

О. К. Колесницький, І. Р. Арсенюк, В. І. Месюра

ЧИСЕЛЬНІ МЕТОДИ

Міністерство освіти і науки України
Вінницький національний технічний університет

О. К. Колесницький, І. Р. Арсенюк, В. І. Месюра

ЧИСЕЛЬНІ МЕТОДИ

Навчальний посібник

Вінниця
ВНТУ
2017

УДК 519.61(075)
ББК 22.193я73
К60

Рекомендовано до видання Вченою радою Вінницького національного технічного університету Міністерства освіти і науки України (протокол № 4 від 27.11.2014 р.)

Рецензенти:

В. П. Кожем'яко, доктор технічних наук, професор

О. М. Роїк, доктор технічних наук, професор

А. Я. Кулик, доктор технічних наук, професор

Колесницький, О. К.

К60 Чисельні методи : навчальний посібник / О. К. Колесницький, І. Р. Арсенюк, В. І. Месюра. – Вінниця : ВНТУ, 2017. – 130 с.

У посібнику розглянуто такі питання, як вступ до чисельних методів; елементарна теорія похибок; деякі ітераційні методи розв'язання систем лінійних алгебраїчних рівнянь; деякі методи розв'язання нелінійних рівнянь та систем нелінійних рівнянь; чисельне диференціювання та інтегрування, а також наближене розв'язання звичайних диференціальних рівнянь.

Посібник розроблений відповідно до плану кафедри та програми дисципліни „Чисельні методи”.

УДК 519.61(075)
ББК 22.193я73

© ВНТУ, 2017

ЗМІСТ

ВСТУП.....	5
1 ВСТУП ДО ЧИСЕЛЬНИХ МЕТОДІВ	6
1.1 Роль чисельних методів у розв’язанні задач навколишнього світу	6
1.2 Структура похибки розв’язку задачі.....	7
1.3 Поняття стійкості та коректності	9
1.4 Основні етапи розв’язання прикладних математичних задач.....	11
1.5 Контрольні запитання.....	13
2 ЕЛЕМЕНТАРНА ТЕОРІЯ ПОХИБОК	14
2.1 Джерела і класифікація похибок	14
2.2 Абсолютна та відносна похибки	15
2.3 Правильні значущі цифри	16
2.4 Похибки арифметичних операцій	17
2.5 Графи обчислювальних процесів	22
2.6 Контрольні запитання.....	25
2.7 Завдання	25
3 РОЗВ’ЯЗАННЯ СИСТЕМ ЛІНІЙНИХ АЛГЕБРАЇЧНИХ РІВНЯНЬ.....	28
3.1 Теорема Кронекера – Капеллі.....	29
3.2 Розв’язання довільних систем лінійних алгебраїчних рівнянь.....	30
3.3 Класифікація методів розв’язання систем лінійних алгебраїчних рівнянь та їх порівняльні характеристики	33
3.4 Ітераційні методи розв’язання систем лінійних алгебраїчних рівнянь.....	35
3.4.1 Метод ітерацій	35
3.4.2 Умови збіжності ітераційного процесу.....	40
3.4.3 Оцінювання похибки методу ітерацій	41
3.4.4 Метод Зейделя	43
3.4.5 Умови збіжності процесу Зейделя.....	45
3.4.6 Зведення системи лінійних рівнянь до вигляду, зручного для ітерацій	47
3.5 Контрольні запитання.....	49
3.6 Завдання	50
4 МЕТОДИ РОЗВ’ЯЗАННЯ НЕЛІНІЙНИХ РІВНЯНЬ ТА СИСТЕМ НЕЛІНІЙНИХ РІВНЯНЬ.....	52

4.1	Постановка задачі	52
4.2	Класифікація нелінійних рівнянь	52
4.3	Відокремлення та уточнення коренів	54
4.4	Методи відокремлення коренів	55
4.4.1	Загальні властивості алгебраїчних рівнянь	58
4.4.2	Визначення кількості дійсних коренів алгебраїчного рівняння.....	59
4.4.3	Знаходження області існування коренів алгебраїчного рівняння.	62
4.5	Методи уточнення коренів	63
4.5.1	Метод хорд.....	63
4.5.2	Метод дотичних і його модифікації	67
4.5.3	Комбінований метод хорд та дотичних	71
4.5.4	Метод ітерацій	73
4.6	Розв'язання систем нелінійних рівнянь.....	77
4.6.1	Метод простої ітерації.	77
4.6.2	Метод Ньютона.....	81
4.6.3	Метод збурення параметрів.....	85
4.7	Вибір методу розв'язання нелінійних рівнянь і систем	86
4.8	Контрольні запитання.....	86
4.9	Завдання	87
5	ЧИСЕЛЬНЕ ДИФЕРЕНЦІЮВАННЯ ТА ІНТЕГРУВАННЯ.	95
5.1	Постановка задачі чисельного диференціювання	95
5.2	Особливості чисельного диференціювання	97
5.3	Постановка задачі чисельного інтегрування	99
5.4	Найпростіші квадратурні формули.....	102
5.5	Формула Сімпсона.....	114
5.6	Порівняння і практичне оцінювання похибки квадратурних формул	118
5.7	Контрольні запитання.....	123
5.8	Завдання	124
	СЛОВНИК ТЕРМІНІВ.....	126
	ЛІТЕРАТУРА.....	128

ВСТУП

Дослідження об'єктів і явищ на їх математичних моделях на сьогоднішній день залишається дуже важливою і актуальною проблемою. Застосування і дослідження таких моделей дозволяє розв'язати багато різноманітних проблем. Математичні методи дослідження об'єктів та систем давно і успішно застосовують майже в усіх сферах діяльності людини: у фізиці, механіці, астрономії, економіці, медицині, техніці тощо. До появи ЕОМ основні зусилля вчених були направлені на пошук розв'язків, поданих у явному вигляді. Проте у математиці часто зустрічаються задачі, розв'язок яких неможливо отримати у вигляді залежностей, що пов'язують шукані величини з вхідними. Для розв'язання таких задач намагаються знайти будь-який нескінченний процес, що збігається до шуканої відповіді. Якщо такий процес знайдено, то, виконуючи певну кількість кроків і обриваючи обчислення, ми отримуємо наближений розв'язок.

Такий підхід до розв'язання задач був відомий ще до появи ЕОМ, проте він застосовувався дуже рідко внаслідок великої складності. Залучення комп'ютерної техніки суттєво розширило клас задач, що можуть бути ефективно розв'язані. Тепер досліднику під час побудови математичної моделі не слід намагатися її максимально спростити, як це було потрібно для отримання розв'язку в явному вигляді. Його увага, перш за все, повинна бути направлена на те, щоб врахувати найсуттєвіші особливості досліджуваного об'єкта або явища. Після цього він вирішує питання щодо розробки і реалізації алгоритму розв'язання відповідної задачі на комп'ютері.

У посібнику наведені основи чисельних методів, розглянута елементарна теорія похибок; деякі ітераційні методи розв'язання систем лінійних алгебраїчних рівнянь; деякі методи розв'язання нелінійних рівнянь та систем нелінійних рівнянь; а також чисельне диференціювання та інтегрування.

Кожен розділ містить теоретичний матеріал, який ілюструється прикладами розв'язання задач, контрольні питання і завдання для лабораторних занять. Такий підхід дозволяє студентам ґрунтовно опанувати і засвоїти основні методи дослідження об'єктів та систем, а також навчитись їх практично реалізовувати. Також у посібнику наведено список літератури, що дозволяє значно глибше вивчити теоретичні основи чисельних методів та їх практичне застосування.

1 ВСТУП ДО ЧИСЕЛЬНИХ МЕТОДІВ

1.1 Роль чисельних методів у розв'язанні задач навколишнього світу

Бурхливий розвиток комп'ютерної техніки сприяв широкому процесу математизації науки, техніки і господарства в цілому. Саме розробка і застосування чисельних методів розв'язання прикладних задач на базі комп'ютерів є предметом сучасної прикладної математики. Математика виникла з практичних потреб людини і її розвиток сприяв загальному науково-технічному прогресу цивілізації, а потреби природознавства, техніки і практичної діяльності людей ставили перед математикою нові задачі і стимулювали її розвиток.

Ступінь розвитку суспільства визначається рівнем розвитку математичного апарату, вмінням отримати всю необхідну інформацію про досліджуваний об'єкт, можливістю описати його найсуттєвіші риси і властивості мовою математичних понять. Тобто, дуже актуальною і важливою задачею постає задача побудови математичної моделі досліджуваного об'єкта.

Математична модель (ММ) будується на основі деякого абстрагування, спрощення та ідеалізації об'єкта, а тому завжди є його наближеним описом. Але завдяки заміні реального об'єкта його ММ виникає можливість сформулювати задачу його вивчення як математичну, для розв'язування якої застосовують математичний апарат, що не залежить від природи досліджуваного об'єкта. ММ – це, як правило, різноманітні рівняння, які описують поведінку досліджуваних об'єктів і явищ [1, 4].

Вивчення реальних явищ або процесів часто приводить до потреби розв'язування диференціального рівняння або їх системи. Тому диференціальні рівняння, які виникають в результаті дослідження цих явищ чи процесів, називають диференціальними моделями. Диференціальні рівняння – окремий випадок тієї множини математичних моделей, які виникають в процесі вивчення реального світу. Математичне моделювання широко використовують у геофізиці, хімії, геології, біології, економіці, соціології, екології, медицині, психології, лінгвістиці та інших науках.

Одна й та сама диференціальна модель може описувати процеси чи явища різної природи. Так, модель, що задається рівнянням виду $y' = ky$, описує і радіоактивний розпад речовини, і зміну атмосферного тиску зі зміною висоти над земною поверхнею, і охолодження тіла внаслідок конвективного теплообміну з навколишнім середовищем, і зміну струму в електричному колі з індуктивністю при розмиканні, і зростання колонії живих організмів, що перебувають у сприятливих умовах тощо. А модель, що задається рівнянням $y'' + \omega^2 y = 0$, може описувати і коливання ку-

лі, підвішеної до пружини, і коливання математичного маятника, і зміну заряду конденсатора, пластини якого замкнені на індуктивність, і крутильні коливання тощо. Коливання різні, а фундаментальна суть їх однакова, тому вони описуються тією самою математичною моделлю. Отже, одна математична модель може мати багато реальних прообразів. Проте для описування різних сторін одного явища іноді виникає потреба у використанні іншої математичної моделі. Такі моделі здебільшого взаємно доповнюють одна одну [1, 4].

Слід зазначити, що вдало розроблена ММ часто дає змогу передбачити існування в природі певних явищ або об'єктів. Таких прикладів відомо багато, і всі вони свідчать про прогнозувальну роль математики у науці, тобто математика стала одним з найважливіших інструментів пізнання матеріального світу.

1.2 Структура похибки розв'язку задачі

Побудувавши математичну модель, намагаються знайти її розв'язок. Для складних прикладних задач, як правило, не існує точного розв'язку у вигляді явних формул або скінченної послідовності арифметичних операцій, кожна з яких виконується точно. У таких випадках застосовують потужний математичний засіб розв'язування задач – чисельні методи. Найпростіші чисельні методи виникли і широко використовувалися задовго до появи ЕОМ. Але є багато прикладних задач, для яких знайти розв'язок без застосування комп'ютерів практично неможливо. Сучасні швидкодіючі обчислювальні машини стали стимулом для розробки нових чисельних методів.

Але слід пам'ятати, що застосовувати чисельні методи для розв'язання прикладних задач на базі комп'ютерів треба обережно, оскільки точність розв'язку залежить від багатьох факторів. При цьому важливим питанням є вміння правильно оцінити похибку обчисленого розв'язку.

Похибка розв'язку задачі складається з [1]:

- похибки математичної моделі;
- неусувної похибки;
- похибки методу;
- обчислювальної похибки.

Похибка математичної моделі обумовлена тим, що модель описує явище наближено, з припущеннями і спрощеннями. І для того, щоб спростити побудову математичної моделі, треба мати уявлення про точність кінцевого результату.

Неусувна похибка зумовлена похибками у вхідних даних задачі і залежить від методу розв'язування задачі. Але, щоб правильно обрати метод і визначити точність обчислень, важливо знати межі неусувної похибки.

Похибка методу пов'язана з необхідністю заміни неперервної моделі дискретною або із завершенням нескінченного ітераційного процесу після скінченної кількості ітерацій. Наприклад, нехай на відрізку $[a;b]$ треба знайти розв'язок задачі Коші [1]

$$y' = f(x, y), y(a) = y_0. \quad (1.1)$$

Знайти чисельно розв'язок $y(x)$ в усіх точках відрізка $[a;b]$ неможливо, оскільки їх безліч. Тому на відрізку $[a;b]$ беруть скінченну кількість точок $(x_i = a + ih), i = 0, 1, \dots, n, a + nh \leq b < a + (n + 1)h)$ і лише в них знаходять значення $y(x)$. Початкове значення $y(a) = y_0$ нам відоме. Для знаходження інших значень $y_i = y(x_i)$ ($i = 1, 2, \dots, n$) диференціальне рівняння розглядають не на всьому відрізку $[a;b]$, а тільки у зазначених точках $y'(x_i) = f(x_i, y(x_i))$.

Замінивши похідну $y'(x)$ її наближеним значенням $(y_{i+1} - y_i)/h$ дістають систему рівнянь

$$y_{i+1} - y_i = hf(x_i, y_i), \quad i = 1, 2, \dots, n - 1. \quad (1.2)$$

Звідси послідовно знаходять y_1, y_2, \dots, y_n .

Якщо у рівняння (1.2) замість y_i і y_{i+1} підставити точні значення розв'язку $y(x_i)$, то рівності задовольняться лише наближено.

Похибка, що виникає під час заміни неперервної моделі дискретною, називається *похибкою дискретизації* (або *похибкою апроксимації*).

Крім похибки дискретизації є інший тип похибки чисельних методів. В основу багатьох методів покладено ідею ітераційного процесу, під час якого за певним правилом будується послідовність наближень до розв'язку задачі. Якщо ця послідовність має границю, коли кількість членів послідовності наближається до нескінченності – ця границя є розв'язком даної задачі. Але на комп'ютері можна обчислити тільки скінченну кількість членів послідовності. Похибку, обумовлену обривом ітераційного процесу, називають *похибкою збіжності*. Похибку методу намагаються звести до величини, яка в кілька разів менша за похибку вхідних даних [1].

Отже, похибку чисельного методу можуть утворювати похибки дискретизації або похибки збіжності, або ж для деяких методів обидва типи похибок одночасно. Всі ці похибки, а також методи їх аналізу і регулювання розглядаються при побудові конкретних чисельних методів.

Обчислювальні похибки пов'язані з похибками округлення чисел. Будь-які обчислення виконують з певною кількістю значущих цифр. Це вносить в результат похибку округлення, яка накопичується під час обчислень. Похибки округлення можуть по-різному впливати на кінцевий ре-

зультат. В результаті виконання мільйонів операцій, кожна з яких вносить невелику похибку, сумарна похибка округлень може значно перевищити шуканий результат обчислень. Але в окремих операціях похибки округлень можуть мати різні знаки і частково компенсувати одна одну. Тому, якщо немає систематичних причин, випадкове накопичення похибок округлення незначне [2, 4, 10].

Систематичною причиною накопичення похибок є, наприклад, віднімання близьких за значеннями чисел, оскільки при малій абсолютній похибці чисел x_1 і x_2 відносна похибка $(\Delta x_1 + \Delta x_2)/|x_1 - x_2|$ результату може стати великою [4].

Обчислювальні похибки виникають і під час перетворення чисел з однієї системи числення в іншу, якщо основа однієї системи числення не є степенем основи іншої. Це може призвести до того, що в новій системі числення число стане ірраціональним.

Втрата точності може виникнути і під час додавання до великого числа дуже малих чисел. Для зменшення похибки додавати числа слід в порядку їх зростання. У машинній арифметиці комутативний і дистрибутивний закони алгебри не завжди виконуються. Обчислювальний алгоритм треба будувати так, щоб похибка округлень була значно меншою за усі інші похибки [1, 4].

1.3 Поняття стійкості та коректності

Похибки у вхідних даних задачі – неусувні. Дослідник не може їх зменшити, але повинен знати, як вони впливають на точність кінцевого результату. Одні задачі мають похибку результату одного порядку з похибками вхідних даних, в інших задачах похибка результату може на кілька порядків перевищувати похибку вхідних даних. Чутливість задачі до неточностей у вхідних даних характеризується поняттям *стійкості*.

Задача є *стійкою* за вхідними даними, якщо її розв'язок постійно залежить від вхідних даних, тобто малому приросту Δx вхідної величини відповідає малий приріст Δu шуканого розв'язку. Іншими словами, малі похибки вхідних даних спричиняють малі похибки розв'язку задачі. Якщо ця умова не виконується – задача є *нестійкою* за вхідними даними. Це означає, що навіть незначні похибки вхідних даних можуть призвести до як завгодно великих похибок розв'язку, тобто розв'язок може бути дуже викривлений. Тому застосовувати безпосередньо до таких задач чисельні методи не можна, оскільки похибки округлень при застосуванні методу будуть катастрофічно накопичуватися під час обчислень.

Наведемо приклад нестійкої задачі, який належить Уілкінсону [1]. Коренями многочлена

$$P(x) = (x - 1)(x - 2) \dots (x - 20) = x^{20} - 210x^{19} + \dots \quad (1.3)$$

є числа $x_1 = 1, x_2 = 2, \dots, x_{20} = 20$. Нехай один з його коренів обчислено з незначною похибкою, наприклад, коефіцієнт -210 при x^{19} замінено на коефіцієнт $-210 + 2^{-23}$ ($2^{-23} \approx 10^{-7}$). Тоді дещо змінений многочлен $P(x) = P(x) - 2^{-23} x^{19}$ має такі округлені до двох десяткових (після коми) знаків корені:

$$x_1 = 1,00; x_2 = 2,00; \dots; x_8 = 8,01; x_9 = 8,92; x_{10,11} = 10,10 \pm 0,64i; \\ x_{12,13} = 11,79 \pm 1,65i; \dots; x_{18,19} = 19,50 \pm 1,94i; x_{20} = 20,85.$$

Незначна похибка в коефіцієнті -210 даного многочлена викликала суттєво інші значення коренів (десять з них стали комплексними). Причиною цього є нестійкість задачі, оскільки корені обчислювали з точністю до 11 значущих цифр і похибка округлень незначна.

Іноді під час розв'язування стійкої за вхідними даними задачі нестійким може бути метод її розв'язування. Нехай треба обчислити інтеграли [1]

$$I_n = \int_0^1 x^n e^{x-1} dx, \quad n = 1, 2, \dots \quad (1.4)$$

Інтегруючи частинами, маємо

$$\int_0^1 x^n e^{x-1} dx = x^n e^{x-1} \Big|_0^1 - \int_0^1 n x^{n-1} e^{x-1} dx = 1 - n \int_0^1 x^{n-1} e^{x-1} dx.$$

Звідси дістанемо $I_1 = \frac{1}{e}, I_2 = 1 - 2I_1, \dots, I_n = 1 - nI_{n-1}$.

Використавши рекурентне співвідношення, обчислимо перші дев'ять інтегралів $I_1 = 0,367879, I_2 = 0,264242, I_3 = 0,207274, I_4 = 0,170904, I_5 = 0,145480, I_6 = 0,127120, I_7 = 0,110160, I_8 = 0,118720, I_9 = 0,0684800$.

Значення інтеграла I_9 помилкове, оскільки підінтегральна функція $x^9 e^{x-1}$ в усіх точках відрізка $[0; 1]$ невід'ємна. Помилка зумовлена похибкою округлення значення I_1 до шести значущих цифр. Ця похибка наближено дорівнює $4,4 \cdot 10^{-7}$. При обчисленні I_1 вона множиться на -3 і т.д. Похибка в I_9 дорівнює $9! \cdot 4,4 \cdot 10^{-7} \approx 0,1601$. Вона спотворила істинне значення I_9 , яке з трьома значущими цифрами дорівнює $0,0916$.

Введемо тепер поняття *коректності* задачі. Задача є *коректно поставленою*, якщо для будь-яких вхідних даних з деякого класу існує єдиний і стійкий за вхідними даними її розв'язок [1]. Наведена вище задача обчислення коренів многочлена (1.3) є некоректно поставленою, а обчислення

інтегралів (1.4) – коректно поставленою задачею. Прикладом некоректно поставленої задачі є також задача чисельного диференціювання функцій.

Для розв'язування некоректно поставлених задач застосовувати класичні чисельні методи не слід, оскільки похибки округлень при розрахунках можуть катастрофічно зростати і призводити до результату, далекого від шуканого розв'язку. Для розв'язування некоректно поставлених задач використовують так звані методи регуляризації, які замінюють дану задачу коректно поставленою [1, 12].

1.4 Основні етапи розв'язання прикладних математичних задач

Метод математичного моделювання – ефективний засіб дослідження реальних об'єктів та явищ. Але з їх ускладненням ускладнюються й математичні моделі, розрахунки яких потребують величезної обчислювальної роботи. Це стримувало використання моделей в науці й техніці, оскільки математичні задачі розв'язували вручну. Застосування швидкодіючих та потужних комп'ютерів для розв'язування складних прикладних задач сформувало новий спосіб проведення теоретичних досліджень на базі математичних моделей – *обчислювальний* (або математичний) *експеримент*. Його основою є математичне моделювання, теоретичною базою – прикладна математика, технічною – комп'ютер.

Поняття розв'язування задачі за допомогою комп'ютера містить в собі набагато більше, ніж просто обчислення на комп'ютері. Обчислювальний експеримент складається з восьми основних етапів (рис. 1.1.)

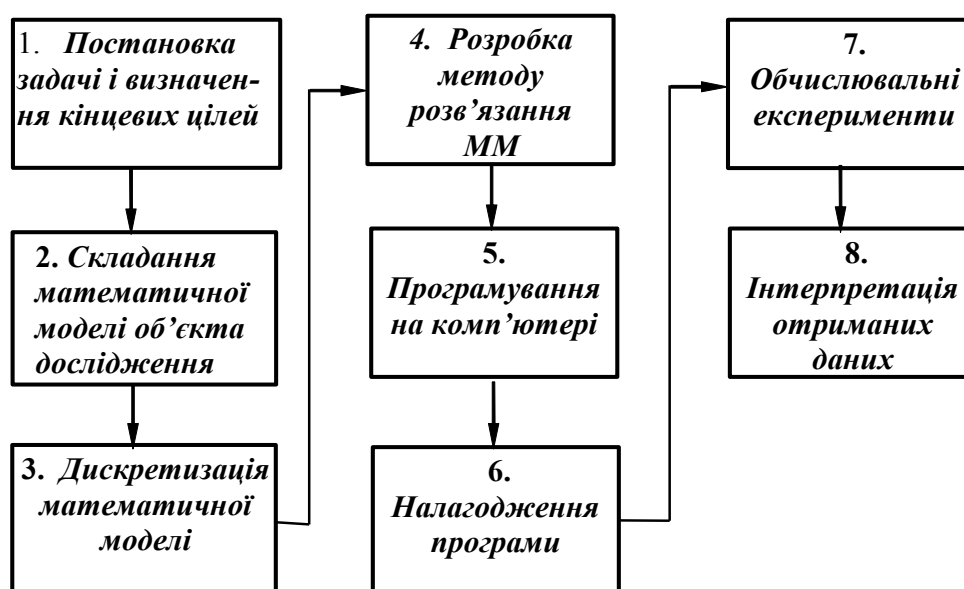


Рисунок 1.1 – Технологічний ланцюг розв'язання прикладних математичних задач

1. *Постановка задачі і визначення кінцевих цілей* – найважливіший етап розв'язування задачі. На ньому реалізується вибір загального підхо-

ду до розв'язування, визначаються основні критерії, які повинна задовольняти розроблювана система і наводиться формальний математичний опис задачі. На цьому етапі необхідне глибоке розуміння задачі.

2. *Складання математичної моделі об'єкта дослідження.* Для цього вивчають суть досліджуваного явища або процесу, з'ясовують його склад, закономірності, особливості і взаємозв'язки окремих його частин. Виділяють основні і другорядні фактори. Другорядними на початковому етапі досліджень нехтують, а основні – записують у вигляді математичних формул. Одночасно встановлюють межі застосування побудованої моделі, оскільки жодна математична модель не адекватна реальності. У прикладних задачах побудова математичної моделі (ММ) – один з найскладніших і найвідповідальніших етапів обчислювального експерименту, оскільки вимагає поєднання математичних і спеціальних знань. Тому над її створенням працюють спільно математики і спеціалісти тієї галузі, до якої належить досліджуваний об'єкт.

Після побудови ММ переходять до її теоретичного аналізу: досліджують, чи коректно поставлено задачу, чи має вона розв'язок, єдиний він чи ні тощо. Часто розв'язувати складні прикладні задачі починають і без вичерпного аналізу їхніх математичних властивостей, оскільки для такого аналізу можна затратити час, який перевищує відведений для розв'язування прикладної задачі термін.

3. *Дискретизація ММ* – викликана конструктивними особливостями цифрових комп'ютерів, що мають обмежену розрядну сітку і пам'ять. Перехід від континуальної до дискретної ММ полягає у заміні функцій безперервного аргументу функціями дискретного аргументу. При цьому, наприклад, інтеграл змінюється кінцевою сумою, похідні – різницеvim відношенням тощо. Внаслідок цього дістають систему з більшою кількістю рівнянь з більшою кількістю змінних.

4. *Розробка методу розв'язання ММ.* Як правило, прикладні задачі описуються моделями, розв'язок яких не можна знайти у вигляді аналітичних формул. Тоді з відомих чисельних методів вибирають найбільш ефективний або розробляють новий метод розв'язання моделі і записують його у вигляді алгоритму. Оскільки для однієї задачі можна скласти, як правило, декілька алгоритмів – слід визначити критерії оцінювання їх якості. Оцінюючи ефективність чисельного методу, враховують такі його якості, як універсальність, простота організації обчислень, контроль точності, швидкості збіжності та стійкості обчислювального процесу.

5. *Програмування на комп'ютері.* На цьому етапі складають програму для реалізації розробленого алгоритму або використовують стандартні пакети прикладних програм, що призначені для розв'язування певного класу математичних задач.

6. *Налагоджування програми.* Тут виконують тестування програми з метою виявлення помилок. Для тестової задачі підбирають такі вхідні дані, які дозволять надійно судити про достовірність розвитку.

7. *Обчислювальні експерименти.* Використовуючи налагоджену програму, виконують повномасштабні виробничі розрахунки на комп'ютері. Тут також передбачається дослідження поведінки об'єкта дослідження (ОД) в умовах, де натурні експерименти ще не проводились або зовсім неможливі.

8. *Інтерпретація отриманих даних.* Отримані результати розрахунків всебічно аналізуються, внаслідок цього або отримується повна відповідність для розв'язуваної задачі, або виявляються деякі особливості поведінки ОД, що потребують нових комп'ютерних експериментів. При цьому часто виникає необхідність у змінюванні постановки задачі, що приводить до повного повторення усіх вище вказаних етапів розв'язання задачі.

Таким чином, основу будь-якого обчислювального експерименту становить тріада: "модель–метод–програма". З розгляду ланцюга розв'язування задач на комп'ютері можна зробити такі висновки:

1) комп'ютер сам задачу не розв'язує, а лише виконує наперед задану послідовність обчислень;

2) використання комп'ютера не звільняє дослідника від всебічного і скрупульозного осмислення своєї роботи та глибокого вивчення ОД.

1.5 Контрольні запитання

1. Поясніть роль математичного моделювання на сучасному етапі розвитку суспільства.

2. Наведіть та коротко охарактеризуйте структуру похибки розв'язку задачі.

3. Наведіть і дайте стисло характеристику основних етапів розв'язання прикладних математичних задач.

4. Дайте означення поняття стійкості задачі. Наведіть відповідні приклади.

5. Дайте означення поняття коректності задачі. Наведіть приклади коректно та некоректно поставлених задач.

6. Наведіть та охарактеризуйте основні етапи розв'язання прикладних математичних задач.

7. Поясніть, які з етапів розв'язання прикладних математичних задач можна виконувати на комп'ютері, а які потребують творчого підходу.

2 ЕЛЕМЕНТАРНА ТЕОРІЯ ПОХИБОК

2.1 Джерела і класифікація похибок

Розв'язання більшості практичних задач з певним ступенем умовності можна навести у вигляді двох послідовних етапів [4]:

- 1) математичний опис проблеми, що розглядається;
- 2) розв'язання сформульованої математичної задачі.

На першому етапі зустрічаються два характерних джерела похибки. По-перше: реальні процеси не завжди можна описати математично, а спрощення дозволяють отримати лише більш-менш ідеалізовані моделі. По-друге: вхідні параметри, отримані, як правило, з експерименту, неточні, що дозволяє отримати лише приблизний результат.

Відповідно до цього сумарна похибка математичної моделі і початкових даних об'єднується у похибку вихідної інформації. Якщо мати на увазі незалежність цієї похибки від другого етапу розв'язання задачі, її часто називають *неусувною похибкою*.

Отримання точного розв'язку математичної задачі (другий етап) незалежно від того, будується воно вручну або на комп'ютері, як правило, неможливо. Так, наприклад, лише для дуже обмеженого класу диференціальних рівнянь можна отримати точний розв'язок. Тому в практичних розрахунках використовуються методи отримання наближених розв'язків – у першу чергу чисельних.

Така вимушена заміна точного розв'язку наближеним і породжує *похибку методу* або, як її часто називають, *похибку апроксимації*.

У процесі розв'язування задачі виконується округлення вхідних даних, проміжних і кінцевих результатів. Ці похибки, а також похибки, які з'являються в результаті виконання арифметичних операцій над наближеними числами, тією чи іншою мірою переносяться на результат розв'язку і утворюють так звану *обчислювальну похибку* (або похибку округлення).

Враховуючи сказане, під час постановки задачі або вказують потрібну точність розв'язку (тобто задають похибку, максимально допустиму в процесі усіх розрахунків), або лише обмежуються потребами підрахунку сумарної похибки результату. Тому при роботі з наближеними числами необхідно вміти розв'язувати такі задачі [4]:

- 1) давати математичні характеристики точності наближених чисел;
- 2) знаючи ступінь точності вихідних даних, оцінювати ступінь точності результату;
- 3) вибирати вихідні дані з таким ступенем точності, щоб забезпечити задану точність результату;
- 4) оптимальним чином будувати процес розв'язування, щоб не виконувати розрахунків, які не впливають на точні числа результату.

2.2 Абсолютна та відносна похибки

Нехай A – точне число, а a – його приблизне значення. Якщо $a < A$, то кажуть, що число a є *наближеним значенням числа A з недостачею*, якщо $a > A$ – *наближеним значенням з надлишком*.

Різницю $A - a$ називають *похибкою* (або *помилкою*). Як правило, величину похибки $A - a$ і навіть її знак виявити неможливо, оскільки невідомо число A . Тому замість самої похибки використовується верхня межа її абсолютної величини.

Абсолютною похибкою наближеного числа a називається величина Δ_a , яка задовольняє нерівність

$$\Delta_a \geq |A - a|. \quad (2.1)$$

Абсолютна похибка – це верхня межа відхилення точного числа A від наближеного:

$$a - \Delta_a \leq A \leq a + \Delta_a. \quad (2.2)$$

Нерівність (2.2) часто записують так:

$$A = a \pm \Delta_a. \quad (2.3)$$

Як абсолютну похибку беруть якомога менше число. Абсолютна похибка відображає лише кількісну сторону похибки, тобто не показує, добре чи погано виконані вимірювання або розрахунки. Для того, щоб оцінити якість виконаних розрахунків чи вимірювань, вводять поняття відносної похибки.

Відотною похибкою наближеного числа a називається величина δ_a , яка задовольняє нерівність

$$\delta_a \geq |(A - a) / a|, \quad a \neq 0. \quad (2.4)$$

Зокрема, відотною похибкою можна вважати

$$\delta_a = \Delta_a / |a|, \quad a \neq 0, \quad (2.5)$$

а відношення (2.3) навести у вигляді

$$A = a \cdot (1 \pm \delta_a). \quad (2.6)$$

2.3 Правильні значущі цифри

Під час розв'язування задач часто ставиться умова: отримати результат з точністю до 0,1; 0,001 тощо. Може скластися враження, що точність розрахунків визначається числом десяткових знаків після коми. Однак це неправильно. Точність розрахунків визначається числом цифр результату, які заслуговують довіри [4].

Значущими цифрами числа називаються всі його цифри, крім нулів, розташованих лівіше першої відмінної від нуля цифри. Нулі в кінці числа – завжди значущі цифри (в іншому випадку їх не пишуть).

Приклад 1. Числа 0,001604 і 30,500 мають відповідно 4 і 5 значущих цифр.

При написанні цілих чисел можуть зустрічатися деякі неточності. Так, якщо ми хочемо показати, що у числа 400 000 останніх три нуля незначущі, то дане число потрібно записати у вигляді двох співмножників: $400 \cdot 10^3$ або $40,0 \cdot 10^4$, або $0,400 \cdot 10^6$. Остання формула запису називається *нормалізованою* і є переважною. У цьому випадку говорять, що 400 є *мантисою* числа, а 6 – його *порядком*.

Нагадаємо, що кожне додатне число може бути наведене у вигляді $a = \alpha_1 \cdot 10^m + \alpha_2 \cdot 10^{m-1} + \dots + \alpha_n \cdot 10^{m-n+1} + \dots$.

Цифра α_n наближеного числа a називається *правильною значущою цифрою* (або просто *правильною*), якщо виконується нерівність

$$|A - a| \leq 0,5 \cdot 10^{m-n+1}, \quad (2.7)$$

тобто, якщо абсолютна величина різниці $|A - a|$ не більше $1/2$ десяткового розряду, у якому стоїть α_n .

Оскільки завжди замість $|A - a|$ розглядається абсолютна похибка Δ_a , то нерівність (2.7) часто записують так:

$$\Delta_a \leq 0,5 \cdot 10^{m-n+1}, \quad (2.8)$$

адже при виконанні цієї нерівності виконується і нерівність (2.7).

З іншого боку, якщо задано число n правильних знаків наближеного числа a , то за абсолютну похибку можна взяти

$$\Delta_a = 0,5 \cdot 10^{m-n+1}. \quad (2.9)$$

Якщо нерівність (2.8) не виконується, то цифра α_n – *сумнівна*. Якщо цифра α_n правильна – то усі попередні (зліва від неї) цифри також правильні.

Приклад 2. Число $a = 23,10$ отримано округленням деякого точного числа. Скільки правильних цифр містить число a ?

При округленні чисел за правилом парної цифри абсолютна похибка не може перевищити половини одиниці останнього розряду, що залишаємо. Це означає, що в округленого числа всі цифри, що залишилися, є правильними. В даному випадку, всі чотири цифри правильні, а похибка $\Delta_a = 0,005$.

Приклад 3. Число $a = 23,071937$ містить п'ять правильних цифр. Визначити його абсолютну похибку.

Скористаємося формулою (2.9). Тут $m=1$, $n=5$, отже за абсолютну похибку можна взяти $\Delta_a = 0,5 \cdot 10^{1-5+1} = 0,0005$.

Приклад 4. Абсолютна похибка числа $a=705,1978$ становить $\Delta_a = 0,3$. Визначити, які цифри числа a є правильними, і округлити число a , залишивши лише правильні цифри.

Скористаємося формулою (2.8). Тут $m=2$, $\Delta_a = 0,3$, а n визначають з нерівності $0,3 \leq 0,5 \cdot 10^{3-n}$. Безпосередньою перевіркою переконуємося, що найбільше n , яке задовольняє цю нерівність, дорівнює 3, і цифра 5 є правильною: $0,3 < 0,5 \cdot 10^{2-3+1}$, а цифра 1 – сумнівною: $0,3 > 0,5 \cdot 10^{2-4+1}$.

Звідси випливає, що число $a=705,1978$ має три правильних цифри. Округляємо його до трьох цифр: $a_1=705$. При цьому сумарна похибка дорівнює сумі вхідної похибки і похибки округлення: $\Delta_a = 0,3 + 0,2 = 0,5$, тому можна записати: $A = 705 \pm 0,5$.

Останнім часом частіше використовують поняття правильних значущих цифр в широкому розумінні. Це розуміння пов'язано з найпростішим правилом округлення.

Цифра α_n наближеного числа

$$a = \alpha_1 \cdot 10^m + \alpha_2 \cdot 10^{m-1} + \dots + \alpha_n \cdot 10^{m-n+1} + \dots$$

є правильною значущою цифрою в широкому розумінні, якщо виконується нерівність

$$\Delta_a \leq 1 \cdot 10^{m-n+1}, \quad (2.10)$$

тобто, якщо абсолютна похибка числа a не перевершує одиниці десятичного розряду, в якому розташовано α_n .

2.4 Похибки арифметичних операцій

Похибки суми і різниці. Розглянемо точні числа A_1, A_2, \dots, A_n і їх наближені значення a_1, a_2, \dots, a_n . Нехай $A = \sum_{i=1}^n A_i$ – сума всіх точних чисел;

$a = \sum_{i=1}^n a_i$ – сума їх наближених значень. Поставимо задачу: при відомих абсолютних похибках $\Delta_{a_1}, \Delta_{a_2}, \dots, \Delta_{a_n}$ всіх наближених чисел оцінити абсолютну похибку їх суми a .

Запишемо різницю $A - a = (A_1 - a_1) + (A_2 - a_2) + \dots + (A_n - a_n)$. Перейшовши до абсолютних величин правої і лівої частин цього відношення і використовуючи властивість абсолютних величин, отримаємо $|A - a| \leq |A_1 - a_1| + |A_2 - a_2| + \dots + |A_n - a_n|$. Відповідно,

$$|A - a| \leq \Delta_{a_1} + \Delta_{a_2} + \dots + \Delta_{a_n} \quad (2.11)$$

і як абсолютну похибку наближеного числа a , тобто суми наближених чисел a_1, a_2, \dots, a_n , можна взяти суму абсолютних похибок доданків:

$$\Delta_a = \Delta_{a_1} + \Delta_{a_2} + \dots + \Delta_{a_n}. \quad (2.12)$$

З останньої формули випливає, що абсолютна похибка алгебраїчної суми, в загальному випадку, не повинна бути меншою абсолютної похибки найнеточнішого з доданків. Тому, щоб не виконувати зайвих розрахунків, не потрібно зберігати зайві знаки і в точніших доданках.

Зауваження. При великій кількості доданків ($n > 10$) оцінка абсолютної похибки суми за формулою (2.12) виявляється дуже завищеною, оскільки завжди виконується частинна компенсація похибок різних знаків. Якщо усі доданки округлені до m -го десяткового розряду, тобто їх похибки оцінюються величиною $0,5 \cdot 10^{-m}$, то статична оцінка абсолютної похибки суми обчислюється за формулою

$$\Delta_a = \sqrt{n} \cdot 0,5 \cdot 10^{-m}.$$

Зауважимо, що під час віднімання близьких чисел часто виникає стан, що називається *втратою точності* [4]. Нехай $x > 0$, $y > 0$ і $a = x - y$, тоді $\delta_a = \Delta_a / |a| = \Delta_x + \Delta_y / |x - y|$.

Таким чином, якщо числа x і y мало відрізняються одне від одного, то навіть при малих похибках Δ_x і Δ_y величина відносної похибки різниці може виявитись значною.

Приклад 1. Нехай $x=5,125$, $y=5,135$ тут $\Delta_x=0,0005$, $\Delta_y=0,0005$, $\delta_x \approx \delta_y \approx 0,01\%$. Відносна похибка різниці $a = x - y$ становить

$$\delta_a = (0,0005 + 0,0005) / 0,01 \cdot 100 = 10 \%.$$

Очевидно, що в результаті ділення двох наближених чисел може виникнути значна втрата точності. Щоб цього не допустити, слід спробувати так перетворити обчислювальну схему, щоб малі різниці величин обраховувались безпосередньо.

Приклад 2. Знайти різницю $\alpha = \sqrt{6,27} - \sqrt{6,26}$ та оцінити відносну похибку результату.

Нехай $a_1 = \sqrt{6,27} = 2,504$; $\Delta_{a_1} = 0,0005$; $a_2 = \sqrt{6,26} = 2,502$; $\Delta_{a_2} = 0,0005$. Тоді $a = 2,504 - 2,502 = 0,2 \cdot 10^{-2}$; $\Delta_a = 0,005 + 0,0005 = 0,001$, звідки $\delta_a = \frac{0,1 \cdot 10^{-2}}{0,2 \cdot 10^{-2}} = 0,5 = 50\%$.

Однак, змінивши обчислювальну схему, можна отримати значно кращу оцінку відносної похибки:

$$\begin{aligned} a &= \sqrt{6,27} - \sqrt{6,26} = \frac{(\sqrt{6,27} - \sqrt{6,26}) \cdot (\sqrt{6,27} + \sqrt{6,26})}{(\sqrt{6,27} + \sqrt{6,26})} = \\ &= \frac{6,27 - 6,26}{\sqrt{6,27} + \sqrt{6,26}} = \frac{0,01}{\sqrt{6,27} + \sqrt{6,26}} \approx 0,2 \cdot 10^{-2}; \\ \delta_a &= \delta \cdot (\sqrt{6,27} + \sqrt{6,26}) = (\Delta_{a_1} + \Delta_{a_2}) / (a_1 + a_2) = 0,001 / 5 = 0,02\%. \end{aligned}$$

Таким чином, при обчисленні з тими ж чотирма правильними знаками a_1 і a_2 ми отримали значно кращий результат в розумінні відносної похибки.

Похибка добутку. Розглянемо два точних числа A_1 і A_2 та їх наближені значення a_1 і a_2 . Нехай $A = A_1 A_2$ і $a = a_1 a_2$. Поставимо задачу: оцінити відносну похибку добутку δ_a , якщо відносні похибки чисел a_1 і a_2 дорівнюють δ_{a_1} і δ_{a_2} , відповідно.

Запишемо A_1 і A_2 у вигляді $A_1 = a_1 + \Delta_1$, $A_2 = a_2 + \Delta_2$, де невідомі Δ_1 і Δ_2 задовольняють нерівності

$$|\Delta_1| \leq |a_1| \delta_{a_1}, \quad |\Delta_2| \leq |a_2| \delta_{a_2}. \quad (2.13)$$

Перемноживши праві і ліві частини цих співвідношень, отримаємо $A_1 A_2 = a_1 a_2 + \Delta_1 a_2 + \Delta_2 a_1 + \Delta_1 \Delta_2$. Перейшовши до абсолютних величин правої і лівої частин цього співвідношення і використовуючи властивості абсолютних величин, отримаємо $|A_1 A_2 - a_1 a_2| \leq |\Delta_2 a_1| + |\Delta_1 a_2| + |\Delta_1 \Delta_2|$.

Відкинемо останній доданок правої частини, оскільки він дуже малий, і поділимо праву та ліву частини нерівності на $|a| = |a_1 a_2|$. Тоді, враховуючи (2.13), маємо $|(A - a) / a| \leq \delta_{a_1} + \delta_{a_2}$.

З отриманого співвідношення випливає, що як відносну похибку до-

бутку $a=a_1a_2$ можна взяти суму відносних похибок співмножників:

$$\delta_a = \delta_{a_1} + \delta_{a_2}. \quad (2.14)$$

Нерівність (2.14) легко поширюється на добуток кількох співмножників, так, якщо $A=A_1A_2..A_n$ і $a=a_1a_2..a_n$, то можна вважати, що

$$\delta_a = \delta_{a_1} + \delta_{a_2} + \dots + \delta_{a_n} \dots \quad (2.15)$$

В тому випадку, коли всі співмножники, крім одного, є точними числами, з формули (2.15) випливає, що відносна похибка добутку збігається з відносною похибкою наближеного співмножника. Таким чином, якщо наближеним числом є лише значення множника a_i , то

$$\delta_a = \delta_{a_i}. \quad (2.16)$$

Зауваження. При множенні наближеного числа a на точний співмножник k відносна похибка добутку дорівнює відносній похибці наближеного числа a , а абсолютна похибка у $|k|$ разів більша за абсолютну похибку наближеного числа.

Число вірних знаків добутку. Нехай даний добуток має k співмножників ($k \leq 10$) $a=a_1a_2..a_k$, де $a_i \neq 0$. Кожний із співмножників містить не менше ніж n правильних цифр ($n > 1$).

Нехай кожний із співмножників має вигляд

$$a_i = \alpha_i \cdot 10^{l_i} + \beta_i \cdot 10^{l_i-1} + \gamma_i \cdot 10^{l_i-2} + \dots \quad (i=1, 2, \dots, k), \quad (2.17)$$

де α_i – перші значущі цифри наближених співмножників, записаних у десятковій системі числення.

Для відносної похибки наближеного числа, що має n правильних знаків, використовуємо формулу

$$\delta_{a_i} = 0,5 / (\alpha_i \cdot 10^{n-1}) \quad (i=1, 2, \dots, k),$$

тоді відносна похибка добутку k наближених чисел, кожне з яких має n правильних значущих цифр, дорівнює

$$\delta_a = \delta_{a_1} + \delta_{a_2} + \dots + \delta_{a_k} = \frac{0.5}{10^{n-1}} \left(\frac{1}{\alpha_1} + \frac{1}{\alpha_2} + \dots + \frac{1}{\alpha_k} \right). \quad (2.18)$$

Враховуючи, що число співмножників не більше 10 ($k \leq 10$), отрима-

ємо $\frac{1}{\alpha_1} + \frac{1}{\alpha_2} + \dots + \frac{1}{\alpha_k} \leq 10$ і, отже, $\delta_a \leq \frac{0,5}{10^{n-2}}$.

Таким чином, якщо всі співмножники мають n правильних цифр і число співмножників не більше десяти, то число правильних знаків добутку на одну або дві одиниці менше, ніж в n . У тому випадку, коли співмножники мають різну точність, під n варто розуміти число правильних знаків найменш точного зі співмножників.

Зауваження. При великому числі співмножників ($k > 10$) зручно користуватися статистичною оцінкою, яка враховує часткову компенсацію похибки різних знаків. Якщо всі числа a_i ($i = 1, 2, \dots, k$) мають приблизно однакову відносну похибку δ , то відносна похибка добутку береться рівною [4]

$$\delta_a = \sqrt{n} \cdot \delta. \quad (2.19)$$

Похибка частки. Розглянемо точні числа A_1, A_2 та їх наближені значення a_1, a_2 з абсолютними похибками $\Delta_{a_1}, \Delta_{a_2}$. Поставимо задачу: оцінити відносну похибку наближеного значення частки $a = a_1/a_2$ для точного значення $A = A_1/A_2$. Нехай $a_1 \neq 0, a_2 \neq 0$. Точні значення A_1 і A_2 наведемо у вигляді

$$A_1 = a_1 + \Delta_1, \quad A_2 = a_2 + \Delta_2, \quad (2.20)$$

де невідомі Δ_1 і Δ_2 задовольняють нерівності

$$|\Delta_1| \leq \Delta_{a_1}, \quad |\Delta_2| \leq \Delta_{a_2}. \quad (2.21)$$

Розглянемо тепер різницю $A - a = \frac{a_1 + \Delta_1}{a_2 + \Delta_2} - \frac{a_1}{a_2} = \frac{a_2\Delta_1 - a_1\Delta_2}{a_2(a_2 + \Delta_2)}$. Поділивши праву і ліву частини на a , розглянемо їх абсолютні величини: $\left| \frac{A - a}{a} \right| = \left| \frac{a_2\Delta_1 - a_1\Delta_2}{a_2(a_2 + \Delta_2)} \right| = \left| \frac{a_2}{a_2 + \Delta_2} \right| \left| \frac{\Delta_1}{a_1} - \frac{\Delta_2}{a_2} \right|$. Враховуючи, що Δ_2 мале порівняно з a_2 , наближено візьмемо, що $a_2 / (a_2 + \Delta_2) \approx 1$. Тоді, використовуючи властивості абсолютних величин і нерівність (2.21), отримаємо $\left| \frac{A - a}{a} \right| = \left| \frac{\Delta_1}{a_1} - \frac{\Delta_2}{a_2} \right| \leq \frac{\Delta_{a_1}}{|a_1|} + \frac{\Delta_{a_2}}{|a_2|} = \delta_{a_1} + \delta_{a_2}$.

Таким чином, за відносну похибку частки $a = a_1 / a_2$ можна взяти суму відносних похибок діленого і дільника:

$$\delta_a = \delta_{a_1} + \delta_{a_2}. \quad (2.22)$$

Число правильних знаків частки. Нехай наближені числа

$$a_1 = \alpha_1 \cdot 10^h + \alpha_2 \cdot 10^{h-1} + \dots; \quad a_2 = \beta_1 \cdot 10^h + \beta_2 \cdot 10^{h-1} + \dots$$

мають по n правильних значущих цифр. Тоді, використовуючи рівності

$$\delta_{a_1} = \frac{0,5}{\alpha_1 \cdot 10^{n-1}}, \quad \delta_{a_2} = \frac{0,5}{\beta_1 \cdot 10^{n-1}}, \quad \text{знайдемо відносну похибку частки } a = \frac{a_1}{a_2}:$$

$$\delta_a = \delta_{a_1} + \delta_{a_2} = \frac{0,5}{\alpha_1 \cdot 10^{n-1}} + \frac{0,5}{\beta_1 \cdot 10^{n-1}} = \frac{0,5}{10^{n-1}} \left(\frac{1}{\alpha_1} + \frac{1}{\beta_1} \right). \quad (2.23)$$

Відповідно, якщо $\alpha_1 \geq 2$ і $\beta_1 \geq 2$, то частка має $n-1$ правильну значущу цифру. Якщо ж $\alpha_1 = 1$ і $\beta_1 = 1$, то частка може мати $n-2$ правильних значущих цифр.

Похибки степеня і кореня. Розглянемо наближене число a_1 , яке має відносну похибку δ_{a_1} . Нехай треба оцінити похибку степеня $a = a_1^m$. Очевидно, що $a = a_1^m = \underbrace{a_1 \cdot a_1 \cdot \dots \cdot a_1}_{m\text{-добутків}}$. Отже, відносна похибка добутку може бути знайдена за формулою

$$\delta_a = \underbrace{\delta_{a_1} + \delta_{a_1} + \dots + \delta_{a_1}}_{m\text{-доданків}} = m\delta_{a_1}. \quad (2.24)$$

У практичних розрахунках при піднесенні до степеня наближеного числа в результаті залишають стільки значущих цифр, скільки їх містило наближене число.

Відносна похибка числа $a = \sqrt[m]{a_1}$ в m разів менша за відносну похибку числа a_1 : $\delta_a = (1/m) \cdot \delta_{a_1}$.

У практичних розрахунках під час обчислення кореня з наближеного числа в результаті залишають стільки значущих цифр, скільки їх містило наближене число.

2.5 Графи обчислювальних процесів

Для оцінювання загальної похибки обчислень деякої послідовності арифметичних операцій існує зручний спосіб обчислення похибок, оснований на графі обчислювального процесу [15].

Граф обчислювального процесу дозволяє наочно зобразити

послідовність арифметичних операцій у деякому обчисленні. З його допомогою легко визначити внесок будь-якої похибки, що виникла в процесі обчислень, у загальну похибку.

Нехай вершини графа є змінні або арифметичні операції. Дуги графа вказують напрямок обчислень, і вони позначені коефіцієнтами, що оцінюють поширення помилок.

Граф обчислювального процесу будується для аналізу процесу поширення відносних похибок в арифметичних виразах. Його слід читати знизу вгору в напрямку дуг графа. Спочатку виконуються операції, розташовані на будь-якому горизонтальному рівні, потім операції, розташовані на більш високому рівні, і т. д. На рис. 2.1. наведені графи обчислювальних процесів основних арифметичних операцій.

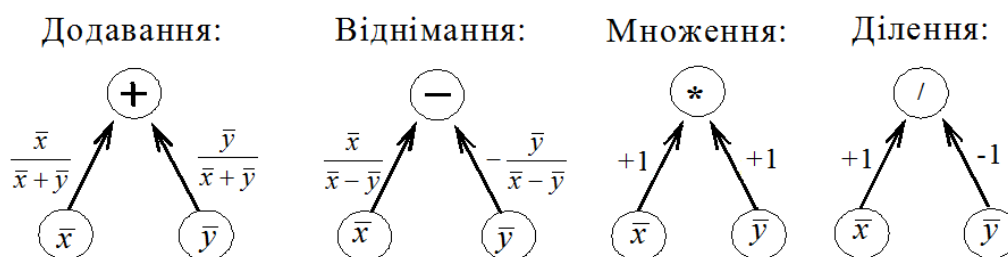


Рисунок 2.1 – Графи обчислювальних процесів основних арифметичних операцій

Правило підрахунку загальної похибки з використанням графа обчислювального процесу таке: відносна похибка результату будь-якої операції входить в результат наступної операції, збільшуючись на коефіцієнт, вказаний коло стрілки, що з'єднує ці дві операції.

Як приклад розглянемо вираз $u=(x+y)z$ і задачу оцінювання загальної похибки результату з урахуванням похибок округлення арифметичних операцій. Припустимо, що $\delta_x, \delta_y, \delta_z$ – відносні похибки округлення чисел x, y, z при їх поданні у комп'ютері, а r_+ і r^* – відносні похибки округлення відповідно операцій додавання і множення.

Для нашого виразу, з огляду на послідовність операцій, граф обчислювального процесу буде мати вигляд, наведений на рис. 2.2.

Спочатку розглянемо операцію додавання (II рівень). Операція додавання використовує числа x і y , задані з відносними похибками δ_x і δ_y . Кожна з похибок входить у результат помноженою на відповідні коефіцієнти

$\frac{x}{x+y}$ і $\frac{y}{x+y}$, тоді похибку операції додавання можна оцінити величиною $\frac{x}{x+y} \cdot \delta_x + \frac{y}{x+y} \cdot \delta_y$, до якої слід додати похибку округлення, тобто відносна похибка операції додавання буде виглядати

$$\frac{x}{x+y} \cdot \delta_x + \frac{y}{x+y} \cdot \delta_y + r_+.$$

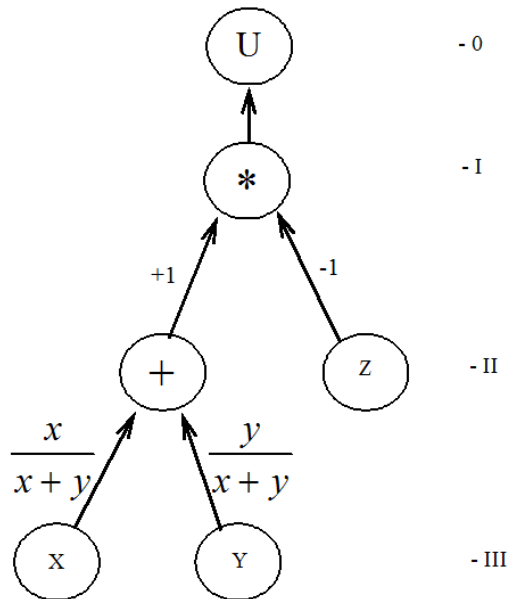


Рисунок 2.2 – Граф обчислювального процесу для виразу $u=(x+y)z$

Вхідні дані (числа та їх похибки) беруть участь ще в одній операції – множення (I рівень графа), куди вони передаються, помножуючись на коефіцієнти +1 і +1, тоді, з урахуванням похибки округлення, маємо

$$\delta_u = +1 \left(\frac{x}{x+y} \cdot \delta_x + \frac{y}{x+y} \cdot \delta_y + r_+ \right) + 1 \cdot \delta_z + r^*.$$

Простішим способом складання формули оцінювання кінцевої похибки є рух від кореня графа до його гілок (метод зверху вниз). Наступна послідовність формул ілюструє ідею цього методу: $\delta_u = +1(\quad) + 1 \cdot \delta_z + r^*$.

$$\delta_u = \left(\frac{x}{x+y} \cdot \delta_x + \frac{y}{x+y} \cdot \delta_y + r_+ \right) + \delta_z + r^*. \quad (2.25)$$

Оскільки природа похибок округлення і подання чисел в комп'ютері одна – обмеженість її розрядної сітки, то для відносних похибок можна покласти $\delta_x = \delta_y = \delta_z = r^* = r_+ = \delta$. Тоді формулу (2.25) можна спростити:

$$\delta_u = \delta \left[\left(\left| \frac{x}{x+y} \right| + \left| \frac{y}{x+y} \right| + 1 \right) + 1 + 1 \right] =$$

$$\left| \text{При } x>0 \text{ та } y>0 \left| \frac{x}{x+y} \right| + \left| \frac{y}{x+y} \right| \leq 1 \right| \leq 4\delta. \quad (2.26)$$

Величину відносної похибки округлення методом «відкидання» молодших розрядів можна оцінити $\delta \leq 10^{-t+1}$ (t – розрядність розмірної сітки комп'ютера), тоді $|\delta_u| \leq 4 \cdot 10^{-t+1}$.

2.6 Контрольні запитання

1. Наведіть основні джерела і класифікацію похибок під час розв'язання практичних задач.
2. Дайте означення абсолютної похибки.
3. Яку похибку називають відносною.
4. Які цифри числа називають значущими та правильними значущими? Наведіть приклад.
5. Наведіть та обґрунтуйте рівняння визначення похибок суми та різниці.
6. Наведіть та обґрунтуйте рівняння похибки добутку. Як визначити число правильних знаків добутку?
7. Наведіть та обґрунтуйте рівняння похибки частки. Визначте число правильних знаків частки.
8. Наведіть рівняння обчислення похибок степеня і кореня.
9. Поясніть, з якою метою застосовують графи обчислювальних процесів.
10. Наведіть графи обчислювальних процесів основних арифметичних операцій.
11. Наведіть граф довільного обчислювального процесу і обрахуйте загальну похибку результату.

2.7 Завдання

Задача 2.1. Дано ряд $\sum_{n=0}^{\infty} a_n$ (таблиця 2.1). Знайти суму ряду аналітично. Обчислити значення часткових сум ряду $S(N) = \sum_{n=0}^N a_n$ і знайти значення похибок при $N = 10, 10^2, 10^3, 10^4, 10^5, 10^6$.

Порядок розв'язування задачі

1. Знайти суму ряду S аналітично як межу часткових сум ряду.
2. Використовуючи функцію $S(N) = \sum_{n=0}^N a_n$, обчислити значення часткових сум ряду при вказаних значеннях N .
3. Для кожного N обчислити величину абсолютної похибки $|S(N) - S|$ і визначити кількість правильних цифр в $S(N)$.

4. Навести результати у вигляді гістограми.

Таблиця 2.1 – Варіанти завдань для задачі 2.1

№	a_n	№	a_n	№	a_n
1	$\frac{2}{n^2 + 5n + 6}$	11	$\frac{60}{11(n^2 + 12n + 35)}$	21	$\frac{24}{7(n^2 + 8n + 15)}$
2	$\frac{36}{11(n^2 + 5n + 4)}$	12	$\frac{144}{5(n^2 + 6n + 8)}$	22	$\frac{36}{n^2 + 5n + 4}$
3	$\frac{9}{n^2 + 7n + 12}$	13	$\frac{36}{n^2 + 7n + 10}$	23	$\frac{46}{n^2 + 5n + 6}$
4	$\frac{48}{5(n^2 + 6n + 8)}$	14	$\frac{48}{n^2 + 8n + 15}$	24	$\frac{96}{n^2 + 9n + 20}$
5	$\frac{48}{5(n^2 + 6n + 5)}$	15	$\frac{20}{n^2 + 4n + 3}$	25	$\frac{60}{n^2 + 6n + 8}$
6	$\frac{72}{5(n^2 + 6n + 8)}$	16	$\frac{32}{n^2 + 5n + 6}$	26	$\frac{72}{n^2 + 7n + 10}$
7	$\frac{24}{n^2 + 8n + 15}$	17	$\frac{144}{n^2 + 18n + 80}$	27	$\frac{24}{n^2 + 4n + 3}$
8	$\frac{32}{n^2 + 9n + 20}$	18	$\frac{24}{n^2 + 4n + 3}$	28	$\frac{96}{n^2 + 8n + 15}$
9	$\frac{216}{7(n^2 + 8n + 15)}$	19	$\frac{180}{n^2 + 20n + 99}$	29	$\frac{72}{n^2 + 6n + 8}$
10	$\frac{84}{13(n^2 + 14n + 48)}$	20	$\frac{112}{15(n^2 + 16n + 63)}$	30	$\frac{12}{5(n^2 + 6n + 8)}$

Задача 2.2. Обчислити значення машинних нуля та нескінченності у режимах одинарної, подвійної і розширеної точності.

Задача 2.3. Дано квадратне рівняння $x^2 + bx + c = 0$. Передбачається, що один з коефіцієнтів рівняння (в індивідуальному варіанті у таблиці 2.2 позначений *) отриманий в результаті округлення. Виконати оцінювання похибок коренів залежно від похибки коефіцієнта. Обчислити корені рівняння при декількох різних значеннях коефіцієнта в межах заданої точності. Порівняти отримані результати.

Задача 2.4. Скласти програму, що моделює обчислення на комп'ютері з обмеженою розрядністю m . Розв'язати задачу 1.1 для випадку $N=10000$, використовуючи цю програму. Скласти графік залежності похибки від кількості розрядів $m = 4, 5, \dots, 8$.

Таблиця 2.2 – Варіанти завдань для задачі 2.3

№	Коефіцієнти	№	Коефіцієнти	№	Коефіцієнти
1	$b^* = -39.6$ $c = -716.85$	2	$b = 27.4$ $c^* = 187.65$	3	$b^* = 37.4$ $c = 187.65$
4	$b = -30.9$ $c^* = 238.7$	5	$b^* = -3.29$ $c = 2.706$	6	$b = -3.29$ $c^* = 2.706$
7	$b^* = -39.6$ $c = -716.85$	8	$b = 27.4$ $c^* = 187.65$	9	$b^* = 37.4$ $c = 187.65$
10	$b = -30.9$ $c^* = 238.7$	11	$b^* = -3.29$ $c = 2.706$	12	$b = -3.29$ $c^* = 2.706$
13	$b^* = -39.6$ $c = -716.85$	14	$b = 27.4$ $c^* = 187.65$	15	$b^* = 37.4$ $c = 187.65$
16	$b = -30.9$ $c^* = 238.7$	17	$b^* = -3.29$ $c = 2.706$	18	$b = -3.29$ $c^* = 2.706$
19	$b^* = -39.6$ $c = -716.85$	20	$b = 27.4$ $c^* = 187.65$	21	$b^* = 37.4$ $c = 187.65$
22	$b = -30.9$ $c^* = 238.7$	23	$b^* = -3.29$ $c = 2.706$	24	$b = -3.29$ $c^* = 2.706$
25	$b^* = -39.6$ $c = -716.85$	26	$b = 27.4$ $c^* = 187.65$	27	$b^* = 37.4$ $c = 187.65$
28	$b = -30.9$ $c^* = 238.7$	29	$b^* = -3.29$ $c = 2.706$	30	$b = -3.29$ $c^* = 2.706$

розширену матрицю \bar{A} системи (3.2): $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n}b_1 \\ a_{21} & a_{22} & \dots & a_{2n}b_2 \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn}b_m \end{bmatrix}$.

З визначення матриці системи A і розширеної матриці \bar{A} зрозуміло, що їх ранги $r(\bar{A})$ і $r(A)$ або рівні між собою, або ранг $r(\bar{A})$ на одиницю більший, ніж $r(A)$.

Питання про сумісність системи (3.1) вирішується за допомогою **теорему Кронекера – Капеллі**: *система лінійних рівнянь (3.1) сумісна тоді і тільки тоді, коли ранг розширеної матриці A дорівнює рангові матриці A , тобто коли $r(\bar{A}) = r(A)$.*

Наслідок 1. Якщо система (3.2) сумісна і ранг матриці системи $r(A)=r$ дорівнює кількості невідомих n , то система має єдиний розв’язок.

Наслідок 2. Якщо система (3.2) сумісна і ранг матриці системи $r(A)=r$ менший за кількість невідомих n , то система має незліченну кількість розв’язків.

3.2 Розв’язання довільних систем лінійних алгебраїчних рівнянь

Нехай

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases} \quad (3.3)$$

довільна СЛАР, де кількість рівнянь системи не дорівнює кількості n невідомих ($m \neq n$). Припустимо, що система (3.3) сумісна, тобто $r(A)=r(\bar{A})=r$, і $r \leq \min\{m, n\}$. Тоді в матрицях A і \bar{A} даної системи знайдуться r лінійно незалежних рядків, а інші $m - r$ рядків будуть їх лінійними комбінаціями. Перестановкою рівнянь можна дійти до того, що ці r лінійно незалежних рядків займуть перші r місць.

Звідси випливає, що кожне з останніх $m - r$ рівнянь системи (3.3) можна записати як суму перших r рівнянь (які називаються *лінійно незалежними* або *базисними*), взятих з деякими коефіцієнтами. Тоді система (3.3) еквівалентна такій системі r рівнянь з n невідомими:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots \\ a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rn}x_n = b_r. \end{cases} \quad (3.4)$$

Припустимо, що міnor r -го порядку, складений з коефіцієнтів при

$$\text{перших } r \text{ невідомих, відмінний від нуля: } M_r = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} & \dots & a_{2r} \\ \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rr} \end{vmatrix} \neq 0,$$

тобто є базисним мінором. Тоді невідомі, з коефіцієнтів при яких складається базисний мінор, називаються *базисними невідомими*, а інші $n-r$ вільними невідомими.

У кожному з рівнянь системи (3.4) перенесемо у праву частину всі члени з вільними невідомими $x_{r+1}, x_{r+2}, \dots, x_n$. Тоді одержимо систему

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1r}x_r = b_1 - a_{1,r+1}x_{r+1} - a_{1,r+2}x_{r+2} - \dots - a_{1n}x_n, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2r}x_r = b_2 - a_{2,r+1}x_{r+1} - a_{2,r+2}x_{r+2} - \dots - a_{2n}x_n, \\ \dots \\ a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rr}x_r = b_r - a_{r,r+1}x_{r+1} - a_{r,r+2}x_{r+2} - \dots - a_{rn}x_n, \end{cases} \quad (3.5)$$

яка містить r рівнянь з r базисними невідомими. Оскільки визначник системи (3.5) є базисним мінором $M_r \neq 0$, то система (3.5) має єдиний розв'язок відносно базисних невідомих $x_{r+1}, x_{r+2}, \dots, x_n$, який можна знайти, наприклад, за формулами Крамера [3, 4]. Якщо вільним невідомим $x_{r+1}, x_{r+2}, \dots, x_n$ призначити довільні числові значення $x_{r+1} = c_1, x_{r+2} = c_2, \dots, x_n = c_{n-r}$, одержимо загальний розв'язок вихідної системи (3.3).

Приклад. Розв'язати систему трьох рівнянь з чотирма невідомими

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

Досліджуємо систему на сумісність. Складемо матриці A і \bar{A} :

$$A = \begin{bmatrix} 3 & -2 & 5 & 4 \\ 6 & -4 & 4 & 3 \\ 9 & -6 & 3 & 2 \end{bmatrix}, \quad \bar{A} = \begin{bmatrix} 3 & -2 & 5 & 4 & 2 \\ 6 & -4 & 4 & 3 & 3 \\ 9 & -6 & 3 & 2 & 4 \end{bmatrix}.$$

Визначимо ранги матриць, застосувавши елементарні перетворення:

- а) 1-й і 2-й стовпці пропорційні, один з них (2-й) відкидаємо;
- б) 1-й стовпець помножимо на 1/3;
- в) 1-й стовпець послідовно помножимо на (-5) і (-4) і додамо відповідно до 2-го і 3-го стовпців;

г) в отриманій матриці два стовпці (2-й і 3-й) пропорційні; один з них (3-й) відкинемо, а 2-й помножимо на $(-1/6)$;

д) 1-ий рядок послідовно помножимо на (-2) і (-3) і додамо відповідно до 2-го і 3-го рядків;

е) 2-й рядок помножимо на (-2) , додамо до 3-го і відкинемо отриманий нульовий рядок.

Маємо:

$$A = \begin{bmatrix} 3 & -2 & 5 & 4 \\ 6 & -4 & 4 & 3 \\ 9 & -6 & 3 & 2 \end{bmatrix} \underset{a)}{\sim} \begin{bmatrix} 3 & 5 & 4 \\ 6 & 4 & 3 \\ 9 & 3 & 2 \end{bmatrix} \underset{б)}{\sim} \begin{bmatrix} 1 & 5 & 4 \\ 2 & 4 & 3 \\ 3 & 3 & 2 \end{bmatrix} \underset{в)}{\sim} \begin{bmatrix} 1 & 0 & 0 \\ 2 & -6 & -5 \\ 3 & -12 & -10 \end{bmatrix} \underset{г)}{\sim} \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{bmatrix} \underset{д)}{\sim} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 2 \end{bmatrix} \underset{е)}{\sim} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Очевидно, що ранг останньої матриці дорівнює 2: $r(A) = 2$.

Матрицю \bar{A} перетворимо аналогічним чином:

$$\bar{A} = \begin{bmatrix} 3 & -2 & 5 & 4 & 2 \\ 6 & -4 & 4 & 3 & 3 \\ 9 & -6 & 3 & 2 & 4 \end{bmatrix} \underset{а)}{\sim} \begin{bmatrix} 3 & 5 & 4 & 2 \\ 6 & 4 & 3 & 3 \\ 9 & 3 & 2 & 4 \end{bmatrix} \underset{б)}{\sim} \begin{bmatrix} 1 & 5 & 4 & 2 \\ 2 & 4 & 3 & 3 \\ 3 & 3 & 2 & 4 \end{bmatrix} \underset{в)}{\sim} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & -6 & -5 & -1 \\ 3 & -12 & -10 & -2 \end{bmatrix} \underset{г)}{\sim} \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & -2 \end{bmatrix} \underset{д)}{\sim} \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ 0 & -2 \end{bmatrix} \underset{е)}{\sim} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \underset{ж)}{\sim} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Отже, $r(\bar{A}) = 2$. Оскільки, $r(A) = r(\bar{A}) = 2$, то дана система сумісна. Оскільки ранг системи дорівнює 2, максимальний порядок мінора, відмінного від нуля, дорівнює 2 і система має два базисних невідомих. Знайдемо будь-який відмінний від нуля мінор другого порядку. Таким, наприклад, є мінор $M_2 = \begin{vmatrix} 3 & 4 \\ 4 & 3 \end{vmatrix} \neq 0$, утворений коефіцієнтами при невідомих x_3 і x_4 .

Отже, невідомі x_3 і x_4 можна вважати базисними, а невідомі x_1 і x_2 – вільними. Система (*) еквівалентна такій:

$$\begin{cases} 3x_1 - 2x_2 + 5x_3 + 4x_4 = 2, \\ 6x_1 - 4x_2 + 4x_3 + 3x_4 = 3. \end{cases} \quad (**)$$

Переносимо вільні невідомі в праву частину:

$$\begin{cases} 5x_3 + 4x_4 = 2 - 3x_1 + 2x_2, \\ 4x_3 + 3x_4 = 3 - 6x_1 + 4x_2. \end{cases} \quad (***)$$

Розв'язуємо систему (3.5) за формулами Крамера [3, 4]:

$$d = \begin{vmatrix} 5 & 4 \\ 4 & 3 \end{vmatrix} = 15 - 16 = -1;$$

$$d_3 = \begin{vmatrix} 2 - 3x_1 + 2x_2 & 4 \\ 3 - 6x_1 + 4x_2 & 3 \end{vmatrix} = 3(2 - 3x_1 + 2x_2) - 4(3 - 6x_1 + 4x_2) = -6 + 15x_1 - 10x_2;$$

$$d_4 = \begin{vmatrix} 5 & 2 - 3x_1 + 2x_2 \\ 4 & 3 - 6x_1 + 4x_2 \end{vmatrix} = 5(3 - 6x_1 + 4x_2) - 4(2 - 3x_1 + 2x_2) = 7 - 18x_1 + 12x_2;$$

$$x_3 = d_3 / d = 6 - 15x_1 + 10x_2, \quad x_4 = d_4 / d = -7 + 18x_1 - 12x_2.$$

Отриманий розв'язок, в якому базисні невідомі x_3 і x_4 виражені через вільні x_1 і x_2 , є загальним розв'язком системи (*). Підставивши в нього довільні значення для вільних невідомих, одержуємо різні окремі розв'язки. Наприклад, якщо $x_1 = 0$, $x_2 = 0$, то $x_3 = 6$, $x_4 = (-7)$; якщо $x_1 = 1$, $x_2 = 2$, то $x_3 = 11$, $x_4 = 13$ тощо. Набори чисел $(0, 0, 6, -7)$ $(1, 2, 11, 13)$ тощо є окремими розв'язками системи (*).

3.3 Класифікація методів розв'язання систем лінійних алгебраїчних рівнянь та їх порівняльні характеристики

Методи розв'язання систем лінійних алгебраїчних рівнянь поділяють на дві групи: *точні* та *ітераційні* [1–4] (рис. 3.1).

Точними називають такі методи, які дозволяють знайти точний розв'язок СЛАР за допомогою виконання скінченної кількості арифметичних операцій у припущенні, що всі обчислення виконуються точно (без округлень), а коефіцієнти і вільні члени системи – точні числа. Проте на практиці всі обчислення виконуються з обмеженою кількістю десяткових розрядів, а ірраціональні коефіцієнти і вільні члени (якщо такі є) замінюються раціональними числами. Тому в процесі обчислень вдаються до округлень, а це означає, що розв'язки, котрі обчислюються за точними методами, є наближеними числами з певними похибками (похибками округлень).

До точних методів розв'язання СЛАР відносять: правило Крамера, метод Гаусса, метод головних елементів, Метод Гаусса – Жордана, схему Халецького, метод квадратних коренів тощо.

Метод Крамера має велике теоретичне значення, проте на практиці його застосовують дуже рідко, оскільки він дуже трудомісткий і під час розв'язування системи з великою кількістю невідомих потребує дуже багато машинного часу.

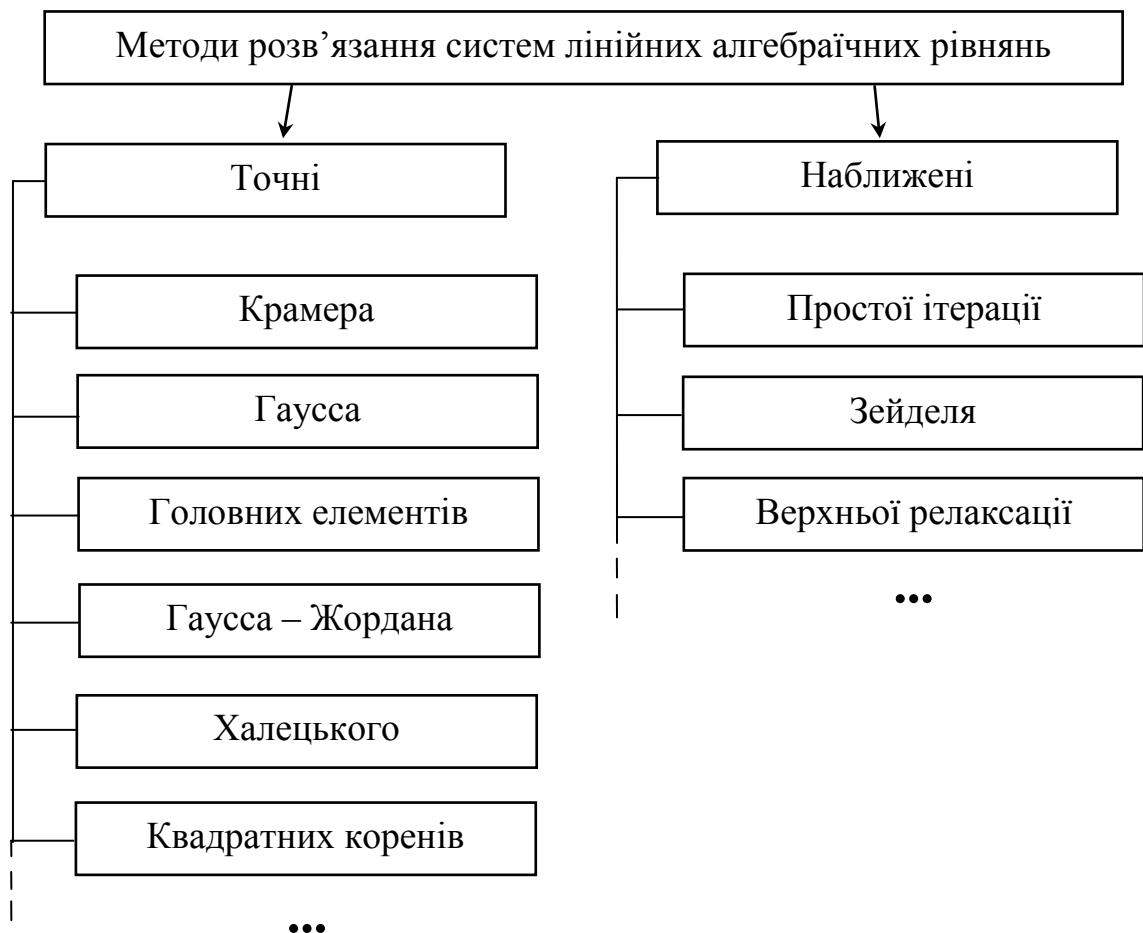


Рисунок 3.1 – Класифікація методів розв'язання систем лінійних алгебраїчних рівнянь

Метод Гаусса потребує виконання $2/3n^3 + O(n^2)$ арифметичних операцій (n – кількість невідомих) і займає n^2 комірок пам'яті. Його застосовують для розв'язання СЛАР з щільно заповненою матрицею.

Метод Гаусса – Жордана дозволяє послідовно виключити невідомі з усіх рівнянь системи, крім одного, і навести матрицю коефіцієнтів розв'язку але, порівняно з методом Гаусса, супроводжується збільшенням обсягу обчислень.

Якщо будь-який елемент провідного рядка дорівнює нулю, застосування методу Гаусса недоцільне. Такої складності можна запобігти, якщо змінити порядок, в якому розташовані рівняння системи. Максимальна точність досягається у випадку, коли провідний елемент має найбільше значення. Вибір за таким правилом головного елемента і покладено в основу методу головних елементів.

В багатьох випадках виникає необхідність розв'язання СЛАР, в якій матриця коефіцієнтів є незмінною, а стовпець вільних членів є змінним. Тоді доцільно застосовувати обчислення за схемою Халецького. В такому випадку один раз обчислюються значення елементів трикутних матриць і багаторазово розв'язуються нижня і верхня трикутні системи, котрі

потребують виконання n^2 операцій множення і ділення.

Якщо матриця коефіцієнтів системи симетрична – доцільно застосувати метод квадратних коренів, який потребує вдвічі меншої кількості комірок пам'яті і арифметичних операцій – відповідно $n/2 \cdot (n+1)$, $1/3n^2 + O(n)$.

Застосування точних методів для розв'язування СЛАР з великою кількістю невідомих є громіздким. Крім того, кількість невідомих може бути настільки великою, що коефіцієнти системи не завжди можна розташувати в оперативній пам'яті комп'ютера. Тоді системи розв'язують за допомогою ітераційних методів [1].

Ітераційними називають такі методи, які дозволяють знайти наближений розв'язок СЛАР із заздалегідь указаною точністю шляхом виконання скінченної кількості арифметичних операцій, хоч самі обчислення можуть виконуватися без округлень, а коефіцієнти і вільні члени системи бути точними числами. Розв'язуючи СЛАР ітераційними методами, крім похибок округлення слід враховувати також і похибку методу.

До ітераційних методів розв'язання СЛАР відносять: метод простої ітерації (або метод ітерації), метод Зейделя, метод верхньої релаксації тощо.

Далі детальніше розглянемо теоретичні та практичні питання розв'язання СЛАР ітераційними методами.

3.4 Ітераційні методи розв'язання систем лінійних алгебраїчних рівнянь

3.4.1 Метод ітерацій

Наближені методи розв'язання СЛАР дозволяють отримувати значення коренів системи з заданою точністю у вигляді межі послідовності декількох векторів. Процес побудови такої послідовності називається *ітераційним* (таким, що повторюється). Ефективність використання наближених методів залежить від вибору початкового вектора і швидкості збіжності процесу [1, 3, 4, 5].

Нехай дана система n лінійних рівнянь з n невідомими:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n. \end{cases} \tag{3.6}$$

Запишемо систему (3.6) у матричному вигляді:

$$Ax = b, \tag{3.7}$$

де

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Вважатимемо, що діагональні елементи $a_{ii} \neq 0$ ($i=1,2,\dots,n$), виразимо x_1 через перше рівняння системи, x_2 – через друге рівняння і т. д. В результаті отримаємо систему, еквівалентну системі (3.6):

$$\begin{cases} x_1 = \frac{b_{11}}{a_{11}} - \frac{a_{12}}{a_{11}}x_2 - \frac{a_{13}}{a_{11}}x_3 - \cdots - \frac{a_{1n}}{a_{11}}x_n; \\ x_2 = \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}}x_1 - \frac{a_{23}}{a_{22}}x_3 - \cdots - \frac{a_{2n}}{a_{22}}x_n; \\ \dots \\ x_n = \frac{b_n}{a_{nn}} - \frac{a_{n1}}{a_{nn}}x_1 - \frac{a_{n2}}{a_{nn}}x_2 - \cdots - \frac{a_{n,n-1}}{a_{nn}}x_{n-1}. \end{cases} \quad (3.8)$$

Позначимо $b_i / a_{ii} = \beta_i$; $-a_{ij} / a_{ii} = \alpha_{ij}$, де $i=1,2,\dots,n$; $j=1,2,\dots,n$. Тоді система (3.8) запишеться так:

$$\begin{cases} x_1 = \beta_1 + \alpha_{12}x_2 + \alpha_{13}x_3 + \dots + \alpha_{1n}x_n, \\ x_2 = \beta_2 + \alpha_{21}x_1 + \alpha_{23}x_3 + \dots + \alpha_{2n}x_n, \\ \dots \\ x_n = \beta_n + \alpha_{n1}x_1 + \alpha_{n2}x_2 + \dots + \alpha_{n,n-1}x_{n-1}. \end{cases} \quad (3.8')$$

Система (3.8') називається системою, зведеною до *нормального вигляду*.

Введемо позначення: $\alpha = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{bmatrix}$, $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$.

Запишемо систему (3.8') у матричній формі: $x = \beta + \alpha x$ або

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \quad (3.9)$$

Розв'яжемо систему (3.9) методом послідовних наближень. За нульове наближення візьмемо стовпець вільних членів:

$$\begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_n^{(0)} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \text{ — нульове наближення.}$$

Далі побудуємо матриці-стовпці

$$\begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ \vdots \\ x_n^{(1)} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{bmatrix} \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_n^{(0)} \end{bmatrix} \text{ — перше наближення;}$$

$$\begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ \vdots \\ x_n^{(2)} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{bmatrix} \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ \vdots \\ x_n^{(1)} \end{bmatrix} \text{ — друге наближення}$$

тощо. Взагалі, будь-яке $(k + 1)$ -те наближення розраховують за формулою

$$x^{(k+1)} = \beta + \alpha x^{(k)} \quad (k = 0, 1, \dots, n). \quad (3.10)$$

Якщо в послідовність наближень $x^{(0)}, x^{(1)}, \dots, x^{(k)}$ має границю $x = \lim_{k \rightarrow \infty} x^{(k)}$, то ця границя є розв'язком системи (3.8), оскільки з врахуванням властивостей границі $\lim_{k \rightarrow \infty} x^{(k+1)} = \beta + \alpha \lim_{k \rightarrow \infty} x^{(k)}$, тобто $x = \beta + \alpha x$.

Приклад 1. Методом ітерації з точністю до 10^{-1} розв'язати систему

$$\begin{cases} 8x_1 + x_2 + x_3 = 26, \\ x_1 + 5x_2 - x_3 = 7, \\ x_1 - x_2 + 5x_3 = 7. \end{cases}$$

1. Зведемо систему до нормального вигляду:

$$\begin{cases} x_1 = 3,25 - 0,125x_2 - 0,125x_3, \\ x_2 = 1,4 - 0,2x_1 + 0,2x_3, \\ x_3 = 1,4 - 0,2x_1 + 0,2x_2; \end{cases} \quad \alpha = \begin{bmatrix} 0 & -0,125 & -0,125 \\ -0,2 & 0 & 0,2 \\ -0,2 & 0,2 & 0 \end{bmatrix}; \quad \beta = \begin{bmatrix} 3,25 \\ 1,4 \\ 1,4 \end{bmatrix}.$$

2. Будуємо послідовні наближення. Нульове наближення:

$$\begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{bmatrix} = \begin{bmatrix} 3,25 \\ 1,4 \\ 1,4 \end{bmatrix}.$$

Перше наближення:

$$\begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} = \begin{bmatrix} 3,25 \\ 1,4 \\ 1,4 \end{bmatrix} + \begin{bmatrix} 0 & -0,125 & -0,125 \\ -0,2 & 0 & 0,2 \\ -0,2 & 0,2 & 0 \end{bmatrix} \begin{bmatrix} 3,25 \\ 1,4 \\ 1,4 \end{bmatrix} = \begin{bmatrix} 2,9 \\ 1,03 \\ 1,03 \end{bmatrix}.$$

Друге наближення:

$$\begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{bmatrix} = \begin{bmatrix} 3,25 \\ 1,4 \\ 1,4 \end{bmatrix} + \begin{bmatrix} 0 & -0,125 & -0,125 \\ -0,2 & 0 & 0,2 \\ -0,2 & 0,2 & 0 \end{bmatrix} \begin{bmatrix} 2,9 \\ 1,03 \\ 1,03 \end{bmatrix} = \begin{bmatrix} 2,992 \\ 1,026 \\ 1,026 \end{bmatrix}.$$

Третє наближення:

$$\begin{bmatrix} x_1^{(3)} \\ x_2^{(3)} \\ x_3^{(3)} \end{bmatrix} = \begin{bmatrix} 3,25 \\ 1,4 \\ 1,4 \end{bmatrix} + \begin{bmatrix} 0 & -0,125 & -0,125 \\ -0,2 & 0 & 0,2 \\ -0,2 & 0,2 & 0 \end{bmatrix} \begin{bmatrix} 2,992 \\ 1,026 \\ 1,026 \end{bmatrix} = \begin{bmatrix} 2,99 \\ 1,01 \\ 1,01 \end{bmatrix}.$$

Таким чином, $x_1 = 2,99$, $x_2 = 1,01$, $x_3 = 1,01$ і з точністю до 10^{-1} отримуємо $x_1 = 3,0$, $x_2 = 1,0$, $x_3 = 1,0$.

Приклад 2. Методом ітерації з точністю до 10^{-3} розв'язати систему

$$\begin{cases} 7,6x_1 + 0,5x_2 + 2,4x_3 = 1,9, \\ 2,2x_1 + 9,1x_2 + 4,4x_3 = 9,7, \\ -1,3x_1 + 0,2x_2 + 5,8x_3 = -1,4. \end{cases} \quad (*)$$

1. Зведемо систему до нормального вигляду:

$$\begin{cases} x_1 = \frac{1,9}{7,6} - \frac{0,5}{7,6}x_2 - \frac{2,4}{7,6}x_3, \\ x_2 = \frac{9,7}{9,1} - \frac{2,2}{9,1}x_1 - \frac{4,4}{9,1}x_3, \\ x_3 = \frac{-1,4}{5,8} + \frac{1,3}{5,8}x_1 - \frac{0,2}{5,8}x_2, \end{cases} \quad \text{або} \begin{cases} x_1 = 0,25 - 0,065x_2 - 0,3158x_3, \\ x_2 = 1,0659 - 0,2418x_1 - 0,4847x_3, \\ x_3 = -0,2414 + 0,2241x_1 - 0,3448x_2; \end{cases}$$

$$\alpha = \begin{bmatrix} 0 & -0,065 & -0,3158 \\ -0,2418 & 0 & -0,4847 \\ 0,2241 & -0,3448 & 0 \end{bmatrix}; \beta = \begin{bmatrix} 0,25 \\ -1,0659 \\ -0,2414 \end{bmatrix}.$$

Зазначимо, що лінійну систему можна звести до нормального вигляду також записавши коефіцієнти при x_1, x_2, x_3 у відповідних рівняннях системи (*) у вигляді kx , де k – число, близьке до коефіцієнта при відповідному невідомому і на яке легко розділити коефіцієнти при невідомих і вільні члени.

Наприклад:

$$\begin{aligned} 10x_1 &= 7,6x_1 + 2,4x_1 \quad (\text{у першому рівнянні}), \\ 10x_2 &= 9,1x_2 + 0,9x_2 \quad (\text{у другому рівнянні}), \\ 10x_3 &= 5,8x_3 + 4,2x_3 \quad (\text{у третьому рівнянні}). \end{aligned}$$

Перепишемо систему (*) так:

$$\begin{cases} 10x_1 = 1,9 + 2,4x_1 - 0,5x_2 - 2,4x_3, \\ 10x_2 = 9,7 - 2,2x_1 + 0,9x_2 - 4,4x_3, \\ 10x_3 = -1,4 + 1,3x_1 - 0,2x_2 + 4,2x_3; \\ \begin{cases} x_1 = 0,19 + 0,24x_1 - 0,05x_2 - 0,24x_3, \\ x_2 = 0,97 - 0,22x_1 + 0,09x_2 - 0,44x_3, \\ 10x_3 = -0,14 + 0,13x_1 - 0,02x_2 + 0,42x_3. \end{cases} \end{cases}$$

Матриця α і вектор β набувають вигляду

$$\alpha = \begin{bmatrix} 0,24 & -0,05 & -0,24 \\ -0,22 & 0,09 & -0,44 \\ 0,13 & -0,02 & 0,42 \end{bmatrix}, \quad \beta = \begin{bmatrix} 0,19 \\ 0,97 \\ -0,14 \end{bmatrix}.$$

2. Послідовно знаходимо

$$\begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \end{bmatrix} = \begin{bmatrix} 0,19 \\ 0,97 \\ -0,14 \end{bmatrix};$$

$$\begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} = \begin{bmatrix} 0,19 \\ 0,97 \\ -0,14 \end{bmatrix} + \begin{bmatrix} 0,24 & -0,05 & -0,24 \\ -0,22 & 0,09 & -0,44 \\ 0,13 & -0,02 & 0,42 \end{bmatrix} \times \begin{bmatrix} 0,19 \\ 0,97 \\ -0,14 \end{bmatrix} = \begin{bmatrix} 0,2207 \\ 1,0771 \\ -0,1935 \end{bmatrix};$$

$$\begin{bmatrix} x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{bmatrix} = \begin{bmatrix} 0,19 \\ 0,97 \\ -0,14 \end{bmatrix} + \begin{bmatrix} 0,24 & -0,05 & -0,24 \\ -0,22 & 0,09 & -0,44 \\ 0,13 & -0,02 & 0,42 \end{bmatrix} \times \begin{bmatrix} 0,2207 \\ 1,0771 \\ -0,1935 \end{bmatrix} = \begin{bmatrix} 0,2359 \\ 1,1034 \\ -0,2141 \end{bmatrix}.$$

Таким чином, з точністю до 10^{-3} отримаємо

$$x_1 = 0,236; x_2 = 1,103; x_3 = -0,214.$$

3.4.2 Умови збіжності ітераційного процесу

Нехай дана зведена до нормального вигляду система лінійних рівнянь $x = \beta + \alpha x$. Умова збіжності ітераційного процесу така: *якщо сума модулів елементів рядків або стовпців менша одиниці, то процес ітерації для даної системи збігається до єдиного розв'язку незалежно від вибору початкового вектора.*

Отже, умову збіжності можна записати так:

$$\sum_{j=1}^n |\alpha_{ij}|, < 1 \quad (i=1,2,\dots,n) \quad \text{або} \quad \sum_{i=1}^n |\alpha_{ij}|, < 1 \quad (j=1,2,\dots,n).$$

Приклад. Для системи

$$\begin{cases} 8x_1 + x_2 + x_3 = 26, \\ x_1 + 5x_2 - x_3 = 7, \\ x_1 - x_2 + 5x_3 = 7 \end{cases} \quad \text{або} \quad \begin{cases} x_1 = 3,25 - 0,125x_2 - 0,125x_3, \\ x_2 = 1,4 - 0,2x_1 + 0,2x_3, \\ 10x_3 = 1,4 - 0,2x_1 + 0,2x_2. \end{cases}$$

Ітераційний процес збіжний, оскільки $\alpha = \begin{bmatrix} 0 & -0,125 & -0,125 \\ -0,2 & 0 & 0,2 \\ -0,2 & 0,2 & 0 \end{bmatrix}$ і

$$|\alpha_{11}| + |\alpha_{21}| + |\alpha_{31}| = 0,2 + 0,2 = 0,4 < 1;$$

$$|\alpha_{12}| + |\alpha_{22}| + |\alpha_{32}| = 0,125 + 0,2 = 0,325 < 1;$$

$$|\alpha_{13}| + |\alpha_{23}| + |\alpha_{33}| = 0,125 + 0,2 = 0,325 < 1.$$

Аналогічно можна було б перевірити виконання збіжності, взявши суми модулів елементів рядків.

Процес ітерації свідомо збігається, якщо елементи матриці α задовольняють нерівності $|\alpha_{ij}| < 1/n$; де n – кількість невідомих даної системи.

В даному прикладі $n = 3$ і всі елементи $|\alpha_{ij}| < 1/3$.

Збіжність ітераційного процесу пов'язана з нормами матриці α такими співвідношеннями:

Якщо виконується одна з умов:

$$\|\alpha\|_1 = \max_i \sum_{j=1}^n |\alpha_{ij}| < 1$$

або

$$\|\alpha\|_2 = \max_j \sum_{i=1}^n |\alpha_{ij}| < 1,$$

або

$$\|\alpha\|_3 = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |\alpha_{ij}|^2} < 1,$$

то процес ітерації лінійної системи зводиться до єдиного розв'язку.

Так, у розглянутому вище прикладі норма

$$\|\alpha\|_2 = \max(0,4; 0,325; 0,325) = 0,4 < 1,$$

тобто ітераційний процес збігається.

3.4.3 Оцінювання похибки методу ітерацій

Якщо задано допустиму похибку обчислень ε і x_i – вектор точних значень невідомих лінійної системи, а $x_i^{(k)}$ є k -е наближене значення невідомих, обчислених методом ітерації, то для оцінювання похибки $\|x_i - x_i^{(k)}\| \leq \varepsilon$ методу застосовується формула

$$\|x_i - x_i^{(k)}\| \leq \frac{\|\alpha\|^{k+1}}{1 - \|\alpha\|} \cdot \|\beta\|, \quad (3.11)$$

де $\|\alpha\|$ – одна із трьох норм матриці α ; $\|\beta\|$ – та ж норма вектора β , а k – кількість ітерацій, необхідна для досягнення заданої точності. При цьому передбачається, що послідовні наближення $x_i^{(j)}$ (де $j=0, 1, \dots; k, i=1, 2, \dots, n$) обчислюються точно, в них відсутні похибки округлення.

Приклад. Показати, що для системи
$$\begin{cases} 9,9x_1 - 1,5x_2 + 2,6x_3 = 0, \\ 0,4x_1 + 13,6x_2 - 4,2x_3 = 8,2, \\ 0,7x_1 + 0,4x_2 + 7,1x_3 = -1,3 \end{cases}$$

ітераційний процес збігається і визначити, скільки ітерацій потрібно виконати, щоб знайти невідомі з точністю до 10^{-4} .

Розв'язання

1. Зведемо систему до нормального вигляду

$$\begin{cases} 10x_1 = 0,1x_1 + 1,5x_2 - 2,6x_3, \\ 20x_2 = -0,4x_1 + 6,4x_2 + 4,2x_3 + 8,2, \\ 10x_3 = 0,7x_1 - 0,4x_2 + 2,9x_3 - 1,3 \end{cases}$$

або
$$\begin{cases} x_1 = 0,01x_1 + 0,15x_2 - 0,26x_3, \\ x_2 = -0,02x_1 + 0,32x_2 + 0,21x_3 + 0,41, \\ x_3 = -0,07x_1 - 0,04x_2 + 0,29x_3 - 0,13. \end{cases}$$

2. Матриця системи

$$\alpha = \begin{bmatrix} 0,01 & 0,15 & -0,26 \\ -0,02 & 0,32 & 0,21 \\ -0,07 & -0,04 & 0,29 \end{bmatrix}.$$

Застосовуючи норму $\|\alpha\|_2$, отримаємо:

$\|\alpha\|_2 = \max(0,1; 0,51; 0,76) = 0,76 < 1$. Отже, ітераційний процес для даної системи збігається.

3. Маємо $\beta = \begin{bmatrix} 0 \\ 0,41 \\ -0,13 \end{bmatrix}$, $\|\beta\|_2 = 0 + 0,41 + 0,13 = 0,54$.

4. Застосовуючи формулу (3.11), знаходимо

$$x_1^{(1)} = 0,19 + 0,24 \cdot 0,19 - 0,05 \cdot 0,97 - 0,24 \cdot (-0,14) = 0,2207,$$

$$x_2^{(1)} = 0,97 - 0,22 \cdot 0,2207 + 0,09 \cdot 0,97 - 0,44 \cdot (-0,14) = 1,0703,$$

$$x_3^{(1)} = -0,14 + 0,13 \cdot 0,2207 - 0,02 \cdot 1,0703 + 0,42 \cdot (-0,14) = -0,1915.$$

Другі наближення:

$$x_1^{(2)} = 0,19 + 0,24 \cdot 0,2207 - 0,05 \cdot 1,0703 - 0,24 \cdot (-0,1915) = 0,2354,$$

$$x_2^{(2)} = 0,97 - 0,22 \cdot 0,2354 + 0,09 \cdot 1,0703 - 0,44 \cdot (-0,1915) = 1,0988,$$

$$x_3^{(2)} = -0,14 + 0,13 \cdot 0,2354 - 0,02 \cdot 1,0988 + 0,42 \cdot (-0,1915) = -0,2118.$$

Розв'язок цього прикладу наведено у табл. 3.1.

Таблиця 3.1 – Розв'язок прикладу

№ ітерації	x_1	x_2	x_3	№ ітерації	x_1	x_2	x_3
0	0,19	0,97	-0,14	5	0,2467	1,1138	-0,2237
1	0,2207	1,0703	-0,1915	6	0,2472	1,1143	-0,2241
2	0,2354	1,0988	-0,2118	7	0,2474	1,1145	-0,2243
3	0,2424	1,1088	-0,2196	8	0,2475	1,1145	-0,2243
4	0,2454	1,1124	-0,2226				

Побудова ітерації закінчується, коли досягнуто заданого ступеня точності, тобто коли ми отримали однакові значення в двох ітераціях підряд. В даному прикладі це ітерації 7 і 8.

Остаточна відповідь: $x_1 \approx 0,248$; $x_2 \approx 1,114$; $x_3 \approx -0,224$.

3.4.5 Умови збіжності процесу Зейделя

Процес Зейделя для лінійної системи $x = \beta + \alpha x$ так, як і процес послідовних наближень, збігається до єдиного розв'язку при будь-якому виборі початкового наближення, якщо хоч би одна із норм матриці α менша одиниці.

Процес Зейделя часто збігається до єдиного розв'язку швидше за метод процесу простої ітерації.

Приклад 2. Перевірити, чи збігається процес Зейделя для системи, яка розглянута в прикладі 1.

Розв'язання

1. Після зведення системи до нормального вигляду отримаємо матрицю

$$\alpha = \begin{bmatrix} 0,24 & -0,05 & -0,24 \\ -0,22 & 0,09 & -0,44 \\ 0,13 & -0,02 & 0,42 \end{bmatrix}.$$

2. Знайдемо $\|\alpha\|_1 = \max_i \sum_j |\alpha_{ij}| < 1 = \max(0,53; 0,75; 0,57) = 0,75 < 1$. Отже, процес ітерації для даної системи збігається до єдиного розв'язку, не дивлячись на то, що $\|\alpha\|_1 = \max_i \sum_j |\alpha_{ij}| < 1 = \max(0,59; 0,16; 1,1) = 1,1 > 1$.

Оцінювання похибки процесу Зейделя.

Нехай дана лінійна система $x = \beta + \alpha x$. Якщо x_i – точні значення її невідомих, а x_i^k – k -е наближення, обчислене за методом Зейделя, то для оцінювання похибки цього методу застосовується формула

$$\|x - x^{(k)}\|_1 \leq \frac{\|\alpha\|_1^{(k)}}{1 - \|\alpha\|_1} \|x^{(1)} - x^{(0)}\|_1. \quad (3.13)$$

Приклад. Підрахувати, скільки ітерацій за методом Зейделя необхідно виконати, щоб з точністю до 10^{-4} знайти розв'язок системи

$$\begin{cases} 9,9x_1 - 1,5x_2 + 2,6x_3 = 0, \\ 0,4x_1 + 13,6x_2 - 4,2x_3 = 8,2, \\ 0,7x_1 + 0,4x_2 + 7,1x_3 = -1,3. \end{cases}$$

Розв'язання. 1. Зведемо систему до нормального вигляду:

$$\begin{cases} x_1 = 0,01x_1 + 0,15x_2 - 0,26x_3, \\ x_2 = 0,41 - 0,02x_1 + 0,32x_2 + 0,21x_3, \\ x_3 = -0,13 - 0,07x_1 - 0,04x_2 + 0,29x_3. \end{cases}$$

2. За нульові наближення візьмемо стовпець вільних членів $x_1^{(0)} = 0$; $x_2^{(0)} = 0,41$; $x_3^{(0)} = -0,13$ і обчислимо перші наближення:

$$x_1^{(1)} = 0,01 \cdot 0 + 0,15 \cdot 0,41 - 0,26 \cdot (-0,13) = 0,0953;$$

$$x_2^{(1)} = 0,41 - 0,02 \cdot 0,0953 + 0,32 \cdot 0,41 + 0,21 \cdot (-0,13) = 0,5120;$$

$$x_3^{(1)} = -0,13 - 0,07 \cdot 0,0953 - 0,04 \cdot 0,5120 + 0,29 \cdot (-0,13) = -0,1948.$$

3. Матриця $\alpha = \begin{bmatrix} 0,01 & 0,15 & -0,26 \\ -0,02 & 0,32 & 0,21 \\ -0,07 & -0,04 & 0,29 \end{bmatrix}$.

Отже, $\|\alpha\|_1 = \max(0,42; 0,55; 0,40) = 0,55$. Оскільки

$$x^{(0)} \begin{bmatrix} 0 \\ 0,41 \\ -0,13 \end{bmatrix} \text{ і } x^{(1)} = \begin{bmatrix} 0,0953 \\ 0,5120 \\ -0,1948 \end{bmatrix}, \text{ то } (x^{(1)} - x^{(0)}) = \begin{bmatrix} 0,0953 \\ 0,1120 \\ -0,0648 \end{bmatrix} = 0,1120$$

4. За формулою (3.13) визначаємо k :

$$\begin{aligned} 10^{-4} &\leq \frac{0,55^k}{0,45} \cdot 0,1120; \quad 10^{-4} \cdot 0,45 \leq 0,55^k \cdot 0,1120; \\ &-4 \lg 10 + \lg 0,45 \leq k \lg 0,55 + \lg 0,1120; \\ &-4 - 0,3468 \leq k(-0,2596 - 0,9508); \quad k \geq \frac{4,3468}{1,2104} = 3,59; \quad k = 4. \end{aligned}$$

Аналогічно можна виконувати оцінювання методу Зейделя за нормою $\|\alpha\|_2$.

3.4.6 Зведення системи лінійних рівнянь до вигляду, зручного для ітерацій

Процеси послідовних наближень і Зейделя для лінійної системи $x = \beta + \alpha x$ збігаються до єдиного розв'язку незалежно від вибору початкового вектора, якщо

$$\sum_{j=1}^n |\alpha_{ij}| < 1 \quad (i = 1, 2, \dots, n) \text{ або } \sum_{i=1}^n |\alpha_{ij}| < 1 \quad (j = 1, 2, \dots, n).$$

Отже, для збіжності вищевказаних ітераційних процесів достатньо, щоб значення елементів α_{ij} матриці α при $i \neq j$ були невеликими за абсолютною величиною. Це рівнозначно тому, що коли для лінійної системи $Ax = b$ модулі діагональних коефіцієнтів кожного рівняння системи більші за суми модулів всіх наступних коефіцієнтів (не враховуючи вільних членів), то ітераційні процеси для цієї системи сходяться, тобто якщо дана система $\sum_{j=1}^n a_{ij} x_j = b_i \quad (i = 1, 2, \dots, n)$, причому $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$, то процеси простої ітерації і Зейделя для цієї системи збігаються [4].

Застосовуючи елементарні перетворення, лінійну систему $Ax = b$ можна змінити такою еквівалентною системою $x = \beta + \alpha x$, для якої умови збіжності будуть виконані.

Приклад. Звести до вигляду, зручного для ітерацій, таку систему:

$$\begin{cases} 0,9x_1 + 2,7x_2 - 3,8x_3 = 2,4, & (A) \\ 2,5x_1 + 5,8x_2 - 0,5x_3 = 3,5, & (B) \\ 4,5x_1 - 2,1x_2 + 3,2x_3 = -1,2. & (B) \end{cases}$$

Розв'язання

1. Із заданої системи виділяємо рівняння з коефіцієнтами, модулі яких більші за суми модулів наступних коефіцієнтів системи. Кожне виділене рівняння вписуємо в такий рядок нової системи, щоб найбільший за модулем коефіцієнт виявився діагональним.

У рівнянні (B) коефіцієнт при x_2 за модулем більший за суму модулів наступних коефіцієнтів. Візьмемо рівняння (B) за друге рівняння нової системи:

$$2,5x_1 + 5,8x_2 - 0,5x_3 = 3,5. \quad (II)$$

2. Із невикористаних рівнянь системи, що залишилися, утворюємо лінійно незалежні між собою комбінації. Так, за перше рівняння нової системи можна взяти лінійну комбінацію $(2B)+(A)$, тоді матимемо

$$9,9x_1 - 1,5x_2 + 2,6x_3 = 0. \quad (I)$$

За третє рівняння нової системи можна прийняти лінійну комбінацію $(2A) - (B)$, тобто

$$0,7x_1 + 0,4x_2 + 7,1x_3 = -1,3. \quad (III)$$

3. Як результат отримаємо перетворену систему лінійних рівнянь (I), (II), (III), яка еквівалентна вихідній і задовольняє умови збіжності ітераційного процесу:

$$\text{цінного процесу: } \begin{cases} 9,9x_1 - 1,5x_2 + 2,6x_3 = 0, \\ 2,5x_1 + 5,8x_2 - 0,5x_3 = 3,5, \\ 0,7x_1 + 0,4x_2 + 7,1x_3 = -1,3. \end{cases}$$

Зводячи цю систему до нормального вигляду, отримаємо

$$\begin{cases} x_1 = 0,1x_1 + 0,15x_2 - 0,26x_3, \\ x_2 = 0,35x_1 - 0,21x_2 + 0,42x_3 + 0,05, \\ x_3 = -0,13x_1 - 0,07x_2 - 0,04x_3 + 0,29; \end{cases} \quad \alpha = \begin{bmatrix} 0,10 & 0,15 & -0,26 \\ 0,35 & -0,21 & 0,42 \\ -0,13 & -0,07 & -0,04 \end{bmatrix};$$

$$\|\alpha\|_2 = \max(0,58; 0,43; 0,72) = 0,72 < 1.$$

Тепер слід лише розв'язати систему одним з ітераційних методів.

3.5 Контрольні запитання

1. Що називають розв'язком системи лінійних алгебраїчних рівнянь?
2. Наведіть означення сумісної та несумісної СЛАР.
3. Яку СЛАР називають визначеною, а яку невизначеною?
4. Які СЛАР називають еквівалентними?
5. Перерахуйте елементарні перетворення над СЛАР.
6. Поясніть, яку СЛАР називають погано обумовленою, а яку добре обумовленою?
7. Наведіть теорему Кронекера–Капеллі та два її наслідки.
8. В якому випадку СЛАР не має розв'язку, має один розв'язок або має більше одного розв'язку?
9. Наведіть класифікацію методів розв'язання СЛАР.
10. Які методи розв'язання СЛАР називають точними? Перерахуйте декілька точних методів.
11. Наведіть загальний алгоритм розв'язання СЛАР за формулами Крамера. Які переваги та недоліки цього методу?
12. Наведіть алгоритм розв'язання СЛАР за методом Гаусса. Які переваги та недоліки цього методу?
13. Наведіть алгоритм розв'язання СЛАР за методом головних елементів. Чим даний метод відрізняється від методу Гаусса?
14. Наведіть алгоритм розв'язання СЛАР за методом Гаусса–Жордана. Які переваги та недоліки цього методу?
15. Наведіть алгоритм розв'язання СЛАР за схемою Халецького. В яких випадках доцільно застосовувати цю схему?
16. Наведіть алгоритм розв'язання СЛАР за методом квадратних коренів. Коли застосовують даний метод та в чому його переваги?
17. Які методи розв'язання СЛАР називають наближеними? Наведіть приклади наближених методів.
18. Наведіть алгоритм розв'язання СЛАР за методом простої ітерації. Коли застосовують даний метод та в чому його переваги?
19. Наведіть умови збіжності ітераційного процесу розв'язання СЛАР.
20. Яким чином оцінюють похибки методу ітерацій?
21. Наведіть алгоритм розв'язання СЛАР за методом Зейделя. В чому його переваги та недоліки?
22. Наведіть умови збіжності процесу Зейделя.
23. Наведіть рівняння оцінювання похибки розв'язання СЛАР методом Зейделя.
24. Наведіть загальний алгоритм зведення системи лінійних рівнянь до вигляду, який зручний для ітерацій.
25. Наведіть порівняльний аналіз відомих Вам методів розв'язання СЛАР.

3.6 Завдання

Задача 1. Напишіть програму розв'язання системи $AX=B$ (таблиця 3.2) методом простої ітерації з точністю до 11 знаків після коми. За початкове наближення прийміть, наприклад, розв'язок, знайдений методом Гаусса та округлений до двох знаків після коми. Підрахуйте теоретично кількість ітерацій, що слід виконати для досягнення заданої точності і порівняйте її з практично отриманими результатами.

Таблиця 3.2 – Варіанти завдань до задачі 1

№	Матриця А				Матриця В	№	Матриця А				Матриця В
1	0.63	1	0.11	0.34	2.08	2	1	0.47	-0.11	0.55	1.33
	0.17	1.18	-0.45	0.11	0.17		0.42	1	0.35	0.17	1.29
	0.31	-0.15	1.17	-2.35	1.28		-0.25	0.67	1	0.36	2.11
	0.58	0.21	-3.45	-1.18	0.05		0.54	-0.32	-0.74	1	0.10
3	0.77	0.04	-0.21	0.18	1.24	4	0.79	-0.12	0.34	0.16	-0.64
	-0.45	1.23	-0.06	0	-0.88		-0.34	1.18	-0.17	0.18	1.42
	-0.26	-0.34	1.11	0	0.62		-0.16	-0.34	0.85	0.31	-0.42
	-0.05	0.26	-0.34	1.12	-1.17		-0.12	0.26	0.08	0.75	0.83
5	-0.68	-0.18	0.02	0.21	-1.83	6	-0.58	-0.32	0.03	0	-0.44
	0.16	-0.88	-0.14	0.27	0.65		0.11	-1.26	-0.36	0	-1.42
	0.37	0.27	-1.02	-0.24	-2.23		0.12	0.08	-1.14	-0.24	0.83
	0.12	0.21	-0.18	-0.75	1.13		0.15	-0.35	-0.18	0	1.42
7	-0.83	0.31	-0.18	0.22	1.71	8	-0.87	0.27	-0.22	-0.18	-1.21
	-0.21	-0.67	0	0.22	-0.62		-0.21	-1	-0.45	0.18	0.33
	0.32	-0.18	-0.95	-0.19	0.89		0.12	0.13	-0.33	0.18	0.48
	0.12	0.28	-0.14	-1	-0.94		0.33	-0.41	0	-1	1.21
9	-0.81	-0.07	0.38	-0.21	0.81	10	-1	0.22	-0.11	0.31	-2.7
	-0.22	-0.92	0.11	0.33	0.64		0.38	-1	-0.12	0.22	1.5
	0.51	-0.07	-0.81	-0.11	1.71		0.11	0.23	1	-0.51	1.2
	0.33	-0.41	0	-1	1.21		0.17	-0.21	0.31	-1	0.17
11	-0.93	-0.08	0.11	-1.18	0.51	12	-0.95	-0.06	-0.12	0.14	2.17
	0.18	-0.48	0	0.21	-1.17		0.04	-1.12	0.08	0.11	1.4
	0.13	0.31	-1	-0.21	1.02		0.11	0.12	0	1.03	0.8
	0.08	0	-0.33	-0.72	0.28		0.34	0.08	-1.06	0.14	2.1
13	0	-0.19	0.27	-0.88	1.2	14	-0.88	-0.23	0.25	-0.16	1.24
	-0.33	-1	-0.07	0.21	0.92		0.33	0.03	-0.84	-0.32	-1.15
	0.11	0	1.03	-0.42	0.92		0.14	-0.66	-0.18	0.24	0.89
	-0.92	-0.03	0	-0.04	1.2		0.12	-0.05	0	-0.85	0.57

Продовження таблиці 3.2

15	0.12	-1	0.32	-0.18	0.72	16	-0.86	0.23	0.18	0.17	1.42
	0.08	-0.12	-0.77	0.32	0.58		0.12	-1.14	0.08	0.09	0.83
	0.25	0.22	0.14	-1	-1.56		0.16	0.24	-1	-0.35	-1.21
	-0.77	-0.14	0.06	-0.12	-1.21		0.23	-0.08	0.05	-0.75	-0.65
17	76	21	6	-34	-142	18	-83	27	-13	-11	-21
	12	-114	8	9	83		5	-68	13	24	26
	16	24	-100	-35	-121		9	54	127	36	23
	23	-8	5	-75	85		13	27	34	156	49
19	1	2	3	9	1.11	20	-1	0.28	-0.17	0.06	142
	2	1	9	4	1.16		0.52	-1	0.12	0.17	117
	3	9	1	4	1.24		0.17	-0.18	-0.79	0	0.81
	9	1	3	4	1.55		0.11	0.22	0.03	-0.95	-0.72
21	76	21	6	-34	142	22	-83	27	-13	-11	142
	12	-114	8	9	83		5	-68	13	24	26
	16	24	-100	35	121		9	54	127	36	23
	23	-8	5	-75	85		13	27	34	156	49
23	25	3	5	4	1.11	24	0.12	-1	0.32	-0.18	0.72
	5	4	3	25	1.16		0.08	-0.12	-0.77	0.32	0.58
	3	25	4	5	1.24		0.25	0.02	0.14	-1	-1.56
	4	5	25	3	1.55		-0.77	-0.14	0.06	-0.12	-1.21
25	-0.86	0.23	0.18	0.17	1.42	26	-0.93	-0.08	0.11	-1.18	0.51
	0.12	-1.14	0.08	0.09	0.83		0.18	-0.48	0	0.21	-1.17
	0.16	0.24	-1	-0.35	-1.21		0.13	0.31	-1	-0.21	1.02
	0.23	-0.08	0.05	-0.75	-0.65		0.08	0	-0.33	-0.72	0.28
27	1	2	3	1	4	28	2	-4	1	1	3
	3	6	-5	3	5		1	2	5	2	5
	3	2	7	1	11		5	3	-1	8	2
	2	4	2	3	6		6	-1	-6	9	1
29	0.79	-0.12	0.34	0.16	-0.64	30	0	-0.19	0.27	-0.88	1.2
	-0.34	1.18	-0.17	0.18	1.42		-0.33	-1	-0.07	0.21	0.92
	-0.16	-0.34	0.85	0.31	-0.42		0.11	0	1.03	-0.42	0.92
	-0.12	0.26	0.08	0.75	0.83		-0.92	-0.03	0	-0.04	1.2

Задача 2. Розв'яжіть попередню задачу, використовуючи замість методу простої ітерації метод Зейделя. Підрахуйте теоретично кількість ітерацій, що слід виконати для досягнення заданої точності і порівняйте її з практично отриманим результатом. За яким з ітераційних методів розв'язок отримано з меншою кількістю ітерацій?

4 МЕТОДИ РОЗВ'ЯЗАННЯ НЕЛІНІЙНИХ РІВНЯНЬ ТА СИСТЕМ НЕЛІНІЙНИХ РІВНЯНЬ

4.1 Постановка задачі

Нехай задано рівняння з однією змінною

$$f(x)=0. \quad (4.1)$$

Знайти точні значення коренів даного рівняння можна лише для найпростіших функцій: алгебраїчних многочленів не вище 4-го степеня, деяких многочленів степеня $n \geq 5$ і деяких трансцендентних функцій [4, 12].

Універсальних методів для знаходження точних значень коренів алгебраїчних рівнянь степеня $n \geq 5$ і трансцендентних рівнянь не існує. Крім того, розв'язуючи практичні задачі, часто отримують рівняння з коефіцієнтами, які є наближеними числами. Тоді постановка задачі знаходження точних коренів немає смислу. Отже, важливого значення набувають наближені методи знаходження коренів рівняння з достатньою для практики точністю. Задача знаходження коренів рівняння (4.1) вважається розв'язаною, якщо корені обчислені із наперед заданою точністю.

Нехай x^* – точний корінь, а \bar{x} – його наближене значення. Кажуть, що корінь \bar{x} обчислено з наперед заданою точністю ε , якщо $|x^* - \bar{x}| \leq \varepsilon$. Нехай, $x^* \in [a, b]$ і $b - a \leq \varepsilon$, тоді a і b – наближені значення кореня x^* , відповідно з недостачею і надлишком з точністю ε . При цьому за наближене значення \bar{x} з точністю ε можна взяти будь-яке число з відрізка $[a, b]$.

Знаходження наближених коренів рівняння (4.1) складається з двох етапів [4, 12]: 1) відокремлення коренів, тобто знаходження достатньо малих відрізків, на кожному з яких міститься тільки один корінь рівняння; 2) обчислення коренів з наперед заданою точністю.

Зазначимо, що корені рівняння (4.1) можуть бути дійсними і комплексними. Далі розглянемо обчислення лише дійсних коренів рівняння (4.1).

4.2 Класифікація нелінійних рівнянь

Класифікація нелінійних рівнянь наведена на рис. 4.1. Розглянемо означення та приклади різних типів нелінійних рівнянь.

Функція є *алгебраїчною*, якщо для отримання значення функції за даним значенням x слід виконати арифметичні операції і операції піднесення до степеня з раціональним показником.

Алгебраїчна функція є *раціональною* відносно змінної x , якщо над x не виконується ніяких інших дій, крім додавання, віднімання, множення, ділення і піднесення до цілого степеня. Наприклад:

$$f_1(x) = x^3 + 15x^2 - 1200x + 4; \quad f_2(x) = (x - 4)(x + 5);$$

$$f_3(x) = \frac{3}{x+7} + \frac{4x+3}{3x^2+5}.$$

Якщо у раціональну функцію змінна x не входить як дільник або не входить у вираз, який є дільником – функція є цілою раціональною. Вона визначена на всій числовій осі. Так, функції: $y = a_0x^n + a_1x^{n-1} + \dots + a_n$ (де n – натуральне число або нуль, a_0, a_1, \dots, a_n – будь-які дійсні числа, причому $a_0 \neq 0$); $f(x) = \frac{x^2}{4} + \frac{x+8}{3}$ є цілими раціональними.

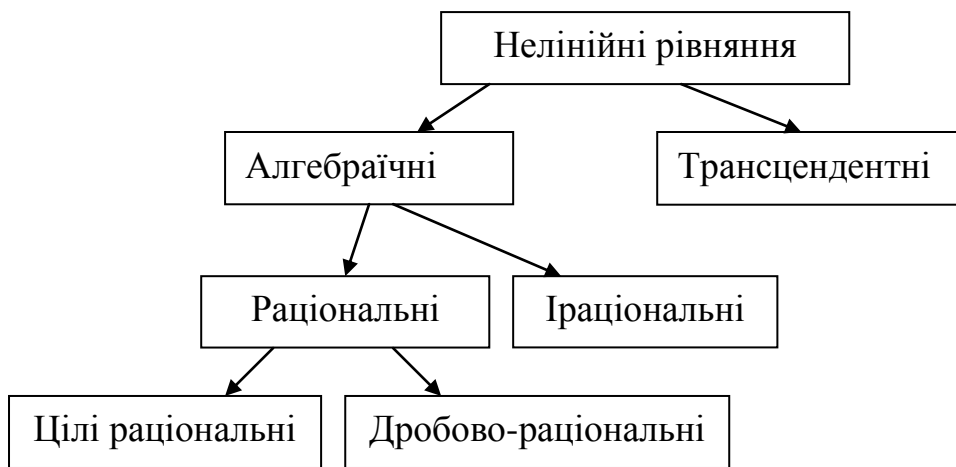


Рисунок 4.1 – Класифікація нелінійних рівнянь

Якщо в раціональній функції хоча б один раз зустрічається ділення на змінну x або змінна x не входить у вираз, який є дільником, то така функція є дробово-раціональною.

Так, наприклад функція

$$y = \frac{b_0x^m + b_1x^{m-1} + \dots + b_m}{a_0x^n + a_1x^{n-1} + \dots + a_n},$$

де m – натуральне число або нуль;

n – натуральне число;

$a_0, a_1, \dots, b_0, b_1, \dots$ – будь-яке дійсне число ($a_0 \neq 0, b_0 \neq 0$) – є дробово-раціональною функцією, що визначена на всій числовій осі, за винятком тих точок, в яких знаменник перетворюється на нуль.

Функція є ірраціональною, якщо для отримання значення функції за даним значенням x необхідно виконати, крім чотирьох арифметичних дій (усіх або деяких), ще й обчислення кореня. При цьому функція буде ірраціональною лише тоді, коли аргумент x знаходиться під знаком радикала.

Так, функція $y = \frac{3x^2 - \sqrt{x-1}}{7x-4}$ – ірраціональна, а $y = \sqrt{\frac{1+\sqrt{3}}{4}}x^2 + \frac{\sqrt{5}}{2}x$ – ні.

Другий великий клас функцій – *трансцендентні* функції. До них відносяться всі неалгебраїчні функції: показникова a^x ; логарифмічна $\log_a x$; тригонометричні $\sin x$, $\cos x$, $\operatorname{tg} x$, $\operatorname{ctg} x$; обернені тригонометричні $\arcsin x$, $\arccos x$, $\operatorname{arctg} x$, $\operatorname{arcctg} x$ тощо.

4.3 Відокремлення та уточнення коренів

На сьогодні відомі точні методи розв'язання алгебраїчних рівнянь (лінійних, квадратних, біквадратних), а також найпростіших трансцендентних рівнянь (тригонометричних, логарифмічних, показникових) [4, 12]. Проте, навіть для алгебраїчних рівнянь степеня n ($n \geq 5$) задача пошуку розв'язку у вигляді виразу з скінченною кількістю алгебраїчних дій над деякими числами є нерозв'язною. У цих випадках застосовують методи, що дозволяють знайти наближений розв'язок із заданою точністю. Наближені методи застосовують і тоді, коли точні методи потребують складних і громіздких обчислень, наприклад, при розв'язуванні алгебраїчних рівнянь третього і вищого степенів. З практичної точки зору такі наближені розв'язки не поступаються точним. Нехай, наприклад, при розв'язанні деякого квадратного рівняння знайдено точне значення одного з коренів $x = \sqrt{3}$. Щоб скористатися цим результатом при розв'язуванні практичної задачі, слід замінити $\sqrt{3}$ його наближеним значенням із необхідною для задачі точністю.

Будь-яке рівняння з одним невідомим може бути записано у вигляді

$$f(x) = 0, \quad (4.2)$$

де $f(x)$ – функція дійсних чи комплексних змінних (тут і нижче будемо вважати, що $f(x)$ – функція дійсних змінних).

Коренем рівняння (4.2) є таке значення $x = x_0$, при якому функція $f(x)$ перетворюється на нуль, тобто $f(x_0) = 0$. Дійсний корінь рівняння (4.2) геометрично є абсцисою точки перетину графіка функції $y = f(x)$ з віссю Ox .

Якщо функція $f(x)$ безперервна на відрізку $[a, b]$, а її значення на кінцях відрізка мають різні знаки ($f(a) \cdot f(b) < 0$), то на даному відрізку знайдеться принаймні одна точка x_0 , така, що $f(x_0) = 0$, тобто на відрізку $[a, b]$ є щонайменше один корінь рівняння $f(x) = 0$, (рис. 4.2). Геометрично це означає, що коли графік безперервної функції розташований по різні сторони від осі Ox , то він перетинає вісь принаймні в одній точці $x = x_0$ (рис. 4.2).

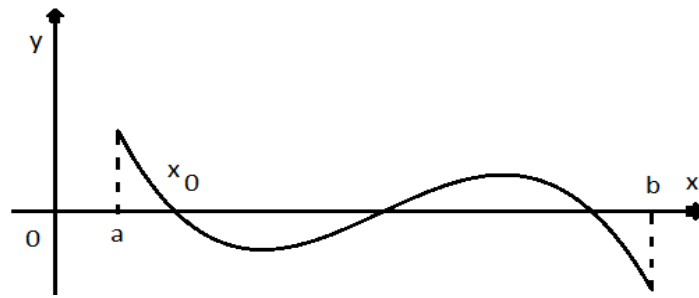


Рисунок 4.2 – Відокремлення коренів

Нехай рівняння (4.2) є алгебраїчним, тобто

$$f(x) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n = 0, \quad (4.3)$$

де a_0, a_1, \dots, a_n – дійсні числа ($a_0 > 0$).

Будь-яке алгебраїчне рівняння n -го степеня має n коренів, враховуючи і комплексні корені. Якщо $a + ib$ – корінь рівняння (4.3), то $a - ib$ – також корінь даного рівняння, тобто рівняння може мати тільки парну кількість комплексних коренів. Алгебраїчне рівняння (4.3) при непарному n має, щонайменше, один дійсний корінь.

Розглянемо деякі з методів пошуку дійсних коренів рівняння, в основу яких покладено ідею послідовного уточнення початкового наближення кореня. Попередньо зупинимось на питанні відокремлення коренів рівнянь.

4.4 Методи відокремлення коренів

Відокремити корінь рівняння – означає знайти такий інтервал, в середині якого є корінь даного рівняння, і цей корінь – єдиний на даному інтервалі.

Для відокремлення коренів рівняння (4.1) застосовують такий критерій: якщо на відрізку $[a, b]$ функція $f(x)$ безперервна і монотонна, а її значення на його кінцях мають різні знаки, то на розглянутому відрізку існує один і тільки один корінь даного рівняння (рис. 4.2) [4, 12]. Достатня ознака монотонності функції на цьому відрізку – збереження знака похідної функції. Якщо $f'(x) > 0$, функція монотонно зростає на даному відрізку (рис. 4.3, а); якщо $f'(x) < 0$ – монотонно спадає (рис. 4.3, б). Якщо на відрізку $[a, b]$ функція $f(x)$ не змінює знак, але у деякій точці x_0 цього відрізку перетворюється на нуль, то або $f'(x_0) = 0$ (рис. 4.4, а), або $f'(x_0)$ не існує (рис. 4.4, б). В обох випадках точка x_0 є точкою екстремуму; корінь рівняння (4.2) – границя проміжку монотонності функції $f(x)$.

Відокремлення коренів рівняння (4.2) можна виконати графічно. Для цього слід побудувати графік функції $y = f(x)$, за яким можна судити

про те, в яких інтервалах знаходяться його точки перетину з віссю Ox . У деяких випадках рівняння (4.2) слід подати в еквівалентному вигляді

$$f_1(x) = f_2(x) \quad (4.4)$$

з таким розрахунком, щоб графіки функцій $y = f_1(x)$, $y = f_2(x)$ будувалися простіше, ніж графік $y = f(x)$.

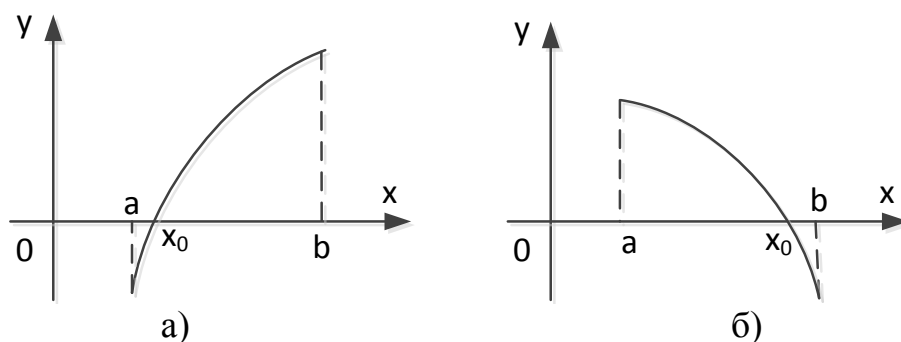


Рисунок 4.3 – Відокремлення коренів безперервної і монотонної функції

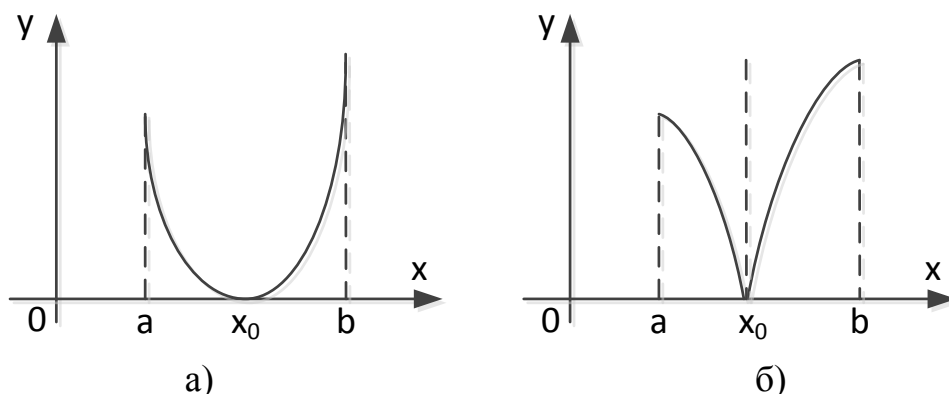


Рисунок 4.4 – Відокремлення коренів функції, що не змінює знак, але у деякій точці x_0 відрізка перетворюється на нуль

Знайти корінь рівняння (4.2) – означає знайти абсцису точки перетину графіка функції $y = f(x)$ з віссю абсцис. Корінь рівняння (4.4) визначається як абсциса точки перетину графіків функцій $y = f_1(x)$ і $y = f_2(x)$. При цьому в силу еквівалентності рівнянь (4.2) і (4.4) абсциси відповідних точок рівні між собою.

Якщо рівняння є алгебраїчним, то під час визначення границь коренів можна застосовувати нижченаведені правила [3, 4, 12].

Усі дійсні корені рівняння (4.4), якщо вони існують, містяться в інтервалі $(-K_0, K_0)$, причому число K_0 визначається формулою

$$K_0 = 1 + A / a_0, \quad (4.5)$$

де A – найбільше з чисел $|a_1|, |a_2|, \dots, |a_n| \dots$

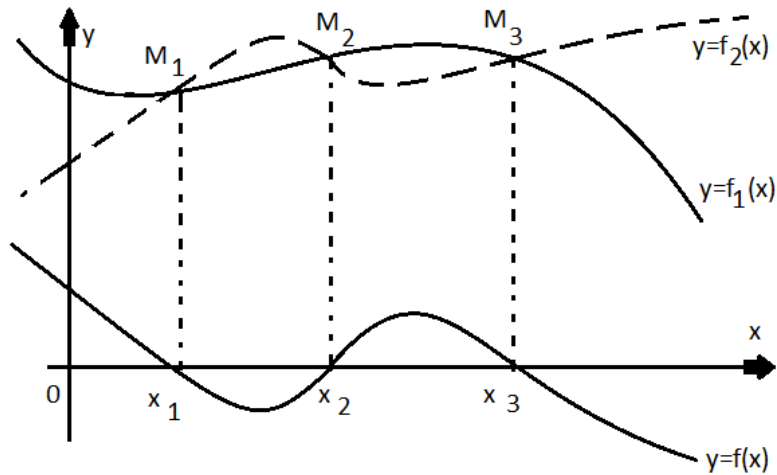


Рисунок 4.5 – Побудова графіка функції $y = f(x)$ у еквівалентному вигляді

Знаючи лише верхню границю додатних коренів будь-якого многочлена, можна вказати інтервали, у яких знаходяться його додатні та від'ємні корені. Разом з многочленом

$$f(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n, \quad (4.6)$$

для якого число K_0 служить верхньою границею його додатних коренів, розглядають багаточлени:

$$\varphi_1(x) = x^n f(1/x); \quad (4.7)$$

$$\varphi_2(x) = f(-x); \quad (4.8)$$

$$\varphi_3(x) = x^n f(-1/x), \quad (4.9)$$

де $f(x)$ визначається за формулою (4.6).

Нехай верхніми границями додатних коренів цих многочленів будуть відповідно числа DO_1 , DO_2 , DO_3 . Тоді всі додатні корені многочлена $f(x)$ задовольняють нерівність

$$1/K_1 < x < K_0, \quad (4.10)$$

а усі від'ємні – нерівність

$$-K_2 < x < -1/K_3. \quad (4.11)$$

Для додатних коренів рівняння (4.2), точніше верхня границя, може бути отримана за формулою Маклорена

$$K' = 1 + \sqrt[m]{M/a_0}, \quad (4.12)$$

де M – максимум модуля від’ємних коефіцієнтів рівняння;
 m – номер першого з від’ємних коефіцієнтів.

Кращий результат порівняно з вищенаведеними двома результатами дає метод Ньютона, згідно з яким число $c > 0$ буде верхньою границею додатних коренів рівняння (5.2), якщо

$$f(c) > 0, f'(c) > 0, \dots, f^{(n-1)}(c) > 0. \quad (4.13)$$

Під час визначення кількості додатних коренів рівняння (4.3) користуються теоремою Декарта [4, 12]. Кількість додатних коренів многочлена, з врахуванням їх кратностей, дорівнює кількості змін знаків у системі коефіцієнтів цього многочлена (коефіцієнти, що дорівнюють нулю не враховуються) або менша за цю кількість на парне число.

Якщо в деякому інтервалі виявляється кілька коренів, для їх відокремлення можна застосовувати табличний спосіб, що потребує обчислення значень многочлена $f(x)$ при різних значеннях x . Найпростіше розв’язок цієї задачі можна отримати за схемою Горнера, яка основана на тому, що при діленні полінома $f(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \dots + a_{n-1}x + a_n$ на $x - x_0$ коефіцієнти частки $q(x) = b_0x^{n-1} + b_1x^{n-2} + b_2x^{n-3} + \dots + b_{n-1}$ і залишок r визначаються формулами:

$$b_0 = a_0, b_k = x_0 \cdot b_{k-1} + a_k \quad (k = 1, 2, \dots, n-1), \quad r = x_0 b_{n-1} + a_n = f(x_0). \quad (4.14)$$

Обчислення при цьому розташовуються у вигляді схеми Горнера:

$$\begin{array}{c|c|c|c|c|c} \frac{a_0}{x_0} & \frac{a_1}{b_1} & \frac{a_2}{b_2} & \dots & \frac{a_{n-1}}{b_{n-1}} & \frac{a_n}{r} \\ \hline & b_1 = x_0 a_0 + a_1 & b_2 = x_0 b_1 + a_2 & \dots & b_{n-1} = x_0 b_{n-2} + a_{n-1} & r = x_0 b_{n-1} + a_n = f(x_0) \end{array}. \quad (4.15)$$

4.4.1 Загальні властивості алгебраїчних рівнянь

Запишемо алгебраїчне рівняння n -ого степеня:

$$P_n(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \dots + a_{n-1}x + a_n = 0, \quad (4.16)$$

де $P_n(x)$ – многочлен n -го степеня;
 n – найбільший степінь при невідомому;
 a_0, a_1, \dots, a_n – дійсні коефіцієнти.

Як відомо, будь-яке число ξ , що перетворює многочлен на нуль (тобто, $P_n(\xi) = 0$), є коренем многочлена. ξ є коренем многочлена $P_n(x)$ тільки тоді, коли $P_n(x)$ ділиться без залишку на $x - \xi$. Якщо при цьому $P_n(x)$ ділиться без залишку на $(x - \xi)^k$ ($k \geq 1$), але вже не ділиться на $(x - \xi)^{k+1}$, то ξ називається k -кратним коренем (або коренем кратності k) многочлена $P_n(x)$. Корені кратності $k=1$ є простими коренями многочлена.

Виникає запитання: чи будь-який многочлен має корені? Відповідь на нього дає нижченаведена теорема.

Теорема 1 (основна теорема алгебри). *Будь-який многочлен з будь-якими числовими коефіцієнтами, степінь якого не менший одиниці, має хоча б один корінь, у загальному випадку комплексний.*

З теореми 1 випливає наслідок: *будь-який многочлен $P_n(x)$ степеня n ($n > 1$) з будь-якими числовими коефіцієнтами має рівно n коренів, дійсних або комплексних, якщо кожний з коренів рахувати стільки разів, скільки дорівнює його кратність.*

Таким чином, корені алгебраїчного рівняння (4.16) можуть бути як дійсними, так і комплексними. Останні мають властивість *парної спряженості*, тобто якщо рівняння (4.16) має комплексний корінь $\xi = \alpha + \beta i$ (де α і β – дійсні числа) кратності k , то воно має і комплексний корінь $\bar{\xi} = \alpha - \beta i$ також кратності k . Модулі цих коренів однакові: $|\xi| = |\bar{\xi}| = \sqrt{\alpha^2 + \beta^2}$.

Якщо рівняння (4.16) має комплексні корені, то їх кількість – парна. Тому будь-яке алгебраїчне рівняння непарного степеня з дійсними коефіцієнтами має, принаймні, один дійсний корінь.

Перш ніж обчислювати корені алгебраїчного рівняння, спочатку необхідно: а) визначити кількість коренів, що має дане рівняння; б) знайти область існування коренів (установити верхню і нижню границі розташування коренів). Після цього можна перейти до відокремлення коренів і їх уточнення.

4.4.2 Визначення кількості дійсних коренів алгебраїчного рівняння

Кількість дійсних додатних коренів алгебраїчного рівняння (4.16) можна визначити приблизно за правилом Декарта: *кількість дійсних додатних коренів алгебраїчного рівняння $P_n(x) = 0$ з дійсними коефіцієнтами (кожен підраховується стільки разів, яка його кратність) або дорівнює кількості змін знака в послідовності коефіцієнтів рівняння $P_n(x) = 0$, або на парне число менша (коефіцієнти, що дорівнюють нулю, не враховуються).*

Кількість від'ємних коренів рівняння дорівнює кількості змін знака в послідовності коефіцієнтів $P_n(-x)$ або на парне число менша.

Якщо рівняння повне, то кількість його додатних коренів дорівнює кількості змін знака в послідовності коефіцієнтів або на парне число менша, а кількість від'ємних коренів – кількості сталості знака або на парне число менша.

Приклад 1. Визначити кількість додатних та від'ємних коренів рівняння $x^5 - 17x^4 + 12x^3 + 7x^2 - x + 1 = 0$.

Згідно з основною теоремою алгебри це рівняння має п'ять коренів (з

них хоча б один є дійсним). Рівняння є повним, послідовність знаків коефіцієнтів така: +, -, +, +, -, +. Знак змінюється чотири рази. Отже, додатних коренів або 4, або 2, або жодного. Кількість сталостей знака дорівнює 1; отже, дане рівняння має один від'ємний корінь.

Приклад 2. Визначити кількість дійсних додатних та від'ємних коренів рівняння $x^6 - 3x^4 + x^3 + x^2 - 1 = 0$.

Дане рівняння має шість коренів. Послідовність знаків: +, -, +, +, -, має місце три зміни знака – отже, додатних коренів або 3, або 1. Далі, для многочлена $P_n(-x) = x^6 - 3x^4 + x^3 + x^2 - 1$ послідовність знаків така: +, -, -, +, -. Тут також маємо три зміни знака, тому від'ємних коренів або 3, або 1.

Точнішу кількість коренів алгебраїчного рівняння можна визначити за допомогою теореми Штурма [4, 12].

Оскільки кратні корені рівняння завжди можуть бути виділені як спільні корені рівнянь $P_n(x)=0$ і $P'_n(x)=0$, без обмеження спільності можна вважати, що задане рівняння $P_n(x)=0$ має лише прості корені.

Нехай для даного алгебраїчного рівняння $P_n(x)=0$ будь-яким способом встановлено, що всі його дійсні корені знаходяться в інтервалі (a, b) (a і b – дійсні числа, які не є коренями рівняння і $a < b$). Знайдемо першу похідну $P'_n(x)$ і поділимо на неї многочлен $P_n(x)$. Залишок від ділення $P_n(x)$ на $P'_n(x)$ візьмемо з протилежним знаком і позначимо його через $R_1(x)$.

Потім так само ділимо $P'_n(x)$ на $R_1(x)$, отриманий залишок беремо з протилежним знаком і позначаємо через $R_2(x)$. Розділивши $R_1(x)$ на $R_2(x)$ і знову взявши залишок з протилежним знаком, одержимо $R_3(x)$. Процес ділення продовжуємо доти, доки не буде отриманий залишок, що є постійною величиною. Цю величину також беремо з протилежним знаком. Одержуємо послідовність функцій

$$P_n(x), P'_n(x), R_1(x), R_2(x), \dots, R_{m-1}(x), R_m = \text{const},$$

яка називається *системою Штурма*. В цю послідовність підставляємо замість x спочатку a , потім b і підраховуємо кількість змін знака в обох випадках (позначимо отримані числа відповідно через $W(a)$ і $W(b)$).

Теорема 2 (теорема Штурма). *Якщо дійсні числа a і b ($a < b$) не є коренями многочлена $P_n(x)$, який не має кратних коренів, то $W(a) > W(b)$ і різниця $W(a) - W(b)$ дорівнює кількості дійсних коренів многочлена $P_n(x)$, розташованих між a і b .*

За допомогою теореми Штурма можна визначити кількість від'ємних коренів рівняння $P_n(x)=0$, тобто кількість його дійсних коренів в інтервалі $(-\infty, 0)$, а також кількість додатних коренів в інтервалі $(0, +\infty)$. Теорема Штурма застосовується також і для відокремлення коренів. Функції, що входять у систему Штурма, можна множити і ділити на довільні додатні числа. Це значно спрощує роботу, коли виконується ділення із залишком.

Приклад 3. Визначити кількість дійсних коренів рівняння $5x^3 - 20x + 3 = 0$, а також відокремити ці корені, користуючись теоремою Штурма.

Складаємо систему функцій Штурма: $P_n(x)=5x^3-20x+33$; $P'_n(x)=15x^2-20$.
Для визначення $R_1(x)$ помножимо $P_n(x)$ на 3 і потім поділимо на $P'_n(x)$:

$$\begin{array}{r|l} 15x^3-60x+9 & 15x^2-20 \\ \hline \mp 15x^3 \pm 20x & x \\ \hline -40x+9 & \end{array}$$

Виходить, $R_1(x)=40x-9$ (залишок взятий з протилежним знаком).
Помножимо $P'(x)$ на 8 і розділимо цей добуток на $R_1(x)$:

$$\begin{array}{r|l} 120x^2-160 & 40x-9 \\ \hline \mp 120x^2 \pm 27x & 3x+27 \\ \hline 40(27x-160) & \\ \hline 40 \cdot 27x - 40 \cdot 160 & \\ \hline \mp 40 \cdot 27x \pm 9 \cdot 27 & \\ \hline - & \end{array}$$

Оскільки останній залишок є постійним числом зі знаком “-” (а нас у цьому випадку цікавить саме знак постійного залишку), змінюємо його на протилежний, тобто на “+”.

Складаємо таку таблицю знаків функцій, що входять у систему Штурма.

Таблиця 4.1 – Таблиця знаків функцій, що входять у систему Штурма

X	$P_n(x)$	$P'(x)$	$R_1(x)$	R_2	$W(x)$
$-\infty$	-	+	-	+	3
0	+	-	-	+	2
$+\infty$	+	+	+	+	0

З таблиці 4.1 видно, що в інтервалі $(-\infty, +\infty)$ містяться три дійсних корені, оскільки $W(-\infty)-W(+\infty)=3-0=3$. З них один корінь – від’ємний ($W(-\infty)-W(0)=3-2=1$), а два – додатні ($W(0)-W(+\infty)=2-0=2$). Використовуючи теорему Штурма, відокремимо корені, скоротивши інтервали до довжини, рівної 1:

Таблиця 4.2 – Інтервали розташування коренів

X	$P_n(x)$	$P'(x)$	$R_1(x)$	R_2	$W(x)$
$-\infty$	-	+	-	+	3
-3	-	+	-	+	3
-2	+	+	-	+	2
-1	+	-	+	+	2
0	+	-	-	+	2
1	-	-	+	+	1
2	+	+	+	+	0

З таблиці 4.2 видно, що корені розташовані в інтервалах $(-3, -2)$, $(0, 1)$, $(1,2)$.

Слід зазначити, що звужувати відрізки існування коренів доцільно, застосовуючи певні правила. Саме такі правила і викладені нижче.

4.4.3 Знаходження області існування коренів алгебраїчного рівняння

Правило кільця. Нехай дано алгебраїчне рівняння

$$P_n(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \dots + a_{n-1}x + a_n = 0,$$

де a_0, a_1, \dots, a_n – дійсні коефіцієнти і нехай $A = \max \{|a_1|, |a_2|, \dots, |a_n|\}$, $B = \max \{|a_0|, |a_1|, \dots, |a_{n-1}|\}$. Тоді корені рівняння розташовані в круговому кільці $r < |x| < R$, де

$$r = \frac{1}{1 + B/|a_n|}; R = 1 + \frac{A}{|a_0|}.$$

При цьому r – нижня, а R – верхня границя додатних коренів алгебраїчного рівняння $P_n(x) = 0$ і $-R$, $-r$ відповідно нижня і верхня границя від'ємних коренів (рис. 4.5).

Приклад 1. Визначити границі коренів рівняння $5x^3 - 20x + 3 = 0$.

Тут $|a_0| = 5$, $A = 20$, $|a_n| = 3$, $B = 20$, тобто

$$R = 1 + \frac{A}{|a_0|} = 1 + \frac{20}{5} = 5; r = \frac{1}{1 + B/|a_n|} = \frac{1}{1 + 20/3} = \frac{3}{23} \approx 0,013.$$

Тоді, якщо дійсні корені рівняння $5x^3 - 20x + 3 = 0$ існують (а вони обов'язково існують, тому що рівняння непарного степеня), то вони розташовані в інтервалі $(-5; 5)$. При цьому від'ємні корені знаходяться в інтервалі $(-5; -0,013)$, а додатні – в інтервалі $(0,013; 5)$.

Під час розв'язування рівнянь зручно спочатку встановити границі коренів, а потім вже застосовувати теорему Штурма. За правилом кільця ці границі визначаються дуже приблизно.

Покажемо прийом точнішого визначення границь дійсних коренів алгебраїчного рівняння $P_n(x) = 0$. Якщо R_1 – верхня границя додатних коренів $P_n(x)$, R_2 – верхня границя додатних коренів $P_n(-x)$, $R_3 > 0$ – верхня границя додатних коренів $x^n P_n(1/x)$ і R_4 – верхня границя додатних коренів $x^n P_n(-1/x)$, то усі ненульові дійсні корені рівняння $P_n(x) = 0$ (якщо вони існують) розташовані усередині інтервалів $(-R_2, -1/R_4)$ і $(1/R_3, R_1)$.

Для визначення верхньої границі додатних коренів алгебраїчного рів-

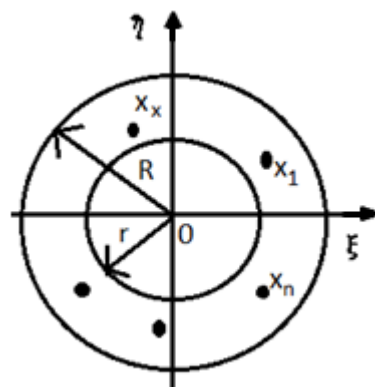


Рисунок 4.5

няння можна скористатися методами Лагранжа або Ньютона [4, 12].

Метод Лагранжа. Якщо коефіцієнти многочлена

$$P_n(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \dots + a_n$$

задовольняють умови $a_0 > 0$, $a_0, a_1, \dots, a_{m-1} \geq 0$, $a_m < 0$, то верхня границя додатних коренів рівняння $P_n(x) = 0$ знаходиться за формулою $R = 1 + \sqrt[n]{B/a_0}$, де B – найбільше з абсолютних значень від’ємних коефіцієнтів $P_n(x)$.

Приклад 2. Методом Лагранжа визначити границі додатних і від’ємних коренів рівняння $8x^4 - 8x^2 - 32x + 1 = 0$.

Тут $a_0 = 8 > 0$; $a_1 = 0$; $a_2 = -8 < 0$; $a_3 = -32$; $a_4 = 1$, $m = 2$ (номер першого з від’ємних коефіцієнтів), $B = 32$. Отже, $R_1 = 1 + \sqrt[4]{32/8} = 3$.

Розглянемо многочлен $P_n(-x) = 8x^4 - 8x^2 + 32x + 1$. Аналогічно знайдемо, що для нього верхньою границею додатних коренів є $R_2 = 1 + \sqrt{8/8} = 2$.

Далі, для многочлена $x_4 \cdot P_n(1/x) = x^4 - 32x^3 - 8x^2 + 8$ маємо: $a_0 = 1 > 0$; $a_1 = -32 < 0$, тобто: $m = 1$, $B = 32$; $R_3 = 1 + 32 = 33$.

Нарешті, для многочлена $x^4 P_n(-1/x) = x^4 - 32x^3 - 8x^2 + 8$ маємо: $a_0 = 1 > 0$; $a_1 = -32 < 0$; $a_2 = -8$; $a_3 = 0$; $a_4 = 8$, тобто $m = 2$. Тому $R_4 = 1 + \sqrt{8} = 1 + 2\sqrt{2} = 3,828$.

Отже, якщо рівняння $8x^4 - 8x^2 + 32x + 1 = 0$ має дійсні корені, то вони обов’язково лежать в інтервалах $(-2, -1/3, 828)$ і $(1/33, 3)$.

Метод Ньютона.

Якщо при $x = c$ многочлен $P_n(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \dots + a_n$ і його похідні $P'_n(x), P''_n(x), \dots$ набувають додатних значень, то c є верхньою границею додатних коренів рівняння $P_n(x) = 0$.

Приклад 3. Методом Ньютона визначити верхню границю додатних коренів рівняння $8x^4 - 8x^2 - 32x + 1 = 0$.

$$\begin{aligned} \text{Знаходимо } P(x) &= 8x^4 - 8x^2 - 32x + 1 = 0; & P'_n(x) &= 32x^3 - 16x - 32; \\ P''_n(x) &= 96x^2 - 16; & P'''_n(x) &= 192x; & P''''_n(x) &= 192. \end{aligned}$$

Перевірці підлягають значення $x > 0$. При $x = c = 1$ маємо $P(1) < 0$. Отже, проводити далі перевірку для $x = 1$ не потрібно. Перевіримо значення $x = c = 2$: $P(2) > 0$; $P'(2) > 0$; $P''(2) > 0$; $P'''(2) > 0$; $P'''' > 0$. Таким чином, верхньою границею додатних коренів є число 2, тобто $R = 2$. За нижню границю можна взяти число, оберне до R тобто $r = 1/2$.

4.5 Методи уточнення коренів

4.5.1 Метод Хорд

Нехай на відрізьку $[a, b]$ знаходиться єдиний дійсний корінь рівняння $f(x) = 0$, причому функція $f(x)$ неперервна на цьому відрізьку. Як перше наближення кореня даного рівняння можна взяти абсцису x_1 точки перетину з віссю Ox хорди, що з’єднує кінці $A[a, f(a)]$, $B[b, f(b)]$ дуги графіка функції $f(x)$ на вказаному відрізьку (рис. 4.6, 4.7).

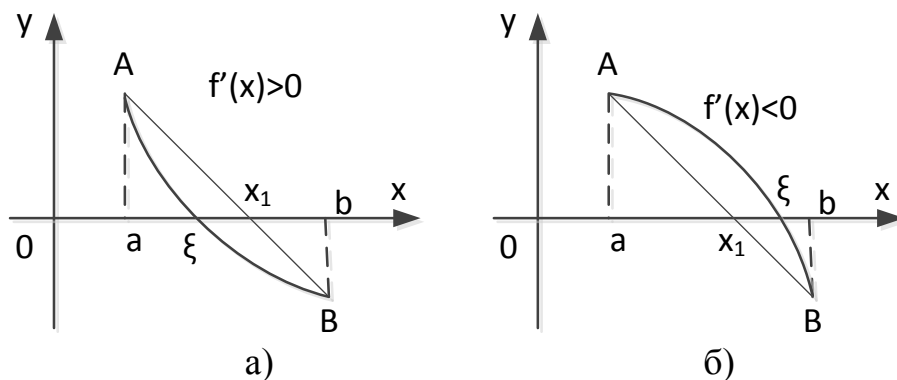


Рисунок 4.6 – Графічна ілюстрація методу хорд

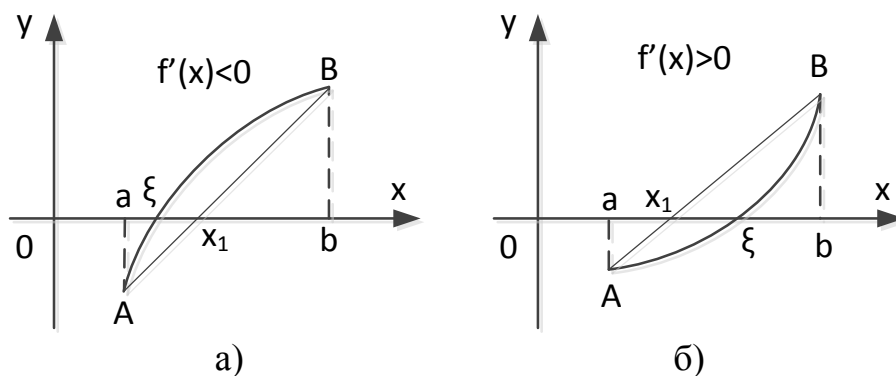


Рисунок 4.7 – Графічна ілюстрація методу хорд

Рівняння хорди AB має вигляд: $\frac{y - f(a)}{f(b) - f(a)} = \frac{x - a}{b - a}$. Поклавши у ньому $y=0$, одержимо перше наближення кореня

$$x_1 = a - \frac{f(a)}{f(b) - f(a)}(b - a). \quad (4.17)$$

Точка x_1 буде тим ближче до шуканого кореня ξ , чим менша кривизна графіка функції, тобто чим менша $|f''(x)|$ і більша $|f'(x)|$ на відрізку $[a, b]$. Взаєморозташування точок x_1 і ξ визначається порівнянням знаків $f(a)$, $f(b)$ і $f(x_1)$. Якщо ж на відрізку $[a, b]$ зберігається знак не тільки першої, а й другої похідної функції $f(x)$, тобто на цьому відрізку функція не лише монотонна, а й її графік не має точок перегину, взаємне розташування точок x_1 і ξ можна установити і без перевірки знака $f(x_1)$. Точка x_1 завжди ближче точки ξ до того кінця відрізка $[a, b]$, в якому знак функції обернений знаку її другої похідної (рис. 4.6, 4.7, а, б).

Для одержання другого наближення x_2 формулу (4.17) необхідно застосувати до того з відрізків $[a, x_1]$, $[x_1, b]$, на кінцях якого функція $f(x)$ має значення протилежних знаків.

Аналогічно обчислюються і наступні наближення. Так, якщо відоме $(n-1)$ -е наближення, то n -е можна обчислити за формулою:

$$x_n = \frac{bf(x_{n-1}) - x_{n-1}f(b)}{f(x_{n-1}) - f(b)} \quad (n=1, 2, 3, \dots) \quad (4.18)$$

(рис. 4.8), коли $f(b)f''(x) > 0$.

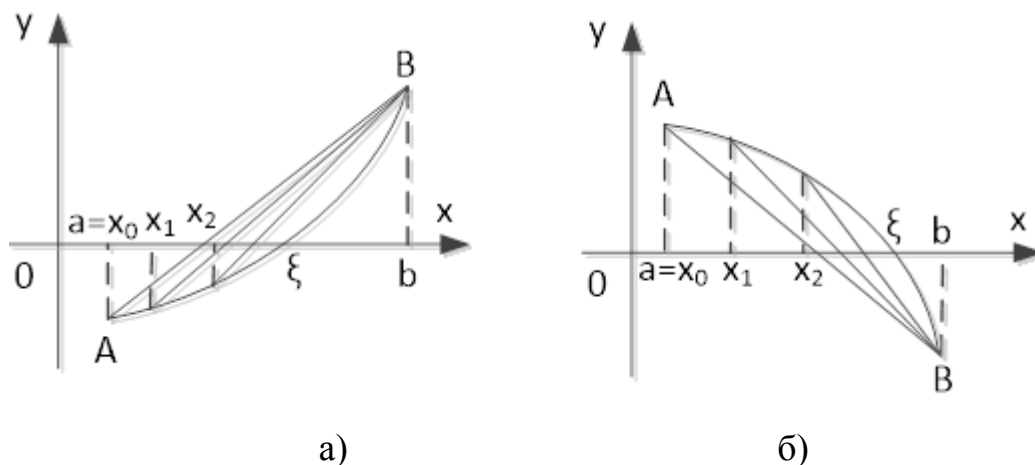


Рисунок 4.8 – Графічна ілюстрація методу хорд

Або за формулою

$$x_n = \frac{af(x_{n-1}) - x_{n-1}f(a)}{f(x_{n-1}) - f(a)} \quad (4.19)$$

(рис. 4.9), коли $f(a)f''(x) > 0$.

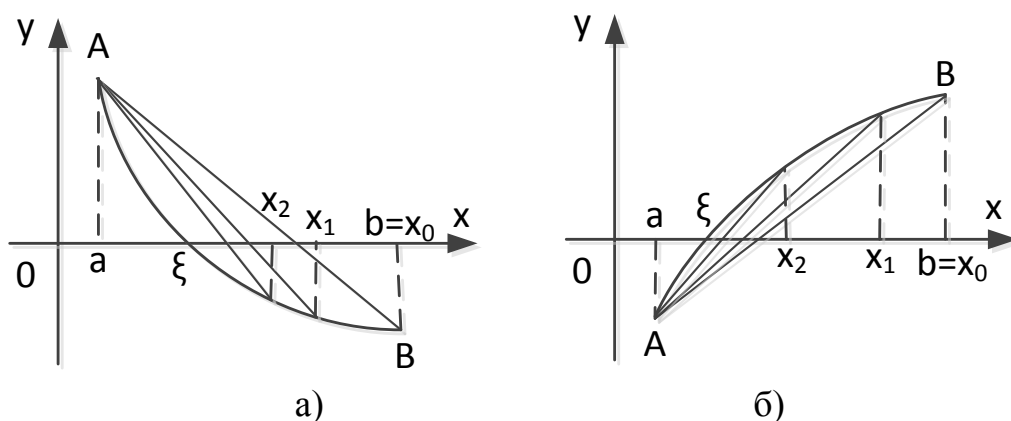


Рисунок 4.9 – Графічна ілюстрація методу хорд

У першому випадку за початкове наближення береться a , тобто $x_0=a$, у другому – b , тобто $x_0=b$.

Послідовність чисел x_n ($n=1, 2, 3, \dots$) збігається до кореня ξ , тобто

$$\lim x_n = \xi.$$

Обчислення наближень x_1, x_2, x_3, \dots слід робити доти, доки два послідовних наближення x_k, x_{k+1} не будуть збігатися на задане число знаків. Для проміжних викладок слід брати один-два запасних знаки.

Якщо функція $f(x)$ має відмінну від нуля похідну $f'(x)$ на відрізку $[a, b]$, то оцінка абсолютної похибки визначається за формулою

$$|\xi - x_n| \leq |f(x_n)| / \mu, \quad (4.20)$$

$$\text{де } \mu = \min_{a \leq x \leq b} |f'(x)|.$$

Приклад 1. Методом хорд знайти корінь рівняння $x^3 - 2x + 7 = 0$.

Відокремлюючи корені даного рівняння вищерозглянутими способами, визначаємо, що єдиний дійсний корінь знаходиться на відрізку $[-3, -2]$; дійсно, $f(-3) = -27 + 6 + 7 < 0$, $f(-2) = -8 + 4 + 7 > 0$. Розглянувши середину цього відрізка, одержимо $f(-2,5) = (-2,5)^3 - 2(-2,5) + 7 = -3,625 < 0$. Це означає, що корінь належить відрізку $[-2,5; -2]$. Продовжуючи аналогічні міркування, знаходимо відрізок $[-2,3; -2,2]$, якому належить корінь, оскільки $f(-2,3) \cdot f(-2,2) < 0$. На цьому відрізку похідні $f'(x) = 3x^2 - 2$ і $f''(x) = 6x$ зберігають знак. Оскільки $f(-2,3) \cdot f''(-2,3) > 0$, використовуємо формулу (4.19), вважаючи $x_0 = -2,2$. Вважаючи, що $f(x_0) = f(-2,2) = 0,752$, $a = -2,3$, $f(a) = f(-2,3) = -0,567$ за даною формулою одержимо результати, наведені у табл. 4.3.

Таблиця 4.3 – Результати розрахунків до прикладу 1

n	x_{n-1}	$f(x_{n-1})$	$A f(x_{n-1})$	$X_{n-1} f(a)$	$\frac{a f(x_{n-1}) - x_{n-1} f(a)}{f(x_{n-1}) - f(a)}$	x_n
1	-2,2	0,752	-1,7296	1,2474	-2,9770	-2,25701
2	-2,25701	0,016558	-0,038083	1,279725	-1,317809	-2,258231
3	-2,258231	0,000371	-0,000853	1,280417	-1,281270	-2,258259

Отже, знайдений корінь рівняння $\xi = -2,258$.

Приклад 2. Розв'язати рівняння $x^3 - 0,9x + 0,6 = 0$ з точністю до 10^{-4} .

Оскільки графіки функцій $y = x^3$ і $y = 0,9x - 0,6$ перетинаються в єдиній точці, дане рівняння має дійсний корінь на відрізку $[-2, 1]$. Оскільки $f(-1,2) < 0$, $f(-1,1) > 0$, корінь належить відрізку $[-1,2; -1,1]$, на якому $f''(x) = 6x < 0$. Тут виконується нерівність $f(-1,2) f''(x) > 0$, тому можна

скористатися формулою (4.19). Щоб скоротити кількість обчислень, укажемо відрізок меншої довжини, на якому знаходиться корінь. Оскільки $f(-1,186) < 0, f(-1,185) > 0$, то корінь рівняння розташований на відріжку $[-1,186; -1,185]$, довжина якого дорівнює 0,001.

За формулою (4.19) обчислюємо значення кореня рівняння (табл. 4.4).

Таблиця 4.4 – Результати розрахунків до прикладу 2

n	I	II(I.I)	III(II.I)	IV(I.(-0,9))	V(III+IV+0,6)
	x_{n-1}	x_{n-1}^2	x_{n-1}^3	$-0,9x_{n-1}$	$f(x_{n-1})$
1	-1,185	1,40422	-1,66400	1,06650	0,00250
2	-1,18570	1,40591	-1,66700	1,06739	0,00014
3	-1,18572	1,40593	-1,66704	1,06715	0,00011
4	-1,18573	1,40596	-1,66709	1,06716	0,00007

n	VI	VII	VIII(VI-VII)	IX	X
	$af(x_{n-1})$	$X_n f(a)$	$af(x_{n-1}) - x_{n-1} f(a)$	$f(a) - f(x_{n-1})$	x_{n-1}
1	-0,00297	0,00955	-0,01252	0,01056	-1,18570
2	-0,00017	0,00956	-0,00973	0,00820	-1,18572
3	-0,00013	0,00956	-0,00969	0,00817	-1,18573

Оцінимо абсолютну похибку за формулою (4.20). Оскільки $f(x_3) = f(-1,18573) = 0,00007, \min_{a \leq x \leq b} f'(x) = \min_{a \leq x \leq b} |3x^2 - 0,9| = 3,31269 (a = -1,186; b = -1,185)$, то $|\xi - x_3| \leq \frac{0,00007}{3,31269} \approx 0,00002 < 0,0001$,

тобто корінь обчислений із заданою точністю.

4.5.2 Метод дотичних і його модифікації

Метод дотичних (або метод Ньютона) [4, 12]. Нехай на відріжку $[a, b]$ знаходиться єдиний корінь рівняння $f(x)=0$. Проведемо дотичну до кривої $y=f(x)$ в точці $A[a, f(a)]$ до перетину з віссю Ox (рис. 4.10). Рівняння дотичної, що проходить через точку A , буде таким: $y=f(a)+f'(a) \cdot (x-a)$. Якщо $f'(a) \neq 0$, то з цього рівняння (при $b=0$) знаходимо абсцису x_1 точки перетину дотичної з віссю Ox :

$$x_1 = a - f(a) / f'(a).$$

Абсцису x_1 точки перетину можна взяти як наближене значення кореня.

Якщо проведемо дотичну через відповідну точку $A_1[x_1, f(x_1)]$ і знайдемо точку пе-

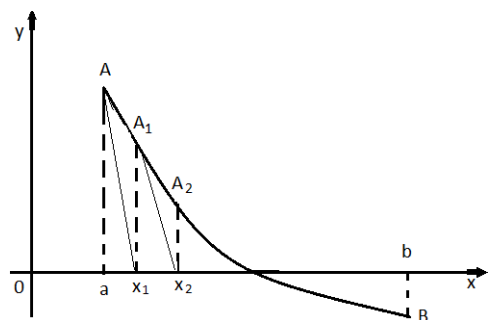


Рисунок 4.10

ретину з віссю Ox , одержимо друге наближення кореня x_2 . Аналогічно визначаються наступні наближення.

Застосовуючи метод дотичних, n -і наближення знаходять за формулою

$$x_{n+1} = x_n - f(x_n) / f'(x_n), \quad (n = 0, 1, 2, \dots), \quad (4.21)$$

причому за нульове наближення x_0 береться таке значення з відрізка $[a, b]$, для якого виконується умова

$$f(x_0) \cdot f''(x_0) > 0. \quad (4.22)$$

Обчислення проводять доти, доки не перестануть змінюватися десяткові знаки, що зберігаються у відповіді. Для проміжних викладок слід брати декілька запасних знаків. Оцінка похибки визначається за формулою

$$|x_n - \xi| \leq f(x_n) / \mu, \quad (4.23)$$

де $\mu = \min_{a \leq x \leq b} |f'(x)|$.

Наближене обчислення кореня за методом дотичних, визначення номера n наближення x_n , при якому буде досягнена задана точність, можна здійснити на основі такої теореми.

Нехай для рівняння $f(x)=0$ і початкового наближення x_0 шуканого кореня виконуються умови:

1) $f'(x_0) \neq 0$ і $|1 / f'(x_0)| \leq B$ 2) $|f(x_0) / f'(x_0)| \leq \eta$; 3) $|f''(x)| \leq K$, на відрізку $|x - x_0| \leq \delta$; 4) $h = BK\eta \leq \frac{1}{2}$; 5) $\frac{1 - \sqrt{1 - 2h}}{h} \eta \leq \delta$.

Тоді дане рівняння має корінь, що належить відрізку $|x - x_0| \leq \delta$, де $\delta = 2\eta / (1 + \sqrt{1 - 2h})$. Цей корінь може бути отриманий, як границя послідовності, кожен член якої знаходиться за формулою (4.21). Оцінка похибки визначається формулою

$$|x_n - \xi| \leq \frac{1}{2^{n-1}} (2h)^{2^n - 1} \eta. \quad (4.24)$$

Недоліком методу дотичних є те, що на кожній ітерації треба обчислювати не тільки значення функції $f(x)$, а й значення її похідної $f'(x)$. У таких випадках похідну $f'(x_k)$ замінюють її наближеним значенням

$$f'(x_n) = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}.$$

Підставивши наближене значення похідної у (4.21), отримаємо

$$x_{n+1} = x_n - \frac{f(x_{n-1})(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}, \quad n = 1, 2, 3, \dots$$

Метод, оснований на застосуванні останньої формули, отримав назву *методу січних*. Геометрично наближення x_{n+1} дістають як абсцису точки перетину осі Ox і січної, що проходить через точки $(x_{n-1}, f(x_{n-1}))$ та $(x_n, f(x_n))$.

Алгоритм методу січних починається із задання двох початкових наближень x_0 і x_1 , які вибирають на відрізку ізоляції шуканого кореня. Тому складність цього методу полягає у знаходженні таких x_0 і x_1 , які є достатньо близькими до кореня, щоб метод був збіжний.

Якщо похідна $f'(x)$ мало змінюється на відрізку $[a, b]$, то кількість обчислень за методом Ньютона можна зменшити, коли значення похідних $f'(x_n)$ у точках x_n ($n=1,2, \dots$) замінити значенням $f'(x_0)$ в точці x_0 . Тоді формула (4.21) набуде вигляду

$$x_{n+1} = x_n - f(x_n) / f'(x_0), \quad (n = 0, 1, 2, \dots).$$

Приклад 1. З точністю до п'яти десяткових знаків після коми обчислити більший від'ємний корінь рівняння $x^3 - 12x - 8 = 0$.

Графічно відокремлюючи корені даного рівняння, бачимо, що рівняння має три дійсних корені, більший від'ємний корінь належить відрізку $[-1, 0]$. Можна вказати відрізок меншої довжини, на якому знаходиться корінь, обчисливши значення многочлена $f(x) = x^3 - 12x - 8$ за допомогою схеми Горнера (табл. 4.5)

Таблиця 4.5 – Значення многочлена $f(x) = x^3 - 12x - 8$ за схемою Горнера

$x \backslash a_k$	1	0	-12	-8	$f(x)$
-0,6	1	-0,6	-11,64	-1,016	-
-0,7	1	0,7	-11,51	-0,057	+
-0,65	1	-0,65	-11,5775	-0,474625	-

Таким відрізком є відрізок $[-0,7; -0,65]$. Оскільки $f''(x) = 6x$, $f''(-0,65) < 0$, $f(-0,65) < 0$, $f(-0,65) \cdot f''(-0,65) > 0$, тобто виконана умова (4.22), то як нульове наближення беремо $x_0 = -0,65$. За формулою (4.21) обчислюємо послідовні наближення, зафіксовані в табл. 4.6.

З таблиці видно, що шуканий корінь $\xi = -0,694593$.

Зауваження. Два інших корені даного рівняння можна знайти аналогічно: спочатку відокремити кожний з них на відрізку достатньо малої довжини, потім обчислити за допомогою методу дотичних. Можна знайти і по-іншому. Розділивши многочлен $x^3 - 12x - 8$ на $(x - \xi)$, одержимо квадратне рівняння, корені якого можна взяти за початкові наближення двох ін-

ших коренів вихідного рівняння, а потім обчислити їх за методом дотичних.

Таблиця 4.6 – Послідовні наближення

n	x_n	$f(x_n)$	$\frac{1}{f'(x_n)}$	$\Delta = \frac{f(x_n)}{f'(x_n)}$	x_{n+1}
0	- 0,65	- 0,474625	- 0,093175	0,044223	- 0,694223
1	- 0,694223	- 0,003902	- 0,094759	0,000370	- 0,694593
2	- 0,694593	- 0,000003	- 0,095711	0,0000003	- 0,694593

Приклад 2. З точністю до п'яти десяткових знаків після коми знайти додатний корінь рівняння $th x+x^2-3,487=0$.

Графічно і методом проб визначаємо, що додатний корінь даного рівняння знаходиться на відрізку $[1,6; 1,7]$, тому що $f(1,6)=-0,00533 < 0$, $f(1,7) = 0,33241 > 0$. Як початкове наближення візьмемо лівий кінець відрізка $[1,6; 1,7]$, тобто $x_0=1,6$. Перевіримо, чи виконуються умови теореми, сформульованої на початку даного підрозділу, для функції $f(x)=thx+x^2-3,487$ і початкового значення кореня $x_0=1,6$.

Випишемо попередньо вирази для першої і другої похідної даної функції: $f'(x) = 1/ch^2x + 2x$, $f''(x) = 2(1 - shx / ch^3x)$. При $x_0=1,6$:

$$f'(x_0)=f'(1,6)=3,35 \neq 0, |1/f'(x_0)| < 0,3;$$

$$y=0,3 \cdot |f(x_0)/f'(x_0)| = 0,0053/3,35 < 0,0053 \cdot 0,3 < 0,0017, \quad \eta=0,0017,$$

$$|f''(x)| \leq 2 \text{ для усіх } x \geq 0; K = 2;$$

$$h = BK\eta = 0,3 \cdot 2 \cdot 0,0017 = 0,00102 < 0,5.$$

Таким чином, усі умови теореми виконано. Отже, на відрізку $|x-x_0| < 2\eta=0,0034$, тобто на відрізку $[1,5966; 1,6034]$, розташовано корінь даного рівняння. Щоб обчислити його із заданою точністю, за допомогою формули (4.24) визначимо число n . За умовою задачі n має бути таким, при якому виконується нерівність $1/2^{n-1} \cdot (2h)^{2^{n-1}} \eta \leq 0,5 \cdot 10^{-5}$ або $1/2^{n-1} \cdot (0,002)^{2^{n-1}} 0,0017 \leq 0,5 \cdot 10^{-5}$, оскільки $\eta = 0,0017$ і $h = 0,001$. Ця нерівність виконується при $n = 2$. За формулою (4.21) знаходимо x_1 і x_2 :

$$x_1 = 1,6 - \frac{f(1,6)}{f'(1,6)} = 1,6 - \frac{0,0053}{3,350524} = 1,601590,$$

$$x_2 = 1,601590 - \frac{0,00003}{3,353298} = 1,601581.$$

Отже, корінь рівняння $\xi = 1,601581$.

4.5.3 Комбінований метод хорд та дотичних

У методах хорд і дотичних вихідний відрізок $[a, b]$, на якому знаходиться корінь рівняння $f(x)=0$ замінюється відрізком $[a, x_1]$ або $[x_1, b]$, де x_1 – точка, що розташована ближче до кореня, ніж кінці відрізка, і т. д. Очевидно, більший ефект можна отримати при наближенні до кореня одночасно з двох сторін. Це досягається застосуванням комбінованого методу, що полягає в одночасному використанні методів хорд і дотичних.

Комбінований метод зручно застосовувати, якщо на вихідному відрізку $[a, b]$ друга похідна $f''(x)$ зберігає знак. У цьому випадку дотична перетинає вісь Ox з боку опуклості, а хорда – з боку увігнутості графіка функції $y=f(x)$ (рис. 4.11). Наближення за методом дотичних будуть розташовуватися з однієї сторони кореня, а наближення за методом хорд – з іншої. Таким чином, одержуємо усе вужчі відрізки, усередині яких розташований корінь. Довжина останнього з відрізків дає величину абсолютної похибки наближеного значення кореня. Одночасне оцінювання похибки одержуваного наближення дає велику зручність під час застосування комбінованого методу.

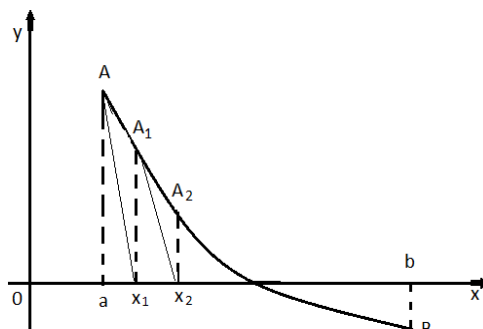


Рисунок 4.11

Приклад 1. Комбінованим методом (з точністю до шести десяткових знаків) розв'язати рівняння $x^3+4x+3=0$.

Графічно відокремлюючи корені рівняння, бачимо, що єдиний дійсний корінь даного рівняння розташований на відрізку $[-1; 0]$. Табличним способом, за допомогою схеми Горнера, знаходимо відрізок меншої довжини, якому належить корінь (табл. 4.7).

Таблиця 4.7 – Уточнення коренів за схемою Горнера

$x \backslash a_k$	1	0	4	3	$f(x)$
0,6	1	-0,6	4,36	0,384	+
0,7	1	-0,7	4,49	-0,143	-

Отже, корінь знаходиться на відрізку $[-0,7; -0,6]$. Перевіримо, для якої з цих двох точок виконується умова $f(x_0) \cdot f''(x_0) > 0$. Оскільки $f''(x)=6x$, $f''(-0,7)=-0,42$, $f(-0,7)=-0,143$ і $f(-0,7) \cdot f''(-0,7) > 0$, то, застосовуючи метод Ньютона, потрібно покласти $x_0 = -0,7$. За методом Ньютона при $x = -0,7$ знаходимо:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = -0,7 - \frac{-0,143}{5,47} = -0,673858,$$

$$x_2 = x_1 - f(x_1) / f'(x_1) = -0,673858 + 0,001421 / 5,362254 = -0,673593.$$

За методом хорд, поклавши $x'_0 = -0,6$, $a = -0,7$, одержимо:

$$x'_1 = \frac{x'_0 f(a) - a f(x'_0)}{f(a) - f(x'_0)} = \frac{-0,6 \cdot (-0,143) + 0,7 \cdot 0,384}{-0,143 - 0,384} = -0,672865,$$

$$x'_2 = \frac{x'_1 f(a) - a f(x'_1)}{f(a) - f(x'_1)} = \frac{-0,672865 \cdot (-0,143) + 0,7 \cdot 0,003902}{-0,143 - 0,003902} = -0,673585.$$

Отже, корінь знаходиться на відрізку $[-0,67350; -0,673585]$. Покладаючи, що $a = -0,673593$, користуючись формулою

$$x'_3 = x'_2 - \frac{f(x'_2)}{f(a) - f(x'_2)} (a - x'_2),$$

одержуємо: $x'_3 = -0,673585 + 1 \cdot 0,000008 = -0,673593$. Таким чином, корінь даного рівняння $\xi = -0,673593$.

Приклад 2. Комбінованим методом знайти менший додатний корінь рівняння $x^3 - 3x^2 + 1 = 0$ (з точністю до шести десяткових знаків).

Шуканий корінь знаходиться на відрізку $[0,6; 0,7]$. Для методу Ньютона $x_0 = 0,7$, для методу хорд $x'_0 = 0,6$, $b = 0,7$.

За відповідними формулами знаходимо:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 0,7 - \frac{-0,127}{-2,73} = 0,653480,$$

$$x'_1 = \frac{x'_0 f(b) - b f(x'_0)}{f(b) - f(x'_0)} = \frac{0,6 \cdot (-0,127) - 0,7 \cdot 0,136}{-0,127 - 0,136} = 0,651711,$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 0,653480 - \frac{-0,002049}{-2,639772} = 0,652704,$$

$$x'_2 = \frac{x'_1 f(b) - b f(x'_1)}{f(b) - f(x'_1)} = \frac{0,651711 \cdot (-0,002049) - 0,653480 \cdot 0,002618}{-0,002049 - 0,002618} = 0,652668,$$

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = 0,652704 - \frac{-0,0000006}{-2,6381579} = 0,6527037,$$

$$x'_3 = x'_2 - \frac{f(x'_2)}{f(x_2) - f(x'_2)} (x_2 - x'_2) = 0,652668 -$$

$$-\frac{0,0000944}{-0,0000006 - 0,0000944} (0,652704 - 0,652668) = 0,65270377.$$

Отже, менший додатний корінь ξ даного рівняння задовольняє нерівність $0,65270377 < \xi < 0,65270378$.

Зауваження. Значення функції $f(x)=x^3-3x^2+1$ і її похідної $f'(x)=3x^2-6x$ у точках $x_0, x_1, x_2, x'_0, x'_1, x'_2$ обчислені за допомогою схеми Горнера (табл. 4.8).

Таблиця 4.8 – Значення функції і її похідної, обчислені за схемою Горнера

$\begin{matrix} a_k \\ x \end{matrix}$	1	-3	0	1
0,7	1 1	-2,3 -1,6	-1,61 -2,73 ($f'(x_0)$)	-0,127 = $f(x_0)$
0,653480	1 1	-2,34652 -1,69304	-1,533404 -2,6397772 ($f'(x_1)$)	-0,002049 = $f(x_1)$
0,652704	1 1	-2,347296 -1,69304	-1,532089 -2,6381579	-0,0000006 = $f(x_2)$
0,6	1	-2,4	-1,44	0,136 = $f(x'_0)$
0,651711	1	-2,348289	-1,530406	0,002618 = $f(x'_1)$
0,652668	1	-2,347332	-1,532028	0,0000944 = $f(x'_2)$

4.5.4 Метод ітерацій

Нехай будь-яким способом, наприклад графічним чи табличним, отримано наближене значення x_0 кореня рівняння

$$f(x) = 0. \quad (4.25)$$

Уточнення кореня можна одержати за так званим методом послідовних наближень чи ітерацій. Для цього рівняння (4.25) подамо у вигляді

$$x = \varphi(x), \quad (4.26)$$

що завжди можна зробити і навіть багатьма способами, наприклад,

$$x = x + cf(x), \quad (4.27)$$

де c – довільна константа.

Нехай число x_1 є результатом підстановки x_0 у праву частину рівності (4.26): $x_1 = \varphi(x_0)$, далі $x_2 = \varphi(x_1)$ і т. д., x_n отримуємо з x_{n-1} за формулою

$$x_n = \varphi(x_{n-1}). \quad (4.28)$$

Цей процес послідовного обчислення чисел x_n ($n = 1, 2, \dots$) за формулою (4.28) називається методом послідовних наближень або ітерацій.

Умови, при яких цей процес збігається (тобто при збільшенні n можна одержати наближення, яке відрізняється від дійсного значення кореня ξ не більше, ніж на величину похибки), визначаються такою теоремою.

Якщо на відрізку, що містить корінь ξ рівняння (4.25), а також усі його послідовні наближення $x_0, x_1, x_2, \dots, x_n$, що обчислюються за методом ітерацій, виконана умова

$$|\varphi'(x)| \leq m < 1, \quad (4.29)$$

то $\lim_{n \rightarrow \infty} x_n = \xi$, тобто процес ітерацій збігається.

Геометрична інтерпретація

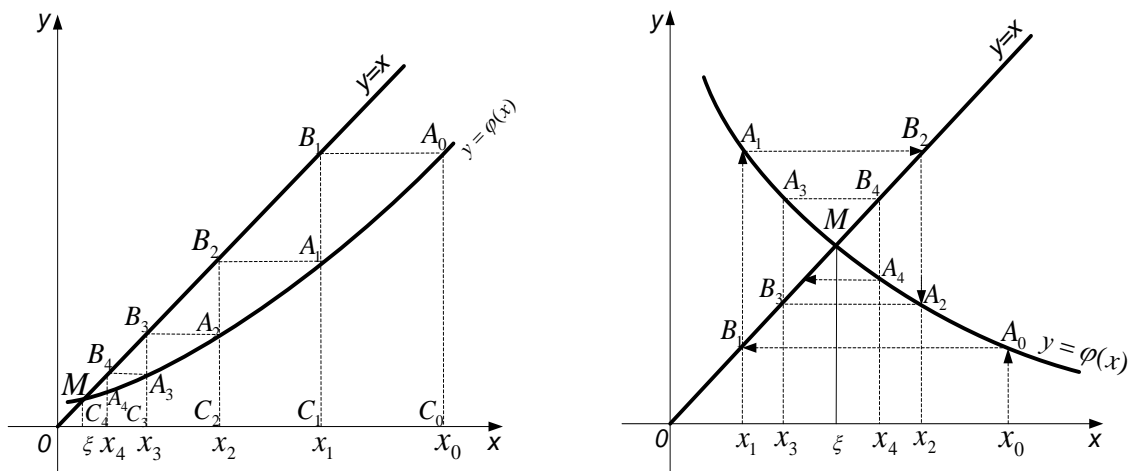
Нехай задано рівняння $f(x)=0$, де $f(x)$ – неперервна функція. Зведемо це рівняння до вигляду $x=\varphi(x)$ і побудуємо графіки функцій $y=x$ і $y=\varphi(x)$. Абсциса точки перетину графіків цих функцій є істинним коренем ξ (рис. 4.12).

Виберемо $x_0 \in [a, b]$ і визначимо $\varphi(x_0)$. Послідовність точок, що знаходяться на кривій $y=\varphi(x)$, позначити через A_i ($i=0,1,2,\dots$), а послідовність точок, що лежать на прямій $y=x$ – через B_i ($i=1,2,3,\dots$)

З точки $A_0(x_0; \varphi(x_0))$ проведемо пряму, паралельну осі ОХ до перетину з прямою $y=x$ і отримаємо точку $B_1(x_1; \varphi(x_0))$.

Дійсно, $A_0C_0 = \varphi(x_0) = B_1C_1$, оскільки $A_0B_1 \parallel OC_0$, $B_1C_1 \parallel A_0C_0$. Але $OC_1 = B_1C_1$ ($\triangle OC_1B_1$ – прямокутний і рівнобедрений, оскільки пряма $y=x$ є бісектрисою координатного кута). Отже, $x_1 = \varphi(x_0)$.

Проведемо $A_1B_2 \parallel OC_1$ і, повторюючи вищесказані судження, переконаємось, що $x_2 = \varphi(x_1)$.



а)

б)

Рисунок 4.12 – Геометрична інтерпретація ітераційного процесу

На рис. 4.12, а) побудовано збіжний ітераційний процес. Крива перетинає бісектрису $y=x$ в точці M з абсцисою ξ і при $x > \xi$ лежить під бісек-

трисою, а $\varphi'(x)$ задовольняє умову $0 < \varphi'(x) < 1$. Послідовність наближень $x_0, x_1, \dots, x_n, \dots$ (загальні абсциси точок графіків обох функцій) *монотонно спадає*. Кожне наступне наближення x_n ближче до істинного кореня, ніж попереднє x_{n-1} . Ламана лінія $A_0B_1A_1B_2A_2\dots$ має вигляд “драбини”.

На рис. 4.12, б) похідна $\varphi'(x) < 0$, але за абсолютною величиною менша одиниці $|\varphi'(x)| < 1$. Ітераційний процес також збігається, але наближення *коливаються* біля точного значення кореня. Ламана лінія $A_0B_1A_1B_2A_2\dots$ має вигляд “спіралі”.

Отже, якщо в деякій ділянці (a, b) кореня ξ рівняння $x = \varphi(x)$ похідна $\varphi'(x)$ зберігає постійний знак і виконана нерівність $|\varphi'(x)| \leq q < 1$, причому $\varphi'(x) > 0$, то послідовні наближення

$$x_n = \varphi(x_{n-1}) (n=1, 2, \dots), \quad x_0 \in [a, b]$$

збігаються до кореня монотонно. В такому випадку, коли $\varphi'(x) < 0$, послідовні наближення коливаються біля ξ .

На рис. 4.13, а) показано розбіжний ітераційний процес. Тут $\varphi'(x) > 1$. Крива перетинає бісектрису $y=x$ в точці M при $x > \xi$, що лежить над бісектрисою.

На рис. 4.13, б) показано розбіжний ітераційний процес для випадку $|\varphi'(x)| > 1$. Послідовні наближення віддаляються від точного значення кореня ξ .

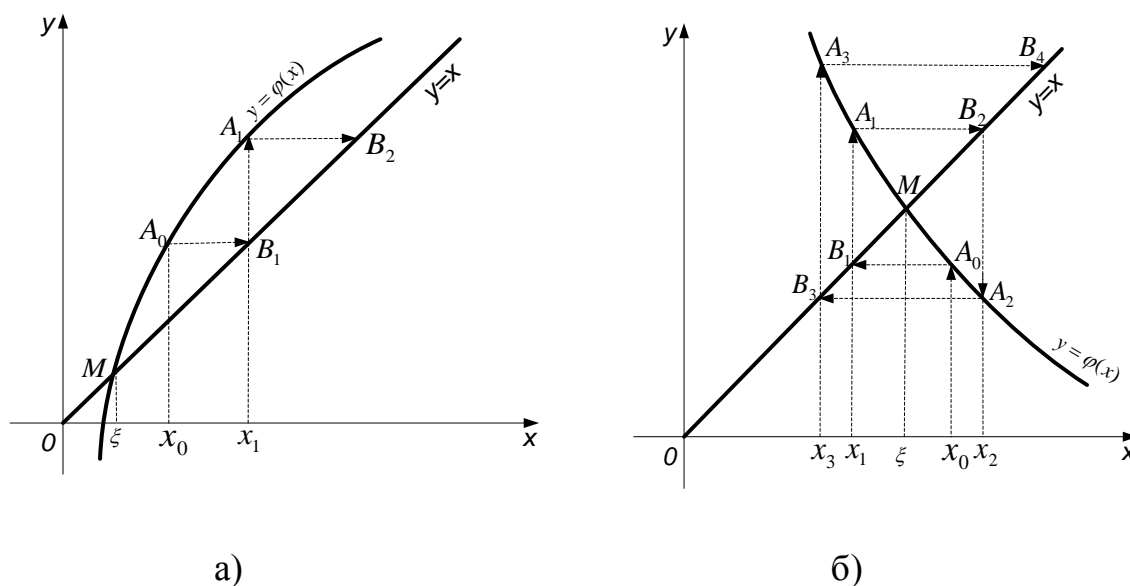


Рисунок 4.13 – Розбіжний ітераційний процес: а) $\varphi'(x) > 1$, б) $|\varphi'(x)| > 1$

Приклад 1. Методом ітерацій знайти менший додатний корінь рівняння $x^3 - 5x + 1 = 0$.

Графічно відокремлюючи корені даного рівняння, знаходимо, що рівняння має три дійсних корені, що належать відрізкам $[-3; -2]$, $[0; 1]$, $[2; 3]$. Знайдемо менший додатний корінь, що належить відрізку $[0; 1]$. Укажемо відрізок меншої довжини, на якому знаходиться корінь. Оскільки $f(0) = 1 > 0$, $f(1/2) = 1/8 - 5/2 + 1 = -11/8 < 0$, корінь знаходиться на відрізку $[0; 0,5]$

Дане рівняння зведемо до вигляду (4.26): $x = 1/5 (x^3 + 1)$ або $x = \varphi(x)$, де $\varphi(x) = 1/5(x^3 + 1)$. Оскільки $\varphi'(x) = 3x^2/5$, $\max_{0 \leq x \leq 0,5} |\varphi'(x)| = 3/20 < 1$, то умова

(4.29) виконується і процес ітерацій буде збігатися. Взавши як нульове наближення середину відрізка $[0; 0,5]$, тобто прийнявши $x_0 = 0,25$, обчислення наступних наближень виконаємо за формулою $x_{n+1} = \frac{1}{5}(x_n^3 + 1)$.

Результати цих обчислень подані у таблиці 4.9, з якої видно, що шуканий корінь $\xi = 0,20164$.

Зауваження. Під час знаходження двох інших коренів методом ітерацій уже не можна застосовувати формулу $x = \varphi(x)$, де $\varphi(x) = \frac{1}{5}(x^3 + 1)$, тому

що $\max_{2 < |x| < 3} |\varphi'(x)| = \max_{2 < |x| < 3} \left| \frac{3x^2}{5} \right| = \frac{27}{5} > 1$. У цьому випадку дане рівняння слід

показати в іншому вигляді, наприклад $x = \sqrt[3]{5x - 1}$; для функції $\varphi(x) = \sqrt[3]{5x - 1}$ на відрізках: $[-3; -2]$, $[2; 3]$ умова (4.29) буде виконуватися.

Таблиця 4.9 – Результати обчислень послідовних наближень

n	x_n	X_n^3	$X_n^3 + 1$	$\frac{1}{5}(X_n^3 + 1) = x_{n+1}$
0	0,25	0,01563	1,01563	0,20313
1	0,20313	0,00838	1,00838	0,20168
2	0,20168	0,00821	1,00821	0,20164
3	0,20164	0,00820	1,00820	0,20164

Приклад 2. Методом ітерацій знайти від'ємний корінь рівняння $x^4 + x - 3 = 0$.

Дане рівняння має два дійсних корені; від'ємний корінь знаходиться на відрізку $[-1,5; -1,4]$, оскільки для його кінців виконується умова $f(-1,5) \cdot f(-1,4) < 0$. Рівняння запишемо у вигляді $x = x + 3(x^4 + x - 3)$, де c – константа.

Виберемо значення постійної так, щоб для функції $\varphi(x)=x+c(x^4+x-3)$ виконувалася умова $\max_{-1,5 < x < -1,4} |\varphi'(x)| = q < 1$. Як таке значення можна взяти $c=0,1$; тоді $\varphi(x) = 0,1x^4 + 1,1x - 0,3$, $\varphi'(x) = 0,4x^3 + 1,1$,

$$\max_{-1,5 < x < -1,4} |\varphi'(x)| = 0,25 < 1$$

Прийнявши $x_0 = -1,45$, обчислимо наступні наближення за формулою

$$x_{n+1} = 0,1x_n^4 + 1,1x_n - 0,3$$

і запишемо результати (табл. 4.10), де $\xi = -1,45262$ – корінь рівняння.

Таблиця 4.10 – Результати обчислень послідовних наближень

N	x_n	x_n^2	x_n^4	$0,1x_n^4$	$1,1x_n$	x_{n+1}
0	-1,45	2,1025	4,4206	0,44205	-1,595	-1,45295
1	-1,45295	2,11106	4,45657	0,44566	-1,59825	-1,45259
2	-1,45259	2,11002	4,45218	0,44522	-1,59785	-1,45263
3	-1,45263	2,11013	4,45265	0,44527	-1,59789	-1,45262
4	-1,45262	2,11011	4,45256	0,44526	-1,59788	-1,45262

4.6 Розв’язання систем нелінійних рівнянь

На відміну від систем лінійних рівнянь для систем нелінійних рівнянь (СНР) не існує прямих методів розв’язання, і тому завжди застосовуються ітераційні методи [8, 11]. У загальному випадку система подається у вигляді

$$\begin{aligned} f_1(x_1, x_2, \dots, x_n) &= 0; \\ f_2(x_1, x_2, \dots, x_n) &= 0; \\ &\dots\dots\dots \\ f_n(x_1, x_2, \dots, x_n) &= 0. \end{aligned}$$

Саме до такого вигляду зводяться багато задач в автоматичі, електроніці, інформаційно-вимірювальній техніці, наприклад, задачі про розробку моделі електронних пристроїв тощо.

4.6.1 Метод простої ітерації

Метод простої ітерації (або послідовних наближень) для розв’язання СНР є розвитком аналогічного методу для розв’язання одного рівняння [8]. Він базується на припущенні, що систему можна звести до вигляду

$$x_1^* = q_1(x_1, x_2, \dots, x_n); x_2^* = q_2(x_1^*, x_2, \dots, x_n); \dots; x_n^* = q_n(x_1^*, x_2^*, \dots, x_n).$$

Структура алгоритму показана на рис. 4.14.

Цей різновид методу простої ітерації побудований аналогічно методу Зейделя для систем лінійних рівнянь, коли для обчислення кожного наступного значення невідомих використовуються нові знайдені значення попередніх невідомих.

Тепер детально зупинимось на основних кроках процесу розв'язання СНР з відповідним прикладом [10]. Розглянемо векторне рівняння

$$F(x)=0, \quad (4.30)$$

де F – вектор-функція,
 x – вектор.

$$F=[F_1, F_2, \dots, F_n]^T.$$

Перетворимо рівняння (4.30) до вигляду

$$x=g(x), \quad (4.31)$$

де $g=[g_1, g_2, \dots, g_n]^T$.

Наближений процес, який розв'язує рівняння (4.31), запишемо так

$$x^{k+1}=g(x^k), \quad k=0, 1, 2, \dots \quad (4.32)$$

Відображення $g: X \rightarrow X$, де X – банахів простір, називається стисненим, якщо для будь-яких $x^1, x^2 \in X$ виконується рівність

$$\|g(x^1)-g(x^2)\| \leq q\|x^1-x^2\|; \quad q < 1. \quad (4.33)$$

Якщо відображення $g: X \rightarrow X$ – стиснене, то рівняння (4.32) має єдиний розв'язок x^* :

$$\|x^*-x^m\| \leq q^m/(1-q)\|x^1-x^0\|. \quad (4.34)$$

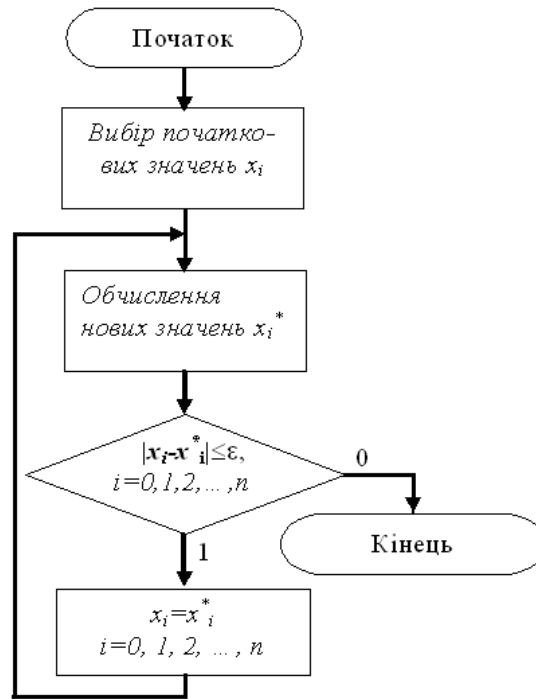


Рисунок 4.14

У достатньо малій області розв’язку x^* рівняння (4.31) для наближень методом простої ітерації

$$x^{m+1} - x^* = g(x^m) - g(x^*) \approx B(x^m - x^*), \quad (4.35)$$

де $B = \left[\frac{\partial g_i}{\partial x_j} \right] \bigg|_{x^*}; i, j = \overline{1, n}$

Таким чином, умова збіжності методу послідовних наближень має вигляд

$$\|B\| = q < 1. \quad (4.36)$$

У цьому випадку відображення g буде стисненим, а наближення x^k будуть збігатися до розв’язку x^* зі швидкістю геометричної прогресії.

Аналогом методу Зейделя при розв’язуванні нелінійних систем рівнянь є наближений процес, згідно з яким послідовні наближення визначаються із відношень

$$\begin{cases} x_1^{k+1} = g_1(x_1^k, x_2^k, \dots, x_n^k); \\ x_2^{k+1} = g_2(x_1^{k+1}, x_2^k, \dots, x_n^k); \\ \dots \\ x_i^{k+1} = g_i(x_1^{k+1}, \dots, x_{i-1}^{k+1}, x_i^k, \dots, x_n^k); \\ \dots \\ x_n^{k+1} = g_n(x_1^{k+1}, x_2^{k+1}, \dots, x_{n-1}^{k+1}, x_n^k). \end{cases} \quad (4.37)$$

Метод послідовних наближень є методом першого порядку, тобто $\|x^{k+1} - x^*\| \leq C \|x^k - x^*\|$, C – стала величина.

Під час розв’язування систем нелінійних рівнянь необхідно визначити прийнятне початкове наближення. Для випадку двох рівнянь з двома невідомими початкові наближення знаходять графічно. Для випадків систем рівнянь з великим числом рівнянь відокремлення коренів може бути виконано іншими методами.

Приклад 1. Розглянемо систему рівнянь

$$\left. \begin{aligned} \sin(x+1) - y &= 0,5; \\ x^2 + y^2 &= 1. \end{aligned} \right\}$$

Дана система описує графіки функцій, що зображені на рис 4.15. Знайдемо її розв’язок за методом Зейделя (4.37). Запишемо систему рівнянь (4.38) у вигляді (4.39):

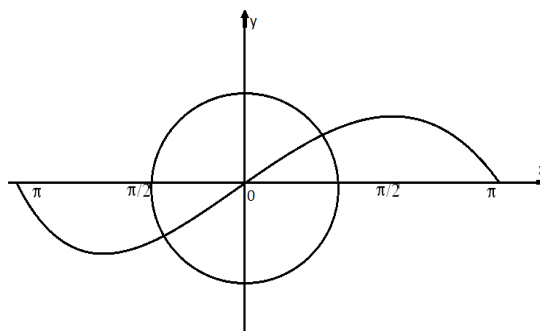


Рисунок 4.15

$$\left. \begin{aligned} x &= g_1(x, y) = \sqrt{1 - y^2}; \\ y &= g_2(x, y) = \sin(x + 1) - 0,5. \end{aligned} \right\} \quad (4.39)$$

Для визначення початкового наближення x^0, y^0 , що належить області збіжності $D = \{a < x < b, c < y < d\}$, побудуємо графіки функцій:

$$\begin{aligned} \text{при } x=1 \quad y &= \sin 2 - 0,5 = 0,41; \\ \text{при } x=0,8 \quad y &= \sin 1,9 - 0,5 = 0,47. \end{aligned}$$

Розв'язок задачі будемо шукати в області $D = \{0,8 < x < 1, 0,41 < y < 0,47\}$.
Визначимо матрицю B :

$$\frac{\partial g_1}{\partial x} = \cos(x + 1); \quad \frac{\partial g_2}{\partial y} = -\frac{y}{\sqrt{1 - y^2}}; \quad \frac{\partial g_1}{\partial y} = \frac{\partial g_2}{\partial x} = 0;$$

$$B = \begin{bmatrix} \cos(x + 1) & 0 \\ 0 & -y / \sqrt{1 - y^2} \end{bmatrix};$$

$$\|B\|_1 = \max_{x, y \in D} \left\{ \cos(x + 1); \frac{y}{\sqrt{1 - y^2}} \right\} = \max \left\{ \cos 1,14; \frac{0,41}{\sqrt{1 - 0,41^2}} \right\} = q_1 = 0,45 < 1.$$

Таким чином, процес послідовних наближень (4.39) збігається в області D . Переходимо до реалізації процесу

$$\begin{aligned} x^{k+1} &= g_1(x^k, y^k) = \sqrt{1 - (y^k)^2}; \\ y^{k+1} &= g_2(x^{k+1}, y^k) = \sin(x^{k+1} + 1) - 0,5. \end{aligned}$$

Вибираємо $x^0 = 0,85; y^0 = 0,46;$

$$x^1 = \sqrt{1 - (y^0)^2} = 0,887; \quad y^1 = \sin(x^1 + 1) - 0,5 = 0,450;$$

$$x^2 = \sqrt{0,7975} = 0,893; \quad y^2 = \sin 1,893 - 0,5 = 0,448;$$

$$x^3 = \sqrt{0,7973} = 0,894; \quad y^3 = 0,448;$$

Перевірка: $x^2_1 + y^2_1 = 0,2007 + 0,7992 = 0,9999$.

Таким чином, з точністю до $\Delta = 0,001$ розв'язок системи (4.38) в області D такий: $x^* = 0,895; y^* = 0,448$.

Зауважимо, що метод простої ітерації має свої недоліки [8]. Наприклад, якщо вихідні значення невідомих значно відрізняються від правильного розв'язку, то процес не буде збіжним. Зі збільшенням числа невідомих область збіжності (область, у якій задані вихідні значення збігаються до розв'язку) зменшується, і у випадку великих систем збіжність

забезпечується лише тоді, коли вихідні значення дуже близькі до шуканого розв'язку.

4.6.2 Метод Ньютона

Метод Ньютона є найрозповсюдженішим методом розв'язання СНР. Він забезпечує набагато швидшу збіжність, ніж метод простої ітерації. В його основу покладено зображення всіх n рівнянь у вигляді рядів Тейлора [8]:

$$f_1(x_1 + \Delta x_1, \dots, x_n + \Delta x_n) = f_1(x_1, \dots, x_n) + \Delta x_1 \frac{\partial f_1}{\partial x_1} + \dots + \Delta x_n \frac{\partial f_1}{\partial x_n} + R_n;$$

$$f_n(x_1 + \Delta x_1, \dots, x_n + \Delta x_n) = f_n(x_1, \dots, x_n) + \Delta x_1 \frac{\partial f_n}{\partial x_1} + \dots + \Delta x_n \frac{\partial f_n}{\partial x_n} + R_n,$$

де R_n – члени другого і вищих порядків.

Якщо прирости змінних Δx_i такі, що f_i набуває значення, близького до кореня, то будемо вважати, що ліві частини цих рівнянь перетворюються на нулі.

Таким чином, задача зводиться до пошуку такої сукупності приростів Δx_i , при якій досягається зазначена мета. Відкинувши R_n , задача зводиться до розв'язання системи лінійних рівнянь:

$$\begin{bmatrix} \partial f_1 / \partial x_1 & \dots & \partial f_1 / \partial x_n \\ \dots & \dots & \dots \\ \partial f_n / \partial x_1 & \dots & \partial f_n / \partial x_n \end{bmatrix} \cdot \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \dots \\ \Delta x_n \end{bmatrix} = \begin{bmatrix} -f_1 \\ -f_2 \\ \dots \\ -f_n \end{bmatrix}.$$

У цій системі матрицю частинних похідних, що називається матрицею Якобі, і вектор-стовпець правої частини можна оцінити для будь-якого наближеного розв'язку. Знайдені Δx_i використовуються як поправки до вихідного наближеного розв'язку

$$x_1 = x_1 + \Delta x_1; x_2 = x_2 + \Delta x_2; \dots; x_n = x_n + \Delta x_n.$$

Кінцева ітераційна формула має вигляд

$$x^{(n+1)} = x^{(n)} + W^{-1}(x^{(n)}) F(x^{(n)}),$$

де $x^{(n)}$, $x^{(n+1)}$ – вектори-стовпці наближених розв'язків системи відповідно на (n) і $(n-1)$ кроці ітераційного процесу;

$F(x^{(n)})$ – вектор-стовпець значень функцій f_1, f_2, \dots, f_n для наближення значень $x^{(n)}$;

$W^{-1}(x)$ – матриця, обернена матриці Якобі $W(x)$.

Якщо всі $\Delta x_i \rightarrow 0$, то розрахунок припиняється, в іншому випадку нові значення використовуються як наближені значення коренів, і процес по-

вторюється доти, доки не буде знайдено розв'язок або стане ясно, що процес є розбіжним.

Структура алгоритму показана на рис. 4.16.

Під час перевірки збіжності слід користуватися умовою

$$\Delta x_i / x_i \leq \varepsilon,$$

де ε – допустима похибка розв'язку.

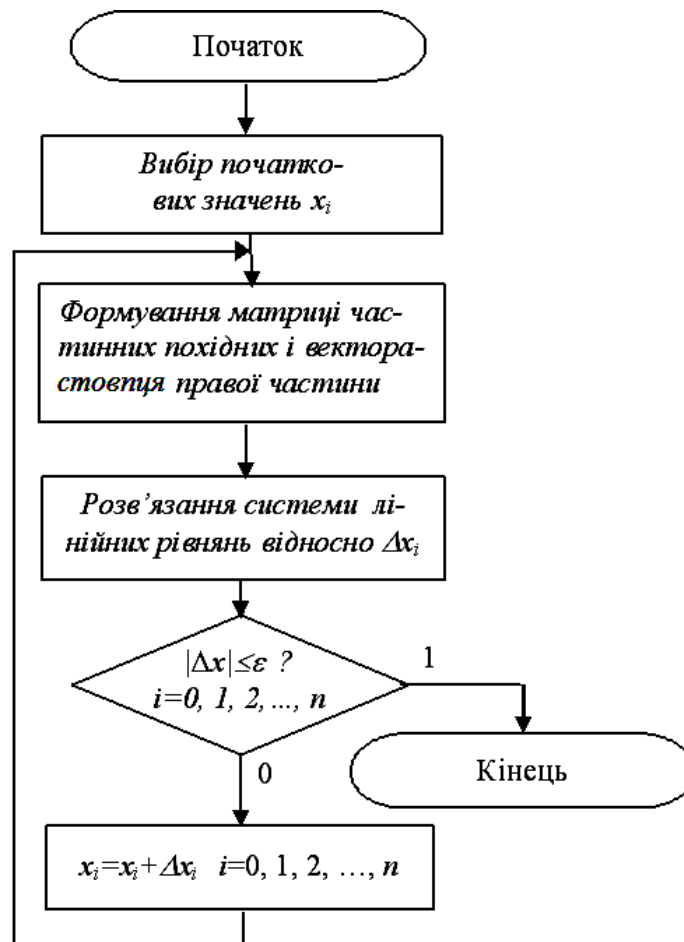


Рисунок 4.16 – Структура алгоритму

Далі детальніше зупинимося на обчислювальній схемі методу Ньютона [10]. Розглянемо нелінійний оператор $F: X \rightarrow X$, X – банахів простір і лінійний оператор $G: X \rightarrow X$.

Означення. Оператор G назвемо похідною оператора $F(x)$ в точці $x \in X$, якщо

$$\lim_{\substack{\|h\| \rightarrow 0 \\ h \in X}} \frac{\|F(x+h) - F(x) - Gh\|_X}{\|h\|_X} = 0. \quad (4.40)$$

Позначимо похідну оператора $F(x)$ через $F'(x)=G$. Нехай x^* – розв'язок рівняння (4.30), x^k – деяке наближення до розв'язку x^* . Якщо значення

$\|x^* - x^k\|/x$ мале, то у припущенні існування $F'(x)$ з урахуванням (4.40) напишемо: $F(x^k) + F'(x^k)(x^* - x^k) \approx F(x^*)$, або $F(x^k) + F'(x^k)(x^* - x^k) \approx 0$.

У припущенні існування $[F'(x)]^{-1}$

$$x^{k+1} = x^k - [F'(x^k)]^{-1} F(x^k). \quad (4.41)$$

Наближений процес (4.41) називається методом Ньютона. Позначимо область $D \subset X$: $D = \{x / \|x - x^*\| < a\}$.

Нехай виконуються такі умови: $F'(x)$ – оператор стиснення, тобто

$$\|F'(x^*) - F'(y)\|/x < q \|x - y\|/x, \quad x, y \in D; \quad (4.42)$$

$$\|F'(x)\|/x < b, \quad x \in D; \quad (4.43)$$

$$\|[F'(x_0)]^{-1} F(x_0)\|/x \leq c_0. \quad (4.44)$$

Тоді процес послідовних наближень (4.41) збігається до розв'язку x^* з оцінкою похибки $\|x^k - x^*\|/x \leq 1/d(d\|x^0 - x^*\|/x)^2$; $d = bq$.

При цьому

$$\|x^* - x_0\|/x \leq 1/d(1 - \sqrt{1 - 2dc_0}) = a_0 \leq a. \quad (4.45)$$

Звідси, визначивши a_0 , оцінюємо необхідну кількість N наближень для досягнення заданої точності Δ розв'язку x^* (4.31):

$$a_0^{2^N} < d\Delta \quad \text{або} \quad N < \log_2 p; \quad p = \frac{\lg(d\Delta)}{\lg a_0}. \quad (4.46)$$

Обчислювальна схема реалізації методу Ньютона (4.41) [10]:

1. Знайти $F'(x)$;
2. Знайти область $D = \{x / \|x - x^*\| < a\}$ графічно або аналітично при $n > 2$;
3. Для обраного початкового наближення $x^0 \in D$ обчислити

$$\|F'(x_0)\|/x = \max_{1 \leq i \leq n} \sum_{j=1}^n \left| \frac{df_i(x^0)}{dx_j^0} \right| = b_0; \quad (4.47)$$

$$4. \text{ Визначити} \quad [F'(x_0)]^{-1} = [\phi_{ij}(x^0)]_1^n; \quad (4.48)$$

$$5. \text{ Знайти} \quad \|[F'(x_0)]^{-1} F(x_0)\|/x = \max_{1 \leq i \leq n} |m_i(x^0)| = c_0; \quad (4.49)$$

6. Визначити перше наближення:

$$x_i^1 = x_i^0 - \sum_{j=1}^n \phi_{ij}(x^0) f_j(x^0); \quad i = \overline{1, n}; \quad (4.50)$$

$$7. \text{ Знайти з (4.42): } q_1 = \max_{l \leq i \leq n} |x_i^1 - x_i^0|; \quad (4.51)$$

$$q_2 = \max_{l \leq i \leq n} \left\{ \sum_{j=1}^n [f_{ij}(x^1) - f_{ij}(x^0)] \right\}; \quad (4.52)$$

$$q = q_2 / q_1; \quad (4.53)$$

8. Знайти коефіцієнт a_0 за формулою (4.45) $a_0 = (1 - \sqrt{1 - 2b_0 q c_0}) / b_0 q$;

9. Оцінити кількість наближень N : $p = \lg \frac{(b_0 q \Delta)}{\lg a_0}$; $N = \log_2 p$;

10. Реалізувати процес послідовних наближень (4.41) для $k=1, 2, \dots, N$.

Приклад 1. З точністю $\Delta=0,001$ знайти за методом Ньютона розв'язок системи рівнянь

$$\begin{cases} f_1(x, y) = x^2 + y^2 + 2x - 3 = 0; \\ f_2(x, y) = x^2 + y^2 + 4y - 5 = 0. \end{cases} \quad (4.54)$$

Знайдемо $F'(x, y)$:
$$F'(x, y) = \begin{bmatrix} 2(x+1)2y \\ 2x2(y-2) \end{bmatrix}.$$

Знайдемо область D . Оскільки графіки функцій $f_1(x, y)$, $f_2(x, y)$ – кола радіусом 2 і 3 відповідно з центрами в точках $(-1; 0)$ і $(0; 2)$, область D має вигляд $D = \{0,8 < x < 0,9; -0,9 > y > -1\}$. Як початкове наближення оберемо $x^0 = 0,8$; $y^0 = -0,9$.

Знайдемо матрицю $[F'(x_0)]^{-1}$:
$$[F'(x, y)]^{-1} = \frac{1}{\Delta(x, y)} \begin{bmatrix} y-2 & -y \\ -x & x+1 \end{bmatrix};$$

$$\Delta(x, y) = 2(y-2x) - 4.$$

Для системи (4.54) процес послідовних наближень (4.31) має вигляд:

$$\begin{aligned} x^{k+1} &= x^k + \frac{1}{\Delta(x^k, y^k)} [y^k f_2(x^k, y^k) - (y^k - 2)f_1(x^k, y^k)]; \\ y^{k+1} &= y^k + \frac{1}{\Delta(x^k, y^k)} [x^k f_1(x^k, y^k) - (x^k + 1)f_2(x^k, y^k)]; \end{aligned} \quad (4.55)$$

Розрахуємо b_0 : $\|F'(x^0)\|_1 = \max\{5, 4; 7, 4\} = 7, 4 = b_0$.

Знайдемо c_0 : $[F'(x^0)]^{-1} F(x^0) = \begin{bmatrix} 0,011 \\ -0,006 \end{bmatrix}$; $c_0 = 0,011$.

Визначимо x^1 і y^1 :

$$x^1 = 0,8 + 1/9(-0,9 \times 0,05 + 2,9 \times 0,5) = 0,8 + 0,0111 = 0,8111;$$

$$y^1 = -0,9 + 1/9(0,8 \times 0,05 - 1,8 \times 0,05) = -0,9 - 0,0056 = -0,9056.$$

Знайдемо q :

$$q_1 = \max\{0,0334; 0,0334\}; q_1 = 0,0334;$$

$$q_2 = \max\{0,0111; 0,0056\} = 0,0111; q = 0,0111 / 0,0334 = 0,332.$$

Розрахуємо радіус збіжності a_0 :

$$d = b_0 q = 2,4568; a_0 = 1/2,4568(1 - \sqrt{1 - 0,0545}) = 0,0113;$$

$$N \geq (-\lg \Delta - \lg d) / \lg a_0 = 1,4; N = 2.$$

Реалізуємо процес послідовних наближень (4.55):

$$f_1(x^1, y^1) = 1,479 + 1,6222 - 3 = 0,1012;$$

$$f_2(x^1, y^1) = 1,479 + 3,6224 - 5 = 0,1014;$$

$$\Delta(x^1, y^1) = -4 - 2 \times 2,5278 = -9,0556;$$

$$x^1 = 0,8111 - 1/9,0556(-0,9056 \times 0,1014 + 2,9056 \times 0,1012) = 0,7889;$$

$$y^1 = -0,9056 - 1/9,0556(0,8111 \times 0,1012 - 1,8111 \times 0,1014) = -0,8845.$$

Таким чином, один з розв'язків системи з точністю $\Delta=0,001$:

$$x^* = 0,789; y^* = -0,894.$$

Аналогічно можна знайти й інший розв'язок системи.

Зауважимо, що для методу Ньютона теж існує проблема збіжності. Величина області збіжності обернено пропорційна ступеню складності і кількості рівнянь. Існують способи оцінювання області збіжності і вибору вихідного наближення [8].

Якщо в цьому методі точні значення похідних знайти не вдається, то можна користуватися їх наближеними значеннями, знайденими методом січних.

4.6.3 Метод збурення параметрів

Метод збурення параметрів реалізовує алгоритм, що дозволяє за допомогою ітерацій одержати розв'язок СНР. Його ефективність не залежить від вибору початкового наближення. Суть цього методу полягає в такому [8]. Спочатку поряд із системою рівнянь

$$f_j(x_i) = 0, \quad i=1,2,\dots,n; \quad j=1,2,\dots,n,$$

розглянемо

$$g_j(x_i) = 0,$$

розв'язок якої відомий. Потім, змінюючи рівняння $g_j = 0$, перетворимо

його у рівняння $f_j = 0$ за допомогою скінченного числа послідовних малих збільшень параметрів:

$$g_j^{(k)}(x_i) = g_j^{(k-1)}(x_i) + [f_j(x_i) - g_j^{(k-1)}(x_i)] \frac{k}{N}, \quad (k=1, 2, \dots, N).$$

Розв'язки $x_i^{(0)}$ вихідної системи рівнянь $g_j^{(0)}(x_i)$ можна використовувати як вихідні значення змінних для ітераційного розв'язування системи $g_j^{(1)}(x_i)$. Ця система мало відрізняється від попередньої, і можна зробити висновок, що збіжність буде забезпечена. У процесі розрахунку розв'язок $x_i^{(k-1)}$ використовується як вихідний для одержання розв'язку x_i^k . Коли $k=N$, система перетворюється у вихідну. Оскільки для цього перетворення часто потрібно багато кроків, метод може потребувати великих затрат комп'ютерного часу. Але його перевага в тому, що при малій кількості кроків збіжність на кожному кроці досягається всього за кілька ітерацій [8].

4.7 Вибір методу розв'язання нелінійних рівнянь і систем

При виборі методу і розробці алгоритму розв'язання нелінійних рівнянь і систем в першу чергу слід враховувати порядок рівняння, характер зміни функції $f(x)$ (гладкість, розривність), вимоги до точності і швидкості отримання розв'язку, наявні ресурси комп'ютера [8].

Найшвидший метод, котрий дозволяє значно скоротити обсяг обчислень як для рівнянь, так і для систем – це метод Ньютона. При складності обчислення похідних він може бути замінений методом січних. Однак ці методи мають ряд обмежень і їх застосування ефективно лише для гладких функцій.

При невідомому характері зміни $f(x)$ найнадійнішим методом є метод половинного ділення.

Залежно від порядку рівняння може бути рекомендоване використання таких методів [8]: при порядку від 3 до 5 – Ньютона; від 6 до 85 – січних; понад 85 – спеціальні методи.

Замість складання власного алгоритму іноді зручно користуватися наявним у розпорядженні стандартним математичним забезпеченням комп'ютера або пакетами прикладних програм.

4.8 Контрольні запитання

1. Які рівняння називають нелінійними? Що означає розв'язати рівняння?
2. Наведіть класифікацію нелінійних рівнянь з прикладами.
3. Наведіть загальний алгоритм розв'язання нелінійних рівнянь.
4. Що означає відокремити корені рівняння? Наведіть методи відокремлення коренів.

кремлення коренів, а також їх переваги та недоліки.

5. Наведіть загальні властивості алгебраїчних рівнянь.
6. Наведіть основну теорему алгебри.
7. Покажіть з відповідними прикладами як визначити кількість дійсних коренів алгебраїчного рівняння.
8. Наведіть теорему Штурма та проілюструйте її застосування на відповідному прикладі.
9. Як за правилом кільця можна знайти область існування коренів алгебраїчного рівняння?
10. Як за правилом Лагранжа знайти область існування коренів алгебраїчного рівняння?
11. Наведіть правило Ньютона для знаходження області існування коренів алгебраїчного рівняння.
12. Наведіть відомі Вам методи уточнення коренів нелінійних рівнянь, а також їх переваги та недоліки.
13. Наведіть геометричну інтерпретацію методів послідовного наближення та половинного ділення. Які недоліки та переваги цих методів?
14. Наведіть загальний алгоритм методів хорд та Ньютона.
15. Наведіть геометричну інтерпретацію методів хорд та Ньютона.
16. Які модифікації методу Ньютона Ви знаєте? Дайте стислу характеристику цих методів та порівняйте з методом Ньютона.
17. Наведіть алгоритм комбінаційного методу уточнення коренів нелінійного рівняння.
18. Наведіть та дайте геометричну інтерпретацію ітераційного методу уточнення коренів нелінійного рівняння. Яка швидкість збіжності цього методу?
19. Дайте означення системи нелінійних рівнянь. Що називають розв'язком системи рівнянь?
20. Які методи розв'язання систем нелінійних рівнянь Ви знаєте?
21. Наведіть загальний алгоритм методу послідовних наближень для розв'язання систем нелінійних рівнянь (СНР).
22. Наведіть обчислювальну схему методу послідовних наближень для розв'язання СНР.
23. Наведіть алгоритм методу Ньютона для розв'язання СНР.
24. Наведіть стислий порівняльний аналіз відомих Вам методів розв'язання нелінійних рівнянь та їх систем.

4.9 Завдання

Задача 1. Напишіть програму відокремлення коренів нелінійного рівняння $f(x)=0$ згідно з варіантом, заданим у таблиці 4.11.

Таблиця 4.11 – Варіанти відокремлення коренів нелінійного рівняння

$f(x) \equiv P_4(x) = x_4 + a_3x^3 + a_2x^2 + a_1x + a_0$				
Вар.	a_3	a_2	a_1	a_0
1	-3,055105	-18,06895	4,002429	4,722482
2	-3,406111	10,89372	-1,752935	-3,423612
3	2,93309	9,274868	-10,32081	0,422098
4	16,28542	-2,771356	-27,95304	-11,33921
5	-3,055105	-18,06895	4,002429	4,722482
6	-3,406111	10,89372	-1,752935	-3,423612
7	2,93309	9,274868	-10,32081	0,422098
8	16,28542	-2,771356	-27,95304	-11,33921
9	-3,055105	-18,06895	4,002429	4,722482
10	-3,406111	10,89372	-1,752935	-3,423612
11	2,93309	9,274868	-10,32081	0,422098
12	16,28542	-2,771356	-27,95304	-11,33921
13	-3,055105	-18,06895	4,002429	4,722482
14	-3,406111	10,89372	-1,752935	-3,423612
15	2,93309	9,274868	-10,32081	0,422098
16	16,28542	-2,771356	-27,95304	-11,33921
17	-3,055105	-18,06895	4,002429	4,722482
18	-3,406111	10,89372	-1,752935	-3,423612
19	2,93309	9,274868	-10,32081	0,422098
20	16,28542	-2,771356	-27,95304	-11,33921
21	-3,406111	10,89372	-1,752935	-3,423612
22	2,93309	9,274868	-10,32081	0,422098
23	16,28542	-2,771356	-27,95304	-11,33921
24	-3,055105	-18,06895	4,002429	4,722482
25	-3,406111	10,89372	-1,752935	-3,423612
26	2,93309	9,274868	-10,32081	0,422098
27	16,28542	-2,771356	-27,95304	-11,33921
28	-3,055105	-18,06895	4,002429	4,722482
29	-3,055105	-18,06895	4,002429	4,722482
30	60,24546	-122,0716	105,6798	-30,19201

Задача 2. Напишіть програму уточнення коренів нелінійного рівняння, згідно з варіантом, заданим у таблиці 4.11, методом проб. Точність обчислення коренів $\varepsilon=10^{-3}$.

Задача 3. Напишіть програму уточнення коренів нелінійного рівняння, згідно з варіантом, заданим у таблиці 4.12, методом половинного ділення. Точність обчислення коренів $\varepsilon=10^{-4}$. Підрахуйте кількість ітерацій,

необхідних для досягнення заданої точності.

Таблиця 4.12 – Варіанти даних до задачі 3

Вар.	$f(x)$	$g(x)$	$[a, b]$
1	$(\sin x)^2 - \frac{5}{6}\sin x + \frac{1}{6}$	$(\sin x)^2 - \sin x + \frac{1}{4}$	[0;1]
2	$(\sin x)^2 + \frac{7}{12}\sin x + \frac{1}{12}$	$(\sin x)^2 + \frac{2}{3}\sin x + \frac{1}{9}$	[-1;0]
3	$(\sin x)^2 - \frac{1}{30}\sin x - \frac{1}{30}$	$(\sin x)^2 - \frac{2}{5}\sin x + \frac{1}{25}$	[-0,5;-0,5]
4	$(\cos x)^2 + \frac{2}{35}\cos x - \frac{1}{35}$	$(\cos x)^2 - \frac{2}{7}\cos x + \frac{1}{49}$	[0;2]
5	$(\cos x)^2 - \left(\frac{1}{\sqrt{2}} + \frac{1}{4}\right)\cos x + \frac{1}{4\sqrt{2}}$	$(\cos x)^2 - \frac{2}{\sqrt{2}}\cos x + \frac{1}{2}$	[0;1,5]
6	$(\cos x)^2 + \frac{1}{2}\cos x + \frac{1}{18}$	$(\cos x)^2 + \frac{1}{3}\cos x + \frac{1}{36}$	[0;2]
7	$(\ln x)^2 - 5\ln x + 6$	$(\ln x)^2 - 4\ln x + 4$	[5;25]
8	$(\ln x)^2 - \ln x - 2$	$(\ln x)^2 + 2\ln x + 1$	[0,1;10]
9	$(\ln x)^2 - \frac{3}{4}\ln x + \frac{1}{8}$	$(\ln x)^2 - \ln x + \frac{1}{4}$	[0,1;2]
10	$(\operatorname{tg} x)^2 + (\sqrt{3} - 1)\operatorname{tg} x - \sqrt{3}$	$(\operatorname{tg} x)^2 - 2\operatorname{tg} x + 1$	[1,2;1]
11	$(\operatorname{tg} x)^2 - 28/9 \cdot \operatorname{tg} x + 1/3$	$(\operatorname{tg} x)^2 - 6\operatorname{tg} x + 9$	[0;1,5]
12	$(\operatorname{tg} x)^2 - \frac{53}{6}\operatorname{tg} x - \frac{3}{2}$	$(\operatorname{tg} x)^2 - \frac{1}{3}\operatorname{tg} x + \frac{1}{36}$	[-0,5;1,5]
13	$x^4 - 7x^2 + 10$	$x^4 - 4x^2 + 4$	[0;3]
14	$x^4 - \frac{10}{3}x^2 + 1$	$x^4 - 6x^2 + 9$	[0;2]
15	$x^4 - \frac{13}{2}x^2 + 3$	$x^4 - x^2 + \frac{1}{4}$	[0;3]
16	$(\sin x)^2 + \frac{5}{6}\sin x + \frac{1}{6}$	$(\sin x)^2 + \frac{2}{3}\sin x + \frac{1}{9}$	[-1;0]
17	$(\sin x)^2 - \frac{7}{12}\sin x + \frac{1}{12}$	$(\sin x)^2 - \frac{1}{2}\sin x + \frac{1}{16}$	[0;1]
18	$(\sin x)^2 + \frac{1}{30}\sin x - \frac{1}{30}$	$(\sin x)^2 + \frac{1}{3}\sin x + \frac{1}{36}$	[-0,5;0,5]

Продовження таблиці 4.12

Вар.	$f(x)$	$g(x)$	$[a, b]$
19	$(\cos x)^2 - \frac{2}{35}\cos x - \frac{1}{35}$	$(\cos x)^2 - \frac{2}{5}\cos x + \frac{1}{25}$	[0;3]
20	$(\cos x)^2 + \left(\frac{1}{\sqrt{2}} - \frac{1}{4}\right)\cos x - \frac{1}{4\sqrt{2}}$	$(\cos x)^2 - \frac{1}{2}\cos x + \frac{1}{16}$	[0;2]
21	$(\cos x)^2 - \frac{1}{2}\cos x + \frac{1}{18}$	$(\cos x)^2 - \frac{2}{3}\cos x + \frac{1}{9}$	[0;2]
22	$(\lg x)^2 + \frac{5}{3}\lg x - \frac{2}{3}$	$(\lg x)^2 - \frac{2}{3}\lg x + \frac{1}{9}$	[0,001;3]
23	$(\lg x)^2 - \lg x - \frac{3}{4}$	$(\lg x)^2 - 3\lg x + \frac{9}{4}$	[0,1;35]
24	$(\lg x)^2 + \frac{3}{4}\lg x - \frac{1}{4}$	$(\lg x)^2 + 2\lg x + 1$	[0,01;3]
25	$(\operatorname{tg} x)^2 - \left(1 + \frac{1}{\sqrt{3}}\right)\operatorname{tg} x + \frac{1}{\sqrt{3}}$	$(\operatorname{tg} x)^2 - 2\operatorname{tg} x + 1$	[0;1]
26	$(\operatorname{tg} x)^2 - \frac{7}{4}\operatorname{tg} x - \frac{1}{2}$	$(\operatorname{tg} x)^2 + \frac{1}{2}\operatorname{tg} x + \frac{1}{16}$	[-0,5;1,5]
27	$(\operatorname{tg} x)^2 + \frac{37}{6}\operatorname{tg} x + 1$	$(\operatorname{tg} x)^2 + 12\operatorname{tg} x + 36$	[-1,5;0]
28	$x^4 - 11x^2 + 24$	$x^4 - 6x^2 + 9$	[1;3]
29	$x^4 - \frac{26}{5}x^2 + 1$	$x^4 - 10x^2 + 25$	[0;3]
30	$x^4 - \frac{21}{2}x^2 + 5$	$x^4 - x^2 + \frac{1}{4}$	[0;5]

Задача 4. Напишіть програму уточнення коренів нелінійного рівняння, заданого у задачі 3, методом хорд. Точність обчислення коренів $\varepsilon=10^{-5}$. Підрахуйте кількість ітерацій, необхідних для досягнення заданої точності.

Задача 5. Напишіть програму уточнення коренів нелінійного рівняння, згідно з варіантом, заданим у таблиці 4.13, методом Ньютона. Точність обчислення коренів $\varepsilon=10^{-5}$. Підрахуйте кількість ітерацій, необхідних для досягнення заданої точності.

Задача 6. Напишіть програму уточнення коренів нелінійного рівняння, згідно з варіантом, заданим у таблиці 4.12, комбінованим методом. Точність обчислення коренів $\varepsilon=10^{-5}$. Підрахуйте кількість ітерацій, необхідних для досягнення заданої точності.

Задача 7. Напишіть програму уточнення коренів нелінійного рівняння, заданого у задачі 5, методом ітерації. Точність обчислення коренів $\varepsilon=10^{-5}$. Підрахуйте кількість ітерацій, необхідних для досягнення заданої точності.

Таблиця 4.13 – Варіанти даних до задачі 5

$f(x) \equiv P_5(x) = x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$					
Вар.	a_4	a_3	a_2	a_1	a_0
1	4,545004	-3,055105	-18,06895	4,002429	4,722482
2	-2,656764	-3,406111	10,89372	-1,752935	-3,423612
3	-4,556062	2,93309	9,274868	-10,32081	0,422098
4	7,809249	16,28542	-2,771356	-27,95304	-11,33921
5	4,545004	-3,055105	-18,06895	4,002429	4,722482
6	-2,656764	-3,406111	10,89372	-1,752935	-3,423612
7	-4,556062	2,93309	9,274868	-10,32081	0,422098
8	7,809249	16,28542	-2,771356	-27,95304	-11,33921
9	4,545004	-3,055105	-18,06895	4,002429	4,722482
10	-2,656764	-3,406111	10,89372	-1,752935	-3,423612
11	-1,556062	2,93309	9,274868	-10,32081	0,422098
12	2,809249	16,28542	-2,771356	-27,95304	-11,33921
13	3,545004	-3,055105	-18,06895	4,002429	4,722482
14	-8,656764	-3,406111	10,89372	-1,752935	-3,423612
15	-4,556062	2,93309	9,274868	-10,32081	0,422098
16	7,809249	16,28542	-2,771356	-27,95304	-11,33921
17	4,545004	-3,055105	-18,06895	4,002429	4,722482
18	-2,656764	-3,406111	10,89372	-1,752935	-3,423612
19	-4,556062	2,93309	9,274868	-10,32081	0,422098
20	4,809249	16,28542	-2,771356	-27,95304	-11,33921
21	2,656764	-3,406111	10,89372	-1,752935	-3,423612
22	4,556062	2,93309	9,274868	-10,32081	0,422098
23	-7,809249	16,28542	-2,771356	-27,95304	-11,33921
24	-4,545004	-3,055105	-18,06895	4,002429	4,722482
25	-2,656764	-3,406111	10,89372	-1,752935	-3,423612
26	-4,556062	2,93309	9,274868	-10,32081	0,422098
27	7,809249	16,28542	-2,771356	-27,95304	-11,33921
28	-4,545004	-3,055105	-18,06895	4,002429	4,722482
29	4,545004	-3,055105	-18,06895	4,002429	4,722482
30	-13,0072	60,24546	-122,0716	105,6798	-30,19201

Задача 8. Написати програму для розв'язання системи нелінійних рівнянь
$$\begin{cases} y = x^2(a_1x + b_1) \\ x = y^2(a_2y + b_2) \end{cases}$$
 з точністю ε .

Значення параметрів даної системи і похибки ε наведені у таблиці 4.14.

Таблиця 4.14 – Варіанти для задачі 8

Варіант	a_1	a_2	b_1	b_2	ε
1	10	20	19	36	0.005
2	100	100	10	20	0.002
3	150	150	5	10	0.001
4	100	50	10	30	0.001
5	50	100	15	20	0.001

Задача 9. Написати програму для розв'язання системи нелінійних рівнянь
$$\begin{cases} x^2 + y^2 = R^2 \\ y = \frac{ax + b}{cx + d} \end{cases}$$

з точністю $\varepsilon = 10^{-4}$. Значення параметрів системи наведено у таблиці 4.15

Таблиця 4.15 – Варіанти для задачі 9

Варіант	R	a	b	c	d
1	5	4	4	1	1
2	5	5	5	1	2
3	5	7	7	2	1
4	5	5	5	2	-1
5	5	8	4	2	1

Задача 10. Написати програму для розв'язання системи нелінійних рівнянь
$$\begin{cases} 3x^2 - 4y + z^2 = 0 \\ 2x^2 + y^2 - 4z = -1 \\ x^2/a^2 + y^2/b^2 + z^2/c^2 = d \end{cases}$$

з точністю $\varepsilon = 10^{-5}$. Значення параметрів системи наведені у таблиці 4.16

Таблиця 4.16 – Варіанти для задачі 10

Варіант	a	b	c	d
1	1	1	1	3
2	1	2	3	49/36
3	2	1	3	49/36
4	2	3	3	17/36
5	2	3	2	11/18

Задача 11. Написати програму для розв'язання системи нелінійних

$$\text{рівнянь } \begin{cases} x + y + z = 0 \\ cx + ay + bz = 0 \\ (x + b)^2 + (y + c)^2 + (z + a)^2 = a^2 + b^2 + c^2 \end{cases} \quad \text{з точністю } \varepsilon = 10^{-4}.$$

Значення параметрів системи наведені в таблиці 4.17.

Таблиця 4.17 – Варіанти для задачі 11

Вар.	a	b	c	Вар.	a	b	c	Вар.	a	b	c
1	5	4	3	3	2,5	4	5	5	2	2	1
2	5	2	4	4	1	2	3,5				

Задача 12. Написати програму для розв'язання нижченаведених систем нелінійних рівнянь із заданими початковими наближеннями та точністю:

$$\text{а) } \begin{cases} x = \lg(y/z) + 1 \\ y = 0,4 + z^2 - 2x^2 \\ z = 2 + xy/20 \end{cases} \quad \text{при } \begin{cases} x^{(0)} = 1,2; y^{(0)} = 2,4 \\ z^{(0)} = 2,1; \varepsilon = 10^{-4} \end{cases}.$$

$$\text{б) } \begin{cases} x + x^2 - 2yz = 0,1 \\ y - y^2 + 3xz = 0,2 \\ z + z^2 + 2xy = 0 \end{cases} \quad \text{при } \begin{cases} x^{(0)} = y^{(0)} = z^{(0)} = 0 \\ \varepsilon = 10^{-5} \end{cases}.$$

$$\text{в) } \begin{cases} 3x^2 + 1,5y^2 + z^2 - 5 = 0 \\ 6xyz - x + 5y + 3z = 0 \\ 5xz - yz - 1 = 0 \end{cases} \quad \text{при } \begin{cases} x^{(0)} = 1,25; y^{(0)} = 0 \\ z^{(0)} = 0,25; \varepsilon = 10^{-4} \end{cases}.$$

Задача 13. Написати програму для розв'язання системи нелінійних

$$\text{рівнянь } \begin{cases} x = \cos y + 0,85 \\ y = \sin x - 1,32 \end{cases} \quad \text{з точністю } \varepsilon = 10^{-4}.$$

Задача 14. Написати програму для розв'язання системи нелінійних

$$\text{рівнянь методом Ньютона } \begin{cases} x_1^2 + x_2^2 + x_3^2 - 1 = 0 \\ 2x_1^2 + x_2^2 + 4x_3 = 0 \\ 3x_1^2 - 4x_2 - x_3^2 = 0 \end{cases} \quad \text{з точністю } \varepsilon = 10^{-5}.$$

За початкове наближення розв'язку взяти $x_1^{(0)} = x_2^{(0)} = x_3^{(0)} = 0,5$

Задача 15. Методом Ньютона з точністю $\varepsilon = 10^{-6}$ знайти усі корені сис-

теми рівнянь $\begin{cases} f_1(x_1; x_2) = 0; \\ f_2(x_1; x_2) = 0. \end{cases}$ Варіанти завдань наведені у таблиці 4.18.

Таблиця 4.18 – Варіанти для задачі 15

Вар.	Система рівнянь	Вар.	Система рівнянь
1	$\sin(x_1 + x_2) - x_2 - 1,2 = 0$ $2x_1 + \cos x_2 - 2 = 0$	16	$\sin(0,5x_1 + x_2) - 1,2x_1 - 1 = 0$ $x_1^2 + x_2^2 - 1 = 0$
2	$\cos(x_1 - 1) + x_2 - 0,5 = 0$ $\sin x_1 + 2x_2 - 2 = 0$	17	$\tan(x_1 x_2 + 0,3) - x_1^2 = 0$ $0,9x_1^2 + 2x_2^2 - 1 = 0$
3	$\sin x_1 + 2x_2 - 2 = 0$ $\cos x_1 + x_2 - 1,5 = 0$	18	$\sin(x_1 + x_2) - 1,3x_1 - 1 = 0$ $x_1^2 + 0,2x_2^2 - 1 = 0$
4	$\cos x_1 + x_2 - 1,5 = 0$ $2x_1 - \sin(x_2 - 0,5) - 1 = 0$	19	$\tan(x_1 x_2) - x_1^2 = 0$ $0,8x_1^2 + 2x_2^2 - 1 = 0$
5	$\sin(x_1 + 1,5) - x_2 + 2,9 = 0$ $\cos(x_2 - 2) + x_1 = 0$	20	$\sin(x_1 + x_2) - 1,5x_1 - 0,1 = 0$ $3x_1^2 + x_2^2 - 1 = 0$
6	$\cos(x_1 + 0,5) + x_2 - 0,8 = 0$ $\sin x_2 - 2x_1 - 1,6 = 0$	21	$\tan(x_1 x_2) - x_1^2 = 0$ $0,7x_1^2 + 2x_2^2 - 1 = 0$
7	$\sin(x_1 - 1) + x_2 - 0,1 = 0$ $x_1 - \sin(x_2 + 1) - 0,8 = 0$	22	$\sin(x_1 + x_2) - 1,2x_1 - 0,1 = 0$ $x_1^2 + x_2^2 - 1 = 0$
8	$\cos(x_1 + x_2) + 2x_2 = 0$ $x_1 + \sin x_2 - 0,6 = 0$	23	$\tan(x_1 x_2 + 0,2) - x_1^2 = 0$ $0,6x_1^2 + 2x_2^2 - 1 = 0$
9	$\cos(x_1 + 0,5) - x_2 - 2 = 0$ $\sin x_2 - 2x_1 - 1 = 0$	24	$\sin(x_1 + x_2) - x_1 + 0,1 = 0$ $x_2 - \cos(3x_1) + 0,1 = 0$
10	$\sin(x_1 + x_2) - x_2 - 1,5 = 0$ $x_1 + \cos(x_2 - 0,5) - 0,5 = 0$	25	$\cos(x_1 + 0,5) + x_2 - 1 = 0$ $\sin x_2 - 2x_1 - 2 = 0$
11	$\sin(x_2 + 1) - x_1 - 1,2 = 0$ $2x_1^2 + x_2 - 2 = 0$	26	$\cos(x_2 - 2) + x_1 = 0$ $\sin(x_1 + 0,5) - x_2 + 2,9 = 0$
12	$\cos(x_2 - 1) + x_1 - 0,5 = 0$ $x_2 - \cos x_1 - 3 = 0$	27	$\sin(x_1 - 1) + x_2 - 1,5 = 0$ $x_1 - \sin(x_2 - 1) - 1 = 0$
13	$\tan(x_1 x_2 + 0,4) - x_1^2 = 0$ $0,6x_1^2 + 2x_2^2 - 1 = 0$	28	$\sin(x_2 + 1) - x_1 - 1 = 0$ $2x_2 + \cos x_1 - 0,5 = 0$
14	$\sin(x_1 + x_2) - 1,6x_1 - 1 = 0$ $x_1^2 + x_2^2 - 1 = 0$	29	$\cos(x_2 - 1) + x_1 - 0,8 = 0$ $x_2 - \cos x_1 - 2 = 0$
15	$\tan(x_1 x_2 + 0,1) - x_1^2 = 0$ $x_1^2 + 2x_2^2 - 1 = 0$	30	$\cos(x_1 - 1) + x_2 - 1 = 0$ $\sin x_2 + 2x_1 - 1,6 = 0$

5 ЧИСЕЛЬНЕ ДИФЕРЕНЦІЮВАННЯ ТА ІНТЕГРУВАННЯ

5.1 Постановка задачі чисельного диференціювання

Під час розв'язування багатьох практичних задач виникає необхідність одержати значення похідних різних порядків функції f , заданої у вигляді таблиці або складного аналітичного виразу. У цих випадках застосувати безпосередньо методи диференціального числення або неможливо, або важко. Тоді використовують наближені методи чисельного диференціювання [1, 4, 5, 11].

Найпростіші вирази для похідних отримуємо як результат диференціювання інтерполяційних формул.

Отже, розглянемо таку задачу. На сітці $a \leq x_0 < x_1 < \dots < x_n \leq b$ у вузлах x_i задані значення f_i функції f , безперервно диференційованої $n+1+m$ разів. Потрібно обчислити похідну $f^{(m)}(x^*)$, $x^* \in [a, b]$ і оцінити похибку.

Один з можливих способів розв'язання цієї задачі. Побудуємо для функції f по вузлах x_i ($i = 0, 1, \dots, n$) інтерполяційний многочлен із залишковим членом R_n так, що

$$f(x) = P_n(x) + R_n(x). \quad (5.1)$$

Продиференціюємо праву і ліву частини співвідношення (5.1) m разів і покладемо $x = x^*$:

$$f^{(m)}(x^*) = P_n^{(m)}(x^*) + R_n^{(m)}(x^*). \quad (5.2)$$

Для достатньо гладких функцій, тобто для функцій з обмеженими похідними, достатньої кількості вузлів і достатньої точності обчислень, величина $R_n^{(m)}(x^*)$ мала і $P_n^{(m)}(x^*)$ є гарним наближенням для $f^{(m)}(x^*)$, тобто можна покласти

$$f^{(m)}(x^*) \approx P_n^{(m)}(x^*). \quad (5.3)$$

У практичних розрахунках чисельне диференціювання є дуже чутливим до похибок у вихідній інформації, відкидання членів ряду і до інших подібних операцій. Крім того, висока точність інтерполяції (мале значення $R_n(x)$) зовсім не гарантує високої точності інтерполяційної формули для похідних (малості $R_n^{(m)}(x)$). Тому чисельне диференціювання слід застосовувати обережно і, як правило, для невеликих m .

Враховуючи сказане, а також те, що обчислення вищих похідних може бути зведене до послідовного обчислення нижчих, детальніше зупинимося на одержанні розрахункових формул для f' і f'' у вузлах рівномірної сітки. Для одержання похідних у вузлових точках доцільно використовувати інтерполяційний многочлен Стірлінга і його залишковий член (див. формули [4, 15]). Так, диференціюючи многочлен Стірлінга і його залишковий член за x і беручи $x^* = x_0$ ($t^* = 0$), одержимо такі вирази для похідної:

$$f'(x_0) = \frac{1}{h^2} \mu f_0^1 \pm \frac{M_3}{6} h^2 \quad (k=1); \quad (5.4)$$

$$f'(x_0) = \frac{1}{h^2} \left(\mu f_0^1 - \frac{\mu f_0^3}{6} \right) \pm \frac{M_3}{30} h^4 \quad (k=2). \quad (5.5)$$

Диференціюючи многочлен Стірлінга два рази за x , і обчислюючи значення другої похідної в точці $x^* = x_0$, маємо

$$f''(x_0) = \frac{1}{h^2} f_0^2 \pm \frac{M_3}{12} h^2 \quad (k=1); \quad (5.6)$$

$$f''(x_0) = \frac{1}{h^2} \left(f_0^2 - \frac{1}{12} f_0^4 \right) \pm \frac{M_6}{90} h^4 \quad (k=2). \quad (5.7)$$

Для обчислення похідної точно в середині між вузлами $x^* = x_0 + h/2$ застосовують многочлен Бесселя. У цьому випадку відповідні формули для похідної мають вигляд

$$f' \left(x_0 + \frac{h}{2} \right) = \frac{1}{h} f_{1/2}^1 \pm \frac{M_3}{24} h^2 \quad (k=1); \quad (5.8)$$

$$f' \left(x_0 + \frac{h}{2} \right) = \frac{1}{h} \left(f_{1/2}^1 - \frac{1}{24} f_{1/2}^3 \right) \pm \frac{3M_5}{640} h^4 \quad (k=2). \quad (5.9)$$

Практичний інтерес викликають також так звані формули одностороннього диференціювання, що дозволяють обчислити $f'(x_0)$ у вузлах $x_i = x_0 + ih$ ($i = 0, 1, \dots, k, \dots$ або $i = 0, -1, \dots, -k, \dots$). Побудову цих формул зручно провести за допомогою першого і другого інтерполяційних многочленів Ньютона.

Диференціюючи перший многочлен Ньютона за x і обчислюючи значення похідної в точці $x = x_0$ ($t = 0$) для $k = 1$ і $k = 2$, одержимо такі формули:

$$f'(x_0) = \frac{1}{h} \Delta f_0 \pm \frac{1}{2} M_2 h, \quad (5.10)$$

$$f'(x_0) = \frac{1}{h} \left(\Delta f_0 - \frac{1}{2} \Delta^2 f_0 \right) \pm \frac{1}{3} M_3 h^2. \quad (5.11)$$

Аналогічно, диференціюючи другий многочлен Ньютона, для $k = -1$ і $k = -2$, відповідно, маємо

$$f'(x_0) = \frac{1}{h} \nabla f_0 \pm \frac{1}{2} M_2 h, \quad (5.12)$$

$$f'(x_0) = \frac{1}{h} \left(\nabla f_0 - \frac{1}{2} \nabla^2 f_0 \right) \pm \frac{1}{3} M_3 h^2. \quad (5.13)$$

5.2 Особливості чисельного диференціювання

Наведемо знову усі формули другого порядку, виразивши скінченні різниці, які входять до їх складу, безпосередньо через значення функції f_i . Враховуючи (5.4), (5.6) та (5.8) маємо:

$$f'(x_0) = \frac{f_1 - f_{-1}}{2h} \pm \frac{M_3}{6} h^2, \quad (5.14)$$

$$f''(x_0) = \frac{f_1 - 2f_0 + f_{-1}}{h^2} \pm \frac{M_4}{12} h^2, \quad (5.15)$$

$$f'(x_0 + \frac{h}{2}) = \frac{f_1 - f_0}{h} \pm \frac{M_3}{24} h^2. \quad (5.16)$$

Співвідношення (5.11) та (5.13) відповідно дають

$$f'(x_0) = \frac{1}{2h} (-3f_0 + 4f_1 - f_2) \pm \frac{M_3}{3} h^2, \quad (5.17)$$

$$f'(x_0) = \frac{1}{2h} (3f_0 - 4f_{-1} + f_{-2}) \pm \frac{M_3}{3} h^2. \quad (5.18)$$

З наведених вище формул видно, що зі зменшенням кроку сітки зменшується і похибка методу. Однак якщо значення функції f_i задані приблизно, наприклад, з однаковою абсолютною похибкою ε , то при використанні формул чисельного диференціювання сумарна похибка буде містити додатковий доданок, обернено пропорційний h_m (m – порядок похідної). Тому зменшення h доцільне лише у певних межах.

Ілюструючи сказане, розглянемо праву частину формули (5.16). Сумарна похибка її становить

$$\Delta = \frac{M_3}{24} h^2 + \frac{2\varepsilon}{h}. \quad (5.19)$$

Прирівнюючи $\Delta'(h)$ до нуля, одержуємо точку екстремуму функції $\Delta(h)$:

$$h_0 = 2\sqrt[3]{\frac{3\varepsilon}{M_3}} \approx 2,9\sqrt[3]{\frac{\varepsilon}{M_3}}. \quad (5.20)$$

Оскільки $\Delta''(h) > 0$, то h_0 – точка мінімуму $\Delta(h)$, причому

$$\Delta(h_0) = \frac{3}{2}\sqrt[3]{\frac{1}{3}M_3\varepsilon^2} \approx \sqrt[3]{M_3\varepsilon^2}. \quad (5.21)$$

Це співвідношення, зокрема, означає, що ні при якому h не можна гарантувати, що похибка результату є величиною $o(\varepsilon^{2/3})$.

Аналогічно, з формули (5.15), для оптимального кроку одержуємо вираз

$$h_0 = 2\sqrt[4]{\frac{3\varepsilon}{M_4}} \approx 2,6\sqrt[4]{\frac{\varepsilon}{M_4}}, \quad (5.22)$$

а з формул (5.17) і (5.18) – вираз

$$h_0 = \sqrt[3]{\frac{6\varepsilon}{M_3}} \approx 1,8\sqrt[3]{\frac{\varepsilon}{M_3}}.$$

Таким чином, при обчисленні похідних попередньо варто визначити оптимальний крок вихідної таблиці значень f_i .

Приклад 1. Обчислити $f'(1,4)$ і $f''(1,4)$ для функції $f = \ln x$, заданої у вигляді таблиці, яка містить значення f_i з усіма правильними (в широкому сенсі) знаками. Оцінити похибку результату.

x	1,2	1,3	1,4	1,5	1,6
f	0,1823	0,2626	0,3364	0,4054	0,4700

Для обчислення необхідних похідних застосуємо відповідно формули (5.5) і (5.2). Тоді, використовуючи рівності (5.10) і (5.9), а також вихідні дані, одержимо такі значення для оптимального кроку:

$$h_{01} = \sqrt[3]{\frac{6 \cdot 10^{-4}}{0,73}} \approx 0,1 \text{ при обчисленні } f'(1,6),$$

$$h_{02} = \sqrt[4]{\frac{48 \cdot 10^{-4}}{2,1}} \approx 0,22 \text{ при обчисленні } f''(1,4).$$

Оскільки табличні дані не дозволяють вибрати величину кроку 0,22, то за h_2 беремо найближче можливе число 0,2. Отже,

$$f'(1,6) = \frac{1}{0,2} (3 \cdot 0,4700 - 4 \cdot 0,4054 + 0,3364) = 0,624,$$

причому сумарна похибка не перевищує

$$\Delta = \frac{0,73}{3} \cdot 0,1^2 + \frac{4 \cdot 10^{-4}}{0,1} = 0,007.$$

$$f''(1,4) = \frac{1}{0,2^2} (0,4700 - 2 \cdot 0,3364 + 0,1823) = -0,512,$$

причому сумарна похибка не перевищує

$$\Delta = \frac{2,9}{12} \cdot 0,2^2 + \frac{4 \cdot 10^{-4}}{0,2^2} = 0,02.$$

Наведені оцінки похибки хоча і є, як правило, занадто завищеними, але вказують на те, що операція пошуку другої похідної є менш надійною, ніж першої.

У деяких практичних випадках для визначення похідної задається тільки таблиця значень функції. Тоді оцінити похибку неможливо. Наближені значення похідних обчислюються безпосередньо за однією з формул (5.4) – (5.13) без врахування похибки.

Приклад 2. Обчислити $f'(1,3)$, $f''(1,4)$ для функції $f(x)$, заданої у вигляді такої таблиці:

x	1,2	1,3	1,4	1,5	1,6
$y = f(x)$	0,18	0,26	0,34	0,41	0,47

На підставі формул (5.4) і (5.6) відповідно одержуємо:

$$f'(1,3) = \frac{1}{0,1} \cdot \frac{1}{2} (0,34 - 0,26 + 0,26 - 0,18) = 0,8,$$

$$f''(1,4) = \frac{1}{0,1^2} (0,41 - 0,34 - 0,34 + 0,26) = -1.$$

5.3 Постановка задачі чисельного інтегрування

Нехай потрібно обчислити інтеграл

$$I = \int_a^b f(x) dx. \quad (5.23)$$

З курсу математичного аналізу відомо, що для неперервної на відрізку $[a, b]$ функції f інтеграл (5.23) існує і дорівнює різниці значень первісної F для функції f у точках b і a :

$$I = \int_a^b f(x) dx = F(b) - F(a). \quad (5.24)$$

Однак, у переважній більшості практичних задач первісну не вдається виразити через елементарні функції [1, 4]. Крім того, функція f часто задається у вигляді таблиці її значень для певних значень аргументу. Усе це породжує потребу в наближених методах обчислення інтеграла (5.24), які можна умовно розділити на аналітичні і чисельні [4]. Перші полягають у наближеній побудові первісної і подальшому використанні формули (5.24). Другі ж дозволяють безпосередньо знайти числове значення інтеграла, ґрунтуючись на відомих значеннях підінтегральної функції (а іноді і її похідних) у заданих точках, які називають вузлами. Далі ми зупинимося лише на чисельних методах інтегрування функцій. Сам процес чисельного визначення інтеграла називається квадратурою, а відповідні формули – квадратурними формулами.

Залежно від способу задання підінтегральної функції будемо розглядати два різних (у сенсі реалізації) випадки чисельного інтегрування.

Задача I. На відрізку $[a, b]$ у вузлах x_i задані значення f_i деякої функції f , що належить певному класу F . Потрібно приблизно обчислити інтеграл (5.23) і оцінити похибку отриманого значення.

Так зазвичай ставиться задача чисельного інтегрування в тому випадку, коли підінтегральна функція задана у вигляді таблиці.

Задача II. На відрізку $[a, b]$ функція $f(x)$ задана у вигляді аналітичного виразу. Потрібно обчислити інтеграл (5.23) із заданою гранично припустимою похибкою ε .

Один з можливих способів розв'язання сформульованих задач оснований на використанні різних квадратурних формул вигляду

$$I \equiv \int_a^b f(x)dx \approx (b-a) \sum_{i=1}^n A_i f(x_i) \equiv I_n \quad (5.25)$$

з відомим залишковим членом $R_n|f| = I - I_n$ або його оцінкою.

У загальному випадку як вузлові точки x_i , так і вагові множники (ваги) A_i наперед не відомі і підлягають визначенню при виведенні кожної конкретної квадратурної формули (5.25) на основі вимог, що висуваються до неї.

За своєю суттю, задача чисельного інтегрування еквівалентна оцінюванню середнього значення функції. Дійсно, середнє значення функції на відрізку $[a, b]$ визначається так: $\bar{f} = \frac{1}{b-a} \int_a^b f(x)dx$. Тому $\int_a^b f(x)dx = (b-a)\bar{f}$.

У свою чергу, визначення середнього значення функції – це статистична задача, що містить проблеми послідовної вибірки і планування експерименту. Зважаючи на складність такої постановки питання, обмежимося лише класичними методами чисельного інтегрування, основаними на попередньому визначенні як вузлових точок, у яких повинна бути задана інформація про інтегровальну функцію, так і на самій інформації.

Перейдемо тепер до алгоритмів розв'язування сформульованих вище задач.

Алгоритм розв'язування задачі I.

1. Вибирають конкретну квадратурну формулу (5.25) і обчислюють I_n . Якщо значення функції f_i задані наближено, то обчислюють лише наближене значення \overline{I}_n для точного I_n .

2. Наближено беруть, що $I \approx \overline{I}_n$.

3. Користуючись конкретним виразом для залишкового члена або оцінкою його для вибраної квадратурної формули, обчислюють похибку методу

$$\Delta_1 = |I - I_n| = |R_n|.$$

4. Визначають похибку обчислення \overline{I}_n : $\Delta_2 = |I_n - \overline{I}_n|$ за похибками наближених значень f_i .

5. Знаходять повну абсолютну похибку наближеного значення I_n :

$$\Delta = |I_n - \overline{I}_n| \leq \Delta_1 + \Delta_2.$$

6. Одержують розв'язок задачі у вигляді $I = \overline{I}_n \pm \Delta$.

Для достатньо гладких функцій, тобто, для функцій з обмеженою змінною похідних, похибка квадратурних формул (5.25) для достатньо великих n , як правило, мала. Тому при достатній точності початкових значень f_i і при достатній точності обчислення \overline{I}_n можна сподіватись, що \overline{I}_n буде хорошим наближенням для I . На цих міркуваннях і оснований наступний алгоритм.

Алгоритм розв'язання задачі II.

1. Подають ε у вигляді суми трьох від'ємних доданків $\varepsilon = \varepsilon_1 + \varepsilon_2 + \varepsilon_3$, де ε_1 – гранично допустима похибка методу; ε_2 – гранично допустима похибка обчислення \overline{I}_n ; ε_3 – гранично допустима похибка округлення результату.

2. Вибирають n у квадратурній формулі так, щоб виконувалась нерівність

$$\Delta_1 = |I - I_n| = |R_n| \leq \varepsilon_1.$$

3. Обчислюють f_i з такою точністю, щоб при підрахунку I_n за формулою (5.25) забезпечити виконання нерівності $\Delta_2 = |I_n - \overline{I}_n| \leq \varepsilon_2$.

Для цього достатньо обчислити усі f_i з абсолютною похибкою $\frac{\varepsilon}{(b-a) \sum_{i=1}^n |A_i|}$.

4. Знайдену в пункті 3 величину \overline{I}_n округляють (якщо $\varepsilon_3 \neq 0$) з гранично допустимою похибкою ε_3 до величини $\overline{\overline{I}_n}$.

5. Одержують розв'язок задачі у вигляді $I = \overline{I}_n \pm \varepsilon$.

Використані в алгоритмах обох задач квадратурні формули будуються, як вже було сказано, на підставі тих або інших критеріїв, що визначають положення вузлових точок і величини вагових множників. Такими критеріями можуть бути: зображення інтеграла у вигляді інтегральної суми; апроксимація підінтегральної функції (наприклад, многочленом) з подальшим інтегруванням апроксимувальної функції; вимога, щоб формула (7.25) була абсолютно точною для певного класу функцій і т. д.

5.4 Найпростіші квадратурні формули

Формули прямокутників. Як відомо, визначений інтеграл є границею інтегральних сум:

$$\int_a^b f(x) \cdot dx = \lim_{\max h_i \rightarrow 0} \sum_{i=1}^n h_i \cdot f(\xi_i), \quad (5.26)$$

кожна з яких відповідає деякому розбиттю D_n : $a=x_0 < x_1 < \dots < x_n=b$ відрізка $[a, b]$ і довільному набору точок $\xi_i \in [x_{i-1}, x_i]$ для кожного розбиття $h_i=x_i-x_{i-1}$.

Обмежуючись скінченною кількістю доданків у правій частині рівності (5.26) і беручи як набір ξ_i ті чи інші значення аргументу з відрізків $[x_{i-1}, x_i]$, можна отримати різні формули наближеного інтегрування. Так, беручи як набір ξ_i значення лівих чи правих кінців відрізків $[x_{i-1}, x_i]$, одержимо відповідно **загальні формули лівих та правих прямокутників** ($h_i=1/n=const$):

$$I \equiv \int_a^b f(x) dx \approx (b-a) \sum_{i=0}^{n-1} \frac{1}{n} \cdot f_i \equiv I_L; \quad (5.27)$$

$$I \equiv \int_a^b f(x) dx \approx (b-a) \sum_{i=0}^n \frac{1}{n} \cdot f_i \equiv I_R. \quad (5.28)$$

Назви цих формул пов'язані з їх геометричною інтерпретацією. Якщо у площині xOy побудувати криву $y=f(x)$, розбити відрізок $[a, b]$ на n частин точками x_i сітки D_n , то формула лівих прямокутників як наближене значення інтеграла дасть сумарну площу заштрихованих прямокутників на рис. 5.1; а формула правих прямокутників – сумарну площу заштрихованих прямокутників на рис. 5.2.

Знайдемо вираз для залишкового члена наближеної формули (5.27). З цією метою зобразимо інтеграл, що входить у ліву частину співвідношення

$$(5.27), \text{ у вигляді суми: } \int_a^b f(x) dx = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f(x) dx. \quad (5.29)$$

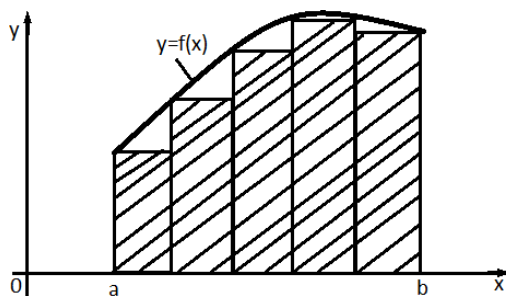


Рисунок 5.1 – Формула лівих прямокутників

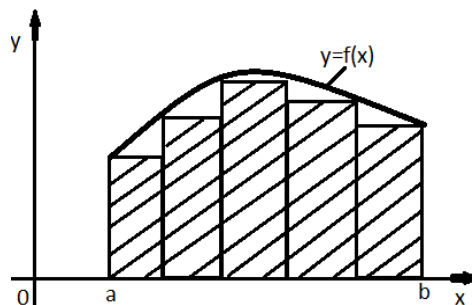


Рисунок 5.2 – Формула правих прямокутників

Припускаючи, що функція $f(x)$ диференційована на відрізку $[a; b]$, запишемо для неї на кожному з відрізків $[x_{i-1}; x_i]$ формулу Тейлора із залишковим членом у формі Лагранжа:

$$f(x) = f_i + (x - x_i) \cdot f'(\eta_i); \quad \eta_i \in (x_{i-1}; x_i). \quad (5.30)$$

Підставимо у праву частину співвідношення (5.29) замість функції f її зображення (5.30) і виконаємо інтегрування, використовуючи другу теорему про середнє значення функції [1]:

$$\int_a^b f(x) dx = (b-a) \frac{1}{n} \sum_{i=1}^n f_i + \frac{h^2}{2} \cdot \sum_{i=1}^n f''(\bar{\eta}_i); \quad \bar{\eta}_i \in (x_{i-1}, x_i).$$

Оскільки друга похідна неперервна, то існує така точка $\eta \in (a, b)$, що

$$\sum_{i=1}^n f''(\bar{\eta}_i) = n f''(\eta) = \frac{b-a}{h} f''(\eta).$$

Використовуючи це співвідношення, остаточно маємо

$$\int_a^b f(x) dx = (b-a) \sum_{i=1}^n \frac{1}{n} \cdot f_i + \frac{b-a}{2} h f''(\eta). \quad (5.31)$$

Порівнюючи формули (5.27) і (5.31), одержуємо вираз для залишкового члена квадратурної формули (5.27):

$$R_n[f] = I - I_n = \frac{b-a}{2} h f''(\eta). \quad (5.32)$$

Таким чином, оцінку похибки квадратурної формули (5.27) можна отримати з такої формули:

$$A_1 = \left| \int_a^b f(x) dx - (b-a) \sum_{i=1}^n \frac{1}{n} f_i \right| \leq \frac{b-a}{2} h M_1, \quad (5.33)$$

де $M_1 = \max_{[a,b]} |f'(x)|$.

Вираз для залишкового члена наближеної формули (5.28) відрізняється від залишкового члена наближеної формули (5.27) протилежним знаком.

Приклад 1. За допомогою формул лівих і правих прямокутників обчислити $\int_1^9 \frac{dx}{x+2}$, беручи $n=4$.

Оскільки $a=1$ і $b=9$, то крок $h=(b-a)/n=2$; тоді точками розбиття є $x_0=1$, $x_1=3$, $x_2=5$, $x_3=7$, $x_4=9$, а значення підінтегральної функції $f(x)=1/(x+2)$ у цих точках такі: $y_0 = f(x_0)=1/3$; $y_1 = f(x_1)=1/5$; $y_2 = f(x_2)=1/7$; $y_3 = f(x_3)=1/9$; $y_4 = f(x_4)=1/11$.

Далі знайдемо числове значення інтеграла, користуючись формулою (5.27): $I_L = \frac{b-a}{n} \cdot (y_0 + y_1 + y_2 + y_3) = 2 \cdot \left(\frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{9} \right) \approx 1,6024$.

Якщо обчислення визначеного інтеграла виконати за формулою (5.28), одержимо $I_R = \frac{b-a}{n} \cdot (y_1 + y_2 + y_3 + y_4) = 2 \cdot \left(\frac{1}{5} + \frac{1}{7} + \frac{1}{9} + \frac{1}{11} \right) \approx 1,1053$.

Найчастіше використовують формулу, основу на ідеї подання визначеного інтеграла у вигляді інтегральної суми, тобто *формулу прямокутників*, де як ξ_i беруть середини відрізків $[x_{i-1}, x_i]$. Для рівномірної сітки ($h_i=h$) ця формула має такий вигляд:

$$I \equiv \int_a^b f(x) dx \approx \frac{b-a}{n} \sum_{i=1}^n f_{i-1/2} \equiv I_n, \quad (5.34)$$

де $f_{i-1/2} = f\left(x_i - \frac{h}{2}\right)$; $x_0 = a$, $x_n = b$

Знайдемо вираз для залишкового члена наближеної формули (5.34). З цією метою зобразимо інтеграл, що входить у ліву частину співвідношення (5.34), у вигляді суми (5.29). Припускаючи, що функція $f(x)$ двічі диференційована, тобто $f \in C^2 [a; b]$, запишемо для функції $f(x)$ на кожному з відрізків $[x_{i-1}; x_i]$ формулу Тейлора із залишковим членом у формі Лагранжа:

$$f(x) = f_{i-1/2} + (x - x_i + h/2) \cdot f'_{i-1/2} + \frac{(x - x_i + h/2)^2}{2} f''(\eta_i); \quad (5.35)$$

де $\eta_i \in (x_{i-1}; x_i)$.

Підставимо у праву частину співвідношення (5.29) замість функції f її зображення (5.35) і виконаємо інтегрування, використовуючи другу теорему про середнє значення функції:

$$\int_a^b f(x) dx = (b-a) \frac{1}{n} \sum_{i=1}^n f_{i-1/2} + \frac{h^3}{24} \cdot \sum_{i=1}^n f''(\eta_i); \quad \bar{\eta}_i \in (x_{i-1}, x_i).$$

Оскільки друга похідна неперервна, то існує така точка $\eta \in (a, b)$, що

$$\sum_{i=1}^n f''(\bar{\eta}_i) = n f''(\eta) = \frac{b-a}{h} f''(\eta).$$

Використовуючи це співвідношення, остаточно маємо

$$\int_a^b f(x) dx = (b-a) \sum_{i=1}^n \frac{1}{n} \cdot f_{i-1/2} + \frac{b-a}{24} h^2 f''(\eta). \quad (5.36)$$

Порівнюючи формули (5.34) і (5.36), одержуємо вираз для залишкового члена квадратурної формули (5.34):

$$R_n[f] = I - I_n = \frac{b-a}{24} h^2 f''(\eta). \quad (5.37)$$

Таким чином, оцінку похибки квадратурної формули (5.34) можна отримати з формули:

$$\Delta_1 = \left| \int_a^b f(x) dx - (b-a) \sum_{i=1}^n \frac{1}{n} f_{i-1/2} \right| \leq \frac{b-a}{24} h^2 M_2, \quad (5.38)$$

де $M_2 = \max_{[a,b]} |f''(x)|$.

Отримані вирази для залишкового члена (5.37) і похибки (5.38) показують, що формула (5.34) є точною для будь-якої лінійної функції, оскільки друга похідна такої функції дорівнює нулю, а отже, залишковий член і похибка також дорівнюють нулю.

Покажемо, що отримана оцінка не може бути поліпшена, тобто що

існує функція, для якої похибка обчислення інтеграла за формулою (5.34) точно дорівнює правій частині (5.38). Для цього як інтегральну функцію розглянемо $f=x^2$ і застосуємо до неї формулу (5.34):

$$I_n = (b-a) \frac{1}{n} \sum_{i=1}^n \left(a + \frac{2i-1}{2} h \right)^2.$$

Розкривши дужки під знаком суми і виконавши необхідні підсумовування, одержимо $I_n = \frac{b^3 - a^3}{3} - (b-a) \frac{h^3}{12}$. З іншого боку, безпосереднє інтегрування функції x^2 дає $I = \int_a^b x^2 dx = \frac{b^3 - a^3}{3}$.

Склавши різницю між точним значенням інтеграла і наближеним, для залишкового члена одержимо вираз $I - I_n = (b-a) \frac{h^2}{12}$.

Повертаючись до оцінювання похибки (5.38) і зауважуючи, що для функції x^2 друга похідна (а отже, і M_2) дорівнює 2, одержуємо для похибки те ж саме значення $\Delta_1 = (b-a) \frac{h^2}{12}$, тобто оцінка похибки (5.38) досягається на параболі $y=x^2$. Цей результат можна поширити і на довільну параболу, зважаючи на лінійність операції інтегрування і той факт, що для лінійних функцій формула (5.34) є точною.

Оцінка (5.38) не враховує похибок, пов'язаних з обчисленням I_n . Похибка Δ_i відбиває розходження між точною формулою Ньютона – Лейбніца і наближеною формулою (5.34), тобто є похибкою методу.

Перейдемо до оцінювання похибки наближеного значення \bar{I}_n . Якщо значення функції, що використовуються в квадратурній формулі, отримані приблизно або обчислення з якихось причин не можуть бути виконані абсолютно точно – це спричиняє появу обчислювальної похибки і похибки округлення. Нехай, наприклад, значення $f_{i-1/2}$ у формулі (5.34) обчислені з однаковою абсолютною похибкою ε ; тоді сумарна обчислювальна похибка \bar{I}_n становитиме

$$\Delta_2 = (b-a) \sum_{i=1}^n \frac{1}{n} \varepsilon = (b-a) \varepsilon. \quad (5.39)$$

Відзначимо характерну властивість цієї похибки: вона не залежить від числа відрізків розбиття усього відрізка інтегрування, а лише пропорційна його довжині.

Приклад 2. Обчислити за допомогою формули прямокутників інтеграл $\int_0^1 \frac{dx}{1+x}$, беручи $n=4$. Оцінити похибку отриманого значення.

За заданими границями інтегрування і кількістю розбивок n визначимо крок: $h=(1-0)/4=0,25$. Далі, на підставі формули (5.34), маємо

$$I_L = 0,25 \left[f\left(\frac{1}{8}\right) + f\left(\frac{3}{8}\right) + f\left(\frac{5}{8}\right) + f\left(\frac{7}{8}\right) \right].$$

Обчисливши необхідні значення функції з трьома правильними, у вузькому сенсі, знаками ($\varepsilon = 0,0005$), одержимо

$$\int_0^1 \frac{dx}{1+x} = 0,25(0,889 + 0,727 + 0,615 + 0,533) = 0,691.$$

Похибку методу оцінимо за формулою (5.38), для чого попередньо знайдемо максимум абсолютної величини другої похідної підінтегральної функції:

$$M_2 = \max_{[0,1]} \left| \left(\frac{1}{1+x} \right)'' \right| = \max_{[0,1]} \frac{2}{(1+x)^3} = 2.$$

Таким чином, похибка методу $\Delta_1 \leq (1/24) \cdot 0,25^2 \cdot 2 \approx 0,0053$.

Користуючись формулою (5.39), знайдемо обчислювальну похибку

$$\Delta_2 \leq 1 \cdot 0,0005 = 0,0005.$$

Отже, за повну похибку наближеного значення інтеграла можна взяти $\Delta = \Delta_1 + \Delta_2 = 0,006$, а остаточну відповідь записати у вигляді

$$\int_0^1 \frac{dx}{1+x} = 0,691 \pm 0,006.$$

Для порівняння наведемо кілька знаків точного значення обчисленого інтеграла: $\ln 2 = 0,693147\dots$

Приклад 3. За допомогою формули прямокутників обчислити інтеграл $\int_0^1 \frac{dx}{1+x}$ з точністю до 10^{-3} .

Застосовуючи алгоритм розв'язання задачі 2 з розділу 5.1, зобразимо сумарну похибку у вигляді суми трьох додатків: $0,0009 + 0,00005 + 0,00005$. Далі виберемо n з умови

$$\Delta = \frac{b-a}{24} h^2 M_2 = \frac{b-a}{24} \left(\frac{b-a}{n} \right)^2 M_2 \leq 0,0009.$$

Розв'язуючи цю нерівність відносно n , при $b-a=1$ і $M_2=2$ отримаємо $n \geq 10$.

Складемо таблицю значень функції $1/(x+1)$ з чотирма правильними знаками у вузькому сенсі:

0,05	0,15	0,25	0,35	0,45	0,55	0,65	0,75	0,85	0,95
0,9254	0,8696	0,8000	0,7407	0,6897	0,6452	0,6061	0,5714	0,5405	0,5128

Використовуючи формулу прямокутників (5.29), отримаємо

$$\bar{I}_{10} = 0,1(0,9254+0,8696+0,8000+0,7407+0,6897+0,6452+0,6061+0,5714+0,5405+0,5128)=0,69284.$$

Округливши отриманий результат, маємо $I=0,6928 \pm 0,001$.

Формула трапецій. Перейдемо тепер до іншого способу побудови квадратурних формул, який полягає у використанні апроксимації підінтегральної функції інтерполяційним многочленом. Розглянемо найпростіший випадок апроксимації многочленом першого степеня з вузлами у точках a і b .

$$f(x) = f(a) + \frac{x-a}{b-a} [f(b) - f(a)] + (x-a)(x-b) \frac{f''(\eta)}{2}; \quad \eta \in (a, b).$$

Інтегруючи праву і ліву частини цієї рівності і використовуючи другу теорему про середнє значення функції при інтегруванні останнього доданка правої частини, одержуємо

$$\int_a^b f(x) dx = \frac{b-a}{2} [f(a) + f(b)] - \frac{(b-a)^3}{12} f''(\eta); \quad \eta \in (a, b).$$

Таким чином, припускаючи, що відрізок інтегрування малий, одержуємо квадратурну формулу, яка називається **формулою трапецій**:

$$I \equiv \int_a^b f(x) dx \approx \frac{b-a}{2} [f(a) + f(b)] \equiv I_2 \quad (5.40)$$

із залишковим членом

$$R_2[f] = I - I_2 = -\frac{(b-a)^3}{12} f''(\eta); \quad \eta \in (a, b) \quad (5.41)$$

Використовуючи вираз (5.41) для залишкового члена, оцінку похибки квадратурної формули (5.40) можна записати у вигляді

$$I = \left| \int_a^b f(x) dx - \frac{b-a}{2} (f(a) + f(b)) \right| \leq \frac{(b-a)^3}{12} M_2, \quad (5.42)$$

де $M_2 = \max_{[a,b]} |f''(x)|$.

Отримані вирази для залишкового члена (5.41) і похибки (5.42) показують, що квадратурна формула (5.40) є точною для всіх лінійних функцій, оскільки друга похідна таких функцій дорівнює нулю, а отже, дорівнюють нулю залишковий член і похибка.

Аналогічно тому, як це було зроблено для оцінювання (5.38), можна показати, що й оцінка (5.42) покращується, тому що вона досягається на довільній параболі.

Оцінка обчислювальної похибки при розрахунках за формулою (5.40) для випадку, коли значення функції обчислені з однаковою точністю ε , має вигляд

$$\Delta_2 \leq \frac{b-a}{2}(\varepsilon + \varepsilon) = (b-a)\varepsilon. \quad (5.43)$$

Відзначимо, що обчислювальні похибки квадратурних формул (5.40) і (5.34) однакові.

Приклад 4. Обчислити за допомогою формули трапецій інтеграл $\int_0^1 \frac{dx}{1+x}$. Оцінити похибку отриманого наближеного значення.

На підставі формули (5.40) маємо $I_2 = 0,5[f(0) + f(1)]$. Обчисливши необхідні значення функції, одержимо $\int_0^1 \frac{dx}{1+x} \approx 0,5(1 + 0,5) = 0,75$.

Похибку методу оцінимо за формулою (5.42), використовуючи значення $M=2$, що отримано у прикладі 2:

$$\Delta_1 \leq \frac{1^3}{12} \cdot 2 \approx 0,17.$$

Обчислювальна похибка дорівнює нулю, оскільки значення функції I_2 знайдені абсолютно точно. Отже, остаточно маємо $\int_0^1 \frac{dx}{1+x} = 0,75 + 0,17$.

Відзначимо, що у прикладі 3 розв'язок був отриманий не такий точний, як у прикладі 2. Однак не слід робити поспішних висновків, тому що використання у прикладі 3 формули трапецій має і свої переваги. По-перше, якщо підінтегральна функція задана у вигляді таблиці її значень у вузлах x_i , то для використання формули прямокутників необхідно визначити значення цієї функції ще й у точках $x_i \pm h/2$, а це вносить додаткові труднощі і додаткову похибку. По-друге, у прикладі 3 значення підінтегральної функції були обчислені лише у двох точках, в той час як у прикладі 2 – у чотирьох точках, що, природно, потребує більшого часу.

Наведені міркування показують, що цінність квадратурної формули визначається не тільки формою її залишкового члена (або похибки), а й іншими факторами, наприклад часом розрахунків.

Інші види квадратурних формул. Розглянемо ще один спосіб побудови квадратурних формул, що полягає у поданні інтеграла як лінійної комбінації значень підінтегральної функції і її похідних у деяких вузлах x_i з подальшим визначенням невідомих коефіцієнтів (вагових множників).

Нехай, наприклад, ми будемо квадратурну формулу вигляду

$$\int_a^b f(x) \cdot dx \approx (b-a) \cdot [A_1 \cdot f(a) + A_2 \cdot f(b) + A_3 \cdot f'(a) + A_4 \cdot f'(b)]. \quad (5.44)$$

Визначимо вагові множники A ($i=1, 2, 3, 4$) так, щоб формула (5.44) була точною для довільних многочленів нульового, першого, другого і третього степеня. Зважаючи на лінійність операцій інтегрування і диференціювання, ця умова буде виконана, якщо вона виконується для многочленів $1, x, x^2, x^3$. Підставляючи ці многочлени замість $f(x)$ у співвідношення (5.44), за умови його точного виконання, одержуємо таку систему лінійних рівнянь відносно A_i :

$$\begin{cases} A_1 + A_2 = 1, \\ aA_1 + bA_2 + A_3 + A_4 = \frac{1}{2}(a+b), \\ a^2A_1 + b^2A_2 + 2aA_3 + 2bA_4 = \frac{1}{3}(a^2 + ab + b^2), \\ a^3A_1 + b^3A_2 + 3a^2A_3 + 3b^2A_4 = \frac{1}{4}(a^3 + a^2b + ab^2 + b^3). \end{cases}$$

Розв'язуючи цю систему, знаходимо $A_1 = A_2 = 1/2$; $A_3 = -A_4 = (b-a)/12$. Таким чином, шукана квадратурна формула має вигляд

$$I \equiv \int_a^b f(x) dx \approx (b-a) \left[\frac{f(a) + f(b)}{2} + (b-a) \frac{f'(a) - f'(b)}{12} \right] \equiv I_4. \quad (5.45)$$

Знайдемо вираз для залишкового члена цієї формули. Зобразимо функцію, що інтегрується, у вигляді суми інтерполяційного многочлена Ерміта [11] третього степеня з двома двократними вузлами a і b та залишковим членом. Потім проінтегруємо праву і ліву частини цього подання на відрізку $[a, b]$:

$$\int_a^b f(x) dx = \int_a^b H_3(x) dx + \int_a^b \frac{(x-a)^2(x-b)^2}{4!} f^{(4)}(\bar{\eta}) dx; \quad \bar{\eta} \in (a, b).$$

Перший доданок правої частини дає праву частину квадратурної формули (5.45), оскільки ця формула є точною для всіх многочленів

третього степеня, а отже, і для многочлена Ерміта $H_3(x)$. Другий доданок правої частини дає вираз для залишкового члена формули (5.45). Використовуючи другу теорему про середнє значення функції і виконуючи інтегрування, одержимо

$$R_4[f] = I - I_4 = \frac{(b-a)^5}{720} f^4(\eta); \quad \eta \in (a, b). \quad (5.46)$$

Отриманий вираз для залишкового члена дозволяє описати оцінку похибки квадратурної формули (5.45) у вигляді

$$\Delta_1 = |I - I_4| \leq \frac{(b-a)^5}{720} M_4, \quad M_4 = \max_{[a,b]} |f^4(x)|. \quad (5.47)$$

Оцінка (5.46) не може покращитися, оскільки вона досягається на довільному многочлені четвертого степеня, що неважко довести, аналогічно тому, як це було зроблено для оцінки (5.38).

Для оцінювання обчислювальної похибки результату, отриманої за формулою (5.45), припустимо, що значення функції задані з точністю ε_1 , а значення похідних з точністю ε_2 . Тоді обчислювальна похибка становить

$$\Delta_2 \leq (b-a)\varepsilon_1 + \frac{(b-a)^2}{6}\varepsilon_2 \quad (5.48)$$

Приклад 5. Обчислити за допомогою квадратурної формули (5.45) інтеграл $\int_0^1 \frac{dx}{1+x}$. Визначити похибку отриманого наближеного значення.

Обчисливши необхідні значення підінтегральної функції і її похідних за формулою (5.45), знаходимо $I_4 = 1 \cdot \left(\frac{1+0,5}{2} + 1 \cdot \frac{-1+0,25}{12} \right) = 0,6875$.

Похибку методу визначимо за формулою (5.47), для цього попередньо знайдемо максимум абсолютної величини четвертої похідної підінтегральної функції $M_4 = 24$: $\Delta_1 \leq \frac{1^5}{720} \cdot 24 \approx 0,034$.

Обчислювана похибка дорівнює нулю, оскільки значення функції і її похідних обчислені абсолютно точно.

Таким чином, округляючи наближені значення інтеграла і похибки, остаточно одержуємо $\int_0^1 \frac{dx}{1+x} = 0,69 \pm 0,04$.

В усіх розглянутих вище квадратурних формулах вузли квадратури були заздалегідь фіксовані. Тепер розглянемо випадки, коли положення усіх вузлів, так само як і усі вагові множники, вважаються вільними пара-

метрами. Щоб наступні викладення були не занадто складними, але в той же час не зовсім тривіальними, будемо шукати значення інтеграла у вигляді

$$\int_a^b f(x) dx = (b-a) \cdot [A_1 f(x_1) + A_2 f(x_2)]. \quad (5.49)$$

Для визначення чотирьох вільних параметрів A_1 , A_2 , x_1 і x_2 висунемо вимогу, щоб формула (5.49) була абсолютно точною для всіх многочленів нульового, першого, другого і третього степенів. Зважаючи на лінійність операції інтегрування і правої частини співвідношення (5.49), для того, щоб квадратурна формула (5.49) була точною для всіх многочленів третього степеня, необхідно і достатньо, щоб вона була точною для функцій 1 , x , x^2 , x^3 . Отже, повинні виконуватися співвідношення

$$\begin{cases} A_1 + A_2 = 1, \\ x_1 A_1 + x_2 A_2 = \frac{1}{2}(a+b), \\ x_1^2 A_1 + x_2^2 A_2 = \frac{1}{3}(a^2 + ab + b^2), \\ x_1^3 A_1 + x_2^3 A_2 = \frac{1}{4}(a^3 + a^2 b + ab^2 + b^3), \end{cases}$$

що є нелінійною системою рівнянь відносно параметрів, які обчислюємо: A_1, A_2, x_1, x_2 .

Розв'язок даної системи

$$A_1 = A_2 = \frac{1}{2}, \quad x_1 = \frac{b+a}{2} - \frac{b-a}{2} \cdot \frac{1}{\sqrt{3}}, \quad x_2 = \frac{b+a}{2} + \frac{b-a}{2} \cdot \frac{1}{\sqrt{3}}. \quad (5.50)$$

Отже, квадратурна формула (5.49) має вигляд

$$I \equiv \int_a^b f(x) dx \approx \frac{b-a}{2} \cdot \left[f\left(\frac{b+a}{2} - \frac{b-a}{2} \cdot \frac{1}{\sqrt{3}}\right) + f\left(\frac{b+a}{2} + \frac{b-a}{2} \cdot \frac{1}{\sqrt{3}}\right) \right] \equiv I_4. \quad (5.52)$$

Формули такого типу, коли не тільки вагові множники, але і вузли заздалегідь не фіксуються, називаються *гауссовими*.

Знайдемо вираз для залишкового члена формули (5.52). Для цього подамо функцію, що інтегрується, у вигляді суми інтерполяційного многочлена Ерміта третього степеня з двома двократними вузлами x_1 і x_2 , що визначаються співвідношеннями (5.50), і залишкового члена. Проінтегрувавши праву і ліву частини цього подання на відрізку (a, b) , одержимо

$$\int_a^b f(x) dx = \int_a^b H_3(x) dx + \int_a^b \frac{(x-x_1)^2 (x-x_2)^2}{4!} f(\bar{\eta}) dx; \quad \bar{\eta} \in (a, b).$$

Перший доданок правої частини дає праву частину квадратурної формули (5.51) через те, що ця формула є точною для всіх многочленів третього степеня, а отже, і для $H_3(x)$. Другий доданок правої частини дає залишковий член формули (5.51). Використовуючи другу теорему про середнє значення функції і виконуючи інтегрування, маємо

$$R_2[f] = I - I_2 = \frac{(b-a)^5}{4320} f^4(\eta); \quad \eta \in (a, b).$$

Отже, оцінка похибки знаходиться із співвідношення

$$\Delta_1 = |I - I_2| = \frac{(b-a)^5}{4320} M_4; \quad M_4 = \max_{[a,b]} |f^4(x)|. \quad (5.53)$$

Отриману оцінку не можна покращити, тому що вона досягається на довільному многочлені четвертого степеня. Це неважко показати прямими викладеннями, як це було зроблено для оцінки (5.38).

Якщо у формулі (5.51) значення вузлів визначені практично точно, а значення функції – з абсолютною похибкою ε , то для обчислюваної похибки при використанні формули (5.27) одержимо той же вираз, що і для обчислюваної похибки при використанні формул (5.52) і (5.53):

Приклад 6. Обчислити за допомогою квадратурної формули (5.51) інтеграл $\int_0^1 \frac{dx}{1+x}$. Оцінити похибку отриманого наближеного значення.

Спочатку визначимо вузли квадратурної формули:

$$x_1 = \frac{1}{2} \left(1 - \frac{1}{\sqrt{3}} \right) = 1.2113249\dots; \quad x_2 = \frac{1}{2} \left(1 + \frac{1}{\sqrt{3}} \right) = 0.7886751\dots$$

Тепер, обчисливши необхідні значення функції, що інтегрується, з точністю до трьох правильних знаків, у вузькому сенсі, скористаємося формулою (5.51): $\int_0^1 \frac{dx}{1+x} \approx \frac{1}{2} (0,826 + 0,559) = 0,6925$.

Похибку методу оцінимо за формулою (5.53), для чого використаємо знайдене у прикладі 2 значення максимуму абсолютної величини четвертої похідної $M_4 = 24$: $\Delta_1 \leq \frac{1^5}{4320} \cdot 24 \approx 0,0056$.

Обчислювану похибку можна знайти за формулою (5.49), враховуючи те, що точність обчислення значень функції, яка інтегрується, дорівнює 0,0005. Таким чином, сумарна похибка є $A = A_1 + A_2 = 0,0061$. Нарешті, округляючи наближене значення інтеграла, остаточно маємо $I = 0,692 \pm 0,007$.

Основна мета даного розділу – показати на простих прикладах, як виводяться різні формули чисельного інтегрування. Зрозуміло, що ми не розглянули всі способи побудови формул. У той же час наведені приклади є характерними, тому, використовуючи їх, можна самостійно побудувати конкретну квадратурну формулу, яка буде найбільше відповідати заданій практичній задачі.

5.5 Формула Сімпсона

Для того, щоб побудувати триточкову квадратурну формулу з рівновіддаленими вузлами для обчислення наближеного значення $\int_{x_0-h}^{x_0+h} f(x)dx$,

де $f(x)$ – неперервна на $[x_0-h; x_0+h]$ разом зі своїми похідними до четвертого порядку включно, можна використати інтерполяційний многочлен Лагранжа 2-го порядку, графік якого проходить через точки $(x_0-h; f(x_0-h))$, $(x_0; f(x_0))$ і $(x_0+h; f(x_0+h))$ та проінтегрувати його за x у межах від x_0-h до x_0+h .

Проте нижче таку квадратичну формулу будуватимемо, користуючись методом невизначених коефіцієнтів [1]. Цей метод, крім того, дає змогу достатньо просто обчислити її залишковий член. Отже, побудуємо квадратичну формулу вигляду

$$\int_{x_0-h}^{x_0+h} f(x)dx = 2h(A(f(x_0-h) + f(x_0+h)) + Bf(x_0)) + R(f), \quad (5.54)$$

де A і B – невідомі коефіцієнти, а $R(f)$ – залишковий член.

Щоб дістати рівняння, з яких можна визначити коефіцієнти A і B , подамо функції $f(x)$, $f(x_0-h)$ і $f(x_0+h)$ в околі точки x_0 за допомогою формули Тейлора. Маємо

$$f(x) = f(x_0) + (x-x_0)f'(x_0) + \frac{(x-x_0)^2}{2!}f''(x_0) + \frac{(x-x_0)^3}{3!}f'''(x_0) + \frac{(x-x_0)^4}{4!}f^{(IV)}(x_0 + \theta h),$$

$$f(x_0-h) = f(x_0) - hf'(x_0) + \frac{h^2}{2!}f''(x_0) - \frac{h^3}{3!}f'''(x_0) + \frac{h^4}{4!}f^{(IV)}(x_0 - \theta_2 h),$$

$$f(x_0+h) = f(x_0) + hf'(x_0) + \frac{h^2}{2!}f''(x_0) + \frac{h^3}{3!}f'''(x_0) + \frac{h^4}{4!}f^{(IV)}(x_0 + \theta_3 h),$$

$$0 < \theta, \theta_2, \theta_3 < 1.$$

Підставляємо ці значення функцій $f(x)$, $f(x_0 - h)$, $f(x_0 + h)$ у (5.54) і зважаючи на те, що

$$\int_{x_0-h}^{x_0+h} f(x)dx = 2hf(x_0) + \frac{2h^3}{3!} f''(x_0) + \frac{2h^5}{5!} f^{(IV)}(x_0 + \theta_1 h), \quad 0 < \theta_1 < 1,$$

тут за загальною теоремою про середнє

$$\int_{x_0-h}^{x_0+h} \frac{(x-x_0)^4}{4!} f^{(IV)}(x_0 + \theta h) dx = \frac{2h^5}{5!} f^{(IV)}(x_0 + \theta_1 h), \quad 0 < \theta_1 < 1,$$

для залишкового члена $R(f)$ отримаємо:

$$R(f) = 2h((1 - 2A - B)f(x_0) + h^2 \left(\frac{1}{3!} - A \right) f''(x_0) + \frac{h^4}{4!} \left(\frac{1}{5} f^{(IV)}(x_0 + \theta_1 h) - A(f^{(IV)}(x_0 - \theta_2 h) + f^{(IV)}(x_0 + \theta_3 h)) \right).$$

Невідомі коефіцієнти A і B доберемо так, щоб

$$\begin{cases} 1 - 2A - B = 0, \\ 1/3! - A = 0. \end{cases}$$

Звідси знаходимо $A=1/6$, $B=2/3$.

За цих значень A і B залишковий член квадратурної формули (5.54)

$$R(f) = \frac{2h^5}{4!} \left(\frac{1}{5} f^{(IV)}(x_0 + \theta_1 h) - \frac{1}{6} (f^{(IV)}(x_0 - \theta_2 h) + f^{(IV)}(x_0 + \theta_3 h)) \right).$$

Але $f^{(IV)}$ неперервна на $[x_0-h, x_0+h]$, тому існує точка $\xi \in [x_0-h, x_0+h]$ така,

$$\text{що } \frac{\frac{1}{5} f^{(IV)}(x_0 + \theta_1 h) - \frac{1}{6} (f^{(IV)}(x_0 - \theta_2 h) + f^{(IV)}(x_0 + \theta_3 h))}{1/5 + 2(-1/6)} = f^{(IV)}(\xi).$$

Отже,

$$R(f) = -\frac{h^5}{90} f^{(IV)}(\xi), \quad x_0 - h \leq \xi \leq x_0 + h. \quad (5.55)$$

Отже, триточкову квадратурну формулу (5.54) можна записати так:

$$\int_{x_0-h}^{x_0+h} f(x)dx = \frac{h}{3} (f(x_0 - h) + 4f(x_0) + f(x_0 + h)) - \frac{h^5}{90} f^{(IV)}(\xi), \quad (5.56)$$

Це і є *квадратурна формула Сімпсона* або формула парабол із залишковим членом. Вона точна для многочлена третього степеня, бо похідна четвертого порядку від такого многочлена дорівнює нулю. З формули (5.55) легко знайти таку оцінку для абсолютної похибки чисельного інтегрування за формулою Сімпсона:

$$|R(f)| \leq \frac{h^5}{90} M_4, \quad M_4 = \max_{[x_0-h; x_0+h]} |f^{(IV)}(x)|.$$

Якщо треба обчислити $\int_a^b f(x)dx$ з достатньою точністю, то відрізок $[a; b]$ ділять на $2n$ рівних відрізків завдовжки $h=(b-a)/2n$ і до кожного з відрізків $[x_{2k}; x_{2k+2}]$; $(k=0, 1, \dots, n-1)$ застосовують формулу Сімпсона (5.56). Тоді

$$\begin{aligned} \int_a^b f(x)dx &= \sum_{k=0}^{n-1} \int_{x_{2k}}^{x_{2k+2}} f(x)dx = \frac{h}{3} \sum_{k=0}^{n-1} (f(x_{2k}) + 4f(x_{2k+1}) + f(x_{2k+2})) - \frac{h^5}{90} \sum_{k=0}^{n-1} f^{(IV)}(\xi_k) = \\ &= \frac{h}{3} (f(a) + f(b) + 2 \sum_{k=0}^{n-1} f(x_{2k}) + 4 \sum_{k=0}^{n-1} f(x_{2k+1})) - \frac{h^5}{90} \sum_{k=0}^{n-1} f^{(IV)}(\xi_k), \end{aligned}$$

де $\xi_k \in [x_{2k}; x_{2k+2}]$; $(k=0, 1, \dots, n-1)$.

Оскільки $f^{(IV)}(x)$ неперервна на відрізку $[a, b]$, то існує точка $\xi \in [a;$

$b]$ така, що $\frac{1}{n} \sum_{k=0}^{n-1} f^{(IV)}(\xi_k) = f^{(IV)}(\xi), \xi \in [a; b]$.

Таким чином отримаємо *узагальнену формулу Сімпсона* (парабол) із залишковим членом вигляду:

$$\int_a^b f(x)dx = \frac{h}{3} (f(a) + f(b) + 2 \sum_{k=0}^{n-1} f(x_{2k}) + 4 \sum_{k=0}^{n-1} f(x_{2k+1})) - \frac{h^4}{180} (b-a) f^{(IV)}(\xi). \quad (5.57)$$

Залишковий член узагальненої формули Сімпсона

$$R(f) = -\frac{h^4}{180} (b-a) f^{(IV)}(\xi) = -\frac{(b-a)^5}{2880n^4} f^{(IV)}(\xi), \quad h = \frac{b-a}{2n}. \quad (5.58)$$

Звідси отримаємо таку оцінку абсолютної похибки чисельного інтегрування за узагальненою формулою Сімпсона:

$$|R(f)| \leq \frac{h^4}{180} (b-a) M_4 = -\frac{(b-a)^5}{2880n^4} M_4; \quad M_4 = \max_{[a;b]} |f^{(IV)}(x)|. \quad (5.59)$$

Якщо наближене значення інтеграла треба обчислити з точністю $\varepsilon > 0$, то відповідний крок інтегрування h визначається нерівністю $h \leq \sqrt[4]{\frac{180\varepsilon}{(b-a)M_4}}$, або, що те саме, відрізок $[a; b]$ треба поділити на n рівних частин, де

$$h \leq \sqrt[4]{\frac{(b-a)^5 M_4}{2880\varepsilon}}. \quad (5.60)$$

За узагальненою формулою Сімпсона обчислимо наближене значення інтеграла $\int_0^1 x \cos x dx = \sin 1 + \cos 1 - 1 = 0,38177334$ з кроком $h=0,1$ і оцінимо повну абсолютну похибку Δ_f . За формулою (5.57), опускаючи проміжні обчислення знайдемо: $I_{cm} = 0,38177448 \approx 0,3817745$.

Для того, щоб оцінити залишковий член $R(f)$ формули Сімпсона за формулою (5.59), треба знайти похідну четвертого порядку від функції $f(x) = x \cdot \cos x$. Маємо $f^{(IV)}(x) = 4 \sin x + x \cos x$, звідси

$$|f^{(IV)}(x)| = 4|\sin x| + |x||\cos x| < 4 + 1 = 5; \quad M_4 = \max_{[0;1]} |f^{(IV)}(x)| = 5.$$

Тому, для залишкового члена $R(f)$ за формулою (5.59) ($a=0$; $b=1$; $h=0,1$; $M_4=5$) дістанемо $|R(f)| \leq 0,278 \cdot 10^{-5}$.

Похибка остаточного округлення $\Delta_0 = 0,2 \cdot 10^{-7}$, а неусувна похибка

$$\tilde{R} = \Delta_f \sum_{k=1}^n |A_k| = \Delta_f \cdot \frac{h}{3} \cdot 6n = 2nh\Delta_f = (b-a)\Delta_f = 0,5 \cdot 10^{-7},$$

оскільки $(b-a)/2n=h$, а значення підінтегральної функції f у вузлах x_k ($k=0, 1, \dots, 10$) обчислювали з точністю $0,5 \cdot 10^{-7}$, тобто $\Delta_f = 0,5 \cdot 10^{-7}$.

Отже, повна абсолютна похибка чисельного інтегрування функції $f(x) = x \cdot \cos x$ дорівнюватиме:

$$\Delta_f = R(f) + \tilde{R} + \Delta_0 = 0,278 \cdot 10^{-5} + 0,5 \cdot 10^{-7} + 0,2 \cdot 10^{-7} = 0,285 \cdot 10^{-5} < 0,3 \cdot 10^{-5}.$$

Отже, обчислене за формулою Сімпсона для $n=10$, $h=0,1$ наближене значення інтеграла (5.57) має п'ять правильних значущих цифр, тобто

$$\int_0^1 x \cos x dx = 0,381774 \pm 0,000003$$

(у відповіді збережено одну сумнівну цифру).

Найбільший внесок у повну абсолютну похибку узагальненої формули Сімпсона вносить залишковий член $R(f)$. Тому для визначення кількості відрізків n розбиття $[a; b]$, яке гарантує обчислення наближеного

значення інтеграла з точністю $\varepsilon > 0$, достатньо скористатися формулою (5.56). Звичайно, всі проміжні обчислення при цьому слід виконувати з точністю, більшою за ε .

Наприклад, щоб обчислити наближене значення інтеграла (5.57) з точністю $\varepsilon = 0,5 \cdot 10^{-4}$ (з чотирма правильними десятковими знаками), треба відрізок $[0; 1]$ поділити не менше як на три рівні частини, бо за формулою (5.56) ($a=0; b=1; M_4=5$) маємо

$$n > \sqrt[4]{\frac{5 \cdot 10^4}{0,5 \cdot 2880}} = 5 \sqrt[4]{\frac{1}{18}} = 5 \cdot 0,48549176 = 2,4274588 \approx 3.$$

Обчислимо інтеграл (5.57) за формулою (5.53), поклавши $n=2, 4, 8, 16$ (це відповідає $h=0,25; 0,125; 0,0625; 0,03125$). Знайдемо $I_2=0,38182200; I_4=0,38177633; I_8=0,38177346; I_{16}=0,38177333$. А це означає, що I_2 має три, I_4 – п'ять, I_8 – шість, а I_{16} – вісім правильних значущих десятичних цифр.

5.6 Порівняння і практичне оцінювання похибки квадратурних формул

Вище розглянуто деякі квадратурні формули. Узагальнені формули трапецій і Сімпсона – це формули замкненого типу, середніх прямокутників – відкритого типу, а лівих і правих прямокутників – напівзамкненого і напіввідкритого типу. Як вже було зазначено, точне значення інтеграла визначають за формулами: лівих і правих прямокутників – якщо підінтегральна функція стала; середніх прямокутників і трапецій – якщо підінтегральна функція лінійна, і за формулою Сімпсона – якщо підінтегральною функцією є многочлен степеня, не вищого за третій [1].

Точність квадратурної формули характеризується порядком залишкового члена $R(f)$ стосовно степеня кроку інтегрування h . З формул (5.32) – (5.37), (5.41) і (5.58) видно, що $R(f)$ квадратурних формул залежить від кроку інтегрування h і $R(f) \rightarrow 0$ ($h \rightarrow 0$). Кажуть, що $R(f)$ має порядок p (p – натуральне число) відносно h , якщо існує скінченна границя

$$\lim_{h \rightarrow 0} \frac{R(f)}{h^p} = M \neq 0,$$

і це записують так: $R(f) = O(h^p)$.

Отже, залишкові члени формул лівих і правих прямокутників відносно h мають перший порядок: $R(f) = O(h)$, середніх прямокутників і трапецій – другий: $R(f) = O(h^2)$, а Сімпсона – четвертий: $R(f) = O(h^4)$. Квадратурну формулу вважають тим точнішою, чим більший порядок її залишкового члена $R(f)$. З вищерозглянутих квадратурних формул найбільш точною є формула Сімпсона, а найбільш неточними – лівих і правих прямокутників. Точність формул середніх прямокутників і трапецій

однакова [1].

Як було зазначено вище, іноді оцінити залишковий член $R(f)$ квадратурної формули дуже важко або взагалі неможливо, наприклад тоді, коли функцію задано графічно чи таблично і аналітичний вираз її невідомий, або коли функцію задано складним аналітичним виразом і її похідні важко оцінити. Але якщо похідну певного порядку знайдено, то оцінити її за модулем на відрізку інтегрування завжди можна, побудувавши за допомогою комп'ютера таблицю значень похідної. Оцінити $R(f)$ квадратурної формули можна й тоді, коли не вдається оцінити зверху модуль похідної підінтегральної функції. Важливо лише знати порядок $R(f)$ відносно кроку інтегрування h . Для цього використовують *метод подвійного перерахунку* [1].

Нехай залишковий член деякої квадратурної формули має порядок p відносно кроку інтегрування h , тобто $R(f) = O(h^p)$, $p \in \mathbb{N}$. Припустимо також, що похідна, яка проходить до залишкового члена $R(f)$, на відрізку інтегрування $[a; b]$ змінюється мало, а тому наближено її можна вважати сталою. Тоді залишковий член $R(f)$ набере вигляду $R(f) = Mh^p$, де M – деяка невідома стала.

Якщо відрізок $[a; b]$ поділити на n і $2n$ рівних частини ($h = (b-a)/n$, $h/2 = (b-a)/2n$) та обчислити за квадратурною формулою наближені значення I_n і I_{2n} інтеграла $I = \int_a^b f(x) dx$, а відповідні їм залишкові члени позначити через $R_n(f)$ і $R_{2n}(f)$, то дістанемо дві рівності

$$I = I_n + R_n(f) = I_n + Mh^p; \quad I = I_{2n} + R_{2n}(f) = I_{2n} + M(h/2)^p. \quad (5.61)$$

Ці рівності можна розглядати як лінійну систему рівнянь відносно I та M . Виключивши з цієї системи I , знайдемо для M значення

$$M = (2^p / (2^p - 1)) \cdot ((I_{2n} - I_n) h^p).$$

Підставивши це значення M у вираз для $R_{2n}(f)$, дістанемо

$$R_{2n}(f) = (I_{2n} - I_n) / (2^p - 1). \quad (5.62)$$

Отже, залишковий член квадратурної формули пропорційний різниці двох наближених значень інтеграла, обчислених за цією ж квадратурною формулою з кроками h і $h/2$. Таке оцінювання похибки квадратурної формули називають *правилом Рунге*. Якщо тепер (5.62) підставити у друге рівняння системи (5.61), то знайдемо уточнене значення інтеграла

$$I_{n,2n} = I_{2n} + (I_{2n} - I_n) / 2^p - 1. \quad (5.63)$$

Обчислення наближеного значення інтеграла за формулою (5.63) на-

зивають *екстраполяцією за Річардсоном*. Якщо $I_n \neq I_{2n}$, то уточнене значення $I_{n,2n}$ ніколи не лежить між I_n і I_{2n} . Якщо $I_{2n} > I_n$, то з формули (5.63) випливає, що $I_{n,2n} > I_{2n} = \max\{I_n, I_{2n}\}$. Отже, наближення $I_{n,2n}$ визначають з наближень I_n та I_{2n} в результаті операції екстраполяції, тому й сам спосіб обчислення $I_{n,2n}$ назвали екстраполяцією.

У таблиці 5.1 для квадратурних формул, розглянутих вище, подано значення порядку p залишкового члена відносно кроку h , формули для обчислення значень залишкового члена $R_{2n}(f)$ і уточненого значення інтеграла $I_{n,2n}$.

Таблиця 5.1 – Формули для обчислення значень

Узагальнена квадратурна формула	p	$R_{2n}(f)$	$I_{n,2n}$
Лівих і правих прямокутників	1	$I_{2n} - I_n$	$I_{2n} + (I_{2n} - I_n)$
Середніх прямокутників і трапецій	2	$1/3(I_{2n} - I_n)$	$I_{2n} + 1/3(I_{2n} - I_n)$
Сімпсона	4	$1/15(I_{2n} - I_n)$	$I_{2n} + 1/15(I_{2n} - I_n)$

З таблиці 5.1 видно, що для обчислення наближеного значення інтеграла з точністю $\varepsilon > 0$ методом подвійного перерахунку треба:

1. Обчислити наближені значення інтеграла I_n та I_{2n} з кроком $h = (b-a)/n$ і $h/2 = (b-a)/2n$.

2. За формулою (5.52) обчислити наближене значення похибки $R_{2n}(f)$ чисельного інтегрування.

3. Порівняти $R_{2n}(f)$ з ε . Якщо $|R_{2n}(f)| < \varepsilon$, то за формулою (5.53) обчислити уточнене значення інтеграла $I_{n,2n}$ і процес обчислень припинити. Якщо $|R_{2n}(f)| \geq \varepsilon$, то, зберігши значення I_{2n} , відрізок $[a; b]$ поділити на $4n$ рівних частин і обчислити $R_{4n}(f)$, яке знову порівнюється з ε . Цей процес послідовного збільшення вдвічі числа вузлів квадратурної формули (зменшення вдвічі кроку інтегрування) продовжують доти, доки на певному кроці k не виконуватиметься нерівність $|R_{2n}^k(f)| < \varepsilon$.

Отже, очевидно, що цей алгоритм має циклічний характер.

Приклад 1. Використавши правило Рунге і екстраполяцію за Річардсоном, обчислити довжину дуги еліпса $(x^2/9) + (y^2/4) = 1$ за квадратурними формулами [1].

Розв'язування. Відомо, що довжина l дуги еліпса

$$l = a \int_0^{2\pi} \sqrt{1 - \varepsilon^2 \cos^2 t} dt = 4a \int_0^{\pi/2} \sqrt{1 - \varepsilon^2 \cos^2 t} dt,$$

де $\varepsilon = \frac{\sqrt{a^2 - b^2}}{a}$ – ексцентриситет еліпса. Якщо в останньому інтегралі

замінити незалежну змінну за формулою $t = \pi/2 - \tau$, то дістанемо:

$$\int_0^{\pi/2} \sqrt{1 - \varepsilon^2 \cos^2 t} dt = a \int_0^{\pi/2} \sqrt{1 - \varepsilon^2 \cos^2 \tau} d\tau = E(\varepsilon^2),$$

де $E(\varepsilon^2)$ – так званий повний еліптичний інтеграл другого роду. Отже, довжина дуги еліпса $l=4aE(\varepsilon^2)$.

За умовою $a=3, b=2, \varepsilon^2 = \frac{5}{9}$ тому $l = 4a \int_0^{\pi/2} \sqrt{1 - \frac{5}{9} \sin^2 t} dt = 4aE\left(\frac{5}{9}\right)$. Для обчислення наближене значення інтеграла $E\left(\frac{5}{9}\right) = \int_0^{\pi/2} \sqrt{1 - \frac{5}{9} \sin^2 t} dt$, знайдемо спочатку значення функції $f(t) = \sqrt{1 - \frac{5}{9} \sin^2 t}$ на відрізку $[0; \pi/2]$ з кроком $h = \pi/24$ (табл. 5.2).

Користуючись цією таблицею, знаходимо:

1. За формулою лівих прямокутників:

а) $n=3, h = \pi/6 = 0,52359876, E_3 = h(f_0 + f_4 + f_8) = \pi/6 \cdot 2,6917233 = 1,4093830$;

б) $n=6, h = \pi/12 = 0,26179938,$

$E_6 = h(f_0 + f_2 + f_4 + f_6 + f_8 + f_{10}) = h(f_0 + \Sigma_2) = \pi/12 \cdot 5,2167929 = 1,3657531,$

$R_6(f) = (E_6 - E_3) = -0,0436299, E_{3,6} = E_6 + E_6(f) = 1,3221232$;

в) $n=12, h = \pi/24 = 0,13089969,$

$E_{12} = h \sum_{k=0}^{11} f_k = h(f_0 + \Sigma_1 + \Sigma_2) = \pi/24 \cdot 10,266919 = 1,3439365,$

$R_{12}(f) = E_{12} - E_6 = -0,0218166, E_{6,12} = E_{12} + R_{12}(f) = 1,3221199.$

Отже, $E_{3,6}$ і $E_{6,12}$ мають по шість правильних значущих цифр, тому $E(5/9) = 1,32212$.

Таблиця 5.2 – Значення функції з прикладу 1

№	t_i	$\sin t_i$	$f_i = f(t_i)$		
0	0	0	1		
1	$\pi/24 = 0,13089969$	0,13052619		0,99525624	
2	$\pi/12 = 0,26179938$	0,25881903			0,98121598
3	$\pi/8 = 0,39269907$	0,38268343		0,95845751	
4	$\pi/6 = 0,52359876$	0,50000000			0,92796072
5	$5\pi/24 = 0,65449845$	0,60876144		0,89113208	
6	$\pi/4 = 0,78539815$	0,70710681			0,84983657
7	$7\pi/24 = 0,91629783$	0,79335334		0,80642922	
8	$\pi/3 = 1,0471975$	0,86602543			0,76376259
9	$3\pi/8 = 1,1780972$	0,92387955		0,72512323	
10	$5\pi/12 = 1,3089969$	0,96592582			0,69401700
11	$11\pi/24 = 1,4398965$	0,99144484		0,67372806	
12	$\pi/2 = 1,5707963$	1	0,66666667		
		$2n = 12$	$\Sigma_0 = 1,66666667$	$\Sigma_1 = 5,0501264$	$\Sigma_2 = 4,21679286$
		$2n = 6$	$\Sigma'_0 = 1,66666667$	$\Sigma'_1 = 2,5250696$	$\Sigma'_2 = 1,6917233$

2. За формулою правих прямокутників:

а) $n=3, h=\pi/6=0,52359876, E_3=h(f_4+f_8+f_{12})=\pi/6 \cdot 2,3583900=1,2348501;$

б) $n=6, h=\pi/12=0,26179938,$

$E_6=h(f_2+f_4+f_6+f_8+f_{10}+f_{12})=h(f_{12}+\Sigma_2)=\pi/12 \cdot 4,8834596=1,2784867,$

$R_6(f)=(E_6-E_3)=0,0436366, E_{3,6}=E_6+E_6(f)=1,3221233;$

в) $n=12, h=\pi/24=0,13089969,$

$E_{12}=h \sum_{k=0}^{12} fk = h(f_{12}+\Sigma_1+\Sigma_2)=\pi/24 \cdot 9,9335859=1,3003033,$

$R_{12}(f)=E_{12}-E_6=0,0218163, E_{6,12}=E_{12}+R_{12}(f)=1,3221196.$

Як бачимо, і за формулою правих прямокутників наближене значення $E(5/9)=1,32212$, де усі цифри правильні.

3. За формулою трапецій:

а) $n=3, h=\pi/6=0,52359876, h/2=\pi/12=0,26179938;$

$E_3=h/2(f_0+2(f_4+f_8)+f_{12})=\pi/12 \cdot 5,0501133=1,3221165;$

б) $n=6, h=\pi/12, h/2=\pi/24=0,13089969,$

$E_6=h/2(f_0+2(f_2+f_4+f_6+f_8+f_{10})+f_{12})=h/2(\Sigma_0+2\Sigma_2)=\pi/24 \cdot 10,10025239=$
 $=1,3221195, R_6(f)=1/3(E_6-E_3)=0,0000011, E_{3,6}=E_6+E_6(f)=1,3221209;$

в) $n=12, h=\pi/24, h/2=\pi/48=0,065449845=6,5449845 \cdot 10^{-2},$

$E_{12}=h/2(f_0+2 \sum_{k=1}^{11} fk +f_{12})=h/2(\Sigma_0+2(\Sigma_1+\Sigma_2))=\pi/48 \cdot 20,2005051=1,3221204,$

$R_{12}(f)=1/3(E_{12}-E_6)=1/3 \cdot 0,00000006=0,00000002 \approx 0,$

$E_{6,12}=E_{12}=1,3221204.$

Отже, $E(5/9)=1,32212$, де всі цифри є правильними.

4. За формулою Сімпсона:

а) $n=3, h=\pi/12, h/3=\pi/36=0,087266461=8,7266461 \cdot 10^{-2},$

$E_3=h/3(f_0+4(f_2+f_6+f_{10})+2(f_4+f_8)+f_{12})=h/3(\Sigma_0+4\Sigma_1+2\Sigma_2)=(\pi/36) \times$
 \times

$(1,6666667+4 \cdot 10,100278+2 \cdot 1,6917233)=\pi/36 \cdot 15,150391=1,3221210;$

б) $n=6, h=\pi/24, h/3=\pi/72=4,3633230 \cdot 10^{-2},$

$E_6=h/3(\Sigma_0+4\Sigma_1+2\Sigma_2)=\pi/72 \cdot 30,300758=1,3221199,$

$R_6(f)=1/15(E_6-E_3)=1/15 \cdot 0,1 \cdot 10^{-6} \approx 0, E_{3,6}=E_6=1,3221199 \approx 1,32212.$

Для порівняння наведено значення інтеграла $E(5/9)$, обчислене за формулою лінійного інтерполювання на основі табличних даних для функції $E(\varepsilon^2)$, взятих з “П’ятизначних математичних таблиць” Б. І. Сегала і К. А. Семендяєва:

k	ε^2	E	ΔE
0	0,55226	1,32384	-0,00303
1	0,55805	1,32081	-0,00303
2	0,56328	1,31778	

Наше значення $\varepsilon^2 = 5/9 = 0,5$ лежить між $\varepsilon_0^2 = 0,55226$ і $\varepsilon_1^2 = 0,55805$.
Значення $E_0 = 1,32384$, $\Delta E_0 = -0,00303$, $h = \varepsilon_1^2 - \varepsilon_0^2 = 0,00579$, $\varepsilon^2 - \varepsilon_0^2 = 5/9 - 0,55226 = -0,00330$. Тоді за формулою лінійної інтерполяції маємо:

$$E(5/9) = E_0 + \Delta E_0 / h (\varepsilon^2 - \varepsilon_0^2) = 1,32384 + (-0,00303/0,00579) \cdot (-0,00330) = 1,322113.$$

Після округлення до п'ятого десяткового розряду маємо: $E(5/9) = 1,32211$, що добре узгоджується з виконаними вище обчисленнями.

Отже, екстраполяція за Річардсоном – потужний і універсальний алгоритм підвищення точності чисельного інтегрування функцій. Навіть у таких методах низького порядку точності, як методи лівих і правих прямокутників, її застосування дає змогу дістати результат достатньо високого порядку точності при мінімальних обсягах обчислювальної роботи. Саме тому метод експлуатації за Річардсоном з розвитком комп'ютерів дістав значного розвитку у багатьох розділах обчислювальної математики [1].

5.7 Контрольні запитання

1. В яких випадках застосовують чисельне диференціювання?
2. Як отримують найпростіші вирази для похідних?
3. Наведіть кілька найпростіших формул чисельного диференціювання.
4. Як залежить точність чисельного диференціювання від кроку сітки?
5. Наведіть постановку чисельного інтегрування.
6. Наведіть та охарактеризуйте два випадки чисельного інтегрування, які розрізняють залежності від способу задання підінтегральної функції.
7. Перерахуйте найпростіші квадратурні формули чисельного інтегрування.
8. Наведіть загальні формули лівих, правих та середніх прямокутників з відповідною геометричною інтерпретацією.
9. Наведіть вирази для визначення похибок при обчисленні інтеграла за формулами лівих, правих та середніх прямокутників.
10. Наведіть загальну формулу трапецій, її геометричну інтерпретацію, та вираз для оцінювання похибки інтегрування за даною формулою.
11. Які формули чисельного інтегрування називають гауссовими?
12. Наведіть узагальнену квадратурну формулу Сімпсона та оцінювання абсолютної похибки чисельного інтегрування за нею.
13. З якою метою застосовують метод подвійного перерахунку? Наведіть алгоритм цього методу.
14. Від яких чинників залежить точність квадратурних формул?
15. Наведіть правило Рунге.
16. Наведіть порівняльну характеристику відомих Вам квадратурних формул.

5.8 Завдання

Задача 1. Методом чисельного диференціювання знайти першу похідну у точці, заданій викладачем, для функції, заданої у вигляді таблиці, згідно з індивідуальним варіантом.

1	X	0,525	0,526	0,527	0,528
	F(X)	0,50121	0,50208	0,50294	0,50381
2	X	50	55	60	65
	F(X)	1,6990	1,7404	1,7782	1,8129
3	X	0,325	0,326	0,327	0,328
	F(X)	0,50121	0,50208	0,50294	0,50381
4	X	20	25	30	35
	F(X)	1,6990	1,7404	1,7782	1,8129
5	X	0,525	0,526	0,527	0,528
	F(X)	0,50121	0,50208	0,50294	0,50381
6	X	5	10	15	20
	F(X)	1,500	1,604	1,782	1,812
7	X	0,825	0,826	0,827	0,828
	F(X)	0,70121	0,70208	0,70294	0,70381
8	X	10	15	20	25
	F(X)	1,6990	1,7404	1,7782	1,8129
9	X	0,825	0,826	0,827	0,828
	F(X)	0,90121	0,90208	0,90294	0,90381
10	X	12	17	22	27
	F(X)	1,6990	1,7404	1,7782	1,8129

Задача 2. За узагальненою формулою середніх прямокутників:

а) обчислити значення інтеграла за вказаним розбиттям відрізка інтегрування $[a; b]$ на N_0 рівних частин і оцінити залишковий член;

б) з точністю $\varepsilon > 0$ визначити n і обчислити наближене значення інтеграла для $N \geq n$ (N вибрати так, якщо це дозволяють умови задачі, щоб $(b-a)$ ділилося на нього точно).

Задача 3. За узагальненою формулою Сімпсона обчислити значення інтеграла із заданою точністю $\varepsilon > 0$, виходячи з вказаного значення N_0 . Крок інтегрування h визначається рівністю $h=(b-a)/(2N_0)$. Значення кроку інтегрування h , що забезпечує вказану точність, визначити за допомогою подвійного перерахунку з поділом відрізка $[a; b]$ на N і $2N$ частин. Надрукувати межі інтегрування, послідовність пар (N, I_N) , $(2N, I_{2N})$ і т. д., а також остаточне значення $R=1/15 \cdot |I_{2N} - I_N|$.

Задача 4. Побудувати алгоритм та програму обчислення наближеного значення інтеграла за узагальненою формулою трапецій. Обчислити наближене значення інтеграла при заданому N_0 (кількість відрізків інтегрування) і оцінити залишковий член; обчислити наближене значення інтеграла з точністю $\varepsilon > 0$ подвоєнням кількості вузлів інтегрування $N_0, 2N_0, 4N_0$ і т. д. до виконання нерівності $|R(f)| \approx |I_{2n} - I_n|/3 < \varepsilon$. Навести межі інтегрування, послідовні пари $(N_0, I_{N_0}), (2N_0, I_{2N_0})$ і т. д., та $|R(f)| \approx |I_{2N} - I_N|/3$. Назвати правильні цифри обчисленого значення інтеграла. Порівняти обчислене значення інтеграла з отриманим за формулою Ньютона–Лейбніца.

Таблиця 5.3 – Варіанти завдань

Вар.	$F(x)$	a	b	N_0	ε
1	$e^x \cos x$	0	1	5	$0,5 \cdot 10^{-6}$
2	$x^2 e^{-x}$	0	1	5	$0,5 \cdot 10^{-8}$
3	$x^2 \operatorname{arctg} x$	0	1	5	$0,5 \cdot 10^{-8}$
4	$e^{\sin x} \sin 2x$	0	1	5	$0,5 \cdot 10^{-8}$
5	$e^{\cos 2x} \cos x$	0	1	5	$0,5 \cdot 10^{-8}$
6	$e^x/(1+x)$	0	1	5	$0,5 \cdot 10^{-8}$
7	$(1+x)^{3/2} e^{-x}$	0	1	5	$0,5 \cdot 10^{-8}$
8	$e^x/(3+2\cos x)$	0	1	5	$0,5 \cdot 10^{-8}$
9	$e^{\cos x}(1+x)$	0	1	5	$0,5 \cdot 10^{-8}$
10	$(1+x) \ln^2(1+x)$	0	1	5	$0,5 \cdot 10^{-6}$
11	$(1+x^2) e^{\operatorname{arctg} x}$	0	1	5	$0,5 \cdot 10^{-7}$
12	$e^{\sin x} \sin(\cos x)$	0	1	6	$0,5 \cdot 10^{-7}$
13	$\sin x \sqrt{1+x^2}$	0	1	6	$0,5 \cdot 10^{-8}$
14	$x^2 + 2 \ln x$	1	2	8	$0,5 \cdot 10^{-8}$
15	$x^2 \cos(x/2)$	1	2	8	$0,5 \cdot 10^{-8}$
16	$x^2 \sin(1+x^2)$	0	1	6	$0,5 \cdot 10^{-6}$
17	$x \cos x^2$	0	1	6	$0,5 \cdot 10^{-6}$
18	$\cos x/(1+\sin^3 x)$	0	1	6	$0,5 \cdot 10^{-8}$
19	$(1-x^2) \ln(1+x)$	0	1	5	$0,5 \cdot 10^{-6}$
20	$\sqrt{1+x} \cos x$	0	1	5	$0,5 \cdot 10^{-8}$
21	$\sqrt{1+x} \sin x$	0	1	5	$0,5 \cdot 10^{-8}$
22	$e^{-x}/(1+x)$	1	2	10	$0,5 \cdot 10^{-6}$
23	$\operatorname{arctg} x/(1+x^2)$	0	1	10	$0,5 \cdot 10^{-4}$
24	$x \ln^2 x$	2	3	8	$0,5 \cdot 10^{-8}$
25	$\sqrt{1-(7/16)\sin^2 x}$	0	$\pi/2$	6	$0,5 \cdot 10^{-8}$
26	$\arcsin e^{-(x/8)}$	1	7	6	$0,5 \cdot 10^{-8}$
27	$\arccos e^{-(x/6)}$	2	7	5	$0,5 \cdot 10^{-3}$
28	$\ln(2+\operatorname{tg}(x+10))$	0	14	7	$0,5 \cdot 10^{-8}$
29	$5x \sin x^2$	0	2	10	$0,5 \cdot 10^{-8}$
30	$3x^3 \cos x^2$	0	1	5	$0,5 \cdot 10^{-8}$

СЛОВНИК ТЕРМІНІВ

Математична модель (*mathematical model* – ММ) будується на основі деякого абстрагування, спрощення та ідеалізації об'єкта, а тому завжди є його наближеним описом. ММ – це, як правило, різноманітні рівняння, які описують поведінку досліджуваних об'єктів і явищ.

Чисельні методи (*numerical methods*) – потужний математичний засіб розв'язування складних прикладних задач, для яких, як правило, не існує точного розв'язку у вигляді явних формул або скінченної послідовності арифметичних операцій, кожна з яких виконується точно.

Похибка математичної моделі (*error of the mathematical model*) – обумовлена тим, що модель описує явище наближено, з припущеннями і спрощеннями.

Неусувна похибка (*fatal error*) – це похибка, яка зумовлена похибками у вхідних даних задачі і залежить від методу розв'язування задачі.

Похибка методу (*error of the method*) – пов'язана з необхідністю заміни неперервної моделі дискретною або із завершенням нескінченного ітераційного процесу після скінченної кількості ітерацій.

Похибка дискретизації (*sampling error*) – похибка, що виникає під час заміни неперервної моделі дискретною (інша назва – похибка апроксимації).

Похибка збіжності (*convergence error*) – це похибка, обумовлена обірванням ітераційного процесу.

Обчислювальні похибки (*computational error*) – це похибки, пов'язані з похибками округлення чисел.

Стійкість задачі (*task stability*) – це чутливість задачі до неточностей у вхідних даних. Задача є стійкою за вхідними даними, якщо її розв'язок безперервно залежить від вхідних даних, тобто малі похибки вхідних даних спричиняють малі похибки розв'язку задачі. Якщо ця умова не виконується – задача є нестійкою за вхідними даними. Це означає, що навіть незначні похибки вхідних даних можуть призвести до як завгодно великих похибок розв'язку.

Коректно поставлена задача (*well-posed task*) – це якщо для будь-яких вхідних даних з деякого класу існує єдиний і стійкий за вхідними даними розв'язок задачі.

Обчислювальний (або математичний) експеримент (*computing (or mathematical) experiment*) – це новий спосіб проведення теоретичних досліджень на базі математичних моделей, що передбачає застосування швидкодійних та потужних комп'ютерів для розв'язування складних прикладних задач.

Абсолютною похибкою (*absolute error*) подання величини A наближеним числом a називається величина, яка задовольняє нерівність $\Delta_a \geq |A - a|$.

Відносною похибкою (*relative error*) подання величини A наближеним числом a називається величина, яка задовольняє нерівність $\delta_a \geq |(A - a)/a|$, $a \neq 0$.

Сумісна система лінійних рівнянь (*compatible system of linear equations*) – така, що має розв'язок, в протилежному випадку вона є несумісною (або суперечливою).

Визначена система лінійних алгебраїчних рівнянь (*determined system of linear algebraic equations*) – це сумісна СЛАР, яка має один єдиний розв'язок.

Невизначена система лінійних алгебраїчних рівнянь (*undetermined system of linear algebraic equations*) – це сумісна СЛАР, яка має більше одного розв'язку.

Елементарними перетвореннями СЛАР називаються такі три типи перетворень: 1) перестановка двох рівнянь системи; 2) множення обох частин рівняння системи на будь-яке відмінне від нуля число; 3) додавання (віднімання) до обох частин одного рівняння відповідних частин іншого рівняння, помножених на будь-яке число.

Погано обумовлена СЛАР (*Poorly conditioned SLAE*) – це така СЛАР, у якій малі збурення коефіцієнтів і (або) вільних членів дуже збурюють її розв'язок. Якщо ж розв'язок збурюється незначно – СЛАР називають **добре обумовленою** (*well-conditioned SLAE*).

Алгебраїчна функція (*algebraic function*) – це якщо для отримання значення функції за даним значенням x слід виконати арифметичні операції і операції піднесення до степеня з раціональним показником.

Раціональна функція (*rational function*) – це алгебраїчна функція від змінної x , в якій над x не виконується ніяких інших дій, крім додавання, віднімання, множення, ділення і піднесення до цілого степеня.

Ціла раціональна функція (*integer rational function*) – раціональна функція, у яку змінна x не входить як дільник або не входить у вираз, який є дільником. Якщо в раціональній функції хоча б один раз зустрічається ділення на змінну x або змінна x входить у вираз, який є дільником, то така функція є **дробово-раціональною** (*fractional rational function*).

Ірраціональна функція (*irrational function*) – це функція від змінної x , в якій для отримання значення функції за даним значенням x необхідно виконати, крім чотирьох арифметичних дій (усіх або деяких), ще й обчислення кореня. При цьому функція буде ірраціональною лише тоді, коли аргумент x знаходиться під знаком радикала.

Трансцендентні функції (*transcendental functions*) – це неалгебраїчні функції: показникова; логарифмічна; тригонометричні; зворотні тригонометричні тощо.

Відокремлення кореня рівняння (*separating root of the equation*) – це знаходження такого інтервалу, всередині якого є корінь даного рівняння, і цей корінь – єдиний на даному інтервалі.

ЛІТЕРАТУРА

1. Фельдман Л. П. Чисельні методи в інформатиці / Фельдман Л. П., Петренко А. І., Дмитрієва О. А. – Київ : Видавнича група ВНУ, 2006.
2. Фельдман Л. П. Чисельні методи та математичні пакети. Лабораторний практикум / Л. П. Фельдман, В. П. Пасько. – Київ : Видавнича група ВНУ, 2007.
3. Возняк Л. С. Чисельні методи : методичний посібник для студентів природничих спеціальностей / Л. С. Возняк, С. В. Шарин. – Івано-Франківськ : “Плай”, 2001. – 64 с.
4. Мотигін В. В. Чисельні методи в інженерних дослідженнях. Частина I. Теоретичні відомості / Мотигін В. В., Роптанов В. І., Дудатьєв А. В. – Вінниця : ВНТУ, 2004. – 102 с.
5. Мотигін В. В. Чисельні методи в інженерних дослідженнях. Частина II. Чисельні методи в електронних таблицях (лабораторний практикум) / В. В. Мотигін, А. В. Дудатьєв. – Вінниця : ВНТУ, 2007. – 95 с.
6. Шаповаленко В. А. Чисельне обчислення функцій, характеристик матриць і розв’язування нелінійних рівнянь та систем рівнянь : навч. посібник / Шаповаленко В. А., Буката Л. М., Трофименко О. Г. – Одеса : ВЦ ОНАЗ, 2010. – 88 с.
7. Ляшенко М. Я. Чисельні методи / М. Я. Ляшенко, М. С. Головань. – К. : Либідь, 1996. – 288 с.
8. Гулин А. В. Численные методы / А. В. Гулин, А. А. Самарський. – М. : Наука, 1989. – 432 с.
9. Гетьманцев В. Д. Лінійна алгебра і лінійне програмування : навчальний посібник / Гетьманцев В. Д. – К. : Либідь, 2001. – 256 с.
10. Вычислительная математика / [Данилина Н. И., Дубровская Н. С., Кваша О. П., Смирнов Г. Л.]. – М. : Высшая школа, 1985. – 472 с.
11. Волков Е. А. Численные методы / Волков Е. А. – М. : Наука, 1987. – 248 с.
12. Методи обчислень : практикум на ЕОМ : навчальний посібник / [В. Л. Бурківська, С. О. Войцехівський, І. П. Гаврилук та ін.]. – К. : Вища школа, 1995. – 303 с.
13. Коварцев А. Н. Численные методы : учебное пособие / Коварцев А. Н. – Самара : Высшая школа, 1998. – 134 с.
14. Маликов В. Т. Вычислительные методы и применение ЭВМ : учебное пособие / В. Т. Маликов, Р. Н. Кветный. – К. : Вища школа, 1989. – 213 с.
15. Хемминг Р. В. Численные методы для научных работников и инженеров / Хемминг Р. В. – М. : Наука, 1968. – 400 с.
16. Краскевич В. Е. Численные методы в инженерных исследованиях / Краскевич В. Е., Зеленский К. Х., Гречко В. И. – К. : Вища шк., 1986. – 263 с.

17. Калитки Н. Н. Численные методы / Калитки Н. Н. – М. : Наука, 1978. – 512 с.
18. Гусак А. А. Элементы методов вычислений / Гусак А. А. – [2-е изд.] – Мн. : Изд – во БГУ, 1982. – 166 с.
19. Иванов В. В. Методы вычислений на ЭВМ : справочное пособие / Иванов В. В. – К. : Наукова думка, 1986. – 584 с.
20. Дубовий В. М. Основи застосування ЕОМ у інженерній діяльності / В. М. Дубовий, Р. Н. Кветний. – К. : ІСД МО України, 1994. – 285 с.
21. Бахвалов Н. С. Численные методы / Бахвалов Н. С., Жидков Н. П., Кобельков Г. М. – М. : Наука, 1987. – 598 с.
22. Жалдак М. І. Чисельні методи математики : посібник для самоосвіти вчителів / М. І. Жалдак, Ю. С. Рамський. – К. : Вища школа, 1984. – 206 с.
23. Орлов О. В. Лабораторные работы по курсу «Численные методы» (ФТИ) / Орлов О. В. [Электронный ресурс]. Режим доступа: http://orloff.am.tpu.ru/chisl_metod/
24. Херсонський Віртуальний Університет Чисельні методи [Електронний ресурс]. Режим доступу: <http://dls.ksu.kherson.ua/DLS/Library/Catalog.aspx?section=790b2934-7126-40bd-8640-198bf5b4aee8>
25. Кравцова Л. В. Лекції з дисципліни Чисельні методи / Кравцова Л. В. [Електронний ресурс]. Режим доступу: <http://dls.ksu.kherson.ua/dls/Library/LibdocView.aspx?id=bbd94fad-a930-49da-952c-df2838670a5b>
26. Чисельні методи. Лабораторний практикум [Електронний ресурс]. Режим доступу: <http://uk.wikibooks.org/wiki> (21.11.2012).

Навчальне видання

**Олег Костянтинович Колесницький
Ігор Ростиславович Арсенюк
Володимир Іванович Месюра**

ЧИСЕЛЬНІ МЕТОДИ

Навчальний посібник

Редактор Т. Старічек

Оригінал-макет підготовлено І. Арсенюком

Підписано до друку 03.05.2017 р.
Формат 29,7×42¼. Папір офсетний.
Гарнітура Times New Roman.
Ум. друк. арк. 7,51
Наклад 50 пр. Зам. № 2017-080

Видавець та виготовлювач
Вінницький національний технічний університет,
інформаційний редакційно-видавничий центр.
ВНТУ, ГНК, к. 114.
Хмельницьке шосе, 95,
м. Вінниця, 21021.
Тел. (0432) 59-85-32, 59-87-38,
press.vntu.edu.ua
E-mail: kivc.vntu@gmail.com.
Свідоцтво суб'єкта видавничої справи
серія ДК № 3516 від 01.07.2009 р.