)} \left[\omega_n^{!}(x_k^{(n)})\right]^2,$$

бо $\omega_n(x_l^{(n)}) = 0$ при $l \neq k$.

Після елементарних перетворень одержимо:

$$K_k^{(n)} = \frac{\int_a^b \psi_{n,k}^2(x)dx}{\left[\omega_n^{!}(x_k^{(n)})\right]^2}. \tag{7.145}$$

Можна отримати точнішу формулу для обчислення $K_k^{(n)}$ порівняно з (7.145). Обчислимо чисельник інтегруванням частинами

$$\int_a^b \psi_{n,k}^2(x)dx = \int_a^b \frac{\omega_n^2(x)dx}{(x - x_k)^2} = -\frac{\omega_n^2(x)}{x - x_k}\Bigg|_{x=a}^{x=b} + 2\int_a^b \frac{\omega_n(x)\omega_n^{!}(x)dx}{x - x_k} =$$

$$= \frac{\omega_n^2(a)}{a - x_k} - \frac{\omega_n^2(b)}{b - x_k} + 2\int_a^b \frac{\omega_n(x)\omega_n^{!}(x)}{x - x_k}dx.$$

Обчислимо значення $\omega_n^2(a)$ та $\omega_n^2(b)$. За формулою Лейбніца

$$\omega_n(x) = \frac{n!}{(2n)!}\frac{d^n}{dx^n}\left[(x-a)^n(x-b)^n\right] = \frac{n!}{(2n)!}\sum_{l=0}^{n} C_n^l \frac{d^n}{dx^n}\left[(x-a)^n\right]\frac{d^{n-l}}{dx^{n-l}}\left[(x-b)^n\right].$$

Звідки відразу отримаємо

$$\omega_n(a) = \frac{n!}{(2n)!}n!(a-b)^n = (-1)^n\frac{(n!)^2}{(2n)!}(b-a)^n,$$

$$\omega_n(b) = \frac{n!}{(2n)!}n!(b-a)^n = \frac{(n!)^2}{(2n)!}(b-a)^n.$$

Таким чином,

$$\frac{\omega_n^2(a)}{a - x_k} - \frac{\omega_n^2(b)}{b - x_k} = \frac{(n!)^4(b-a)^{2n}}{\left[(2n)!\right]^2}\left[\frac{1}{x_k - b} - \frac{1}{x_k - a}\right] = \frac{(n!)^4(b-a)^{2n-1}}{(x_k - a)(x_k - b)\left[(2n)!\right]^2}.$$

Далі слід зауважити, що $\frac{\omega_n(x)\omega_n^!(x)}{x-x_l}$ є многочленом степеня $2n$-2, тому КФ є точною:

$$K_k^{(n)}\left[\omega_n^!(x)\right]^2=\int_a^b \psi_{n,k}^2(x)dx=\frac{(n!)^4(b-a)^{2n-1}}{\left[(2n)!\right]^2(x_k^{(n)}-a)(x_k^{(n)}-b)}+2K_k^{(n)}\left[\omega_n^!(x_k^{(n)})\right]^2$$

Звідки маємо (7.143).

Теорема 5 (про похибку квадратурної формули Гауса). Якщо функція $f\in C^{2n}([a,b])$, то похибка обчислюється за формулою:

$$R_n(f)\equiv\frac{(n!)^4(b-a)^{2n+1}}{\left[(2n)!\right]^2(2n+1)}f^{(2n)}(\xi), \tag{7.146}$$

де $\xi\in[\xi_1,\xi_2]$, $\xi_1=\inf\left\{x_i^{(n)},i=\overline{1,n}\right\}$, $\xi_2=\sup\left\{x_i^{(n)},i=\overline{1,n}\right\}$.

Наслідок 4. Корені рівняння $\omega_n(x)=0$ симетричні відносно $x=\frac{a+b}{2}$, тобто $x=\frac{a+b}{2}x_l^{(n)}=x_{n-l+1}$, $l=\overline{1,n}$.

Доведення. У многочлені $\varphi_n(x)$ проведемо заміну $x=\frac{a+b}{2}+y$. Одержимо $\varphi_n(y)=\frac{n!}{(2n)!}\left[y^2-\left(\frac{b-a}{2}\right)^2\right]^n$. А це означає, що $\varphi_n(x)$ симетричний відносно $x=\frac{a+b}{2}$, тобто корені $\omega_n(x)=0$, симетричні відносно цієї точки.

Наслідок 5. Коефіцієнти КФ Гауса симетричні відносно середини $[a,b]$, тобто $K_l^{(n)}=K_{n-l+1}^{(n)}$, $l=\overline{1,n}$.

Доведення. Пронумеруємо корені в порядку зростання: $x_1^{(n)}<x_2^{(n)}<\ldots<x_n^{(n)}$. Тоді $(x_k^{(n)}-a)(x_k^{(n)}-b)(\omega_n^!(x_k^{(n)}))^2=(x_{n-k+1}^{(n)}-a)(x_{n-k+1}^{(n)}-b)(\omega_n^!(x_{n-k+1}^{(n)}))^2$, тобто одержимо твердження наслідку 5: коефіцієнти для $f(x_k^{(n)})$ і $f(x_{n-k+1}^{(n)})$ будуть співпадати.

7.12.4. Застосування формули Гауса

Одержані КФ вперше винайшов Гаус, тому й названо їх *квадратурними формулами Гауса.*

$$\int_a^b f(x)dx = \sum_{i=1}^{n} Kf(x) + R(\mathbf{f}). \qquad (7.147)$$

Заміна $x = \frac{a+b}{2} + \frac{b-a}{2}t$, переведе $[a,b]$ у $[-1,1]$, тоді квадратурна формула матиме вигляд:

$$\int_{-1}^{1} f(x)dx = \sum_{i=1}^{n} Kf(x) + \frac{2(n!)}{[(2n!)](2n+1)} f(\xi). \qquad (7.148)$$

де $K = 2K_i^{(n)}$.

Приклад 7.4. Побудувати квадратурні формули для різних n.

Розв'язок. 1. Якщо n=1, то $x_1^{(1)} = 0$; $K_1^{(1)} = 2$; $R_1 = \frac{1}{3} f^{(2)}(\xi)$, а це означає, що квадратурна формула Гауса для одного вузла має вигляд: $\int_{-1}^{1} f(x)dx = 2f(0) + R_1(\mathbf{f})$.

2. Якщо n=2, то $-x_1^{(2)} = -x_2^{(2)} = 0.5773502691896258$; $K_1^{(2)} = K_2^{(2)} = \frac{1}{2}$; $R_2 = \frac{1}{135} f^{(4)}(\xi)$, а відповідна КФ: $\int_{-1}^{1} f(x)dx = 2\left[K_1^{(2)} f(x_1^{(2)}) + K_2^{(2)} f(x_2^{(2)})\right] + R_2(\mathbf{f})$.

3. Якщо n=3, то $-x_1^{(3)} = x_2^{(3)} = 0.7745966692414834$; $x_3^{(3)} = 0$; $K_1^{(3)} = K_3^{(3)} = \frac{5}{9}$; $K_2^{(3)} = \frac{8}{9}$; $R_3 = \frac{1}{15750} f^{(6)}(\xi)$ та відповідна КФ:

$$\int_{-1}^{1} f(x)dx = 2\left[K_1^{(3)} f(x_1^{(3)}) + K_2^{(3)} f(x_2^{(3)}) + K_3^{(3)} f(x_3^{(3)})\right] + R_3(\mathbf{f}).$$

4. Якщо $n = 4$, то $R_4 = \frac{1}{3472875} f^{(8)}(\xi)$ та

$x_1^{(4)} = x_4^{(4)} = 0.8611363115940492$; $-x_2^{(4)} = x_3^{(4)} = 0.3399810435848646$;

$K_1^{(4)} = K_4^{(4)} = 0.1739274225687284$; $K_2^{(4)} = K_3^{(4)} = 0.3260725774312716$.

Далі для n => 5 можна записати КФ Гауса, як це записано для n=1,2,3.

5. Якщо n=5, то $R_5 = \frac{1}{1237732650} f^{(10)}(\xi)$, $x_3^{(5)} = 0$, $K_3^{(5)} = \frac{64}{225}$ та

$x_1^{(5)} = x_5^{(5)} = 0.9061798459386640$; $-x_2^{(5)} = x_4^{(5)} = 0.5384693101056830$;

$K_1^{(5)} = K_5^{(5)} = 0.1184634425280945$; $K_2^{(5)} = K_4^{(5)} = 0.2393143352496832$.

6. Якщо $n = 6$, то $R_6 = \frac{1}{648984486159} f^{(12)}(\xi)$ та

$$x_1^{(6)} = x_6^{(6)} = 0.932465142031520\,;\ -x_2^{(6)} = x_5^{(6)} = 0.6612093864662644\,;$$
$$x_3^{(6)} = x_4^{(6)} = 0.2386191860831970\,;\ K_1^{(6)} = K_6^{(6)} = 0.0856622461895852\,;$$
$$K_2^{(6)} = K_5^{(6)} = 0.1803807865240693\,;\ K_3^{(6)} = K_4^{(6)} = 0.2339569672873455\,.$$

Можна довести побудову КФ Гауса. Зауважимо, що вузли та коефіцієнти належать довільній $f(x)$.

7.13. Квадратурні формули чисельного інтегрування Чебишова

7.13.1. Постановка задачі

Нехай значення $\{f(x_i),\ i=\overline{1,n}\}$ знаходиться під впливом випадкових похибок (наприклад, одержуємо їх з експерименту). Тоді вираз $\sum_{i=1}^{n} K_i f(x_i^{(n)})$ буде мати найменшу випадкову похибку для $K_1=K_2=\ \ldots\ =K_n = K$. В зв`язку з цим, П.Л. Чебишов поставив задачу побудувати КФ вигляду

$$\int_{-1}^{1} f(x)dx = K\sum_{i=1}^{n} f(x_i^{(n)}) + R_n(\mathbf{f}), \tag{7.149}$$

тобто знайти $\{x_i^{(n)}, i=\overline{1,n}\}\in[-1;+1]$ та K так, щоб у формулі (7.149) $R(f)=0$, коли $f(x)=P(x)$ довільний многочлен найвищого степеня n:

$$P_n(x) = \sum_{j=0}^{n} d_j x^j\,, \tag{7.150}$$

де $d\in\mathbf{R}$. Дійсно, маємо шуканий K і вузли $\{x_i^{(n)}, i=\overline{1,n}\}\in[-1;+1]$, тобто всього $(n+1)$ невідомих. Тоді степінь $f(x)$ не менший n. Підстановкою (7.150) у (7.149) одержимо

$$\sum_{j=0}^{n} d_j \int_{-1}^{1} x^j dx = K\sum_{j=0}^{n} d_j\left[\sum_{i=1}^{n}\left(x_i^{(n)}\right)^j\right]. \tag{7.151}$$

Позначимо

$$m_j = \int_{-1}^{1} x^j dx\,. \tag{7.152}$$

Для $j=0$: $m_0 = K*n$, а значить

$$K = \frac{m_0}{n}, \text{ а } m_0 = \int_{-1}^{1} dx = 2. \tag{7.153}$$

Прирівнюючи коефіцієнти при невідомих d_j, $j = \overline{1,m}$, будемо мати наступну систему нелінійних рівнянь відносно вузлів квадратичної формули $\{x_i^{(n)}, i = \overline{1,n}\}$ при

$$\mu_i = \frac{m_i}{K}, \; i = \overline{1,n}. \tag{7.154}$$

$$\left\{ \begin{array}{l} \sum\limits_{i=1}^{n} x_i^{(n)} = \mu_1 \\ \sum\limits_{i=1}^{n} \left[x_i^{(n)}\right]^2 = \mu_2 \\ \ldots\ldots\ldots\ldots\ldots \\ \sum\limits_{i=1}^{n} \left[x_i^{(n)}\right]^n = \mu_n \end{array} \right\}. \tag{7.155}$$

Одержали систему нелінійних рівнянь для знаходження вузлів $\{x_i^{(n)}, i = \overline{1,n}\}$. Але ця система відносно $\{x_i^{(n)}, i = \overline{1,n}\}$ розв`язується складно. Запропонуємо інший шлях знаходження вузлів квадратурної формули.

7.13.2. Метод Чебишова знаходження вузлів квадратурних формул

Теорема 6. $\{x_i^{(n)}, i = \overline{1,n}\} \in [-1;+1]$ визначаються як корені многочлена n-ого степеня

$$\omega_n(x) = x^n + b_1 x^{n-1} + b_2 x^{n-2} + \ldots + b_{n-1} x + b_n, \tag{7.156}$$

де b_1 визначають за формулами (7.161), (7.164) (див. нижче).

Доведення. За означенням $\{x_i^{(n)}, i = \overline{1,n}\} \in [-1;+1]$ є дійсними, тоді $\omega_n(x)$ подамо у вигляді добутку двочленів за теоремою Вієта

$$\omega_n(x) = (x - x_1^{(n)})(x - x_2^{(n)})\ldots(x - x_n^{(n)}). \tag{7.157}$$

Знайдемо $b_1, b_2, \ldots, b_{n-1}, b_n$. Похідну $\omega_n(x)$ за (7.156) записуємо у вигляді

$$\omega_n(x) = nx^{n-1} + (n-1)b_1 x^{n-3} + (n-1)b_2 x^{n-4} + \ldots + b_{n-1}. \tag{7.158}$$

Інакше похідна $\omega_n(x)$ з (7.157) має вигляд

$$\omega_n(x)=\sum_{i=1}^{n}\frac{\omega_n(x)}{x-x_i^{(n)}}=\sum_{i=1}[x^{n-1}+(b_1+x_i^{(n)})x^{n-2}+(b_2+b_1x_i^{(n)}+(x_i^{(n)})^2)x^{n-3}+$$
$$+\ldots+(b_{n-1}+b_{n-2}x_i^{(n)}+\ldots+b_1(x_i^{(n)})^{n-2}+(x_i^{(n)})^{n-1})]=nx^{n-1}+(nb_1+m)x^{n-2}+ \quad (7.159)$$
$$+(nb_2+mb_1+m)x^{n-3}+\ldots+(nb_{n-1}+nb_{n-2}m_1+\ldots+b_1m_{n-2}+m_{n-1}).$$

Прирівнюємо відповідні коефіцієнти при x у (7.158) і (7.159). Тоді одержимо систему лінійних алгебраїчних рівнянь для знаходження b_j при $j=\overline{1,n}$:

$$\left\{\begin{array}{l} nb_1+\mu_1=(n-1)b_1, \\ nb_2+b_1\mu_1=(n-2)b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ nb_{n-1}+b_{n-2}\mu_1+\ldots+b_1\mu_{n-2}+\mu_{n-1}=b_{n-1}. \end{array}\right\}. \quad (7.160)$$

Звідки

$$\left\{\begin{array}{l} b_1=-\mu_1, \\ b_2=-\dfrac{\mu_2-\mu_1}{2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ b_{n-1}=b_{n-2}\mu_1+\ldots+b_1\mu_{n-2}+\mu_{n-1}. \end{array}\right\}. \quad (7.161)$$

Знайдемо b_n. Зауважимо, що з вигляду (7.157) маємо

$$\omega_n(x_1^{(n)})+\omega_n(x_2^{(n)})+\ldots+\omega_n(x_n^{(n)})=0. \quad (7.162)$$

Враховуючи вигляд $\omega_n(x)$ за формулою (7.156), маємо з попередньої рівності

$$nb_{n-1}+b_{n-2}\mu_1+\ldots+b_1\mu_{n-2}+b_{n-1}=0. \quad (7.163)$$

Звідки обчислюємо

$$b_n=\frac{\mu_n+b_1\mu_{n-1}+\ldots+b_{n-1}\mu_1}{n}. \quad (7.164)$$

Таким чином, отримуємо так званий многочлен Чебишова, з якого знаходимо вузли квадратурної формули.

7.13.3. Побудова квадратурних формул Чебишова

Зауважимо, що залишковий член КФ Чебишова одержується достатньо складно, тому надалі будемо записувати їх вигляд без доведення.

Приклад 7.5. Побудувати квадратурні формули Чебишова для різних n.

Розв'язок. Для n=1 маємо: x=0 при $\int_{-1}^{1} f(x)dx = K\sum_{i-1}^{n} f(x_i^{(n)}) + R_n(\mathbf{f})$ формула Чебишова має вигляд: $\int_{-1}^{1} f(x)dx = 2f(0) + R_1(\mathbf{f})$, $k = \int_{-1}^{1} dx = 2$, $R_1(\mathbf{f}) = \frac{1}{3} f^{(2)}(\xi)$.

$$n = 2:\ x^2 - \frac{1}{3} = 0,\ -x_1^{(2)} = x_2^{(2)} = 0.577350,\ R_2(\mathbf{f}) = \frac{1}{135} f^{(4)}(\xi),$$

$$\int_{-1}^{1} f(x)dx = 2\left[f(x_1^{(2)}) + f(x_2^{(2)})\right] + \frac{1}{135} f^{(4)}(\xi)$$

$$n = 3:\ x^3 - \frac{1}{2}x = 0,\ x_2^{(3)} = 0,\ -x_1^{(3)} = x_3^{(3)} = 0.707107,\ R_3(\mathbf{f}) = \frac{1}{360} f^{(6)}(\xi),$$

$$\int_{-1}^{1} f(x)dx = 2\left[f(x_1^{(3)}) + f(x_2^{(3)}) + f(x_3^{(3)})\right] + \frac{1}{360} f^{(6)}(\xi).$$

Очевидна за аналогією n=1,2,3 формула Чебишова для n=4,5,6 і 9

$$n = 4:\ x^4 - \frac{2}{3}x^2 + \frac{1}{45} = 0,\ -x_1^{(4)} = x_4^{(4)} = 0.794654,\ -x_2^{(4)} = x_3^{(4)} = 0.187592,$$

$$R_4(\mathbf{f}) = \frac{1}{42525} f^{(8)}(\xi),\ \int_{-1}^{1} f(x)dx = 2\left[f(x_1^{(4)}) + f(x_2^{(4)}) + f(x_3^{(4)}) + f(x_4^{(4)})\right] + \frac{1}{42525} f^{(8)}(\xi).$$

$$n = 5:\ x^5 - \frac{5}{6} = 0,\ -x_1^{(5)} = x_5^{(5)} = 0.832498,\ x_3^{(5)} = 0,\ -x_2^{(5)} = x_4^{(5)} = 0.53846,$$

$$R_5(\mathbf{f}) = \frac{1}{544320} f^{(10)}(\xi),$$

$$\int_{-1}^{1} f(x)dx = 2\left[f(x_1^{(5)}) + f(x_2^{(5)}) + f(x_3^{(5)}) + f(x_4^{(5)}) + f(x_5^{(5)})\right] + \frac{1}{544320} f^{(10)}(\xi).$$

$$n = 6:\ x^6 - \frac{1}{2}x = 0,\ -x_1^{(6)} = x_6^{(6)} = 0.866247,\ x_3^{(6)} = x_4^{(0)} = 0,$$

$$-x_2^{(6)} = x_5^{(6)} = 0.661209,$$

$$\int_{-1}^{1} f(x)dx = 2\left[f(x_1^{(5)}) + f(x_2^{(5)}) + f(x_3^{(5)}) + f(x_4^{(5)}) + f(x_5^{(5)}) + f(x_6^{(6)})\right] + R_6(\mathbf{f}).$$

$$n = 9:\ x^9 - \frac{3}{2}x^7 + \frac{27}{40}x^5 - \frac{57}{560}x^3 + \frac{2217}{22400}x = 0,\ -x_1^{(9)} = x_9^{(9)} = 0.866247911589,$$

$$-x_3^{(9)} = x_5^{(9)} = 0.528762,\ -x_4^{(9)} = x_5^{(9)} = 0.167906,\ -x_2^{(9)} = x_8^{(9)} = 0.601019,$$

$$R_7(\mathbf{f}) = \frac{281}{1959552000} f^{(8)}(\xi),\ R_9(\mathbf{f}) = \frac{74747}{3657830400} f^{(10)}(\xi).$$

Наведемо без доведення важливий результат про дійсні корені многочлена $\omega_n(x) = 0$.

Теорема 7 (теорема С.Н.Бернштейна). Многочлен $\omega_n(x)=0$ для n = 8, n => 10 мають комплексні корені. Це означає що визначати КФ неможливо.

7.14. Збіжність квадратурних формул

За попередніми міркуваннями будують КФ для обчислення визначеного інтеграла

$$I=\int_a^b f(x)dx=I_n(f)+R(\mathbf{f}), \tag{7.165}$$

$$I_n(f)=\sum_{i=1}^{n} K_i^{(n)} f(x), \tag{7.166}$$

а $R(\mathbf{f})$ похибка КФ

$$R(\mathbf{f})=I-I_n(f) \tag{7.167}$$

є похибкою методу чисельного інтегрування.

За якими накладеними умовами на $\{K_i^{(n)}, i=\overline{1,n}\}$, $\{x_i^{(n)}, i=\overline{1,n}\}$ для довільної $f\in C([a,b])$ має місце

$$\lim_{x\to\infty} I(f)=\int_a^b f(x)dx. \tag{7.168}$$

При такому процесі маємо справу зі зліченими трикутними значеннями коефіцієнтів та вузлів квадратурної формули:

$$K=\begin{Bmatrix} K_1^{(1)} & 0 & \dots & 0 \\ K_1^{(2)} & K_2^{(2)} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ K_1^{(n)} & K_2^{(n)} & \dots & K_n^{(n)} \\ \dots & \dots & \dots & \dots \end{Bmatrix} \text{ та } X=\begin{Bmatrix} x_1^{(1)} & 0 & \dots & 0 \\ x_1^{(2)} & x_2^{(2)} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ x_1^{(n)} & x_2^{(n)} & \dots & x_n^{(n)} \\ \dots & \dots & \dots & \dots \end{Bmatrix}.$$

Визначимо, які умови необхідно накласти на K та X, щоб виконувалася рівність (7.167).

Теорема 8 (про необхідні та достатні умови збіжності КФ). Для $f\in C([a,b])$ квадратурна формула (7.167) збігається до інтеграла тоді і тільки тоді:

1) квадратурна формула збігається для будь-якого многочлена $P(x)$ при $n\in N$, тобто

$$\lim_{x\to\infty} I(f)=\int_a^b f(x)dx. \tag{7.169}$$

2) для $\forall n \in N$

$$\sum_{i=1}^{n} \left| K_i^{(n)} \right| \le M \,,\; M > 0 \,. \tag{7.170}$$

Доведення. Достатність базується на наступній теоремі Вейєрштрасе.

Теорема 9. (Теорема Вейєрштрасе).

$$f \in C([a,b]), \varepsilon_1 > 0, \left| P(x) - f(x) \right| < \varepsilon_1, x \in [a,b] \,. \tag{7.171}$$

Доведення. Достатність. Нехай виконується умови 1) з теореми 8. Доведемо умову (7.168), яка означає, що для $\varepsilon > 0$ існує таке $n \in N$, що виконується нерівність

$$\left| \int_a^b f(x)dx - I_n(P) \right| < \varepsilon \,. \tag{7.172}$$

Доведемо цю нерівність. Є зрозумілим ланцюжок нерівностей

$$I - I_n(f) = \int_a^b (f - P)dx + \int_a^b P dx - I_n(P) - I_n(f - P) \,. \tag{7.173}$$

$$\left| \int_a^b f(x)dx - I_n(f) \right| = \left| \int_a^b \left[f(x) - P(x) \right] dx \right| + \left| \int_a^b P(x)dx - I_n(P) \right| + \tag{7.174}$$
$$+ \left| I_n(f - P) \right| \le \varepsilon_1 [b - a] + \varepsilon_2 + \varepsilon M .$$

Дійсно, з (7.171) випливає, що $\forall \varepsilon > 0$ існує $n_0 \in N$ таке, що при $n > n_0$ виконується

$$\left| \int_a^b P(x)dx - I_n(P) \right| < \varepsilon_2 \,, \tag{7.175}$$

а з лінійної суми маємо

$$\left| I_n(f - P) \right| \le \sum_{i=1}^{n} \left| K_i^{(n)} \right| \left| f(x_i^{(n)}) - P(x_i^{(n)}) \right| \le \varepsilon M \,. \tag{7.176}$$

Необхідність. Нехай виконується умова (7.168). Доведемо (7.171) та (7.172). Дійсно, для будь-якої $f \in C([a,b])$ існує $P \in C([a,b])$. Зрозуміло що

$$I_n(\varphi_n) = \sum_{i=1}^{n} \left| K_i^{(n)} \right| \equiv M_n \,. \tag{7.177}$$

Візьмемо деяку з побудованих функцій $\{ \varphi_{n_i}(x) \}$, для якої

$$\lim_{n\to\infty} I_n(\varphi_{n_1}) = \int_a^b \varphi_{n_1}(x)dx\,. \tag{7.178}$$

Але за властивістю 3) функції $\{\varphi_{n_1}(x)\}$ маємо:

$$\left|\int_a^b \varphi_{n_1}(x)dx\right| \le b-a\,. \tag{7.179}$$

Далі знайдеться таке $n_2 > n_1$, що $M > 2!2$. Розглянемо функцію

$$\psi_2(x) \equiv \frac{\varphi_{n_1}(x)}{1!} + \frac{\varphi_{n_2}(x)}{2!}\,, \tag{7.180}$$

яка є неперервною. Отже,

$$\lim_{n\to\infty} I_n\left(\frac{\varphi_{n_1}(x)}{1!} + \frac{\varphi_{n_2}(x)}{2!}\right) = \int_a^b \left[\frac{\varphi_{n_1}(x)}{1!} + \frac{\varphi_{n_2}(x)}{2!}\right]dx\,, \tag{7.181}$$

$$\left|\int_a^b \left[\frac{\varphi_{n_1}(x)}{1!} + \frac{\varphi_{n_2}(x)}{2!}\right]dx\right| \le \left(\frac{1}{1!} + \frac{2}{2!}\right)(b-a) < (e-1)(b-a)\,, \tag{7.182}$$

тому знайдеться таке N, що для $n > N$ матимемо

$$\left|I_n\left(\frac{\varphi_{n_1}(x)}{1!} + \frac{\varphi_{n_2}(x)}{2!}\right)\right| < e(b-a)\,. \tag{7.183}$$

Далі знайдемо таке $n_3 > n_2 > n_1 > 0$, що $M > 3!3$. Продовжуємо: нехай знайдемо $N > 0$. Тоді існує $n > N$ таке, що $M > (m+1)(m+1)!$. Побудуємо функцію

$$\psi_{m+1}(x) \equiv \frac{\varphi_{n_1}(x)}{1!} + \frac{\varphi_{n_2}(x)}{2!} + \ldots + \frac{\varphi_{n_{m+1}}(x)}{(m+1)!}\,, \tag{7.184}$$

яка є неперервною. Отже,

$$\lim_{n\to\infty} I_n(\psi_{m+1}) = \int_a^b \psi_{m+1}(x)dx\,. \tag{7.185}$$

$$\left|\int_a^b \psi_{m+1}(x)dx\right| \le \left(\frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{(m+1)!}\right)(b-a) < (e-1)(b-a)\,. \tag{7.186}$$

Знову знаходимо N таке, що для $n > N$ маємо $|I_n(\psi_{m+1})| < e(b-a)$ і продовжуємо побудову далі. Таким чином, одержимо ряд

$$\frac{\varphi_{n_1}(x)}{1!}+\frac{\varphi_{n_2}(x)}{2!}+\ldots+\frac{\varphi_{n_{m+1}}(x)}{(m+1)!}+\ldots. \tag{7.187}$$

Цей ряд рівномірно збігається, а отже, його сума є неперервною. Позначимо її через $f(x)$. Подамо $f(x)$ у вигляді

$$f(x)=\sum_{i=1}^{k-1}\frac{\varphi_{n_i}(x)}{i!}+\frac{\varphi_{n_k}(x)}{k!}+\sum_{i=k+1}^{\infty}\frac{\varphi_{n_i}(x)}{i!}. \tag{7.189}$$

При цьому КФ, яка побудована для $f(x)$ вигляду (7.189) задовольняє рівність

$$I_{n_k}(f)=I_{n_k}\left(\sum_{i=1}^{k-1}\frac{\varphi_{n_i}(x)}{i!}\right)+I_{n_k}\left(\frac{\varphi_{n_k}(x)}{k!}\right)+I_{n_k}\left(\sum_{i=k+1}^{\infty}\frac{\varphi_{n_i}(x)}{i!}\right). \tag{7.190}$$

Далі оцінимо кожне з доданків у (7.190). Так як $n > N > 0$, то

$$\left|I_{n_k}\left(\sum_{i=1}^{k-1}\frac{\varphi_{n_i}(x)}{i!}\right)\right|<e(b-a). \tag{7.191}$$

Зауважимо, що очевидний ланцюжок нерівностей

$$\left|\sum_{i=k+1}^{\infty}\frac{\varphi_{n_i}(x)}{i!}\right|\le\sum_{i=k+1}^{\infty}\frac{1}{i!}=\frac{1}{(k+1)!}\left(1+\frac{1}{k+2}+\frac{1}{(k+2)(k+3)}+\ldots\right)\le\left|\sum_{i-1}^{k-1}\frac{\varphi_{n_i}(x)}{i!}\right|. \tag{7.192}$$

$$\left|I_{n_k}\left(\sum_{i=k+1}^{\infty}\frac{\varphi_{n_i}(x)}{i!}\right)\right|<\frac{1}{k!k}M_{n_k}. \tag{7.193}$$

Далі КФ від середнього доданка буде рівна

$$I_{n_k}\left(\frac{\varphi_{n_k}(x)}{k!}\right)=\frac{M_{n_k}}{k!}. \tag{7.194}$$

Співставляючи співвідношення, одержуємо

$$I_n(f)>\frac{M_{n_k}}{k!}-\frac{M_{n_k}}{k!k}-e(b-a). \tag{7.195}$$

Але $M > k!k$, тоді

$$I_k(f)>k!k\left(\frac{1}{k!}-\frac{1}{k!k}\right)-e(b-a)-k-1-e(b-a). \tag{7.196}$$

Наслідок 6. Довести, що квадратурна формула Гауса збігається. Слід перевірити умови теореми (7.172). Дійсно, (7.171) має місце в силу теореми 2 (многочлен $\omega_n(x)g_{n-1}(x)$ степеня $2n$-1). Якщо в (7.165) взяти $f(x) = 1$, тоді КФ $I_n(1)$ є точною.

Наслідок 7. Чи збігається квадратурна формула Ньоютона-Котеса? Можна сказати, що умова $\sum_{i=1}^{n}\left|K_i^{(n)}\right| < M$ не виконується.

Наслідок 8. Довести, що квадратурний процесс Чебишова не збігається. Питання про збіжність не можна ставити, так як при $n = 8$, $n > 10$ не існує формул.

Зауваження 8. Про виконання квадратурних формул.

1. Якщо $f(x)$ задана таблично, то використовують формули Ньоютона-Котеса, Гауса.
2. Якщо $f(x)$ задана графічно або має випадкові похибки, то використовуються формули Чебишова
3. Якщо $f(x)$ малої гладкості, то найбільш практичними будуть узагальнені формули прямокутників, трапецій, Сімпсона.

7.15. Чисельні методи обчислення кратних інтегралів

Нехай для функції n–змінних $f(P): \mathbf{R}^n \to \mathbf{R}^1$ визначено n кратний інтеграл $I(f) \equiv \iint\limits_G \ldots \int f(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_n$. Тоді можна вивести, так звані, кубатурні форми

$$I(f) \equiv \sum_{i=1}^{n} K_i f(P_i) + R_n(f). \tag{7.197}$$

де $P_1, P_2, \ldots, P_n \subset G$, $P_i = (x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)}) \in G$ – сітка вузлів кубатурної формули; $R_n(f)$ – залишковий член кубатурної формули; $\{K_i\}$, $i = \overline{1,n}$ – коефіцієнт кубатурної формули. Зауважимо, що розглядаються функції обмежені, тобто $|f(P)| < M$ в $G \subset \mathbf{R}^n$. Методику обчислення n-кратних інтегралів покажемо на обчисленні двократного інтегралу.

7.15.1. Методи повторного інтегрування для двократного інтегрування

Кратні інтеграли обчислюються повторним обчисленням двократних інтегралів. Для обчислення однократних інтегралів використовуються формули Ньютона-Котеса, Гауса, Чебишова. Нехай

$$I = \iint\limits_G f(x, y) dx dy. \tag{7.198}$$

$$G \equiv \left\{(x, y) \in \mathbf{R}^2 \middle| x \in [a, b], y \in [c, d]\right\}, \tag{7.199}$$

$f \in C(G)$. Тоді відомо, що

$$I = \int_a^b dx \int_c^d f(x,y)dy = \int_a^b F(x)dx, \tag{7.200}$$

$$F(x) \equiv \int_c^d (x,y)dy. \tag{7.201}$$

Застосуємо, наприклад, формулу Сімпсона для обчислення $I \equiv \int_a^b F(x)dx$:

$$I = \frac{b-a}{6}\left(F(a) + 4F\left(\frac{a+b}{2}\right) + F(b)\right) R_1(F). \tag{7.202}$$

$$R_1(F) = -\frac{(b-a)^5(d-c)}{2^5 \cdot 90} F^{(4)}(\xi) = -\frac{(b-a)^5}{2^5 \cdot 90} \int_c^d \frac{\partial^4 f(\xi,y)}{\partial x^4} dy = \\ = \frac{(b-a)^5(d-c)}{2^5 \cdot 90} \frac{\partial^4 f(\xi,\eta)}{\partial x^4}, \tag{7.203}$$

де $\xi \in (a,b)$, $\eta \in (c,d)$.

Далі використаємо формули Сімпсона для обчислення інтегралу

$$F(x) = \int_c^d (x,y)dy = \frac{c-d}{6}\left[f(x,y) + 4f\left(x, \frac{c+d}{2}\right) + f(x,d)\right] + R_y(f). \tag{7.204}$$

$$R_y(f) = \int_c^d (y-c)\left(y - \frac{c+d}{2}\right)(y-d) f\left(x; y; c; \frac{c+d}{2}; \frac{c+d}{2}; d\right) dy. \tag{7.205}$$

Якщо підставити (7.205) у (7.204), то матимемо

$$I = \frac{b-a}{6}\frac{d-c}{6}\{ f(a,c) + f(a,d) + f(b,c) + f(b,d) + \\ + 4\left[f\left(\frac{a+b}{2}, c\right) + f\left(\frac{a+b}{2}, d\right) + f\left(a, \frac{c+d}{2}\right) + f\left(b, \frac{c+d}{2}\right)\right] + \\ + 16 f\left(\frac{a+b}{2}, \frac{c+d}{2}\right) + R(f)\}. \tag{7.206}$$

Цю похибку можна одержати за формулою Сімпсона, де

$$R(f) = -\frac{(b-a)^5(d-c)}{2^5 \cdot 90} \frac{\partial^4 f(\xi_0,\eta_0)}{\partial x^4} - \frac{(b-a)(d-c)^5}{2^5 \cdot 90} \frac{\partial^4 f(\xi_1,\eta_1)}{\partial y^4} - \frac{(b-a)^5(d-c)^5}{2^{10} \cdot 90^2} \frac{\partial^8 f(\xi_2,\eta_2)}{\partial x^4 \partial y^4}. \quad (7.207)$$

де $\xi_i \in (a,b)$, $\eta_i \in (c,d)$, $i = \overline{0,2}$.

Цю похибку можна одержати за формулою Сімпсона та теореми про середнє значення.

7.15.2. Метод заміни підінтегральної функції інтерполяційним многочленом

Нехай потрібно обчислити

$$I \equiv \iint_G f(x,y)dxdy. \quad (7.208)$$

Для $f(x,y)$ побудуємо інтерполяційний многочлен такий, що

$$f(x,y) = L(x,y) + R(x,y), \quad (7.209)$$

$$L(x,y) = \sum_{i=1}^{N} f(x_i,y_i)L_i(x,y). \quad (7.210)$$

А многочлен $L_i(x,y)$ не залежить від $f(x,y)$ і володіє властивістю

$$L_i(x_i,y_i) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \quad (7.211)$$

$$I = \iint_G f(x,y)dxdy \approx \iint_G L(x,y)dxdy = \sum_{i-1}^{N} K_i f(x_i,y_i), \quad (7.212)$$

$$K_i = \iint_G L_i(x,y)dxdy. \quad (7.213)$$

Для простої області G обчислення $\{K_i\}$, $i = \overline{1,n}$, не викликає особливих проблем. Розглянемо побудову кубатурної формули для вигляду прямокутної області інтегрування. Розглянемо сітку вузлів, яка утворюється на перетині прямих

$$\left.\begin{aligned} x_i &= a + i \cdot h; \quad h = \frac{b-a}{n}; \quad i = \overline{0,n} \\ y_j &= c + l \cdot l; \quad l = \frac{d-c}{m}; \quad j = \overline{0,m} \end{aligned}\right\}. \quad (7.214)$$

За сіткою (7.214) будуємо інтерполяційну формулу Лагранжа для функцій двох змінних $f:[a,b]\times[c,d]\to\mathbf{R}^1$:

$$f(x,y)=\sum_{i=0}^{n}\sum_{j=0}^{m}f(x_i,y_i)\frac{\omega_n(x)\omega_m(y)}{(x-x_i)(y-y_i)\omega_n'(x_i)\omega_m'(y_i)}+ \tag{7.215}$$
$$+R(x,y)\equiv L(x,y)+R(x,y),$$

$$\omega_n(x)=\prod_{i=0}^{n}(x-x_i),\ \omega_m(y)=\prod_{j=0}^{m}(y-y_j). \tag{7.216}$$

Інтегруємо формулу (7.216) по прямокутнику $(a\le x\le b, c\le y\le d)$, в результаті одержимо

$$\int_a^b\int_c^d f(x,y)dxdy=\sum_{i=0}^{n}\sum_{j=0}^{m}K_{ij}f(x_i,y_j)+R(f), \tag{7.217}$$

$$K_{ij}\equiv\int_a^b\frac{\omega_n(x)}{(x-x_i)\omega_n'(x_i)}dx\cdot\int_c^d\frac{\omega_m(y)}{(y-y_i)\omega_m'(y_i)}dy, \tag{7.218}$$

$$R(f)\equiv\int_a^b\int_c^d R(x,y)dxdy, \tag{7.219}$$

$$K_{ij}=(b-a)(d-c)I_{i,1}^{(n+1)}\cdot I_{j,1}^{(m+1)}, \tag{7.220}$$

де $I_{i,1}^{(n+1)}, I_{j,1}^{(m+1)}$ – коефіцієнт формули Ньютона-Котеса відкритого типу.

Оскільки похибка інтерполяційної формули Лагранжа $R(x,y)$ має вигляд:

$$R(x,y)=\omega_n(x)f(x;x_0;x_1;...;x_n;y)+\omega_m(y)f(x;y;y_0;y_1;...;y_m)- \tag{7.221}$$
$$-\omega_n(x)\omega_m(y)f(x;x_0;x_1;...;x_n;y;y_0;y_1;...;y_m),$$

то після інтегрування (7.220) по прямокутній області $(a\le x\le b, c\le y\le d)$ одержимо вигляд похибки $R(f)$ кубатурної формули (7.217).

Зауважимо, що формула (7.204) є частинним випадком формули (7.220) для $n=m=2$. Для обчислення інтегралів великої кратності (більше 3) при похибці біля 15% широко використовують методи статистичного моделювання.

Теорема 10. Для сітки вузлів (7.203) функцію $f(x,y)$ можна подати за допомогою інтерполяційного многочлена Лагранжа формулою (7.216).

Доведення. Шукати $L(x,y)$ у (7.216) будемо у вигляді

$$L(x,y)=\sum_{i=0}^{n}\sum_{j=0}^{m}f(x_i,y_i)L_{ij}(x,y), \tag{7.222}$$

$$L_{ij}(x_k, y_l) = \begin{cases} 1, & i = k, j = l; \\ 0, & i \neq k, j \neq l, k = \overline{0,n}, l = \overline{0,m}. \end{cases} \quad (7.223)$$

Це випливає з умови інтерполювання $L(x_k, y_l) = f(x_k, y_l)$. Будемо шукати $L_{ij}(x, y)$ у такому вигляді:

$$L_{ij}(x, y) = L_i(x) L_j(y), \quad (7.224)$$

де $L_i(x)$ – многочлен степеня n: $L_i(x) = A(x - x_0)...(x - x_{i-1})(x - x_{i+1})...(x - x_n)$, для якого

$$L_i(x_j) = 0, i \neq j, L_i(x_i) = 1. \quad (7.225)$$

Але за умовою (7.225) маємо $A = \dfrac{1}{A(x_i - x_0)...(x_i - x_{i-1})(x_i - x_{i-1})...(x_i - x_n)}$,

$$L_i(x) = \frac{\omega_n(x)}{(x - x_i)\omega_n'(x_i)}. \quad (7.226)$$

Аналогічно знаходимо

$$L_i(y) = \frac{\omega_m(y)}{(y - y_i)\omega_m'(y_i)}. \quad (7.227)$$

Тоді з (7.224) за умов (7.226) і (7.227) функція (7.216) $f \in \mathbf{R}^2$ набуде вигляду:

$$f(x, y) = \sum_i \sum_j f(x_i, y_j) \frac{\omega_n(x)\omega_m(y)}{(x - x_i)(y - y_j)\omega_n'(x_i)\omega_m'(y_j)} + R(x, y). \quad (7.228)$$

7.16. Контрольні питання

1. Дайте визначення поняттю *чисельного інтегрування*.
2. Які особливості чисельного інтегрування функцій?
3. Запишіть *узагальнену формулу трапецій*.
4. Графічна інтерпретація формули прямокутників.
5. Графічна інтерпретація методу парабол (Сімпсона).
6. Для чого призначені кубатурні формули?
7. В чому суть методу комірок?
8. Запишіть квадратурні формули найкращого степеня точності.
9. Чи можна застосовувати метод послідовного інтегрування до області довільної форми?
10. В чому полягає ідея побудови квадратурних формул Гауса?
11. Сформулюйте теорему Гауса.

12. Сформулюйте теорему про вибір коефіцієнтів квадратурної формули Гауса.
13. Опишіть метод Чебишова знаходження вузлів КФ.
14. Сформулюйте теорему про необхідні та достатні умови збіжності КФ.
15. В якому випадку використовують формули Ньоютона-Котеса?
16. В якому випадку використовують формули Чебишова?
17. Для якої функції найбільш практичними будуть узагальнені формули прямокутників, трапецій, Сімпсона?

7.17. Задачі для самостійної роботи

Приклад 7.17.1. Скласти таблицю для $I_{i,m}^{(n)}$, $m=0$, $2\le n\le 5$ (ІКФ закритого типу).

Таблиця 1 Умова задачі 7.17.1 (m=0, ІКФ закритого типу)

№	m	n	I	№	m	n	I
1	0	1	$I_{1,0}^{(1)}$	3	0	3	$I_{1,0}^{(3)}$, $I_{2,0}^{(3)}$, $I_{3,0}^{(3)}$
2	0	2	$I_{1,0}^{(2)}$, $I_{2,0}^{(2)}$	4	0	4	$I_{1,0}^{(4)}$, $I_{2,0}^{(4)}$, $I_{3,0}^{(4)}$, $I_{4,0}^{(4)}$

Аналогічно у випадку коли n=k, буде $I_{1,0}^{(k)}$, $I_{2,0}^{(k)}$, $I_{3,0}^{(k)}$, …, $I_{k,0}^{(k)}$ при $k\le n$.

Задача 7.17.2. Скласти таблицю: $I_{i,m}^{(n)}$, $m=1$, $2\le n\le 5$ (ІКФ відкритого типу).

Таблиця 3 Умова задачі 7.17.2 (m=0, ІКФ закритого типу)

№	m	n	I	№	m	n	I
1	1	1	$I_{1,1}^{(1)}$	3	1	3	$I_{1,1}^{(3)}$, $I_{2,1}^{(3)}$, $I_{3,1}^{(3)}$
2	1	2	$I_{1,1}^{(2)}$, $I_{2,1}^{(2)}$	4	1	4	$I_{1,1}^{(4)}$, $I_{2,1}^{(4)}$, $I_{3,1}^{(4)}$, $I_{4,1}^{(4)}$

Аналогічно у випадку коли n=k, буде $I_{1,0}^{(k)}$, $I_{2,0}^{(k)}$, $I_{3,0}^{(k)}$, …, $I_{k,0}^{(k)}$ при $k\le n$. Обчислити коефіцієнти ІКФ для n=5, m=1, $i=\overline{1,5}$ та n=6, m=1, $i=\overline{1,6}$ і результат звести в таблицю.

Задача 7.17.3. Отримати наступні формули наближеного інтегрування

a) $$\int_{-1}^{1}\sqrt{1-x^2}\,f(x)dx\approx\frac{\pi}{n}\sum_{k=1}^{n}f\left(\cos\frac{k\pi}{n}\right)\sin^2\frac{k\pi}{n}.$$

b) $$\int_{-1}^{1}\sqrt{\frac{1-x}{1+x}}f(x)dx\approx\frac{4\pi}{n}\sum_{k=1}^{n}f\left(\cos\frac{2k\pi}{n}\right)\sin\frac{2k\pi}{n}.$$

c) $$\int_{0}^{1}\frac{f(x)dx}{\sqrt{x(1-x)}}\approx\frac{\pi}{n}\sum_{k=1}^{n}f\left(\cos^2\frac{(2\pi+1)\pi}{2n}\right).$$

d) $\int_0^1 \sqrt{x(1-x)}f(x)dx \approx \frac{\pi}{4n}\sum_{k-1}^{n} f\left(\cos^2\frac{2k\pi}{n}\right)\sin^2\frac{2k\pi}{n}$.

e) $\int_0^1 \sqrt{\frac{1-x}{x}}f(x)dx \approx \frac{2\pi}{n}\sum_{k-1}^{n} f\left(\cos^2\frac{k\pi}{n}\right)\sin\frac{2k\pi}{n}$.

f) Знайти залишкові члени.

Задача 7.17.4. Показати, що формула

$$\int_0^{\infty} e^{-x}f(x)dx \approx \frac{2+\sqrt{2}}{4}f(2-\sqrt{2})+\frac{2-\sqrt{2}}{4}f(2+\sqrt{2}).$$

Дати точне значення, якщо $f(x)$ – многочлен степеня не більше 3.

Задача 7.17.5. Показати, що формула

$$\int_{-\infty}^{\infty} e^{-x}f(x)dx \approx \frac{\sqrt{\pi}}{6}\left[f\left(-\sqrt{\frac{3}{2}}\right)+4f(0)+f\left(\sqrt{\frac{3}{2}}\right)\right].$$

Дати точне значення, якщо $f(x)$ – многочлен степеня не більше 3.

Задача 7.17.6. Обчислити наступні інтеграли:

1) $\int_0^1 \frac{dx}{1+x}$; 3) $\int_0^1 \frac{dx}{1+x^2}$; 5) $\int_0^1 \frac{\ln(1+x)}{1+x^2}dx$; 7) $\int_0^1 \frac{\ln(1+x)}{x}dx$.

2) $\int_1^3 \frac{dx}{1+x}$; 4) $\int_0^{\frac{\pi}{2}} \frac{\sin x}{x}dx$; 6) $\int_0^1 \frac{\ln(1+x)}{1+x}dx$;

з точністю 10^{-3} за допомогою наближених формул трапецій, Сімпсона, Чебишова (&5), Гауса (&4).

Задача 7.17.7. Всі нулі многочлена $\omega_n(x)=\frac{d^n}{dx^n}\left[(x-c)^n(d-x)^n\right]$ є дійсними і різними $x_i^{(n)} \in [c,d]$ (перевірити послідовним використанням теорем Ролля).

Задача 7.17.8. Одержати кубатурну формулу прямокутників.

Задача 7.17.9. Одержати кубатурну формулу трапецій.

7.18. Завдання до лабораторної роботи

Перший рівень. Написати програмну реалізацію всіх вищенаведених квадратурних формул для наступних функцій:

1) $f(x)=x$; 2) $f(x)=4x^2$; 3) $f(x)=\sqrt{x^2+r^2}, r>1$.

Другий рівень. Знайти абсолютну і відносну похибки обчислень інтегралів.

Третій рівень. Розробити алгоритм розв'язування задачі згідно індивідуального завдання та написати програму. Текст програми має бути з коментарями. Результати виконання програми мають бути з поясненнями.

7.19.Варіанти завдання (рівняння для розв'язування)

Індивідуальні завдання до лабораторної роботи

№ варіанта	Підінтегральна функція $f(x)$	a	b	ε
1.	$3x^9 e^{\sqrt{x^6+3}} \sin(20x^8)$	0	1,5	0,001
2.	$2^{\sin(10x)}(0,2x^2+0,12x+0,83)^8$	-2	2	0,001
3.	$tg^3(0,1x)\sqrt[5]{x^{10}+0,1x^6+4x}$	-2	3	0,002
4.	$2^{\cos(3x^4)}(0,01x^{12}+x^4\log_4(x^4+2)-0,121)$	-2	2,5	0,001
5.	$\cos^4(0,1x^2)(0,1x^{12}+e^{4x^2}x^5-3x^2+10)$	0	1,5	0,025
6.	$\ln^2(x^4+2)((0,2x)^{14}+4x^5-3x^3)\cos^2(10x)$	-2	3	0,001
7.	$\arcsin(0,1x^3)(2^{x^2+x}x^{-4}+0,012x^{10})$	-2	1,5	0,025
8.	$7^{\cos(7x^4)}\sin^6(x^5)(7x^8+3x^2+2)^2$	-3	3	0,05
9.	$\frac{x^{12}+6x^6-7x^3+8x^2+8}{x^6+8x^4+12x^2+8}\sin(3^{x^4})$	-1	2	0,01
10.	$6^{3\sin x tg(0,1x^2)}(2^{-x^3}x^6+7x^5+3x^3-1)$	-2	3	0,001
11.	$7^{\sin^3 x}(0,35x^{12}+0,4x^8-0,01x^2+0,125)$	-1	4	0,01
12.	$\log_5(0,001x^{10}+0,4x^6+1)(0,1x^{12}+2,5x)$	-3	3	0,1
13.	$tg(\cos^4 x+1)(x^{10}+0,1x^4+1)$	-2	2,5	0,02
14.	$2x^{18}e^{\sqrt{x^4+3}}\cos^3(2x^4)\log_2(x^6+0,15)$	0	1,5	0,01
15.	$(2,2x^2+1,12x+1,83)^8 2^{\cos^4(12x)}$	-2	2	0,04
16.	$ctg^4(0,15x)\sqrt[7]{0,01x^{16}+0,1x^4+4}$	2	4	0,05
17.	$2^{\sin^3(3x^3)}(x^{14}+\log_4(x^6+2)x^4-0,12)$	-2	2,5	0,1
18.	$\cos^2(0,2x^4)(0,15x^{14}+e^{4x^2}x^4-3x^3+1)$	0	1,5	0,05
19.	$\log_3(x^4+2)(6x^{10}+4x^5-3x^2+1)\sin^4(10x)$	-2	3	0,054
20.	$tg^4(\sin^4 x-0,5)(x^{12}+3,1x^6+1)$	-2	2,5	0,02
21.	$\lg(0,001x^8+4x^4+1)(0,01x^{10}+2,5x)$	-2	2	0,02
22.	$7^{x^6\cos^4 x}(0,135x^{18}+0,4x^{12}-0,12x^2+0,125)$	-1	1,5	0,025
23.	$5^{3\cos^4 x tg(0,1x^2)}(2^{x^2}x^8+7x^5+3x-1)$	-2	3	0,001

24.	$\frac{11x^{11}+6x^5-7x^3+8x+8}{4x^6+12x^4+17x^2+8}\cos(2^{-x^2})$	-1	2	0,01
25.	$8^{\cos(10x^4)}\cos^4(x^5)(7x^5+3x^2+2)^2$	-3	3	0,05
26.	$arctg(0,1x^5)(2^{x^2}+x^4+0,12x^8)$	-2	4	0,025
27.	$\arccos^4(0,1x^2)(2^{0,1x^2+x}x^{-5}+0,0125x^{11})$	-2	4	0,025
28.	$3x^{19}e^{\sqrt{x^6+3}}\cos(20x^8)\lg(x^4+0,1)$	0	1,5	0,001
29.	$2^{\cos^2(10x)}(0,2x^2+0,12x+0,83)^9$	-2	2	0,001
30.	$ctg^3(0,1x)\sqrt[5]{0,001x^{10}+0,1x^6+4x}$	2	4	0,002
31.	$2^{\sin(3x^3)}(x^{12}+\lg(x^4+2)x^4-0,12)$	-2	2,5	0,001
32.	$\sin^4(0,2x^2)(0,1x^{12}+e^{4x^2}x^5-3x^3+10)$	0	1,5	0,025
33.	$\ln(x^4+2)(6x^8+4x^5-3x^3)\sin^2(10x)$	-2	3	0,001

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 – С. 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. 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.
4. 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.
5. 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.
6. 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.
7. 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.
8. 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.

9. 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.
10. 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.
11. 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.
12. 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.
13. 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.
14. 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.
15. 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.
16. 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.
17. D'Agunno B. CHPACK: A package for manipulation Chebyshev approximation / B. D'Agunno, A. Nobile // Comput. Commun. – 1983. – 29, N4. – P.361 – 374.
18. 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.
19. Everitt B. A handbook of statistical analyses using R / B. Everitt, T. Hothorn. – 2–nd ed. – Chapman and HALL/CRC, 2009. – 376 p.
20. 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.
21. Geddes K.O. A Package for Numerical Approximation / K.O. Geddes. – Maple Tech, 10. – 1993. – P. 28–36.
22. Kabacoff Robert I. Quick–R [Electronic resource] / Robert I. Kabacoff. – Access mode: http://www.statmethods.net/index.html.
23. 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.
24. 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.
25. 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.
26. Netlib. Class K2. Minimax (L–infinity) approximation [Electronic resource]. – Access mode: http://gams.nist.gov/serve.cgi/Class/K2.

27. R Site Search [Electronic resource]. – Access mode : http://finzi.psych.upenn.edu/nmz.html.
28. Rtips. Revival 2014! [Electronic resource]. – Access mode : http://pj.freefaculty.org/R/Rtips.html.
29. 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.
30. 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.
31. 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.
32. Solution of Overdetermined Linear System in the Chebychev Norm [Electronic resource]. – Access mode: http://wwwasdoc.web.cern.ch/wwwasdoc/shortwrupsdir/e222/top.html.
33. Statistics with R [Electronic resource]. – Access mode : http://zoonek2.free.fr/UNIX/48_R/all.html.
34. The Comprehensive R Archive Network [Electronic resource]. – Access mode: http://cran.r–project.org.
35. 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.
36. 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.
37. 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.

38. 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.

39. 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.

40. 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.

41. 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.

42. 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.

43. 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.

44. 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.
45. 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.
46. 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.
47. 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.
48. 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.
49. 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.
50. 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 – С. 328– 332.

51. 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.
52. 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.
53. 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.
54. 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.
55. 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.
56. 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.
57. 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.
58. 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.
59. 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.
60. 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.
61. 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.
62. 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.
63. 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.
64. Zadachyn V. Calculation of optimal path for parallel car parking / V. Zadachyn, O. Dorokhov // Transport and Telecommunication. – Volume 13. – 2012. – P. 303–309.
65. 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.
66. 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.
67. 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.

68. 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.

69. 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/.

70. 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.

71. Амосов А.А. Вычислительные методы для инженеров: учебн. пособ. / А.А. Амосов, Ю.А. Дубинський, Н.В. Копченова. – М. : Высшая школа, 1994. – 544 с.

72. 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.

73. 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.

74. Ахиезер Н.И. Лекции по теори апроксимации / Н.И. Ахиезер. – Москва: Наука, 1965.

75. Базара М. Нелинийное программирование. Теория и алгоритмы / М. Базара, К. Шетти. – М: Мир, 1982. – 584 с.

76. Барахнин В.Б. Введение в численный анализ / В.Б. Барахнин, В.П. Шапеев. – Новосибирск, 1997. – 112 с.

77. Бахвалов Н.С. Численные методы / Н.С. Бахвалов, Н.П. Жидков, Г.М. Кобельков. – М: Бином, 2007. – 636 с.

78. Бахвалов Н.С. Численные методы. Анализ, алгебра, обычные диференциальные уравнения / Н.С. Бахвалов. – М.: Наука, 1975. – 631 с.

79. Березин И.С. Методы вичислений / И.С. Березин, Н.П. Житков. – М.: Наука, 1966. т. 1. – 632 с.

80. 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.

81. 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.
82. 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.
83. 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.
84. 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.
85. 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.
86. 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.
87. 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.
88. Биллингсли П. Схоимость вероятностных мер / П. Биллингсли. – М.: Наука, 1977. – 352 с.
89. 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/.
90. 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.
91. 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/.

92. Боглаев Ю.П. Вычислительная математика и программирование / Ю.П. Боглаев. – М: Высшая школа, 1990. – 544 с.

93. Бронтейн И.Н. Справочник по математике для инженеров и учащихся втузов / И.Н. Бронштейн, К.А. Семендяев. – М.: Нука, 1986. – 544 с.

94. 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.

95. 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.

96. 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.

97. 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.

98. 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.

99. 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.

100. 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.

101. 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.

102. 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.

103. 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.

104. 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.

105. 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.

106. 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.

107. Волков Е.А. Численные методы / Е.А. Волков. – М: Высшая школа, 1987. – 312 с.

108. Выгодский М.Я. Справочник по высшей математике / М.Я. Выгодский. – М.: Наука, 1977. – 872 с.

109. 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.

110. Гаврілюк І.П. Методи обчислень / І.П. Гаврілюк, В.Л. Макаров.– К: Вища школа, в 2–х частинах, 1995.

111. 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.

112. Геращенко О.А. Температурные измерения / Геращенко О.А., Гордов А.И., Еремина А.К. и др. – Киев: Наук. думка, 1989. – 704 с.

113. 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.
114. Гмурман В.Е. Теория вероятностей и математическая статистика / В.Е. Гмурман. – М: Высшая кола, 1999. – 479 с.
115. 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.
116. Грановский В.А. Методы обработки экспериментальных данных при измерениях / В.А. Грановский, Т.Н. Сирая // Ленинград: Энергоатомиздат. Ленингр. отделение, 1990. – 288с.
117. 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.
118. Данилина Н.И. Вычислительная математика / Н.И. Данилина, Н.С. Дубровская, О.П. Кваша, Г.Л. Смирнов. – М.: Высшая школа, 1985. –368 с.
119. Де Бор К. Практическое руководство по сплайнам / К. Де Бор. – М: Радио и связь, 1985. – 304 с.
120. Демидович Б.П. Основы вычислительной математики / Б.П. Демидович. – М: Наука, 1994. – 664 с.
121. Демидович Б.П. Численные методы анализа / Б.П. Демидович, И.А. Марон, Е.З. Шувалова. – М: Мир, 1967
122. Демьянов В.Ф. Введение в минимакс / В.Ф. Демьянов, В.Н. Малоземов. – М: Наука, 1972. – 368 с.
123. 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.
124. Денежкина И.Е. Численные методы. Курс лекций: Учебное пособие / И.Е. Денежкина. – М: Финакадемия, 2010. – 132 с.
125. 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.
126. Дэннис Дж. Численные методы безусловной оптимизации и решения нелинейных уравнений / Дж. Дэннис, Р. Шнабель. – М: Мир, 1988. – 40 с.

127. Дьяконов В.П. Математическая система MAPLE V R3/R4/R5 / В.П. Дьяконов. – М: Солон, 1998. – 399 с.
128. Емелин А. Высшая математика для заочников и не только [Electronic resource] / А. Емелин. – 2010–2016. – Access mode: http://mathprofi.ru.
129. Емец С.В. Алгоритм корекции статических характеристик измерительных преобразователей / С.В. Емец, С.А. Емец // Межвуз. науч. сб. – Уфа: УГАТУ, 1996. – С. 152–159.
130. Жалдак М.І. Чисельні методи математики / М.І. Жалдак, Ю.С. Рамський. – К:Радянська школа, 1984. – 206 с.
131. Заварыкин В.М. Численные методы / В.М. Заварыкин, В.Г. Житомирский, М.П. Лапчик. – М: Просвещение, 1990. – 176 с.
132. 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.
133. 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.
134. 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.
135. 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.
136. 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.
137. Калиткин Н.Н. Численные методы / Н.Н. Калиткин. – М: Наука, 1978. – 512 с.
138. Катренко А.В. Теорія прийняття рішень: підручник для вищих навчальних закладів / А.В. Катренко, В.В. Пасічник, В.П. Пасько. – К: Видавнича група BHV, 2009. – 448 с.
139. 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.
140. Кононюк А.Е. Дискретно–непрерывная математика. Матрицы / А.Е. Кононюк. – К: Освіта України, 2012. – 508 с.
141. Копченова Н.В. Вычислительная математика в примерах и задачах / Н.В. Копченова, И.А. Марон. – М: Наука, 1972. – 368 с.

142. 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.
143. Корнейчук А.А. Итерационный метод отыскания многочлена наилучшего приближения / А.А. Корнейчук, Н.Ю. Ширикова. – Дубна, 1967.
144. Коссак О. Методи наближених обчислень, навч. посібн. / О. Коссак, О. Тумашова, О. Коссак.– Л: БаК, 2003. – 168 с.
145. 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.
146. Кривоносов А.И. Статические характеристики поликристалических резисторов / А.И. Кривоносов, В.Я. Кауфман. – М: Энергия, 1976. – 120 с.
147. 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, – 304С.
148. 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.
149. Ляшко И.И. Методы вичислений / И.И. Ляшко, В.Л. Макаров, А.А. Скоробагатько. – Киев: Вища школа, 1977. – 406 с.
150. Лященко М.Я. Чисельні методи / М.Я. Лященко, М.С. Голованъ. – К: Либідь, 1996. – 287 с.
151. 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.
152. Малачівський П. Рівномірне сплайн–наближення / П. Малачівський, В. Андруник // "Комп'ютерні технології друкарства", № 7. – Львів: Українська академія друкарства, 2002. – С. 107–115.
153. Малачівський П. Модифікований алгоритм Валле–Пуссена / П. Малачівський // Фізико–математичне моделювання та інформаційні технології. 2005. – Вип. 2. – С. 159–166.
154. 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.

155. Малачівський П.С. Програмування в середовищі Visual Basic / П.С. Малачівський. – Львів: Видавництво "Бескид Біт", 2004. – 260 с.
156. 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.
157. 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.
158. 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.
159. 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.
160. Малоземов В.Н. Наилучшее равномерное приближение функций нескольких аргументов / В.Н. Малоземов. – Киев: Ин–т кибернетики АН УССР. – 1970.
161. Маркушевич А.И. Краткий курс теории аналитических функцій / А.И. Маркушевич. – М.: Наука, 1978. – 416 с.
162. Марчук Г.И. Методы вычислительной математики / Г.И. Марчук. – М. : Наука, 1989. – 608 с.
163. 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.
164. Иванов В.В. Методы вычислений на ЭВМ: Справочное пособие / В.В. Иванов. – Киев: Наук. думка, 1986. — 564 с.
165. 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.
166. 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.
167. 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.

168. 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.
169. Мэйндоналд Дж. Вычислительные алгоритмы в прикладной статистике / Дж. Мэйндоналд // Под ред. Е.З. Демиденко. – М: Финансы и статистика, 1988. – 350 с.
170. Носач В.В. Решение задач аппроксимации с помошью персональных компьютеров / В.В. Носач // М: Микап, 1993. – С. 6–7.
171. Окуненко В.М. Чисельні методи в моделюванні систем / В.М. Окуненко, В.К. Ясинський. – Чернівці: Золоті литаври, 2006. – 592 с.
172. Олександренко В.Л. Побудова чебишовського найкращого поліноміального наближення функції одної змінної за методом підвищуючої дії / В.Л. Олександренко, А.О. Порханова // Автоматика. – №4. – 1967.
173. Ортега Дж. Введение в численные методы решения дифференциальных уравнений / Дж. Ортега, У. Пул. – М: Наука, 1986. – 56 с.
174. 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.
175. Полиа Г. Задачи и теоремы из анализа / Г. Полиа, Г. Сеге. – Том I, II. – М.: Наука, 1978. – 392 с., 432 с.
176. Попов Б.А. Некоторые способы наилучшего приближения функцій / Б.А. Попов. – К: Наук. думка, 1973.
177. Попов Б.А. Наилучшие чебышевские приближения суммой многочлена и нелинейных функций / Б.А. Попов, П.С. Малачивский. – Львов, 1984. – 79 с.
178. Попов Б.А. Приближение функций для технических приложений / Б.А. Попов, Г.С. Теслер. – К: Наук. думка, 1980. – 352 с.
179. Попов Б.А. Равномерное приближение сплайнами / Б.А. Попов. – К: Наук. думка, 1989. – 272 с.
180. Попов Б.О. Розв'язування задач у системі комп'ютерної алгерби MAPLE V / Б.О. Попов. – К: ViP, 2001. – 312 с.
181. Попов Б.О. Чисельні методи рівномірного наближення сплайнами / Б.О. Попов. – Л: Львів. націон. ун–т, 1992. – 92 с.
182. Попов Б.А. Приближение функций для технических приложений / Б.А. Попов, Г.С. Теслер. – К: Наукова думка, 1980. – 352 с.
183. Правила устройства элелектроустановок. – М: Энергоатомиздат, 1985.
184. Пулькин С.П. Вычислительная математика / С.П. Пулькин, Л.Н. Никольская, А.С. Дьячков. – М.: Просвещение, 1980. – 176 с.
185. 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.

186. Рего К.Г. Метрологічна обробка результатів технічних вимірювань / К.Г. Рего. – Київ: Вища школа, 1987. – 128 с.
187. Ремез Е.Я. Общие вычислительные методы чебишевского приближения / Е.Я. Ремез. – К: Ин–т кибернетики АН УССР, 1957.
188. Ремез Е.Я. Основы численных методов чебышевского приближения / Е.Я. Ремез. – Киев: Наук. думка, 1969. – 623 с.
189. 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.
190. 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.
191. 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.
192. Самарский А.А. Введение в численные методы / А.А. Самарский. – М: Наука, 1997. – 240 с. 16.
193. Самарский А.А. Численные методы / А.А. Самарский, А.В. Гулин. - М: Наука, 1989. – 432 с.
194. Сборник задач по методам вычислений / под ред. П. И. Монастырного. – М: Наука, 1994. – 320 с.
195. 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.
196. 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.
197. Тейлор Дж. Введение в теорию ошибок / Дж. Тейлор. – М: Мир, 1985. – 16 с.
198. Теслер Г.С. Аппроксимация элементарных функций с помощью полиномов нулевой и первой степени / Г.С. Теслер. – Киев: Ин–т кибернетики АН УССР, 1969.
199. Ткачук К.В. Захисні плівкові екрани для моніторів ЕОМ / К.В. Ткачук та ін. // Охорона праці, №4, 1997.
200. Турецкий В.Я. Математика и информатка / В.Я. Турецкий. – М: ИНФРА–М, 2000. – 560 с.
201. Турчак Л.И. Основы численных методов / Л.И. Турчак. – М: Наука, 1997. – 320 с.
202. Федоренко Р.П. Введение в вычислительную физику / Р.П. Федоренко. – М: МФТИ, 1994. – 528 с.

203. 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.
204. Форсайт Дж., Малькольм М., Моулер К. Машинные методы математических вычислений. – М.: Мир, 1980, – 280 с.
205. Форсайт Дж. Численное решение систем линейных алгебраических уравнений / Дж. Форсайт, К. Молер.– М: Мир, 1969.–168 с.
206. 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.
207. Холл Дж. Современные численне методы решения обыкновенных дифференциальных уравнений / Дж. Холл, Дж. Уатт. – М: Мир, 1979. – 312с.
208. Цегелик Г.Г. Чисельні методи: підручник / Г.Г. Цегелик. – Л: Видавничий центр Львівського національного університету імені Івана Франка, 2004. – 408 с.
209. Чабан В. Чисельні методи / В.Чабан. – Львів: В–во НУ «Львівська політехніка», 2001. – 186 с.
210. 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.
211. 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.
212. 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.
213. 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.
214. 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.

215. 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.
216. 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.
217. 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.
218. Чирун Л.В. Основи програмування: конспект лекцій / Л.В. Чирун, В.А. Висоцька. – Л: В–во НУ «Львівська політехніка», 2010. – 174 с.
219. 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.
220. 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.
221. 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.
222. 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»CЋPlP.PDF.
223. 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.
224. 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.
225. Шмойлов В.И. Непрерывные дроби и комплексные числа: монография / В.И. Шмойлов, Л.В. Чирун. – Л: Меркатор, 2001. – 564 с.

226. Шунда Н.М. Функції та їх графіки / Н.М. Шунда. – К: Радянська школа, 1976. – 192 с.
227. Шуп Т. Решение инженерных задач на ЭВМ: практическое руководство : учебн. изд. / Т. Шуп; пер. с англ. – М: Мир, 1982. – 237 с.
228. Шураков В.В. Задачник по основам алгоритмизации, алгоязыкам и машинной обработке: учебное пособие / В.В. Шураков, В.П. Морозов. – М: Статистика, 1978. – С. 198.
229. Ястрембский А.Л. Об одном алгоритме рациональной аппроксимации таблично заданой функции / А.Л. Ястрембский, П.Г. Македонская. – М: ФТИНТ, 1971.
230. 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 1

Навчальний посібник

Підписано до друку 23.10.2017 р.
Формат 70x100/16. Папір друк. №2. Гарнітура Times New Roman Умовн. друк. арк. 28,12.

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'ютер та інформаційних технологій Національного університету "Львівська політехніка". В 2002 році з відзнакою Українську академію друкарства за спеціальністю «Комп'ютерно-інтегровані технологічні процеси та вироб Працює на посаді асистента кафедри ІСМ, веде лабораторні заняття з курсів «Комп'ютерна графіка», «Си аналіз», «Моделювання складних систем», «Комп'ютерні системи обробки текстової та графічної інф Публікує статті у фахових виданнях, приймає активну участь у міжнародних науково-технічних конф виставках та семінарах.

Область наукових інтересів – математичне моделювання, неперервна та гладка мінімаксна апроксимація, комп'ютерна графіка.



Висоцька Вікторія Анатоліївна – доцент кафедри «Інформаційні системи та мережі» НУ «Львівська пол Автор понад 210 наукових праць, співавтор 3-х монографій та 3-х навчальних посібників. В 2014 р. з дисертаційну роботу на тему «Методи та засоби опрацювання інформаційних ресурсів в системах еле контент-комерції». Бере активну участь у роботі міжнародних конференцій (http://victana.lviv.ua). Читає лекційн таких дисциплін: «Дискретна математика», «Математична лінгвістика», «Математична структурна та п лінгвістика», «Чисельні методи», «Алгоритмізація та програмування», «Методи опрацювання природної мови платформне програмування та хмарні сервіси», «Технології електронної комерції та Інтернет-маркетингу», «Т створення програмних продуктів», «Системний аналіз», «Технології комп'ютерного проектування». Наукові ін системи електронної контент-комерції, методи опрацювання природної мови, математична лінгвістика, N технології, інтелектуальні системи прийняття рішень, технологія програмування та створення програмних п штучний інтелект, інтелектуальний аналіз даних.



Пасічник Володимир Володимирович — професор, доктор технічних наук, професор кафедри інфор систем та мереж, Інституту комп'ютерних наук та інформаційних технологій Національного університету «Л політехніка». Автор понад 200 наукових праць, понад 10 монографій та 20 підручників та навчальних посібни року захистив кандидатську дисертацію на здобуття наукового ступеня кандидата фізико-математичних наук спеціальність «Математична кібернетика». 1994 року захистив дисертацію на здобуття наукового ступеня технічних наук за спеціальністю «Теоретичні основи інформатики». Відзначений державною премією Україні науки та техніки. Область наукових інтересів – створення та розроблення інформаційних систем для економіки, виробничої та невиробничої сфери. Керівник наукової школи з інформаційного моделювання, си даних і знань. Під науковим керівництвом професора В. Пасічника захищено більше 10 докторських та б кандидатських дисертацій. Науковий редактор освітньої серії підручників і посібників «КОМП'ЮТИНГ», відпов редактор серії «Інформаційні системи та мережі» Вісника Національного університету «Львівська політехніка»



Чирун Лілія Борисівна – провідний спеціаліст інституту комп'ютерних наук та інформаційних техно «Львівська політехніка». Працює у НУ «Львівська політехніка» з 1979 р., після закінчення Львівського де університету ім. І. Франка за спеціальністю «Прикладна математика» факультету прикладної математики та З жовтня 2005 р. працювала старшим викладачем кафедри інформаційних систем та мереж Інституті комп наук та інформаційних технологій. Є співавтором рацпропозиції та винаходу, понад 60-ти наукових методичних розробок. Читала лекційні курси: «Комп'ютерні мережі», «Сховища даних та довідково-інфо фонди», «Комп'ютерно-технологічні засоби інформаційної діяльності», «Системи автоматизованого перекладу

Наукові інтереси – сховища даних, веб-технології, прикладна лінгвістика, застосування чисельних м інформаційних технологіях.



Чирун Любомир Вікторович – доцент кафедри «Системи автоматизованого проектування» НУ «Л політехніка», кандидат технічних наук. З 2001 року працював в НУ «Львівська політехніка» в Інституті комп наук та інформаційних технологій. на посаді доцента кафедри інформаційних систем і мереж. В 2007р. кандидатську дисертацію за спеціальністю 01.05.02 - «Математичне моделювання та обчислювальні ме співавтором монографії «Непрерывные дроби и комплексные числа». Автор понад 90 наукових публікац лекційні курси з наступних дисциплін: «Системне програмування та операційні системи», «Основи програму алгоритмічні мови», «Технологія програмування та створення програмних продуктів», «Чисельні м інформатиці».

Наукові інтереси – розпізнавання образів, застосування чисельних методів в інформаційних тех об'єктно-орієнтоване програмування, веб-технології.

Вища освіта в Україні