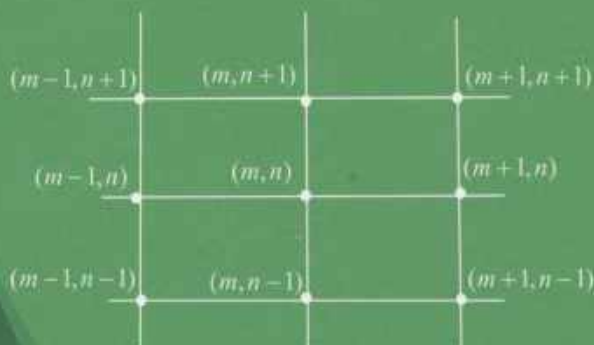


С.М. Шахно,  
А.Т. Дудикевич,  
С.М. Левицька

# Практикум з чисельних методів



$$B_{m-1,k}(x) = B_{m-1}(x_k; x_{k+1}, \dots, x_{k+m}; x) = m \sum_{i=k}^{k+m} \frac{(x_i - x)_+^{m-1}}{\omega'_{m,k}(x_i)},$$

$$(x_i - x)_+^{m-1} = [\max \{0; (x_i - x)\}]^{m-1};$$

$$\omega_{m,k}(x) = (x - x_k)(x - x_{k+1}) \dots (x - x_{k+m}).$$

Міністерство освіти і науки, молоді та спорту України  
Львівський національний університет імені Івана Франка

**М. ШАХНО, А.Т. ДУДИКЕВИЧ, С.М. ЛЕВИЦЬКА**

# **П РА К Т И К У М**

## **З ЧИСЕЛЬНИХ МЕТОДІВ**

**Навчальний посібник**

*Рекомендовано*

*Міністерством освіти і науки, молоді та спорту України*

Львів  
ЛНУ імені Івана Франка  
2013

УДК 519.6(075.8)  
ББК В 22.192я73  
Ш 81

**Рецензенти:**

д-р фіз.-мат. наук, ст. наук. співроб. **О.М. Хімич**  
(Інститут кібернетики НАН України);

д-р фіз.-мат. наук, проф. **Я.М. Чабанюк**  
(Національний університет "Львівська політехніка");

канд. фіз.-мат. наук, доц. **В.О. Оніщук**  
(Луцький національний технічний університет)

**Рекомендовано**

*Міністерством освіти і науки, молоді та спорту України  
як навчальний посібник для студентів вищих навчальних закладів  
(Лист № 1/11-6147 від 03.05 2012 року)*

**Шахно С.М.**

Ш 81 Практикум з чисельних методів навч. посібник / Шахно С.М.,  
Дудикевич А.Т., Левицька С.М. – Львів ЛНУ імені Івана Франка,  
2013. – 432 с.  
ISBN 978-617-10-0036-0.

Розглянуто головні чисельні методи розв'язування задач лінійної алгебри, математичного аналізу, звичайних диференціальних рівнянь, рівнянь у частинних похідних, інтегральних рівнянь та різні аспекти їхнього практичного використання. Наведено результати тестових розрахунків та набори завдань для самостійної роботи.

Для студентів вищих навчальних закладів, які вивчають чисельні методи.

**УДК 519.6(075.8)  
ББК В 22.192я73**

© Шахно С.М., Дудикевич А.Т.,  
Левицька С.М., 2013  
© Львівський національний університет  
імені Івана Франка, 2013

ISBN 978-617-10-0036-0

# Зміст

Вступ .....	7
Розділ 1. УРАХУВАННЯ ПОХИБОК НАБЛИЖЕНИХ ОБЧИСЛЕНЬ ..	9
1.1. Елементи теорії похибок .....	9
1.2. Абсолютна та відносна похибки .....	10
1.3. Пряма задача теорії похибок .....	11
1.4. Обернена задача теорії похибок .....	18
Завдання до розділу 1 .....	20
Розділ 2. ПРЯМІ МЕТОДИ РОЗВ'ЯЗУВАННЯ СИСТЕМ ЛІНІЙНИХ АЛГЕБРИЧНИХ РІВНЯНЬ .....	24
2.1. Метод Гаусса розв'язування систем лінійних алгебричних рівнянь .....	24
2.2. Розв'язування систем лінійних алгебричних рівнянь за допомогою LU-розкладу .....	29
2.3. Метод квадратних коренів .....	34
2.4. Метод ортогоналізації .....	37
2.5. Метод прогонки розв'язування СЛАР з тридіагональною матрицею .....	41
2.6. Метод поворотів розв'язування СЛАР .....	43
Розділ 3. ІТЕРАЦІЙНІ МЕТОДИ РОЗВ'ЯЗУВАННЯ СИСТЕМ ЛІНІЙНИХ АЛГЕБРИЧНИХ РІВНЯНЬ .....	49
3.1. Принципи побудови ітераційних процесів .....	49
3.2. Метод простих ітерацій .....	51
3.3. Метод Зейделя .....	57
3.4. Метод релаксації .....	65
3.5. Ітераційні методи варіаційного типу .....	68
3.5.1. Метод мінімальних відхилів .....	68
3.5.2. Метод мінімальних поправок .....	71
3.5.3. Метод спряжених градієнтів .....	74
Завдання до розділів 2, 3 .....	78
Розділ 4. МЕТОДИ РОЗВ'ЯЗУВАННЯ ЗАДАЧ НА ВЛАСНІ ЗНАЧЕННЯ .....	95
4.1. Класичний спосіб .....	96
4.2. Степеневий метод знаходження максимального за модулем власного значення і відповідного йому власного вектора .....	99
4.3. Метод скалярних добутків .....	103
4.4. Метод зворотних ітерацій .....	107
4.5. Зворотні ітерації зі змінними зсувами. RQI-алгоритм .....	111
4.6. Метод поворотів Якобі для знаходження власних пар симетричної матриці .....	117
4.7. LU-алгоритм для несиметричних задач .....	122

4.8. QR-алгоритм з використанням перетворень Хаусхолдера.....	128
4.9. QR-алгоритм з використанням плоских поворотів Гівенса.....	134
4.10. Метод Ланцоша зведення симетричної матриці до тридіагонального вигляду.....	140
4.11. Метод бісекцій.....	143
4.12. QL-метод знаходження власних значень симетричної тридіагональної матриці.....	147
Завдання до розділу 4.....	153

## Розділ 5. МЕТОДИ РОЗВ'ЯЗУВАННЯ ОДНОГО НЕЛІНІЙНОГО

РІВНЯННЯ З ОДНІЄЮ ЗМІННОЮ.....	158
5.1. Методи відокремлення коренів.....	158
5.1.1. Табуляція функції.....	159
5.1.2. Графічний спосіб.....	161
5.1.3. Аналітичний метод відділення коренів.....	162
5.2. Ітераційні методи розв'язування алгебричних і трансцендентних рівнянь.....	164
5.2.1. Метод дихотомії (половинного ділення).....	164
5.2.2. Метод ітерацій (послідовних наближень). Принцип стискальних відображень.....	166
5.2.3. Метод хорд.....	173
5.2.4. Метод дотичних (Ньютона).....	177
5.2.5. Комбінований метод хорд і дотичних.....	180
5.3. Ітерації вищих порядків.....	184
5.3.1. Метод Чебишова.....	185
5.3.2. Метод Ейткена.....	186
Завдання до розділу 5.....	193

## Розділ 6. ІТЕРАЦІЙНІ МЕТОДИ РОЗВ'ЯЗУВАННЯ СИСТЕМ НЕЛІНІЙНИХ РІВНЯНЬ

6.1. Формулювання задачі.....	197
6.2. Метод простої ітерації.....	200
6.3. Градієнтний метод розв'язування систем нелінійних рівнянь.....	202
6.4. Метод Ньютона та аспекти його практичної реалізації.....	204
6.5. Локальна збіжність методу Ньютона.....	206
6.6. Модифікації методу Ньютона.....	208
6.7. Методи зі скінченнорізницеви́ми похідними.....	209
6.8. Інші різни́цеві аналоги методу Ньютона.....	212
6.9. Квазіньютонівські методи для розв'язування систем нелінійних рівнянь.....	213
6.10. Числова апробація методів.....	216
6.11. Тестові завдання.....	224
Завдання до розділу 6.....	226

<b>Розділ 7. ПОЛІНОМІАЛЬНЕ ІНТЕРПОЛЮВАННЯ.....</b>	<b>228</b>
7.1. Інтерполяційний багаточлен Лагранжа .....	229
7.2. Схема Ейткена.....	231
7.3. Інтерполяційний багаточлен Ньютона .....	233
7.4. Інтерполяційний багаточлен Ерміта.....	237
7.5. Найліпше наближення і поліноми Чебишова, що найменше відхиляються від нуля .....	240
7.6. Обернене інтерполювання.....	242
<i>Завдання до розділу 7.....</i>	<i>245</i>
<b>Розділ 8. МЕТОД НАЙМЕНШИХ КВАДРАТІВ І НАЙЛІПШІ     СЕРЕДНЬОКВАДРАТИЧНІ НАБЛИЖЕННЯ.....</b>	<b>249</b>
8.1. Просте опрацювання емпіричних даних методом найменших квадратів .....	249
8.2. Середньоквадратичне наближення за допомогою тригонометричних багаточленів .....	255
<i>Завдання до розділу 8.....</i>	<i>259</i>
<b>Розділ 9. ІНТЕРПОЛЯЦІЙНІ СПЛАЙНИ .....</b>	<b>264</b>
9.1. Кусково-поліноміальна апроксимація.....	264
9.2. Інтерполяційний сплайн .....	266
9.3. Базисні сплайни.....	270
9.4. Ермітові (локальні) сплайни.....	273
<i>Завдання до розділу 9.....</i>	<i>277</i>
<b>Розділ 10. ЧИСЕЛЬНЕ ІНТЕГРУВАННЯ.....</b>	<b>278</b>
10.1. Задача чисельного інтегрування. Квадратурні формули інтерполяційного типу .....	278
10.2. Квадратурні формули Ньютона–Котеса.....	279
10.3. Принцип Рунге практичної оцінки похибки чисельного інтегрування .....	282
10.4. Квадратурна формула Гаусса .....	284
10.5. Квадратурні формули з ваговими функціями .....	287
10.6. Квадратурна формула Чебишова.....	288
10.7. Формули Ейлера–Маклорена .....	290
10.8. Обчислення невласних інтегралів .....	293
10.9. Наближене обчислення кратних інтегралів .....	302
10.9.1. Кубатурні формули.....	302
10.9.2. Метод Монте-Карло.....	304
<i>Завдання до розділу 10.....</i>	<i>307</i>
<b>Розділ 11. МЕТОДИ РОЗВ'ЯЗУВАННЯ ЗАДАЧІ КОШІ ДЛЯ     ЗВИЧАЙНИХ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ.....</b>	<b>315</b>
11.1. Формулювання задачі .....	315
11.2. Метод Ейлера та його модифікації .....	316
11.3. Методи Рунге–Кутта.....	317
11.4. Методи Адамса.....	320

11.5. Розв'язування задачі Коші для систем звичайних диференціальних рівнянь .....	325
11.6. Оцінка похибки наближеного розв'язку задачі Коші .....	327
<i>Завдання до розділу 11</i> .....	330
<b>Розділ 12. МЕТОДИ РОЗВ'ЯЗУВАННЯ ЛІНІЙНИХ КРАЙОВИХ ЗАДАЧ ДЛЯ ЗВИЧАЙНИХ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ</b> .....	341
12.1. Методи зведення крайової задачі до задачі Коші .....	342
12.1.1. Метод варіації сталих .....	342
12.1.2. Метод стрільби .....	343
12.1.3. Метод диференціальної прогонки .....	345
12.2. Сіткові методи .....	346
12.2.1 Метод сіток розв'язування крайових задач для звичайних диференціальних рівнянь .....	346
12.2.2. Метод різницевої прогонки .....	349
12.3. Варіаційні методи .....	351
12.3.1. Метод Рітца .....	351
12.3.2. Метод найменших квадратів .....	356
12.4. Проекційні методи .....	359
12.4.1. Метод колокації .....	359
12.4.2. Метод моментів. Метод Гальоркіна .....	362
<i>Завдання до розділу 12</i> .....	368
<b>Розділ 13. ЧИСЕЛЬНЕ РОЗВ'ЯЗУВАННЯ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ У ЧАСТИННИХ ПОХІДНИХ</b> .....	371
13.1. Метод сіток розв'язування задачі Діріхле для рівняння Пуассона .....	374
13.2. Метод сіток розв'язування змішаної задачі для рівняння теплопровідності .....	381
13.3. Метод сіток розв'язування змішаної задачі для рівняння коливань струни .....	389
<i>Завдання до розділу 13</i> .....	400
<b>Розділ 14. МЕТОДИ НАБЛИЖЕНОГО РОЗВ'ЯЗУВАННЯ ІНТЕГРАЛЬНИХ РІВНЯНЬ</b> .....	411
14.1. Формулювання задачі .....	411
14.2. Метод послідовних наближень розв'язування одновимірного рівняння другого роду .....	412
14.3. Наближене розв'язування інтегрального рівняння Фредгольма методом заміни інтеграла скінченною сумою (метод механічних квадратур) .....	415
14.4. Метод заміни ядра на вироджене .....	420
<i>Завдання до розділу 14</i> .....	426
<i>Список літератури</i> .....	430

## **Вступ**

*Математика як наука виникла з огляду на потребу розв'язування практичних задач: вимірювань на місцевості, навігації тощо. Тому вона була чисельною математикою, мета якої – отримання розв'язку у вигляді числа.*

*Визначні вчені минулого поєднували в працях побудову математичних моделей та їхнє дослідження. Аналіз складних моделей потребував створення нових, зазвичай, чисельних або чисельно-аналітичних методів розв'язування задач. Назви деяких із методів – методи Гаусса, Ньютона, Чебишова, Ейлера – свідчать про те, що їх розробляли видатні вчені.*

*Для останнього півстоліття характерний бурхливий розвиток обчислювальної техніки і теорії чисельних методів. У підсумку відбувається швидка зміна поглядів на весь комплекс питань, пов'язаних із застосуванням комп'ютерів, зокрема на вимоги до чисельних методів. Тому актуальним є видання посібників з розв'язування задач із обчислювальної математики з використанням комп'ютерів.*

*Ця книга написана на підставі досвіду викладання курсів "Чисельні методи лінійної алгебри" та "Чисельні методи" на факультетах прикладної математики та інформатики і механіко-математичному Львівського національного університету імені Івана Франка.*

*Кожен розділ починається з коротких теоретичних відомостей. Далі розглянуто типові задачі з*



детальними розв'язками. Завершують розділ вправи для самостійних занять та завдання для практичного виконання на комп'ютерах.

Автори щиро вдячні рецензентам за корисні зауваження, викладачам і студентам кафедри обчислювальної математики Львівського національного університету імені Івана Франка за допомогу, особливо доценту Б.А. Остудіну, який дав цінні поради.

## **Розділ 1**

# **УРАХУВАННЯ ПОХИБОК НАБЛИЖЕНИХ ОБЧИСЛЕНЬ**

## **1.1. Елементи теорії похибок**

Під похибкою розумітимемо величину, що характеризує точність результату. Похибки, що виникають під час розв'язування задачі, можна розділити на три групи:

- 1) похибка задачі (неусувна похибка);
- 2) похибка методу;
- 3) похибка заокруглень (похибка операцій).

Неусувна похибка є наслідком

а) неточності вхідних даних, що є в математичному описі задачі;

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

Похибка методу виникає тому, що для розв'язування математичної задачі доводиться використовувати наближені схеми, оскільки отримання точного розв'язку потребує необмеженої або неприйнятно великої кількості арифметичних операцій. У багатьох випадках це просто неможливо. Похибка заокруглень виникає в разі введення–виведення даних до комп'ютера та в ході виконання математичних операцій.

Основна задача теорії похибок – знаходження області невизначеності результату. Розрізняють такі задачі теорії похибок:

- пряма задача – визначення точності результату розв'язування задачі залежно від різних видів похибок;
- обернена задача – визначення того, з якою точністю треба взяти початкові дані, щоб неусувна похибка результату розв'язування задачі була менша від заданого значення.

Розглянемо процес заокруглення чисел. Якщо число  $x = 2,128492$  і його потрібно заокруглити до п'яти десяткових знаків після коми, то матимемо  $\tilde{x} = 2,12849$ . Тобто якщо старший розряд,

який відкидають, менше 5, то попередня цифра не змінюється. Якщо  $x = 2,128492$  потрібно заокруглити до чотирьох знаків після коми, то  $\bar{x} = 2,1285$ . Тобто якщо старший розряд, який відкидають, дорівнює або більше 5, то попередню цифру в числі збільшують на 1.

**Зауваження 1.1.** Іноколи вважають: якщо старший розряд, який відкидають, дорівнює 5, а попередня цифра парна, то її не змінюють. Якщо ж попередня цифра непарна, то її збільшують на одиницю.

Розглянемо приклади заокруглення чисел:

$$x = 4,6495843; \quad x = 237,472355;$$

$$\tilde{x} = 4,649584; \quad \hat{x} = 237,47236;$$

$$\bar{x} = 4,64958; \quad \bar{x} = 237,4724;$$

$$\tilde{x} = 4,6496; \quad \tilde{x} = 237,472;$$

$$\tilde{x} = 4,650; \quad \tilde{x} = 236,47$$

$$\tilde{x} = 4,65; \quad \tilde{x} = 237,5;$$

$$\tilde{x} = 4,6; \quad \tilde{x} = 238;$$

$$\tilde{x} = 5; \quad \tilde{x} = 2,4 \cdot 10^2;$$

$$\tilde{x} = 2 \cdot 10^3$$

## 1.2. Абсолютна та відносна похибки

Нехай  $x$  – точне значення деякої величини, а  $\bar{x}$  – її відоме наближене значення.

**Абсолютною похибкою** числа  $\bar{x}$  називають величину  $\Delta\bar{x} = |x - \bar{x}|$ . **Відносною похибкою** числа  $\bar{x}$  називають величину

$$\delta\bar{x} = \frac{\Delta\bar{x}}{|\bar{x}|}, \text{ якщо } x \neq 0. \text{ Оскільки в більшості випадків точне значення}$$

числа невідоме, то на практиці використовують граничні абсолютну  $\Delta_{\bar{x}}$  і відносну  $\delta_{\bar{x}}$  похибки, які дещо більші, ніж значення  $\Delta\bar{x}$  та  $\delta\bar{x}$ , відповідно.

Зазначимо, що точність результату ліпше характеризує відносна похибка. Інформацію про абсолютну та відносну похибки можна використати для такого відображення числа  $x$ :

$$x = \tilde{x} \pm \Delta\tilde{x},$$

$$x = \tilde{x}(1 \pm \delta\tilde{x}).$$

**Значущими цифрами числа** називають усі цифри в його записі, починаючи з першої ненульової цифри зліва.

Наприклад:

$x = 5,650123$  – усі цифри в записі цього числа значущі;

$x = 0,003487$  – значущі цифри тільки 3, 4, 8, 7;

$x = 0,02204900$  – значущі цифри 2, 2, 0, 4, 9, 0, 0 (два останні нулі в записі числа є значущими);

$x = 1\,230\,000$  – усі цифри значущі;

$x = 7,16 \cdot 10^6$  – значущі цифри тільки 7, 1, 6.

**Значущу цифру називають правильною (точною)**, якщо абсолютна похибка числа не перевищує половини одиниці розряду, що відповідає цій цифрі (або одиниці розряду, що відповідає цій цифрі).

**П р и к л а д 1.1.** Нехай  $\tilde{x} = 34,519$  і відомо, що  $\Delta\tilde{x} = 0,03$ . Скільки правильних значущих цифр має число  $\tilde{x}$ ?

*Розв'язування.* Маємо  $\Delta\tilde{x} = 0,03 = 0,3 \cdot 10^{-1} > 0,5 \cdot 10^{-2}$  і  $\Delta\tilde{x} = 0,3 \cdot 10^{-1} < 0,5 \cdot 10^{-1}$ . Отже, у числа  $\tilde{x}$  правильними будуть значущі цифри 3, 4, 5, а цифри 1 і 9 – сумнівні.

**П р и к л а д 1.2.** Нехай  $\tilde{x} = 9,535155$  і  $\Delta\tilde{x} = 2 \cdot 10^{-4}$ . Скільки правильних значущих цифр має число  $\tilde{x}$ ?

*Розв'язування.* Оскільки  $\Delta\tilde{x} = 2 \cdot 10^{-4} = 0,2 \cdot 10^{-3} < 0,5 \cdot 10^{-3}$  то  $\tilde{x}$  має правильні три значущі цифри після коми, тобто правильними будуть значущі цифри 9, 5, 3, 5.

**П р и к л а д 1.3.** Нехай  $\tilde{x} = 0,054289$  і  $\Delta\tilde{x} = 0,006$ . Скільки правильних значущих цифр має число  $\tilde{x}$ ?

*Розв'язування.* Маємо  $\Delta\tilde{x} = 0,6 \cdot 10^{-2} > 0,5 \cdot 10^{-2}$ . Отже, у числа  $\tilde{x}$  всі значущі цифри сумнівні.

### **1.3. Пряма задача теорії похибок**

Пряма задача теорії похибок полягає в оцінці точності результату розв'язування задачі залежно від похибок вхідних даних.

Нехай потрібно визначити абсолютну та відносну похибки функції  $y = f(x_1, x_2, \dots, x_n)$ , якщо задані похибки аргументів. Припустимо, що функція  $y$  є неперервно диференційовною в області невизначеності аргументів, тобто в області

$$\bar{x}_i - \Delta \bar{x}_i \leq x_i \leq \bar{x}_i + \Delta \bar{x}_i \quad (i = 1, 2, \dots, n),$$

де  $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n$  – наближені значення відповідних аргументів,  $\Delta \bar{x}_i$  – гранична абсолютна похибка наближеного значення  $\bar{x}_i$  ( $i = 1, 2, \dots, n$ ).

Нехай  $\Delta \bar{x}_i = |x_i - \bar{x}_i|$  ( $i = 1, 2, \dots, n$ ) – абсолютна похибка наближеного значення аргументу  $\bar{x}_i$ ;  $\Delta \tilde{y} = |y - \tilde{y}|$  – абсолютна похибка функції, де  $y$  – точне значення функції;  $\tilde{y}$  – її наближене значення. Тоді, використовуючи теорему Лагранжа про скінченний приріст, одержимо

$$y - \tilde{y} = f(x_1, x_2, \dots, x_n) - f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) = \sum_{i=1}^n \frac{\partial f(\xi_1, \xi_2, \dots, \xi_n)}{\partial x_i} (x_i - \bar{x}_i).$$

Отже, обмежуючись лінійною частиною приросту  $\Delta y$ , якою є повний диференціал  $dy$ , отримаємо

$$\begin{aligned} |dy| &= |df(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)| = \\ &= \left| \sum_{i=1}^n \frac{\partial f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)}{\partial x_i} \Delta \bar{x}_i \right| \leq \sum_{i=1}^n \left| \frac{\partial f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)}{\partial x_i} \right| \Delta \bar{x}_i. \end{aligned} \quad (1.1)$$

У практичних розрахунках, окрім оцінки (1.1), використовують оцінку

$$\Delta_{\bar{y}} = \sum_{i=1}^n \left| \frac{\partial f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)}{\partial x_i} \right| \Delta_{\bar{x}_i}, \quad (1.2)$$

яку називають лінійною оцінкою похибки. З оцінки (1.2) знайдемо відносну похибку

$$\delta_{\bar{y}} = \sum_{i=1}^n \left| \frac{\partial}{\partial x_i} \ln f(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n) \right| \cdot |\bar{x}_i| \delta_{\bar{x}_i}. \quad (1.3)$$

За допомогою формул (1.2) і (1.3) визначимо похибки результатів арифметичних операцій.

Нехай  $y = \pm x_1 \pm x_2 \pm \dots \pm x_n$ , тоді з (1.2) матимемо

$$\Delta_{\tilde{y}} = \sum_{i=1}^n \Delta_{\tilde{x}_i}, \quad (1.4)$$

а з (1.3), відповідно, –

$$\delta \tilde{y} \leq \delta^*, \quad (1.5)$$

де  $\delta^* = \max \delta \tilde{x}_i$ .

Тобто відносна похибка алгебричної суми  $n$  наближених чисел не перевищує максимальної відносної похибки доданків.

Аналогічно знайдемо похибки для інших арифметичних операцій.

Нехай тепер  $y = x_1 \cdot x_2 \cdot \dots \cdot x_n$ , де всі множники можемо вважати додатними. Оскільки  $\ln y = \ln x_1 + \dots + \ln x_n$  і  $\frac{\partial \ln y}{\partial x_i} = \frac{1}{x_i}$ , то, згідно з (1.3), одержимо

$$\delta_{\tilde{y}} = \sum_{i=1}^n \frac{1}{x_i} \Delta_{\tilde{x}_i} = \sum_{i=1}^n \delta_{\tilde{x}_i} \quad (1.6)$$

Якщо ж  $y = \frac{x_1}{x_2}$ , де  $x_1, x_2 > 0$ , то  $\ln y = \ln x_1 - \ln x_2$ ,  $\left| \frac{\partial \ln y}{\partial x_i} \right| = \frac{1}{x_i}$ .

Отже,

$$\delta_{\tilde{y}} = \frac{\Delta_{\tilde{x}_1}}{x_1} + \frac{\Delta_{\tilde{x}_2}}{x_2} = \delta_{\tilde{x}_1} + \delta_{\tilde{x}_2} \quad (1.7)$$

Рівності (1.6), (1.7) означають, що граничні відносні похибки під час множення та ділення наближених чисел додають.

Зазначимо, що для суми та різниці абсолютні похибки додають. Якщо два близькі числа віднімають, то відносна похибка результату може значно зрости. А в разі ділення на досить мале число може значно зрости абсолютна похибка.

Розглянемо деякі приклади.

**П р и к л а д 1.4.** Закруглюючи числа до трьох значущих цифр, визначити абсолютну та відносну похибки отриманих наближених чисел:

1) 0,5235; 2) 5,3244; 3) -154,82.

**Розв'язування:**

1)  $x = 0,5235$ . Заокруглення до трьох значущих цифр дає  $\tilde{x} = 0,524$ , тоді  $\Delta\tilde{x} = 0,0005 = 5 \cdot 10^{-4}$ , а відносна похибка

$$\delta\tilde{x} = 5 \cdot 10^{-4} / 0,524 \approx 0,95 \cdot 10^{-3};$$

2)  $x = 5,324$ . Тоді  $\tilde{x} = 5,32$ ,  $\Delta\tilde{x} = |x - \tilde{x}| = 0,004$ . Відповідно, відносна похибка  $\delta\tilde{x} = 4 \cdot 10^{-3} / 5,32 = 0,75 \cdot 10^{-3}$ ;

3)  $x = -154,82$ . Тоді  $\tilde{x} = -155$ ,  $\Delta\tilde{x} = 0,18$ , а  $\delta\tilde{x} = 0,18 / 155 = 1,2 \cdot 10^{-3}$ .

**П р и к л а д 1.5.** Визначити кількість правильних цифр у числі  $\tilde{x}$ , якщо відома його відносна похибка:

1)  $\tilde{x} = 23,924$ ,  $\delta\tilde{x} = 0,1$ ;

2)  $\tilde{x} = 9,4698$ ,  $\delta\tilde{x} = 0,1 \cdot 10^{-2}$ ;

3)  $\tilde{x} = 51782$ ,  $\delta\tilde{x} = 0,01$ .

**Розв'язування:**

1) обчислимо абсолютну похибку  $\Delta\tilde{x} = \tilde{x} \cdot \delta\tilde{x} = 2,3924$ . Тоді будемо мати, що в числі  $\tilde{x}$  правильною є тільки цифра 2, тобто одна цифра вірна;

2) обчислимо абсолютну похибку

$\Delta\tilde{x} = \tilde{x} \cdot \delta\tilde{x} = 9,4698 \cdot 0,1 \cdot 10^{-2} = 0,0094698$ . Тоді в числі  $\tilde{x}$  будуть вірними дві цифри – 9 та 4;

3) абсолютна похибка дорівнюватиме  $\Delta\tilde{x} = \tilde{x} \cdot \delta\tilde{x} = 51782 \cdot 0,01 = 517,82$ . Отже, в числі  $\tilde{x}$  будуть правильними дві цифри – 5 та 1.

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

**П р и к л а д 1.6.** На гіпотетичній обчислювальній машині з мантиєю довжини чотири знайти суму  $S = 0,2764 + 0,3944 + 1,475 + 26,46 + 1364$ , підсумовуючи від меншого доданка до більшого; від більшого доданка до меншого.

*Розв'язування:*

а) маємо  $S_2 = 0,2764 + 0,3944 = 0,6708$ ,  $S_3 = S_2 + 1,475$

Заокругливши  $S_2$  до трьох знаків після коми, матимемо

$$S_3 = 1,475 + 0,671 = 2,146.$$

Аналогічно далі

$$S_4 = S_3 + 26,46 = 2,15 + 26,46 = 28,61;$$

$$S = S_5 = S_4 + 1364 = 1393;$$

б) маємо  $S_2 = 1364 + 26,46 = 1364 + 26 = 1390$ ;

$$S_3 = S_2 + 1,475 = 1390 + 1 = 1391;$$

$$S_4 = S_3 + 0,3944 = 1391;$$

$$S = S_5 = S_4 + 0,2764 = 1391.$$

З урахуванням точного значення  $S = 1392,6058$  бачимо: підсумовування потрібно виконувати, починаючи з менших доданків. У протилежному випадку може бути значна втрата значущих цифр.

**П р и к л а д 1.7.** Нехай числа  $\sqrt{2,01} = 1,417744688$  та  $\sqrt{2} = 1,414213562$  задані з десятима правильними значущими цифрами. Скільки правильних значущих цифр матиме число  $\bar{x} = \sqrt{2,01} - \sqrt{2}$ ?

*Розв'язування.* Віднімемо від числа  $\sqrt{2,01}$  число  $\sqrt{2}$ . Отримаємо  $\bar{x} = 0,003531126$ . Позначимо  $\bar{x}_1 = 1,417744688$ ,  $\bar{x}_2 = 1,414213562$ . Тоді абсолютні похибки  $\Delta x_1 = \Delta x_2 = 0,5 \cdot 10^{-9}$ . Абсолютна похибка різниці  $\bar{x} = \bar{x}_1 - \bar{x}_2$  буде  $\Delta x = \Delta x_1 + \Delta x_2 = 10^{-9}$ . Оскільки абсолютна похибка числа задовольняє нерівність  $10^{-9} < 0,5 \cdot 10^{-8}$  то, згідно з означенням кількості точних значущих цифр, робимо висновок, що число  $\bar{x}$  має шість правильних значущих цифр 3, 5, 3, 1, 1, 2.



Те ж саме значення можна отримати, замінивши віднімання близьких чисел їхнім додаванням. Для цього запишемо  $\tilde{x}$  у вигляді

$$\tilde{x} = \frac{(\sqrt{2,01} - \sqrt{2})(\sqrt{2,01} + \sqrt{2})}{\sqrt{2,01} + \sqrt{2}} = \frac{0,01}{\sqrt{2,01} + \sqrt{2}},$$

причому достатньо взяти величини  $\tilde{x}_1$  і  $\tilde{x}_2$  з сімома правильними значущими цифрами.

В обчислювальних алгоритмах бажано уникати віднімання близьких за значенням чисел.

**Приклад 1.8.** Оцінити похибку обчислення функції

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

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

*Розв'язування.* Згідно з формулою (1.2), для абсолютної похибки результату отримаємо

$$\begin{aligned} \Delta \tilde{f} &= \left| \frac{2\tilde{x}}{\tilde{z}} \right| \Delta \tilde{x} + \left| \frac{1}{\tilde{z}} \right| \Delta \tilde{y} + \left| -\frac{\tilde{x}^2 + \tilde{y}}{\tilde{z}^2} \right| \Delta \tilde{z} = \frac{2 \cdot 0,15}{1,14} \cdot 0,005 + \\ &+ \frac{1}{1,14} \cdot 0,01 + \frac{(0,15)^2 + 2,13}{(1,14)^2} \cdot 0,007 = 0,00131579 + 0,00877193 + \\ &+ 0,01159395 \approx 0,02168167 = 2,2 \cdot 10^{-2}. \end{aligned}$$

$$\text{Знайдемо } f(\tilde{x}, \tilde{y}, \tilde{z}) = \frac{(0,15)^2 + 2,13}{1,14} = 1,88815789.$$

$$\text{Тоді } \delta \tilde{f} = \frac{2,2 \cdot 10^{-2}}{1,88815789} = 0,01165157$$

**Приклад 1.9.** Висота  $h$  та радіус основи циліндра  $R$  виміряні з точністю до 0,5 %. Яка відносна похибка в разі обчислення об'єму циліндра, якщо  $\tilde{\pi} = 3,14$ ?

**Розв'язування.**  $V = \pi R^2 h$ . Точніше значення  $\pi = 3,14159265$ , отже,  $\Delta\pi = 0,16 \cdot 10^{-2}$ , а

$$\delta\pi = \frac{0,16 \cdot 10^{-2}}{3,14} = 0,0005 = 0.05 \%$$

Тоді, згідно з формулою про відносну похибку добутку, матимемо

$$\delta\bar{V} \approx \delta\pi + 2\delta\bar{R} + \delta\bar{h} = 1,55 \%$$

**Приклад 1.10.** Ребро куба, виміряне з точністю до 0,02 см, дорівнює 8 см. Знайти абсолютну та відносну похибки в разі обчислення об'єму куба.

**Розв'язування.** Позначимо сторону куба через  $a$ . Тоді  $V = a^3$   
 $V = a^3 = 512 \text{ см}^3$ . Застосуємо формулу (1.2), матимемо  
 $\Delta\tilde{V} = 3\bar{a}^2 \cdot \Delta\bar{a} = 3 \cdot 8^2 \cdot 0,02 = 3,84$ , а  $\delta\tilde{V} = 3,84 / 512 = 0,0075$

**Приклад 1.11.** Визначити відносну похибку числа, що записане на комп'ютері з системою числення  $\beta$  та довжиною мантиси  $\ell$ .

**Розв'язування.** Число  $\tilde{x}$  можна записати на комп'ютері у вигляді

$$\tilde{x} = \pm(d_1\beta^{-1} + d_2\beta^{-2} + \dots + d_\ell\beta^{-\ell})\beta^\ell,$$

де  $\ell$  визначає порядок числа,  $d_i$  – цілі, причому  $0 \leq d_i \leq \beta - 1$ ,  $d_1 \neq 0$ . Нехай точне значення числа

$$x = \pm(d_1\beta^{-1} + d_2\beta^{-2} + \dots + d_\ell\beta^{-\ell} + d_{\ell+1}\beta^{-(\ell+1)})\beta^\ell$$

Тоді

$$\frac{|\tilde{x} - x|}{|\tilde{x}|} = \frac{d_{\ell+1}\beta^{-(\ell+1)}}{|\tilde{x}|} = \frac{d_{\ell+1}}{|d_1\beta^\ell + d_2\beta^{\ell-1} + \dots + d_\ell\beta|} \leq \frac{d_{\ell+1}}{d_1\beta^\ell} \leq \frac{d_{\ell+1}}{\beta^\ell} \leq \beta^{1-\ell}$$

Отже,  $\delta\tilde{x} \leq \beta^{1-\ell}$ . Якщо ж числа вводять за правилами заокруглення, то  $d_{\ell+1} \leq 0,5\beta$  і тоді матимемо

$$\delta\tilde{x} \leq \frac{1}{2}\beta^{1-\ell}$$

## 1.4. Обернена задача теорії похибок

Обернена задача теорії похибок полягає в такому: з якою точністю потрібно задати значення аргументів  $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n$  функції  $f(x_1, x_2, \dots, x_n)$ , щоб на виході отримати результат заданої точності?

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

$$\Delta \tilde{x} = \frac{\Delta \tilde{y}}{|f'(\tilde{x})|}, \quad f'(\tilde{x}) \neq 0 \quad (1.8)$$

Для функції декількох змінних  $y = f(x_1, x_2, \dots, x_n)$  задачу розв'язують з використанням таких рекомендацій.

- А. Принцип однакових впливів, тобто вважаємо, що всі доданки  $c_i = \left| \partial f / \partial x_i \right| \Delta \tilde{x}_i$ ,  $i = \overline{1, n}$  рівні між собою. Тоді абсолютні похибки всіх аргументів визначені формулами

$$\Delta \tilde{x}_i = \frac{\Delta \tilde{y}}{n \left| \frac{\partial f}{\partial x_i} \right|}, \quad i = \overline{1, n}. \quad (1.9)$$

- Б. Уважаємо, що всі похибки дорівнюють максимально можливій похибці аргументів, тобто приймаємо

$$\Delta \tilde{x}_1 = \Delta \tilde{x}_2 = \dots = \Delta \tilde{x}_n = \delta,$$

де

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

**П р и к л а д 1.12.** Сторона квадрата дорівнює 2 м. З якою точністю її потрібно виміряти, щоб похибка обчислення площі не перевищувала  $0,5 \text{ см}^2$ ?

*Розв'язування.* Позначимо сторону квадрата через  $x$ ;  $S = x^2$ ,  $S' = 2x$ . Тоді за формулою (1.8) отримаємо

$$\Delta \tilde{x} = \frac{0,5}{2 \cdot 200} = \frac{1}{8} 10^{-2} \text{ см}.$$

**Приклад 1.13.** З якою кількістю правильних значущих цифр потрібно взяти вільний член квадратного рівняння  $x^2 - 2x + \lg 2 = 0$ , щоб отримати корені рівняння з чотирма правильними значущими цифрами?

*Розв'язування.* Для коренів рівняння маємо  $x_1 = 1 + \sqrt{1 - \lg 2}$ ,  $x_2 = 1 - \sqrt{1 - \lg 2}$ . Оскільки  $\lg 2 \approx 0,3\dots$ , то  $x_1 = 1,8\dots$ ,  $x_2 = 0,1\dots$ . Отже, за змістом задачі  $\tilde{x}_1$  потрібно визначити так, щоб  $\Delta \tilde{x}_1 \leq 0,5 \cdot 10^{-3}$ , а для  $\tilde{x}_2$ , – щоб  $\Delta \tilde{x}_2 \leq 0,5 \cdot 10^{-4}$ . Позначимо  $z = \lg 2$  і розглянемо функцію  $f(z) = 1 + \sqrt{1 - z}$ . З'ясуємо, з якою точністю потрібно обчислити  $\tilde{z}$  в околі точки 0,3. Оскільки  $f'(z) = -\frac{1}{2\sqrt{1-z}}$ , то, використовуючи формулу (1.8), отримуємо

$$\Delta \tilde{z} \leq 0,5 \cdot 10^{-3} \cdot \frac{1}{2\sqrt{0,7}} = 0,000299.$$

Звідси робимо висновок, що для знаходження кореня  $x_1$  потрібно обчислити  $\lg 2$  з трьома правильними значущими цифрами після коми, тобто  $\lg 2 = 0,301$ .

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  $\lg 2$  з чотирма правильними значущими цифрами після коми, тобто  $\lg 2 = 0,3010$ .

**Приклад 1.14.** У п'ятизначних логарифмічних таблицях наведено значення десяткових логарифмів з точністю до  $\varepsilon = 0,5 \cdot 10^{-6}$ . Оцінити можливу похибку в разі знаходження числа за його логарифмом, якщо саме число є між 300 та 400.

*Розв'язування.* Позначимо  $y = \lg x$ ,  $x \in [100; 200]$ . За умовою задачі  $\Delta \tilde{y} \leq 0,5 \cdot 10^{-6}$  і потрібно знайти  $\Delta \tilde{x}$ . Маємо  $y' = \frac{1}{x \ln 10}$ . Тоді за формулою (1.8) матимемо

$$\Delta \tilde{x} = x(\ln 10) \cdot \Delta \tilde{y} \leq 200 \cdot 2,30 \cdot 0,5 \cdot 10^{-6} = 0,00023.$$

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

## Завдання до розділу 1

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

5,1634; 0,34359; 0,01306; 1,335; -0,0015993; 0,002811; 315,55;  
81,543.

1.2. Визначити кількість правильних цифр у числі  $x$ , якщо задано його відносну похибку:

1)  $x = 4,7892$ ,  $\delta\bar{x} = 0,1 \cdot 10^{-2}$ ;

2)  $x = 23,9470$ ,  $\delta\bar{x} = 1\%$ ;

3)  $x = 0,2239$ ,  $\delta\bar{x} = 0,2 \cdot 10^{-1}$ ;

4)  $x = 483,9$ ,  $\delta\bar{x} = 2\%$ ;

5)  $x = 43,874$ ,  $\delta\bar{x} = 0,1$ ;

6)  $x = 28164$ ,  $\delta\bar{x} = 1\%$ ;

7)  $x = 0,01595$ ,  $\delta\bar{x} = 0,5 \cdot 10^{-2}$

8)  $x = 0,44257$ ,  $\delta\bar{x} = 10\%$ ;

9)  $x = 0,000115$ ,  $\delta\bar{x} = 0,15$ ;

10)  $x = 6,5376$ ,  $\delta\bar{x} = 0,1\%$ .

1.3. 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)  $49/13 = 3,77$ ,  $\sqrt{14} = 3,74$ .

1.4. Якою буде відносна похибка, якщо число  $\pi$  наблизити числом 3,14?

1.5. Записати число  $\pi$  з сімома правильними значущими цифрами та визначити відносну похибку отриманого наближення.

1.6. Знайти  $\sqrt{5,03} - \sqrt{5}$  з трьома правильними значущими цифрами.

1.7. Під час вимірювання радіуса круга з точністю до 0,5 см отримали число 14 см. Знайти абсолютну та відносну похибки під час обчислення площі круга.

1.8. Виявилось, що кожне ребро куба, виміряне з точністю 0,02 см, дорівнює 15 см. Знайти абсолютну та відносну похибки в разі обчислення площі куба.

1.9. 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, такі:  $R = 23,64$  см,  $r = 17,31$  см,  $l = 10,21$  см.

1.10. Обчислити значення функції  $f$ . Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних правильними:

$$f = x_1 + x_2.$$

де

- 1)  $x_1 = 5,39$ ,  $x_2 = 6,7$
- 2)  $x_1 = 14,2$ ,  $x_2 = 1,432$ ;
- 3)  $x_1 = 0,02$ ,  $x_2 = 10,3$ .

1.11. Обчислити значення функції  $f$ . Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних правильними:

$$f = x_1 - x_2 - x_3.$$

де

- 1)  $x_1 = 324,45$ ,  $x_2 = 1287$ ,  $x_3 = 0,0056$ ;
- 2)  $x_1 = 0,431$ ,  $x_2 = 432$ ,  $x_3 = 87,4$ ;
- 3)  $x_1 = 2,48$ ,  $x_2 = 7,4$ ,  $x_3 = 3,136$ .

1.12. Обчислити значення функції  $f$ . Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних правильними:

$$f = x_1 x_2 + x_2 x_3.$$

де  $x_1 = 2,305$ ,  $x_2 = 2,873$ ,  $x_3 = 0,875$

1.13. Обчислити значення функції  $f$ . Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних правильними:

$$f = x_1 / x_2$$

- 1)  $x_1 = 623,685, x_2 = 453$ ;
- 2)  $x_1 = 745,8371, x_2 = 336,2$ ;
- 3)  $x_1 = 10,5, x_2 = 554$ ;
- 4)  $x_1 = 5,684, x_2 = 5,032$ .

1.14. Обчислити значення функції  $f$ . Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних правильними:

$$f = \ln(3x_1 + 2x_2^2),$$

де  $x_1 = 0,35, x_2 = 2,324$ .

1.15. Обчислити значення функції  $f$ . Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних правильними:

$$f = \frac{x_1 x_3 + x_2}{x_1},$$

де  $x_1 = 4,43, x_2 = 0,752, x_3 = 2,321$ .

1.16. Оцінити абсолютну та відносну похибки обчислення функції:

$$1) \quad f(x, y, z) = \ln\left(\frac{x}{z} + xy\right),$$

при  $x = 1,54 \pm 0,02, y = 1,34 \pm 0,03, z = 1,02 \pm 0,02$ ;

$$2) \quad f(x, y, z) = \sqrt{\frac{x}{z}} + \frac{y}{z},$$

при  $x = 5,8 \pm 0,01, y = 0,65 \pm 0,02, z = 1,1753 \pm 0,0002$ ;

$$3) \quad f(x, y, z) = \frac{\sqrt{yx} + yz}{\sqrt[3]{z}},$$

при  $x = 5 \pm 0,1, y = 2 \pm 0,03, z = 1 \pm 0,07$ ;

$$4) \quad f(x, y, z) = \ln\left(xy + \sqrt{\frac{z}{x}}\right),$$

при  $x = 2,03 \pm 0,02, y = 2,48 \pm 0,02, z = 3,04 \pm 0,03$ ;

$$5) \quad f(x, y, z) = \frac{(z + y)(2x - 1)^2}{(x - y)^2},$$

при  $x = 5,7 \pm 0,01, y = 0,78 \pm 0,02, z = 1,3253 \pm 0,0001$ ;

$$6) \quad f(x, y, z) = \frac{(x^2 + 4xy + y^2)(2z - 1)^2}{(x + y)^2} \cdot \frac{z^2}{18},$$

при  $x = 37,64 \pm 0,002$ ,  $y = 43,62 \pm 0,003$ ,  $z = 41,5 \pm 0,07$ ;

$$7) \quad f(x, y) = \frac{1}{64} \pi \sqrt{x^3 - y^2}$$

при  $x = 46,2 \pm 0,02$ ,  $y = 23,48 \pm 0,003$ ,  $\pi = 3,14$ .

1.17. Знайти межі абсолютної та відносної похибки аргументів, які дають змогу обчислити з чотирма правильними знаками функцію

$$f = \frac{\sqrt{x_1} + 4x_2^2}{2},$$

де  $x_1 = 1,20326$ ,  $x_2 = 1,83241$ ,  $x_3 = 0,74521$ .

1.18. Оцінити похибку в визначенні кута  $x = 45^\circ$  за п'ятизначною таблицею синусів.

1.19. З якою кількістю правильних значущих цифр потрібно взяти значення аргументу  $x$ , щоб обчислити значення функції  $f(x) = \sin x^3 \cdot \lg x$  з точністю до  $0,2 \cdot 10^{-5}$ ?

1.20. З якою точністю потрібно обчислити  $\cos \frac{\pi}{7}$ , щоб відносна похибка обчислення коренів рівняння

$$x^2 - 4x + \cos \frac{\pi}{7} = 0$$

не перевищувала  $10^{-3}$ ?

1.21. З якою відносною похибкою треба виміряти висоту  $h = 0,7$  м та радіус основи  $r = 8$  м для того, щоб відносна похибка обчислення об'єму конуса не перевищувала  $0,05\%$ ?



## Розділ 2

### ПРЯМІ МЕТОДИ РОЗВ'ЯЗУВАННЯ СИСТЕМ ЛІНІЙНИХ АЛГЕБРИЧНИХ РІВНЯНЬ

#### 2.1. Метод Гаусса розв'язування систем лінійних алгебричних рівнянь

Розглянемо систему лінійних алгебричних рівнянь (СЛАР) вигляду

$$Ax = b \quad (2.1)$$

Метод Гаусса, або метод послідовного вилучення невідомих, розв'язування системи (2.1) складається з двох ходів: прямого і зворотного. Внаслідок прямого ходу систему (2.1) зводимо до системи з трикутною матрицею за допомогою таких послідовних перетворень  $k = 1, 2, \dots, n-1$

$$m_{ik} = -\frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}, \quad b_i^{(k)} = b_i^{(k-1)} + m_{ik}b_k^{(k-1)}, \quad i = k+1, \dots, n; \quad (2.2)$$

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} + m_{ik}a_{kj}^{(k-1)}, \quad i, j = k+1, \dots, n; \quad (2.3)$$

$$a_{ii}^{(0)} = a_{ii}, \quad b_i^{(0)} = b_i, \quad i, j = 1, \dots, n.$$

На зворотньому ході методу знаходимо розв'язки

$$x_n = \frac{b_n^{(n)}}{a_{nn}^{(n)}}; \quad (2.4)$$

$$x_k = \frac{b_k - \sum_{j=k+1}^n a_{kj}^{(k)} x_j}{a_{kk}^{(k)}}, \quad k = n-1, \dots, 2, 1. \quad (2.5)$$

Визначник матриці  $A$

$$\det A = \prod_{k=1}^n a_{kk}^{(k-1)} \quad (2.6)$$

Для отримання матриці  $A^{-1}$  оберненої до  $A = (a_{i,j})_{i,j=1}^n$ , враховуємо те, що вона є розв'язком матричного рівняння

$$AX = E, \quad (2.7)$$

де  $E$  – одинична матриця;  $X = A^{-1}$   $X = (x_{i,j})_{i,j=1}^n$  – шукана обернена матриця.

Якщо зобразимо матрицю  $X$  як набір векторів-стовпців

$$x_1 = \begin{pmatrix} x_{1,1} \\ x_{2,1} \\ \vdots \\ x_{n,1} \end{pmatrix}, x_2 = \begin{pmatrix} x_{1,2} \\ x_{2,2} \\ \vdots \\ x_{n,2} \end{pmatrix}, \dots, x_n = \begin{pmatrix} x_{1,n} \\ x_{2,n} \\ \vdots \\ x_{n,n} \end{pmatrix},$$

а одиничну матрицю  $E$  як набір одиничних векторів

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

то матричне рівняння (2.7) запишемо як  $n$  систем лінійних алгебричних рівнянь

$$\begin{cases} Ax_1 = e_1; \\ Ax_2 = e_2; \\ \dots\dots\dots \\ Ax_n = e_n \end{cases} \quad (2.8)$$

Специфікою (2.8) є те, що всі СЛАР мають одну й ту ж матрицю коефіцієнтів, що дає змогу прямий хід методу Гаусса виконати лише один раз.

**Приклад 2.1.****А. Методом Гаусса розв'язати систему**

$$\begin{pmatrix} [2] & 2 & -1 & 1 \\ 4 & 3 & -1 & 2 \\ 8 & 5 & -3 & 4 \\ 3 & 3 & -2 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ 12 \\ 6 \end{pmatrix}.$$

*Розв'язування:* Спочатку виконаємо прямий хід методу Гаусса за чотири кроки згідно з формулами (2.2), (2.3).

**Крок 1:**

$$\left( \begin{array}{cccc|c} 1 & 1 & -0,5 & 0,5 & 2 \\ 4 & 3 & -1 & 2 & 6 \\ 8 & 5 & -3 & 4 & 12 \\ 3 & 3 & -2 & 2 & 6 \end{array} \right) \sim \left( \begin{array}{cccc|c} 1 & 1 & -0,5 & 0,5 & 2 \\ 0 & [-1] & 1 & 0 & -2 \\ 0 & -3 & 1 & 0 & -4 \\ 0 & 0 & -0,5 & 0,5 & 0 \end{array} \right)$$

**Крок 2:**

$$\left( \begin{array}{cccc|c} 1 & 1 & -0,5 & 0,5 & 2 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & -3 & 1 & 0 & -4 \\ 0 & 0 & -0,5 & 0,5 & 0 \end{array} \right) \sim \left( \begin{array}{cccc|c} 1 & 1 & -0,5 & 0,5 & 2 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & [-2] & 0 & 2 \\ 0 & 0 & -0,5 & 0,5 & 0 \end{array} \right) \sim$$

**Крок 3:**

$$\left( \begin{array}{cccc|c} 1 & 1 & -0,5 & 0,5 & 2 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & -0,5 & 0,5 & 0 \end{array} \right) \sim \left( \begin{array}{cccc|c} 1 & 1 & -0,5 & 0,5 & 2 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & [0,5] & -0,5 \end{array} \right) \sim$$

**Крок 4:**

$$\sim \left( \begin{array}{cccc|c} 1 & 1 & -0,5 & 0,5 & 2 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \end{array} \right).$$

Ми отримали систему з верхньою трикутною матрицею, яка містить одиниці на головній діагоналі. На зворотному ході методу Гаусса за формулами (2.4), (2.5) знаходимо розв'язки заданої системи:

$$x_4 = -1, \quad x_3 = -1, \quad x_2 = 2 + x_3 = 1, \quad x_1 = 2 - x_2 + 0,5x_3 - 0,5x_4 = 1.$$

Отже, розв'язок системи

$$x = (1, 1, -1, -1)^T$$

**Б.** За формулою (2.6) легко обчислити визначник матриці  $A$ :

$$\det A = a_{11} \cdot a_{22}^{(1)} \cdot a_{33}^{(2)} \cdot a_{44}^{(3)} = 2 \cdot (-1) \cdot (-2) \cdot 0,5 = 2.$$

**В.** Для знаходження оберненої матриці  $A^{-1}$  за формулами (2.8) прямий хід методу Гаусса виконуємо тільки один раз.

**Крок 1:**

$$\left( \begin{array}{cccc|cccc} [2] & 2 & -1 & 1 & 1 & 0 & 0 & 0 \\ 4 & 3 & -1 & 2 & 0 & 1 & 0 & 0 \\ 8 & 5 & -3 & 4 & 0 & 0 & 1 & 0 \\ 3 & 3 & -2 & 2 & 0 & 0 & 0 & 1 \end{array} \right) \sim \left( \begin{array}{cccc|cccc} 1 & 1 & -0,5 & 0,5 & 0,5 & 0 & 0 & 0 \\ 4 & 3 & -1 & 2 & 0 & 1 & 0 & 0 \\ 8 & 5 & -3 & 4 & 0 & 0 & 1 & 0 \\ 3 & 3 & -2 & 2 & 0 & 0 & 0 & 1 \end{array} \right) \sim$$

$$\left( \begin{array}{cccc|cccc} 1 & 1 & -0,5 & 0,5 & 0,5 & 0 & 0 & 0 \\ 0 & [-1] & 1 & 0 & -2 & 1 & 0 & 0 \\ 0 & -3 & 1 & 0 & -4 & 0 & 1 & 0 \\ 0 & 0 & -0,5 & 0,5 & -1,5 & 0 & 0 & 1 \end{array} \right) \sim$$

**Крок 2:**

$$\sim \left( \begin{array}{cccc|cccc} 1 & 1 & -0,5 & 0,5 & 0,5 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 2 & -1 & 0 & 0 \\ 0 & -3 & 1 & 0 & -4 & 0 & 1 & 0 \\ 0 & 0 & -0,5 & 0,5 & -1,5 & 0 & 0 & 1 \end{array} \right) \sim$$

$$\sim \left( \begin{array}{cccc|ccc} 1 & 1 & -0,5 & 0,5 & 0,5 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 2 & -1 & 0 & 0 \\ 0 & 0 & [-2] & 0 & 2 & -3 & 1 & 0 \\ 0 & 0 & -0,5 & 0,5 & -1,5 & 0 & 0 & 1 \end{array} \right) \sim$$

**Крок 3:**

$$\sim \left( \begin{array}{cccc|ccc} 1 & 1 & -0,5 & 0,5 & 0,5 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 2 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1,5 & -0,5 & 0 \\ 0 & 0 & -0,5 & 0,5 & -1,5 & 0 & 0 & 1 \end{array} \right) \sim$$

$$\sim \left( \begin{array}{cccc|ccc} 1 & 1 & -0,5 & 0,5 & 0,5 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 2 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1,5 & -0,5 & 0 \\ 0 & 0 & 0 & [0,5] & -2 & 0,75 & -0,25 & 1 \end{array} \right) \sim$$

**Крок 4:**

$$\sim \left( \begin{array}{cccc|ccc} 1 & 1 & -0,5 & 0,5 & 0,5 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 2 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1,5 & -0,5 & 0 \\ 0 & 0 & 0 & 1 & -4 & 1,5 & -0,5 & 2 \end{array} \right).$$

Згідно зі зворотним ходом методу Гаусса, за формулами (2.4), (2.5) знаходимо розв'язки чотирьох систем:

$$x_{4,1} = -4; \quad x_{4,2} = 1,5; \quad x_{4,3} = -0,5; \quad x_{4,4} = 2;$$

$$x_{3,1} = -1; \quad x_{3,2} = 1,5; \quad x_{3,3} = -0,5; \quad x_{3,4} = 0;$$

$$x_{2,1} = 1; \quad x_{2,2} = 0,5; \quad x_{2,3} = -0,5; \quad x_{2,4} = 0;$$

$$x_{1,1} = 1; \quad x_{1,2} = -0,5; \quad x_{1,3} = 0,5; \quad x_{1,4} = -1.$$

Отже, обернена матриця має вигляд

$$A^{-1} = \begin{pmatrix} 1 & -0,5 & 0,5 & -1 \\ 1 & 0,5 & -0,5 & 0 \\ -1 & 1,5 & -0,5 & 0 \\ -4 & 1,5 & -0,5 & 2 \end{pmatrix}.$$

Виконаємо перевірку  $A \cdot A^{-1} = E$ , тобто

$$A \cdot A^{-1} = \begin{pmatrix} 2 & 2 & -1 & 1 \\ 4 & 3 & -1 & 2 \\ 8 & 5 & -3 & 4 \\ 3 & 3 & -2 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & -0,5 & 0,5 & -1 \\ 1 & 0,5 & -0,5 & 0 \\ -1 & 1,5 & -0,5 & 0 \\ -4 & 1,5 & -0,5 & 2 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

## 2.2. Розв'язування систем лінійних алгебричних рівнянь за допомогою LU-розкладу

Розглянемо СЛАР вигляду (2.1).

**Теорема.** Нехай усі головні мінори матриці  $A$  відмінні від нуля,  $\Delta_j \neq 0$ ,  $j = 1, 2, \dots, n$ . Тоді матрицю  $A$  можна записати, причому єдиним способом, у вигляді добутку

$$A = LU, \quad (2.9)$$

де  $L$  – нижня трикутна матриця з одиничними діагональними коефіцієнтами;  $U$  – верхня трикутна матриця з ненульовими діагональними елементами.

**Зауваження 2.2.1.** Якщо хоча б один з головних мінорів матриці  $A$  дорівнює нулю, то описаний  $LU$ -розклад неможливий.

Перепишемо (2.9) у вигляді

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{n1} & l_{n2} & l_{n3} & 1 \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{12} & u_{1n} \\ 0 & u_{22} & u_{2n} \\ 0 & 0 & u_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{1n} \\ a_{21} & a_{22} & a_{2n} \\ a_{n1} & a_{n2} & a_{nn} \end{pmatrix}. \quad (2.10)$$

Перемножимо матриці в (2.10) і прирівняємо поелементно ліву і праву частини рівності, отримаємо формули для визначення  $l_{ij}$  (при  $i > j$ ) та  $u_{ij}$  (при  $i \leq j$ ):

$$u_{1j} = a_{1j} \quad (j = 1, \dots, n);$$

$$l_{i1} = \frac{a_{i1}}{a_{11}} \quad (i = 2, \dots, n);$$

$$u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj} \quad (i \leq j); \quad (2.11)$$

$$l_{ij} = \frac{1}{u_{jj}} \left( a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj} \right) \quad (i > j).$$

Підставимо розклад (2.9) у систему (2.1):

$$LUx = b \quad (2.12)$$

Одержимо такі дві системи:

$$Ly = b; \quad (2.13)$$

$$Ux = y. \quad (2.14)$$

Спочатку розв'яжемо систему (2.13):

$$y_i = b_i - \sum_{k=1}^{i-1} l_{ik} y_k \quad (i = 1, 2, \dots, n), \quad (2.15)$$

Далі з системи (2.14) знайдемо шуканий розв'язок

$$x_i = \frac{1}{u_{ii}} \left( y_i - \sum_{k=i+1}^n u_{ik} x_k \right) \quad (i = n, n-1, \dots, 1). \quad (2.16)$$

Зазначимо, що обчислення визначника  $LU$ -факторизованої матриці  $A$  зводиться до перемноження  $n$  чисел

$$\det A = \det L \cdot \det U = u_{11} \cdot u_{22} \cdot \dots \cdot u_{nn}. \quad (2.17)$$

Для обертання  $LU$ -факторизованої матриці позначимо шукані елементи матриці  $A^{-1}$  через  $x_{ij}$ , які визначимо за формулами

$$x_{jj} = \frac{1}{u_{jj}} \left( 1 - \sum_{k=j+1}^n u_{jk} x_{kj} \right); \quad (2.18)$$

$$x_{ij} = -\frac{1}{u_{ii}} \sum_{k=i+1}^n u_{ik} x_{kj} \quad (i < j); \quad (2.19)$$

$$x_{ij} = -\sum_{k=j+1}^n x_{ik} l_{kj} \quad (i > j). \quad (2.20)$$

## Приклад 2.2.

**А.** Виконати  $LU$ -розклад матриці та розв'язати за його допомогою систему рівнянь  $Ax = b$ , де

$$A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 2 & 1 \\ -1 & 5 & 3 \end{pmatrix}, \quad b = (1, 5, 12)^T$$

*Розв'язування.*

1. Виконуємо  $LU$ -розклад:

$$\begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 2 & 1 \\ -1 & 5 & 3 \end{pmatrix}.$$

Поступово знаходимо елементи  $l_{ij}$  та  $u_{ij}$ :

$$u_{11} = -1; \quad u_{12} = 1; \quad u_{13} = 0;$$



$$l_{21} = \frac{0}{-1} = 0; \quad l_{31} = \frac{-1}{-1} = 1;$$

$$u_{22} = a_{22} - l_{21} \cdot u_{12} = 2 - 0 \cdot 1 = 2;$$

$$u_{23} = a_{23} - l_{21} \cdot u_{13} = 1 - 0 \cdot 0 = 1;$$

$$l_{32} = \frac{1}{u_{22}} \cdot (a_{32} - l_{31} \cdot u_{12}) = \frac{1}{2} (5 - 1 \cdot 1) = 2;$$

$$u_{33} = a_{33} - l_{31} \cdot u_{13} - l_{32} \cdot u_{23} = 3 - 1 \cdot 0 - 2 \cdot 1 = 1.$$

Одержано такі  $L$ - та  $U$ -матриці:

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}; \quad U = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. Переходимо до розв'язування систем (2.13) та (2.14):

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 12 \end{pmatrix}.$$

Звідси  $y_1 = 1$ ;  $y_2 = 5$ ;  $y_3 = 12 - 1 - 2 \cdot 5 = 1$ ;

$$\begin{pmatrix} -1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 1 \end{pmatrix}.$$

Звідси  $x_1 = 1$ ;  $x_2 = \frac{1}{2} (5 - 1) = 2$ ;  $x_3 = \frac{1}{-1} (1 - 1 \cdot 2) = 1$ .

Отже, шуканий розв'язок заданої системи

$$x = (1, 2, 1)^T$$

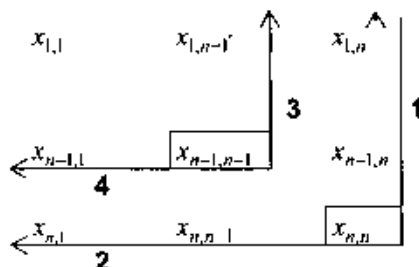
**Б.** Знайти визначник заданої матриці  $A$ .

*Розв'язування.* За формулою (2.17) маємо

$$\det A = u_{11} \cdot u_{22} \cdot u_{33} = (-1) \cdot 2 \cdot 1 = -2.$$

В. Обчислити обернену матрицю  $A^{-1}$

*Розв'язування.* На підставі формул (2.18)–(2.20) послідовність обчислень виконуємо за схемою



$$x_{3,3} = 1;$$

$$x_{2,3} = -\frac{1}{u_{2,2}} \cdot u_{2,3} \cdot x_{3,3} = -\frac{1}{2} \cdot 1 \cdot 1 = -0,5;$$

$$x_{1,3} = -\frac{1}{u_{1,1}} (u_{1,2} \cdot x_{2,3} + u_{1,3} \cdot x_{3,3}) = -\frac{1}{(-1)} (1 \cdot (-0,5) + 0 \cdot 1) = -0,5;$$

$$x_{3,2} = -x_{3,3} \cdot l_{3,2} = -1 \cdot 2 = -2;$$

$$x_{3,1} = -(x_{3,2} \cdot l_{2,1} + x_{3,3} \cdot l_{3,1}) = -((-2) \cdot 0 + 1 \cdot 1) = -1;$$

$$x_{2,2} = \frac{1}{u_{2,2}} (1 - u_{2,3} \cdot x_{3,2}) = \frac{1}{2} (1 - 1 \cdot (-2)) = 1,5;$$

$$x_{1,2} = -\frac{1}{u_{1,1}} (u_{1,2} \cdot x_{2,2} + u_{1,3} \cdot x_{3,2}) = -\frac{1}{(-1)} (1 \cdot 1,5 + 0 \cdot (-2)) = 1,5;$$

$$x_{2,1} = -(x_{2,2} \cdot l_{2,1} + x_{2,3} \cdot l_{3,1}) = -(1,5 \cdot 0 + (-0,5) \cdot 1) = 0,5;$$

$$x_{1,1} = \frac{1}{u_{1,1}} (1 - u_{1,2} \cdot x_{2,1} - u_{1,3} \cdot x_{3,1}) = \frac{1}{(-1)} (1 - 1 \cdot 0,5 - 0 \cdot (-1)) = -0,5.$$

Отже, обернена матриця

$$A^{-1} = \begin{pmatrix} -0,5 & 1,5 & -0,5 \\ 0,5 & 1,5 & -0,5 \\ -1 & -2 & 1 \end{pmatrix}.$$

Перевірка:

$$A \cdot A^{-1} = I$$

$$\begin{pmatrix} -1 & 1 & 0 \\ 0 & 2 & 1 \\ -1 & 5 & 3 \end{pmatrix} \cdot \begin{pmatrix} -0,5 & 1,5 & -0,5 \\ 0,5 & 1,5 & -0,5 \\ -1 & -2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

## 2.3. Метод квадратних коренів

Нехай задано систему лінійних алгебричних рівнянь (2.1), де  $A = (a_{ij})_{i,j=1}^n$  – симетрична додатньо визначена матриця, тобто  $a_{ij} = a_{ji}$ , головні мінори заданої матриці додатні. Тоді матрицю  $A$  можна записати у вигляді добутку двох транспонованих між собою трикутних матриць:

$$A = U^T \cdot U \quad (2.21)$$

Матрицю  $U$  визначаємо за формулами

$$u_{11} = \sqrt{a_{11}}; \quad u_{i1} = \frac{a_{i1}}{u_{11}}; \quad u_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} u_{ki}^2} \quad \text{для } i=2, \dots, n; \quad (2.22)$$

(у силу наших припущень на матрицю  $A$  підкореневий вираз додатний)

$$u_{ij} = \frac{a_{ij} - \sum_{k=1}^{i-1} u_{ik} u_{kj}}{u_{ii}} \quad \text{для } j=2, \dots, n, (j > i); \quad (2.23)$$

$$u_{ii} \neq 0 \quad (i=1, 2, \dots, n); \quad u_{ij} = 0 \quad \text{для } u_{ij} = 0 \quad (j < i). \quad (2.24)$$

Розв'язок симетричної системи (2.1) зводиться до послідовного розв'язування двох трикутних систем

$$U^T y = b; \quad (2.25)$$

$$Ux = y, \quad (2.26)$$

причому допоміжні невідомі  $y_1, y_2, \dots, y_n$  отримаємо за формулою

$$y_i = \frac{b_i - \sum_{k=1}^{i-1} u_{ki} y_k}{u_{ii}}, \quad (i = 1, 2, \dots, n). \quad (2.27)$$

Із системи (2.26) шукані значення  $x_i$  знаходимо у зворотному порядку, тобто для  $i = n, n-1, \dots, 1$  за формулою

$$x_i = \frac{y_i - \sum_{k=i+1}^n u_{ik} x_k}{u_{ii}}. \quad (2.28)$$

Розв'язування симетричних СЛАР за формулами (2.27), (2.28) називають методом квадратних коренів.

**Приклад 2.3.** Методом квадратних коренів розв'язати СЛАР із симетричною матрицею

$$\begin{cases} 4x_1 + 2x_2 + 2x_3 + x_4 = 9; \\ 2x_1 + 5x_2 + x_3 + 2x_4 = 10; \\ 2x_1 + x_2 + 5x_3 + x_4 = 9; \\ x_1 + 2x_2 + x_3 + 4,875x_4 = 8,875. \end{cases}$$

*Розв'язування.* Знаходимо елементи матриці  $U$  за формулами (2.22)–(2.24):

$$u_{11} = \sqrt{a_{11}} = 2; \quad u_{12} = \frac{a_{12}}{u_{11}} = \frac{2}{2} = 1; \quad u_{13} = \frac{a_{13}}{u_{11}} = 1; \quad u_{14} = \frac{a_{14}}{u_{11}} = 0,5;$$

$$u_{22} = \sqrt{a_{22} - u_{12}^2} = \sqrt{5 - 1^2} = 2; \quad u_{23} = \frac{a_{23} - u_{12} \cdot u_{13}}{a_{22}} = \frac{1 - 1 \cdot 1}{5} = 0;$$

$$u_{24} = \frac{a_{24} - u_{12} \cdot u_{14}}{a_{22}} = \frac{2 - 1 \cdot 0,5}{2} = 0,75;$$

$$u_{33} = \sqrt{a_{33} - u_{13}^2 - u_{23}^2} = \sqrt{5 - 1 - 0} = 2;$$

$$u_{34} = \frac{a_{34} - u_{13} \cdot u_{14} - u_{23} \cdot u_{24}}{u_{33}} = \frac{1 - 1 \cdot 0,5 - 0 \cdot 0,75}{2} = 0,25;$$

$$u_{44} = \sqrt{a_{44} - \sum_{k=1}^3 u_{k4}^2} = \sqrt{4,875 - 0,5^2 - 0,75^2 - 0,25^2} = 2.$$

Виконуємо перевірку  $U^T \cdot U = A$ , тобто

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 0,5 & 0,75 & 0,25 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & 1 & 0,5 \\ 0 & 2 & 0 & 0,75 \\ 0 & 0 & 2 & 0,25 \\ 0 & 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 2 & 1 \\ 2 & 5 & 1 & 2 \\ 2 & 1 & 5 & 1 \\ 1 & 2 & 1 & 4,875 \end{pmatrix}.$$

За формулами (2.27), відповідно, знаходимо розв'язки системи (2.25)

$$y_1 = \frac{b_1}{u_{11}} = \frac{9}{2} = 4,5;$$

$$y_2 = \frac{b_2 - u_{12}y_1}{u_{22}} = \frac{10 - 1 \cdot 4,5}{2} = \frac{5,5}{2} = 2,75;$$

$$y_3 = \frac{b_3 - u_{13}y_1 - u_{23}y_2}{u_{33}} = \frac{9 - 1 \cdot 4,5 - 0 \cdot 2,75}{2} = \frac{4,5}{2} = 2,25;$$

$$\begin{aligned} y_4 &= \frac{b_4 - u_{14}y_1 - u_{24}y_2 - u_{34}y_3}{u_{44}} = \frac{8,875 - 0,5 \cdot 4,5 - 0,75 \cdot 2,75 - 0,25 \cdot 2,25}{2} = \\ &= \frac{8,875 - 4,875}{2} = 2, \end{aligned}$$

а за формулами (2.28) – шукані розв'язки заданої системи:

$$x_4 = \frac{y_4}{u_{44}} = \frac{2}{2} = 1;$$

$$x_3 = \frac{y_3 - u_{34}x_4}{u_{33}} = \frac{2,25 - 0,25 \cdot 1}{2} = 1;$$



$$\lambda_1^{(2)} = -\frac{(Ay^{(2)}, x^{(1)})}{(Ax^{(1)}, x^{(1)})}. \quad (2.34)$$

Аналогічно вектор  $x^{(3)}$  побудуємо у вигляді

$$x^{(3)} = y^{(3)} + \lambda_1^{(3)} x^{(1)} + \lambda_2^{(3)} x^{(2)} \quad (2.35)$$

причому

$$\lambda_i^{(3)} = -\frac{(Ay^{(3)}, x^{(i)})}{(Ax^{(i)}, x^{(i)})} \quad (i = 1, 2). \quad (2.36)$$

Так само будемо вектори  $x^{(4)}, \dots, x^{(n)}$

**П р и к л а д 2.4.** Методом ортогоналізації розв'язати систему

$$\begin{cases} 10x_1 + x_2 + 0,5x_3 = 7; \\ x_1 + 20x_2 + 2x_3 = 24,5; \\ 0,5x_1 + 2x_2 + 15x_3 = 32,25. \end{cases}$$

*Розв'язування.* Результати обчислень записуємо в табл. 2.1.

Вибираємо систему лінійно незалежних векторів

$$\begin{cases} y^{(1)} = (1, 0, 0)^T \\ y^{(2)} = (0, 1, 0)^T \\ y^{(3)} = (0, 0, 1)^T \end{cases}$$

За формулами (2.33), (2.34) маємо

$$\begin{aligned} x^{(1)} &= (1, 0, 0); \\ x^{(2)} &= y^{(2)} + \lambda_1^{(2)} x^{(1)}, \end{aligned}$$

$$\text{де} \quad \lambda_1^{(2)} = -\frac{\left( \begin{pmatrix} 1 \\ 20 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right)}{\left( \begin{pmatrix} 10 \\ 1 \\ 0,5 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right)} = -0,1.$$

Таблиця 2.1

Обчислення за методом ортогоналізації

	$a_{i1}$	$a_{i2}$	$a_{i3}$	$b_i$
	10	1	0,5	7
	1	20	2	24,5
	0,5	2	15	32,25
$x^{(1)}$	1	$\lambda_7^{(2)} = -0,1;$	$\lambda_1^{(3)} = -0,05;$	$\lambda_1$
$x^{(2)}$	-0,1	1	$\lambda_7^{(3)} = -0,1;$	$\lambda_2$
$x^{(3)}$	-0,04	-0,1	1	
$Ax^{(1)}$	10	1	0,5	$(b, x^{(k)})$
$Ax^{(2)}$	0	19,9	1,95	
$Ax^{(3)}$	0	-0,04	14,78	
$(Ax^{(1)}, x^{(1)})$	10			7
$(Ax^{(2)}, x^{(2)})$		19,9		23,8
$(Ax^{(3)}, x^{(3)})$			14,784	29,52
$\alpha_k$	0,7	1,2	2,0	

Отже,

$$x^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - 0,1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -0,1 \\ 1 \\ 0 \end{pmatrix}.$$

Вектор  $x^{(3)}$  знаходимо за формулою (2.35), де за (2.36)

$$\lambda_1^{(3)} = -\frac{(Ay^{(3)}, x^{(1)})}{(Ax^{(1)}, x^{(1)})} = -\frac{\left( \begin{pmatrix} 0,5 \\ 2 \\ 15 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right)}{10} = -0,05;$$

$$\lambda_2^{(3)} = -\frac{(Ay^{(3)}, x^{(2)})}{(Ax^{(2)}, x^{(2)})} = -\frac{\left( \begin{pmatrix} 0,5 \\ 2 \\ 15 \end{pmatrix}, \begin{pmatrix} -0,1 \\ 1 \\ 0 \end{pmatrix} \right)}{\left( \begin{pmatrix} 0 \\ 19,9 \\ 1,95 \end{pmatrix}, \begin{pmatrix} -0,1 \\ 1 \\ 0 \end{pmatrix} \right)} = \frac{-1,95}{19,9} \approx -0,1;$$



$$x^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - 0,05 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - 0,1 \begin{pmatrix} -0,1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -0,04 \\ -0,1 \\ 1 \end{pmatrix}.$$

Далі знаходимо

$$(b, x^{(1)}) = 7; (b, x^{(2)}) = \left( \begin{pmatrix} 7 \\ 24,5 \\ 32,25 \end{pmatrix}, \begin{pmatrix} -0,1 \\ 1 \\ 0 \end{pmatrix} \right) = 23,8;$$

$$(b, x^{(3)}) = \left( \begin{pmatrix} 7 \\ 24,5 \\ 32,25 \end{pmatrix}, \begin{pmatrix} -0,04 \\ -0,1 \\ 1 \end{pmatrix} \right) = 29,52.$$

За формулами (2.30) обчислюємо  $\alpha_i$  ( $i = \overline{1,3}$ ):

$$\alpha_1 = \frac{(b, x^{(1)})}{(Ax^{(1)}, x^{(1)})} = \frac{7}{10} = 0,7$$

$$\alpha_2 = \frac{(b, x^{(2)})}{(Ax^{(2)}, x^{(2)})} = \frac{23,8}{19,9} \approx 1,2;$$

$$\alpha_3 = \frac{(b, x^{(3)})}{(Ax^{(3)}, x^{(3)})} = \frac{29,52}{14,784} \approx 2.$$

Остаточню

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0,7 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 1,2 \begin{pmatrix} -0,1 \\ 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} -0,04 \\ -0,1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0,5 \\ 1 \\ 2 \end{pmatrix}.$$

## 2.5. Метод прогонки розв'язування СЛАР з тридіагональною матрицею

Розглянемо різницеву крайову задачу вигляду

$$A_i y_{i-1} - C_i y_i + B_i y_{i+1} = -f_i, \quad (i=1, 2, \dots, n-1); \quad (2.37)$$

$$y_0 = \chi_1 y_1 + \mu_1; \quad (2.38)$$

$$y_n = \chi_2 y_{n-1} + \mu_2,$$

де  $A_i \neq 0$ ;  $B_i \neq 0$ ;  $i=1, 2, \dots, n-1$ ,  $\chi_1, \chi_2, \mu_1, \mu_2$  – задані числа.

Цю систему можна записати у вигляді СЛАР

$$Dy = f \quad (2.39)$$

з тридіагональною матрицею розмірності  $(n+1) \times (n+1)$

$$D = \begin{pmatrix} 1 & -\chi_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ A_1 & -C_1 & B_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & A_2 & -C_2 & B_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & A_{n-2} & -C_{n-2} & B_{n-2} & 0 \\ 0 & 0 & 0 & 0 & 0 & A_{n-1} & -C_{n-1} & B_{n-1} \\ 0 & 0 & 0 & 0 & 0 & 0 & -\chi_2 & 1 \end{pmatrix},$$

$$y = (y_0, y_1, \dots, y_n)^T \quad f = (\mu_1, -f_1, \dots, -f_{n-1}, \mu_2)^T$$

Згідно з прямим ходом методу прогонки, обчислюємо прогонні коефіцієнти

$$\alpha_{i+1} = \frac{B_i}{C_i - A_i \alpha_i}, \quad i=1, 2, \dots, n-1;$$

$$\beta_{i+1} = \frac{f_i + A_i \beta_i}{C_i - A_i \alpha_i}, \quad i=1, 2, \dots, n-1, \quad (2.40)$$

причому

$$\alpha_1 = \chi_1, \quad \beta_1 = \mu_1. \quad (2.41)$$

За зворотним ходом прогонки спочатку знаходимо

$$y_n = \frac{\mu_2 + \chi_2 \beta_n}{1 - \chi_2 \alpha_n}, \quad (2.42)$$

а далі при  $i = n-1, \dots, 0$  обчислюємо шуканий розв'язок системи (2.40) за формулами

$$y_i = \alpha_{i+1} \cdot y_{i+1} + \beta_{i+1} \quad (2.43)$$

Це так званий *метод правої прогонки* розв'язування СЛАР.

Формули прогонки можна застосовувати, коли знаменники у формулах (2.40) та (2.42) відмінні від нуля.

**Теорема.** Достатніми умовами коректності і стійкості методу прогонки (2.40)–(2.43) для розв'язування системи лінійних алгебраїчних рівнянь (2.37)–(2.38) є виконання нерівностей

$$|C_i| \geq |A_i| + |B_i|; \quad i = 1, 2, \dots, n-1; \quad |\chi_j| \leq 1; \quad j = 1, 2, \quad (2.44)$$

де хоча б одна з них строга.

**П р и к л а д 2.5.** Методом правої прогонки розв'язати СЛАР такого вигляду:

$$A_i = B_i = 1; C_i = 2; f_i = 0, \quad (i = 1, 2, 3, 4);$$

$$\chi_1 = 0; \mu_1 = 0; \chi_2 = 0; \mu_2 = 2.$$

*Розв'язування:* Матриця системи (2.40) має вигляд

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

а вектори  $y$  та  $f$  такі:  $y = (y_0, y_1, \dots, y_5)^T$ ,  $f = (1, 0, 0, 0, 0, 2)^T$

Обчислимо прогонні коефіцієнти:

$$\alpha_1 = \chi_1 = 0; \quad \beta_1 = \mu_1 = 1;$$

$$\alpha_2 = \frac{1}{2 - 0 \cdot 1} = 0,5; \quad \beta_2 = \frac{0 + 1 \cdot 1}{2 - 0 \cdot 1} = 0,5;$$

$$\alpha_3 = \frac{1}{2 - 1 \cdot 0,5} = \frac{2}{3}; \quad \beta_3 = \frac{0 + 1 \cdot 0,5}{1,5} = \frac{1}{3};$$

$$\alpha_4 = \frac{1}{2 - 1 \cdot \frac{2}{3}} = \frac{3}{4}; \quad \beta_4 = \frac{0 + 1 \cdot \frac{1}{3}}{\frac{4}{3}} = \frac{1}{4};$$

$$\alpha_5 = \frac{1}{2 - 1 \cdot \frac{3}{4}} = \frac{4}{5}; \quad \beta_5 = \frac{0 + 1 \cdot \frac{1}{4}}{\frac{5}{4}} = \frac{1}{5}$$

За допомогою зворотного ходу методу знайдемо розв'язки:

$$y_5 = \frac{2 + 0 \cdot \frac{1}{5}}{1 - 0 \cdot \frac{4}{5}} = 2;$$

$$y_4 = \frac{4}{5} \cdot 2 + \frac{1}{5} = 1,8;$$

$$y_3 = \frac{3}{4} \cdot 1,8 + \frac{1}{4} = 1,6;$$

$$y_2 = \frac{2}{3} \cdot 1,6 + \frac{1}{3} = 1,4;$$

$$y_1 = 0,5 \cdot 1,4 + 0,5 = 1,2;$$

$$y_0 = 0 \cdot 1,2 + 1 = 1.$$

Отже, одержано розв'язок  $y = (1; 1,2; 1,4; 1,6; 1,8; 2)^T$

## 2.6. Метод поворотів розв'язування СЛАР

Метод Гаусса може допускати в процесі перетворень зростання елементів матриці. У разі великих розмірностей системи рівнянь це може призвести якщо не до переповнення розрядної сітки комп'ютерів, то до сильного впливу похибок заокруглень, унаслідок

чого можлива втрата стійкості методу. Тому доцільною є побудова прямих методів розв'язування лінійних систем, можливо, складніших, ніж метод Гаусса, однак таких, що не допускають великого зростання елементів у процесі перетворень і, як наслідок, чисельно стійкіших. До таких методів належить метод поворотів.

Як і в методі Гаусса, мета прямого ходу перетворень у методі поворотів – зведення системи до трикутного вигляду за допомогою послідовного занулення піддіагональних елементів, спочатку першого стовпця, потім другого і так далі. Роблять це так.

Нехай  $c_1$  і  $s_1$  – деякі відмінні від нуля числа. Перше рівняння системи (2.1) помножимо на  $c_1$ , друге – на  $s_1$  і додамо їх. Отриманим рівнянням замінимо перше рівняння системи (2.1). Далі перше рівняння системи (2.1) множимо на  $-s_1$ , друге – на  $c_1$  і результатом їхнього додавання замінимо друге рівняння.

На введені параметри  $c_1$  і  $s_1$  накладемо дві умови:

занулення  $-s_1 a_{11} + c_1 a_{21} = 0$

та нормування  $c_1^2 + s_1^2 = 1$ , (2.45)

звідки знаходимо числа

$$c_1 = \frac{a_{11}}{\sqrt{a_{11}^2 + a_{21}^2}}, \quad s_1 = \frac{a_{21}}{\sqrt{a_{11}^2 + a_{21}^2}}. \quad (2.46)$$

Система (2.1) набуде вигляду

$$\begin{cases} a_{11}^{(1)} x_1 + a_{12}^{(1)} x_2 + \dots + a_{1n}^{(1)} x_n = b_1^{(1)}, \\ \quad \quad \quad a_{22}^{(1)} x_2 + \dots + a_{2n}^{(1)} x_n = b_2^{(1)}, \\ a_{31} x_1 + a_{32} x_2 + \dots + a_{3n} x_n = b_3, \\ \quad \quad \quad a_{n1} x_1 + a_{n2} x_2 + \dots + a_{nn} x_n = b_n, \end{cases} \quad (2.47)$$

де

$$a_{1j}^{(1)} = c_1 a_{1j} + s_1 a_{2j}, \quad (j = \overline{1, n}), \quad b_1^{(1)} = c_1 b_1 + s_1 b_2; \quad (2.48)$$

$$a_{2j}^{(1)} = -s_1 a_{1j} + c_1 a_{2j}, \quad (j = \overline{2, n}), \quad b_2^{(1)} = -s_1 b_1 + c_1 b_2. \quad (2.49)$$



Знаходження розв'язків  $x_n, x_{n-1}, \dots, x_1$  з (2.55) збігається зі зворотним ходом методу Гаусса.

**П р и к л а д 2.6.** Методом поворотів розв'язати таку систему рівнянь:

$$\begin{cases} 10x_1 + x_2 + 0,5x_3 = 11,5; \\ x_1 + 20x_2 + 2x_3 = 23; \\ 0,5x_1 + 2x_2 + 15x_3 = 17,5. \end{cases}$$

*Розв'язування:* Виконаємо такі перетворення за формулами (2.52), (2.53).

**Крок 1:**

$$c_1 = \frac{a_{11}}{\sqrt{a_{11}^2 + a_{21}^2}}, \quad s_1 = \frac{a_{21}}{\sqrt{a_{11}^2 + a_{21}^2}};$$

$$a_{1j}^{(1)} = c_1 a_{1j} + s_1 a_{2j} \quad (j = \overline{1, 3}); \quad b_1^{(1)} = c_1 b_1 + s_1 b_2;$$

$$a_{2j}^{(1)} = -s_1 a_{1j} + c_1 a_{2j} \quad (j = \overline{2, 3}); \quad b_2^{(1)} = -s_1 b_1 + c_1 b_2;$$

$$c_1 = \frac{10}{\sqrt{10^2 + 1}} = \frac{10}{\sqrt{101}}, \quad s_1 = \frac{1}{\sqrt{101}};$$

$$a_{11}^{(1)} = \frac{10}{\sqrt{101}} \cdot 10 + \frac{1}{\sqrt{101}} \cdot 1 = \sqrt{101}; \quad a_{12}^{(1)} = \frac{10}{\sqrt{101}} \cdot 1 + \frac{1}{\sqrt{101}} \cdot 20 = \frac{30}{\sqrt{101}};$$

$$a_{13}^{(1)} = \frac{10}{\sqrt{101}} \cdot 0,5 + \frac{1}{\sqrt{101}} \cdot 2 = \frac{7}{\sqrt{101}}; \quad b_1^{(1)} = \frac{10}{\sqrt{101}} \cdot 11,5 + \frac{1}{\sqrt{101}} \cdot 23 = \frac{138}{\sqrt{101}};$$

$$a_{22}^{(1)} = -\frac{1}{\sqrt{101}} \cdot 1 + \frac{10}{\sqrt{101}} \cdot 20 = \frac{199}{\sqrt{101}}; \quad a_{23}^{(1)} = -\frac{1}{\sqrt{101}} \cdot 0,5 + \frac{10}{\sqrt{101}} \cdot 2 = \frac{19,5}{\sqrt{101}};$$

$$b_2^{(1)} = -\frac{1}{\sqrt{101}} \cdot 11,5 + \frac{10}{\sqrt{101}} \cdot 23 = \frac{218,5}{\sqrt{101}}.$$

На першому кроці одержимо таку систему:

$$\begin{cases} 101x_1 + 30x_2 + 7x_3 = 138; \\ 199x_2 + 19,5x_3 = 218,5; \\ 0,5x_1 + 2x_2 + 15x_3 = 17,5. \end{cases}$$

**Крок 2:**

$$c_2 = \frac{a_{11}^{(1)}}{\sqrt{a_{11}^{(1)2} + a_{31}^2}}; \quad s_2 = \frac{a_{31}}{\sqrt{a_{11}^{(1)2} + a_{31}^2}};$$

$$a_{1j}^{(2)} = c_2 a_{1j}^{(1)} + s_2 a_{3j}, \quad (j = \overline{1, 3}); \quad b_1^{(2)} = c_2 b_1^{(1)} + s_2 b_3;$$

$$a_{3j}^{(1)} = -s_2 a_{1j}^{(1)} + c_2 a_{3j}, \quad (j = \overline{2, 3}); \quad b_3^{(1)} = -s_2 b_1^{(1)} + c_2 b_3;$$

$$c_2 = \frac{101}{\sqrt{101^2 + 0,5^2}} = \frac{101}{\sqrt{10201,25}}; \quad s_2 = \frac{0,5}{\sqrt{10201,25}};$$

$$a_{11}^{(2)} = \frac{101}{\sqrt{10201,25}} \cdot 101 + \frac{0,5}{\sqrt{10201,25}} \cdot 0,5 = \sqrt{10201,25};$$

$$a_{12}^{(2)} = \frac{101}{\sqrt{10201,25}} \cdot 30 + \frac{0,5}{\sqrt{10201,25}} \cdot 2 = \frac{3031}{\sqrt{10201,25}};$$

$$a_{13}^{(2)} = \frac{101}{\sqrt{10201,25}} \cdot 7 + \frac{0,5}{\sqrt{10201,25}} \cdot 15 = \frac{714,5}{\sqrt{10201,25}};$$

$$b_1^{(2)} = \frac{101}{\sqrt{10201,25}} \cdot 138 + \frac{0,5}{\sqrt{10201,25}} \cdot 17,5 = \frac{13946,75}{\sqrt{10201,25}};$$

$$a_{32}^{(1)} = -\frac{0,5}{\sqrt{10201,25}} \cdot 30 + \frac{101}{\sqrt{10201,25}} \cdot 2 = \frac{187}{\sqrt{10201,25}};$$

$$a_{33}^{(1)} = -\frac{0,5}{\sqrt{10201,25}} \cdot 7 + \frac{101}{\sqrt{10201,25}} \cdot 15 = \frac{1511,5}{\sqrt{10201,25}};$$

$$b_3^{(1)} = -\frac{0,5}{\sqrt{10201,25}} \cdot 138 + \frac{101}{\sqrt{10201,25}} \cdot 17,5 = \sqrt{10201,25}$$

Отримано систему

$$\begin{cases} 10201,25x_1 + 3031x_2 + 714,5x_3 = 13946,75; \\ \phantom{10201,25x_1} 199x_2 + 19,5x_3 = 218,5; \\ \phantom{10201,25x_1} 187x_2 + 1511,5x_3 = 1698,5. \end{cases}$$



**Крок 3:**

$$c_1^{(1)} = \frac{199}{\sqrt{199^2 + 187^2}} = \frac{199}{\sqrt{74570}}; \quad s_1^{(1)} = \frac{187}{\sqrt{199^2 + 187^2}};$$

$$a_{2j}^{(2)} = c_1^{(1)} a_{2j}^{(1)} + s_1^{(1)} a_{3j}^{(1)} \quad (j = 2, 3); \quad b_2^{(2)} = c_1^{(1)} b_2^{(1)} + s_1^{(1)} b_3^{(1)};$$

$$a_{3j}^{(2)} = -s_1^{(1)} a_{2j}^{(1)} + c_1^{(1)} a_{3j}^{(1)} \quad (j = 3); \quad b_3^{(2)} = -s_1^{(1)} b_2^{(1)} + c_1^{(1)} b_3^{(1)};$$

$$a_{22}^{(2)} = \frac{199}{\sqrt{74570}} \cdot 199 + \frac{187}{\sqrt{74570}} \cdot 187 = \frac{74570}{\sqrt{74570}};$$

$$a_{23}^{(2)} = \frac{199}{\sqrt{74570}} \cdot 19,5 + \frac{187}{\sqrt{74570}} \cdot 1511,5 = \frac{286531,0}{\sqrt{74570}};$$

$$b_2^{(2)} = \frac{199}{\sqrt{74570}} \cdot 218,5 + \frac{187}{\sqrt{74570}} \cdot 1698,5 = \frac{361101,0}{\sqrt{74570}};$$

$$a_{33}^{(2)} = -\frac{187}{\sqrt{74570}} \cdot 19,5 + \frac{199}{\sqrt{74570}} \cdot 1511,5 = \frac{297142,0}{\sqrt{74570}};$$

$$b_3^{(2)} = -\frac{187}{\sqrt{74570}} \cdot 218,5 + \frac{199}{\sqrt{74570}} \cdot 1698,5 = \frac{297142,0}{\sqrt{74570}}.$$

Остаточно отримано систему

$$\begin{cases} 10201,25x_1 + 3031x_2 + 714,5x_3 = 13946,75; \\ \phantom{10201,25x_1} 74570x_2 + 286531x_3 = 361101; \\ \phantom{10201,25x_1} \phantom{74570x_2} 297142x_3 = 297142. \end{cases}$$

Зворотним ходом знаходимо розв'язки:

$$x_3 = 1;$$

$$x_2 = \frac{361101 - 286531x_3}{74570} = 1;$$

$$x_1 = \frac{13946,75 - 3031x_2 - 714,5x_3}{10201,25} = 1.$$

## **Розділ 3**

### **ІТЕРАЦІЙНІ МЕТОДИ РОЗВ'ЯЗУВАННЯ СИСТЕМ ЛІНІЙНИХ АЛГЕБРИЧНИХ РІВНЯНЬ**

Опишемо ітераційні методи розв'язування систем лінійних алгебричних рівнянь. Ці методи дають розв'язок у вигляді границі послідовності деяких векторів. Такі вектори будують шляхом виконання одноманітного процесу, який називають процесом ітерацій.

У сучасній літературі наведено багато ітераційних методів, що ґрунтуються на різних принципах. Зазвичай обчислювальні схеми таких методів прості й зручні у використанні. Проте ітераційний процес може виявитись розбіжним для системи або збіжність процесу може бути настільки повільною, що практично важко досягти задовільної точності.

Важливою особливістю ітераційних методів є їхні само-уточнення і простота реалізації на комп'ютерах. Ітераційний метод для початку обчислень потребує задання одного або декількох початкових наближень. Як уже зазначено, умови і швидкість збіжності кожного ітераційного процесу суттєво залежать від властивостей матриці системи і вибору початкових наближень.

Розглянемо лінійні стаціонарні методи: метод простих ітерацій та метод Зейделя, релаксаційні та градієнтні методи, методи варіаційного типу.

#### **3.1. Принципи побудови ітераційних процесів**

Головні ітераційні процеси можна описати за допомогою такої загальної схеми.

Нехай задано систему лінійних алгебричних рівнянь

$$Ax = b \tag{3.1}$$

з неособливою матрицею  $A$  розмірності  $n \times n$ . Побудуємо послідовність векторів  $x^{(1)}, x^{(2)}, \dots, x^{(k)}$  за рекурентною формулою

$$x^{(k)} = x^{(k-1)} + H^{(k)}(b - Ax^{(k-1)}), \quad (3.2)$$

де  $H^{(1)}, H^{(2)}, \dots$  – деяка послідовність матриць;  $x^{(0)}$  – початкове наближення. Різний вибір матриць  $H^{(k)}$  приводить до різних ітераційних процесів.

Ітераційні процеси (3.2) мають ту властивість, що для кожного з них точний розв'язок  $x^*$  є нерухомою точкою. Це означає таке: якщо за початкове наближення  $x^{(0)}$  взяти  $x^*$  то всі наступні наближення теж дорівнюватимуть  $x^*$ . Навпаки, якщо довільний ітераційний процес, для якого  $x^*$  є нерухомою точкою, реалізується за формулою

$$x^{(k)} = C^{(k)}x^{(k-1)} + d^{(k)}; \quad k = 1, 2, \dots, \quad (3.3)$$

де  $C^{(k)}$  – послідовність матриць;  $d^{(k)}$  – послідовність векторів, то його можна записати у вигляді (3.2).

Неважко визначити необхідну і достатню умову того, щоб ітераційний процес (3.2) збігався до розв'язку за довільного початкового вектора. Справді,

$$x^* - x^{(k)} = x^* - x^{(k-1)} - H^{(k)}(Ax^* - Ax^{(k-1)}) = (E - H^{(k)}A)(x^* - x^{(k-1)}).$$

Звідси

$$x^* - x^{(k)} = (E - H^{(k)}A)(E - H^{(k-1)}A) \dots (E - H^{(1)}A)(x^* - x^{(0)}).$$

Для того, щоб  $x^* - x^{(k)} \rightarrow 0$  при  $k \rightarrow \infty$  для довільного початкового вектора  $x^{(0)}$ , необхідно і достатньо, щоб матриця

$$T^{(k)} = (E - H^{(k)}A)(E - H^{(k-1)}A) \dots (E - H^{(1)}A)$$

прямувала до нуль-матриці. Для цього достатньо, щоб довільна норма матриці  $T^{(k)}$  прямувала до нуля. Ця умова дає лише загальний погляд для побудови збіжних конкретних ітераційних методів.

Найпростіші ітераційні методи – стаціонарні ітераційні процеси, у яких матриці  $H^{(k)}$  не залежать від номера кроку  $k$ . При  $H^{(k)} = E$  отримуємо класичний процес послідовних наближень. Вибір матриці  $H$  для стаціонарного процесу і  $H^{(k)}$  для нестаціонарного можна виконати багатьма різними способами.

Важливим способом побудови ітераційних процесів є спосіб релаксації. Під ним розуміють такий вибір матриць  $H^{(k)}$  з деякого наперед заданого класу матриць, щоб на кожному кроці ітераційного процесу зменшувалось певне значення, що характеризує точність розв'язку системи.

Серед релаксаційних методів найбільше поширені *координатні*, у яких матриці  $H^{(k)}$  підібрані так, що на кожному кроці змінюється одна або декілька компонент послідовних наближень, і *градієнтні*.

## 3.2. Метод простих ітерацій

Оскільки метод простих ітерацій є стаціонарним, то систему рівнянь (3.1), згідно з формулою (3.3), перетворимо до вигляду

$$x = Cx + d. \quad (3.4)$$

Це можна виконати, записавши  $A = B + D$ , і тоді  $C = -B^{-1}D$ ;  $d = B^{-1}b$ .

Розв'язок (3.4) знаходимо як межу послідовності

$$x^{(k+1)} = Cx^{(k)} + d. \quad (3.5)$$

**Теорема 3.1** (критерій збіжності). Для збіжності методу простих ітерацій з довільним початковим вектором  $x^{(0)}$  необхідно і достатньо, щоб усі власні числа матриці  $C$  були за модулем менші від 1, тобто всі корені характеристичного рівняння

$$\det(C - \lambda E) = 0 \quad (3.6)$$

були за модулем менші від одиниці.

Оскільки знаходження коренів рівняння (3.6) є непростою проблемою, то на практиці часто використовують достатню умову збіжності методу простих ітерацій.

**Теорема 3.2** (ознака збіжності). Якщо  $\|C\| < 1$ , то система (3.4) має єдиний розв'язок, і ітераційний процес (3.5) збігається до розв'язку зі швидкістю геометричної прогресії.

Найчастіше обмежуються перевіркою однієї з двох умов:

$$\begin{aligned} 1) \quad \|C\|_I < 1, \text{ де } \|C\|_I &= \max_{1 \leq i \leq n} \sum_{j=1}^n |c_{ij}|; \\ 2) \quad \|C\|_{II} < 1, \text{ де } \|C\|_{II} &= \max_{1 \leq j \leq n} \sum_{i=1}^n |c_{ij}|. \end{aligned} \quad (3.7)$$

Точність отримуваних розв'язків характеризує величина  $\frac{\|C\|}{1-\|C\|}$ . Щоб оцінити кількість ітерацій для досягнення необхідної точності  $\varepsilon$  розв'язку, можна скористатися нерівностями

$$\|x^{(k)} - x^*\| \leq \frac{\|C\|}{1-\|C\|} \|x^{(k+1)} - x^{(k)}\| \quad (3.8)$$

або

$$\|x^{(k+1)} - x^*\| \leq \frac{\|C\|^{k+1}}{1-\|C\|} \|x^{(1)} - x^{(0)}\|.$$

Якщо за початкове наближення вибирають стовпець вільних членів системи (3.4), то

$$\|x^{(k+1)} - x^*\| \leq \frac{\|C\|^{k+1}}{1-\|C\|} \|d\|.$$

Нехай  $\|C\|_I = q$ . Тоді  $\frac{q^{k+1}}{1-q} \|d\|_I \leq \varepsilon$ ; звідси можна знайти

апостеріорну оцінку кількості ітерацій

$$k > \frac{1}{\lg q} [\lg \|d\|_I - \lg(\varepsilon(1-q))] - 1.$$

З умови збіжності методу простих ітерацій бачимо, що вираз  $A = B + D$  не може бути зовсім довільним. Якщо матриця  $A$  має діагональне переважання, то за матрицю  $B$  можна вибрати нижню трикутну матрицю з  $A$ :

$$B = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{nn} \end{bmatrix}, \quad a_{ii} \neq 0.$$

Якщо в матриці  $A$  немає діагонального переважання, то його можна домогтися яким-небудь способом, наприклад, перетворивши систему (3.1) за допомогою лінійної комбінації її рівнянь.

Наприклад, щоб систему рівнянь

$$\begin{cases} 5x_1 - x_2 + 2x_3 = 8; \\ 3x_1 - 2x_2 + 9x_3 = -3; \\ 6x_1 - 5x_2 + 2x_3 = -10 \end{cases}$$

звести до вигляду, придатного для застосування методу простих ітерацій, треба виконати такі перетворення: перше рівняння залишити без змін, від першого рівняння відняти третє, і це буде другим рівнянням нової системи. Друге рівняння вважати третім у новій системі. Одержимо систему з діагональним переважанням

$$\begin{cases} 5x_1 - x_2 + 2x_3 = 8; \\ -x_1 + 4x_2 = -2; \\ 3x_1 - 2x_2 + 9x_3 = -3. \end{cases}$$

Ще один спосіб зведення рівняння (3.1) до вигляду, зручного для застосування методу простих ітерацій у випадку додатно визначеної матриці  $A$ , такий. Позначимо першу норму матриці  $A$  через  $\mu$ :  $\|A\|_1 = \mu$ , отримаємо, що всі власні значення матриці  $A$  є в інтервалі  $(0, \mu)$ . Тоді систему  $Ax = b$  зводимо до вигляду

$$x = (E - \frac{2}{\mu}A)x + \frac{2b}{\mu},$$

Тобто

$$C = E - \frac{2}{\mu}A; \quad d = \frac{2}{\mu}b.$$

Власні значення матриці  $C$  будуть у відкритому інтервалі  $(-1, 1)$ , тому метод простих ітерацій збігатиметься.

Якщо для системи (3.1) виконується умова  $a_{ii} \neq 0 \quad (i = \overline{1, n})$ , то з кожного  $i$ -го рядка визначаємо  $x_i$ . Отримаємо рівняння

$$x_i = - \sum_{j=1, j \neq i}^n \frac{a_{ij}}{a_{ii}} x_j + \frac{b_i}{a_{ii}}, \quad (i = \overline{1, n}).$$

Цю систему називають системою, зведеною до *нормального* вигляду. Для неї ітераційний процес матиме вигляд

$$x_i^{(k+1)} = - \sum_{j=1, j \neq i}^n \frac{a_{ij}}{a_{ii}} x_j^{(k)} + \frac{b_i}{a_{ii}}, \quad (i = \overline{1, n}), \quad (3.9)$$

або у матричному зображенні:  $x = Cx + d$ , де  $C = -D^{-1}(L + R)$ ;

$d = D^{-1}b$ ;  $D$  – діагональна матриця;  $L, R$  – ліва і права строго трикутні матриці (по діагоналі нульові елементи).

Побудований ітераційний процес називають **методом Якобі**. Він збігається, якщо виконується одна з таких умов:

$$|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}| \quad \text{для всіх } i = \overline{1, n};$$

$$|a_{jj}| > \sum_{i=1, i \neq j}^n |a_{ji}| \quad \text{для всіх } j = \overline{1, n}.$$

Отже, достатньою умовою збіжності ітераційного процесу є переважання діагональних елементів у матриці  $A$ .

**Теорема 3.3.** *Метод Якобі (3.9) збігається до розв'язку системи (3.1) тоді й тільки тоді, коли всі корені рівняння*

$$\begin{vmatrix} a_{11}\lambda - 1 & a_{12} & a_{1n} \\ a_{21} & a_{22}\lambda - 1 & a_{2n} \\ a_{n1} & a_{n2} & a_{nn}\lambda - 1 \end{vmatrix} = 0$$

за модулем менші від одиниці.

**П р и к л а д 3.1.** Дослідити збіжність методу Якобі для системи  $Ax = b$ , де

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 3 & 2 \\ -1 & 1 & 1 \end{pmatrix}; \quad b = \begin{pmatrix} 8 \\ 7 \\ 1 \end{pmatrix}.$$

**Розв'язування.** За теоремою 3.3 запишемо рівняння

$$\begin{vmatrix} 4\lambda & 2 & 2 \\ 2 & 3\lambda & 2 \\ -1 & 1 & \lambda \end{vmatrix} = 12\lambda^3 + 4 - 4 + 6\lambda - 8\lambda - 4\lambda = \\ = 12\lambda^3 - 6\lambda = 6\lambda(2\lambda^2 - 1) = 0.$$

Звідси  $\lambda_1 = 0, \lambda_2 = \frac{1}{\sqrt{2}}, \lambda_3 = -\frac{1}{\sqrt{2}}$ .

Оскільки  $|\lambda_i| < 1$  для всіх  $i = 1, 2, 3$ , то метод Якобі збіжний.

**П р и к л а д 3.2.** З точністю до  $\varepsilon = 10^{-3}$  методом Якобі розв'язати систему лінійних алгебричних рівнянь четвертого порядку

$$\begin{cases} 8x_1 + x_2 + x_3 - x_4 = 9; \\ x_1 + 5x_2 - x_3 - x_4 = 4; \\ x_1 - x_2 + 5x_3 + x_4 = 18; \\ 2x_1 + x_2 - x_3 + 10x_4 = 41. \end{cases} \quad (3.10)$$

**Розв'язування.** Зведемо цю систему до нормального вигляду:

$$\begin{cases} x_1 = 1,125 - \frac{1}{8}x_2 - \frac{1}{8}x_3 + \frac{1}{8}x_4; \\ x_2 = 0,8 - \frac{1}{5}x_1 + \frac{1}{5}x_3 + \frac{1}{5}x_4; \\ x_3 = 3,6 - \frac{1}{5}x_1 + \frac{1}{5}x_2 - \frac{1}{5}x_4; \\ x_4 = 4,1 - \frac{1}{5}x_1 - \frac{1}{10}x_2 + \frac{1}{10}x_3. \end{cases}$$

У матричному вигляді

$$C = \begin{pmatrix} 0 & -0,125 & -0,125 & 0,125 \\ -0,2 & 0 & 0,2 & 0,2 \\ -0,2 & 0,2 & 0 & -0,2 \\ -0,2 & -0,1 & 0,1 & 0 \end{pmatrix}; \quad d = \begin{pmatrix} 1,125 \\ 0,8 \\ 3,6 \\ 4,1 \end{pmatrix}.$$



Перевіримо достатню умову збіжності:

$$\sum_{j=1}^4 |c_{1j}| = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = 0,375 < 1; \quad \sum_{j=1}^4 |c_{2j}| = 0,2 + 0,2 + 0,2 = 0,6 < 1;$$

$$\sum_{j=1}^4 |c_{3j}| = 0,2 + 0,2 + 0,2 = 0,6 < 1; \quad \sum_{j=1}^4 |c_{4j}| = 0,2 + 0,1 + 0,1 = 0,4 < 1.$$

Отже, за норму матриці  $C$  можна вибрати найбільшу зі знайдених сум, тобто  $\|C\| = 0,6$ .

Ітераційний процес закінчується, коли оцінки, згідно з нерівністю (3.8), свідчать про досягнення заданої точності:

$$|x_i^* - x_i^{(k)}| \leq \frac{\|C\|}{1 - \|C\|} \max_{j=1,4} |x_j^{(k+1)} - x_j^{(k)}| \quad i = 1, 2, 3, 4.$$

Знайдемо потрібну кількість ітерацій:

$$\frac{\|C\|^{k+1}}{1 - \|C\|} \|d\| < \varepsilon: \quad \frac{0,6^{k+1}}{0,4} \cdot 4,1 < 0,001; \quad 0,6^{k+1} < 0,0000975;$$

$$(k+1) \lg 0,6 < \lg 0,0000975; \quad \text{тому} \quad k+1 > \frac{\lg(975 \cdot 10^{-7})}{\lg 0,6};$$

$$k+1 > \frac{-7+3}{-0,2147}; \quad k > 18,677-1; \quad k > 17,677. \quad \text{Отже, } k=18. \quad \text{Хоч за}$$

апостеріорною оцінкою отримано 18 ітерацій, на практиці ця кількість може бути іншою, оскільки така оцінка, зазвичай, завищена.

Побудуємо послідовні наближення, узявши за початкове наближення вектор вільних членів  $d$ :

$$x^{(0)} = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \\ x_4^{(0)} \end{pmatrix} = \begin{pmatrix} 1,125 \\ 0,8 \\ 3,6 \\ 4,1 \end{pmatrix}.$$

Перше наближення:

$$\begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \\ x_4^{(1)} \end{pmatrix} = \begin{pmatrix} 1,125 \\ 0,8 \\ 3,6 \\ 4,1 \end{pmatrix} + \begin{pmatrix} 0 & -0,125 & -0,125 & 0,125 \\ -0,2 & 0 & 0,2 & 0,2 \\ -0,2 & 0,2 & 0 & -0,2 \\ -0,2 & -0,1 & 0,1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1,125 \\ 0,8 \\ 3,6 \\ 4,1 \end{pmatrix} = \begin{pmatrix} 1,0875 \\ 2,1150 \\ 2,7150 \\ 4,1550 \end{pmatrix}.$$

Друге наближення:

$$\begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \\ x_4^{(2)} \end{pmatrix} = \begin{pmatrix} 1,125 \\ 0,8 \\ 3,6 \\ 4,1 \end{pmatrix} + \begin{pmatrix} 0 & -0,125 & -0,125 & 0,125 \\ -0,2 & 0 & 0,2 & 0,2 \\ -0,2 & 0,2 & 0 & -0,2 \\ -0,2 & -0,1 & 0,1 & 0 \end{pmatrix} \begin{pmatrix} 1,0875 \\ 2,1150 \\ 2,7150 \\ 4,1550 \end{pmatrix} = \begin{pmatrix} 1,0406 \\ 1,9565 \\ 2,9745 \\ 3,9425 \end{pmatrix}.$$

Аналогічно отримаємо такі наближення до розв'язку:

$$\bar{x}^{(3)} = \begin{pmatrix} 1,0014 \\ 1,9753 \\ 2,9947 \\ 3,9937 \end{pmatrix}; \quad \bar{x}^{(4)} = \begin{pmatrix} 1,0030 \\ 1,9974 \\ 2,9960 \\ 4,0017 \end{pmatrix}; \quad \bar{x}^{(5)} = \begin{pmatrix} 1,0010 \\ 1,9989 \\ 2,9986 \\ 3,9993 \end{pmatrix}; \quad \bar{x}^{(6)} = \begin{pmatrix} 1,0002 \\ 1,9994 \\ 2,9997 \\ 3,9998 \end{pmatrix}.$$

Отже, розв'язок  $x_1 = 1; x_2 = 2; x_3 = 3; x_4 = 4$  одержано з точністю до  $10^{-3}$  за шість ітерацій, оскільки

$$\frac{\|C\|}{1 - \|C\|} \|\bar{x}^{(6)} - \bar{x}^{(5)}\| = 1,5 \|\bar{x}^{(6)} - \bar{x}^{(5)}\| < 10^{-3}$$

що не суперечить теоретичним оцінкам кількості ітерацій ( $k = 18$ ).

### 3.3. Метод Зейделя

Метод Зейделя відрізняється від методу простих ітерацій лише тим, що під час обчислення  $(k+1)$ -го наближення  $i$ -ї компоненти  $x_i$  враховують уже визначене раніше  $(k+1)$ -ше наближення компонент  $x_1, x_2, \dots, x_{i-1}$ , тобто обчислення виконують за формулами

$$x_i^{(k+1)} = d_i + \sum_{j=1}^{i-1} c_{ij} x_j^{(k+1)} + \sum_{j=i+1}^n c_{ij} x_j^{(k)}, \quad (i = \overline{1, n}), \quad (3.11)$$

або в матричному вигляді

$$x^{(k+1)} = Bx^{(k+1)} + Dx^{(k)} + d, \quad k = 1, 2, \dots \quad (3.12)$$

$$\text{де } B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ c_{21} & 0 & 0 & 0 \\ c_{31} & c_{32} & 0 & 0 \\ c_{n1} & c_{n2} & c_{n,n-1} & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & c_{12} & c_{1,n-1} & c_{1,n} \\ 0 & 0 & c_{2,n-1} & c_{2,n} \\ 0 & 0 & c_{3,n-1} & c_{3,n} \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

або

$$x^{(k+1)} = (E - B)^{-1} x^{(k)} + (E - B)^{-1} d.$$

Із останнього рівняння випливає, що ітераційний процес Зейделя збігається, якщо виконується умова  $|\lambda_i| < 1$ , де  $\lambda_i$  – власні числа матриці  $(E - B)^{-1} D$ . Ці власні числа визначають із рівняння

$$\det((E - B)^{-1} D - \lambda E) = 0.$$

**Теорема 3.4.** Для того, щоб метод Зейделя (3.12) збігався за довільного початкового наближення  $x^{(0)}$ , необхідно і достатньо, щоб усі корені рівняння

$$\begin{vmatrix} c_{11} - \lambda & c_{12} & c_{1n} \\ c_{21}\lambda & c_{22} - \lambda & c_{2n} \\ c_{n1}\lambda & c_{n2}\lambda & c_{nn} - \lambda \end{vmatrix} = 0$$

були за модулем менші від одиниці.

Під час обчислень стаціонарні ітераційні процеси збігаються до розв'язку з точністю  $\varepsilon$  у разі виконання нерівності [3]

$$\|x^{(k+1)} - x\| \leq \mu \|x^{(k+1)} - x^{(k)}\|,$$

$$\text{де } \mu = \max \frac{\gamma_i}{1 - \beta_i}, \quad \gamma_i = \sum_{j=1}^n |c_{ji}|, \quad \beta_i = \sum_{j=1}^{i-1} |c_{ij}|.$$

**П р и к л а д 3.2.** Методом Зейделя з точністю до  $\varepsilon = 10^{-3}$  розв'язати систему лінійних алгебричних рівнянь (3.10).

*Розв'язування.* Запишемо цю систему у вигляді (3.11):

$$\begin{cases} x_1^{(k+1)} = 1,125 - 0,125x_2^{(k)} - 0,125x_3^{(k)} + 0,125x_4^{(k)}; \\ x_2^{(k+1)} = 0,8 - 0,2x_1^{(k+1)} + 0,2x_3^{(k+1)} + 0,2x_4^{(k)}; \\ x_3^{(k+1)} = 3,6 - 0,2x_1^{(k+1)} + 0,2x_2^{(k+1)} - 0,2x_4^{(k)}; \\ x_4^{(k+1)} = 4,1 - 0,2x_1^{(k+1)} - 0,1x_2^{(k+1)} + 0,1x_3^{(k+1)} \end{cases}$$

Норма цієї перетвореної матриці  $q = 0,6 < 1$ . Виберемо за початкове наближення вектор вільних членів, отримаємо наближені значення розв'язку, наведені в табл. 3.1.

Таблиця 3.1

**Послідовність наближених розв'язків за методом Зейделя**

Ітерації	0	1	2	3	4	5
$x_1^{(k)}$	1,125	1,0875	0,9824	1,0002	0,9997	1,0000
$x_2^{(k)}$	0,8	2,1225	1,9947	2,0027	1,9999	2,0000
$x_3^{(k)}$	3,6	2,9870	3,0001	2,9995	3,0001	3,0000
$x_4^{(k)}$	4,1	3,9689	4,0049	3,9996	4,0001	4,0000

На п'ятій ітерації задана точність досягнута, бо  $\|\bar{x}^{(5)} - \bar{x}^{(4)}\| < \frac{2}{3} \times 10^{-3}$ .

Часто для поліпшення збіжності зручніше змінити порядок рівнянь у заданій системі або навіть нумерацію невідомих. Можна піти й далі, а саме – на кожній ітерації вибирати свій порядок. Ця ідея стала основою методу, який є видозміною методу Зейделя.

Як і в попередніх методах, вибираємо початкове наближення  $x^{(0)}$  і обчислюємо відхили:

$$\delta_i = a_{i1}x_1^{(0)} + a_{i2}x_2^{(0)} + \dots + a_{in}x_n^{(0)} - b_i, \quad i = \overline{1,4}$$

Відшукуємо максимальний за модулем відхил  $\max_{i=\overline{1,4}} |\delta_i|$  і з рівняння з цим відхилом знаходимо перше наближення першої невідомої, тобто  $x_1^{(1)}$ . Потім знову перераховуємо всі відхили

$$\delta_j = a_{j1}x_1^{(1)} + a_{j2}x_2^{(1)} + \dots + a_{jn}x_n^{(1)} - b_j, \quad j \neq i,$$

і підберемо  $x_2^{(1)}$ , що задовольняє рівняння

$$a_{j1}x_1^{(l)} + a_{j2}x_2^{(l)} + a_{j3}x_3^{(l)} + \dots + a_{jn}x_n^{(l)} = b_j,$$

де  $j$  – номер рівняння з найбільшим за модулем відхилом. Так робимо доти, поки не знайдемо перше наближення всіх змінних, тобто вектор  $(x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)})$ . На наступному циклі замість вектора  $x^{(0)}$  беремо  $x^{(1)}$  і т. д. Повторення циклів триває до досягнення потрібної точності. Іноді в разі вибору рівняння, з якого обчислюють поліпшене значення невідомої, керуються не принципом максимального за модулем відхилу, а яким-небудь іншим.

**П р и к л а д 3.3.** Видозміненим методом Зейделя розв'язати систему з точністю  $\varepsilon = 10^{-2}$ :

$$\begin{cases} 10x_1 - 2x_2 - 2x_3 = 6; \\ -x_1 + 10x_2 - 2x_3 = 7; \\ -x_1 - x_2 + 10x_3 = 8. \end{cases}$$

*Розв'язування.* Для зручності кожне рівняння поділимо на діагональний коефіцієнт і вільний член перенесемо ліворуч. У всіх коефіцієнтах поміняємо знаки:

$$\begin{cases} -x_1 + 0,2x_2 + 0,2x_3 + 0,6 = 0; \\ 0,1x_1 - x_2 + 0,2x_3 + 0,7 = 0; \\ 0,1x_1 + 0,1x_2 - x_3 + 0,8 = 0. \end{cases}$$

Для заданого прикладу  $q = 0,4 < 1$ , тому ітераційний процес збігається. За початкове наближення виберемо нуль-вектор  $x^{(0)} = \{0; 0; 0\}$ .

Знайдемо відповідні відхили  $\delta_1^{(0)} = 0,6$ ;  $\delta_2^{(0)} = 0,7$ ;  $\delta_3^{(0)} = 0,8$ .

Далі  $\max_{i=1,3} |\delta_i| = 0,8$  є у третьому рівнянні і

$$x_3^{(1)} = 0,1x_1^{(0)} + 0,1x_2^{(0)} + 0,8; \quad x_3^{(1)} = 0,8.$$

Підставимо  $x_3^{(1)}$  у перше і друге рівняння й знову визначимо відхили:

$$\begin{cases} \delta_1^{(1)} = 0 + 0,2 \cdot 0 + 0,2 \cdot 0,8 + 0,6 = 0,76; \\ \delta_2^{(1)} = 0 - 0 + 0,2 \cdot 0,8 + 0,7 = 0,86; \\ \delta_3^{(1)} = 0. \end{cases}$$

Знайдемо  $\max_{i=1,3} |\delta_i^{(1)}| = 0,86$  у другому рівнянні. Тому

$$x_2^{(1)} = 0,1x_1^{(0)} + 0,2x_1^{(1)} + 0,7 = 0,1 \cdot 0 + 0,2 \cdot 0,8 + 0,7 = 0,86.$$

З першого рівняння

$$x_1^{(1)} = 0,2 \cdot 0,86 + 0,2 \cdot 0,8 + 0,6 = 0,932.$$

Отже, перше наближення  $x^{(1)} = (0,932; 0,86; 0,8)$ .

Перейдемо на наступний цикл для знаходження компонент вектора другого наближення:

$$\begin{cases} \delta_1^{(2)} = -0,932 + 0,2 \cdot 0,86 + 0,2 \cdot 0,8 + 0,6 = 0; \\ \delta_2^{(2)} = 0,1 \cdot 0,932 - 0,86 + 0,2 \cdot 0,8 + 0,7 = 0,0932; \\ \delta_3^{(2)} = 0,1 \cdot 0,932 + 0,1 \cdot 0,86 - 0,8 + 0,8 = 0,1792. \end{cases}$$

Тоді  $\max_{i=1,3} |\delta_i^{(2)}| = 0,1792$  є у третьому рівнянні. Тому знайдемо

$$x_3^{(2)} \text{ за формулою } x_3^{(2)} = 0,1x_1^{(1)} + 0,1x_2^{(1)} + 0,8 = 0,9792.$$

Підставимо його в перші два рівняння і знайдемо відповідні відхили

$$\begin{aligned} \delta_1^{(2)} &= -x_1^{(1)} + 0,2x_2^{(1)} + 0,2x_3^{(2)} + 0,6 = -0,932 + 0,2 \cdot 0,86 + \\ &\quad + 0,2 \cdot 0,9792 + 0,6 = 0,03584; \\ \delta_2^{(2)} &= 0,1x_1^{(1)} - x_2^{(1)} + 0,2x_3^{(2)} + 0,7 = 0,1 \cdot 0,932 - 0,86 + \\ &\quad + 0,2 \cdot 0,9792 + 0,7 = 0,12904. \end{aligned}$$

Тоді  $\max_{i=1,2} |\delta_i^{(2)}| = 0,12904$  у другому рівнянні,

$$x_2^{(2)} = 0,1x_1^{(1)} + 0,2x_3^{(2)} + 0,7 = 0,1 \cdot 0,86 + 0,2 \cdot 0,9792 + 0,7 = 0,98184.$$

Знайдемо відхил у першому рівнянні

$$\begin{aligned} \delta_1^{(2)} &= -x_1^{(1)} + 0,2x_2^{(2)} + 0,2x_3^{(2)} + 0,6 = -0,932 + 0,2 \times \\ &\quad \times 0,9792 + 0,6 = 0,06021. \end{aligned}$$

Отже,  $x_1^{(2)} = 0,99221$ .

Знайдемо друге наближення до розв'язку

$$x^{(2)} = (0,99221; 0,98184; 0,97920);$$

$$\max_{i=1,3} |x_i^{(2)} - x_i^{(1)}| = 0,1792.$$

Точності не досягнуто, бо не виконується умова  $\|x^{(2)} - x^{(1)}\| \leq \frac{1-q}{q} \varepsilon$ . Тому виконаємо ще одну ітерацію. Шукаємо

відхили

$$\begin{cases} \delta_1^{(3)} = -x_1^{(2)} + 0,2x_2^{(2)} + 0,2x_3^{(2)} + 0,6 = -0,99211 + 0,2 \cdot 0,98184 + \\ \quad + 0,2 \cdot 0,9792 + 0,6 = -0,000002; \\ \delta_2^{(3)} = 0,1x_1^{(2)} - x_2^{(2)} + 0,2x_3^{(2)} + 0,7 = 0,1 \cdot 0,99221 - 0,98184 + 0,2 \cdot 0,9792 + \\ \quad + 0,7 = 0,013221; \\ \delta_3^{(3)} = 0,1x_1^{(2)} + 0,1x_2^{(2)} - x_3^{(2)} + 0,8 = 0,1 \cdot 0,99221 + 0,1 \cdot 0,98184 - 0,9792 + \\ \quad + 0,8 = 0,018205. \end{cases}$$

Оскільки  $\max_{i=1,3} |\delta_i^{(3)}| = 0,018205$  у третьому рівнянні, то уточнимо спочатку  $x_3$

$$\begin{aligned} x_3^{(3)} &= 0,1x_1^{(2)} + 0,1x_2^{(2)} + 0,8 = 0,1 \cdot 0,99221 + 0,1 \cdot 0,98184 + \\ &\quad + 0,8 = 0,997405 \approx 0,99741. \end{aligned}$$

Знайдемо відхили в першому і другому рівнянні

$$\begin{aligned} \delta_1^{(3)} &= -x_1^{(2)} + 0,2x_2^{(2)} + 0,2x_3^{(3)} + 0,6 = -0,99211 + 0,2 \cdot 0,98184 + \\ &\quad + 0,2 \cdot 0,99741 + 0,6 = 0,00364; \\ \delta_2^{(3)} &= 0,1x_1^{(2)} - x_2^{(2)} + 0,2x_3^{(3)} + 0,7 = 0,1 \cdot 0,99221 - 0,98184 + 0,2 \cdot 0,99741 + \\ &\quad + 0,7 = 0,01686. \end{aligned}$$

Більший відхил у другому рівнянні, тому уточнимо  $x_2$

$$\begin{aligned} x_2^{(3)} &= 0,1x_1^{(2)} + 0,2x_3^{(3)} + 0,7 = 0,1 \cdot 0,99221 + 0,2 \cdot 0,99741 + \\ &\quad + 0,7 = 0,998703. \end{aligned}$$

Знайдемо відхил першого рівняння

$$\delta_1^{(3)} = -x_1^{(2)} + 0,2x_2^{(3)} + 0,2x_3^{(3)} + 0,6 = -0,99211 + 0,2 \cdot 0,99870 + \\ + 0,2 \cdot 0,99741 + 0,6 = 0,0070126.$$

Отже,

$$x_1^{(3)} = 0,9992226 \approx 0,99922 \mid x^{(3)} = (0,99922; 0,99870; 0,99741).$$

Перевіримо, чи досягнуто потрібної точності:

$$\|x^{(3)} - x^{(2)}\| = 0,01821 > \frac{1-q}{q} \varepsilon.$$

Ні, тому зробимо ще один крок. Обчислимо відхили

$$\left\{ \begin{array}{l} \delta_1^{(4)} = -x_1^{(3)} + 0,2x_2^{(3)} + 0,2x_3^{(3)} + 0,6 = -0,99922 + 0,2 \cdot 0,99870 + \\ \quad + 0,2 \cdot 0,99741 + 0,6 = 0,000002; \\ \delta_2^{(4)} = 0,1x_1^{(3)} - x_2^{(3)} + 0,2x_3^{(3)} + 0,7 = 0,1 \cdot 0,99922 - 0,9987 + 0,2 \cdot 0,99741 + \\ \quad + 0,7 = 0,000704; \\ \delta_3^{(4)} = 0,1x_1^{(3)} + 0,1x_2^{(3)} - x_3^{(3)} + 0,8 = 0,1 \cdot 0,99922 + 0,1 \cdot 0,9987 - 0,99741 + \\ \quad + 0,8 = 0,002382. \end{array} \right.$$

Найбільший відхил у третьому рівнянні, тому уточнимо  $x_3$

$$x_3^{(4)} = 0,1x_1^{(3)} + 0,1x_2^{(3)} + 0,8 = 0,1 \cdot 0,99922 + 0,1 \cdot 0,9987 + \\ + 0,8 = 0,99979.$$

Знаходимо відхили першого і другого рівнянь

$$\delta_1^{(4)} = -x_1^{(3)} + 0,2x_2^{(3)} + 0,2x_3^{(4)} + 0,6 = -0,99922 + 0,2 \cdot 0,9987 + \\ + 0,2 \cdot 0,99979 + 0,6 = 0,000478;$$

$$\delta_2^{(4)} = 0,1x_1^{(3)} - x_2^{(3)} + 0,2x_3^{(4)} + 0,7 = 0,1 \cdot 0,99922 - 0,9987 + \\ + 0,2 \cdot 0,99979 + 0,7 = 0,00118.$$

Знову найбільший відхил у другому рівнянні, тому

$$x_2^{(4)} = x_2^{(3)} + \delta_2^{(4)} = 0,9987 + 0,00118 = 0,99988.$$



Визначимо першу компоненту вектора  $x^{(4)}$

$$x_1^{(4)} = 0,2x_2^{(4)} + 0,2x_3^{(4)} + 0,6 = 0,2 \cdot 0,99988 + 0,2 \cdot 0,99979 + 0,6 = 0,999934.$$

Четверте наближення до розв'язку

$$x^{(4)} = (0,99993; 0,99988; 0,99979).$$

Перевіримо точність:

$$\max_{i=1,3} |x_i^{(4)} - x_i^{(3)}| = \max \{0,0007; 0,001; 0,002\} = 0,002 < \frac{1-q}{q} \varepsilon.$$

З точністю до  $\varepsilon = 0,01$  розв'язок системи  $x^* = (1,00; 1,00; 1,00)$ .

Розглянемо метод Зейделя як **варіант методу Якобі**.

Запишемо матрицю системи  $Ax=b$  у вигляді

$$A = L + D + R,$$

де  $D$  – діагональна матриця,  $L$  і  $R$  – відповідно, ліва і права трикутні матриці з нульовими діагональними елементами (строго нижня і строго верхня трикутні матриці). Уважатимемо, що всі діагональні елементи  $a_{ii}$  відмінні від нуля, а отже, матриця  $D + L$  має обернену. Тоді метод Зейделя записуємо у вигляді

$$x^{(k+1)} = -(L + D)^{-1} R x^{(k)} + (L + D)^{-1} b, \quad (3.13)$$

$$k = 0, 1, 2, \dots$$

або в компонентному вигляді

$$\begin{cases} x_1^{(k+1)} = -(a_{12}x_2^{(k)} + a_{13}x_3^{(k)} + \dots + a_{1n}x_n^{(k)} - b_1) / a_{11}; \\ x_2^{(k+1)} = -(a_{21}x_1^{(k+1)} + a_{23}x_3^{(k)} + \dots + a_{2n}x_n^{(k)} - b_2) / a_{22}; \\ \vdots \\ x_n^{(k+1)} = -(a_{n1}x_1^{(k+1)} + a_{n2}x_2^{(k+1)} + \dots + a_{nn-1}x_{n-1}^{(k+1)} - b_n) / a_{nn}. \end{cases}$$

**Теорема 3.4.** Для збіжності методу Зейделя (3.13) для довільного вибору початкового наближення  $x^{(0)}$  і довільного значення  $b$  необхідно і достатньо, щоб усі корені рівняння

$$\begin{vmatrix} a_{11}\lambda & a_{12} & a_{1n} \\ a_{21}\lambda & a_{22}\lambda & a_{2n} \\ a_{n1}\lambda & a_{n2}\lambda & a_{nn}\lambda \end{vmatrix} = 0$$

були за модулем менші від одиниці.

**П р и к л а д.** Знайти область збіжності методу Зейделя для системи  $Ax = b$  з матрицею

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 3 & 2 \\ -1 & 1 & 1 \end{pmatrix}.$$

**Розв'язування.** За теоремою 3.4 про необхідні й достатні умови збіжності методу Зейделя складемо рівняння

$$\begin{vmatrix} 4\lambda & 2 & 2 \\ 2\lambda & 3\lambda & 2 \\ -\lambda & \lambda & \lambda \end{vmatrix} = 12\lambda^3 + 4\lambda^2 - 4\lambda + 6\lambda^2 - 8\lambda^2 - 4\lambda^2 =$$

$$= 12\lambda^3 - 2\lambda^2 - 4\lambda = 2\lambda(6\lambda^2 - \lambda - 2) = 0.$$

$$\lambda_1 = 0; \quad \lambda_2 = \frac{2}{3}; \quad \lambda_3 = -\frac{1}{2}$$

Отже, метод Зейделя збіжний для довільного початкового наближення  $x^{(0)}$  і довільного значення  $b$ , оскільки  $|\lambda_i| < 1$ ,  $i = 1, 3$ .

### 3.4. Метод релаксації

Часто для пришвидшення збіжності ітераційних методів використовують параметри. Метод, що використовує параметр, називають методом релаксації

$$x_i^{(k+1)} = (1-\omega)x_i^{(k)} + \frac{\omega}{a_{ii}}(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)}). \quad (3.14)$$

Тут початкова матриця адитивно зображена у вигляді  $A=L+D+R$ . Метод релаксації у векторно-матричній формі має вигляд

$$(D + \omega L)x^{(k+1)} = (1 - \omega)Dx^{(k)} - \omega R x^{(k)} + \omega b.$$

Для значень  $\omega \in (1, 2)$  метод (3.14) називають методом послідовної верхньої релаксації (ПВР, або SOR-методом), для  $\omega \in (0, 1)$  – методом нижньої релаксації (він не є ефективним).

Ітераційний процес (3.14) доцільно завершити в разі виконання умови

$$\frac{\|x^{(k+1)} - x^{(k)}\|}{\|x^{(k)}\|} \leq \frac{\varepsilon}{1 - \varepsilon} \cdot \frac{1}{\|A^{-1}\| \cdot \|L + \frac{1}{\omega} D\|}.$$

Для нормальних систем ( $A$  – додатно визначена симетрична матриця) метод релаксації збігається для довільного початкового наближення  $x^{(0)}$  і довільного  $\omega \in (0, 2)$ .

**П р и к л а д 3.4.** Методом ПВР з точністю  $\varepsilon = 10^{-3}$  розв'язати систему (3.10).

*Розв'язування:* Запишемо систему рівнянь у вигляді (3.14)

$$\left\{ \begin{aligned} x_1^{(k+1)} &= (1 - \omega)x_1^{(k)} + \frac{\omega}{a_{11}}(b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - a_{14}x_4^{(k)}) = \\ &= (1 - \omega)x_1^{(k)} + \frac{\omega}{8}(9 - x_2^{(k)} - x_3^{(k)} + x_4^{(k)}); \\ x_2^{(k+1)} &= (1 - \omega)x_2^{(k)} + \frac{\omega}{a_{22}}(b_2 - a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} - a_{24}x_4^{(k)}) = \\ &= (1 - \omega)x_2^{(k)} + \frac{\omega}{5}(4 - x_1^{(k+1)} + x_3^{(k)} + x_4^{(k)}); \\ x_3^{(k+1)} &= (1 - \omega)x_3^{(k)} + \frac{\omega}{a_{33}}(b_3 - a_{31}x_1^{(k+1)} - a_{32}x_2^{(k+1)} - a_{34}x_4^{(k)}) = \\ &= (1 - \omega)x_3^{(k)} + \frac{\omega}{5}(18 - x_1^{(k+1)} + x_2^{(k+1)} - x_4^{(k)}); \\ x_4^{(k+1)} &= (1 - \omega)x_4^{(k)} + \frac{\omega}{a_{44}}(b_4 - a_{41}x_1^{(k+1)} - a_{42}x_2^{(k+1)} - a_{43}x_3^{(k+1)}) = \\ &= (1 - \omega)x_4^{(k)} + \frac{\omega}{10}(41 - 2x_1^{(k+1)} - x_2^{(k+1)} + x_3^{(k+1)}). \end{aligned} \right.$$

Виберемо за початкове наближення нуль-вектор  $x^{(0)} = (0; 0; 0; 0)$ . Результати обчислень для різних значень параметра релаксації наведені в табл. 3.2.

Таблиця 3.2

**Залежність між параметрами релаксації і кількістю ітерацій**

$\varepsilon = 10^{-3}$	$\omega$	Кількість ітерацій
- " -	0,3	25
- " -	0,5	15
- " -	0,8	8
- " -	0,9	6
- " -	1,1	8
- " -	1,5	59
- " -	1,6	Розбіжний
- " -	0,85	7
- " -	0,89	7
- " -	0,91	6
- " -	0,95	6
$\varepsilon = 10^{-2}$	0,9	5
- " -	0,8	6
- " -	1,1	6
- " -	1,3	12
$\varepsilon = 10^{-4}$	0,8	10
- " -	0,9	8
- " -	1	7
- " -	1,1	10

Для нашого прикладу оптимальним за кількістю ітерацій є значення параметра  $\omega = 0,9$

Для  $\varepsilon = 10^{-3}$  і  $\omega = 0,9$  отримані наближення до розв'язку наведені в табл. 3.3.

Таблиця 3.3

## Послідовність наближень розв'язків методу ПВР

Ітерації	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$
1	1,0125	0,53775	3,15454	3,74326
2	1,11948	1,81387	3,00666	3,97017
3	1,02878	1,97204	2,99582	3,99398
4	1,00582	1,99432	2,99860	3,99874
5	1,00124	1,99873	2,99964	3,99973
6	1,00028	1,99971	2,99991	3,99994

Точний розв'язок  $x^* = (1,000; 2,000; 3,000; 4,000)$ .

### 3.5. Ітераційні методи варіаційного типу

До розв'язування системи лінійних алгебричних рівнянь  $Ax = b$  застосуємо ітераційні методи вигляду

$$B \frac{x^{(k+1)} - x^{(k)}}{\tau_{k+1}} + Ax^{(k)} = b, \quad (3.15)$$

у яких параметр  $\tau_{k+1}$  вибирають з умови мінімуму похибки  $\|x^{(k+1)} - x^*\|_D$ . Тут  $D$  – задана симетрична, додатно визначена матриця, для якої  $\|v\|_D = \sqrt{(Dv, v)}$ . Залежно від вибору матриць  $D$  і  $B$  отримують різні ітераційні процеси.

#### 3.5.1. Метод мінімальних відхилів

Нехай матриця  $A$  – симетрична, додатно визначена. Прийmemo за

$$r_k = Ax^{(k)} - b \quad (3.16)$$

відхил, який отримують унаслідок підстановки  $k$ -го наближення до розв'язку в початкову систему. Очевидно, що похибка і відхил пов'язані рівністю

$$A \cdot Z_k = r_k,$$

де  $Z_k = x^{(k)} - x^*$

Розглянемо явний ітераційний процес ( $B = E$ )

$$\frac{x^{(k+1)} - x^{(k)}}{\tau_{k+1}} + Ax^{(k)} = b \quad (3.17)$$

і перетворимо його до вигляду

$$x^{(k+1)} = x^{(k)} - \tau_{k+1} r_k. \quad (3.18)$$

У методі мінімальних відхилів (3.17) параметр  $\tau_{k+1}$  вибирають з умови мінімуму  $\|r_{k+1}\|$  за заданої норми  $\|r_k\|$ , тоді

$$\tau_{k+1} = \frac{(Ar_k, r_k)}{\|Ar_k\|^2}. \quad (3.19)$$

Отже, у цьому методі за значенням  $x^{(k)}$  обчислюють вектор відхилів  $r_k = Ax^{(k)} - b$ , потім за формулою (3.19) знаходять параметр  $\tau_{k+1}$  і остаточно за формулою (3.18) шукають наступне наближення до розв'язку.

Для похибки методу мінімальних відхилів виконується оцінка

$$\|A(x^{(n)} - x^*)\| \leq \rho_0^n \|A(x^{(0)} - x^*)\|, \quad n = 0, 1, \dots,$$

де  $\rho_0 = \frac{1-\xi}{1+\xi}$ ,  $\xi = \frac{\lambda_{\min}(A)}{\lambda_{\max}(A)}$ ,  $\lambda_{\min}(A)$ ,  $\lambda_{\max}(A)$  – найменше та найбільше власне значення матриці  $A$ .

**П р и к л а д 3.5.** Методом мінімальних відхилів розв'язати з точністю до  $\varepsilon = 10^{-2}$  систему

$$\begin{cases} 8x_1 + x_2 + x_3 = 13; \\ x_1 + 5x_2 - x_3 = 8; \\ x_1 - x_2 + 5x_3 = 14. \end{cases}$$

*Розв'язування.* Виберемо за початкове наближення вектор вільних членів  $x^{(0)} = (13; 8; 14)$ . Тоді за формулою (3.16) знайдемо початковий відхил

$$r_0 = Ax^{(0)} - b = \begin{pmatrix} 8 & 1 & 1 \\ 1 & 5 & -1 \\ 1 & -1 & 5 \end{pmatrix} \begin{pmatrix} 13 \\ 8 \\ 14 \end{pmatrix} - \begin{pmatrix} 13 \\ 8 \\ 14 \end{pmatrix} = \begin{pmatrix} 113 \\ 31 \\ 61 \end{pmatrix}.$$

Обчислимо  $\tau_1 = \frac{(Ar_0, r_0)}{\|Ar_0\|^2}$ ;

$$Ar_0 = \begin{pmatrix} 8 & 1 & 1 \\ 1 & 5 & -1 \\ 1 & -1 & 5 \end{pmatrix} \begin{pmatrix} 113 \\ 31 \\ 61 \end{pmatrix} = \begin{pmatrix} 996 \\ 207 \\ 387 \end{pmatrix}; \quad \|Ar_0\| = \max |Ar_0| = 996;$$

$$(Ar_0, r_0) = \left( \begin{pmatrix} 996 \\ 207 \\ 387 \end{pmatrix}, \begin{pmatrix} 113 \\ 31 \\ 61 \end{pmatrix} \right) = 1,4257 \cdot 10^5; \quad \tau_1 = 0,1437.$$

Тоді  $x^{(1)} = x^{(0)} - \tau_1 r_0 = \begin{pmatrix} 13 \\ 8 \\ 14 \end{pmatrix} - 0,1437 \begin{pmatrix} 113 \\ 31 \\ 61 \end{pmatrix} = \begin{pmatrix} -3,2403 \\ 3,5447 \\ 5,2331 \end{pmatrix}.$

Наступний відхил

$$r_1 = Ax^{(1)} - b = \begin{pmatrix} 8 & 1 & 1 \\ 1 & 5 & -1 \\ 1 & -1 & 5 \end{pmatrix} \begin{pmatrix} -3,24 \\ 3,54 \\ 5,23 \end{pmatrix} - \begin{pmatrix} 13 \\ 8 \\ 14 \end{pmatrix} = \begin{pmatrix} -30,1446 \\ 1,2501 \\ 5,3806 \end{pmatrix};$$

$$Ar_1 = \begin{pmatrix} 8 & 1 & 1 \\ 1 & 5 & -1 \\ 1 & -1 & 5 \end{pmatrix} \begin{pmatrix} -30,1446 \\ 1,2501 \\ 5,3806 \end{pmatrix} = \begin{pmatrix} -234,53 \\ -29,27 \\ -4,49 \end{pmatrix};$$

$$\|Ar_1\| = 234,53; \quad (Ar_1, r_1) = 7 \cdot 10^3; \quad \tau_2 = 0,1274.$$

Друге наближення до розв'язку

$$x^{(2)} = x^{(1)} - \tau_2 r_1 = \begin{pmatrix} -3,2403 \\ 3,5447 \\ 5,2331 \end{pmatrix} - 0,1274 \begin{pmatrix} -30,1446 \\ 1,2501 \\ 5,3806 \end{pmatrix} = \begin{pmatrix} -0,6010 \\ 3,3854 \\ 4,5475 \end{pmatrix}.$$

Знову знайдемо відхил

$$r_2 = Ax^{(2)} - b = \begin{pmatrix} 8 & 1 & 1 \\ 1 & 5 & -1 \\ 1 & -1 & 5 \end{pmatrix} \begin{pmatrix} 0,601 \\ 3,385 \\ 4,548 \end{pmatrix} - \begin{pmatrix} 13 \\ 8 \\ 14 \end{pmatrix} = \begin{pmatrix} -0,2591 \\ 4,9805 \\ 5,9530 \end{pmatrix}.$$

Для обчислення  $\tau_3$  знайдемо

$$Ar_2 = \begin{pmatrix} 8 & 1 & 1 \\ 1 & 5 & -1 \\ 1 & -1 & 5 \end{pmatrix} \begin{pmatrix} -0,2591 \\ 4,9805 \\ 5,9530 \end{pmatrix} = \begin{pmatrix} 8,8608 \\ 18,6906 \\ 24,5252 \end{pmatrix};$$

$$\|Ar_2\| = 24,5252; \quad (Ar_2, r_2) = 2,368 \cdot 10^2;$$

тоді  $\tau_3 = 0,3937$ .

Нарешті за формулою (3.18) маємо

$$x^{(3)} = (0,7030; 1,42478; 2,2039).$$

Аналогічно  $r_4 = (-3,7474; -2,3776; -3,7021)$ ;  $\tau_4 = 0,1823$ ;

$$x^{(4)} = (1,3863; 1,8582; 2,8790).$$

Обчислені значення та наступні наближення наведено в табл. 3.4.

Таблиця 3.4

Значення наближеного розв'язку на різних ітераціях

Ітерації	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\tau_{k+1}$
0	13	8	14	0,1437
1	-3,2403	3,5447	5,2331	0,1274
2	0,6010	3,3854	4,5475	0,3937
3	0,7030	1,4247	2,2039	0,1823
4	1,3863	1,8582	2,8790	0,1254
5	1,0318	1,8835	2,8886	0,4514
6	1,0199	2,0819	3,0731	0,2069
7	0,9545	2,0082	3,0103	0,1251
8	0,9977	2,0100	3,0105	0,4474
9	0,9967	1,9934	2,9925	0,1752
10	1,0038	1,9984	2,9985	

### 3.5.2. Метод мінімальних поправок

Ітераційний процес за формулами (3.15) перепишемо у вигляді

$$x^{(k+1)} = x^{(k)} - \tau_{k+1} B^{-1} r_k, \quad (3.21)$$

де  $r_k$  – відхил.



Вектор  $\omega_k = B^{-1}r_k$  називають *поправкою*  $(k+1)$ -ї ітерації. Поправки  $\omega_k$  задовольняють таке ж рівняння, що і похибка  $\varepsilon_k = x^{(k)} - x^*$  тобто рівняння (3.15):

$$B \frac{\omega_{k+1} - \omega_k}{\tau_{k+1}} + A\omega_k = 0. \quad (3.22)$$

Для  $B$  – симетричної, додатно визначеної матриці – методом мінімальних поправок називають неявний ітераційний метод (3.15), у якому параметр  $\tau_{k+1}$  вибирають з умови мінімуму норми  $\|\omega_{k+1}\|_B^2 = (B\omega_{k+1}, \omega_{k+1})$  для заданого вектора  $\omega_k$ . Коли  $B = E$ , то метод мінімальних поправок збігається з методом мінімальних відхилів.

Норма  $\|\omega_{k+1}\|_B^2$  буде мінімальною при

$$\tau_{k+1} = \frac{(A\omega_k, \omega_k)}{(B^{-1}A\omega_k, A\omega_k)} \quad (3.23)$$

Для похибки методу мінімальних поправок виконується оцінка

$$\|(x^{(n)} - x^*)\|_B \leq \rho_0^n \|(x^{(0)} - x^*)\|_B, \quad n = 0, 1, \dots,$$

де 
$$\rho_0 = \frac{1 - \xi}{1 + \xi}, \quad \xi = \frac{\lambda_{\min}(B^{-1}A)}{\lambda_{\max}(B^{-1}A)},$$

Ітераційний процес методу мінімальних поправок на кожній ітерації складається з таких етапів:

- 1) зі СЛАР  $B\omega_k = r_k$  знаходять поправки  $\omega_k$ ;
- 2) зі СЛАР  $B\dot{v}_k = A\omega_k$  визначають вектор  $v_k = B^{-1}A\omega_k$ ;
- 3) за формулою (3.23) знаходять параметр  $\tau_{k+1}$ ;
- 4) за формулою (3.21) обчислюють наступне наближення  $x^{(k+1)}$

**П р и к л а д 3.6.** Методом мінімальних поправок з точністю до  $10^{-2}$  розв'язати систему (3.20).

*Розв'язування:* Виберемо матрицю  $B$  у вигляді

$$B = \begin{pmatrix} 3,1 & 1 & 2,1 \\ 1 & 3,6 & 2,1 \\ 2,1 & 2,1 & 4,1 \end{pmatrix}. \text{ Як і в попередньому прикладі, за початкове}$$

наближення до розв'язку беремо вектор  $x^{(0)} = (13, 8, 14)$ . Знайдемо відхил  $r_0 = Ax^{(0)} - b = (113, 31, 61)$ . Розв'яжемо систему  $B\omega_k = r_k$

$$\begin{pmatrix} 3,1 & 1 & 2,1 \\ 1 & 3,6 & 2,1 \\ 2,1 & 2,1 & 4,1 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \begin{pmatrix} 113 \\ 31 \\ 61 \end{pmatrix}; \quad \omega_0 = (47,252; 7,506; 4,547);$$

$$A\omega_0 = \begin{pmatrix} 8 & 1 & 1 \\ 1 & 5 & -1 \\ 1 & -1 & 5 \end{pmatrix} \begin{pmatrix} 47,252 \\ 7,506 \\ 4,547 \end{pmatrix} = \begin{pmatrix} 390,072 \\ 80,234 \\ 62,482 \end{pmatrix}.$$

Згідно з етапом 2, розв'яжемо систему  $Bv_0 = A\omega_0$ :

$$\begin{pmatrix} 3,1 & 1 & 2,1 \\ 1 & 3,6 & 2,1 \\ 2,1 & 2,1 & 4,1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 390,072 \\ 80,234 \\ 62,482 \end{pmatrix}; \quad v_0 = (188,282; 35,599; -50,192).$$

За формулами (3.23) знайдемо  $\tau_1 = \frac{(A\omega_0, \omega_0)}{(v_0, A\omega_0)} = 0,264$ . Тоді

перше наближення шуканого розв'язку

$$x^{(1)} = x^{(0)} - \tau_1 B^{-1} r_0 = \begin{pmatrix} 13 \\ 8 \\ 14 \end{pmatrix} - 0,264 \begin{pmatrix} 47,252 \\ 7,506 \\ 4,547 \end{pmatrix} = \begin{pmatrix} 0,524 \\ 6,018 \\ 12,800 \end{pmatrix}.$$

Виконаємо ще одну ітерацію.

$r_1 = Ax^{(1)} - b = (10,007; 9,815; 44,503)$ . З рівняння  $B\omega_1 = r_1$  знайдемо

$$A\omega_1 = (-3,787; -29,730; 88,429).$$

Тоді зі СЛАР  $Bv_1 = A\omega_1$  визначимо

$$v_1 = (-18,270; -24,707; 46,401). \quad \text{Далі} \quad \tau_2 = \frac{(A\omega_1, \omega_1)}{(v_1, A\omega_1)} = 0,334.$$

Наступне наближення

$$x^{(2)} = x^{(1)} - \tau_2 B^{-1} r_1 = \begin{pmatrix} 0,524 \\ 6,018 \\ 12,800 \end{pmatrix} - 0,334 \begin{pmatrix} -2,462 \\ -1,894 \\ 17,799 \end{pmatrix} = \begin{pmatrix} 1,346 \\ 6,651 \\ 6,852 \end{pmatrix}.$$

Продовжимо цей процес, на 13-й ітерації отримаємо розв'язок  $x^{(11)} = (0,993; 2,020; 3,025)$ .

Отже, за розв'язок системи можна прийняти вектор

$$x^* = (0,99; 2,02; 3,03).$$

Швидкість збіжності методу залежить від вибору допоміжної матриці  $B$

### 3.5.3. Метод спряжених градієнтів

Як і в попередніх методах, уважатимемо, що матриця  $A$  є симетричною і додатно визначеною. Виберемо початкове наближення  $x^{(0)}$ , побудуємо послідовності векторів

$$x^{(n)} = x^{(0)} + \sum_{k=0}^{n-1} \alpha_k A^k (Ax^{(0)} - b). \quad (3.24)$$

Завдяки вибору параметрів  $\alpha_k$  мінімізуємо величину

$$F(x^{(n)}) = (Ax^{(n)}, x^{(n)}) - 2(b, x^{(n)}).$$

Алгоритм методу спряжених градієнтів (МСГ) такий.

1. Вибираємо початкове наближення  $x^{(0)}$ . Обчислюємо відхил  $r_0 = Ax^{(0)} - b$ .
2. Будуємо наступні наближення розв'язків за схемою ( $p_0 = r_0, k = 0$ ):
  - 2.1)  $\alpha_k = \frac{(r_k, p_k)}{(p_k, Ap_k)}$ ;
  - 2.2)  $x^{(k+1)} = x^{(k)} + \alpha_k p_k$ ;
  - 2.3)  $r_{k+1} = r_k - \alpha_k Ap_k$ ;
  - 2.4) перевірка збіжності:  $\|r_{k+1}\| \leq \varepsilon$  – ітерації припинити, інакше продовжити обчислення.
3. Підготовка до наступної ітерації:
  - 3.1)  $\beta_k = \frac{(r_{k+1}, p_k)}{(p_k, Ap_k)}$ ;
  - 3.2)  $p_{k+1} = r_{k+1} - \beta_k p_k$ ;
  - 3.3)  $k := k + 1$  – перехід на наступну ітерацію з пункту 2.1.

Центральною частиною алгоритму є визначення напрямів  $p_k$ . Вони повинні задовольняти умови

$$(p_i, Ap_j) = 0, \quad i \neq j; \quad i, j = \overline{1, n-1}.$$

Такі напрями називають спряженими щодо матриці  $A$ . Хоча метод МСГ належить до прямих методів (згідно з (3.24)), однак на практиці розв'язок отримують значно швидше, ніж за  $n$  ітерацій.

Для цього методу правильна оцінка похибки

$$\|x^{(n)} - x^*\|_A \leq q_n \|x^{(0)} - x^*\|_A,$$

де  $q_n = \frac{2\rho_1^n}{1+\rho_1^{2n}}, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \frac{\lambda_{\min}(A)}{\lambda_{\max}(A)}.$

**П р и к л а д 3.7.** Методом СГ з точністю до  $10^{-5}$  розв'язати СЛАР:

$$\begin{pmatrix} 16 & 0 & 1 & 2 & -3 \\ 0 & 20 & 4 & -4 & 1 \\ 1 & 4 & 10 & 0 & 1 \\ 2 & -4 & 0 & 20 & 2 \\ -3 & 1 & 1 & 2 & 18 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 12 \\ 41 \\ 44 \\ 90 \\ 100 \end{pmatrix}.$$

*Розв'язування.* Виберемо початкове наближення

$x^{(0)} = (1, 1, 1, 1, 1)$ . Згідно з алгоритмом, при  $k = 0$

$$r_0 = b - Ax^{(0)} = \begin{pmatrix} 12 \\ 41 \\ 44 \\ 90 \\ 100 \end{pmatrix} - \begin{pmatrix} 16 & 0 & 1 & 2 & -3 \\ 0 & 20 & 4 & -4 & 1 \\ 1 & 4 & 10 & 0 & 1 \\ 2 & -4 & 0 & 20 & 2 \\ -3 & 1 & 1 & 2 & 18 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -4 \\ 20 \\ 28 \\ 67 \\ 81 \end{pmatrix}.$$

Прийmemo  $p_0 = r_0$ . Для обчислення  $\alpha_k$  знайдемо

$$Ap_0 = \begin{pmatrix} 16 & 0 & 1 & 2 & -3 \\ 0 & 20 & 4 & -4 & 1 \\ 1 & 4 & 10 & 0 & 1 \\ 2 & -4 & 0 & 20 & 2 \\ -3 & 1 & 1 & 2 & 18 \end{pmatrix} \begin{pmatrix} -4 \\ 20 \\ 28 \\ 67 \\ 81 \end{pmatrix} = \begin{pmatrix} -145 \\ 325 \\ 437 \\ 1474 \\ 1652 \end{pmatrix};$$

$$\begin{aligned}\alpha_0 &= \frac{(r_0, p_0)}{(p_0, Ap_0)} = \\ &= \frac{((-4, 20, 28, 67, 81), (-4, 20, 28, 67, 81))}{((-4, 20, 28, 67, 81), (-145, 325, 437, 1474, 1652))} = \\ &= \frac{12250}{251886} \approx 0,0486.\end{aligned}$$

$$\text{Тоді } x^{(1)} = x^{(0)} + \alpha_0 p_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + 0,0486 \begin{pmatrix} -4 \\ 20 \\ 28 \\ 67 \\ 81 \end{pmatrix} = \begin{pmatrix} 0,8055 \\ 1,9727 \\ 2,3617 \\ 4,2584 \\ 4,9393 \end{pmatrix}.$$

Подальші обчислення наведені в табл. 3.5.

Таблиця 3.5

Наближені розв'язки за методом спряжених градієнтів

$k$	$x^{(k)}$	$\alpha_k - \frac{(r_k, p_k)}{(p_k, Ap_k)}$
1	2	3
0	(1, 1, 1, 1, 1)	0,0486
1	(0,805468; 1,972662; 2,361727; 4,25818; 4,939282)	0,055722
2	(0,975519; 2,206373; 2,737701; 3,997350; 4,975953)	0,069157
3	(1,016125; 1,994419; 2,865367; 4,022000; 5,005043)	0,059706
4	(1,007035; 2,038190; 2,945820; 3,992869; 5,008256)	0,068579
5	(1,005708; 1,998150; 2,971451; 4,003171; 5,001587)	0,060173
6	(1,001835; 2,007916; 2,988636; 3,998366; 5,002346)	0,068492
7	(1,001308; 1,999577; 2,993964; 4,000574; 5,000291)	0,060293

1	2	3
8	(1,000394; 2,001664; 2,997609; 3,999664; 5,000532)	0,068563
9	(1,000281; 1,999909; 2,998729; 4,000112; 5,000052)	0,060350
10	(1,000082; 2,000350; 2,999498; 3,999931; 5,000115)	0,068627
11	(1,000060; 1,999981; 2,999733; 4,000023; 5,000010)	0,060381
12	(1,000017; 2,000073; 2,999895; 3,999986; 5,000025)	0,068669
13	(1,000013; 1,999996; 2,999944; 4,000005; 5,000002)	0,060399
14	(1,000004; 2,000015; 3,999978; 3,999997; 5,000005)	

Отже, на 14-й ітерації за розв'язок можна прийняти  $x = (1; 2; 3; 4; 5)$ .

## Завдання до розділів 2, 3

### 3.1.

А. Методом Гаусса знайти розв'язки заданих систем лінійних алгебричних рівнянь.

Б. Розв'язати системи ітераційними методами з точністю  $\varepsilon = 10^{-6}$ , попередньо перетворивши їх до вигляду, зручного для ітерацій. Дати порівняльну характеристику цих методів за кількістю ітерацій.

$$1) \begin{cases} 4x_1 + 3x_2 + 2x_3 + x_4 = 12; \\ 3x_1 + 6x_2 + 4x_3 + 2x_4 = 19; \\ 2x_1 + 4x_2 + 6x_3 + 3x_4 = 21; \\ x_1 + 2x_2 + 3x_3 + 4x_4 = 13 \end{cases} \quad x^* = (1, 1, 2, 1);$$

$$2) \begin{cases} x_1 + 2x_2 + x_3 = 6; \\ -x_1 - 2x_2 + 2x_3 = -3; \\ 3x_1 + x_2 + x_3 = 6 \end{cases} \quad x^* = (1, 2, 1);$$

$$3) \begin{cases} x_1 + x_2 + 2x_3 + 3x_4 = 8; \\ 3x_1 - x_2 - x_3 - 2x_4 = 2; \\ 2x_1 + 3x_2 - x_3 - x_4 = 5; \\ x_1 + 2x_2 + 3x_3 - x_4 = 6 \end{cases} \quad x^* = (2, 1, 1, 1);$$

$$4) \begin{cases} 5x_1 + 8x_2 - x_3 = 16; \\ x_1 + 2x_2 + 3x_3 = 10; \\ 2x_1 - 3x_2 + 2x_3 = 5 \end{cases} \quad x^* = (2, 1, 2);$$

$$5) \begin{cases} x_1 + 5x_2 + 3x_3 - 4x_4 = 5; \\ 3x_1 + x_2 - 2x_3 = 2; \\ 5x_1 - 7x_2 + 10x_4 = 8; \\ 3x_2 - 5x_3 = -2 \end{cases} \quad x^* = (1, 1, 1, 1);$$

$$6) \begin{cases} 2x_1 - x_2 - x_3 = 1; \\ 3x_1 + 4x_2 - 2x_3 = 12; \\ 3x_1 - 2x_2 + 4x_3 = 6 \end{cases} \quad x^* = (2, 2, 1);$$

$$7) \quad \begin{cases} x_1 + 3x_2 + 5x_3 + 7x_4 = 22; \\ 3x_1 + 5x_2 + 7x_3 + x_4 = 26; \\ 5x_1 + 7x_2 + x_3 + 3x_4 = 22; \\ 7x_1 + x_2 + 3x_3 + 5x_4 = 26 \end{cases} \quad x^* = (2, 1, 2, 1);$$

$$8) \quad \begin{cases} 4x_1 - 3x_2 + 2x_3 = 9; \\ 2x_1 + 5x_2 - 3x_3 = 3; \\ 5x_1 + 6x_2 - 2x_3 = 12 \end{cases} \quad x^* = (2, 1, 2);$$

$$9) \quad \begin{cases} x_1 - 2x_2 + 3x_3 + 4x_4 = 11; \\ x_1 - 2x_3 + 3x_4 = 6; \\ 3x_1 + 2x_2 - 5x_4 = -2; \\ 4x_1 + 3x_2 - 5x_3 = 6 \end{cases} \quad x^* = (2, 1, 1, 2);$$

$$10) \quad \begin{cases} x_1 + 2x_2 + 4x_3 = 15; \\ 5x_1 + x_2 + 2x_3 = 12; \\ 3x_1 - x_2 + x_3 = 2 \end{cases} \quad x^* = (1, 3, 2);$$

$$11) \quad \begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 10; \\ 2x_1 + x_2 + 2x_3 + 3x_4 = 8; \\ 3x_1 + 2x_2 + x_3 + 2x_4 = 8; \\ 4x_1 + 3x_2 + 2x_3 + x_4 = 10 \end{cases} \quad x^* = (1, 1, 1, 1);$$

$$12) \quad \begin{cases} 3x_1 + 2x_2 + x_3 = 10; \\ 2x_1 + 3x_2 + x_3 = 11; \\ 2x_1 + x_2 + 3x_3 = 13 \end{cases} \quad x^* = (1, 2, 3);$$

$$13) \quad \begin{cases} x_1 + 2x_2 + 3x_3 - 2x_4 = 5; \\ x_1 - x_2 - 2x_3 - 3x_4 = -4; \\ 3x_1 + 2x_2 - x_3 + 2x_4 = 9; \\ 2x_1 - 3x_2 + 2x_3 + x_4 = 4 \end{cases} \quad x^* = (2, 1, 1, 1);$$



$$14) \begin{cases} x_1 + 2x_2 + x_3 = 7; \\ 3x_1 - 5x_2 + 3x_3 = -1; \\ 2x_1 + 7x_2 - x_3 = 17 \end{cases} \quad x^* = (2, 2, 1);$$

$$15) \begin{cases} 2x_1 - x_2 + x_3 - x_4 = 0; \\ 2x_1 - x_2 - 3x_4 = -7; \\ 3x_1 - x_3 + x_4 = 8; \\ 2x_1 + 2x_2 - 2x_3 + 5x_4 = 21 \end{cases} \quad x^* = (2, 2, 1, 3);$$

$$16) \begin{cases} 2x_1 + x_2 - x_3 = 2; \\ x_1 + x_2 + x_3 = 6; \\ 3x_1 - x_2 + x_3 = 8 \end{cases} \quad x^* = (2, 1, 3);$$

$$17) \begin{cases} 4x_1 + x_2 - x_4 = 4; \\ x_1 - 3x_2 + 4x_3 = -1; \\ 3x_2 - 2x_3 + 4x_4 = 12; \\ x_1 + 2x_2 - x_3 - 3x_4 = -2 \end{cases} \quad x^* = (1, 2, 1, 2);$$

$$18) \begin{cases} x_1 + x_2 + x_3 = 6; \\ 2x_1 - x_2 - 6x_3 = -18; \\ 3x_1 - 2x_2 = -1 \end{cases} \quad x^* = (1, 2, 3);$$

$$19) \begin{cases} x_1 + 2x_2 - x_3 + x_4 = 6; \\ 2x_1 + x_2 + x_3 + x_4 = 9; \\ x_1 - x_2 + 2x_3 + x_4 = 3; \\ x_1 + x_2 - x_3 + 3x_4 = 4 \end{cases} \quad x^* = (3, 2, 1, 0);$$

$$20) \begin{cases} 3x_1 - x_2 + x_3 = 7; \\ 2x_1 - 5x_2 - 3x_3 = -7; \\ x_1 - x_2 - x_3 = -1 \end{cases} \quad x^* = (2, 1, 2);$$

$$21) \begin{cases} 2x_1 - x_2 + 3x_3 + 2x_4 = 11; \\ 3x_1 + 3x_2 + 3x_3 + 2x_4 = 16; \\ 3x_1 - x_2 - x_3 + 2x_4 = 4; \\ 3x_1 - x_2 + 3x_3 - x_4 = 6 \end{cases} \quad x^* = (1, 1, 2, 2);$$

$$22) \begin{cases} 3x_1 - x_2 + 4x_3 = 5; \\ -2x_1 + x_2 + x_3 = 1; \\ 2x_1 - x_2 + 4x_3 = 4 \end{cases} \quad x^* = (1, 2, 1);$$

$$23) \begin{cases} 2x_1 + x_2 - 5x_3 + x_4 = 0; \\ x_1 - 3x_2 + x_3 + 6x_4 = 11; \\ 2x_1 + x_2 - x_3 + 2x_4 = 6; \\ x_1 + 4x_2 - 7x_3 + 6x_4 = 10 \end{cases} \quad x^* = (1, 1, 1, 2);$$

$$24) \begin{cases} x_1 + x_2 + 2x_3 = 5; \\ 2x_1 - x_2 + 2x_3 = 5; \\ 4x_1 + x_2 + 4x_3 = 13 \end{cases} \quad x^* = (2, 1, 1);$$

$$25) \begin{cases} x_1 + x_2 - x_3 - x_4 = -2; \\ \quad x_2 + 2x_3 - x_4 = 0; \\ x_1 - x_2 \quad \quad -x_4 = -1; \\ -x_1 + 3x_2 - 2x_3 \quad \quad = -3 \end{cases} \quad x^* = (1, 0, 1, 2);$$

$$26) \begin{cases} 5x_1 + x_2 \quad \quad -x_4 = 7; \\ 3x_1 - 3x_2 + x_3 + 4x_4 = -2; \\ 3x_1 - 2x_2 + x_3 + x_4 = 0; \\ x_1 - 4x_2 \quad \quad + x_4 = -7 \end{cases} \quad x^* = (1, 2, 1, 0);$$

$$27) \begin{cases} 2x_1 - 6x_2 + 2x_3 + 2x_4 = 6; \\ x_1 + 3x_2 + 5x_3 + 7x_4 = 13; \\ 3x_1 + 5x_2 + 7x_3 + x_4 = 11; \\ 5x_1 + 7x_2 + x_3 + 3x_4 = 9 \end{cases} \quad x^* = (1, 0, 1, 1);$$

$$28) \begin{cases} x_1 + x_2 - x_3 - x_4 = 2; \\ x_1 \quad \quad + 2x_3 - 2x_4 = -1; \\ x_1 - x_2 \quad \quad - x_4 = -2; \\ -x_1 + 3x_2 - 2x_3 = 5 \end{cases} \quad x^* = (1, 2, 0, 1);$$

$$29) \begin{cases} 5x_1 + 3x_2 - 7x_3 + 3x_4 = 1; \\ x_1 - 2x_2 \quad \quad - 3x_4 = -2; \\ \quad x_2 - 3x_3 + 4x_4 = 1; \\ 4x_1 + 3x_2 - 5x_3 = -1 \end{cases} \quad x^* = (1, 0, 1, 1);$$

$$30) \begin{cases} -x_1 + x_2 + x_3 + x_4 = 2; \\ 2x_1 + x_2 + 2x_3 + 3x_4 = 8; \\ 3x_1 + 2x_2 + x_3 + 2x_4 = 8; \\ 4x_1 + 3x_2 + 2x_3 + x_4 = 10 \end{cases} \quad x^* = (1, 1, 1, 1).$$

3.2. Методом Гаусса обчислити визначник матриці й знайти обернену матрицю.

$$1) \quad A = \begin{pmatrix} 1,00 & 0,47 & -0,11 & 0,55 \\ 0,42 & 1,00 & 0,35 & 0,17 \\ -0,25 & 0,67 & 1,00 & 0,36 \\ 0,54 & -0,32 & -0,74 & 1,00 \end{pmatrix};$$

$$2) \quad A = \begin{pmatrix} 1 & 1 & 3 & 4 \\ -1 & 0 & 3 & -2 \\ 2 & 1 & 2 & -3 \\ 1 & 2 & -1 & 1 \end{pmatrix};$$

$$3) \quad A = \begin{pmatrix} 2,11 & 3,01 & 4,02 & 0,22 \\ 0,18 & 3,41 & 0,15 & 1,43 \\ 2,14 & 0,17 & 0,26 & 0,18 \\ 1,28 & 0,42 & 0,54 & 1,00 \end{pmatrix};$$

$$4) \quad A = \begin{pmatrix} 1,5 & 2,7 & -1,3 & 5,2 \\ 2,7 & -3,4 & 1,8 & 2,2 \\ -1,3 & 0,16 & 0,82 & 1,05 \\ 5,2 & 2,2 & 1,05 & 3,4 \end{pmatrix};$$

$$5) \quad A = \begin{pmatrix} 1,17 & 2,13 & 0,32 & 0,56 \\ 2,13 & 0,82 & -0,72 & 1,10 \\ 0,32 & 0,25 & -0,42 & 0,16 \\ 0,56 & 1,10 & -0,25 & -0,44 \end{pmatrix};$$

$$6) \quad A = \begin{pmatrix} 1,13 & 2,15 & 0,83 & 0,77 \\ 0,64 & -0,43 & 0,62 & -0,32 \\ 2,32 & 1,15 & 1,84 & 0,68 \\ -0,72 & 0,53 & 0,64 & -0,57 \end{pmatrix};$$

$$7) \quad A = \begin{pmatrix} 0,75 & 0,18 & 0,63 & -0,32 \\ 0,92 & 0,38 & -0,14 & 0,56 \\ 0,63 & -0,42 & 0,18 & 0,37 \\ -0,65 & 0,52 & 0,47 & 0,27 \end{pmatrix};$$

$$8) \quad A = \begin{pmatrix} 0,72 & 3,54 & 7,28 & 0,33 \\ -0,28 & -0,72 & 3,04 & 0,22 \\ 1,00 & 0,35 & -0,78 & 1,00 \\ 7,03 & -5,04 & -3,75 & 3,41 \end{pmatrix};$$

$$9) \quad A = \begin{pmatrix} -0,33 & 0,42 & 0,51 & -0,11 \\ 2,71 & -0,92 & -2,17 & 0,81 \\ 0,75 & 0,68 & 0,33 & 0,17 \\ 0,28 & -3,71 & 2,17 & 0,16 \end{pmatrix};$$

$$10) \quad A = \begin{pmatrix} 0,54 & 0,32 & 1,00 & 0,22 \\ 0,66 & 0,44 & 0,22 & 1,00 \\ 1,00 & 0,42 & 0,54 & 0,66 \\ 0,42 & 1,00 & 0,32 & 0,44 \end{pmatrix};$$

$$11) \quad A = \begin{pmatrix} 2,00 & 0,17 & 3,02 & 0,11 \\ 0,28 & 0,13 & 0,54 & 3,12 \\ 0,54 & 0,18 & 2,11 & 3,08 \\ 2,33 & 0,11 & 0,22 & 2,22 \end{pmatrix};$$

$$12) \quad A = \begin{pmatrix} 4,20 & 0,32 & 0,11 & 0,13 \\ 0,17 & 0,25 & 0,48 & 0,52 \\ 0,12 & 0,08 & 0,72 & 0,61 \\ 0,54 & 0,13 & 0,81 & 0,17 \end{pmatrix};$$

$$13) \quad A = \begin{pmatrix} 1,0 & 2,14 & 0,42 & -1,13 \\ 0,23 & 0,42 & -1,5 & 0,16 \\ 0,34 & -0,12 & 0,18 & 0,57 \\ 0,83 & -0,17 & 0,62 & -0,83 \end{pmatrix};$$

$$14) \quad A = \begin{pmatrix} 1,00 & 0,13 & 0,25 & 0,82 \\ -0,15 & 0,27 & 0,35 & -0,44 \\ 0,83 & 0,11 & 0,72 & -0,32 \\ 0,94 & 0,08 & 0,32 & 0,12 \end{pmatrix};$$

$$15) \quad A = \begin{pmatrix} 1,00 & 0,27 & 0,64 & 0,83 \\ 0,27 & 0,35 & -0,81 & 0,16 \\ 0,64 & -0,81 & -0,14 & 0,15 \\ 0,83 & -0,14 & 0,25 & 0,37 \end{pmatrix};$$

$$16) \quad A = \begin{pmatrix} 0,2 & -7,7 & 3,1 & -1,6 \\ 0,7 & 1,4 & -8,5 & 4,8 \\ 8,2 & -2,3 & 0,3 & -7,9 \\ 0,55 & 1,00 & 0,32 & 0,4 \end{pmatrix};$$

$$17) \quad A = \begin{pmatrix} -2,41 & 7,55 & 0,82 & 0,33 \\ 0,28 & -3,44 & 0,75 & 0,23 \\ 0,17 & 0,28 & 0,05 & 3,48 \\ -1,00 & 0,23 & 2,00 & 7,00 \end{pmatrix};$$

$$18) \quad A = \begin{pmatrix} 4,5 & 4,8 & -3,7 & 2,1 \\ 4,5 & -3,7 & 5,6 & 3,3 \\ 4,8 & 7,5 & 8,3 & 9,2 \\ -1,5 & 2,3 & 4,8 & 3,1 \end{pmatrix};$$

$$19) \quad A = \begin{pmatrix} 1,5 & 0,84 & 0,63 & -0,18 \\ 0,15 & 0,36 & -0,16 & 0,88 \\ -0,27 & 0,45 & 0,64 & -0,38 \\ 0,41 & -0,83 & 0,62 & 0,27 \end{pmatrix};$$

$$20) \quad A = \begin{pmatrix} 0,8 & 1,3 & -0,12 & 0,25 \\ -1,2 & 0,18 & 0,72 & 0,13 \\ 1,6 & 0,2 & 0,12 & -0,11 \\ 1,4 & 0,15 & -0,83 & 0,41 \end{pmatrix};$$

$$21) \quad A = \begin{pmatrix} 2,5 & 0,35 & 0,4 & -0,8 \\ 0,2 & -1,5 & 0,61 & 2,3 \\ 0,16 & -0,42 & 0,57 & 0,63 \\ 0,23 & 0,15 & -0,08 & 3,1 \end{pmatrix};$$

$$22) \quad A = \begin{pmatrix} 5,5 & 3,7 & -8,3 & 9,1 \\ -4,5 & 6,8 & 7,2 & 3,4 \\ 7,5 & -4,9 & 3,5 & 7,1 \\ 5,6 & -4,8 & 7,3 & 5,3 \end{pmatrix};$$

$$23) \quad A = \begin{pmatrix} 1,8 & 1,02 & 1,03 & 1,05 \\ 7,03 & 8,04 & 9,05 & 6,08 \\ 1,11 & -2,02 & 2,03 & -3,04 \\ -3,41 & 4,52 & 7,28 & 5,51 \end{pmatrix};$$

$$24) \quad A = \begin{pmatrix} 8,2 & 1,4 & -2,3 & 0,2 \\ -1,6 & 5,4 & -7,7 & 3,1 \\ 0,7 & 1,9 & -8,5 & 4,8 \\ 5,3 & -5,9 & 2,7 & -7,9 \end{pmatrix};$$

$$25) \quad A = \begin{pmatrix} 1,6 & 5,4 & -7,7 & -3,1 \\ 8,2 & 1,4 & -2,3 & 0,2 \\ 5,3 & -5,9 & 2,7 & -7,9 \\ 0,7 & 1,9 & -8,5 & 4,8 \end{pmatrix};$$

$$26) \quad A = \begin{pmatrix} 1 & 1 & 2 & 3 \\ 3 & -1 & -1 & -2 \\ 2 & 3 & -1 & -1 \\ 1 & 2 & 3 & -1 \end{pmatrix};$$

$$27) \quad A = \begin{pmatrix} 1 & -2 & 3 & 4 \\ 1 & 0 & -2 & 3 \\ 3 & 2 & 0 & -5 \\ 4 & 3 & -5 & 0 \end{pmatrix};$$

$$28) \quad A = \begin{pmatrix} 1 & 3 & 5 & 7 \\ 3 & 5 & 7 & 1 \\ 5 & 7 & 1 & 3 \\ 7 & 1 & 3 & 5 \end{pmatrix};$$

$$29) \quad A = \begin{pmatrix} 1 & 5 & 3 & -4 \\ 3 & 1 & -2 & 0 \\ 5 & -7 & 0 & 10 \\ 0 & 3 & -5 & 0 \end{pmatrix};$$

$$30) \quad A = \begin{pmatrix} 1 & 2 & 3 & -2 \\ 1 & -1 & -2 & -3 \\ 3 & 2 & -1 & 2 \\ 2 & -3 & 2 & 1 \end{pmatrix};$$

$$31) \quad A = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 1 & 1 & 1 \\ 1 & -1 & 2 & 1 \\ 1 & 1 & -1 & 3 \end{pmatrix}.$$

3.3. Системи лінійних алгебричних рівнянь розв'язати такими методами.

А. Методом квадратних коренів і методом ортогоналізації.

Б. Методами варіаційного типу. Виконати порівняльну характеристику за кількістю ітерацій і різною точністю.

$$1) \begin{cases} 10x_1 + x_2 + 2x_3 + 3x_4 = 18; \\ x_1 + 11x_2 + 3x_3 + x_4 = 19; \\ 2x_1 + 3x_2 + 15x_3 + x_4 = 36; \\ 3x_1 + x_2 + x_3 + 14x_4 = 20 \end{cases} \quad x^* = (1, 1, 2, 1);$$

$$2) \begin{cases} 7x_1 + x_2 + 3x_3 - x_4 = 10; \\ x_1 + 9x_2 - 2x_3 = 8; \\ 3x_1 - 2x_2 + 10x_3 + 2x_4 = 13; \\ -x_1 + 2x_3 + 5x_4 = 6 \end{cases} \quad x^* = (1, 1, 1, 1);$$

$$3) \begin{cases} 13x_1 + 5x_2 + 3x_3 - 4x_4 = 21; \\ 5x_1 + 13x_2 - x_3 + 2x_4 = 17; \\ 3x_1 - x_2 + 13x_3 + 2x_4 = 15; \\ -4x_1 + 2x_2 + 2x_3 + 13x_4 = 0 \end{cases} \quad x^* = (1, 1, 1, 0);$$

$$4) \begin{cases} 20x_1 + 3x_2 + 5x_3 + 7x_4 = 32; \\ 3x_1 + 15x_2 + 7x_3 + x_4 = 11; \\ 5x_1 + 5x_2 + 17x_3 + 2x_4 = 24; \\ 7x_1 + x_2 + 2x_3 + 15x_4 = 24 \end{cases} \quad x^* = (1, 0, 1, 1);$$

$$5) \begin{cases} 10x_1 - 2x_2 + 3x_3 + 4x_4 = 12; \\ -2x_1 + 11x_2 + x_3 + 2x_4 = 11; \\ 3x_1 + x_2 + 15x_3 - x_4 = 3; \\ 4x_1 + 2x_2 - x_3 + 17x_4 = 23 \end{cases} \quad x^* = (1, 1, 0, 1);$$

$$6) \begin{cases} 11x_1 + 2x_2 + 3x_3 + 4x_4 = 20; \\ 2x_1 + 13x_2 - x_3 + 2x_4 = 16; \\ 3x_1 - x_2 + 13x_3 - 3x_4 = 12; \\ 4x_1 + 2x_2 - 3x_3 + 20x_4 = 23 \end{cases} \quad x^* = (1, 1, 1, 1);$$



$$7) \quad \begin{cases} 15x_1 + 2x_2 - x_3 + 2x_4 = 18; \\ 2x_1 + 11x_2 + 2x_3 - x_4 = 14; \\ -x_1 + 2x_2 + 13x_3 + 3x_4 = 17; \\ 2x_1 - x_2 + 3x_3 + 17x_4 = 21 \end{cases} \quad x^* = (1, 1, 1, 1);$$

$$8) \quad \begin{cases} 10x_1 + 2x_2 + x_3 - 2x_4 = 9; \\ 2x_1 + 13x_2 - 2x_3 - 3x_4 = -3; \\ x_1 - 2x_2 + 15x_3 + 4x_4 = 20; \\ -2x_1 - 3x_2 + 4x_3 + 10x_4 = 12 \end{cases} \quad x^* = (1, 0, 1, 1);$$

$$9) \quad \begin{cases} 10x_1 + 2x_2 + 3x_3 - 2x_4 = 10; \\ 2x_1 + 11x_2 - 2x_3 - 3x_4 = 10; \\ 3x_1 - 2x_2 + 15x_3 + x_4 = 2; \\ -2x_1 - 3x_2 + x_3 + 15x_4 = 10 \end{cases} \quad x^* = (1, 1, 0, 1);$$

$$10) \quad \begin{cases} 12x_1 + 3x_2 + x_3 - 2x_4 = 14; \\ 3x_1 + 13x_2 - 4x_3 + 3x_4 = 15; \\ x_1 - 4x_2 + 14x_3 + x_4 = 12; \\ -2x_1 + 3x_2 + x_3 + 12x_4 = 14 \end{cases} \quad x^* = (1, 1, 1, 1);$$

$$11) \quad \begin{cases} 16x_1 - x_2 + 3x_3 + 2x_4 = 20; \\ -x_1 + 14x_2 + 2x_3 - 5x_4 = 10; \\ 3x_1 + 2x_2 + 20x_3 + 3x_4 = 28; \\ 2x_1 - 5x_2 + 3x_3 + 15x_4 = 15 \end{cases} \quad x^* = (1, 1, 1, 1);$$

$$12) \quad \begin{cases} 13x_1 + x_2 - 5x_3 + x_4 = 10; \\ x_1 + 16x_2 + x_3 + 2x_4 = 20; \\ -5x_1 + x_2 + 15x_3 - x_4 = 10; \\ x_1 + 2x_2 - x_3 + 13x_4 = 15 \end{cases} \quad x^* = (1, 1, 1, 1);$$

$$13) \quad \begin{cases} 17x_1 + x_2 - x_3 - x_4 = 16; \\ x_1 + 15x_2 - 4x_3 + 2x_4 = 14; \\ -x_1 - 4x_2 + 15x_3 + 2x_4 = 12; \\ -x_1 + 2x_2 + 2x_3 + 23x_4 = 26 \end{cases} \quad x^* = (1, 1, 1, 1);$$

$$14) \quad \begin{cases} 19x_1 + x_2 - x_4 = 19; \\ x_1 + 17x_2 + x_3 + 4x_4 = 23; \\ x_2 + 11x_3 - 3x_4 = 9; \\ -x_1 + 4x_2 - 3x_3 + 13x_4 = 13 \end{cases} \quad x^* = (1, 1, 1, 1);$$

$$15) \quad \begin{cases} 15x_1 - 6x_2 + 2x_3 + 2x_4 = 13; \\ -6x_1 + 21x_2 - 3x_3 - 5x_4 = 7; \\ 2x_1 - 3x_2 + 17x_3 + x_4 = 17; \\ 2x_1 - 5x_2 + x_3 + 12x_4 = 10 \end{cases} \quad x^* = (1, 1, 1, 1);$$

$$16) \quad \begin{cases} 13x_1 + x_2 - 3x_3 - x_4 = 10; \\ x_1 + 20x_2 + x_3 - 4x_4 = 18; \\ -3x_1 + x_2 + 13x_3 + 2x_4 = 13; \\ -x_1 - 4x_2 + 2x_3 + 20x_4 = 17 \end{cases} \quad x^* = (1, 1, 1, 1);$$

$$17) \quad \begin{cases} 17x_1 - x_2 + x_3 - x_4 = 16; \\ -x_1 + 19x_2 - 3x_3 + 2x_4 = 17; \\ x_1 - 3x_2 + 21x_3 - 3x_4 = 16; \\ -x_1 + 2x_2 - 3x_3 + 10x_4 = 8 \end{cases} \quad x^* = (1, 1, 1, 1);$$

$$18) \quad \begin{cases} 12x_1 + x_2 - x_3 - 2x_4 = 10; \\ x_1 + 13x_2 + 2x_3 + 3x_4 = 19; \\ -x_1 + 2x_2 + 14x_3 - x_4 = 14; \\ -2x_1 + 3x_2 - x_3 + 15x_4 = 15 \end{cases} \quad x^* = (1, 1, 1, 1);$$

$$19) \quad \begin{cases} 15x_1 + x_2 - 3x_3 + 2x_4 = 15; \\ x_1 + 16x_2 - x_3 + 3x_4 = 19; \\ -3x_1 - x_2 + 17x_3 - 4x_4 = 9; \\ 2x_1 + 3x_2 - 4x_3 + 18x_4 = 19 \end{cases} \quad x^* = (1, 1, 1, 1);$$

$$20) \quad \begin{cases} 11x_1 + 3x_2 - 4x_3 + x_4 = 8; \\ 3x_1 + 12x_2 - x_3 + 2x_4 = 4; \\ -4x_1 - x_2 + 13x_3 + 3x_4 = 12; \\ x_1 + 2x_2 + 3x_3 + 14x_4 = 18 \end{cases} \quad x^* = (1, 0, 1, 1);$$

$$21) \quad \begin{cases} 14x_1 - x_2 - x_3 + x_4 = 14; \\ -x_1 + 15x_2 - 2x_3 + 3x_4 = 17; \\ -x_1 - 2x_2 + 16x_3 - x_4 = -4; \\ x_1 + 3x_2 - x_3 + 17x_4 = 21 \end{cases} \quad x^* = (1, 1, 0, 1);$$

$$22) \quad \begin{cases} 16x_1 + 2x_2 - 3x_3 + x_4 = 19; \\ 2x_1 + 17x_2 - x_3 + 2x_4 = 21; \\ -3x_1 - x_2 + 18x_3 + 3x_4 = -1; \\ x_1 + 2x_2 + 3x_3 + 19x_4 = 22 \end{cases} \quad x^* = (1, 1, 0, 1);$$

$$23) \quad \begin{cases} 11,65x_1 - 1,76x_2 + 0,77x_3 = 10,66; \\ -1,76x_1 + 11,04x_2 - 2,61x_3 = 6,67; \\ 0,77x_1 - 2,61x_2 + 13,18x_3 = 11,34 \end{cases} \quad x^* = (1, 1, 1);$$

$$24) \quad \begin{cases} 2,56x_1 + 0,67x_2 - 1,78x_3 = 1,45; \\ 0,67x_1 + 2,67x_2 - 1,35x_3 = 1,99; \\ -1,78x_1 - 1,35x_2 + 11,55x_3 = 8,42 \end{cases} \quad x^* = (1, 1, 1);$$

$$25) \quad \begin{cases} 12,16x_1 + 1,45x_2 - 0,89x_3 = 12,72; \\ 1,45x_1 + 10,44x_2 + 1,18x_3 = 13,07; \\ -0,89x_1 + 1,18x_2 + 12,07x_3 = 12,36 \end{cases} \quad x^* = (1, 1, 1);$$

$$26) \quad \begin{cases} 9,87x_1 + 1,35x_2 - 0,44x_3 = 10,78; \\ 1,35x_1 + 11,22x_2 - 2,32x_3 = 10,25; \\ -0,44x_1 - 2,32x_2 + 13,73x_3 = 10,97 \end{cases} \quad x^* = (1, 1, 1);$$

$$27) \quad \begin{cases} 9,36x_1 + 0,92x_2 - 1,87x_3 = 8,41; \\ 0,92x_1 + 8,24x_2 - 0,77x_3 = 8,39; \\ -1,87x_1 - 0,77x_2 + 9,16x_3 = 6,52 \end{cases} \quad x^* = (1, 1, 1);$$

$$28) \quad \begin{cases} 8,17x_1 - 0,65x_2 + 1,54x_3 = 9,06; \\ -0,65x_1 + 11,16x_2 - 1,73x_3 = 8,78; \\ 1,54x_1 - 1,73x_2 + 9,15x_3 = 8,96 \end{cases} \quad x^* = (1, 1, 1);$$

$$29) \quad \begin{cases} 12,32x_1 + 1,17x_2 - 0,28x_3 = 13,21; \\ 1,17x_1 + 9,43x_2 - 0,88x_3 = 9,72; \\ -0,28x_1 - 0,88x_2 + 14,55x_3 = 13,39 \end{cases} \quad x^* = (1, 1, 1);$$

$$30) \quad \begin{cases} 9,48x_1 + 0,75x_2 - 1,23x_3 = 9; \\ 0,75x_1 + 9,61x_2 + 1,64x_3 = 12; \\ -1,23x_1 + 1,64x_2 + 5,51x_3 = 5,92 \end{cases} \quad x^* = (1, 1, 1).$$

3.4. Методом прогонки розв'язати системи рівнянь:

$$1) \quad a_i x_{i-1} - c_i x_i + b_i x_{i+1} = f_i \quad \forall i = 1, 2, \dots, n;$$

$$a_1 = 0, \quad b_n = 0,$$

де

$$a_i = -(1+i), \quad i = 2, 3, \dots, n;$$

$$c_i = 12+i, \quad i = 1, 2, \dots, n;$$

$$b_i = i, \quad i = 1, 2, \dots, n-1;$$

$$f_i = -(i^2 + 14i - 1), \quad i = 1, \dots, n-1;$$

$$n=7 \quad f_7 = -202; \quad n=10 \quad f_{10} = -600;$$

$$2) \quad a_i y_{i-1} - c_i y_i + b_i y_{i+1} = f_i, \quad i = \overline{1, n-1};$$

$$y(0) = y_0 = 1; \quad y(1) = y_n = 0,5;$$

де

$$a_i = \frac{1}{h^2} + \frac{1+ih}{2h}, \quad b_i = \frac{1}{h^2} - \frac{1+ih}{2h},$$

$$c_i = 1 + \frac{2}{h^2}, \quad f_i = \frac{2}{(1+ih)^3}, \quad h = 0,1; 0,01;$$

$$h = \frac{1}{n}; \quad y^*(x) = \frac{1}{x+1}; \quad x = x_i = ih;$$

$$3) \quad a_i y_{i-1} - c_i y_i + b_i y_{i+1} = f_i, \quad i = \overline{1, n-1};$$

$$y(0) = y_0 = -0,5; \quad y(1) = y_n = -1;$$

де

$$a_i = \frac{1}{h^2} - \frac{1}{h(ih-2)}, \quad b_i = \frac{1}{h^2} + \frac{1}{h(ih-2)},$$

$$c_i = \frac{2}{h^2} + 2 - ih; \quad f_i = 1; \quad h = \frac{1}{n}; \quad h = 0,1; 0,01;$$

$$y^*(x) = \frac{1}{x-2}; \quad x = x_i = ih;$$

$$4) \quad a_i y_{i-1} - c_i y_i + b_i y_{i+1} = f_i, \quad i = \overline{1, n-1};$$

$$y(0) = 0; \quad y(1) = 2,71828;$$

де

$$a_i = \frac{1}{h^2} + \frac{1}{h}, \quad b_i = \frac{1}{h^2} - \frac{1}{h},$$

$$c_i = \frac{2}{h^2} + 1, \quad f_i = -2ih \cdot e^{ih}, \quad h = \frac{1}{n}, \quad h = 0,1; 0,01;$$

$$y^*(x) = x \cdot e^x, \quad x = x_i = ih;$$

$$5) \quad a_i y_{i-1} - c_i y_i + b_i y_{i+1} = f_i, \quad i = \overline{1, n-1};$$

$$y(0) = 0; \quad y(1) = 1,38294;$$

де

$$a_i = \frac{1}{h^2} + \frac{ih+1}{2h}, \quad b_i = \frac{1}{h^2} - \frac{ih+1}{2h},$$

$$c_i = \frac{2}{h^2} + 1, \quad f_i = \frac{x_i^2 + 2x_i + 2}{x_i + 1},$$

$$x_i = ih, \quad y^*(x_i) = (x_i + 1) \ln(x_i + 1); \quad h = 0,1; 0,01.$$

3.5. З використанням теореми про необхідні й достатні або достатні умови збіжності методів, розв'язати такі задачі.

- A. Знайти області збіжності методів Якобі та Зейделя для систем  $Ax = b$  з матрицями вигляду

$$A = \begin{pmatrix} \alpha & \beta & 0 \\ \beta & \alpha & \beta \\ 0 & \beta & \alpha \end{pmatrix}.$$

- B. Довести, що для СЛАР другого порядку ( $n = 2$ ) методи Якобі і Зейделя є збіжними та розбіжними одночасно.  
 C. Нехай задано невідроджену матрицю  $A$  з діагональною перевагою, тобто для всіх  $i$  маємо

$$\sum_{i,j} |a_{ij}| \leq q |a_{ii}|, \quad q < 1.$$

Довести, що для похибки в методі Зейделя справджується нерівність

$$\|x - x^{(k)}\|_{\infty} \leq q^k \|x - x^{(0)}\|_{\infty}.$$

- D. Довести збіжність методу Зейделя для матриць з такими елементами:

a)  $a_{ij} = 3^{-|i-j|};$

б)  $a_{ij} = \begin{cases} 2, & i = j; \\ 1, & |i - j| = 1; \\ 0, & |i - j| > 1. \end{cases}$

- E. Знайти  $\alpha, \beta$ , за яких метод Зейделя є збіжним для СЛАР  $Ax = b$  з матрицями  $A$  вигляду:

a)  $\begin{pmatrix} \alpha & 0 & \beta \\ 0 & \alpha & 0 \\ \beta & 0 & \alpha \end{pmatrix};$

$$\text{б) } \begin{pmatrix} \alpha & \beta & 0 \\ \beta & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix};$$

$$\text{в) } \begin{pmatrix} \alpha & \alpha & 0 \\ \alpha & \beta & \beta \\ 0 & \beta & \alpha \end{pmatrix};$$

$$\text{г) } \begin{pmatrix} \alpha & \beta & 0 \\ \beta & \alpha & \beta \\ 0 & \beta & \alpha \end{pmatrix}.$$

Г. Дослідити збіжність методу Якобі для СЛАР з матрицею

$$\begin{pmatrix} 2 & -0,2 & 0,3 & 0,4 \\ 0,3 & -3 & 1 & -1,4 \\ 0,4 & 0,8 & 4 & 2,4 \\ -0,5 & 1,2 & -2,5 & -5 \end{pmatrix}.$$

## Розділ 4

### МЕТОДИ РОЗВ'ЯЗУВАННЯ ЗАДАЧ НА ВЛАСНІ ЗНАЧЕННЯ

Задача визначення власних значень і векторів матриць є важливою для вирішення широкого кола питань обчислювальної математики, її застосовують у диференціальних рівняннях (зокрема, у проблемі стійкості), механіці, радіофізиці й інших галузях. Цю задачу формулюють так: нехай задано дійсну  $(n \times n)$ -матрицю  $A$ . Знайти такі пари чисел  $\lambda$  і ненульових векторів  $x$ , що

$$Ax = \lambda x. \quad (4.1)$$

Задачу (4.1) називають задачею на власні значення. Розрізняють повну (алгебричну, або матричну) проблему власних значень, яка передбачає знаходження всіх власних пар  $\{\lambda, x\}$  матриці  $A$ , та часткові проблеми власних значень, які полягають, як звичайно, у відшуванні одного чи декількох власних чисел  $\lambda$  і, можливо, відповідних їм власних векторів  $x$ . В останньому випадку найчастіше йдеться про знаходження найбільшого і найменшого за модулем власних чисел. Знання таких характеристик матриці дає змогу, наприклад, робити висновки про збіжність тих чи інших ітераційних методів, оптимізувати параметри ітераційних методів, урахувати вплив на результати розв'язування алгебричних задач похибок вхідних даних і обчислювальних похибок. Є й інші формулювання часткових проблем.

Теоретично задачу (4.1) легко розв'язати: необхідно знайти корені так званого характеристичного, чи, інакше кажучи, вікового рівняння

$$\det(A - \lambda E) = 0, \quad (4.2)$$

а потім, підставляючи їх почергово в (4.1), отримати з відповідних систем власні вектори. Практична реалізація цього за суттю простого підходу пов'язана з низкою труднощів, що зростають зі збільшенням розмірності розв'язуваної задачі. Ці труднощі



зумовлені розгортанням характеристичного визначника  $\det(A - \lambda E)$  і обчисленням коренів отриманого в цьому разі полінома  $n$ -го степеня, а також пошуком лінійно незалежних розв'язків вироджених СПАР. З огляду на це такий безпосередній підхід до розв'язування алгебричної проблеми власних значень застосовують лише за малих розмірностей матриці  $A$  ( $n = 2, 3$ ), вже при  $n \geq 4$  на перший план виходять спеціальні чисельні методи розв'язування таких задач.

Зазначимо, що недавно чисельні методи розв'язування задач на власні значення ґрунтувались, як звичайно, на класичному підході, тобто на розгортанні характеристичних визначників. Методи розв'язування проблеми власних значень, які ґрунтуються на розгортанні характеристичного визначника, детально описані в [7].

Сучасні методи розв'язування повної проблеми орієнтовані на алгоритмічну побудову з матриці  $A$  такої матриці, певні елементи якої були б наближеними значеннями власних чисел  $A$ , причому формувались би і її власні вектори.

Нижче розглянуто деякі з цих методів так, щоб можна було зрозуміти ідеї, які є в їхній основі, й одночасно практично розв'язано сформульовані задачі.

## 4.1. Класичний спосіб

**П р и к л а д 4.1.** Визначити власні числа і власні вектори матриці

$$A = \begin{pmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & 8 & -2 \end{pmatrix}.$$

*Розв'язування.* Щоб визначити власні числа матриці  $A$ , потрібно обчислити визначник  $|A - \lambda E|$ . Розкриємо його й отримаємо

$$\begin{vmatrix} 3-\lambda & 1 & 0 \\ -4 & -1-\lambda & 0 \\ 4 & 8 & -2-\lambda \end{vmatrix} = -\lambda^3 + 3\lambda - 2.$$

Власні числа є коренями рівняння

$$-\lambda^3 + 3\lambda - 2 = (\lambda - 1)^2 \cdot (\lambda + 2) = 0,$$

тобто  $\lambda_1 = -2$ ,  $\lambda_2 = 1$ ,  $\lambda_3 = 1$ .

Для знаходження власного вектора, що відповідає власному числу  $\lambda_1 = -2$ , отримаємо систему лінійних алгебричних рівнянь з визначником, що дорівнює нулю:

$$\begin{cases} 5v_1 + v_2 = 0; \\ -4v_1 + v_2 = 0; \\ 4v_1 + 8v_2 = 0. \end{cases}$$

Легко переконатися, що ранг матриці цієї системи  $r = 2$ . Надамо вільній змінній  $v_3$ , наприклад, значення 1. Розв'яжемо систему перших двох рівнянь, отримаємо  $v_2 = 0$ ,  $v_1 = 0$ . Отже, власному числу  $\lambda_1 = -2$  відповідає власний вектор  $x_1 = (0, 0, 1)^T$ . Для знаходження власного вектора, відповідного кратному власному числу  $\lambda = 1$ , отримаємо систему

$$\begin{cases} 2v_1 + v_2 = 0; \\ -4v_1 - 2v_2 = 0; \\ 4v_1 + 8v_2 - 3v_3 = 0. \end{cases}$$

Як бачимо, її ранг дорівнює двом. Надамо вільній невідомій  $v_1 = 1$ , знайдемо наступні значення  $v_3 = -4$ ,  $v_2 = -2$ .

Отже, власному числу  $\lambda_2 = \lambda_3 = 1$  відповідає такий власний вектор:  $x_2 = (1, -2, -4)^T$ .

**П р и к л а д 4.2 [9].** Знайти всі власні числа і власні вектори

матриці  $A = \begin{pmatrix} 5 & -2 & 1 \\ 3 & -2 & 3 \\ 3 & -6 & 7 \end{pmatrix}$ .

*Розв'язування.* Складемо визначник  $|A - \lambda E|$  і розкриємо його:

$$\begin{vmatrix} 5-\lambda & -2 & 1 \\ 3 & -2-\lambda & 3 \\ 3 & -6 & 7-\lambda \end{vmatrix} = -\lambda^3 + 10\lambda^2 - 32\lambda + 32.$$

Власні числа є коренями рівняння

$$-\lambda^3 + 10\lambda^2 - 32\lambda + 32 = (\lambda - 2)(\lambda - 4)^2 = 0,$$

тобто  $\lambda_1 = 2$ ,  $\lambda_2 = \lambda_3 = 4$ .

Для знаходження власного вектора, відповідного власному числу  $\lambda_1 = 2$ , отримаємо систему лінійних алгебричних рівнянь з визначником, що дорівнює нулю:

$$\begin{cases} 3v_1 - 2v_2 + v_3 = 0; \\ 3v_1 - 4v_2 + 3v_3 = 0; \\ 3v_1 - 6v_2 + 5v_3 = 0. \end{cases}$$

Для розв'язування цієї системи визначимо ранг її матриці. Легко переконатися, що ранг матриці цієї системи  $r = 2$ . Надамо вільній змінній  $v_1$ , наприклад, значення 1, розв'яжемо систему перших двох рівнянь

$$\begin{cases} -2v_2 + v_3 = -3; \\ -4v_2 + 3v_3 = -3 \end{cases}$$

і отримаємо  $v_2 = 3$ ,  $v_3 = 3$ . Отже, власному числу  $\lambda_1 = 2$  відповідає власний вектор  $x_1 = (1, 2, 3)^T$ . Для знаходження власного вектора, відповідного кратному власному числу  $\lambda_2 = 4$ , отримаємо систему

$$\begin{cases} v_1 - 2v_2 + v_3 = 0; \\ 3v_1 - 6v_2 + 3v_3 = 0; \\ 3v_1 - 6v_2 + 3v_3 = 0. \end{cases}$$

Ранг її дорівнює одиниці, тому розв'язок наведемо у вигляді  $v_1 = 2v_2 - v_3$ . Вільним невідомим  $v_2$  і  $v_3$  необхідно надати такі значення, за яких отримують два лінійно незалежні вектори. Для цього візьмемо такі значення:

$$v_2 = 1, \quad v_3 = 0, \quad \text{і} \quad v_2 = 0, \quad v_3 = 1.$$

Отже, власному числу  $\lambda_2 = \lambda_3 = 4$  відповідають два власні вектори

$$x_2 = (2, 1, 0)^T$$

$$x_3 = (-1, 0, 1)^T$$

## 4.2. Степеневий метод знаходження максимального за модулем власного значення і відповідного йому власного вектора

Для знаходження максимального за модулем власного значення і відповідного йому власного вектора квадратної матриці простої структури<sup>1</sup> можна використовувати степеневий метод.

Спочатку розглянемо реалізацію цього методу для дійсної матриці  $A$  простої структури, яка має власні значення, розташовані з урахуванням їхньої кратності в такому порядку:

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|.$$

Задаванням довільного початкового наближення  $y^{(0)}$  для власного вектора  $y^{(1)}$  і значення  $\varepsilon > 0$ , за яким закінчуватиметься ітераційний процес, степеневий метод реалізують за формулами

$$\bar{y}^{(k)} = Ay^{(k-1)}, \quad y^{(k)} = \frac{\bar{y}^{(k)}}{\max_i |\bar{y}_i^{(k)}|}, \quad \mu_k = \frac{1}{m} \sum_{s=1}^m \frac{\bar{y}_s^{(k)}}{y_s^{(k-1)}}, \quad (4.3)$$

$$k = 0, 1, \dots$$

перевіряючи на кожній ітерації умову закінчення процесу

$$|\mu_k - \mu_{k-1}| \leq \varepsilon |\mu_k|. \quad (4.4)$$

Зазначимо, що  $\mu_k$  можна обчислити також за формулою

<sup>1</sup> Матрицею простої структури називають матрицю  $A$  порядку  $n$ , яка має  $n$  лінійно незалежних власних векторів.

$$\mu_k = \max \left\{ \bar{y}_i^{(k)} \right\}.$$

Тут і далі ми використовуємо позначення  $\max \{x\}$  для максимального за модулем елемента вектора  $x$ .

У (4.3) і (4.4)  $k$  – номер ітерації;  $s$  – номер компоненти;  $m$  – кількість ненульових компонент вектора  $y^{(1)}$ ;  $\mu_k$  – наближення до власного числа  $\lambda_1$ .

Доведемо збіжність степеневого методу, коли значення  $\lambda_1$  дійсне і не кратне. Оскільки матриця  $A$  є матрицею простої структури, то систему власних векторів  $\{x_j\}, j=1,2,\dots,n$  можна взяти за базис у  $n$ -вимірному просторі. Відповідно, довільний вектор  $y^{(0)}$  можна розкласти за власними векторами

$$y^{(0)} = c_1 x_1 + c_2 x_2 + \dots + c_n x_n.$$

З урахуванням того факту, що власний вектор визначений з точністю до сталого множника, під час дослідження збіжності не братимемо до уваги умови нормування в (4.3).

З послідовних ітерацій отримаємо вектори

$$y^{(1)} = Ay^{(0)} = c_1 \lambda_1 x_1 + c_2 \lambda_2 x_2 + \dots + c_n \lambda_n x_n,$$

$$y^{(2)} = Ay^{(1)} = c_1 \lambda_1^2 x_1 + c_2 \lambda_2^2 x_2 + \dots + c_n \lambda_n^2 x_n,$$

$$y^{(k)} = Ay^{(k-1)} = c_1 \lambda_1^k x_1 + c_2 \lambda_2^k x_2 + \dots + c_n \lambda_n^k x_n =$$

$$= \lambda_1^k \left[ c_1 x_1 + c_2 \left( \frac{\lambda_2}{\lambda_1} \right)^k x_2 + \dots + c_n \left( \frac{\lambda_n}{\lambda_1} \right)^k x_n \right] = \lambda_1^k \left[ c_1 x_1 + O \left( \left( \frac{\lambda_2}{\lambda_1} \right)^k \right) \right],$$

тобто вектор  $y$  наближатиметься до власного вектора  $x_1$  за напрямом.

**Приклад 4.3.** За допомогою степеневому методу знайти

максимальне власне число матриці  $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix}$ , починаючи з

початкового вектора  $y^{(0)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ .

*Розв'язування.* За наведеними вище формулами степеневому методу, прийнявши  $\varepsilon = 0,001$ , виконаємо наше завдання:

$$\tilde{y}^{(1)} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \\ 6 \end{pmatrix}; y^{(1)} = \begin{pmatrix} 2/3 \\ 1 \\ 2/3 \end{pmatrix} = \begin{pmatrix} 0,66 \\ 1 \\ 0,66 \end{pmatrix}; \mu_1 = 7$$

$$\tilde{y}^{(2)} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2/3 \\ 1 \\ 2/3 \end{pmatrix} = \begin{pmatrix} 14/3 \\ 23/3 \\ 14/3 \end{pmatrix};$$

$$y^{(2)} = 3/23 \begin{pmatrix} 14/3 \\ 23/3 \\ 14/3 \end{pmatrix} = \begin{pmatrix} 14/23 \\ 1 \\ 14/23 \end{pmatrix} = \begin{pmatrix} 0,6087 \\ 1 \\ 0,6087 \end{pmatrix}; \mu_2 = 7,22.$$

Умова  $|\mu_2 - \mu_1| \leq \varepsilon |\mu_2|$  не виконується, тому продовжуємо ітераційний процес:

$$\tilde{y}^{(3)} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 14/23 \\ 1 \\ 14/23 \end{pmatrix} = \begin{pmatrix} 102/23 \\ 171/23 \\ 102/23 \end{pmatrix};$$

$$y^{(3)} = 23/171 \begin{pmatrix} 102/23 \\ 171/23 \\ 102/23 \end{pmatrix} = \begin{pmatrix} 102/171 \\ 1 \\ 102/171 \end{pmatrix} = \begin{pmatrix} 0,596 \\ 1 \\ 0,596 \end{pmatrix}; \mu_3 = 7,3354.$$

Умова  $|\mu_3 - \mu_2| \leq \varepsilon |\mu_3|$  не виконується, знову продовжимо ітераційний процес:

$$\bar{y}^{(4)} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 102/171 \\ 1 \\ 102/171 \end{pmatrix} = \begin{pmatrix} 750/171 \\ 1263/171 \\ 750/171 \end{pmatrix}; \quad y^{(4)} = 171/1263 \begin{pmatrix} 750/171 \\ 1263/171 \\ 750/171 \end{pmatrix} = \begin{pmatrix} 0,5938 \\ 1 \\ 0,5938 \end{pmatrix};$$

$\mu_4 = 7,3639$ . Умова  $|\mu_4 - \mu_3| \leq \varepsilon |\mu_4|$  не виконується, продовжимо обчислення далі:

$$\bar{y}^{(5)} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 250/421 \\ 1 \\ 250/421 \end{pmatrix} = \begin{pmatrix} 1842/421 \\ 3105/421 \\ 1842/421 \end{pmatrix};$$

$$y^{(5)} = 421/3105 \begin{pmatrix} 1842/421 \\ 3105/421 \\ 1842/421 \end{pmatrix} = \begin{pmatrix} 1842/3105 \\ 1 \\ 1842/3105 \end{pmatrix} = \begin{pmatrix} 0,5932 \\ 1 \\ 0,5932 \end{pmatrix};$$

$$\mu_5 = 7,3704.$$

Умова  $|\mu_5 - \mu_4| \leq \varepsilon |\mu_5|$  виконується, отже, ми знайшли наближене власне значення і власний вектор. Дані занесемо в табл. 4.1.

Таблиця 4.1

Наближення до власних значень і векторів за степеневим методом

$\mu, y_i$	$k$					
	0	1	2	3	4	5
$\mu$	–	7	7,2222	7,3354	7,3639	7,3704
$y_1$	1	0,6667	0,6087	0,5965	0,5938	0,5932
$y_2$	1	1	1	1	1	1
$y_3$	1	0,6667	0,6087	0,5965	0,5938	0,5932

Для матриці  $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix}$  отримано власне значення 7,3704

і власний вектор  $\begin{pmatrix} 0,5932 \\ 1 \\ 0,5932 \end{pmatrix}$ .

На комп'ютері степеневий метод можна реалізувати також у вигляді такого *РМ*-алгоритму.

1. Увести  $(n \times n)$ -матрицю  $A$ , задати  $n$ -вимірний вектор  $y^{(0)}$

обчислити  $\|y^{(0)}\|$  і вектор  $x^{(0)} = \frac{y^{(0)}}{\|y^{(0)}\|}$ ; прийняти  $k = 1$ .

2. Обчислити вектор  $y^{(k)} = Ax^{(k-1)}$

3. Обчислити  $\|y^{(k)}\|$  і  $x^{(k)} = \frac{y^{(k)}}{\|y^{(k)}\|}$ .

4. Обчислити співвідношення  $\lambda_i^{(k)} = \frac{y_i^{(k)}}{x_i^{(k-1)}}$  (координат векторів  $y^{(k)}$  і  $x^{(k-1)}$ ) при  $i=1,2,\dots,n$  таких, що  $|x_i^{(k-1)}| > \delta$ , де  $\delta > 0$  – деяке задане мале число (допуск).

5. Випробувати число  $\lambda_j^{(k)}$  тестом на збіжність. Якщо виявиться збіг необхідної кількості знаків  $\lambda_j^{(k)}$  і  $\lambda_j^{(k-1)}$  ( $\lambda_j^{(0)}$  можна задавати довільно), то роботу алгоритму припинити і за старше власне число  $\lambda_j$  прийняти усереднене (за  $i$ ) значення  $\lambda_j^{(k)}$ , а за нормований старший власний вектор  $x_j$  – вектор  $x^{(k)}$ . Інакше повернутися до кроку 2.

### 4.3. Метод скалярних добутків

Метод скалярних добутків доцільно використовувати під час знаходження максимального за модулем власного числа і відповідного до нього власного вектора *симетричної матриці*. Він має швидшу збіжність, ніж степеневий метод.



Обчислювальна схема методу скалярних добутків має вигляд [17]

$$y^{(k)} = Ay^{(k-1)} \quad \tilde{y}^{(k)} = \frac{1}{\sqrt{(y^{(k)}, y^{(k)})}} y^{(k)} \quad \mu_k = \frac{(y^{(k)}, y^{(k)})}{(y^{(k)}, \tilde{y}^{(k-1)})}.$$

Умовою завершення процесу може слугувати умова

$$|\mu_k - \mu_{k-1}| \leq \varepsilon |\mu_{k-1}|,$$

де  $\varepsilon > 0$  – задана точність.

**П р и к л а д 4.4.** Методом скалярних добутків знайти максимальне за модулем власне число матриці

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix}, \text{ вибравши за початкове наближення } y^{(0)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

*Розв'язування.* Зробимо декілька кроків цим методом.

**Крок 1:**

$$\tilde{y}^{(0)} = \frac{1}{\sqrt{(y^{(0)}, y^{(0)})}} y^{(0)} = \frac{1}{\sqrt{(1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1)}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} = \begin{pmatrix} 0,57735 \\ 0,57735 \\ 0,57735 \end{pmatrix};$$

$$y^{(1)} = A\tilde{y}^{(0)} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0,57735 \\ 0,57735 \\ 0,57735 \end{pmatrix} = \begin{pmatrix} 1 \cdot 0,57735 + 2 \cdot 0,57735 + 3 \cdot 0,57735 \\ (2 + 5 + 2) \cdot 0,57735 \\ (3 + 2 + 1) \cdot 0,57735 \end{pmatrix} = \begin{pmatrix} 3,46410 \\ 5,19615 \\ 3,46410 \end{pmatrix};$$

$$\begin{aligned} \mu_1 &= \frac{(y^{(1)}, y^{(1)})}{(y^{(1)}, \tilde{y}^{(0)})} = \\ &= \frac{(3,46410 \cdot 3,46410 + 5,19615 \cdot 5,19615 + 3,46410 \cdot 3,46410)}{(3,46410 \cdot 0,57735 + 5,19615 \cdot 0,57735 + 3,46410 \cdot 0,57735)} = \\ &= \frac{50,99995}{6,99999} = 7,28571. \end{aligned}$$

Крок 2:

$$\begin{aligned}\tilde{y}^{(1)} &= \frac{1}{\sqrt{(y^{(1)}, y^{(1)})}} y^{(1)} = \\ &= \frac{1}{\sqrt{(3,46410 \cdot 3,46410 + 5,19615 \cdot 5,19615 + 3,46410 \cdot 3,46410)}} \times \\ &\times \begin{pmatrix} 3,46410 \\ 5,19615 \\ 3,46410 \end{pmatrix} = \begin{pmatrix} 3,46410/7.14143 \\ 5,19615/7.14143 \\ 3,46410/7.14143 \end{pmatrix} = \begin{pmatrix} 0,48507 \\ 0,72761 \\ 0,48507 \end{pmatrix};\end{aligned}$$

$$\begin{aligned}y^{(2)} &= Ay^{(1)} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0,48507 \\ 0,72761 \\ 0,48507 \end{pmatrix} = \\ &= \begin{pmatrix} 1 \cdot 0,48507 + 2 \cdot 0,72761 + 3 \cdot 0,48507 \\ 2 \cdot 0,48507 + 5 \cdot 0,72761 + 2 \cdot 0,48507 \\ 3 \cdot 0,48507 + 2 \cdot 0,72761 + 1 \cdot 0,48507 \end{pmatrix} = \begin{pmatrix} 3,3955 \\ 5,57833 \\ 3,3955 \end{pmatrix};\end{aligned}$$

$$\begin{aligned}\mu_2 &= \frac{(y^{(2)}, y^{(2)})}{(y^{(2)}, \tilde{y}^{(1)})} = \\ &= \frac{(3,3955 \cdot 3,3955 + 5,57833 \cdot 5,57833 + 3,3955 \cdot 3,3955)}{(3,3955 \cdot 0,48507 + 5,57833 \cdot 0,72761 + 3,3955 \cdot 0,48507)} = \\ &= \frac{54,17940}{7,35334} = 7,368.\end{aligned}$$

Результати цих та наступних кроків занесемо у табл. 4.2.

Таблиця 4.2

## Результати обчислень за методом скалярних добутків

$k$	$y_1^{(k)}$	$y_2^{(k)}$	$y_3^{(k)}$	$\mu_k$
1	3,46410	5,19615	3,46410	7,28571
2	3,39549	5,57831	3,39549	7,36800
3	3,36101	5,63464	3,36101	7,37207
4	3,35245	5,64554	3,35245	7,37227

З табл. 4.2 бачимо, що для отримання максимального власного числа з похибкою  $\varepsilon = 0,0001$  достатньо чотирьох ітерацій. Отже, отримано такі власне число та ненормований і нормований власні вектори:

$$\mu_4 = 7,37227; \quad y^{(4)} = \begin{pmatrix} 3,35052 \\ 5,64787 \\ 3,35052 \end{pmatrix}; \quad \tilde{y}^{(4)} = \begin{pmatrix} 0,45593 \\ 0,76436 \\ 0,45593 \end{pmatrix}.$$

Реалізувати метод скалярних добутків на комп'ютері можна також у вигляді *SP*-алгоритму.

**Крок 1.** Увести: задану  $n \times n$ -матрицю  $A$ , довільний  $n$ -вимірний початковий вектор  $y^{(0)} \neq 0$ , мале число  $\varepsilon > 0$  (допустиму абсолютну похибку шуканого власного числа  $\lambda_1$ ), число  $\lambda^{(0)}$  для початкового порівнювання (наприклад, 0). Задати  $k=1$  (включити лічильник ітерацій).

**Крок 2.** Обчислити скаляри  $s^{(0)} = (y^{(0)}, y^{(0)})$ ,

$$\|y^{(0)}\|_2 = \sqrt{s^{(0)}} \text{ і вектор } x^{(0)} = \frac{y^{(0)}}{\|y^{(0)}\|_2}.$$

**Крок 3.** Обчислити  $y^{(k)} = Ax^{(k-1)}$  (ітерація нормованого вектора).

**Крок 4.** Обчислити  $s^{(k)} = (y^{(k)}, y^{(k)})$  і  $t^{(k)} = (y^{(k)}, x^{(k-1)})$

$$\|y^{(k)}\|_2 = \sqrt{s^{(k)}} \quad x^{(k)} = \frac{y^{(k)}}{\|y^{(k)}\|_2} \quad (\text{наближення до нормованого}$$

власного вектора),  $\lambda^{(k)} = \frac{s^{(k)}}{t^{(k)}} \quad (\text{наближення до власного числа } \lambda_1).$

**Крок 5.** Якщо  $|\lambda^{(k)} - \lambda^{(k-1)}| > \varepsilon$ , то задати  $k := k+1$  і повторити крок 3, інакше завершити роботу алгоритму, вважаючи, що  $\lambda_1 \approx \lambda^{(k)} \quad x_1 \approx x^{(k)}$

**Зауваження.** Цей алгоритм дає змогу швидше, ніж *РМ*-алгоритм, знайти з потрібною точністю найбільше власне число матриці, проте точність наближеної рівності  $x_1 \approx x^{(k)}$  для відповідного власного вектора може виявитися недостатньою.

## 4.4. Метод зворотних ітерацій

У деяких практичних задачах виникає потреба в знаходженні мінімального за модулем власного значення і відповідного йому власного вектора. Для розв'язування задачі можна використати модифікацію степеневому методу, ітеруючи вектор оберненою матрицею

$$y^{(k)} = A^{-1} y^{(k-1)} \quad (4.5)$$

Отже, можна обчислювати максимальне власне значення матриці  $A^{-1}$ . Його обернене значення дорівнює мінімальному власному значенню матриці  $A$ . У випадку реалізації цього методу на практиці нема потреби знаходити обернену до  $A$  матрицю, оскільки це передбачає велику кількість арифметичних операцій. Тому замість ітерацій за формулою (4.5) на кожному кроці ітераційного процесу можна розв'язувати систему лінійних алгебричних рівнянь

$$A y^{(k)} = y^{(k-1)}$$

У цьому разі, як і в степеневому методі, на кожному кроці необхідне нормування певного вектора.

Метод зворотних ітерацій, як і степеневий, має лінійну збіжність, тому на практиці використовують ті чи інші прийоми, що прискорюють його збіжність.

Один із можливих варіантів прискорення збіжності методу зворотних ітерацій відомий під назвою метод Віландта. В цьому методі за вихідною матрицею  $A$  будують деяку допоміжну

матрицю  $B = (A - \mu_i E)^{-1}$  з власними числами  $v^{(k)} = \frac{1}{(\lambda_k - \mu_i)}$ , де

$\mu_i$  – наближення до  $i$ -го власного значення  $\lambda_i$  матриці  $A$ . Потім степеневим методом знаходять максимальне власне значення матриці  $B$ . Оскільки за доброго початкового наближення  $\mu_i$  значення  $\lambda_i$  домінує над іншими власними значеннями, то степеневий метод збігається за декілька ітерацій. Після отримання власного значення  $\lambda_i^{(k)}$  ( $k$  – номер ітерації) матриці  $B$  власне значення матриці  $A$  знаходять за формулою

$$\lambda_i = \mu + \frac{1}{v^{(k)}}, \quad (4.6)$$

а  $y_i^{(k)}$  є власним вектором, що відповідає цьому власному значенню. У цьому випадку, як і в методі зворотних ітерацій, немає потреби попередньо знаходити обернену матрицю до  $(A - \mu_i E)$ , а достатньо на кожному кроці степеневого методу розв'язати СЛАР

$$(A - \mu E) y^{(k)} = y^{(k-1)} \quad (4.7)$$

Обчислювальну схему методу зворотних ітерацій для уточнення  $i$ -го власного значення матриці  $A$  й обчислення відповідного йому власного вектора можна реалізувати так.

1. Перед початком ітераційного процесу задають наближення  $\mu$  і  $\lambda_i$ , початкове наближення  $y^{(0)}$  до власного вектора  $y_i$ , відносну похибку отримуваного власного значення  $\varepsilon$ .
2. Тоді для  $k=1, 2, \dots$

$$(A - \mu E) \tilde{y}^{(k)} = y^{(k-1)}, \quad v^{(k)} = \max_{1 \leq j \leq n} \left\{ \tilde{y}_j^{(k)} \right\},$$

$$y^{(k)} = \frac{\tilde{y}^{(k)}}{v^{(k)}}. \quad (4.8)$$

3. Якщо виявиться, що всі значення  $y_i^{(k)} < 0$ , то за  $v^{(k)}$  беруть мінімальну компоненту  $\tilde{y}_i^{(k)}$ . Якщо  $\left| v^{(k)} - v^{(k-1)} \right| < \varepsilon \left| v^{(k)} \right|$ , то приймають

$$\lambda_i = \mu + \frac{1}{v^{(k)}}, \text{ а } y_i = y_i^{(k)};$$

У разі невиконання нерівності ітераційний процес продовжують.

Розглянутий нами метод Віланда можна використовувати в різних варіантах. Ним можна уточнювати довільні власні значення матриці або власні значення і відповідні їм власні вектори. Цей метод дає змогу за відомим власним значенням знайти власний вектор, який йому належить. Методом Віланда можна обчислити мінімальне власне значення і відповідний йому вектор матриці  $A$ . Для цього попередньо виконують декілька зворотних ітерацій без зсуву; отримавши таким способом початкове наближення до мінімального власного значення, його уточнюють методом Віланда.

Перевагами методу є достатньо швидка збіжність, можливість збереження вихідної структури матриці, простота реалізації на комп'ютерах. До недоліків треба зачислити необхідність розв'язування системи лінійних алгебричних рівнянь на кожному кроці процесу. Проте якщо відомий трикутний розклад матриці, то на кожному кроці потрібно виконувати лише зворотний хід.

**П р и к л а д 4.5.** Нехай задано матрицю  $A = \begin{pmatrix} -2 & 14 & 10 \\ -1 & -6 & -3 \\ 0 & 2 & 1 \end{pmatrix}$ .

відоме наближення до третього власного значення  $\mu_3 = 0,7$ . Потрібно уточнити третє власне значення і відповідний власний вектор.

*Розв'язування.* Побудуємо матрицю

$$(A - \mu_3 E) = \begin{pmatrix} 32,3 & 16,0 & 72,0 \\ -24,0 & -10,7 & -57,0 \\ -8,0 & -4,0 & -17,7 \end{pmatrix}.$$

Виберемо початкове наближення до другого власного вектора

$y^{(0)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$  і, розв'язавши систему лінійних алгебричних рівнянь

$$(A - \mu_3 E) \tilde{y}^{(1)} = y^{(0)} \quad \text{отримаємо} \quad \tilde{y}^{(1)} = \begin{pmatrix} -2,06254 \\ 1,65652 \\ 0,55730 \end{pmatrix} \times 10^2 \quad \text{Виконаємо}$$

нормування  $\tilde{y}^{(1)}$ , поділивши всі його компоненти на максимальну за модулем компоненту  $v^{(1)} = -2,06254 \cdot 10^2$  Отримаємо

$$y^{(2)} = \begin{pmatrix} 1,00000 \\ -0,80022 \\ -0,26723 \end{pmatrix}.$$

Далі  $v^{(2)} = 3,90819$ , розв'яжемо систему рівнянь

$(A - \mu_3 E) \tilde{y}^{(2)} = y^{(1)}$  і пронормуємо її розв'язок. Тоді отримаємо

$$y^{(6)} = \begin{pmatrix} 1,00000 \\ -0,79999 \\ -0,26666 \end{pmatrix}. \quad \text{і т. д. У підсумку матимемо} \quad v^{(5)} = 3,33693, \text{ а}$$

$$v^{(6)} = 3,33411, \quad y^{(6)} = \begin{pmatrix} 1,00000 \\ -0,79999 \\ -0,26666 \end{pmatrix}.$$

Отже, з точністю до трьох десяткових знаків значення  $v^{(6)}$  є максимальним власним значенням матриці  $(A - \mu_3 E)^{-1}$ . Третім власним значенням матриці  $A$  є

$$\lambda_3 = \mu_3 + \frac{1}{v^{(6)}} = 0,7 + 0,29992 = 0,99992,$$

а  $y^{(6)}$  – наближений власний вектор, що відповідає  $\lambda_3$ .

Зазначимо, що точними власними значеннями матриці є  $\lambda_1 = 3$ ,  $\lambda_2 = 2$ ,  $\lambda_3 = 1$ .

#### 4.5. Зворотні ітерації зі змінними зсувами. **RQI-алгоритм**

Практика обчислень засвідчує, що збіжність процесу зворотних ітерацій зі зсувами має велику швидкість (порівняно зі звичайним степеневим методом). Однак ще більшу швидкість збіжності можна отримати введенням змінних зсувів, які визначені деякою послідовністю чисел  $\sigma_0, \sigma_1, \sigma_2, \dots$ , збіжною до шуканого власного числа. За такі числа доцільно використовувати наближення  $\lambda_j^{(k)}$  до власного  $\lambda_j$ , отримувані за формулою (4.6).

Отже, зворотні ітерації зі змінними зсувами можна визначити сукупністю рівностей

$$\begin{aligned}(A - \lambda_j^{(k-1)} E) y^{(k)} &= x^{(k-1)}; \\ x^{(k)} &= \frac{y^{(k)}}{\|y^{(k)}\|}; \\ \lambda_j^{(k)} &= \lambda_j^{(k-1)} + \left\langle \frac{x_i^{(k-1)}}{y_j^{(k)}} \right\rangle,\end{aligned}\tag{4.9}$$

де  $k=1, 2, \dots$ , а число  $\lambda_j^{(0)} \approx \lambda_j$  і нормований вектор  $x^{(0)}$  задані,  $\langle \cdot \rangle$  - знак усереднення по всіх  $i$ , для яких  $y_i \neq 0$ .

Швидкість збіжності такого процесу – квадратична [11], тоді як у випадку постійних зсувів – лише лінійна, хоча з малими, як звичайно, знаменниками геометричної прогресії. Часто буває достатньо зробити дві-три ітерації, щоб отримати задану власну пару з реально можливою точністю. Однак у першому випадку при кожному  $k$  розв'язують СЛАР з однією й тією ж матрицею коефіцієнтів, а в другому – на різних кроках треба розв'язувати цілком різні системи.

У методі (4.9) часто не відомо, як підбирати початковий зсув  $\sigma = \lambda_j^{(0)}$  хіба що треба знайти власне число, найближче до заданого значення, і відповідний йому власний вектор.

Швидше збіжною є така модифікація методу (4.9): зворотні ітерації зі співвідношеннями Релея, які застосовують для розв'язування симетричних задач на власні значення. Їхня основа – RQI-алгоритм (Rayleigh quotient iteration) [5, 18].



**RQI-алгоритм.**

**Крок 0.** Задати вектор  $x^{(0)}$  такий, що  $\|x^{(0)}\| = 1$ .

**Крок 1.** Для  $k = 1, 2, \dots$ :

1.1) обчислити  $\rho_{k-1} = (Ax^{(k-1)}, x^{(k-1)}) / (x^{(k-1)}, x^{(k-1)})$ ;

1.2) знайти  $y^{(k)}$  з рівняння  $(A - \rho_{k-1}E)y^{(k)} = x^{(k-1)}$ ;

а) нормувати  $y^{(k)}$ , тобто прийняти  $x^{(k)} = y^{(k)} / \|y^{(k)}\|$ ;

б) перевірити  $\rho_{k-1}, x^{(k)}$  на збіжність. Перейти на крок 1.1 або зупинитися.

Після закінчення ітераційного процесу при деякому  $k = k_0$  за власну для заданої матриці  $A$  беруть пару  $(\rho_{k_0-1}, x^{(k_0)})$ . Зазначимо, що як критерій припинення роботи алгоритму можна використати умову  $\|y^{(k)}\| > C$ , де  $C$  – достатньо велике число.

Зсуви на співвідношення Релея за наявності ортонормованого базису із власних векторів  $x_1, x_2, \dots, x_n$  забезпечують асимптотично кубічну швидкість збіжності послідовності Релея  $x^{(0)}, x^{(1)}, x^{(2)}, \dots$  до деякого з векторів цього базису. До якого – залежить від вибору початкового вектора цієї послідовності.

Зазначимо, що RQI-алгоритм потребує доброго початкового наближення. Тому іноді корисно зробити спочатку кілька кроків РМ-алгоритму.

**П р и к л а д 4.6.** До симетричної матриці  $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix}$

застосувати RQI-алгоритм, узявши за початковий вектор  $x^{(0)} = (1, 1, 1)^T$  та прийнявши  $C = 10^5$

*Розв'язування.* Для  $k = 1$ .

**Крок 1.1.** Потрібно знайти спочатку добуток матриці  $A$  на

$$\text{початковий вектор } x^{(0)}: Ax^{(0)} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \\ 6 \end{pmatrix}.$$

Далі обчислимо скалярний добуток векторів:

$$\{Ax^{(0)}, x^{(0)}\} = \left( \begin{pmatrix} 6 \\ 9 \\ 6 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right) = 21; \{x^{(0)}, x^{(0)}\} = \left( \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right) = 3.$$

Тоді, маючи необхідні дані, знайдемо

$$\rho_0 = \{Ax^{(0)}, x^{(0)}\} / \{x^{(0)}, x^{(0)}\} = 21/3 = 7$$

**Крок 1.2.** Із заданого рівняння  $(A - \rho_0 E)y^{(1)} = x^{(0)}$ , тобто з рівняння

$$\begin{pmatrix} -6 & 2 & 3 \\ 2 & -2 & 2 \\ 3 & 2 & -6 \end{pmatrix} \begin{pmatrix} y_1^{(1)} \\ y_2^{(1)} \\ y_3^{(1)} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

отримаємо вектор  $y^{(1)} = \begin{pmatrix} 2 \\ 3,5 \\ 2 \end{pmatrix}$ .

**Крок 1.3.** Знайдемо  $x^{(1)} = y^{(1)} / \|y^{(1)}\|$ . Норму  $\|y^{(1)}\|$  шукаємо так:

$$\|y^{(1)}\| = \sqrt{\left( (y_1^{(1)})^2 + (y_2^{(1)})^2 + (y_3^{(1)})^2 \right)} \quad \text{Тоді}$$

$$\|y^{(1)}\| = \sqrt{4 + 12,25 + 4} = 4,5 \quad \text{Тепер можемо знайти вектор}$$

$$x^{(1)} = \begin{pmatrix} y_1^{(1)} / 4,5 \\ y_2^{(1)} / 4,5 \\ y_3^{(1)} / 4,5 \end{pmatrix} = \begin{pmatrix} 0,444444444444 \\ 0,777777777778 \\ 0,444444444444 \end{pmatrix}.$$

**Крок 1.4.** Перевіримо на збіжність: чи виконується нерівність  $\|y^{(k)}\| > C$ . Якщо так, то виходимо з алгоритму, інакше переходимо на крок 1.1. Очевидно,  $\|y^{(1)}\| < C$

Тепер виконаємо цей алгоритм для  $k = 2$ .

**Крок 1.1.** Потрібно визначити  $\rho_1 = (Ax^{(1)}, x^{(1)}) / (x^{(1)}, x^{(1)})$ .

Знайдемо спочатку добуток матриці  $A$  на вектор:

$$x^{(1)} : Ax^{(1)} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0,444444444444 \\ 0,777777777778 \\ 0,444444444444 \end{pmatrix} = \begin{pmatrix} 3,333333333333 \\ 5,666666666667 \\ 3,333333333333 \end{pmatrix}.$$

Далі обчислимо добутки векторів:

$$\begin{aligned} (Ax^{(1)}, x^{(1)}) &= \left( \begin{pmatrix} 3,333333333333 \\ 5,666666666667 \\ 3,333333333333 \end{pmatrix}, \begin{pmatrix} 0,444444444444 \\ 0,777777777778 \\ 0,444444444444 \end{pmatrix} \right) = \\ &= 7,37037037037037; \end{aligned}$$

$$\begin{aligned} (x^{(1)}, x^{(1)}) &= \left( \begin{pmatrix} 0,444444444444 \\ 0,777777777778 \\ 0,444444444444 \end{pmatrix}, \begin{pmatrix} 0,444444444444 \\ 0,777777777778 \\ 0,444444444444 \end{pmatrix} \right) = \\ &= 1,000000000000000. \end{aligned}$$

Тоді, маючи необхідні дані, знайдемо

$$\rho_1 = (Ax^{(1)}, x^{(1)}) / (x^{(1)}, x^{(1)}) = 7,37037037037037$$

**Крок 1.2.** З матрично-векторного рівняння  $(A - \rho_1 E)y^{(2)} = x^{(1)}$ , тобто з системи рівнянь

$$\begin{aligned} \left( \begin{array}{ccc} -6,370370370370 & 2,000000000000 & 3,000000000000 \\ 2,000000000000 & -2,370370370370 & 2,000000000000 \\ 3,000000000000 & 2,000000000000 & -6,370370370370 \end{array} \right) y^{(2)} &= \\ &= \begin{pmatrix} 0,444444444444 \\ 0,777777777778 \\ 0,444444444444 \end{pmatrix} \end{aligned}$$

знаходимо  $y^{(2)}$ :

$$y^{(2)} = \begin{pmatrix} 2,3774999999776 \\ 4,00874999999623 \\ 2,37749999999776 \end{pmatrix} \times 10^2$$

**Крок 1.3.** Шукаємо  $\|y^{(2)}\| = 5,232120895243966 \times 10^2$ . Тепер можемо знайти вектор  $x^{(2)}$ :

$$x^{(2)} = \begin{pmatrix} 0,45440463773666 \\ 0,76618069044242 \\ 0,45440463773666 \end{pmatrix}.$$

**Крок 1.4.** Перевіримо на збіжність виконанням нерівності  $\|y^{(k)}\| > C$ . Задана умова не виконується.

Продовжимо алгоритм для  $k=3$ .

**Крок 1.1.** Знайдемо  $\rho_2 = (Ax^{(2)}, x^{(2)}) / (x^{(2)}, x^{(2)})$  та добуток матриці  $A$  на вектор  $x^{(2)}$ :

$$Ax^{(2)} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0,45440463773666 \\ 0,76618069044242 \\ 0,45440463773666 \end{pmatrix} = \begin{pmatrix} 3,34997993183149 \\ 5,64852200315875 \\ 3,34997993183149 \end{pmatrix}.$$

Далі обчислимо добуток векторів

$$(Ax^{(2)}, x^{(2)}) = \begin{pmatrix} 3,34997993183149 \\ 5,64852200315875 \\ 3,34997993183149 \end{pmatrix} \cdot \begin{pmatrix} 0,45440463773666 \\ 0,76618069044242 \\ 0,45440463773666 \end{pmatrix} = 7,37228132305734;$$

$$(x^{(2)}, x^{(2)}) = 1,000000000000000 \text{ і знайдемо}$$

$$\rho_2 = (Ax^{(2)}, x^{(2)}) / (x^{(2)}, x^{(2)}) = 7,37228132305734.$$

**Крок 1.2.** З рівняння  $(A - \rho_2 E) y^{(3)} = x^{(2)}$ , тобто з

$$\begin{pmatrix} -6,37228132305734 & 2,00000000000000 & 3,00000000000000 \\ 2,00000000000000 & -2,37228132305734 & 2,00000000000000 \\ 3,00000000000000 & 2,00000000000000 & -6,37228132305734 \end{pmatrix} \times \\ \times y^{(3)} = \begin{pmatrix} 0,45440463773666 \\ 0,76618069044242 \\ 0,45440463773666 \end{pmatrix}$$

знайдемо:

$$y^{(3)} = \begin{pmatrix} 2,14669210717966 \\ 3,61962484992548 \\ 2,14669210717966 \end{pmatrix} \times 10^9$$

**Крок 1.3.** Шукаємо  $\|y^{(3)}\| = 4,724220365335739 \cdot 10^9$  та  $x^{(3)}$ :

$$x^{(3)} = \begin{pmatrix} 0,45440134904187 \\ 0,76618459132108 \\ 0,45440134904187 \end{pmatrix}.$$

**Крок 1.4.** Перевіримо на збіжність.

Оскільки  $\|y^{(3)}\| = 4,724220365335739 \cdot 10^9 > C$ , то ітераційний процес завершений. У цьому разі  $\rho_3 = (Ax^{(3)}, x^{(3)}) / (x^{(3)}, x^{(3)}) = -7,37228132326901$ . На підставі порівняння цього значення з попереднім  $\rho_2$  бачимо, що в них збігаються десять десяткових знаків. Це свідчить про високу швидкість збіжності розглянутого алгоритму.

## 4.6. Метод поворотів Якобі для знаходження власних пар симетричної матриці

Для розв'язування повної проблеми власних значень для симетричних матриць розроблено класичний метод Якобі, який має квадратичну швидкість збіжності.

Класичний ітераційний метод поворотів, запропонований Якобі (1846), передбачає побудову послідовності матриць

$$B_0 (:= A), B_1, B_2, \dots, B_k, \dots$$

за допомогою перетворень типу

$$B_k = T_{ij}^T B_{k-1} T_{ij} \quad (4.10)$$

такої, що на  $k$ -му кроці перетворюється в нуль максимальний за модулем елемент матриці  $B_{k-1}$  попереднього кроку (а отже, і симетричний до нього елемент). На кожному кроці таких перетворень відбуваються перетворення тільки двох рядків (або двох стовпців, що неважливо з огляду на симетрію) матриці попереднього кроку. Хоча, на жаль, не треба сподіватися, що таким способом за скінченну кількість кроків можна точно знайти діагональну матрицю  $\Lambda$ , бо отримані на деякому етапі перетворень нульові елементи на наступному етапі стануть ненульовими, проте

$$B_{k \rightarrow \infty} \longrightarrow \Lambda,$$

де  $\Lambda$  – діагональна матриця, на діагоналі якої стоять власні числа матриці  $A$ .

У  $n$ -вимірному випадку запишемо сукупність формул, які визначають *один крок методу поворотів Якобі*.

Для простоти вважатимемо, що перетворюємо матрицю  $A$  в матрицю  $B$  згідно з (4.10), хоча насправді на  $k$ -му кроці потрібно застосовувати перетворення (4.10) до матриці  $B_{k-1} = \left( b_{ml}^{(k-1)} \right)_{m,l=1}^n$  з результатом  $B_k = \left( b_{ml}^{(k)} \right)_{m,l=1}^n$

Нехай  $a_{ij}$  – головний елемент перетворюваної матриці  $A$ . Матрицю  $B$ , подібну до  $A$ , формують так.

1. Обчислюють  $p := 2a_{ij}$ ,  $q := a_{ii} - a_{jj}$ ,  $d := \sqrt{p^2 + q^2}$

2. Якщо  $q \neq 0$ , то  $r := \frac{|q|}{2d}$ ,  $c := \sqrt{0,5+r}$ ,  $s := \sqrt{0,5-r} \cdot \text{sign}(pq)$

(якщо  $p \ll q$ , то ліпше  $s := |p| \cdot \text{sign}(pq)/(2cd)$ ), якщо ж  $q = 0$ , то  $c = s := \sqrt{2}/2$ ).

3. Обчислюють нові діагональні елементи

$$b_{ii} := c^2 a_{ii} + s^2 a_{jj} + 2csa_{ij},$$

$$b_{jj} := s^2 a_{ii} + c^2 a_{jj} - 2csa_{ij}.$$

4. Приймають  $b_{ij} = b_{ji} := 0$  (або для контролю обчислюють

$$b_{ij} = b_{ji} := (c^2 - s^2)a_{ij} + cs(a_{ji} - a_{ii}).$$

5. При  $m=1,2,\dots,n$  таких, що  $m \neq i$ ,  $m \neq j$ , обчислюють змінювані позадіагональні елементи:

$$b_{im} = b_{mi} := ca_{mi} + sa_{mj},$$

$$b_{jm} = b_{mj} := -sa_{mi} + ca_{mj}.$$

6. Для решти пар індексів  $m, l$  приймають  $b_{ml} := a_{ml}$ .

Звичайно, у реальних обчисленнях, якщо це вважати основою алгоритму, не все записане тут треба виконувати, зокрема, останні переприсвоювання. Також потрібно враховувати симетрію отримуваної матриці  $B$ .

Зазначимо, що найбільші вимоги до точності в описуваному методі ставлять саме на стадії обчислення  $c$  і  $s$ , оскільки тут можливі найбільші втрати точності, а спотворення  $c$  і  $s$  порушує ортогональність матриць  $T_{ij}$ . Це веде до неусувних похибок (метод поворотів, ітераційний за формою, не є ітераційним за суттю: йому невластиве самовиправлення методів послідовних наближень).

**П р и к л а д 4.7.** Методом поворотів Якобі розв'язати задачу

знаходження усіх власних пар матриці

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix}.$$

**Розв'язування.** Для заданої симетричної матриці  $A$  ми поетапно застосовуємо алгоритм, що переводить її в матрицю  $B$ , яка подібна до  $A$  і має меншу суму квадратів позадіагональних елементів. У цьому разі ключовий позадіагональний елемент  $a_{ij}$  братимемо у піддіагональній частині матриці  $A$  (можна і навпаки, у цьому разі принципово нічого не зміниться) і припускати, що всі обчислення будемо проводити точно (тобто розглядаємо ідеальний процес).

**Перший етап.** Виберемо ключовий елемент  $a_{31}$  (максимальний у зазначеній частині матриці) і зафіксуємо індекси  $i = 3, j = 1$ . Знайдемо

$$p = 2a_{31} = 6; \quad q = a_{33} - a_{11} = 1 - 1 = 0 \Rightarrow \quad c = s = \frac{1}{\sqrt{2}}.$$

Далі знайдемо елементи нової матриці

$$b_{33} = c^2 a_{33} + s^2 a_{11} + 2csa_{31} = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 1 + 2 \cdot \frac{1}{2} \cdot 3 = 4;$$

$$b_{11} = s^2 a_{33} + c^2 a_{11} - 2csa_{31} = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 1 - 2 \cdot \frac{1}{2} \cdot 3 = -2$$

і при  $m = 2$

$$b_{32} = b_{23} = ca_{23} + sa_{21} = \frac{1}{\sqrt{2}}(2 + 2) = 2\sqrt{2};$$

$$b_{12} = b_{21} = -sa_{23} + ca_{21} = \frac{1}{\sqrt{2}}(-2 + 2) = 0.$$

Отже,  $A \approx B = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 5 & 2\sqrt{2} \\ 0 & 2\sqrt{2} & 4 \end{pmatrix}.$

**Другий етап.** Отриману на попередньому етапі матрицю  $B$  вважаємо матрицею  $A$ , тобто будемо обчислювати за тими ж формулами, прийнявши  $A := B$ . Ключовий елемент тут  $a_{32} = 2\sqrt{2}$ .



тобто  $i=3, j=2$ . Згідно з алгоритмом,  $p=2a_{32}=4\sqrt{2}$ ;  
 $q=a_{33}-a_{22}=4-5=-1$  (зазначимо, що  $p \cdot q < 0$ ),

$$d = \sqrt{p^2 + q^2} = \sqrt{32+1} = \sqrt{33}, \quad r = \frac{|q|}{2d} = \frac{1}{2\sqrt{33}};$$

$$c = \sqrt{0,5+r} = \sqrt{0,5 + \frac{1}{2\sqrt{33}}} = \sqrt{\frac{\sqrt{33}+1}{2\sqrt{33}}};$$

$$s = \text{sign}(p \cdot q) \sqrt{0,5-r} = -\sqrt{0,5 - \frac{1}{2\sqrt{33}}} = -\sqrt{\frac{\sqrt{33}-1}{2\sqrt{33}}}.$$

Знаючи числа  $c$  і  $s$ , які визначають перетворення повороту, обчислимо

$$b_{33} = c^2 a_{33} + s^2 a_{22} + 2csa_{32} = \frac{\sqrt{33}+1}{2\sqrt{33}} \cdot 4 + \\ + \frac{\sqrt{33}-1}{2\sqrt{33}} \cdot 5 + 2 \cdot \left( \sqrt{\frac{\sqrt{33}+1}{2\sqrt{33}}} \right) \left( -\sqrt{\frac{\sqrt{33}-1}{2\sqrt{33}}} \right) \cdot 2\sqrt{2} \approx 1,6277;$$

$$b_{22} = s^2 a_{33} + c^2 a_{22} - 2csa_{32} = \frac{\sqrt{33}-1}{2\sqrt{33}} \cdot 4 + \\ + \frac{\sqrt{33}+1}{2\sqrt{33}} \cdot 5 - 2 \cdot \left( \sqrt{\frac{\sqrt{33}+1}{2\sqrt{33}}} \right) \left( -\sqrt{\frac{\sqrt{33}-1}{2\sqrt{33}}} \right) \cdot 2\sqrt{2} \approx 7,3723$$

і далі при  $m=1$  маємо

$$b_{31} = b_{13} = ca_{13} + sa_{12} = \sqrt{\frac{\sqrt{33}+1}{2\sqrt{33}}} \cdot 0 + \left( -\sqrt{\frac{\sqrt{33}-1}{2\sqrt{33}}} \right) \cdot 0 = 0;$$

$$b_{12} = b_{21} = -sa_{23} + ca_{21} = \sqrt{\frac{\sqrt{33}-1}{2\sqrt{33}}} \cdot 0 + \sqrt{\frac{\sqrt{33}+1}{2\sqrt{33}}} \cdot 0 = 0.$$

Отже, уже після другого етапу отримано діагональну матрицю

$$B = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 7,3723 & 0 \\ 0 & 0 & 1,6277 \end{pmatrix}.$$

подібну до початкової матриці  $A$ , з огляду на це можна стверджувати, що власними значеннями матриці  $A$  є числа

$$\lambda_1 = -2; \quad \lambda_2 = 7,3723; \quad \lambda_3 = 1,6277$$

Щоб знайти відповідні їм власні вектори, випишемо матрицю плоских поворотів першого і другого етапів (відповідно до їхньої структури):

$$T_y^{(1)} = T_{31} = \begin{pmatrix} c^{(1)} & 0 & s^{(1)} \\ 0 & 1 & 0 \\ -s^{(1)} & 0 & c^{(1)} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix};$$

$$T_y^{(2)} = T_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c^{(2)} & s^{(2)} \\ 0 & -s^{(2)} & c^{(2)} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{\frac{\sqrt{33}+1}{2\sqrt{33}}} & -\sqrt{\frac{\sqrt{33}-1}{2\sqrt{33}}} \\ 0 & \sqrt{\frac{\sqrt{33}-1}{2\sqrt{33}}} & \sqrt{\frac{\sqrt{33}+1}{2\sqrt{33}}} \end{pmatrix}.$$

Їхній добуток, тобто матриця

$$T = T_{31} \cdot T_{32} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \cdot \sqrt{\frac{\sqrt{33}-1}{2\sqrt{33}}} & \frac{1}{\sqrt{2}} \cdot \sqrt{\frac{\sqrt{33}+1}{2\sqrt{33}}} \\ 0 & \sqrt{\frac{\sqrt{33}+1}{2\sqrt{33}}} & -\sqrt{\frac{\sqrt{33}-1}{2\sqrt{33}}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \cdot \sqrt{\frac{\sqrt{33}-1}{2\sqrt{33}}} & \frac{1}{\sqrt{2}} \cdot \sqrt{\frac{\sqrt{33}+1}{2\sqrt{33}}} \end{pmatrix} =$$

$$= \begin{pmatrix} 0,7071 & 0,4544 & 0,5418 \\ 0 & 0,7662 & -0,6426 \\ -0,7071 & 0,4544 & 0,5418 \end{pmatrix},$$

згідно з теоретичними твердженнями, має стовпцями власні вектори матриці  $A$ .

Легко перевірити за означенням, що її стовпці у природному порядку, тобто вектори

$$x_1 = \begin{pmatrix} 0,7071 \\ 0 \\ -0,7071 \end{pmatrix}; \quad x_2 = \begin{pmatrix} 0,4544 \\ 0,7662 \\ 0,4544 \end{pmatrix}; \quad x_3 = \begin{pmatrix} 0,5418 \\ -0,6426 \\ 0,5418 \end{pmatrix} \quad \text{утворюють}$$

власні пари з числами  $\lambda_1 = -2$ ,  $\lambda_2 = 7,3723$ ,  $\lambda_3 = 1,6277$ , відповідно.

## 4.7. LU-алгоритм для несиметричних задач

Часто (принаймні в несиметричному випадку) алгоритми наближеного розв'язування повних проблем власних значень ґрунтуються на зведенні заданих матриць до подібних їм матриць не діагонального, а трикутного вигляду. Найпростіший з таких алгоритмів обчислення власних чисел ґрунтується на добре відомому  $LU$ -розкладі матриць.

Нехай задано  $(n \times n)$ -матрицю  $A$  у вигляді  $A = LU$ , де  $L$  і  $U$  – відповідно, нижня і верхня трикутні матриці. Позначимо  $A_1 = UL$  (нагадаємо, що добуток матриць не комутативний), тоді  $U = A_1 L^{-1}$ . Підставимо цей вираз матриці  $U$  в рівняння  $A = LU$  отримаємо нове зображення  $A$ :

$$A = L A_1 L^{-1} \quad (4.11)$$

яке свідчить про подібність матриць  $A$  і  $A_1$ , тобто рівність їхніх власних чисел  $\lambda_A$  і  $\lambda_{A_1}$ .

Якщо матрицю  $A_1$  можна, як і  $A$ , записати у вигляді добутку нижньої  $L_1$  і верхньої  $U_1$  трикутних, тобто  $A_1 = L_1 U_1$ , то, прийнявши  $A_2 = U_1 L_1$  і визначивши звідси  $U_1 = A_2 L_1^{-1}$  аналогічно до попереднього отримаємо

$$A_1 = L_1 A_2 L_1^{-1} \quad (4.12)$$

Відповідно,  $A_1$  подібна до  $A_2$  і, отже,  $\lambda_{A_1} = \lambda_{A_2}$ .

Суперпозиція цих двох перетворень, тобто підстановка (4.12) в (4.11), дає  $A$  через  $A_2$ :

$$A = LL_1 A_2 L_1^{-1} L^{-1} = LL_1 A_2 (L_1 L)^{-1},$$

звідки випливає рівність власних чисел  $\lambda_A$  і  $\lambda_{A_1}$ .

Такий процес побудови теоретично нескінченної послідовності подібних матриць і є основою  $LU$  (інакше  $LR$ ) -алгоритму. Він визначений фактично двома формулами:

$$A_k = L_k U_k, \quad A_{k+1} = U_k L_{k+1}, \quad (4.13)$$

де  $A_0 = A$ ,  $k = 0, 1, 2, \dots$ , причому перша з цих формул означає процедуру трикутної факторизації матриці  $A_k$  на  $k$ -му кроці, а друга – просте множення верхньої трикутної матриці на нижню.

**П р и к л а д 4.8.** Розглянемо, як поводить ся  $LU$ -алгоритм (4.13), застосований для знаходження власних чисел матриці

$$A = \begin{pmatrix} 1 & 5 \\ 4 & 2 \end{pmatrix}.$$

*Розв'язування.* Виконаємо  $LU$ -розклад за формулами

$$l_{ij} = a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj} \quad (i \geq j), \quad u_{ij} = \frac{1}{l_{ii}} \left( a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj} \right) \quad (i < j), \quad \text{отримаємо}$$

$$A_0 := A = L_0 U_0 = \begin{pmatrix} 1 & 0 \\ 4 & -18 \end{pmatrix} \begin{pmatrix} 1 & 5 \\ 0 & 1 \end{pmatrix}, \quad A_1 := U_0 L_0 = \begin{pmatrix} 21 & -90 \\ 4 & -18 \end{pmatrix}. \quad \text{Після}$$

факторизації цієї матриці аналогічно до попередньої маємо

$$A_1 = L_1 U_1 = \begin{pmatrix} 21 & 0 \\ 4 & -0,8571 \end{pmatrix} \begin{pmatrix} 1 & -4,2857 \\ 0 & 1 \end{pmatrix}, \quad \text{звідки}$$

$$A_2 := U_1 L_1 = \begin{pmatrix} 3,8571 & 3,6734 \\ 4 & -0,8571 \end{pmatrix}.$$

$$\text{Наступний крок дає } A_2 = L_2 U_2 = \begin{pmatrix} 3,8571 & 0 \\ 4 & -4,6667 \end{pmatrix} \begin{pmatrix} 1 & 0,9524 \\ 0 & 1 \end{pmatrix},$$

далі

$$A_3 := U_2 L_2 = \begin{pmatrix} 7,6667 & -4,4445 \\ 4 & -4,6667 \end{pmatrix};$$

$$A_3 = L_3 U_3 = \begin{pmatrix} 7,6667 & 0 \\ 4 & -2,3478 \end{pmatrix} \begin{pmatrix} 1 & -0,5797 \\ 0 & 1 \end{pmatrix};$$

$$A_4 := U_3 \tilde{L}_3 = \begin{pmatrix} 5,3478 & 1,3611 \\ 4 & -2,3478 \end{pmatrix};$$

$$A_4 = L_4 U_4 = \begin{pmatrix} 5,3478 & 0 \\ 4 & -3,3659 \end{pmatrix} \begin{pmatrix} 1 & 0,2545 \\ 0 & 1 \end{pmatrix};$$

$$A_5 := U_4 \tilde{L}_4 = \begin{pmatrix} 6,3659 & -0,8566 \\ 4 & -3,3659 \end{pmatrix}.$$

Якщо в тому ж прикладі фіксувати одиничну діагональ у матриць  $L_k$ , то процес (4.13), для якого використовують формули

$$u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj} \quad \text{при} \quad i \geq j, \quad l_{ij} = \frac{1}{u_{jj}} (a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj}) \quad \text{при} \quad i < j,$$

розвиватиметься так:

$$\tilde{A}_0 := A = \tilde{L}_0 \tilde{U}_0 = \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 5 \\ 0 & -18 \end{pmatrix}; \quad \tilde{A}_1 := \tilde{U}_0 \tilde{L}_0 = \begin{pmatrix} 21 & 5 \\ -72 & -18 \end{pmatrix};$$

$$\tilde{A}_1 = \tilde{L}_1 \tilde{U}_1 = \begin{pmatrix} 1 & 0 \\ -3,4286 & 1 \end{pmatrix} \begin{pmatrix} 21 & 5 \\ 0 & -0,8571 \end{pmatrix};$$

$$\tilde{A}_2 := \tilde{U}_1 \tilde{L}_1 = \begin{pmatrix} 3,8571 & 5 \\ 2,9388 & -0,8571 \end{pmatrix};$$

$$\tilde{A}_2 = \tilde{L}_2 \tilde{U}_2 = \begin{pmatrix} 1 & 0 \\ 0,7619 & 1 \end{pmatrix} \begin{pmatrix} 3,8571 & 5 \\ 0 & -4,6667 \end{pmatrix};$$

$$\tilde{A}_3 := \tilde{U}_2 \tilde{L}_2 = \begin{pmatrix} 7,6667 & 5 \\ -3,5556 & -4,6667 \end{pmatrix};$$

$$\tilde{A}_3 = \tilde{L}_3 \tilde{U}_3 = \begin{pmatrix} 1 & 0 \\ -0,4638 & 1 \end{pmatrix} \begin{pmatrix} 7,6667 & 5 \\ 0 & -2,3478 \end{pmatrix};$$

$$\tilde{A}_4 := \tilde{U}_3 \tilde{L}_4 = \begin{pmatrix} 5,3478 & 5 \\ 1,0888 & -2,3478 \end{pmatrix};$$

$$\bar{A}_4 = \bar{L}_4 \bar{U}_4 = \begin{pmatrix} 1 & 0 \\ 0,2036 & 1 \end{pmatrix} \begin{pmatrix} 5,3478 & 5 \\ 0 & -3,3659 \end{pmatrix};$$

$$\tilde{A}_5 := \tilde{U}_4 \tilde{L}_5 = \begin{pmatrix} 6,3659 & 5 \\ -0,6853 & -3,3659 \end{pmatrix}.$$

Діагоналі матриць  $A_k$  і  $\tilde{A}_k$ , які несуть наближення до власних чисел  $A$ , за одних і тих же значень  $k$  повністю збігаються.

**П р и к л а д 4.9.** Розглянемо, як поводить ся  $LU$ -алгоритм (4.13), застосований для знаходження власних чисел матриці

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 2 & 1 \\ 5 & 2 & 2 \end{pmatrix}.$$

*Розв'язування.* Виконаємо  $LU$ -розклад за формулами

$$l_{ij} = a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj}, \quad (i \geq j), \quad u_{ij} = \frac{1}{l_{ii}} (a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}), \quad (i < j), \quad \text{отримаємо}$$

$$A_0 = A = L_0 U_0 = \begin{pmatrix} 1 & 0 & 0 \\ 4 & -6 & 0 \\ 5 & -8 & 1,6666 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1,8333 \\ 0 & 0 & 1 \end{pmatrix};$$

$$A_1 = U_0 L_0 = \begin{pmatrix} 24 & -36 & 5 \\ 13,1666 & -20,6667 & 3,0555 \\ 5 & -8 & 1,6667 \end{pmatrix}.$$

Після факторизації цієї матриці аналогічно до попередньої маємо

$$A_1 = I_1 U_1 = \begin{pmatrix} 24 & 0 & 0 \\ 13,1667 & -0,9167 & 0 \\ 5 & -0,5 & 0,4545 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1,5 & 0,26 \\ 0 & 1 & -0,3 \\ 0 & 0 & 1 \end{pmatrix},$$

звідки

$$A_2 = U_1 L_1 = \begin{pmatrix} 5,2917 & 1,2708 & 0,0947 \\ 11,4621 & -0,5757 & 4,7004 \\ 5 & -0,5 & 0,4545 \end{pmatrix}.$$

Наступний крок дає

$$A_2 = L_2 U_2 = \begin{pmatrix} 5,2917 & 0 & 0 \\ 11,4621 & -3,3285 & 0 \\ 5 & -1,7008 & 8,0106 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0,2401 & 0,0179 \\ 0 & 1 & 4,4953 \\ 0 & 0 & 1 \end{pmatrix},$$

$$A_3 = U_2 L_2 = \begin{pmatrix} 8,1338 & -0,8119 & 0,1433 \\ 33,9419 & -10,9740 & 36,0101 \\ 5 & -1,7008 & 8,0106 \end{pmatrix}.$$

Для отримання остаточного результату обчислимо матрицю  $A$  на 10- і 20-му кроках, за допомогою відповідної програми. Отримаємо такий результат:

$$A_{10} = \begin{pmatrix} 7,1741 & 0,0007 & 0 \\ 11,9051 & -2,6917 & 0 \\ 5 & -1,2989 & 0,5176 \end{pmatrix};$$

$$A_{20} = \begin{pmatrix} 7,1751 & 0 & 0 \\ 11,9051 & -2,6927 & 0 \\ 5 & -1,2989 & 0,5176 \end{pmatrix}.$$

Якщо в тому ж прикладі фіксувати одиничну діагональ у матриць  $L_k$ , то процес (4.13), для якого використовують формули

$$u_{ij} = a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj} \quad (i \leq j); \quad l_{ij} = \frac{1}{u_{ij}} (a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj}), \quad (i < j), \quad \text{буде}$$

розвиватися так:

$$\tilde{A}_0 = A = \tilde{L}_0 \tilde{U}_0 = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 5 & 1,3333 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 0 & -6 & -11 \\ 0 & 0 & 1,6667 \end{pmatrix};$$

$$\tilde{A}_1 = \tilde{U}_0 \tilde{L}_0 = \begin{pmatrix} 24 & 5,9999 & 3 \\ -79 & -20,6667 & -11 \\ 8,3331 & 2,2223 & 1,6667 \end{pmatrix};$$

$$\tilde{A}_1 = \tilde{L}_1 \tilde{U}_1 = \begin{pmatrix} 1 & 0 & 0 \\ -3,2917 & 1 & 0 \\ 0,3472 & 0,1515 & 1 \end{pmatrix} \cdot \begin{pmatrix} 24 & 5,9999 & 3 \\ 0 & 0,9167 & -1,1249 \\ 0 & 0 & 0,7955 \end{pmatrix};$$

$$\tilde{A}_2 = \tilde{U}_1 \tilde{L}_1 = \begin{pmatrix} 5,2416 & 6,0454 & 3 \\ -3,408 & 0,7462 & -1,125 \\ 0,2762 & 0,1515 & 0,7955 \end{pmatrix};$$

$$\tilde{A}_2 = \tilde{L}_2 \tilde{U}_2 = \begin{pmatrix} 1 & 0 & 0 \\ -0,6512 & 1 & 0 \\ 0,0527 & -0,0357 & 1 \end{pmatrix} \cdot \begin{pmatrix} 5,2416 & 6,0454 & 3 \\ 0 & 4,6768 & 0,8255 \\ 0 & 0 & 0,6669 \end{pmatrix};$$

$$\tilde{A}_3 = \tilde{U}_2 \tilde{L}_2 = \begin{pmatrix} 8,1338 & 5,9383 & 3 \\ -3,0455 & -10,9740 & 0,8255 \\ 0,0351 & -0,0238 & 8,0106 \end{pmatrix}.$$

Для отримання остаточного результату обчислимо матрицю  $A$  на 10- і 20-му кроках за допомогою відповідної програми. Отримаємо такий результат:



$$\tilde{A}_{10} = \begin{pmatrix} 7,1741 & 5,7208 & 3 \\ 0,0016 & -2,6917 & -2,1108 \\ 0 & 0 & 0,5176 \end{pmatrix};$$

$$\tilde{A}_{20} = \begin{pmatrix} 7,1751 & 5,7208 & 3 \\ 0 & -2,6927 & -2,1112 \\ 0 & 0 & 0,5176 \end{pmatrix}.$$

Як бачимо, діагоналі матриць  $A_k$  і  $\tilde{A}_k$ , які становлять наближення до власних чисел матриці  $A$ , за одних і тих же значень  $k$  збігаються, а відносна швидкість спадання модулів наддіагональних елементів  $A_k$  і піддіагональних елементів  $\tilde{A}_k$  приблизно однакова.

#### 4.8. QR-алгоритм з використанням перетворень Хаусхолдера

Ідейно близьким до описаного  $LU$ -алгоритму є  $QR$ -алгоритм (В. Н. Кублановська, 1961; Дж. Френсіс, 1962). При  $k = 0, 1, 2, \dots$ , починаючи з  $A_0 := A$ , тут будують послідовність матриць  $A_k$  за формулами

$$A_k = Q_k R_k, \quad A_{k+1} = R_k Q_k, \quad (4.14)$$

перша з яких означає розклад матриці  $A_k$  в добуток ортогональної  $Q_k$  і правої трикутної  $R_k$  (такий розклад існує для довільної квадратної матриці), а друга – перемноження отриманих унаслідок факторизації  $A_k$  матриць  $Q_k$  і  $R_k$  у зворотному порядку.

Аналогічно до параграфа 4.7, на підставі властивості ортогональних матриць  $Q_k^T = Q_k^{-1}$ , відповідно до (4.14), матрицю  $A$  можна записати у вигляді

$$A = Q_0 Q_1 \dots Q_{k-1} Q_k A_{k+1} Q_k^T Q_{k-1}^T \dots Q_1^T Q_0^T$$

або

$$A = (Q_0 Q_1 \dots Q_{k-1} Q_k) A_{k+1} (Q_0 Q_1 \dots Q_{k-1} Q_k)^{-1}$$

Це означає, що кожна з матриць послідовності  $A_k$  ортогонально подібна до матриці  $A$ .

Якщо матриця  $A$  не має однакових за модулем власних значень, то за певних обмежень утворена процесом (4.14) послідовність матриць  $\{A_k\}$  збігається до матриці правої трикутної форми з діагоналлю із власних чисел. Швидкість занулення піддіагональних частин матриць  $A_k$  лінійна і залежить, як і в

багатьох інших методах, від співвідношення  $\frac{|\lambda_1|}{|\lambda_j|}$  при  $i > j$

(уважаємо  $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$ ).

У випадку наявності комплексно-спряжених пар власних чисел у заданій дійсній матриці  $A$  граничною матрицею для  $A_k$  буде матриця квазітрикутного (блоково-трикутного) вигляду. Кожній комплексній парі власних чисел у такій матриці відповідатиме діагональний  $2 \times 2$ -блок, причому збіжність тут простежується за формою матриці, а не поелементно (тобто елементи всередині цих блоків можуть змінюватися без помітної залежності від  $k$  зі збереженням незмінними їхніх власних чисел).

Як звичайно,  $QR$ -алгоритм (4.14) застосовують не до вихідної матриці  $A$ , а до подібної їй правої майже трикутної матриці  $B$ , яку називають матрицею Хессенберга:

$$B = \begin{pmatrix} b_{11} & b_{12} & & b_{1,n-1} & b_{1,n} \\ b_{21} & b_{22} & & b_{2,n-1} & b_{2,n} \\ 0 & b_{32} & & b_{3,n-1} & b_{3,n} \\ & 0 & & b_{n-1,n-1} & b_{n-1,n} \\ 0 & 0 & & b_{n,n-1} & b_{n,n} \end{pmatrix}.$$

В основі перетворення  $A$  до вигляду  $B$  (де  $b_{ij} = 0$  при  $j < i-1$ ) є перетворення Хаусхолдера, або перетворення відбиття, виконуване за допомогою матриці відбиття (Хаусхолдера)  $H = E - 2ww^T$ , де  $w$  – довільний вектор-стовпець, але такий, що його евклідова норма дорівнює одиниці. Як відомо, матриця відбиття ортогональна, тому матриці  $A$  і  $B$ , пов'язані співвідношенням  $B = HAH$  ( $= HAH^T = HAH^{-1}$ ), є подібними.

У разі побудови матриць відбиття елементи векторів  $w$  вибирають так, щоб за скінченну кількість кроків перетворень Хаусхолдера довільну задану матрицю  $A$  звести до форми Хессенберга  $B$ .

Справді, за початкованим (з  $B_1 = A$ ) процесом

$$B_{m+1} = H_m B_m H_m^T, \quad m=1, 2, \dots, n-2,$$

де

$$H_m = E - 2w_m w_m^T$$

задана  $n \times n$ -матриця  $A$  за  $n-2$  кроки буде зведена до вигляду  $B$ , тобто  $B := B_{n-1}$  подібна до  $A$ , якщо вектори  $w_m$ , які задають матриці Хаусхолдера  $H_m$ , за заданою матрицею  $A$  будувати так [14].

При  $m=1$  вектор  $w_1$  визначений рівністю

$$w_1^T = \mu_1 (0, a_{21} - s_1, a_{31}, \dots, a_{n1})$$

$$\text{де } s_1 = \text{sign}(-a_{21}) \sqrt{\sum_{i=2}^n a_{i1}^2}, \quad \mu_1 = \frac{1}{\sqrt{2s_1(s_1 - a_{21})}}.$$

Таке задання  $w_1$  забезпечує ортогональність симетричної матриці  $H_1 = E - 2w_1 w_1^T$  й одночасне отримання за її допомогою потрібних  $n-2$  нулів у першому стовпці матриці  $B_2 = H_1 B_1 H_1^T (= H_1 A H_1^T)$ .

Вектор  $w_2$  за матрицею  $B_2$  будують аналогічно, тільки фіксують нульовими не одну, а дві перші його координати. Визначальну роль тут відіграють не перший, а другий стовпець матриці  $B_2$  і його третій елемент. У цьому разі в матриці  $B_3 = H_2 B_2 H_2^T$  виявляються  $n-3$  нульові елементи у другому стовпці і зберігаються отримані на попередньому кроці нулі в першому стовпці.

Цей процес аналогічно можна продовжити до вичерпання матриці.

Зазначимо, що для симетричних матриць цей процес приводить до тридіагональної матриці.

Розглянемо на прикладі, як матрицю зводять до форми Хессенберга, коли для цього потрібний лише один крок перетворень Хаусхолдера.

**П р и к л а д 4.10.** Задано матрицю  $A = \begin{pmatrix} 4 & 2 & 1 \\ 2 & 1 & -2 \\ 5 & 1 & -1 \end{pmatrix}$ . Знайти

матрицю  $B$ , подібну до матриці  $A$ , яка має форму Хессенберга.

*Розв'язування.* Шукану матрицю знайдемо за такими формулами для  $n = 3$

$$s = \text{sign}(-a_{21})\sqrt{a_{21}^2 + a_{11}^2}; \mu = \frac{1}{\sqrt{2s(s-a_{21})}};$$

$$w^T = \mu(0; a_{21} - s; a_{31}); H = E - 2ww^T \quad B = HAH$$

Маємо

$$s = -\sqrt{4 + 25} = -\sqrt{29};$$

$$\mu = \frac{1}{\sqrt{-2\sqrt{29}(-\sqrt{29} - 2)}} = \frac{1}{\sqrt{58 + 4\sqrt{29}}};$$

$$ww^T = \frac{1}{29 + 2\sqrt{29}} \times$$

$$\times \begin{pmatrix} 0 \\ 2 + \sqrt{29} \\ 5 \end{pmatrix} \begin{pmatrix} 0, & 2 + \sqrt{29}, & 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{(2 + \sqrt{29})^2}{29 + 2\sqrt{29}} & \frac{10 + 5\sqrt{29}}{29 + 2\sqrt{29}} \\ 0 & \frac{10 + 5\sqrt{29}}{29 + 2\sqrt{29}} & \frac{25}{29 + 2\sqrt{29}} \end{pmatrix};$$

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{2}{\sqrt{29}} & -\frac{5}{\sqrt{29}} \\ 0 & -\frac{5}{\sqrt{29}} & \frac{2}{\sqrt{29}} \end{pmatrix}; \quad HA = \begin{pmatrix} 4 & 2 & 1 \\ -\sqrt{29} & -\frac{7}{\sqrt{29}} & \frac{9}{\sqrt{29}} \\ 0 & -\frac{3}{\sqrt{29}} & \frac{8}{\sqrt{29}} \end{pmatrix};$$

$$B = HAH = \begin{pmatrix} 4 & -\frac{9}{\sqrt{29}} & -\frac{8}{\sqrt{29}} \\ -\sqrt{29} & -\frac{31}{29} & \frac{53}{29} \\ 0 & -\frac{34}{29} & \frac{31}{29} \end{pmatrix} \approx \begin{pmatrix} 4 & -1,67 & -1,48 \\ -5,37 & -1,06 & 1,81 \\ 0 & -1,16 & -1,06 \end{pmatrix}.$$

Ми знайшли матрицю  $B$ , подібну до матриці  $A$ , яка має форму Хессенберга.

**Приклад 4.11.** Задано матрицю  $A = \begin{pmatrix} 4 & 2 & 1 \\ 2 & 1 & -2 \\ 1 & -2 & -1 \end{pmatrix}$ . Знайти

матрицю  $B$ , подібну до матриці  $A$ , яка має тридіагональну форму.

*Розв'язування.* Аналогічно до попереднього прикладу знайдемо

$$s = -\sqrt{4+1} = -\sqrt{5}; \quad \mu = \frac{1}{\sqrt{-2\sqrt{5}(-\sqrt{5}-2)}} = \frac{1}{\sqrt{10+4\sqrt{5}}};$$

$$ww^T = \frac{1}{5+2\sqrt{5}} \begin{pmatrix} 0 \\ 2+\sqrt{5} \\ 1 \end{pmatrix} \begin{pmatrix} 0, & 2+\sqrt{5}, & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{(2+\sqrt{5})^2}{5+2\sqrt{5}} & \frac{2+\sqrt{5}}{5+2\sqrt{5}} \\ 0 & \frac{2+\sqrt{5}}{5+2\sqrt{5}} & \frac{1}{5+2\sqrt{5}} \end{pmatrix};$$

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ 0 & -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}; \quad HA = \begin{pmatrix} 4 & 2 & 1 \\ -\sqrt{5} & 0 & \sqrt{5} \\ 0 & -\sqrt{5} & 0 \end{pmatrix};$$

$$B = \begin{pmatrix} 4 & -\sqrt{5} & 0 \\ -\sqrt{5} & -1 & 2 \\ 0 & 2 & 1 \end{pmatrix} \approx \begin{pmatrix} 4 & -2,2361 & 0 \\ -2,2361 & -1 & 2 \\ 0 & 2 & 1 \end{pmatrix}.$$

Ми знайшли матрицю  $B$ , подібну до матриці  $A$ , яка має тридіагональну форму.

Для обчислення власних значень симетричної матриці  $A$  теж використовують метод Хаусхолдера. Цю задачу розв'язують у два етапи: спочатку вихідну матрицю зводять до тридіагонального вигляду; після цього знаходять власні значення симетричної тридіагональної матриці. Якщо  $z$  – власний вектор тридіагональної матриці  $A_{n-1}$ , то  $Q_1 Q_2 \dots Q_{n-2} z$  – власний вектор матриці  $A$ .

Для виконання відображення елементи матриці  $Q$  не обов'язково записувати в явному вигляді, оскільки матриця повністю визначена вектором  $w$ . Метод Хаусхолдера для реалізації потребує  $\frac{2}{3}n^3$  операцій множення і  $n-2$  знаходження квадратного кореня.

Обчислена за методом Хаусхолдера тридіагональна матриця  $A_{n-1}$  ортогонально подібна до матриці  $A + d \cdot A$ , де  $\|d \cdot A\|_E \leq K n^2 2^{-t} \|A\|_E$ ;  $K$  – константа, яка залежить від способу заокруглення;  $t$  – кількість двійкових знаків мантиси числа.

Якщо матриця має кратні корені, то тридіагональна матриця має у випадку точних обчислень нульові недіагональні елементи. Водночас наявність нульового недіагонального елемента в тридіагональній матриці не є свідченням того, що матриця має кратні корені.

## 4.9. QR-алгоритм з використанням плоских поворотів Гівенса

Часто на другому етапі – зведенні матриці Хессенберга до верхньої трикутної матриці – застосовують інше перетворення, а саме: перетворення плоских поворотів Гівенса. Матриця  $G_{ij} = (g_{ml})_{m,l=1}^n$  яка визначає ці перетворення, у разі фіксованих індексів  $i, j$  головного елемента перетворюваної матриці має таку ж структуру, як і матриця плоских поворотів Якобі  $T_{ij}$ . Однак тут двовимірну підматрицю з елементів, які стоять на перетині  $i$ -ї  $j$ -х рядків та стовпців, візьмемо у вигляді [5]

$$\tilde{G}_{ij} = \begin{pmatrix} s & c \\ -c & s \end{pmatrix}.$$

Як і раніше, числа  $s$  і  $c$  пов'язані співвідношенням  $s^2 + c^2 = 1$  (це дає змогу інтерпретувати їх як синус та косинус деякого кута  $\theta$ , що забезпечує ортонормованість матриць  $G_{ij}$ ).

Перший повний крок перетворення Гівенса, застосованого до матриці Хессенберга  $B$   $n$ -го порядку, за QR-алгоритмом

$$A_k = Q_k R_k; \quad A_{k+1} = R_k Q_k \quad (4.15)$$

складається з  $n-1$  елементарних підкроків, кожен з яких занулює піддіагональні елементи в  $i$ -ствпцях від першого до  $n-1$ -го. У підсумку одержимо розклад матриці  $B$  в добуток ортогональної і трикутної, що потрібно за першою формулою (4.7) при  $k=1$ ,  $A_1 := B$ .

Оскільки в кожному стовпці матриці Хессенберга треба занулювати тільки один елемент, то в цьому випадку матриці поворотів Гівенса можна позначати тільки одним індексом  $G_i$ .

Прийmemo  $c = \cos \theta$  і  $s = \sin \theta$  такими, що  $\operatorname{tg} \theta = \frac{b_{11}}{b_{21}}$ , матимемо

$b_{11} \cos \theta = b_{21} \sin \theta$ , тобто  $-cb_{11} + sb_{21} = 0$ . Тому результат першого проміжного кроку – матриця  $B_1 = G_1 B$ , отримана за таких  $s$  і  $c$ , – не міститиме ненульових елементів під діагоналлю в першому стовпці.

Другий проміжний крок роблять аналогічно. Матрицю  $B_2 = G_2 B_1$  отримують з попередньої  $B_1$  за допомогою матриці Гівенса  $G_2$ , яка відрізняється від  $G_1$  тим, що матриця  $\begin{pmatrix} s & c \\ -c & s \end{pmatrix}$  зсувається на одну позицію вздовж діагоналі, а кут повороту підбирають так, щоб у матриці  $B_2$  занулити елемент  $b_{32}^{(2)}$ .

Якщо процес перетворень Гівенса продовжити, то в підсумку отримаємо праву трикутну матрицю  $B_{n-1} = G_{n-1} G_{n-2} \dots G_1 B$ . Цю рівність можна переписати у вигляді  $(G_{n-1} G_{n-2} \dots G_1)^{-1} B_{n-1} = B$ , який дає змогу вважати виконаним необхідний у (4.15) розклад  $B = Q_1 R_1$ , де  $Q_1 := (G_{n-1} G_{n-2} \dots G_2 G_1)^{-1} = G_1^T G_2^T \dots G_{n-2}^T G_{n-1}^T$  – ортогональна, а  $R_1 := B_{n-1}$  – права трикутна матриці. У цьому разі матриця

$$A_2 = R_1 Q_1 = (G_{n-1} \dots G_1) B (G_{n-1} \dots G_1)^{-1}$$

яка є підсумком першого повного кроку  $QR$ -алгоритму (застосованого до матриці  $B$ ), зберігає не тільки спектр заданої матриці, а й форму Хессенберга, завдяки чому зведення вихідної матриці  $A$  до майже трикутного вигляду  $B$  достатньо виконати тільки один раз.

Скалярні параметри  $c_j = \cos \theta_j$  і  $s_j = \sin \theta_j$  матриць Гівенса  $G_j$ , за допомогою яких виконують перехід від матриці Хессенберга  $B = (b_{ij})_{i,j=1}^n$  через матриці Хессенберга  $B_j = (b_{im}^{(j)})_{i,m=1}^n$  до матриці Хессенберга  $A_2$ , можна обчислити на  $j$ -му проміжному кроці ( $j = 1, 2, \dots, n-1$ ) за формулами

$$c_j = \frac{1}{\sqrt{1+t_j^2}}; \quad s_j = t_j c_j, \quad (4.16)$$

де  $t_j = \frac{b_{jj}^{(j-1)}}{b_{j+1,j}^{(j-1)}} \quad (= \tan \theta_j)$ ,  $b_y^{(0)} := b_y$  (якщо знаменник у виразі  $t_j$

дорівнює нулю або за модулем менший від деякого суттєво малого порогового значення, то можна вважати  $c_j = 0$ ,  $s_j = 1$ , тобто

$$G_j := E).$$



**Приклад 4.12.** Перетвореннями Гівенса виконати два кроки *QR*-алгоритму для матриці

$$B = \begin{pmatrix} 4 & -\frac{9}{\sqrt{29}} & -\frac{8}{\sqrt{29}} \\ -\sqrt{29} & \frac{31}{29} & \frac{53}{29} \\ 0 & -\frac{34}{29} & \frac{31}{29} \end{pmatrix}.$$

отриманої внаслідок перетворень Хаусхолдера з матриці

$$A = \begin{pmatrix} 4 & 2 & 1 \\ 2 & 1 & -2 \\ 5 & 1 & -1 \end{pmatrix} \text{ (див. приклад 4.10).}$$

*Розв'язування.* Для розв'язування цієї задачі використаємо формули (4.16).

При  $j = 1$  поступово знайдемо

$$\begin{aligned} t_1 &= -\frac{4}{\sqrt{29}}, \\ c_1 &= \frac{1}{\sqrt{1 + \frac{16}{29}}} = \frac{\sqrt{29}}{\sqrt{45}}, \\ s_1 &= -\frac{4}{\sqrt{29}} \cdot \frac{\sqrt{29}}{\sqrt{45}} = -\frac{4}{\sqrt{45}}; \end{aligned}$$

$$G_1 = \begin{pmatrix} -\frac{4}{\sqrt{45}} & \frac{\sqrt{29}}{\sqrt{45}} & 0 \\ \frac{\sqrt{29}}{\sqrt{45}} & -\frac{4}{\sqrt{45}} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -0,596285 & 0,802773 & 0 \\ -0,802773 & -0,596285 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$B_1 = G_1 B = \begin{pmatrix} -\frac{4}{\sqrt{45}} & \frac{\sqrt{29}}{\sqrt{45}} & 0 \\ \frac{\sqrt{29}}{\sqrt{45}} & \frac{4}{\sqrt{45}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 4 & -\frac{9}{\sqrt{29}} & -\frac{8}{\sqrt{29}} \\ -\sqrt{29} & -\frac{29}{31} & \frac{53}{29} \\ 0 & -\frac{29}{34} & \frac{29}{29} \end{pmatrix} =$$

$$= \begin{pmatrix} -\frac{45}{\sqrt{45}} & \frac{5}{\sqrt{45 \cdot 29}} & \frac{85}{\sqrt{45 \cdot 29}} \\ 0 & \frac{385}{29 \cdot \sqrt{45}} & \frac{20}{29 \cdot \sqrt{45}} \\ 0 & -\frac{34}{29} & \frac{31}{29} \end{pmatrix}.$$

При  $j = 2$  (заокруглюючи до  $10^{-6}$ )

$$t_2 = \frac{385}{\sqrt{45} \cdot 29} \left( -\frac{29}{34} \right) = -\frac{385}{34 \cdot \sqrt{45}} = -\frac{385}{228,0789} = -1,688012;$$

$$c_2 = \frac{1}{\sqrt{1 + \left( -\frac{385}{34 \cdot \sqrt{45}} \right)^2}} = \frac{1}{\sqrt{1 + \frac{148225}{1156 \cdot 45}}} =$$

$$= \frac{1}{\sqrt{1 + \frac{148225}{52020}}} = \frac{1}{1,961985} = 0,509687;$$

$$s_2 = -1,688012 \cdot 0,509687 = -0,860358.$$

Отже, отримаємо таку матрицю:

$$G_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -0,860358 & 0,509687 \\ 0 & -0,509687 & -0,860358 \end{pmatrix}.$$

Далі

$$R_1 = B_2 = G_2 \cdot B_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -0,860358 & 0,509687 \\ 0 & -0,509687 & -0,860358 \end{pmatrix} \times \\ \times \begin{pmatrix} -6,708204 & 0,138409 & 2,352961 \\ 0 & 1,979049 & 0,102808 \\ 0 & -1,172414 & 1,068965 \end{pmatrix} = \begin{pmatrix} -6,708204 & 0,138409 & 2,352961 \\ 0 & -2,300255 & 0,456385 \\ 0 & 0 & -0,972092 \end{pmatrix}.$$

Одержимо

$$Q_1 = G_1^T G_2^T = \begin{pmatrix} -0,596285 & -0,802773 & 0 \\ 0,802773 & -0,596285 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \\ \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & -0,860358 & -0,509687 \\ 0 & 0,509687 & -0,860358 \end{pmatrix} = \begin{pmatrix} -0,596285 & 0,690672 & 0,409163 \\ 0,802773 & 0,513018 & 0,303919 \\ 0 & 0,509687 & -0,860358 \end{pmatrix}; \\ A_2 = R_1 Q_1 = \begin{pmatrix} -6,708204 & 0,138409 & 2,352961 \\ 0 & -2,300255 & 0,456385 \\ 0 & 0 & -0,972092 \end{pmatrix} \times \\ \times \begin{pmatrix} -0,596285 & 0,690672 & 0,409163 \\ 0,802773 & 0,513018 & 0,303919 \\ 0 & 0,509687 & -0,860358 \end{pmatrix} = \begin{pmatrix} 4,111112 & -3,362889 & -4,727073 \\ -1,846583 & -0,947459 & -1,091745 \\ 0 & -0,495463 & 0,836347 \end{pmatrix}.$$

Перейдемо на наступний крок  $QR$ -алгоритму, беручи за матрицю  $B$  отриману матрицю  $A_2$ .

При  $j = 1$  поступово знайдемо

$$t_1 = -2,226335; \quad c_1 = 0,4097344; \quad s_1 = -0,912204;$$

$$G_1 = \begin{pmatrix} -0,912204 & 0,409734 & 0 \\ -0,409734 & -0,9122048 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$B_1 = G_1 B = \begin{pmatrix} -4,506785 & 2,679443 & 3,864733 \\ -0,000000 & 2,242174 & 2,932744 \\ -0,000000 & -0,495464 & 0,836350 \end{pmatrix}.$$

При  $j = 2$

$$t_2 = -4,525395; \quad c_2 = 0,215769; \quad s_2 = -0,976444.$$

Отже, отримаємо таку матрицю:  $G_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -0,976444 & 0,215769 \\ 0 & -0,215769 & 0,976444 \end{pmatrix}.$

Далі

$$R_1 = B_2 = G_2 \cdot B_1 = \begin{pmatrix} -4,506785 & 2,679443 & 3,864733 \\ 0,000000 & -2,296264 & -2,683202 \\ 0,000000 & -0,000000 & -1,449447 \end{pmatrix}.$$

Остаточно

$$Q_1 = G_1^T G_2^T = \begin{pmatrix} -1,449447 & 0,400082 & 0,088408 \\ 0,409734 & 0,890717 & 0,196826 \\ 0 & 0,215769 & -0,976444 \end{pmatrix};$$

$$A_2 = R_1 Q_1 = \begin{pmatrix} 5,208971 & 1,417432 & -3,644749 \\ -0,940858 & -2,624276 & 2,168032 \\ -0,000000 & 0,312747 & 1,415304 \end{pmatrix}.$$

Отже, ми за допомогою перетворень Гівенса зробили два кроки  $QR$ -алгоритму для матриці, отриманої внаслідок перетворень Хаусхолдера, і одержали матрицю  $A_2$ . Вона має ті ж власні значення, що й  $B$  і  $A$ , зберігає форму Хессенберга, і модулі її піддіагональних елементів менші, ніж у матриці  $B$ , однак вона ближча до подібної  $B$  матриці трикутного вигляду, на діагоналі якої повинні бути власні числа заданої матриці

$$(\lambda_1=5,000000000000000; \lambda_2 \approx -2,30277563773199; \\ \lambda_3 \approx 1,30277563773199).$$

#### 4.10. Метод Ланцоша зведення симетричної матриці до тридіагонального вигляду

Метод Ланцоша – це алгоритм зведення симетричної матриці  $A$  до тридіагональної матриці  $T$ . Основою алгоритму є рекурентне співвідношення

$$q_{j+1}\beta_j = Aq_j - q_j d_j - q_{j-1}\beta_{j-1}, \quad j=1,2,\dots,n, \quad q_0=0. \quad (4.17)$$

Це співвідношення впливає з такої матричної рівності:

$$Q_n T_n = A Q_n,$$

де  $Q_n = \{q_1, q_2, \dots, q_n\}$  – ортогональна матриця;  $T_n$  – тридіагональна матриця

$$T_n = \begin{pmatrix} \alpha_1 & \beta_1 & 0 & 0 & 0 \\ \beta_1 & \alpha_2 & \beta_2 & 0 & 0 \\ 0 & \beta_2 & \alpha_3 & 0 & 0 \\ 0 & 0 & 0 & \alpha_{n-1} & \beta_{n-1} \\ 0 & 0 & 0 & \beta_{n-1} & \alpha_n \end{pmatrix}.$$

Якщо  $T_n$  нерозкладна (тридіагональну матрицю називають нерозкладною, якщо всі  $\beta_j \neq 0$ ), то, не обмежуючи загальності,

можна вважати, що  $\beta_j > 0$ . Тоді співвідношення (4.17) дає змогу записати алгоритм побудови матриці  $T_n$  за матрицею  $A$

Задамо початковий вектор  $\|q_1\|=1$ , потім для  $j=1,2,\dots,n$  повторимо таке:

$$1) u_j = Aq_j;$$

$$2) \alpha_j = (q_j, Aq_j);$$

$$3) r_j = u_j - q_j \alpha_j - q_{j-1} \beta_{j-1}; \quad q_0 = 0;$$

$$4) \beta_j = \|r_j\|;$$

$$5) q_{j+1} = r_j / \beta_j.$$

Після виконання  $n$  кроків такого алгоритму в точній арифметиці отримаємо тридіагональну матрицю  $T_n$ , ортогонально подібну до вихідної матриці  $A$ . Отже, задача знаходження власних значень симетричної матриці зводиться до знаходження власних значень тридіагональної матриці.

Припинення процесу може відбутися, якщо  $\beta_j = 0$ . Це трапиться в тому випадку, коли вектор  $q_j$  ортогональний принаймні до одного власного вектора  $A$ . Якщо  $\beta_j = 0$ , то за  $q_{j+1}$  можна прийняти будь-який одиничний вектор, ортогональний до попередніх векторів  $q_j$ , і продовжити процес.

**П р и к л а д 4.13.** Методом Ланцоша звести до тридіагонального вигляду матрицю  $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix}$ .

**Розв'язування.** За початковий вектор  $q_1$  приймемо одиничний вектор  $q_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ . Обчислимо  $u_1 = Aq_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ ;  $\alpha_1 = 1$ .

За відповідною формулою знайдемо  $r_1$ :

$$r_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}; \quad \beta_1 = \sqrt{4+9} = \sqrt{13}$$

Аналогічно, як на попередньому кроці, визначимо  $q_2, u_2, \alpha_2, r_2, \beta_2$ :

$$q_2 = \begin{pmatrix} 0 \\ \frac{2}{\sqrt{13}} \\ \frac{3}{\sqrt{13}} \end{pmatrix}, \quad u_2 = \begin{pmatrix} \frac{13}{\sqrt{13}} \\ \frac{16}{\sqrt{13}} \\ \frac{7}{\sqrt{13}} \end{pmatrix}, \quad \alpha_2 = \frac{32}{\sqrt{13}} + \frac{21}{\sqrt{13}} = \frac{53}{\sqrt{13}};$$

$$r_2 = \begin{pmatrix} \frac{13}{\sqrt{13}} \\ \frac{16}{\sqrt{13}} \\ \frac{7}{\sqrt{13}} \end{pmatrix} - \begin{pmatrix} 0 \\ \frac{106}{13} \\ \frac{159}{13} \end{pmatrix} - \begin{pmatrix} \sqrt{13} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{16\sqrt{13}-106}{13} \\ \frac{7\sqrt{13}-159}{13} \end{pmatrix} \approx \begin{pmatrix} 0 \\ -3,72 \\ -10,29 \end{pmatrix};$$

$$\beta_2 \approx \sqrt{13,8384 + 105,8841} = \sqrt{119,7225} \approx 10,94.$$

На третьому кроці обчислимо  $q_3, u_3, \alpha_3$ :

$$q_3 \approx \begin{pmatrix} 0 \\ -0,34 \\ -0,94 \end{pmatrix}; \quad u_3 = \begin{pmatrix} -3,5 \\ -3,58 \\ -1,62 \end{pmatrix}; \quad \alpha_3 = 1,2172 + 1,5228 \approx 2,74.$$

Отже, отримано таку тридіагональну матрицю:

$$T = \begin{pmatrix} 1 & \sqrt{13} & 0 \\ \sqrt{13} & \frac{53}{\sqrt{13}} & 10,94 \\ 0 & 10,94 & 2,74 \end{pmatrix} \approx \begin{pmatrix} 1 & 3,6 & 0 \\ 3,6 & 14,72 & 10,94 \\ 0 & 10,94 & 2,74 \end{pmatrix}.$$

Однак у разі реалізації розглянутого алгоритму на практиці для великих  $n$  строга ортогональність векторів  $q_i$ , зазвичай, втрачається. Головною причиною цього є скінченна розрядність комп'ютерів, що призводить до похибок заокруглень. У цьому

випадку можна використати метод Ланцоша з дуже точною переортогоналізацією. Однак така процедура досить трудомістка.

## 4.11. Метод бісекцій

Для обчислення власних значень тридіагональних матриць розроблено досить ефективний метод бісекцій (половинного поділу). Його також можна вважати одним зі способів локалізації власних чисел. Розглянемо цей метод детальніше.

Нехай задано тридіагональну матрицю

$$A = \begin{pmatrix} a_1 & b_2 & 0 & & 0 & 0 & 0 \\ b_2 & a_2 & b_3 & & 0 & 0 & 0 \\ 0 & 0 & 0 & & b_{n-1} & a_n & b_n \\ 0 & 0 & 0 & & 0 & b_n & a_n \end{pmatrix}.$$

Не зменшуючи загальності, можна припустити, що жоден з елементів  $b_i$  не дорівнює нулю. Справді, нехай для деякого  $q$  елемент  $b_q = 0$ . Тоді можна розглядати тридіагональні матриці:  $A_1$  порядку  $q-1$  і  $A_2$  порядку  $n-q+1$ , для них

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}.$$

Власні значення матриць  $A_1$  і  $A_2$  є власними значеннями матриці  $A$ . Якщо  $v$  – власний вектор матриці  $A_1$ , то відповідний власний вектор матриці  $A$  буде  $(v^T, \underbrace{0, \dots, 0}_{n-q+1})^T$ .

Можна довести, що в симетричній тридіагональній матриці, яка не має ненульових недіагональних елементів, усі власні значення різні. Якщо симетрична матриця має кратну кількість власних значень, то тридіагональна матриця, побудована з неї методом Хаусхолдера, матиме декілька близьких до нуля недіагональних елементів. Тому метод бісекцій можна розглядати для симетричних тридіагональних матриць з ненульовими недіагональними елементами.



Метод половинного поділу ґрунтується на властивості послідовності Штурма. Якщо інтервал  $(A_0, B_0)$ , на якому шукають власне значення, невідомий, то прийнятними для  $A_0$  і  $B_0$  є значення  $\pm \|A\|_\infty$ . Оскільки матриця має тридіагональний вигляд, то цю норму легко знайти:

$$\|A\|_\infty = \max \{ |b_i| + |a_i| + |b_{i+1}| \}; \quad b_1 = b_{n+1} = 0.$$

Схема обчислення методу половинного поділу для визначення  $k$ -го власного значення може бути організована так. Перед обчисленням задають значення  $A_0, B_0$  і  $\varepsilon$  – умова закінчення ітераційного процесу. Потім ділять інтервал  $(A_{r-1}, B_{r-1})$  ( $r$  – номер ітерації,  $r = 1, 2, \dots$ ) навпіл. Точка  $C_r = \frac{1}{2}(A_{r-1} + B_{r-1})$  є його серединою. Обчислюють послідовність  $p_0(C_r), p_1(C_r), \dots, p_n(C_r)$  за правилом

$$\begin{aligned} p_0(C_r) &= 1; & p_1(C_r) &= a_1 - C_r; \\ p_k(C_r) &= (a_k - C_r) \cdot p_{k-1} - b_k \cdot p_{k-2}; & 2 \leq k \leq n, \end{aligned}$$

і визначають кількість збігів знаків  $s(C_r)$

Якщо  $s(C_r) \geq k$ , то вважають, що  $A_r = C_r, B_r = B_{r-1}$ , а у випадку  $s(C_r) < k$  вважають, що  $A_r = A_{r-1}, B_r = C_r$ . Тепер  $\lambda_k \in (A_r, B_r)$ . Цей процес повторюють доти, доки не буде виконана умова

$$\frac{B_r - A_r}{2} < \varepsilon C_{r+1}$$

У разі виконання цієї умови значення  $C_{r+1}$  приймають за наближене значення  $\lambda_k$ .

На кожному кроці методу бісекцій необхідно  $2n$  операцій множення і  $2n$  операцій віднімання.

Опишемо застосування цього методу на прикладах.

**П р и к л а д 4.14.** Нехай потрібно знайти третє власне значення матриці

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$$

методом половинного поділу.

*Розв'язування.* Оскільки інтервал, у якому є третє власне значення, невідомий, то визначимо  $\|A\|_{\infty} = 4$  і власні значення будемо шукати в інтервалі  $(-4, 4)$ . Справді, підставивши значення  $-4$  в послідовність Штурма, отримаємо

$$p_0(-4) = 1; \quad p_1(-4) = 6; \quad p_2(-4) = 35; \quad p_3(-4) = 204; \quad p_4(-4) = 1184.$$

Отже, кількість збігів знака  $s(-4) = 4$ . Це означає, що всі чотири власні значення матриці  $A$  більші, ніж  $-4$ . Підставимо тепер у послідовність Штурма число  $4$ , отримаємо

$$p_0(4) = 1, \quad p_1(4) = -2, \quad p_2(4) = 3, \quad p_3(4) = -4, \quad p_4(4) = 5.$$

Кількість збігів знаків  $s(4) = 0$ , і всі власні значення матриці  $A$  не більші від чотирьох.

Тоді ділимо інтервал  $(-4, 4)$  навпіл, отримаємо

$$C_1 = \frac{1}{2}(-4 + 4) = 0$$

і побудуємо послідовність Штурма для цього значення  $C_1$ :

$$p_0(0) = 1; \quad p_1(0) = 2; \quad p_2(0) = 3; \quad p_3(0) = 4.$$

Оскільки маємо три збіги знаків, то подальші обчислення значень поліномів можна не виконувати. Отже,  $s(0) \geq 3$  означає, що  $\lambda_3 \in (0, 4)$ . Далі маємо

$$C_2 = \frac{1}{2}(0 + 4) = 2,$$

$$p_0(2) = 1; \quad p_1(2) = -0; \quad p_2(2) = -1; \quad p_3(2) = +0; \quad p_4(2) = 1.$$

Знак нуля в  $p_1(2)$  беруть протилежний до знака  $p_0(2)$ , а знак  $p_2(2)$  – протилежний до знака  $p_1(2)$ . У побудованій послідовності Штурма  $s(2) = 2 < 3$ . Це означає  $\lambda_3 \in (0, 2)$ . Тоді ділимо  $(0, 2)$  навпіл:

$$C_3 = \frac{1}{2}(0 + 2) = 1;$$

$$p_0(1) = 1; \quad p_1(1) = 1; \quad p_2(1) = -0; \quad p_3(1) = -1; \quad p_4(1) = -1.$$

У розглянутій послідовності Штурма  $s(1) = 3$  означає  $\lambda_3 \in (1, 2)$  і т. д. Продовживши цей процес, отримуємо  $\lambda_3 \approx 1,382$ , що добре узгоджується з істинним значенням  $\lambda_3 = 1,381966\dots$

**П р и к л а д 4.15.** Знайти друге власне число матриці  $A$  методом бісекцій, якщо

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 1 \end{pmatrix}.$$

*Розв'язування.* Аналогічно до попереднього прикладу визначаємо  $\|A\|_{\infty} = 7$ . Тому початковий інтервал, на якому міститься друге власне значення, є  $(-7, 7)$ . Справді, підставивши значення  $-7$  у послідовність Штурма, маємо

$$p_0(-7) = 1; \quad p_1(-7) = 8; \quad p_2(-7) = 76; \quad p_3(-7) = 576.$$

Отже, кількість збігів знака  $s(-7) = 3$ . Це означає, що всі три власні значення матриці  $A$  більші, ніж  $-7$ . Підставимо тепер у послідовність Штурма число 7, отримуємо

$$p_0(7) = 1; \quad p_1(7) = -6; \quad p_2(7) = 20; \quad p_3(7) = -96.$$

Кількість збігів знаків  $s(7) = 0$ , і всі власні значення матриці  $A$  не більші від семи.

Тоді поділимо інтервал  $(-7, 7)$  навпіл, отримуємо

$$C_1 = \frac{1}{2}(-7 + 7) = 0$$

і побудуємо послідовність Штурма для цього значення  $C_1$ :

$$p_0(0) = 1; \quad p_1(0) = 1; \quad p_2(0) = -1; \quad p_3(0) = -5.$$

Оскільки  $s(0) = 2$ , то  $\lambda_2 \in (0, 7)$ . Далі маємо

$$C_2 = \frac{1}{2}(0 + 7) = 3,5;$$

$$p_0(3,5) = 1; \quad p_1(3,5) = -2,5; \quad p_2(3,5) = -2,75; \quad p_3(3,5) = 16,875.$$

Оскільки  $s(3,5) = 1 < 2$ , то  $\lambda_2 \in (0; 3,5)$ ;

$$C_3 = \frac{1}{2}(0 + 3,5) = 1,75;$$

$$p_0(1,75) = 1; \quad p_1(1,75) = -0,75; \quad p_2(1,75) = -4,9375;$$

$$p_3(1,75) = 6,703125.$$

Тут  $s(1,75) = 1 < 2$  тому  $\lambda_2 \in (0; 1,75)$  і т. д. Продовживши цей процес, отримаємо  $\lambda_2 \approx 1,001$  (з точністю  $\varepsilon = 0,01$ ), а оскільки істинне  $\lambda_2 = 1$ , то побачимо, що обчислення виконано правильно.

Зазначимо, що розглянутий метод визначення власних значень матриці універсальний. Його можна використовувати не тільки для знаходження заданого за номером власного значення, а й для обчислення всіх чи частини власних значень, для дослідження загального розподілу власних значень тощо. На його реалізацію не впливає наявність близьких і кратних власних значень, і навіть дуже велике їхнє скупчення. У цьому разі точність не залежить від розмірів матриці.

#### **4.12. QL-метод знаходження власних значень симетричної тридіагональної матриці**

Цей метод знаходження власних значень симетричної тридіагональної матриці зводиться до перетворення в ітераційному процесі вихідної тридіагональної матриці  $A$  в діагональну за допомогою перетворень подібності

$$A_{k+1} = Q_k A_k Q_k^T = Q_k Q_{k-1} A_{k-1} Q_{k-1}^T Q_k^T = Q_k \dots Q_1 A_1 Q_1^T \dots Q_k^T$$

Тут  $A_1 = A$ .

По суті,  $QL$ -метод реалізують так:

$$A_{k+1} = L_k Q_k^T \quad k = 1, 2, \dots,$$

де  $Q_k$  – ортогональна матриця, що перетворює  $A_k$  в ліву трикутну матрицю  $L_k$ ;  $k$  – номер ітерації. За досить великого  $k$  матриця  $A_k$  стає близькою до діагональної. Її діагональні елементи є наближеними власними значеннями матриці  $A$ . Матриця перетворення  $Q_k$  на кожному кроці ітерації є добутком елементарних матриць поворотів чи відбиттів.

Розглянемо варіант методу, що ґрунтується на використанні матриць поворотів. Нехай задано матрицю

$$A_k = \begin{pmatrix} d_1^{(k)} & l_1^{(k)} & 0 & 0 & 0 & 0 \\ l_1^{(k)} & d_2^{(k)} & l_2^{(k)} & 0 & 0 & 0 \\ 0 & 0 & 0 & l_{n-2}^{(k)} & d_{n-1}^{(k)} & l_n^{(k)} \\ 0 & 0 & 0 & 0 & l_{n-1}^{(k)} & d_n^{(k)} \end{pmatrix}.$$

Анулюємо елемент  $l_{n-1}^{(k)}$  використовуючи матрицю повороту, яка має вигляд

$$P_{n-1}^{(k)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c & -s \\ 0 & 0 & s & c \end{pmatrix}.$$

Унаслідок перетворення  $P_{n-1}^{(k)} A_k$  відбувається занулювання елемента  $l_{n-1}^{(k)}$ , а перетворення  $P_{n-1}^{(k)} A_k (P_{n-1}^{(k)})^T$  веде, з одного боку, до появи нового елемента  $l_{n-1}^{(k)}$ , меншого за значенням, ніж попередній, а з іншого, – до появи на місці  $(n-2, n)$  і, з огляду на симетрію матриці, на місці  $(n, n-2)$  ненульового елемента  $g_i^{(k)}$ . На схемі 1 елементи матриці  $A_k$  позначені значком  $\times$ , елементи  $(n, n-1)$  і  $(n-1, n)$ , які знову з'явились, –  $\bullet$ , а елемент  $g_1^{(k)}$  –  $\otimes$ .

$$\begin{pmatrix} \times & \times & & & & \\ \times & \times & \times & & & \\ & \times & \times & \times & & \\ & & \times & \times & \times & \otimes \\ & & & \times & \times & \bullet \\ & & & & \otimes & \times \end{pmatrix}$$

Схема 1

$$\begin{pmatrix} \times & \times & & & & \\ \times & \times & \times & & & \\ & \times & \times & \times & \otimes & \\ & & \times & \times & \times & 0 \\ & & \otimes & \times & \times & \bullet \\ & & & 0 & \bullet & \times \end{pmatrix}$$

Схема 2

Отже, на цьому етапі обчислень матриця  $A_k$  перестає бути тридіагональною, тому її необхідно звести до тридіагональної. Для цього обчислюють коефіцієнти поворотів

$$c_{n-1}^{(k)} = \frac{l_{n-2}^{(k)}}{\sqrt{(l_{n-2}^{(k)})^2 + (g_1^{(k)})^2}} \quad s_{n-1}^{(k)} = \frac{g_1^{(k)}}{\sqrt{(l_{n-2}^{(k)})^2 + (g_1^{(k)})^2}}$$

і перетворюють рядки й стовпці з номерами  $n-1$  і  $n-2$  за формулами методу Гівенса.

У підсумку  $g_1$  дорівнює нулю, а на місці елементів з індексами  $(n-3, n-1)$  і  $(n-1, n-3)$  з'являється елемент  $g_2^{(k)}$  (схема 2).

Для анулювання елемента, що виникає, знову застосовують формули Гівенса і т. д. Лише анулювання елемента  $g_{n-2}^{(k)}$ , який має індекси  $(1, 3)$ ,  $(3, 1)$ , не спричиняє появи елемента  $g$ . На цьому крок процесу вважають закінченим. Кроки ітераційного процесу повторюють доти, доки елемент  $l_{n-1}^{(k)}$  не стане достатньо малим. У цьому випадку елемент з індексами  $(n, n)$  можна взяти за наближення до власного значення матриці  $A$ . Після отримання наближеного власного значення розглядають матрицю, отриману з вихідної викреслюванням останнього стовпця й останнього рядка, порядок якого на одиницю менший від попереднього, і повторюють знову всю ітераційну процедуру обчислення наступного власного значення. І так продовжують доти, доки не обчислять усі власні значення з заданим ступенем точності. Отже, замість тридіагональної матриці отримують наближено діагональну.

**П р и к л а д 4.16.** Знайти  $QL$ -методом власні значення тридіагональної матриці

$$A = \begin{pmatrix} 3 & -2 & 0 \\ -2 & 4 & -2 \\ 0 & -2 & 3 \end{pmatrix}.$$

*Розв'язування.*

**Крок 1.** Занулимо елемент  $a_{23}$ . Для цього знайдемо  $s$  і  $c$  в матриці поворотів  $P_2$ :

$$s = \frac{a_{23}}{\sqrt{a_{23}^2 + a_{33}^2}} = -0,5547; \quad c = \frac{a_{33}}{\sqrt{a_{23}^2 + a_{33}^2}} = 0,8321.$$

Матриця поворотів матиме вигляд

$$P_2 = \begin{pmatrix} 1,0000 & 0 & 0 \\ 0 & 0,8321 & 0,5547 \\ 0 & -0,5547 & 0,8321 \end{pmatrix}.$$

Нескладно переконатись, що

$$P_2 A P_2^T = \begin{pmatrix} 3,0000 & -1,6640 & 1,1094 \\ -1,6641 & 1,8462 & -1,2308 \\ 1,1094 & -1,2308 & 5,1538 \end{pmatrix}.$$

Тепер занулимо елемент  $a_{13}$  матриці  $P_2 A P_2^T$ . Для цього обчислимо  $s$  і  $c$ :

$$s = \frac{a_{13}}{\sqrt{a_{23}^2 + a_{13}^2}} = 0,6695; \quad c = \frac{a_{23}}{\sqrt{a_{23}^2 + a_{13}^2}} = -0,7428,$$

після чого отримуємо нову матрицю

$$P_3 = \begin{pmatrix} -0,7428 & -0,6695 & 0 \\ 0,6695 & -0,7428 & 0 \\ 0 & 0 & 1,0000 \end{pmatrix}$$

і

$$P_3 A P_3^T = \begin{pmatrix} 0,8276 & -0,7460 & 0 \\ 0,7102 & 4,0186 & 1,6570 \\ 0 & 1,6570 & 5,1538 \end{pmatrix}.$$

**Крок 2.** Прийmemo  $A = P_3 A P_3^T$  виконаємо операції аналогічно до кроку 1.

Занулимо елемент  $a_{23}$ . Для цього знайдемо

$$s = \frac{a_{23}}{\sqrt{a_{23}^2 + a_{33}^2}} = 0,3061; \quad c = \frac{a_{33}}{\sqrt{a_{23}^2 + a_{33}^2}} = 0,9520.$$

Матриця поворотів матиме вигляд

$$P_2 = \begin{pmatrix} 1,0000 & 0 & 0 \\ 0 & 0,9520 & 0,3061 \\ 0 & 0,3061 & 0,9520 \end{pmatrix}$$

i

$$P_2 A P_2^T = \begin{pmatrix} 0,8276 & -0,7102 & -0,2283 \\ -0,7102 & 3,1593 & 1,0157 \\ -0,2283 & 1,0157 & 6,0131 \end{pmatrix}.$$

Тепер занулимо елемент  $a_{13}$  матриці  $P_2 A P_2^T$ . Для цього обчислимо

$$s = \frac{a_{13}}{\sqrt{a_{23}^2 + a_{13}^2}} = -0,2193; \quad c = \frac{a_{23}}{\sqrt{a_{23}^2 + a_{13}^2}} = 0,9757$$

після чого отримаємо нову матрицю

$$P_3 = \begin{pmatrix} 0,9757 & 0,2193 & 0 \\ -0,2193 & 0,9757 & 0 \\ 0 & 0 & 1,0000 \end{pmatrix}$$

i

$$P_3 A P_3^T = \begin{pmatrix} 0,6358 & -0,1429 & 0,0000 \\ -0,1429 & 3,3511 & 1,0411 \\ -0,0000 & 1,0411 & 6,0131 \end{pmatrix}.$$



Після 12 кроків ми отримаємо матрицю

$$\begin{pmatrix} 0,6277 & 0,0000 & 0,0000 \\ 0,0000 & 3,0000 & 0,0006 \\ 0,0000 & 0,0006 & 6,3723 \end{pmatrix},$$

з якої бачимо, що з точністю до трьох десяткових знаків ми обчислили наближені власні значення  $\lambda_1 = 6,3723$ ;  $\lambda_2 = 3,0000$ ;  $\lambda_3 = 0,6277$ .

## Завдання до розділу 4

### 4.1. Задано матриці:

$$1) \begin{pmatrix} 6 & -2 & 0 & 3 \\ -2 & 6 & 1 & 2 \\ 0 & 1 & 4 & 1 \\ 3 & 2 & 0 & 6 \end{pmatrix};$$

$$2) \begin{pmatrix} 7 & -1 & -2 & -3 \\ -1 & 6 & 0 & 2 \\ -2 & 0 & 5 & 1 \\ 3 & 2 & 1 & 7 \end{pmatrix};$$

$$3) \begin{pmatrix} 8 & 1 & 2 & 3 & 0 \\ 1 & 6 & 1 & 2 & 1 \\ 2 & 1 & 5 & 1 & 1 \\ 3 & 2 & 9 & 1 & 2 \\ 0 & 1 & 1 & 2 & 4 \end{pmatrix};$$

$$4) \begin{pmatrix} 6 & -1 & 1 & 2 & 1 \\ -1 & 7 & -2 & 1 & 2 \\ 1 & -2 & 5 & -1 & 0 \\ 2 & 1 & -1 & 7 & -2 \\ 1 & 2 & 0 & -2 & 6 \end{pmatrix}.$$

Знайти найбільше за модулем власне число і відповідний йому власний вектор:

а) степеневим методом;

б) методом скалярних добутків.

(За початкове наближення взяти вектор  $(1; 1; \dots; 1)$ ).

4.2. Для матриць із вправи 4.1 знайти найменше за модулем власне число і відповідний йому власний вектор:

а) методом зворотних ітерацій;

б) методом зворотних ітерацій зі співвідношенням Релея.

4.3. Для матриць із вправи 4.1 розв'язати повну проблему власних значень методом поворотів Якобі.

Точність  $\varepsilon = 10^{-6}$  (у евклідовій нормі).

4.4. Знайти грубі наближення до власних чисел матриці

$$A = \begin{pmatrix} 5 & 2 & -3 \\ 4 & 5 & -4 \\ 6 & 4 & -4 \end{pmatrix}$$

степеневим методом, уточнити ці значення зворотними ітераціями зі зсувами.

4.5. А. Проаналізувати збіжність степеневого методу у випадку, якщо  $\lambda_1$  – найбільше кратно за модулем власне число  $n \times n$ -

матриці простої структури. Як можна знайти всі відповідні йому власні вектори залежно від показника кратності?

**Б.** Що можна сказати про поведінку послідовності відношень, якщо  $\lambda_1 = -\lambda_2$ , і  $|\lambda_2| > |\lambda_3| \geq \dots \geq |\lambda_n|$  ( $\lambda_i \in \mathbb{R}$ )?

**В.** Розглянути і пояснити поведінку степеневого методу у випадку, коли задана матриця  $A$  – діагональна.

4.6. Знайти всі власні пари матриці

$$A = \begin{pmatrix} 5 & 3 & -1 \\ 3 & 5 & 1 \\ -1 & 1 & 3 \end{pmatrix}$$

а) методом скалярних добутків;

б) *RQI*-алгоритмом, починаючи його з різних векторів.

4.7. Методом поворотів Якобі знайти власні пари матриці  $A$ , якщо:

$$\text{а) } A = \begin{pmatrix} 3 & -2 \\ -2 & 3 \end{pmatrix}; \quad \text{б) } A = \begin{pmatrix} 5 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{pmatrix}.$$

4.8. Порівняти два підходи для знаходження всіх власних чисел матриці

$$A = \begin{pmatrix} 4 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

методом поворотів Якобі:

а) застосовуючи його безпосередньо до заданої матриці;

б) попередньо звівши її до тридіагонального вигляду перетвореннями Хаусхолдера.

4.9. Для знаходження власних пар симетричних додатно визначених матриць побудувати *LU*-алгоритм на базі  $U^T U$  (або  $LL^T$ ) – розкладу Холецького. Випробувати його на матрицях:

$$A = \begin{pmatrix} 4 & -2 \\ -2 & 2 \end{pmatrix} ; B = \begin{pmatrix} 5 & 1 & 0 \\ 1 & 5 & -2 \\ 0 & -2 & 1 \end{pmatrix}.$$

Чи зберігають отримані на кожному кроці такого алгоритму подібні до  $B$  матриці тридіагональну структуру?

4.10. Задано матрицю  $A = \begin{pmatrix} 5 & -1 \\ -2 & 3 \end{pmatrix}$ . Зробити по три кроки:

- $LU$ -алгоритму;
- $QR$ -алгоритму на підставі перетворень Гівенса;
- $QR$ -алгоритму на підставі перетворень Хаусхолдера.

Порівняти отримані наближення до власних чисел матриці  $A$  за точністю (знайшовши спочатку її точні значення за допомогою характеристичного рівняння) і за обчислювальними затратами.

4.11. Матрицю  $A = \begin{pmatrix} 3 & 2 & -1 \\ 2 & 4 & 1 \\ -1 & 1 & 3 \end{pmatrix}$  звести до тридіагонального

вигляду:

- перетворенням Гівенса;
- методом Ланцоша.

Порівняти ці способи за обчислювальними затратами.

4.12. Для  $n \times n$ -матриці приблизно підрахувати кількість арифметичних операцій, потрібних для:

- одного кроку степеневого методу;
- одного кроку методу зворотних ітерацій (без зсувів);
- одного повного циклу методу поворотів Якобі;
- повного циклу зведення матриці до форми Хессенберга перетвореннями Хаусхолдера.

4.13. Знайти власні значення тридіагональних матриць

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 2 \\ 0 & 2 & 4 \end{pmatrix}; B = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 3 & 2 & 0 \\ 0 & 2 & 3 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix};$$

$$C = \begin{pmatrix} 6 & -1 & 0 & 0 & 0 \\ -1 & 7 & -2 & 0 & 0 \\ 0 & -2 & 5 & -1 & 0 \\ 0 & 0 & -1 & 7 & -2 \\ 0 & 0 & 0 & -2 & 6 \end{pmatrix},$$

використовуючи:

- а) метод бісекцій;
- б)  $QL$ -метод.

#### 4.14. Матриця Уілкінсона

$$A = \begin{pmatrix} 20 & 20 & 0 & 0 & 0 & 0 \\ 0 & 19 & 20 & 0 & 0 & 0 \\ 0 & 0 & 18 & 20 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 20 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

має найменше за модулем власне число, що дорівнює 1. Як воно зміниться внаслідок збурення першого елемента останнього рядка на  $\epsilon = 20^{-19} 20! \approx 5 \cdot 10^{-7}$ ?

4.15. Довести, що коли  $A = QR$ , то

$$\frac{1}{n} \operatorname{cond}_1 A \leq \operatorname{cond}_1 R \leq n \cdot \operatorname{cond}_2 A.$$

Довести також, що  $\operatorname{cond}_2 A = \operatorname{cond}_2 R$ . Тут використано позначення:  $\operatorname{cond}_1 B$  і  $\operatorname{cond}_2 B$  – числа зумовленості матриці  $B$  за нормами  $\|\cdot\|_1$  і  $\|\cdot\|_2$ , відповідно.

4.16. Знайти матрицю поворотів Якобі, яка б давала нуль у

позиції  $(3, 1)$  матриці  $\begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 5 & 9 \end{pmatrix}$ .

4.17. Довести, що для квадратних матриць  $A$  і  $B$  однакового розміру спектри матриць  $AB$  і  $BA$  збігаються.

4.18. Нехай  $A = A^T > 0$ . Довести, що

$$\lambda_{\max} A = \max R_A(x); \quad \lambda_{\min} A = \min R_A(x),$$

де  $R_A(x) = \frac{(Ax, x)}{(x, x)}$  – співвідношення Релея.

4.19. Нехай  $P = I - \mu w w^T$ , де  $w$  – вектор такий, що  $w^T w = 1$

А. Довести, що матриця  $P$  буде ортогональною тоді й тільки тоді, коли  $\mu = 2$ .

Б. Довести, що коли  $Q$  – матриця Хессенберга, а  $R$  – верхня трикутна, то добуток  $RQ$  є матрицею Хессенберга.

4.20. Довести, що коли  $A$  і  $B$  – дійсні ортогональні матриці розмірності  $n \times n$ , то матриця  $AB$  також є ортогональною. Після цього довести, що коли кожен з векторів  $w_i$  задовольняє умову  $w_i^T w_i = 1$ , то матриця

$$Q = (I - 2w_1^T w_1)(I - 2w_2^T w_2) \dots (I - 2w_{n-1}^T w_{n-1})(I - 2w_n^T w_n)$$

ортогональна.

## Розділ 5

### МЕТОДИ РОЗВ'ЯЗУВАННЯ ОДНОГО НЕЛІНІЙНОГО РІВНЯННЯ З ОДНІЄЮ ЗМІННОЮ

Нехай задано рівняння

$$f(x) = 0,$$

де  $f(x)$  – задана функція дійсного або комплексного аргументу, визначена і неперервна на проміжку  $[a, b]$ . Зокрема,  $f(x)$  може бути алгебричним багаточленом.

Універсальні обчислювальні алгоритми ґрунтуються на тому, що відоме наближене значення  $x_0$  одного з коренів, а потім його поліпшують до заданої точності.

Обчислювальні алгоритми наближеного знаходження коренів нелінійного рівняння складаються з таких етапів:

- I) відшукування достатньо малих відрізків (інтервалів), що належать області визначення рівняння, у кожному з яких є один і тільки один корінь рівняння  $f(x) = 0$ . Це етап *відокремлення коренів рівняння, або визначення відрізків ізоляції коренів*;
- II) обчислення кореня з заданою точністю, якщо відоме деяке початкове його наближення в інтервалі, який не містить інших коренів. Це *етап уточнення наближених коренів*.

Перший етап складніший від другого, бо в загальному випадку не існує достатньо ефективних методів відокремлення коренів.

#### 5.1. Методи відокремлення коренів

Для відокремлення коренів рівняння

$$f(x) = 0 \tag{5.1}$$

використовують відому теорему з математичного аналізу.

**Теорема 5.1.** Якщо неперервна функція  $f(x)$  набуває значення різних знаків на кінцях відрізка  $[a, b]$ , тобто  $f(a) \cdot f(b) < 0$ , то всередині цього відрізка міститься принаймні один корінь рівняння  $f(x) = 0$ , тобто знайдеться хоча б одне число  $\xi \in (a, b)$  таке, що  $f(\xi) = 0$

Корінь  $\xi$  явно буде єдиний, якщо похідна  $f'(x)$  існує і зберігає постійний знак усередині проміжку, тобто  $f'(x) > 0$  (або  $f'(x) < 0$ ) при  $a < x < b$

Процес відокремлення коренів починається з визначення знаків функції  $f(x)$  у граничних точках  $x = a$  і  $x = b$  в області її існування. Далі необхідно тим чи іншим способом упевнитися, що корінь на цьому проміжку єдиний. Для відокремлення коренів часто буває достатньо використати процес половинного ділення, наближено ділячи проміжок  $[a, b]$  на два, чотири, вісім і так далі однакових частин, і шукати зміни знаків функції. Корисно пам'ятати, що алгебричний багаточлен  $n$ -го степеня має не більше  $n$  дійсних коренів. Тому якщо для такого рівняння ми отримали  $n+1$  зміну знаків, то корені всі його відділені.

### 5.1.1. Табуляція функції

Для того, щоб знайти проміжок існування коренів рівняння (5.1), скористаємося теоремою 5.1. Побудуємо таблицю значень функції  $y = f(x)$  на достатньо великому проміжку  $[a, b]$  області визначення з великим кроком  $h_1$ . З'ясуємо, на якому підпроміжку  $[a_1, b_1]$  функція змінює знак. Знову табулюємо її на цьому проміжку з довільним кроком  $h_2$ . На практиці достатньо мати проміжок у межах одиниць, наприклад,  $[-3, 2]$

На підставі теореми 5.1 можна записати такий алгоритм знаходження проміжку  $[a, b]$ .

1. Знайти область визначення функції  $f(x)$  рівняння (5.1).
2. Обчислити похідну  $f'(x)$  і знайти точки, в яких похідна перетворюється в нуль.
3. Записати інтервали монотонності.
4. Дослідити знак функції на кінцях інтервалів монотонності.
5. Визначити проміжки ізоляції коренів.
6. Звузити ці проміжки, протабулювавши функцію.



**Приклад 5.1.** Відокремити корені рівняння

$$f(x) = x^3 - 4x^2 + 2 = 0.$$

*Розв'язування:*

1. О.В:  $(-\infty, \infty)$

2.  $f'(x) = 3x^2 - 8x$ . Знайдемо розв'язки рівняння  $3x^2 - 8x = 0$ ;

тоді  $x_1 = 0$ ;  $x_2 = \frac{8}{3}$ .

3. Інтервали монотонності:  $(-\infty, 0)$ ,  $\left(0, \frac{8}{3}\right)$ ,  $\left(\frac{8}{3}, \infty\right)$ .

4. Дослідження знаків функції на кінцях інтервалів монотонності:

$$\lim_{x \rightarrow -\infty} f(x) = -\infty; f(0) = 2 > 0; f\left(\frac{8}{3}\right) = -\frac{202}{27} < 0;$$

$$\lim_{x \rightarrow \infty} f(x) = +\infty.$$

5. Проміжки ізоляції коренів:  $(-\infty, 0)$ ,  $\left(0, \frac{8}{3}\right)$ ,  $\left(\frac{8}{3}, \infty\right)$ .

6. Протабулюємо функцію  $y = f(x)$  на кожному з проміжків:

а)  $(-1000; 0)$  при  $h = 100$ ;  $f(-100) < 0$ ;  $f(0) > 0$ ;

$(-100; 0)$  при  $h = 10$ ;  $f(-10) < 0$ ;  $f(0) > 0$ ;

$(-10; 0)$  при  $h = 1$ ;  $f(-1) < 0$ ;  $f(0) > 0$ ;

$(-1; 0)$  – це вже проміжок одиничної довжини;

б)  $\left(0, \frac{8}{3}\right)$  при  $h = 1$ ;  $f(0) > 0$ ;  $f(1) < 0$ ;

$(0, 1)$  – звужений другий проміжок;

в)  $\left(\frac{8}{3}, 1000\right)$  при  $h_1 = 100$ ;  $f\left(\frac{8}{3}\right) < 0$ ;  $f(100) > 0$

$\left(\frac{8}{3}, 100\right)$  при  $h_2 = 10$ ;  $f\left(\frac{8}{3}\right) < 0$ ;  $f(10) > 0$ ;

$\left(\frac{8}{3}, 10\right)$  при  $h = 1$ ;  $f(3) < 0$ ;  $f(4) > 0$ ;

$(3, 4)$  – проміжок, де є останній корінь.

Отже, корені є на проміжках  $x_1 \in (-1; 0)$ ;  $x_2 \in (0; 1)$ ;  $x_3 \in (3; 4)$

### 5.1.2. Графічний спосіб

Якщо функція  $f(x)$  нескладного вигляду, то знаходимо точки перетину графіка функції з віссю абсцис. Якщо це важко зробити, то рівняння (5.1) записуємо у вигляді

$$\varphi(x) = \psi(x),$$

де  $\varphi(x)$ ,  $\psi(x)$  – достатньо прості функції, графіки яких будуть просто. Точки перетину графіків  $y = \varphi(x)$  та  $y = \psi(x)$  дадуть наближене значення розв'язку рівняння (5.1).

**Приклад 5.2.** Відокремити корені рівняння

$$e^{-x} + x^2 - 2 = 0.$$

*Розв'язування.* Перетворимо це рівняння до вигляду  $e^{-x} = -x^2 + 2$ . Тоді  $\varphi(x) = e^{-x}$ ;  $\psi(x) = -x^2 + 2$ . Графіки цих функцій будувати нескладно (рис. 5.1).

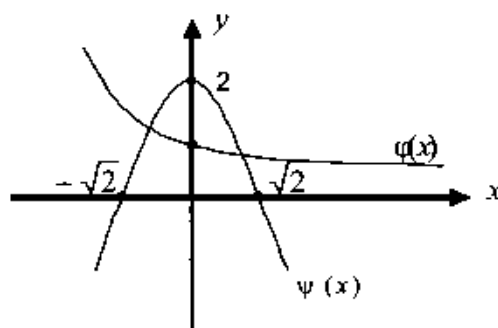


Рис. 5.1.

Із графіка бачимо, що рівняння має два корені, розміщені в проміжках  $(-1,5 ; 0)$  та  $(1 ; 1,5)$

Графічно функцію  $y = f(x)$  можна побудувати на комп'ютері й наближено знайти проміжки, де містяться корені рівняння. Можна скласти програму, яка згущує сітку по  $x$  для побудови графіка, що поєднує в собі графічний спосіб і табуляцію.

### 5.1.3. Аналітичний метод відділення коренів

Якщо  $f(x)$  – багаточлен, то для локалізації коренів можна використати відомі з алгебри теореми про кількість коренів, про межі додатних і від’ємних коренів.

Для алгебричного багаточлена  $f(z) = A_n z^n + A_{n-1} z^{n-1} + \dots + A_0$  відома теорема про кільце, в якому є корені рівняння  $f(z) = 0$

**Теорема 5.2.** Корені багаточлена  $f(z)$  з дійсними чи комплексними коефіцієнтами містяться у кільці

$$\frac{A_0}{|A_0| + B} \leq |z| \leq 1 + \frac{A}{|A_n|}, \quad (5.2)$$

де  $A_0 \cdot A_n \neq 0$ ;  $A = \max\{|A_0|, |A_1|, \dots, |A_{n-1}|\}$ ,

$$A_n > 0, \quad B = \max\{|A_1|, |A_2|, \dots, |A_n|\}.$$

**П р и к л а д 5.3.** Знайти кільце, у якому містяться корені рівняння

$$z^4 + 3z^3 + 8z^2 - 5 = 0$$

*Розв’язування.* Згідно з позначеннями теореми,

$$A_0 = -5; \quad A = \max\{|-5|, 8, 3\} = 8; \quad B = \max\{3, 8, 0, 1\} = 8.$$

$$\text{Тоді } \frac{5}{5+8} \leq |z| \leq 1 + \frac{8}{1}; \quad \text{отже, } \frac{5}{13} < |z| < 9.$$

Нехай задано багаточлен з дійсними коефіцієнтами

$$f(x) = \sum_{i=0}^n A_i x^i = 0, \quad (5.3)$$

де  $A_0 \cdot A_n \neq 0$ .

**Теорема 5.3.** Усі дійсні корені рівняння (5.3) містяться у проміжку

$$x \leq 1 + \sqrt[k]{\frac{B}{A_n}}, \quad (5.4)$$

де  $B = \max_{i=0, \dots, k-1} |A_i|$ ;  $A_{n-k}$  ( $k \geq 1$ ) – перший з від’ємних коефіцієнтів рівняння, рахуючи зліва;  $A_n > 0$ .

Формулу (5.4) називають формулою Маклорена для відшукування верхньої межі коренів.

**П р и к л а д 5.4.** Знайти верхню і нижню межі додатних і від'ємних коренів рівняння  $x^3 - 8x + 1 = 0$ .

*Розв'язування.* Скористаємося формулою (5.4). Знайдемо верхню межу додатних коренів рівняння  $\alpha \leq 1 + \sqrt[k]{\frac{B}{A_n}}$ . У нашому прикладі  $A_n = 1$ ;  $B = 8$ ;  $k = 2$ . Тоді  $\alpha = 1 + \sqrt{8} = 1 + 2\sqrt{2}$

Відшукаємо нижню межу додатних коренів рівняння. Для цього зробимо заміну  $z = \frac{1}{x}$ . Тоді рівняння набуде вигляду  $z^3 - 8z^2 + 1 = 0$ . Знову, згідно з формулою (5.4), визначимо  $A_n = 1$ ;  $B = 8$ ;  $k = 1$ ,  $b^* = 1 + 8 = 9$  – верхня межа додатних коренів перетвореного рівняння. Отже, нижня межа додатних коренів початкового рівняння  $-\frac{1}{9}$ .

Перейдемо до знаходження від'ємних коренів. Зробимо заміну змінних  $u = -x$ . Одержимо рівняння  $u^3 - 8u - 1 = 0$ . Згідно з теоремою 5.3, маємо  $c^* = 1 + \sqrt{8} = 1 + 2\sqrt{2}$  – верхня межа додатних коренів перетвореного рівняння;  $c = -1 - 2\sqrt{2}$  – нижня межа від'ємних коренів.

Заміною  $v = -\frac{1}{x}$  знайдемо верхню межу від'ємних коренів рівняння:

$$v^3 + 8v^2 - 1 = 0; k = 3; B = 1; A_n = 1; d^* = 1 + \sqrt[3]{1} = 2; d = -\frac{1}{2}$$

Отже, додатні корені є в інтервалі  $\left(\frac{1}{9}; 1 + 2\sqrt{2}\right)$ ; від'ємні – у  $\left(-1 - 2\sqrt{2}; -\frac{1}{2}\right)$ .

## 5.2. Ітераційні методи розв'язування алгебричних і трансцендентних рівнянь

### 5.2.1. Метод дихотомії (половинного ділення)

Нехай на проміжку  $[a, b]$  є корінь рівняння (5.1), тобто виконується умова  $f(a) \cdot f(b) < 0$ . Звужуватимемо цей проміжок так, щоб його довжина стала не більшою від наперед заданої точності обчислення кореня  $\varepsilon$ , тобто

$$|b - a| \leq \varepsilon. \quad (5.5)$$

Відрізок  $[a, b]$  ділимо навпіл  $c = \frac{a+b}{2}$ . Якщо  $f(c) = 0$ , то  $c = \frac{a+b}{2}$  є коренем рівняння (5.1). Якщо  $f(c) \neq 0$ , то вибираємо той з проміжків  $[a, c]$  або  $[c, b]$ , на кінцях якого функція  $f(x)$  має протилежні знаки. Звужений проміжок знову ділимо навпіл і виконуємо ті ж дослідження на знаки функції, як і на попередньому кроці. Отримаємо послідовність вкладених один в одного відрізків  $[a_1, b_1], [a_2, b_2], \dots$ , для яких  $f(a_n) \cdot f(b_n) < 0$  ( $n = 1, 2, \dots$ ).

Умовою закінчення обчислень слугує умова (5.5) або  $f(c) = 0$ .

**Наведемо алгоритм методу дихотомії.**

1. Ввести межі коренів  $a, b$  і точність  $\varepsilon$ .
2. Знайти середину проміжку  $c = (a + b) / 2$ .
3. Якщо  $f(c) = 0$ , то вивести  $c$  і закінчити алгоритм. Якщо ж  $f(c) \neq 0$ , то перейти до пункту 4.
4. Перевірити умову:  $f(a) \cdot f(c) < 0$ . Якщо умова виконується, то  $b = c$ , інакше  $a = c$ .
5. Якщо  $|b - a| \leq \varepsilon$ , то закінчити обчислення. Наближеним значенням кореня буде  $c$ . Якщо проміжок ще великий, то перейти до пункту 2.

**Приклад 5.5.** Уточнити корені рівняння  $x^3 - 4x^2 + 2 = 0$  методом дихотомії.

**Розв'язування.** За наведеним алгоритмом складена програма мовою Pascal. З точністю до  $\varepsilon$  на проміжках  $[a, b]$  знайдено корені рівняння  $x^3 - 4x^2 + 2 = 0$  за  $n$  кроків, які наведено в табл. 5.1.

Таблиця 5.1

**Результати обчислення кореня для різних проміжків  
та різної точності**

$a$	$b$	$\varepsilon$	$x_n$	$n$
3	4	0,01	3,86718	7
3	4	0,0001	3,86615	14
3	4	0,000001	3,86620	20
0	1	0,01	0,789063	7
0	1	0,0001	0,789246	14
0	1	0,000001	0,789245	20
-1	0	0,01	-0,656250	5
-1	0	0,0001	-0,655457	14
-1	0	0,000001	-0,655442	20

**П р и к л а д 5.6.** З точністю  $\varepsilon$  знайти корінь нелінійного рівняння  $x^3 - 2 \sin x - 1 = 0$ .

*Розв'язування.* Побудуємо графіки функцій  $y = x^3$  та  $y = 2 \sin x + 1$  і локалізуємо корінь рівняння на проміжку  $[1, 2]$ .

У табл. 5.2 наведено послідовні наближення до кореня цього рівняння, отримані за методом дихотомії.

Таблиця 5.2

**Результати обчислень кореня нелінійного рівняння  
за методом дихотомії**

$\varepsilon$	$n$	$x_n$
0,000001	1	1,50000000
	2	1,25000000
	3	1,37500000
	4	1,43750000
	5	1,46875000
	6	1,45312500
	7	1,44531250
	8	1,44140625
	9	1,43945313
	10	1,44042969
	11	1,43994141
	12	1,43969727
	13	1,43957520
	14	1,43951416

## 5.2.2. Метод ітерацій (послідовних наближень).

### Принцип стискувальних відображень

Для розв'язування алгебричних і трансцендентних рівнянь (5.1) розроблено багато різних методів. Суть їхня така. Нехай відома достатньо мала область, у якій є єдиний корінь  $x = \alpha$  рівняння  $f(x) = 0$ . Вибирають точку  $x_0$  – початкове наближення, і за допомогою деякого рекурентного співвідношення

$$x_k = \varphi_k(x_0, x_1, \dots, x_{k-1}) \quad (5.6)$$

відшукують послідовність точок  $x_1, x_2, \dots, x_n, \dots$ , яка збігається до  $x = \alpha$ . Збіжність послідовності точок  $\{x_n\}$  забезпечена відповідним вибором функцій  $\varphi_k$  і початкового наближення  $x_0$ . Найчастіше на основі функції  $f(x)$  будують функцію  $\varphi(x)$  таку, що шуканий корінь  $x = \alpha$  рівняння (5.1) є і коренем рівняння

$$x = \varphi(x). \quad (5.7)$$

Після того будують послідовність точок  $\{x_n\}$  за допомогою співвідношення

$$x_n = \varphi(x_{n-1}), \quad (n = 1, 2, \dots) \quad (5.8)$$

для деякого початкового наближення  $x_0$ .

Методи типу (5.6) називають *нестационарними ітераційними процесами*, а (5.8) – *стационарними*.

Процес послідовного отримання чисел  $x_1, x_2, \dots, x_n$  за формулою (5.8) називають методом ітерацій, або послідовних наближень. Він є однокроковим. У разі виконання певних умов, накладених на функцію  $\varphi(x)$ , послідовність (5.8) збігається до кореня  $\alpha$ :

$$\lim_{n \rightarrow \infty} x_n = \alpha.$$

Умови збіжності ітераційного процесу впливають з принципу стискувальних відображень.

**Теорема 5.4** (ознака збіжності методу ітерацій). *Якщо в околі  $R = \{x : |x - \alpha| \leq r\}$  функція  $\varphi(x)$  задовольняє умову Ліпшиця з константою  $k < 1$*

$$|\varphi(x) - \varphi(y)| < k|x - y|, \quad x \in R; y \in R, \quad (5.9)$$

то послідовність (5.8) збігається до кореня  $x = \alpha$  за будь-якого  $x_0 \in R$ , причому швидкість збіжності характеризує нерівність

$$|x_n - \alpha| \leq k^n |x_0 - \alpha| \quad (5.10)$$

**Теорема 5.5** (ознака існування кореня рівняння). Якщо в околі  $R = \{x : |x - \alpha| \leq r\}$  функція  $\varphi(x)$  задовольняє умову Ліпшиця (5.9) з константою  $k < 1$  і в точці  $x_0$  справджується нерівність

$$|x_0 - \varphi(x_0)| \leq (1 - k)r, \quad (5.11)$$

то в  $R$  рівняння (5.7) має єдиний корінь  $x = \alpha$ , який можна отримати як границю послідовності

$$x_n = \varphi(x_{n-1}), \quad n = 1, 2, \dots,$$

де  $x_0$  – довільна точка з околу  $R$

Наведемо умову збіжності ітераційного процесу, яка впливає з теореми 5.4.

**Теорема 5.6.** Нехай на проміжку  $[a, b]$  є один простий корінь рівняння  $x = \varphi(x)$  і в усіх точках цього проміжку  $|\varphi'(x)| \leq k < 1$ . Якщо в цьому разі виконується умова  $a \leq \varphi(x) \leq b$ , то ітераційний процес (5.8) збіжний, а за початкове наближення  $x_0$  можна взяти довільне число з проміжку  $[a, b]$ .

Чим менше значення  $|\varphi'(x)|$ , тим швидше збігається ітераційний процес. Із теореми випливає, що умова  $|\varphi'(x)| < 1$  є достатньою умовою збіжності ітераційного процесу (5.8).

Оцінка методу простої ітерації така

$$|x_n - \alpha| \leq \frac{k^n}{1 - k} |\varphi(x_0) - x_0|; \quad 0 < k < 1.$$

За критерій кількості ітерацій можна взяти умову

$$|x_n - x_{n-1}| \leq \frac{\varepsilon(1 - k)}{k}; \quad 0 < k < 1. \quad (5.12)$$

Як бачимо з теорем, рівняння  $f(x) = 0$  зводять до вигляду  $x = \varphi(x)$  так, щоб  $|\varphi'(x)| < 1$ ,  $x \in [a, b]$ .



Розглянемо геометричну інтерпретацію методу ітерацій. На рис. 5.2 зображено випадок, коли  $0 < \varphi'(x) \leq k < 1$ .

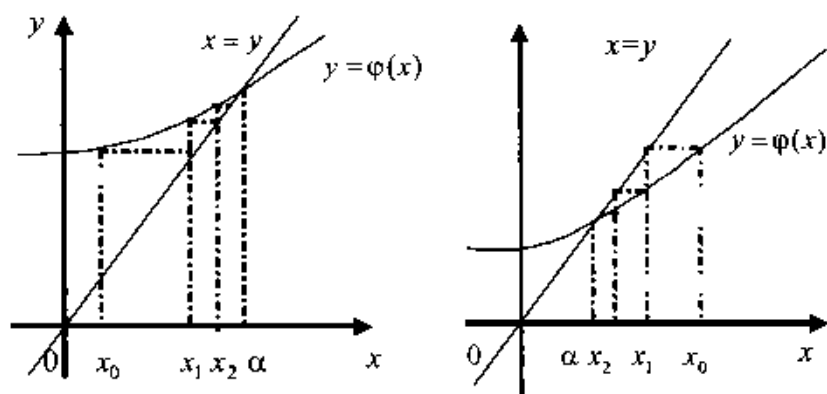


Рис. 5.2.

Послідовність  $x_1, x_2, x_3, \dots$  монотонно збігається до кореня з того боку, з якого розташовано початкове наближення  $x_0$ .

Розглянемо випадок  $-1 < \varphi'(x) < 0$  (рис. 5.3).

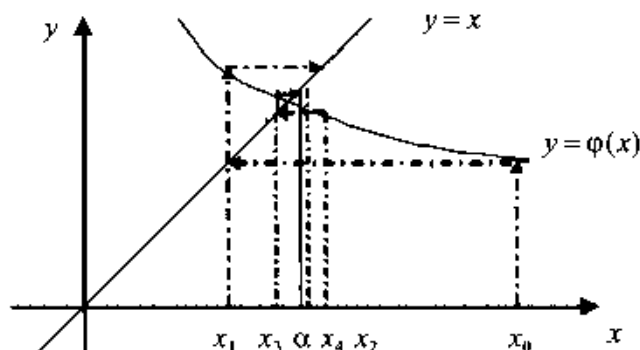


Рис. 5.3.

Послідовні наближення розташовані по чергово з різних боків від кореня  $x = \alpha$ .

У випадку  $|\varphi'(x)| > 1$  ітераційний процес розбіжний (рис. 5.4).

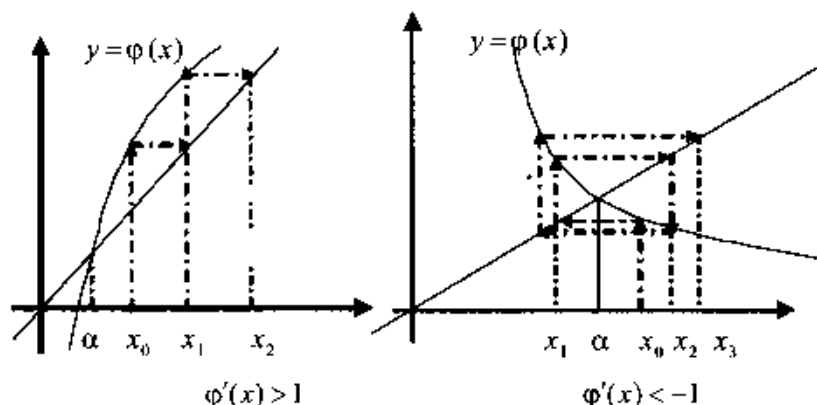


Рис. 5.4.

Наведемо один достатньо загальний спосіб зведення рівняння  $f(x) = 0$  до вигляду  $x = \varphi(x)$  в разі виконання умови  $|\varphi'(x)| \leq k < 1$ . Нехай шуканий корінь  $\alpha \in [a, b]$  і

$$0 < m_1 \leq f'(x) \leq M_1, \quad x \in [a, b],$$

$$\text{де } m_1 = \min_{[a, b]} f'(x);$$

$$M_1 = \max_{[a, b]} f'(x)$$

Замінімо рівняння (5.1) еквівалентним йому рівнянням  $x = x - \lambda f(x)$ . Підберемо  $\lambda$  так, щоб виконувалася умова  $0 \leq \varphi'(x) = 1 - \lambda f'(x) \leq k < 1$ . Звідси  $0 \leq 1 - \lambda M_1 \leq 1 - \lambda m_1 \leq k$ . Отже,

$$\text{можна вибрати } \lambda = \frac{1}{M_1}, \quad k = 1 - \frac{m_1}{M_1} < 1.$$

Якщо ж обчислення похідної  $f'(x)$  та її оцінка є громіздкими, то дослідження збіжності процесу ітерацій можна доручити обчислювальній машині, обмеживши кількість ітерацій деяким числом, наприклад,  $n \leq 100$ . Якщо після великої кількості ітерацій точність не досягнута, то в обчисленнях треба поміняти функцію  $\varphi(x)$ .

**П р и к л а д 5.7.** Методом ітерацій з точністю до  $\varepsilon$  знайти корені рівняння  $x^3 - 4x^2 + 2 = 0$

*Розв'язування.* Запишемо нашу функцію у вигляді (5.7), тобто знайдемо  $\varphi(x)$ .

Рівняння (5.1) домножимо на константу  $M > 0$  і додамо до обох частин  $x$ . Одержимо

$$x = x + M f(x)$$

Отже,  $\varphi(x) = x + M f(x)$ .

Знайдемо константу  $M$  з умови, що  $|\varphi'(x)| < 1$ :

$$\varphi'(x) = 1 + M f'(x);$$

$$-1 < 1 + M f'(x) < 1; \quad -2 < M f'(x) < 0;$$

$$0 > f'(x) > -\frac{2}{M} \quad (\text{якщо } f'(x) > 0 \text{ на } [a, b], \text{ то виберемо } M < 0).$$

Отже, умова вибору  $M$  така:

$$f'(x) > -\frac{2}{M}; \quad (f'(x) < 0; \quad M > 0);$$

$$f'(x) < -\frac{2}{M}; \quad (f'(x) > 0; \quad M < 0).$$

Для нашого рівняння  $f'(x) = 3x^2 - 8x$ .

Для проміжку  $[3, 4]$ ,  $f'(x) > 0$ , функція монотонно зростає,  $f'(4) = 16$ .

$$\text{Отже, } 16 < -\frac{2}{M}; \quad M \in \left(-\frac{1}{8}; 0\right).$$

При  $M = -0,1$  одержимо правильну нерівність  $16 < 20$

Для проміжку  $[3, 4]$  виберемо  $\varphi_1(x) = x - 0,1(x^3 - 4x^2 + 2)$ ,

$$|\varphi'_1(x)| \leq 0,7, \quad \text{критерій кількості ітерацій } |x_n - x_{n-1}| < \frac{3}{7} \varepsilon$$

Для проміжку  $[0, 1]$  маємо  $f'(x) < 0$ , тому функція монотонно спадає.

$$\text{Отже, } M > 0; \quad f'(x) > -\frac{2}{M};$$

$$f'(1) = -5; \quad -5 > -\frac{2}{M}; \quad 5 < \frac{2}{M} \Rightarrow M < \frac{2}{5} = 0,4.$$

При  $M = 0,1$  одержимо  $\varphi_2(x) = x + 0,1(x^3 - 4x^2 + 2)$ .

Оскільки  $k = 0,99$ , то  $|x_n - x_{n-1}| < \varepsilon \cdot 0,01$ .

Для проміжку  $[-1, 0]$  маємо  $f'(x) > 0$ , тому функція монотонно зростає. Отже,  $M < 0$  і  $f'(x) < -\frac{2}{M}$ ,

$$f'(-1) = 11; \quad 11 < -\frac{2}{M}; \quad M \in \left(-\frac{2}{11}; 0\right).$$

За  $M$  можна вибрати  $-0,1$ , тоді

$$\varphi_3(x) = x - 0,1(x^3 - 4x^2 + 2).$$

Звідси  $\varphi_1 = \varphi_3$ . Критерій завершення обчислень  $|x_n - x_{n-1}| < 0,01\varepsilon$ , де  $\varepsilon \approx 0,9$ . Дані обчислень наведені в табл. 5.3.

**П р и к л а д 5.8.** Методом ітерацій з точністю до  $\varepsilon$  знайти корінь рівняння  $x^3 - 2 \sin x - 1 = 0$ .

*Розв'язування.* Перетворимо наше нелінійне рівняння до вигляду (5.7).

А. Додамо  $x$  до обох частин рівняння. Одержимо  $x = x + x^3 - 2 \sin x - 1 = 0$ . Отже,  $\varphi(x) = x + x^3 - 2 \sin x - 1$ .

Таблиця 5.3

Результати ітераційного процесу за формулою (5.8)

$a$	$b$	$\varepsilon$	$\varphi(x)$	$x_0$	$x_n$	$n$
1	2	3	4	5	6	7
3	4	0,01	$\varphi_1(x)$	3	3,911250	1
					3,844642	2
					3,873633	3
					3,863247	4
					3,867346	5
				3,5	3,700000	1
					3,910700	2
					3,847272	3
					3,873333	4
					3,863368	5
					3,867300	6
				4	3,800000	1
					3,888800	2
					3,856965	3
					3,869745	4
					3,864800	5
					3,866744	6

Продовження табл. 5.3

1	2	3	4	5	6	7
3	4	0,0001	$\varphi_1(x)$	3 3,5 4	3,866188 3,866188 3,866193	11 10 11
3	4	0,000001	$\varphi_1(x)$	3 3,5 4	3,866199 3,866198 3,866198	15 16 15
0	1	0,01	$\varphi_2(x)$	0 0,5 1	0,784113 0,785972 0,792287	10 8 7
0	1	0,0001	$\varphi_2(x)$	0 0,5 1	0,789198 0,789191 0,789294	18 15 14
0	1	0,000001	$\varphi_2(x)$	0 0,5 1	0,789244 0,789244 0,789245	26 23 22
-1	0	0,01	$\varphi_1(x)$	-1 -0,5 0	-0,657101 -0,654272 -0,654463	4 5 8
-1	0	0,0001	$\varphi_1(x)$	-1 -0,5 0	-0,655451 -0,655425 -0,655428	9 9 12
-1	0	0,000001	$\varphi_1(x)$	-1 -0,5 0	-0,655443 -0,655442 -0,655442	13 14 16

Перевіримо, чи  $\varphi(x)$  задовольняє умову Ліпшиця. Для цього знайдемо похідну  $\varphi'(x) = 1 + 3x^2 - 2\cos x$ . Тоді  $\varphi'(1) = 1 + 3 - 2\cos 1 \approx 3$ . Оскільки  $|\varphi'(1)| > 1$ , то таке перетворення нас не задовольняє.

Б. У заданому рівнянні перенесемо  $x^3$  у правий бік і знайдемо  $x = \sqrt[3]{2\sin x + 1}$

$$\varphi(x) = \sqrt[3]{2\sin x + 1}; \quad \varphi'(x) = \frac{2}{3} \frac{\cos x}{\sqrt[3]{(2\sin x + 1)^2}};$$

$$\varphi'(1) \approx \frac{2}{3} \frac{1/2}{\sqrt[3]{(\sqrt{3} + 1)^2}} \approx 0,25 < 1, \quad \varphi'(2) \approx \frac{2}{3} \frac{-0,4}{\sqrt[3]{7,99}} \approx -0,137.$$

Отже, для вибраної функції виконується умова  $|\varphi'(x)| < 1$  на проміжку  $[1, 2]$ . Результати уточнення кореня рівняння  $x = \sqrt[3]{2\sin x + 1}$  методом ітерацій наведено в табл. 5.4.

Таблиця 5.4

**Уточнення коренів трансцендентного рівняння  
методом ітерацій**

$\varepsilon$	$x_0$	$x_n$	$n$
0,01	1	1,439379	3
	1,5	1,439567	2
	2	1,439432	3
0,0001	1	1,439485	5
	0,5	1,439485	4
	2	1,439485	5
0,000001	1	1,4394851	6
	0,5	1,4394851	5
	2	1,4394851	6

### 5.2.3. Метод хорд

Цей метод ще називають методом лінійної інтерполяції, методом пропорційних частин, методом січних.

Якщо рівняння  $f(x) = 0$  має корінь  $x = \alpha$ , а функція  $\psi(x)$  неперервна в околі  $x = \alpha$ , то рівняння

$$x = \varphi(x), \quad (5.13)$$

де  $\varphi(x) = x - \psi(x)f(x)$ , також має корінь  $x = \alpha$ . Функцію  $\psi(x)$  можна підібрати так, що ітераційний процес для рівняння  $x = \varphi(x)$  буде збіжним.

Ідея методу хорд полягає в тому, що на достатньо малому проміжку  $[a, b]$  дугу кривої  $y = f(x)$  замінює хорда. За наближене значення кореня приймають точку перетину хорди з віссю  $OX$ .

Якщо рівняння (5.1) має дійсний корінь  $\alpha \in [a, b]$ , а функція  $\psi(x)$  неперервна на  $[a, b]$  або хоча б в околі точки  $\alpha$ , то рівняння  $x = \varphi(x)$  має також корінь  $\alpha$ . Функцію  $\psi(x)$  підбираємо з умови збіжності ітераційного процесу.

У методі хорд  $\psi(x) = \frac{x - x_0}{f(x) - f(x_0)}$ , де  $x_0$  – точка з околу  $\alpha$ .

Тоді ітераційний процес матиме вигляд

$$x_n \approx x_{n-1} - \frac{(x_{n-1} - x_0)f(x_{n-1})}{f(x_{n-1}) - f(x_0)}. \quad (5.14)$$

Формула (5.14) відображає метод хорд.

Припустимо, що в деякому околі  $S$  точки  $x = \alpha$  функція  $f(x)$  є неперервна і в цьому околі має неперервні похідні  $f'(x)$  і  $f''(x)$ . Крім того,  $f'(x)$  і  $f''(x)$  в околі  $S$  не змінюють знака і не набувають значення нуль. Це означає, що в разі переходу через точку  $x = \alpha$  функція  $f(x)$  змінює знак і має точку  $x = \alpha$  простим коренем.

Для вибору початкового наближення в методі хорд користуються таким правилом: нерухомим кінцем проміжку є той, для якого знак функції збігається зі знаком другої похідної

$$f(x_0)f''(x_0) > 0. \quad (5.15)$$

Із рис. 5.5 бачимо, що метод хорд дає наближення кореня лише з одного боку (у нашому випадку  $x_0 = b$  – нерухома точка в формулі (5.14)). Ідемо до кореня з точки  $a$ . Формула (5.14) набуде вигляду

$$x_n = x_{n-1} - \frac{(x_{n-1} - a)f(x_{n-1})}{f(x_{n-1}) - f(a)}.$$

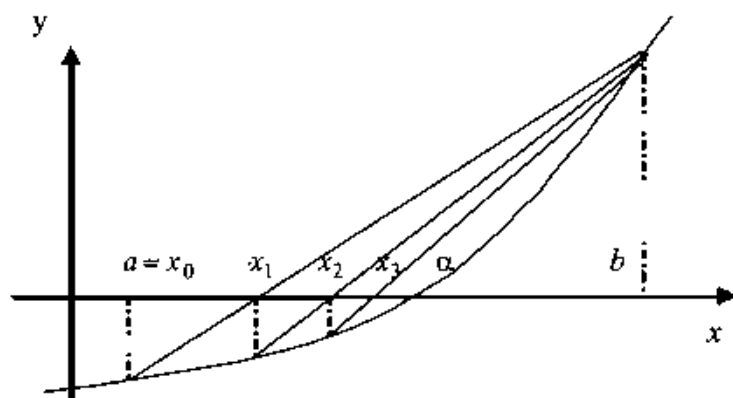


Рис. 5.5.

Якщо ж функція опукла вгору, то нерухомим кінцем буде точка  $a$  (рис. 5.6).

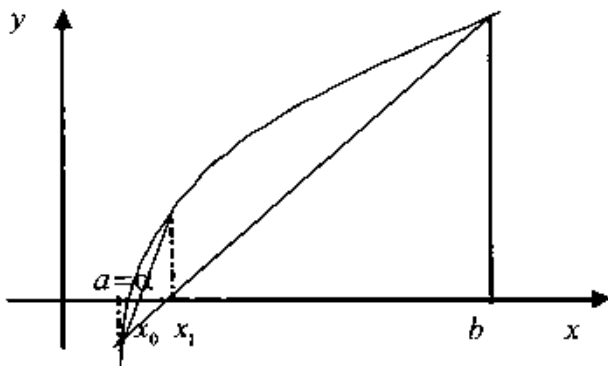


Рис. 5.6.

Тоді ітераційний процес матиме вигляд

$$x_n = x_{n-1} - \frac{(x_{n-1} - a)f(x_{n-1})}{f(x_{n-1}) - f(a)}.$$

Метод хорд є ітераційним методом першого порядку. Для достатньо малих проміжків можна використати оцінку похибки

$$|x_n - a| \leq \frac{|f(x_n)|}{\inf_{x \in S} |f'(x)|}.$$

**П р и к л а д 5.9.** Методом хорд з точністю до  $\varepsilon$  знайти корені рівняння  $x^3 - 4x^2 + 2 = 0$ .

*Розв'язування.* Знайдемо похідні

$$f'(x) = 3x^2 - 8x,$$

$$f''(x) = 6x - 8.$$

Розглянемо ітераційний процес на кожному з проміжків локалізації коренів.

$$\text{а) } [3, 4]: f(3) = -7 < 0; \quad f(4) = 2 > 0;$$

$$f''(3) > 0; \quad f''(4) > 0.$$

Нерухомим кінцем буде точка  $x_0 = 4$ .

Ітераційна формула матиме вигляд



$$x_n = x_{n-1} - \frac{(x_{n-1} - 4)f(x_{n-1})}{f(x_{n-1}) - 2}$$

б)  $[0, 1]: f(0) = 2 > 0; f(1) = -1 < 0;$

$f''(0) = -8 < 0; f''(1) = -2 < 0.$

Отже, нерухомим кінцем буде точка  $x_0 = 1.$

Розрахунки виконаємо за формулою

$$x_n = x_{n-1} - \frac{(x_{n-1} - 1)f(x_{n-1})}{f(x_{n-1}) + 1}.$$

в)  $[-1, 0]: f(-1) = -3 < 0; f(0) = 2 > 0;$

$f''(-1) = -14 < 0; f''(0) = -8 < 0.$

Нерухомою буде точка  $x_0 = -1.$

Формула (5.14) набуде вигляду

$$x_n = x_{n-1} - \frac{(x_{n-1} + 1)f(x_{n-1})}{f(x_{n-1}) + 3}.$$

Результати обчислень наведені в табл. 5.5.

Таблиця 5.5

Знаходження коренів алгебричного рівняння за методом хорд

	$a$	$b$	$\varepsilon$	Нерухома точка	$x_n$	$n$
а)	3	4	0,01	4	3,866168	3
	3	4	0,0001	4	3,866196	5
	3	4	0,000001	4	3,866198	7
б)	0	1	0,01	1	0,788686	3
	0	1	0,0001	1	0,783242	5
	0	1	0,000001	1	0,789244	6
в)	-1	0	0,01	-1	-0,654215	5
	-1	0	0,0001	-1	-0,655423	8
	-1	0	0,000001	-1	-0,655442	11

#### 5.2.4. Метод дотичних (Ньютона)

Нехай на проміжку  $[a, b]$  функція  $f(x)$  неперервна і на  $[a, b]$  існують неперервні похідні  $f'(x)$  і  $f''(x)$ , які не змінюють знака і не набувають значення нуля на  $[a, b]$ .

Якщо в формулі (5.13) прийняти  $\psi(x) \equiv \frac{1}{f'(x)}$ , то отримаємо ітераційний процес за методом Ньютона

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})} \quad (n = 1, 2, \dots). \quad (5.16)$$

За початкове наближення доцільно вибирати той край проміжку  $[a, b]$ , для якого

$$f(x_0)f''(x_0) > 0.$$

Геометрично методом дотичних знаходять абсциси точок перетину дотичних до графіка функції  $y = f(x)$  у точках  $x = x_n$  ( $n = 1, 2, \dots$ ) (рис. 5.7). Якщо почати проводити дотичні з точки  $x_0 = a$ , то точка  $x_1$  вийде за межі проміжку  $[a, b]$ . Отже, вибір початкового наближення суттєво залежить від вигляду функції  $f(x)$ .

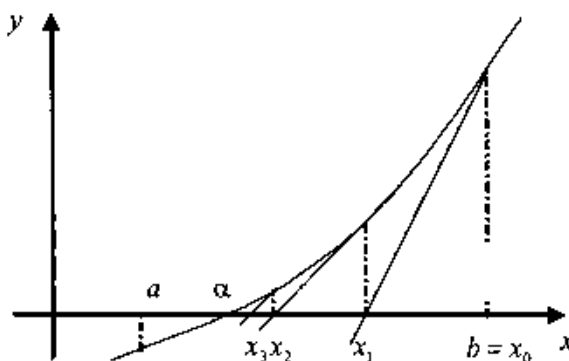


Рис. 5.7.

Метод Ньютона є ітераційним методом другого порядку. Швидкість збіжності методу

$$|x_n - \alpha| \leq \frac{M_2}{2m_1} |x_{n-1} - \alpha|^2$$

де  $m_1 = \min_{[a,b]} |f'(x)|$ ;  $M_2 = \max_{[a,b]} |f''(x)|$ . Це свідчить про швидку збіжність методу Ньютона. В методі хорд, як і в методі дотичних, ми наближаємо корінь  $\alpha$  лише з одного боку.

Якщо  $f'(x)$  мало змінюється на проміжку  $[a, b]$ , то для спрощення обчислень можна використати формулу

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)}. \quad (5.17)$$

Геометрично це означає, що дотичні в точках  $(x_n, f(x_n))$  замінюють прямими, паралельними до дотичних, які проведені до кривої  $y = f(x)$  в точці  $(x_0, f(x_0))$  (рис. 5.8).

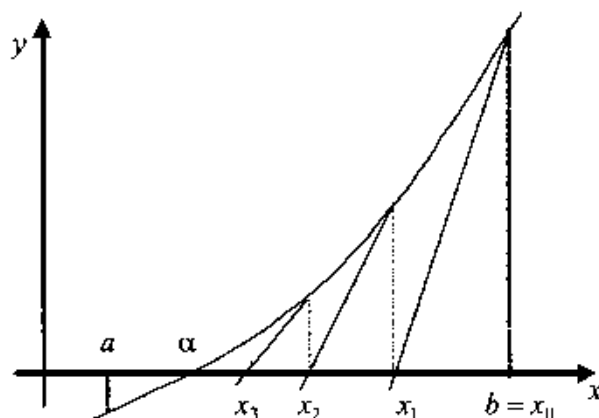


Рис. 5.8.

**Приклад 5.10.** З точністю до  $\varepsilon$  методом дотичних знайти корені рівняння  $x^3 - 4x^2 + 2 = 0$

*Розв'язування.* Під час аналізу цього прикладу в методі дотичних ми з'ясували, що початковими наближеннями будуть відповідні значення  $x_0$  на проміжках

$$[3, 4]: x_0 = 4; \quad [0, 1]: x_0 = 1; \quad [-1, 0]: x_0 = -1.$$

Результати обчислень наведені в табл. 5.6.

Таблиця 5.6

Результати обчислень кореня рівняння для методу дотичних

$a$	$b$	$\varepsilon$	$x_0$	$x_n$	$n$
3	4	0,01	4	3,866240	2
3	4	0,0001	4	3,866198	3
3	4	0,000001	4	3,866198	3
0	1	0,01	1	0,789244	3
0	1	0,0001	1	0,789244	3
0	1	0,000001	1	0,789244	3
-1	0	0,01	-1	-0,655459	3
-1	0	0,0001	-1	-0,655442	4
-1	0	0,000001	-1	-0,655442	4

**П р и к л а д 5.11.** Методом Ньютона з точністю до  $\varepsilon$  розв'язати нелінійне рівняння  $x^3 - 2\sin x - 1 = 0$ .

*Розв'язування.* З'ясуємо, яке початкове наближення вибирати з проміжку  $[1; 2]$ . Знайдемо значення функції та других похідних на кінцях проміжку

$$f''(x) = 6x + 2\sin x;$$

$$f''(1) = 6 + 2\sin 1 > 0;$$

$$f''(2) = 12 + 2\sin 2 > 0;$$

$$f(1) = 1 - 2\sin 1 - 1 = -2\sin 1 < 0; \quad f(2) = 8 - 2\sin 2 - 1 = 7 - 2\sin 2 > 0.$$

Отже, умова  $f(x_0)f''(x_0) > 0$  виконується для точки  $x_0 = 2$ .

З використанням формули (5.16) одержимо наближене значення кореня. Розрахунки наведені в табл. 5.7.

Таблиця 5.7

Ітераційний процес за методом дотичних

$\varepsilon$	$n$	$x_n$	$f(x_n)$
0,01	1	1,596221	1,06769589
	2	1,457463	0,10876887
	3	1,439766	0,00167465
0,0001	1	1,596221	1,06769589
	2	1,457463	0,10876887
	3	1,439766	0,00167465
0,000001	1	1,596221	1,06769549
	2	1,457463	0,10876887
	3	1,439766	0,00167465
	4	1,439485	0,00000042

### 5.2.5. Комбінований метод хорд і дотичних

Методи хорд і дотичних дають наближення кореня з різних боків. Тому часто їх застосовують у комбінації один з одним, і уточнення кореня відбувається швидше.

З урахуванням типу графіка функції ці методи комбінують так.

Якщо  $f'(x_0)f''(x_0) > 0$ , то метод хорд дає наближене значення кореня з недостаткою, а метод дотичних – з надлишком.

Якщо ж  $f'(x_0)f''(x_0) < 0$ , то за методом хорд отримаємо значення кореня  $\alpha$  з надлишком, а методом дотичних – з недостаткою.

У всіх випадках справжній корінь розміщений між наближеними значеннями, отриманими за методами хорд і дотичних, тобто

$$a < \underline{x}_n < \alpha < \overline{x}_n < b,$$

де  $\underline{x}_n$  – наближене значення кореня з недостаткою;  $\overline{x}_n$  – наближене значення кореня з надлишком.

• Нехай  $f'(x_0)f''(x_0) > 0$ , тоді з кінця  $a$  лежать наближені значення, отримані за методом хорд, а з кінця  $b$  – за методом дотичних. Ітераційний процес матиме вигляд

$$\begin{aligned} \underline{x}_0 &= a; \quad \overline{x}_0 = b; \\ \underline{x}_n &= \underline{x}_{n-1} - \frac{f(\underline{x}_{n-1})(\overline{x}_{n-1} - \underline{x}_{n-1})}{f(\overline{x}_{n-1}) - f(\underline{x}_{n-1})}; \\ \overline{x}_n &= \overline{x}_{n-1} - \frac{f(\overline{x}_{n-1})}{f'(\overline{x}_{n-1})}; \quad n = 1, 2, \dots \end{aligned} \quad (5.18)$$

• Нехай  $f'(x_0)f''(x_0) < 0$ , тоді, навпаки, з кінця  $a$  є наближені значення кореня за методом дотичних, а з кінця  $b$  – за методом хорд

$$\begin{aligned} \underline{x}_0 &= a; \quad \overline{x}_0 = b; \\ \underline{x}_n &= \underline{x}_{n-1} - \frac{f(\underline{x}_{n-1})}{f'(\underline{x}_{n-1})}; \\ \overline{x}_n &= \overline{x}_{n-1} - \frac{f(\overline{x}_{n-1})(\overline{x}_{n-1} - \underline{x}_{n-1})}{f(\underline{x}_{n-1}) - f(\overline{x}_{n-1})}; \quad n = 1, 2, \dots \end{aligned} \quad (5.19)$$

Ітераційний процес закінчується, коли  $\left| \bar{x}_n - \underline{x}_n \right| < \varepsilon$ . За значення кореня вибираємо середину звуженого проміжку

$$a = \frac{1}{2}(\bar{x}_n + \underline{x}_n).$$

Комбінований метод є нестационарним методом уточнення дійсних коренів рівняння  $f(x) = 0$ . Збігається він значно швидше, ніж метод дотичних.

**П р и к л а д 5.12.** Комбінованим методом знайти корені рівняння  $x^3 - 4x^2 + 2 = 0$

*Розв'язування.* Знайдемо похідні:

$$f'(x) = 3x^2 - 8x; \quad f''(x) = 6x - 8;$$

$$\text{інтервал } [3; 4]: \quad f'(3) = 3 > 0; \quad f''(3) = 10 > 0;$$

$$f'(4) = 16 > 0; \quad f''(4) = 16 > 0.$$

Маємо випадок, коли  $f'(x_0)f''(x_0) > 0$ . Розрахунки виконаємо за формулами (5.18);

$$\text{інтервал } [0, 1]: \quad f'(0,1) = -0,77 < 0; \quad f''(1) = -5 < 0;$$

$$f''(0,1) = -7,6 < 0; \quad f''(1) = -2 < 0.$$

Знову  $f'(x_0)f''(x_0) > 0$  і обчислення виконаємо за формулами (5.18);

$$\text{інтервал } [-1, 0]: \quad f'(-1) = 11 > 0; \quad f'(-0,1) = 0,83 > 0;$$

$$f''(-1) = -14 < 0; \quad f''(-0,1) = -8,6 < 0.$$

У цьому випадку  $f'(x_0)f''(x_0) < 0$  і корінь шукаємо за формулами (5.19). Результати обчислень для різної точності наведені в табл. 5.8.

Таблиця 5.8

Результати розв'язування прикладу 5.12 комбінованим методом

$a$	$b$	$\varepsilon$	$x_n$	$n$
3	4	0,01	3,866000	2
3	4	0,0001	3,866198	3
3	4	0,000001	3,866198	3
0	1	0,01	0,789065	3
0	1	0,0001	0,789244	3
0	1	0,000001	0,789244	3
-1	0	0,01	-0,657703	2
-1	0	0,0001	-0,655451	3
-1	0	0,000001	-0,655451	3

**П р и к л а д 5.13.** Комбінованим методом хорд і дотичних з точністю до  $\varepsilon$  розв'язати нелінійне рівняння  $x^3 - 2\sin x - 1 = 0$ .

*Розв'язування.* Єдиний корінь рівняння є на проміжку  $[1,2]$ , згідно з прикладом 5.6. Вибір нерухокої точки для методу хорд і початкового наближення для методу дотичних виконують програмним способом. Результати розв'язування заданого рівняння комбінованим методом наведені в табл. 5.9.

Таблиця 5.9

Наближені розв'язки комбінованим методом

$\varepsilon$	$n$	$x_n$	$f(x_n)$
0,01	1	1,42069644	-0,11000930
	2	1,43461491	-0,02887370
	3	1,43940203	-0,00049468
0,0001	1	1,42069644	-0,11000930
	2	1,43461491	-0,02887370
	3	1,43940203	-0,00049468
	4	1,43948509	-0,00000012
0,000001	1	1,42069644	-0,11000930
	2	1,43461491	-0,02887370
	3	1,43940203	-0,00049468
	4	1,43948509	-0,00000012

Метод хорд є ітераційним методом першого порядку, а метод Ньютона – другого порядку. Метод Ньютона можна застосовувати і для знаходження комплексних коренів. Тоді для випадку дійсної функції  $f(x)$  початкове наближення  $x_0$  повинно бути комплексним числом.

Швидкість збіжності розглянутих методів така:

- ✓ метод ітерацій:  $|x_n - \alpha| \leq M^n |x_0 - \alpha|$ , де

$$M = \max_{[a,b]} |\varphi'(x)|;$$

- ✓ метод хорд:  $|x_n - \alpha| \leq \frac{|f(x_n)|}{m}$ , де  $m = \min_{[a,b]} |f'(x)|$ ;

- ✓ метод дотичних:  $|x_n - \alpha| \leq \frac{M_2}{2m} |x_{n-1} - \alpha|^2$ , де

$$M_2 = \max_{[a,b]} |f''(x)|.$$

**Зауваження 1.** За умов теореми 5.6 метод ітерацій збігається для довільного початкового значення  $x_0 \in [a, b]$ . Завдяки цьому він є самовиправним, тобто окрема помилка в обчисленнях не впливає на кінцевий результат, бо помилкове значення можна розглядати як нове початкове наближення  $x_0$ . Ця властивість робить метод ітерацій найнадійнішим методом обчислень.

На практиці часто виникає ситуація, коли проміжок  $[a, b]$  досить великий і умова  $|\varphi'(x)| < 1$  виконується лише в деякому околі кореня, тому в разі невдалого вибору початкового наближення  $x_0$  послідовні наближення  $x_n = \varphi(x_{n-1})$  ( $n = 1, 2, \dots$ ) можуть вийти за межі проміжку  $[a, b]$ .

**Зауваження 2.** У методі дотичних з формули (5.16) бачимо, що чим більше числове значення похідної  $f'(x)$  в околі заданого кореня, тим менша поправка, яку треба додати до  $n$ -го наближення, щоб отримати  $(n+1)$ -ше наближення. Тому метод Ньютона особливо зручно застосовувати тоді, коли в околі заданого кореня графік функції має велику стрімкість. Однак коли числове значення похідної  $f'(x)$  біля кореня мале, то поправки будуть великими й обчислення кореня за цим методом може виявитись довгим, а іноді й зовсім неможливим. Отже, якщо крива



$y = f(x)$  поблизу точки перетину з віссю  $Ox$  майже горизонтальна, то застосовувати метод Ньютона для розв'язування рівняння  $f(x) = 0$  не рекомендуємо.

### 5.3. Ітерації вищих порядків

Нехай задано нелінійне рівняння

$$f(x) = 0, \quad (5.20)$$

яке на проміжку  $[a, b]$  має єдиний корінь  $\alpha$ .

Зведемо рівняння (5.20) до вигляду

$$x = \varphi(x), \quad (5.21)$$

причому шуканий корінь  $x = \alpha$  є і коренем рівняння (5.21).

Побудуємо стаціонарний ітераційний процес

$$x_{n+1} = \varphi(x_n) \quad (n = 0, 1, 2, \dots), \quad (5.22)$$

починаючи з деякого початкового наближення  $x_0$ , близького до  $\alpha$ .

Ітераційний процес (5.22) називають ітераційним процесом  $m$ -го порядку, якщо

$$\varphi'(\alpha) = \varphi''(\alpha) = \dots = \varphi^{(m-1)}(\alpha) = 0; \quad \varphi^{(m)}(\alpha) \neq 0. \quad (5.23)$$

У [12] наведено таку оцінку збіжності ітераційного процесу (5.22):

$$|x_n - \alpha| \leq \left( \frac{M_m}{m!} \right)^{\frac{m^n - 1}{m-1}} \cdot |x_0 - \alpha|^{m^n} \quad (5.24)$$

де  $M_m = \max_{x \in [a, b]} |\varphi^{(m)}(x)|$ .

При  $\frac{M_m}{m!} < 1$  і  $|x_0 - \alpha| < 1$  швидкість збіжності ітераційного процесу  $m$ -го порядку (5.22) буде набагато більша, ніж швидкість збіжності простих ітераційних процесів за методами хорд і дотичних.

Є декілька способів побудови ітераційних процесів вищих порядків: метод Чебишова, метод Ейткена, ітераційний процес за теоремою Кеніга та ін.

### 5.3.1. Метод Чебишова

Зазначений метод призначений для знаходження дійсних коренів рівняння (5.20). В його основі є принцип зображення функції, оберненої до  $f(x)$ , за формулою Тейлора.

Нехай рівняння (5.20) на проміжку  $[a, b]$  має єдиний корінь  $x = \alpha$ . Припустимо, що функція  $f(x)$  є неперервною на відрізку  $[a, b]$ , має неперервні похідні достатньо високого порядку і  $f'(x) \neq 0$  на  $[a, b]$ . Тоді функція  $y = f(x)$  має обернену функцію

$$x = F(y),$$

що визначена на відрізку  $[c, d]$ , який є областю значень  $f(x)$  при  $x \in [a, b]$ . Функція  $F(y)$  має стільки неперервних похідних, скільки має функція  $f(x)$ , тобто

$$x \equiv F(f(x)), \quad x \in [a, b],$$

$$y \equiv f(F(y)), \quad y \in [c, d],$$

і при  $x = \alpha$

$$\alpha \equiv F(0).$$

У [14] доведено, що  $x = \alpha$  є коренем рівняння

$$x = \Phi_{r+1}(x),$$

де

$$\Phi_{r+1}(x) \equiv x + \sum_{k=1}^r (-1)^k \frac{F^{(k)}(f(x))}{k!} (f(x))^k$$

У цьому разі можна з'ясувати, що

$$\Phi_{r+1}^{(l)}(\alpha) = 0 \quad (l = 1, 2, \dots, r).$$

Побудуємо ітераційний процес не нижче  $(r+1)$ -го порядку за формулою

$$x_{n+1} = \Phi_{r+1}(x_n) \quad (n = 0, 1, \dots; x_0 \in [a, b]), \quad (5.25)$$

де  $n$  – номер ітерації.

Ітераційний процес (5.25) при  $r = 1$  запишемо у вигляді

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},$$

тобто він збігається з методом Ньютона.

При  $r = 2$  маємо ітераційний процес третього порядку

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} - \frac{f''(x_n)f^2(x_n)}{2(f'(x_n))^3}, \quad (5.26)$$

який називають методом Чебишова.

У [14] наведено таку оцінку збіжності цього ітераційного процесу:

$$|x_n - \alpha| \leq q^{\frac{(r+1)^n - 1}{r}} \cdot |x_0 - \alpha|^{(r+1)^n} \quad (5.27)$$

де

$$q = \frac{M_{r+1}L^{r+1}}{(r+1)!}; \quad L = \max_{x \in [a, b]} |f'(x)|; \quad M_{r+1} = \max_{x \in [a, b]} |F^{(r+1)}(f(x))|.$$

Якщо  $q < 1$  і  $|x_0 - \alpha| < 1$ , то ітераційний процес (5.26) дуже швидко збігається.

### 5.3.2. Метод Ейткена

Ейткен запропонував спосіб одержання ітерацій вищого порядку на підставі заданої ітерації або двох заданих ітерацій одного й того ж порядку.

Нехай маємо ітерації

$$x_n^{(1)} = \varphi_1(x_{n-1}^{(1)}), \quad x_n^{(2)} = \varphi_2(x_{n-1}^{(2)}), \quad n = 1, 2, \dots$$

порядку  $r$ , які збігаються до розв'язку  $x = \alpha$ .

За допомогою функцій  $\varphi_1(x)$  і  $\varphi_2(x)$  побудуємо функцію

$$\Phi(x) \equiv \frac{x\varphi_1[\varphi_2(x)] - \varphi_1(x)\varphi_2(x)}{x - \varphi_1(x) - \varphi_2(x) + \varphi_1[\varphi_2(x)]}.$$

Тоді ітераційний процес

$$x_n = \Phi(x_{n-1}), \quad n = 1, 2, \dots$$

згідно з [3], має порядок  $r$ , якщо виконується умова  $[\varphi_1'(\alpha) - 1] \cdot [\varphi_2'(\alpha) - 1] \neq 0$ . Якщо  $\varphi(x) = \varphi_1(x) \neq \varphi_2(x)$ , то

$$\Phi(x) \equiv \frac{x\varphi[\varphi(x)] - \varphi^2(x)}{x - 2\varphi(x) + \varphi[\varphi(x)]}.$$

В ітераційному процесі

$$x_n = \Phi(x_{n-1}), \quad n = 1, 2, \dots$$

немає потреби шукати  $\Phi(x)$  в явному вигляді, а можна чинити так. Знаючи початкове наближення  $x_0$ , знаходимо

$$x_1 = \varphi(x_0), \quad x_2 = \varphi(x_1).$$

Далі обчислюємо  $x_3$  за формулою

$$x_3 = \frac{x_0 \cdot x_2 - x_1^2}{x_0 - 2x_1 + x_2} = x_0 - \frac{(\Delta x_0)^2}{\Delta^2 x_0},$$

де

$$\Delta x_i = x_{i+1} - x_i;$$

$$\Delta^2 x_i = x_{i+2} - 2x_{i+1} + x_i$$

Одержуємо нестационарний ітераційний процес

$$\left. \begin{aligned} x_{3i+1} &= \varphi(x_{3i}) \\ x_{3i+2} &= \varphi(x_{3i+1}), \quad i = 0, 1, 2, \dots \\ x_{3(i+1)} &= x_{3i} - \frac{(\Delta x_{3i})^2}{\Delta^2 x_{3i}} \end{aligned} \right\}. \quad (5.28)$$

Ітераційний процес (5.28) визначає ітерацію не нижче другого порядку, якщо  $\varphi(x)$  визначає ітерацію першого порядку, і не нижче  $(2r-1)$ -го порядку, якщо  $\varphi(x)$  визначає ітерацію  $r$ -го порядку.

**П р и к л а д 5.14.** За допомогою ітераційних процесів Чебишова та Ейткена знайти корені рівняння  $f(x) = x^3 + 5x^2 - 15x - 7 = 0$ , розміщені на проміжках  $[-1, 1]$ ,  $[-8, -6]$ ,  $[1, 3]$  із заданою точністю  $\varepsilon$ , тобто до виконання умови

$$|x_{n+1} - x_n| < \varepsilon$$

*Розв'язування.* Для методу Чебишова знайдемо похідні

$$f'(x) = 3x^2 + 10x - 15;$$

$$f''(x) = 6x + 10$$

та запишемо ітераційний процес третього порядку у вигляді

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} - \frac{f''(x_n)f^2(x_n)}{2[f'(x_n)]^3} =$$

$$= x_n - \frac{x_n^3 + 5x_n^2 - 15x_n - 7}{3x_n^2 + 10x_n - 15} - \frac{(3x_n + 5)(x_n^3 + 5x_n^2 - 15x_n - 7)^2}{(3x_n^2 + 10x_n - 15)^3}, \quad n = 0, 1, \dots$$

Результати обчислень на різних проміжках та з різними значеннями точності  $\varepsilon$  наведені в табл. 5.10.

Таблиця 5.10

Результати ітераційного методу Чебишова

$x_0$	$a$	$b$	$\varepsilon$	$x_n$	$n$
1	1	3	0,01	2,41421356	11
			0,0001	2,41421356	12
			0,000001	2,41421356	12
3	1	3	0,000001	2,41421356	4
1	-1	1	0,01	-0,41421356	3
			0,0001	-0,41421356	3
			0,000001	-0,41421356	3
-1	-1	1	0,000001	-0,41421356	4
-8	-8	-6	0,01	-6,99999992	3
			0,0001	-7,0	4
			0,000001	-7,0	4
-6	-8	-6	0,000001	-7,0	4

У табл. 5.11 методом Чебишова для початкового наближення  $x_0$ ,  $\varepsilon = 0,0001$  одержано на кожній ітерації значення  $n$ -го наближення  $x_n$  і значення  $f(x_n)$ .

Таблиця 5.11

Значення розв'язку і функції на кожній ітерації

$n$	$x_n$	$f(x_n)$
0	0	-7
1	-0,39407407	-0,37361449
2	-0,41421242	-0,00002119
3	-0,41421356	0

Для методу Ейткена рівняння  $f(x) = 0$  перетворимо до вигляду

$$x = \varphi(x), \text{ тобто } x = \frac{-7 + x^3 + 5x^2}{15}.$$

Задамо початкове наближення  $x_0$ , обчислимо  $x_1$  і  $x_2$ :

$$x_1 = \varphi(x_0) = \frac{-7 + x_0^3 + 5x_0^2}{15}; \quad x_2 = \varphi(x_1) = \frac{-7 + x_1^3 + 5x_1^2}{15},$$

далі побудуємо ітераційний процес

$$x_{3i+1} = \varphi(x_{3i}) = \frac{-7 + x_{3i}^3 + 5x_{3i}^2}{15};$$

$$x_{3i+2} = \varphi(x_{3i+1}) = \frac{-7 + x_{3i+1}^3 + 5x_{3i+1}^2}{15};$$

$$x_{3i+3} = x_{3i} - \frac{(x_{3i+1} - x_{3i})^2}{x_{3i+2} - 2x_{3i+1} + x_{3i}}, \quad i = 0, 1, 2, \dots$$

до виконання умови

$$|x_{3i+3} - x_{3i+2}| < \varepsilon.$$

Результати обчислень за методом Ейткена наведені в табл. 5.12.

Таблиця 5.12  
Наближені значення розв'язку за методом Ейткена

$x_0$	$\varepsilon$	$x_n$	$n$
3	0,01	2,41425458	4
	0,0001	2,41421356	5
	0,000001	2,41421356	5
1	0,01	-0,40947458	2
	0,0001	-0,41421247	3
	0,000001	-0,41421247	3
-8	0,01	-7,00028110	5
	0,0001	-7,00000010	6
	0,000001	-7,0	7

У табл. 5.13 наведено значення наближеного розв'язку  $x_n$  і функції  $f(x_n)$  при  $x_0 = -0,1$ ;  $\varepsilon = 0,0001$ , знайдені за методом Ейткена.

Таблиця 5.13

Значення розв'язку і функції за методом Ейткена

$n$	$x_n$	$f(x_n)$
0	-0,1	-5,451
1	-0,41067050	-0,06595075
2	-0,41421295	-0,00001138

**П р и к л а д 5.15.** Знайти корінь нелінійного рівняння

$$f(x) = 3x - \cos x - 1 = 0$$

розміщеного на відрізьку  $[0, 1]$ *Розв'язування.* За методом Чебишова ітераційний процес запишемо формулою

$$x_{n+1} = x_n - \frac{3x_n - \cos x_n - 1}{3 + \sin x_n} - \frac{\cos x_n (3x_n - \cos x_n - 1)^2}{2[3 + \sin x_n]^3}, \quad n = 0, 1, 2, \dots$$

Результати обчислень наведені в табл. 5.14.

Таблиця 5.14

Значення розв'язку нелінійного рівняння  
за методом Чебишова

$x_0$	$\varepsilon$	$x_n$	$n$
0	0,01	0,60710165	2
	0,0001	0,60710165	3
	0,000001	0,60710165	3

Для початкового наближення  $x_0 = 0,1$  та  $\varepsilon = 0,0001$  у табл. 5.15 за методом Чебишова одержано такі значення  $x_n$  і  $f(x_n)$ .

Таблиця 5.15

Значення розв'язку і функції  
за методом Чебишова

$n$	$x_n$	$f(x_n)$
0	0,1	-1,69500417
1	0,59881817	-0,02954785
2	0,60710162	-0,00000011
3	0,60710164	0

Для методу Ейткена зведене рівняння  $x = \varphi(x)$  запишемо так:

$$x = \frac{1 + \cos x}{3},$$

а ітераційний процес Ейткена має вигляд

$$x_{3i+1} = \frac{1 + \cos x_{3i}}{3};$$

$$x_{3i+2} = \frac{1 + \cos x_{3i+1}}{3};$$

$$x_{3i+3} = x_{3i} - \frac{(x_{3i+1} - x_{3i})^2}{x_{3i+2} - 2x_{3i+1} + x_{3i}}, \quad i = 0, 1, 2, \dots$$

Результати за методом Ейткена наведені в табл. 5.16.

Таблиця 5.16

Значення розв'язку за методом Ейткена

$x_0$	$\varepsilon$	$x_n$	$n$
1	0,01	0,60331992	1
	0,0001	0,60799532	2
	0,000001	0,60693190	3

У табл. 5.17 за методом Ейткена записано для кожної  $n$ -ї ітерації значення  $x_n$  і  $f(x_n)$  для  $\varepsilon = 0,0001$  і  $x_0 = 0,1$ .

Таблиця 5.17

Значення розв'язку і функції за методом Ейткена

$n$	$x_n$	$f(x_n)$
0	0,1	-1,69500417
1	0,60322292	-0,01384279
2	0,60710132	-0,00000117

**П р и к л а д 5.16.** Порівняти розглянуті методи за кількістю ітерацій.

*Розв'язування.*

1. Дослідимо кількість ітерацій у разі обчислення кореня рівняння  $f(x) = 3x - \cos x - 1 = 0$ , розміщеного на проміжку  $[-1, 1]$ , різними методами. Результати наведені в табл. 5.18.



Таблиця 5.18

## Результати порівняння кількості ітерацій

$\varepsilon$	Метод						
	дихото- мії	ітера- цій	хорд	Нью- тона	комбі- нований	Ейтке- на	Чеби- шова
$10^{-2}$	5	4	2	1	1	1	2
$10^{-4}$	13	7	4	2	2	2	3
$10^{-6}$	15	10	5	2	3	3	3

2. Знайдемо корені кубічного багаточлена

$$f(x) = x^3 + 5x^2 - 15x - 7 = 0$$

різними методами. Для різної точності виконано порівняльні розрахунки. Результати наведені в табл. 5.19.

Таблиця 5.19

## Порівняльна таблиця кількості ітерацій

[a, b]	$\varepsilon$	Метод						
		дихо- томії	ітера- цій	хорд	Нью- тона	комбі- нований метод	Ейтке- на	Чеби- шова
[-1,1]	$10^{-2}$	7	4	4	2	2	2	3
	$10^{-4}$	15	7	6	2	3	3	3
	$10^{-6}$	21	10	8	3	3	3	3
[-8,-6]	$10^{-2}$	1	4	4	3	2	5	3
	$10^{-4}$	1	6	7	3	2	6	4
	$10^{-6}$	1	9	10	3	4	7	4
[1,3]	$10^{-2}$	7	5	5	2	3	4	11
	$10^{-4}$	15	8	8	3	4	5	12
	$10^{-6}$	21	11	11	3	4	5	12

### Завдання до розділу 5

5.1. Локалізувати й обчислити корені нелінійного рівняння з точністю  $\varepsilon = 10^{-5}$

- 1)  $x^2 + 4\sin(x) - 1 = 0$ ;
- 2)  $(x-3)^3 + 4\sin(x) = 0$ ;
- 3)  $e^x - 3x - 2 = 0$ ;
- 4)  $2^x - x - 2 = 0$ ;
- 5)  $x^2 + 2\cos(x) - 4 = 0$ ;
- 6)  $(x+1)^3 + \frac{1}{4}e^x = 0$ ;
- 7)  $e^{-x} + x^2 - 2 = 0$ ;
- 8)  $\lg(2x) + x^2 = 0$ ;
- 9)  $\lg(x) + x - 0,5 = 0$ ;
- 10)  $2(x-3)\ln(x) - 1 = 0$ ;
- 11)  $x^2 - \sin(\pi x) = 0$ ;
- 12)  $2\lg(x) - 0,5x + 1 = 0$ ;
- 13)  $\sin(x) - 2x + 0,5 = 0$ ;
- 14)  $2^x - 2x = 0$ ;
- 15)  $\lg(x) - 3x + 5 = 0$ ;
- 16)  $2\sin(x - 0,6) + x - 1,5 = 0$ ;
- 17)  $x^3 - \sin(x) - 2 = 0$ ;
- 18)  $(x-1)^2 - 0,5e^x = 0$ ;
- 19)  $\operatorname{tg}(0,58x + 0,1) - x^2 = 0$ ;
- 20)  $x^3 - \ln(x+2) = 0$ ;
- 21)  $(2-x)e^x - 0,5 = 0$ ;
- 22)  $2x + \lg(2x+3) - 1 = 0$ ;
- 23)  $5^x - 6x - 3 = 0$ ;
- 24)  $3x - \cos(x) - 1 = 0$ ;
- 25)  $5\sin(x) - x + 1 = 0$ ;
- 26)  $x\lg(x+1) - 1 = 0$ ;
- 27)  $e^{-2x} - 2x + 1 = 0$ .

Відповіді:

- 1)  $-2,1069; 0,2381$ ;

- 2) 1,4187; 3,1423; 4,5829;
- 3) -0,4553; 2,1254;
- 4) -1,6901; 2,0;
- 5) -2,3136; 2,3136;
- 6) -1,3956;
- 7) -0,5373; 1,3116;
- 8) 0,3668;
- 9) 0,6724;
- 10) 0,7969; 3,4078;
- 11) 0; 0,7872;
- 12) 0,3975; 4,6816;
- 13) 0,4816;
- 14) 1; 2;
- 15) 0,00001; 1,7475;
- 16) 0,9031;
- 17) 1,4390;
- 18) 0,2133;
- 19) -0,1391; 0,7905;
- 20) -0,5876; 1,0571;
- 21) 1,9272;
- 22) 0,2304;
- 23) 1,5633;
- 24) 0,6071;
- 25) -0,2550;
- 26) 2,0592;
- 27) 0,63972.

**5.2.** Знайти найменші відмінні від нуля корені нелінійного рівняння.

- 1)  $x \sin(x) + 1 = 0$ .
- 2)  $x \cos(x) - \sin(x) = 0$ .
- 3)  $\operatorname{tg}(0,55x + 0,1) - x^2 = 0$ .

*Відповіді:*

- 1) 3,4368;
- 2) 4,4934;
- 3) -0,1441.

**5.3.** Знайти найменший за модулем дійсний корінь алгебричного рівняння з точністю  $\varepsilon = 10^{-5}$

- 1)  $x^3 - 9x + 3 = 0$ ;
- 2)  $x^3 + 10x - 9 = 0$ ;
- 3)  $x^3 - 8x + 2 = 0$ .

- 4)  $x^3 - 15x + 10 = 0$ ;
- 5)  $x^4 + 4x^3 + 4,8x^2 + 16x + 1 = 0$ ;
- 6)  $x^4 + 4x^3 + 4x^2 + 4x - 1 = 0$ ;
- 7)  $8x^4 - 8x^2 + 32x + 1 = 0$ ;
- 8)  $x^4 - 2x^3 + x - 1,5 = 0$ ;
- 9)  $x^4 - 3x^3 + 8x^2 - 5 = 0$ ;
- 10)  $x^3 + 3x^2 - 24x + 1 = 0$ ;
- 11)  $x^3 - 3x^2 - 3 = 0$ ;
- 12)  $x^3 - x - 1 = 0$ ;
- 13)  $5x^3 - 20x + 3 = 0$ ;
- 14)  $x^3 - 2x^2 + 2x - 1 = 0$ ;
- 15)  $x^3 - 3x^2 + 9x - 10 = 0$ ;
- 16)  $x^3 + 0,4x^2 + 0,6x - 1,6 = 0$ ;
- 17)  $5x^3 + 2x^2 - 15x - 6 = 0$ ;
- 18)  $x^6 - 3x^2 + x - 1 = 0$ ;
- 19)  $x^5 - 3x^2 + 1 = 0$ ;
- 20)  $x^6 - x^2 + 0,5x - 2 = 0$ ;
- 21)  $2x^3 - 3x^2 - 12x + 8 = 0$ ;
- 22)  $x^3 - 4x^2 + 10x - 10 = 0$ ;
- 23)  $x^3 + 2x - 11 = 0$ ;
- 24)  $x^5 + 3x^4 - 2x^3 + x^2 - x + 1 = 0$ ;
- 25)  $x^5 - 2x^4 - 7x^3 - 33x^2 - 30x - 25 = 0$ ;
- 26)  $x^3 - 2x^2 + 7x + 3 = 0$ ;
- 27)  $x^3 - 2x^2 - 4x + 7 = 0$ ;
- 28)  $x^3 - 0,2x^2 + 0,5x + 1,5 = 0$ ;
- 29)  $x^3 + 3x^2 - 1 = 0$ ;
- 30)  $x^4 + 10x^3 - 1 = 0$ .

*Βιθνοβιθι:*

- |             |             |
|-------------|-------------|
| 1) 0,3378;  | 8) -0,9397; |
| 2) 0,841;   | 9) -0,6887; |
| 3) 0,2520;  | 10) 0,0419; |
| 4) 3,4826;  | 11) 3,2790; |
| 5) -0,0637; | 12) 1,3247; |
| 6) 0,2010;  | 13) 0,1514; |
| 7) -0,0310; | 14) 1,0000; |

15) 1,4814;  
16) 0,9018;  
17) 1,7320;  
18) 1,2962;  
19) 0,5992;  
20) 1,1885;  
21) 0,6113;  
22) 1,6293;

23) 1,9263;  
24) -3,6493;  
25) 5,0000;  
26) -0,3796;  
27) 1,4625;  
28) -0,9464;  
29) 0,5321;  
30) 0,4573.

## ІТЕРАЦІЙНІ МЕТОДИ РОЗВ'ЯЗУВАННЯ СИСТЕМ НЕЛІНІЙНИХ РІВНЯНЬ

Розглянемо систему нелінійних рівнянь

[illegible]

Систему (6.1) часто зручно вивчати як операторне рівняння в евклідовому просторі  $R^n$ . Позначимо  $x = (x_1, x_2, \dots, x_n)^T \in R^n$ ,  $F(x) = (f_1(x), f_2(x), \dots, f_n(x))^T$  і запишемо (6.1) у вигляді операторного рівняння

$$F(x) = 0, \quad 0 = (0, 0, \dots, 0)^T \in R^n, \quad (6.2)$$

де  $F: R^n \rightarrow R^n$  – нелінійний оператор.

Розглянемо ітераційний метод, який генерує послідовність точок  $\{x_k\}$ ,  $x_k \in R^n$ ,  $k=1,2,\dots$ , починаючи з початкового наближення  $x_0 \in R^n$

**Означення 6.1.** Нехай  $x^* \in R^n$ ,  $x_k \in R^n$ ,  $k=0,1,2,\dots$ . Тоді кажуть, що послідовність  $\{x_k\}$  збігається до  $x^*$ , якщо

$$\lim_{k \rightarrow \infty} \|x_k - x^*\| = 0$$

Якщо, крім того, існують константа  $c$  і ціле  $\bar{k} \geq 0$  такі, що для всіх  $k \geq \bar{k}$

$$\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|,$$

то кажуть, що  $\{x_k\}$  є лінійно збіжною до  $x^*$

Якщо для деякої послідовності дійсних чисел  $\{c_k\}$ , збіжної до 0, виконується нерівність

$$\|x_{k+1} - x^*\| \leq c_k \|x_k - x^*\|,$$

то кажуть, що  $\{x_k\}$  збігається надлінійно до  $x^*$

Якщо  $\{x_k\}$  збігається до  $x^*$  та існують сталі  $p > 1$ ,  $c \geq 0$  і  $\bar{k} \geq 0$  такі, що для всіх  $k \geq \bar{k}$

$$\|x_{k+1} - x^*\| \leq c \|x_k - x^*\|^p \quad (6.3)$$

то кажуть, що  $\{x_k\}$  збігається до  $x^*$  з порядком, що принаймні дорівнює  $p$ . Якщо  $p = 2$  або  $p = 3$ , то кажуть, що швидкість збіжності є квадратичною або кубічною. Якщо  $\{x_k\}$  збігається до  $x^*$  і замість (6.3) для  $k \geq \bar{k}$  виконується

$$\|x_{k+j} - x^*\| \leq c \|x_k - x^*\|^p$$

за деякого фіксованого  $j$ , то кажуть, що  $\{x_k\}$  має  $j$ -крокову збіжність з порядком, що принаймні дорівнює  $p$

**Означення 6.2.** Нехай  $n > 0$ ,  $G: R^n \rightarrow R^{n \times n}$ ,  $x \in R^n$ , і нехай  $\|\cdot\|$  – норма в  $R^n$ , а  $\|\cdot\|$  – норма в  $R^{n \times n}$ . Будемо говорити, що функція  $G$  неперервна за Ліпшицем у точці  $x$ , якщо існують відкрита множина  $D \subset R^n$ ,  $x \in D$  і константа  $\gamma$  такі, що для всіх  $v \in D$

$$\|G(v) - G(x)\| \leq \gamma \|v - x\|. \quad (6.4)$$

Константу  $\gamma$  називають константою Ліпшиця для  $G$  в точці  $x$ . Для довільної конкретної  $D$ , яка містить  $x$  і для якої виконується

(6.4), кажуть, що  $G$  неперервна за Ліпшицем в околі  $D$  точки  $x$ . Якщо (6.4) справджується для довільного  $x \in D$ , то  $G \in Lip_\gamma(D)$ .

**Лема 6.1.** Нехай функція  $F: R^n \rightarrow R^n$  неперервно диференційовна в опуклій відкритій області  $D \subset R^n$ . Тоді для довільних  $x, x+p \in D$

$$F(x+p) - F(x) = \int_0^1 J(x+tp) p dt = \int_x^{x+p} F'(z) dz.$$

**Лема 6.2.** Нехай функція  $F: R^n \rightarrow R^n$  неперервно диференційовна в опуклій відкритій області  $D \subset R^n$ . Нехай для заданої векторної норми, породженої нею матричної норми і константи  $\gamma$  якобіан  $J$  неперервний за Ліпшицем в околі  $D$  точки  $x$ . Тоді для довільних  $x+p \in D$

$$\|F(x+p) - F(x) - J(x)p\| \leq \frac{\gamma}{2} \|p\|^2 \quad (6.5)$$

Для оцінки швидкості збіжності ітераційних процесів ми використовуватимемо норми векторів і матриць. Зазначимо, що для нас найліпше підходять такі векторні норми для  $v = (v_1, v_2, \dots, v_n)^T \in R^n$ :

$$\|v\|_1 = \sum_{i=1}^n |v_i|$$

( $l_1^{(n)}$ -норма, або норма найменших абсолютних різниць),

$$\|v\|_2 = \left( \sum_{i=1}^n (v_i)^2 \right)^{1/2}$$

( $l_2^{(n)}$ -евклідова, або найменших квадратів, норма).

Відповідно, породжені ними матричні  $l_1^{(n)}$ - та  $l_2^{(n)}$ -норми для матриць  $A \in R^{n \times n}$  задають такими формулами:

$$\|A\|_1 = \max_{1 \leq j \leq n} \{\|a_{\cdot j}\|\}, \quad \text{де через } a_{\cdot j} \text{ позначено } j\text{-й стовпець}$$

матриці  $A$ ;

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}, \quad \text{де } \lambda_{\max}(A^T A) \text{ — максимальне власне}$$

значення матриці  $A^T A$ .



Для розв'язування рівняння (6.2) розроблено багато методів. Ми розглянемо низку найвідоміших, серед яких метод простої ітерації, градієнтний метод, метод Ньютона та його модифікації.

## 6.2. Метод простої ітерації

Метод простої ітерації призначений для розв'язування нелінійного рівняння вигляду

$$x = g(x), \quad (6.6)$$

або

$$x_i = g_i(x_1, x_2, \dots, x_n), \quad i = 1, \dots, n$$

Рівняння у вигляді (6.6) часто отримують у разі апроксимації нелінійних крайових задач скінченними різницями. До такого вигляду можна звести і рівняння (6.1).

Ітерації методу виконують за формулою

$$x^{(k+1)} = g(x^{(k)}), \quad (6.7)$$

або

$$x_i^{(k+1)} = g_i(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}), \quad i = 1, \dots, n; \quad k = 0, 1, \dots$$

де  $x^{(0)}$  – задане початкове наближення.

Для побудованої таким способом послідовності векторів  $\{x^{(k)}\}_{k \geq 0}$  справджується така теорема.

**Теорема 6.1.** Нехай  $D$  – прямокутний паралелепіпед в  $n$ -вимірному евклідовому просторі

$$D = \{x \in R^n \mid a_k \leq x_k \leq b_k, 1 \leq k \leq n\}$$

Нехай вектор-функція  $g(x)$  визначена на  $D$  і задовольняє таку умову:

існує дійсна константа  $q$ ,  $0 < q < 1$ , така, що

$$\|g(x) - g(y)\| \leq q \|x - y\| \quad (6.8)$$

для всіх  $x, y \in D$ .

Тоді:

1) рівняння (6.6) має єдиний розв'язок  $x^*$  у  $D$ ;



### 6.3. Градієнтний метод розв'язування систем нелінійних рівнянь

Припустимо, що функції  $f_i(x_1, \dots, x_n)$  ( $i=1, \dots, n$ ) – дійсні функції дійсних змінних  $x_1, \dots, x_n$ , неперервні з першими частинними похідними в деякій області  $G$ , а також у цій області система (6.1) має єдиний розв'язок.

Розглянемо функцію

$$u(x_1, \dots, x_n) = \sum_{i=1}^n f_i^2(x_1, \dots, x_n). \quad (6.11)$$

Кожному розв'язку  $\alpha = \{\alpha_1, \dots, \alpha_n\}$  системи (6.1) відповідає нульовий мінімум функції  $u(x)$ , і, навпаки, кожна точка нульового мінімуму  $u(x)$  дає розв'язок системи (6.1). Отже, задача відшукування розв'язку системи (6.1) зводиться до задачі відшукування точок нульового мінімуму допоміжної функції  $u(x)$ .

Для розв'язування цієї задачі й використовують методи спуску.

Нехай  $x^{(0)}$  – початкове наближення до деякого розв'язку системи  $f(x)=0$ , тобто до точки мінімуму функції  $u(x)$ . Розглянемо поверхню  $u(x) = u(x^{(0)})$ . Нехай  $v_0$  – вектор, який не є дотичним до поверхні в точці  $x^{(0)}$ . Розглянемо пряму  $x = x^{(0)} + \lambda v_0$  і будемо рухатися з точки  $x^{(0)}$  по цій прямій у напрямі спадання  $u(x)$  доти, доки функція  $F(\lambda) = u(x^{(0)} + \lambda v_0)$  не досягне найближчого до  $\lambda=0$  мінімуму. Нехай це буде при  $\lambda = \lambda_0$ . Одержимо точку  $x^{(1)} = x^{(0)} + \lambda_0 v_0$ . Розглянемо поверхню  $u(x) = u(x^{(1)})$  і вектор  $v_1$ , що не дотикається до поверхні в точці  $x^{(1)}$ . Знову шукаємо мінімум  $F_1(\lambda) = u(x^{(1)} + \lambda v_1)$ . Нехай він досяжний при  $\lambda = \lambda_1$ , тоді приймемо  $x^{(2)} = x^{(1)} + \lambda_1 v_1$ . Кожного разу рухаємося в напрямі спадання  $u(x)$ , тобто відбувається спуск до точки мінімуму.

На  $k$ -му кроці необхідно знайти мінімум функції

$$F_k(\lambda) = u(x^{(k)} + \lambda v_k), \quad (6.12)$$

тобто розв'язати рівняння

$$F'_k(\lambda) = 0. \quad (6.13)$$

Позначимо розв'язок рівняння (6.13) через  $\lambda_k$ , тоді одержимо

$$x^{(k+1)} = x^{(k)} + \lambda_k v_k. \quad (6.14)$$

Вектори  $v_k$ , які визначають напрям спуску на  $k$ -му кроці, можна вибирати різними способами. Найліпше за вектор  $v_k$  брати

$$-\text{grad } u(x^{(k)}) = \left\{ -\frac{\partial u(x^{(k)})}{\partial x_1}, \dots, -\frac{\partial u(x^{(k)})}{\partial x_n} \right\}, \quad (6.15)$$

бо цей вектор дає напрям найшвидшого спадання  $u(x)$  в точці  $x^{(k)}$ . Методи, у яких беруть  $v_k = -\text{grad } u(x^{(k)})$ , називають градієнтними методами, або методами найшвидшого спуску.

Часто за  $v_k$  вибирають вектори в напрямі осей координат, тоді кажуть про методи покоординатного спуску. Доцільно за вектор  $v_k$  брати напрям уздовж тієї осі координат, яка відповідає максимальній за абсолютною величиною похідній функції  $u(x)$  в заданій точці. Такого типу методи називають релаксаційними.

Розглянемо один варіант вибору  $\lambda$  для градієнтного методу. В цьому варіанті за значення  $\lambda_k$  приймають абсцису точки перетину дотичної до кривої  $y = F_k(\lambda) = u(x^{(k)} - \lambda \text{grad } u(x^{(k)}))$  в точці  $(0, F_k(0))$  з віссю  $\lambda$ .

Очевидно, що

$$\lambda_k = -\frac{F_k(0)}{F'_k(0)} = \frac{u(x^{(k)})}{\sum_{i=1}^n \left[ \frac{\partial u(x^{(k)})}{\partial x_i} \right]^2}. \quad (6.16)$$

Тоді послідовні наближення визначають зі співвідношення

$$x^{(k+1)} = x^{(k)} - \frac{u(x^{(k)})}{\sum_{i=1}^n \left[ \frac{\partial u(x^{(k)})}{\partial x_i} \right]^2} \text{grad } u(x^{(k)}). \quad (6.17)$$

Як звичайно, застосування градієнтних методів дає змогу знайти розв'язок системи (6.1) і для менш точних початкових наближень, ніж це потрібно в разі застосування методу Ньютона, однак збіжність цих методів дуже повільна.

## 6.4. Метод Ньютона та аспекти його практичної реалізації

Нехай потрібно знайти розв'язок рівняння

$$F(x) = 0, \quad (6.18)$$

де  $F: R^n \rightarrow R^n$  – нелінійна за  $x$  функція. Треба знайти  $x^*$  таке, що  $F(x^*) = 0$ .

Ми розглянемо виведення формули методу Ньютона для задачі (6.18) і обговоримо умови його застосування, питання його реалізації та особливості його збіжності. Не розглядатимемо питання існування і єдиності розв'язку задачі (6.18), вважаючи, що на практиці з цим, зазвичай, не виникає проблем.

Метод Ньютона для задачі (6.18) отримаємо зі знаходження лінійної апроксимації для  $F(x)$ , яка відповідає поточному наближенню  $x_k$ . (Оскільки розглядатимемо рівняння (6.18) у векторному вигляді, то надалі для позначення номера ітерації будемо використовувати нижній індекс). Цю апроксимацію будемо за формuloю Ньютона–Лейбніца, а саме:

$$F(x_k + p) = F(x_k) + \int_{x_k}^{x_k+p} J(z) dz, \quad (6.19)$$

де  $J(z) = F'(z)$ . Щоб отримати лінійну апроксимацію для  $F(x)$  щодо приросту  $p$  до  $x_k$ , наближаємо інтеграл у (6.19) лінійним членом  $J(x_k)p$ . Тоді ньютонівський напрям  $s_k$  знаходимо з рівняння

$$F(x_k) + J(x_k)s_k = 0. \quad (6.20)$$

Отже, розв'язуємо систему

$$\begin{aligned} J(x_k)s_k &= -F(x_k), \\ x_{k+1} &= x_k + s_k \end{aligned} \quad (6.21)$$

Оскільки не очікуємо, що  $x_{k+1}$  дорівнюватиме  $x^*$  а насправді буде лише наближенням до  $x^*$ , більш вдалим, ніж  $x_k$ , то побудуємо на базі ньютонівського кроку (6.21) алгоритм, застосовуючи його ітеративно з початкового наближення  $x_0$ .

**Алгоритм 6.1.** Задані: неперервно диференційовна функція  $F : R^n \rightarrow R^n$   $x_0 \in R^n$  Потрібно на кожній  $k$ -й ітерації розв'язати рівняння (6.21).

Що стосується збіжності, то природно, що метод Ньютона буде поводитися не ліпше, ніж у задачах з однією змінною. Тому особливості збіжності методу Ньютона свідчать про те, як його використовувати в багатовимірних алгоритмах. Для нас завжди буде бажаним застосувати його принаймні на завершальних етапах будь-якого нелінійного алгоритму, щоб використати його швидку локальну збіжність. Однак для того, щоб він збігався глобально, його треба модифікувати.

Є дві головні проблеми, які стосуються реалізації ітерації алгоритму 6.1. По-перше, якобіан функції  $F$  може не бути заданий аналітично. Це, зазвичай, трапляється в реальних застосуваннях, коли, наприклад,  $F$  сама не задана аналітично. Отже, апроксимація  $J(x_k)$  скінченними різницями або іншим, менш дорогим способом, є важливою задачею і буде розглянута нижче. По-друге, матриця  $J(x_k)$  може бути виродженою або погано зумовленою, і тоді нема змоги отримати достовірний розв'язок  $s_k$  системи лінійних рівнянь

$$J(x_k)s_k = -F(x_k).$$

Тому можлива модифікація алгоритму 6.1 полягає у внесенні деякого збурення в  $J(x_k)$ , достатнього, щоб зробити її добре зумовленою і перейти до наступних дій на ітерації. Однак така модифікація  $J(x_k)$  не має переконливого обґрунтування в термінах вихідної задачі (6.1), і тому її не рекомендують. Замість цього в разі поганої зумовленості  $J(x_k)$  надають перевагу переходу безпосередньо до такого глобального методу розв'язування нелінійних рівнянь, який відповідає збуренню лінійної моделі (6.20), цілком осмисленому з погляду вихідної нелінійної задачі [9].

Розглянуті вище переваги і недоліки методу Ньютона слугують ключем до розвитку багатовимірних алгоритмів, оскільки вони свідчать про властивості, які бажано зберегти, і про області, які потребують поліпшення і модифікації. Назвемо їх.

*Переваги.*

1. Квадратична збіжність з доброго наближення за умови, що  $J(x_k)$  не вироджена.
2. Точний розв'язок за одну ітерацію для лінійної  $F$  (точний на кожній ітерації для всіх лінійних компонент  $F$ ).

*Недоліки.*

1. Відсутність глобальної збіжності для багатьох задач.
2. Необхідність обчислення  $J(x_k)$  на кожній ітерації.
3. Необхідність розв'язування на кожній ітерації системи лінійних рівнянь, яка може бути виродженою або погано зумовленою.

## 6.5. Локальна збіжність методу Ньютона

Для заданої векторної норми  $\|\cdot\|$  через  $N(x, r)$  будемо позначати відкритий окіл радіусом  $r$  з центром в  $x$ , тобто  $N(x, r) = \{\tilde{x} \in R^n \mid \|\tilde{x} - x\| \leq r\}$

**Теорема 6.2.** Нехай функція  $F: R^n \rightarrow R^n$  неперервно диференційовна у відкритій опуклій області  $D \subset R^n$ . Припустимо, що існують  $x^* \in R^n$  і  $r, \beta > 0$  такі, що  $N(x, r) \subset D$ ,  $F(x^*) = 0$ , існує  $J(x^*)^{-1}$ , причому  $\|J(x^*)^{-1}\| \leq \beta$  і  $J(x)$  задовольняє умову Ліпшиця з константою  $\gamma > 0$ . Тоді існує  $\varepsilon > 0$  таке, що для всіх  $x_0 \in N(x^*, \varepsilon)$  послідовність  $\{x_k\}$ , породжена методом Ньютона

$$x_{k+1} = x_k - J(x_k)^{-1} F(x_k), \quad k = 0, 1, \dots, \quad (6.22)$$

коректно визначена, збігається до  $x^*$  і задовольняє нерівність

$$\|x_{k+1} - x^*\| \leq \beta \gamma \|x_k - x^*\|^2 \quad k = 0, 1, \dots \quad (6.23)$$

Константи  $\gamma$  і  $\beta$  можна об'єднати в одну  $\gamma_R = \gamma\beta$ , яка, по суті, є константою Ліпшиця для вимірювання відносної нелінійності  $F$  в  $x^*$ , бо

$$\|J(x^*)^{-1} [J(x) - J(x^*)]\| \leq \|J(x^*)^{-1}\| \|J(x) - J(x^*)\| \leq \beta \gamma \|x - x^*\| = \gamma_R \|x - x^*\|$$

для  $x \in N(x^*, r)$ . Отже, теорему 6.2 також можна трактувати як

твердження про те, що радіус збіжності методу Ньютона обернено пропорційний до відносної нелінійності  $F$  в  $x^*$ . Відносна нелінійність функції є головним чинником, який визначає поведінку цієї функції в наших алгоритмах.

Може природно постати запитання, чи достатньо близькі на практиці до реальної області збіжності методу Ньютона її оцінки з теореми 6.2. Виявляється, що оцінка  $\varepsilon$  в теоремі 6.2 розрахована на найгірший випадок. Вона досить добре відображає те, наскільки далеко поширюється область квадратичної збіжності в тому напрямі з  $x_k$ , у якому  $F$  найбільше нелінійна. З іншого боку, в тих напрямках з  $x_k$ , у яких  $F$  менш нелінійна, область збіжності може простягатися набагато далі.

Глибоке дослідження методу Ньютона для функціональних рівнянь виконав Л. Канторович. Тому цей метод ще називають методом Ньютона–Канторовича. Теорема Канторовича відрізняється від теореми 6.3 головню тим, що не припускає існування розв'язку рівняння  $F(x) = 0$ . Замість цього вона доводить таке: якщо яacobian  $J(x_0)$  не вироджений,  $J(x)$  неперервний за Ліпшицем у деякій області, яка містить  $x_0$ , і перший крок методу Ньютона достатньо малий щодо нелінійності  $F$ , то в цій області повинен існувати корінь і бути єдиним. Ця теорема має і важливе методичне значення.

**Теорема 6.3 (Канторовича).** Нехай  $r > 0$ ,  $x_0 \in R^n$ ,  $F: R^n \rightarrow R^n$ , і припустимо, що функція  $F$  неперервно диференційована у  $N(x_0, r)$ . Нехай для заданої векторної норми та породженої нею операторної норми  $J \in Lip_\gamma(N(x_0, r))$ , причому  $J(x_0)$  не вироджений, і існують  $\beta, \eta \geq 0$  такі, що

$$\|J(x_0)^{-1}\| \leq \beta, \quad \|J(x_0)^{-1} F(x_0)\| \leq \eta.$$

Якщо  $\alpha = \beta\eta \leq \frac{1}{2}$  і  $r \geq r_0 \equiv (1 - \sqrt{1 - 2\alpha})/(\beta\gamma)$ , то послідовність  $\{x_k\}$ , генерована методом Ньютона

$$x_{k+1} = x_k - J(x_k)^{-1} F(x_k), \quad k = 0, 1, \dots,$$

коректно визначена, збігається до  $x^*$ , який є єдиним нулем  $F$  у  $N(x_0, r_0)$  разом з її границею. Якщо ж у цьому випадку  $\alpha < \frac{1}{2}$ , то



$x^*$  буде єдиним нулем  $F$  в  $N(x_0, r_1)$ , де  $r_1 \equiv \min \left[ r, \frac{1 + \sqrt{1 - 2\alpha}}{\beta\gamma} \right]$ , і

$$\|x_k - x^*\| \leq (2\alpha)^{2^k} \frac{\gamma}{\alpha}, \quad k = 0, 1, \dots$$

Зазначимо, що для більшості методів отримання найліпших оцінок швидкості збіжності потребує в кожному випадку окремого доведення.

## 6.6. Модифікації методу Ньютона

Метод Ньютона (6.22) передбачає на кожній ітерації обчислення якобіана  $J$  в точці  $x_k$  та розв'язування системи лінійних алгебричних рівнянь, що потребує  $O(n^3)$  арифметичних операцій. Для зменшення кількості обчислень під час розв'язування нелінійних систем часто використовують різноманітні модифікації методу Ньютона. Найпростіша з них – *модифікований метод Ньютона*, ітераційна формула якого має вигляд

$$x_{k+1} = x_k - J(x_0)^{-1} F(x_k), \quad k = 0, 1, \dots \quad (6.24)$$

У цьому методі якобіан обчислюють тільки на першій ітерації, і він є незмінним на наступних. Це дає змогу зробити прямий хід методу Гаусса ( $LU$ -розклад матриці-якобіана) лише один раз, на першій ітерації, а на наступних ітераціях виконувати тільки зворотний хід методу Гаусса (розв'язувати дві системи лінійних рівнянь з трикутними матрицями). Отже, часто можна отримати розв'язок з заданою точністю за меншу кількість арифметичних операцій, хоча теоретична швидкість збіжності методу (6.24) лише лінійна.

Зазначимо, що є методи з порядком збіжності, вищим від квадратичного. Зокрема, метод з порядком збіжності  $1 + \sqrt{2}$  має вигляд [8]

$$\begin{aligned} x_{k+1} &= x_k - J\left(\frac{x_k + y_k}{2}\right)^{-1} F(x_k), \\ y_{k+1} &= x_{k-1} - J\left(\frac{x_k + y_k}{2}\right)^{-1} F(x_{k-1}), \quad x_0 = y_0, \quad k = 0, 1, \dots \end{aligned} \quad (6.25)$$

Оскільки в обох формулах методу використано якобіан, обчислений в одній точці, то на одній ітерації треба розв'язувати фактично одну систему з двома правими частинами. Отже, одна ітерація методу (6.25) за кількістю обчислень майже така, як ітерація методу Ньютона. Тому часто метод (6.25) за кількістю обчислень ефективніший, ніж метод Ньютона.

Розглянемо ще раз рівняння вигляду

$$F(x) = x - g(x) = 0 \quad (6.26)$$

Для його розв'язування опишемо ітераційний процес [7]

$$x_{k+1} = x_k - J((1-\mu)x_k + \mu g(x_k))^{-1} F(x_k), \quad k = 0, 1, \dots \quad (6.27)$$

Цей метод є комбінацією методу простої ітерації та методу Ньютона і має квадратичну збіжність. При  $\mu = 0$  з (6.27) отримаємо класичний метод Ньютона, при  $\mu = 1$  – метод Стірлінга. На одній ітерації методу (6.27) потрібна така ж кількість обчислень, як і в методі Ньютона. Однак якщо функція  $g(x)$  задовольняє умову

$$\|g'(x)\| \leq q < 1, \quad (6.28)$$

то метод (6.27) має найшвидшу збіжність при  $\mu = 0,5$ . Ітераційна формула цього методу має вигляд

$$x_{k+1} = x_k - J\left(\frac{x_k + g(x_k)}{2}\right)^{-1} F(x_k), \quad k = 0, 1, \dots \quad (6.29)$$

Будемо називати цю модифікацію методу Ньютона *прискореним методом Ньютона*.

Методи третього і вищих порядків передбачають на кожній ітерації розв'язування кількох систем лінійних рівнянь з різними якобіанами, а то й обчислення похідних другого і вищих порядків. Такі методи досить громіздкі, і ми їх не розглядатимемо.

## 6.7. Методи зі скінченнорізницевиими похідними

Розглянемо вплив на метод Ньютона заміни аналітично заданого якобіана  $J(x)$  скінченнорізницевою апроксимацією. Можна довести таке: якщо довжина скінченно різницевого кроку вибрана належно, то зберігається квадратична збіжність методу

Ньютона. Наведемо також практичні способи вибору довжини кроку з урахуванням впливу машинної арифметики. Для більшості задач метод Ньютона з використанням аналітичних похідних і метод Ньютона з належним вибором кроку скінченних різниць фактично збігаються.

Позначимо апроксимацію  $(i, j)$ -ї компоненти  $J(x)$  за допомогою апроксимації за різницями вперед

$$(A)_{ij} = \frac{F_i(x + h e_j) - F_i(x)}{h},$$

де  $e_j$  позначає  $j$ -й одиничний вектор. Це еквівалентно апроксимації  $j$ -го стовпця  $J(x)$  за формулою

$$A_{\cdot j} = \frac{F(x + h e_j) - F(x)}{h}. \quad (6.30)$$

**Лема 6.3.** Нехай  $F : R^n \rightarrow R^n$  задовольняє умови леми 6.2 у векторній нормі  $\|\cdot\|$ , для якої  $\|e_j\| = 1$ ,  $j = 1, \dots, n$ . Нехай далі  $A \in R^{n \times n}$ , де  $A_{\cdot j}$  визначені за (6.30),  $j = 1, \dots, n$ . Тоді

$$\|A_{\cdot j} - J(x)_{\cdot j}\| \leq \frac{\gamma}{2} |h|. \quad (6.31)$$

**Теорема 6.4.** Нехай  $F$  та  $x^*$  задовольняють умови теореми 6.2, де  $\|\cdot\|$  позначає векторну  $l_1$ -норму та узгоджену з нею матричну норму. Тоді існують  $\varepsilon, h > 0$  такі, що коли  $\{h_k\}$  є послідовність дійсних чисел, причому  $0 < |h_k| \leq h$  і  $x_0 \in N(x^*, \varepsilon)$ , то послідовність  $\{x_k\}$ , задана формулами

$$(A_k)_{\cdot j} = \frac{F(x_k + h_k e_j) - F(x_k)}{h_k}, \quad j = 1, 2, \dots, n, \quad (6.32)$$

$$x_{k+1} = x_k - A_k^{-1} F(x_k), \quad k = 0, 1, \dots, \quad (6.33)$$

коректно визначена і збігається лінійно до  $x^*$ . Якщо  $\lim_{k \rightarrow \infty} h_k = 0$ , то збіжність надлінійна. Якщо існує деяка константа  $c_1$  така, що

$$\|h_k\| \leq c_1 \|x_k - x^*\|, \quad (6.34)$$

або, що еквівалентно, константа  $c_2$  така, що

$$\|h_k\| \leq c_2 \|F(x_k)\|, \quad (6.35)$$

то збіжність квадратична.

Для доведення надлінійної і квадратичної швидкості треба отримати поліпшену оцінку величин  $\|A_k - J(x_k)\|$ .

Теорема 6.4 не визначає, як точно треба вибирати довжину скінченнорізницевого кроку на практиці. Це зумовлено тим, що тут знову є два взаємно суперечливі джерела похибок, які виникають у разі обчислення скінченнорізничевих похідних. Розглянемо скінченнорізничеве співвідношення (6.32). Для зручності опустимо номер ітерації  $k$ .

За лемою 6.3 похибка апроксимації  $(J(x))_j$  за допомогою  $A_j$  дорівнює  $O(\gamma h)$  у точній арифметиці, що наводить на думку вибирати  $h$  настільки малим, наскільки це можливо. Однак за малих значень  $h$  суттєво впливають інші похибки в (6.32), пов'язані з обчисленням в арифметиці скінченної точності чисельника цієї формули. Оскільки там чисельник ділиться на  $h$ , то похибка  $\delta$  в чисельнику приводить у (6.23) до похибки  $\delta/h$ . Стосовно тієї похибки в чисельнику, яка походить з неточних значень функції та із втрати точності в разі їхнього віднімання, можна з достатньою підставою сказати, що вона становить лише малу частину від  $F(x)$ . Справді, якщо  $h$  надто мале, то  $x + he_j$  може бути настільки близьким до  $x$ , що  $F(x + he_j) \cong F(x)$ , і тому  $A_j$  може зовсім не мати достовірних розрядів або мати їх дуже мало. Отже,  $h$  потрібно вибирати так, щоб збалансувати в (6.23) похибки  $O(\gamma h)$  і  $O(F(x)/h)$ .

Якщо  $F(x)$  має  $t$  достовірних розрядів, то доцільно прагнути до того, щоб  $F(x + he_j)$  відрізнялась від  $F(x)$  правою частиною розрядів. Точніше кажучи, якщо відносна похибка в разі обчислення  $F(x)$  дорівнює  $\eta$  ( $\eta$  відображає відносний рівень шуму), то ми хотіли б мати

$$\frac{F(x + he_j) - F(x)}{F(x)} \leq \sqrt{\eta}, \quad j = 1, 2, \dots, n. \quad (6.36)$$

Без якої-небудь додаткової інформації розумний спосіб задовольнити (6.36) полягає в використанні для збурення кожної компоненти  $x_j$  своєї власної довжини кроку  $h_j = \sqrt{\eta} x_j$ , яка відповідає відносній зміні  $\sqrt{\eta}$  на величину  $x_j$ . Тоді  $A_j$  обчислюють за формулою

$$A_{\cdot j} = \frac{F(x + h e_j) - F(x)}{h_j}. \quad (6.37)$$

Для зручності ми припускали, що використовують однакову довжину кроку  $h$  для всієї скінченнорізницевої апроксимації. Однак на практиці не треба так чинити. Єдина довжина кроку може призвести до згубних наслідків, якщо компоненти вектора  $x$  сильно відрізняються за значенням.

У випадку, якщо  $F(x)$  задана простою формулою, розумно прийняти  $\eta \approx \text{macheps}$ , і тоді довжина кроку  $h_j = \sqrt{\text{macheps}} \cdot x_j$ .

Якщо  $F(x)$  обчислюють дуже довгою програмою або ітераційною процедурою, то  $\eta$  може бути значно більшим, ніж  $\text{macheps}$ . Коли  $\eta$  настільки велике, що можна очікувати неточність у скінченнорізничевих похідних (6.37), то варто скористатися замість цього апроксимацією  $J(x)$  за січними. Це стосується і випадку, коли недопустимі затрати пов'язані з  $n$  додатковими обчисленнями  $F(x)$  на кожній ітерації для отримання  $A$ .

## 6.8. Інші різницеві аналоги методу Ньютона

У розглянутих вище скінченнорізничевих апроксимаціях для похідних виникає проблема вибору кроку. Цю проблему автоматично вирішують у методі Стеффенсена, який є різницеvim аналогом методу (6.29). Метод Стеффенсена для рівняння (6.26) має вигляд

$$x_{k+1} = x_k - F(x_k, g(x_k))^{-1} F(x_k), \quad k = 0, 1, \dots, \quad (6.38)$$

де через  $F(u, v)$  позначено поділену різницю функції  $F(x)$  за вузлами  $u \neq v$ , яку визначають, наприклад, за формулою [8]

$$F(u, v)_{ij} = \frac{F(u_1, \dots, u_j, v_1, \dots, v_n) - F(u_1, \dots, u_{j-1}, v_j, \dots, v_n)}{u_j - v_j}. \quad (6.39)$$

Цей метод ефективніший, ніж метод Ньютона для рівняння (6.26), якщо справджується оцінка (6.28), хоча також має квадратичний порядок збіжності.

Для рівняння загального вигляду  $F(x) = 0$  використовують метод

$$x_{k+1} = x_k - F(x_k, x_{k-1})^{-1} F(x_k), \quad k=1, 2, \dots, \quad (6.40)$$

який є узагальненням на багатовимірний випадок відомого методу хорд. Цей метод простий у реалізації, однак його швидкість збіжності має нижчий порядок,  $\frac{1+\sqrt{5}}{2} \approx 1,618$ .

Квадратичний порядок збіжності має різницевий аналог методу Ньютона, запропонований В. Курчатовим [8]:

$$x_{k+1} = x_k - F(2x_k - x_{k-1}, x_{k-1})^{-1} F(x_k), \quad k=1, 2, \dots, \quad (6.41)$$

Однак у цьому методі потрібне додаткове обчислення функції  $F(x)$  порівняно з методом (6.40). Назвемо також як недолік методів (6.40) та (6.41) потребу задання двох початкових наближень  $x_0, x_1$ .

## 6.9. Квазіньютонівські методи для розв'язування систем нелінійних рівнянь

У попередніх розділах описано методи, у яких на кожній ітерації похідні чи їхні різницеві аналоги обчислювали заново. Тепер розглянемо дещо інший клас апроксимацій, які не потребують додаткових обчислень функції або пам'яті і які досить дієві на практиці. Зокрема, з цього класу проаналізуємо метод Бroyдена, що подібний до методу Ньютона, однак використовує певну апроксимацію замість аналітично заданого якобіана.

Розглянемо лінійну модель для функції  $F(x)$ :

$$M_+(x) = F(x_+) + A_+(x - x_+), \quad (6.42)$$

Вона задовольняє рівність  $M_+(x_+) = F(x_+)$  для довільних  $A_+ \in R^{n \times n}$ . У методі Ньютона  $A_+ = J(x_+)$ . Якщо  $J(x_+)$  недоступне, то вимогою, яка в одновимірному випадку приводила б до методу січних (хорд), є  $M_+(x) = F(x)$ , тобто

$$F(x_c) = F(x_+) + A_+(x_c - x_+)$$

або

$$A_+(x_c - x_+) = F(x_+) - F(x_-). \quad (6.43)$$

Ми називатимемо (6.43) *співвідношенням січних*, або *квазіньютонівською умовою*. Позначимо  $s_c = x_+ - x_-$  для поточного кроку і  $y_c = F(x_+) - F(x_-)$  для результату поточного кроку. Тоді співвідношення січних можна записати так:

$$A_+ s_c = y_c. \quad (6.44)$$

В одновимірному випадку  $A_+$  легко визначити з умови (6.43). Головна складність у разі узагальнення методу січних на  $n$ -вимірний випадок пов'язана з тим, що (6.44) не повністю визначає  $A_+$  при  $n > 1$ . Один зі способів отримання  $A_+$  пов'язаний з мінімізацією змін у лінійній моделі (6.42) за умови виконання рівності  $A_+ s_c = y_c$ . Відмінність між новою і старою лінійними моделями для довільного  $x \in R^n$  така:

$$\begin{aligned} M_+(x) - M_-(x) &= F(x_+) + A_+(x - x_-) - F(x_-) - A_-(x - x_-) = \\ &= F(x_+) - F(x_-) - A_-(x - x_-) + (A_+ - A_-)(x - x_-) = (A_+ - A_-)(x - x_-), \end{aligned}$$

де в останній рівності використано співвідношення січних (6.44). Мінімізацією різниці  $M_+(x) - M_-(x)$  отримують формулу [9], яка задовольняє рівність (6.44):

$$A_+ = A_- + \frac{(y_- - A_- s_-) s_-^T}{s_-^T s_-}. \quad (6.45)$$

Формула для перерахунку  $A_+$  запропонована 1965 р. С. Бroyденом, її називають *формулою перерахунку Бройдена* ("добра" формула Бройдена), або, простіше, *формулою січних*. Слово *перерахунок* означає, що тут не апроксимують повністю  $J(x_+)$ , а швидше перераховують апроксимацію  $A_-$  матриці  $J(x_-)$  в апроксимацію  $A_+$  матриці  $J(x_+)$ . Ця властивість перерахунку характерна для всіх вдалих способів апроксимації за січними. З використанням (6.45) можна отримати такий алгоритм.

*Алгоритм 6.2. Метод Бройдена*

Задані:  $F : R^n \rightarrow R^n$ ,  $x_0 \in R^n$ ,  $A_0 \in R^{n \times n}$

Для  $k = 0, 1, \dots$  виконати:

розв'язати  $A_k s_k = -F(x_k)$  відносно  $s_k$ .

$$\begin{aligned}x_{k+1} &= x_k + s_k; \\y_k &= F(x_{k+1}) - F(x_k); \\A_{k+1} &= A_k + \frac{(y_k - A_k s_k) s_k^T}{s_k^T s_k}.\end{aligned}\tag{6.46}$$

Безпосереднє використання формули (6.46) для обчислення  $A_{k+1}$  потребує на кожній ітерації розв'язування системи лінійних рівнянь (знаходження  $LU$ - чи  $QR$ -розкладу матриці  $A_{k+1}$ ). Таку реалізацію формули перерахунку Бroyдена називають формулою перерахунку в нефакторизованій формі. Факторизована форма передбачає безпосередній перерахунок множників  $Q_{k+1}$  та  $R_{k+1}$  розкладу матриці  $A_{k+1}$  через множники  $Q_k$  та  $R_k$  матриці  $A_k$ . Факторизована форма робить алгоритм ефективнішим, проте вона складніша в програмній реалізації.

Постає запитання: як вибирати початкову апроксимацію  $A_0$  матриці  $J(x_0)$ ? На практиці для забезпечення хорошого початку ітераційного процесу тут єдиний раз використовують скінченні різниці або приймають  $A_0 = J(x_0)$ .

Для послідовності  $\{x_k\}$ , визначеної формулами (6.46), Дж. Денніс і Д. Море (1974) довели надлінійну збіжність до  $x^*$  і  $F(x^*) = 0$  тоді й тільки тоді, коли

$$\lim_{k \rightarrow \infty} \frac{\|(A_k - J(x^*))s_k\|}{\|s_k\|} = 0,\tag{6.47}$$

де  $s_k = x_{k+1} - x_k$ . Пізніше Д. Гей (1979) довів  $2n$ -крокову квадратичну збіжність методу Бroyдена для загального випадку. Для довільної лінійної  $F$  метод Бroyдена дає  $x_{2n+1} = x^*$ .

Якщо відомо  $A_k^{-1}$  то цим можна скористатися, записавши формулу перерахунку Бroyдена як

$$A_{k+1}^{-1} = A_k^{-1} + \frac{(s_k - A_k^{-1} y_k) s_k^T A_k^{-1}}{s_k^T A_k^{-1} y_k},\tag{6.48}$$

що потребує  $O(n^2)$  операцій. (Цю формулу називають "поганою" формулою Бroyдена). Тоді напрям методу січних  $-A_{k+1}^{-1} F(x_{k+1})$  можна обчислити множенням матриці на вектор, що потребує



додатково  $O(n^2)$  операцій. У цьому випадку в алгоритмі 6.2 треба замінити розв'язування системи лінійних рівнянь на обчислення

$$x_{k+1} = x_k - A_k^{-1} F(x_k)$$

і обчислення  $A_{k+1}$  на обчислення  $A_{k+1}^{-1}$  за формулою (6.48). Однак ця реалізація має недолік: у цьому разі важко виявити погану зумовленість  $A_{k+1}$ . Оскільки факторизована формула перерахунку (6.46) з  $A_k$  в  $A_{k+1}$  не має цього недоліку і потребує такої ж кількості арифметичних обчислень, що й (6.48), то в програмній реалізації перевагу надають їй, а не (6.48).

Зазначимо таке: уже не гарантують, що квазіньютонівський напрям  $s_k = -A_k^{-1} F(x_k)$  буде напрямом спуску для  $f(x) = \frac{1}{2} \|F(x)\|^2$

Визначимо

$$s_k^T \nabla f(x_k) = -F(x_k)^T (A_k^{-1})^T J(x_k)^T F(x_k).$$

Якщо  $A_k = J(x_k)$ , то ця величина недодатна, що не обов'язково в протилежному випадку. Крім того, нема способу перевірити, чи є  $s_k$  напрямом спадання. Тому можливо, що крок виявиться невдалим. В алгоритмах варто передбачити відновлення  $A_k$  з використанням скінченно різницевої апроксимації  $J(x_k)$ .

## 6.10. Числова апробація методів

Для виявлення властивостей основних з розглянутих вище методів проведено значний числовий експеримент. Тестування виконували на багатьох системах нелінійних рівнянь. Методи реалізовані в середовищі MATLAB. Для всіх методів застосовували однакові критерії припинення обчислювального процесу. Розв'язки задач шукали з різних початкових наближень і з різною точністю. Наведені в табл. 6.1–6.12 результати відображають, який характер збіжності ітераційних процесів, як на кількість ітерацій впливає вибір початкового наближення та задана точність.

### Приклад 6.1:

$$\begin{cases} f_1(x, y) = 2x - \sin 0,5(x - y) = 0; \\ f_2(x, y) = 2y - \cos 0,5(x + y) = 0. \end{cases}$$

Розв'язування. Розв'язок системи шукали методом простої ітерації з точністю  $\varepsilon = 10^{-6}$

Таблиця 6.1

Результати за методом ітерації

$n$	$x_n/y_n$	$F(x_n, y_n)$
0	0,00000000000000 0,50000000000000	0,24740395925452 0,03108757828936
1	-0,12370197962726 0,48445621085532	0,05201068654705 -0,01486368627232
2	-0,14970732290079 0,49188805399148	0,01590900877948 -0,00162360247951
3	-0,15766182729053 0,49269985523124	0,00415650040598 -0,00060176178168
4	-0,15974007749352 0,49300073612208	0,00112699763710 -0,00014778689395
5	-0,16030357631207 0,49307462956905	0,30185659548132e-003 -0,04057349704267e-003
6	-0,16045450460981 0,49309491631757	0,81078567736370e-004 -0,10816249702006e-004
7	-0,16049504389368 0,49310032444242	0,21757909883291e-004 -0,02907907932448e-004
8	-0,16050592284862 0,49310177839639	0,58400898532707e-005 -0,07800829002758e-005
9	-0,16050884289355 0,49310216843784	0,15674466082305e-005 -0,02093981010809e-005
10	-0,16050962661685 0,49310227313689	0,42070023331675e-006 -0,05619980802951e-006

Таблиця 6.2

Результати за методом Ньютона

$n$	$x_n/y_n$	$F(x_n, y_n)$
0	0,00000000000000 0,50000000000000	0,24740395925452 0,03108757828936
1	-0,16157349074847 0,49477298710696	-0,00083268614150 0,00339164334075
2	-0,16051012047714 0,49310234157224	-0,30075298884658e-006 0,04546042264852e-006

Таблиця 6.3

## Результати за модифікованим методом Ньютона

$n$	$x_n / y_n$	$F(x_n, y_n)$
0	0,000000000000000	0,24740395925452
	0,500000000000000	0,03108757828936
1	-0,16157349074847	-0,00083268614150
	0,49477298710696	0,00339164334075
2	-0,16049344201748	0,30221158960286e-004
	0,49311303314888	0,23693996680674e-004
3	-0,16051012708198	-0,00709786328046e-005
	0,49310284809402	0,10998798084527e-005
4	-0,16050991066538	0,08156471609233e-007
	0,49310231758127	0,12858011322336e-007

Таблиця 6.4

## Результати за методом Бройдена ("добра" формула Бройдена)

$n$	$x_n / y_n$	$F(x_n, y_n)$
0	0,000000000000000	0,24740395925452
	0,500000000000000	0,03108757828936
1	-0,16157349074847	-0,00083268614150
	0,49477298710696	0,00339164334075
2	-0,16050025232154	0,24781941794283e-004
	0,49312350007584	0,44930518593445e-004
3	-0,16051001372505	0,02176318936176e-006
	0,49310267853269	0,75610537053628e-006

Таблиця 6.5

## Результати за методом Бройдена ("погана" формула Бройдена)

$n$	$x_n / y_n$	$F(x_n, y_n)$
0	0,000000000000000	0,24740395925452
	0,500000000000000	0,03108757828936
1	-0,16157349074847	-0,00083268614150
	0,49477298710696	0,00339164334075
2	-0,16050025232154	0,24781941794283e-004
	0,49312350007584	0,44930518593445e-004
3	-0,16051001372505	0,02176318936176e-006
	0,49310267853269	0,75610537053628e-006

Таблиця 6.6

Порівняльна таблиця кількості ітерацій

$x_0$ $y_0$	$\varepsilon$	Метод				
		Ньютона	модифікований Ньютона	Бройдена ("добрий")	Бройдена ("поганий")	простої ітерації
0	$10^{-3}$	2	2	2	2	5
0,5	$10^{-6}$	2	4	3	3	10
0	$10^{-1}$	3	7	5	5	7
5	$10^{-6}$	4	12	7	7	13
0	$10^{-3}$	5	6	7	7	6
50	$10^{-6}$	6	8	9	9	12

Таблиця 6.7

Кількість ітерацій для отримання розв'язку методом простої ітерації залежно від точності та початкового наближення

Приклад	$x_0$	$y_0$	$\varepsilon$	Кількість ітерацій
6.1	0	0,5	$10^{-3}$	5
	0	0,5	$10^{-6}$	10
	0	5	$10^{-3}$	7
	0	5	$10^{-6}$	13
	0	50	$10^{-3}$	6
	0	50	$10^{-6}$	12
6.2	0,3	0,6	$10^{-1}$	4
	0,3	0,6	$10^{-6}$	5
	3	6	$10^{-1}$	6
	3	6	$10^{-6}$	7
	30	60	$10^{-3}$	6
	30	60	$10^{-6}$	7
6.3	-0,8	0,4	$10^{-3}$	5
	-0,8	0,4	$10^{-6}$	9
	-8	4	$10^{-1}$	7
	-8	4	$10^{-6}$	11
	-80	40	$10^{-3}$	6
	-80	40	$10^{-6}$	11

**Приклад 6.2:**

$$\begin{cases} \sin(x+1) - y = 1; \\ 2x + \cos y = 2. \end{cases}$$

**Приклад 6.3:**

$$\begin{cases} \cos \frac{1}{3}(x-y) - 2y = 0; \\ \sin \frac{1}{3}(x+y) - 2x = 0. \end{cases}$$

Для тестування інших методів вибрано суттєво нелінійну систему рівнянь. Розв'язки шукали з вищою точністю  $\varepsilon = 10^{-10}$

**Приклад 6.4: (Box-3D)**

$$F_i(x) = e^{-x_i/4} - e^{-x_i/3} - x_i(e^{-x_i/4} - e^{-x_i/3});$$

$$t_i = 0,1i;$$

$$x_0 = (0, 10, 20), \quad n = 3.$$

Таблиця 6.8

**Результати для методу Ньютона**

$k$	$x_k$	$F(x_k)$
0	0,000000000000000	-10,10703897846179
	10,000000000000000	-12,80324468006400
	20,000000000000000	-12,87041011464494
1	0,73140661840952	0,00238503304962
	10,35235406498873	0,00986766305003
	1,06515277411398	0,02213617686076
2	0,97285320166996	0,00008888501503
	10,03576399920959	0,00073299075519
	1,00686111436599	0,00184796342413
3	0,99968417296274	0,00969442147003e-004
	10,00041082505432	0,08464319049106e-004
	1,00007956175671	0,21350554461175e-004
4	0,99999995670238	0,01408034799866e-008
	10,00000005602048	0,11762544271932e-008
	1,00000001087200	0,29465394568717e-008
5	1,000000000000000	0,000000000000000
	9,999999999999999	0,000000000000000
	1,000000000000000	0,000000000000000

Таблиця 6.9

## Результати для модифікованого методу Ньютона

$k$	$x_k$	$F(x_k)$
0	0,000000000000000	-10,10703897846179
	10,000000000000000	-12,80324468006400
	20,000000000000000	-12,87041011464494
1	0,73140661840952	0,00238503304962
	10,35235406498873	0,00986766305003
	1,06515277411398	0,02213617686076
2	0,88255396506644	0,00136061109407
	10,14173399449066	0,00477899890502
	1,02701574685707	0,00997098887793
3	0,94450106397030	0,00056912503035
	10,06792550837326	0,00211023211759
	1,01295624577085	0,00448874795160
4	0,97301934403361	0,00026717891037
	10,03308679400520	0,00100469123042
	1,00631820740130	0,00214632743399
5	0,98671594855834	0,00012936540977
	10,01630589514298	0,00048974301541
	1,00311531930616	0,00104838025205
6	0,99342015210592	0,06356486554415e-003
	10,00808014854500	0,24141256274979e-003
	1,00154412884526	0,51728439153143e-003
----	-----	-----
24	0,99999997662342	0,02240788665731e-008
	10,00000002871859	0,08536675810689e-008
	1,00000000548948	0,18308820015633e-008
28	0,99999999856007	0,01380262570905e-009
	10,00000000176899	0,05258338209302e-009
	1,00000000033814	0,11277712097524e-009
29	0,99999999928265	0,06876277325318e-010
	10,00000000088127	0,26196378399845e-010
	1,00000000016845	0,56183835361878e-010

## Результати для методу Бройдена ("добра" формула Бройдена)

$k$	$x_k$	$F(x_k)$
0	0,000000000000000	-10,10703897846179
	10,000000000000000	-12,80324468006400
	20,000000000000000	-12,87041011464494
1	0,73140661840948	0,00238503304964
	10,35235406498895	0,00986766305004
	1,06515277411398	0,02213617686078
2	0,88287380135203	0,00135849683538
	10,14128831057684	0,00476880728243
	1,02693504676268	0,00994674728495
3	0,98100993018913	0,00011004647975
	10,02442083992161	0,00058392806481
	1,00466922025877	0,00136934487098
4	0,99990281883009	-0,42982964209015e-005
	10,00015956870613	-0,39008354660863e-005
	1,00003531340191	-0,04209911764752e-005
5	1,00001621042363	-0,01627426496897e-005
	9,99998015531846	-0,06027182691826e-005
	0,99999621184190	-0,12813525788191e-005
6	1,00000220223403	-0,02119376474319e-006
	9,99999729506027	-0,08053903910454e-006
	0,99999948304749	-0,17260740070579e-006
7	0,99999999464401	0,05149036752528e-009
	10,00000000657836	0,19581258836610e-009
	1,00000000125735	0,41973213793511e-009
8	0,99999999999989	0,09992007221626e-014
	10,00000000000013	0,39968028886506e-014
	1,00000000000003	0,85487172896137e-014

**Приклад 6.5:** Функція Розенброка [33]:

$$F_{2i-1}(x) = 10(x_{2i} - x_{2i-1}^2); \quad F_{2i}(x) = 1 - x_{2i-1}, \quad i = 1, 2, \dots, \frac{n}{2};$$

$$x_0 = (-1, 2; 1; \dots; -1, 2; 1); \quad x^* = (1; 1; \dots; 1; 1); \quad n = 4, \quad F = 0.$$

**Приклад 6.6:** Функція Вуда [9]:

$$F_1(x) = 10(x_2 - x_1^2); \quad F_2(x) = 1 - x_1; \quad F_3(x) = \sqrt{90}(x_4 - x_3^2);$$

$$F_4(x) = 1 - x_1, \quad x_0 = (-3; -1; -3; -1); \quad x^* = (1; 1; 1; 1); \quad F = 0.$$

Таблиця 6.11

Результати для методу Бройдена ("погана" формула Бройдена)

$k$	$x_k$	$F(x_k)$
0	0,00000000000000	-10,10703897846179
	10,00000000000000	-12,80324468006400
	20,00000000000000	-12,87041011464494
1	0,73140661840952	0,00238503304962
	10,35235406498873	0,00986766305003
	1,06515277411398	0,02213617686076
2	0,88287380135204	0,00135849683538
	10,14128831057685	0,00476880728243
	1,02693504676268	0,00994674728495
3	0,98100993018919	0,00011004647975
	10,02442083992152	0,00058392806481
	1,00466922025875	0,00136934487098
4	0,99990281883005	-0,42982964204574e-005
	10,00015956870620	-0,39008354653092e-005
	1,00003531340192	-0,04209911742548e-005
5	1,00001621042366	-0,01627426498008e-005
	9,99998015531839	-0,06027182696267e-005
	0,99999621184189	-0,12813525807065e-005
6	1,00000220223401	-0,02119376463217e-006
	9,99999729506031	-0,08053903899352e-006
	0,99999948304750	-0,17260740003966e-006
7	0,99999999464403	0,05149003445837e-009
	10,00000000657831	0,19581192223228e-009
	1,00000000125734	0,41973069464518e-009
8	0,99999999999989	0,11102230246252e-014
	10,00000000000013	0,42188474935756e-014
	1,00000000000003	0,91038288019263e-014

Таблиця 6.12

Порівняння кількості ітерацій для функцій Розенброка, Вох-3D і Вуда

Функція	$n$	Метод				
		Ньютона	модифікований Ньютона	Бройдена ("добра" формула)	Бройдена ("погана" формула)	градієнтний
Розенброка	2	2	2	3	3	795
	4	2	2	3	3	1131
	8	2	2	3	3	1119
Вох-3D	3	5	29	8	8	4583
Вуда	2	2	2	3	3	924
	4	2	2	3	3	940
	8	2	2	3	3	1073



## 6.11. Тестові завдання

Наведемо низку поширених тестових завдань, які використовуються для апробації методів розв'язування нелінійних систем рівнянь та порівняння їх між собою.

**Приклад 6.7.** Функція Ковалика та Осборна [33]:

$$\begin{aligned}F_1(x) &= y_1 - \frac{x_1(u_1^2 + u_1x_2)}{u_1^2 + u_1x_3 + x_4}; \\y &= (0,1957; 0,1947; 0,1735; 0,1600); \\u &= (4,0000; 2,0000; 1,0000; 0,5000); \\x_0 &= (0,25; 0,39; 0,415; 0,39); \\x^* &= (0,1995; 0,2085; 0,2171; 0,2954); \quad F = 0.\end{aligned}$$

**Приклад 6.8.** Вох-3D функція [9]:

$$\begin{aligned}F_i(x) &= e^{-t_1t_i} - e^{-t_1t_2} - x_3(e^{-t_1} - e^{-10t_1}); \\t_1 &= 0,1i, \quad x_0 = (0; 10; 20), \quad x^* = (1; 10; 1); \quad F = 0.\end{aligned}$$

**Приклад 6.9.** Функція [33]:

$$\begin{aligned}F_1(x) &= x_1 - 13 + x_2(x_2(5 - x_2) - 2); \\F_2(x) &= x_1 - 29 + x_2(x_2(x_2 + 1) - 14); \\x_0 &= (15; -2), \quad x^* = (5; 4), \quad F = 0.\end{aligned}$$

**Приклад 6.10.** З необхідної умови мінімуму функції

$$\Phi(x) = x_1^2 + 100(x_1^2 - x_1 - x_2)^2$$

отримаємо таку систему нелінійних рівнянь:

$$\begin{aligned}F_1(x) &= 2x_1 + 200(x_1^2 - x_1 - x_2)(2x_1 - 1) = 0; \\F_2(x) &= -200(x_1^2 - x_1 - x_2) = 0; \\x_0 &= (1; 1); \quad x^* = (0; 0).\end{aligned}$$

**Приклад 6.11.** “Погана” функція Пауелла [33]:

$$F_1(x) = 10^4 x_1 x_2 - 1; \quad F_2(x) = e^{-x_1} + e^{-x_2} - 1,0001;$$

$$x_0 = (0; 1); \quad x^* = (1,098...10^{-5}; 9,106...); \quad F = 0.$$

**Приклад 6.12.** Сингулярна функція Пауелла [33]:

$$F_1(x) = x_1 + 10x_2, \quad F_2(x) = \sqrt{5}(x_3 - x_4);$$

$$F_3(x) = (x_2 - 2x_1)^2, \quad F_4(x) = \sqrt{10}(x_1 - x_4)^2;$$

$$x_0 = (3; -1; 0; 1), \quad F = 0.$$

**Приклад 6.13.** Тригонометрична функція [33]:

$$F_i(x) = n - \sum_{j=1}^n \cos x_j + i(1 - \cos x_i) - \sin x_i;$$

$$x_0 = (\frac{1}{n}; \frac{1}{n}; \dots; \frac{1}{n}), \quad n = 4, \quad F = 0.$$

**Приклад 6.14.** Функція скінченнорізницевої крайової задачі [33]:

$$F_i(x) = 2x_i - x_{i-1} - x_{i+1} + h^2(x_i + t_i + 1)^3/2;$$

$$h = 1/(n+1), \quad t_i = ih, \quad x_0 = x_{n+1} = 0;$$

$$x_0 = (\xi_i), \quad \xi_i = t_i(t_i - 1), \quad F = 0.$$

**Приклад 6.15.** Тридіагональна функція Бройдена [33]:

$$F_i(x) = (3 - 2x_i)x_i - x_{i-1} - 2x_{i+1} + 1, \quad x_0 = x_{n+1} = 0;$$

$$x_0 = (-1; -1; \dots; -1), \quad F = 0, \quad n = 4, 8, 16, 32, 64, 128.$$

**Приклад 6.16.** Об'єднана функція Бройдена [33]:

$$F_i(x) = x_i(2 + 5x_i^2) + 1 - \sum_{j \in J_i} x_j(1 + x_j);$$

$$J_i = \{j: j \neq i, \max(1, i - m_i) \leq j \leq \min(n, i + m_n)\};$$

$$m_i = 5, \quad m_n = 1, \quad x_0 = (-1; -1; \dots; -1), \quad F = 0, \quad n = 4.$$

## Завдання до розділу 6

6.1. Ітераційними методами з точністю  $\varepsilon = 10^{-4}$  розв'язати такі системи. Початкове наближення знайти графічно.

$$1) \quad \begin{cases} \sin(x+1) - y - 1 = 0; \\ 2x + \cos(y) - 2 = 0. \end{cases}$$

$$2) \quad \begin{cases} \sin(x) - y - 1,32 = 0; \\ \cos(y) - x + 0,85 = 0. \end{cases}$$

$$3) \quad \begin{cases} \sin(x-0,6) - y - 1,6 = 0; \\ 3x - \cos(y) - 0,9 = 0. \end{cases}$$

$$4) \quad \begin{cases} \cos(x-1) + y - 1 = 0; \\ \sin(y) + 2x - 1,6 = 0. \end{cases}$$

$$5) \quad \begin{cases} x - \cos(y) - 2 = 0; \\ \cos(x-1) + y - 0,8 = 0. \end{cases}$$

$$6) \quad \begin{cases} y - 0,5x^2 + x - 0,5 = 0; \\ 2x + y - \frac{1}{6}y^3 - 1,6 = 0. \end{cases}$$

$$7) \quad \begin{cases} \sin(x-y) - xy + 1 = 0; \\ x^2 - y^2 - 0,75 = 0. \end{cases}$$

$$8) \quad \begin{cases} x^3 - y^2 - 1 = 0; \\ xy^3 - y - 4 = 0. \end{cases}$$

$$9) \quad \begin{cases} \cos(0,4y + x^2) + x^2 + y^2 - 1,6 = 0; \\ 1,5x^2 - 2y^2 - 1 = 0. \end{cases}$$

$$10) \quad \begin{cases} \sin(x+y) - 1,3x - 0,1 = 0; \\ x^2 + y^2 - 1 = 0. \end{cases}$$

$$11) \quad \begin{cases} 2x^2 + 3y^2 - 6y - 4 = 0; \\ x^2 - 3y^2 + 4x - 2 = 0. \end{cases}$$

$$12) \quad \begin{cases} xy + \cos(x) = 0; \\ x + y - \sin(x) = 0. \end{cases}$$

- 13) 
$$\begin{cases} \sin(x+y) - 1,1x + 0,2 = 0; \\ x^2 + y^2 - 1 = 0. \end{cases}$$
- 14) 
$$\begin{cases} \lg(xy + 0,4) - x^2 = 0; \\ 0,5x^2 + 2y^2 - 1 = 0, \quad x > 0, \quad y > 0. \end{cases}$$
- 15) 
$$\begin{cases} e^{xy} - x^2 + y - 1,5 = 0; \\ (x + 0,5)^2 + y^2 - 1 = 0. \end{cases}$$
- 16) 
$$\begin{cases} 2x^2 - xy - 5x + 1 = 0; \\ x + 3\lg(x) - y^2 = 0. \end{cases}$$
- 17) 
$$\begin{cases} x^2y^2 - 3x^3 - 6y^3 + 8 = 0; \\ x^4 - 9y + 2 = 0. \end{cases}$$
- 18) 
$$\begin{cases} \sin(2x - y) - 1,2x - 0,4 = 0; \\ 0,8x^2 + 1,5y^2 - 1 = 0. \end{cases}$$
- 19) 
$$\begin{cases} \sin(x) - y - 1,32 = 0; \\ \cos(y) - x + 0,85 = 0. \end{cases}$$
- 20) 
$$\begin{cases} y(x-1) - 1 = 0; \\ x^2 - y^2 - 1 = 0. \end{cases}$$
- 21) 
$$\begin{cases} 2x^2 + y^2 - 1 = 0; \\ x^3 + 6x^2y - 1 = 0. \end{cases}$$
- 22) 
$$\begin{cases} 2x - y - 6\lg(x) - 3 = 0; \\ 15x - 10y - 60\lg(y) - 6 = 0. \end{cases}$$
- 23) 
$$\begin{cases} 5x - 6y + 20\lg(x) + 16 = 0; \\ 2x + y - 10\lg(x) - 4 = 0. \end{cases}$$
- 24) 
$$\begin{cases} x - y - 6\lg(x) - 1 = 0; \\ x - 3y - 6\lg(y) - 2 = 0. \end{cases}$$

## Розділ 7

### ПОЛІНОМІАЛЬНЕ ІНТЕРПОЛЮВАННЯ

Однією з головних задач чисельного аналізу є задача про інтерполювання функцій. У математиці під інтерполюванням розуміють задачу знаходження невідомого значення певної величини за декількома відомими її значеннями. Крім того, може виявитися, що функція задана формулою, і обчислення її значень за цією формулою дуже трудомістке, тому бажано мати для функції простішу формулу, яка б давала змогу відшукати наближене значення функції з потрібною точністю в будь-якій точці. У підсумку виникає така математична задача.

Нехай на відрізку  $[a, b]$  задана функція  $y = f(x)$  такими  $n + 1$  значеннями:

$$y_0 = f(x_0), \quad y_1 = f(x_1), \quad \dots, \quad y_n = f(x_n)$$

в точках  $x_0 = a < x_1 < x_2 < \dots < x_n = b$ , які називають вузлами інтерполяції. Треба побудувати інтерполянту – функцію  $F(x)$ , яка збігається з функцією  $f(x)$  у вузлах інтерполяції

$$F(x_i) = y_i, \quad i = \overline{0, n}.$$

Головна мета інтерполяції – отримати швидкий алгоритм обчислення значень функції  $F(x)$  для значень  $x$ , що не містяться в таблиці даних.

Якщо аргумент  $x$ , для якого обчислюють наближене значення функції, не належить заданому відрізку  $[a, b]$ , то таку задачу називають *екстраполяцією*. Геометрично задача інтерполювання для функції однієї змінної  $y = f(x)$  означає побудову кривої, що проходить через точки площини  $(x_i, y_i)$ ,  $y_i = f(x_i)$ ,  $i = \overline{0, n}$ . Таких кривих  $F(x)$  може бути безліч, тому задача в такому формулюванні досить невизначена.

Ця задача стає однозначною, якщо за інтерполяційну функцію вибрати багаточлен  $F_n(x)$  степеня не вище  $n$ , для якого  $F_n(x_i) = y_i, i = \overline{0, n}$ .

Багаточлен  $F_n(x)$ , що задовольняє ці умови, називають *інтерполяційним багаточленом*, а відповідні формули – *інтерполяційними формулами*.

Якщо інтерпольована функція  $f(x)$  періодична, то  $F(x)$  вибирають з класу тригонометричних функцій.

У разі інтерполювання виникають такі задачі:

- 1) вибір найзручнішого способу побудови інтерполяційної функції для кожного конкретного випадку;
- 2) оцінка похибки в разі заміни  $f(x)$  інтерполяційною функцією  $F(x)$  на відрізьку  $[a, b]$ ;
- 3) оптимальний вибір вузлів інтерполяції для отримання мінімальної похибки.

## 7.1. Інтерполяційний багаточлен Лагранжа

Найзагальнішою формулою інтерполювання є інтерполяційна формула Лагранжа.

Задачу формулюють так. Треба побудувати поліном  $L_n(x)$  степеня  $n$ , який в  $n+1$  заданих точках – вузлах інтерполяції  $x_0, x_1, \dots, x_n$  – набуває наперед заданого значення  $y_0, y_1, \dots, y_n$ .

Багаточлен Лагранжа має вигляд

$$L_n(x) = \sum_{k=0}^n \frac{\omega_{n+1}(x)}{(x-x_k)\omega'_{n+1}(x_k)} f(x_k), \quad (7.1)$$

де

$$\omega_{n+1}(x) = \prod_{k=0}^n (x - x_k) \\ \omega'_{n+1}(x_k) = (x_k - x_0) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_n). \quad (7.2)$$

Різницю  $r_n(x) = f(x) - L_n(x)$  називають похибкою інтерполювання, або залишковим членом. У вузлах інтерполювання похибка перетворюється в нуль, в інших точках відмінна від нуля.

Похибку багаточлена Лагранжа виражають формулою

$$r_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(x), \text{ де } \xi \in [a, b]. \quad (7.3)$$

Якщо  $\max |f^{(n+1)}(x)| = M_{n+1}$ ,  $x \in [a, b]$ , то для похибки справджується оцінка

$$|r_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |\omega_{n+1}(x)|. \quad (7.4)$$

Інтерполяційний багаточлен Лагранжа зручно записувати у вигляді

$$L_n(x) = \sum_{k=0}^n f(x_k) \frac{(x-x_0)(x-x_1)\dots(x-x_{k-1})(x-x_{k+1})\dots(x-x_n)}{(x_k-x_0)(x_k-x_1)\dots(x_k-x_{k-1})(x_k-x_{k+1})\dots(x_k-x_n)}. \quad (7.5)$$

**П р и к л а д 7.1.** Побудувати багаточлен третього порядку, що набуває значень, заданих такою таблицею:

$x$	1	2	3	5
$y$	1	5	17	89

*Розв'язування.* Скористаємося формулою (7.5):

$$\begin{aligned} L_3(x) &= 1 \frac{(x-2)(x-3)(x-5)}{(1-2)(1-3)(1-5)} + 5 \frac{(x-1)(x-3)(x-5)}{(2-1)(2-3)(2-5)} + \\ &+ 17 \frac{(x-1)(x-2)(x-5)}{(3-1)(3-2)(3-5)} + 89 \frac{(x-1)(x-2)(x-3)}{(5-1)(5-2)(5-3)} = \\ &= -\frac{1}{8}(x^3 - 10x^2 + 31x - 30) + \frac{5}{3}(x^3 - 9x^2 + 23x - 15) - \\ &- \frac{17}{4}(x^3 - 8x^2 + 17x - 10) + \frac{89}{24}(x^3 - 6x^2 + 11x - 6) = \\ &= x^3 \left(-\frac{1}{8} + \frac{5}{3} - \frac{17}{4} + \frac{89}{24}\right) + x^2 \left(\frac{5}{4} - \frac{45}{3} + 34 - \frac{89}{4}\right) + \\ &+ x \left(-\frac{31}{8} + \frac{115}{3} - \frac{289}{4} + \frac{979}{24}\right) + \left(\frac{15}{4} - 25 + \frac{170}{4} - \frac{89}{4}\right) = \\ &= x^3 - 2x^2 + 3x - 1. \end{aligned}$$

**Приклад 7.2.** З якою точністю можна обчислити  $\sqrt{125}$  за допомогою інтерполяційної формули Лагранжа для функції  $y = \sqrt{x}$ , вибравши вузли інтерполювання 100, 121, 144 ?

*Розв'язування.* Скористаємося формулою (7.4). Знайдемо похідні для функції  $f(x) = \sqrt{x}$

$$f'(x) = \frac{1}{2\sqrt{x}}; \quad f''(x) = -\frac{1}{4\sqrt{x^3}}; \quad f'''(x) = \frac{3}{8\sqrt{x^5}}.$$

$$\text{Тоді } M_3 = \max |f'''(x)| = \frac{3}{8} \frac{1}{\sqrt{100^5}} = \frac{3}{8} 10^{-5}$$

Отже,

$$|r_2| \leq \frac{3}{8} 10^{-5} \cdot \frac{1}{3!} |(125-100)(125-121)(125-144)| \approx 1,2 \cdot 10^{-3}$$

Кількість арифметичних дій під час обчислення за формулою Лагранжа пропорційна до  $n^2$ . Якість інтерполяції за Лагранжем залежить від двох чинників: вибору вузлів  $x_i$  та властивостей функції  $f(x)$ .

Як видно з (7.1), останні два чинники відокремлені: багаточлени  $\omega_{n+1}(x)$  залежать лише від вузлів, а властивості функції враховує множник  $f(x_i)$ . Тому завдяки вибору вузлів інтерполювання можна зменшити похибку в (7.4).

Обчислювальне застосування (7.1) має суттєвий недолік, оскільки треба наперед визначити  $n+1$  вузлів інтерполювання, необхідних для досягнення прийнятної точності.

Якщо ж потрібно додавати нові вузли інтерполювання, то обчислення треба виконувати заново. Тобто у формулі (7.1) не тільки з'являються нові доданки, а й необхідно перераховувати всі раніше знайдені члени суми, бо в них з'являються нові множники.

## 7.2. Схема Ейткена

У випадку, коли нас цікавить значення інтерполяційного багаточлена Лагранжа в деякій точці  $x$ , а не загальний його вигляд, зручно користуватись інтерполяційною схемою Ейткена.

Обчислювальний алгоритм має вигляд, як у табл. 7.1, де



$$L_{012\dots n}(x) = \frac{1}{x_n - x_0} \begin{vmatrix} L_{012\dots (n-1)}(x) & x_0 - x \\ L_{123\dots n}(x) & x_n - x \end{vmatrix}$$

– інтерполяційний багаточлен Лагранжа за вузлами  $x_0, x_1, \dots, x_n$ . Кожен з  $L_{012\dots k}(x)$  отримують з  $L_{012\dots (k-1)}(x)$  та  $L_{123\dots k}(x)$  перехресним множенням та діленням. Із застосуванням схеми Ейткена поступово можна залучати щораз нові значення вузлів  $x_k$  доти, доки обчислення не засвідчать, що точність уже не зростає.

Таблиця 7.1

Схема Ейткена

$x_k$	$f_k$	$x_k - x$	$L_{k-1,k}$	$L_{k-2,k-1,k}$	$L_{k-3,k-2,k-1,k}$	$L_{k-4,k-3,k-2,k-1,k}$
$x_0$	$f_0$	$x_0 - x$				
$x_1$	$f_1$	$x_1 - x$	$L_{01}(x)$			
$x_2$	$f_2$	$x_2 - x$	$L_{12}(x)$	$L_{012}(x)$		
$x_3$	$f_3$	$x_3 - x$	$L_{23}(x)$	$L_{123}(x)$	$L_{0123}(x)$	
$x_4$	$f_4$	$x_4 - x$	$L_{34}(x)$	$L_{234}(x)$	$L_{1234}(x)$	$L_{01234}(x)$
$x_5$	$f_5$	$x_5 - x$	$L_{45}(x)$	$L_{345}(x)$	$L_{2345}(x)$	$L_{12345}(x)$

**П р и к л а д 7.3.** Користуючись таблицею значень  $\sin x$  для  $x \in [0,65; 0,7]$  з кроком  $h = 0,01$ , знайти наближене значення  $\sin x$  при  $x = 0,676$ .

*Розв'язування.* Обчислення наведено в табл. 7.2.

Таблиця 7.2

Схема Ейткена для функції  $\sin x$

$x_k$	$f_k$	$x_k - x$	$L_{k-1,k}$	$L_{k-2,k-1,k}$	$L_{k-3,k-2,k-1,k}$	$L_{k-4,k-3,k-2,k-1,k}$
0,65	0,6051864	-0,026				
0,66	0,6131169	-0,016	0,6528056			
0,67	0,6209860	-0,006	0,6257075	0,6256780		
0,68	0,6287930	0,004	0,6256702	0,6256777	0,6256777	
0,69	0,6365372	0,014	0,6256954	0,6256778	0,6256777	0,6256777
0,70	0,6442177	0,024	0,6257845	0,6256775	0,6256777	0,6256777

Отже,  $\sin x = 0,625678$ .

### 7.3. Інтерполяційний багаточлен Ньютона

Розглянемо інше зображення формули Лагранжа, яке допускає послідовне уточнення результатів обчислень у разі додавання нових вузлів. Крім того, часто не потрібно зазначати степінь інтерполювання. Це інтерполяційний багаточлен Ньютона

$$P(x) = f(x_0) + (x - x_0)f(x_0, x_1) + (x - x_0)(x - x_1)f(x_0, x_1, x_2) + \dots + (x - x_0)(x - x_1)\dots(x - x_{n-1})f(x_0, x_1, \dots, x_n) + \dots, \quad (7.6)$$

де  $f(x_0, x_1, \dots, x_k)$  – поділена різниця  $k$ -го порядку. Для неї відома рекурентна формула

$$f(x_{j-1}, x_j, \dots, x_{i+k}) = \frac{f(x_i, x_{i+1}, \dots, x_{i+k}) - f(x_{j-1}, x_j, \dots, x_{i+k-1})}{x_{i+k} - x_{j-1}} = \sum_{j=i}^{i+k} \frac{f(x_j)}{\omega'(x_j)} \quad (7.7)$$

$$\omega(x) = (x - x_i)(x - x_{i+1})\dots(x - x_{i+k}). \quad (7.8)$$

Багаточлен

$$P_n(x) = f(x_0) + (x - x_0)f(x_0, x_1) + (x - x_0)(x - x_1)f(x_0, x_1, x_2) + \dots + (x - x_0)(x - x_1)\dots(x - x_{n-1})f(x_0, x_1, \dots, x_n) \quad (7.9)$$

називають *інтерполяційним багаточленом Ньютона* інтерполювання вперед для нерівновіддалених вузлів. Його зазвичай використовують для наближення функції біля початкового вузла. Залишковий член має вигляд

$$R_{n+1}(x) = (x - x_0)\dots(x - x_n)f(x_0, x_1, \dots, x_n). \quad (7.10)$$

Якщо ж відома похідна функції, то

$$f(x_0, x_1, \dots, x_n) = \frac{f^{(n+1)}(\xi)}{(n+1)!}, \quad (7.11)$$

де  $\xi$  – деяка точка, яка належить найменшому проміжку, що містить усі точки  $x, x_0, \dots, x_n$ .

На відміну від полінома Лагранжа, в якого кожний член залежить від усіх вузлів інтерполювання, будь-який  $k$ -й член багаточлена Ньютона залежить лише від  $k$  перших вузлів інтерполювання і від значення функції в цих вузлах.

Якщо вибрати порядок вузлів  $x_n, x_{n-1}, \dots, x_0$ , то одержимо інтерполяційний багаточлен Ньютона інтерполювання назад для нерівновіддалених вузлів:

$$P_n(x) = f(x_n) + (x - x_n)f(x_n, x_{n-1}) + \\ + (x - x_n)(x - x_{n-1})f(x_n, x_{n-1}, x_{n-2}) + \\ + (x - x_n)(x - x_{n-1}) \dots (x - x_1)f(x_n, x_{n-1}, \dots, x_0). \quad (7.12)$$

У випадку рівновіддалених вузлів  $x_i = x_{i-1} + h$  у формулах (7.9) та (7.12) поділені різниці замінюють скінченними різницями

$$f(x_k, x_{k-1}, \dots, x_0) = \frac{\Delta^k y_0}{k! h^k}. \quad (7.13)$$

Одержимо, відповідно, інтерполяційну формулу Ньютона інтерполювання вперед для рівновіддалених вузлів

$$P_n(x) = y_0 + \frac{t}{1!} \Delta y_0 + \frac{t(t-1)}{2!} \Delta^2 y_0 + \frac{t(t-1)(t-2)}{3!} \Delta^3 y_0 + \dots + \\ + \frac{t(t-1) \dots (t-k+1)}{k!} \Delta^k y_0; \quad (7.14)$$

$$R_n = h^{k+1} \frac{t(t-1) \dots (t-k)}{(k+1)!} f^{(k+1)}(\xi); \quad t = \frac{x - x_0}{h};$$

та інтерполяційну формулу Ньютона інтерполювання назад для рівновіддалених вузлів

$$P_n(x) = y_n + \frac{t}{1!} \Delta y_{n-1} + \frac{t(t+1)}{2!} \Delta^2 y_{n-2} + \frac{t(t+1)(t+2)}{3!} \Delta^3 y_{n-3} + \dots + \\ + \frac{t(t+1) \dots (t+k-1)}{k!} \Delta^k y_{n-k}; \quad (7.15)$$

$$R_n = h^{k+1} \frac{t(t+1) \dots (t+k)}{(k+1)!} f^{(k+1)}(\xi).$$

Зокрема, для  $n=1$ :  $f(x) \approx y_0 + t \Delta y_0$  – лінійна інтерполяція; для

$n=2$ :  $f(x) \approx y_0 + t \Delta y_0 + \frac{t(t-1)}{2} \Delta^2 y_0$  – квадратична інтерполяція.

Формулу Ньютона інтерполювання вперед використовують у разі інтерполювання функції для значень  $x$ , близьких до початкового вузла  $x_0$ , а інтерполювання назад – до  $x_n$ .

З оцінок похибки багаточлена Ньютона видно таке: якщо перейти від таблиць з великим кроком, до таблиць з дрібним кроком, то похибка методу буде спадати як  $h^{n+1}$ . Тому говорять, що багаточлен Ньютона має похибку  $O(h^{n+1})$  і забезпечує  $(n+1)$ -й порядок точності інтерполяції.

Інтерполяцію застосовують у багатьох задачах, а не лише для обчислення табульованої функції за будь-яких значень аргументу.

За допомогою поділених різниць контролюють точність таблиць. Для цього складають таблиці поділених різниць різних порядків для сусідніх вузлів і аналізують їхню поведінку. Якщо, наприклад, другі скінченні різниці змінюються хаотично, то лінійну інтерполяцію застосовувати не можна.

Інтерполяцію використовують для субтабулювання – згущення таблиць. Алгоритми безпосереднього обчислення багатьох функцій досить складні. Тому в разі математичного табулювання функцію безпосередньо обчислюють у невеликій кількості вузлів, тобто з великим кроком. Потім за допомогою інтерполяції високого порядку точності сітку вузлів згущують і складають детальну таблицю. Крок цієї таблиці вибирають таким, щоб найпростіша інтерполяція (двоточкова) забезпечила потрібну точність.

У випадку розрахунків на комп'ютерах задання детальних таблиць не вигідне, бо вони займають багато місця в пам'яті, а час пошуку стає більшим від часу виконання алгебричних дій. Вигідніші таблиці з великим кроком, хоч у цьому разі потрібні складніші й точніші способи інтерполювання.

**П р и к л а д 7.4.** 3 тисячі народжених до віку 20, 30, 40, 50, 60 років доживають, відповідно, 800, 770, 720, 640, 520 людей. Скільки людей із тисячі доживають до 25 і 55 років?

*Розв'язування.* Спочатку відшукаємо значення  $f(25)$ . Оскільки  $20 < 25 < 30$ , то застосуємо інтерполяційну формулу Ньютона інтерполювання вперед (7.14):

$$t = \frac{x - x_0}{h} = \frac{25 - 20}{10} = 0,5.$$

Побудуємо таблицю скінченних різниць:

$i$	$x_i$	$y_i$	$\Delta y$	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
0	20	800	-30			
1	30	770	-50	-20	-10	
2	40	720	-80	-30	-10	0
3	50	640	-120	-40		
4	60	520				

Тут  $\Delta^k y_i = \Delta^{k-1} y_i - \Delta^{k-1} y_{i-1}$ ,

Використаємо верхній рядок таблиці:

$$f(25) = 800 + 0,5 \cdot (-30) + \frac{0,5 \cdot (-0,5)}{2} (-20) + \frac{0,5 \cdot (-0,5) \cdot (-1,5)}{6} (-10) = \\ = 800 - 15 + 2,5 - 0,625 = 786,875 \approx 787.$$

Для знаходження значення функції в точці  $x = 55$  скористаємося інтерполяційним багаточленом Ньютона інтерполювання назад (7.15) та нижніми значеннями в кожному стовпці таблиці скінченних різниць:

$$t = \frac{x - x_n}{h} = \frac{55 - 60}{10} = -0,5;$$

$$f(55) = 520 - 0,5 \cdot (-120) + \frac{(-0,5) \cdot 0,5}{2} (-40) + \frac{(-0,5) \cdot 0,5 \cdot 1,5}{6} (-10) = \\ = 520 + 60 + 5 + 0,625 = 585,625 \approx 586.$$

**П р и к л а д 7.5.** З використанням таблиці значень функції  $y = \sin x$  для  $x = 30^\circ, 35^\circ, 40^\circ, 45^\circ, 50^\circ, 55^\circ$  відшукати значення  $\sin 54^\circ$  і  $\sin 56^\circ$ . Визначити точність обчислень.

*Розв'язування.* Побудуємо таблицю скінченних різниць:

$n$	$x$	$y$	$\Delta y$	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
0	$30^\circ$	0,5	0,0736			
1	$35^\circ$	0,5735	0,0692	-0,0044	-0,0005	
2	$40^\circ$	0,6428	0,0643	-0,0049	-0,0005	0
3	$45^\circ$	0,7071	0,0589	-0,0054	-0,0004	0,0001
4	$50^\circ$	0,7660	0,0532	-0,0057		
5	$55^\circ$	0,8192				

Використаємо інтерполяційний багаточлен Ньютона інтерполювання назад (7.15). Обчислимо  $t = \frac{54^\circ - 55^\circ}{5^\circ} = -0,2$ . Тоді

$$\begin{aligned}\sin 54^\circ &\approx 0,8192 + (-0,2) \cdot 0,0532 - \frac{(-0,2) \cdot 0,8}{2} 0,0057 - \\ &- \frac{(-0,2) \cdot 0,8 \cdot 1,8}{6} 0,0004 \approx 0,809016.\end{aligned}$$

Залишок визначимо за другою з формул (7.15)

$$R_3(x) = h^4 \frac{t(t+1)(t+2)(t+3)}{4!} f^{(IV)}(\xi).$$

У нашому прикладі  $h = 5^\circ \approx 0,0873$ ;  $f^{(IV)}(\xi) = \sin \xi \leq 1$ . Тоді

$$R_3(54^\circ) \leq 0,0873^4 \frac{0,2 \cdot 0,8 \cdot 1,8 \cdot 2,8}{24} \approx 0,2 \cdot 10^{-5}$$

Отже,  $\sin 54^\circ \approx 0,8090$ .

Аналогічно знайдемо  $\sin 56^\circ$ :

$$t = \frac{56^\circ - 55^\circ}{5^\circ} = 0,2;$$

$$\begin{aligned}\sin 56^\circ &\approx 0,8192 + 0,2 \cdot 0,0532 - \frac{0,2 \cdot 1,2}{2} 0,007 - \frac{0,2 \cdot 1,2 \cdot 2,2}{6} 0,0004 \approx \\ &\approx 0,829038;\end{aligned}$$

$$|R_3(56^\circ)| \leq \frac{0,0873^4 \cdot 0,2 \cdot 1,2 \cdot 2,2 \cdot 3,2}{24} \approx 0,4 \cdot 10^{-5}$$

Отже,  $\sin 56^\circ \approx 0,8290$ .

## 7.4. Інтерполяційний багаточлен Ерміта

Нехай табульована не лише функція, а також її похідні до деякого порядку. Тоді можна домагатися, щоб у вузлах інтерполяції збігалися не лише значення шуканої функції  $f(x)$  та інтерполяційної функції  $L_n(x)$ , а й значення похідних до певного

порядку. Таку інтерполяцію називають ермітовою, а якщо  $L_n(x)$  – алгебричний багаточлен  $n$ -го степеня, то його називають *інтерполяційним багаточленом Ерміта*. Він має вигляд

$$H_m(x) = L_n(x) + \omega_{n+1}(x)H_{m-n-1}(x), \quad (7.16)$$

де  $m+1 = \alpha_0 + \alpha_1 + \dots + \alpha_n$  – кількість умов, що накладаються на функцію  $f(x)$ , тобто

$\alpha_0$  – кількість значень функції у вузлах;

$\alpha_1$  – кількість значень першої похідної у певних вузлах;

$\alpha_2$  – кількість значень другої похідної у певних вузлах і так далі;

$L_n(x)$  – інтерполяційний багаточлен Лагранжа;

$$\omega_{n+1}(x) = (x - x_0)(x - x_1)\dots(x - x_n).$$

Формула (7.16) є рекурентною, тобто кожного разу треба шукати багаточлен Ерміта нижчого порядку  $H_{m-n-1}$ , для якого потрібно задавати свої умови. Їх знаходять диференціюванням (7.16). У цьому разі порядок заданих похідних буде зменшуватись весь час на одиницю. Цей процес повторюють доти, доки залишиться побудувати інтерполяційний багаточлен Лагранжа за деякими точками  $x_i$ .

### Приклад 7.6.

Нехай функція задана таблицею

$x$	0	1	2
$y$	1	2	129
$y'$	0	7	448
$y''$	0		1344

Побудувати інтерполяційний багаточлен Ерміта.

*Розв'язування.* Всіх умов є вісім, тобто  $m+1 = 8$ . Отже, степінь багаточлена Ерміта  $m = 7$ . Оскільки задано три вузли, то степінь багаточлена Лагранжа  $n = 2$ . Тобто шукаємо

$$H_7(x) = L_2(x) + \omega_3(x)H_4(x).$$

Побудуємо інтерполяційний багаточлен Лагранжа

$$\begin{aligned} L_2(x) &= 1 \cdot \frac{(x-1)(x-2)}{2} + 2 \cdot \frac{x(x-2)}{-1} + 129 \cdot \frac{x(x-1)}{2} = \\ &= 63x^2 - 62x + 1. \end{aligned}$$

Тоді

$$H_7(x) = 63x^2 - 62x + 1 + x(x-1)(x-2)H_4(x).$$

Продиференціюємо цей вираз:

$$H_7'(x) = 126x - 62 + (3x^2 - 6x + 2)H_4(x) + (x^3 - 3x^2 + 2x)H_4'(x).$$

Підставимо сюди значення  $x=0, 1, 2$  (з урахуванням, що  $H_7'(x_i) = y_i'$ ), отримаємо

$$H_4(0) = 31; \quad H_4(1) = 57; \quad H_4(2) = 129$$

Друга похідна від  $H_7(x)$  має вигляд

$$H_7''(x) = 126 + (6x - 6)H_4(x) + 2(3x^2 - 6x + 2)H_4'(x) + (x^3 - 3x^2 + 2x)H_4''(x).$$

Звідси знайдемо значення  $H_4'(x)$  в точках  $x=0$  та  $x=2$ :

$$H_4'(0) = 15 \quad H_4'(2) = 111.$$

Задача звелася до відшукування багаточлена  $H_4(x)$ , що задовольняє такі умови:

$x$	0	1	2
$H_4(x)$	31	57	129
$H_4'(x)$	15		111

Знову використаємо формулу (7.16):

$$H_4(x) = \bar{L}_2(x) + \omega_1(x)H_1(x).$$

$$\bar{L}_2(x) = 31 \cdot \frac{(x-1)(x-2)}{2} + 57 \cdot \frac{x(x-2)}{-1} + 129 \cdot \frac{x(x-1)}{2} = 23x^2 + 3x + 31.$$

Тобто

$$H_4(x) = 23x^2 + 3x + 31 + (x^3 - 3x^2 + 2x)H_1(x).$$

Продиференціюємо останній вираз:

$$H_4'(x) = 46x + 3 + (3x^2 - 6x + 2)H_1(x) + (x^3 - 3x^2 + 2x)H_1'(x).$$

Підставимо сюди  $x=0$  та  $x=2$ , одержимо  $H_1(0) = 6$ ;  $H_1(2) = 8$ .



Для  $H_1(x)$  маємо таблицю

$x$	0	2
$H_1(x)$	6	8

Отже,  $H_1(x) = L_1(x) = x + 6$ .

Остаточню

$$H_7(x) = 63x^2 - 62x + 1 + x(x-1)(x-2)[23x^2 + 3x + 31 + x(x-1)(x-2)(x+6)] = x^7 + 1.$$

## 7.5. Найліпше наближення і поліноми Чебишова, що найменше відхиляються від нуля

Абсолютна похибка (7.3) методу інтерполювання на заданому проміжку  $[a, b]$  для заданої функції  $f(x)$ , тобто за фіксованої величини  $f^{(n+1)}(\xi)$ , залежить лише від  $\omega_{n+1}(x)$ . Однак поліноми  $\omega_{n+1}(x)$  визначені повністю вузлами інтерполяції  $x_0, x_1, \dots, x_n$ . Природно тепер сформулювати таку задачу: як потрібно на  $[a, b]$  розташувати вузли інтерполяції, щоб похибка методу була найменшою? Розв'язком її є багаточлен, що найменше відхиляється від нуля на  $[a, b]$ . На відрізку  $[-1, 1]$  таким буде багаточлен Чебишова першого роду  $(n+1)$ -го степеня

$$T_{n+1}(t) = \frac{\cos[(n+1) \cdot \arccos t]}{2^n} \quad (7.17)$$

з простими коренями  $t_n^k = \cos \frac{2k+1}{2(n+1)}\pi$  ( $k = 0, 1, \dots, n$ ). Всі ці корені лежать усередині відрізка  $[-1, 1]$ . Точки  $t_n^k$  скупчуються на краях відрізка  $[-1, 1]$ . Всі максимуми дорівнюють  $\frac{1}{2^n}$ , а мінімуми дорівнюють  $-\frac{1}{2^n}$ , тобто  $|T_{n+1}(t)| < \frac{1}{2^n}$ .

Отже,  $\omega_{n+1}(x)$  матиме найменше можливе значення  $\sup |\omega_{n+1}(x)|$  за умови, що за нулі інтерполювання взято корені

багаточлена Чебишова, і в цьому випадку оцінка похибки матиме вигляд

$$r_{n+1} \leq \frac{f^{(n+1)}(\xi)}{2^n (n+1)!} \quad (7.18)$$

Якщо інтерполювання виконують на довільному проміжку  $[a, b]$ , то лінійною заміною

$$x = \frac{a+b}{2} + \frac{b-a}{2}t$$

його можна перевести в  $[-1, 1]$ . Тоді корені багаточлена  $T_{n+1}(x)$  перейдуть у

$$x_n^* = \frac{1}{2}[(b-a) \cdot \cos \frac{2n+1}{2n+2} \pi + (b+a)]. \quad (7.19)$$

Оцінку похибки в цьому випадку виражають так:

$$r_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(x) \leq \frac{f^{(n+1)}(\xi)}{(n+1)!} \frac{(b-a)^{n+1}}{2^{2n+1}}, \quad (7.20)$$

оскільки  $|\omega_{n+1}(x)| \leq 2 \left(\frac{b-a}{4}\right)^{n+1}$

Багаточлени Чебишова можна отримати за рекурентними формулами

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \quad (7.21)$$

де  $T_1(x) = x$ ,  $T_2(x) = 2x^2 - 1$

Багаточлени Чебишова є ортогональними на  $[a, b]$  з вагою

$$p(x) = \frac{1}{\sqrt{1-x^2}}.$$

**П р и к л а д 7.7.** Оцінити похибку інтерполяції функції  $f(x)$  багаточленом Лагранжа п'ятого порядку в точках  $x \in [0; \frac{\pi}{3}]$ , якщо вузли інтерполяції

$$x_m^n = \frac{\pi}{6} \left( 1 + \cos \frac{\pi(2m+1)}{12} \right); \quad m = \overline{0,5}; \quad |f^{(IV)}(x)| < 1.$$

**Розв'язування.** За формулою (7.20) для нашого прикладу маємо

$$|r_5(x)| \leq \frac{1}{6!} \frac{\left(\frac{\pi}{3}\right)^6}{2^{11}} = \frac{\pi^6}{3^6 \cdot 360 \cdot 2^{11}} = \frac{\pi^6}{3^6 \cdot 368640} < 0,0000006.$$

## 7.6. Обернене Інтерполювання

Задача оберненого інтерполювання полягає в такому. Нехай на відрізку  $[a, b]$  задані точки  $x_0, x_1, \dots, x_n$  і значення деякої функції  $f(x)$  в цих точках:

$$y_0 = f(x_0), \quad y_1 = f(x_1), \quad \dots, \quad y_n = f(x_n)$$

За заданими значеннями функції  $f(x)$  потрібно знайти відповідне значення аргументу  $x$ .

У випадку нерівновіддалених вузлів цю задачу легко розв'язати за допомогою інтерполяційного багаточлена Лагранжа. Для цього достатньо взяти  $y$  за незалежну змінну, а  $x$  уважати функцією, тобто

$$x = \sum_{i=0}^n x_i \frac{(y - y_0)(y - y_1) \dots (y - y_{i-1})(y - y_{i+1}) \dots (y - y_n)}{(y_i - y_0)(y_i - y_1) \dots (y_i - y_{i-1})(y_i - y_{i+1}) \dots (y_i - y_n)}. \quad (7.22)$$

У разі рівновіддалених вузлів інтерполювання звичайно використовують метод послідовних наближень. Якщо задана монотонна функція з однаковим кроком інтерполювання  $h$  і значення  $\bar{x}$  міститься на відрізку  $x_0 < \bar{x} < x_1$  (тобто на початку таблиці), то можна застосувати першу інтерполяційну формулу Ньютона, яку запишемо у вигляді

$$y = y_0 + \frac{\Delta y_0}{1!} t + \frac{\Delta^2 y_0}{2!} t(t-1) + \dots + \frac{\Delta^n y_0}{n!} t(t-1) \dots (t-n+1), \quad (7.23)$$

$$t = \frac{x - x_0}{h}. \quad (7.24)$$

Якщо відома кількість кроків  $h$ , тоді з (7.24) можна визначити

$$x = x_0 + th. \quad (7.25)$$

З рівняння (7.23) визначимо

$$t = \frac{y - y_0}{\Delta y_0} - \frac{\Delta^2 y_0}{2! \Delta y_0} t(t-1) - \dots - \frac{\Delta^n y_0}{n! \Delta y_0} t(t-1) \dots (t-n+1). \quad (7.26)$$

За початкове наближення  $t_0$  беруть  $t_0 = (y - y_0) / \Delta y_0$  і будують ітераційний процес

$$t_i = t_0 - \frac{\Delta^2 y_0}{2! \Delta y_0} t_{i-1}(t_{i-1}-1) - \dots - \frac{\Delta^n y_0}{n! \Delta y_0} t_{i-1}(t_{i-1}-1) \dots (t_{i-1}-n+1). \quad (7.27)$$

Тобто значення  $t$  визначають методом послідовних наближень як границю послідовності

$$t = \lim_{i \rightarrow \infty} t_i, \quad (7.28)$$

де  $t_i = \Phi(t_{i-1})$  ( $i = 1, 2, \dots$ )

Кількість членів у формулі (7.27) залежить від кількості вузлів інтерполювання. На практиці процес ітерацій (7.27) триває до виконання умови

$$|t_i - t_{i-1}| < \varepsilon,$$

де  $\varepsilon$  – задана точність. Шукане значення  $x$ , що відповідає заданому значенню  $y$ , знаходять за формулою (7.25).

**Приклад 7.8.** Функція  $y = f(x)$  задана таблицею

$x_i$	10	15	17	20
$y_i$	3	7	11	17

Знайти значення аргументу  $x$ , для якого  $y = 10$ .

*Розв'язування.* За формулою (7.22) знайдемо

$$\begin{aligned} x = \Phi(10) = & 10 \frac{(10-7)(10-11)(10-17)}{(3-7)(3-11)(3-17)} + 15 \frac{(10-3)(10-11)(10-17)}{(7-3)(7-11)(7-17)} + \\ & + 17 \frac{(10-3)(10-7)(10-17)}{(11-3)(11-7)(11-17)} + 20 \frac{(10-3)(10-7)(10-11)}{(17-3)(17-7)(17-11)} = 16,64. \end{aligned}$$

**Приклад 7.9.** Функція  $y = f(x)$  задана таблицею

$x_i$	1,0	1,2	1,4
$y_i$	35	55	63

За якого значення  $x$  значення  $y$  дорівнює 40?

*Розв'язування.* Складемо таблицю скінченних різниць

$x_i$	$y_i$	$\Delta y_i$	$\Delta^2 y_i$
1,0	35		
1,2	55	20	
1,4	63	8	-12

За початкове значення  $y$  приймемо  $y_0 = 35$ . Звідси

$$t_0 = \frac{y - y_0}{\Delta y_0} = \frac{40 - 35}{20} = 0,25.$$

З використанням методу ітерацій знайдемо

$$t_1 = t_0 - \frac{\Delta^2 y_0}{2! \Delta y_0} t_0(t_0 - 1) = 0,25 - \frac{-12}{2 \cdot 20} 0,25(-0,75) = 0,194;$$

$$t_2 = t_1 - \frac{\Delta^2 y_0}{2! \Delta y_0} t_1(t_1 - 1) = 0,25 - 0,3 \cdot 0,194 \cdot 0,806 = 0,203;$$

$$t_3 = t_0 - \frac{\Delta^2 y_0}{2! \Delta y_0} t_2(t_2 - 1) = 0,25 - 0,3 \cdot 0,203 \cdot 0,797 = 0,201;$$

$$t_4 = t_0 - \frac{\Delta^2 y_0}{2! \Delta y_0} t_3(t_3 - 1) = 0,25 - 0,3 \cdot 0,201 \cdot 0,799 = 0,201.$$

За формулою (7.25)

$$x = x_0 + th = 1 + 0,201 \cdot 0,2 = 1,04$$

## Завдання до розділу 7

7.1. Відшукати значення функції в точці  $x=13,5$ , якщо вона задана таблицею

$x$	11	13	14	18	19	21
$y$	1342	2210	2758	5850	6878	9282

7.2. Задані значення  $\lg 56, \lg 58, \lg 60, \lg 62$ . Відшукати значення  $\lg 57$

7.3. Функція задана таблицею

$x$	-0,5	0	0,5	1
$y$	0,57735	1	1,73205	3

Побудувати інтерполяційний багаточлен Лагранжа. Відшукати значення функції в точках  $x=0,2; 0,25; 0,333$ .

7.4. Відомі значення функції  $y = \ln x$ :

$x$	2	3	4	5
$y$	0,693147	1,098613	1,386295	1,609438

Знайти  $\ln 2,5$ .

7.5. Відомі значення функції  $y = e^x$ :

$x$	$y$
1	2,7183
1,05	2,8577
1,10	3,0042
1,15	3,1582

$x$	$y$
1,20	3,3201
1,25	3,4903
1,30	3,6693

Побудувати інтерполяційний багаточлен та за ним відшукати значення функції при  $x=1,13$  та  $x=1,17$

7.6. Узявши крок  $h=0,05$ , побудувати на проміжку  $[4,5; 4,8]$  інтерполяційний багаточлен Ньютона для функції  $e^x$  заданої таблицею

$x$	4,50	4,55	4,60	4,65	4,70	4,75	4,80
$y$	90,0171	94,6324	99,4843	104,5850	109,9472	115,5843	121,5104

7.7. У таблиці наведені значення інтеграла ймовірностей

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-x^2} dx :$$

x	y
1,0	0,8427
1,1	0,8802
1,2	0,9103
1,3	0,9340
1,4	0,9523

x	y
1,5	0,9661
1,6	0,9763
1,7	0,9838
1,8	0,9891
1,9	0,9928
2,0	0,9953

Застосовуючи першу інтерполяційну формулу Ньютона, наближено знайти  $\Phi(1,3)$ .

7.8. Задано таблицю значень  $y = \lg(x)$  семизначних логарифмів

x	1000	1010	1020	1030
y	3,0000000	3,0043214	3,0086002	3,0128372
x	1040	1050		
y	3,0170333	3,0211893		

Знайти  $\lg 1044$ .

7.9. Маючи таблицю значень повного еліптичного інтеграла

$$K(p) = \int_0^{\pi/2} \frac{dx}{\sqrt{1 - \sin^2 p \cdot \sin^2 x}}, \text{ знайти } K(80^\circ 30').$$

p	K(p)
75°	2,76806
76°	2,83267
77°	2,90256
78°	2,97657
79°	3,06173

p	K(p)
80°	3,15339
81°	3,25530
82°	3,36987
83°	3,50042
84°	3,65186

7.10. Побудувати інтерполяційний багаточлен Ньютона інтерполювання вперед для значень функції Бесселя  $y = J_0(x)$ , заданої таблицею

x	2,4	2,5	2,6
y	0,0025	-0,0484	-0,0968

7.11. Для функції  $y = \sin \pi x$  побудувати інтерполяційний поліном Лагранжа, вибравши вузли  $x_0 = 0$ ;  $x_1 = \frac{1}{6}$ ;  $x_2 = \frac{1}{2}$ .

7.12. Побудувати інтерполяційний багаточлен для функції  $y = \cos x$ , заданої таблицею значень для  $x \in [0,28; 0,32]$  з кроком  $h = 0,01$ .

7.13. Дослідити таблицю  $y = \cos x$  для  $x \in [0,28; 0,32]$  з кроком  $h = 0,01$  на можливість лінійної інтерполяції. Знайти  $y = \cos 0,30333$ .

7.14. Побудувати інтерполяційний багаточлен і відшукати значення функції  $y = f(x)$  в точках  $a$  та  $b$ . Функція задана таблицею

7.15.

А)

$x$	1	1,5	2	2,5	3	3,5
$y$	-1	-1,84453	-3,30685	-5,33371	-7,90139	-10,999724

$a = 1,7$ ;  $b = 3,4$ .

Б)

$x$	-1	-0,5	0	0,5	1	1,5
$y$	2,09861	1,31093	0,69315	0,31093	0,09861	-0,05308

$a = -0,7$ ;  $b = 1,2$ .

В)

$x$	-2	-1,5	-1	-0,5	0	0,5
$y$	-1,64165	-1,21062	-0,78028	-0,33781	0,13863	0,66219

$a = -1,58$ ;  $b = 0,3$ .

Г)

$x$	$y$
-0,5	-0,29727
-0,4	-0,16500
-0,3	-0,03469
-0,2	0,09389
-0,1	0,22093

$x$	$y$
0	0,34657
0,1	0,47097
0,2	0,59423
0,3	0,71645
0,4	0,83773
0,5	0,95815

$a = -0,45$ ;  $b = -0,45$ .



7.16. Побудувати інтерполяційний багаточлен Ерміта для функції, заданої таблицею:

А)

$x$	-0,5	1	1,5	3
$y$	-0,22943	1,84147	3,24749	9,14112
$y'$	-0,12242	2,54030	3,07074	5,01001
$y''$	2,47943		1,00251	

Б)

$x$	0	1	1,5	2
$y$	1	2,36788	4,72313	8,13534
$y'$	-1	3,63212	5,77687	7,86466
$y''$	5	4,36788		

В)

$x$	0,5	1	2	3
$y$	-0,81815	-1,00000	-7,30685	-25,90139
$y'$	1,25000	-2,00000	-11,50000	-26,66667
$y''$		-7,00000		-18,11111

Г)

$x$	0	0,5	1	1,5
$y$	0	-1,38014	-2,15853	-2,25564
$y'$	-3	-2,30118		0,15164
$y''$	0	2,59416		0,17505

## Розділ 8

# МЕТОД НАЙМЕНШИХ КВАДРАТІВ І НАЙЛІПШІ СЕРЕДНЬОКВАДРАТИЧНІ НАБЛИЖЕННЯ

### 8.1. Просте опрацювання емпіричних даних методом найменших квадратів

Розглянемо дві найпростіші задачі, які приводять до відомого методу найменших квадратів: розв'язування перевизначених систем лінійних алгебричних рівнянь (СПАР) і підбір параметрів емпіричних залежностей.

Характерна особливість таких задач та, що вхідні дані, зазвичай, є наближеними внаслідок неточності вимірювальних приладів або випадкових похибок.

**Задача А.** Нехай потрібно відшукати значення  $m$  різних величин  $x_1, x_2, \dots, x_m$ , які неможливо виміряти безпосередньо, проте відомо, що їх знаходять за лінійним законом, а коефіцієнти можна одержати наближено за допомогою вимірювання. Виконавши  $n$  таких вимірювань, одержимо СПАР

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m = b_1; \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m = b_2; \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m = b_n. \end{cases} \quad (8.1)$$

Для того, щоб ця система мала єдиний розв'язок  $x_1, x_2, \dots, x_m$ , потрібно, щоб кількість вимірювань була не менша, ніж кількість невідомих  $n \geq m$ , і ранг матриці системи (8.1) повинен дорівнювати  $m$ .

Якщо  $n > m$ , то система рівнянь (8.1) перевизначена. В загальному випадку така система не має розв'язку, тому її розв'язують у сенсі найменших квадратів.

Для цього шукаємо наближений набір  $x_1^*, \dots, x_m^*$  значень змінних  $x_1, \dots, x_m$  не такий, який би перетворював кожне з рівнянь системи (8.1) в тотожність, а за якого була б мінімальна сума квадратів відхилень значень лівих частин рівнянь (8.1) від правих (інакше, квадратів відхилів). Це означає, що розв'язок  $(x_1^*, \dots, x_m^*)$  повинен бути розв'язком задачі мінімізації функції вигляду

$$r_1^2 + r_2^2 + \dots + r_n^2 \rightarrow \min,$$

де  $r_i = a_{i1}x_1 + \dots + a_{im}x_m - b_i$ ,  $i = \overline{1, n}$ .

Функція

$$\varphi(x_1, \dots, x_m) = \sum_{i=1}^n (a_{i1}x_1 + \dots + a_{im}x_m - b_i)^2 \quad (8.2)$$

будучи невід'ємною квадратичною, має єдину точку мінімуму  $(x_1^*, x_2^*, \dots, x_m^*)$ , яку знаходимо з необхідних умов гладкого екстремуму

$$\frac{\partial \varphi(x_1, \dots, x_m)}{\partial x_i} = 0 \quad \text{для } i = \overline{1, m}, \quad (8.3)$$

тобто як розв'язок системи

$$\begin{cases} 2 \sum_{i=1}^n (a_{i1}x_1 + \dots + a_{im}x_m - b_i) a_{i1} = 0, \\ 2 \sum_{i=1}^n (a_{i1}x_1 + \dots + a_{im}x_m - b_i) a_{i2} = 0, \\ \dots \\ 2 \sum_{i=1}^n (a_{i1}x_1 + \dots + a_{im}x_m - b_i) a_{im} = 0. \end{cases}$$

Ця СЛАР має  $m$  рівнянь і  $m$  невідомих

$$\begin{cases} \left( \sum_{i=1}^n a_{i1}^2 \right) x_1 + \left( \sum_{i=1}^n a_{i2} a_{i1} \right) x_2 + \dots + \left( \sum_{i=1}^n a_{im} a_{i1} \right) x_m = \sum_{i=1}^n b_i a_{i1}, \\ \left( \sum_{i=1}^n a_{i1} a_{i2} \right) x_1 + \left( \sum_{i=1}^n a_{i2}^2 \right) x_2 + \dots + \left( \sum_{i=1}^n a_{im} a_{i2} \right) x_m = \sum_{i=1}^n b_i a_{i2}, \\ \dots \\ \left( \sum_{i=1}^n a_{i1} a_{im} \right) x_1 + \left( \sum_{i=1}^n a_{i2} a_{im} \right) x_2 + \dots + \left( \sum_{i=1}^n a_{im}^2 \right) x_m = \sum_{i=1}^n b_i a_{im}. \end{cases} \quad (8.4)$$

Якщо розв'яжемо систему (8.4) із симетричною матрицею, наприклад, методом квадратного кореня, то одержимо розв'язок  $(x_1^*, x_2^*, \dots, x_m^*)$

Такий спосіб розв'язування перевизначених систем називають методом найменших квадратів (МНК), а вектор  $(x_1^*, \dots, x_m^*)$  стосовно вхідної системи (8.1) – нормальним псевдорозв'язком.

**П р и к л а д 8.1.** Знайти два значення  $x_1$  і  $x_2$ , які задані трьома рівняннями:

$$\begin{cases} x_1 + 2x_2 = 8,5, \\ 2x_1 + x_2 = 6,7, \\ x_1 + 3x_2 = 11,1. \end{cases}$$

**Розв'язування.** Нехай праві частини цієї перевизначеної системи є наближеними числами. Розв'яжемо цю систему методом найменших квадратів, для чого відшукаємо мінімум функції

$$\varphi(x_1, x_2) = \sum_{i=1}^3 (a_{i1}x_1 + a_{i2}x_2 - b_i)^2$$

тобто

$$\begin{cases} \frac{\partial \varphi(x_1, x_2)}{\partial x_1} = 6x_1 + 7x_2 - 33 = 0, \\ \frac{\partial \varphi(x_1, x_2)}{\partial x_2} = 7x_1 + 14x_2 - 57 = 0. \end{cases}$$

Розв'яжемо цю систему:

$$x_1^* \approx 1,80; \quad x_2^* \approx 3,17$$

**Задача Б.** Нехай функція  $y = f(x)$  задана таблицею

$x$	$x_0$	$x_1$		$x_n$
$y$	$y_0$	$y_1$		$y_n$

наближених значень  $y_i \approx f(x_i)$ , які одержані експериментально. Потрібно знайти функцію  $\varphi(x)$  таку, яка б на відрізку  $[x_0, x_n]$  апроксимувала функцію  $f(x)$ , що задана окремими наближеними

значеннями  $y_i$ . Функцію  $\varphi(x)$ , яка апроксимує  $f(x)$ , беремо з ви-  
значеної  $m$ -параметричної сім'ї функцій, і її параметри підбираємо  
так, щоб сума квадратів відхилень обчислених значень  $\varphi(x_i)$  від  
заданих наближених значень  $y_i$  була мінімальною. Така функція  
(тобто за такого оптимального набору параметрів) буде найліпшою  
апроксимацією  $f(x)$  серед функцій вибраної сім'ї в сенсі методу  
найменших квадратів.

Отже, згідно з МНК, задаємо сім'ю  $y = \varphi(x, a_1, a_2, \dots, a_m)$  і шукаємо  
значення параметрів  $a_1, a_2, \dots, a_m$ , де  $m \leq n+1$ , розв'язуючи  
екстремальну задачу

$$\varphi(a_1, a_2, \dots, a_m) = \sum_{i=0}^n [\varphi(x_i, a_1, \dots, a_m) - y_i]^2 \rightarrow \min. \quad (8.5)$$

Оптимальний набір параметрів  $a_1^*, \dots, a_m^*$  знаходимо з системи

$$\begin{cases} \sum_{i=0}^n [\varphi(x_i, a_1, \dots, a_m) - y_i] \frac{\partial \varphi}{\partial a_1} \Big|_{x=x_i} = 0, \\ \sum_{i=0}^n [\varphi(x_i, a_1, \dots, a_m) - y_i] \frac{\partial \varphi}{\partial a_2} \Big|_{x=x_i} = 0, \\ \dots \\ \sum_{i=0}^n [\varphi(x_i, a_1, \dots, a_m) - y_i] \frac{\partial \varphi}{\partial a_m} \Big|_{x=x_i} = 0. \end{cases} \quad (8.6)$$

Якщо функція  $\varphi(x, a_1, \dots, a_m)$  є лінійною функцією відносно своїх  
параметрів  $a_1, \dots, a_m$ , то система (8.6) також буде лінійною. В  
загальному випадку система (8.6) нелінійна, а це зумовлює певні  
труднощі під час її розв'язування.

**Приклад 8.2.** Функція задана таблично:

$x$	2,0	2,2	2,4	2,6	2,8	3,0
$y$	0,301	0,3424	0,3802	0,4150	0,4472	0,4771

Підібрати багаточлен другого степеня  $y = ax^2 + bx + c$ , знаходячи  
 $a, b, c$  методом найменших квадратів.

*Розв'язування.* Параметри  $a, b, c$  виберемо з умови

$$S = \sum_{i=0}^5 \epsilon_i^2 = \sum_{i=0}^5 [y_i - (ax_i^2 + bx_i + c)]^2 \rightarrow \min.$$

Одержимо систему

$$\begin{cases} \frac{\partial S}{\partial a} = 2 \sum_{i=0}^5 [y_i - (ax_i^2 + bx_i + c)](-x_i^2) = 0, \\ \frac{\partial S}{\partial b} = 2 \sum_{i=0}^5 [y_i - (ax_i^2 + bx_i + c)](-x_i) = 0, \\ \frac{\partial S}{\partial c} = 2 \sum_{i=0}^5 [y_i - (ax_i^2 + bx_i + c)](-1) = 0. \end{cases}$$

Після перетворень отримаємо таку систему:

$$\begin{cases} a \sum_{i=0}^5 x_i^4 + b \sum_{i=0}^5 x_i^3 + c \sum_{i=0}^5 x_i^2 = \sum_{i=0}^5 x_i^2 y_i, \\ a \sum_{i=0}^5 x_i^3 + b \sum_{i=0}^5 x_i^2 + c \sum_{i=0}^5 x_i = \sum_{i=0}^5 x_i y_i, \\ a \sum_{i=0}^5 x_i^2 + b \sum_{i=0}^5 x_i + 6c = \sum_{i=0}^5 y_i. \end{cases}$$

Для нашого випадку

$$\begin{cases} 260,7664a + 99b + 38,2c = 15,6565, \\ 99a + 38,2b + 15c = 6,0302, \\ 38,2a + 15b + 6c = 2,3629. \end{cases}$$

Звідси

$$a = -0,035762; \quad b = 0,35448; \quad c = -0,2647.$$

Тоді шуканий багаточлен набуде вигляду

$$y = -0,035762x^2 + 0,35448x - 0,2647$$

Залежно від характеру табличних даних для опрацювання результатів експериментів часто використовують, наприклад, такі двопараметричні сім'ї функцій:

$$y = ax + b; \quad y = a + b \ln x \quad (y = a + b \lg x); \quad y = ax^b;$$

$$y = ae^{bx}; \quad y = a + \frac{b}{x}; \quad y = \frac{1}{ax + b}; \quad y = \frac{b}{ax + b};$$

рідше застосовують трипараметричні сім'ї

$$y = ax^2 + bx + c; \quad y = ax^b + c;$$

$$y = ae^{bx} + c \quad (y = a \cdot 10^{bx} + c).$$

Для дослідження періодичних явищ використовують тригонометричні функції.

**Приклад 8.3.** Нехай значення  $x_i$  і відповідні їм наближені значення  $y_i$  виражають деяку лінійну залежність  $\phi(x, a, b) = ax + b$

Потрібно знайти параметри  $a$  і  $b$

*Розв'язування.* Знайдемо  $\frac{\partial \phi}{\partial a} = x_i$ ;  $\frac{\partial \phi}{\partial b} = 1$ . Для обчислення параметрів цієї функції  $\phi$  складемо систему типу (8.6)

$$\begin{cases} \sum_{i=0}^n (ax_i + b - y_i) x_i = 0, \\ \sum_{i=0}^n (ax_i + b - y_i) \cdot 1 = 0. \end{cases}$$

Звідси одержимо СЛАР щодо коефіцієнтів  $a$  і  $b$

$$\begin{cases} \left( \sum_{i=0}^n x_i^2 \right) a + \left( \sum_{i=0}^n x_i \right) b = \sum_{i=0}^n x_i y_i \\ \left( \sum_{i=0}^n x_i \right) a + (n+1)b = \sum_{i=0}^n y_i \end{cases}$$

Ця система з симетричною квадратною матрицею має єдиний розв'язок  $(a^*, b^*)$ .

**Зауваження.** Замість того, щоб розв'язувати систему нелінійних рівнянь (8.6) для відшукування параметрів конкретних наборів функцій, якщо ці параметри входять нелінійним способом, іноді можна зробити лінеаризацію підібраної залежності.

**Приклад 8.4.** Нехай за апроксимувальну функцію  $f(x)$  вибрано експоненціальну

$$\phi(x, a, b) = a \cdot e^{bx}$$

Знайти параметри  $a$  і  $b$

*Розв'язування.* Знайдемо частинні похідні

$$\frac{\partial \phi}{\partial a} = e^{bx}, \quad \frac{\partial \phi}{\partial b} = ab \cdot e^{bx}.$$

У цьому випадку система (8.6) буде нелінійною. Щоб не розв'язувати нелінійні рівняння, прологарифмуємо рівність

$$ae^{bx} = y;$$

$$\ln a + bx = \ln y$$

Уведемо позначення  $A = \ln a$ . За цими значеннями  $y_i = f(x_i)$ ,  $i = \overline{0, n}$  складемо таблицю значень  $Y_i = \ln y_i$ .

$x$	$x_0$	$x_1$	...	$x_n$
$Y$	$Y_0$	$Y_1$	...	$Y_n$

Методом найменших квадратів знайдемо оптимальні параметри  $A^*$  і  $b^*$  для лінійної функції  $Y = A + bx$ . Після цього обчислимо оптимальне значення  $a^* = e^{A^*}$  й остаточний результат  $f(x) \approx a^* e^{b^* x}$ .

## 8.2. Середньоквадратичне наближення за допомогою тригонометричних багаточленів

Нехай задано неперервну функцію  $f(x)$  з періодом  $2\pi$ . Апроксимувальний багаточлен  $T_m(x)$  беремо у вигляді

$$T_m(x) = \sum_{k=0}^m (a_k \cos kx + b_k \sin kx). \quad (8.7)$$

Коефіцієнти Фур'є  $a_k$  і  $b_k$  запишемо так:

$$\begin{cases} a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x) dx, \\ a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx dx, \\ b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx dx. \end{cases} \quad (8.8)$$

Якщо функція  $f(x)$  парна, то

$$a_0 = \frac{1}{\pi} \int_0^{\pi} f(x) dx; \quad a_k = \frac{2}{\pi} \int_0^{\pi} f(x) \cos kx dx, \quad b_k = 0, \quad k = 1, m, \quad (8.9)$$

і якщо функція  $f(x)$  непарна, то

$$a_0 = 0; \quad a_k = 0; \quad b_k = \frac{2}{\pi} \int_0^{\pi} f(x) \sin kx dx, \quad (8.10)$$



Нехай у випадку періодичної з періодом  $2\pi$  функції відомі  $n$  її значень у рівновіддалених точках проміжку  $[0, 2\pi]$

$$x_j = \frac{2\pi j}{n}, \quad j = 0, 1, \dots, n-1. \quad (8.11)$$

Коефіцієнти  $a_k$  і  $b_k$  тригонометричного багаточлена (8.7) при  $n > 2m$  підбираємо так, щоб значення

$$S = \sum_{j=0}^{n-1} [f(x_j) - T_m(x_j)]^2 \quad (8.12)$$

було мінімальним.

Прирівнюємо до нуля частинні похідні від  $S$  за всіма  $a_k$  і  $b_k$  та, зробивши відповідні перетворення, знайдемо коефіцієнти Фур'є:

$$\begin{cases} a_0 = \frac{1}{\pi} \sum_{j=0}^{n-1} f(x_j), \\ a_k = \frac{2}{\pi} \sum_{j=0}^{n-1} f(x_j) \cos kx_j, \\ b_k = \frac{2}{\pi} \sum_{j=0}^{n-1} f(x_j) \sin kx_j \quad (k = \overline{1, m}). \end{cases} \quad (8.13)$$

**Зауваження.** У формулах (8.13) підсумовування можна виконувати тільки на половині відрізка  $[0, 2\pi]$ , тоді остаточний результат подвоюється.

Нехай функція  $f(x)$  має період  $T$ . Тоді коефіцієнти Фур'є знаходять за формулами

$$\begin{cases} a_0 = \frac{1}{T} \int_0^T f(x) dx, \\ a_k = \frac{2}{T} \int_0^T f(x) \cos k \frac{2\pi x}{T} dx, \\ b_k = \frac{2}{T} \int_0^T f(x) \sin k \frac{2\pi x}{T} dx \quad (k = \overline{1, m}). \end{cases} \quad (8.14)$$

**П р и к л а д 8.5.** За допомогою тригонометричного багаточлена дев'ятого степеня потрібно одержати наближення періодичної функції

$$f(x) = \begin{cases} x(\pi - x) & 0 \leq x \leq \pi, \\ x(\pi + x) & -\pi \leq x \leq 0. \end{cases}$$

*Розв'язування.* З урахуванням непарності цієї функції за формулами (8.10) маємо

$$a_k = 0, \quad k = 0, 1, 2, \dots,$$

$$\begin{aligned} b_k &= \frac{2}{\pi} \int_0^{\pi} x(\pi - x) \sin kx dx = \\ &= \frac{2}{\pi} \left[ -x(\pi - x) \frac{\cos kx}{k} \Big|_0^{\pi} + \frac{1}{k} \int_0^{\pi} (\pi - 2x) \cos kx dx \right] = \\ &= \frac{2}{\pi k} \left[ (\pi - 2x) \frac{\sin kx}{k} \Big|_0^{\pi} + \frac{2}{k} \int_0^{\pi} \sin kx dx \right] = \frac{4}{\pi k^2} \int_0^{\pi} \sin kx dx = \\ &= -\frac{4}{\pi k^3} \cos kx \Big|_0^{\pi} = \frac{4}{\pi k^3} [1 - (-1)^k]. \end{aligned}$$

Отже,

$$b_{2n} = 0; \quad b_{2n-1} = \frac{8}{\pi(2n-1)^3}$$

Шуканий тригонометричний багаточлен запишемо так:

$$T_9(x) = \frac{8}{\pi} \left[ \sin x + \frac{1}{27} \sin 3x + \frac{1}{125} \sin 5x + \frac{1}{343} \sin 7x + \frac{1}{729} \sin 9x \right]$$

**П р и к л а д 8.6.** З використанням тригонометричного полінома третього степеня знайти наближення непарної періодичної функції  $f(x)$ , яка задана на відрізку  $[0, \pi]$  відповідними значеннями:

$x_i$	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	$\pi$
$y_i$	0	5	7	6	0

*Розв'язування.* Тут  $n = 8$ . Оскільки функція  $f(x)$  непарна, то всі  $a_k = 0$ , а для визначення коефіцієнтів  $b_k$  за формулою (8.10) маємо

$$b_k = \frac{1}{4} \sum_{j=0}^7 f(x_j) \sin kx_j = \frac{1}{2} \sum_{j=0}^3 f\left(\frac{\pi j}{4}\right) \sin k \frac{\pi j}{4}$$

або

$$b_1 = 3,5 + \frac{11\sqrt{2}}{4}; \quad b_2 = -0,5; \quad b_3 = -3,5 + \frac{11\sqrt{2}}{4}.$$

Тоді шуканий тригонометричний багаточлен запишемо так:

$$T_3(x) = \left(3,5 + \frac{11\sqrt{2}}{4}\right) \sin x - 0,5 \sin 2x + \left(-3,5 + \frac{11\sqrt{2}}{4}\right) \sin 3x.$$

**Приклад 8.7.** Побудувати тригонометричний багаточлен для парної функції, заданої таблицею

$i$	$x_i$	$f(x_i)$
0	0	9,55
1	$\pi/12$	9,46
2	$\pi/6$	9,25
3	$\pi/4$	8,96
4	$\pi/3$	8,58
5	$5\pi/12$	8,10
6	$\pi/2$	7,59
7	$7\pi/12$	7,00
8	$2\pi/3$	6,34
9	$3\pi/4$	5,56
10	$5\pi/6$	4,80
11	$11\pi/12$	4,24
12	$\pi$	4,00

*Розв'язування.* Беремо  $m = 12$ , тобто відрізок  $[-\pi, \pi]$  ділимо на 24 частини. Обмежимося шістьма гармоніками (і нульовою),  $n = 6$ ,  $b_k = 0$ , бо функція парна:

$$\sum_{i=0}^n y_i \cos kx = S_k \quad \text{для } k = 0, 1, 2, \dots, 6$$

Поділимо  $S_k$  на  $\frac{m+1}{2} = 6,5$ . Одержимо наближені значення коефіцієнтів  $a_k$ . Шуканий поліном запишемо так:

$$Q_6(x) = 7,22 + 2,59 \cos x - 0,41 \cos 2x + \\ + 0,18 \cos 3x - 0,04 \cos 4x + 0,01 \cos 6x.$$

Для перевірки:

$$Q_6(0) = 7,22 + 2,59 - 0,41 + 0,18 - 0,04 + 0,01 = 9,55$$

### Завдання до розділу 8

8.1. Методом найменших квадратів розв'язати такі перевизначені системи:

$$1) \begin{cases} x_1 + 3x_2 - 2x_3 = 5; \\ 3x_1 + 4x_2 - 5x_3 = 6; \\ -2x_1 - 5x_2 + 3x_3 = -13; \\ x_2 - 2x_3 = 4; \\ -2x_1 - 3x_2 + 2x_3 = -10; \end{cases}$$

$$2) \begin{cases} 3x_1 + x_2 - x_3 = 2; \\ -5x_1 + x_2 + 3x_3 = -10; \\ 2x_1 + x_3 = 1; \\ x_1 - 5x_2 + 3x_3 = 7; \\ 2x_1 - 4x_2 - x_3 = 11; \end{cases}$$

$$3) \begin{cases} 2x_1 - x_2 + 3x_3 = 8; \\ x_1 + x_2 - 4x_3 = -3; \\ 3x_1 - x_2 + x_3 = 8; \\ x_1 - 3x_2 + 3x_3 = 8; \\ x_1 - x_2 + x_3 = 4; \end{cases}$$

$$4) \begin{cases} 2x_1 + 2x_2 - x_3 = -2; \\ 4x_1 + 3x_2 - x_3 = -1; \\ 8x_1 + 5x_2 - 3x_3 = -3; \\ 3x_1 + 3x_2 - 2x_3 = -4; \\ x_1 + 2x_2 + 4x_3 = 7; \end{cases}$$

- $$5) \begin{cases} x_1 + x_2 + 2x_3 = -3; \\ 3x_1 - x_2 - x_3 = -6; \\ 2x_1 + 3x_2 - x_3 = 0; \\ x_1 + 2x_2 + 3x_3 = -3; \\ 3x_1 - 2x_2 - x_3 = -7; \end{cases}$$
- $$6) \begin{cases} x_1 + 2x_2 + 3x_3 = 8; \\ 2x_1 - x_2 - 2x_3 = 3; \\ 3x_1 + 2x_2 - x_3 = 10; \\ 2x_1 - 3x_2 + 2x_3 = 5; \\ -2x_1 - 3x_2 + 2x_3 = -7; \end{cases}$$
- $$7) \begin{cases} 3x_1 - 2x_2 + 2x_3 = -1; \\ 2x_1 + x_2 + x_3 = 6; \\ 3x_1 - x_2 + 2x_3 = 2; \\ x_1 + 2x_2 - x_3 = 6; \\ -2x_1 + x_2 - x_3 = 0; \end{cases}$$
- $$8) \begin{cases} x_1 + 4x_2 + x_3 = 8; \\ -x_2 + 3x_3 = 8; \\ 3x_1 + x_2 = 4; \\ x_1 - 2x_2 + x_3 = 2; \\ 3x_1 - x_2 + 2x_3 = 8; \end{cases}$$
- $$9) \begin{cases} x_1 + x_2 - x_3 = 4; \\ 2x_1 + x_2 + 2x_3 = 10; \\ x_1 - x_2 + 3x_3 = 4; \\ 3x_1 + x_2 - x_3 = 10; \\ x_1 + 2x_2 + 3x_3 = 10; \\ 4x_1 - x_2 + x_3 = 11; \end{cases}$$

$$10) \begin{cases} x_1 + 2x_2 + 3x_3 = 8; \\ -x_1 + x_2 - 4x_3 = -3; \\ 2x_1 - x_2 + 5x_3 = 5; \\ 3x_1 + 2x_2 - x_3 = 6; \\ x_1 - x_2 + 5x_3 = 4; \\ 7x_1 + x_2 - 10x_3 = -1. \end{cases}$$

8.2. Для функції, заданої таблично, підібрати багаточлен другого степеня  $y = ax^2 + bx + c$ , знаходячи параметри  $a, b, c$  методом найменших квадратів.

1)

$x$	0,41	0,46	0,52	0,60	0,65	0,72
$y$	2,5741	2,3251	2,0934	1,8620	1,7493	1,6210

2)

$x$	0,35	0,41	0,47	0,51	0,56	0,64
$y$	2,7395	2,3008	1,9686	1,7878	1,5950	1,3431

3)

$x$	0,05	0,10	0,17	0,25	0,30	0,36
$y$	0,0500	0,1003	0,1717	0,2553	0,3093	0,3764

4)

$x$	0,62	0,67	0,74	0,80	0,87	0,96	0,99
$y$	0,5379	0,5117	0,4771	0,4493	0,4189	0,3829	0,3716

5)

$x$	1,03	1,08	1,16	1,23	1,26	1,33	1,39
$y$	2,8011	2,9447	3,1899	3,4212	3,5254	3,7810	4,0149

6)

$x$	0,15	0,20	0,25	0,30	0,35	0,40	0,45
$y$	0,8607	0,8187	0,7788	0,7408	0,7047	0,6703	0,6376

7)

$x$	0,01	0,06	0,11	0,16	0,21	0,26	0,31
$y$	0,9918	0,9519	0,9137	0,8769	0,8416	0,8078	0,7753

8)

$x$	0,15	0,17	0,18	0,20	0,21	0,23
$y$	4,4817	5,4739	6,0496	7,3891	8,1662	9,9742

9)

$x$	0,45	0,46	0,48	0,49	0,51	0,53	0,56
$y$	2,7946	2,6133	2,1746	1,7301	1,5198	1,2551	0,9265

10)

$x$	0,63	0,64	0,65	0,66	0,67	0,68
$y$	0,5891	0,5972	0,6052	0,6131	0,6210	0,6288

8.3. Для функції, яка задана таблично, побудувати багаточлен

$$y = a + \frac{b}{x},$$
 параметри  $a$  і  $b$  знайти методом найменших квадратів.

1)

$x$	1,1	1,3	1,4	1,8	1,9
$y$	1,3421	2,2107	2,7583	5,5803	6,8789

2)

$x$	2,0	2,3	2,5	3,0	3,5	3,7
$y$	5,7343	6,1274	6,3007	6,6943	7,0474	7,2434

3)

$x$	0,90	0,96	1,00	1,04	1,09	1,13
$y$	1,1384	1,2804	1,4705	1,7076	1,9916	2,0195

4)

$x$	1,0	1,1	1,3	1,5	1,6	1,8
$y$	1,1752	1,3357	1,6984	2,1293	2,3756	2,9422

5)

$x$	1,0	1,1	1,3	1,4	1,6	1,9
$y$	0,6827	0,7287	0,8064	0,8385	0,8664	0,9426

6)

$x$	0,3	0,4	0,6	0,7	0,9	1,0
$y$	0,2913	0,3799	0,5370	0,6044	0,7163	0,7616

7)

$x$	1,0	1,1	1,3	1,4	1,6	1,7
$y$	1,0205	0,9514	0,8975	0,8873	0,8935	0,9086

8)

$x$	1,2	1,4	1,6	2,0	2,4	2,6
$y$	0,6711	0,5669	0,4554	0,2239	0,1104	0,0025

9)

$x$	0,20	0,24	0,26	0,29	0,32	0,38
$y$	1,2214	1,2712	1,2969	1,3364	1,3771	1,4623

10)

$x$	0,1	0,15	0,18	0,22	0,28	0,30
$y$	1,1052	1,1618	1,1972	1,2461	1,3231	1,3499



## Розділ 9

# ІНТЕРПОЛЯЦІЙНІ СПЛАЙНИ

### 9.1. Кусково-поліноміальна апроксимація

У випадку, коли функція  $f(x)$  не є достатньо гладкою на проміжку  $[a, b]$ , замість апроксимації її багаточленами високих степенів перспективнішим є застосування кусково-поліноміальної апроксимації  $f(x)$ . У цьому разі апроксимувальна функція  $\phi(x)$  складається з окремих багаточленів, зазвичай, однакового невеликого степеня, кожен з яких визначений на своїй частині відрізка  $[a, b]$ .

Використання низьких степенів багаточленів, які становлять  $\phi(x)$ , дає змогу легко знаходити їхні коефіцієнти як з інтерполяційних, так і з інших умов.

Зокрема, якщо задані значення  $y_i$  функції  $y = f(x)$  на системі вузлів  $x_i$  такій, що

$$a \leq x_0 < x_1 < \dots < x_n \leq b, \quad (9.1)$$

і необхідно апроксимувати  $f(x)$  кусково-лінійною функцією  $\phi(x)$  з огляду на умови інтерполяції

$$\phi(x_i) = y_i, \quad i = 0, 1, \dots, n,$$

то, узявши функцію  $\phi(x)$  у вигляді

$$\phi(x) = \begin{cases} a_1 x + b_1 & \text{при } x \in [x_0, x_1], \\ a_2 x + b_2 & \text{при } x \in [x_1, x_2], \\ \dots & \dots \\ a_n x + b_n & \text{при } x \in [x_{n-1}, x_n], \end{cases} \quad (9.2)$$

для знаходження  $n$  пар її коефіцієнтів  $a_k, b_k$  ( $k = 1, 2, \dots, n$ ) маємо систему з  $2n$  лінійних рівнянь

$$\begin{cases} a_1 x_0 + b_1 = y_0, \\ a_1 x_1 + b_1 = y_1; \\ \dots \\ \begin{cases} a_2 x_1 + b_2 = y_1, \\ a_2 x_2 + b_2 = y_2; \\ \dots \\ a_n x_{n-1} + b_n = y_{n-1}, \\ a_n x_n + b_n = y_n. \end{cases} \end{cases} \quad (9.3)$$

Кожну пару сусідніх рівнянь (9.3) можна розв'язувати окремо. Аналогічно, кожна ланка кусково-квадратичної функції (при  $n = 2m$  у (9.2))

$$\phi(x) = \begin{cases} a_1 x^2 + b_1 x + c_1 & \text{при } x \in [x_0, x_2], \\ a_2 x^2 + b_2 x + c_2 & \text{при } x \in [x_2, x_4], \\ \dots \\ a_n x^2 + b_n x + c_n & \text{при } x \in [x_{2m-2}, x_{2m}], \end{cases} \quad (9.4)$$

визначена трійкою коефіцієнтів  $a_k, b_k, c_k$  ( $k = 1, 2, \dots, m$ ), які можна знайти послідовним розв'язуванням (при  $k = 1, 2, \dots, m$ ) тривимірних лінійних систем

$$\begin{cases} a_k x_{2k-2}^2 + b_k x_{2k-2} + c_k = y_{2k-2}, \\ a_k x_{2k-1}^2 + b_k x_{2k-1} + c_k = y_{2k-1}, \\ a_k x_{2k}^2 + b_k x_{2k} + c_k = y_{2k}, \end{cases} \quad (9.5)$$

які відповідають інтерполяційним умовам.

**Приклад 9.1.** Для функції  $y = f(x)$ , заданої таблицею

$x_j$	-2	-1	0	1	1.5	2	3
$y_i$	3	1	1	2	3	1	4

виконати найпростіші кусково-лінійне і кусково-квадратичне інтерполювання.

**Розв'язування.** Використавши лінійне інтерполювання заданої функції на кожному з елементарних проміжків, визначених сусідніми вузлами верхнього рядка таблиці, отримаємо  $f(x) \approx \phi_1(x)$ , де

$$\phi_1(x) = \begin{cases} -2x - 1 & \text{при } x \in [-2, -1], \\ 1 & \text{при } x \in [-1, 0], \\ x + 1 & \text{при } x \in [0, 1], \\ 2x & \text{при } x \in [1, 1.5], \\ -4x + 9 & \text{при } x \in [1.5, 2], \\ -3x - 5 & \text{при } x \in [2, 3]. \end{cases}$$

Квадратичне інтерполювання за трійками вузлових точок  $[-2; 0]$ ,  $[0; 1.5]$  і  $[1.5; 3]$  приводить до наближеної рівності  $f(x) \approx \phi_2(x)$ , де

$$\phi_2(x) = \begin{cases} x^2 + x + 1 & \text{при } x \in [-2, 0], \\ \frac{2}{3}x^2 + \frac{1}{3}x + 1 & \text{при } x \in [0; 1.5], \\ \frac{1}{3}(14x^2 - 61x + 69) & \text{при } x \in [1.5; 3]. \end{cases}$$

## 9.2. Інтерполяційний сплайн

Нехай на відрізку  $[a, b]$  задано впорядковану систему незбіжних точок  $x_k$  ( $k = 0, 1, \dots, n$ ).

**Означення 9.1.** Сплайном  $S_n(x)$  називають визначену на  $[a, b]$  функцію, яка належить класу  $C^l[a, b]$  і разів неперервно диференційовних функцій, таку, що на кожному проміжку  $[x_{k-1}, x_k]$  ( $k = 1, 2, \dots, n$ ) – це багаточлен  $m$ -го степеня. Різницю  $d = m - l$  між степенем сплайна  $m$  і показником його гладкості  $l$  називають **дефектом** сплайна.

Якщо для сплайна  $S_m(x)$  виконуються умови  $S_m(x_i) = f(x_i)$ , то такий сплайн називають *інтерполяційним сплайном* для функції  $f(x)$ ; у цьому разі вузли сплайна  $x_k$ , взагалі кажучи, можуть не збігатися з вузлами інтерполяції.

Розглянуті вище кусково-лінійна функція  $\phi_1(x)$  і кусково-квадратична функція  $\phi_2(x)$ , визначені, відповідно, формулами (9.2) і (9.4), є інтерполяційним сплайном степеня 1 з дефектом 1 та інтерполяційним сплайном степеня 2 з дефектом 2.

Збіг дефекту сплайна з його степенем забезпечує лише неперервність сплайна. Важливою є побудова сплайнів з більшою гладкістю, тобто з малим дефектом.

Широковживаний інтерполяційний сплайн степеня 3 з дефектом 1. У цьому разі будемо вважати, що вузли сплайна

$$a = x_0, x_1, x_2, \dots, x_{n-1}, x_n = b \quad (9.6)$$

одночасно є вузлами інтерполяції, тобто в них відомі значення функції  $f_k = f(x_k)$ ,  $k = 0, 1, \dots, n$ .

**Означення 9.2.** Кубічним сплайном дефекту 1, що інтерполює на відрізку  $[a, b]$  задану функцію  $f(x)$ , називають функцію

$$g(x) = \{g_k(x) = a_k + b_k(x - x_k) + c_k(x - x_k)^2 + d_k(x - x_k)^3 \text{ при } x \in [x_{k-1}, x_k]\}_{k=1}^n, \quad (9.7)$$

яка задовольняє сукупності таких умов:

- а)  $g(x_k) = f(x_k)$  (умова інтерполювання у вузлах сплайна);
- б)  $g(x) \in C_{[a,b]}^2$  (двічі неперервна диференційованість);
- в)  $g''(a) = g''(b) = 0$  (крайові умови).

Визначений так сплайн ще називають природним. Зазначимо, що крайові умови в означенні 9.2 можна замінити на інші. Наприклад, можна накласти додаткові умови на першу похідну функції  $g(x)$  в точках  $a$  і  $b$ . Однак у цьому випадку кубічний сплайн (9.7) втрачає властивість природності.

Для побудови за заданою функцією  $f(x)$  її інтерполяційного сплайна (9.7) необхідно знайти  $4n$  його коефіцієнтів  $a_k, b_k, c_k, d_k$  ( $k = 1, 2, \dots, n$ ). Їх можна визначити

а) з умов інтерполяції

$$g_1(x_0) = f_0, \quad g_k(x_k) = f_k \quad \text{при } k = 1, 2, \dots, n; \quad (9.8)$$

б) з умов гладкості з'єднання ланок сплайна

$$\left. \begin{aligned} g_{k-1}(x_{k-1}) &= g_k(x_{k-1}), \\ g'_{k-1}(x_{k-1}) &= g'_k(x_{k-1}), \\ g''_{k-1}(x_{k-1}) &= g''_k(x_{k-1}) \end{aligned} \right\} \text{ при } k = 2, 3, \dots, n; \quad (9.9)$$

в) із крайових умов

$$g'_1(x_0) = 0, \quad g''_n(x_n) = 0. \quad (9.10)$$

Таких умов виявилось  $4n$  – стільки, скільки є в записі сплайна (9.7) невідомих коефіцієнтів. У кінцевому підсумку для визначення невідомих коефіцієнтів  $c_k$  сплайна (9.7) необхідно розв'язати триточкове різницеве рівняння другого порядку

$$h_{k-1}c_{k-2} + 2(h_{k-1} + h_k)c_{k-1} + h_k c_k = 3f(x_{k-1}; x_k) - 3f(x_{k-2}; x_{k-1}), \quad (9.11)$$

де  $k = 2, 3, \dots, n$ ,  $c_0 = 0$ ,  $c_n = 0$ ,  $h_k = x_k - x_{k-1}$ ,  $f(x_{k-1}; x_k) = \frac{f_k - f_{k-1}}{h_k}$  –

поділена різниця першого порядку. Інші коефіцієнти сплайна визначають за формулами

$$d_k = \frac{c_k - c_{k-1}}{3h_k} \quad (k = 1, 2, \dots, n). \quad (9.12)$$

$$b_k = f(x_{k-1}; x_k) + \frac{2}{3}h_k c_k + \frac{1}{3}h_k c_{k-1} \quad (k = 1, 2, \dots, n), \quad (9.13)$$

$$a_k = f_k \text{ при довільному } k \in \{1, 2, \dots, n\}. \quad (9.14)$$

Оскільки  $|2h_{k-1} + 2h_k| > |h_{k-1}| + |h_k|$ , то ця нерівність, згідно з умовами (2.44), забезпечує стійкість методу прогонки для розв'язування системи рівнянь (9.11).

**П р и к л а д 9.2.** Для функції, заданої таблицею (див. приклад 9.1), побудувати інтерполяційний сплайн степеня 3 з дефектом 1 у вигляді (9.7).

**Розв'язування.** За даними таблиці, згідно з формулою (9.11), знайдемо

$$h_1 = 1; h_2 = 1; h_3 = 1; h_4 = 0,5; h_5 = 0,5; h_6 = 1;$$

$$f(x_0; x_1) = -2, \quad f(x_1; x_2) = 0,$$

$$f(x_2; x_3) = 1, \quad f(x_3; x_4) = 2, \quad f(x_4; x_5) = -4, \quad f(x_5; x_6) = 3.$$

Далі сформуємо тридіагональну матрицю  $H$  і вектор правих частин  $e$ :

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 4 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 & 0,5 & 0 & 0 \\ 0 & 0 & 0 & 0,5 & 2 & 0,5 & 0 \\ 0 & 0 & 0 & 0 & 0,5 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad e = \begin{pmatrix} 0 \\ 6 \\ 3 \\ 3 \\ -18 \\ 21 \\ 0 \end{pmatrix}.$$

Розв'яжемо систему  $Hc = e$  методом прогонки, отримаємо вектор

$$c = (0; 1,610245; -0,440980; 3,153675; -12,040089; 9,006681; 0)^T$$

Потім за формулами (9.12)–(9.14) знайдемо значення інших коефіцієнтів сплайна  $a_k, b_k, c_k$ . Запишемо їх у табл. 9.1.

Таблиця 9.1

Значення коефіцієнтів кубічного сплайна

$k$	$x_k$	$a_k = y_k$	$b_k$	$c_k$	$d_k$
0	-2	3	—	0	—
1	-1	1	-0,926503	1,610245	0,536748
2	0	1	0,242762	-0,440980	-0,683742
3	1	2	<b>2,955457</b>	<b>3,153675</b>	<b>1,198218</b>
4	1,5	3	-1,487751	-12,040089	-10,129176
5	2	1	-3,004454	9,006682	14,033180
6	3	4	6,002227	0	-3,002227

Цими коефіцієнтами сплайн (9.7) для заданого прикладу визначений повністю. Зокрема, для  $x = 1,5$  за допомогою елементів виокремленого в таблиці рядка отримаємо

$$\begin{aligned} \phi_3(1,5) &= 2 + 2,95545 \cdot (-0,5) + \\ &+ 3,15367 \cdot 0,25 - 1,19822 \cdot 0,125 = 1,160915. \end{aligned}$$

### 9.3. Базисні сплайни

Важливу роль у разі побудови чисельно-аналітичних методів розв'язування диференціальних та інтегральних рівнянь відіграють так звані базисні сплайни (*B*-сплайни). Зокрема, лінійні *B*-сплайни є в основі популярного методу скінченних елементів.

**Означення 9.3.** Базисним сплайном, або *B*-сплайном, степеня  $m-1$  (з дефектом 1) відносно вузлів  $x_k < x_{k+1} < \dots < x_{k+m}$  називають функцію

$$B_{m-1,k}(x) = B_{m-1}(x_k; x_{k+1}; \dots, x_{k+m}; x) = m \sum_{i=k}^{k+m} \frac{(x_i - x)_+^{m-1}}{\omega'_{m,k}(x_i)}, \quad (9.15)$$

$$\text{де} \quad (x_i - x)_+^{m-1} = [\max\{0; (x_i - x)\}]^{m-1}; \quad (9.16)$$

$$\omega_{m,k}(x) = (x - x_k)(x - x_{k+1}) \dots (x - x_{k+m}). \quad (9.17)$$

Розглянемо, чим є *B*-сплайни від нульового до третього степеня у простішому, проте й у більш уживаному випадку рівновіддалених вузлів *B*-сплайна, тобто коли  $x_{k+i} = x_k + i\bar{h}$ , де  $i = 0, 1, \dots, m$ .

У цьому разі врахуємо позначення у формулі (9.17):  $\omega_{m,k}(x_i)$  – добуток різниць між вузлом  $x_i$  і рештою вузлів від  $x_k$  до  $x_{k+m}$ .

1. Нехай  $m = 1$ . Тоді з (9.15) маємо

$$B_{0,k}(x) = B_0(x_k; x_{k+1}; x) = \begin{cases} 0 & \text{при } x < x_k, \\ \frac{1}{h} & \text{при } x_k \leq x \leq x_{k+1}, \\ 0 & \text{при } x \geq x_{k+1}. \end{cases} \quad (9.18)$$

Отже, *B*-сплайн нульового степеня є функцією-“сходинкою” (рис. 9.1).

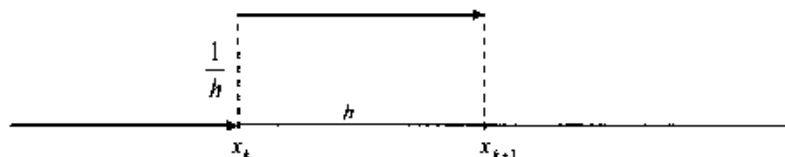


Рис. 9.1.

2. При  $m = 2$  з (9.15) випливає вираз лінійного  $B$ -сплайна:

$$B_{1,k}(x) = B_1(x_k; x_{k-1}; x_{k+2}; x) = \begin{cases} 0 & \text{при } x < x_k, \\ \frac{1}{h} + \frac{x - x_{k+1}}{h^2} & \text{при } x_k \leq x \leq x_{k+1}, \\ \frac{1}{h} - \frac{x - x_{k+1}}{h^2} & \text{при } x_{k+1} \leq x \leq x_{k+2}, \\ 0 & \text{при } x \geq x_{k+2}. \end{cases} \quad (9.19)$$

Геометричне зображення лінійного  $B$ -сплайна (рис. 9.2) – це так звана функція-кришка (інакше, функція-шалочка).

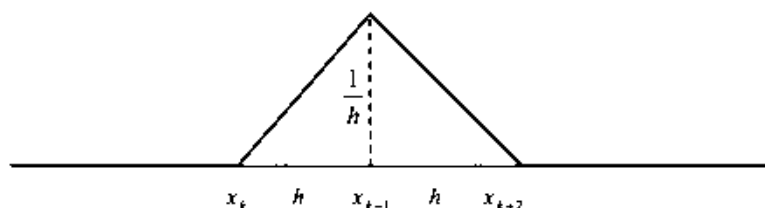


Рис. 9.2.

3. Приймаючи в (9.15)  $m = 3$  і виконавши нескладні перетворення, отримаємо квадратичний  $B$ -сплайн:

$$B_{2,k}(x) = B_2(x_k; x_{k+1}; x_{k+2}; x_{k+3}; x) = \begin{cases} 0 & \text{при } x \leq x_k, \\ \frac{1}{2h} - \frac{1}{h^2}(x_{k+1} - x) + \frac{1}{2h^3}(x_{k+1} - x)^2 & \text{при } x_k \leq x \leq x_{k+1}, \\ \frac{1}{2h} + \frac{1}{h^2}(x_{k+2} - x) + \frac{1}{h^3}(x_{k+2} - x)^2 & \text{при } x_{k+1} \leq x \leq x_{k+2}, \\ \frac{1}{2h^3}(x_{k+3} - x)^2 & \text{при } x_{k+2} \leq x \leq x_{k+3}, \\ 0 & \text{при } x \geq x_{k+3}. \end{cases} \quad (9.20)$$

Графік такого сплайна зображений на рис. 9.3.

4. Аналогічно при  $m = 4$  можна отримати з (9.15) кубічний  $B$ -сплайн



$$B_{3,k}(x) = B_3(x_k; x_{k+1}; x_{k+2}; x_{k+3}; x_{k+4}; x) =$$

$$= \begin{cases} 0 & \text{при } x \leq x_k, \\ \frac{1}{6h^4}(x-x_k)^3 & \text{при } x_k \leq x \leq x_{k+1}, \\ \frac{1}{6h} + \frac{1}{2h^2}(x-x_{k+1}) + \frac{1}{2h^3}(x-x_{k+1})^2 - \\ \quad - \frac{1}{2h^4}(x-x_{k+1})^3 & \text{при } x_{k+1} \leq x \leq x_{k+2}, \\ \frac{1}{6h} + \frac{1}{2h^2}(x_{k+3}-x) + \frac{1}{2h^3}(x_{k+3}-x)^2 - \\ \quad - \frac{1}{2h^4}(x_{k+3}-x)^3 & \text{при } x_{k+2} \leq x \leq x_{k+3}, \\ \frac{1}{6h^4}(x_{k+4}-x)^3 & \text{при } x_{k+3} \leq x \leq x_{k+4}, \\ 0 & \text{при } x \geq x_{k+4}. \end{cases} \quad (9.21)$$

Базисний сплайн третього степеня має неперервними не тільки перші похідні, а й другі. Графік цього сплайна зображений на рис. 9.4.

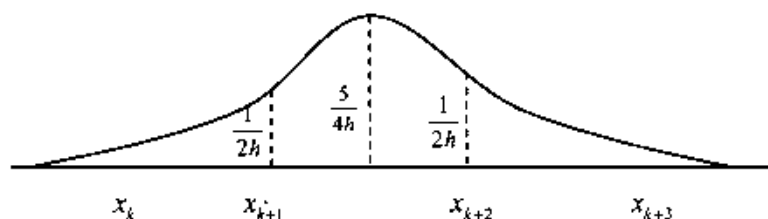


Рис. 9.3.

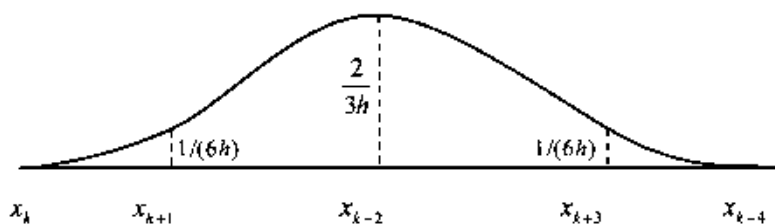


Рис. 9.4.

Як лінійно незалежні функції,  $B$ -сплайни певного степеня утворюють базис у відповідних функціональних просторах, що пояснює їхню назву і свідчить про можливість зображення через  $B$ -сплайни інших функцій цього простору.

## 9.4. Ермітові (локальні) сплайни

Припустимо, що в апроксимованій функції  $y = f(x)$  в точках  $x_0, x_1, \dots, x_n \in [a, b]$  відомі не тільки значення  $y_0, y_1, \dots, y_n$ , а й значення похідних до  $m$ -ї включно, тобто функція  $y = f(x)$  задана таблицею.

$x$	$y$	$y'$	...	$y^{(m)}$
$x_0$	$y_0$	$y'_0$	...	$y_0^{(m)}$
$x_1$	$y_1$	$y'_1$	...	$y_1^{(m)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_n$	$y_n$	$y'_n$	...	$y_n^{(m)}$

За цією інформацією можна побудувати єдиний інтерполяційний багаточлен Ерміта степеня  $(n+1)(m+1)-1$ . Навіть за порівняно невеликих значень  $n$  і  $m$  цей степінь може виявитися невідповідно високим.

Застосування в таких випадках інтерполяційних сплайнів високих степенів і малих дефектів дає потрібну гладкість інтерпольованої функції, тобто самого сплайна, однак не забезпечує узгодження його похідних із заданими похідними функції  $f(x)$ .

Тому виникає потреба в утворенні гібрида інтерполяційного сплайна й інтерполяційного багаточлена Ерміта. Такі гібриди називають *ермітовими сплайнами*.

Якщо дивитись на таблицю, то легко уявити, що за заданими в точках  $x_0$  і  $x_1$  значеннями функції  $f(x)$  та її  $m$  похідних можна побудувати єдиний багаточлен Ерміта  $H_{2m+1}^1(x)$  степеня  $2m+1$ , за заданими в точках  $x_1$  і  $x_2$  – багаточлен такого ж степеня  $H_{2m+1}^2(x)$  і так далі. Оскільки кожен з цих багаточленів-ланок ермітового сплайна

$$S_{2m+1}(x) := \left\{ H_{2m+1}^i(x), x \in [x_{i-1}, x_i] \right\}_{i=1}^n \quad (9.22)$$

побудований незалежно від решти, то такі сплайни називають *локальними сплайнами*. Властивість локальності ермітового сплайна дає певні переваги в разі його побудови, зокрема, у випадку великої кількості  $n$  елементарних проміжків  $[x_{i-1}, x_i]$ , а також дає змогу складати ермітів сплайн із ланок різних степенів, якщо на якихось ділянках проміжку інтерполяції  $[a, b]$  функція  $f(x)$  відображена похідними одного порядку, а на інших – іншого порядку. У цьому разі в точках стику таких ділянок можливе використання односторонніх похідних.

Розглянемо для прикладу ермітів сплайн третього степеня.

Нехай на  $i$ -му елементарному проміжку функція  $y = f(x)$  задана таблицею

$x$	$y$	$y'$
$x_{i-1}$	$y_{i-1}$	$y'_{i-1}$
$x_i$	$y_i$	$y'_i$

Прийнемо в (9.22)  $m=1$  і відшукаємо  $i$ -ту ланку сплайна  $S_3(x)$  у вигляді

$$H_3(x) := a_0^i + a_1^i(x - x_{i-1}) + a_2^i(x - x_{i-1})^2 + a_3^i(x - x_{i-1})^3 \quad (9.23)$$

Два рівняння для знаходження багаточлена  $H_3(x)$  отримаємо з інтерполяційних умов  $H_3(x_{i-1}) = y_{i-1}$ ,  $H_3(x_i) = y_i$ , а саме:

$$a_0^i = y_{i-1}, \quad (9.24)$$

$$a_0^i + a_1^i h_i + a_2^i h_i^2 + a_3^i h_i^3 = y_i, \quad (9.25)$$

де  $h_i = x_i - x_{i-1}$ . Далі продиференціюємо (9.23):

$$(H_3^i)'(x) = a_1^i + 2a_2^i(x - x_{i-1}) + 3a_3^i(x - x_{i-1})^2,$$

а з умов  $(H_3^i)'(x_{i-1}) = y'_{i-1}$ ,  $(H_3^i)'(x_i) = y'_i$  отримаємо ще два рівняння

$$a_1' = y_{i-1}', \quad (9.26)$$

$$a_1' + 2a_2'h_i + 3a_3'h_i^2 = y_i'. \quad (9.27)$$

З рівностей (9.24), (9.26) визначають перші коефіцієнти багаточлена  $H_3^i(x)$ . Підставивши їх у рівності (9.25), (9.27), знайдемо третій і четвертий коефіцієнти

$$a_2' = \frac{1}{h_i^2}(3y_i' - 3y_{i-1}' - 2h_i y_{i-1}' - h_i y_i'), \quad (9.28)$$

$$a_3' = \frac{1}{h_i^3}(2y_{i-1}' - 2y_i' + h_i y_{i-1}' - h_i y_i'). \quad (9.29)$$

Зазначимо, що формула (9.22) визначає ермітів сплайн непарного степеня. Для побудови ермітових сплайнів парного степеня формулювання задачі дещо інше.

**П р и к л а д 9.3.** Наблизити кубічним ермітовим сплайном функцію  $y = \cos \frac{\pi}{2}x$  на відрізку  $[-1, 1]$ , інформація про яку задана такою таблицею:

$x$	$y$	$y'$
-1	0	$\frac{\pi}{2}$
0	1	0
1	0	$-\frac{\pi}{2}$

*Розв'язування.* Прийmemo  $i=1$  та  $i=2$ , за формулами (9.24), (9.26), (9.28), (9.29) знаходимо коефіцієнти кубічного ермітового сплайна, підстановка яких у (9.23) дає

$$S_3(x) = \begin{cases} H_3^1(x) = \frac{\pi}{2}(x+1) + (3-\pi)(x+1)^2 + \left(-2 + \frac{\pi}{2}\right)(x+1)^3 & \text{при } x \in [-1, 0], \\ H_3^2(x) = 1 + \left(-3 + \frac{\pi}{2}\right)x^2 + \left(2 - \frac{\pi}{2}\right)x^3 & \text{при } x \in [0, 1]. \end{cases}$$

Точність апроксимації функції  $\cos \frac{\pi}{2}x$  та її похідної  $-\frac{\pi}{2} \sin \frac{\pi}{2}x$  за допомогою ермітового сплайна  $S_3(x)$  відобразимо в трьох контрольних точках  $x = -\frac{1}{2}, x = \frac{1}{3}, x = \frac{2}{3}$  такою таблицею:

$x$	$\cos \frac{\pi}{2}x$	$S_3(x)$	$\left( \cos \frac{\pi}{2}x \right)'$	$S_3'(x)$
$-1/2$	0,707	0,696	1,111	1,107
$1/3$	0,866	0,857	-0,785	-0,810
$2/3$	0,500	0,492	-1,360	-1,333

## Завдання до розділу 9

9.1. Побудувати лінійний інтерполяційний сплайн за значеннями  $f(0)$ ,  $f(1)$ .

9.2. Нехай відомі такі значення деякої функції  $f$ :

$$f(1) = 2, \quad f(2) = 3, \quad f(3) = 5, \quad f(4) = 3;$$

а) побудувати за цими даними інтерполяційний поліном третього степеня і записати його у вигляді  $a_0 + a_1x + a_2x^2 + a_3x^3$ ;

б) побудувати за цими даними квадратичний сплайн, який задовольняє умову  $p'(1) = 0$  (вказівка: почніть зліва);

в) побудувати за цими даними кубічний сплайн, який задовольняє умови  $p''(1) = 6$ ,  $p''(4) = -9$  (вказівка: попробуйте використати поліном пункту а).

9.3. Для функції з вправи 9.2 побудувати кубічний сплайн, який задовольняє умови  $p'(1) = 1$ ,  $p'(4) = -1$ .

9.4. Скласти програму побудови природного кубічного сплайна за заданим набором точок  $x_1 < x_2 < \dots < x_n$  і відповідним значенням функції  $y_1, y_2, \dots, y_n$ . Програма повинна спочатку розв'язувати тридіагональну систему (9.11), а потім використовувати формули (9.12), (9.13), (9.14). Скласти також програму обчислення значення цього кубічного сплайна в заданій точці  $x$ .

9.5. Побудувати лінійний сплайн  $S_1(x)$ , який інтерполює функцію  $f(x)$ , задану таблицею

$x$	0	1	2	3
$f(x)$	0	1	1	3

Виразити  $S_1(x)$  через лінійні  $B$ -сплайни.

9.6. За даними вправи 9.5 побудувати кубічний і квадратичний інтерполяційний сплайни дефекту 1, приймаючи  $f''(0) = f''(3) = 0$

## Розділ 10

### ЧИСЕЛЬНЕ ІНТЕГРУВАННЯ

#### 10.1. Задача чисельного інтегрування.

##### Квадратурні формули інтерполяційного типу

Нехай потрібно обчислити інтеграл

$$I = \int_a^b F(x) dx = \int_a^b p(x) f(x) dx, \quad (10.1)$$

де  $[a, b]$  – скінченний або нескінченний проміжок числової осі;  $p(x)$  – деяка функція, яку називають ваговою. Уважатимемо функцію  $p(x)$  фіксованою і такою, що добуток  $p(x)f(x)$  є абсолютно інтегровна функція на  $[a, b]$ .

На практиці у випадках, коли підінтегральна функція задана таблично або функція  $F(x)$  має особливості чи складний аналітичний вигляд, постає питання наближеного обчислення інтеграла (10.1).

Наближені формули обчислення інтеграла (10.1) ґрунтуються на апроксимації підінтегральної функції іншою функцією або алгебричним багаточленом, або раціональною функцією, або іншим способом.

У разі заміни підінтегральної функції деяким інтерполяційним багаточленом одержимо таку квадратурну формулу:

$$\int_a^b p(x) f(x) dx = \sum_{k=0}^n A_k^{(n)} f(x_k) + R_n(f), \quad (10.2)$$

де  $x_k$  – квадратурні вузли, або абсциси квадратурної формули ( $a \leq x_0 < x_1 < \dots < x_n \leq b$ ); числа  $A_k^{(n)}$  ( $k = \overline{0, n}$ ) – коефіцієнти квадра-

турної формули, що залежать тільки від вибору вузлів;  $R_n(f)$  – залишковий член квадратурної формули. Таку квадратурну формулу називають *інтерполяційною*. У випадку інтерполяційного багаточлена Лагранжа коефіцієнти  $A_k^{(n)}$  мають вигляд

$$A_k^{(n)} = \int_a^b \frac{p(x)\omega_{n+1}(x)}{(x-x_k)\dot{\omega}_{n+1}(x_k)} dx, \quad k = \overline{0, n} \quad (10.3)$$

$$\omega_{n+1}(x) = (x-x_0)(x-x_1)\dots(x-x_n). \quad (10.4)$$

Для залишкового члена формули (10.2) справджується оцінка

$$|R_n(f)| \leq \frac{M_{n+1}}{(n+1)!} \int_a^b |p(x)\omega_{n+1}(x)| dx, \quad (10.5)$$

де  $M_{n+1} = \max_{x \in [a,b]} |f^{(n+1)}(x)|$ . Оцінка (10.5) є точною, якщо добуток  $p(x)\omega_{n+1}(x)$  зберігає знак на  $[a, b]$ .

## 10.2. Квадратурні формули Ньютона–Котеса

Квадратурні формули інтерполяційного типу для рівновіддалених вузлів інтерполювання називають *формулами Ньютона–Котеса*. Вони мають вигляд

$$\int_a^b f(x) dx \approx (b-a) \sum_{k=0}^n C_k^{(n)} f(x_k), \quad (10.6)$$

де  $x_k = a + kh$ ,  $h = \frac{b-a}{n}$ ,  $C_k^{(n)}$  – коефіцієнти квадратурної формули, які обчислюють за формулою

$$C_k^{(n)} = \int_0^1 \frac{t \left(t - \frac{1}{n}\right) \dots \left(t - \frac{k-1}{n}\right) \left(t - \frac{k+1}{n}\right) \dots (t-1)}{n \left(\frac{k}{n} - \frac{1}{n}\right) \dots \left(\frac{k}{n} - \frac{k-1}{n}\right) \left(\frac{k}{n} - \frac{k+1}{n}\right) \dots \left(\frac{k}{n} - 1\right)} dt \quad (10.7)$$

На практиці найчастіше використовують формули Ньютона–Котеса при  $n = 0, 1, 2, 3$ ; їх називають, відповідно, формулами прямокутників, трапецій, Сімпсона (парабол), "трьох восьмих". Для



підвищення точності обчислення визначеного інтеграла відрізок  $[a, b]$ , як звичайно, розбивають на  $N$  однакових частин довжиною

$h = \frac{b-a}{N}$  і застосовують відповідну формулу до кожної частини.

Такі квадратурні формули називають узагальненими. Наведемо відповідні квадратурні формули та їхні залишкові члени.

*Квадратурна формула прямокутників ( $n = 0$ )* має вигляд

$$\int_a^b f(x) dx = (b-a) f\left(\frac{a+b}{2}\right) + \frac{(b-a)^3}{24} f''(\eta), \quad (10.8)$$

де  $\eta \in (a, b)$ .

*Узагальнена квадратурна формула середніх прямокутників* така:

$$\begin{aligned} \int_a^b f(x) dx = \frac{b-a}{N} & \left( f\left(a + \frac{h}{2}\right) + f\left(a + \frac{3h}{2}\right) + \dots + \right. \\ & \left. + f\left(a + \frac{2N-1}{2}h\right) \right) + \frac{(b-a)^3}{24N^2} f''(\eta), \end{aligned} \quad (10.9)$$

де  $\eta \in (a, b)$ ,  $N$  – кількість однакових частин, на які поділено  $[a, b]$ .

*Квадратурна формула трапецій ( $n = 1$ )*

$$\int_a^b f(x) dx = \frac{b-a}{2} (f(a) + f(b)) - \frac{1}{12} (b-a)^3 f''(\xi), \quad (10.10)$$

де  $\xi \in (a, b)$ .

*Узагальнена квадратурна формула трапецій*

$$\begin{aligned} \int_a^b f(x) dx = \frac{b-a}{2N} & (f(a) + 2f(a+h) + \dots + 2f(a+(N-1)h) + \\ & + f(b)) - \frac{1}{12N^2} (b-a)^3 f''(\xi), \end{aligned} \quad (10.11)$$

де  $\xi \in (a, b)$ .

Квадратурна формула Сімпсона, або формула парабол  
( $n = 2$ ),

$$\int_a^b f(x) dx = \frac{b-a}{6} \left( f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right) - \left(\frac{b-a}{2}\right)^5 \frac{f^{(4)}(\xi)}{90}, \quad (10.12)$$

де  $\xi \in (a, b)$ .

Узагальнена квадратурна формула парабол має вигляд

$$\begin{aligned} \int_a^b f(x) dx = \frac{b-a}{6M} & \left( f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \right. \\ & \left. + 2f(x_4) + \dots + 2f(x_{2M-2}) + 4f(x_{2M-1}) + f(x_{2M}) \right) - \\ & - \left( \frac{b-a}{2} \right)^5 \frac{f^{(4)}(\xi)}{90M^4}, \end{aligned} \quad (10.13)$$

де  $\xi \in (a, b)$ ,  $x_0 = a$ ,  $x_{2M} = b$ ,  $x_i = x_0 + ih$  ( $i = 0, 1, \dots, 2M$ )

Квадратурна формула "трьох восьмих"

$$\begin{aligned} \int_a^b f(x) dx = \frac{b-a}{8} & \left[ f(a) + 3f(a+h) + 3f(a+2h) + f(b) \right] - \\ & - \frac{(b-a)^5}{3^4 \cdot 80} f^{(4)}(\xi), \end{aligned} \quad (10.14)$$

де  $\xi \in (a, b)$ ,  $h = \frac{b-a}{3}$ .

Узагальнена квадратурна формула "трьох восьмих"

$$\begin{aligned} \int_a^b f(x) dx = \frac{3b-a}{8N} & \left[ (f(x_0) + f(x_N)) + 2(f(x_3) + f(x_6) + \dots + \right. \\ & \left. + f(x_{N-3})) + 3(f(x_1) + f(x_2) + f(x_4) + f(x_5) + \dots + f(x_{N-2}) + \right. \\ & \left. + f(x_{N-1})) \right] - \frac{(b-a)^5}{3^4 \cdot 80N^4} f^{(4)}(\xi), \end{aligned} \quad (10.15)$$

де  $\xi \in (a, b)$ ,  $N = 3k$ ,  $k \in \mathbb{Z}$ ,  $x_i = x_0 + ih$  ( $i = 0, 1, \dots, N$ ).

### 10.3. Принцип Рунге практичної оцінки похибки чисельного інтегрування

На практиці часто використовують так званий принцип Рунге для оцінки похибки формул чисельного інтегрування. Залишковий член кожної з квадратурних формул (10.8)–(10.15) можна записати у вигляді  $R = h^k M$ , де  $k$  – фіксоване число,  $h$  – крок інтегрування;  $M$  – добуток сталої на похідну підінтегральної функції  $(k-1)$ -го порядку в деякій точці проміжку інтегрування.

Нехай  $I$  – точне значення інтеграла,  $I_k(h)$  – наближене, тоді

$$I = I_k(h) + Ch^k \quad (10.16)$$

За тією ж формулою обчислимо інтеграл  $I$  проте за крок інтегрування візьмемо величину  $\frac{h}{2}$ , тоді

$$I = I_k\left(\frac{h}{2}\right) + C_1\left(\frac{h}{2}\right)^k \quad (10.17)$$

Якщо за малого  $h$  значення величин  $C$  і  $C_1$  близькі, то із формул (10.16) і (10.17) одержимо

$$I_k(h) + C(h)^k = I_k\left(\frac{h}{2}\right) + C_1\left(\frac{h}{2}\right)^k \approx I_k\left(\frac{h}{2}\right) + C\left(\frac{h}{2}\right)^k$$

Далі

$$C \approx C_1 \approx \frac{I_k\left(\frac{h}{2}\right) - I_k(h)}{h^k - \left(\frac{h}{2}\right)^k} \quad (10.18)$$

Підставимо значення (10.18) у формулу (10.17) й отримаємо

$$I \approx I_k\left(\frac{h}{2}\right) + \frac{I_k\left(\frac{h}{2}\right) - I_k(h)}{2^k - 1}. \quad (10.19)$$

На підставі подвійного розрахунку з кроком  $h$  і кроком  $\frac{h}{2}$  можна контролювати точність чисельного інтегрування, порівнюючи з заданою похибкою  $\varepsilon$  абсолютне значення виразу

$$\left| R_k \left( \frac{h}{2} \right) \right| \leq \left| I - I_k \left( \frac{h}{2} \right) \right| \approx \frac{\left| I_k \left( \frac{h}{2} \right) - I_k(h) \right|}{2^k - 1}. \quad (10.20)$$

У цьому і полягає принцип Рунге, який часто використовують під час практичного обчислення інтеграла із заданою точністю.

**П р и к л а д 10.1.** Обчислити з точністю  $\varepsilon = 10^{-6}$  такі інтеграли:

$$I_1 = \int_{0,5}^1 \frac{\cos^3 x}{\sin x} dx, \quad I_{1,\text{точне}} = \frac{1}{2} (\cos^2 1 - \cos^2 0,5) + \ln \left| \frac{\sin 1}{\sin 0,5} \right|,$$

$$I_2 = \int_3^4 \frac{dx}{(x+a)\sqrt{x^2-a^2}} \quad I_{2,\text{точне}} = \frac{1}{a} \left( \sqrt{\frac{4-a}{4+a}} - \sqrt{\frac{3-a}{3+a}} \right).$$

*Розв'язування.* Обидва інтеграли обчислені за різними квадратурними формулами і порівняні з точними значеннями. Результати обчислень наведені в табл. 10.1 і 10.2.

Таблиця 10.1

Значення інтеграла за різними квадратурними формулами

Інтеграл	Значення інтеграла за квадратурною формулою			
	прямокут- ників	трапецій	Сімпсона	"трьох восьмих"
$I_1$	0,3234505	0,3234508	0,3234506	0,3234506
$I_2, a=1$	0,0674896	0,0674899	0,0674898	0,0674898
$I_2, a=2$	0,0650680	0,0650684	0,0650683	0,0650683

Точні значення:  $I_1 = 0,3234506\dots$ ,  $I_2 = 0,674898\dots$  ( $a=1$ ),  
 $I_2 = 0,0650683\dots$  ( $a=2$ ).

**Кількість етапів поділу навіпіл кроку  $h$  за різними  
квадратурними формулами**

Інтеграл	Кількість кроків за квадратурною формулою			
	прямо- кутників	трапецій	Сімпсона	"трьох восьмих"
$I_1$	7	7	2	2
$I_2, a=1$	4	5	2	2
$I_2, a=2$	4	5	2	2

## 10.4. Квадратурна формула Гаусса

Нехай у формулі чисельного інтегрування

$$\int_a^b p(x)f(x)dx = \sum_{k=1}^n A_k^{(n)} f(x_k) + R_n(f) \quad (10.21)$$

вагова функція  $p(x)$  така, що

$$p(x) > 0, \text{ для } x \in [a, b] \quad \text{і} \quad \int_a^b p(x)|f(x)|dx - \text{існує.}$$

Формула (10.21) містить  $2n$  параметрів  $A_k^{(n)}, x_k$ , і вибрати їх можна так, щоб рівність (10.21) виконувалась точно для всіх алгебричних багаточленів степеня  $2n-1$ :

$$\int_a^b p(x)P_{2n-1}(x)dx = \sum_{k=1}^n A_k^{(n)} P_{2n-1}(x_k). \quad (10.22)$$

Виберемо  $p(x) \equiv 1$ . З умови (10.22) визначимо коефіцієнти

$$A_k^{(n)} = \frac{(n!)^4 (b-a)^{2n-1}}{((2n)!)^2 (x_k - a)(b - x_k) \omega_n'^2(x_k)} \quad (10.23)$$

й абсциси  $x_k$  як корені багаточлена

$$\omega_n(x) \equiv (x - x_1)(x - x_2) \dots (x - x_n).$$

Проміжок інтегрування  $[a, b]$  переводимо в  $[-1, 1]$  заміною змінних

$$x = \frac{1}{2}(a+b) + \frac{1}{2}(b-a)z \quad (10.24)$$

Тоді квадратурна формула Гаусса матиме вигляд

$$\int_a^b f(x)dx = \frac{b-a}{2} \int_{-1}^1 f(z)dz = \frac{b-a}{2} \sum_{k=1}^n A_k^{(n)} f(z_k) + R_n(f), \quad (10.25)$$

$$\text{де } R_n(f) = \frac{(b-a)^{2n+1} (n!)^4}{((2n)!)^3 (2n+1)} f^{(2n)}(\xi), \quad \xi \in (a, b).$$

Для кожного  $n$  необхідно знаходити коефіцієнти й абсциси квадратурної формули (10.25). Оскільки вони не залежать від підінтегральної функції, то їх можна протабулювати. В табл. 10.3 наведені абсциси і коефіцієнти формули Гаусса.

**Приклад 10.2.** З точністю  $\varepsilon = 10^{-6}$  обчислити інтеграли

$$J_1 = \int_3^4 \frac{dx}{(x+2)\sqrt{x^2-4}}, \quad J_2 = \int_{0,5}^1 \frac{\cos^2 x}{\sin x} dx$$

і порівняти з точними значеннями

$$J_1 = \frac{1}{2} \sqrt{\frac{x-2}{x+2}} \Big|_3^4 \approx 0,0650683368;$$

$$J_2 = \frac{\cos^2 x}{2} + \frac{1}{2} \ln |\sin x| \Big|_{0,5}^1 \approx 0,3234506545.$$

**Розв'язування.** Для обчислення інтегралів з точністю до  $\varepsilon$  скористаємося таким алгоритмом:

- 1) зафіксуємо порядок квадратурної формули, наприклад,  $n = 6$ ;
- 2) обчислимо інтеграл за формулою (10.25);
- 3) проміжок інтегрування  $[a, b]$  поділимо навпіл. До кожного з підпроміжків застосуємо квадратурну формулу (10.25) при  $n = 6$ . Обчислені значення підсумуємо, оскільки інтеграл можна розбити на суму інтегралів по підпроміжках;

4) порівняємо отриманий результат з подвоєною кількістю проміжків із попереднім. Якщо вони збігаються з точністю до  $\varepsilon$ , то вважаємо, що інтеграл обчислений із заданою точністю. Якщо ж ні, то знову подвоюємо кількість проміжків інтегрування  $m := 2m$ .

Таблиця 10.3

Абсциси і коефіцієнти квадратурної формули Гаусса

$n=1$	$z_1 = 0,5$	$A_1^{(1)} = 2$
$n=2$	$-z_1 = z_2 = 0,5773502$	$A_1^{(2)} = A_2^{(2)} = 1$
$n=3$	$-z_1 = z_3 = 0,7745967;$ $z_2 = 0$	$A_1^{(3)} = A_3^{(3)} = 0,55555556;$ $A_2^{(3)} = 0,88888889$
$n=4$	$-z_1 = z_4 = 0,8611363;$ $-z_2 = z_3 = 0,3399810$	$A_1^{(4)} = A_4^{(4)} = 0,3478548;$ $A_2^{(4)} = A_3^{(4)} = 0,6521452$
$n=5$	$-z_1 = z_5 = 0,9061798;$ $-z_2 = z_4 = 0,5384693;$ $z_3 = 0$	$A_1^{(5)} = A_5^{(5)} = 0,2369269;$ $A_2^{(5)} = A_4^{(5)} = 0,4786287;$ $A_3^{(5)} = 0,5688889;$
$n=6$	$-z_1 = z_6 = 0,9324695;$ $-z_2 = z_5 = 0,6612094;$ $-z_3 = z_4 = 0,2386192;$	$A_1^{(6)} = A_6^{(6)} = 0,1713245;$ $A_2^{(6)} = A_5^{(6)} = 0,3607616;$ $A_3^{(6)} = A_4^{(6)} = 0,4679139;$
$n=7$	$-z_1 = z_7 = 0,949108;$ $-z_2 = z_6 = 0,741531;$ $-z_3 = z_5 = 0,405845;$ $z_4 = 0$	$A_1^{(7)} = A_7^{(7)} = 0,129485;$ $A_2^{(7)} = A_6^{(7)} = 0,279705;$ $A_3^{(7)} = A_5^{(7)} = 0,381830;$ $A_4^{(7)} = 0,417960$
$n=8$	$-z_1 = z_8 = 0,960290;$ $-z_2 = z_7 = 0,796666;$ $-z_3 = z_6 = 0,525532;$ $-z_4 = z_5 = 0,183434$	$A_1^{(8)} = A_8^{(8)} = 0,101228;$ $A_2^{(8)} = A_7^{(8)} = 0,222381;$ $A_3^{(8)} = A_6^{(8)} = 0,313707;$ $A_4^{(8)} = A_5^{(8)} = 0,362684$

Зазначимо, що порядок квадратурної формули є фіксованим на кожному проміжку розбиття і його зручно вибирати однаковим на всіх підпроміжках.

Для обчислення першого інтеграла виконано чотири розбиття проміжку  $[3,4]$  на підпроміжки й отримано наближене значення  $J_1 = 0,065068336$ . Для другого інтеграла маємо  $J_2 \approx 0,323450655$ . Наближені значення збігаються з точними до восьми знаків після коми, оскільки були використані квадратурні формули Гауса найвищого алгебричного порядку точності.

## 10.5. Квадратурні формули з ваговими функціями

Квадратурну формулу (10.21) ефективно застосовують для функцій, що мають на кінцях проміжку  $[a,b]$  степеневі особливості, або коли апіорі відомо, що точки  $a$  і  $b$  є нулями функції  $f(x)$ , і відома кратність цих нулів. Лінійним перетворенням (10.24) відрізок  $[a,b]$  переведемо в  $[-1,1]$  і обмежимося розглядом інтеграла

$$I_1 = \int_{-1}^1 (1-x)^\alpha (1+x)^\beta f(x) dx \quad (\alpha, \beta > -1). \quad (10.26)$$

Системою багаточленів, ортогональних на  $[-1,1]$  з вагою  $p(x) = (1-x)^\alpha (1+x)^\beta$ , є система багаточленів Якобі. Для побудови квадратурної формули найвищої алгебричної точності

$$\int_{-1}^1 (1-x)^\alpha (1+x)^\beta f(x) dx \approx \sum_{k=1}^n C_k^{(\alpha)} f(x_k) \quad (10.27)$$

виберемо за вузли  $x_k$  корені багаточлена Якобі степеня  $n$ :  $P_n^{(\alpha, \beta)}(x_k) = 0$ ;  $(k = \overline{1, n})$ . Коефіцієнти  $C_k^{(\alpha)}$  можна відшукати за формулою

$$C_k^{(\alpha)} = 2^{\alpha+\beta+1} \frac{\Gamma(\alpha+n+1)\Gamma(\beta+n+1)}{n! \Gamma(\alpha+\beta+n+1) (1-x_k^2) \left[ P_n^{(\alpha, \beta)'}(x_k) \right]^2},$$

де  $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$ ,  $x > 0$  – інтеграл Ейлера другого роду.

Частковим випадком (10.27) є квадратурна формула Гауса при



$\alpha = \beta = 0$ . У випадку  $\alpha = \beta = -\frac{1}{2}$  вагова функція матиме вигляд

$p(x) = \frac{1}{\sqrt{1-x^2}}$ . Тоді вузлами квадратурної формули є нулі

багаточлена Чебишова першого роду  $x_k = \cos\left(\frac{2k-1}{2n}\pi\right)$

( $k=1, 2, \dots, n$ ). Коефіцієнти  $C_k^{(n)}$  будуть однакові й дорівнюватимуть  $\frac{\pi}{n}$

Отже, отримаємо квадратурну формулу Ерміта найвищого алгебричного порядку точності з вагою  $p(x) = \frac{1}{\sqrt{1-x^2}}$  на  $[-1, 1]$ :

$$\int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx = \frac{\pi}{n} \sum_{k=1}^n f\left(\cos \frac{2k-1}{2n} \pi\right) + R_n(f), \quad (10.28)$$

$$R_n(f) = \frac{\pi}{2^{2n-1} (2n)!} f^{(2n)}(\xi); \quad \xi \in [-1; 1].$$

**П р и к л а д 10.3.** Обчислити за квадратурною формулою Ерміта інтеграл

$$J_3 = \int_1^{\frac{1}{2}} \frac{dx}{\sqrt{1+x^2} \sqrt{1-x^2}}$$

**Розв'язування.** Цей інтеграл обчислено з точністю до  $\varepsilon = 10^{-6}$  за формулою (10.28) з подвоєнням кількості абсцис квадратурної формули і для  $n=18$  отримано наближене значення  $J_3 \approx 2,221441469$

## 10.6. Квадратурна формула Чебишова

Квадратурні формули наближеного інтегрування, всі коефіцієнти яких однакові, мають вигляд

$$\int_a^b p(x) f(x) dx \approx a_n \sum_{k=1}^n f(x_k), \quad (10.29)$$

Невідомі  $x_k$  ( $k = \overline{1, n}$ ) та  $a_n$  в (10.29) можна обчислити з умови точної рівності для всіх багаточленів степеня  $n$ , що рівносильно виконанню співвідношення

$$\int_a^b p(x)x' dx = a_n \sum_{k=1}^n x'_k, \quad i = \overline{1, n}. \quad (10.30)$$

Таблиця 10.4

Абсциси і коефіцієнти квадратурної формули Чебишова

$n=1$	$x_1 = 0$
$n=2$	$-x_1 = x_2 = 0,5773503;$
$n=3$	$-x_1 = x_3 = 0,7071068; \quad x_2 = 0;$
$n=4$	$-x_1 = x_4 = 0,7946545; \quad -x_2 = x_3 = 0,1875925;$
$n=5$	$-x_1 = x_5 = 0,8324975; \quad -x_2 = x_4 = 0,3745414; \quad x_3 = 0;$
$n=6$	$-x_1 = x_6 = 0,8662468; \quad -x_2 = x_5 = 0,4225187;$ $-x_3 = x_4 = 0,2666354;$
$n=7$	$-x_1 = x_7 = 0,8838617; \quad -x_2 = x_6 = 0,5296568;$ $-x_3 = x_5 = 0,3239118; \quad x_4 = 0;$
$n=9$	$-x_1 = x_9 = 0,9115893; \quad -x_2 = x_8 = 0,6010187;$ $-x_3 = x_7 = 0,5287618; \quad -x_4 = x_6 = 0,167906; \quad x_5 = 0$

Для проміжку інтегрування  $[-1, 1]$  формула (10.29) матиме вигляд

$$\int_{-1}^1 f(x) dx \approx a_n \sum_{k=1}^n f(x_k), \quad (10.31)$$

де  $a_n = \frac{2}{n}$ . Цю формулу називають *квадратурною формулою Чебишова*. Абсциси  $x_k$  обчислюють для кожного конкретного  $n$ . Їхні значення наведені в табл. 10.4. При  $n = 8, 10, \dots$  серед  $x_k$  будуть комплексні значення, тому ці випадки недоцільно розглядати.

**П р и к л а д 10.4.** За квадратурною формулою Чебишова обчислити з точністю до  $\varepsilon = 10^{-6}$  значення інтегралів

$$J_1 = \int_3^4 \frac{dx}{(x+2)\sqrt{x^2-4}} \quad \text{та} \quad J_4 = \int_0^1 \frac{dx}{1+x^2}.$$

**Розв'язування.** Для обчислення інтеграла  $J_1$  виконано розбиття на  $m=4$  підпроміжки з кількістю абсцис для  $n=6$  і отримано наближене значення  $J_1 \approx 0,650683$ . Для обчислення другого інтеграла виконано  $m=8$  розбиттів на підпроміжки. Одержано  $J_4 \approx 0,7853981808$ , яке сімома знаками збігається з точним значенням цього інтеграла  $\frac{\pi}{4} \approx 0,78539816134$ .

## 10.7. Формули Ейлера–Маклорена

Точність розглянутих вище квадратурних формул залежить від властивостей підінтегральної функції та вибраного квадратурного правила. Квадратурні формули побудовані за допомогою заміни функції  $f(x)$  алгебричним багаточленом і розраховані на інтегровність функцій з певним порядком гладкості. Наприклад, похибка квадратурної формули Гаусса з  $n$  вузлами залежить від того, наскільки точно  $f(x)$  на проміжку  $[a, b]$  може бути наближена багаточленом степеня  $2n-1$ .

Якщо ж розглядати похибку, зумовлену недостатньою точністю квадратурної формули, то до основної суми  $\sum_{k=1}^n a_k f(x_k)$  треба додати доданок, який є головною частиною залишку  $R_n(f)$ . Цей доданок повинен бути достатньо простим і легко обчислюваним. Така ідея закладена у формулі Ейлера–Маклорена. В різних задачах необхідно виділяти різну кількість головних частин і для побудови теорії уточнення правил ми повинні передбачити розклад  $R_n(f)$  у ряд із головних частин щораз більших порядків, а також уміти знайти залишок після виділення з  $R_n(f)$  довільної скінченної кількості членів такого розкладу.

Формула Ейлера–Маклорена має вигляд [12]

$$\begin{aligned} \int_a^{a+nh} f(x) dx &= \frac{h}{2} [f(a+nh) + f(a)] + h \sum_{k=1}^{n-1} f(a+kh) - \\ &- \sum_{v=1}^m \frac{h^{2v} B_{2v}}{(2v)!} [f^{(2v-1)}(a+nh) - f^{(2v-1)}(a)] + R_{2m}(f), \end{aligned} \quad (10.32)$$

де  $B_{2v}$  – числа Бернуллі;

$$R_{2m}(f) = -nh^{2m+3} \frac{B_{2m+2}}{(2m+2)!} f^{(2m+2)}(\xi), \quad \xi \in [a, a+nh]. \quad (10.33)$$

Розглянемо декілька доданків для похідних у (10.32) і для практичного використання одержимо таку формулу:

$$\begin{aligned} \int_a^{a+nh} f(x) dx \approx h \left( \frac{1}{2} f_1 + f_2 + f_3 + \dots + f_n + \frac{1}{2} f_{n+1} \right) - \\ - \frac{h^2}{12} (f'_{n+1} - f'_1) + \frac{h^4}{720} (f'''_{n+1} - f'''_1) - \\ - \frac{h^6}{30240} (f^{(5)}_{n+1} - f^{(5)}_1) + \frac{h^8}{1209600} (f^{(7)}_{n+1} - f^{(7)}_1). \end{aligned} \quad (10.34)$$

У формулі (10.34) доданків з непарними вищими степенями  $h$  нема, бо числа Бернуллі з непарними номерами дорівнюють нулю.

Часто формули Ейлера–Маклорена використовують для підсумовування значень функцій у точках рівномірного розбиття проміжку  $[a; a+nh]$ :

$$\begin{aligned} \sum_{k=0}^n f(a+kh) = \frac{1}{h} \int_a^{a+nh} f(x) dx + \frac{1}{2} (f_1 + f_{n+1}) + \\ + \sum_{v=1}^m \frac{B_{2v}}{(2v)!} h^{2v-1} (f^{(2v-1)}_{n+1} - f^{(2v-1)}_1) + R_{2m}(f), \end{aligned} \quad (10.35)$$

де

$$R_{2m}(f) = nh^{2m+2} \frac{B_{2m+2}}{(2m+2)!} f^{(2m+2)}(\xi). \quad (10.36)$$

**П р и к л а д 10.5.** За формулою Ейлера–Маклорена обчислити інтеграл  $F = \int_1^{13} \frac{dx}{x}$ .

**Розв'язування.** Проміжок інтегрування  $[1, 13]$  розіб'ємо на частини і на кожному підпроміжку застосуємо формулу Ейлера–Маклорена з певною кількістю доданків функції та її похідних. Величини  $n$  та  $m$  залежать від поведінки функції на конкретному підпроміжку. Для нашого прикладу використаємо три підпроміжки

$$F = \int_1^{13} \frac{dx}{x} = \int_1^2 \frac{dx}{x} + \int_2^4 \frac{dx}{x} + \int_4^{13} \frac{dx}{x}.$$

На перших двох з них виберемо  $n=10$ ,  $m=3$  з кроками  $h=0,1$  та  $h=0,2$ ; для третього  $n=9$ ,  $m=3$ ,  $h=1$ . Обчислимо похідні

$$f'(x) = -\frac{1}{x^2}; \quad f''(x) = -\frac{6}{x^4}; \quad f^{(5)}(x) = -\frac{120}{x^6}$$

Тоді за формулою (10.35) одержимо

$$\begin{aligned} F_1 &\approx \int_1^2 \frac{dx}{x} \approx 0,1 \left[ \frac{1}{2} f(1) + f(1,1) + f(1,2) + \dots + f(1,9) + \frac{1}{2} f(2) \right] - \\ &\quad - \frac{1}{1200} (f'(2) - f'(1)) + \frac{1}{7200000} (f''(2) - f''(1)) = \\ &= 0,1 \left[ \frac{1}{2} + \frac{1}{1,1} + \frac{1}{1,2} + \frac{1}{1,3} + \dots + \frac{1}{4} \right] + \frac{1}{1200} \left[ \frac{1}{2^2} - 1 \right] - \\ &\quad - \frac{1}{1200000} \left[ \frac{1}{2^4} - 1 \right] \approx 0,6937714 - 0,000625 + 0,0000007 \approx \\ &\quad \approx 0,6931471; \end{aligned}$$

$$\begin{aligned} F_2 &\approx \int_2^4 \frac{dx}{x} \approx 0,2 \left[ \frac{1}{2} f(2) + f(2,2) + f(2,4) + \dots + f(3,8) + \frac{1}{2} f(4) \right] - \\ &\quad - \frac{1}{300} (f'(4) - f'(2)) + \frac{1}{450000} (f''(4) - f''(2)) = \\ &= 0,1 \left[ \frac{1}{4} + \frac{1}{2,2} + \frac{1}{2,4} + \frac{1}{2,6} + \dots + \frac{1}{3,8} + \frac{1}{2} \cdot \frac{1}{4} \right] + \frac{1}{300} \left[ \frac{1}{4^2} - \frac{1}{2^2} \right] - \\ &\quad - \frac{1}{450000} \left[ \frac{1}{4^4} - \frac{1}{2^4} \right] \approx 0,6937713 - 0,000625 + 0,0000007 \approx 0,693147; \end{aligned}$$

$$\begin{aligned} F_3 &\approx \int_4^{13} \frac{dx}{x} \approx 1 \cdot \left[ \frac{1}{2} f(4) + f(5) + f(6) + \dots + f(12) + \frac{1}{2} f(13) \right] - \\ &\quad - \frac{1}{12} (f'(13) - f'(4)) + \frac{1}{720} (f''(13) - f''(4)) = \\ &= \left[ \frac{1}{2} \cdot \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \dots + \frac{1}{12} + \frac{1}{2} \cdot \frac{1}{13} \right] + \frac{1}{12} \left[ \frac{1}{13^2} - \frac{1}{4^2} \right] = \\ &= -\frac{1}{120} \left[ \frac{1}{13^4} - \frac{1}{4^4} \right] \approx 1,1833386 - 0,0047152 + 0,0000322 = 1,1785912. \end{aligned}$$

Остаточню  $F = F_1 + F_2 + F_3 \approx 2,5648853$ .

Точне значення інтеграла  $F = \ln 13 \approx 2,56494935$ .

**Приклад 10.6.** Обчислити суму значень функції  $f(x) = \frac{1}{x^2}$  на проміжку  $[51, 99]$  з кроком  $h = 2$ , тобто

$$S = \sum_{i=0}^{24} \frac{1}{x_i^2} = \sum_{i=0}^{24} \frac{1}{(51 + ih)^2}.$$

*Розв'язування.* Для обчислення  $S$  скористаємося формулою (10.35). Спочатку обчислимо похідні

$$f'(x) = -\frac{2}{x^3}; \quad f''(x) = -\frac{4!}{x^5}; \quad f^{(r)} = -\frac{6!}{x^7}.$$

Тоді

$$\begin{aligned} S &\approx \frac{1}{2} \left[ \int_{51}^{99} \frac{dx}{x^2} + \frac{1}{51^2} + \frac{1}{99^2} \right] + \frac{B_2}{2!} \cdot 2 \left[ -\frac{2}{99^3} + \frac{2}{51^3} \right] + \\ &+ \frac{B_4}{4!} \cdot 2^3 \left[ -\frac{4!}{99^5} + \frac{4!}{51^5} \right] + \frac{B_6}{6!} \cdot 2^5 \left[ -\frac{6!}{99^7} + \frac{6!}{51^7} \right] = \\ &= \frac{1}{2} \left[ -\frac{1}{x} \Big|_{51}^{99} + \frac{1}{51^2} + \frac{1}{99^2} \right] + \frac{1}{3} \left[ \frac{1}{51^3} - \frac{1}{99^3} \right] - \\ &- \frac{8}{30} \left[ \frac{1}{51^5} - \frac{1}{99^5} \right] + \frac{1}{42} \cdot 32 \left[ \frac{1}{51^7} - \frac{1}{99^7} \right] = \\ &= \frac{1}{2} \left[ -\frac{1}{x} \Big|_{51}^{99} + \frac{1}{51^2} + \frac{1}{99^2} \right] + \frac{1}{3} \left[ \frac{1}{51^3} - \frac{1}{99^3} \right] - \\ &- \frac{4}{15} \left[ \frac{1}{51^5} - \frac{1}{99^5} \right] + \frac{16}{21} \left[ \frac{1}{51^7} - \frac{1}{99^7} \right] \approx 4,996665 \cdot 10^{-3} + \\ &+ 2,16932 \cdot 10^{-6} - 7,448949 \cdot 10^{-10} + 8,40829 \cdot 10^{-13} \approx \\ &\approx 0,0049966. \end{aligned}$$

## 10.8. Обчислення невластних інтегралів

На практиці часто трапляються задачі, пов'язані з обчисленням невластних інтегралів. Це можуть бути інтеграли з нескінченними межами або інтеграли зі скінченними межами, проте

з підінтегральною функцією, яка перетворюється у нескінченність на відрізьку інтегрування.

Невласний інтеграл з нескінченними межами завжди можна перетворити у невласний інтеграл чи навіть власний зі скінченними межами. Для цього достатньо зробити відповідну заміну змінної під знаком інтеграла чи взяти інтеграл у скінченних, але достатньо великих межах так, щоб відкинута частина інтеграла була значно менша, ніж задана нами точність обчислення інтеграла.

Для обчислення невласних інтегралів зі скінченними межами інтегрування найзручніше використовувати *метод виділення особливостей*. Є кілька методів виділення особливостей. Ми розглянемо *мультиплікативний* та *адитивний* способи.

Суть мультиплікативного способу виділення особливостей полягає в такому. Нехай нам треба обчислити інтеграл

$$I = \int_a^b f(x) dx, \quad (10.37)$$

де функція  $f(x)$  перетворюється в нескінченність в одній або декількох точках відрізка  $[a, b]$ . Ми запишемо цю функцію у вигляді

$$f(x) = \phi(x)p(x), \quad (10.38)$$

де  $\phi(x)$  — обмежена функція на  $[a, b]$ , яка має там достатню кількість неперервних похідних, а  $p(x) > 0$  на  $[a, b]$ . Розглядають  $p(x)$  як вагову функцію і будують відповідну формулу чисельного інтегрування за допомогою тих же способів, які описані вище (наприклад, квадратурні формули типу Гаусса). За наближене значення інтеграла (10.37) приймають результат застосування отриманої формули чисельного інтегрування до функції  $\phi(x)$ .

Розглянемо *квадратурну формулу Гаусса–Кристоффеля* (або *типу Гаусса*, або просто *Гаусса*). Нею називають формулу найвищого алгебричного порядку точності вигляду

$$\int_a^b p(x)\phi(x)dx \approx \sum_{j=1}^n A_j^{(n)}\phi(x_j), \quad (10.39)$$

де межі інтегрування  $a$  і  $b$  можуть бути як скінченними, так і нескінченними, а вагова функція  $p(x)$  повинна задовольняти такі умови: вона повинна бути неперервна і додатна на інтервалі  $(a, b)$ ;

при  $x=a$  і  $x=b$  вона може перетворюватися в нуль або в нескінченність. У цьому разі повинен існувати  $\int_a^b p(x)dx$ . Очевидно,

на скінченному проміжку  $[a,b]$  ці вимоги задовольняє функція  $p(x) \equiv 1$ , з якою квадратурна формула Гаусса є частковим випадком формули (10.39).

Проаналізуємо конкретні квадратурні формули Гаусса-Кристоффеля, отримані на базі класичних ортогональних багаточленів.

Прийmemo у формулі (10.39)  $a=-1$ ,  $b=1$ ,  $p(x) = \frac{1}{\sqrt{1-x^2}}$  та візьmemo за вузли  $x_i$  корені багаточлена Чебишова  $T_n(x)$ . Тоді отримаемо формулу

$$\int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx \approx \sum_{i=1}^n A_i^{(n)} f(x_i), \quad (10.40)$$

яку називають *квадратурною формулою Ерміта*. Всі коефіцієнти у формулі (10.40) однакові між собою:  $A_i^{(n)} = \frac{\pi}{n}$ ,  $i=1,2,\dots,n$ , а

$x_i = \cos \frac{2i-1}{2n} \pi$  – корені багаточлена  $T_n(x)$ .

Якщо одна або обидві межі інтегрування нескінченні, то на базі ортогональних багаточленів Лагерра й Ерміта отримаємо, відповідно, формули

$$\int_0^{\infty} e^{-x} f(x) dx \approx \sum_{i=1}^n A_i^{(n)} f(x_i) \quad (10.41)$$

$$\int_{-\infty}^{+\infty} e^{-x^2} f(x) dx \approx \sum_{i=1}^n A_i^{(n)} f(x_i), \quad (10.42)$$

які іноді називають також *квадратурними формулами Лагерра й Ерміта*. За вузли в них приймають корені багаточленів Лагерра  $L_n(x)$  у (10.41) та Ерміта  $H_n(x)$  у (10.42), а вагові коефіцієнти  $A_i^{(n)}$  обчислюють за формулами

$$A_i^{(n)} = \int_0^{\infty} \frac{L_n(x)}{(x-x_i)L_n'(x_i)} e^{-x} dx = \left[ \frac{(n-1)!}{nL_{n-1}(x_i)} \right]^2 x_i$$



для квадратури (10.41) і

$$A_i^{(n)} = \int_{-\infty}^{+\infty} \frac{H_n(x)}{(x-x_i)H'_n(x_i)} e^{-x^2} dx = \frac{2^{n-1}(n-1)!\sqrt{\pi}}{nH_{n-1}(x_i)}$$

для (10.42). Формули (10.41), (10.42) точні для багаточленів степеня не вище  $2n-1$ .

Залишкові члени формул (10.41) і (10.42) мають, відповідно, вигляд

$$R_n^I = \frac{(n!)^2 \sqrt{\pi}}{(2n)!} f^{(2n)}(\xi_{IJ}) \quad \text{і} \quad R_n^E = \frac{n! \sqrt{\pi}}{2^n (2n)!} f^{(2n)}(\xi_E),$$

де  $\xi_{IJ} = (0, +\infty)$ ,  $\xi_E = (-\infty, +\infty)$

Числові дані про вузли і ваги квадратур Лагерра й Ерміта для значень від 1 до 6 наведені в табл. 10.5.

**П р и к л а д 10.7.** Обчислити інтеграли

$$I_1 = \int_0^{\infty} \frac{\sin 1,5x}{1,5+x} e^{-x} dx, \quad I_2 = \int_{-\infty}^{\infty} e^{-x^2} \cos x dx,$$

використовуючи квадратурні формули з вагами.

**Розв'язування.** Використаємо для обчислення інтеграла  $I_1$  квадратурну формулу Лагерра (10.41), а для  $I_2$  — квадратурну формулу Ерміта (10.42). Необхідні значення вузлів та коефіцієнтів беремо з табл. 10.5. У результаті обчислень отримано: для першого інтеграла: у випадку  $n=3$  наближене значення  $I_1 \approx 0,195078$ , а для  $n=5$  —  $I_1 \approx 0,218690$ ; для другого інтеграла у випадку  $n=3$  значення  $I_2 \approx 1,382033$ , а для  $n=6$  —  $I_2 \approx 1,3803884$ . Точне значення цього інтеграла  $I_2 = \sqrt{\pi} e^{-1/4} \approx 1,3803885$ .

Для обчислення інтеграла  $\int_0^{\infty} f(x) dx$  можна використати

квадратурну формулу Лагерра (10.41), виділивши під інтегралом відповідну вагову функцію

$$\int_0^{\infty} f(x) dx = \int_0^{\infty} e^{-x} e^x f(x) dx \approx \sum_{i=1}^n A_i^{(n)} e^{\xi_i} f(x_i). \quad (10.43)$$

Вузли і ваги квадратурних формул Лагерра й Ерміта

n	i	Формула Лагерра		Формула Ерміта	
		$x_i$	$A_i^{(n)}$	$x_i$	$A_i^{(n)}$
1	1	1	1	0	1,77245385
2	1	0,58578644	0,85355339	-0,70710678	0,88622693
	2	3,41421356	0,14644661	0,70710678	0,88622693
3	1	0,41577557	0,71109301	-1,22474487	0,29540896
	2	2,29428036	0,27851773	0	1,18163590
	3	6,28994508	0,01038926	1,22474487	0,29540896
4	1	0,32254769	0,60315410	-1,65068012	0,08131284
	2	1,74576101	0,35741869	-0,52464762	0,80491409
	3	4,53662030	0,03888791	0,52464762	0,80491409
	4	9,39507091	0,00053929	1,65068012	0,08131284
5	1	0,26356032	0,52175556	-2,02018287	0,01995324
	2	1,41340306	0,39866681	-0,5857246	0,39361932
	3	3,59642577	0,07594250	0	0,94530872
	4	7,08581001	0,00361176	0,95857246	0,39361932
	5	12,6408008	0,00002337	2,02018287	0,01995323
6	1	0,22284660	0,45896474	-2,35060497	0,00453001
	2	1,88893210	0,41700083	-1,3584907	0,15706732
	3	2,99273632	0,11337338	-0,43607741	0,72462960
	4	5,77514356	0,01039919	0,43607741	0,72462960
	5	9,83746741	0,00026101	1,33584907	0,15706732
	6	15,98287398	0,00000089	2,35060497	0,00453001

Коефіцієнти  $A_i^{(n)}$  і  $x_i$  відповідають формулі Лагерра. Загальніший спосіб інтегрування на проміжку  $[a, \infty)$  зводиться до розглянутого вище або лінійною заміною змінних, або зображенням

інтеграла  $\int_a^\infty f(x)dx$  у вигляді суми двох інтегралів

$$\int_a^0 f(x)dx + \int_0^\infty f(x)dx, \text{ перший з яких означений.}$$

Аналогічно для інтеграла  $\int_{-\infty}^{+\infty} f(x)dx$  можна отримати

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty} e^{-x^2} e^{x^2} f(x)dx \approx \sum_{i=1}^n A_i^{(n)} e^{x_i^2} f(x_i), \quad (10.44)$$

де  $A_i^{(n)}$  і  $x_i$  повинні відповідати квадратурній формулі Ерміта.

**Приклад 10.8.** Обчислити значення невласного інтеграла

$$I_3 = \int_0^{\infty} \ln \operatorname{th} \frac{x}{2} dx.$$

*Розв'язування.* Застосуємо до цього інтеграла формулу (10.43)

$$\text{у вигляді} \quad \int_0^{\infty} \ln \operatorname{th} \frac{x}{2} dx \approx \sum_{i=1}^n A_i^{(n)} \phi(x_i), \quad \text{де} \quad \phi(x) = e^x \ln \operatorname{th} \frac{x}{2} = \\ = \left( \ln \frac{1 - e^{-x}}{1 + e^{-x}} \right) / e^{-x} \quad \text{коефіцієнти } A_i^{(n)} \text{ і } x_i \text{ взяті з табл. 10.5.}$$

Для  $n = 5$  отримано результат  $I_3 \approx -2,353028$ . Точне значення інтеграла  $I_3 = -\frac{\pi^2}{4} \approx -2,467401$ . Отже, точність отриманого наближеного розв'язку невелика.

Для обчислення невласного збіжного інтеграла вигляду

$$\int_a^b f(x) dx$$

за наявності нескінченних розривів у функції  $f(x)$  або в точці  $x = a$ , або в точці  $x = b$ , або в загальному випадку в деякій точці  $x = c \in [a, b]$  Л. Канторович запропонував *адитивний спосіб виділення особливостей*. Суть способу полягає в тому, що цей невласний інтеграл записують у такому вигляді:

$$\int_a^b f(x) dx = \int_a^b g(x) dx + \int_a^b [f(x) - g(x)] dx, \quad (10.45)$$

де функцію  $g(x)$  підбирають так, щоб вона мала на  $[a, b]$  таку ж особливість, як і  $f(x)$ , і значення  $\int_a^b g(x) dx$  відшукували аналітичним способом, функція  $f(x) - g(x)$  повинна бути достатньо гладкою, тобто інтеграл  $\int_a^b [f(x) - g(x)] dx$  можна знайти звичайними квадратурами.

Для функцій деякого класу зі степеневими особливостями розроблено технологію підбирання таких функцій  $g(x)$ , які

переймають на себе особливість не тільки функції  $f(x)$ , а й деякої кількості її похідних, що дає змогу до другого інтеграла в (10.45) застосувати квадратурні формули заданого порядку. Розглянемо, як це можна зробити.

Нехай  $c \in [a, b]$  – особлива точка функції  $f(x)$  такої, що  $f(x) = \frac{\phi(x)}{(x-c)^\gamma}$ , де  $\gamma \in (0, 1)$ , а  $g(x) \in C^{k+1}[a, b]$ , і нехай для обчислення другого інтеграла в (10.45) хочемо використати квадратурну формулу, яка потребує від функції  $f(x) - g(x)$  неперервності  $k$ -ї похідної. Розглянемо функцію

$$\phi_{k+1}(x) = \phi(c) + \phi'(c)(x-c) + \dots + \frac{\phi^{(k+1)}(c)}{(k+1)!}(x-c)^{k+1} + \dots$$

Оскільки ця функція є відрізком розкладу Тейлора функції  $\phi(x)$  в околі особливої точки  $x = c$ , то, приймаючи  $g(x) = \frac{\phi_{k+1}(x)}{(x-c)^\gamma}$ , отримаємо рівність

$$f(x) - g(x) = \frac{\phi(x) - \phi_{k+1}(x)}{(x-c)^\gamma} = o[(x-c)^k],$$

яка означає неперервну диференційовність функції  $f(x) - g(x)$  не менше  $k$  разів.

**Приклад 10.9.** Обчислити інтеграл  $I = \int_0^{\frac{1}{2}} \frac{dx}{\sqrt{x(1-x)}}$

**Розв'язування.** Підінтегральна функція  $f(x) = \frac{1}{\sqrt{x(1-x)}}$  має степеневу особливість  $\left(\gamma = \frac{1}{2}\right)$  у точці  $x=0$  і задовольняє перелічені вище умови. Розкладемо функцію  $\phi(x) = \frac{1}{\sqrt{1-x}}$  в ряд Тейлора в околі точки  $x=0$

$$\phi(x) = 1 + \frac{1}{2}x + \frac{3}{8}x^2 + \frac{5}{16}x^3 + \frac{35}{128}x^4 + \dots$$

і прийемо

$$\phi_4(x) = 1 + \frac{1}{2}x + \frac{3}{8}x^2 + \frac{5}{16}x^3 + \frac{35}{128}x^4$$

Тоді

$$g(x) = \frac{\phi_4(x)}{x^{1/2}} = x^{-1/2} + \frac{1}{2}x^{1/2} + \frac{3}{8}x^{3/2} + \frac{5}{16}x^{5/2} + \frac{35}{128}x^{7/2},$$

а

$$\begin{aligned} f(x) - g(x) &= x^{1/2} [\phi(x) - \phi_4(x)] = \\ &= x^{1/2} \left[ \frac{1}{\sqrt{1-x}} - \left( 1 + \frac{1}{2}x + \frac{3}{8}x^2 + \frac{5}{16}x^3 + \frac{35}{128}x^4 \right) \right] \end{aligned} \quad (10.46)$$

Звідси

$$\begin{aligned} I_1 = \int_0^{1/2} g(x) dx &= \int_0^{1/2} x^{-1/2} dx + \frac{1}{2} \int_0^{1/2} x^{1/2} dx + \frac{3}{8} \int_0^{1/2} x^{3/2} dx + \frac{5}{16} \int_0^{1/2} x^{5/2} dx + \\ &+ \frac{35}{128} \int_0^{1/2} x^{7/2} dx = 1,5691585; \end{aligned}$$

а до інтеграла  $I_2 = \int_0^{1/2} [f(x) - g(x)] dx$  можна застосувати, наприклад, ефективний алгоритм Сімпсона, оскільки функція має, принаймні, неперервну третю похідну. З останнього зображення цієї функції в (10.46) бачимо  $f(0) - g(0) = 0$ . Обчислення

$$I_2 = \int_0^{1/2} [f(x) - g(x)] dx \text{ наведеним алгоритмом із заданою точністю } \varepsilon = 10^{-5} \text{ дає } I_2 = 0,0016385, \text{ звідки } I = I_1 + I_2 = 1,5707970.$$

Відзначимо, що інтеграл  $I = \int_0^{1/2} \frac{dx}{\sqrt{x(1-x)}}$  знаходиться елементарно і його точне значення є

$$I = \frac{\pi}{2} = 1,5707963.$$

Часто ефекту адитивного виділення особливості досягають інтегруванням частинами чи за допомогою заміни змінних.

У разі наближеного обчислення збіжних невласних інтегралів  $\int_a^{\infty} f(x)dx$ ,  $\int_a^b f(x)dx$  ( $c \in [a, b]$  – особлива точка) може виявитися корисним такий простий прийом, який ґрунтується на самому означенні цих інтегралів. Невласні інтеграли  $\int_a^{\infty} f(x)dx$  і  $\int_a^b f(x)dx$  можуть набути вигляду

$$\int_a^{\infty} f(x)dx = \int_a^A f(x)dx + \int_A^{\infty} f(x)dx,$$

$$\int_a^b f(x)dx = \int_a^{c-\delta_1} f(x)dx + \int_{c-\delta_1}^b f(x)dx + \int_{c-\delta_1}^{c+\delta_2} f(x)dx,$$

причому  $A$  вибирають настільки великим, а  $\delta_1$  і  $\delta_2$  – настільки малими, щоб у межах заданої точності інтеграли  $\int_A^{\infty} f(x)dx$ ,  $\int_{c-\delta_1}^{c+\delta_2} f(x)dx$  не впливали на результат. Інтеграли  $\int_a^A f(x)dx$ ,  $\int_a^{c-\delta_1} f(x)dx$ ,  $\int_{c+\delta_2}^b f(x)dx$  обчислюють уже відомими методами.

**Приклад 10.10.** Обчислити наближено інтеграл  $\int_a^{\infty} \frac{dx}{e^x + 1}$  з

точністю до  $\varepsilon = 10^{-6}$

**Розв'язування.** Виберемо число  $A$  так, щоб виконувалась нерівність  $\int_A^{\infty} \frac{dx}{e^x + 1} < \frac{\varepsilon}{2}$ .

Зауваживши, що  $\int_A^{\infty} \frac{dx}{e^x + 1} dx < \int_A^{\infty} \frac{dx}{e^x} < e^{-A}$ , вибираємо  $A$  з умови

$\frac{1}{e^A} = \frac{10^{-6}}{2}$ , звідки отримаємо  $A = 15$ . Приймаємо наближено

$\int_0^{\infty} \frac{dx}{e^x + 1} \approx \int_0^{15} \frac{dx}{e^x + 1} = I$  й обчислюємо означений інтеграл  $I$  з

точністю  $\frac{10^{-6}}{2}$  за формулою Сімпсона:  $\int_0^{\infty} \frac{dx}{e^x + 1} \approx 0,693147$ .

## 10.9. Наближене обчислення кратних інтегралів

### 10.9.1. Кубатурні формули

Розглянемо деякі методи наближеного обчислення кратних інтегралів. Вони ґрунтуються на побудові формул вигляду

$$I = \iint_G \dots \int f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n = \sum_{i=1}^N C_i f(P_i) + R(f),$$

які називають *кубатурними формулами*. Ці формули дають змогу відшукувати наближене значення інтеграла за допомогою лінійної комбінації значень підінтегральної функції в скінченній кількості точок  $P_1, P_2, \dots, P_N \in G$ .

Одним з найпростіших способів отримання кубатурних формул є прийом повторного застосування розглянутих раніше квадратурних формул для наближеного обчислення однократних інтегралів.

Нехай треба обчислити подвійний інтеграл

$$I = \iint_G f(x, y) dx dy, \quad (10.47)$$

де область інтегрування – прямокутник  $G = \{a \leq x \leq A, b \leq y \leq B\}$ , сторони якого паралельні до осей координат. Кожен із проміжків  $[a, A]$  і  $[b, B]$  розіб'ємо навпіл точками  $x_0 = a$ ,  $x_1 = a + h$ ,  $x_2 = a + 2h = A$  і, відповідно,  $y_0 = b$ ,  $y_1 = b + k$ ,  $y_2 = b + 2k = B$ , де

$h = \frac{A-a}{2}$ ,  $k = \frac{B-b}{2}$ . Усього таким способом отримаємо дев'ять

точок  $(x_i, y_i)$  ( $i, j = 0, 1, 2$ ). Тоді  $I = \iint_G f(x, y) dx dy = \int_a^A dx \int_b^B f(x, y) dy$

Звідси, обчислюючи внутрішній, а потім зовнішній інтеграл за квадратурною формулою Сімпсона, отримаємо

$$I = \iint_G f(x, y) dx dy \approx \frac{hk}{9} \{ [f(x_0, y_0) + f(x_2, y_0) + f(x_0, y_2) + f(x_2, y_2)] + 4[f(x_1, y_0) + f(x_0, y_1) + f(x_2, y_1) + f(x_1, y_2)] + 16f(x_1, y_1) \}. \quad (10.48)$$

Цю формулу називають кубатурною формулою Сімпсона.

Якщо розміри прямокутника  $G = \{a \leq x \leq A, b \leq y \leq B\}$  великі, то для збільшення точності кубатурної формули (10.48) область  $G$  розбивають на систему прямокутників, до кожного з яких застосовують кубатурну формулу Сімпсона. У підсумку отримують узагальнену кубатурну формулу Сімпсона

$$I = \iint_G f(x, y) dx dy \approx \frac{hk}{9} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} [ (f_{2i,2j} + f_{2i+2,2j} + f_{2i+2,2j+2} + f_{2i,2j+2}) + 4(f_{2i+1,2j} + f_{2i+2,2j+1} + f_{2i+1,2j-2} + f_{2i,2j-1}) + 16f_{2i-1,2j+1} ], \quad (10.49)$$

де

$$f_{i,j} = f(x_i, y_j), \quad x_i = x_0 + ih \quad \left( x_0 = a, h = \frac{A-a}{2n}; i = 0, \dots, 2n \right), \\ y_j = y_0 + jk \quad \left( y_0 = b, k = \frac{B-b}{2m}; j = 0, \dots, 2m \right).$$

**Приклад 10.11.** Із застосуванням кубатурної формули Сімпсона обчислити подвійний інтеграл

$$I = \int_2^6 \int_2^5 \frac{dx dy}{xy}$$

**Розв'язування.** Відповідні наближені значення цього інтеграла при різних  $m$  і  $n$  наведені в табл. 10.6. Точне значення  $I = \ln 1,5 \cdot \ln 2,5 \approx 0,37152392$ .



Таблиця 10.6

Наближене значення інтеграла  
для різних  $m$  і  $n$

$n$	$m$	$I_{\text{нобл}}$
1	1	0,37369048
2	3	0,37157483
4	6	0,37152739
10	15	0,37152401
20	30	0,37152393

Зазначимо, що на цьому шляху отримують кубатурні формули з великою кількістю вузлів. Тому доцільно використовувати квадратурні формули максимальної точності (з мінімальною кількістю вузлів), наприклад, формули Гаусса.

Недоліки методу:

- його зручно застосовувати тільки для прямокутних областей;
- зі зростанням кратності інтеграла різко збільшується обсяг обчислень;  
збільшення точності завдяки зменшенню кроків інтегрування значно збільшує обсяг обчислень.

### 10.9.2. Метод Монте-Карло

Для обчислення інтегралів великої кратності з малою точністю широко використовують метод статистичних випробувань, або метод Монте-Карло.

Нехай треба обчислити інтеграл

$$I = \iiint_G f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \quad (10.50)$$

за областю  $G$ , визначеною нерівностями

$$\begin{aligned} 0 \leq x_1 \leq 1, \\ 0 \leq \phi_i(x_1, x_2, \dots, x_{i-1}) \leq x_i \leq \psi_i(x_1, x_2, \dots, x_{i-1}) \leq 1 \\ (i = 2, 3, \dots, n). \end{aligned} \quad (10.51)$$

Припустимо, що ми вміємо отримати значення випадкової величини, рівномірно розподіленої на відріzkі  $[0, 1]$ . Випадкову величину  $\eta$  називають *рівномірно розподіленою* на  $[0, 1]$ , якщо вона набуває значення з цього інтервалу і ймовірність того, що вона набуде значення з інтервалу  $(t_i, t_{i+1}) \in [0, 1]$ , дорівнює  $t_{i+1} - t_i$ . Нехай потрібно

обчислити інтеграл (10.50) за областю (10.51), де функція  $f(x_1, x_2, \dots, x_n)$  обмежена й абсолютно інтегровна. Знову розглянемо сукупність  $N$  незалежних випадкових точок  $P_k = (\xi_1^{(k)}, \xi_2^{(k)}, \dots, \xi_n^{(k)})$ , координати яких є незалежними випадковими величинами, рівномірно розподіленими на відрізьку  $[0, 1]$ . Якщо точка  $P_k$  потрапила в область  $G$ , то обчислюють  $f(P_k)$ , якщо  $P_k \notin G$ , то приймаємо  $f(P_k) = 0$ .

Нехай із загальної кількості  $N$  випадкових точок  $n$  точок потрапили в область  $G$ , решта  $N - n$  виявилась поза  $G$ . Тоді за достатньо великого  $N$  справджується наближена формула

$$I \approx \frac{V_G}{n} \sum_{i=1}^n f(P_i), \text{ де під } V_G \text{ розуміємо } m\text{-вимірний об'єм області інтегрування.}$$

Якщо об'єм  $V_G$  обчислити важко, то можна прийняти

$$V_G \approx \frac{n}{N}, \text{ і для наближеного обчислення інтеграла отримаємо}$$

$$I \approx \frac{1}{N} \sum_{i=1}^n f(P_i).$$

Зазначимо, що в цьому способі ми ніби продовжуємо функцію  $f(x_1, x_2, \dots, x_n)$  на весь куб  $\{0 \leq x_i \leq 1, i = 1, 2, \dots, n\}$ , приймаючи, що вона дорівнює нулю поза областю  $G$ . У разі застосування цього методу виникає трудність у визначенні необхідного мінімального числа  $N$  випробувань [6].

Недоліками методу Монте-Карло є те, що для обчислення інтеграла з точністю  $\varepsilon$  необхідно обчислити значення підінтегральної функції в кількості точок  $N$  яка має порядок  $O(\varepsilon^{-2})$ . Ця кількість швидко зростає зі зменшенням  $\varepsilon$ . Крім того, збільшення гладкості підінтегральної функції не зменшує обсягу обчислень у випадку застосування методу Монте-Карло, що є його суттєвим недоліком у разі інтегрування гладких функцій.

Описаний спосіб можна застосувати до обчислення кратних інтегралів і для довільної області  $G$ , якщо існує така заміна змінних, за якої нова область інтегрування буде розміщена в  $m$ -вимірному одиничному кубі.

**П р и к л а д 10.12.** Методом Монте-Карло обчислити інтеграл  $I = \iint_G (x^2 + y) dx dy$ , де область  $G$  визначена такими нерівностями

(рис. 10.1):  $0 \leq x \leq 1, 0 \leq y \leq 3x - 2$ .

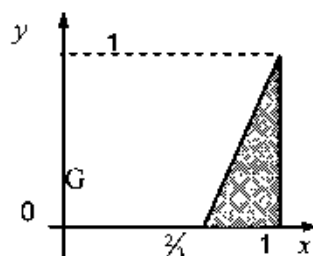


Рис. 10.1

**Розв'язування.** Область інтегрування належить одиничному квадрату  $0 \leq x \leq 1$ ,  $0 \leq y \leq 1$ .

Узявши загальну кількість точок  $N=100$  та виконавши десять разів обчислення інтегралу за методом Монте-Карло з використанням псевдовипадкових чисел, генерованих комп'ютером, ми отримали наближене значення інтеграла  $I_{\text{набл}} \approx 0,18889$ , яке відрізняється на 3 % від точного:

$$I = \iint_G (x^2 + y) dx dy = \frac{61}{324} \approx 0,18827$$

У табл. 10.7 наведено кількість точок  $n$ , які потрапили в область, та обчислене значення інтеграла  $I_k$  у разі  $k$ -го запуску програми. Як бачимо, точність отриманих результатів невисока. Вона практично не підвищується й зі збільшенням  $N$ .

Таблиця 10.7

**Значення інтеграла в разі  $k$ -го запуску програми**

$k$	$n$	Значення інтеграла $I_k$
1	23	0,20008
2	23	0,21005
3	32	0,18274
4	22	0,17209
5	29	0,18773
6	28	0,18755
7	22	0,19188
8	22	0,18038
9	20	0,17739
10	21	0,19898
$I_{\text{набл}} = \frac{1}{10} \sum_{k=1}^{10} I_k = 0,18889$		

## Завдання до розділу 10

10.1. Обчислити інтеграли за квадратурними формулами прямокутників, трапецій, Сімпсона і "трьох восьмих" з точністю  $\varepsilon = 10^{-6}$  і порівняти з точним значенням інтеграла для заданої первісної функції  $F(x)$ :

- 1)  $\int_1^2 \frac{dx}{x}, \quad F(x) = \ln|x|;$
- 2)  $\int_0^1 \frac{dx}{1+x^2}, \quad F(x) = \arctg x;$
- 3)  $\int_1^9 \sqrt{6x-5} \, dx, \quad F(x) = \frac{1}{9} \sqrt{(6x-5)^3};$
- 4)  $\int_{0,7}^{1,7} \frac{dx}{\sqrt{2x^2+0,3}}, \quad F(x) = \frac{1}{\sqrt{2}} \ln \left( x + \sqrt{x^2+0,15} \right);$
- 5)  $\int_0^{1,2} \ln(1+x^2) \, dx, \quad F(x) = x \ln(1+x^2) - 2(x - \arctg(x));$
- 6)  $\int_1^2 \ln x \, dx, \quad F(x) = x \ln x - x;$
- 7)  $\int_2^3 \ln^2 x \, dx, \quad F(x) = x \ln^2 x - 2x \ln x + 2x;$
- 8)  $\int_1^2 \frac{4xe^{2x}}{(1+2x)^2} \, dx, \quad F(x) = \frac{e^{2x}}{1+2x};$
- 9)  $\int_2^3 e^x \sin 2x \, dx, \quad F(x) = \frac{e^x}{5} (\sin 2x - 2 \cos 2x);$
- 10)  $\int_2^3 e^x \cos x \, dx, \quad F(x) = \frac{e^x}{2} (\cos x + \sin x);$
- 11)  $\int_2^3 \frac{dx}{1+e^{3x}}, \quad F(x) = \frac{1}{3} \ln \frac{e^{3x}}{1+e^{3x}};$
- 12)  $\int_2^3 \frac{e^x dx}{4+5e^x}, \quad F(x) = \frac{1}{5} \ln(4+5e^x);$

- 13)  $\int_0^{0,1} \lg^2 5x \, dx, \quad F(x) = \frac{\lg 5x}{5} - x;$
- 14)  $\int_0^1 \frac{dx}{1+x+x^2}, \quad F(x) = \frac{2}{\sqrt{3}} \operatorname{arctg} \frac{2x+1}{\sqrt{3}};$
- 15)  $\int_{1,4}^2 \frac{dx}{\sqrt{2x^2+0,7}}, \quad F(x) = \frac{1}{\sqrt{2}} \ln(x + \sqrt{x^2+0,35});$
- 16)  $\int_{1,3}^{2,1} \frac{dx}{\sqrt{5x^2-0,4}}, \quad F(x) = \frac{1}{\sqrt{5}} \ln(x + \sqrt{x^2-0,08});$
- 17)  $\int_1^2 x \sin(x) \, dx, \quad F(x) = \sin x - x \cos x;$
- 18)  $\int_1^2 \lg^3 x \, dx, \quad F(x) = \frac{1}{2} \lg^2 x + \ln |\cos x|;$
- 19)  $\int_0^{0,1} \frac{dx}{\cos^2 5x}, \quad F(x) = \frac{1}{5} \lg 5x;$
- 20)  $\int_0^2 \frac{\sin(2x)}{7 + \cos(2x)} \, dx, \quad F(x) = -\frac{1}{2} \ln(7 + \cos 2x);$
- 21)  $\int_0^{0,1} \frac{s \, dx}{9 \cos^2(3x) - \sin^2(3x)}, \quad F(x) = \frac{1}{18} \ln \left| \frac{\lg 3x + 3}{\lg 3x - 3} \right|;$
- 22)  $\int_0^1 \sin^2(5x) \cos^2(5x) \, dx, \quad F(x) = \frac{x}{8} - \frac{\sin 20x}{160};$
- 23)  $\int_0^1 \frac{e^{2x} \, dx}{4 + e^{2x}}, \quad F(x) = \frac{1}{2} \ln(4 + e^{2x});$
- 24)  $\int_{0,7}^{1,5} \frac{x+2}{\sqrt{x^2+1}} \, dx, \quad F(x) = \sqrt{x^2+1} + 2 \ln(x + \sqrt{x^2+1});$
- 25)  $\int_1^2 \frac{x \, dx}{(x^2+4)^2}, \quad F(x) = -\frac{1}{2(x^2+2)^2};$
- 26)  $\int_1^2 \frac{x^3 \, dx}{(4+x^2)^2}, \quad F(x) = \frac{2}{4+x^2} + \frac{1}{2} \ln(4+x^2);$

$$27) \int_1^2 x \sqrt{(8-x^2)^3} \, dx, \quad F(x) = -\frac{1}{5} \sqrt{(8-x^2)^5}$$

$$28) \int_3^4 \frac{x \, dx}{1+x}, \quad F(x) = x - \ln|1+x|;$$

$$29) \int_1^2 \cos^2\left(\frac{x}{2}\right) dx, \quad F(x) = \frac{1}{2}(x + \sin x);$$

$$30) \int_2^3 e^x \sqrt{e^x - 1} \, dx, \quad F(x) = \frac{2}{3} \sqrt{(e^x - 1)^3}$$

10.2. Обчислити такі інтеграли за квадратурними формулами Гаусса та Чебишова. Отримані розв'язки порівняти з точними значеннями для заданої первісної  $F(x)$ :

$$1) \int_2^3 \frac{x dx}{x+3}, \quad F(x) = x - 3 \ln|x+3|;$$

$$2) \int_1^2 \frac{x dx}{(x+2)^3}, \quad F(x) = -\frac{1}{x+2} + \frac{1}{(x+2)^2};$$

$$3) \int_5^{10} \frac{dx}{x(x+3)}, \quad F(x) = \frac{1}{3} \ln \left| \frac{x}{x+3} \right|;$$

$$4) \int_5^7 \frac{(x+5) dx}{x+2}, \quad F(x) = x - 3 \ln|x+2|;$$

$$5) \int_0^1 \frac{dx}{(2x+3)^{3/2}}, \quad F(x) = -\frac{1}{(2x+3)^{1/2}};$$

$$6) \int_6^7 \frac{x^3 dx}{\sqrt{x^2-25}}, \quad F(x) = \frac{1}{3} \sqrt{(x^2-25)^3} + 25 \sqrt{x^2-25};$$

$$7) \int_{-\pi/2}^1 \frac{dx}{\sin^3 x}, \quad F(x) = -\frac{1}{2} \frac{\cos x}{\sin^2 x} + \frac{1}{2} \ln \left| \tan \frac{x}{2} \right|;$$

$$8) \int_0^1 \frac{dx}{\cos^4 x}, \quad F(x) = \frac{1}{3} \operatorname{tg}^3 x + \operatorname{tg} x;$$

$$9) \int_{-1}^0 \frac{\sin^4 x}{\cos x} dx, \quad F(x) = -\frac{1}{3} \sin^3 x - \sin x + \ln \left| \operatorname{tg} \left( \frac{x}{2} + \frac{\pi}{4} \right) \right|;$$

- 10)  $\int_{-3}^{-2} \frac{\cos x}{\sin^3 x} dx,$   $F(x) = -\frac{1}{2 \sin^2 x};$
- 11)  $\int_1^2 \frac{\sin 3x}{\sin^3 x} dx,$   $F(x) = -3 \operatorname{ctg} x - 4x;$
- 12)  $\int_{0,1}^{0,3} \operatorname{tg}^3 4x dx,$   $F(x) = \frac{1}{8} \operatorname{tg}^2 4x + \frac{1}{4} \ln |\cos 4x|;$
- 13)  $\int_{-1}^1 \frac{\ln x}{(x+2)^2} dx,$   $F(x) = -\frac{\ln x}{x+2} + \frac{1}{2} \ln \frac{x}{x+2};$
- 14)  $\int_{0,1}^{0,2} \operatorname{ctg}^2(3x) dx$   $F(x) = -\frac{1}{3} \operatorname{ctg} 3x - x;$
- 15)  $\int_2^4 \frac{x^4 dx}{1+x^2},$   $F(x) = \operatorname{arctg} x + \frac{x^3}{3} - x;$
- 16)  $\int_0^3 \frac{x^3 dx}{(\sqrt{x^2+5^2})^5},$   $F(x) = -\frac{1}{\sqrt{x^2+5^2}} + \frac{a^2}{3(\sqrt{x^2+5^2})^3};$
- 17)  $\int_{10}^{11} \frac{dx}{x \ln x},$   $F(x) = \ln |\ln x|;$
- 18)  $\int_2^3 x^3 \ln x dx,$   $F(x) = \frac{x^4}{4} \ln x - \frac{x^4}{16};$
- 19)  $\int_1^2 \frac{(\ln 3x)^2}{x} dx$   $F(x) = \frac{1}{3} (\ln 3x)^3;$
- 20)  $\int_{0,5}^1 \operatorname{tg}^2 x dx,$   $F(x) = \operatorname{tg} x - x;$
- 21)  $\int_{0,5}^{1,5} \sin x \cos^2 x dx,$   $F(x) = -\frac{\cos^3 x}{3};$
- 22)  $\int_{0,4}^1 \frac{dx}{(1+\cos x)^2},$   $F(x) = \frac{1}{2} \operatorname{tg} \frac{x}{2} + \frac{1}{6} \operatorname{tg}^3 \frac{x}{2};$
- 23)  $\int_0^2 \frac{\cos x dx}{(1+\cos x)^2},$   $F(x) = \frac{1}{2} \operatorname{tg} \frac{x}{2} - \frac{1}{6} \operatorname{tg}^3 \frac{x}{2};$

- 24)  $\int_1^2 \frac{\cos x dx}{(1 - \cos x)^2}, \quad F(x) = \frac{1}{2} \operatorname{ctg} \frac{x}{2} - \frac{1}{6} \operatorname{ctg}^3 \frac{x}{2};$
- 25)  $\int_2^4 \frac{\cos x dx}{1 - \sin x}, \quad F(x) = -\ln(1 - \sin x);$
- 26)  $\int_{0.1}^{0.4} \frac{\sin x dx}{\cos x(1 + \cos x)}, \quad F(x) = \ln \left| \frac{1 + \cos x}{\cos x} \right|.$

10.3. Обчислити із заданою точністю невластні інтеграли:

- 1)  $\int_0^{\infty} \frac{\operatorname{arctg} x}{1+x^3} dx, \quad \varepsilon = 10^{-3}$
- 2)  $\int_0^{\infty} \frac{x e^{-x^2}}{2 + \sin x} dx, \quad \varepsilon = 10^{-6}$
- 3)  $\int_{-1}^{\infty} \frac{dx}{\sqrt[3]{x}}, \quad \varepsilon = 10^{-3};$
- 4)  $\int_{\frac{\pi}{2}}^{\infty} \frac{1}{x^2} \sin \frac{1}{x} dx, \quad \varepsilon = 10^{-3};$
- 5)  $\int_0^1 \ln x \ln(1+x) dx, \quad \varepsilon = 10^{-4};$
- 6)  $\int_0^1 \frac{\ln x}{1-x} dx, \quad \varepsilon = 10^{-3};$
- 7)  $\int_0^{\frac{\pi}{2}} \log \sin x dx, \quad \varepsilon = 10^{-3};$
- 8)  $\int_0^{\infty} \frac{dx}{\sqrt{1+ax^2}}, \quad \varepsilon = 10^{-4}, \quad a = 0,5 + 0,1k, \quad k = 0,1,\dots,10;$
- 9)  $\int_0^{\infty} \frac{e^{-x}}{a+x} dx, \quad \varepsilon = 10^{-6}, \quad a = 0,4 + 0,2k, \quad k = 0,1,\dots,12.$



10.4. Обчислити інтеграли, використовуючи квадратурні формули з вагою.

$$1) \int_0^{\infty} \frac{x e^{-x}}{x+2} dx, \quad n=5;$$

$$2) \int_0^{\infty} \frac{x+1}{x+2} e^{-x^2} dx, \quad n=5$$

$$3) \int_0^{\infty} \frac{\sin ax}{a+x} e^{-x} dx, \quad n=5, \quad a=0,5+0,1k, \quad k=0,1,\dots,10;$$

$$4) \int_0^{\infty} \frac{e^{-x}}{a+\sqrt{x}} dx, \quad n=5, \quad a=0,6+0,2k, \quad k=0,1,\dots,10;$$

$$5) \int_{-\infty}^{\infty} \frac{e^{-x^2}}{\sqrt{2+\sin ax}} dx, \quad n=5, \quad a=1,5+0,2k, \quad k=0,1,\dots,10;$$

$$6) \int_{-\infty}^{\infty} \frac{\cos ax}{a+x^2} e^{-x^2} dx, \quad n=5, \quad a=1,0+0,2k, \quad k=0,1,\dots,10;$$

$$7) \int_0^{\infty} \frac{2e^{-x}}{3+x^2} dx, \quad \varepsilon=10^{-5}$$

$$8) \int_{-\infty}^{\infty} \frac{e^{-x^2}}{\sqrt{1+x^2}} dx, \quad \varepsilon=10^{-4}$$

$$9) \int_{-\infty}^{\infty} \frac{e^{-x^2}}{1+x^2} dx, \quad \varepsilon=10^{-4}$$

$$10) \int_{-\infty}^{\infty} \frac{e^{-x} \sin x}{2+x} dx, \quad \varepsilon=10^{-3}$$

$$11) \int_{-1}^1 \frac{e^{2x}}{\sqrt{1-x^2}} dx, \quad n=5;$$

$$12) \int_0^1 \frac{e^{2x}}{(1+x)\sqrt{x}} dx, \quad n=5;$$

$$13) \int_0^1 \frac{dx}{\sqrt{x}(e^{0,4x}+1,5)}, \quad n=4;$$

$$14) \int_{-1}^1 \frac{dx}{(1+x)^2 \sqrt{1-x^2}}, \quad n=5;$$

- 15)  $\int_0^1 \frac{e^{a(x-1)}}{\sqrt{x}(x+b)} dx, \quad n=5, \quad a=0,60+0,07k, \quad k=0,1,\dots,8,$   
 $b=2,00+0,25k, \quad k=0,1,2,\dots,6;$
- 16)  $\int_{-1}^1 \frac{\cos ax}{\sqrt{1-x^2}(0,3+x^2)} dx, \quad n=12, \quad a=2,60+0,04k, \quad k=0, 1, \quad 10.$

10.5. Методом повторного інтегрування, застосовуючи різні квадратурні формули, обчислити інтеграли. Порівняти отримане значення з точним значенням інтеграла.

- 1)  $\int_0^2 dy \int_0^1 (x^2 + 2y) dx$  за формулою трапецій для  $n_x = 4, n_y = 8;$
- 2)  $\int_3^4 dx \int_1^2 \frac{dy}{(x+y)^2}$  за формулою Сімпсона для  $n_x = 4, n_y = 8;$
- 3)  $\int_3^4 dx \int_1^2 \frac{dy}{(x+y)^2}$  за формулою Сімпсона для  $n_x = 4, n_y = 8.$

10.6. Обчислити інтеграли методом Монте-Карло.

- 1)  $\iint_G \sqrt{x^2 - y^2} dx dy$ , де  $G$  – трикутник з вершинами  $O(0,0), A(1,0), B(1,1);$
- 2)  $\iint_G e^{y^2} dx dy$ , де  $G$  – криволінійний трикутник, обмежений параболою  $y^2 = x$  і прямими  $x=0, y=1;$
- 3)  $\iiint_G \frac{dx dy dz}{(x+y+z+1)^2},$  де  $G$  – область інтегрування, обмежена координатними площинами і площиною  $x+y+z=1.$

*Відповіді до заведань розділу 10*

*Відповіді до 10.1*

1. 0,69315;
2. 0,78540;
3. 32,90309;
4. 0,40413;
5. 0,42251;
6. 0,38629;
7. 0,84086;
8. 8,45661;
9. -9,65017;
10. -10,34699;
11. 0,00078;
12. 0,18725;
13. 0,00926;
14. 0,60460;
15. 0,23755;
16. 0,21782;
17. 1,44042;
18. 0,59713;
19. 0,10926;
20. 0,01837;
21. 0,01150;
22. 0,11929;
23. 0,41161;
24. 1,66631;
25. 0,03750;
26. 0,08500;
27. 19,52836;
28. 0,77686;
29. 0,53391;
30. 44,81975;

*Відповіді до 10.2*

1. 0,45304;
2. 0,03472;
3. 0,06921;
4. 2,75394;
5. 0,05602;
6. 66,58975;
7. -1,15699;
8. 2,81858;
9. 0,18611;
10. 24,50216;
11. -0,70075;
12. 0,57143;
13. 0,02945;
14. 0,49034;
15. 16,88534;
16. 0,00387;
17. 0,10791;
18. 15,41181;
19. 1,47543;
20. 0,51111;
21. 0,22517;
22. 0,19758;
23. 0,14912;
24. 0,38391;
25. -2,96366;
26. 0,03945.

## Розділ 11

### МЕТОДИ РОЗВ'ЯЗУВАННЯ ЗАДАЧІ КОШІ ДЛЯ ЗВИЧАЙНИХ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ

Клас задач Коші, інтегрованих у явному вигляді, є дуже малою часткою задач, які потрібно розв'язувати. Клас реально розв'язуваних диференціальних рівнянь, а отже, і сфера застосування математики суттєво розширилися з розробкою чисельних методів і активним використанням комп'ютерів.

Сьогодні затрати людської праці в разі розв'язування на комп'ютерах задачі Коші для звичайних диференціальних рівнянь складаються із затрат на вибір потрібних методів і програм, а також на підготовку даних, які необхідні для роботи програми. За бажанням можна безпосередньо з оперативної пам'яті комп'ютера отримати графік чи інше зображення розв'язку на екрані дисплея або на папері. Тому для багатьох категорій наукових працівників інтерес до вивчення часткових способів інтегрування звичайних диференціальних рівнянь у явному вигляді суттєво зменшився, а до розв'язування їх на комп'ютерах збільшився.

#### 11.1. Формулювання задачі

Нехай потрібно на відрізку  $[x_0, x_0 + a]$ , де  $a > 0$ , знайти розв'язок диференціального рівняння

$$y' = f(x, y), \quad (11.1)$$

який задовольняє початкову умову

$$y(x_0) = y_0. \quad (11.2)$$

Задачу (11.1)–(11.2) називають задачею Коші для звичайного диференціального рівняння першого порядку.

Будемо припускати, що функція  $f(x, y)$  неперервна і задовольняє умову Ліпшиця за  $y$ , тобто

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2|,$$

де  $L$  – деяка додатна стала. В цьому випадку задача Коші має єдиний розв'язок на проміжку  $[x_0, x_0 + a]$ .

## 11.2. Метод Ейлера та його модифікації

Розіб'ємо проміжок  $[a, b]$ , на якому шукаємо розв'язок, на відрізки рівномірного розбиття  $[x_{n-1}, x_n]$ ,  $n = 1, \dots, N$ , причому  $x_0 = a$ ,  $x_N = b$ .

Розрахункова формула методу Ейлера має вигляд

$$y_{n+1} = y_n + hf(x_n, y_n), \quad n = 0, 1, \dots, N-1, \quad (11.3)$$

де  $h = x_{n+1} - x_n$ . У випадку рівномірного розбиття відрізка  $[a, b]$  точками  $x_n = x_0 + nh$  ( $n = 0, 1, \dots, N$ ) отримаємо  $h = (b - a)/N$ . Похибка обчислень на кожному кроці цієї формули має порядок  $O(h)$ . Метод Ейлера є найпростішим чисельним методом інтегрування диференціального рівняння. Його недолік – мала точність. Проте доведено: якщо права частина  $f(x, y)$  рівняння (11.1) неперервна, то послідовність (11.3) при  $h \rightarrow 0$  на достатньо малому відрізку  $[a, b]$  рівномірно збігається до шуканої інтегральної кривої  $y = y(x)$ .

Розглянемо модифікації методу Ейлера, а саме – так званий метод Ейлера–Коші, за яким обчислення наближеного розв'язку  $y_n \approx y(x_n)$  виконують так: спочатку визначають грубе наближення

$$\bar{y}_{n+1} = y_n + hf(x_n, y_n)$$

далі знаходять точніше наближення

$$y_{n+1} = y_n + h \frac{f(x_n, y_n) + f(x_{n+1}, \bar{y}_{n+1})}{2}.$$

Метод Ейлера–Коші можна ще більше уточнити, застосовуючи ітераційне опрацювання кожного знайденого значення  $y_n$ . Спочатку обчислюють грубе наближення

$$y_{n+1}^{(0)} = y_n + hf(x_n, y_n),$$

а потім будують ітераційний процес за формулою

$$y_{n+1}^{(s+1)} = y_n + \frac{h}{2} [f(x_{n+1}, y_{n+1}^{(s)}) + f(x_n, y_n)],$$

де  $s$  – номер ітерації,  $s=0,1,\dots$ . Ітерації тривають доти, доки два послідовні наближення  $y_{n+1}^{(s+1)}$  і  $y_{n+1}^{(s)}$  не збіжаться із заданою точністю. Після цього за розв'язок приймають наближене значення  $y_{n+1} \approx y_{n+1}^{(s+1)}$ . Якщо після проведення трьох–чотирьох ітерацій у разі вибраного значення  $h$  не відбувається збігу потрібних знаків, то треба зменшити крок  $h$ .

### 11.3. Методи Рунге–Кутти

На практиці для розв'язування задачі Коші найчастіше використовують методи Рунге–Кутта підвищеної точності. Цими методами можна розв'язати задачу Коші для звичайного диференціального рівняння першого порядку, для диференціальних рівнянь вищих порядків і системи диференціальних рівнянь першого порядку.

Ідея побудови методів Рунге–Кутта  $p$ -го порядку полягає в отриманні наближень значень функції  $y(x_{i+1})$  за формулою

$$y_{i+1} = y_i + hf(x_i, y_i, h), \quad (11.4)$$

де  $\varphi(x, y, h)$  – деяка функція, яка наближає відрізок ряду Тейлора

$$y(x_{i+1}) = y(x_i) + hy'(x_i) + \frac{h^2}{2!} y''(x_i) + \dots + \frac{h^p}{p!} y^{(p)}(x_i) + O(h^{p+1}) \quad \text{до } p\text{-го}$$

порядку і не містить частинних похідних функції  $f(x, y)$ . Прийнемо у формулі (11.4)  $\varphi(x, y, h) = f(x, y)$ , отримаємо метод Ейлера ( $p=1$ ). Для побудови методів Рунге–Кутта порядку  $p > 1$  функцію  $\varphi(x, y, h)$  беремо багатопараметричною і параметри підбираємо за допомогою порівняння виразу (11.4) з багаточленом Тейлора для  $y(x)$  до  $p$ -го порядку.

Перевага методів Рунге–Кутта та, що обчислювальні алгоритми є однорідними, тобто не змінюються в разі переходу від однієї точки до

іншої; крок змінюється відповідно до потреби точності обчислень без ускладнення обчислювального алгоритму.

Методи Рунге–Кутта мають високу точність, обчислення можна виконувати зі змінним кроком: неважко зменшити крок там, де функція швидко змінюється, і збільшити крок у протилежному випадку.

Недоліком методів Рунге–Кутта є те, що наближений розв'язок у точці заданого відрізка знаходять за допомогою декількох обчислень значень функцій.

Наведемо рекурентні формули методу Рунге–Кутта різних порядків точності.

**А. Формули методу Рунге–Кутта другого порядку:**

$$1) \quad y_{n+1} = y_n + \frac{k_1 + k_2}{2}, \quad (11.5)$$

де

$$k_1 = hf(x_n, y_n), \quad k_2 = hf(x_n + h, y_n + k_1);$$

$$2) \quad y_{n+1} = y_n + k_2, \quad (11.6)$$

де

$$k_1 = hf(x_n, y_n), \quad k_2 = hf\left(x_n + \frac{h}{2}, y_n + \frac{k_1}{2}\right).$$

**Б. Формули методу Рунге–Кутта третього порядку:**

$$1) \quad y_{n+1} = y_n + \frac{1}{6}(k_1 + 4k_2 + k_3), \quad (11.7)$$

де

$$k_1 = hf(x_n, y_n), \quad k_2 = hf\left(x_n + \frac{h}{2}, y_n + \frac{k_1}{2}\right), \\ k_3 = hf(x_n + h, y_n - k_1 + 2k_2);$$

$$2) \quad y_{n+1} = y_n + \frac{1}{4}(k_1 + 3k_3), \quad (11.8)$$

де

$$k_1 = hf(x_n, y_n), \quad k_2 = hf\left(x_n + \frac{h}{3}, y_n + \frac{k_1}{3}\right), \\ k_3 = hf\left(x_n + \frac{2}{3}h, y_n + \frac{2}{3}k_2\right);$$

$$3) \quad y_{n+1} = y_n + \frac{1}{9}(2k_1 + 3k_2 + 4k_3), \quad (11.9)$$

де

$$k_1 = hf(x_n, y_n), \quad k_2 = hf\left(x_n + \frac{h}{2}, y_n + \frac{k_1}{2}\right),$$

$$k_3 = hf(x_n + \frac{3}{4}h, y_n + \frac{3}{4}k_2).$$

**В. Формули методу Рунге–Кутта четвертого порядку:**

$$1) \ y_{n+1} = y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \quad (11.10)$$

де

$$k_1 = hf(x_n, y_n), \quad k_2 = hf\left(x_n + \frac{h}{2}, y_n + \frac{k_1}{2}\right),$$

$$k_3 = hf\left(x_n + \frac{h}{2}, y_n + \frac{k_2}{2}\right), \quad k_4 = hf(x_n + h, y_n + k_3);$$

$$2) \ y_{n+1} = y_n + \frac{1}{6}(k_1 + 4k_3 + k_4), \quad (11.11)$$

де

$$k_1 = hf(x_n, y_n), \quad k_2 = hf\left(x_n + \frac{h}{4}, y_n + \frac{k_1}{4}\right),$$

$$k_3 = hf\left(x_n + \frac{h}{2}, y_n + \frac{k_2}{2}\right), \quad k_4 = hf(x_n + h, y_n + k_1 - 2k_2 + 2k_3);$$

$$3) \ y_{n+1} = y_n + \frac{1}{8}(k_1 + 3k_2 + 3k_3 + k_4), \quad (11.12)$$

де

$$k_1 = hf(x_n, y_n), \quad k_2 = hf\left(x_n + \frac{h}{3}, y_n + \frac{k_1}{3}\right),$$

$$k_3 = hf\left(x_n + \frac{2}{3}h, y_n - \frac{1}{3}k_1 + k_2\right), \quad k_4 = hf(x_n + h, y_n + k_1 - k_2 + k_3).$$

У випадку використання формул методу Рунге–Кутта виникають питання: якими з формул (11.5)–(11.12) доцільно користуватись у кожному конкретному разі і як вибирати крок сітки? Якщо права частина диференціального рівняння неперервна й обмежена разом зі своїми четвертими похідними, то добрі результати дає метод Рунге–Кутта четвертого порядку завдяки швидкому зростанню точності зі зменшенням кроку сітки. Якщо права частина не має обмежених четвертих похідних, то максимального порядку точності цієї схеми не можна досягнути. В цьому разі доцільно використовувати обчислювальні схеми методу меншого порядку точності, який відповідає порядку похідних.



## 11.4. Методи Адамса

Серед методів розв'язування задачі Коші для звичайного диференціального рівняння часто використовують різницеві методи Адамса. Для їхнього застосування поряд із заданими початковими умовами необхідно знати значення шуканого розв'язку  $y_k$  у деяких додаткових точках  $x_1, x_2, \dots, x_p$ .

Такі методи називають багатокроковими. Їх отримують з (11.1) на підставі квадратурних формул

$$y_{n+1} = y_n + \int_{x_n}^{x_{n+1}} p(x) dx, \quad (11.13)$$

де  $p(x)$  – поліном  $p$ -го степеня, який у  $p+1$  точках збігається зі значеннями функції  $p(x_k) = f(x_k, y_k) = f_k$ .

Отже, застосування різницевого методу розпадається на два етапи.

Необхідно спочатку знайти деяку кількість найближчих значень  $y_k$  і  $f_k$ , що можна зробити за допомогою методів Ейлера чи Рунге–Кутти;

З використанням багатокрокового методу обчислити наступні значення шуканої функції.

Якщо  $p(x)$  буде інтерполяційним багаточленом Ньютона інтерполювання назад, то отримаємо екстраполяційний метод Адамса:

$$y_{n+1} = y_n + h \sum_{k=0}^p \alpha_k \Delta^k f_{n-k}, \quad (11.14)$$

де

$$\alpha_k = \frac{1}{k!} \int_0^1 t(t+1)\dots(t+(k-1)) dt. \quad (11.15)$$

На практиці часто використовують екстраполяційну формулу Адамса вигляду

$$y_{n+1} = y_n + h \left( f_n + \frac{1}{2} \Delta f_n + \frac{5}{12} \Delta^2 f_{n-2} + \frac{3}{8} \Delta^3 f_{n-3} \right). \quad (11.16)$$

Зазначимо, що в екстраполяційній формулі Адамса (11.16) скінченні різниці третього порядку  $\Delta^3 f_n$  вважають сталими. Тому величину  $h$  початкового кроку обчислень можна визначати з нерівності  $h^p < E$ . На практиці стежать за значеннями скінченних різниць,

вибираючи  $h$  так, щоб сусідні різниці  $h\Delta^3 f_i$  і  $h\Delta^3 f_{i+1}$  відрізнялись між собою не більше, ніж на одну-дві одиниці заданого розряду (не враховуючи запасних знаків).

Заміною скінченних різниць у (11.14) через значення функцій у вузлах інтерполяції отримують іншу формулу екстраполяційного методу Адамса:

$$y_{n+1} = y_n + h \sum_{k=0}^p \beta_k f_{n-k}. \quad (11.17)$$

Залежно від кількості членів інтерполяційного багаточлена отримують формули Адамса різного ступеня точності. Залишковим членом є величина  $O(h^p)$ . Коефіцієнти методу найвищого порядку апроксимації для кожного конкретного  $p$  знаходять з системи

$$\sum_{k=1}^p k^{l-1} \beta_k = \frac{1}{l}; \quad l=1, 2, \dots, p.$$

Уживаними є формули Адамса при  $p=2, 3, 4, 5$ .

$$y_{n+1} = y_n + h \left( \frac{3}{2} f_n - \frac{1}{2} f_{n-1} \right), \quad p=2,$$

$$y_{n+1} = y_n + \frac{h}{12} (23 f_n - 16 f_{n-1} + 5 f_{n-2}), \quad p=3, \quad (11.18)$$

$$y_{n+1} = y_n + \frac{h}{24} (55 f_n - 59 f_{n-1} + 37 f_{n-2} - 9 f_{n-3}), \quad p=4,$$

$$y_{n+1} = y_n + \frac{h}{720} (1901 f_n - 2774 f_{n-1} + 2616 f_{n-2} - 1274 f_{n-3} + 251 f_{n-4}), \quad p=5.$$

Аналогічно отримують інтерполяційний метод Адамса. За вузли інтерполяції вибирають точки  $x_1, x_2, \dots, x_{p+1}$ . Формула (11.14) набуде вигляду

$$y_{n+1} = y_n + h \sum_{k=0}^p \alpha_k^* \Delta^k f(x_{n-k+1}, y_{n-k+1}) \quad (11.19)$$

або

$$y_{n+1} = y_n + h \sum_{k=0}^p \beta_k^* f(x_{n-k+1}, y_{n-k+1}) \quad (11.20)$$

Формули (11.19) та (11.20) визначають неявну схему методу Адамса. Надалі будемо вважати степені багаточлена сталим. Ці формули містять шукане значення як у лівій, так і в правій частині, причому  $f(x_{n-k+1}, y_{n-k+1})$  може бути нелінійною функцією. Тому для їхньої реалізації необхідно застосувати ітераційний процес:

$$y_{n+1}^{(s+1)} = y_n + h \alpha_0^* f(x_{n+1}, y_{n+1}^{(s)}) + h \sum_{k=1}^n \alpha_k^* \Delta^k f(x_{n-k+1}, y_{n-k+1}) \quad (11.21)$$

або

$$y_{n+1}^{(s+1)} = y_n + h \beta_0^* f(x_{n+1}, y_{n+1}^{(s)}) + h \sum_{k=1}^n \beta_k^* f(x_{n-k+1}, y_{n-k+1}) \quad (11.22)$$

$$s = 0, 1, 2, \dots$$

Ітераційний процес (11.21) та (11.22) збіжний.

Розрахункові формули неявного методу Адамса найвищого  $m = n+1$  порядку апроксимації для практичних обчислень отримують з (11.19), при  $m = 3$  вони мають вигляд

$$y_{n+1} = y_n + hf(x_{n+1}, y_{n+1}) - \frac{1}{2} \Delta f(x_n, y_n) - \frac{1}{24} \Delta^2 f(x_{n-1}, y_{n-1}) - \frac{19}{720} \Delta^3 f(x_{n-2}, y_{n-2}).$$

Із (11.20) одержують найуживаніші формули інтерполяційного методу Адамса:

$$y_{n+1} = y_n + \frac{h}{2} (f_{n+1} + f_n), \quad m=2,$$

$$y_{n+1} = y_n + \frac{h}{12} (5f_{n+1} + 8f_n - f_{n-1}), \quad m=3, \quad (11.23)$$

$$y_{n+1} = y_n + \frac{h}{24} (9f_{n+1} + 19f_n - 5f_{n-1} + f_{n-2}), \quad m=4,$$

$$y_{n+1} = y_n + \frac{h}{720} (25f_{n+1} + 646f_n - 264f_{n-1} + 106f_{n-2} - 19f_{n-3}), \quad m=5.$$

Як приклад, наведемо неявну схему методу Адамса четвертого порядку:

$$y_{n+1}^{(s+1)} = y_n + \frac{h}{24} \left( 9f(x_{n+1}, y_{n+1}^{(s)}) + 19f(x_n, y_n) - 5f(x_{n-1}, y_{n-1}) + f(x_{n-2}, y_{n-2}) \right), \quad (11.24)$$

де  $s$  – номер ітерації,  $s = 0, 1, \dots$ . За початкове значення  $y_{n+1}^{(0)}$  можна прийняти розв'язок, отриманий за допомогою явного методу Адамса третього порядку:

$$y_{n+1}^{(0)} = y_n + \frac{h}{12} \left( 23f(x_n, y_n) - 16f(x_{n-1}, y_{n-1}) + 5f(x_{n-2}, y_{n-2}) \right). \quad (11.25)$$

Запишемо (11.24) у вигляді

$$y_{n+1}^{(s+1)} = \frac{3h}{8} f(x_{n+1}, y_{n+1}^{(s)}) + F,$$

отримаємо таке: якщо  $\left| \frac{\partial f}{\partial y} \right| \leq M$ , то ітераційний процес збіжний за

умови  $\frac{3hM}{8} < 1$ , яка виконується за досить малих значень  $h$ .

Якщо в (11.21) чи (11.22) обмежимося лише однією ітерацією, то отримаємо метод, який називають предиктор-коректор. Хоча інтерполяційний метод Адамса більш трудомісткий порівняно з екстраполяційним, проте він дає вищу точність.

Для звичайних диференціальних рівнянь вищих порядків застосовують методику зниження порядку. В підсумку отримують систему диференціальних рівнянь першого порядку, для якої можна застосувати методи Ейлера, Рунге-Кутта чи Адамса.

Для диференціального рівняння другого порядку

$$y'' = f(x, y);$$

$$y(x_0) = y_0;$$

$$y'(x_0) = y'_0$$

часто застосовують методи Адамса-Штермера:  
*екстраполяційний*

$$y_{n+1} = 2y_n - y_{n-1} + h^2 \sum_{k=0}^n \bar{\alpha}_k \Delta^k f_{n-k}$$

або

$$y_{n+1} = 2y_n - y_{n-1} + h^2 \sum_{k=0}^n \bar{\beta}_k f_{n-k};$$

$$y_{n+1} = 2y_n - y_{n-1} + h^2 \sum_{k=0}^n \bar{\alpha}_k \Delta^k f_{n-k+1}$$

або

$$y_{n+1} = 2y_n - y_{n-1} + h^2 \sum_{k=0}^n \bar{\beta}_k f_{n-k+1}. \quad (11.26)$$

Для знаходження розв'язку за формулою інтерполяційного методу Адамса–Штермера треба застосувати ітераційний процес

$$y_{n+1}^{(s+1)} = 2y_n - y_{n-1} + h^2 \bar{\alpha}_0 f(x_{n+1}, y_{n+1}^{(s)}) + h^2 \sum_{k=1}^n \bar{\alpha}_k \Delta^k f_{n+1-k} \quad (11.27)$$

Ітераційний процес збіжний, якщо  $h < \frac{1}{\sqrt{|\bar{\alpha}_0|} \cdot L}$ ,

де  $L$  константа Ліпшиця для функції  $f(x, y')$ .

**П р и к л а д 11.1.** Розв'язати диференціальне рівняння

$y' = 1 + 2\frac{y}{x}$ , що задовольняє початкову умову  $y(1) = 0$ , і порівняти з

точним значенням  $y^* = x^2 - x$ .

*Розв'язування.*

x	Метод				
	Ейлера (11.4)	Рунге– Кутта (11.7)	Рунге– Кутта (11.10)	Адамса (11.18)- (11.23)	точне значення $x^2 - x$
1,0	0	0	0	0	0
1,1	0,1000	0,11601	0,10999	0,10999	0,11
1,2	0,21818	0,25385	0,23999	0,23999	0,24
1,3	0,35454	0,41352	0,38999	0,38999	0,39
1,4	0,50909	0,59504	0,55999	0,55999	0,56
1,5	0,68181	0,79843	0,74999	0,74999	0,75
1,6	0,87272	1,02368	0,95998	0,95999	0,96
1,7	1,08181	1,27081	1,18998	1,18999	1,19
1,8	1,30909	1,53982	1,43998	1,43999	1,44
1,9	1,55454	1,83072	1,70998	1,70999	1,71
2,0	1,81818	1,94351	1,99998	1,99999	2

## 11.5. Розв'язування задачі Коші для систем звичайних диференціальних рівнянь

Розглянемо задачу Коші для системи двох звичайних диференціальних рівнянь першого порядку

$$\begin{cases} y' = f(x, y, z), \\ z' = g(x, y, z), \end{cases} \quad (11.28)$$

$$\begin{cases} y(x_0) = y_0, \\ z(x_0) = z_0. \end{cases} \quad (11.29)$$

Тобто треба відшукати розв'язок системи рівнянь (11.28)  $y = y(x)$ ,  $z = z(x)$ , який задовольняє початкові умови (11.29). Відомо [29], що задача (11.28)–(11.29) має єдиний розв'язок на відрізку  $[x_0, x_0 + a]$ , якщо в деякій області  $G$   $\{x_0 \leq x \leq x_0 + a, |y - y_0| \leq b, |z - z_0| \leq c\}$  функції  $f(x, y, z)$  і  $g(x, y, z)$  неперервні і задовольняють умови Ліпшиця по  $y$  і  $z$ .

Для розв'язування задачі (11.28)–(11.29) можна використати ті ж методи, які застосовують у разі задачі Коші для звичайного диференціального рівняння. Зокрема, розрахункові формули методу Ейлера для задачі (11.28)–(11.29) мають вигляд

$$\begin{cases} y_{n+1} = y_n + hf(x_n, y_n, z_n), \\ z_{n+1} = z_n + hg(x_n, y_n, z_n), \end{cases} \quad (11.30)$$

де  $y_n \approx y(x_n)$ ,  $z_n \approx z(x_n)$ ,  $n = 0, 1, \dots, m-1$ .

Ці формули мають також похибку порядку  $O(h)$ .

Аналогічно використовують для систем рівнянь і формули методу Ейлера–Коші та методу Ейлера з ітераційним уточненням.

Наведемо формули методу Рунге–Кутта різних порядків для системи двох диференціальних рівнянь першого порядку.

А. Другого порядку:

$$\begin{cases} y_{n+1} = y_n + \frac{k_1 + k_2}{2}, \\ z_{n+1} = z_n + \frac{l_1 + l_2}{2}, \end{cases} \quad (11.31)$$

де

$$\begin{aligned}
k_1 &= hf(x_n, y_n, z_n), \\
l_1 &= hg(x_n, y_n, z_n), \\
k_2 &= hf(x_n + h, y_n + k_1, z_n + l_1), \\
l_2 &= hg(x_n + h, y_n + k_1, z_n + l_1).
\end{aligned}$$

Б. Третьего порядку:

$$\begin{cases} y_{n+1} = y_n + \frac{1}{6}(k_1 + 4k_2 + k_3), \\ z_{n+1} = z_n + \frac{1}{6}(l_1 + 4l_2 + l_3), \end{cases} \quad (11.32)$$

де

$$\begin{aligned}
k_1 &= hf(x_n, y_n, z_n), \\
l_1 &= hg(x_n, y_n, z_n), \\
k_2 &= hf(x_n + \frac{h}{2}, y_n + \frac{k_1}{2}, z_n + \frac{l_1}{2}), \\
l_2 &= hg(x_n + \frac{h}{2}, y_n + \frac{k_1}{2}, z_n + \frac{l_1}{2}), \\
k_3 &= hf(x_n + h, y_n - k_1 + 2k_2, z_n - l_1 + 2l_2), \\
l_3 &= hg(x_n + h, y_n - k_1 + 2k_2, z_n - l_1 + 2l_2).
\end{aligned}$$

В. Четвертого порядку:

$$\begin{cases} y_{n+1} = y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \\ z_{n+1} = z_n + \frac{1}{6}(l_1 + 2l_2 + 2l_3 + l_4), \end{cases} \quad (11.33)$$

де

$$\begin{aligned}
k_1 &= hf(x_n, y_n, z_n), \\
l_1 &= hg(x_n, y_n, z_n), \\
k_2 &= hf(x_n + \frac{h}{2}, y_n + \frac{k_1}{2}, z_n + \frac{l_1}{2}), \\
l_2 &= hg(x_n + \frac{h}{2}, y_n + \frac{k_1}{2}, z_n + \frac{l_1}{2}), \\
k_3 &= hf(x_n + \frac{h}{2}, y_n + \frac{k_2}{2}, z_n + \frac{l_2}{2}), \\
l_3 &= hg(x_n + \frac{h}{2}, y_n + \frac{k_2}{2}, z_n + \frac{l_2}{2}), \\
k_4 &= hf(x_n + h, y_n - k_1 + 2k_2 - k_3, z_n - l_1 + 2l_2 - l_3), \\
l_4 &= hg(x_n + h, y_n - k_1 + 2k_2 - k_3, z_n - l_1 + 2l_2 - l_3).
\end{aligned}$$

$$k_4 = hf(x_n + h, y_n + k_3, z_n + l_3),$$

$$l_4 = hg(x_n + h, y_n + k_3, z_n + l_3).$$

Для системи (11.28) можна застосувати екстраполяційний та інтерполяційний методи Адамса:

$$\begin{cases} y_{n+1} = y_n + h \sum_{k=0}^n \alpha_k \Delta^k f_{n-k}, \\ z_{n+1} = z_n + h \sum_{k=0}^n \alpha_k \Delta^k g_{n-k} \end{cases} \quad (11.34)$$

$$\begin{cases} y_{n+1} = y_n + h \sum_{k=0}^n \beta_k^* f_{n-k+1}, \\ z_{n+1} = z_n + h \sum_{k=0}^n \beta_k^* g_{n-k+1}. \end{cases} \quad (11.35)$$

Для інтерполяційного методу Адамса можна побудувати ітераційний процес:

$$\begin{cases} y_{n+1}^{(s+1)} = y_n + h \beta_0^* f(x_{n+1}, y_{n+1}^{(s)}, z_{n+1}^{(s)}) + h \sum_{k=1}^n \beta_k^* f_{n-k+1}, \\ z_{n+1}^{(s+1)} = z_n + h \beta_0^* g(x_{n+1}, y_{n+1}^{(s)}, z_{n+1}^{(s)}) + h \sum_{k=1}^n \beta_k^* g_{n-k+1}, \end{cases} \quad (11.36)$$

$s = 0, 1, 2, \dots$

## 11.6. Оцінка похибки наближеного розв'язку задачі Коші

Для методів Ейлера та його модифікацій, а також методів Рунге-Кутта, Адамса застосовують апріорні оцінки похибки наближеного розв'язку задачі Коші (11.1)–(11.2). Однак ці оцінки здебільшого значно завищені. Тому їхнє значення не стільки практичне, скільки теоретичне, бо з них безпосередньо випливає висновок про збіжність цих методів. Крім того, апріорні оцінки містять у собі низку сталих, для знаходження яких часто треба виконувати досить складні обчислення.

Тому, щоб оцінити похибку наближеного розв'язку задачі (11.1)–(11.2), намагаються використати інформацію, яку отримують у процесі чисельного розрахунку (такі оцінки називають апостеріорними). Найефективнішими є оцінки, що використовують подвійний перерахунок (принцип Рунге).

Розглянемо детальніше метод подвійного перерахунку, зокрема, такі три випадки.



1. Задано крок інтегрування  $h$  і треба визначити точні цифри наближеного розв'язку в кожній вузловій точці  $x_k$ .

2. Задано точність  $\varepsilon > 0$ , з якою треба обчислити наближений розв'язок задачі, добираючи належно як сам метод, так і крок інтегрування  $h$ .

3. Оцінити похибку  $\varepsilon_k = y_k - y(x_k)$ , де  $y_k$  і  $y(x_k)$  – відповідно, наближений і точний розв'язок задачі в кожній вузловій точці  $x_k$ .

Для реалізації подвійного перерахунку розв'язок (11.1)–(11.2) у кожній вузловій точці обчислюють двічі: з кроком  $h$  і  $\frac{h}{2}$ .

Позначатимемо їх, відповідно,  $y_k$  і  $y_k^*$ . Десяткові розряди наближень  $y_k$  і  $y_k^*$ , які збігаються між собою, вважають точними цифрами наближеного розв'язку в точці  $x_k$ .

Якщо наближений розв'язок задачі (11.1)–(11.2) треба обчислити з наперед заданою точністю  $\varepsilon > 0$ , то, використовуючи метод певного порядку точності, інтегрування з кроками  $h$  і  $\frac{h}{2}$  доцільно вести паралельно, щоби вчасно виявити неузгодженість між значеннями  $y_k$  і  $y(x_k)$  і, можливо, перейти до нового кроку. Якщо в точці  $x_k$  значення  $y_k$  і  $y_k^*$  задовольняють нерівність  $|y_k^* - y_k| < \varepsilon$ , то крок інтегрування для наступної точки  $x_{k+1}$  треба збільшити, наприклад, подвоїти. Якщо  $|y_k^* - y_k| \geq \varepsilon$ , то крок інтегрування ділять навпіл. Цим забезпечують автоматичний вибір кроку інтегрування.

Нарешті, наявність наближених значень  $y_k$  і  $y_k^*$ , обчислених, відповідно, з кроком  $h$  і  $\frac{h}{2}$ , дає змогу наближено оцінити похибку методу  $\varepsilon_k = y_k^* - y(x_k)$  у точці  $x_k$ . Щоб вивести таку оцінку похибки, припустимо, що виконуються такі умови:

1) на кожному кроці інтегрування  $h$  похибка методу приблизно пропорційна до  $h^{s+1}$  ( $s \geq 1$ ), де  $s$  – порядок точності методу;

2) похибка методу на кожному кроці інтегрування однакова;

3) на кожному наступному кроці інтегрування сумарна похибка методу охоплює також усі похибки, зроблені на попередніх кроках. Тому якщо  $y_1 - y(x_1) = Mh^{s+1}$ , де  $M$  – невідомий коефіцієнт пропорційності, то

$$y_2 - y(x_2) = 2Mh^{s+1},$$

$$y_3 - y(x_3) = 3Mh^{s+1},$$

$$y_n - y(x_n) = nMh^{s+1}$$

Отже, для похибки в точці  $x_k$  у разі інтегрування з кроком  $h$  маємо рівність

$$y_k - y(x_k) = kMh^{s+1}, \quad (11.37)$$

а в разі інтегрування з кроком  $\frac{h}{2}$  – рівність

$$y_k^* - y(x_k) = 2kM\left(\frac{h}{2}\right)^{s+1} \quad (11.38)$$

Якщо віднімемо почленно (11.38) від рівності (11.37) і розв'яжемо отриману рівність щодо невідомого коефіцієнта  $M$ , то знайдемо

$$M = \frac{2^s(y_k - y_k^*)}{kh^{s+1}(2^s - 1)}$$

Підставивши це значення  $M$  в (11.38), отримаємо

$$y_k^* - y(x_k) = \frac{y_k - y_k^*}{2^s - 1}.$$

## Завдання до розділу 11

### 11.1. Задача Коші для одного рівняння першого порядку.

А. Із застосуванням чисельних методів розв'язування задачі Коші розв'язати диференціальне рівняння першого порядку з точністю  $\varepsilon = 0,0001$  із заданими початковими умовами  $x_0, y_0$  на відрізку  $[x_0, x_0 + 0,5]$  з кроком  $h = 0,05$ , використовуючи принцип Рунге, і порівняти з точним розв'язком  $y^*$ .

Б. Екстраполяційним методом Адамса знайти розв'язок у точці  $x_i$ , уточнити цей розв'язок інтерполяційним методом Адамса з точністю до  $\varepsilon$ , а тоді шукати розв'язок у точці  $x_{i+1}$  за такою ж схемою.

1)  $y' = \cos(y - x),$

$$y(1) = 2,57,$$

$$y^* = x + 2 \operatorname{arctg} \frac{1}{x};$$

2)  $y' = -\frac{2xy^2}{x^2 - 1},$

$$y(0) = 1,$$

$$y^* = \frac{1}{\ln|1 - x^2| + 1};$$

3)  $y' = (2 - y) \cdot \operatorname{tg} x,$

$$y\left(\frac{\pi}{3}\right) = 0,$$

$$y^* = 2 - 4 \cos x;$$

4)  $y' = \frac{y^2}{x} - \frac{y}{x},$

$$y(1) = 0,5,$$

$$y^* = \frac{1}{1 + x}$$

5)  $y' = y + 2x - 3,$

$$y(0) = 3,$$

$$y^* = 2e^x - 2x + 1$$

6)  $y' = \frac{1}{x + 2y},$

$$y(0) = -1,$$

$$y^* = -\frac{x+2}{2}$$

7)  $y' = \frac{2xy}{x^2 + y^2},$

$$y(0)=1,$$

$$8) \quad y' = \frac{y}{x} - e^{\frac{y}{x}}$$

$$y(e)=0,$$

$$9) \quad y' = \frac{y(2x^2 - y^2)}{2x^3},$$

$$y(2)=2,$$

$$10) \quad y' = \frac{y}{2x} + \frac{y^3}{2x^2},$$

$$y(1)=0,5,$$

$$11) \quad y' = y^2 - \frac{2}{x^2},$$

$$y(1)=-2,$$

$$12) \quad y' - y^2 e^x - 2y,$$

$$y(0)=0,5,$$

$$13) \quad y' = \frac{x}{y^2} + \frac{y}{x},$$

$$y(1)=2,$$

$$14) \quad y' = \frac{4x+2y}{2x+1},$$

$$y(0)=2,$$

$$15) \quad y' = y + \frac{e^x}{x},$$

$$y(1)=0,$$

$$16) \quad y' = \frac{y}{x+y^2},$$

$$y(2)=1,$$

$$y' = (1 + \sqrt{1+4x^2})/2;$$

$$y' = -x \ln \ln x;$$

$$y' = \frac{x}{\sqrt{\ln \frac{x}{2} + 1}};$$

$$y' = \sqrt{\frac{x}{4 - \ln x}};$$

$$y' = -\frac{2}{x};$$

$$y' = \frac{1}{e^x + e^{2x}};$$

$$y' = \sqrt[3]{12x^3 - 3x^2}$$

$$y' = (2x+1)(1 + \ln |2x+1|) + 1;$$

$$y' = e^x \cdot \ln |x|;$$

$$y' = (-1 + \sqrt{4x+1})/2;$$

$$17) y' = \frac{y^2 + x}{xy},$$

$$y(1) = 2,$$

$$y^* = \sqrt{2x(3x-1)};$$

$$18) y' = y\left(\frac{1}{x} - 1\right),$$

$$y(1) = 1,$$

$$y^* = \frac{2x}{x^2 + 1};$$

$$19) y' = \frac{1 - y \sin x}{\cos x},$$

$$y(0) = 2,$$

$$y^* = 2 \cos x + \sin x;$$

$$20) y' = xy^3 - y,$$

$$y(0) = 0.4,$$

$$y^* = \frac{1}{\sqrt{5,75e^{2x} + x + 0,5}};$$

$$21) y' = 2\sqrt{y} \frac{\sin x}{x} - 2\frac{y}{x},$$

$$y\left(\frac{\pi}{6}\right) = \left(\frac{3}{\pi}\right)^2$$

$$y^* = \left(\frac{\sin x}{x}\right)^2;$$

$$22) y' = -1,5\frac{y}{x} - 3,5\frac{1}{x^3y},$$

$$y(1) = -6,$$

$$y^* = -\frac{7}{x^2} + \frac{1}{x^3};$$

$$23) y' = \frac{y}{x} - \dot{x} - y,$$

$$y(1) = 0,$$

$$y^* = x(e^{-x+1} - 1);$$

$$24) y' = \frac{1 - y^2}{2x},$$

$$y(1) = 0,$$

$$y^* = \frac{x-1}{x+1};$$

$$25) y' = xy^3 - xy,$$

$$y(1) = 0,5,$$

$$y^* = \frac{1}{\sqrt{3e^{x^2-1} + 1}};$$

## 11.2. Задачі Коші для системи рівнянь першого порядку.

Із застосуванням чисельних методів розв'язування задачі Коші розв'язати системи диференціальних рівнянь з точністю  $\varepsilon = 0,01$  із заданими початковими умовами  $x_0, y_0, z_0$  на відрізку  $[a, b]$  з кроком  $h$ , використовуючи принцип Рунге, і порівняти з точним розв'язком  $y^*, z^*$ .

$$1) \quad \begin{cases} y' = 2y + z \\ z' = 3y + 4z \end{cases} \quad y(0) = 0, \quad z(0) = 4, \quad [0; 0,5], \quad h = 0,01; \\ \begin{cases} y^* = -e^x + e^{5x}, \\ z^* = e^x + 3e^{5x}; \end{cases}$$

$$2) \quad \begin{cases} y' = y - z \\ z' = z - 4y \end{cases} \quad y(0) = 0, \quad z(0) = -4, \quad [0; 0,5], \quad h = 0,01; \\ \begin{cases} y^* = -e^{-x} + e^{3x}, \\ z^* = -2e^{-x} - 2e^{3x}; \end{cases}$$

$$3) \quad \begin{cases} y' = 8z - y \\ z' = y + z \end{cases} \quad y(0) = 6, \quad z(0) = 0, \quad [0; 0,1], \quad h = 0,01; \\ \begin{cases} y^* = 2e^{3x} + 4e^{-3x}, \\ z^* = e^{3x} - e^{-3x}; \end{cases}$$

$$4) \quad \begin{cases} y' = y + z \\ z' = 3z - 2y \end{cases} \quad y(0) = 1, \quad z(0) = 2, \quad [0; 0,1], \quad h = 0,01; \\ \begin{cases} y^* = e^{2x} (\cos x + \sin x), \\ z^* = 2e^{2x} \cos x; \end{cases}$$

$$5) \quad \begin{cases} y' = y - 3z \\ z' = 3y + z \end{cases} \quad y(0) = 2, \quad z(0) = -1, \quad [0; 0,1], \quad h = 0,01; \\ \begin{cases} y^* = e^x (2 \cos 3x + \sin 3x), \\ z^* = e^x (2 \sin 3x - \cos 3x); \end{cases}$$

$$6) \quad \begin{cases} y' = -y - 5z \\ z' = y + z \end{cases} \quad y(0) = 3, \quad z(0) = -1, \quad [0; 0,1], \quad h = 0,01; \\ \begin{cases} y^* = 3 \cos 2x + \sin 2x, \\ z^* = -\cos 2x + \sin 2x; \end{cases}$$

- 7)  $\begin{cases} y' = 2y + z \\ z' = 4z - y \end{cases} \quad y(0) = 1, \quad z(0) = 2, \quad [0; 0, 1], \quad h = 0, 01;$
- $$\begin{cases} y^* = (1+x)e^{3x}, \\ z^* = (2+x)e^{3x}; \end{cases}$$
- 8)  $\begin{cases} y' = 3y - z \\ z' = 4y - z \end{cases} \quad y(0) = 2, \quad z(0) = 5, \quad [0; 0, 1], \quad h = 0, 01;$
- $$\begin{cases} y^* = (2-x)e^x, \\ z^* = (5-2x)e^x; \end{cases}$$
- 9)  $\begin{cases} y' = 2z - 3y \\ z' = z - 2y \end{cases} \quad y(0) = 2, \quad z(0) = 3, \quad [0; 0, 1], \quad h = 0, 01;$
- $$\begin{cases} y^* = 2(1+x)e^{-x}, \\ z^* = (3+2x)e^{-x}; \end{cases}$$
- 10)  $\begin{cases} y' = 5y + 3z \\ z' = -3y - z \end{cases} \quad y(0) = 3, \quad z(0) = 1, \quad [0; 0, 1], \quad h = 0, 01;$
- $$\begin{cases} y^* = 3(1+4x)e^{2x}, \\ z^* = (1-12x)e^{2x}; \end{cases}$$
- 11)  $\begin{cases} y' = z + 2e^x \\ z' = y + x^2 \end{cases} \quad y(0) = 4, \quad z(0) = 1, \quad [0; 0, 1], \quad h = 0, 01;$
- $$\begin{cases} y^* = (4+x)e^x + 2e^{-x} - x^2 - 2, \\ z^* = (3+x)e^x - 2e^{-x} - 2x; \end{cases}$$
- 12)  $\begin{cases} y' = z - 5\cos x \\ z' = 2y + z \end{cases} \quad y(0) = 1, \quad z(0) = 10, \quad [0; 0, 1], \quad h = 0, 01;$
- $$\begin{cases} y^* = 3e^{2x} - e^{-x} - 2\sin x - \cos x, \\ z^* = 6e^{2x} + e^{-x} + \sin x + 3\cos x; \end{cases}$$
- 13)  $\begin{cases} y' = 3y + 2z + 4e^{5x} \\ z' = y + 2z \end{cases} \quad y(0) = 6, \quad z(0) = 4, \quad [0; 0, 1], \quad h = 0, 01;$
- $$\begin{cases} y^* = -e^x + 4e^{4x} + 3e^{5x}, \\ z^* = e^x + 2e^{4x} + e^{5x}; \end{cases}$$
- 14)  $\begin{cases} y' = 2y - 4z + 4e^{2x} \\ z' = 2y - 2z \end{cases} \quad y(0) = 3, \quad z(0) = 2, \quad [0; 0, 1], \quad h = 0, 01;$

- $$\begin{aligned}
 & \begin{cases} y^* = 3 \cos 2x + \sin 2x, \\ z^* = \cos 2x + 2 \sin 2x + e^{-2x}; \end{cases} \\
 15) & \begin{cases} y' = 4y + z - e^{2x} \\ z' = z - 2y \end{cases} \quad y(0) = 2, \quad z(0) = 0, \quad [0; 0, 1], \quad h = 0, 01; \\
 & \begin{cases} y^* = x e^{2x} + 2 e^{3x}, \\ z^* = 2(1-x) e^{2x} - 2 e^{3x}; \end{cases} \\
 16) & \begin{cases} y' = 2z - y + 1 \\ z' = 3z - 2y \end{cases} \quad y(0) = 1, \quad z(0) = 3, \quad [0; 0, 1], \quad h = 0, 01; \\
 & \begin{cases} y^* = 2(2+x) e^x - 3, \\ z^* = (5+2x) e^x - 2; \end{cases} \\
 17) & \begin{cases} y' = 5y - 3z + 2 e^{3x} \\ z' = y + z + 5 e^{-x} \end{cases} \quad y(0) = 4, \quad z(0) = 1, \quad [0; 0, 1], \quad h = 0, 01; \\
 & \begin{cases} y^* = 3 e^{2x} + 6 e^{4x} - e^{-x} - 4 e^{3x}, \\ z^* = 3 e^{2x} + 2 e^{4x} - 2 e^{-x} - 2 e^{3x}; \end{cases} \\
 18) & \begin{cases} y' = 2y + z + e^x \\ z' = -2y + 2x \end{cases} \quad y(0) = 3, \quad z(0) = 5, \quad [0; 0, 1], \quad h = 0, 01; \\
 & \begin{cases} y^* = (1 + \cos x + 9 \sin x) e^x + x + 1, \\ z^* = (8 \cos x - 10 \sin x - 2) e^x - 2x - 1; \end{cases} \\
 19) & \begin{cases} y' = 2y - 4z \\ z' = y - 3z + 3 e^x \end{cases} \quad y(0) = 3, \quad z(0) = 1, \quad [0; 0, 1], \quad h = 0, 01; \\
 & \begin{cases} y^* = 4(1-x) e^x - e^{-2x}, \\ z^* = (2-x) e^x - e^{-2x}; \end{cases} \\
 20) & \begin{cases} y' = y + 2z \\ z' = y - 5 \sin x \end{cases} \quad y(0) = 3, \quad z(0) = 7, \quad [0; 0, 1], \quad h = 0, 01; \\
 & \begin{cases} y^* = -2 e^{-x} + 6 e^{2x} - \cos x + 3 \sin x, \\ z^* = 2 e^{-x} + 3 e^{2x} + 2 \cos x - \sin x; \end{cases} \\
 21) & \begin{cases} y' = 2y - z \\ z' = z - 2y + 18x \end{cases} \quad y(0) = 1, \quad z(0) = 6, \quad [0; 0, 1], \quad h = 0, 01;
 \end{aligned}$$



- $$\begin{cases} y^* = -2e^{3x} + 3x^2 + 2x + 3, \\ z^* = 2e^{3x} + 6x^2 - 2x + 4; \end{cases}$$
- 22)  $\begin{cases} y' = y + 2z + 16xe^x & y(0) = 1, \quad z(0) = 6, \quad [0; 0,1], \quad h = 0,01; \\ z' = 2y - 2z \end{cases}$
- $$\begin{cases} y^* = 16e^{2x} - 2e^{-3x} - (12x + 13)e^x \\ z^* = 8e^{2x} + 4e^{-3x} - (8x + 6)e^x \end{cases}$$
- 23)  $\begin{cases} y' = 2y + 4z - 8 & y(0) = 2, \quad z(0) = 3,5, \quad [0; 0,1], \quad h = 0,01; \\ z' = 3y + 6z \end{cases}$
- $$\begin{cases} y^* = 2e^{8x} - 6x, \\ z^* = 3e^{8x} + 3x + 0,5; \end{cases}$$
- 24)  $\begin{cases} y' = 2y - 3z & y(0) = 4, \quad z(0) = 1, \quad [0; 0,1], \quad h = 0,01; \\ z' = y - 2z + 2\sin x \end{cases}$
- $$\begin{cases} y^* = 3e^x + e^{-x} + 3\sin x, \\ z^* = e^x + e^{-x} - \cos x + 2\sin x; \end{cases}$$
- 25)  $\begin{cases} y' = y - z + 2\sin x & y(0) = 2, \quad z(0) = 1, \quad [0; 0,1], \quad h = 0,01; \\ z' = 2y - z \end{cases}$
- $$\begin{cases} y^* = 2(\cos x + \sin x) + x(\sin x - \cos x), \\ z^* = 5\sin x + (1 - 2x)\cos x; \end{cases}$$

11.3. Знайти розв'язки задач Коші на відрізку  $[a, b]$  з точністю  $\varepsilon = 10^{-4}$ . Для перевірки результатів скористатись наведеними точними розв'язками  $y^*$ .

1)  $y'' + 2y' + 2y = 6x^2 + 14x + 6,$

$y(0) = 0, \quad y'(0) = 1,$

$a = 0, \quad b = 3; \quad y^* = e^{-x}(\cos x + \sin x) + 3x^2 + x - 1;$

2)  $y'' + 2y' + y = 4x^2 - 3x - 5,$

$y(0) = 2, \quad y'(0) = 5,$

$a = 0, \quad b = 2; \quad y^* = e^{-x}(-23 + x) + 4x^2 - 19x + 25;$

$$3) y'' - y = 3x^2 - 7x + 9,$$

$$y(0) = 1, y'(0) = 7,$$

$$a = 0, b = 1; \quad y^* = 8(e^{-x} + e^x) - 3x^2 + 7x - 15;$$

$$4) y'' + y' - 2y = 8x^2 - 4x,$$

$$y(0) = 7, y'(0) = -3,$$

$$a = 0, b = 1; \quad y^* = e^{-2x} + e^x - 4x^2 - 2x + 5;$$

$$5) y'' + y' = 9x^2 + 4x - 6,$$

$$y(0) = 2, y'(0) = 7,$$

$$a = 0, b = 2; \quad y^* = e^{-x} + 3x^3 - 7x^2 + 8x + 1;$$

$$6) y'' - y' = 5x^2 - 12x + 8,$$

$$y(0) = 1, y'(0) = -1,$$

$$a = 0, b = 3; \quad y^* = 5e^x - \frac{5}{3}x^3 + x^2 - 6x - 4;$$

$$7) y'' - 3y' = 18x^2 - 2x,$$

$$y(0) = 2, y'(0) = 2\frac{1}{3},$$

$$a = 0, b = 0,5; \quad y^* = e^{3x} - 2x^3 - 2x^2 - \frac{2}{3}x + 1;$$

$$8) y'' + 3y' = 9x^2 + 3x + 5,$$

$$y(0) = 2, y'(0) = -1,$$

$$a = 0, b = 3; \quad y^* = e^{-3x} + x^3 - \frac{1}{2}x^2 + 2x + 1;$$

$$9) y'' + 2y' + y = 2x^3 + 4x^2 - 6x + 1,$$

$$y(0) = -6, y'(0) = 14,$$

$$a = 0, b = 3; \quad y^* = e^{-x}(5 + 5x) + 2x^3 - 8x^2 + 14x - 11;$$

$$10) y'' - 6y' + 5y = 9e^{2x},$$

$$y(0) = 1, y'(0) = 6,$$

$$a = 0, b = 0,3; \quad y^* = 2(e^x + e^{5x}) - 3e^{2x};$$

$$11) y'' - 5y' + 4y = 8e^{3x},$$

$$y(0) = 0, y'(0) = -2,$$

$$a = 0, b = 0,5; \quad y^* = 2(e^x + e^{4x}) - 4e^{3x};$$

$$12) y'' - 6y' + 8y = 14e^{2x},$$

$$y(0) = 2, y'(0) = -1,$$

$$a = 0, b = 0,6; \quad y^* = e^{2x} + e^{4x} - 7xe^{2x};$$

$$13) y'' - 2y' - 3y = 12e^{3x},$$

$$y(0) = 2, y'(0) = 5,$$

$$a = 0, b = 0,5; \quad y^* = e^{-x} + e^{3x} + 3xe^{3x};$$

$$14) y'' - 2y' + y = 10e^x,$$

$$y(0) = 1, y'(0) = 2,$$

$$a = 0, b = 1,5; \quad y^* = (1+x)e^x + 5x^2e^x;$$

$$15) y'' + 2y' + y = 6e^{-x},$$

$$y(0) = 1, y'(0) = 0,$$

$$a = 0, b = 4; \quad y^* = (1+x)e^{-x} + 3x^2e^{-x};$$

$$16) y'' - 4y' + 3y = 20\cos 2x + 27\sin 2x,$$

$$y(0) = 4, y'(0) = -10,$$

$$a = 0, b = 0,5; \quad y^* = 2(e^x - e^{3x}) + 4\cos 2x - 3\sin 2x;$$

$$17) y'' + y' - 2y = 11 \cos \frac{x}{2} - 7 \sin \frac{x}{2},$$

$$y(0) = 0, \quad y'(0) = 0,$$

$$a = 0, \quad b = 2; \quad y^* = 2(e^x + e^{-2x}) + 4 \sin \frac{x}{2} - 4 \cos \frac{x}{2};$$

$$18) y'' + 5y' + 6y = -50 \sin 4x,$$

$$y(0) = 4, \quad y'(0) = -1,$$

$$a = 0, \quad b = 1; \quad y^* = e^{-2x} + e^{-3x} + \sin 4x + 2 \cos 4x;$$

$$19) y'' - 5y' + 2y = -85 \cos 3x,$$

$$y(0) = 1, \quad y'(0) = 2,$$

$$a = 0, \quad b = 0,8; \quad y^* = (-6 \cos x - 10 \sin x)e^{-x} + 7 \cos 3x + 6 \sin 3x;$$

$$20) y'' + 3y' + 2y = 26 \sin \frac{x}{3} - 8 \cos \frac{x}{3},$$

$$y(0) = 0, \quad y'(0) = -9,$$

$$a = 0, \quad b = 2; \quad y^* = 6e^{-x} + 3e^{-2x} - 9 \cos \frac{x}{3} + 9 \sin \frac{x}{3};$$

$$21) y'' - y' - 2y = 22 \cos 5x - 32 \sin 5x,$$

$$y(0) = 2, \quad y'(0) = 5,$$

$$a = 0, \quad b = 0,6; \quad y^* = 2e^{-x} + e^{2x} + \sin 5x - \cos 5x;$$

$$22) y'' + y' = 14 \cos x + 6 \sin x,$$

$$y(0) = 1, \quad y'(0) = -2,$$

$$a = 0, \quad b = 3; \quad y^* = \cos x + \sin x + x(7 \sin x - 3 \cos x);$$

$$23) y'' + 9y = 18 \cos 3x - 12 \sin 3x,$$

$$y(0) = 1, \quad y'(0) = 5,$$

$$a = 0, \quad b = 1; \quad y^* = \cos 3x + \sin 3x + x(2 \cos 3x + 3 \sin 3x);$$

$$24) y'' + 4y = 18 \sin 2x,$$

$$y(0) = 1, \quad y'(0) = 0,$$

$$a = 0, \quad b = 1, 5;$$

$$y' = \cos 2x + \sin 2x - 2x \cos 2x;$$

$$25) y'' + 9y = -18 \sin 3x,$$

$$y(0) = 1, \quad y'(0) = 6,$$

$$a = 0, \quad b = 1;$$

$$y' = \cos 3x + \sin 3x + 3x \cos 3x.$$

## Розділ 12

### МЕТОДИ РОЗВ'ЯЗУВАННЯ ЛІНІЙНИХ КРАЙОВИХ ЗАДАЧ ДЛЯ ЗВИЧАЙНИХ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ

Розглядатимемо доволі поширену крайову задачу для лінійного диференціального рівняння другого порядку з лінійними крайовими (граничними) умовами.

Нехай на проміжку  $[a, b]$  потрібно знайти розв'язок лінійного диференціального рівняння другого порядку

$$L(y) = y'' + p(x)y' - q(x)y = f(x), \quad a < x < b, \quad (12.1)$$

який задовольняє такі крайові умови:

$$\alpha_0 y(a) + \beta_0 y'(a) = \gamma_0, \quad (12.2)$$

$$\alpha_1 y(b) + \beta_1 y'(b) = \gamma_1,$$

де  $p(x), q(x), f(x)$  – відомі функції;  $\alpha_0, \beta_0, \gamma_0, \alpha_1, \beta_1, \gamma_1$  – задані сталі, причому

$$|\alpha_0| + |\beta_0| \neq 0, \quad |\alpha_1| + |\beta_1| \neq 0. \quad (12.3)$$

Питання існування і єдиності розв'язку  $y(x)$  задачі (12.1)–(12.2) розглядають у курсі диференціальних рівнянь. Тут вважаємо, що розв'язок  $y(x)$  існує, єдиний і існують похідні від  $y(x)$  достатньо високого порядку. Для цього необхідно виконання умов (12.3).

Методи розв'язування крайових задач для звичайних диференціальних рівнянь можна класифікувати так.

1. Методи зведення крайової задачі до задачі Коші. Сюди належать методи варіації сталих, стрільби, диференціальної прогонки.

2. Сіткові методи, до яких належать метод скінченних елементів, метод скінченних різниць, де розв'язок знаходять у вузлах сітки.
3. Варіаційні, коли розв'язування задачі (12.1)–(12.2) можна замінити задачею мінімізації відповідного функціонала (метод Рітца, метод найменших квадратів).
4. Проекційні (метод колокації, метод Бубнова–Гальоркіна).

## 12.1. Методи зведення крайової задачі до задачі Коші

### 12.1.1. Метод варіації сталих

Розв'язок рівняння (12.1) можна записати у вигляді

$$y(x) = Z(x) + c_1 Z_1(x) + c_2 Z_2(x), \quad a < x < b, \quad (12.4)$$

де  $c_1$  і  $c_2$  – довільні сталі;  $Z(x)$ ,  $Z_1(x)$ ,  $Z_2(x)$  – розв'язки таких задач Коші:

$$Z''(x) + p(x)Z'(x) - q(x)Z(x) = f(x); \quad (12.5)$$

$$Z(a) = 0, \quad Z'(a) = 0; \quad (12.6)$$

$$Z_1''(x) + p(x)Z_1'(x) - q(x)Z_1(x) = 0; \quad (12.7)$$

$$Z_1(a) = 0, \quad Z_1'(a) = 1; \quad (12.8)$$

$$Z_2''(x) + p(x)Z_2'(x) - q(x)Z_2(x) = 0; \quad (12.9)$$

$$Z_2(a) = 1, \quad Z_2'(a) = 0. \quad (12.10)$$

Систему диференціальних рівнянь другого порядку (12.5), (12.7), (12.9) зводимо до системи шести диференціальних рівнянь першого порядку, зробивши відповідні заміни:

$$T(x) = Z'(x), \quad T_1(x) = Z_1'(x), \quad T_2(x) = Z_2'(x).$$

Одержимо задачі Коші для систем диференціальних рівнянь першого порядку:

$$\begin{cases} T'(x) + p(x)T(x) - q(x)Z(x) = f(x), \\ Z'(x) = T(x); \end{cases} \quad (12.11)$$

$$Z(a)=0, \quad T(a)=0; \quad (12.12)$$

$$\begin{cases} T_1'(x) + p(x)T_1(x) - q(x)Z_1(x) = 0, \\ Z_1'(x) = T_1(x); \end{cases} \quad (12.13)$$

$$Z_1(a)=0; \quad T_1(a)=1; \quad (12.14)$$

$$\begin{cases} T_2'(x) + p(x)T_2(x) - q(x)Z_2(x) = 0, \\ Z_2'(x) = T_2(x); \end{cases} \quad (12.15)$$

$$Z_2(a)=1; \quad T_2(x)=0. \quad (12.16)$$

Розв'яжемо системи (12.11)–(12.16) одним із чисельних методів розв'язування задачі Коші для системи звичайних диференціальних рівнянь першого порядку і використаємо крайові умови (12.2) та (12.3), одержимо систему двох лінійних алгебричних рівнянь для визначення невідомих  $c_1$  і  $c_2$ :

$$\begin{cases} c_1 [\alpha_0 Z_1(a) + \beta_0 Z_1'(a)] + c_2 [\alpha_0 Z_2(a) + \beta_0 Z_2'(a)] = \\ \quad = \gamma_0 - \alpha_0 Z(a) - \beta_0 Z'(a), \\ c_1 [\alpha_1 Z_1(b) + \beta_1 Z_1'(b)] + c_2 [\alpha_1 Z_2(b) + \beta_1 Z_2'(b)] = \\ \quad = \gamma_1 - \alpha_1 Z(b) - \beta_1 Z'(b). \end{cases} \quad (12.17)$$

Тоді розв'язок  $y(x)$  на  $(a, b)$  знайдемо за формулою (12.4).

### 12.1.2. Метод стрільби

Розв'язок задачі (12.1)–(12.3) методом стрільби шукаємо у вигляді

$$y(x) = (1-c)Z_1(x) + cZ_2(x), \quad (12.18)$$

де  $c$  – довільна стала.

Під час розв'язування задачі (12.1)–(12.3) спочатку знаходимо розв'язки  $Z_1(x)$  і  $Z_2(x)$  таких задач Коші:

$$Z_1''(x) + p(x)Z_1'(x) - q(x)Z_1(x) = f(x), \quad a \leq x \leq b, \quad (12.19)$$

$$\alpha_0 Z_1(a) + \beta_0 Z_1'(a) = \gamma_0, \quad (12.20)$$

$$Z_1'(a) = t_0, \quad \text{якщо } \alpha_0 \neq 0, \quad (12.21)$$

або



$$Z_1(a) = t_0, \text{ якщо } \alpha_0 = 0 \quad (12.22)$$

i

$$Z_2''(x) + p(x)Z_2'(x) - q(x)Z_2(x) = f(x), \quad a \leq x \leq b, \quad (12.23)$$

$$\alpha_0 Z_2(a) + \beta_0 Z_2'(a) = \gamma_0, \quad (12.24)$$

$$Z_2'(a) = t_1, \text{ якщо } \alpha_0 \neq 0, \quad (12.25)$$

або

$$Z_2(a) = t_1, \text{ якщо } \alpha_0 = 0, \quad (12.26)$$

де  $t_0$  і  $t_1$  – довільні числа;  $t_0 \neq t_1$ .

Зробимо заміну

$$T_1(x) = Z_1'(x), \quad T_2(x) = Z_2'(x).$$

Задачу Коші для системи чотирьох звичайних диференціальних рівнянь першого порядку запишемо так:

$$\begin{cases} Z_1'(x) = T_1(x), \\ T_1'(x) + p(x)T_1(x) - q(x)Z_1(x) = f(x), \end{cases} \quad (12.27)$$

$$\alpha_0 Z_1(a) + \beta_0 T_1(a) = \gamma_0,$$

$$T_1(a) = t_0, \text{ якщо } \alpha_0 \neq 0, \text{ або } Z_1(a) = t_0, \text{ якщо } \alpha_0 = 0, \quad (12.28)$$

i

$$\begin{cases} Z_2'(x) = T_2(x), \\ T_2'(x) + p(x)T_2(x) - q(x)Z_2(x) = f(x), \end{cases} \quad (12.29)$$

$$\alpha_0 Z_2(a) + \beta_0 T_2(a) = \gamma_0.$$

$$T_2(a) = t_1, \text{ якщо } \alpha_0 \neq 0, \text{ або } Z_2(a) = t_1, \text{ якщо } \alpha_0 = 0 \quad (12.30)$$

Розв'яжемо задачу Коші (12.27) і (12.29), знайдемо сталу  $c$  за формулою

$$c = [\gamma_1 - \alpha_1 Z_1(b) - \beta_1 T_1(b)] / \{ \alpha_1 [Z_2(b) - Z_1(b)] + \beta_1 [T_2(b) - T_1(b)] \}.$$

На проміжку  $[a, b]$  з кроком  $h$  знайдемо розв'язок граничної задачі (12.1)–(12.3) за формулою (12.18).

### 12.1.3. Метод диференціальної прогонки

За допомогою методу диференціальної прогонки розв'язують задачу Коші не для початкового рівняння (12.1), а для відповідних диференціальних рівнянь нижчого порядку. Якщо  $\beta_0 \neq 0$ , то алгоритм методу диференціальної прогонки для задачі (12.1)–(12.3) складається з двох етапів.

1. Розв'язування задачі Коші для двох диференціальних рівнянь першого порядку щодо  $Z_1(x)$  і  $Z_2(x)$  при  $x \in [a, b]$ :

$$\begin{cases} Z_1'(x) = -Z_1^2(x) - p(x)Z_1(x) + q(x), \\ Z_1(a) = -\alpha_0/\beta_0; \end{cases} \quad (12.31)$$

$$\begin{cases} Z_2'(x) = -Z_2(x)(Z_1(x) + p(x)) + f(x), \\ Z_2(a) = \gamma_0/\beta_0; \end{cases} \quad (12.32)$$

2. За відомими значеннями  $Z_1(x)$  і  $Z_2(x)$  знаходимо  $y(x)$  як розв'язок такої задачі Коші:

$$\begin{cases} y'(x) = Z_1(x)y(x) + Z_2(x), a \leq x \leq b, \\ y(b) = (\gamma_1 - \beta_1 Z_2(b)) / (\alpha_1 + \beta_1 Z_1(b)). \end{cases} \quad (12.33)$$

Розв'язування задач (12.31) і (12.32) називають *прямою прогонкою*, а задачі (12.33) – *оберненою прогонкою*.

При  $\alpha_0 \neq 0$  метод диференціальної прогонки полягає в розв'язуванні таких задач Коші.

1. Пряма прогонка:

$$\begin{cases} U_1'(x) = -U_1^2(x)q(x) + U_1(x)p(x) + 1, \quad a \leq x \leq b, \\ U_1(a) = -\beta_0/\alpha_0; \end{cases} \quad (12.34)$$

$$\begin{cases} U_2'(x) = -U_1(x)(U_2(x)q(x) + f(x)), \quad a \leq x \leq b, \\ U_2(a) = \gamma_0/\alpha_0; \end{cases} \quad (12.35)$$

2. Обернена прогонка:

$$\begin{cases} y(x) = U_1(x)y'(x) + U_2(x), & a \leq x \leq b, \\ y(b) = (\gamma_1 U_1(b) + \beta_1 U_2(b)) / (\beta_1 + \alpha_1 U_1(b)). \end{cases} \quad (12.36)$$

У разі виконання умов  $\alpha_0 \beta_0 \leq 0$ ,  $\alpha_1 \beta_1 \geq 0$ ,  $q(x) \geq 0$ ,  $x \in [a, b]$ , метод диференціальної прогонки стійкий до похибок обчислень.

## 12.2. Сіткові методи

### 12.2.1. Метод сіток розв'язування крайових задач для звичайних диференціальних рівнянь

Метод сіток розв'язування задачі (12.1)–(12.3) полягає в такому.

1. На відрізку  $[a, b]$  вибирають скінченну множину точок  $a = x_0 < x_1 < \dots < x_N = b$ . Сукупність цих точок називають сіткою, кожна з точок  $x_i$  – вузлом сітки.
2. Крайову задачу (12.1)–(12.3) замінюють системою алгебричних рівнянь, невідомими в якій є величини, наближені до розв'язку крайової задачі в точках  $x_0, x_1, \dots, x_N$ .
3. Систему алгебричних рівнянь розв'язують яким-небудь чисельним методом, розв'язок якої й буде наближеним розв'язком крайової задачі.

Кількість і розташування вузлів сітки вибирають залежно від необхідної точності розв'язування задачі. Часто сітку вибирають рівномірною, а саме:

$$x_n = nh, \quad n = 0, 1, \dots, N, \quad h = \frac{b-a}{N}, \quad x_0 = a; \quad x_N = b. \quad (12.37)$$

Для побудови системи алгебричних рівнянь використовують заміну похідних через значення функції у вузлах сітки. Заміну можна зробити різними способами. Наприклад, використавши розвинення  $y(x)$  у ряд Тейлора, у випадку сітки (12.37) запишемо

$$y'(x_n) = \frac{y(x_{n+1}) - y(x_n)}{h} - \frac{h}{2} y''(\xi_n), \quad \xi_n \in [x_n, x_{n+1}]; \quad (12.38)$$

$$y'(x_n) = \frac{y(x_n) - y(x_{n-1}))}{h} + \frac{h}{2} y''(\xi_{2n}), \quad \xi_{2n} \in [x_{n-1}, x_n], \quad (12.39)$$

$$y'(x_n) = \frac{y(x_{n+1}) - y(x_{n-1}))}{2h} - \frac{h^2}{6} y'''(\xi_{3n}), \quad \xi_{3n} \in [x_{n-1}, x_{n+1}], \quad (12.40)$$

$$y''(x_n) = \frac{y(x_{n+1}) - 2y(x_n) + y(x_{n-1}))}{h^2} - \frac{h^2}{12} y^{(4)}(\xi_{4n}), \quad (12.41)$$

$$\xi_{4n} \in [x_{n-1}, x_{n+1}].$$

У цьому випадку ми припускали, що  $y(x)$  має неперервні похідні до четвертого порядку включно.

На підставі (12.40), (12.41) з рівняння (12.1) отримаємо систему

$$\frac{1}{h^2} [a_n y(x_{n-1}) + b_n y(x_n) + c_n y(x_{n+1})] = f_n + R_n, \quad (12.42)$$

$$n = 1, 2, \dots, N-1,$$

де

$$a_n = 1 + \frac{h}{2} p(x_n), \quad b_n = -(2 + \frac{h}{2} q(x_n)), \quad c_n = 1 - \frac{h}{2} p(x_n),$$

$$f = f(x_n), \quad R_n = \frac{h^2}{12} y^{(4)}(\xi_{4n}) + \frac{h^2}{6} p(x_n) y'''(\xi_{3n}).$$

Аналогічно за допомогою (12.38) і (12.39) з рівнянь (12.2) і (12.3) одержимо рівняння

$$\left( \alpha_0 - \frac{\beta_0}{h} \right) y(x_0) + \frac{\beta_0}{h} y(x_1) = \gamma_0 + R_0, \quad (12.43)$$

$$-\frac{\beta_1}{h} y(x_{N-1}) + \left( \alpha_1 + \frac{\beta_1}{h} \right) y(x_N) = \gamma_1 + R_N. \quad (12.44)$$

Тут  $R_0 = \frac{h}{2} \beta_0 y''(\xi_{1,0}), \quad R_N = -\frac{h}{2} \beta_1 y''(\xi_{2,N}).$

За достатньо малих  $h$  із (12.42)–(12.44), відкинувши  $R_n, R_0, R_N$ , отримаємо різницеву схему

$$\frac{1}{h^2} [a_n y_{n-1} + b_n y_n + c_n y_{n+1}] = f_n; \quad (12.45)$$

$$\left( \alpha_0 - \frac{\beta_0}{h} \right) y_0 + \frac{\beta_0}{h} y_1 = \gamma_0; \quad (12.46)$$

$$-\frac{\beta_1}{h} y_{N-1} + \left( \alpha_1 + \frac{\beta_1}{h} \right) y_N = \gamma_1. \quad (12.47)$$

Якщо  $\|R(h)\| \leq Mh^k$  де  $M$  – деяка додатна стала,  $R(h) = (R_0, R_1, \dots, R_N)$ , то кажуть, що різницева схема (12.45)–(12.47) апроксимує задачу (12.1)–(12.3) з похибкою  $k$ -го порядку відносно  $h$ . Для розглянутої різницевої схеми (12.45)–(12.47) одержимо  $\|R(h)\| \leq Mh$ . Щоб підвищити порядок апроксимації різницевих рівнянь (12.46)–(12.47), необхідно підвищити порядок апроксимації похідних  $y'(x_0)$ ,  $y'(x_N)$ . Використовуючи триточковий шаблон, це можна зробити за допомогою формул

$$y'(x_0) = \frac{-y(x_2) + 4y(x_1) - 3y(x_0)}{2h} + O(h^2); \quad (12.48)$$

$$y'(x_N) = \frac{3y(x_N) - 4y(x_{N-1}) + y(x_{N-2})}{2h} + O(h^2). \quad (12.49)$$

Недоліком апроксимацій (12.48), (12.49) є порушення тридіагональної структури матриці. Проте існує інший спосіб апроксимації крайових умов. Суть його така. Замінімо  $y'(x_0)$  та  $y'(x_N)$  у крайових умовах скінченними різницями

$$y'(x_0) \approx \frac{y(x_1) - y(x_{-1})}{2h}, \quad y'(x_N) \approx \frac{y(x_{N+1}) - y(x_{N-1})}{2h}$$

та запишемо основне різницеве рівняння (12.42) на границі при  $i=0$  та  $i=N$  з використанням фіктивних вузлів  $x_{-1}$ ,  $x_{N+1}$ . Вилучивши з отриманих рівнянь значення сіткових функцій у фіктивних вузлах, ми отримаємо апроксимацію крайових умов з другим порядком точності, причому без зміни шаблону.

За похибкою апроксимації в загальному випадку не можна зробити висновку про величину  $\varepsilon_n = y(x_n) - y_n$ . Однак у разі виконання умов  $\alpha_0 \beta_0 \leq 0$ ,  $\alpha_1 \beta_1 \leq 0$ ,  $q(x) \geq 0$  і описаних замін у

задачі (12.1)–(12.3) норма  $\|\varepsilon(h)\| = \|(\varepsilon_0, \varepsilon_1, \dots, \varepsilon_N)\|$  матиме той же порядок відносно  $h$ , що й  $\|R(h)\| \leq M_1 h^k$ , де  $M_1 = \text{const}$ .

### 12.2.2. Метод різницевої прогонки

У випадку застосування методу сіток до крайових задач для диференціальних рівнянь другого порядку отримують системи лінійних алгебричних рівнянь, кожне з яких містить лише три сусідні змінні. Для розв'язування таких систем розроблено ефективний метод – *метод прогонки*. Він є реалізацією методу Гаусса для систем рівнянь з тридіагональною матрицею.

Розглянемо задачу

$$\begin{aligned} A_i y_{i-1} - C_i y_i + B_i y_{i+1} &= -F_i, \quad i=1, 2, \dots, N-1, \\ y_0 &= \chi_1 y_1 + \mu_1, \quad y_N = \chi_2 y_{N-1} + \mu_2, \end{aligned} \quad (12.50)$$

причому  $A_i \neq 0$ ,  $B_i \neq 0$  для всіх  $i=1, 2, \dots, N-1$ . Матриця системи (12.50) є тридіагональною. Вона має вигляд

$$\begin{pmatrix} 1 & -\chi_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ A_1 & -C_1 & B_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & A_2 & -C_2 & B_2 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & A_{N-2} & -C_{N-2} & B_{N-2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & A_{N-1} & -C_{N-1} & B_{N-1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\chi_2 & 1 \end{pmatrix}.$$

Нехай виконуються умови (2.44). Тоді мають зміст такі розрахункові формули.

**Метод правої прогонки.** Прогонні коефіцієнти  $\alpha_i, \beta_i$  і розв'язок  $y_i$  шукаємо за формулами

$$\alpha_{i+1} = \frac{B_i}{C_i - \alpha_i A_i}, \quad i=1, 2, \dots, N-1, \quad \alpha_1 = \chi_1, \quad (12.51)$$

$$\beta_{i+1} = \frac{A_i \beta_i + F_i}{C_i - \alpha_i A_i}, \quad i = 1, 2, \dots, N-1, \beta_1 = \mu_1, \quad (12.52)$$

$$y_N = \frac{\mu_2 + \chi_2 \beta_N}{1 - \alpha_N \chi_2}, \quad (12.53)$$

$$y_i = \alpha_{i+1} y_{i+1} + \beta_{i+1}, \quad i = N-1, N-2, \dots, 1, 0. \quad (12.54)$$

**Метод лівої прогонки.** Прогонні коефіцієнти  $\xi_i, \eta_i$  і розв'язок  $y_i$  шукаємо за формулами

$$\xi_i = \frac{A_i}{C_i - \xi_{i+1} A_i}, \quad i = N-1, N-2, \dots, 2, 1, \quad (12.55)$$

$$\xi_N = \chi_2;$$

$$\eta_i = \frac{B_i \eta_{i+1} + F_i}{C_i - \xi_{i+1} A_i}, \quad i = N-1, N-2, \dots, 2, 1, \quad (12.56)$$

$$\eta_N = \mu_2;$$

$$y_0 = \frac{\mu_1 + \chi_1 \eta_1}{1 - \xi_1 \chi_1}; \quad (12.57)$$

$$y_{i+1} = \xi_{i+1} y_i + \eta_{i+1}, \quad i = 0, 1, \dots, N-1. \quad (12.58)$$

**Метод зустрічних прогонки.** Комбінуючи ліву і праву прогонки, отримаємо метод зустрічних прогонки. Нехай  $i = i_0, 0 < i_0 < N$ , – деякий внутрішній вузол. Тоді в області  $0 \leq i \leq i_0$  за формулами (12.51)–(12.54) обчислюють прогонні коефіцієнти  $\alpha_i, \beta_i$ :

$$\alpha_{i+1} = \frac{B_i}{C_i - \alpha_i A_i}, \quad i = 1, 2, \dots, i_0 - 1, \alpha_1 = \chi_1; \quad (12.59)$$

$$\beta_{i+1} = \frac{A_i \beta_i + F_i}{C_i - \alpha_i A_i}, \quad i = 1, 2, \dots, i_0 - 1, \beta_1 = \mu_1, \quad (12.60)$$

а в області  $i_0 \leq i \leq N$  за формулами (12.55)–(12.58) –  $\xi_i, \eta_i$ :

$$\xi_i = \frac{A_i}{C_i - \xi_{i+1} A_i}, \quad i = i_0 - 1, i_0 - 2, \dots, 2, 1, \quad (12.61)$$

$$\xi_N = \chi_2;$$

$$\eta_i = \frac{B_i \eta_{i+1} + F_i}{C_i - \xi_{i+1} A_i}, \quad i = i_0 - 1, i_0 - 2, \dots, 2, 1, \quad (12.62)$$

$$\eta_N = \mu_2.$$

При  $i = i_0$  зшиваємо розв'язки у вигляді (12.54) і (12.58). З формул

$$y_{i_0} = \alpha_{i_0+1} y_{i_0+1} + \beta_{i_0+1}, \quad y_{i_0+1} = \xi_{i_0+1} y_{i_0} + \eta_{i_0+1}$$

знаходимо

$$y_{i_0} = \frac{\alpha_{i_0+1} \eta_{i_0+1} + \beta_{i_0+1}}{1 - \alpha_{i_0+1} \xi_{i_0+1}}$$

Знаючи  $y_{i_0}$ , можна за формулою (12.54) знайти всі  $y_i$  при  $i < i_0$ , а за формулою (12.58) – значення  $y_i$  при  $i > i_0$ .

Метод зустрічних прогонок може виявитись корисним, наприклад, коли потрібно знайти  $y_i$  тільки в одному вузлі  $i = i_0$ .

## 12.3. Варіаційні методи

### 12.3.1. Метод Рітца

Розглянемо варіаційні методи розв'язування крайових задач для звичайних диференціальних рівнянь, що дають змогу отримати наближений розв'язок в аналітичному вигляді.

Метод Рітца для знаходження наближених розв'язків крайових задач ґрунтується на відомій теоремі про те, що коли розв'язок операторного рівняння

$$Lu = f$$

із самоспряженим і додатно визначеним у дійсному гільбертовому просторі  $H$  оператором існує, то цей розв'язок надає функціоналу енергії

$$F(u) = (Lu, u) - 2(f, u)$$



найменшого значення і, навпаки, елемент гільбертового простору, що реалізує мінімум функціонала, задовольняє відповідне рівняння.

Нехай на проміжку  $[a, b]$  треба знайти розв'язок лінійного самоспряженого диференціального рівняння другого порядку із крайовими умовами

$$-\frac{d}{dx}(p(x)y') + q(x)y = f(x); \quad (12.63)$$

$$\begin{aligned} p(x) \geq p > 0, \quad p(x) \in C^1[a, b], \quad q(x), f(x) \in C[a, b]; \\ y(a) = A; \quad y(b) = B, \end{aligned} \quad (12.64)$$

Сформульовану задачу, заміняємо варіаційною про відшукування функції  $y(x) \in C^2[a, b]$ , що задовольняє умови (12.64) і мінімізує функціонал

$$J(y) = \int_a^b [p(x)y'^2 + q(x)y^2 - 2f(x)y] dx. \quad (12.65)$$

Одним із найпоширеніших методів мінімізації функціонала є метод Рітца.

Мінімізувальну послідовність для функціонала (12.65) шукають у вигляді

$$y_n(x) = \phi_0(x) + \sum_{k=1}^n a_k \phi_k(x), \quad (12.66)$$

де базисні функції  $\{\phi_k(x)\}$ ,  $k = 0, 1, 2, \dots, n$ , повинні задовольняти такі умови:

- 1)  $\phi_k(x) \in C^1[a, b]$ ,  $k = 0, 1, 2, \dots, n$
- 2)  $\phi_0(a) = A$ ;  $\phi_0(b) = B$ ;  
 $\phi_k(a) = \phi_k(b) = 0$ ;  $k = 1, 2, \dots, n$
- 3) система функцій  $\{\phi_k(x)\}$ ,  $k = 1, 2, \dots, n$ , повинна бути лінійно незалежна на  $[a, b]$  для будь-якого скінченного  $n$ ;
- 4) система функцій  $\{y_n(x)\}$  є повна в класі  $C^2[a, b]$ .

Підставимо (12.66) у (12.65), з умови мінімуму функціонала отримаємо систему лінійних алгебричних рівнянь для визначення невідомих  $a_1, a_2, \dots, a_n$ :

$$\frac{\partial J(y_n)}{\partial a_k} = \sum_{i=1}^n A_{ki} a_i + A_k = 0, \quad k=1, 2, \dots, n, \quad (12.67)$$

де

$$A_k = \int_a^b (p(x)\phi'_0(x)\phi'_k(x) + q(x)\phi_0(x)\phi_k(x) + f(x)\phi_k(x))dx, \quad (12.68)$$

$$A_{ki} = A_{ik} = \int_a^b (p(x)\phi'_k(x)\phi'_i(x) + q(x)\phi_k(x)\phi_i(x))dx. \quad (12.69)$$

У [29] доведено, що система (12.67) має єдиний розв'язок.

Оскільки крайову задачу з ненульовими умовами можна звести до крайової задачі з нульовими крайовими умовами  $y(a)=0$ ;  $y(b)=0$ , то в цьому випадку за систему координатних функцій вибирають функції вигляду

$$\text{а) } \phi_k(x) = (x-a)^k(x-b) \quad \text{або} \quad \phi_k(x) = (x-a)(x-b)^k \quad (12.70) \\ k=1, 2, \dots, n;$$

$$\text{б) } \phi_k(x) = \sin(k\pi \frac{x-a}{b-a}), \quad k=1, 2, \dots, n. \quad (12.71)$$

Якщо ж умови (12.64) ненульові, то за  $\phi_0(x)$  доцільно вибрати лінійну функцію

$$\phi_0(x) = A + \frac{B-A}{b-a}(x-a).$$

**П р и к л а д 12.1.** Знайти розв'язок диференціального рівняння

$$y'' + y + x = 0, \quad 0 \leq x \leq 1, \quad (12.72)$$

що задовольняє крайові умови

$$y(0) = y(1) = 0. \quad (12.73)$$

**Розв'язування.** Побудуємо розв'язок методом Рунца. Виберемо систему координатних функцій, що задовольняють умови 1) – 4), у вигляді

$$\begin{aligned}\phi_1(x) &= x(1-x); \\ \phi_2(x) &= x^2(1-x); \\ \phi_3(x) &= x^3(1-x).\end{aligned}\tag{12.74}$$

Тоді розв'язок

$$y_3(x) = \sum_{i=1}^3 a_i \phi_i(x) = a_1 x(1-x) + a_2 x^2(1-x) + a_3 x^3(1-x).\tag{12.75}$$

Невідомі  $a_i$  визначимо з системи лінійних алгебричних рівнянь (12.67). Для цього обчислимо коефіцієнти за формулами (12.69). Для нашого прикладу  $p(x) \equiv 1$ ;  $q(x) \equiv 1$ ;  $f(x) \equiv -x$ ;

$$\begin{aligned}\phi'_1(x) &= 1 - 2x; \\ \phi'_2(x) &= 2x - 3x^2; \\ \phi'_3(x) &= 3x^2 - 4x^3\end{aligned}$$

Тоді

$$\begin{aligned}A_{11} &= \int_0^1 ((\phi'_1(x))^2 + (\phi_1(x))^2) dx = \int_0^1 ((1-2x)^2 - (x-x^2)^2) dx = \\ &= \int_0^1 (x^4 - 2x^3 + 5x^2 - 4x + 1) dx = \frac{11}{30} = 0,366666;\end{aligned}$$

$$\begin{aligned}A_{22} &= \int_0^1 ((\phi'_2(x))^2 + (\phi_2(x))^2) dx = \\ &= \int_0^1 ((2x-3x^2)^2 - (x^2-x^3)^2) dx = \\ &= \int_0^1 (x^6 - 2x^5 + 10x^4 - 12x^3 + 4x^2) dx = \frac{1}{7} \approx 0,142857;\end{aligned}$$

$$\begin{aligned}
 A_{33} &= \int_0^1 ((\phi_3'(x))^2 + (\phi_3(x))^2) dx = \\
 &= \int_0^1 ((3x^2 - 4x^3)^2 - (x^3 - x^4)^2) dx = \\
 &= \int_0^1 (x^8 - 2x^7 + 17x^6 - 24x^5 + 9x^4) dx = \frac{113}{1260} \approx 0,0896825;
 \end{aligned}$$

$$\begin{aligned}
 A_{12} = A_{21} &= \int_0^1 (\phi_1'(x)\phi_2'(x) + \phi_1(x)\phi_2(x)) dx = \\
 &= \int_0^1 ((1 - 2x)(2x - 3x^2) + x^3(1 - x)^2) dx = \\
 &= \int_0^1 (x^5 - 2x^4 + 7x^3 - 7x^2 + 2x) dx = \frac{11}{60} \approx 0,183333;
 \end{aligned}$$

$$\begin{aligned}
 A_{13} = A_{31} &= \int_0^1 (\phi_1'(x)\phi_3'(x) + \phi_1(x)\phi_3(x)) dx = \\
 &= \int_0^1 ((1 - 2x)(3x^2 - 4x^3) + x^4(1 - x)^2) dx = \\
 &= \int_0^1 (x^6 - 2x^5 + 9x^4 - 10x^3 + 3x^2) dx = \frac{23}{210} \approx 0,1095238;
 \end{aligned}$$

$$\begin{aligned}
 A_{23} = A_{32} &= \int_0^1 (\phi_2'(x)\phi_3'(x) + \phi_2(x)\phi_3(x)) dx = \\
 &= \int_0^1 ((2x - 3x^2)(3x^2 - 4x^3) + x^5(1 - x)^2) dx = \\
 &= \int_0^1 (x^7 - 2x^6 + 13x^5 - 17x^4 + 6x^3) dx = \frac{178}{1680} \approx 0,105952;
 \end{aligned}$$

$$A_1 = \int_0^1 f(x)\phi_1(x) dx = - \int_0^1 x(x - x^2) dx = -\frac{1}{12} \approx -0,0833333;$$

$$A_2 = \int_0^1 f(x)\phi_2(x) dx = - \int_0^1 x(x^2 - x^3) dx = -\frac{1}{20} = -0,05;$$

$$A_3 = \int_0^1 f(x)\phi_3(x)dx = -\int_0^1 x(x^3 - x^4)dx = -\frac{1}{30} \approx -0,0333333.$$

Система алгебричних рівнянь набуде вигляду

$$\begin{cases} 0,366666a_1 + 0,183333a_2 + 0,109523a_3 = -0,083333; \\ 0,183333a_1 + 0,142857a_2 + 0,105952a_3 = -0,05; \\ 0,109523a_1 + 0,105952a_2 + 0,0896825a_3 = -0,033333. \end{cases}$$

Розв'яжемо її методом Гаусса за схемою єдиного ділення. Отримаємо значення невідомих коефіцієнтів

$$a_1 = -0,1496505; \quad a_2 = -0,1440745; \quad a_3 = -0,0187068.$$

Шуканим наближенням розв'язком буде

$$y_3(x) = -x(1-x)(0,1496505 + 0,1440745x + 0,0187068x^2).$$

### 12.3.2. Метод найменших квадратів

Розглянемо диференціальне рівняння

$$L(y) \equiv y'' + p(x)y' + q(x)y = f(x) \quad (12.76)$$

з нульовими крайовими умовами

$$y(a) = y(b) = 0, \quad (12.77)$$

де функції  $p(x)$ ,  $q(x)$ ,  $f(x)$  – неперервні на проміжку  $[a, b]$ .

Вважатимемо, що розв'язок цієї крайової задачі  $y(x) \in C^2[a, b]$ .

Наближений розв'язок шукаємо у вигляді

$$y_n(x) = \sum_{k=1}^n a_k \phi_k(x), \quad (12.78)$$

де координатні функції  $\{\phi_k(x)\}$  задовольняють такі умови:

- I)  $\phi_k(x) \in C_1[a, b]$ ;
- II)  $\{\phi_k(x)\}$  – лінійно незалежні на  $[a, b]$ ;
- III)  $\phi_k(a) = \phi_k(b) = 0$ ;  $k = 1, 2, \dots$ ;
- IV)  $\{\phi_k(x)\}$  – повна система функцій в  $C^1[a, b]$ .

Виберемо в (12.78) коефіцієнти  $a_k$  так, щоб наближений розв'язок якомога точніше задовольняв рівняння (12.76). Для цього побудуємо відхили

$$\varepsilon_n^2 = \int_a^b \left( \sum_{k=1}^n a_k L(\phi_k(x)) - f(x) \right)^2 dx \quad (12.79)$$

і мінімізуємо їх:

$$\frac{\partial \varepsilon_n^2}{\partial a_i} = 0; \quad i = 1, \dots, n. \quad (12.80)$$

Легко побачити, що (12.80) – це система лінійних алгебричних рівнянь

$$\sum_{k=1}^n A_{ik} a_k = b_i, \quad (12.81)$$

де

$$A_{ik} = A_{ki} = \int_a^b L(\phi_k(x)) L(\phi_i(x)) dx;$$

$$b_i = \int_a^b f(x) L(\phi_i(x)) dx.$$

**П р и к л а д 12.2.** Приклад (12.72), (12.73) розв'язати за методом найменших квадратів.

*Розв'язування.* Обчислимо коефіцієнти системи (12.81):

$$A_{11} = \int_0^1 (L(\phi_1(x)))^2 dx = \int_0^1 (-2 + x - x^2)^2 dx =$$

$$= \int_0^1 (x^4 - 2x^3 + 5x^2 - 4x + 4) dx = \frac{101}{30} \approx 3,366666;$$

$$A_{12} = A_{21} = \int_0^1 L(\phi_1(x)) L(\phi_2(x)) dx = \int_0^1 (-2 + x - x^2)(2 - 6x + x^2 - x^3) dx =$$

$$= \int_0^1 (x^5 - 2x^4 + 9x^3 - 10x^2 + 14x - 4) dx = \frac{101}{60} \approx 1,683919;$$

$$A_{13} = A_{31} = \int_0^1 L(\phi_1(x))L(\phi_{32}(x))dx = \int_0^1 (-2+x-x^2)(6x-12x^2+x^3-x^4)dx = \\ = \int_0^1 (x^6-2x^5+15x^4-20x^3+30x^2-12x)dx = \frac{38}{21} \approx 1,8095238;$$

$$A_{22} = \int_0^1 (L(\phi_2(x)))^2 dx = \int_0^1 (2-6x+x^2-x^3)^2 dx = \\ = \int_0^1 (x^6-2x^5+13x^4-16x^3+40x^2-24x+4)dx = \frac{131}{35} \approx 3,742857;$$

$$A_{33} = \int_0^1 (L(\phi_3(x)))^2 dx = \int_0^1 (6x-12x^2+x^3-x^4)^2 dx = \\ = \int_0^1 (x^8-2x^7+25x^6-36x^5+156x^4-144x^3+36x^2)dx = \frac{5837}{1260} \approx 4,6325539;$$

$$b_1 = \int_0^1 f(x)L(\phi_1(x))dx = \int_0^1 (-x(-2+x-x^2))dx = -\int_0^1 (-2x+x^2-x^3)dx = \\ = \frac{11}{12} \approx 0,9666666;$$

$$b_2 = \int_0^1 f(x)L(\phi_2(x))dx = \int_0^1 (-x(2-6x+x^2-x^3))dx = \\ = \frac{19}{20} = 0,95;$$

$$b_3 = \int_0^1 f(x)L(\phi_3(x))dx = \int_0^1 (-x(6x-12x^2+x^3-x^4))dx = \\ = \frac{29}{30} \approx 0,9666666.$$

Система (12.81) матиме вигляд

$$\begin{cases} 3,366666a_1 + 1,683333a_2 + 1,809524a_3 = 0,9666666; \\ 1,683333a_1 + 3,742857a_2 + 3,805952a_3 = 0,95; \\ 1,809523a_1 + 3,80592a_2 + 4,632539a_3 = 0,9666666. \end{cases}$$

Її розв'язки такі:  $a_1 = 0,20767$ ;  $a_2 = 0,18665$ ;  $a_3 = -0,02580$ .

Отже, за методом найменших квадратів наближений розв'язок сформульованої задачі запишемо так:

$$y_3(x) = x(1-x)(0,20767 + 0,18665x - 0,02580x^2).$$

## 12.4. Проекційні методи

### 12.4.1. Метод колокації

Метод сіток (метод скінченних різниць) для розв'язування крайових задач дає змогу отримати чисельний розв'язок задачі, тобто тільки таблицю значень шуканої функції. Розглянемо *метод колокації*, за допомогою якого можна знайти наближений розв'язок крайової задачі у вигляді аналітичного виразу.

Нехай потрібно визначити функцію  $y = y(x)$ , яка задовольняє лінійне диференціальне рівняння

$$L[y] \equiv y'' + p(x)y' + q(x)y = f(x) \quad (12.82)$$

і лінійні крайові умови

$$\Gamma_a[y] \equiv \alpha_0 y(a) + \alpha_1 y'(a) = A; \quad (12.83)$$

$$\Gamma_b[y] \equiv \beta_0 y(b) + \beta_1 y'(b) = B, \quad (12.84)$$

де  $|\alpha_0| + |\alpha_1| \neq 0$ ,  $|\beta_0| + |\beta_1| \neq 0$ . Виберемо деяку сукупність лінійно незалежних функцій  $u_0(x), u_1(x), \dots, u_n(x)$  (*базисні функції*), з яких  $u_0(x)$  задовольняє неоднорідні крайові умови

$$\Gamma_a[u_0] = A, \quad \Gamma_b[u_n] = B, \quad (12.85)$$

а решта функцій  $u_i(x)$  ( $i = 1, 2, \dots, n$ ) – відповідні однорідні умови

$$\Gamma_a[u_i] = 0, \quad \Gamma_b[u_i] = 0. \quad (12.86)$$

Класичним прикладом набору базисних функцій, що задовольняють умови (12.86), є функції

$$u_i(x) = (x-a)^i (x-b), \quad i = 1, 2, \dots, n.$$



Іншим прикладом може бути набір функцій

$$u_i(x) = \sin \frac{\pi i(x-a)}{b-a}, \quad i=1,2,\dots,n.$$

Якщо крайові умови однорідні ( $A=0, B=0$ ), то можна прийняти  $u_0(x) \equiv 0$  і розглядати лише систему функцій

$$u_i(x) \quad (i=1,2,\dots,n).$$

Розв'язок крайової задачі (12.82)–(12.84) шукатимемо у вигляді лінійної комбінації базисних функцій

$$y = u_0(x) + \sum_{i=1}^n c_i u_i(x) \quad (12.87)$$

Очевидно, функція  $y(x)$  задовольняє крайові умови (12.83). Підставимо вираз (12.87) у рівняння (12.82), отримаємо

$$R(x, c_1, c_2, \dots, c_n) = L[y] - f(x) = L[u_0] - f(x) + \sum_{i=1}^n c_i L[u_i]. \quad (12.88)$$

Якщо за деякого набору коефіцієнтів  $c_i$  ( $i=1,2,\dots,n$ ) виконується рівність  $R(x, c_1, c_2, \dots, c_n) \equiv 0$  для  $a \leq x \leq b$ , то функція  $y$  є точним розв'язком задачі (12.82)–(12.84). Однак так вдало підібрати коефіцієнти  $c_i$ , зазвичай, неможливо. Тому обмежуються тим, що домагаються, щоб функція  $R(x, c_1, c_2, \dots, c_n)$  перетворювалася в нуль у заданій достатньо густій системі точок  $x_1, x_2, \dots, x_n$  із  $[a, b]$  (*точки колокації*), у яких диференціальне рівняння (12.82) задовольнятиметься точно. За точки колокації можна, наприклад, вибирати точки, які ділять відрізок  $[a, b]$  на однакові частини. У підсумку отримаємо систему лінійних рівнянь

$$\begin{cases} R(x_1, c_1, c_2, \dots, c_n) = 0; \\ \dots\dots\dots \\ R(x_n, c_1, c_2, \dots, c_n) = 0. \end{cases} \quad (12.89)$$

Із системи (12.70) у випадку її сумісності можна визначити коефіцієнти  $c_1, c_2, \dots, c_n$ , після чого наближений розв'язок задачі (12.82)–(12.84) знаходять за формулою (12.87).

**П р и к л а д 12.3.** Методом колокації розв'язати крайову задачу

$$y'' + (1+x^2)y + 1 = 0, \quad y(\pm 1) = 0. \quad (12.90)$$

*Розв'язування.* За базисні функції виберемо поліноми  $u_n(x) = x^{2n-2}(1-x^2)$  ( $n=1, 2, \dots$ ), які, очевидно, задовольняють крайові умови  $y(\pm 1) = 0$ . За точки колокації виберемо  $x_0 = 0$ ,  $x_{\pm} = \frac{1}{2}$ . Обмежившись двома базовими функціями, отримаємо

$$y = c_1(1-x^2) + c_2x^2(1-x^2). \quad (12.91)$$

Підстановка в диференціальне рівняння (12.91) дає

$$R(x) = -2c_1 + c_2(2-12x^2) + (1+x^2)[c_1(1-x^2) + c_2x^2(1-x^2)] + 1 = 1 - c_1(1+x^4) + c_2(2-11x^2-x^6). \quad (12.92)$$

У точках колокації маємо  $R(x_0) = 0$ ,  $R(x_{\pm}) = 0$ . Звідси, використовуючи формулу (12.91), для визначення коефіцієнтів  $c_1$ ,  $c_2$  отримаємо лінійну систему рівнянь

$$\begin{cases} 1 - c_1 + 2c_2 = 0; \\ 1 - \frac{17}{16}c_1 - \frac{49}{64}c_2 = 0. \end{cases} \quad (12.93)$$

Розв'яжемо систему (12.93) і знайдемо  $c_1 \approx 0,95676$ ,  $c_2 \approx -0,02162$ . Отже, маємо наближений розв'язок

$$y = 0,95676(1-x^2) - 0,02162x^2(1-x^2). \quad (12.94)$$

У випадку трьох базисних функцій розв'язок шукатимемо у вигляді

$$y = c_1(1-x^2) + c_2x^2(1-x^2) + c_3x^4(1-x^2).$$

Тоді

$$R_1(x) = -2(c_1 + c_2x^2 + c_3x^4) - 8x^2(c_2 + 2c_3x^2) + 2(1-x^2)(c_2 + 6c_3x^2) + (1-x^4)(c_1 + c_2x^2 + c_3x^4) + 1 = 0. \quad (12.95)$$

Виберемо за точки колокації  $x_0 = 0$ ,  $x_{\pm 1} = \pm \frac{1}{3}$ ,  $x_{\pm 2} = \pm \frac{2}{3}$  і з (12.95)

для визначення невідомих коефіцієнтів  $c_1, c_2, c_3$  отримаємо систему рівнянь

$$\begin{cases} -c_1 + 2c_2 + 1 = 0; \\ -6642c_1 + 5094c_2 + 6398c_3 + 6561 = 0; \\ -7857c_1 - 19530c_2 - 2848c_3 + 6561 = 0. \end{cases}$$

Розв'язок цієї системи такий:  $c_1 \approx 0,93150$ ,  $c_2 \approx -0,03425$ ,  $c_3 \approx -0,03119$ . Отже, наближений розв'язок задачі (12.90) в цьому випадку матиме вигляд

$$y = 0,93150(1 - x^2) - 0,03425x^2(1 - x^2) - 0,03119x^4(1 - x^2). \quad (12.96)$$

З урахуванням симетричності розв'язку задачі (12.90) в табл. 12.1 наведено чисельні розв'язки цієї задачі, отримані методом колокації (формули (12.94) та (12.96)) та методом сіток з використанням методу правої прогонки тільки на відрізку  $[-1, 0]$ .

Таблиця 12.1

Чисельний розв'язок задачі (12.90)

x	Метод колокації		Метод сіток	
	(12,94)	(12,96)	h = 0,1	h = 0,5
-1,0	0	0	0	0
-0,9	0,17846	0,16783	0,16849	0,16820
-0,8	0,33945	0,32285	0,32394	0,32341
-0,7	0,48254	0,46269	0,46407	0,46335
-0,6	0,60735	0,58568	0,58728	0,58642
-0,5	0,71352	0,69074	0,69251	0,69155
-0,4	0,80077	0,77719	0,77908	0,77804
-0,3	0,86888	0,84463	0,84662	0,84552
-0,2	0,91766	0,89288	0,89492	0,89380
-0,1	0,94698	0,92184	0,92392	0,92278
0,0	0,95676	0,93150	0,93359	0,93244

### 12.4.2. Метод моментів. Метод Гальоркіна

Якщо розглядати диференціальне рівняння в загальному вигляді

$$F(x, y, y', y'') = 0, \quad (12.97)$$

загалом кажучи, нелінійне, то для розв'язування крайової задачі доцільно застосувати метод моментів чи його модифікацію – метод Гальоркіна.

У методі моментів для зображення розв'язку вибирають дві системи функцій:

- I.
  - 1)  $\psi_k(x) \in C[a, b]$ ,  $k = 1, 2, \dots, n$ ;
  - 2)  $\{\psi_k(x)\}$  – замкнена система функцій на  $[a, b]$ ;
- II.
  - 1)  $\phi_k(x) \in C_2[a, b]$ ,  $k = 0, 1, 2, \dots, n$ ;
  - 2)  $\{\phi_k(x)\}$ ,  $k = 0, 1, 2, \dots, n$ , – лінійно незалежна;
  - 3)  $\phi_0(a) = A$ ;  $\phi_0(b) = B$ ;
  - 4)  $\phi_k(a) = \phi_k(b) = 0$ ;  $k = 0, 1, 2, \dots, n$ .
  - 5)  $\{\phi_k(x)\}$  – повна система функцій у класі  $C^2[a, b]$ .

Розв'язок наведемо у вигляді

$$y_n(x) = \phi_0(x) + \sum_{k=1}^n a_k \phi_k(x). \quad (12.98)$$

Параметри  $a_k$  виберемо так, щоб перші  $n$  моментів дорівнювали нулю, тобто

$$\mu_k(a_1, a_2, \dots, a_n) = \int_a^b F(x, y_n, y_n', y_n'') \psi_k(x) dx = 0; \quad (k = 1, 2, \dots, n). \quad (12.99)$$

Остання рівність – це система рівнянь для визначення невідомих коефіцієнтів  $a_1, a_2, \dots, a_n$ .

Якщо ж системи функцій  $\{\phi_k(x)\}$  та  $\{\psi_k(x)\}$  збігаються, то метод моментів перетворюється в метод Гальоркіна.

У разі застосування методу моментів до лінійного рівняння

$$F(x, y, y', y'') \equiv L(y) - f(x) \equiv y'' + p(x)y' + q(x)y - f(x) = 0 \quad (12.100)$$

отримаємо систему лінійних алгебричних рівнянь, яка має вигляд

$$\sum_{k=1}^n C_{ki} a_k = d_i; \quad i = 1, 2, \dots, n, \quad (12.101)$$

де

$$C_{ki} = \int_a^b L(\phi_k(x)) \psi_i(x) dx; \quad (12.102)$$

$$d_i = \int_a^b [f(x) - L(\phi_0(x))] \psi_i(x) dx; \quad (12.103)$$

У методі Гальоркіна в (12.102)–(12.103) замість  $\psi_i(x)$  підставимо  $\phi_i(x)$ ,  $i = 1, \dots, n$

Метод Гальоркіна не є варіаційним. Він не потребує попереднього зведення крайової задачі для диференціального рівняння до варіаційної, тому є універсальнішим. За невеликої кількості членів цей метод дає грубе наближення. Зі збільшенням кількості координатних функцій  $n$  не тільки додаються нові коефіцієнти, а й змінюються старі, що не дуже зручно. Зазначимо: якщо задача (12.97) лінійна, а система функцій  $\{\phi_k(x)\}$  ортогональна, то вже знайдені  $C_{ki}$  зі збільшенням  $n$  не будуть змінюватись. Тому ортогональні системи зручніші, ніж неортогональні. Метод Гальоркіна для нелінійної задачі (12.97) використовують для грубого наближення розв'язку, для лінійної задачі (12.98) цим методом можна знайти розв'язок із заданою точністю. Результат дуже суттєво залежить від вдалого вибору функцій  $\{\phi_k(x)\}$

Для лінійних диференціальних рівнянь другого порядку метод Гальоркіна приводить до таких же систем лінійних алгебричних рівнянь, що й метод Рітца.

**П р и к л а д 12.3.** Знайти розв'язок задачі (12.72), (12.73) методом Гальоркіна.

*Розв'язування.* Координатні функції  $\{\phi_k(x)\}$  виберемо такі ж, як і в методі Рітца. Тоді

$$\phi_1''(x) = -2; \quad \phi_2''(x) = 2 - 6x; \quad \phi_3''(x) = 6x - 12x^2$$

Обчислимо коефіцієнти системи за формулами (10.102) та (10.103):

$$\begin{aligned} C_{11} &= \int_0^1 L(\phi_1(x)) \phi_1(x) dx = \int_0^1 (-2 + x - x^2)(x - x^2) dx = \\ &= \int_0^1 (x^4 - 2x^3 + 3x^2 - 2x) dx = -\frac{3}{10} \approx -0,3; \end{aligned}$$

$$C_{12} = \int_0^1 L(\phi_1(x))\phi_2(x)dx = \int_0^1 (-2+x-x^2)(x^2-x^3)dx = \\ = \int_0^1 (x^5-2x^4+3x^3-2x^2)dx = -\frac{3}{20} \approx -0,15;$$

$$C_{13} = \int_0^1 L(\phi_1(x))\phi_3(x)dx = \int_0^1 (-2+x-x^2)(x^3-x^4)dx = \\ = \int_0^1 (x^6-2x^5+3x^4-2x^3)dx = -\frac{19}{210} \approx -0,0904762;$$

$$C_{21} = \int_0^1 L(\phi_2(x))\phi_1(x)dx = \int_0^1 (2-6x+x^2-x^3)(x-x^2)dx = \\ = \int_0^1 (x^5-2x^4+7x^3-8x^2+2x)dx = -\frac{3}{20} \approx -0,15;$$

$$C_{22} = \int_0^1 L(\phi_2(x))\phi_2(x)dx = \int_0^1 (2-6x+x^2-x^3)(x^2-x^3)dx = \\ = \int_0^1 (x^6-2x^5+7x^4-8x^3+2x^2)dx = -\frac{13}{105} \approx -0,1238095;$$

$$C_{23} = \int_0^1 L(\phi_2(x))\phi_3(x)dx = \int_0^1 (2-6x+x^2-x^3)(x^3-x^4)dx = \\ = \int_0^1 (x^7-2x^6+7x^5-8x^4+2x^3)dx = -\frac{158}{1680} \approx -0,0940476;$$

$$C_{31} = \int_0^1 L(\phi_3(x))\phi_1(x)dx = \int_0^1 (6x-12x^2+x^3-x^4)(x-x^2)dx = \\ = \int_0^1 (x^6-2x^5+13x^4-18x^3+6x^2)dx = -\frac{19}{210} \approx -0,09047619;$$

$$C_{32} = \int_0^1 L(\phi_3(x)) \phi_2(x) dx = \int_0^1 (6x - 12x^2 + x^3 - x^4)(x^2 - x^3) dx = \\ = \int_0^1 (x^7 - 2x^6 + 13x^5 - 18x^4 + 6x^3) dx = -\frac{79}{840} \approx -0,09404762;$$

$$C_{33} = \int_0^1 L(\phi_3(x)) \phi_3(x) dx = \int_0^1 (6x - 12x^2 + x^3 - x^4)(x^3 - x^4) dx = \\ = \int_0^1 (x^8 - 2x^7 + 13x^6 - 18x^5 + 6x^4) dx = -\frac{103}{1250} \approx -0,08174603;$$

$$d_1 = \int_0^1 f(x) \phi_1(x) dx = -\int_0^1 x(x - x^2) dx = -\frac{1}{12} \approx -0,0833333;$$

$$d_2 = \int_0^1 f(x) \phi_2(x) dx = -\int_0^1 x(x^2 - x^3) dx = -\frac{1}{20} = -0,05;$$

$$d_3 = \int_0^1 f(x) \phi_3(x) dx = -\int_0^1 x(x^3 - x^4) dx = -\frac{1}{30} \approx -0,0333333.$$

Отже, система матиме вигляд

$$\begin{cases} 0,3a_1 + 0,15a_2 + 0,904762a_3 = 0,0833333, \\ 0,15a_1 - 0,123809a_2 + 0,094047a_3 = 0,05, \\ 0,0904762a_1 + 0,0940476a_2 + 0,0817460a_3 = 0,0333333. \end{cases}$$

Розв'яжемо її методом Гаусса, отримаємо

$$a_1 = 0,18775; \quad a_2 = 0,19422; \quad a_3 = -0,02348.$$

Розв'язок задачі матиме вигляд:

$$y_3(x) = 0,18775x(1-x) + 0,19422x^2(1-x) - 0,02348x^3(1-x) = \\ = x(1-x)(0,18775 + 0,19422x - 0,02348x^2).$$

Розв'язки розглянутого прикладу протабільовано на проміжку  $[0,1]$  з кроком 0,1 і наведено в табл. 12.2.

## Розв'язки задачі (12.72)–(12.73)

$x$	Метод			Точне значення
	Рітца	Гальоркіна	найменших квадратів	
0,10	0,018619	0,018606	0,020346	0,018642
0,20	0,036102	0,036073	0,039032	0,036098
0,30	0,051221	0,051178	0,054876	0,051195
0,40	0,062809	0,062755	0,066759	0,062783
0,50	0,069754	0,069698	0,073624	0,069747
0,60	0,071004	0,070951	0,074475	0,071018
0,70	0,065562	0,065520	0,068379	0,065585
0,80	0,052492	0,052464	0,054464	0,052502
0,90	0,030912	0,030900	0,031920	0,030902



## Завдання до розділу 12

Знайти розв'язки крайових задач на відрізку  $[a, b]$  з точністю  $\varepsilon = 10^{-4}$  і порівняти з точним значенням  $y^*$

- 1)  $y'' + 2y' - 3y = -6x^2 + 5x + 6, \quad 0 \leq x \leq 1,$   
 $y(0) + y'(0) = 1, \quad y'(1) = 5; \quad y^* = 2x^2 + 1;$
- 2)  $y'' + 2y' / (x - 4) + (x - 4)y = 1, \quad 0 \leq x \leq 1,$   
 $y(0) = -0,25, \quad y(1) - 3y'(1) = 0, \quad y^* = 1 / (x - 4);$
- 3)  $y'' + xy' - 4y = x + 1/x - 3x \ln(x), \quad 1 \leq x \leq 2,$   
 $y(1) = 0, \quad y'(2) = \ln(2) + 2, \quad y^* = x \ln(x);$
- 4)  $y'' + y' / x - 2y = -2x + 1/x - 2 \ln(x), \quad 1 \leq x \leq 2,$   
 $y(1) = 1, \quad y'(2) = 1,5, \quad y^* = x + \ln(x);$
- 5)  $y'' + xy' - y = -\sin(2 + x^2) - x \cos(x), \quad 0 \leq x \leq 1,$   
 $y(0) + y'(0) = 1, \quad y(1) = \cos 1, \quad y^* = x \cos(x);$
- 6)  $y'' + y' / x - 4y = -4x + 1/x - 4 \ln(x), \quad 1 \leq x \leq 2,$   
 $y(1) = 1, \quad y'(2) = 1,5, \quad y^* = x + \ln(x);$
- 7)  $y'' + y' - 2y = -2x + 1, \quad 0 \leq x \leq 1,$   
 $y'(0) = 1, \quad y(1) - y'(1) = 0, \quad y^* = x + e^x;$
- 8)  $y'' + xy' - y = e^x + x^2 + 1, \quad 0 \leq x \leq 1,$   
 $y(0) = 1, \quad y'(1) = e, \quad y^* = xe^x;$
- 9)  $y'' + 2y' - y / x = -2x - 1, \quad 0 \leq x \leq 1,$   
 $y(0) = 1, \quad y(1) - y'(1) = 1, \quad y^* = xe^{-x};$
- 10)  $y'' + 2xy' - y = 2 \cos(x)(x^2 + 1), \quad 0 \leq x \leq 0,5,$   
 $y'(0) = 0, \quad y(0,5) = 0,5 \sin(0,5), \quad y^* = x \sin(x);$

- 11)  $y'' + 2y' - 4y/x = 1, \quad 0,5 \leq x \leq 1,$   
 $y'(0,5) = 1,5, \quad y(1) + y'(1) = 4, \quad y^* = x^2 + 0,5x;$
- 12)  $y'' + (x+3)y' - 2y/(x+3)^2 =$   
 $= -6x/(x+3)^2 + 3(x+3)^2 + 1, \quad 0 \leq x \leq 1,$   
 $y(0) = -1/3, \quad y(1) + 4y'(1) = 15, \quad y^* = 3x - 1/(x+3);$
- 13)  $y'' - y' - 2y = -3e^{-x}, \quad 0 \leq x \leq 1,$   
 $y(0) = 1, \quad y(1) + 2y'(1) = 0, \quad y^* = (x+1)e^{-x};$
- 14)  $y'' - y' - 2y = -e^{-x}, \quad 0 \leq x \leq 1,$   
 $y'(0) = 3, \quad 2y(1) - y'(1) = e, \quad y^* = e^x + e^{2x};$
- 15)  $y'' + y' - 2y = 2(-x^2 + x + 1), \quad 0 \leq x \leq 1,$   
 $y(0) = 1, \quad y(1) - y'(1) = -1, \quad y^* = e^x + x^2;$
- 16)  $y'' + 3y' - 4y = -2\cos(x) - 8\sin(x), \quad 0 \leq x \leq 0,5,$   
 $y'(0) = 1, \quad y(0,5) + y'(0,5) = 2\cos(0,5), \quad y^* = \cos(x) + \sin(x);$
- 17)  $y'' + y'/x - y/x^2 = 8x + 1, \quad -1 \leq x \leq 0,$   
 $y'(-1) + y'(-1) = 1, \quad y(0) = 0, \quad y^* = x^3 + x^2;$
- 18)  $y'' + xy' - 2y = -x - 3/x + 2/x^3, \quad 0,5 \leq x \leq 1,$   
 $y(0,5) = 2,5, \quad 2y(1) + y'(1) = 4, \quad y^* = x + 1/x;$
- 19)  $y'' + xy' - x^2y = 1/x + x\ln(x) + x - x^3\ln(x), \quad 1 \leq x \leq 2,$   
 $y'(1) = 1, \quad y(2) = 2\ln(2), \quad y^* = x\ln(x);$
- 20)  $y'' + y' - xy = 2/x^3 - 1/x^2 - 1, \quad 1 \leq x \leq 2,$   
 $y'(1) = 1, \quad y(2) = 2\ln(2), \quad y^* = 1/x;$
- 21)  $y'' - y' - xy = -x(x + e^x) - 1, \quad 1 \leq x \leq 2,$   
 $y(1) - y'(1) = 0, \quad y'(2) = 1 + e^2, \quad y^* = x + e^x;$

- 22)  $y'' + y' - y/x = 3x^3 + 12x^2 + x + 2, \quad 0 \leq x \leq 0,5,$   
 $y'(0) = 0, \quad y(0,5) = 0,875, \quad y^* = x^4 + x^2;$
- 23)  $y'' - y'/x - y/x^2 = (-2 + 2\ln(x))/x^3, \quad 0,5 \leq x \leq 1,$   
 $y(0,5) = 2\ln 0,5, \quad y'(1) = 1, \quad y^* = \ln x/x;$
- 24)  $y'' - 4xy'/(x^2 + 1) - y/(x^2 + 1) = -3/(x^2 + 1)^2, \quad 0 \leq x \leq 1,$   
 $y(0) + y'(0) = 1, \quad y'(1) = -0,5, \quad y^* = 1/(x^2 + 1);$
- 25)  $y'' - xy' - 3y = 2 + 6x - x^2, \quad 0,5 \leq x \leq 1,$   
 $y(0,5) = 0,375, \quad y(1) - y'(1) = -3, \quad y^* = x^2(x + 1).$

## Розділ 13

### ЧИСЕЛЬНЕ РОЗВ'ЯЗУВАННЯ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ У ЧАСТИННИХ ПОХІДНИХ

Нехай в області  $G$  потрібно знайти розв'язок  $u = u(x)$  лінійного диференціального рівняння

$$Lu(x) = f(x), \quad x = (x_1, x_2, \dots, x_n) \in G, \quad (13.1)$$

який задовольняє додаткові (граничні, початкові) умови

$$lu(x) = \mu(x), \quad x \in \Gamma, \quad (13.2)$$

де  $\Gamma$  – межа області  $G$ ,  $K(x, \xi)$ ,  $f(x)$ ,  $\mu(x)$  – задані функції. Замінімо область  $G + \Gamma$  дискретною множиною вузлів  $\bar{\omega}_h = \omega_h + \gamma_h$  (сітка), де  $\omega_h$  – множина внутрішніх вузлів,  $\gamma_h$  – множина граничних вузлів сітки. Похідні, які входять в рівняння (13.1)–(13.2), замінімо деякими різницевиими співвідношеннями на сітці  $\bar{\omega}_h$ . У підсумку отримаємо систему різницевих рівнянь

$$L_h y_h(x) = \varphi_h(x), \quad x \in \omega_h; \quad (13.3)$$

$$l_h y_h(x) = \chi_h(x), \quad x \in \gamma_h, \quad (13.4)$$

де  $y_h(x)$ ,  $\varphi_h(x)$ ,  $\chi_h(x)$  – сіткові функції,  $L_h$  та  $l_h$  – різницеві оператори.

Розв'язок задачі (13.3)–(13.4) залежить від параметра  $h$ . Змінюючи  $h$ , одержуємо послідовність розв'язків  $\{y_h\}$ . Сім'ю рівнянь (13.3)–(13.4), що залежить від параметра  $h$ , називають *різничевою схемою*. Нехай  $u_h$  – проєкція розв'язку  $u = u(x)$  задачі

(13.1)–(13.2) на простір функцій, які задані на сітці  $\bar{\omega}_h$ . Позначимо через  $z_h = y_h - u_h$  похибку схеми (13.3)–(13.4). Підставимо  $y_h = u_h + z_h$  у (13.3)–(13.4) і припустимо, що  $L_h$  і  $I_h$  – лінійні оператори, одержимо для похибки  $z_h$  таку задачу:

$$L_h z_h = \psi_h, \quad x \in \omega_h; \quad I_h z_h = v_h, \quad x \in \gamma_h, \quad (13.5)$$

де  $\psi_h$  і  $v_h$  – похибки апроксимації на розв'язку  $u(x)$  різницевою схемою диференціальної задачі (13.1) і (13.2), відповідно. Для оцінки  $z_h, \psi_h, v_h$  вводять норми  $\|\cdot\|_{(1_h)}, \|\cdot\|_{(2_h)}, \|\cdot\|_{(3_h)}$  в просторах сіткових функцій, які задані на  $\bar{\omega}_h, \omega_h$  і  $\gamma_h$ , відповідно.

Різницева схема (13.3)–(13.4)

1а) апроксимує задачу (13.1)–(13.2), якщо

$$\|\psi_h\|_{(2_h)} \rightarrow 0, \quad \|v_h\|_{(3_h)} \rightarrow 0 \quad \text{при } |h| \rightarrow 0;$$

1б) має  $m$ -й порядок апроксимації, якщо

$$\|\psi_h\|_{(2_h)} = O(|h|^m), \quad \|v_h\|_{(3_h)} = O(|h|^m);$$

2а) збігається, якщо  $\|y_h - u_h\|_{(1_h)} \rightarrow 0$  при  $h \rightarrow 0$ ;

2б) має  $m$ -ий порядок точності або збігається зі швидкістю  $O(|h|^m)$ , якщо

$$\|y_h - u_h\|_{(1_h)} = O(|h|^m).$$

Різницева схема (13.3)–(13.4) коректна, якщо вона за всіх достатньо малих  $|h| \leq h_0$ :

1) має єдиний розв'язок для довільних вхідних даних  $\phi_h, \chi_h$ ;

2) розв'язок  $y_h$  рівномірно по  $h$  неперервно залежить від вхідних даних  $\phi_h, \chi_h$ .

Властивість неперервної залежності  $y_h$  від  $\phi_h, \chi_h$  (рівномірної по  $h$ ) називають *стійкістю* схеми (13.3)–(13.4). Якщо  $L_h$  і  $I_h$  – лінійні оператори, то стійкість схеми (13.3)–(13.4) означає існування таких сталих  $M_1 > 0, M_2 > 0$ , які не залежать від  $h$  і від

вибору вхідних даних  $\varphi_h$  і  $\chi_h$ , що для розв'язку задачі (13.3)–(13.4) при  $|h| \leq h_0$  виконується нерівність

$$\|y_h\|_{(1_h)} \leq M_1 \|\varphi_h\|_{(2_h)} + M_2 \|\chi_h\|_{(3_h)}. \quad (13.6)$$

Якщо схема (13.3)–(13.4) стійка, а  $z_h$  – розв'язок задачі (13.5), то

$$\|z_h\|_{(1_h)} = \|y_h - u_h\|_{(1_h)} \leq M_1 \|\psi_h\|_{(2_h)} + M_2 \|v_h\|_{(3_h)}. \quad (13.7)$$

Із нерівності (13.7) випливає твердження: якщо схема (13.3)–(13.4) стійка й апроксимує задачу (13.1)–(13.2), то вона збігається або зі *стійкості й апроксимації схеми випливає її збіжність*. Порядок точності різницевої схеми (13.3)–(13.4) визначений порядком апроксимації.

Для дослідження розв'язності різницевої схеми (13.3)–(13.4), перепишемо її у вигляді рівняння

$$A(P)y(P) = \sum_{Q \in \Pi'(P)} B(P, Q)y(Q) + f(P), \quad P \in \bar{\omega}_h, \quad (13.8)$$

де  $A(P) > 0$ ,  $B(P, Q)$ ,  $f(P)$  – сіткові функції, які визначені для всіх  $P, Q \in \bar{\omega}_h$ . Підмножина  $\Pi'(P)$  не містить самого вузла  $P$ .

Уведемо позначення

$$L[y] = A(P)y(P) - \sum_{Q \in \Pi'(P)} B(P, Q)y(Q). \quad (13.9)$$

Тоді рівняння (13.8) перепишемо так:

$$L[y] = f(P).$$

Припускаємо, що виконуються такі умови:

- 1)  $A(P) > 0$  для всіх  $P \in \bar{\omega}_h$ ,
- 2)  $B(P, Q) > 0$  для всіх  $P, Q \in \bar{\omega}_h$ , (13.10)
- 3)  $D(P) = A(P) - \sum_{Q \in \Pi'(P)} B(P, Q) \geq 0$  для всіх  $P \in \bar{\omega}_h$ .

**Теорема 13.1** (принцип максимуму).

Нехай  $y(P) \neq \text{const}$  – сіткова функція, яка визначена на  $\bar{\omega}_h$  і нехай виконуються умови (13.10). Якщо  $L[y] \leq 0$  ( $L[y] \geq 0$ ) для

всіх  $P \in \bar{\omega}_h$ , тоді  $y(p) \leq 0$  ( $y(P) \geq 0$ ). Тобто  $y(P)$  не може набувати найбільшого додатного (найменшого від'ємного) значення у внутрішніх вузлах сітки  $P \in \omega_h$ .

Принцип максимуму дає достатню умову стійкості за крайовими умовами.

### 13.1. Метод сіток розв'язування задачі Діріхле для рівняння Пуассона

Нехай  $\bar{G}_0 = \{0 \leq x \leq a, 0 \leq y \leq b\}$  — прямокутник зі сторонами  $a$  і  $b$ ;  $\Gamma$  — його межа. Розглянемо в  $\bar{G} = G_0 + \Gamma$  задачу Діріхле для рівняння Пуассона

$$\Delta u = -f(x, y), \quad (x, y) \in G_0, \quad (13.11)$$

$$u|_{\Gamma} = \mu(x, y). \quad (13.12)$$

Побудуємо в  $\bar{G}_0$  сітку  $\bar{\omega}_h$  з кроками  $h = \frac{a}{M}$ ,  $\tau = \frac{b}{N}$ , де  $M > 0$ ,  $N > 0$  — цілі числа. Для цього побудуємо дві сім'ї прямих

$$x_m = mh, \quad m = \overline{0, M}, \quad y_n = n\tau, \quad n = \overline{0, N}.$$

Точки перетину цих прямих  $(mh, n\tau)$  називають *вузлами сітки*. Якщо  $(mh, n\tau)$  лежить усередині прямокутника (тобто  $1 \leq m \leq M-1$ ,  $1 \leq n \leq N-1$ ), то такий вузол називають *внутрішнім*. Нехай  $\omega_h$  — множина внутрішніх вузлів. Загальна кількість внутрішніх вузлів дорівнює  $(M-1) \times (N-1)$ .

Вузли, які лежать на межі прямокутника (при  $m = 0, M$ ,  $n = 0, N$ ), називають *граничними*. Вони утворюють множину  $\chi_h = \{(mh, n\tau)\}$ . Сукупність усіх внутрішніх та граничних вузлів називати *сіткою*  $\bar{\omega}_h = \omega_h + \chi_h$  у прямокутнику  $\bar{G}_0$ .

Замінімо другі похідні різницевиими співвідношеннями в кожному внутрішньому вузлі, отримаємо різницеве рівняння

$$\frac{u_{m+1,n} - 2u_{m,n} + u_{m-1,n}}{h^2} + \frac{u_{m,n+1} - 2u_{m,n} + u_{m,n-1}}{\tau^2} = -f(x_m, y_n), \quad (13.13)$$

$$m = \overline{1, M-1}, \quad n = \overline{1, N-1}.$$

Це співвідношення апроксимує рівняння (13.11) з похибкою

$$\frac{M_4}{12}(h^2 + \tau^2), \text{ де } M_4 = \max_{G_0} \left\{ \left| \frac{\partial^4 u}{\partial x^4} \right|, \left| \frac{\partial^4 u}{\partial y^4} \right| \right\}.$$

Оскільки область  $\overline{G_0}$  – прямокутник, то граничні умови можна записати у вигляді

$$u(x, 0) = \alpha_0(x), \quad u(x, b) = \alpha_1(x), \quad 0 \leq x \leq a; \quad (13.14)$$

$$u(0, y) = \beta_0(y), \quad u(a, y) = \beta_1(y), \quad 0 \leq y \leq b. \quad (13.15)$$

У підсумку отримаємо СЛАР

$$u_{m+1,n} + \sigma u_{m,n+1} - 2(1 + \sigma)u_{m,n} + \sigma u_{m,n-1} + u_{m-1,n} = -h^2 f(x_m, y_n); \quad (13.16)$$

$$\sigma = h^2 / \tau^2 \quad m = \overline{1, M-1}, \quad n = \overline{1, N-1};$$

$$u_{m,0} = \alpha_0(x_m), \quad u_{m,N} = \alpha_1(x_m), \quad m = \overline{0, M}; \quad (13.17)$$

$$u_{0,n} = \beta_0(y_n), \quad u_{M,n} = \beta_1(y_n), \quad n = \overline{0, N} \quad (13.18)$$

Різницеві рівняння (13.16)–(13.18) називають *різницевою схемою*. Цю СЛАР, зазвичай, розв'язують прямим методом Гаусса або ітераційними методами простої ітерації та Зейделя.

**П р и к л а д 13. 1.** Знайти розв'язок рівняння Пуассона

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{2}{1+y} \left[ 1 + \left( \frac{x+5}{1+y} \right)^2 \right], \quad (13.19)$$

який задовольняє граничні умови

$$u(x, 0) = (x+5)^2, \quad u(x, 1) = \frac{(x+5)^2}{2}, \quad 0 \leq x \leq 1,$$

$$u(0, y) = \frac{25}{1+y}, \quad u(1, y) = \frac{36}{1+y}, \quad 0 \leq y \leq 1. \quad (13.20)$$



де  $\bar{G} = G + \Gamma = \{0 \leq x \leq 1, 0 \leq y \leq 1\}$

Точний розв'язок цієї задачі

$$u_T = (x+5)^2 / (y+1). \quad (13.21)$$

**Розв'язування.** У квадраті  $\{0 \leq x \leq 1, 0 \leq y \leq 1\}$  проведемо сукупність прямих  $x_m = mh, y_n = n\tau$  ( $m = 0, 1, \dots, M; n = 0, 1, \dots, N$ ), де  $h$  – крок сітки за напрямом осі  $OX$ ;  $\tau$  – крок за напрямом осі  $OY$  ( $h > 0, \tau > 0$ ).

Замінімо похідні  $\frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial y^2}$  відповідними різницевиими співвідношеннями, отримаємо різницеве рівняння

$$\frac{u_{m+1,n} - 2u_{m,n} + u_{m-1,n}}{h^2} + \frac{u_{m,n+1} - 2u_{m,n} + u_{m,n-1}}{\tau^2} = \frac{2}{1+y_n} \left[ 1 + \frac{(x_m+5)^2}{1+y_n} \right], \quad (13.22)$$

$$x_m = mh, \quad y_n = n\tau.$$

Граничні умови

$$u_{m,0} = (x_m + 5)^2, \quad u_{m,N} = \frac{(x_m + 5)^2}{2}, \quad u_{0,n} = \frac{25}{1+y_n}, \quad u_{M,n} = \frac{36}{1+y_n}. \quad (13.23)$$

Рівняння (13.22), (13.23) є СЛАР відносно значення розв'язку  $u_{m,n}$  у внутрішніх вузлах сітки.

**А.** Метод Гаусса розв'язування СЛАР (13.22), (13.23).

При  $h = \tau = 0,2$  для цієї задачі отримаємо точний розв'язок (табл. 13.1), наближений розв'язок (табл. 13.2) та похибку наближеного розв'язку (табл. 13.3).

Таблиця 13.1

Точний розв'язок

y	x					
	0	0,2	0,4	0,6	0,8	1
0	25,00000	27,04000	29,16000	31,36000	33,64000	36,00000
0,2	20,83333	22,53333	24,30000	26,13333	28,03333	30,00000
0,4	17,85714	19,31429	20,82857	22,40000	24,02857	25,71429
0,6	15,62500	16,90000	18,22500	19,60000	21,02500	22,50000
0,8	13,88889	15,02222	16,20000	17,42222	18,68889	20,00000
1	12,50000	13,52000	14,58000	15,68000	16,82000	18,00000

Таблиця 13.2

**Наближений розв'язок у внутрішніх  
вузлах сітки за методом Гаусса**

y	x			
	0,2	0,4	0,6	0,8
0,2	22,55536	24,33113	26,16592	28,05852
0,4	19,33550	20,85990	22,43253	24,05225
0,6	16,91506	18,24773	19,62349	21,04157
0,8	15,03015	16,21201	17,43461	18,69757

Таблиця 13.3

**Похибка**

y	x			
	0,2	0,4	0,6	0,8
0,2	0,02203	0,03113	0,03258	0,02519
0,4	0,02121	0,03133	0,03253	0,02368
0,6	0,01506	0,02273	0,02349	0,01657
0,8	0,00793	0,01201	0,01239	0,00868

Максимальне значення похибки – 0,0325825556.

Мінімальне значення похибки – 0,0079281650.

Розв'язок, отриманий при  $h = \tau = 0,1$ , наведено в табл. 13.4.

Таблиця 13.4

**Наближений розв'язок у внутрішніх вузлах сітки  
за методом Гаусса**

y	x			
	0,2	0,4	0,6	0,8
0,2	22,53919	24,30821	26,14194	28,04006
0,4	19,31985	20,83680	22,4085	24,03479
0,6	16,90392	18,23094	19,60613	21,02931
0,8	15,02428	16,20312	17,42544	18,69114

Таблиця 13.5

**Похибка**

y	x			
	0,2	0,4	0,6	0,8
0,2	0,00586	0,00821	0,00861	0,00673
0,4	0,00556	0,00822	0,00854	0,00622
0,6	0,00392	0,00594	0,00613	0,00431
0,8	0,00206	0,00312	0,00322	0,00225

Максимальне значення похибки – 0,0091957230.

Мінімальне значення похибки – 0,0006100639.

**Б.** Ітераційні методи розв'язування СЛАР (13.22), (13.23):

**а)** метод простої ітерації

$$u_{m,n}^{(s)} = \left( \frac{u_{m+1,n}^{(s-1)} + u_{m-1,n}^{(s-1)}}{h^2} + \frac{u_{m,n+1}^{(s-1)} + u_{m,n-1}^{(s-1)}}{\tau^2} - f_{m,n} \right) / \left( \frac{2}{h^2} + \frac{2}{\tau^2} \right), \quad (13.24)$$

де  $s$  – номер ітерації,  $u_{m,n}^{(0)} = 0$  у внутрішніх вузлах, а на границі – граничні значення.

Нехай  $h = 0,2$  і  $\tau = 0,2$ . Розв'язок наведено в табл. 13.6.

Таблиця 13.6

**Значення наближеного розв'язку  
за методом простої ітерації**

y	x			
	0,2	0,4	0,6	0,8
0,2	22,55536	24,33112	26,16591	28,05852
0,4	19,33549	20,85990	22,43253	24,05225
0,6	16,91506	18,24773	19,62348	21,04157
0,8	15,03015	16,21201	17,43461	18,69757

Таблиця 13.7

**Похибка**

y	x			
	0,2	0,4	0,6	0,8
0,2	0,02202	0,03112	0,03258	0,02519
0,4	0,02121	0,03132	0,03253	0,02368
0,6	0,01506	0,02273	0,02348	0,01657
0,8	0,00793	0,01201	0,01239	0,00868

Кількість ітерацій – 75.

Максимальне значення похибки – 0,0325803736.

Мінімальне значення похибки – 0,0079268184.

Нехай  $h = 0,1$  і  $\tau = 0,1$ . Розв'язок наведено в табл. 13.8, його похибка – у табл. 13.9.

Таблиця 13.8

**Значення наближеного розв'язку  
за методом простої Ітерації**

y	x			
	0,2	0,4	0,6	0,8
0,2	22,53919	24,30820	26,14193	28,04006
0,4	19,31984	20,83678	22,40853	24,03478
0,6	16,90391	18,23092	19,60611	21,02930
0,8	15,02427	16,20311	17,42543	18,69113

Таблиця 13.9

**Похибка**

y	x			
	0,2	0,4	0,6	0,8
0,2	0,00585	0,00820	0,00860	0,00672
0,4	0,00555	0,00821	0,00853	0,00621
0,6	0,00391	0,00592	0,00611	0,00430
0,8	0,00205	0,00311	0,00321	0,00224

Кількість ітерацій – 286.

Максимальне значення похибки – 0,0091801477.

Мінімальне значення похибки – 0,0091801477.

б) метод Зейделя

$$u_{m,n}^{(s)} = \left( \frac{u_{m+1,n}^{(s-1)} + u_{m,n+1}^{(s)} + u_{m,n}^{(s-1)}}{h^2} + \frac{u_{m,n+1}^{(s-1)} + u_{m,n}^{(s)}}{\tau^2} - f_{m,n} \right) / \left( \frac{2}{h^2} + \frac{2}{\tau^2} \right). \quad (13.25)$$

Ітераційні процеси закінчуються, коли

$$\left| u_{m,n}^{(s)} - u_{m,n}^{(s-1)} \right| < \varepsilon = 0,000001; \quad \forall (m, n) \in \omega_h,$$

де  $\omega_h$  — множина внутрішніх вузлів.

Нехай  $h = 0,2$  і  $\tau = 0,2$ . Розв'язок наведено в табл. 13.10, його похибка у табл. 13.11.

Таблиця 13.10

Значення наближеного розв'язку за методом Зейделя

y	x			
	0,2	0,4	0,6	0,8
0,2	22,55536	24,33113	26,16591	28,05852
0,4	19,33549	20,85990	22,43253	24,05225
0,6	16,91506	18,24773	19,62348	21,04157
0,8	15,03015	16,21201	17,43461	18,69757

Таблиця 13.11

Похибка

y	x			
	0,2	0,4	0,6	0,8
0,2	0,02202	0,03112	0,03258	0,02519
0,4	0,02121	0,03132	0,03253	0,02368
0,6	0,01506	0,02273	0,02348	0,01657
0,8	0,00793	0,01201	0,01239	0,00868

Кількість ітерацій – 40.

Максимальне значення похибки – 0,0325814800.

Мінімальне значення похибки – 0,0079276272.

Нехай  $h = 0,1$  і  $\tau = 0,1$ . Розв'язок наведено в табл. 13.12, його похибка – у табл. 13.13.

Таблиця 13.12

Значення наближеного розв'язку за методом Зейделя

y	x			
	0,2	0,4	0,6	0,8
0,2	22,55574	24,33151	26,16634	28,05901
0,4	19,33575	20,86027	22,43293	24,05255
0,6	16,91520	18,24797	19,62373	21,04172
0,8	15,03022	16,21212	17,43473	18,69764

Таблиця 13.13

Похибка

y	x			
	0,2	0,4	0,6	0,8
0,2	0,00586	0,00820	0,00860	0,00673
0,4	0,00556	0,00822	0,00854	0,00621
0,6	0,00391	0,00593	0,00612	0,00430
0,8	0,00205	0,00312	0,00322	0,00225

Кількість ітерацій – 151.

Максимальне значення похибки – 0,0091877718.

Мінімальне значення похибки – 0,0006092150.

**Висновок.** Зі збільшенням розбиття похибка зменшується.

### 13.2. Метод сіток розв'язування змішаної задачі для рівняння теплопровідності

Розглянемо диференціальне рівняння вигляду

$$Lu = \frac{\partial u}{\partial t} - a \frac{\partial^2 u}{\partial x^2} = f(x, t), \quad (13.26)$$

де  $a = a(x, t) > 0$ .

Нехай необхідно знайти розв'язок цього рівняння  $u(x, t)$  в області  $\pi = \{c \leq x \leq d; 0 \leq t \leq T\}$ , що задовольняє початкову

$$u(x, 0) = \phi(x), \quad x \in [c, d] \quad (13.27)$$

та граничні умови

$$u(c, t) = \gamma_0(t), \quad u(d, t) = \gamma_2(t), \quad 0 \leq t \leq T \quad (13.28)$$

Для знаходження наближеного розв'язку задачі (13.26)–(13.28) методом сіток розглянемо прямокутну сітку вузлів, яку утворює сім'я прямих

$$x = mh, \quad y = n\tau, \quad m, n = 0, \pm 1, \pm 2, \dots$$

Тоді позначимо

$$u(x, t) = u(mh, n\tau) = u_{m,n},$$

$$u(x + h, t) = u(mh + h, n\tau) = u_{m+1,n}.$$

Для кожного вузла  $(m, n)$  запишемо різницеве рівняння, яке апроксимує рівняння (13.26) з деякою точністю. Залежно від апроксимації похідної за часом отримаємо різні різницеві схеми.

А. Похідні в (13.26) замінимо різницевими аналогами

$$\left. \frac{\partial^2 u}{\partial x^2} \right|_{m,n} \approx \frac{u_{m+1,n} - 2u_{m,n} + u_{m-1,n}}{h^2}, \quad (13.29)$$

$$\left. \frac{\partial u}{\partial t} \right|_{m,n} \approx \frac{u_{m,n+1} - u_{m,n}}{\tau}. \quad (13.30)$$

Тоді

$$Lu|_{m,n} = \frac{u_{m,n+1} - u_{m,n}}{\tau} - a_{m,n} \frac{u_{m+1,n} - 2u_{m,n} + u_{m-1,n}}{h^2} = f_{m,n}. \quad (13.31)$$

Перетворимо останнє рівняння, отримаємо різницеве рівняння

$$Lu|_{m,n} = \frac{u_{m+1,n} + u_{m-1,n}}{h^2} + \left( \frac{1}{r} - 2 \right) \frac{u_{m,n}}{h^2} - \frac{1}{r} \frac{u_{m,n+1}}{h^2} = -f_{m,n} / a_{m,n}, \quad (13.32)$$

де  $r = \frac{\tau a_{m,n}}{h^2}$ .

Початкова умова  $u_{m,0} = \phi_m$ . Граничні умови  $u_{m,n}|_{\gamma_n} = \psi_{m,n}$ .

Із (13.32) визначимо  $u_{m,n+1}$ :

$$u_{m,n+1} = r(u_{m+1,n} + u_{m-1,n}) + (1 - 2r)u_{m,n} + \tau f_{m,n}. \quad (13.33)$$

Шаблон різницевої схеми (13.32)–(13.33), яку називають явною, зображений на рис. 13.1. Ця схема апроксимує диференціальне рівняння (13.26) з точністю  $O(\tau + h^2)$ .

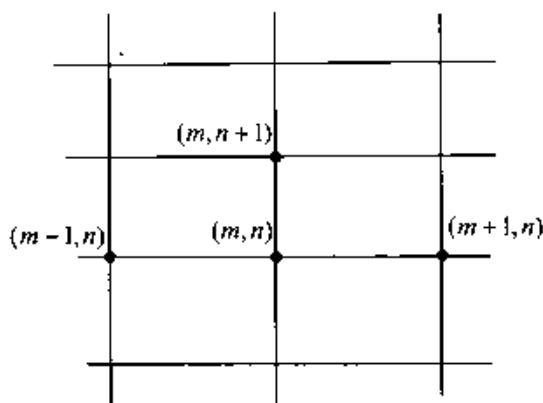


Рис. 13.1.

Різницева схема стійка при  $r \leq \frac{1}{2}$ . Найбільш уживаними є різницеві схеми при  $r = \frac{1}{6}$  та  $r = \frac{1}{2}$ . Наведемо відповідні розрахункові формули:

$$u_{m,n+1} = \frac{1}{6}(u_{m+1,n} + 4u_{m,n} + u_{m-1,n}) + \frac{h^2}{6a_{m,n}} f_{m,n} \quad (13.34)$$

та

$$u_{m,n+1} = \frac{1}{2}(u_{m-1,n} + u_{m+1,n}) + \frac{h^2}{2a_{m,n}} f_{m,n} \quad (13.35)$$

з початковими умовами  $u_{m,0} = \phi_m$ .

Для  $r > \frac{1}{2}$  збіжності різницевої схеми (13.33) до розв'язку  $u(x,t)$  при  $h \rightarrow 0$  не буде.

Б. Виконаємо апроксимацію похідної за часом як  $\frac{u_{m,n} - u_{m,n-1}}{\tau}$ , отримаємо неявне різницеве рівняння, яке стійке для всіх  $r$ :

$$r(u_{m+1,n} + u_{m-1,n}) - (1 + 2r)u_{m,n} = -u_{m,n-1} - \tau f_{m,n}. \quad (13.36)$$

Найпростіший вигляд (13.36) має при  $r = 2$ ,  $f = 0$

$$2u_{m+1,n} - 5u_{m,n} + 2u_{m-1,n} = -u_{m,n-1}. \quad (13.37)$$

Рівняння (13.36) апроксимує диференціальне рівняння (13.26) з точністю  $O(\tau + h^2)$  і використовує чотири вузли (рис. 13.2).

Для різницевого рівняння (13.34) похибка  $O(\tau + h^4)$  при  $f(x,t) = 0$ . У випадку розрахунку за формулою (13.36) виникають труднощі: ми повинні з яких-небудь міркувань знайти значення  $u$  на першому шарі, бо початкова умова визначає розв'язок лише на нульовому шарі. Для різницевого рівняння (13.36) похибка швидко зростає, тому на практиці її не застосовують.



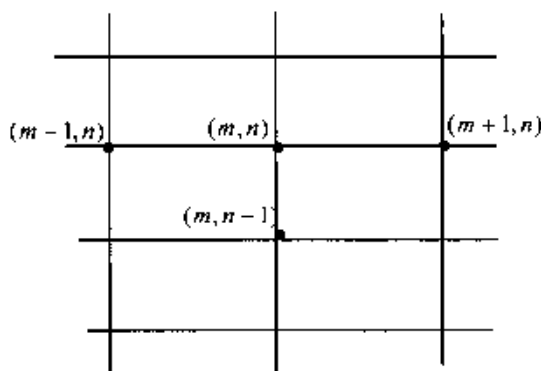


Рис. 13.2.

В. Застосуємо для апроксимації похідної за часом різницеве співвідношення

$$\frac{\partial u}{\partial t} \approx \frac{u_{m,n+1} - u_{m,n-1}}{2\tau},$$

отримаємо

$$\frac{u_{m,n+1} - u_{m,n-1}}{2\tau} - a_{m,n} \frac{u_{m+1,n} - 2u_{m,n} + u_{m-1,n}}{h^2} = f_{m,n}.$$

Звідси

$$u_{m,n+1} = \frac{2\tau a_{m,n}}{h^2} (u_{m+1,n} - 2u_{m,n} + u_{m-1,n}) + u_{m,n-1} + 2\tau f_{m,n}. \quad (13.38)$$

У це рівняння входять значення п'яти вузлів (рис.13.3), і апроксимація рівняння (13.26) в цьому випадку буде  $O(\tau^2 + h^2)$ . Для вузлів нульового шару з початкової умови маємо значення  $u_{m,0} = \phi_m$  ( $m = 0, \pm 1, \pm 2, \dots$ )

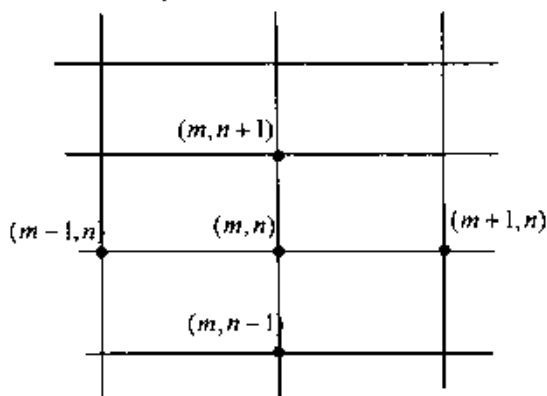


Рис. 13.3.

Різницева схема (13.38) є явною.

На практиці різницеву схему (13.38) не використовують, оскільки вона є нестійкою.

**Приклад 13.2.** Розв'язати рівняння

$$\frac{\partial u}{\partial t} - a(x,t) \frac{\partial^2 u}{\partial x^2} = f(x,t),$$

де  $f(x) = 2 \sin x \cdot e^{-t}(xt - 1)$ ,  $a(x,t) = xt$  для  $x \in [0,1], t \in [0, \dots, 5]$ , яке задовольняє умови

$$\begin{cases} u(0,t) = 0; \\ u(1,t) = 2 \sin 1 \cdot e^{-t}; \\ u(x,0) = 2 \sin x. \end{cases}$$

Точний розв'язок  $u_T = 2e^{-t} \cdot \sin x$ .

**Розв'язування.** Точний розв'язок  $u_T$  протабульовано в табл. 13.14.

Таблиця 13.14

Точний розв'язок

x	t					
	0	1	2	3	4	5
0	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
0,1	0,19967	0,07345	0,02702	0,00994	0,00366	0,00135
0,2	0,39734	0,14617	0,05377	0,01978	0,00728	0,00268
0,3	0,59104	0,21743	0,07999	0,02943	0,01083	0,00398
0,4	0,77884	0,28652	0,10540	0,03878	0,01426	0,00525
0,5	0,95885	0,35274	0,12977	0,04774	0,01756	0,00646
0,6	1,12928	0,41544	0,15283	0,05622	0,02068	0,00761
0,7	1,28844	0,47399	0,17437	0,06415	0,02360	0,00868
0,8	1,43471	0,52780	0,19417	0,07143	0,02628	0,00967
0,9	1,56665	0,57634	0,21202	0,07800	0,02869	0,01056
1	1,68294	0,61912	0,22776	0,08379	0,03082	0,01134

#### А. Явна схема

Нехай  $h = 0,1$  і  $\tau = 1$ . Схема нестійка, бо не виконується умова

$r \leq \frac{1}{2}$ . Результати запишемо в табл. 13.15.

Таблиця 13.15

Значення наближеного розв'язку за явною схемою

x	t				
	1	2	3	4	5
0,1	0,19967	0,17311	-0,86447	2,11761	147,58241
0,2	0,39734	0,29455	-1,62946	7,87185	217,22538
0,3	0,59104	0,36822	-2,23602	16,24304	191,10649
0,4	0,77884	0,39880	-2,63723	26,07139	60,68366
0,5	0,95885	0,15034	0,05415	0,02031	0,00777

**Б. Неявна схема**Нехай  $h = 0,1$  і  $\tau = 1$ .

Результати запишемо в табл. 13.16, а похибку наближеного розв'язку – у табл. 13.17.

Таблиця 13.16

Значення наближеного розв'язку за неявною схемою

x	t				
	1	2	3	4	5
0,1	0,10257	0,03590	0,01223	0,00444	0,00168
0,2	0,19608	0,06868	0,02374	0,00870	0,00330
0,3	0,28071	0,09860	0,03459	0,01279	0,00487
0,4	0,35651	0,12580	0,04475	0,01668	0,00637
0,5	0,42347	0,15034	0,05415	0,02031	0,00777
0,6	0,48149	0,17214	0,06268	0,02359	0,00902
0,7	0,53037	0,19105	0,07015	0,02642	0,01007
0,8	0,56985	0,20685	0,07634	0,02867	0,01085
0,9	0,59957	0,21921	0,08098	0,03019	0,01130

Таблиця 13.17

Похибка

x	t				
	1	2	3	4	5
0,1	0,02911	0,00888	0,00229	0,00078	0,00033
0,2	0,04991	0,01491	0,00396	0,00142	0,00062
0,3	0,06328	0,01861	0,00516	0,00197	0,00089
0,4	0,06999	0,02040	0,00597	0,00242	0,00113
0,5	0,07073	0,02057	0,00641	0,00275	0,00131
0,6	0,06605	0,01930	0,00646	0,00291	0,00141
0,7	0,05638	0,01668	0,00600	0,00282	0,00139
0,8	0,04205	0,01268	0,00491	0,00239	0,00118
0,9	0,02323	0,00719	0,00298	0,00150	0,00075

Максимальне значення похибки – 0,0707285193.

Мінімальне значення похибки – 0,00033.

#### А. Явна схема

Нехай  $h = 0,1$  і  $\tau = 0,001$ . Схема стійка, бо виконується умова

$$r \leq \frac{1}{2}.$$

Результати запишемо в табл. 13.18, а похибку отриманого розв'язку – у табл.13.19.

Таблиця 13.18

**Значення наближеного розв'язку за явною схемою**

$x$	$t$				
	1	2	3	4	5
0,1	0,11671	0,02924	0,01059	0,00406	0,00157
0,2	0,21686	0,05707	0,02092	0,00804	0,00312
0,3	0,30244	0,08369	0,03098	0,01194	0,00464
0,4	0,37529	0,10915	0,04074	0,01573	0,00611
0,5	0,43700	0,13339	0,05010	0,01933	0,00750
0,6	0,48896	0,15627	0,05890	0,02267	0,00876
0,7	0,53235	0,17753	0,06693	0,02563	0,01312
0,8	0,56807	0,19683	0,07396	0,02808	0,01331
0,9	0,59683	0,21373	0,07968	0,02987	0,01275

Таблиця 13.19

**Похибка**

$x$	$t$				
	1	2	3	4	5
0,1	0,04325	0,00222	0,00065	0,00040	0,00023
0,2	0,07068	0,00330	0,00113	0,00076	0,00044
0,3	0,08501	0,00370	0,00155	0,00112	0,00066
0,4	0,08873	0,00374	0,00196	0,00146	0,00086
0,5	0,08426	0,00363	0,00236	0,00177	0,00104
0,6	0,07352	0,00344	0,00267	0,00199	0,00115
0,7	0,05836	0,00316	0,00279	0,00203	0,00116
0,8	0,04027	0,00266	0,00253	0,00180	0,00102
0,9	0,02049	0,00171	0,00168	0,00117	0,00065

Максимальне значення похибки – 0,0707285193.

Мінімальне значення похибки – 0,0001995670.

**В. Неявна схема**

Нехай  $h = 0,1$  і  $\tau = 0,001$ . Результати запишемо в табл. 13.20, а похибку наближеного розв'язку – у табл. 13.21.

Таблиця 13.20

**Значення наближеного розв'язку за неявною схемою**

$x$	$t$				
	1	2	3	4	5
0,1	0,11666	0,02925	0,01060	0,00406	0,00157
0,2	0,21679	0,05710	0,02092	0,00804	0,00312
0,3	0,30239	0,08372	0,03098	0,01194	0,00464
0,4	0,37524	0,10918	0,04074	0,01573	0,00611
0,5	0,43697	0,13342	0,05010	0,01933	0,00750
0,6	0,48895	0,15630	0,05890	0,02267	0,00876
0,7	0,53235	0,17756	0,06694	0,02563	0,00985
0,8	0,56807	0,19685	0,07396	0,02808	0,01069
0,9	0,59684	0,21374	0,07968	0,02987	0,01121

Таблиця 13.21

**Похибка**

$x$	$t$				
	1	2	3	4	5
0,1	0,04321	0,00223	0,00066	0,00040	0,00023
0,2	0,07062	0,00332	0,00114	0,00076	0,00044
0,3	0,08495	0,00373	0,00156	0,00112	0,00066
0,4	0,08877	0,00377	0,00197	0,00146	0,00086
0,5	0,08423	0,00366	0,00236	0,00177	0,00104
0,6	0,07351	0,00347	0,00268	0,00199	0,00115
0,7	0,05836	0,00319	0,00279	0,00204	0,00116
0,8	0,04027	0,00268	0,00253	0,00180	0,00102
0,9	0,02050	0,00172	0,00168	0,00117	0,00065

Максимальне значення похибки – 0,0887696670.

Мінімальне значення похибки – 0,00023.

### 13.3. Метод сіток розв'язування змішаної задачі для рівняння коливань струни

Нехай потрібно знайти  $u(x, t)$  – розв'язок рівняння

$$\frac{\partial^2 u(x, t)}{\partial t^2} = a^2(x, t) \frac{\partial^2 u(x, t)}{\partial x^2} + \phi(x, t) \quad (13.39)$$

$$0 \leq x \leq 1, 0 \leq t \leq T$$

який задовольняє граничні умови

$$u(0, t) = \gamma_0(t), \quad u(1, t) = \gamma_1(t), \quad 0 \leq t \leq T \quad (13.40)$$

і початкові умови

$$u(x, 0) = \alpha(x), \quad \frac{\partial u(x, 0)}{\partial t} = \beta(x), \quad 0 \leq x \leq 1. \quad (13.41)$$

За сітку виберемо сукупність точок  $(x_m, t_n)$  з координатами  $x_m = mh$ ,  $t_n = n\tau$ ,  $m = 0, 1, \dots, M$ ,  $n = 0, 1, \dots, N$ ,  $h = 1/M$ ,  $\tau = 1/N$ .

Замінімо другі похідні різницевиими співвідношеннями з (13.39), отримаємо різницеве рівняння

$$u_{m,n+1} = s_{m,n} u_{m+1,n} + 2(1 - s_{m,n}) u_{m,n} + s_{m,n} u_{m-1,n} - u_{m,n-1} + \tau \phi(x_m, t_n) \quad (13.42)$$

$$m = 1, 2, \dots, M-1, \quad n = 1, 2, \dots, N-1,$$

$$s_{m,n} = \frac{\tau^2}{h^2} a_{m,n}^2,$$

яке апроксимує на п'ятиточковому шаблоні (рис. 13.3) диференціальне рівняння (13.39) з похибкою  $M_4(\tau^2 + Ah^2)/12$ , де

$$M_4 = \max_{x,t} \left\{ \left| \frac{\partial^4 u}{\partial t^4} \right|, \left| \frac{\partial^4 u}{\partial x^4} \right| \right\}, \quad A = \max_{x,t} a^2(x, t)$$

З умов (13.40) маємо

$$u_{0,n} = \gamma_0(t_n), \quad u_{M,n} = \gamma_1(t_n), \quad n = 0, 1, 2, \dots, N \quad (13.43)$$

Ми зможемо знайти  $u_{m,n+1}$  при  $n = 1, \dots, N-1$  за формулами (13.42) і (13.43), якщо будуть відомі  $u_{m,0}$  і  $u_{m,1}$  при  $m = 1, 2, \dots, M-1$ .

Для обчислення  $u_{m,0}$  і  $u_{m,1}$  використовують умови (13.41).

### Спосіб I.

Скористаємось зображенням похідної за часом

$$\frac{\partial u(x, t_0)}{\partial t} = \frac{u(x, t_1) - u(x, t_0)}{\tau} - \frac{\tau}{2} \frac{\partial^2 u(x, \tilde{t}_0)}{\partial t^2},$$

$$t_0 \leq \tilde{t}_0 \leq t_1,$$

з (13.41) отримаємо

$$u_{m,0} = \alpha_m, \quad u_{m,1} = \alpha_m + \tau \beta_m, \quad m = 1, 2, \dots, M-1. \quad (13.44)$$

Похибку апроксимації початкових умов (13.41) рівностями

$$(13.44) \text{ оцінюють величиною } \frac{\tau M_t^{(2)}}{2}, \text{ де } M_t^{(2)} = \max_{x, t} \left| \frac{\partial^2 u}{\partial t^2} \right|$$

### Спосіб II.

За формулою Тейлора маємо

$$u(x, t_1) = u(x, t_0) + \frac{\tau}{1!} \frac{\partial u(x, t_0)}{\partial t} + \frac{\tau^2}{2!} \frac{\partial^2 u(x, t_0)}{\partial t^2} + \frac{\tau^3}{3!} \frac{\partial^3 u(x, \tilde{t}_1)}{\partial t^3}, \quad t_0 \leq \tilde{t}_1 \leq t_1.$$

Звідси

$$\frac{\partial u(x, t_0)}{\partial t} = \frac{u(x, t_1) - u(x, t_0)}{\tau} - \frac{\tau}{2} \frac{\partial^2 u(x, t_0)}{\partial t^2} - \frac{\tau^2}{6} \frac{\partial^3 u(x, \tilde{t}_1)}{\partial t^3}. \quad (13.45)$$

Якщо  $\alpha(x)$  має скінченну другу похідну, то з (13.39) і (13.41) отримаємо

$$\begin{aligned} \frac{\partial^2 u(x, t_0)}{\partial t^2} &= a^2(x, t_0) \frac{\partial^2 u(x, t_0)}{\partial x^2} + \phi(x, t_0) = \\ &= a^2(x, t_0) \alpha''(x) + \phi(x, t_0). \end{aligned} \quad (13.46)$$

Підставимо значення з (13.46) в (13.45), знайдемо

$$u_{m,0} = \alpha_m,$$

$$u_{m,1} = \alpha(x_m) + \tau \beta(x_m) + \frac{\tau^2}{2} \left[ a^2(x_m, 0) \alpha''(x_m) + \phi(x_m, 0) \right] \quad (13.47)$$

$$m = 1, 2, \dots, M-1.$$

Похибку апроксимації початкових умов рівняннями (13.47) оцінюють величиною  $\frac{\tau^2 M_t^{(3)}}{6}$ , де  $M_t^{(3)} = \max_{x,t} \left| \frac{\partial^3 u}{\partial t^3} \right|$ .

### Спосіб III. Рівняння

$$\frac{u_{m,1} - u_{m,-1}}{2\tau} = \beta_m, \quad u_{m,-1} = u(x_m, -\tau) \quad (13.48)$$

апроксимує другу початкову умову  $u'_t(x, 0) = \beta(x)$  з другим порядком. Щоб знайти значення  $u_m^1$ , запишемо різницеве рівняння для (13.42) при  $n = 0$ :

$$\frac{u_{m,1} - 2u_{m,0} + u_{m,-1}}{\tau^2} = \alpha_{m,0}^2 \frac{u_{m+1,0} - 2u_{m,0} + u_{m-1,0}}{h^2} + \phi(x_m, 0)$$

і врахуємо, що  $u_{m,0} = \alpha_m$ . Звідси отримаємо, що

$$\frac{u_{m,1} - 2u_{m,0} + u_{m,-1}}{\tau^2} = s_{m,0} \frac{u_{m+1,0} - 2u_{m,0} + u_{m-1,0}}{h^2} + \phi(x_m, t_0).$$

Підставимо цей вираз для  $u_{m,-1}$  у рівняння (13.48), отримаємо вираз для знаходження  $u_{m,1}$ . Отже,

$$u_{m,0} = \alpha_m,$$

$$u_{m,1} = \frac{2\alpha_m + 2\tau\beta_m + s_{m,0}(\alpha_{m+1} - 2\alpha_m + \alpha_{m-1}) + \tau^2\phi_{m,0}}{2}, \quad (13.49)$$

$$m = 1, 2, \dots, M-1.$$

Різницеві схеми (13.42)–(13.44); (13.42), (13.43), (13.47); (13.42), (13.43), (13.49) будуть стійкі в разі виконання умови  $\frac{\tau^2}{h^2} \leq \frac{1}{A}$ .

**Неявна схема** (Схема Кранка–Нікольсона). Розглянемо сім'ю рівнянь з вагами на дев'ятиточковому шаблоні (рис. 13.4):



$$\frac{u_{m,n+1} - 2u_{m,n} + u_{m,n-1}}{\tau^2} =$$

$$= A_{m,n} \left[ \sigma \frac{u_{m+1,n+1} - 2u_{m,n+1} + u_{m-1,n+1}}{h^2} + (1-2\sigma) \frac{u_{m+1,n} - 2u_{m,n} + u_{m-1,n}}{h^2} \right] +$$

$$+ A_{m,n} \sigma \frac{u_{m+1,n-1} - 2u_{m,n-1} + u_{m-1,n-1}}{h^2} + \phi(x_m, t_0).$$

Для знаходження  $u_{m,n+1}$ ,  $m=1,2,\dots,M-1$  потрібно розв'язати СЛАР:

$$\sigma \gamma^2 (u_{m+1,n+1} + u_{m-1,n+1}) - (1+2\sigma \gamma^2) u_{m,n+1} = -F_m,$$

$$F_m = 2u_{m,n} - u_{m,n-1} + \tau^2 S_{m,n} \left[ (1-2\sigma) \frac{u_{m+1,n} - 2u_{m,n} + u_{m-1,n}}{h^2} \right] +$$

$$+ \tau^2 S_{m,n} \left[ \sigma \frac{u_{m+1,n-1} - 2u_{m,n-1} + u_{m-1,n-1}}{h^2} \right] + \tau^2 \phi(x_m, t_n), \quad (13.50)$$

$$\gamma^2 = A_{m,n} \frac{\tau^2}{h^2}, \quad \text{де } A_{m,n} = a^2(x_m, t_n) = a_{m,n}^2.$$

Отриману СЛАР можна розв'язати методом прогонки. Умови стійкості методу прогонки виконуються при  $\sigma > 0$  (див. 2.5).

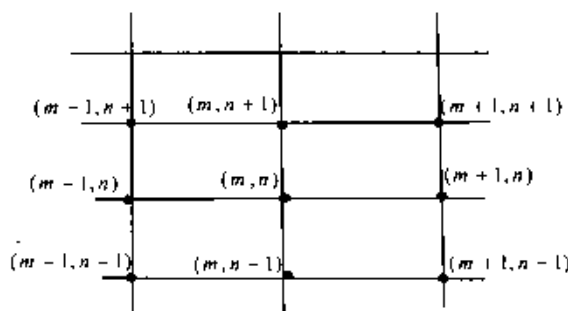


Рис. 13.4.

**П р и к л а д 13.3.** Знайти розв'язок рівняння гіперболічного типу:

$$\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{\partial^2 u(x,t)}{\partial x^2} + \frac{2\alpha}{[\alpha(x+t)+2]^2},$$

$$0 \leq x \leq 1, \quad 0 \leq t \leq 1,$$

$$u(0,t) = 0, \quad u(1,t) = \frac{1}{\alpha(t+1)+2}, \quad 0 \leq t \leq 1,$$

$$u(x,0) = \frac{x}{\alpha x + 2}, \quad \frac{\partial u(x,0)}{\partial t} = -\frac{\alpha x}{(\alpha x + 2)^2}, \quad 0 \leq x \leq 1,$$

$$\alpha = 0,2k, \quad k = 3, \quad u_T = \frac{x}{\alpha(x+t)+2}.$$

**Розв'язування.** Нехай  $M = N = 5$ . Точний і наближені розв'язки цієї задачі на дев'ятиточковому шаблоні (рис.13.4) за різними способами записані в табл. 13.22–13.39.

Таблиця 13.22

**Точний розв'язок**

$t$	$x$					
	0	0,2	0,4	0,6	0,8	1
0	0	0,09434	0,17857	0,25424	0,32258	0,38462
0,2	0	0,08929	0,16949	0,24194	0,30769	0,36765
0,4	0	0,08475	0,16129	0,23077	0,29412	0,35211
0,6	0	0,08065	0,15385	0,22059	0,28169	0,33784
0,8	0	0,07692	0,14706	0,21127	0,27027	0,32468
1	0	0,07353	0,14085	0,20270	0,25974	0,31250

**Явна схема**

**Спосіб I**

Таблиця 13.23

**Наближений розв'язок  
за формулами (13.44), (13.43), (13.42)**

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,08929	0,16951	0,24196	0,30772
0,4	0,08473	0,16129	0,23079	0,29412
0,6	0,08062	0,15382	0,22056	0,28167
0,8	0,07690	0,14699	0,21119	0,27022
1	0,07347	0,14075	0,20261	0,25967

Таблиця 13.24

## Похибка

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,00000	0,00001	0,00002	0,00003
0,4	0,00001	0,00000	0,00002	0,00001
0,6	0,00002	0,00002	0,00003	0,00002
0,8	0,00003	0,00007	0,00008	0,00005
1	0,00006	0,00009	0,00010	0,00007

Максимальне значення похибки –  $9,6283372666e-5$

Мінімальне значення похибки – 0

Спосіб II

Таблиця 13.25

## Наближений розв'язок

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,09182	0,17165	0,24379	0,30930
0,4	0,08688	0,16566	0,23452	0,29596
0,6	0,08246	0,15755	0,22493	0,28382
0,8	0,07848	0,14883	0,21334	0,27276
1	0,07347	0,14075	0,20261	0,25967

Таблиця 13.26

## Похибка

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,00254	0,00216	0,00186	0,00161
0,4	0,00213	0,00437	0,00375	0,00184
0,6	0,00181	0,00371	0,00434	0,00213
0,8	0,00155	0,00177	0,00207	0,00249
1	0,00006	0,00009	0,00010	0,00007

Максимальне значення похибки – 0,0043704092.

Мінімальне значення похибки –  $6,0631238561e-5$ .

Спосіб III

Таблиця 13.27

Наближений розв'язок за формулами (13.49), (13.43), (13.42)

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,09182	0,17165	0,24379	0,30930
0,4	0,08860	0,16505	0,23408	0,29672
0,6	0,08414	0,15825	0,22462	0,28439
0,8	0,07874	0,15086	0,21499	0,27225
1	0,07329	0,14283	0,20498	0,26053

Таблиця 13.28

Похибка

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,00254	0,00216	0,00186	0,00161
0,4	0,00385	0,00376	0,00331	0,00260
0,6	0,00349	0,00440	0,00403	0,00270
0,8	0,00181	0,00381	0,00373	0,00198
1	0,00024	0,00199	0,00228	0,00079

Максимальне значення похибки – 0,0044028182.

Мінімальне значення похибки – 0,00024244391.

Схема Кранка–Ніколсона для  $\sigma = 0,5$ ;  $M = N = 5$ .

Спосіб I

Таблиця 13.29

Наближений розв'язок  
за формулами (13.44), (13.43), (13.50)

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,09182	0,17165	0,24379	0,30930
0,4	0,08819	0,16512	0,23417	0,29655
0,6	0,08315	0,15824	0,22472	0,28400
0,8	0,07774	0,15039	0,21478	0,27185
1	0,07312	0,14166	0,20411	0,26039

Таблиця 13.30

## Похибка

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,00254	0,00216	0,00186	0,00161
0,4	0,00344	0,00383	0,00340	0,00244
0,6	0,00251	0,00440	0,00413	0,00231
0,8	0,00082	0,00333	0,00351	0,00158
1	0,00041	0,00082	0,00140	0,00065

Максимальне значення похибки – 0,0043973878.

Мінімальне значення похибки – 0,0004099400.

Нехай  $M=10$ ,  $N=10$ .

Таблиця 13.31

## Точний розв'язок

$t$	$x$					
	0	0,2	0,4	0,6	0,8	1
0	0	0,09434	0,17857	0,25424	0,32258	0,38462
0,2	0	0,08929	0,16949	0,24194	0,30769	0,36765
0,4	0	0,08475	0,16129	0,23077	0,29412	0,35211
0,6	0	0,08065	0,15385	0,22059	0,28169	0,33784
0,8	0	0,07692	0,14706	0,21127	0,27027	0,32468
1	0	0,07353	0,14085	0,20270	0,25974	0,31250

Спосіб II

Таблиця 13.32

## Наближений розв'язок

за формулами (13.37), (13.33), (13.40)

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,09055	0,17057	0,24286	0,30849
0,4	0,08582	0,16348	0,23264	0,29504
0,6	0,08156	0,15570	0,22277	0,28276
0,8	0,07771	0,14796	0,21232	0,27153
1	0,07351	0,14082	0,20268	0,25972

Таблиця 13.33

## Похибка

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,00127	0,00108	0,00092	0,00080
0,4	0,00107	0,00219	0,00187	0,00092
0,6	0,00091	0,00186	0,00218	0,00107
0,8	0,00078	0,00090	0,00105	0,00126
1	0,00002	0,00002	0,00002	0,00002

Максимальне значення похибки – 0,0025517738.

Мінімальне значення похибки –  $8,5126388165e-6$ .

## Спосіб III

Таблиця 13.34

Наближений розв'язок  
за формулами (13.49), (13.43), (13.50)

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,08929	0,16949	0,24194	0,30770
0,4	0,08474	0,16129	0,23077	0,29412
0,6	0,08064	0,15384	0,22058	0,28169
0,8	0,07692	0,14704	0,21125	0,27026
1	0,07351	0,14082	0,20268	0,25972

Таблиця 13.35

## Похибка

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,00000	0,00000	0,00001	0,00001
0,4	0,00000	0,00000	0,00001	0,00000
0,6	0,00001	0,00001	0,00001	0,00000
0,8	0,00001	0,00002	0,00002	0,00001
1	0,00002	0,00002	0,00002	0,00002

Максимальне значення похибки –  $2,4680303601e-5$ .

Мінімальне значення похибки – 0.

Схема Кранка–Ніколсона  $\sigma = 0,5$

Таблиця 13.36

Наближений розв'язок  
за формулами (13.49), (13.43), (13.50)

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,09044	0,17056	0,24285	0,30844
0,4	0,08612	0,16325	0,23251	0,29518
0,6	0,08136	0,15592	0,22266	0,28263
0,8	0,07749	0,14809	0,21265	0,27119
1	0,07382	0,14062	0,20267	0,26019

Таблиця 13.37

Похибка

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,00115	0,00107	0,00092	0,00075
0,4	0,00138	0,00196	0,00174	0,00106
0,6	0,00071	0,00207	0,00207	0,00094
0,8	0,00056	0,00103	0,00139	0,00092
1	0,00029	0,00022	0,00004	0,00045

Максимальне значення похибки – 0,0022587747.

Мінімальне значення похибки –:  $2,0600711464e-5$ .

Неявна схема  $\sigma = 1$

Таблиця 13.38

Наближений розв'язок  
за формулами (13.49), (13.43), (13.50)

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,09043	0,17054	0,24284	0,30844
0,4	0,08625	0,16318	0,23246	0,29523
0,6	0,08157	0,15594	0,22261	0,28270
0,8	0,07717	0,14837	0,21280	0,27100
1	0,07362	0,14074	0,20293	0,26014

Таблиця 13.39

## Похибка

$t$	$x$			
	0,2	0,4	0,6	0,8
0,2	0,00115	0,00105	0,00091	0,00074
0,4	0,00151	0,00189	0,00169	0,00112
0,6	0,00093	0,00210	0,00203	0,00101
0,8	0,00025	0,00131	0,00153	0,00073
1	0,00009	0,00011	0,00023	0,00040

Максимальне значення похибки – 0,0021948721.

Мінімальне значення похибки –  $7,2953224401e-5$ .



### Завдання до розділу 13

13.1. Побудувати різницеву схему, яка апроксимує граничну задачу для рівняння

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y), \quad x, y \in G,$$

$$u|_{\Gamma} = \varphi(x, y),$$

де  $G$  і  $\Gamma$  – область і границя області. Знайти розв'язок цієї задачі на сітці  $x_i = ih$ ,  $y_j = j\tau$ .

$$1) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{2}{1+y} \left[ 1 + \left( \frac{1+x}{1+y} \right)^2 \right], \quad 0 \leq x, y \leq 1,$$

$$u(x, 0) = (x+1)^2 \quad u(x, 1) = \frac{(x+1)^2}{2}, \quad 0 \leq x \leq 1;$$

$$u(0, y) = \frac{1}{1+y}, \quad u(1, y) = \frac{4}{1+y}, \quad 0 \leq y \leq 1;$$

$$u_i = \frac{(x+1)^2}{1+y};$$

$$2) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 8,$$

$$u(x, 0) = x^2 \quad u(x, 1) = x^2 + 3, \quad 0 \leq x \leq 1;$$

$$u(0, y) = 3y^2 \quad u(1, y) = 1 + 3y^2 \quad 0 \leq y \leq 1;$$

$$u_i = x^2 + 3y^2$$

$$3) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 2(x^2 + y^2),$$

$$u(x, 0) = 0, \quad u(x, 1) = x^2$$

$$u(0, y) = 0, \quad u(1, y) = y^2;$$

$$u_{\tau} = x^2 y^2;$$

$$4) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 4, \quad 0 \leq x \leq 0,8, \quad 0 \leq y \leq 0,8;$$

$$u(x, 0) = x^2, \quad u(x, 0,8) = x^2 + 0,64, \quad 0 \leq x \leq 0,8;$$

$$u(0, y) = y^2, \quad u(0,8; y) = y^2 + 0,64, \quad 0 \leq y \leq 0,8;$$

$$u_{\tau} = x^2 + y^2$$

$$5) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 10x, \quad 0 \leq x, y \leq 1;$$

$$u(x, 0) = 0, \quad u(x, 1) = 5x, \quad 0 \leq x \leq 1;$$

$$u(0, y) = 0, \quad u(1, y) = 5y^2, \quad 0 \leq y \leq 1;$$

$$u_{\tau} = 5xy^2$$

$$6) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 6(x + y), \quad 0 \leq x, y \leq 1;$$

$$u(x, 0) = x^3, \quad u(x, 1) = 1 + x^3, \quad 0 \leq x \leq 1;$$

$$u(0, y) = y^3, \quad u(1, y) = 1 + y^3, \quad 0 \leq y \leq 1;$$

$$u_{\tau} = x^3 + y^3$$

$$7) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -\frac{1}{4} \left( \frac{1}{x\sqrt{x}} + \frac{1}{y\sqrt{y}} \right), \quad 1 \leq x, y \leq 2;$$

$$u(x, 1) = \sqrt{x} + 1, \quad u(x, 2) = \sqrt{2} + \sqrt{x}, \quad 1 \leq x \leq 2;$$

$$u(1, y) = 1 + \sqrt{y}, \quad u(2, y) = \sqrt{2} + \sqrt{y}, \quad 1 \leq y \leq 2;$$

$$u_{\tau} = \sqrt{x} + \sqrt{y};$$

$$8) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 12(x^2 + y^2), \quad 0 \leq x, y \leq 1;$$

$$u(x, 0) = x^4 \quad u(x, 1) = 1 + x^4 \quad 0 \leq x \leq 1;$$

$$u(0, y) = y^4 \quad u(1, y) = 1 + y^4 \quad 0 \leq y \leq 1;$$

$$u_{\tau} = x^4 + y^4$$

$$9) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 6xy, \quad 0 \leq x, y \leq 1;$$

$$u(x, 0) = 0, \quad u(x, 1) = x, \quad 0 \leq x \leq 1;$$

$$u(0, y) = 0, \quad u(1, y) = y^3 \quad 0 \leq y \leq 1;$$

$$u_{\tau} = xy^3$$

$$10) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 6xy, \quad 0 \leq x, y \leq 1;$$

$$u(x, 0) = 0, \quad u(x, 1) = x^3 \quad 0 \leq x \leq 1;$$

$$u(0, y) = 0, \quad u(1, y) = y, \quad 0 \leq y \leq 1;$$

$$u_{\tau} = x^3 y;$$

$$11) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 2(1 + 6y^2), \quad 0 \leq x, y \leq 1;$$

$$u(x, 0) = x^2 \quad u(x, 1) = 1 + x^2 \quad 0 \leq x \leq 1;$$

$$u(0, y) = y^4 \quad u(1, y) = 1 + y^4 \quad 0 \leq y \leq 1;$$

$$u_{\tau} = x^2 + y^4$$

$$12) \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -2 \sin x \sin y, \quad 0 \leq x, y \leq \frac{\pi}{2};$$

$$u(x, 0) = 0, \quad u\left(x, \frac{\pi}{2}\right) = \sin x, \quad 0 \leq x \leq \frac{\pi}{2};$$

$$u(0, y) = 0, \quad u\left(\frac{\pi}{2}, y\right) = \sin y, \quad 0 \leq y \leq \frac{\pi}{2};$$

$$u_{\tau} = \sin x \sin y;$$

$$\begin{aligned}
 13) \quad & \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -x \sin y - y \sin x, \Omega = \left( 0 \leq x, y \leq \frac{\pi}{2} \right); \\
 & u(x, 0) = 0, \quad u\left(x, \frac{\pi}{2}\right) = x + \frac{\pi}{2} \sin x, 0 \leq x \leq \frac{\pi}{2}; \\
 & u(0, y) = 0, \quad u\left(\frac{\pi}{2}, y\right) = \frac{\pi}{2} \sin y + y, \quad 0 \leq y \leq \frac{\pi}{2}; \\
 & u_t = x \sin y + y \sin x;
 \end{aligned}$$

$$\begin{aligned}
 14) \quad & \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 6(x + y), 0 \leq x, y \leq 1; \\
 & u(0, y) = y^3, \quad u(1, y) = 1 + y^3, \quad 0 \leq y \leq 1; \\
 & u(x, 0) = x^3, \quad u(x, 1) = 1 + x^3, \quad 0 \leq x \leq 1; \\
 & u_t = x^3 + y^3
 \end{aligned}$$

$$\begin{aligned}
 15) \quad & \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{12(x+5)}{(1+y)^5}, \quad 0 \leq x, y \leq 1; \\
 & u(0, y) = \frac{5}{(1+y)^3}, \quad u(1, y) = \frac{6}{(1+y)^3}, \quad 0 \leq y \leq 1; \\
 & u(x, 0) = x + 5, \quad u(x, 1) = \frac{1}{8}(x + 5), \quad 0 \leq x \leq 1; \\
 & u_t = \frac{x+5}{(1+y)^3}.
 \end{aligned}$$

13.2. З використанням методу сіток знайти розв'язок рівняння теплопровідності

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(x, t) \quad 0 \leq x \leq 1, \quad 0 \leq t \leq 0,02$$

який задовольняє умови

$$\begin{aligned}
 & u(x, 0) = \varphi(x), \quad 0 \leq x \leq 1; \\
 & u(0, t) = \gamma_0(t), \quad u(1, t) = \gamma_1(t), \quad 0 \leq t \leq 0,02;
 \end{aligned}$$

- а) за явною різницевою схемою для  $h = 0,1$ ;  $\tau = 0,005$ ;  
 б) за неявною різницевою схемою.  
 в) порівняти одержані розв'язки з точним.

- 1) 
$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + 0,5(x^2 - 2t),$$

$$u(x, 0) = 0, \quad 0 \leq x \leq 1;$$

$$u(0, t) = 0, \quad u(1, t) = 0,5t, \quad 0 \leq t \leq 0,02;$$

$$u_T(x, t) = 0,5x^2t;$$
- 2) 
$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + (1 + 3x)^4 - 108t(1 + 3x)^2$$

$$u(x, 0) = 0, \quad 0 \leq x \leq 1;$$

$$u(0, t) = t, \quad u(1, t) = 256t, \quad 0 \leq t \leq 0,02;$$

$$u_T(x, t) = t(1 + 3x)^4$$
- 3) 
$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + e^{t-x}$$

$$u(x, 0) = 0, \quad 0 \leq x \leq 1;$$

$$u(0, t) = te^t, \quad u(1, t) = te^{t-1}, \quad 0 \leq t \leq 0,02;$$

$$u_T(x, t) = te^{(t-x)};$$
- 4) 
$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + (x + 1)^2 - 2(t + 1),$$

$$u(x, 0) = (x + 1)^2 \quad 0 \leq x \leq 1;$$

$$u(0, t) = (t + 1), \quad u(1, t) = 4(t + 1), \quad 0 \leq t \leq 0,02;$$

$$u_T(x, t) = (t + 1)(x + 1)^2$$
- 5) 
$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + 2t(3x^3 + x^2 + 4x + 5) - 2t^2(9x + 1),$$

$$u(x, 0) = 0, \quad 0 \leq x \leq 1;$$

$$u(0, t) = 5t^2, \quad u(1, t) = 13t^2, \quad 0 \leq t \leq 0,02;$$

$$u_T(x, t) = t^2(3x^3 + x^2 + 4x + 5);$$

- 6)  $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + (1+2x)^2 \left[ (1+2x)^2 - 48t \right],$   
 $u(x, 0) = 0, \quad 0 \leq x \leq 1;$   
 $u(0, t) = t, \quad u(1, t) = 81t, \quad 0 \leq t \leq 0,02;$   
 $u_T(x, t) = t(1+2x)^4;$
- 7)  $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + e^{-x}(2t - t^2),$   
 $u(x, 0) = 0, \quad 0 \leq x \leq 1;$   
 $u(0, t) = t^2, \quad u(1, t) = t^2 e^{-1} \quad 0 \leq t \leq 0,02;$   
 $u_T(x, t) = t^2 e^{-x}$
- 8)  $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + 2 \left[ t(x+0,5)^2 - (t^2 + 2) \right],$   
 $u(x, 0) = 2(x+0,5)^2 \quad 0 \leq x \leq 1;$   
 $u(0, t) = 0,25(t^2 + 2), \quad u(1, t) = 2,25(t^2 + 2),$   
 $0 \leq t \leq 0,02;$   
 $u_T(x, t) = (t^2 + 2)(x+0,5)^2$
- 9)  $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} - (2t+1)e^{x-t} \quad 0 \leq x \leq 1; \quad 0 \leq t \leq 0,02;$   
 $u(x, 0) = e^x \quad 0 \leq x \leq 1;$   
 $u(0, t) = (t+1)e^{-t} \quad u(1, t) = (t+1)e^{1-t} \quad 0 \leq t \leq 0,02;$   
 $u_T(x, t) = (t+1)e^{x-t}$
- 10)  $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + 0,4t(x^2 - t); \quad 0 \leq x \leq 1; \quad 0 \leq t \leq 0,02;$   
 $u(x, 0) = 0, \quad 0 \leq x \leq 1;$   
 $u(0, t) = 0, \quad u(1, t) = 0,2t^2, \quad 0 \leq t \leq 0,02;$   
 $u_T(x, t) = 0,2x^2t^2;$

- 11)  $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} - 4(t-x+1)e^{t-x} \quad 0 \leq x \leq 1; \quad 0 \leq t \leq 0,02;$   
 $u(x, 0) = 2x^2 e^{-x} \quad 0 \leq x \leq 1;$   
 $u(0, t) = 2t^2 e^t \quad u(1, t) = 2(t-1)^2 e^{t-1} \quad 0 \leq t \leq 0,02;$   
 $u_T(x, t) = 2(t-x)^2 e^{t-x}$
- 12)  $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + (1+x)^2 - 2(t+2); \quad 0 \leq x \leq 1; \quad 0 \leq t \leq 0,02;$   
 $u(x, 0) = 2(x+1)^2 \quad 0 \leq x \leq 1;$   
 $u(0, t) = (t+2), \quad u(1, t) = 4(t+2), \quad 0 \leq t \leq 0,02;$   
 $u_T(x, t) = (t+2)(1+x)^2$
- 13).  $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + 2(t+2)(1-2x) \left[ (1-2x)^2 - 24(t+2) \right];$   
 $0 \leq x \leq 1; \quad 0 \leq t \leq 0,02;$   
 $u(x, 0) = 4(1-2x)^3 \quad 0 \leq x \leq 1;$   
 $u(0, t) = (t+2)^2 \quad u(1, t) = -(t+2)^2 \quad 0 \leq t \leq 0,02;$   
 $u_T(x, t) = (t+2)^2 (1-2x)^3$
- 14)  $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + 3t^2 - 4; \quad 0 \leq x \leq 1; \quad 0 \leq t \leq 0,02;$   
 $u(x, 0) = 2x^2, \quad 0 \leq x \leq 1;$   
 $u(0, t) = t^3 \quad u(1, t) = 2+t^3, \quad 0 \leq t \leq 0,02;$   
 $u_T(x, t) = 2x^2 + t^3;$
- 15)  $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + 3t^2 e^{-3x} (1-3t); \quad 0 \leq x \leq 1; \quad 0 \leq t \leq 0,02;$   
 $u(x, 0) = 0, \quad 0 \leq x \leq 1;$   
 $u(0, t) = t^3 \quad u(1, t) = t^3 e^{-3}, \quad 0 \leq t \leq 0,02;$   
 $u_T(x, t) = t^3 e^{-3x}$

13.3. Методом сіток в області  $G = \{0 \leq x \leq 1; 0 \leq t \leq 1\}$  знайти розв'язок змішаної задачі для рівняння коливань струни на сітці  $(x_m, t_n)$ , де  $x_m = 0,1m$ ;  $t_n = 0,1n$ ;  $m, n = 0, 1, \dots, 10$ .

$$1) \quad \frac{\partial^2 u(x, t)}{\partial t^2} = \frac{\partial^2 u(x, t)}{\partial x^2} - \frac{1,5}{(x + 0,5t + 1)^3},$$

$$u(0, t) = \frac{1}{0,5t + 1}, \quad u(1, t) = \frac{1}{0,5t + 2}, \quad 0 \leq t \leq 1;$$

$$u(x, 0) = 0, \quad \frac{\partial u(x, 0)}{\partial t} = -\frac{0,5}{(1+x)^2}, \quad 0 \leq x \leq 1;$$

$$u_t(x, t) = \frac{1}{x + 0,5t + 1};$$

$$2) \quad \frac{\partial^2 u(x, t)}{\partial t^2} = \frac{\partial^2 u(x, t)}{\partial x^2},$$

$$u(0, t) = \frac{1}{t + 0,6}, \quad u(1, t) = \frac{1}{t + 1,6}, \quad 0 \leq t \leq 1;$$

$$u(x, 0) = \frac{1}{x + 0,6}, \quad \frac{\partial u}{\partial x} = -\frac{1}{(x + 0,6)^2}, \quad 0 \leq x \leq 1;$$

$$u_t(x, t) = \frac{1}{x + t + 0,6};$$

$$3) \quad \frac{\partial^2 u(x, t)}{\partial t^2} = \frac{\partial^2 u(x, t)}{\partial x^2} + \frac{0,4}{[0,2(x+t) + 2]^2}$$

$$u(0, t) = 0, \quad u(1, t) = \frac{1}{0,2(x+t) + 2}, \quad 0 \leq t \leq 1;$$

$$u(x, 0) = \frac{x}{2 + 0,2x}, \quad \frac{\partial u(x, 0)}{\partial t} = -\frac{0,2x}{(0,2x + 2)^2},$$

$$0 \leq x \leq 1;$$

$$u_t(x, t) = \frac{x}{0,2(x+t) + 2};$$

$$4) \quad \frac{\partial^2 u(x, t)}{\partial t^2} = \frac{\partial^2 u(x, t)}{\partial x^2},$$



$$u(0, t) = e^{0,5t} \quad u(1, t) = e^{0,5(1+t)}, \quad 0 \leq t \leq 1;$$

$$u(x, 0) = e^{0,5x}, \quad \frac{\partial u(x, 0)}{\partial t} = 0,5e^{0,5x} \quad 0 \leq x \leq 1;$$

$$u_T(x, t) = e^{0,5(x+t)};$$

$$5) \quad \frac{\partial^2 u(x, t)}{\partial t^2} = \frac{\partial^2 u(x, t)}{\partial x^2} - \frac{240}{(5x+t+2)^3},$$

$$u(0, t) = \frac{5}{t+2}, \quad u(1, t) = \frac{5}{t+7}, \quad 0 \leq t \leq 1;$$

$$u(x, 0) = \frac{5}{5x+2}, \quad \frac{\partial u(x, 0)}{\partial t} = \frac{-5}{(5x+2)^2}, \quad 0 \leq x \leq 1;$$

$$u_T(x, t) = \frac{5}{5x+t+2};$$

$$6) \quad \frac{\partial^2 u(x, t)}{\partial t^2} = \frac{\partial^2 u(x, t)}{\partial x^2} - \frac{0,14[0,7(x+0,7)-t]}{[x+0,7(t+1)]^3},$$

$$u(0, t) = \frac{t}{t+1}, \quad u(1, t) = \frac{0,7t}{1+0,7(t+1)}, \quad 0 \leq t \leq 1;$$

$$u(x, 0) = 0, \quad \frac{\partial u(x, 0)}{\partial t} = \frac{0,49x}{(x+0,7)^2}, \quad 0 \leq x \leq 1;$$

$$u_T(x, t) = \frac{0,7t}{x+0,7(t+1)};$$

$$7) \quad \frac{\partial^2 u(x, t)}{\partial t^2} = \frac{\partial^2 u(x, t)}{\partial x^2} - \frac{x+0,5t+2}{[0,5x+t+1]^3},$$

$$u(0, t) = \frac{t}{t+1}, \quad u(1, t) = \frac{t}{t+1,5}, \quad 0 \leq t \leq 1;$$

$$u(x, 0) = 0, \quad \frac{\partial u(x, 0)}{\partial t} = 1, \quad 0 \leq x \leq 1;$$

$$u_T(x, t) = \frac{t}{0,5x+t+1};$$

$$8) \quad \frac{\partial^2 u(x, t)}{\partial t^2} = \frac{\partial^2 u(x, t)}{\partial x^2} + 2(4+3x)e^{2(x+t)},$$

$$\begin{aligned}
u(0, t) &= 0, \quad u(1, t) = e^{2t} \quad 0 \leq t \leq 1; \\
u(x, 0) &= x e^{2x}, \quad \frac{\partial u(x, 0)}{\partial t} = 2x e^{2x} \quad 0 \leq x \leq 1; \\
u_T(x, t) &= x e^{2(x+t)}
\end{aligned}$$

$$\begin{aligned}
9) \quad \frac{\partial^2 u(x, t)}{\partial t^2} &= \frac{\partial^2 u(x, t)}{\partial x^2} + 4(x+t), \\
u(0, t) &= t^3 \quad u(1, t) = t(1+t)^2, \quad 0 \leq t \leq 1; \\
u(x, 0) &= 0, \quad \frac{\partial u(x, 0)}{\partial t} = x^2 \quad 0 \leq x \leq 1; \\
u_T(x, t) &= t(x+t)^2;
\end{aligned}$$

$$\begin{aligned}
10) \quad \frac{\partial^2 u(x, t)}{\partial t^2} &= \frac{\partial^2 u(x, t)}{\partial x^2} + \left(\frac{x}{2} + \frac{t}{5}\right) \left[6 \frac{x^2}{25} - 3\left(x + \frac{t}{5}\right)\right], \\
u(0, t) &= 0, \quad u(1, t) = \left(\frac{1}{2} + \frac{t}{5}\right)^3 \quad 0 \leq t \leq 1; \\
u(x, 0) &= \frac{x^4}{8}, \quad \frac{\partial u(x, 0)}{\partial t} = \frac{3x^3}{20}, \quad 0 \leq x \leq 1; \\
u_T(x, t) &= x \left(\frac{x}{2} + \frac{t}{5}\right)^3
\end{aligned}$$

$$\begin{aligned}
11) \quad \frac{\partial^2 u(x, t)}{\partial t^2} &= \frac{\partial^2 u(x, t)}{\partial x^2} + (2-3t)e^{-(x+t)} \\
u(0, t) &= t^2 \quad u(1, t) = t^2 e^{-(1+t)}, \quad 0 \leq t \leq 1; \\
u(x, 0) &= 0, \quad \frac{\partial u(x, 0)}{\partial t} = 0, \quad 0 \leq x \leq 1; \\
u_T(x, t) &= t^2 e^{-(x+t)}
\end{aligned}$$

$$12) \quad \frac{\partial^2 u(x, t)}{\partial t^2} = \frac{\partial^2 u(x, t)}{\partial x^2} + e^{-t} \left[ \left(x + \frac{t}{2}\right)^2 - t \left(x + \frac{t}{2}\right) - 2 \right],$$

$$u(0, t) = \frac{t^2}{4} e^{-t} \quad u(1, t) = \left(1 + \frac{t}{2}\right)^2 e^{-t} \quad 0 \leq t \leq 1;$$

$$u(x, 0) = x^2 \quad \frac{\partial u(x, 0)}{\partial t} = -x^2 \quad 0 \leq x \leq 1;$$

$$u_t(x, t) = \left(x + \frac{t}{2}\right)^2 e^{-t}$$

$$13) \quad \frac{\partial^2 u(x, t)}{\partial t^2} = \frac{\partial^2 u(x, t)}{\partial x^2} + (2 - 4x + 2x^2 - 2t^2) e^{-t}$$

$$u(0, t) = t^2 \quad u(1, t) = (2 + t^2) e^{-1}, \quad 0 \leq t \leq 1;$$

$$u(x, 0) = 2x^2 e^{-x} \quad \frac{\partial u(x, 0)}{\partial t} = 0, \quad 0 \leq x \leq 1;$$

$$u_t(x, t) = (2x^2 + t^2)^2 e^{-x}.$$

$$14) \quad \frac{\partial^2 u(x, t)}{\partial t^2} = \frac{\partial^2 u(x, t)}{\partial x^2} +$$

$$+ 3(x + te^t) \{ e^t [2(1 + t^2)e^t + (x + te^t)(2 + t)] - 2 \},$$

$$u(0, t) = (te^t)^3, \quad u(1, t) = (1 + te^t)^3 \quad 0 \leq t \leq 1;$$

$$u(x, 0) = x^3 \quad \frac{\partial u(x, 0)}{\partial t} = 3x^2 \quad 0 \leq x \leq 1;$$

$$u_t(x, t) = (x + e^t t)^3 (x + te^t)^3$$

$$15) \quad \frac{\partial^2 u(x, t)}{\partial t^2} = \frac{\partial^2 u(x, t)}{\partial x^2} + 2t [3(x + 1)^2 - t^2]$$

$$u(0, t) = t^3 \quad u(1, t) = 4t^3 \quad 0 \leq t \leq 1;$$

$$u(x, 0) = 0, \quad \frac{\partial u(x, 0)}{\partial t} = 0, \quad 0 \leq x \leq 1;$$

$$u_t(x, t) = (x + 1)^2 t^3$$

## Розділ 14

# МЕТОДИ НАБЛИЖЕНОГО РОЗВ'ЯЗУВАННЯ ІНТЕГРАЛЬНИХ РІВНЯНЬ

### 14.1. Формулювання задачі

Інтегральним називають рівняння, у якому шукана функція  $u(x)$  серед іншого є під знаком інтеграла. Одновимірне нелінійне інтегральне рівняння має вигляд

$$\int_a^b K(x, \xi, u(\xi)) d\xi = F(x, u(x)), \quad a \leq x \leq b, \quad (14.1)$$

де ядро  $K(x, \xi, u(\xi))$  і права частина  $F(x, u(x))$  – задані функції.

До інтегрального рівняння (ІР) приводять багато фізичних задач. ІР в багатьох випадках зручніші, ніж диференціальні. По-перше, ІР містять повне формулювання задачі. Наприклад, ІР

$$u(x) = u_0 + \int_{x_0}^x f(x, u(\xi)) d\xi \quad \text{еквівалентне задачі Коші для}$$

диференціального рівняння  $\frac{du(x)}{dx} = f(x, u), \quad u(x_0) = u_0$ . Легко

побачити, що для ІР не треба задавати ніяких додаткових умов, початкових чи граничних. По-друге, в ІР перехід від однієї змінної до багатьох є природним. Наприклад, багатовимірним аналогом (14.1) є рівняння

$$\int_G K(\bar{x}, \xi, u(\xi)) d\xi = F(\bar{x}, u(\bar{x})), \quad \bar{x} = (x_1, x_2, \dots, x_p)^T \in G(\bar{x}),$$

яке відрізняється від (14.1) тільки тим, що інтегрування відбувається за багатовимірною областю  $G \subset R^p$ . Отже, теоретичне обґрунтування формулювань і методів розв'язування

одновимірних задач безпосередньо узагальнюють на випадок багатьох вимірів.

Розглянемо ІР вигляду

$$u(x) - \lambda \int_a^b K(x, \xi) u(\xi) d\xi = f(x), \quad a \leq x \leq b, \quad (14.2)$$

яке переважно називають рівнянням Фредгольма другого роду. Ядро  $K(x, \xi)$  визначене на квадраті  $a \leq x, \xi \leq b$ . Якщо  $K(x, \xi)$  визначене лише на трикутнику  $a \leq \xi \leq x \leq b$ , то (14.2) трактуємо як рівняння Вольтерра другого роду

$$u(x) - \lambda \int_a^x K(x, \xi) u(\xi) d\xi = f(x), \quad a \leq x \leq b. \quad (14.3)$$

Якщо в рівнянні (14.2) чи (14.3) відкинути член  $u(x)$ , а залишити тільки  $u(\xi)$  під знаком інтеграла, то отримаємо рівняння першого роду.

## 14.2. Метод послідовних наближень розв'язування одновимірною рівняння другого роду

Розв'язок ІР (14.2) шукаємо у вигляді степеневому ряду

$$u(x) = \sum_{k=0}^{\infty} \lambda^k \varphi_k(x). \quad (14.4)$$

Підставимо (14.4) у рівняння (14.2) і прирівняємо коефіцієнти при однакових степенях  $\lambda$ , отримаємо

$$\begin{aligned} \varphi_0(x) &= f(x); \\ \varphi_1(x) &= \int_a^b K(x, \xi) \varphi_0(\xi) d\xi; \\ \varphi_2(x) &= \int_a^b K(x, \xi) \varphi_1(\xi) d\xi; \end{aligned} \quad (14.5)$$

$$\varphi_n(x) = \int_a^b K(x, \xi) \varphi_{n-1}(\xi) d\xi.$$

Інтеграли в (14.5) обчислюємо або аналітично, або за квадратурними формулами.

Якщо функція  $K(x, \xi)$  в  $[a, b] \times [a, b]$  обмежена константою  $M$   $|K(x, \xi)| \leq M$ , то рівномірна збіжність ряду (14.4) буде забезпечена, якщо  $|\lambda| < \frac{1}{M(b-a)}$ .

Формули (14.5) можна переписати у вигляді

$$\varphi_k(x) = \int_a^b K_k(x, \xi) f(\xi) d\xi, \quad (14.6)$$

$$\text{де } K_k(x, \xi) = \int_a^b K(x, t) K_{k-1}(t, \xi) dt, \quad K_1(x, \xi) = K(x, \xi).$$

Тоді розв'язок (14.4) можна записати як

$$u(x) = f(x) + \lambda \sum_{k=1}^{\infty} \lambda^{k-1} \int_a^b K_k(x, \xi) f(\xi) d\xi, \quad (14.7)$$

або

$$u(x) = f(x) + \lambda \int_a^b R(x, \xi, \lambda) f(\xi) d\xi, \quad (14.8)$$

$$\text{де } R(x, \xi, \lambda) = \sum_{k=1}^{\infty} \lambda^{k-1} K_k(x, \xi). \quad (14.9)$$

Коефіцієнти  $K_k(x, \xi)$  називають ітерованими ядрами, а функцію  $R(x, \xi, \lambda)$  – резольвентою IP, яка для малих  $|\lambda|$  визначена степеневим рядом (14.9).

**П р и к л а д 14.1.** Методом послідовних наближень знайти розв'язок рівняння

$$u(x) - \lambda \int_0^{\infty} e^{-(x+\xi)} u(\xi) d\xi = x. \quad (14.10)$$

**Розв'язування.** Застосуємо (14.5), отримаємо

$$\varphi_0(x) = x;$$

$$\varphi_1(x) = \int_0^{\infty} e^{-x} e^{-\xi} \xi d\xi = e^{-x} \int_0^{\infty} \xi e^{-\xi} d\xi = e^{-x}; \quad (14.11)$$

$$\varphi_2(x) = \int_0^{\infty} e^{-x} e^{-\xi} e^{-\xi} d\xi = e^{-x} \int_0^{\infty} e^{-2\xi} d\xi = \frac{1}{2} e^{-x};$$

$$\varphi_3(x) = \frac{1}{2} \int_0^{\infty} e^{-x} e^{-\xi} e^{-\xi} d\xi = \frac{1}{2} e^{-x} \int_0^{\infty} e^{-2\xi} d\xi = \frac{1}{4} e^{-x}$$

Тоді розв'язок рівняння (14.10), згідно з (14.4), матиме вигляд

$$u(x) = x + \lambda e^{-x} + \frac{1}{2} \lambda^2 e^{-x} + \frac{1}{4} \lambda^3 e^{-x} + \dots = x + e^{-x} \left( \lambda + \frac{1}{2} \lambda^2 + \frac{1}{4} \lambda^3 + \dots \right).$$

У нашому прикладі вдається знайти точний розв'язок

$$u(x) = x + \frac{2\lambda}{2-\lambda} e^{-x}$$

Послідовні наближення збігаються тільки при  $|\lambda| < 2$ .

**Приклад 14.2.** Методом послідовних наближень розв'язати рівняння Вольтерра

$$u(x) = 1 - \int_0^x (x - \xi) u(\xi) d\xi.$$

**Розв'язування.** Нехай  $\varphi_0(x) = x$ ; Тоді

$$\varphi_1(x) = \int_0^x (x - \xi) d\xi = \frac{x^2}{2};$$

$$\varphi_2(x) = \int_0^x (x - \xi) \left(1 - \frac{\xi^2}{2}\right) d\xi = \frac{x^2}{2} - \frac{x^4}{4!}.$$

Для  $n$ -го наближення отримаємо

$$\varphi_n(x) = \frac{x^2}{2} - \frac{x^4}{4!} + \frac{x^6}{6!} + \dots + (-1)^{n+1} \frac{x^{2n}}{(2n)!} = \sum_{k=1}^n (-1)^{k+1} \frac{x^{2k}}{(2k)!}.$$

Звідси, згідно з (14.7),

$$u(x) = 1 - \sum_{k=1}^n (-1)^{k+1} \frac{x^{2k}}{(2k)!} = \sum_{k=0}^n (-1)^k \frac{x^{2k}}{(2k)!} \approx \cos x.$$

### **14.3. Наближене розв'язування інтегрального рівняння Фредгольма методом заміни інтеграла скінченною сумою (метод механічних квадратур)**

Якщо ядро ІР досить складне, то аналітично обчислювати інтеграли (14.5) в методі послідовних наближень важко. Тому для наближеного розв'язування рівняння (14.2) можна застосувати інші методи, не пов'язані з перетворенням ядер. Серед таких методів найвідоміший метод квадратур. Основна його ідея – це заміна інтеграла в (14.2) скінченною сумою. Для цього використовують квадратурні формули

$$\int_a^b F(x) dx = \sum_{k=1}^n A_k F(x_k) + R_n(F), \quad (14.12)$$

де  $x_k$  – задані наперед вузли;  $A_k$  – коефіцієнти, які не залежать від підінтегральної функції;  $R_n(F)$  – залишковий член квадратурної формули.

Правило (14.12) застосовують для наближеного запису (14.2) у такий спосіб:

$$u(x) - \lambda \left\{ \sum_{k=1}^n A_k K(x, x_k) u(x_k) + R_n(Ku) \right\} = f(x). \quad (14.13)$$

Якщо величина  $R_n(Ku)$  мала, то нею можна знехтувати, а за аналітичний вираз наближеного розв'язку  $U(x)$  ІР (14.2) прийняти функцію

$$U(x) = f(x) + \lambda \sum_{k=1}^n A_k K(x, x_k) U(x_k). \quad (14.14)$$



Права частина формули (14.14) містить невідомі величини  $U(x_k)$ . Їхні наближені значення  $U_k$  можна знайти, приймаючи в (14.14)  $x = x_i$  ( $i = \overline{1, n}$ ),  $x_i \in [a, b]$ , що дає

$$U_i - \lambda \sum_{k=1}^n A_k K(x_i, x_k) U_k = f_i, \text{ де } f_i = f(x_i). \quad (14.15)$$

Легко побачити, що це система лінійних алгебричних рівнянь  $n$ -го порядку відносно невідомих наближених значень розв'язку  $U_1, U_2, \dots, U_n$ .

Визначник системи має вигляд

$$\Delta = \begin{vmatrix} 1 - \lambda A_1 K_{11} & -\lambda A_2 K_{12} & \dots & -\lambda A_n K_{1n} \\ -\lambda A_1 K_{21} & 1 - \lambda A_2 K_{22} & \dots & -\lambda A_n K_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -\lambda A_1 K_{n1} & -\lambda A_2 K_{n2} & \dots & 1 - \lambda A_n K_{nn} \end{vmatrix}. \quad (14.16)$$

де  $K_{ik} = K(x_i, x_k)$ .

СПАР (14.15) має єдиний розв'язок, якщо  $\lambda \neq \lambda_i^{(n)}$  де  $\lambda_i^{(n)}$  – власні значення матриці системи. Отже, описаний алгоритм добре зумовлений, якщо параметр  $\lambda$  не лежить у малому околі одного з власних значень  $\lambda_i$  ядра.

Якщо  $\lambda \approx \lambda_i$ , то система (14.15) стає погано зумовленою. За деяких  $n$  можливий збій алгоритму, якщо випадково значення  $\lambda_i^{(n)}$  близько підходить до одного з власних значень ядра. Тоді наближений розв'язок  $U_n(x)$  може сильно відрізнятись від точного  $u(x)$ . Зазвичай, нам невідомі власні значення ядра. Тому для виявлення й уникнення останнього випадку всі розрахунки треба виконувати в разі послідовного збільшення кількості вузлів квадратурної формули. Якщо зі збільшенням  $n$  розв'язок  $\{U_i\}_{i=1}^n$  відтворює деяку граничну функцію  $U(x)$ , то ця функція і є шуканим наближеним розв'язком (14.2). Якщо за якогось  $n$  розрахунок випадає із загальної закономірності, тоді є випадковим збіг  $\lambda \approx \lambda_i^{(n)}$ . Отже, розв'язавши (14.15), знайдемо значення  $U_1, U_2, \dots, U_n$ , і наближений розв'язок рівняння (14.2) в аналітичному вигляді запишемо так:

$$U_n(x) = f(x) + \lambda \sum_{k=1}^n A_k K(x, x_k) U_k.$$

Більшість задач розв'язують, використовуючи порівняно невелику кількість вузлів  $n$ . Тому для отримання доброго результату доцільно вибирати квадратурні формули високого порядку точності, уважаючи, що  $K(x, \xi)$  і  $f(x)$  достатньо гладкі функції.

Зазвичай, найліпші результати дають квадратурні формули Гаусса або Гаусса–Кристоффеля; для кількості вузлів  $k$  їхній порядок точності  $p = 2k$ . Можна також використовувати формулу трапецій, послідовно згущуючи сітку вдвічі від  $n_1 = 2$  до  $n_k = 2^k$  і уточнюючи розв'язок за принципом Рунге. Це теж дає результат з порядком точності  $p = 2k$ , однак потребує використання суттєво більшої кількості вузлів, ніж у формулах Гаусса.

Досить часто ядро  $K(x, \xi)$  або права частина  $f(x)$  недостатньо гладкі й навіть мають розриви. Найтипівішим є розрив ядра або його похідних за  $x = \xi$ ; трапляються особливості й на інших лініях у площині  $(x, \xi)$ . У цих випадках використовувати квадратурні формули Гаусса недоцільно. Зручніше вибирати вузли  $x_n$  так, щоб особливі лінії перетинали лінії сітки  $x = x_n$  лише у вузлах  $\xi = x_m \neq x_n$ . Як (14.12) можна вибирати узагальнену формулу трапецій, причому в інтервалах, які прилягають до особливих ліній, використовують відповідні односторонні границі функцій.

Розглянемо деякі конкретні квадратурні формули.

А. Якщо вибрана узагальнена квадратурна формула трапецій, то маємо

$$A_1 = A_n = \frac{1}{2}h; \quad A_k = h; \quad k = \overline{2, n-1};$$

$$x_k = a + (k-1)h; \quad k = \overline{1, n}; \quad h = \frac{b-a}{n-1};$$

$$R_n(Ku) = -\frac{(b-a)^3}{12(n-1)^2} \left[ \frac{\partial^2}{\partial \xi^2} (Ku) \right]_{\xi=\eta} \quad \eta \in [a, b].$$

Б. Для узагальненої квадратурної формули Сімпсона, прийнявши  $n = 2m+1$ , одержимо

$$A_1 = A_{2m+1} = \frac{1}{3}h; \quad A_2 = A_4 = \dots = A_{2m} = \frac{4}{3}h;$$

$$A_3 = A_5 = \dots = A_{2m-1} = \frac{2}{3}h;$$

$$x_k = a + (k-1)h; \quad k = \overline{1, 2m+1}; \quad h = \frac{b-a}{2m};$$

$$R_\pi(Ku) = -\frac{(b-a)^5}{90(2m)^4} \left[ \frac{\partial^5}{\partial \xi^5} (Ku) \right]_{\xi=\eta}, \quad \eta \in [a, b].$$

В. Шукану функцію в (14.12) апроксимуємо інтерполяційним багаточленом Лагранжа

$$u(x) = \sum_{k=1}^n l_k(x) u(x_k) + r(x); \quad (14.17)$$

де  $l_k(x) = \frac{\omega(x)}{(x-x_k)\omega'(x_k)}$ ;  $\omega(x) = (x-x_1)\dots(x-x_n)$ ;  $r(x)$  – залишок інтерполювання.

Підставимо вираз (14.17) у рівняння (14.2). Отримаємо

$$u(x) = f(x) + \lambda \sum_{k=1}^n L_k(x) u(x_k) + \lambda \rho(x), \quad (14.18)$$

де

$$L_k(x) = \int_a^b K(x, \xi) l_k(\xi) d\xi;$$

$$\rho(x) = \int_a^b K(x, \xi) r(\xi) d\xi.$$

Відкинемо в (14.18) залишковий член  $\lambda \rho(x)$  і в отриманій наближеній рівності приймемо  $x = x_1, x_2, \dots, x_n$ :

$$u_i = f_i + \lambda \sum_{k=1}^n L_k(x_i) u_k, \quad i = \overline{1, n}. \quad (14.19)$$

Це СЛАР  $n$ -го порядку з невідомими  $u_1, u_2, \dots, u_n$ . Після її розв'язування знайдено наближені значення розв'язку  $u(x)$  в точках  $x_i, i = \overline{1, n}$ .

Наближене значення розв'язку  $U(x)$  на  $[a, b]$  знаходимо за формулою інтерполювання через ядро  $K(x, \xi)$ :

$$U(x) = f(x) + \sum_{k=1}^n I_k(x) u(x_k). \quad (14.20)$$

**П р и к л а д 14.3.** Методом механічних квадратур знайти розв'язок інтегрального рівняння

$$u(x) + \int_0^1 x(e^{x\xi} - 1)u(\xi)d\xi = e^x - x.$$

*Розв'язування.* Застосуємо квадратурну формулу Сімпсона. Одержимо

$$\int_0^1 f(x)dx \approx \frac{1}{6}[f(0) + 4f(0,5) + f(1)].$$

Тоді для відшукування наближеного розв'язку в точках  $x = 0; 0,5; 1$  отримаємо систему

$$\begin{cases} u_1 = 1; \\ \frac{1}{3}(e^{0,25} - 1)u_2 + \frac{1}{12}(e^{0,5} - 1)u_3 = e^{0,5} - 0,5; \\ \frac{2}{3}(e^{0,5} - 1)u_2 + \frac{1}{6}(e - 1)u_3 = e - 1, \end{cases}$$

або

$$\begin{cases} u_1 = 1; \\ 1,0947u_2 + 0,0541u_3 = 1,1487; \\ 0,4325u_2 + 1,2864u_3 = 1,7183. \end{cases}$$

Розв'язки її

$$u_1 = 1; \quad u_2 = 0,9999; \quad u_3 = 0,9996.$$

Точний розв'язок  $u(x) \equiv 1$ . Як бачимо, результат достатньо добрий.

Для комп'ютерної реалізації можна використати узагальнену квадратурну формулу Сімпсона і застосувати принцип Рунге, щораз порівнюючи знайдені розв'язки в точках  $x_i$ .

#### 14.4. Метод заміни ядра на вироджене

Ядро IP (14.2) називають виродженим, якщо його можна записати у вигляді

$$K(x, \xi) = \sum_{i=1}^n A_i(x) B_i(\xi), \quad (14.21)$$

де  $\{A_i(x)\}_{i=1}^n, \{B_i(\xi)\}_{i=1}^n$  – відомі системи лінійно незалежних функцій.

У цьому випадку розв'язок  $u(x)$ , якщо він існує, можна знайти в явному вигляді

$$u(x) = f(x) + \lambda \sum_{i=1}^n c_i A_i(x), \quad (14.22)$$

де  $c_i = \int_a^b B_i(\xi) u(\xi) d\xi$  – деякі числові коефіцієнти.

Справді, підставивши (14.22) у попередню формулу, для знаходження  $c_i$  отримаємо СЛАР

$$c_i - \lambda \sum_{j=1}^n c_j \alpha_{ij} = f_i, \quad i = \overline{1, n}, \quad (14.23)$$

$$\text{де } \alpha_{ij} = \int_a^b A_j(\xi) B_i(\xi) d\xi; \quad \beta_i = \int_a^b f(\xi) B_i(\xi) d\xi. \quad (14.24)$$

Якщо визначник системи (14.23) відмінний від нуля, то система має єдиний розв'язок  $c_1, c_2, \dots, c_n$ , і розв'язок інтегрального рівняння  $u(x)$  буде знайдений явно. Якщо ж для заданого значення  $\lambda$  визначник дорівнює нулю, то  $\lambda$  буде власним значенням ядра  $K(x, \xi)$ . У цьому випадку, знаходячи лінійно незалежні розв'язки відповідної однорідної системи, ми в явному вигляді знайдемо всі лінійно незалежні між собою власні функції ядра  $K(x, \xi)$ , що відповідають заданому власному значенню  $\lambda$ .

Якщо  $c_i^*$  ( $i = \overline{1, n}$ ) – розв'язок системи (14.23), то шуканий розв'язок IP (14.2) можна також обчислити за формулою

$$u^*(x) = f(x) + \lambda \sum_{i=1}^n c_i^* A_i(x).$$

Така ситуація можлива, якщо вироджене ядро близьке до ядра  $K(x, \xi)$ . У цьому випадку чим ліпше сума  $\sum_{k=1}^n A_k(x) B_k(\xi)$  апроксимує  $K(x, \xi)$ , тим ближче  $u^*(x)$  до розв'язку IP (14.2). Це підтверджує така теорема.

**Теорема.** Якщо

$$y(x) - \lambda \int_a^b K(x, \xi) y(\xi) d\xi = f(x);$$

$$z(x) - \lambda \int_a^b H(x, \xi) z(\xi) d\xi = f_1(x),$$

– два IP,  $R(x, \xi, \lambda)$  – резольвента другого з цих рівнянь і існують такі сталі  $\delta, \varepsilon, M$ , що виконуються нерівності

$$\int_a^b |K(x, \xi) - H(x, \xi)| d\xi < \delta;$$

$$|f(x) - f_1(x)| < \varepsilon, \quad \forall x \in [a, b];$$

$$\int_a^b |R(x, \xi, \lambda)| d\xi < M;$$

$$|\lambda| \delta (1 + |\lambda| M) < 1,$$

то перше рівняння має єдиний розв'язок  $y(x)$ , і

$$|y(x) - z(x)| < \frac{N |\lambda| (1 + |\lambda| M)^2 \delta}{1 - |\lambda| \delta (1 + |\lambda| M)} + \varepsilon (1 + |\lambda| M),$$

де  $N = \max_{a \leq x \leq b} |f(x)|$ .

З наведеної теореми випливає таке: якщо побудувати досить близьке до ядра  $K(x, \xi)$  вироджене ядро  $H(x, \xi)$ , то, розв'язавши

інтегральне рівняння з виродженим ядром  $H(x, \xi)$ , одержимо розв'язок  $z(x)$ , який буде близьким до розв'язку  $y(x)$  інтегрального рівняння з ядром  $K(x, \xi)$  за тієї ж правої частини.

Існує декілька способів побудови для ядра  $H(x, \xi)$  виродженого і близького ядра до  $K(x, \xi)$ .

1. Якщо ядро  $K(x, \xi)$  має високий порядок гладкості, то розкладаємо його за формулою Тейлора. Можливі розклади за одним із аргументів  $x$  або  $\xi$ , а також розклад за двома аргументами одночасно. Наприклад, якщо ми розкладемо  $K(x, \xi)$  за першим аргументом, і за початкову точку візьмемо середину відрізка  $[a, b]$ , тобто приймемо  $c = 0,5(a + b)$ , тоді для виродженого ядра  $H(x, \xi)$  одержимо такий вираз

$$H(x, \xi) = K(c, \xi) + \frac{x-c}{1!} \frac{\partial K(c, \xi)}{\partial x} + \dots + \frac{(x-c)^n}{n!} \frac{\partial^n K(c, \xi)}{\partial x^n}.$$

Функції  $A_j(x)$  дорівнюватимуть 1,  $\frac{x-c}{1!}, \dots, \frac{(x-c)^n}{n!}$ , а функції

$$B_j(\xi) - \text{відповідно } K(c, \xi), \frac{\partial K(c, \xi)}{\partial x}, \dots, \frac{\partial^n K(c, \xi)}{\partial x^n}.$$

2. Для побудови  $H(x, \xi)$  можна використати відомі ортогональні розклади функцій. Наприклад, якщо ядро  $K(x, \xi)$  – достатньо гладка періодична функція за одним із аргументів  $x$  або  $\xi$  чи за обома аргументами, тоді можна скористатися відрізком однократного або подвійного ряду Фур'є. Припустимо, що відрізок інтегрування  $[a, b]$  є проміжок  $[-\pi, \pi]$  і  $K(x, \xi)$  є  $2\pi$  періодичною функцією за  $x$ . Тоді

$$H(x, \xi) = a_0(\xi) + \sum_{i=1}^n (a_i(\xi) \cos ix + b_i(\xi) \sin ix),$$

$$a_0(\xi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} K(x, \xi) dx,$$

$$a_i(\xi) = \frac{1}{\pi} \int_{-\pi}^{\pi} K(x, \xi) \cos ix dx,$$

$$b_i(\xi) = \frac{1}{\pi} \int_{-\pi}^{\pi} K(x, \xi) \sin ix dx.$$

Використовують також інші ортогональні розклади, такі як розклад за багаточленами Лежандра, Чебишова тощо.

3. Можна використовувати інтерполювання функції  $K(x, \xi)$  за одним або двома аргументами. Наприклад, для аргументу  $x$  на проміжку  $[a, b]$  виберемо  $n$  вузлів інтерполювання  $x_1, \dots, x_n$  і запишемо інтерполяційний багаточлен Лагранжа у вигляді

$$H(x, \xi) = \sum_{i=1}^n \frac{\omega(x)}{(x - x_i) \omega'(x_i)} K(x_i, \xi),$$

$$\omega(x) = (x - x_1) \dots (x - x_n).$$

**Приклад 14.4.** Знайти розв'язок інтегрального рівняння

$$u(x) - 0,5 \int_0^1 e^{\frac{x\xi}{25}} u(\xi) d\xi = \sin 0,1x \quad (14.25)$$

методом заміни ядра на вироджене.

**Розв'язування.** Замінімо ядро IP (14.25) на вироджене, розклавши  $K(x, \xi) = e^{\frac{x\xi}{25}}$  в ряд Тейлора:

$$K(x, \xi) = \sum_{0 \leq p+q \leq r} \frac{(x - x_0)^p}{p!} \frac{(\xi - \xi_0)^q}{q!} \frac{\partial^{p+q} K(x_0, \xi_0)}{\partial x^p \partial \xi^q}.$$

Для простоти приймемо  $r=1$  і за точку розкладу —

$$(x_0, \xi_0) = \left( \frac{1}{2}, \frac{1}{2} \right).$$

Отримаємо

$$K(x, \xi) = e^{\frac{1}{100}} + \frac{1}{25} \left( x - \frac{1}{2} \right) \xi e^{\frac{1}{100}} + \frac{1}{25} \left( \xi - \frac{1}{2} \right) x e^{\frac{1}{100}} + R_1(K),$$

де

$$|R_1(K)| \leq \frac{1}{2} \frac{1}{625} \approx 0,8 \cdot 10^{-3}$$



Відкинемо залишковий член розкладу в ряд і величини порядку  $R_1(x, \xi)$ , отримаємо формулу для виродженого ядра  $K(x, \xi)$ :

$$\begin{aligned} K(x, \xi) &= 1,001 + 1,001 \cdot \frac{1}{25} \left( x\xi - \frac{1}{2} \xi \right) + 1,001 \cdot \frac{1}{25} \left( x\xi - \frac{1}{2} x \right) = \\ &= \left( 1,001 - \frac{1}{50} \cdot 1,001 \xi \right) + x \left( -0,001 \cdot \frac{1}{50} + \frac{2,002}{25} \xi \right). \end{aligned} \quad (14.26)$$

Тут враховано, що  $e^{\frac{1}{100}} \approx 1,001$ . Ядро, згідно з (14.21), можна записати у вигляді

$$K(x, \xi) = \sum_{i=1}^2 A_i(x) B_i(\xi),$$

приймавши

$$\begin{aligned} A_1(x) &= 1; & A_2(x) &= x; \\ B_1(\xi) &= 1,001 - \frac{1,001}{50} \xi; & B_2(\xi) &= -\frac{1,001}{50} + \frac{2,002}{25} \xi. \end{aligned}$$

Розкладемо праву частину IP (14.25) в ряд Тейлора, відкинувши члени порядку  $10^{-7}$ , отримаємо

$$f(x) \approx 0,1x - 0,000167x^3$$

Розв'язок рівняння

$$u(x) - 0,5 \int_0^1 K(x, \xi) u(\xi) d\xi = 0,1x - 0,000167x^3,$$

де  $K(x, \xi)$  визначають за формулою (14.26), знайдемо у вигляді

$$u(x) = f(x) + c_1 A_1(x) + c_2 A_2(x) = 0,1x - 0,000167x^3 + c_1 + c_2 x.$$

При  $n = 2$  отримаємо СЛАР (14.23), у якій

$$\begin{aligned} f_1 &= \int_0^1 f(x) B_1(x) dx = \\ &= \int_0^1 (0,1x - 0,000167x^3) \cdot \left( 1,001 - \frac{1,001}{50} x \right) dx = 0,05076; \end{aligned}$$

$$\begin{aligned}
f_2 &= \int_0^1 f(x) B_2(x) dx = \\
&= \int_0^1 (0,1x - 0,000167x^3) \cdot \left(-\frac{1,001}{50} + \frac{2,002}{25}x\right) dx = -0,00166; \\
\alpha_{11} &= \int_0^1 A_1(x) B_1(x) dx = \int_0^1 \left(0,001 - \frac{1,001}{50}x\right) dx = 0,99099; \\
\alpha_{12} &= \int_0^1 A_2(x) B_1(x) dx = \int_0^1 x \left(0,001 - \frac{1,001}{50}x\right) dx = 0,49383; \\
\alpha_{21} &= \int_0^1 A_1(x) B_2(x) dx = \int_0^1 \left(-\frac{1,001}{50} + \frac{2,002}{25}x\right) dx = 0,02002; \\
\alpha_{22} &= \int_0^1 A_2(x) B_2(x) dx = \int_0^1 x \left(-\frac{1,001}{50} + \frac{2,002}{25}x\right) dx = 0,01668.
\end{aligned}$$

Визначник системи відмінний від нуля, тому при  $n = 2$  СПАР має єдиний розв'язок  $c_1 = 0,049$ ,  $c_2 = -0,01$ .

Отже, для наближеного розв'язку рівняння (14.25) отримаємо формулу

$$u(x) \approx 0,049 + 0,099x + 0,000167x^3$$

## Завдання до розділу 14

14.1. Розв'язати інтегральні рівняння методом послідовних наближень.

$$1) \quad u(x) = 1 - \int_0^x u(t) dt, \quad u_0(x) = 0;$$

$$2) \quad u(x) = \frac{x^2}{2} + x - \int_0^x u(t) dt,$$

$$a) \quad u_0(x) = 1; \quad б) \quad u_0(x) = \frac{x^2}{2} + x;$$

$$3) \quad u(x) = 1 - x^2 - \int_0^x x u(t) dt,$$

$$a) \quad u_0(x) = 1 - x^2; \quad б) \quad u_0(x) = 1;$$

$$4) \quad u(x) = 1 + \int_0^x x u(t) dt, \quad u_0(x) = 0;$$

$$5) \quad u(x) = 1 + \int_0^x t u(t) dt, \quad u_0(x) = 1;$$

$$6) \quad u(x) = x + \int_0^x x t u(t) dt, \quad u_0(x) = x;$$

$$7) \quad u(x) = \sin x + 2 \int_0^x e^{x-t} u(t) dt, \quad u_0(x) = \sin x;$$

$$8) \quad u(x) = \operatorname{ch} x - \int_0^x \frac{\operatorname{ch} x}{\operatorname{ch} t} u(t) dt, \quad u_0(x) = \operatorname{ch} x;$$

$$9) \quad u(x) = \frac{1}{1+x^2} + \int_0^x \frac{1+t^2}{1+x^2} u(t) dt, \quad u_0(x) = \frac{1}{1+x^2};$$

$$10) \quad u(x) = 2x + \int_0^1 x t u(t) dt, \quad u_0(x) = 2x;$$

$$11) \quad u(x) = 1 - \frac{1}{\pi} \int_0^\pi (\cos^2 t) u(t) dt, \quad u_0(x) = 1;$$

$$12) u(x) = \frac{1}{2}(1-x) + \pi \int_0^1 (1-x)(\sin 2\pi t)u(t)dt, \quad u_0(x) = \frac{1}{2}(1-x);$$

$$13) u(x) = 2 \sin x + \frac{1}{2\pi} \int_0^\pi (\sin x)tu(t)dt, \quad u_0(x) = 2 \sin x;$$

$$14) u(x) = \cos x - \frac{1}{2\pi} \int_0^\pi (\cos(x+t) + \cos(x-t))u(t)dt, \quad u_0(x) = \cos x;$$

$$15) u(x) = 1 - \frac{1}{\pi} \int_0^\pi (\cos^2 t)u(t)dt, \quad u_0(x) = 1.$$

14.2. Методом механічних квадратур розв'язати інтегральне рівняння, використовуючи узагальнену квадратурну формулу трапецій, Сімпсона чи Гаусса.

$$1) u(x) = 1 + x + e^x - \frac{1}{2} \int_0^\pi \sin \frac{xt}{5} u(t)dt;$$

$$2) u(x) = \sin \pi x - \int_0^{0.5} \frac{1}{5 + \cos(x+t)} u(t)dt;$$

$$3) u(x) = 1 + e^x + 0,3 \int_0^1 \frac{1}{\ln(5+xt)} u(t)dt;$$

$$4) u(x) = \cos \pi x + \int_0^1 e^{-\frac{x+t}{5}} u(t)dt;$$

$$5) u(x) = \ln(1+x) + 0,1 \int_0^1 \lg \frac{x}{5+t} u(t)dt;$$

$$6) u(x) = \cos x + \frac{1}{7} \int_0^\pi \frac{1}{5 + \sin^2(x+t)} u(t)dt;$$

$$7) u(x) = e^{-x} + \frac{1}{2} \int_0^1 \frac{\sin xt}{5+t} u(t)dt;$$

$$8) u(x) = 1 - x - \frac{1}{6} \int_0^1 \arcsin \frac{xt}{5} u(t)dt;$$

$$9) u(x) = \frac{1}{1+x} + \frac{1}{3} \int_0^1 \operatorname{arctg} \frac{x}{5+t} u(t)dt;$$

$$10) u(x) = \frac{\sin x}{1+x^2} - \frac{1}{9} \int_0^1 \cos x(5+t)^2 u(t)dt;$$

$$11) \quad u(x) = 1 + \sin x + \int_0^x \operatorname{tg} \frac{x+t}{10} u(t) dt;$$

$$12) \quad u(x) = e^x - \frac{1}{4} \int_0^1 \cos \frac{x}{5+t} u(t) dt;$$

$$13) \quad u(x) = 1 + x^2 + 7 \int_0^1 \sin \frac{x-t}{50} u(t) dt;$$

$$14) \quad u(x) = 1 - x - 9 \int_0^1 e^{-(xt+10)} u(t) dt;$$

$$15) \quad u(x) = \frac{1}{1+x^2} + \frac{1}{8} \int_0^1 \frac{x-t}{10t+13} u(t) dt;$$

14.3. Знайти розв'язок заданих інтегральних рівнянь за методом заміни ядра на вироджене.

$$1) \quad u(x) - \lambda \int_0^1 \frac{1}{10-xt} u(t) dt = 1 + x^2; \quad \lambda = \frac{1}{i}; \quad i = \overline{5, 10};$$

$$2) \quad u(x) - 0,1 \int_0^1 \sin \frac{xt}{i} u(t) dt = xe^x; \quad i = \overline{5, 10};$$

$$3) \quad u(x) - 0,5 \int_0^1 \frac{x+t}{10+t} u(t) dt = 1 + ix; \quad i = \overline{0, 10};$$

$$4) \quad u(x) - 0,7 \int_0^1 \sin(x + \frac{t}{i}) u(t) dt = \sin x; \quad i = \overline{1, 10};$$

$$5) \quad u(x) - \int_0^1 \operatorname{tg} \frac{x}{i+t} u(t) dt = 1 + x; \quad i = \overline{1, 10};$$

$$6) \quad u(x) - 0,25 \int_0^1 \frac{1}{ix + 0,1t + 5} u(t) dt = e^{-x}; \quad i = \overline{5, 10};$$

$$7) \quad u(x) - \int_0^x \sin(x-t) u(t) dt = \sin x;$$

$$8) \quad u(x) + \int_0^x (x-t) u(t) dt = x + 2 \sin x;$$

$$9) \quad u(x) - \int_0^x \frac{\operatorname{ch} t}{\operatorname{ch} x} u(t) dt = 1;$$

$$10) \quad u(x) - \int_0^x \sin(x-t) u(t) dt = \frac{1}{1+x^2};$$

$$11) \quad u(x) + \int_0^x \cos x e^{-(x-t)} u(t) dt = e^{-x} \cos x;$$

$$12) \quad u(x) + \int_0^x (x-t) u(t) dt = 4e^x + 3x - 4;$$

$$13) \quad u(x) - \int_0^x (x-t) u(t) dt = x - 1;$$

$$14) \quad u(x) - \frac{1}{2} \int_0^x (x-t)^2 u(t) dt = \sin x;$$

$$15) \quad u(x) - \int_0^x ((x-t)^2 - (x-t)) u(t) dt = 1.$$

## Список літератури

1. Бахвалов Н.С., Лапин А.В., Чижонков Е.В. Численные методы в задачах и упражнениях. – М.: Высшая школа, 2000. – 190 с.
2. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы. – М.: Наука, 1987. – 598 с.
3. Березин Н.С., Жидков Н.П. Методы вычислений: В 2 т. – М.: Наука, 1966. – Т. 1. – 632 с.; Т. 2. – 620 с.
4. Бурківська В.Л., Войцехівський С.О., Гаврилюк І.П. та ін. Методи обчислень: Практикум на ЕОМ. – К.: Вища школа, 1995. – 303 с.
5. Вержбицкий В.М. Основы численных методов. – М.: Высшая школа, 2002. – 840 с.
6. Гаврилюк І.П., Макаров В.Л. Методи обчислень: У 2 ч. – К.: Вища школа, 1995. – Ч. 1. – 376 с.; Ч. 2. – 431 с.
7. Демидович Б.П., Марон К.А. Основы вычислительной математики. – М.: Наука, 1970. – 660 с.
8. Дудикевич А.Т., Левицька С.М., Шахно С.М. Практична реалізація методів розв'язування нелінійних рівнянь і систем. – Львів: ВЦ ЛНУ ім. І. Франка, 2007. – 78 с.
9. Дэннис Дж., мл., Шнабель Р. Численные методы безусловной оптимизации и решения нелинейных уравнений. – М.: Мир, 1988. – 440 с.
10. Ильин В.А., Позняк Э.Г. Линейная алгебра. – М.: Наука, 1984. – 304 с.
11. Калиткин Н.Н. Численные методы. – М.: Наука, 1978. – 512 с.
12. Крылов В.И., Бобков В.В., Монастырский П.И. Вычислительные методы высшей математики / Под ред. И.П. Мыковских. В 2 т. – Минск: Вышэйш. школа, 1972. – Т. 1.; – 1975. – Т. 2.
13. Канторович Л.В., Акилов Г.П. Функциональный анализ. – М.: Наука, 1977. – 742 с.
14. Крылов И.И., Бобков В.В., Монастырский П.И. Начала теории вычислительных методов. Линейная алгебра и нелинейные уравнения. – Минск: Наука и техника, 1985.
15. Крылов В.И., Бобков В.В., Монастырский П.И. Начала теории вычислительных методов. Интерполирование и интегрирование. – Минск: Наука и техника, 1983. – 287 с.
16. Лященко М. Я., Головань М.С. Чисельні методи. – К.: Либідь, 1996. – 288 с.
17. Молчанов И.Н. Машинные методы решения прикладных задач. Алгебра, приближение функций. – Киев: Наук. думка, 1987. – 288 с.
18. Парлетт Б. Симметричная проблема собственных значений. – М.: Мир, 1983. – 382 с.
19. Положий Г.И., Пахарева И.А., Степаненко И.З. и др. Математический практикум. – М.: Физматгиз, 1960. – 328 с.
20. Самарский А.А. Теория разностных схем. – М.: Наука, 1983. – 616 с.
21. Самарский А.А., Гулин А.В. Численные методы. – М.: Наука, 1989. – 432 с.

22. Сборник задач по методам вычислений / Азаров А.И., Басик В.А., Мелешко И.Н. и др.; Под ред. П.И. Монастырного. – М.: Наука, 1983. – 287 с.
23. Уилкинсон Дж., Райнш К. Справочник алгоритмов на языке АЛГОЛ. Линейная алгебра. – М.: Машиностроение, 1976. – 389 с.
24. Уилкинсон Дж. Х. Алгебраическая проблема собственных значений. – М.: Наука, 1970. – 564 с.
25. Фаддеев Д.К., Фаддеева В.Н. Вычислительные методы линейной алгебры. – М.: Физматгиз, 1963. – 734 с.
26. Форсайт Дж., Малькольм М., Моупер К. Численное решение систем линейных алгебраических уравнений. – М.: Мир, 1969. – 168 с.
27. Хейгеман Л., Янг Д. Прикладные итерационные методы. – М.: Мир, 1986. – 448 с.
28. Химич А.Н., Молчанов И.Н., Попов А.В., Чистяков Т.В., Яковлев М.Ф. Параллельные алгоритмы решения задач вычислительной математики. – Киев: Наук. думка, 2008. – 248 с.
29. Цегелик Г.Г. Чисельні методи. – Львів: ВЦ ЛНУ ім. І. Франка, 2004. – 408 с.
30. Шахно С.М. Чисельні методи лінійної алгебри. – Львів: ВЦ ЛНУ ім. І. Франка, 2007. – 245 с.
31. Шахно С.М., Дудикевич А.Т., Левицька С.М. Практична реалізація чисельних методів лінійної алгебри. – Львів: ВЦ ЛНУ ім. І. Франка, 2009. – 144 с.
32. Kelley C. Iterative Methods for Linear and Nonlinear Equations. – Philadelphia: SIAM Publications, 1995. – 166 p.
33. More J.J., Garbow B.S. and Hillstom K.E. Testing Unconstrained Optimization Software // ACM Transaction on Mathematical Software. — 1981. – Vol. 7. № 1. – PP. 17—41.
34. Quarteroni A., Sacco R., Saleri F. Numerical Mathematics. – New York Berlin Heidelberg: Springer, 2000. – 655 p.



Навчальне видання

**ШАХНО Степан Михайлович**  
**ДУДИКЕВИЧ Анна Теодорівна**  
**ЛЕВИЦЬКА Софія Михайлівна**

## **П РА К Т И К У М**

### **З ЧИСЕЛЬНИХ МЕТОДІВ**

Навчальний посібник

*Рекомендовано*  
*Міністерством освіти і науки, молоді та спорту*

Редактор *М.М. Мартиняк*  
Технічний редактор *С.З. Сеник*  
Коректор *Ю. Р. Глиняна*  
Комп'ютерне верстання *Л.М. Семенович*

Формат 60×90<sup>1</sup>/<sub>16</sub>. Умовн. друк. арк. 28,37. Тираж 300 прим. Зам.

Львівський національний університет імені Івана Франка,  
вул. Університетська, 1, м. Львів, 79000

#### Свідоцтво

про внесення суб'єкта видавничої справи до Державного реєстру  
видавців, виготівників і розповсюджувачів видавничої продукції.

Серія ДК № 3059 від 13.12.2007 р.

Видруковано з готових діалозитивів  
у книжковій друкарні "Коло"  
вул. Бориславська, 8, м. Дрогобич, Львівська обл., 82100

#### Свідоцтво

про внесення суб'єкта видавничої справи до Державного реєстру  
видавців, виготівників і розповсюджувачів видавничої продукції.

Серії ДК № 498 від 20.06.2001 р.

ISBN 978-617-10-0036-0



9 786171 000360