



Міністерство освіти і науки України  
Сумський державний університет

Дворниченко А. В.

## **ЧИСЕЛЬНІ МЕТОДИ ОБРОБКИ ДАНИХ**

Конспект лекцій

У двох частинах  
Частина I

Суми  
Сумський державний університет  
2024

Міністерство освіти і науки України  
Сумський державний університет

## ЧИСЕЛЬНІ МЕТОДИ ОБРОБКИ ДАНИХ

Конспект лекцій  
для здобувачів спеціальності 113 *«Прикладна математика»*  
очної форми здобуття вищої освіти

У двох частинах  
Частина I

Затверджено на засіданні  
кафедри прикладної математики та моделювання складних систем як конспект лекцій з дисципліни «Чисельні методи обробки даних».  
Протокол №5 від 09.10.2023.

Суми  
Сумський державний університет  
2024

Чисельні методи обробки даних : конспект лекцій / укладач  
А. В. Дворниченко. — Суми : Сумський державний університет,  
2024. — Ч.1. - 116 с.

Кафедра прикладної математики та моделювання складних  
систем

# Зміст

<b>Лекція 1 Вступ. Чисельні методи як інструмент сучасних наукових досліджень</b> . . . . .	5
1.1 Приклади математичних моделей . . . . .	6
1.2 Етапи розв'язання практичних задач на комп'ютері . . . . .	8
1.3 Методи реалізації математичних моделей . . . . .	10
1.4 Характеристики чисельних методів . . . . .	11
<b>Лекція 2 Елементи теорії похибок</b> . . . . .	14
2.1 Абсолютна та відносна похибки . . . . .	15
2.2 Значущі цифри числа. Правильні та сумнівні цифри числа . . . . .	20
2.3 Правила округлення . . . . .	22
<b>Лекція 3 Пряма та обернена задачі теорії похибок</b> . . . . .	26
3.1 Пряма задача теорії похибок . . . . .	26
3.2 Обернена задача теорії похибок . . . . .	30
3.3 Правила підрахунку цифр . . . . .	31
<b>Лекція 4 Відокремлення коренів рівняння. Уточнення коренів рівняння методом дихотомії та хорд</b> . . . . .	33
4.1 Поставлення задачі . . . . .	33
4.2 Етапи наближеного розв'язання нелінійних рівнянь. Локалізація коренів . . . . .	35
4.2.1 Графічна локалізація кореня . . . . .	35
4.2.2 Аналітична локалізація кореня . . . . .	38
4.2.3 Таблична локалізація кореня . . . . .	39
4.3 Етапи наближеного розв'язання нелінійних рівнянь. Уточнення коренів нелінійного рівняння . . . . .	40
4.4 Метод половинного ділення (бісекції, дихотомії) . . . . .	40
4.5 Метод хорд . . . . .	45
<b>Лекція 5 Метричні простори. Метод простих ітерацій розв'язання нелінійних рівнянь</b> . . . . .	48
5.1 Метричні простори . . . . .	48
5.2 Метод простих ітерацій . . . . .	50
<b>Лекція 6 Метод Ньютона уточнення кореня нелінійного рівняння. Модифікації методу Ньютона</b> . . . . .	58
6.1 Метод Ньютона уточнення кореня нелінійного рівняння . . . . .	58

6.2	Спрощений метод Ньютона (метод однієї дотичної) . . . . .	64
6.3	Метод січних . . . . .	65
6.4	Метод Стефенсена . . . . .	67
<b>Лекція 7 Класифікація методів розв’язання систем лінійних алгебраїчних рівнянь та їх порівняльні характеристики. Точні методи розв’язання СЛАР . .</b>		
7.1	Розв’язання систем лінійних алгебраїчних рівнянь . . . . .	68
7.2	Поставлення задачі . . . . .	68
7.3	Точні методи розв’язання СЛАР . . . . .	73
<b>Лекція 8 Прямі методи розв’язання СЛАР. LU-розкладання квадратної невивродженої матриці, алгоритм Краута, алгоритм Холецького. . . . .</b>		
8.1	<i>LU</i> -розкладання квадратної невивродженої матриці . . . . .	87
8.2	Алгоритм Краута . . . . .	88
8.3	Алгоритм Холецького . . . . .	92
<b>Лекція 9 Ітераційні методи розв’язання СЛАР . . . . .</b>		
9.1	Ітераційні методи розв’язання СЛАР . . . . .	98
9.2	Метод простих ітерацій . . . . .	98
9.3	Метод Зейделя розв’язання СЛАР . . . . .	104
<b>Лекція 10 Чисельні методи розв’язання систем нелінійних рівняньq . . . . .</b>		
10.1	Поставлення задачі . . . . .	107
10.2	Метод Ньютона . . . . .	108
<b>Список літератури . . . . .</b>		<b>115</b>

# Лекція 1

## Вступ. Чисельні методи як інструмент сучасних наукових досліджень

Реальні інженерні та фізичні задачі у всіх галузях науки і техніки зазвичай вирішуються за допомогою використання двох підходів: фізичного експерименту та попереднього аналізу конструкцій, схем, явищ із метою вибору будь-яких їх оптимальних параметрів.

**Перший підхід** пов'язаний з великими та не завжди виправданими витратами матеріальних і тимчасових ресурсів.

**Другий підхід** пов'язаний із математичним моделюванням, в основі якого закладені знання фундаментальних законів природи і побудова на їх основі математичних моделей для довільних технічних і наукових завдань. Математичні моделі розробляються з дотриманням коректності та адекватності до реальних процесів, але зазвичай з урахуванням простоти їх технічної реалізації.

Математична модель, заснована на спрощенні, ідеалізації, не тотожна реальному явищу, об'єкту, а є його наближеним описом. Проте завдяки заміні реального об'єкта наближеною моделлю стає можливим його математичний опис і застосування математичного апарату щодо його аналізу. Математика дозволяє провести детальний аналіз явища, що розглядається, передбачити його поведінку в різних умовах та в майбутньому.

Складність математичної моделі та її дослідження залежить від складності досліджуваного об'єкта. Якщо раніше математичні методи та моделі застосовувалися лише в механіці, фізиці, астрономії, що вивчають найпростіші форми руху, то з появою ЕОМ і розвитком обчислювальної математики математичні методи знаходять застосування і в інших галузях діяльності людини.

Побудова моделі об'єкта, явища починається з виділення його найбільш істотних рис і властивостей та опису їх за допомогою математичних співвідношень. Потім, після створення математичної моделі, її досліджують математичними методами, тобто вирішують сформульоване математичне завдання.

## 1.1 Приклади математичних моделей

Як приклад розглянемо задачу визначення площі поверхні столу. Моделлю цієї поверхні, здавалося б, може бути прямокутник зі сторонами, такими, що дорівнюють сторонам столу. Якщо ж довжини протилежних сторін столу та його діагоналей виявляться не однаковими, як модель потрібно взяти чотирикутник. Для більш точного визначення площі столу необхідно врахувати ще заокруглення його кутів. Отже, з підвищенням вимог до точності визначення площі столу його математичні моделі постійно уточнюються.

Математична модель не визначається однозначно досліджуванним об'єктом. Вибір конкретної моделі визначається вимогами до її точності. Побудова математичної моделі одна з найскладніших і найвідповідальніших етапів дослідження об'єкта. Математична модель ніколи не буває тотожна об'єкту, що розглядається, не передає всіх його властивостей, оскільки ґрунтується на спрощенні та ідеалізації об'єкта. Тому результати, одержувані з урахуванням цієї моделі, мають завжди наближений характер. Їхня точність визначається ступенем відповідності, адекватності моделі та об'єкта. Питання точності є найважливішим у прикладній математиці. Проте воно не є суто математичним питанням і не може бути вирішене математичними методами. Основним критерієм істини є експеримент, тобто зіставлення результатів, одержуваних на основі математичної моделі, з об'єктом, що розглядається. Лише практика дозволяє порівняти різні гіпотетичні моделі й вибрати з них найбільш просту й достовірну, зазначити галузі застосування різних моделей і напрямків їх удосконалення.

Однією з типових задач обчислювального експерименту – проблема лазерного термоядерного синтезу – невичерпного джерела одержання енергії. Взяти під контроль процес термоядерного синтезу і навчитися ним керувати – означає розв'язати енергетичну проблему.

Розглянемо розвиток моделі на прикладі відомої задачі балістики визначення траєкторії тіла, випущеного з початковою швидкістю  $v_0$  під кутом  $\alpha_0$  до горизонту. Для початку припустимо, що швидкість  $v$  і дальність польоту тіла невеликі. Тоді для цього завдання буде справедлива математична модель Галілея, заснована на таких припущеннях:

1. Земля – інерційна система.
2. Прискорення вільного падіння  $g$  постійне.
3. Земля – плоске тіло.
4. Опір повітря відсутній.

Проте, як свідчить практика, результати, одержані з урахуванням цієї моделі, виявляються справедливими лише на малих початкових швидкостях руху тіла  $v_0 < 30 \text{ m/s}$ .

Подальше уточнення моделі балістичного завдання щодо обліку опору повітря було зроблено Ньютоном. Це дозволило з достатньою точністю розраховувати траєкторії руху гарматних ядер, що вистрелюються зі значними початковими швидкостями.

Перехід від гладкоствольної до нарізної зброї дозволив збільшити швидкість, дальність та висоту польоту снарядів, що зажадало подальшого уточнення математичної моделі задачі. У новій математичній моделі було переглянуто всі припущення, взяті моделі Галілея, тобто Земля вже не вважалася плоскою та інерційною системою, і сила земного тяжіння не бралася постійною.

Наступне вдосконалення математичної моделі завдання пов'язане з використанням методів теорії ймовірності. Це було викликано тим, що параметри снарядів, зарядів і навколишнього середовища через пуски оброблення деталей та інших причин не залишаються незмінними, а підпорядковуються випадковим коливанням.

Унаслідок послідовних уточнень і вдосконалень була створена математична модель, що найбільш повно й точно описує задачу зовнішньої балістики. Зіставлення її даних із результатами стрільб показало добрий їх збіг.

На цьому прикладі показані етапи створення, розвитку та уточнення математичної моделі об'єкта, які супроводжуються постійно зіставленням і перевіреною практикою, тобто з самим реальним об'єктом або явищем. Саме недостатньо хороший збіг результатів, що надаються моделлю, з об'єктом викликає подальше вдосконалення моделі.

Нарешті зазначимо, що вибір конкретної математичної моделі об'єкта для його аналізу необхідно робити з умови забезпечення



достатньої точності одержуваних результатів і простоти моделі. Водночас завжди потрібно пам'ятати, що не можна використовувати дуже точну й складну модель об'єкта, коли потрібна невелика точність результатів.

## 1.2 Етапи розв'язання практичних задач на комп'ютері

У зв'язку зі швидким розвитком обчислювальної техніки обчислювальний експеримент набув широкого застосування для проведення наукових досліджень та інженерного проектування. Він ґрунтується на побудові та аналізі за допомогою комп'ютера математичних моделей досліджуваного об'єкта чи явища.



Рисунок 1.1– Етапи розв'язання практичних задач на комп'ютері

*Об'єкт, явище або процес дослідження.* Нехай потрібно дослідити певний об'єкт, явище або процес (рис. 1.1). Спочатку формулюють основні закони та взаємозв'язки, що описують цей об'єкт, реалізується вибір загального підходу до розв'язування, визначаються

основні критерії, які повинна задовольняти розроблена система і наводиться формальний математичний опис задачі. На цьому етапі необхідне глибоке розуміння задачі.

*Математична модель.* На основі поставленої задачі розробляють математичну модель, що становить собою зазвичай запис цих законів у вигляді системи рівнянь (алгебраїчних, диференціальних, інтегральних та ін.). Після того, як задачу сформульовано, її потрібно розв'язати. Лише в досить простих випадках можливо одержати розв'язок в явному вигляді. Варто зазначити, що якщо математична модель вибрана недостатньо коректно, то які б методи не застосовувалися для розрахунків з її використанням, одержані висновки будуть ненадійні, або й зовсім неправильні.

*Чисельний метод.* Після побудови математичної моделі 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. Здебільшого виникає необхідність використання того чи іншого наближеного методу: обчислювального методу або дискретної моделі. На основі одержаної дискретної моделі будують обчислювальний алгоритм, результатом реалізації якого є число або таблиця чисел.

*Програмування.* На цьому етапі складають програму для реалізації розробленого алгоритму або використовують стандартні пакети прикладних програм, що призначені для розв'язування певного класу математичних задач, далі виконують тестування програми з метою виявлення помилок. Для тестової задачі підбирають такі вхідні дані, які дозволять спостерігати достовірність розвитку.

*Розрахунки.* Після розроблення та відладання програми настає етап проведення обчислень.

*Аналіз одержаних результатів.* Одержані результати детально аналізують з точки зору їх відповідності досліджуваному явищу і за необхідності вносять зміни в математичну модель або обирають інший обчислювальний метод. Цей цикл повторюють доти, поки не буде одержано результати з заданою точністю. Обчислювальна математика забезпечує лише один з етапів обчислювального експери-

менту, а саме етап вибору (побудови) обчислювального методу, від якого значною мірою залежить ефективність усього експерименту.

Отже, комп'ютер сам задачу не розв'язує, а лише виконує наперед задану послідовність обчислень; використання комп'ютера не звільняє дослідника від всебічного та скрупульозного осмислення своєї роботи та глибокого вивчення ОД.

### 1.3 Методи реалізації математичних моделей

Методи реалізації математичних моделей можна поділити на три групи: 1) графічні; 2) аналітичні; 3) чисельні. Зазначені методи використовуються як самостійно, так і спільно.

*Графічні методи* дозволяють оцінювати порядок шуканих величин і напрямок розрахункових алгоритмів.

*Аналітичні методи* (точні, наближені) спрощують фрагментарні розрахунки й дозволяють успішно вирішувати завдання оцінювання коректності та точності чисельних рішень.

Основним інструментом реалізації математичних моделей є *чисельні методи*. Чисельні методи, що дозволяють звести розв'язування задачі до виконання скінченного числа арифметичних і логічних дій з числами. Водночас розв'язок визначається як набір чисел, які надалі можуть бути інтерпретовані різним способом (наприклад, подані у вигляді таблиць, графіків, анімації тощо). Їх перевагами є: абсолютна універсальність, бо теоретично можуть бути застосовані для розв'язання будь-яких задач; добре пристосовані для реалізації на комп'ютері. Недоліком є велика трудомісткість у ході ручного рахунку, що зазвичай не є проблемою, оскільки вони призначені для використання на комп'ютері. Отже, чисельні методи є основним апаратом розв'язання математичних задач, а їх значущість лише збільшуватиметься через удосконалення комп'ютерної техніки. Чисельні методи бувають двох типів: прямі та ітераційні.

Метод називається *прямим*, якщо він дозволяє одержувати розв'язок після виконання скінченного числа елементарних операцій. Інколи прямі методи називають точними, маючи на увазі, що за відсутності похибок у вхідних даних і під час виконання елементарних операцій результат буде точним.

Далі розглядатимемо *ітераційні методи*, суть яких полягає в побудові послідовних наближень до розв'язку задачі. Спочатку ви-

бирають одне чи декілька початкових наближень, а потім послідовно, використовуючи знайдені раніше наближення й однотипну процедуру розрахунку, будують нові наближення. Внаслідок такого ітераційного процесу теоретично можна побудувати нескінченну послідовність наближень. Якщо ця послідовність збігається (що не гарантовано), то це ітераційний метод збіжний. Окремий крок ітераційного процесу називається ітерацією.

Проте під час комп'ютерної реалізації будь-якого методу неминучі похибки заокруглення і, як наслідок, наявність обчислювальної похибки.

Відхилення наближеного розв'язку від дійсного називається *похибкою*.

## 1.4 Характеристики чисельних методів

Для оцінювання чисельних методів, тобто порівняння між собою методів для розв'язання однієї задачі, вводять такі їх основні характеристики:

- трудомісткість;
- порядок методу;
- збіжність;
- швидкість збіжності;
- стійкість до погрешностей обчислень;
- стійкість до погрешностей у відправних даних.

Під *трудомісткістю методу* розуміють кількість та якість обчислень, необхідних для досягнення достатньо близького наближення розв'язку задачі. Під *порядком методу* розуміють вимоги до знань про функції, що входять у математичне формулювання задачі (наприклад, використання в методі похідних цих функцій):

- метод нульового порядку, якщо він використовує лише значення цих функцій;
- метод першого порядку, якщо він використовує значення функцій і їх перших похідних;

- метод другого порядку, якщо він використовує значення функцій та їх перших і других похідних тощо.

Чисельний метод називається таким, що *збігається*, якщо наближення  $x^k$  прямує до розв'язку  $x^*$  зі збільшенням  $k$ .

Очевидно, що методи, які не збігаються, не цікаві з прикладної точки зору. Тому одним із найважливіших етапів під час введення нового чисельного методу є теоретичне доведення його збіжності, тобто формулювання умов, за яких метод гарантовано збігається. В основному розрізняють три швидкості збіжності методів.

### 1 Лінійна збіжність.

Послідовність  $x^{(k)}$ ,  $k = 0, 1, 2, \dots$  лінійно збігається до розв'язку  $x^*$  (або зі швидкістю геометричної прогресії), якщо існують числа  $q \in (0; 1)$ ,  $k_0 > 0$  такі, що

$$\|x^{(k+1)} - x^*\| \leq q \|x^{(k)} - x^*\|, k \geq k_0. \quad (1.1)$$

Тут норма  $x - y$  означає відстань між  $x$  і  $y$ .

### 2 Надлінійна збіжність.

Послідовність  $x^{(k)}$ ,  $k = 0, 1, 2, \dots$  надлінійно збігається до розв'язку  $x^*$ , якщо існує послідовність  $q_k$ ,  $k = 0, 1, 2, \dots$ ,  $q_k \in (0; 1)$  така, що

$$\|x^{(k+1)} - x^*\| \leq q_k \|x^{(k)} - x^*\|, q_k \rightarrow 0, k \rightarrow \infty. \quad (1.2)$$

### 3 Квадратична збіжність.

Послідовність  $x^{(k)}$ ,  $k = 0, 1, 2, \dots$  квадратично збігається до розв'язку  $x^*$ , якщо існують числа  $C > 0$ ,  $k_0 > 0$  такі, що

$$\|x^{(k+1)} - x^*\| \leq C \|x^{(k)} - x^*\|^2, k \geq k_0. \quad (1.3)$$

Під *стійкістю до погрешностей обчислень* розуміють те, що застосування чисельного методу приводить до розв'язку задачі на комп'ютері, незважаючи на помилки округлень і обчислень. Для цього в чисельних методах, якщо потрібно, передбачаються додаткові операції, що не змінюють суть методу, але забезпечують його стійкість до помилок обчислень.

Під *стійкістю до погрешностей у відправних даних* розуміють те, що за невеликих погрешностей у відправних даних застосування чисельного методу дозволяє одержати наближений розв'язок задачі з не дуже великою погрешністю. Стійкість до погрешностей у відправних даних досягається зазвичай шляхом модифікації чисельного методу, тобто внесенням змін до суті методу.

## Лекція 2

### Елементи теорії похибок

Розв'язки задач, одержаних за допомогою технології комп'ютерного експерименту, завжди будуть мати похибку. Чинники появи похибок можуть бути такими:

- 1 Побудова математичної моделі реальної задачі. Похибка може з'явитися відповідно до степені коректності та обумовленості математичної моделі задачі.
- 2 Одержання вихідних даних. Похибка може з'явитися внаслідок вимірювання чи округлення вихідних даних.
- 3 Розроблення обчислювального алгоритму. Похибку може обумовлювати степінь збіжності наближеного методу розв'язання задачі до точного.
- 4 Похибка обчислювальної техніки, на якій моделюється комп'ютерний експеримент. Сюди належать похибки обмеженості розрядної сітки комп'ютера і похибки машинної арифметики.

Похибка, що з'являється через чинники 1 і 2 наведеного переліку носить назву неусувної похибки комп'ютерного експерименту. Чинник 3 дає так звану похибку методу і чинник 4 – похибку наближених обчислень.

**Похибка математичної моделі** виникає через прагнення забезпечити порівняльну простоту її технічної реалізації та доступності дослідження. Потрібно мати на увазі, що конкретна математична модель (ММ), що чудово працює в одних умовах, може бути абсолютно непридатною в інших. З погляду споживача, важливим є правильне оцінювання галузі її (ММ) застосування.

**Похибка чисельного методу (похибка апроксимації)**, пов'язана, наприклад, із заміною інтеграла сумою, з усіченням рядів під час обчислення функцій, з інтерполюванням табличних значень функціональних залежностей тощо. Зазвичай похибка чисельного методу регульована і може бути зменшена до будь-якого розумного значення шляхом зміни певного параметра.

**Обчислювальна похибка** виникає через округлення чисел, проміжних та остаточних результатів рахунку. Вона залежить від правил і необхідності округлення, а також від алгоритмів чисельного рішення.

## 2.1 Абсолютна та відносна похибки

Розглянемо більш детально види похибок і рекомендації для їх усунення або зменшення.

**Означення 2.1** *Наближеним числом називається таке число, яке відрізняється від точного числа на незначну величину і заміняє його в процесі обчислень.*

Працюючи з наближеними величинами обчислювач повинен вміти:

- а) давати математичні характеристики точності наближених величин;
- б) знаючи степінь точності вихідних даних, оцінити степінь точності результатів;
- в) брати вихідні дані з таким ступенем точності, щоб забезпечити задану точність результату. У цьому разі не потрібно надмірно завищувати точність вихідних даних, щоб позбавити обчислювача від непотрібних розрахунків;
- г) вміти правильно побудувати обчислювальний процес, щоб позбавити його від тих обчислень, які не вплинуть на точні цифри результату.

Основними характеристиками точності обчислень є абсолютна і відносна похибки наближених величин.

**Означення 2.2** *Абсолютною похибкою наближеної величини  $\hat{x}$  називається модуль її відхилення від відповідної точної величини  $x$ , тобто*

$$\Delta_x = |\hat{x} - x|. \quad (2.1)$$



Абсолютна похибка  $\Delta_x$  є величиною вимірною і не підходить для порівняння похибок різних величин, які є вхідними, або такими, що обчислюються в процесі комп'ютерного експерименту.

---

## ПРИКЛАД 1

---

Якщо для плаского об'єкта, який моделюється, ввести площу з точністю до  $\text{см}^2$  (тобто ввести  $\hat{S} = 999 \text{ см}^2$  замість  $S = 1\,000 \text{ см}^2$ ), одержимо абсолютну похибку  $\Delta_S = 1 \text{ см}^2$ , а товщину ввести з точністю до  $10^{-3} \text{ см}$  (тобто замість  $l = 0.011 \text{ см}$  ввести  $\hat{l} = 0.01 \text{ см}$ , то одержимо абсолютну похибку  $\Delta_l = 0.001 \text{ см}$ . Виглядає так, що площа об'єкта була введена з більшою похибкою, ніж його товщина. Хоча насправді це не так. ■

Для порівняння відхилень різних наближених числових даних моделі від їх точного значення застосовують відносну похибку наближеної величини.

**Означення 2.3** *Відносною похибкою  $\delta_x$  наближеної величини  $\hat{x}$  називається відношення абсолютної похибки  $\Delta_x$  наближеної величини до модуля точного значення величини  $x$*

$$\delta_x = \frac{|\hat{x} - x|}{|x|} = \frac{\Delta_x}{|x|}. \quad (2.2)$$

Відносна похибка відображає долю викривленої інформації у наближеній величині щодо істинного значення цієї величини.

Оскільки часто дослідників цікавить відсоток похибки внаслідок комп'ютерного експерименту, то часто результат (2.3) множать на 100 %, тобто розглядають відносну похибку за такою формулою

$$\delta_x \% = \frac{\Delta_x}{|x|} \cdot 100 \%. \quad (2.3)$$

Якщо повернутися до попереднього приклада і порівняти відносні похибки площі та товщини пласкої фігури, то одержимо, що

$$\delta_S = \frac{1}{1\,000} = 0.001\delta; \quad \delta_l = \frac{0.001}{0.011} = \frac{1}{11}0.0(9) \approx 0.1$$

або у відсотковому відношенні

$$\delta_S = 0.1\%; \quad \delta_l = 10\%.$$

Отже, площу об'єкта було введено із значно меншим викривленням, ніж товщину.

Абсолютна і відносна похибки легко визначаються в разі відомого точного значення величини  $x$ . На практиці дуже часто такого знання в дослідника немає. В такому разі визначається максимальне допустиме відхилення  $\Delta_x^*$  від точного значення для цієї задачі і розглядається діапазон значень наближеної величини  $\hat{x}$ :

$$|x - x^*| \leq \Delta_x^* \quad (2.4)$$

або

$$\hat{x} - \Delta_x^* \leq x \leq \hat{x} + \Delta_x^*.$$

За такого визначення величина  $\Delta_x^*$  називається **гранично допустимою абсолютною похибкою** величини  $x$  і вважається, що істинне значення величини  $x$  лежить у діапазоні  $[\hat{x} - \Delta_x^*; \hat{x} + \Delta_x^*]$ .

Цей факт відображають записом  $x = x \pm \Delta_x^*$  і кажуть, що величина  $x$  задана з точністю до  $\Delta_x^*$ .

У відносній похибці за таких умов заміняється  $|x|$  на  $|\hat{x}|$ , відносна похибка заміняється на **гранично допустиму відносну похибку** наближеної величини і формула (2.3), набирає вигляду

$$\delta_x^* = \frac{\Delta_x^*}{|\hat{x}|}. \quad (2.5)$$

та відповідно

$$\delta_x \% = \frac{\Delta_x^*}{|\hat{x}|} \cdot 100 \%.$$

**Зауваження!** Загалом за гранично допустимі абсолютну і відносні похибки можуть бути взяті будь-які величини, що перевищують відомі похибки. Але за можливості потрібно визначати найменші з таких.

Наприклад, якщо для одержання вихідних даних для математичної моделі проводиться серія експериментів, унаслідок якої дослідник одержує ряд абсолютних похибок для величини, що досліджується, то за граничне значення краще за все взяти найбільшу

абсолютну похибку з цього ряду. Або якщо передчасно відомо максимально допустиме відхилення від точного значення величини, то саме його і необхідно брати за граничне значення абсолютної похибки.



## ПРИКЛАД 2

Після серії вимірювань довжин двох об'єктів одержали, що довжина об'єкта  $A$  дорівнює  $l = 43.3 \pm 0.1$  см, а довжина об'єкта  $B : L = 934.5 \pm 0.1$  см. Знайти гранично допустимі відносні похибки вимірювання об'єктів і оцінити точність одержаних даних у відсотках.

### Розв'язування

Використаємо формулу (2.5) для обчислення граничної відносної похибки у відсотках:

$$\delta_l^* = \frac{0.1}{43.3} \cdot 100 \% \approx 0.23\%; \quad \delta_L^* = \frac{0.1}{934.5} \cdot 100 \% \approx 0.01 \%$$

Одержали, що гранично допустима відносна похибка вимірювання об'єкта  $A$  становить 0.23 %, а об'єкта  $B$  – 0.01 % для однакової гранично допустимої абсолютної похибки. Відповідно точність задання даних по об'єкту  $A$  не менше ніж  $100 \% - 0.23 \% = 99.77 \%$ , а по об'єкту  $B$  –  $100 \% - 0.01 \% = 99.99 \%$ . ■



## ПРИКЛАД 3

Визначити граничну абсолютну похибку наближеного числа  $a = 12.937$ , якщо відомо, що граничне значення відносної похибки цього числа становить 0,01%.

### Розв'язування

$$0.01 \% = \frac{\Delta_a^*}{|12.937| \cdot 100 \%}$$

Звідси

$$\Delta_a^* = \frac{\delta_a^* \% \cdot |\hat{a}|}{100 \%} = \frac{12.937 \cdot 0.01}{100} = 0.001 = 10^{-3}. \quad \blacksquare$$



З'ясувати, яка рівність точніша  $x = \frac{13}{19} \approx 0,684$  або  $y = \sqrt{52} \approx 7.21$ .

**Розв'язування**

Обчислюємо обидва вирази з більшою кількістю знаків (щонайменше на 2 знаки більше)

$$x = \frac{13}{19} \approx 0.68421; \quad y = \sqrt{52} \approx 7.2111.$$

Обчислюємо граничну абсолютну похибку, округлюючи одержані результати з надлишком.

$$\Delta_x^* = |0.68421 - 0.684| = 0.00022; \quad \Delta_y^* = |7.2111 - 7.21| = 0.0012.$$

Обчислюємо граничні відносні похибки

$$\delta_x^* = \frac{0.00022}{0.684} \cdot 100 \% = 0.033 \%;$$

$$\delta_y^* = \frac{0.0012}{7.21} \cdot 100 \% = 0.017 \%.$$

Одержали  $\delta_y^* < \delta_x^*$ , отже, вираз  $y = \sqrt{52} \approx 7.21$  точніший. ■

У будь-якій позиційній системі числення з основою  $p$  довільне дійсне число  $a$  можна подати як лінійну комбінацію цифр, які відповідають цій основі, помножених на степені основи  $p^k$ ,  $k = 1, \dots, n$ , тобто

$$a = a_1 p^n + a_2 p^{n-1} + \dots + a_k p^{n-k+1}, \quad (2.6)$$

де  $a_i$  – цифри системи числення за основою  $p$ , множник  $p^j$  біля цифри  $a_i$  – позиція цієї цифри у числі;  $a_1 \neq 0$ ,  $p^n$  – старший розряд числа за цією основою, порядок числа.

Наприклад, у десятковій системі числення числа

$$a = 21.003 = 2 \cdot 10^1 + 1 \cdot 10^0 + 0 \cdot 10^{-1} + 0 \cdot 10^{-2} + 3 \cdot 10^{-3};$$

$$b = 0.00023 = 2 \cdot 10^{-4} + 3 \cdot 10^{-5}$$

мають відповідно порядок 1 і  $-4$ .

Для стандартизації запису чисел застосовують нормалізовану та стандартизовану форму запису числа з рухомою комою в системі числення з основою  $p$ :

$a = a_1, a_2, \dots, a_k \cdot p^n$  – стандартизований запис числа з рухомою комою;

$a = 0, a_1, a_2, \dots, a_k \cdot p^{n+1}$  – нормалізований запис числа з рухомою комою.

Наприклад, для десяткової системи числення

$$0.00023 = 2.3 \cdot 10^{-4} = 0.23 \cdot 10^{-3}; \quad 1234000 = 1.234 \cdot 10^6 = 0.1234 \cdot 10^7.$$

## 2.2 Значущі цифри числа. Правильні та сумнівні цифри числа

**Означення 2.4** *Всі цифри стандартизованої форми числа, враховуючи праві нулі, якщо вони характеризують розряд, називаються **значущими цифрами** числа (мантисою).*

У наведеному прикладі значущі цифри виділені:

$$\mathbf{123.0420} = 1.230420 \cdot 10^2; \mathbf{0.304057} = 3.04057 \cdot 10^{-1};$$

$$\mathbf{0.00230} = 2.30 \cdot 10^{-3}.$$

Для аналізу розрахунків на похибку більш істотним є не питання значущих цифр у наближеному числі, а питання правильної чи сумнівної цифри наближеного числа.

**Означення 2.5** *Для наближеного числа*

$$\hat{x} = a_1 \cdot 10^n + a_2 \cdot 10^{n-1} + \dots + a_k \cdot 10^{n-k+1}$$

*з  $k$  значущими цифрами цифра  $a_k$  буде **правильною**, якщо абсолютна похибка числа  $\hat{x}$  відповідає умові*

$$\Delta_x \leq \omega \cdot 10^{n-k+1}, \quad 0.5 \leq \omega \leq 1. \quad (2.7)$$

У разі порушення умови цифру  $a_k$  називають **сумнівною**.

Якщо  $\omega = 0.5$ , то це свідчить, що  $a_k$  є **правильною у вузькому сенсі**. В такому разі абсолютна похибка числа не перевищує половину одиниці розряду  $10^{n-k+1}$ .

Якщо  $\omega = 1$ , то це свідчить, що  $a_k$  є **правильною у широкому сенсі**. В такому разі абсолютна похибка числа не перевищує одиниці розряду  $10^{n-k+1}$ .

Наприклад, для точного числа  $x = 7.138$  в наближеному числі  $\hat{x} = 7.14$  права цифра  $a_3 = 4$  є **правильною у вузькому сенсі**, оскільки

$$\Delta_x = |7.14 - 7.138| = 0.002 < 0.005 = 0.5 \cdot 10^{-2} = 0.5 \cdot 10^{0-3+1}.$$

А в наближеному  $\hat{x} = 7.13$ ,  $a_3 = 3$  є сумнівною у вузькому сенсі, оскільки

$$\Delta_x = |7.13 - 7.138| = 0.008 > 0.005.$$

Але ця цифра є **правильною в широкому сенсі**, оскільки

$$\Delta_x = |7.13 - 7.138| = 0.008 < 0.01 = 1 \cdot 10^{-2}.$$

Якщо наближені числа з'являються внаслідок обчислень із точними вихідними даними, має сенс розглядати наближення у вузькому сенсі.

Якщо вихідні дані вже є наближеними величинами, то краще брати точність у широкому сенсі, тобто у (2.7) брати  $\omega = 1$ .

## ПРИКЛАД 5

- Скільки правильних знаків має наближене число  $a = 2.324 \pm 0.0076$  у вузькому сенсі, у широкому сенсі?
- Визначити граничну абсолютну похибку чисел  $a = 96.384$ ;  $b = 9.18$  якщо вони мають усі правильні цифри у вузькому та широкому сенсі відповідно.

### Розв'язування

а)  $0.0076 > 0.005 = 0.5 \cdot 10^{-2=0-3+1}$ , отже, 3-тя цифра  $a_3 = 2$  у вузькому сенсі є сумнівною, правильних знаків є 2:  $a_1 = 2$ ,  $a_2 = 3$ .

Водночас  $0.0076 < 0.01 = 1 \cdot 10^{-2=0-3+1}$ , тобто в широкому сенсі цифра  $a_3 = 2$  є **правильною** і правильних цифр у числі 3.

б) У вузькому сенсі:  $a = 96.384 \pm (0.0005 = 0.5 \cdot 10^{1-5+1=-3})$ ; у широкому сенсі:  $b = 9.18 \pm (0.01 = 1 \cdot 10^{0-3+1=-2})$ . ■

## 2.3 Правила округлення

У процесі обчислень не потрібно забувати правил округлення чисел, бо округлення теж додає помилки.

- 1 Якщо перша ліва цифра з тих, що відкидаються, більша ніж 5, то остання із збережених підсилюється (додається 1)  
( $5.348 \approx 5.35$ ).
- 2 Якщо перша ліва цифра з тих, що відкидаються, менша ніж 5, то остання із збережених залишається без змін ( $5.343 \approx 5.34$ ).
- 3 Якщо перша ліва цифра з тих, що відкидаються, дорівнює 5, то остання із збережених залишається без змін, **якщо вона парна**, і підсилюється, **якщо вона непарна**

$$(5.345 \approx 5.34; \quad 5.375 \approx 5.38).$$

Отже, застосування наведених правил округлення дає абсолютну похибку, яка не перевищує половини одиниці розряду, який визначає остання залишена цифра.



### ПРИКЛАД 6

Округлити сумнівні цифри числа  $a = 26.7255 \pm 0.0026$ , залишивши правильні знаки у широкому сенсі. Знайти відносну точність вихідного і округленого чисел.

#### Розв'язування

$\Delta_a = 0.0026 < 0.01 = 1 \cdot 10^{-2=1-k+1}$ ;  $k = 4 \Rightarrow 2, 6, 7, 2$  – правильні числа в широкому сенсі. Округлюємо за правилом 3:  $a = 26.7255 \approx 26.72$ .

$$\Delta_{r-a} = |26.7255 - 26.72| = 0.0055;$$

$$\Delta_{c-a} = \Delta_a + \Delta_{r-a} = 0.0026 + 0.0050 < 0.01.$$

Тобто після округлення кількість правильних цифр у широкому сенсі не змінилася.

$$\hat{a} = 26.72 \pm 0.0081;$$

$$\delta_a = \frac{0.0026}{26.7255} = 9.728 \cdot 10^{-5} \approx 0.0001 = 0.01 \%$$

$$\delta_{\hat{a}} = \frac{0.0081}{26.72} = 3 \cdot 10^{-4} \approx 0.0003 = 0.03 \%.$$

Як бачимо, порядок відносної похибки не змінився. ■

## ПРИКЛАД 7

Округлити сумнівні цифри числа  $a = 35.8436$ , залишивши правильні знаки у вузькому сенсі. Визначити абсолютну похибку числа, якщо відносна похибка дорівнює  $\delta_a = 0.03\%$ .

### Розв'язування

Для визначення правильних цифр потрібна абсолютна похибка. Знайдемо її, використавши (2.5):

$$0.0003 = \frac{\Delta_a}{35.8436} \Rightarrow \Delta_a = 35.8436 \cdot 0.0003 = 0.01075308 \approx 0.010754$$

$$\Delta_a^* = 0.010754 < 0.05 = 0.5 \cdot 10^{-1=1-k+1}; k = 3; \hat{a} = 35.8;$$

$$\Delta_{r-a} = |35.8 - 35.8436| = 0.0436;$$

$$\Delta_{c-a} = \Delta_a + \Delta_{r-a} = 0.010754 + 0.0436 = 0.054354 \approx 0.05.$$

Одержали, що спільна помилка завдання числа + округлення числа дещо перебільшила 0.5, отже, точність у вузькому сенсі порушена, але похибка потрапила в діапазон  $(0.05; 0.1)$ , тобто округлене число має правильні числа. ■

Розглянемо зв'язок між кількістю правильних цифр числа та відносною похибкою числа. Число візьмемо з (2.6):

- 1 Вважаємо, що число має  $k$  правильних цифр. Правильні цифри числа пов'язані з абсолютною похибкою за (2.7), тобто за граничну абсолютну похибку беремо  $\Delta_a^* = \omega \cdot 10^{n-k+1}$ . Тоді згідно (2.5)

$$\begin{aligned} \delta_a &= \frac{\omega \cdot 10^{n-k+1}}{a_1 \cdot 10^n + a_2 \cdot 10^{n-1} + \dots + a_k \cdot 10^{n-k+1}} \leq \\ &\leq \frac{\omega \cdot 10^{n-k+1}}{a_1 \cdot 10^n} = \frac{\omega}{a_1 \cdot 10^{k-1}} = \delta_a^*. \end{aligned}$$

Одержали, що за відомих  $k$  правильних цифр числа гранична відносна похибка дорівнює

$$\delta_a^* = \frac{\omega}{a_1 \cdot 10^{k-1}}, \quad (2.8)$$



де  $a_1$  – перша ліва значуща цифра числа;  $0.5 \leq \omega \leq 1$ ,  $k$  – кількість правильних цифр числа.

**Зауваження!** Відносну похибку можна одержати і за формулою (2.5), але наведений спосіб простіший.

2 У разі, коли відносна гранична похибка числа відома, зв'язок одержують з (4.4) так:

$$\delta_a^* = \frac{\omega}{a_1 \cdot 10^{k-1}} \Rightarrow \delta_a^* = \frac{10\omega}{a_1 \cdot 10^k} \Rightarrow 10^k = \frac{10\omega}{a_1 \cdot \delta_a^*}.$$

Щоб знайти  $k$ , логарифмується вираз за основою 10:

$$\hat{k} = \lg(10\omega) - \lg(a_1 \delta_a^*) = 1 + \lg\left(\frac{\omega}{a_1 \delta_a^*}\right).$$

Одержали

$$\hat{k} = 1 + \lg\left(\frac{\omega}{a_1 \delta_a^*}\right) \quad (2.9)$$

і береться найближче до  $\hat{k}$  ціле число.

У разі пошуку кількості правильних цифр у вузькому сенсі, формула набере вигляду

$$\hat{k} = 1 - \lg(2a_1 \delta_a^*). \quad (2.10)$$

Для широкого сенсу

$$\hat{k} = 1 - \lg(a_1 \delta_a^*) \quad (2.11)$$

і береться найближче до  $\hat{k}$  ціле число.



## ПРИКЛАД 8

Яку граничну відносну похибку в широкому сенсі має число 435.849, якщо воно має лише правильні цифри.

**Розв'язування**

Використаємо (4.4).  $\omega = 1$ ;  $a_1 = 4$ ;  $k = 6$ .

$$\delta_a^* = \frac{\omega}{a_1 \cdot 10^{k-1}} = \frac{1}{4 \cdot 10^6 - 1} = 0.25 \cdot 10^{-5} = 0.25 \cdot 10^{-3}\%. \quad \blacksquare$$

---

 ПРИКЛАД 9

---

Зі скількома правильними цифрами у вузькому сенсі треба взяти  $a = \sqrt{\pi} \approx \sqrt{3.1415927} \approx 1.77245385$ , щоб відносна похибка не перевищувала 0.1 %.

**Розв'язування**

Візьмемо першу цифру для числа  $a$ :  $a = \sqrt{\pi} \approx \sqrt{3.1415927} \approx 1$ , відносна похибка  $\delta_a = 0.001$ ;  $\omega = 0.5$ . Скористаємося (2.10)

$$\hat{k} = 1 - \lg(2 \cdot 1 \cdot 0.001) = 1 - (\lg 2 - 3) \approx 4 - 0.301 \approx 3.699 \approx 3.7.$$

Найближче ціле число  $k = 4$ . Отже,  $a = \sqrt{\pi} \approx 1.772$  – округлення виконане за правилами.  $\Delta_r = 0.00045385 < 0.5 \cdot 10^{-3=0-4+1}$ , тоді можна записати, що у вузькому сенсі

$$a = \sqrt{\pi} = 1.772 \pm 5 \cdot 10^{-4}. \quad \blacksquare$$

## Лекція 3

### Пряма та обернена задачі теорії похибок

Працюючи з наближеними числами, важливим завданням є оцінювання ступеня впливу похибок вихідних даних на точність остаточного результату. Це необхідно не лише для правильного обліку обчислювальних похибок, й для визначення можливих шляхів їх зменшення.

Зокрема, під час обчислення значень функцій, аргументами яких є наближені числа, виникає питання похибки значень, що обчислюються. Визначення величини похибки результату відомих похибок вихідних даних становить завдання теорії похибок.

#### 3.1 Пряма задача теорії похибок

**Прямою задачею** теорії похибок є задача визначення похибки обчислення функції за відомими похибками аргументів.

Нехай в області  $G$  задана диференційована функція  $y = f(x_1, x_2, \dots, x_n)$  і нехай відомі  $\Delta_{x_i}, i = \overline{1, n}$  – абсолютні похибки аргументів. Тоді абсолютну похибку функції розраховують за формулою Лагранжа

$$\Delta_y = |y - \hat{y}| \leq \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i} \Delta_{x_i} \right|. \quad (3.1)$$

Для функції однієї змінної

$$\Delta_y = |f'(\hat{x})| \Delta_x. \quad (3.2)$$

**Означення 3.1** *Граничною абсолютною похибкою функції багатьох змінних називають оцінку*

$$\Delta_y^* = \sup_{(x_1, x_2, \dots, x_n) \in G} \Delta_y. \quad (3.3)$$

Нехай для функції, заданої вище, відомі **відносні похибки аргументів**  $\delta_{x_i}, i = 1, n$ , тоді

$$\delta_y = \sum_{i=1}^n \left| \frac{\partial}{\partial x_i} \Delta_{x_i} \right| = \sum_{i=1}^n \left| \hat{x}_i \frac{\partial}{\partial x_i} \ln \hat{y} \right| \delta_{x_i}. \quad (3.4)$$

Для функції однієї змінної

$$\delta_y = \left| \hat{x} \frac{f'(\hat{x})}{f(\hat{x})} \right| \delta_{x_i}. \quad (3.5)$$

**Означення 3.2** *Граничною відносною похибкою функції багатьох змінних називають оцінку*

$$\delta_y^* = \frac{\Delta_y^*}{|\hat{y}|}. \quad (3.6)$$

У таблиці 3.1 наведені оцінки для деяких функцій від однієї та двох змінних.

Таблиця 3.1– Оцінки для деяких функцій від однієї та двох змінних

Функція	Абсолютна похибка	Відносна похибка
$y = \hat{x}_1 + \hat{x}_2$	$\Delta_y = \Delta_{x_1} + \Delta_{x_2}$	$\delta_y = \frac{ \hat{x}_1  \cdot \delta_{x_1} +  \hat{x}_2  \cdot \delta_{x_2}}{ \hat{x}_1 + \hat{x}_2 }$
$y = \hat{x}_1 - \hat{x}_2$	$\Delta_y = \Delta_{x_1} + \Delta_{x_2}$	$\delta_y = \frac{ \hat{x}_1  \cdot \delta_{x_1} +  \hat{x}_2  \cdot \delta_{x_2}}{ \hat{x}_1 - \hat{x}_2 }$
$y = \hat{x}_1 \cdot \hat{x}_2$	$\Delta_y = \hat{x}_2 \Delta_{x_1} + \hat{x}_1 \Delta_{x_2}$	$\delta_y = \delta_{x_1} + \delta_{x_2}$
$y = \frac{\hat{x}_1}{\hat{x}_2}$	$\Delta_y = \frac{1}{\hat{x}_2} \Delta_{x_1} + \frac{\hat{x}_1}{\hat{x}_2^2} \Delta_{x_2}$	$\delta_y = \delta_{x_1} + \delta_{x_2}$
$\hat{x}^n$	$\Delta_y = n \hat{x}^{n-1} \Delta_x$	$\delta_y =  n  \delta_x$
$e^{\hat{x}}$	$\Delta_y = e^{\hat{x}} \Delta_x$	$\delta_y = \hat{x} \delta_x$
$\ln \hat{x}$	$\Delta_y = \frac{\Delta_x}{\hat{x}}$	$\delta_y = \frac{\delta_x}{ \ln \hat{x} }$

## ПРИКЛАД 1

Знайти абсолютну та відносну похибки розрахунку об'єма кругового конуса  $V = \frac{1}{3} \pi r^2 h$ , якщо  $r = 5.23 \pm 0.005$ ;  $h = 12.3 \pm 0.1$ ;  $\pi = 3.14 \pm 0.01$ .

### Розв'язування

Виконаємо аналіз на кількість правильних цифр у кожному з даних.  $r = 5.23 \pm 0.005$ . У вузькому сенсі:

$$\Delta_r^* = 0.005 = 0.5 \cdot 10^{-2} \Rightarrow -2 = 0 - k + 1 \Rightarrow k = 3.$$

Це  $r$  має усі 3 правильні цифри у вузькому сенсі, отже, і в широкому теж.

$$h = 12.3 \pm 0.1.$$

У вузькому сенсі:

$$\Delta_h^* = 0.1 < 0.5 \cdot 10^0 \Rightarrow 0 = 1 - k + 1 \Rightarrow k = 2.$$

Це  $h$  має 2 правильні цифри у вузькому сенсі.

У широкому сенсі:

$$\Delta_h^* = 0.1 = 1 \cdot 10^{-1} \Rightarrow -1 = 1 - k + 1 \Rightarrow k = 3.$$

Це  $h$  має усі 3 правильні цифри у широкому сенсі

$$\pi = 3.14 \pm 0.01.$$

У вузькому сенсі:

$$\Delta_\pi^* = 0.01 < 0.5 \cdot 10^{-1} \Rightarrow -1 = 0 - k + 1 \Rightarrow k = 2.$$

Це  $\pi$  має 2 правильні цифри у вузькому сенсі.

У широкому сенсі:

$$\Delta_\pi^* = 0.01 = 1 \cdot 10^{-2} \Rightarrow -2 = 0 - k + 1 \Rightarrow k = 3.$$

Це  $\pi$  має усі 3 правильні цифри в широкому сенсі.

Отже, має сенс розглядати абсолютну похибку результату в широкому сенсі.

**Перший шлях – через абсолютні похибки.**

Використаємо для обчислення абсолютної похибки формулу Лагранжа та обчислюємо частинні похідні:

$$\begin{aligned} \frac{\partial V}{\partial \pi} &= \frac{1}{3} r^2 h = \frac{1}{3} (5.23)^2 \cdot 12.3 = 112.147; \\ \frac{\partial V}{\partial r} &= \frac{2}{3} \pi r h = \frac{2}{3} \cdot 3.14 \cdot 5.23 \cdot 12.3 = 134.662; \\ \frac{\partial V}{\partial h} &= \frac{1}{3} \pi r^2 = \frac{1}{3} \cdot 3.14 \cdot (5.23)^2 = 28.629. \end{aligned}$$

$$\Delta_V = V'_\pi \Delta_\pi + V'_r \Delta_r + V'_h \Delta_h = 112.15 \cdot 0.01 + 134.66 \cdot 0.005 + 28.63 \cdot 0.1 \approx 4.658 \text{ см}^3.$$

Отже,  $\Delta_V^* = 5$ . Обчислимо безпосередньо об'єм конуса

$$V = \frac{1}{3} \cdot 3.14 \cdot (5.23)^2 \cdot 12.3 = 352.141 \Rightarrow V = 352.141 \pm 5.$$

З'ясуємо кількість правильних цифр результату в широкому сенсі

$$\Delta_V^* = 5 < 10 = 10^1 \Rightarrow 1 = 2 - k + 1 \Rightarrow k = 2.$$

Унаслідок обчислень є 2 правильні цифри 3 і 5. Якщо результат округлити до правильних цифр, одержимо

$$V = 350 \pm 5.$$

Водночас похибка округлення буде  $\Delta_V = |352.141 - 350| = 2.141$ ; повна похибка:  $\Delta_V = 4.658 + 2.141 = 6.799 < 10$ , тобто за граничну абсолютну похибку результату в широкому сенсі потрібно брати  $\Delta_V^* = 10$ .

Остаточо:  $V = 350 \pm 10$ . Відносна гранична похибка може бути розрахована як

$$\delta_V^* \frac{\Delta_V^*}{|V|} = \frac{10}{350} \approx 0.029$$

або 3 %.

Об'єм кругового конуса обрахований із точність 97 %.

**Другий шлях — через відносні похибки.**

Об'єм конуса знаходимо як

$$V = \frac{1}{3} \pi r^2 h.$$

Розрахунок будемо вести в широкому сенсі, де усі дані мають по 3 правильні цифри:

$$\delta_V^* = \delta_\pi^* + \delta_{r,2}^* + \delta_h^* \Rightarrow \Delta_V^* = \delta_V^* \cdot V;$$

$$\delta_\pi^* = \frac{1}{3 \cdot 10^2} = 0.003; \delta_{r,2}^* = 2\delta_r^* = 2 \frac{1}{5 \cdot 100} = 0.004;$$

$$\delta_h^* = \frac{1}{10^2} = 0.01 \Rightarrow \delta_V^* = 0.017.$$

Абсолютна похибка:

$$\Delta_V^* = \delta_V^* V; V = 352.141 \Rightarrow \Delta_V^* = 352.141 \cdot 0.017 = 5.98.$$

Ця абсолютна похибка знову, як і за першим шляхом, може бути апроксимована граничною абсолютною похибкою в широкому сенсі:

$$\Delta_V^* = 5.98 < 10 \Rightarrow k = 2 \Rightarrow V = 350 \Rightarrow \Delta_{r,V} = 2.141 \Rightarrow$$

$$\Rightarrow \Delta_V^* = 5.98 + 2.141 = 8.121 < 10.$$

Отже, за другим шляхом граничною абсолютною похибкою обчислення об'єму є 10, тобто

$$V = 350 \pm 10.$$

Оскільки  $\delta_V^* = 3\% > \delta_V^* = 2\%$ , то за остаточну граничну відносну похибку результату потрібно взяти 3 %. ■

## 3.2 Обернена задача теорії похибок

Обернена задача теорії похибок вирішує питання, які повинні бути похибки аргументів функції, щоб абсолютна похибка функції не перевищувала заданої величини. Це завдання математично не визначено, оскільки рішення можна забезпечити, по-різному встановлюючи похибки аргументів.

Обернена задача полягає в такому: за яких значень аргументу відома функція  $y = f(x_1, x_2, \dots, x_n)$  буде мати похибку, яка не перебільшує задану величину.

Найпростіший розв'язок оберненої задачі дається за принципом **рівних впливів**. Вважається, що усі частинні диференціали функції однаково впливають на створення загальної абсолютної похибки:

$$\Delta x_i = \frac{\Delta y^*}{n \frac{\partial f}{\partial x_i}}. \quad (3.7)$$

Розглянемо обернену задачу, використовуючи приклад 10.



### ПРИКЛАД 2

З якою точністю потрібно вимірювати вихідні дані для обчислення об'єму кругового конуса з заданою абсолютною граничною похибкою  $\Delta V^* = \pm 1$ . Для обчислення об'єму попередньо була визначена точка

$$(\pi, r, h) = (3.14; 5.23; 12.3).$$

#### Розв'язування

Застосуємо формулу 3.7 принципу однакових впливів похибок аргументів. Тобто вважаємо, що усі координати точки з простору визначення функції об'єму однаково впливають на формування похибки результату.

Попередньо потрібно обчислити частинні похідні за аргументами у заданій точці. З попереднього приклада маємо

$$\frac{\partial V}{\partial \pi} = 112.147; \quad \frac{\partial V}{\partial r} = 134.662; \quad \frac{\partial V}{\partial h} = 28.629.$$

Тоді

$$\Delta \pi = \frac{1}{3 \cdot 112.147} = 0.00297 \approx 0.003;$$
$$\Delta r = \frac{1}{3 \cdot 134.662} = 0.00248 \approx 0.002;$$

$$\Delta_h = \frac{1}{3 \cdot 28.629} = 0.0116 \approx 0.01.$$

Відносні похибки відповідно будуть

$$\delta_\pi = \frac{0.003}{3.14} \cdot 100 \% = 0.1 \% ; \delta_r = \frac{0.002}{5.23} \cdot 100 \% = 0.04 \% ; \delta_h = \frac{0.01}{12.3} \cdot 100 \% = 0.08 \%$$

Тобто вихідні дані повинні бути задані з такою точністю:

$$\pi - 99.9 \% ; r - 99.96 \% ; h - 99.92 \% . \quad \blacksquare$$

### 3.3 Правила підрахунку цифр

Під час обчислень, якщо не проводиться строгий підрахунок похибок, рекомендується користуватися правилами підрахунку цифр. Ці правила показують, як потрібно проводити округлення всіх результатів, щоб, по-перше, забезпечити задану точність остаточного результату і, по-друге, не виконувати обчислень із зайвими знаками, що не надають вплив на правильні знаки результату.

#### Правила підрахунку цифр, подані В. М. Брадїсом

- 1 Додаючи та віднімаючи наближені числа, необхідно зберегти стільки десяткових знаків, скільки їх у наближених даних із найменшим числом десяткових знаків.
- 2 Від множення та ділення потрібно зберегти стільки значущих цифр, скільки їх у наближених даних із найменшим числом правильних значущих цифр.
- 3 Підносячи наближене число до квадрата або куба, потрібно зберегти стільки значущих цифр, скільки їх в основі степеня.
- 4 Беручи квадратний або кубічний корінь з наближеного числа, необхідно зберегти стільки значущих цифр, скільки їх у підкорінному числі.
- 5 Розраховуючи проміжні результати, потрібно зберегти на одну цифру більше, ніж рекомендують правила 1–4. В остаточному результаті ця «запасна цифра» відкидається.



- 6 Якщо деякі дані мають більше десяткових знаків (під час додавання й віднімання) або більше значущих цифр (за інших дій), ніж інші, то їх попередньо потрібно округлити, зберігаючи лише одну «запасну цифру».
- 7 Для обчислення за допомогою логарифмів одночленного виразу рекомендується підрахувати число значущих цифр у наближених даних, що має найменше число значущих цифр, і використати таблицю логарифмів із числом десяткових знаків на одиницю більшим. В остаточному результаті остання значуща цифра відкидається.
- 8 Якщо дані можна брати з довільною точністю, то для одержання результату з  $k$  правильними цифрами вихідні дані потрібно брати з таким числом цифр, які згідно з попередніми правилами забезпечують  $k + 1$  цифру в результаті.

Ці правила даються в припущенні, що компоненти дій містять лише правильні цифри і кількість дій невелика.

## Лекція 4

# Відокремлення коренів рівняння. Уточнення коренів рівняння методом дихотомії та хорд

### 4.1 Поставлення задачі

Розв'язання нелінійних рівнянь з одним невідомим є однією з найважливіших математичних задач, що виникають у різних розділах фізики, біології, хімії, економіки та інших галузях науки і техніки

$$F(x) = 0, \quad (4.1)$$

де  $F(x)$  – деяка неперервна функція аргумента  $x$ , аргумент  $x$  належить деякій області  $X$ .

Нехай  $x^*$  – точний корінь, а  $x$  – його наближене значення. Зазначають, що корінь  $x$  обчислено з наперед заданою точністю  $\varepsilon$ , якщо  $|x^* - x| \leq \varepsilon$ .

Нехай  $x^* \in [a; b]$ ,  $b - a \leq \varepsilon$ , тоді  $a$  і  $b$  – наближені значення кореня  $x^*$ , відповідно з недостачею і надлишком з точністю  $\varepsilon$ . Водночас за наближене значення  $x$  з точністю  $\varepsilon$  можна взяти будь-яке число з відрізка  $[a; b]$ .

Розв'язати рівняння означає знайти множину його коренів, тобто таких значень  $x \in (a; b)$ , за яких рівняння (4.1) перетворюється в тотожність.

У загальному випадку рівняння (4.1) може мати багато коренів. Чисельні методи розв'язання нелінійних рівнянь, які розглянуто далі, дозволяють знаходити один корінь на заданому відрізку  $[a; b]$ . Водночас на інтервалі повинен існувати лише один корінь.

Знайти відрізок, що задовольняє умові можна різними способами:

- а) з фізичних міркувань, тобто на основі фізичних знань про задачу;
- б) на основі досвіду розв'язання аналогічних задач;
- в) за допомогою графічних методів;

г) шляхом відокремлення коренів.

Задачею розв'язання нелінійного рівняння є **задача знаходження такого невідомого**  $x^* \in X$ , який зводить рівняння (4.1) до тотожності  $0 \equiv 0$ , або доведення відсутності такого  $x^*$ .

Традиційно нелінійні рівняння розподіляють на алгебраїчні та трансцендентні рівняння.

*Алгебраїчним рівнянням* з одним невідомим називається рівняння, в лівій частині якого є суперпозиція алгебраїчних функцій (степеневих, раціональних, ірраціональних).

Загальний вигляд алгебраїчної функції можна подати так:

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

де  $a_i$  – числові коефіцієнти над деяким полем чисел (цілих, раціональних, дійсних);  $n$  – раціональний степінь, який визначає степінь алгебраїчної функції.

Значення аргумента  $x^*$ , яке зводить алгебраїчне рівняння виду (4.1) до тотожності  $0 \equiv 0$ , називаються коренем алгебраїчного рівняння. Числа, які можуть бути коренями алгебраїчних рівнянь складають множину алгебраїчних чисел.

Приклади алгебраїчних рівнянь

$$\begin{aligned} x^5 - 17x^4 + \sqrt{3}x^3 - 11x + 21 &= 0; \\ \sqrt{15}x^3 + 21x^2 - \frac{4\sqrt{x}}{\sqrt[3]{3x^2}} + \sqrt[5]{3x - 10} &= 0. \end{aligned}$$

*Трансцендентним рівнянням* називається рівняння, в лівий бік якого входять трансцендентні функції (логарифмічні, показникові, тригонометричні, обернені тригонометричні) і числа.

Наприклад,

$$\ln x + e^{2x^2-3} - \arctan 3x = 0.$$

Значення аргументу  $x^*$ , які зводять трансцендентне рівняння виду (4.1) до тотожності  $0 \equiv 0$ , називаються коренями трансцендентного рівняння.

Якщо значення аргумента  $x^*$  окрім функції  $F(x)$  зводить до 0 усі похідні цієї функції до  $k-1$ -го порядку включно, то  $x^*$  називають  $k$

– кратним коренем рівняння (4.1). Однократний корінь називають простим.

Методи розв'язання нелінійних рівнянь поділяють на прямі (точні, аналітичні) та ітераційні.

Прямі методи дозволяють записати рівняння у вигляді формули із скінченною кількістю операцій.

Наприклад, для рівняння  $x^2 + px + q = 0$  такими формулами будуть  $x_{1,2} = -p/2 \pm \sqrt{p^2/4 - 4}$ . Формулу розв'язку можна знайти і для багатьох трансцендентних рівнянь.

На практиці в більшості математичних моделей дослідник стикається з рівняннями більш складної структури, які вміщують у собі як трансцендентні, так і алгебраїчні функції. Аналітичний розв'язок знайти для таких рівнянь практично неможливо. У всіх таких випадках використовують чисельні методи, які дозволяють знайти корінь рівняння наближено з будь-якою наперед заданою точністю.

Чисельний підхід до розв'язання нелінійних рівнянь розбиває процес розв'язання на 2 етапи: локалізація коренів і уточнення коренів.

## 4.2 Етапи наближеного розв'язання нелінійних рівнянь. Локалізація коренів

Суть процесу локалізації коренів нелінійного рівняння полягає у виділенні з області неперервності функції  $F(x)$  такого інтервалу  $[a; b]$ , на якому функція має єдиний корінь. Цей процес відповідно до задачі називають ще процесом відокремлення кореня, знайдений інтервал  $[a; b]$  – відрізком локалізації кореня.

Локалізацію кореня можна виконувати графічно, аналітично та таблично.

### 4.2.1 Графічна локалізація кореня

Для графічної локалізації кореня будується графік функції  $y = f(x)$  і визначаються проміжки, на яких графік перетинає вісь абсцис.

Необхідно локалізувати графічно корені функції

$$F(x) = 4 - 3x^2 + 0.3x^3.$$

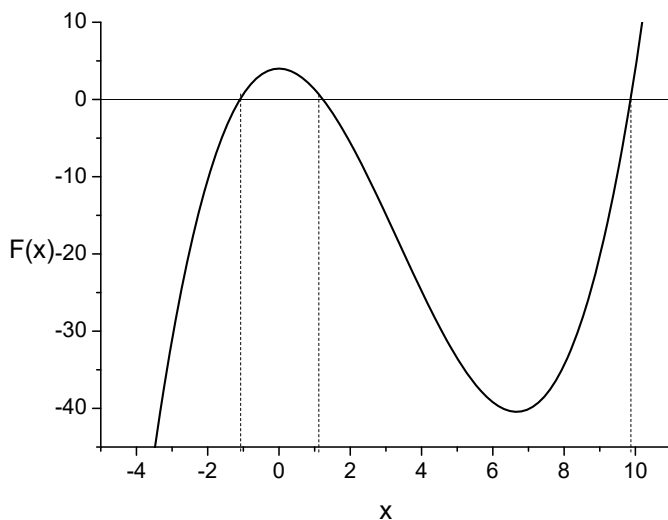


Рисунок 4.1– Графік функції  $F(x) = 4 - 3x^2 + 0.3x^3$

З рисунка 4.1 видно, що рівняння на заданому інтервалі має 3 корені. Розміщені корені на інтервалах  $[-2; 0]$ ,  $[0; 2]$  та  $[9; 11]$ . Перший етап наближеного розв'язання рівняння (4.1) виконаний. ■

На практиці частіше використовують подання нелінійного рівняння у вигляді

$$\varphi(x) = \psi(x). \quad (4.2)$$

Тобто розділяють вихідну функцію на дві, будують графіки функцій  $y_1 = \varphi(x)$ ;  $y_2 = \psi(x)$  і аналізують точки перетину цих функцій.



## ПРИКЛАД 2

Графічно локалізувати корені рівняння на інтервалі  $[0; 2]$

$$\ln x + e^{2x^2-3} - \arctan 3x = 0.$$

### Розв'язування

Перепишемо рівняння в такому вигляді

$$\ln x + e^{2x^2-3} = \arctan 3x.$$

Одержали

$$\varphi(x) = \ln x + e^{2x^2-3}; \psi(x) = \arctan 3x.$$

Будуємо графіки обох функцій, що подані на рисунку 4.2.

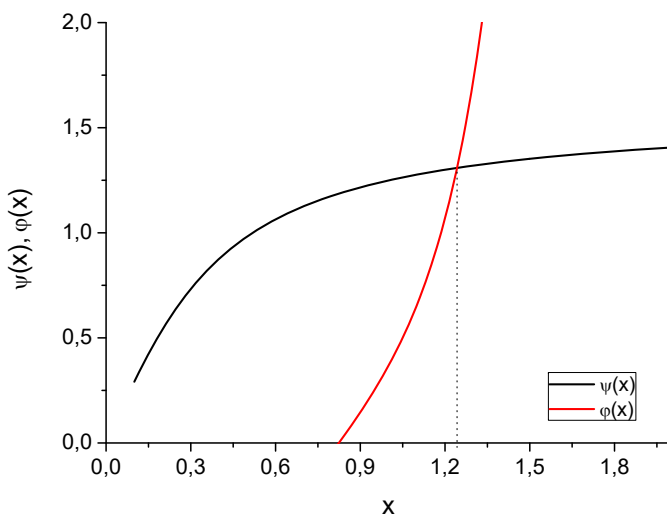


Рисунок 4.2– Графік функцій  $\varphi(x) = \ln x + e^{2x^2-3}; \psi(x) = \arctan 3x$

З графіка бачимо, що інтервал локалізації кореня може бути обраним як  $[1; 1.5]$ , але враховуючи наближеність побудови графіка і близькість кореня до правого краю, можна обрати відрізок локалізації кореня  $[1; 1.3]$ .

■

**Зауваження!** Чим щільніше локалізований корінь, тим швидше відбувається його уточнення, але занадто звуживши інтервал за графіком, можна проскочити корінь.

#### 4.2.2 Аналітична локалізація кореня

Для аналітичної локалізації кореня використовують відомості про поведінку диференційованої функції однієї змінної з математичного аналізу. Зокрема, для відокремлення простих коренів застосовують теорему про поведінку монотонної диференційованої функції в замкненому інтервалі.

##### ТЕОРЕМА 4.1: Теорема Больцано – Коші

Нехай диференційована і монотонна функція  $F(x)$  задана на інтервалі  $[a; b]$ . Якщо на кінцях цього інтервалу функція набуває значення різних знаків, тобто

$$F(a) \cdot F(b) < 0, \quad (4.3)$$

то в середині інтервалу існує єдиний корінь.

Монотонність є достатньою умовою існування єдиного кореня на замкненому інтервалі за виконання інших умов теореми. В разі немонотонної функції (перша похідна змінює знак у середині інтервалу) мова йде лише про існування хоча б одного кореня.

##### ТЕОРЕМА 4.2: Існування та єдиності кореня

Якщо функція  $F(x)$  неперервна і монотонна на відрізку  $[a; b]$  і набуває на кінцях цього відрізка значення різних знаків, тобто  $F(a)F(b) < 0$ , то в середині відрізка  $[a; b]$  рівняння (4.1) має корінь, причому єдиний.

Неперервна функція називається *монотонною* на проміжку  $[a; b]$ , якщо вона або лише зростає, або лише спадає на ньому.

### ТЕОРЕМА 4.3: Існування та єдиності кореня

Якщо функція  $F(x)$  неперервна та диференційована на відрізку  $[a; b]$  і набуває на кінцях цього відрізка значення різних знаків, тобто  $F(a)F(b) < 0$ , а похідна  $F'(x)$  зберігає знак усередині відрізка  $[a; b]$ , то рівняння (4.1) має на проміжку  $[a; b]$  корінь, причому єдиний.

Для того щоб довести, що похідна зберігає знак усередині відрізка  $[a; b]$ , потрібно показати, що  $[a; b]$  повністю міститься в одному з проміжків знакосталості функції  $y = F'(x)$ .

Відповідно до вищезазначених теорем, алгоритм відокремлення коренів можна сформулювати так:

- 1) знайти область визначення рівняння;
- 2) знайти критичні точки функції;
- 3) записати інтервали монотонності функції;
- 4) визначити знак функції на кінцях інтервалів монотонності;
- 5) визначити відрізки, на кінцях яких функція набуває значень протилежних знаків;
- 6) знайдені відрізки ізоляції коренів за необхідності звузити.

#### 4.2.3 Таблична локалізація кореня

Табличний спосіб локалізації кореня рівняння застосовують, наприклад, коли наперед заданий інтервал дослідження функції і базується на тих самих засадах, що і аналітичний метод.

Для аналізу обирають крок обчислення значення функції на заданому відрізку. Чим менший крок, тим краще буде описана функція. Для точок  $\{a, a + h, a + 2h, \dots, a + kh, \dots, b\}$  визначають знак функції  $F(x)$ .

Зміна знаку функції в сусідніх точках відрізка сигналізує про існування хоча б одного кореня рівняння на інтервалі, який замикають ці точки. Після визначення інтервалу з коренями перевіряють монотонність функції на цьому інтервалі й роблять висновок про



єдність кореня. Кількість змін знаку функції на відрізку свідчить про мінімальну кількість інтервалів локалізації кореня.

Але надійність цього методу залежить як від складності функції, так і від величини кроку. Для деяких функцій можливе існування декілька коренів на одному знайденому інтервалі, крім того не зникає і проблема кратних коренів.

Тому найпростішим методом локалізації коренів за умови використання сучасних комп'ютерів і алгоритмічних мов є графічний метод.

### 4.3 Етапи наближеного розв'язання нелінійних рівнянь. Уточнення коренів нелінійного рівняння

Розглядається нелінійне рівняння (4.1).

Необхідно знайти таке наближене значення  $\hat{x}$  кореня рівняння  $x^*$ , щоб  $\Delta_x = |\hat{x} - x^*| < \varepsilon$ , де  $\varepsilon$  – наперед задана абсолютна похибка наближеного значення кореня.

Нехай на поточному етапі за допомогою вищенаведених методів був визначений відрізок локалізації кореня  $[a; b]$ .

Уточнення коренів на відрізках локалізації відбувається ітераційно.

**Методи ітерацій** – методи, які на основі попередньо відомого початкового значення наближеної величини уточнюють її значення до заданої точності, використовуючи на кожному кроці (ітерації) наближення шуканої величини з попереднього кроку (ітерації).

Унаслідок виконання методу ітерацій одержуємо послідовність наближених значень  $\{x_0, x_1, \dots, x_n\}$ , яка повинна наближатися до шуканої величини  $x^*$  на задану точність  $0 < \varepsilon \ll 1$ .

### 4.4 Метод половинного ділення (бісекції, дихотомії)

Найпростішим з ітераційних методів є метод ділення відрізка навпіл (метод бісекції, метод дихотомії).

Суть методу полягає в звужуванні відрізка локалізації кореня  $[a; b]$  шляхом ділення його навпіл із наступною перевіркою на кінцях відрізка виконання умови (4.3), а саме  $F(a) \cdot F(b) < 0$ .

Нехай відомо, що корінь рівняння  $F(x) = 0$  належить відрізку, причому функція  $F(x)$  неперервна на цьому відрізку і набуває на його кінцях різні за знаком значення, тобто  $F(a_0) \cdot F(b_0) < 0$ .

- 1 На першому кроці ( $k = 1$ ) розділимо відрізок  $[a_0; b_0]$  навпіл точкою  $x_1 = (a_0 + b_0)/2$ , як подано на рисунку 4.3.

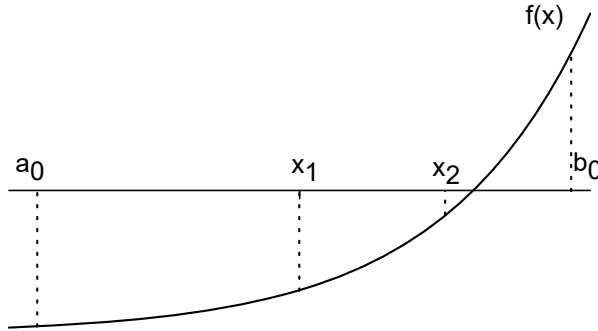


Рисунок 4.3– Ілюстрація методу половинного ділення

- 2 Якщо  $F(x_1) = 0$ , то  $x_1$  – шуканий корінь і задача розв’язана.
- 3 Якщо  $F(x_1) \neq 0$ , то потрібно визначити, на кінцях якого з відрізків  $[a_0; x_1]$  чи  $[x_1; b_0]$  функція набуває різні за знаком значення.
- 4 Позначимо цей відрізок  $[a_1; b_1]$ . Оскільки  $F(a_1) \cdot F(b_1) < 0$ , то  $x^* \in [a_1; b_1]$  і довжина відрізка  $[a_1; b_1]$  удвічі менша, ніж довжина попереднього відрізка  $[a_0; b_0]$ .
- 5 На  $k$ -му кроці обчислення проводять за такою схемою:

$$x_k = \frac{a_{k-1} + b_{k-1}}{2}; \quad F(a_{k-1}) \cdot F(b_{k-1}) < 0,$$

$x_k$  – наближення до шуканого кореня  $x^*$ ;  $[a_{k-1}; b_{k-1}]$  – попередній проміжок ізоляції кореня.

Оскільки  $x^* \in [a_k; b_k]$ , то маємо оцінку

$$|x_k - x^*| \leq b_k - a_k = (b_0 - a_0)/2^k,$$

яка характеризує похибку методу ділення відрізка навпіл і свідчить про швидкість збіжності: *метод збігається зі швидкістю геометричної прогресії, знаменник якої  $q = 1/2$ .*

Оскільки  $|x_k - x^*| \leq b_k - a_k = |x_k - x_{k-1}|$ , то обчислення закінчують, коли буде виконана нерівність  $|x_k - x_{k-1}| < \varepsilon$ , де  $\varepsilon > 0$  – задана точність.

**Зауваження!** Необхідну кількість ітерацій для досягнення заданої точності  $\varepsilon > 0$  можна оцінити попередньо: зі співвідношення  $|x_k - x_{k-1}| = ((b_0 - a_0))/2^k < \varepsilon$  випливає  $k > \log_2((b_0 - a_0)/\varepsilon)$ .

Для одержання кожних 3 правильних десяткових знаків необхідно зробити близько 10 ітерацій. На практиці ітераційний процес закінчують, коли два рази підряд одержують однакові з заданою точністю числа (перестають змінюватися десяткові знаки, які потрібно зберегти у відповіді). Проміжні обчислення здійснюють з одним чи двома запасними десятковими знаками.

**Зауваження!** Метод дихотомії є досить простим і надійним способом розв’язування нелінійних рівнянь. Він не накладає ніяких вимог на гладкість функції  $F(x)$ : функція  $F(x)$  може бути недиференційовною, достатньо лише її неперервності. Якщо під час локалізації кореня допущена помилка і на відрізку  $[a; b]$  є кілька коренів, то процес дихотомії збігається до одного з них.

**Зауваження!** До недоліків методу можна віднести невисоку швидкість збіжності. Крім того, він не застосовний для знаходження коренів парної кратності. Хоча у разі кореня високої непарної кратності метод збігається, але водночас має зменшену точність і невисоку стійкість до похибок обчислень.

Метод дихотомії застосовують тоді, коли необхідна висока надійність, а швидкість збіжності несуттєва.

## Алгоритм методу половинного ділення (бісекції, дихотомії)

1 Встановлюємо лічильник  $i = 0$ .

2 Вводимо координати кінців відрізка локалізації кореня  $[a; b]$  і точність обчислення  $\varepsilon$ .

3 Обчислюємо середину відрізка  $[a; b]$ , ( $a < b$ )

$$\hat{x}_i = \frac{a + b}{2}.$$

4 Перевіряємо умову (4.3):

Якщо  $F(a) \cdot F(\hat{x}_i) < 0$ , то перепозначаємо  $b = \hat{x}_i$ , інакше  $a = \hat{x}_i$ .

5 Порівнюємо довжину нового відрізка з заданою точністю.

Якщо  $|b - a| < 2\varepsilon$ , то

$$\hat{x}_i = \frac{a + b}{2}.$$

**Кінець алгоритму.**

Інакше рахуємо номер наступної ітерації  $i = i + 1$ .

**Перехід до пункту 3.**



## ПРИКЛАД 3

---

Знайти корінь рівняння

$$\ln x + e^{2x^2 - 3} = \arctan 3x.$$

методом половинного ділення.

### Розв'язування

На рисунку 4.2 графічно локалізовано корінь рівняння на відрізку  $[1; 1.3]$ .

Виконаємо уточнення кореня методом бісекції з точністю  $\varepsilon = 0.0001$ . Для обраного інтервалу локалізації попередньо проводимо аналіз:

- оцінюємо інтервал  $|1.3 - 1| = 0.0 > 2 \cdot 0.0001$  – точність не задовольняється;
- оцінимо необхідну кількість ітерацій:

$$N \approx \log_2 \left( \frac{1.3 - 1}{10^{-4}} \right) = \log_2 3 + 3 \log_2 10 \approx 12.$$

Тобто для наближення кореня з точністю до 4-го знака нам необхідно виконати 12 ітерацій.

Переходимо до обчислень згідно алгоритму:

1 Призначаємо лічильнику ітерацій  $i = 0$ , та кінці відрізка  $a = 1.0; b = 1.3$ .

2 Обчислюємо наближення

$$x_i = \frac{a + b}{2} = \frac{2.3}{2} \approx 1.15.$$

3 Визначаємо знак добутку значень функції в точках  $a$  і  $x_i$ .

$$\begin{aligned} F(1) \cdot F(1.15) &= (\ln 1 + e^{2 \cdot 1^2 - 3} - \arctan(3 \cdot 1)) (\ln 1.15 + e^{2 \cdot 1.15^2 - 3} - \arctan(3 \cdot 1.15)) = \\ &= (-0.8812) \cdot (-0.447737) > 0. \end{aligned}$$

Згідно з умовою п. 4 алгоритму переприсвоюємо  $a = x_i = 1.15$ .

4 Обчислюємо  $|b - a| = |1.3 - 1.15| = 0.15 > 2 \cdot 0.0001$ , точність не досягнена, тоді  $i = i + 1 = 0 + 1 = 1$ .

Переходимо до п. 2 із зміненими даними.

Метод збігся за 12 ітерацій.

Наближене значення кореня:  $\hat{x} \approx 1.242505$ .

Для заданої точності точне значення кореня  $x^* = 1.2425 \pm 0.0001$ .

**Зауваження!** Під час реалізації методу потрібно брати до уваги той факт, що функція  $F(x)$  обчислюється з деякою абсолютною похибкою  $\Delta F(x)$ . Біля точного значення кореня значення функції за абсолютною величиною наближаються до 0 і можуть стати порівняними з похибкою її обчислення. Тобто біля точного значення кореня значення функції можуть потрапити у смугу «шумів» шириною  $2\Delta F(x)$  і подальше уточнення кореня стає неможливим. Тому рекомендується заздалегідь задавати смугу «шумів» і припиняти процес наближення кореня одразу після потрапляння в неї. Якщо взяти  $\Delta F = \varepsilon$ , то процес ітерації потрібно припиняти, як тільки

$$|F(x_n)| < \varepsilon. \quad (4.4)$$

Для визначення ширини смуги можна визначити граничну абсолютну похибку обчислення функції  $F(x)$ . Тоді умова (4.4) набере вигляду

$$|F(x_n)| < \Delta^* F(x). \quad (4.5)$$

Для нашого прикладу

$$\hat{x} = x_{n+1} = 1.242505; F(x_{n+1}) \approx 0.00029 > 0.0001.$$

$$\begin{aligned} \Delta F = F'(\hat{x})\Delta x &= \left( \frac{1}{\hat{x}} + 4\hat{x}e^{2\hat{x}^2_{n+1}-3} - \frac{3}{1+9\hat{x}^2} \right) \Bigg|_{\hat{x} \approx 1.24251} \cdot 0.0001 \approx \\ &\approx 6.02864 \cdot 0.0001 \approx 0.0006 > 0.0001. \end{aligned}$$

Тобто задана точність наближення кореня рівняння (1) була досягнута раніше, ніж наближене значення функції стало менше заданої точності. ■

## 4.5 Метод хорд

Розглянемо рівняння 4.1  $F(x) = 0$ .

Необхідно знайти його корінь із точністю  $\varepsilon$  на відрізку  $[a; b]$ , на якому функція безперервна і на кінцях має значення різних знаків, тобто  $F(a) \cdot F(b) < 0$ .

Ідея методу хорд полягає в заміні на відрізку  $[a; b]$  функції  $F(x)$  на пряму, що проходить через кінці її графіка. Шуканим коренем буде перетин  $F(x)$  з віссю  $x$ .

Проведемо пряму з точки  $F(a)$  у точку  $F(b)$ , як показано на рисунку 4.4, з'єднаємо їх та одержуємо пряму  $AB$ .

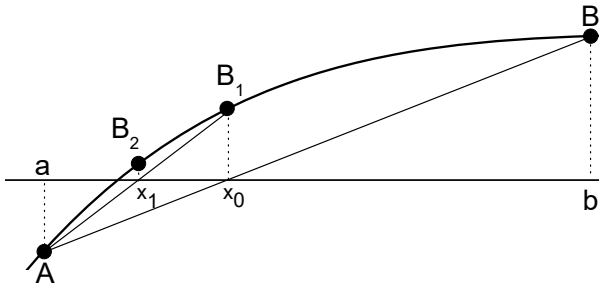


Рисунок 4.4— Ілюстрація наближень методу хорд:  $A = F(a)$ ;  $B = F(b)$ ;  $B_1 = F(x_0)$

Рівняння такої прямої має вигляд

$$\frac{x - a}{x - b} = \frac{F(x) - F(a)}{F(b) - F(a)}. \quad (4.6)$$

Оскільки на відрізку  $[a; b]$  виконується умова (4.3), побудована пряма обов'язково перетне вісь  $Ox$ . У точці перетину рівняння набере вигляду

$$\frac{x - a}{b - x} = \frac{-F(a)}{F(b) - F(a)}. \quad (4.7)$$

Позначимо точку перетину через  $x_0$ , тоді

$$x_0 = a - F(a) \frac{b - a}{F(b) - F(a)}. \quad (4.8)$$

Обчислюємо  $F(x_0)$ , як і у методі дихотомії, перевіряємо умову (4.3), визначаємо новий відрізок, на якому умова виконується.

Ітераційний процес має загальну формулу

$$x_i = a_i - F(a_i) \frac{b_i - a_i}{F(b_i) - F(a_i)}. \quad (4.9)$$

Процес потрібно припиняти або за умови (4.4), або (якщо  $\Delta F^*(x)$  менше заданої точності обчислення кореня) за умови

$$|x_i - x_{i-1}| < \varepsilon. \quad (4.10)$$

За умови, що на відрізку  $[a_i; b_i]$  перша похідна відмінна від 0  $f'(x) \neq 0$ , а друга похідна  $f''(x)$  зберігає знак, оцінка абсолютної похибки наближення визначається за формулою

$$|x_i - x_{i-1}| \leq \frac{|f(x_i)|}{\mu}; \quad \mu = \min_{x \in [a_i; b_i]} |f'(x)|. \quad (4.11)$$

**Зауваження!** Метод хорд аналогічний методу поділу навпіл, але забезпечує більш швидку збіжність.

**Зауваження!** Якщо на відрізку локалізації  $[a; b]$  друга похідна  $f''(x)$  зберігає знак, то під час обчислень за методом хорд один з кінців відрізків локалізації  $[a_i; b_i]$ , де знак функції  $f(x)$  збігається зі знаком другої похідної, залишається нерухомим. Водночас послідовні наближення  $x_1, x_2, \dots, x_i$  лежать по ту сторону від кореня, де знак функції протилежний знакові  $f''(x)$ .

## Переваги методу:

- більш висока швидкість збіжності, ніж метод половинного поділу;
- можливість заздалегідь визначити кількість кроків;
- універсальність (накладення мінімальних обмежень).

## Недолік методу:

- неможливість оцінити кількість кроків;
- великі вимоги до функції (порівняно з методом половинного поділу);
- неможливість узагальнення методу (не можна використовувати для систем).

За складністю алгоритму метод хорд майже не відрізняється від методу половинного ділення. Але для багатьох рівнянь метод хорд збігається швидше.

Розглянуті методи є не найшвидшими методами чисельного розв'язання нелінійних рівнянь. Але в них є одна істотна перевага – для неперервних монотонних функцій вони завжди збігаються до точного кореня з заданою точністю і окрім неперервності й монотонності не використовують ніяких інших властивостей функції  $F(x)$ , яка стоїть у лівій частині рівняння. Як наслідок – ці методи будуть збіжними навіть за розривної лівої частини рівняння. Усі інші методи, які мають більшу швидкість збіжності, базуються на додатковому аналізі властивостей функції  $F(x)$ .



## Лекція 5

# Метричні простори. Метод простих ітерацій розв'язання нелінійних рівнянь

### 5.1 Метричні простори

Низка питань, пов'язаних з існуванням й одиничністю розв'язків рівнянь того чи іншого типу (наприклад, диференціальних), можна сформулювати у вигляді питання про існування й одиничність нерухомої точки деякого відображення відповідного метричного простору в себе. Серед різних критеріїв існування й одиничності нерухомої точки таких відображень одним з найпростіших, і водночас найбільш важливим, є так званий принцип стискальних відображень.

**Означення 5.1** *Нехай* – деяка непуста множина ( $X \neq \emptyset$ ). *Метрикою на множині* називається функція  $\rho : X \times X \Rightarrow R$ , яка задовольняє 4 властивостям, які беруть, як аксіоми:

1 *Невід'ємність*: метрика є невід'ємною функцією, тобто

$$\rho(M_1, M_2) \geq 0 \forall M_1, M_2 \in X.$$

2 *Відокремленість*: відстань між двома точками дорівнює 0 лише в разі їх збігу, тобто

$$\rho(M_1, M_2) = 0 \iff M_1 = M_2.$$

3 *Симетричність*: відстань між  $M_1$  і  $M_2$  дорівнює відстані між  $M_2$  і  $M_1$ , тобто

$$\forall M_1, M_2 \in X : \rho(M_1, M_2) = \rho(M_2, M_1).$$

4 *Нерівність трикутника*: для будь-яких трьох точок  $M_1, M_2, M_3 \in X$  сторона трикутника, створеного на цих точках, не перевищує суми двох інших сторін, тобто

$$\forall M_1, M_2, M_3 \in X : \rho(M_1, M_2) \leq \rho(M_1, M_3) + \rho(M_3, M_2).$$

**Означення 5.2** Множина разом із заданою на ній метрикою  $\rho$  (тобто впорядкована пара  $(X, \rho)$ ) називається **метричним простором** або **простором**.

**Означення 5.3** Елементи метричного простору  $x \in X$  називають точками простору, метрика простору  $\rho(x, y) \forall x, y \in X$  задає відстань між двома точками цього простору.

Прикладами метричних просторів можуть бути, наприклад:

1 Пара  $(R, \rho) : \forall x, y \in R : \rho(x, y) = |x - y|$ .

2 Пара

$$\begin{aligned} (R^2, \rho) : \forall M_1, M_2 \in R : \rho(M_1, M_2) &= |M_2 - M_1| = \\ &= \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}. \end{aligned}$$

3 Множина  $n$ -вимірних векторів із дійсними координатами і відстанню

$$\rho(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

є метричним простором, який називається  $n$ -вимірним евклідовим простором, позначають  $R_n$ .

4 Множина  $n$ -вимірних векторів із відстанню

$$\rho_1(x, y) = \max_{1 \leq i \leq n} |x_i - y_i|$$

є метричним простором, який позначають  $R_n^0$ .

5 Множина  $n$ -вимірних векторів і відстанню

$$\rho_2(x, y) = \sum_{i=1}^n |x_i - y_i|$$

є метричним простором, який позначають  $R_n^1$ .

**Означення 5.4** Послідовність елементів  $(M_n)_{n \in \mathbb{N}} \subset X$  називається **фундаментальною** (збіжною в собі, послідовністю Коші), якщо

$$\forall \varepsilon > 0 \exists n_0(\varepsilon) : k, l > n_0 \implies \rho(M_k, M_l) < \varepsilon,$$

тобто для великих номерів відстань між елементами прямує до 0.

**Означення 5.5** Розглядається метричний простір  $(X, \rho)$ . Відображення  $\varphi : X \rightarrow X$  називається **стискаючим**, якщо існує таке додатне число  $q < 1$ , що

$$\forall M_1, M_2 \in X : \rho(\varphi(M_1), \varphi(M_2)) < q \cdot \rho(M_1, M_2). \quad (5.1)$$

**Означення 5.6** Точка  $M$  метричного простору  $(X, \rho)$  називається **нерухомою** точкою відображення  $\varphi : X \rightarrow X$ , якщо

$$M = \varphi(M). \quad (5.2)$$

Для операторів стиску справедлива теорема Банаха про нерухому точку.

### ТЕОРЕМА 5.1: Принцип стискаючих відображень

- 1 Будь-яке стискаюче відображення  $\varphi : X \rightarrow X$  у повному метричному просторі  $(X, \rho)$  має лише одну нерухому точку  $M^* \in X : M^* = \varphi(M^*)$ .
- 2 Послідовність  $M_{n+1} = \varphi(M_n), M_0 \in X, n \in \mathbb{N}$  за будь-якого початкового наближення  $M_0 \in X$  збігається до цієї точки.
- 3 Є оцінка наближення

$$\rho(M_n, M^*) \leq \frac{\rho(M_1, M_0)}{1 - q} q^n. \quad (5.3)$$

## 5.2 Метод простих ітерацій

Розглядається нелінійне рівняння (4.1) з монотонною диференційованою функцією  $F(x)$ .

Необхідно обчислити наближення  $\hat{x}$  кореня рівняння  $x^*$  із заданою точністю  $\varepsilon$ , тобто повинна виконуватися нерівність

$$|x^* - \hat{x}| \leq \varepsilon.$$

Основний ітераційний процес наближення може бути заданий формулою

$$x_{i+1} = \varphi(x_i). \quad (5.4)$$

Оскільки функція  $F(x)$  є монотонною диференційованою функцією  $R \rightarrow R$ ,  $R$  з метрикою  $\rho(M_1, M_2) = |M_2 - M_1|$  є метричним повним простором, то процес, який задається (5.4), є процесом наближення до нерухомої точки (5.2) функції  $\varphi(x)$ . Нерухомою точкою буде корінь рівняння (4.1).

У разі, **якщо функція  $\varphi(x)$  буде відповідати умовам стискуючого відображення (5.1)**, то результатом ітераційного процесу (5.4) буде послідовність  $\{x_0, x_1, \dots, x_n\}$ , елементи якої з ростом  $n$  будуть наближатися один до одного і до точного значення кореня  $x^*$  на відстань  $< \varepsilon$ . У такому разі, використовуючи теорему про принцип стискуючого відображення, можна було б зробити висновок про збіжність ітераційного процесу до нерухомої точки, яка є коренем вихідного рівняння (4.1) і оцінити степінь наближення результату до  $x^*$ .

Отже, основні задачі методу простих ітерацій – виразити з рівняння (4.1)  $x$  через функцію  $\varphi(x)$  таку, яка відповідала б умовам стискуючого відображення і визначити параметри оцінки наближення кореня  $x^*$  послідовністю ітерацій.

### Розглянемо спочатку оцінку наближення

Для наочності розглянемо оцінку наближення ітерацій до кореня  $x^*$ .

Нехай  $x_n = x^* + \varepsilon_n$ ;  $x_{n+1} = x^* + \varepsilon_{n+1}$ , де  $\varepsilon_n, \varepsilon_{n+1}$  – похибки наближення  $n$ -ї та  $n + 1$ -ї ітерацій відповідно.

Тоді згідно (5.4)

$$x^* + \varepsilon_{n+1} = \varphi(x^* + \varepsilon_n).$$

Використавши розкладення функції  $\varphi(x^* + \varepsilon_n)$  у ряд Тейлора в околі  $x^*$ , одержимо

$$x^* + \varepsilon_{n+1} = \varphi(x^* + \varepsilon_n) = \varphi(x^*) + \varphi'(x^*)\varepsilon_n + 0(\varepsilon_n^2), \quad (5.5)$$

де  $0(\varepsilon_n^2)$  – величина більшого порядку малості, ніж  $\varepsilon_n$ . Якщо відкинути цю величину як неістотну, то з урахуванням нерухомості точки  $x^*$  ( $\varphi(x^*) = x^*$ ) одержимо

$$\varepsilon_{n+1} \approx \varphi'(x^*)\varepsilon_n. \quad (5.6)$$

- 1 Якщо  $|\varphi'(x^*)| > 1$ , то похибка  $\varepsilon_{n+1}$  за абсолютним значенням більша, ніж  $\varepsilon_n$ . Наближення  $x_{n+1}$  віддалене від далі, ніж  $x_n$ . Отже, точка  $x^*$  буде точкою «відштовхування» для наближень  $x_n$ , достатньо близьких до неї. Послідовність наближень  $(x_n)_{n \in \mathbb{N}}$  не буде збігатися до  $x^*$ .
- 2 Якщо  $|\varphi'(x^*)| < 1$ , то  $|\varepsilon_{n+1}| < |\varepsilon_n|$  і можна сподіватися, що послідовність  $(x_n)_{n \in \mathbb{N}}$  у разі, коли початкове наближення  $x_0$  взяте досить близько до  $x^*$ , буде збігатися до  $x^*$  зі швидкістю геометричної прогресії зі знаменником  $q = |\varphi'(x^*)|$ .

Тобто

$$\varepsilon_{n+1} = |x_{n+1} - x^*| = |\varphi(x_n) - \varphi(x^*)|; \varepsilon_n = |x_n - x^*|;$$

$$|\varphi(x_n) - \varphi(x^*)| < q|x_n - x^*| \Rightarrow |\varphi(x_n) - \varphi(x^*)| < q^n|x_0 - x^*|$$

**Функція  $\varphi(x)$  за таких умов є стискуючим відображенням.**

- 3 Якщо  $0 < \varphi'(x^*) < 1$ , то  $\varepsilon_n$  і  $\varepsilon_{n+1}$  будуть мати однакові знаки і збіжність буде монотонною.
- 4 У разі, коли  $-1 < \varphi'(x^*) < 0$ ,  $\varepsilon_n$  і  $\varepsilon_{n+1}$  будуть мати різні знаки і наближення будуть збігатися до  $x^*$  коливаючись навколо нього.
- 5 Якщо  $\varphi'(x^*) = 0$ , маємо справу з кратним коренем. Якщо корінь  $x^*$  має кратність  $m$ , тобто  $\varphi'(x^*) = \varphi''(x^*) = \dots = \varphi^{(m-1)}(x^*) = 0$ , розкладання у ряд Тейлора функції  $\varphi(x_n)$  буде таким:

$$\varphi(x^* + \varepsilon_n) = \varphi(x^*) + \frac{\varphi^{(m)}(x^*)}{m!}\varepsilon_n^m + 0(\varepsilon_n^{m+1}).$$

Підставивши його в (5.5) і відкинувши малу величину більшого порядку, ніж  $\varepsilon_n^m$ , одержимо

$$\varepsilon_{n+1} \approx \frac{\varphi^{(m)}(x^*)}{m!}\varepsilon_n^m. \quad (5.7)$$

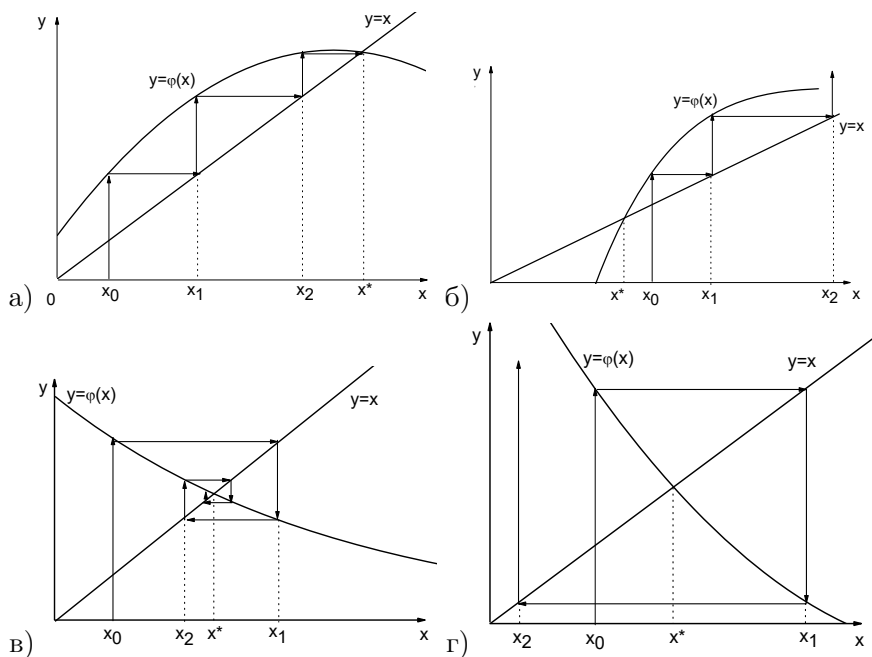


Рисунок 5.1- а)  $0 < \varphi'(\hat{x}) < 1$ ; б)  $\varphi'(\hat{x}) > 1$ ; в)  $-1 < \varphi'(\hat{x}) < 0$ ; г)  $\varphi'(\hat{x}) < -1$

З останнього виразу видно, що похибка  $\varepsilon_{n+1}$  буде малою величиною порядку  $m$  (кратність кореня) відносно  $\varepsilon_n$ .

Отже, можна зробити висновок, що для того, щоб послідовність наближень  $(x_n)_{n \in \mathbb{N}}$  збігалася до кореня  $x^*$  необхідно, щоб поблизу кореня  $x^*$  виконувалася умова

$$|\varphi'(x)| < 1. \quad (5.8)$$

Якщо ця умова виконується всюди на відрізку локалізації кореня, то за початкове наближення можна брати будь-яку точку з цього відрізка. Чим ближче похідна до 0, тим швидше збігається ітераційний процес.

На рисунку 5.1 ілюструється перебіг ітераційного процесу методу простих ітерацій за різних значень.

З рисунку 5.1 видно, що в разі  $|\varphi'(x)| < 1$  ітераційна послідовність  $(x_n)_{n \in \mathbb{N}}$  за скінчену кількість кроків збігається до  $x^*$ , а якщо  $|\varphi'(x)| > 1$  – віддаляється від нього.

**Зауваження!** Для з'ясування питання виконання умови (5.8) іноді достатньо аналізувати найбільше  $M$  і найменше  $m$  значення похідної  $\varphi'(x)$  на інтервалі локалізації кореня  $[a; b]$ . Оскільки функція  $\varphi(x)$  неперервна і диференційована, то такі значення завжди існують. Для з'ясування монотонності поведінки  $\varphi'(x)$  буває зручно провести аналіз другої похідної  $\varphi''(x)$  на інтервалі  $[a; b]$ .

Якщо  $\varphi(x)$  у (5.4) обрана такою, що виконується умова (5.8), ітераційний процес наближення до кореня рівняння (4.1) збігається зі швидкістю геометричної прогресії зі знаменником  $q = |\varphi'(x^*)|$ . Послідовність  $(x_n)_{n \in \mathbb{N}}$  є фундаментальною і збіжною до  $x^*$ . Функція  $\varphi(x)$  є стискующим відображенням  $R \rightarrow R$  і для нього виконується теорема 5.1, тобто є оцінка

$$|x_{n+1} - x^*| \leq \frac{q^{n+1}}{1 - q} |x_1 - x_0|. \quad (5.9)$$

### Перетворення рівняння до ітераційного вигляду

Будь-яку нелінійну функцію можна перетворити до вигляду (5.4) не одним способом. Головним критерієм перетворення є виконання умови (5.8) на відрізьку локалізації кореня.



### ПРИКЛАД 1

Привести рівняння до застосування методу простих ітерацій

$$e^{2x} + 3x = 4.$$

#### Розв'язування

Рівняння  $e^{2x} + 3x = 4$  має інтервал локалізації кореня  $[0.4; 0.6]$ .

Це рівняння можна привести до вигляду (5.4) двома способами:

$$1 \quad x = \frac{4 - e^{2x}}{3};$$

$$2 \quad x = \frac{1}{2} \ln(4 - 3x).$$

Для першого подання маємо

$$\varphi(x) = \frac{4 - e^{2x}}{3}; \varphi'(x) = -\frac{2}{3}e^{2x};$$

$$\varphi'(0.4) \approx -1.5; \varphi''(x) = -\frac{4}{3}e^{2x} < 0 \quad \forall x \in R;$$

Отже,  $\varphi'(x)$  є спадною функцією, на інтервалі  $[0.4; 0.6]$ , найбільше значення набуває в точці  $x = 0.4$ ;  $|\varphi'(0.4)| = |-1.5| > 1$  в усіх інших точках інтервалу похідна за модулем буде лише більшою. Тому така функція не є стискующим відображенням і не підходить до методу простих ітерацій.

Для другого подання

$$\varphi(x) = \frac{1}{2} \ln(4 - 3x); \varphi(x) = \frac{-3}{8 - 6x};$$

$$\varphi''(x) = -\frac{18}{(8 - 6x)^2} < 0; |\varphi'(0.6)| = \left| \frac{-3}{8 - 3.6} \right| < 1.$$

Отже,  $\varphi'(x)$  є спадною від'ємною функцією на інтервалі  $[0.4; 0.6]$ . Найбільше за модулем значення похідна набуває в точці  $x = 0.6 < 1$ . Тож функція  $\varphi(x) = \frac{1}{2} \ln(4 - 3x)$  є стискующим відображенням на інтервалі  $[0.4; 0.6]$ . Ітераційний процес, побудований на другому поданні, буде збіжним. ■

*Якщо важко оцінити значення похідної на інтервалі локалізації кореня, ітераційний процес будують за такою схемою:*

Розглянемо рівняння  $F(x) = 0$ .

Помножимо обидві частини рівняння на параметр  $\tau$ :  $\tau F(x) = 0$ . В одержаному рівнянні додамо до обох частин  $x$ . Одержимо рівняння

$$x + \tau F(x) = x \quad \text{або} \quad x = x - \tau F(x)$$

еквівалентне вихідному.

У цьому разі

$$\varphi(x) = x - \tau F(x).$$

Тоді ітераційний процес буде мати вигляд

$$x_{n+1} = x_n - \tau F(x_n). \quad (5.10)$$

Підбираючи  $\tau$ , потрібно перевіряти виконання умови

$$|\varphi'(x)| = |1 - \tau F'(x)| < 1.$$



Очевидно, що параметр  $\tau$  потрібно обирати одного знака з  $F'(x)$ .

У такому разі вираз  $1 - \tau F'(x)$  буде розміщуватись у межах

$$1 - \tau m < 1 - \tau F'(x) < 1 - \tau M,$$

де  $M, m$  – найбільше і найменше значення  $F'(x)$  на інтервалі  $[a; b]$ . Якщо відомі обидві ці величини, параметр  $\tau = 2/(M + m)$ , якщо лише  $M$ , то  $\tau = 1/M$ .

- Якщо параметр підібраний так, що  $-1 < \varphi'(x) < 0$ , то через те, що наближення коливаються навколо кореня,  $|x_{n+1} - x_n| \geq |x_{n+1} - x^*|$  завжди і можна для зупинення ітераційного процесу використати умову

$$|x_{n+1} - x_n| \leq \varepsilon. \quad (5.11)$$

- У разі, коли  $0 < \varphi'(x) < 1$ , процес наближення до кореня односторонній і умова  $|x_{n+1} - x_n| \leq \varepsilon$  може виконатися раніше, ніж умова  $|x_{n+1} - x^*| \leq \varepsilon$ , тому за критерій зупинення процесу наближення краще використовувати умову

$$\frac{1}{1 - q} |x_{n+1} - x_n| \leq \varepsilon, \quad (5.12)$$

де  $q = \max_{[a; b]}(\varphi'(x))$ .

У разі, коли не вдається обчислити значення  $q$ , критерій закінчення ітераційного процесу має вигляд

$$|x_n - x_{n+1}| \leq \left| \frac{1 - \alpha_n}{\alpha_n} \right| \cdot \varepsilon, \quad (5.13)$$

де  $\alpha_n = \frac{|x_n - x_{n-1}|}{|x_{n-1} - x_{n-2}|}$ .



## ПРИКЛАД 2

Знайти розв'язок рівняння  $\ln x + e^{2x^2-3} - \arctan 3x = 0$  методом простих ітерацій на інтервалі  $[1; 1.3]$ .

**Розв'язування**

Побудуємо ітераційну формулу.

Функція  $F(x) = \ln x + e^{2x^2-3} - \arctan 3x$  за будь-якого перетворення до вигляду (5.4)  $x_{i+1} = \varphi(x_i)$  дає досить складну похідну, тому побудуємо ітераційний процес, використавши (5.10)

$$x_{n+1} = x_n - \tau F(x_n).$$

Похідна функції  $F(x)$

$$F'(x) = \frac{1}{x} + 4xe^{2x^2-3} - \frac{3}{1+9x^2}$$

на інтервалі  $[1; 1.3]$  є функцією зростаючою, тому її найбільше і найменше значення досягаються у точках  $a = 1; b = 1.3$ . За  $\tau$  візьмемо величину

$$\tau = \frac{2}{M+m} = \frac{2}{F'(a) + F'(b)}.$$

Ітераційний процес набирає вигляду

$$x_{n+1} = x_n - \frac{2F(x_n)}{F'(a) + F'(b)}.$$

Тоді

$$\varphi(x) = x_n - \frac{2F(x_n)}{F'(a) + F'(b)}; \varphi'(x) = 1 - \frac{2F'(x)}{F'(a) + F'(b)}.$$

$\tau \approx 0.19; F'(x) \in (0.2; 0.85)$ .

Якщо за початкове наближення взяти  $x_0 = 1.3$ , то виконується умова  $-1 < \varphi'(x) < 0$ , отже, ітераційний процес можна припинити за умови  $|x_{n+1} - x_n| \leq \varepsilon$ .

Результат ітераційного процесу:

$i = 1$	$x_i = 1.22183928406$	$f = -0.118224323891$
$i = 2$	$x_i = 1.24466348759$	$f = 0.0131141998029$
$i = 3$	$x_i = 1.24213168063$	$f = -0.00221951660973$
$i = 4$	$x_i = 1.24256017705$	$f = 0.000361624780762$

корінь  $x = 1.24249036233; f(x_i) = -5.93070225468e - 05$ .

Ітераційний процес зійшовся за 6 ітераціями (враховуючи нульову і останню). З ітераційного процесу видно, що значення наближень кореня і функції коливаються навколо точних значень, що відповідає тому факту, що  $-1 < \varphi'(x) < 0$ . Функція наближається до 0 лінійно, тобто ітераційний процес має лінійну збіжність. ■

## Лекція 6

### Метод Ньютона уточнення кореня нелінійного рівняння. Модифікації методу Ньютона

#### 6.1 Метод Ньютона уточнення кореня нелінійного рівняння

Розглядається рівняння  $f(x) = 0$ , де  $f(x)$  – диференційована функція.

Задача, як і в попередніх випадках, полягає в знаходженні з заданим ступенем точності  $\varepsilon$  наближеного значення  $\hat{x}$  до кореня рівняння  $x^*$ .

Нехай відомо початкове наближення  $x_0$  кореня. З точки  $(x_0, f(x_0))$  проведемо дотичну до графіка функції, як показано на рисунку 6.1.

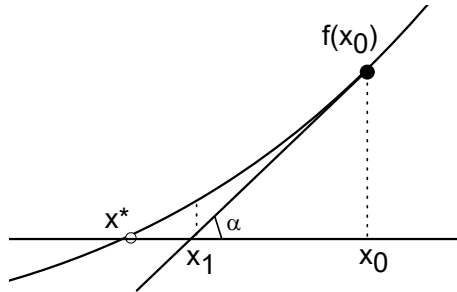


Рисунок 6.1– Ілюстрація побудови наближень методом Ньютона

Дотична перетне вісь абсцис у точці  $x_1$  під кутом  $\alpha$ . Значення  $x_1$  легко знайти, використовуючи формулу дотичної

$$f(x) = f(x_0) + f'(x_0)(x - x_0).$$

Якщо виразити звідси значення  $x$  з урахуванням того, що в точці перетину осі абсцис  $f(x) = 0$ , одержимо формулу обчислення наближення  $x_1$  :

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Будуємо дотичну до точки  $(x_1, f(x_1))$ . Вона перетне вісь абсцис у точці  $x_2$  :

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

Загальна формула ітераційного процесу

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \quad (6.1)$$

Одержану ітераційну формулу можна вивести таким чином. Нехай маємо деяке наближення  $x_{k-1}$  до кореня  $x^*$ . Лінеаризуємо функцію  $f(x)$  в околі точки  $x^*$ , дістанемо:

$$f(x^*) = f(x_{k-1}) + f'(\xi)(x^* - x_{k-1}) = 0; \xi \in (x^*; x_{k-1}).$$

Для метода Ньютона необов'язково попередньо визначати інтервал локалізації кореня, але початкове значення необхідно обирати, виходячи з властивостей функції  $f(x)$ .

Аналізуючи рисунок 6.1, можна помітити, що для того, щоб точки перетину дотичними осі абсцис мали змогу підходити до кореня  $x^*$  у разі, коли функція розміщена над віссю абсцис, тобто  $f(x) > 0$ , початкове наближення для ітераційного процесу потрібно обирати на угнутій частині функції, тобто  $f''(x_0) > 0$ . Причому необхідно, щоб такий тип викривлення функція зберігала на всьому інтервалі від  $x_0$  до  $x^*$ . Коли ж початкове наближення обирається на відрізку, де  $f(x) < 0$  (під віссю абсцис), то для збіжності ітераційного процесу на цьому відрізку функція повинна бути опуклою, тобто  $f''(x_0) < 0$ . Результатом такого аналізу стає умова вибору початкового наближення  $x_0$  кореня  $x^*$ :

$$f(x)f''(x) > 0. \quad (6.2)$$

**Достатню умову збіжності** методу дотичних встановлює така теорема

### ТЕОРЕМА 6.1

Нехай  $[a; b]$  – відрізок, який містить корінь  $x^*$  рівняння

$$f(x) = 0.$$

Якщо  $f(a) \cdot f(b) < 0$ , причому  $f'(x)$  і  $f''(x)$  відмінні від нуля та зберігають знак на  $[a; b]$ , то виходячи з довільного початкового наближення  $x_0 \in [a; b]$ , що задовольняє умову  $f(x) \cdot f''(x_0) > 0$ , ітераційний процес методу Ньютона

$$x_k = x_{k-1} - \frac{f(x_{k-1})}{f'(x_{k-1})}$$

монотонно збігається до єдиного на відрізку  $[a; b]$  кореня  $x^*$ , який можна обчислити з будь-яким ступенем точності.

**Зауваження!** Збіжність методу Ньютона істотно залежить від того, чи достатньо близько до кореня взято початкове наближення.

Функція  $f(x)$  повинна зберігати **строгу** монотонність на всьому проміжку від  $x_0$  до  $x^*$ . На рисунку (6.2) графічно ілюструються випадки виконання і порушення умови

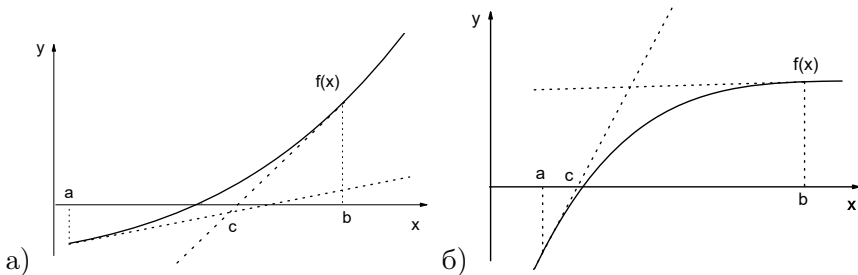


Рисунок 6.2– а)  $f''(x) > 0 \forall x \in [a; b]$ ; б)  $f''(x) < 0 \forall x \in [a; b]$

Графік 6.2 а) має функцію  $f(x)$  з  $f''(x) > 0$  на інтервалі  $[a; b]$ . Якщо за початкове наближення обрано точку  $b = x_0$ , в якій  $f(b) > 0$ , то дотична перетинає вісь абсцис у точці  $x \in [a; b]$  і наближається до кореня  $\xi$ , для початкового значення  $x_0 = a; f(a) < 0$ , умова (6.2) порушена і дотична до функції в точці  $(a, f(a))$  перетинає вісь абсцис поза  $[a; b]$ , тобто ітераційний процес може збігатися (в разі, якщо для  $x > b$  поведінка функції не зміниться), але за більшу кількість ітерацій, або зовсім не збігатися.

Графік 6.2 б) демонструє таку саму побудову ітераційного процесу для  $f''(x) < 0 \forall x \in [a; b]$ .

Для аналітичного виведення формули (6.1) наблизимо функцію  $f(x)$  відрізком ряду Тейлора в околі початкового наближення  $x_0$ .

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0).$$

Замість рівняння (4.1) розв'яжемо лінеаризоване рівняння

$$f(x_0) + f'(x_0)(x - x_0) = 0.$$

Звідси маємо  $x = x_0 - f(x_0)/f'(x_0)$ . Взявши такий  $x$  за перше наближення, одержимо  $x_1 = x_0 - f(x_0)/f'(x_0)$ , звідки автоматично випливає (6.1).

### Оцінка збіжності методу Ньютона

Виводячи (6.1) з лінеаризованого вигляду рівняння, ми відкинули усі члени ряду Тейлора, починаючи з третього. Зробимо аналіз похибки такого обмеження на  $n$ -му кроці ітераційного процесу. Розглянемо розкладання функції у ряд Тейлора в околі  $x_n$  з урахуванням відкинутого члена.

$$f(x) \approx f(x_n) + f'(x_n)(x - x_n) + \frac{f''(x_n)}{2!}(x - x_n)^2.$$

Для кореня рівняння (4.1) можна записати

$$f(x^*) = 0 \approx f(x_n) + f'(x_n)(x^* - x_n) + \frac{f''(x_n)}{2!}(x^* - x_n)^2.$$

Оскільки  $f'(x) \neq 0$  (через строгу монотонність  $f(x)$  на інтервалі побудови ітераційного процесу), то одержаний вираз можна перетворити так:

$$\begin{aligned} 0 &\approx \frac{f(x_n)}{f'(x_n)} + (x^* - x_n) + \frac{f''(x_n)}{2f'(x_n)}(x^* - x_n)^2 \Rightarrow \\ &\Rightarrow \left[ x_n - \frac{f(x_n)}{f'(x_n)} \right] - x^* \approx \frac{f''(x_n)}{2f'(x_n)}(x^* - x_n)^2. \end{aligned}$$

Враховуючи, що  $x_n - f(x_n)/f'(x_n) = x_{n+1}$ , запишемо остаточно

$$x_{n+1} - x^* \approx \frac{f''(x_n)}{2f'(x_n)}(x^* - x_n)^2.$$

Позначивши  $|x_{n+1} - x^* = \varepsilon_{n+1}|$ ;  $|x_n - x^* = \varepsilon_n|$  і замінивши  $f''(x_n)$  на максимальне значення її модуля  $|f''(x)|$  з інтервалу  $|x^* - x_0|$  (позначимо його  $M_2$ ) і  $f'(x_n)$  на мінімальне значення її модуля  $|f'(x)|$  з цього ж інтервалу (позначимо його  $m_1$ ), одержимо оцінку збіжності ітераційного процесу

$$\varepsilon_{n+1} < \frac{M_2}{2m_1} \varepsilon_n^2. \quad (6.3)$$

Очевидно, що похибка буде зменшуватися, якщо на першому кроці

$$\frac{M_2}{2m_1} |x^* - x_0| < 1. \quad (6.4)$$

За умов (6.2) і (6.4) ітераційний процес (6.1) буде збіжним. Причому кожне наближення буде мати похибку, меншу за квадрат попередньої. Тобто процес ітерації за методом Ньютона має **квадратичну** збіжність до кореня рівняння.



## ПРИКЛАД 1

Знайти розв'язок рівняння  $\ln x + e^{2x^2-3} - \arctan 3x = 0$  методом Ньютона на інтервалі  $[1; 1.3]$ .

### Розв'язування

Виконаємо дослідження виконання умови (6.2) для вибору початкового значення наближення  $x_0$ :

$$f'(x) = \frac{1}{x} + 4e^{2x^2-3} - \frac{3}{1+9x^2};$$

$$f''(x) = -\frac{1}{x^2} + 4e^{2x^2-3}(1+4x^2) + \frac{54x}{(1+9x^2)^2}.$$

$f(1) = -0.8812$ ;  $f''(1) = 5.4261 \Rightarrow f(1)f''(1) < 0$  – умова не виконується;

$f(1.3) = 0.4049$ ;  $f''(1.3) = 39.2156 \Rightarrow f(1.3)f''(1.3) > 0$  – умова виконується, тож  $x_0 = 1.3$  може виступати початковим значенням.

Як показано на рисунку 6.3, функція  $f(x)$  монотонно зростає, угнута на всьому інтервалі  $[1; 1.3]$ , виходячи із значень  $f''(x)$ , перша похідна теж монотонно зростає.

Отже, найбільше значення 2-ї похідної розміщена в точці 1.3, а найменше значення 1-ї похідної – у точці 1. Маємо

$$M_2 = 45.1; \quad m_1 = 2.2; \quad \frac{M_2}{2m_1} = \frac{45.1}{4.4} \approx 10.4,$$

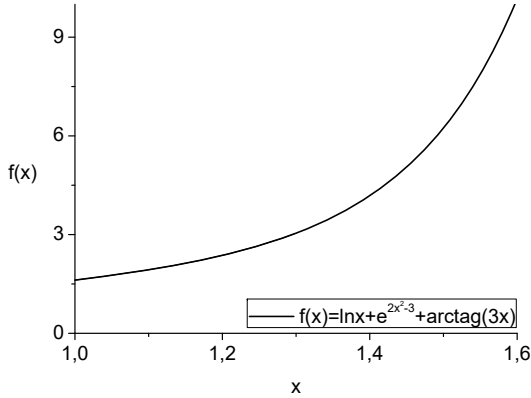


Рисунок 6.3– Графік функції на інтервалі [1; 1.6]

отже, ітераційний процес буде збіжним, якщо початкове наближення розміщено в околі кореня з радіусом, меншим, ніж  $1/10.4$ , тобто

$$|x^* - x_0| < \frac{1}{10.4} \approx 0.096.$$

У такому разі наближення буде мати похибку  $\varepsilon_{n+1} = 0.0001$ , якщо попередня ітерація матиме

$$\varepsilon_n = \sqrt{\frac{\varepsilon_{n+1} \cdot 2 \cdot m_1}{M}} = \sqrt{\frac{0.0001}{0.96}} \approx 0.03.$$

Процес збігся за 5 ітерацій (з урахуванням 0-ї).

Відстань між наближенням необхідної точності та початковим наближенням  $|1.3 - 1.2425| = 0.0575 < 0.096$ , тобто початкове значення вибрано вдало

$$f(\hat{x}) = f(1.2425) = 7.77 \cdot 10^{-14}.$$

Степені малості функції в ітераціях показують квадратичну збіжність до 0.

Для чисельного розв'язання цього рівняння за методом Ньютона за початкове наближення необхідно взяти правий край визначеного інтервалу відокремлення кореня. Для нього умова збіжності виконана, розв'язок одержаний за 5 ітерацій. ■



## 6.2 Спрощений метод Ньютона (метод однієї дотичної)

Одним із значних недоліків методу Ньютона є необхідність обчислювати похідну функції на кожному кроці ітераційного процесу. Така процедура може витратити значний час у складних обчисленнях. У зв'язку з цим часто застосовують цей метод, взявши в кожній ітерації за поточну похідну, значення похідної функції для початкового наближення. Тобто ітераційна формула набирає вигляду

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)}. \quad (6.5)$$

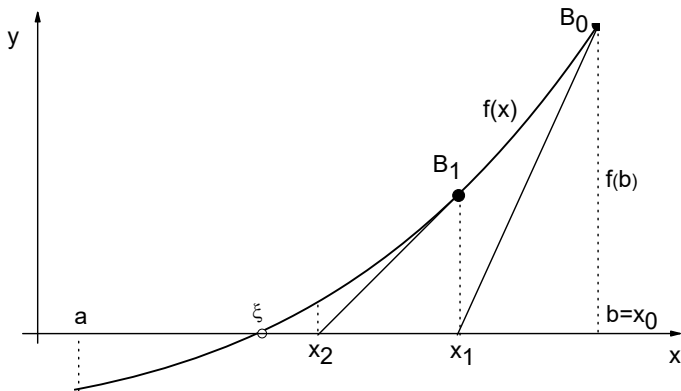


Рисунок 6.4– Ілюстрація побудови наближень до кореня  $\xi$  рівняння  $f(x) = 0$  методом однієї дотичної

Як подано на рисунку 6.4, початкове наближення взяте в точці  $b$ , дотична до  $B_0 = f(b)$  визначає перше наближення  $x_1$ , далі в точці  $B_1$  будуємо пряму, паралельну дотичній в точці  $b_0$ . Перетин цієї прямої з віссю абсцис дає  $x_2$  і точку  $B_2 = f(x_2)$  і т. д.

Цей метод фактично є методом простих ітерацій з

$$\varphi(x) = x - \tau f(x); \tau = 1/f'(x_0).$$

Похідна  $\varphi'(x) = 1 - f'(x)/f'(x_0)$ .

Отже, початкове значення потрібно обирати так, щоб на інтервалі від  $x^*$  до  $x_0$  похідна мала сталий знак. Це досягається **близь-**

кістю початкового наближення до кореня. В цьому разі, якщо

$$\left| 1 - \frac{f'(x - n)}{f'(x_0)} \right| < 1,$$

ітераційний процес збігається лінійно, як і в методі простих ітерацій.

Зупинення ітераційного процесу відбувається за тими самими критеріями, що і в методі простих ітерацій (5.11).

### 6.3 Метод січних

Метод Ньютона потребує для своєї реалізації обчислення похідної, що обмежує його застосування. Цей недолік усуває його модифікація – метод січних.

Якщо обрати початкове наближення кореня  $x^*$  рівняння  $f(x) = 0$  досить близько, то уникнути процедури обчислення похідної на кожному кроці можна, застосовуючи формулу побудови розділеної різниці функції за двома попередніми наближеннями.

Розділена різниця функції  $y = f(x)$  за двома точками  $x_0, x_1$  має вигляд

$$f(x_0, x_1) = \frac{f(x_0) - f(x_1)}{x_0 - x_1}$$

і є узагальнювальним поняттям похідної.

Обравши за початкове наближення точку  $x_0$  за умовами методу Ньютона, а за наближення  $x_1$  точку  $x_1 = x_0 + \varepsilon$  або  $x_1 = x_0 - \varepsilon$  (із міркувань зсуву у бік кореня), проведемо січну через точки  $(x_0, f(x_0))$  та  $(x_1, f(x_1))$ . Січна має рівняння

$$\frac{x - x_1}{x_0 - x_1} = \frac{f(x) - f(x_1)}{f(x_0) - f(x_1)}.$$

Звідси точку  $x_2$  перетину січною осі абсцис з урахуванням того, що можна знайти

$$x_2 = x_1 - \frac{(x_0 - x_1)}{f(x_0) - f(x_1)} f(x_1).$$

Загальний ітераційний процес має формулу

$$x_i = x_{i-1} - \frac{(x_{i-2} - x_{i-1})}{f(x_{i-2}) - f(x_{i-1})} f(x_{i-1}). \quad (6.6)$$

Оскільки на малих відстанях від кореня

$$\frac{f(x_{i-2}) - f(x_{i-1})}{x_{i-2} - x_{i-1}} \approx f'(x_i) \Rightarrow \frac{(x_{i-2} - x_{i-1})}{f(x_{i-2}) - f(x_{i-1})} f(x_{i-1}) \approx \frac{1}{f'(x_i)},$$

тобто метод є наближенням до методу Ньютона.

З формули видно, що метод двокроковий, тобто для чергової ітерації потрібні два попередні кроки.

Альтернативою до вибору  $x_1$  шляхом додавання заданої похибки до  $x_0$  в разі, коли визначений інтервал локалізації кореня  $[a; b]$ , за початкові наближення беруть межі інтервалу. В цьому разі початкова ітерація нагадує ітерацію *методу хорд*.

Ілюстрація методу січних наведена на рисунку 6.5

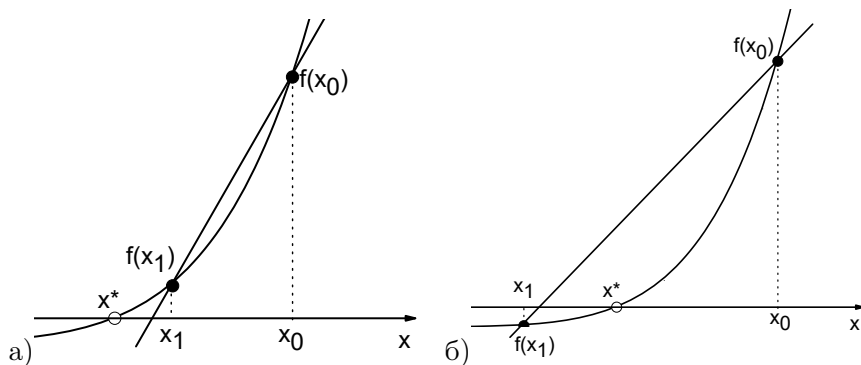


Рисунок 6.5– Ілюстрація методу січних: а) обидва наближення з одного боку від кореня; б) обидва наближення по різні боки від кореня

Метод січних збігається дещо повільніше, ніж метод Ньютона, але для нього немає потреби обчислювати  $f'(x)$  та аналізувати поведінку  $f''(x)$  на інтервалі  $[x^*; x_0]$ , що істотно спрощує і прискорює ітераційний процес для складних функцій. Крім цього такий процес стає можливим для функції, заданої таблично, без попередньої інтерполяції.

Ітераційний процес зупиняється за виконання умов (4.4) та (4.10).

За своїм алгоритмом метод січних схожий на метод хорд, але на відміну від останнього початкові наближення в методі січних можуть бути розміщеними як з обох боків (див. рисунок 6.5 б)), так

і з одного боку (див. рисунок 6.5 а)). Окрім того немає необхідності перевіряти знаки функції на кінцях новостворених відрізків.

## 6.4 Метод Стефенсена

Нехай маємо дві точки  $z_n \approx x_n$ , які належать малому околу кореня  $x^*$  і  $z_n = x_n + f(x_n)$ . Тоді розділена різниця функції для цих двох точок буде

$$f(x_n, z_n) = \frac{f(z_n) - f(x_n)}{z_n - x_n} = \frac{f(f(x_n) + x_n) - f(x_n)}{f(x_n)}.$$

За умови малої відстані між  $x_n$  і  $z_n$  наведена розділена різниця наближається до похідної в точці  $x_n$ . Підставивши таке наближення у формулу (6.1), одержуємо *ітераційну формулу методу Стефенсена*

$$x_i = x_{i+1} - \frac{f^2(x_{i-1})}{f(x_{i-1} + f(x_{i-1})) - f(x_{i-1})}. \quad (6.7)$$

### Переваги методу

Однокроковий, збігається квадратично.

### Недоліки методу

Залежність збіжності від вибору початкового наближення  $x_0$ . Біля самого кореня виникає полоса «шумів». Зупинення ітераційного процесу відбувається за (4.4) або за (4.10).

## Лекція 7

# Класифікація методів розв'язання систем лінійних алгебраїчних рівнянь та їх порівняльні характеристики. Точні методи розв'язання СЛАР

### 7.1 Розв'язання систем лінійних алгебраїчних рівнянь

У методах обчислення, які застосовуються до комп'ютерного моделювання задач більшості наукових галузей, дуже велике значення має розділ, в якому розглядаються методи лінійної алгебри. Крім окремого використання в різних наукових галузях (наприклад, моделювання задач фізики, економіки, механіки і т. ін.) методи лінійної алгебри є як допоміжними для інших чисельних методів.

Можна назвати такі методи, які є складовими методів лінійної алгебри:

- методи розв'язання систем лінійних алгебраїчних рівнянь;
- методи обчислення визначників квадратних матриць;
- методи обернення матриці;
- методи обчислення власних чисел і власних векторів матриці;

### 7.2 Постановлення задачі

Розглядаємо систему з  $n$  лінійних рівнянь з  $n$  невідомими.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1; \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2; \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n; \end{cases} \quad (7.1)$$

Коротко система (7.1) може бути записана так:

$$\sum_{j=1}^n a_{ij}x_j = b_i. \quad (7.2)$$

Розв'язком системи лінійних рівнянь (7.1) є будь-яка сукупність чисел  $a_1, a_2, \dots, a_n$ , яка під час підстановки на місце невідомих  $x_1, x_2, \dots, x_n$  у рівняння цієї системи перетворює усі ці рівняння в тотожності.

**Означення 7.1** Система лінійних алгебраїчних рівнянь (СЛАР) (7.1) є **сумісною**, якщо вона має розв'язок; у протилежному разі вона є **несумісною** (або **суперечливою**).

Сумісна СЛАР може мати один або кілька розв'язків.

**Означення 7.2** СЛАР називається **визначеною**, якщо має один єдиний розв'язок, і **невизначеною**, якщо має більше одного розв'язку.

**Означення 7.3** Дві СЛАР з однаковою кількістю невідомих є **еквівалентними**, якщо вони або обидві несумісні, або сумісні й мають однаковий розв'язок.

Елементарними перетвореннями СЛАР є:

- 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.

Елементарні перетворення над будь-якою СЛАР перетворюють її в еквівалентну. Виконання елементарних перетворень рівносильне вираженню одного невідомого через інші. Система, в якій вільні члени  $b_1, b_2, \dots, b_m$  дорівнюють нулю, називається **однорідною**.

Звернемо увагу на ще одне важливе питання. У процесі вирішення різних питань техніки, економіки тощо доводиться розв'язувати СЛАР. У таких системах коефіцієнти і вільні члени є наближеними, що призводить до появи додаткових, неусувних похибок, які потрібно враховувати як у процесі обчислень, так і в остаточному округленні результату. Коефіцієнти СЛАР, які виникають під час оброблення результатів, містять помилки спостережень. Якщо

СЛАР записати в пам'ять комп'ютера навіть точно, то обчислення призводить до похибок округлення. Проте, якщо матриця системи (7.1) майже вироджена – можна сподіватися, що малі зміни в коефіцієнтах і (або) вільних членах також призведуть до значних змін у її розв'язку.

Якщо малі збурення коефіцієнтів і (або) вільних членах СЛАР дуже збурюють її розв'язок – таку систему називають погано обумовленою.

Якщо ж розв'язок збурюється незначно – СЛАР називають добре обумовленою. Ознакою поганої обумовленості СЛАР є її майже виродженість (коли значення визначника системи наближається до нуля).

Для встановлення умови сумісності системи необхідно ввести поняття матриці системи і розширеної матриці системи.

**Означення 7.4** *Матрицею системи (7.2) називається матриця, складена з коефіцієнтів за невідомих цієї системи.*

У матричному вигляді система (7.1) має вигляд

$$AX = B, \tag{7.3}$$

де  $A$  – матриця коефіцієнтів системи з  $n$  рядків і  $n$  стовпців;

$X$  –  $n$ -компонентний вектор невідомих;

$B$  –  $n$ -компонентний вектор правих частин.

Розширеною матрицею системи буде матриця

$$\hat{A} = (A|B) = \left( \begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{array} \right) \tag{7.4}$$

розміром у  $n$  рядків і  $n + 1$  стовпець.

Розв'язком системи називається такий вектор  $X^* = (x_1^*, x_2^*, \dots, x_n^*)$ , який зводить кожне рівняння (7.1) до тотожності.

Умова існування єдиного розв'язку системи за теоремою Кронекера – Капеллі.

### ТЕОРЕМА 7.1: Теорема Кронекера – Капеллі

Для того щоб система лінійних рівнянь була сумісною, необхідно і достатньо, щоб ранг її основної матриці дорівнював рангу розширеної матриці.

$$\text{rang}A = \text{rang}\hat{A} = n. \quad (7.5)$$

Ця умова еквівалентна умові нерівності 0 визначника матриці коефіцієнтів системи

$$\Delta = \det A \neq 0. \quad (7.6)$$

**Наслідок 1** Якщо система (7.2) сумісна і ранг матриці системи  $r(A)$  дорівнює кількості невідомих  $n$ , то система має єдиний розв'язок.

**Наслідок 2** Якщо система (7.2) сумісна і ранг матриці системи  $r(A)$  менший за кількість невідомих  $n$ , то система має незліченну кількість розв'язків.

Усю множину існуючих методів розв'язання СЛАР можна поділити на 2 типи:

- 1) точні (прямі) методи;
- 2) наближені (ітераційні) методи.

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 та 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, які обчислюються за точними методами, є наближеними числами з певними похибками (похибками округлень).

До точних методів розв'язання СЛАР належать: правило Крамера, метод Гаусса, метод головних елементів, Метод Гаусса – Жордана, схема Халевського, 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.

Метод Гаусса застосовують для розв'язання СЛАР із 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 розміщені рівняння системи. Максимальна точність досягається у разі, коли провідний елемент має найбільше значення. Вибір за таким правилом головного елемента і покладено в основу методу головних елементів.

У 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 разі один раз обчислюються значення елементів трикутних матриць і багаторазово розв'язуються нижня та верхня трикутні системи, які потребують виконання операцій множення і ділення. Якщо матриця коефіцієнтів системи симетрична – доцільно застосовувати метод квадратних коренів, який потребує вдвічі меншої кількості комірок пам'яті та арифметичних операцій – відповідно.

Застосування точних методів для розв'язування СЛАР із великою кількістю невідомих є громіздким. Крім того, кількість невідомих може бути настільки великою, що коефіцієнти системи не завжди можна розміщувати в оперативній пам'яті комп'ютера.

Тоді системи розв'язують за допомогою ітераційних методів. Ітераційними називають такі методи, які дозволяють знайти наближений розв'язок СЛАР із заздалегідь зазначеною 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 потрібно врахо-

увати також і похибку методу. До ітераційних методів розв'язання СЛАР належать: метод простої ітерації (або метод ітерації), метод Зейделя, метод верхньої релаксації тощо. Далі детальніше розглянемо теоретичні та практичні питання розв'язання СЛАР ітераційними методами.

### 7.3 Точні методи розв'язання СЛАР

Точні(прямі) методи розв'язання СЛАР можуть бути визначені як такі, що за умови точних вихідних даних і обчислення без округлень дають точний розв'язок СЛАР, вектор  $X^* = (x_1^*, x_2^*, \dots, x_n^*)$ , за скінчену кількість кроків. До таких методів належать, наприклад:

- правило Крамера

$$x_i^* = \frac{\Delta_i}{\Delta} \quad (7.7)$$

$\Delta_i$  визначник матриці, утвореної з матриці коефіцієнтів  $A$  заміною  $i$ -го стовпця на вектор правих частин. Вимагає виконання приблизно  $2n^4/3$  додавань/віднімань і множень/ділень;

- розв'язання матричного рівняння

$$X = A^{-1}B \quad (7.8)$$

де  $A^{-1}$  – обернена матриця до матриці коефіцієнтів  $A$ :

$$AA^{-1} = A^{-1}A = E,$$

$$A^{-1} = \frac{1}{\Delta}A' = \frac{1}{\Delta} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix} \quad (7.9)$$

$A_{ij}$  – алгебраїчні доповнення до відповідних елементів  $a_{ij}$ ,  $A'$  – приєднана матриця матриці  $A$ .

Обидва наведені методи вміщують у собі велику кількість операцій для обчислення визначників  $n - \text{го}$  та  $n - 1 - \text{го}$  порядків. Обчислення визначника  $n - \text{го}$  порядку вимагає виконати додавання  $n!$  доданків, кожний з яких є добутком  $n$  множників.

Отже, кількість множень і додавань для обчислення одного визначника  $n$ -го порядку дорівнює  $n \cdot n!$ . Вже система навчального рівня  $10 \times 10$  вимагає 36288000 операцій для одного визначника. Для правила Крамера таких обчислень потрібно зробити 11, а для оберненої матриці – 1 визначник  $10 \times 10$  і 100 визначників  $9 \times 9$ .

Отже, не дивлячись на сучасні потужності комп'ютерів, для великих систем ці методи є нераціональними;

- методи виключення змінних (методи Гауса та Жордана - Гауса);
- наближені (ітераційні) методи визначають як такі, що розв'язок СЛАР (7.1) знаходять, як послідовність наближень

$$\{X^{(i)} : \lim_{i \rightarrow \infty} X^{(i)} = X^*\}.$$

У реальному обчислювальному процесі обчислення розв'язку припиняється за умови  $|X^{(i)} - X^*| < \varepsilon$ , де  $\varepsilon$  – задана точність обчислень. До таких методів належать, наприклад, **метод простих ітерацій, метод Зейделя.**

## Метод Гауса

Розглянемо метод Гауса більш детально.

Розглядаємо СЛАР, задану своєю розширеною матрицею 7.4 або у вигляді таблиці 7.1

Таблиця 7.1– Табличний запис розширеної матриці  $A$

	$x_1$	$x_2$	$\dots$	$x_n$	$B$	$KC$
1	$a_{11}$	$a_{12}$	$\dots$	$a_{1n}$	$b_1$	$\sum_1$
2	$a_{21}$	$a_{22}$	$\dots$	$a_{2n}$	$b_2$	$\sum_2$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$n$	$a_{n1}$	$a_{n2}$	$\dots$	$a_{nn}$	$b_n$	$\sum_n$

Вважаємо, що ця система визначена, тобто  $\det A \neq 0$ .

У цій таблиці стовбець із номером  $n + 2$  є контролвальним. У цей стовбець записується сума усіх елементів відповідних рядків, тобто  $\sum_i = \sum_{j=1}^n a_{ij} + b_i$ .

Метод Гауса розв'язання СЛАР має порівняно невелику кількість операцій, приблизно  $2n^3/3$  і має досить простий алгоритм.

Складається метод Гауса з 2 частин. Метод Гауса і його модифікації ще називають методами одного ділення.

Перша частина методу Гауса є поступовим виключенням невідомих із рівнянь системи, щоб після завершення алгоритму матриця коефіцієнтів системи мала трикутний вигляд. Ця частина методу носить назву **прогонка вниз**.

У другій частині методу відбувається поступове обчислення компонентів розв'язку  $X^*$ , починаючи з  $x_n$ . Назва другої частини – **прогонка вгору**.

Враховуючи знання звичайного методу Гауса, розглянемо модифікований метод – **метод Гауса з вибором головного елемента за стовпцем** як такий, що дозволяє обійти такі неприємні моменти розв'язання СЛАР, як наявність на діагоналі матриці коефіцієнтів нульового або досить малого коефіцієнтів.

### **Прогонка вниз (пряма прогонка).**

1 Аналізується перший стовбець матриці коефіцієнтів. Знаходимо найбільший елемент  $\max_{i=1,n} |a_{i1}|$ . Переставляємо увесь рядок з обраним елементом на перше місце.

*Результат дії* – елемент  $a_{11}$  матриці коефіцієнтів – найбільший за модулем з усіх коефіцієнтів біля  $x_1$ .

Визначаємо перший рядок, перший стовбець і елемент  $a_{11}$  як головні.

2 Ділимо усі елементи головного рядка (включаючи  $n + 1$  – й та  $n + 2$  – й елементи) на головний  $a_{11}$ :

$$a_{1j}^{(1)} = \frac{a_{1j}}{a_{11}}; b_1^{(1)} = \frac{b_1}{a_{11}}; \sum_1^{(1)} = \frac{\sum_1}{a_{11}}. \quad (7.10)$$

3 Виключаємо змінну  $x_1$  з усіх рівнянь, окрім першого. Для цього додаємо до кожного неголовного рядка з номером  $i$  таблиці головний рядок, помножений на  $-a_{i1}$ ,  $i = 2, n$ . Перерахунок елементів

$i$ -го рядка матриці відбувається за такими формулами:

$$a_{ij}^{(1)} = a_{ij} + a_{1j}^{(1)} \cdot (-a_{i1}) = a_{ij} - \frac{a_{i1}a_{1j}}{a_{11}}, \quad (7.11)$$

$$b_i^{(1)} = b_i - b_1^{(1)} \cdot a_{i1} = b_i - \frac{a_{i1}b_1}{a_{11}}, \quad (7.12)$$

$$\sum_i^{(1)} = \sum_i - \sum_1^{(1)} = \sum_i - \frac{a_{i1} \sum_1}{a_{11}}, i = 2, n; j = 1, n. \quad (7.13)$$

*Результат дії* – таблиця 7.2 В разі, якщо алгоритм працює пра-

Таблиця 7.2– Результат дії 3-го пункту алгоритму

	$x_1$	$x_2$	...	$x_n$	$B$	$KC$
1	1	$a_{12}^{(1)}$	...	$a_{1n}^{(1)}$	$b_1^{(1)}$	$\sum_1^{(1)}$
2	0	$a_{22}^{(1)}$	...	$a_{2n}^{(1)}$	$b_2^{(1)}$	$\sum_2^{(1)}$
...	...	...	...	...	...	...
n	0	$a_{n2}^{(1)}$	...	$a_{nn}^{(1)}$	$b_n^{(1)}$	$\sum_n^{(1)}$

вильно, то виконується рівність

$$\sum_i^{(1)} = \sum_{j=1}^n a_{ij}^{(1)} + b_i^{(1)}. \quad (7.14)$$

У реальній роботі чисельного алгоритму ми ніколи не будемо мати точної рівності (7.14), тому рівність (7.14) замінюють наближенням

$$\left| \sum_i^{(1)} - \left( \sum_{j=1}^n a_{ij}^{(1)} + b_i^{(1)} \right) \right| < \varepsilon, \quad (7.15)$$

де  $\varepsilon$  – задана точність розв'язання системи.

- 4 Аналізується другий стовбець матриці коефіцієнтів, **починаючи з другого елемента**.

Знаходимо найбільший елемент  $\max_{i=2,n} |a_{i2}|$ . Переставляємо увесь рядок з обраним елементом на **друге** місце.

*Результат дії* – елемент  $a_{22}$  матриці коефіцієнтів – найбільший за модулем із коефіцієнтів біля  $x_2$  у рівняннях з 2-го по  $n$ -е.

Визначаємо другий рядок і елемент  $a_{22}$  як *головні*.

5 Розраховуємо значення коефіцієнтів еквівалентної системи, в якій невідома змінна  $x_2$  наявна лише в 1-му та 2-му рівняннях. Для розрахунку використовуємо формули (7.10), (7.11–7.13). Вихідні дані беремо з таблиці 7.2.

Складаємо результувальну таблицю, перевіряємо правильність розрахунку.

Пункти 4, 5 виконуються до  $n$ -го рядка таблиці включно кожного разу, беручи за головні рядок  $i$  та елемент  $a_{ii}$ .

*Результат дії* – трикутна матриця коефіцієнтів системи, подана в таблиці 7.3.

Таблиця 7.3– Результат дії 5-го пункту алгоритму.

	$x_1$	$x_2$	...	$x_n$	$B$	$KC$
1	1	$a_{12}^{(1)}$	...	$a_{1n}^{(1)}$	$b_1^{(1)}$	$\sum_1^{(1)}$
2	0	1	...	$a_{2n}^{(2)}$	$b_2^{(2)}$	$\sum_2^{(2)}$
...	...	...	...	...	...	...
n	0	0	...	1	$b_n^{(n)}$	$\sum_n^{(n)}$

Прогонка вниз закінчена.

Далі виконуємо **прогонку вгору**.

З останнього рядка таблиці 7.3 маємо:  $x_n = b_n^{(n)}$ .

Підставляємо одержане значення в рядок  $n - 1$ , маємо:

$$x_{n-1} = b_{n-1}^{n-1} - a_{n-1,n}^{(n-1)} b_n^{(n)}.$$

Загальна формула для обчислення  $x_i$  буде:

$$x_i = b_i^{(i)} - \sum_{j=i+1}^n a_{i,j+1}^{(i)} x_{j+1}.$$

Розв'язати визначену СЛАР методом Гауса з вибором головного елемента за стовпцем

$$\begin{cases} 0.123x_1 + 0.392x_2 + 0.546x_3 - 0.249x_4 = -0.383, \\ -0.385x_1 + 0.002x_2 + 0.846x_3 - 0.221x_4 = -1.89, \\ 0.948x_1 - 0.35x_2 - 0.21x_3 - 0.001x_4 = 0.455, \\ -0.248x_1 + 2.5x_2 - 0.0013x_3 - 0.0003x_4 = 0.2524. \end{cases} \quad (7.16)$$

**Розв'язування**

Це СЛАР – система 4 рівнянь із чотирма невідомими. Не замислюючись над типом системи за кількістю розв'язків, застосуємо до її розв'язання метод Гауса з вибором головного елемента за стовпцем. Запишемо систему в таблицю 7.4.

Таблиця 7.4– Табличний запис системи (7.16)

	$x_1$	$x_2$	$x_3$	$x_4$	$B$	$КС$
1	0.123	0.392	0.546	-0.249	-0.383	0.43
2	-0.385	0.002	0.846	-0.221	-1.89	-1.648
3	0.948	-0.35	-0.21	-0.001	0.455	0.842
4	-0.248	0.25	-0.0013	-0.0003	0.2524	0.2528

Виконуємо прогонку униз.

- 1 Шукаємо максимальний за модулем елемент 1-го стовпця. Це елемент  $a_{31} = 0.948$ . Переміщуємо 3-й рядок на перше місце. Перші рядок і стовбець – головні.

Таблиця 7.5– Перший пункт алгоритму прогонки вниз

	$x_1$	$x_2$	$x_3$	$x_4$	$B$	$КС$
1	<b>0.948</b>	-0.35	-0.21	-0.001	0.455	0.842
2	0.123	0.392	0.546	-0.249	-0.383	0.43
3	-0.385	0.002	0.846	-0.221	-1.89	-1.648
4	-0.248	0.25	-0.0013	-0.0003	0.2524	0.2528

- 2 Ділимо перший рядок на 0.948.

3 Усі інші елементи таблиці обчислюємо за (7.11) та (7.13). Одержимо таблицю 7.6.

У таблиці додатковий стовпець – стовпець сум коефіцієнтів кожного рядка. Стовпець, що стоїть передостаннім, є стовпець контрольної суми КС, які перераховувалися за (7.11) і (7.13), як і всі інші елементи матриці коефіцієнтів. Збіг двох останніх стовпців показує нормальну роботу алгоритму методу.

Таблиця 7.6– Третій пункт алгоритму прогонки вниз

	$x_1$	$x_2$	$x_3$	$x_4$	$B$	КС обч.	КС
1	1	-0.3691	-0.2215	-0.00105	0.47995	0.88818	0.88818
2	0	<b>0.4374</b>	0.5732	-0.2478	-0.4420	0.32075	0.32075
3	0	-0.1401	0.7607	-0.22141	-1.70522	-1.30605	-1.30605
4	0	0.1584	-0.0562	-0.00056	0.37143	0.47307	0.47307

4 Обираємо найбільший за модулем елемент у стовпці № 2 з елементів 2–4. Це буде  $a_{22}^{(1)} = 0.4374$ . Другі рядок і стовпець – головні.

5 Ділимо 2-й рядок на  $a_{22}^{(2)}$ .

6 Перераховуємо елементи 3-го і 4-го рядків згідно (7.11) і (7.13). Результат запишемо в таблицю 7.7

Таблиця 7.7– Шостий пункт алгоритму прогонки вниз

	$x_1$	$x_2$	$x_3$	$x_4$	$B$	КС обч.	КС
1	1	-0.3691	-0.2215	-0.00105	0.47995	0.88818	0.88818
2	0	1	1.31054	-0.56668	-1.01057	0.73329	0.73329
3	0	0	<b>0.94437</b>	-0.30082	-1.84684	-1.20328	-1.20328
4	0	0	-0.26388	0.08922	0.53154	0.35688	0.38688

7 У третьому стовпці найбільшим за модулем буде  $a_{33}^{(2)} = 0.94437$ .

8 Третій рядок і стовпець – головні. Ділимо 3-й рядок на  $a_{33}^{(2)}$ .

9 Перераховуємо елементи і 4-го рядків згідно (7.11) та (7.13). Результат запишемо в таблицю 7.8.

10 Ділимо останній рядок на  $a_{44}^{(3)} = 0.005166$ . Результат запишемо в таблицю 7.9.



Таблиця 7.8– Дев'ятий пункт алгоритму прогонки вниз

	$x_1$	$x_2$	$x_3$	$x_4$	$B$	$КС$ обч.	$КС$
1	1	-0.3691	-0.2215	-0.00105	0.47995	0.88818	0.88818
2	0	1	1.31054	-0.56668	-1.01057	00.73329	0.73329
3	0	0	1	-0.31854	-1.95562	-1.27416	-1.27416
4	0	0	0	0.005166	0.015499	0.02066	0.02066

Таблиця 7.9– Десятий пункт алгоритму прогонки вниз

	$x_1$	$x_2$	$x_3$	$x_4$	$B$	$КС$ обч.	$КС$
1	1	-0.3691	-0.2215	-0.00105	0.47995	0.88818	0.88818
2	0	1	1.31054	-0.56668	-1.01057	00.73329	0.73329
3	0	0	1	-0.31854	-1.95562	-1.27416	-1.27416
4	0	0	0	1	3	4	4

Матриця коефіцієнтів системи набрала трикутного вигляду.

Прогонка вниз виконана.

**Прогонка вгору (зворотний хід алгоритму):**

- з четвертого рядка останньої таблиці маємо  $x_4 = 3$ ;
- з третього рядка  $x_3 = -1.95562 + 0.31854 \cdot 3 = -1$ ;
- з 2-го рядка  $x_2 = -1.01057 - 1.31054 \cdot (-1) + 0.56668 \cdot 3 = 2$ ;
- з 1-го рядка  $x_1 = 0.479958 + 0.369198 \cdot 2 + 0.22152 \cdot (-1) + 0.00150 \cdot 3 = 1$ .

Одержали розв'язок  $X(1; 2; -1; 3)$ . ■

### Інші модифікації методу Гауса

Окрім метода Гауса і метода Гауса з вибором головного елемента за стовпцем таку саму складність мають метод Гауса з вибором головного елемента матриці коефіцієнтів СЛАР і метод Жордана – Гауса.

**Метод головного елемента матриці** коефіцієнтів СЛАР полягає у виборі на кожному кроці найбільшого елемента матриці коефіцієнтів і призначення цього елемента, його рядка і стовпця за головні.

Ідея цього методу виникла у зв'язку з тим, що коефіцієнти СЛАР є параметрами реальних інженерних систем і здебільшого є наближеними значеннями, тому що одержані звичайно внаслідок вимірювання або як статистичні дані.

Проілюструємо цей метод на довільній СЛАР  $4 \times 4$ .  
Розглянемо систему

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 = b_1; \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 = b_2; \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 = b_3; \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 + a_{44}x_4 = b_4; \end{cases}$$

Вважаємо, що система визначена, тобто  $\det A \neq 0$ .

Запишемо її у табличному вигляді. Знаходимо максимальний за

Таблиця 7.10– Табличний запис матриці  $A$

	$x_1$	$x_2$	$x_3$	$x_4$	$B$	$КС$
1	$a_{11}$	$a_{12}$	$a_{13}$	$a_{14}$	$b_1$	$\sum_1$
2	$a_{21}$	$a_{22}$	$a_{23}$	$a_{24}$	$b_2$	$\sum_2$
3	$a_{31}$	$a_{32}$	$a_{33}$	$a_{34}$	$b_3$	$\sum_3$
4	$a_{41}$	$a_{42}$	$a_{43}$	$a_{44}$	$b_4$	$\sum_4$

модулем елемент матриці коефіцієнтів. Нехай це буде елемент  $a_{34}$ .  
Тоді головним рядком буде рядок 3, а стовпцем – стовпець 4.

Ділимо головний рядок № 3 на головний елемент  $a_{34}$

$$a_{3j}^{(1)} = \frac{a_{3j}}{a_{34}}; b_3^{(1)} = \frac{b_3}{a_{34}}; \sum_3^{(1)} = \frac{\sum_3}{a_{34}}. \quad (7.17)$$

Кожний елемент неголовного рядка  $i$  перераховуємо за формулою

$$a_{ij}^{(1)} = a_{ij} - \frac{a_{3j}}{a_{34}}; b_i^{(1)} = b_i - \frac{b_3}{a_{34}}a_{i4}; \sum_i^{(1)} = \sum_i - \frac{\sum_3}{a_{34}}a_{i4}. \quad (7.18)$$

Легко помітити, в перерахуванні беруть участь елементи, які створюють прямокутник (таблиця 7.10, позначені зеленим кольором).

Наприклад, для перерахунку  $a_{12}$  потрібно взяти елементи  $a_{12}$ ,  $a_{14}$ ,  $a_{34}$ ,  $a_{32}$ . Тоді формулу перерахунку можна подати так:

$$a_{12}^{(1)} = \frac{a_{12}a_{34} - a_{14}a_{32}}{a_{34}},$$

яка повністю відповідає загальній формулі за змістом і відображає процес перерахунку як різницю добутків елементів, які стоять на

діагоналях прямокутника. Причому перший добуток **завжди** має за множника елемент, який перераховується.

Тобто усі елементи неголовних рядків перераховуються за правилом прямокутника, яке відповідає (7.18).

Одержали матрицю коефіцієнтів системи, еквівалентної вихідній. У цій матриці 4-й стовпець буде складатися з 1 на місці (3, 4) і нулів на усіх інших місцях.

Таблиця 7.11– Запис матриці  $A$  після перерахунку елемента  $a_{34}$

	$x_1$	$x_2$	$x_3$	$x_4$	$B$	$KC$
1	$a_{11}^{(1)}$	$a_{12}^{(1)}$	$a_{13}^{(1)}$	0	$b_1^{(1)}$	$\sum_1^{(1)}$
2	$a_{21}^{(1)}$	$a_{22}^{(1)}$	$a_{23}^{(1)}$	0	$b_2^{(1)}$	$\sum_2^{(1)}$
3	$a_{31}^{(1)}$	$a_{32}^{(1)}$	$a_{33}^{(1)}$	1	$b_3^{(1)}$	$\sum_3^{(1)}$
4	$a_{41}^{(1)}$	$a_{42}^{(1)}$	$a_{43}^{(1)}$	0	$b_4^{(1)}$	$\sum_4^{(1)}$

Запам'ятовуємо головний рядок в окрему матрицю і викидаємо його і головний стовпець із подальшого процесу. Позначимо матрицю з головних рядків через  $U$ :

$$U = \begin{pmatrix} a_{31}^{(1)} & a_{32}^{(1)} & a_{33}^{(1)} & 1 & b_3^{(1)} & \sum_3^{(1)} \end{pmatrix}.$$

Розрахункова таблиця набере вигляду таблиці 7.12.

Таблиця 7.12– Розрахункова таблиця

	$x_1$	$x_2$	$x_3$	$B$	$KC$
1	$a_{11}^{(1)}$	$a_{12}^{(1)}$	$a_{13}^{(1)}$	$b_1^{(1)}$	$\sum_1^{(1)}$
2	$a_{21}^{(1)}$	$a_{22}^{(1)}$	$a_{23}^{(1)}$	$b_2^{(1)}$	$\sum_2^{(1)}$
3	$a_{41}^{(1)}$	$a_{42}^{(1)}$	$a_{43}^{(1)}$	$b_4^{(1)}$	$\sum_4^{(1)}$

Знову обираємо найбільший за модулем елемент матриці коефіцієнтів. Нехай це буде  $a_{41}^{(1)}$ . Головним рядком є рядок № 3 таблиці, головним стовпцем є стовпець № 1 таблиці 7.12.

Перераховуємо коефіцієнти відповідно до наведених вище загальних формул. Одержимо нову таблицю 7.13

Таблиця 7.13– Перерахунок коефіцієнтів

	$x_1$	$x_2$	$x_3$	$B$	$KC$
1	0	$a_{12}^{(2)}$	$a_{13}^{(2)}$	$b_1^{(2)}$	$\sum_1^{(2)}$
2	0	$a_{22}^{(2)}$	$a_{23}^{(2)}$	$b_2^{(2)}$	$\sum_2^{(2)}$
3	1	$a_{42}^{(2)}$	$a_{43}^{(2)}$	$b_4^{(2)}$	$\sum_4^{(2)}$

Заносимо головний рядок у матрицю  $U$  і видаляємо головний рядок і стовпець із таблиці

$$U = \begin{pmatrix} a_{31}^{(1)} & a_{32}^{(1)} & a_{31}^{(1)} & 1 & b_3^{(1)} & \sum_3^{(1)} \\ 1 & a_{42}^{(2)} & a_{43}^{(2)} & 0 & b_4^{(2)} & \sum_4^{(2)} \end{pmatrix}.$$

Таблиця набере вигляду:

	$x_2$	$x_3$	$B$	$KC$
1	$a_{12}^{(2)}$	$a_{13}^{(2)}$	$b_1^{(2)}$	$\sum_1^{(2)}$
2	$a_{22}^{(2)}$	$a_{23}^{(2)}$	$b_2^{(2)}$	$\sum_2^{(2)}$

Черговий максимальний за модулем елемент призначимо на  $a_{23}^{(2)}$ .  
Головний рядок – 2, стовпець – 3.

Перерахована за формулами (7.17) і (7.18) таблиця набирає вигляду 7.14

Таблиця 7.14– Розрахункова таблиця після перерахунку за формулами (7.17) і (7.18)

	$x_2$	$x_3$	$B$	$KC$
1	$a_{12}^{(3)}$	0	$b_1^{(3)}$	$\sum_1^{(3)}$
2	$a_{22}^{(3)}$	1	$b_2^{(3)}$	$\sum_2^{(3)}$

Поповнюємо  $U$ :

$$U = \begin{pmatrix} a_{31}^{(1)} & a_{32}^{(1)} & a_{31}^{(1)} & 1 & b_3^{(1)} & \sum_3^{(1)} \\ 1 & a_{42}^{(2)} & a_{43}^{(2)} & 0 & b_4^{(2)} & \sum_4^{(2)} \\ 0 & a_{22}^{(3)} & 1 & 0 & b_2^{(3)} & \sum_2^{(3)} \end{pmatrix}.$$

Видаляємо головні рядок і стовпець:

	$x_2$	$B$	$КС$
1	$a_{12}^{(3)}$	$b_1^{(3)}$	$\sum_1^{(3)}$

Поповнюємо  $U$ :

$$U = \begin{pmatrix} a_{31}^{(1)} & a_{32}^{(1)} & a_{31}^{(1)} & 1 & b_3^{(1)} & \sum_3^{(1)} \\ 1 & a_{42}^{(2)} & a_{43}^{(2)} & 0 & b_4^{(2)} & \sum_4^{(2)} \\ 0 & a_{22}^{(3)} & 1 & 0 & b_2^{(3)} & \sum_2^{(3)} \\ 0 & 1 & 0 & 0 & b_1^{(4)} & \sum_1^{(4)} \end{pmatrix}.$$

Матриця  $U$ , складена з головних рядків, має структуру трикутної матриці. В разі перестановки рядків і стовпців і перенумерації змінних матриця набере явно трикутного вигляду. Тобто ми виконали прямий хід методу виключення змінних, тобто **прогонку вниз**.

Останній рядок матриці  $U$  відповідає рівнянню  $x_2 = b_1^{(4)}$ .

Підставляємо значення  $x_2$  в рівняння, яке відповідає рядку 3:

$$x_3 = b_2^{(3)} - a_{22}^{(3)} x_2.$$

З другого рядка

$$x_1 = b_2^{(2)} - a_{42}^{(2)} x_2 - a_{43}^{(2)} x_3.$$

З першого рядка

$$x_4 = b_3^{(2)} - a_{31}^{(1)} x_1 - a_{32}^{(1)} x_2 - a_{33}^{(1)} x_3.$$

Розв'язок системи –  $X = (x_1; x_2; x_3; x_4)$ .

**Зауваження!** Звертаємо увагу на те, що в разі невиродженості матриці коефіцієнтів ненульовий елемент знайдеться на кожному кроці. В іншому випадку  $\det A = 0$  і система не є визначеною. В цьому разі необхідно зробити аналіз стовпця правих частин системи. Якщо хоча б проти одного нульового рядка матриці коефіцієнтів стоїть ненульове значення правої частини – СЛАР несумісна, якщо усім нульовим рядкам відповідають нульові значення правих частин – система невизначена, має базисні та вільні змінні, матриця коефіцієнтів базисних змінних набрала трикутного вигляду, та частина матриці коефіцієнтів,

які залишилися не зведеними до трикутної матриці, є коефіцієнтами вільних змінних. Отже, з останньої таблиці можна виписати загальний розв'язок невизначеної системи.

Якщо зробити порівняльний аналіз методу головного елемента стовпця та методу головного елемента матриці, то можна помітити, що перший метод є частинним випадком другого за умови, що найбільший елемент завжди буде в крайньому лівому стовпці таблиці (матриці), яка підлягає перетворенню на черговому кроці методу.

**Метод Гауса – Жордана (метод повного виключення)** відрізняється від методу Гауса тим, що шляхом еквівалентних перетворень квадратна матриця коефіцієнтів СЛАР зводиться до одичної. Це означає, що вихідна СЛАР зводиться до системи, в якій кожне рівняння визначає лише одну змінну.

Алгоритм методу Жордана – Гауса (з вибором головного елемента стовпця) можна подати так:

- 1 П. 1 – п. 2 повністю збігаються з алгоритмом методу Гауса.
- 2 Цей пункт можна подати так: перший (головний) стовець, починаючи з 2-го елемента заповнюємо «0». Всі елементи неголовних рядків таблиці розраховуються за методом прямокутника (7.18).
- 3 Перевіряємо стовець КС.
- 4 Обираємо найбільший за модулем елемент у 2-му стовпці серед елементів із номерами 2, 3, ...,  $n$ . Пересуваємо рядок з максимальним елементом на друге місце.
- 5 Елемент  $a_{22}^{(1)}$ , 2-й рядок і 2-й стовець – головні. Ділимо головний рядок на головний елемент – формула (7.17).
- 6 Головний стовець заповнюємо «0» всюди, окрім  $a_{22}^{(2)} = 1$ . Усі елементи неголовних рядків розраховуємо за (7.18).
- 7 Перевіряємо стовець КС.

**Для довільного  $i$ -го рядка:**

- 8 Обираємо найбільший за модулем елемент у  $i$ -му стовпці серед елементів із номерами  $i, i + 1, \dots, n$ . Пересуваємо рядок із максимальним за модулем елементом на  $i$ -те місце.

- 9 Елемент  $a_{ii}^{(i-1)}$ ,  $i$ -й рядок та  $i$ -й стовпець – головні. Ділимо головний рядок на головний елемент – формула (7.17).
- 10 Головний стовпець заповнюємо «0» всюди, окрім  $a_{ii}^{(i)} = 1$ . Усі елементи неголовних рядків розраховуємо за (7.18).
- 11 Перевіряємо стовпець КС.
- 12 Після виконання  $4n$  кроків алгоритму одержуємо результат у вигляді таблиці з  $n$  рядками та  $n + 2$ -ма стовпцями. В  $n + 1$  – му стовпці стоять компоненти розв'язку системи, у  $n + 2$ -му – контрольна сума.

Ще однією модифікацією методу Гауса є **метод прогонки**, який застосовується для трьох діагональних матриць.

## Лекція 8

### Прямі методи розв'язання СЛАР. LU-розкладання квадратної невинродженої матриці, алгоритм Краута, алгоритм Холецкого

#### Прямі методи розв'язання СЛАР

##### 8.1 LU-розкладання квадратної невинродженої матриці

Здебільшого неможливо одержати рішення диференціальних рівнянь у частинних похідних в аналітичному вигляді за допомогою елементарних або спеціальних функцій. Тому важливо чисельне розв'язання диференціальних рівнянь у частинних похідних, яке дає можливість одержати розв'язок у вигляді значень шуканої функції у визначених точках координат для визначених моментів часу (часових прошарків). Найчастіше використовують для цього метод сіток, або метод скінченних різниць.

Дуже широке застосування в чисельних схемах задач лінійної алгебри з невинродженими квадратними матрицями  $A$  має розкладання такої матриці на 2 трикутні: нижню трикутну матрицю  $L$  та верхню трикутну матрицю  $U$ .

Визначимо нижню трикутну матрицю  $L$  з одиничною головною діагоналлю, щоб вище діагоналі були розміщені нульові елементи, а нижче діагоналі – ненульові елементи, значення яких одержані під час приведення матриці.

Визначимо верхню трикутну матрицю  $U$ , на головній діагоналі якої розміщені ненульові та неединичні елементи, а нижче діагоналі – нульові елементи

$$A = L \cdot U,$$
$$L = \begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ \dots & \dots & \dots & \dots \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{pmatrix}; \quad U = \begin{pmatrix} 1 & u_{12} & \dots & u_{1n} \\ 0 & 1 & \dots & u_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}. \quad (8.1)$$

Таке подання  $A$  застосовується для розв'язання СЛАР, обчислення визначників, обертання квадратних матриць тощо, де  $L$  –



нижня трикутна матриця; а  $U$  – верхня трикутна матриця, усі діагональні елементи якої дорівнюють 1.

Називається таке розкладання квадратної невідродженої матриці  **$LU$ -розкладанням** і базується на еквівалентних перетвореннях методу Гауса.

Нагадаємо, що внаслідок прямого ходу методу Гауса ми одержуємо верхню трикутну матрицю

$$U = \begin{pmatrix} 1 & u_{12} & \dots & u_{1n} \\ 0 & 1 & \dots & u_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix} = \begin{pmatrix} 1 & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & 1 & \dots & a_{2n}^{(1)} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}. \quad (8.2)$$

**Зауваження!** Головна діагональ матриці  $L$  складається з діагональних елементів матриці Гаусовських перетворень, тобто

$$l_{ii} = a_{ii}^{(i-1)}, i = 1, n. \quad (8.3)$$

Існує декілька алгоритмів розкладання. Розглянемо 2 з них.

## 8.2 Алгоритм Краута

Основна ідея методу Краута, або  $LU$ -розкладання полягає в тому, що це своєрідний перезапис методу Гауса. Суть у цьому, що можна явно виділити два етапи, зокрема один робить перетворення над матрицею  $A$  системи, інший – з вектором правих елементів  $b$ .

Для полегшення узагальнення алгоритму на матрицю  $n \times n$  розглянемо частинний випадок для невідродженої  $A_{4 \times 4}$ .

Нехай розкладання існує, тобто

$$L = \begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{pmatrix}; \quad U = \begin{pmatrix} 1 & u_{12} & u_{13} & u_{14} \\ 0 & 1 & u_{23} & u_{24} \\ 0 & 0 & 1 & u_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Тоді

$$\begin{aligned} A &= \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} = LU = \\ &= \begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{pmatrix} \cdot \begin{pmatrix} 1 & u_{12} & u_{13} & u_{14} \\ 0 & 1 & u_{23} & u_{24} \\ 0 & 0 & 1 & u_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} = \\ &= \begin{pmatrix} l_{11} & l_{11}u_{12} & l_{11}u_{13} & l_{11}u_{14} \\ l_{21} & l_{21}u_{12} + l_{22} & l_{21}u_{13} + l_{22}u_{23} & l_{21}u_{14} + l_{22}u_{24} \\ l_{31} & l_{31}u_{12} + l_{32} & l_{31}u_{13} + l_{32}u_{23} + l_{33} & l_{31}u_{14} + l_{32}u_{24} + l_{33}u_{34} \\ l_{41} & l_{41}u_{12} + l_{42} & l_{41}u_{13} + l_{42}u_{23} + l_{43} & l_{41}u_{14} + l_{42}u_{24} + l_{43}u_{34} + l_{44} \end{pmatrix}. \end{aligned}$$

1 Якщо зіставити початкове значення  $A$  і результат добутку, то можна визначити, що

**1-й стовпець матриці  $L$  дорівнює 1-му стовпцю  $A$ , тобто**

$$l_{i1} = a_{i1}, i = 1, 4(i = 1, n).$$

2 Порівняння **перших рядків** дає нам співвідношення

$$a_{1j} = l_{11}u_{1j}, j = 2, 3, 4; (j = 2, n).$$

З цього співвідношення знаходимо  $u_{1j}$  – **перший рядок  $U$**

$$u_{ij} = \frac{a_{1j}}{l_{11}}, j = 2, 3, 4; (j = 2, n).$$

3 З порівняння **других стовпців** маємо співвідношення

$$a_{i2} = l_{i1}u_{12} + l_{i2}, i = 2, 3, 4; (i = 2, n).$$

Отже, можемо знайти **елементи 2-го стовпця  $L$** :

$$l_{i2} = a_{i2} - l_{i1}u_{12}, i = 2, 3, 4; (i = 2, n).$$

4 Прирівнюючи 3-ті рядки, починаючи з третіх елементів, одержимо співвідношення:

$$a_{2j} = l_{21}u_{1j} + l_{22}u_{2j}, j = 3, 4; (j = 3, n).$$

звідси визначаються **елементи другого рядка  $U$** :

$$u_{2j} = \frac{1}{l_{22}}(a_{2j} - l_{21}u_{1j}), j = 3, 4; (j = 3, n).$$

Так, прирівнюючи по черзі спочатку стовпці, потім рядки вихідного вигляду  $A$  і добутку  $LU$ , одержуємо всі елементи трикутних матриць.

Загальні формули побудови елементів трикутних матриць розкладання  $A = LU$  такі:

$$\begin{aligned} l_{i1} &= a_{i1}; \\ u_{1j} &= \frac{a_{1j}}{l_{11}}; \\ l_{ij} &= a_{ij} - \sum_{k=1}^{j-1} l_{ik}u_{kj}, i \geq j, \\ u_{ij} &= \frac{1}{l_{ii}} \left( a_{ij} - \sum_{k=1}^{i-1} l_{ik}u_{kj} \right), i < j. \end{aligned} \quad (8.4)$$



## ПРИКЛАД 1

Виконати  $LU$ -розкладання матриці  $A$  за методом Краута.

$$A = \begin{pmatrix} 10 & -7 & 0 \\ -3 & 6 & 2 \\ 5 & -1 & 5 \end{pmatrix}.$$

### Розв'язування

Загальний вигляд трикутних матриць розкладання

$$L = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix}; \quad U = \begin{pmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{pmatrix}.$$

Застосовуємо формули алгоритму Краута (8.4) для заповнення трикутних матриць значеннями для вихідних даних.

1 Для першого стовпця  $L$  беремо перший стовпець  $A$

$$L = \begin{pmatrix} 10 & 0 & 0 \\ -3 & l_{22} & 0 \\ 5 & l_{32} & l_{33} \end{pmatrix}.$$

2 Обчислюємо перший рядок  $U$ :

$$u_{12} = \frac{a_{12}}{l_{11}} = \frac{-7}{10} = -0.7; u_{13} = \frac{a_{13}}{l_{11}} = \frac{0}{10} = 0.$$

Отже,

$$U = \begin{pmatrix} 1 & -0.7 & 0 \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{pmatrix}.$$

3 Використовуючи знайдені значення, розрахуємо другий стовпець  $L$ :

$$\begin{aligned} l_{i2} &= a_{ij} - \sum_{k=1}^{j-1} l_{ik}u_{kj} = a_{i2} - \sum_{k=1}^1 l_{ik}u_{k2}, i = 2, 3 (i \geq j) \Rightarrow \\ l_{22} &= a_{22} - l_{21}u_{12} = 6 - (-3) \cdot (-0.7) = 6 - 2.1 = 3.9; \\ l_{32} &= a_{32} - l_{31}u_{12} = -1 - 5 \cdot (-0.7) = -1 + 3.5 = 2.5. \end{aligned}$$

Маємо

$$L = \begin{pmatrix} 10 & 0 & 0 \\ -3 & 3.9 & 0 \\ 5 & 2.5 & l_{33} \end{pmatrix}.$$

4 Розраховуємо елементи 2-го рядка  $U$ , тобто елемент  $u_{23}$ :

$$u_{23} = \frac{1}{l_{ii}} \left( a_{ij} - \sum_{k=1}^{i-1} l_{ik}u_{kj} \right) = \frac{1}{l_{22}} (a_{23} - l_{21}u_{12}) = \frac{1}{3.9} (2 - (-3) \cdot 0) = \frac{2}{3.9} \approx 0.513.$$

Отже,

$$U = \begin{pmatrix} 1 & -0.7 & 0 \\ 0 & 1 & 0.513 \\ 0 & 0 & 1 \end{pmatrix}.$$

5 Розраховуємо третій стовпець  $L$ , тобто  $l_{33}$ :

$$l_{33} = a_{33} - \sum_{k=1}^2 l_{3k}u_{k3} = a_{33} - l_{31}u_{13} - l_{32}u_{23} = 5 - 5 \cdot 0 - 2.5 \cdot 0.513 \approx 3.718.$$

Маємо

$$L = \begin{pmatrix} 10 & 0 & 0 \\ -3 & 3.9 & 0 \\ 5 & 2.5 & 3.718 \end{pmatrix}.$$

Одержали дві трикутні матриці

$$L = \begin{pmatrix} 10 & 0 & 0 \\ -3 & 3.9 & 0 \\ 5 & 2.5 & 3.718 \end{pmatrix}; U = \begin{pmatrix} 1 & -0.7 & 0 \\ 0 & 1 & 0.513 \\ 0 & 0 & 1 \end{pmatrix},$$

добуток яких є матриця  $A$ .

*Перевірка:*

$$A = LU = \begin{pmatrix} 10 & 0 & 0 \\ -3 & 3.9 & 0 \\ 5 & 2.5 & 3.718 \end{pmatrix} \cdot \begin{pmatrix} 1 & -0.7 & 0 \\ 0 & 1 & 0.513 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 10 & -7 & 0 \\ -3 & 6 & 2.0007 \\ 5 & -1 & 5.0005 \end{pmatrix}.$$

Елементи вихідної  $A$  і розрахованої  $A$  як добуток трикутних матриць збіглися з точністю до 0.001, тобто з точністю до округлень. ■

### 8.3 Алгоритм Холецького

Алгоритм Холецького застосовують для симетричних знакопозитивних невідроджених матриць  $A$ . Інша назва методу – **метод квадратного кореня**.

Симетричну знакопозитивну невідроджену матрицю можна розкласти на дві трикутні матриці, які є транспонованими одна відносно одній. Тобто

$$A = LL^T.$$

Фактично, достатньо обчислити лише трикутну матрицю  $L$ . Іншу матрицю одержуємо транспонуванням. Для доступності перетворень розглянемо знову матрицю  $A_{4 \times 4}$ .

Розглядається матриця  $A$  – симетрична знакопозитивна матриця.

$$A = \begin{pmatrix} a_{11} & a_{21} & a_{31} & a_{41} \\ a_{21} & a_{22} & a_{32} & a_{42} \\ a_{31} & a_{32} & a_{33} & a_{43} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}.$$

Вважаємо, що існують матриці

$$L = \begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{pmatrix}; \quad L^T = \begin{pmatrix} l_{11} & l_{21} & l_{31} & l_{41} \\ 0 & l_{22} & l_{32} & l_{42} \\ 0 & 0 & l_{33} & l_{43} \\ 0 & 0 & 0 & l_{44} \end{pmatrix}.$$

такі, що  $A = LL^T$ , тобто

$$A = \begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{pmatrix}; \quad L^T = \begin{pmatrix} l_{11} & l_{21} & l_{31} & l_{41} \\ 0 & l_{22} & l_{32} & l_{42} \\ 0 & 0 & l_{33} & l_{43} \\ 0 & 0 & 0 & l_{44} \end{pmatrix} =$$

$$= \begin{pmatrix} l_{11}^2 & l_{11}l_{21} & l_{11}l_{31} & l_{11}l_{41} \\ l_{11}l_{21} & l_{21}^2 + l_{22}^2 & l_{21}l_{31} + l_{22}l_{32} & l_{21}l_{41} + l_{22}l_{42} \\ l_{11}l_{31} & l_{21}l_{31} + l_{22}l_{32} & \sum_{i=1}^3 l_{3i}^2 & \sum_{i=1}^3 l_{3i}l_{4i} \\ l_{11}l_{41} & l_{21}l_{41} + l_{22}l_{42} & \sum_{i=1}^3 l_{3i}l_{4i} & \sum_{i=1}^3 l_{4i}^2 \end{pmatrix}.$$

Тоді можна розрахувати  $l_{ij}, i \geq j$ :

1 Розраховуємо лівий верхній елемент

$$l_{11}^2 = a_{11} \Rightarrow l_{11} = \sqrt{a_{11}}.$$

2 Розраховуємо 1-й стовпець  $L$ :

$$a_{i1} = l_{11}l_{i1} \Rightarrow l_{i1} = \frac{a_{i1}}{l_{11}}, i = 2, 3, 4.$$

3 Розраховуємо діагональний елемент 2-го стовпця  $l_{22}$ :

$$a_{22} = l_{21}^2 + l_{22}^2 \Rightarrow l_{22} = \sqrt{a_{22} - l_{21}^2}.$$

4 Розраховуємо інші піддіагональні елементи 2-го стовпця:

$$a_{32} = l_{21}l_{31} + l_{22}l_{32} \Rightarrow l_{32} = \frac{1}{l_{22}}(a_{32} - l_{21}l_{31});$$

$$a_{42} = l_{21}l_{41} + l_{22}l_{42} \Rightarrow l_{42} = \frac{1}{l_{22}}(a_{42} - l_{21}l_{41}).$$

5 Розраховуємо діагональний елемент 3-го стовпця  $l_{33}$ :

$$a_{33} = l_{31}^2 + l_{32}^2 + l_{33}^2 \Rightarrow l_{33} = \sqrt{a_{33} - l_{31}^2 - l_{32}^2}.$$

6 Розраховуємо інші піддіагональні елементи 3-го стовпця. Це елемент  $l_{43}$ :

$$a_{43} = l_{31}l_{41} + l_{32}l_{42} + l_{33}l_{43} \Rightarrow l_{43} = \frac{1}{l_{33}}(a_{43} - l_{31}l_{41} - l_{32}l_{42}).$$

7 Розрахуємо діагональний елемент 4-го стовпця  $l_{44}$ :

$$a_{44} = l_{41}^2 + l_{42}^2 + l_{43}^2 + l_{44}^2 \Rightarrow l_{44} = \sqrt{a_{44} - l_{41}^2 - l_{42}^2 - l_{43}^2}.$$

Узагальнення алгоритму може бути таким:

1 У кожному стовпці на першому кроці обчислюється діагональний елемент, потім усі інші елементи стовпця, які стоять під діагональним.

2 Формули обчислення елементів першого стовпця:

$$l_{11} = \sqrt{a_{11}}; l_{i1} = \frac{a_{i1}}{l_{11}}, i = 2, n. \quad (8.5)$$

3 Формула обчислення діагональних елементів інших стовпців

$$l_{ii} = \sqrt{a_{ii} - \sum_{j=1}^{i-1} l_{ij}^2}, i = 2, n. \quad (8.6)$$

4 Формула обчислення елементів  $j$ -го стовпця, які стоять під діагональним ( $i > j$ )

$$l_{ij} = \frac{1}{l_{jj}} \left( a_{ij} - \sum_{k=1}^{j-1} l_{jk} l_{ik} \right), j = 2, n-1; i = j+1, n. \quad (8.7)$$



## ПРИКЛАД 2

Виконати розкладання матриці  $A$  за методом Холецького

$$A = \begin{pmatrix} 9 & 6 & -3 \\ 6 & 5 & -2 \\ -3 & -2 & 10 \end{pmatrix}.$$

### Розв'язування

Матриця симетрична, діагональні мінори

$$M_1^1 = 9 > 0; M_{12}^{12} = 9 > 0; M_{123}^{123} = \Delta = 81 > 0.$$

Отже, матриця позитивно визначена. Можемо застосовувати метод Хо-лецького (метод квадратного кореня) (8.5–8.7). Будемо матрицю  $L$ :

$$l_{11} = \sqrt{a_{11}} = \sqrt{9} = 3; l_{21} = \frac{a_{21}}{l_{11}} = \frac{6}{3} = 2; l_{31} = \frac{a_{31}}{l_{11}} = \frac{-3}{3} = -1;$$

$$l_{22} = \sqrt{a_{22} - l_{21}^2} = \sqrt{5 - 4} = 1; l_{32} = \frac{1}{l_{22}}(a_{32} - l_{21}l_{31}) = -2 - 2 \cdot (-1) = 0;$$

$$l_{33} = \sqrt{a_{33} - l_{31}^2 - l_{32}^2} = \sqrt{10 - 1 - 0} = 3.$$

Одержали

$$L = \begin{pmatrix} 3 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 3 \end{pmatrix}; L^T = \begin{pmatrix} 3 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Перевірка:

$$\begin{pmatrix} 3 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 9 & 6 & -3 \\ 6 & 5 & -2 \\ -3 & -2 & 10 \end{pmatrix} = A.$$

## Застосування $LU$ -розкладання матриці для задач лінійної алгебри

**1 Обчислення визначника квадратної невідродженої матриці  $A$ .** Розглянемо розкладання (8.1) квадратної невідродженої матриці  $A$ .

Очевидно, що

$$\det A = \det(L \cdot U) = \det L \cdot \det U.$$

Зі структури матриці  $U$  випливає, що  $\det U = 1$ , отже,

$$\det A = \det L = \prod_{i=1}^n l_{ii}. \quad (8.8)$$

Якщо згадаємо зв'язок елементів  $l_{ii}$  з еквівалентними перетвореннями методу Гауса), то ще одним записом формули обчислення визначника квадратної не виродженої матриці  $A$  буде

$$\det A = \prod_{i=1}^n a_{ii}^{(i-1)}, \quad (8.9)$$

де індекс  $(i - 1)$  є індексом кроку еквівалентних перетворень Гауса.



**2 Розв'язання СЛАР.** Нехай для матриці коефіцієнтів  $A$  відоме  $LU$ -розкладання. Тоді система набере вигляду

$$L \cdot U \cdot X = B.$$

Якщо позначити  $U \cdot X = Y$ , то одержимо СЛАР  $L \cdot Y = B$  з трикутною матрицею коефіцієнтів. З цієї системи звичайною Гаусовською прогонкою вниз знаходимо  $Y$ :

$$y_1 = \frac{b_1}{l_{11}}; y_i = \frac{1}{l_{ii}} \left( b_i - \sum_{j=1}^{i-1} l_{ij} y_j \right), i = 2, n. \quad (8.10)$$

Далі розв'язуємо СЛАР  $U \cdot X = Y$ , використовуючи прогонку догори:

$$x_n = y_n; x_i = y_i - \sum_{j=i+1}^n u_{ij} y_j, i = n-1, 1.$$

**3 Обертання матриці  $A$ .** Якщо матриця  $A$  є квадратною не виродженою матрицею, то для неї існує єдина матриця  $A^{-1}$  така, що  $AA^{-1} = A^{-1}A = E$ .

Позначимо елементи матриці  $A^{-1}$  через  $x_{ij}, i = 1, n, j = 1, n$ .

Для наочності будемо знову розглядати матриці розміром  $4 \times 4$ .

Розглянемо більш докладно добуток  $AA^{-1}$ :

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \cdot \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Це матричне рівняння можна розглянути як 4 СЛАР з різними правими частинами, які складають стовпці матриці  $E$  і векторами-

	$x_{1j}$	$x_{2j}$	$x_{3j}$	$x_{4j}$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	КС
1	$a_{11}$	$a_{12}$	$a_{13}$	$a_{14}$	1	0	0	0	$\sum_1$
2	$a_{21}$	$a_{22}$	$a_{23}$	$a_{24}$	0	1	0	0	$\sum_2$
3	$a_{31}$	$a_{32}$	$a_{33}$	$a_{34}$	0	0	1	0	$\sum_3$
4	$a_{41}$	$a_{42}$	$a_{43}$	$a_{44}$	0	0	0	1	$\sum_1$

невідомими, які складають стовпці матриці  $A^{-1}$ :

$$S_1 : \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \cdot \begin{pmatrix} x_{11} \\ x_{21} \\ x_{31} \\ x_{41} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix};$$

$$S_2 : \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \cdot \begin{pmatrix} x_{12} \\ x_{22} \\ x_{32} \\ x_{42} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix};$$

$$S_3 : \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \cdot \begin{pmatrix} x_{13} \\ x_{23} \\ x_{33} \\ x_{43} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix};$$

$$S_4 : \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \cdot \begin{pmatrix} x_{14} \\ x_{24} \\ x_{34} \\ x_{44} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Для зручності перетворення чотирьох систем можна їх поєднати в одній таблиці.

У разі розкладання матриці  $A = L \cdot U$  обернена матриця

$$A^{-1} = U^{-1}L^{-1},$$

тобто можна побудувати процедуру обертання верхньої та нижньої трикутних матриць і обернену до  $A$  знайти як їх добуток.

## Лекція 9

### Ітераційні методи розв'язання СЛАР

#### 9.1 Ітераційні методи розв'язання СЛАР

У теорії чисельного розв'язання СЛАР відомо, що прямі методи розв'язання СЛАР є ефективними для систем порядку  $n \leq 10^3$ . Для систем більшого порядку Гаусовські розрахунки стають громіздкими. Одним із класів методів розв'язання СЛАР з  $10^3 < n \leq 10^6$  є клас ітераційних методів.

#### 9.2 Метод простих ітерацій

Згадаємо принцип побудови ітераційного процесу для нелінійних рівнянь.

Якщо за невідому рівняння взяти вектор  $X = (x_1, x_2, \dots, x_n)^T$  лінійного векторного простору  $R^n$ , то ітераційний процес розв'язання системи рівнянь як матричного рівняння вигляду

$$AX = B \quad (9.1)$$

може бути поданий у загальному вигляді:

$$X^{(k+1)} = F_k(X^{(0)}, X^{(1)}, \dots, X^{(k)}), \quad (9.2)$$

де функція  $F_k$  у загальному випадку може бути функцією, що залежить від вигляду матриці коефіцієнтів  $A$  і вектора правих частин  $B$ , усіх попередніх наближень  $X^{(0)}, X^{(1)}, \dots, X^{(k)}$  і від номера ітерації  $k$ .

Критерій збіжності ітераційного процесу для матричного рівняння буде виглядати так:

$$\frac{q}{1-q} \|X^{(k+1)} - X^{(k)}\| < \varepsilon, \quad (9.3)$$

де  $q$  – параметр, який визначається для певного методу,  $0 < q < 1$ .

Будемо називати ітераційний метод розв'язання СЛАР методом **першого порядку**, якщо функція  $F_k$  залежить лише від останнього наближення, тобто  $X^{(k+1)} = F_k(X^k)$  і **стаціонарним методом**,

якщо функція не залежить від  $k$ , тобто  $F_k = F$ . У цьому разі ітераційний процес набирає вигляду

$$X^{(k+1)} = F_k(X^k) \quad (9.4)$$

з критерієм зупинення (9.3).

Якщо задача полягає в розв'язанні лінійної системи рівнянь, то функція  $F(X)$  набирає лінійного вигляду та ітераційне рівняння для стаціонарного методу першого порядку буде таким:

$$X^{(k+1)} = C(X^{(k)}) + P, \quad (9.5)$$

де  $C, P$  – матриці чисел розміром  $n \times n$  та  $n \times 1$  відповідно.

У цьому разі правильними будуть оцінки, що випливають із стискуючих властивостей матричного оператора  $C$  у разі, якщо  $\|C\| \leq q < 1$ .

$$\|X^{(k+1)} - X^*\| \leq \frac{\|C\|}{1 - \|C\|} \|X^{(k+1)} - X^{(k)}\|, \quad (9.6)$$

$$\|X^{(k+1)} - X^*\| \leq \|C\|^k \|X^{(0)} - X^*\|. \quad (9.7)$$

Отже, за критерій зупинення стаціонарного методу беруть нерівність

$$\frac{\|C\|}{1 - \|C\|} \|X^{(k+1)} - X^{(k)}\| \leq \varepsilon. \quad (9.8)$$

Критерієм зупинення нестационарного процесу є нерівність

$$\|X^{(k+1)} - X^{(k)}\| \leq \varepsilon. \quad (9.9)$$

Виходячи з наведеного критерію і загальної теорії побудови збіжного стаціонарного ітераційного процесу збіжності методу ітерацій розв'язання СЛАР, необхідно виконання умови, накладеної на норму матриці

$$\|C\| \leq q < 1. \quad (9.10)$$

За норму матриці  $C$  можна взяти

$$\|C\| = \max_{j=1,n} \sum_{i=1}^n |c_{ij}| \text{ або } \|C\| = \max_{i=1,n} \sum_{j=1}^n |c_{ij}|.$$

**Висновок:** стаціонарний метод першого порядку реалізує ітераційний процес, заданий рекурентною формулою (9.5) з критерієм зупинення (9.8). Ітераційний процес збігається до точного розв'язку за умови (9.10) для будь-якого початкового значення  $X^{(0)}$ .

Розглянемо тепер більш детально побудову складових процесу (9.5) для методу простих ітерацій (методу Якобі).

**Метод простих ітерацій розв'язання СЛАР (метод Якобі).**

Розглядаємо систему з  $n$  рівнянь та  $n$  невідомих:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1; \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2; \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n. \end{cases} \quad (9.11)$$

Обмеженнями, що накладає на систему побудова ітераційного процесу є умови  $\det A \neq 0$  та  $a_{ii} \neq 0$ , тобто матриця не вироджена і діагональні елементи не дорівнюють нулю.

Якщо умови виконані, то можемо з кожного рівняння виразити відповідну компоненту вектора невідомих  $X$ .

$$\begin{cases} x_1 = -\frac{a_{12}}{a_{11}}x_2 - \dots - \frac{a_{1n}}{a_{11}}x_n + \frac{b_1}{a_{11}} \\ x_2 = -\frac{a_{21}}{a_{22}}x_1 - \dots - \frac{a_{2n}}{a_{22}}x_n + \frac{b_2}{a_{22}} \\ \dots \\ x_n = -\frac{a_{n1}}{a_{nn}}x_1 - \frac{a_{n2}}{a_{nn}}x_2 - \dots + \frac{b_n}{a_{nn}}. \end{cases} \quad (9.12)$$

Якщо ввести позначення  $c_{ij} =_{i \neq j} -\frac{a_{ij}}{a_{ii}}$ ,  $c_{ii} = 0$ ;  $p_i = \frac{b_i}{a_{ii}}$ , то цю систему можна подати матричним рівнянням

$$x = D \cdot x + C.$$

Взявши за початкове наближення довільний вектор  $X^{(0)}$  (наприклад, за початкове наближення часто беруть вектор  $P$ ), ми можемо одержати з цього матричного рівняння ітераційний процес

$$X^{(k+1)} = C(X^{(k)}) + P, k = 0, 1, 2, \dots \quad (9.13)$$

Формулу (9.16) можна використовувати для обчислення ітераційної формули, тобто значення кожного нового  $(k + 1)$ -го наближення розв'язку системи рівня обчислюватиме попереднє:

$$x^{(k+1)} = D \cdot x^k + C.$$

За початкове наближення оберемо вектор  $x^{(0)} = C$ .

Ітераційний процес вважається завершеним, якщо виконується умова

$$\frac{\|x^{(k+1)} - x^{(k)}\|}{x^{(k)}} \leq \varepsilon. \quad (9.14)$$

Достатня умова збіжності методу ітерацій: процес ітерації для системи лінійних рівнянь збігається до єдиного її розв'язку, якщо кожна норма матриці  $D$  є меншою за одиницю, тобто

$$\|D\| < 1.$$

Збіжність цього процесу буде забезпечена, якщо

$$\sum_{i=1}^n \left| \frac{a_{ij}}{a_{ii}} \right| \leq q < 1, j = 1, n \quad \text{або} \quad \sum_{j=1}^n \left| \frac{a_{ij}}{a_{ii}} \right| \leq q < 1, i = 1, n$$

(впливає з (9.6)). Також умови збіжності будуть виконані в разі, якщо

$$|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}|, i = 1, n. \quad (9.15)$$

Це означає, що СЛАР має сенс розв'язувати методом простих ітерацій у разі, коли діагональні елементи матриці коефіцієнтів системи за модулем більші, ніж сума модулів усіх інших коефіцієнтів у відповідному рядку. Вектор правих частин на збіжність впливу не має.

---

## ПРИКЛАД 1

---

Розв'язати систему лінійних алгебраїчних рівнянь методом простих ітерацій з точністю  $\varepsilon = 10^{-4}$

$$\begin{cases} -4x_1 + 0.5x_2 - 1.5x_3 - x_4 = -1.67 \\ 0.55x_1 + 6x_2 + 2.4x_3 + 1.3x_4 = 3.15 \\ 0.33x_1 - 0.52x_2 + 3x_3 + 0.21x_4 = -0.21 \\ -0.22x_1 + 0.98x_2 - 2.5x_3 + 5x_4 = 5.54 \end{cases}.$$

### Розв'язування

Запишемо систему в табличному вигляді

$x_1$	$x_2$	$x_3$	$x_4$	$B$
<b>-4</b>	0.5	-1.5	-1	-1.67
0.55	<b>6</b>	2.4	1.3	3.15
0.33	-0.52	<b>3</b>	0.21	-0.21
-0.22	0.98	-2.5	<b>5</b>	5.54

Перевіряємо критерій збіжності методу ітерацій:

$$0.5 + 1.5 + 1 = 3 < 4; \quad 0.55 + 2.4 + 1.3 = 4.25 < 6;$$

$$0.33 + 0.52 + 0.21 = 1.06 < 3; \quad 0.22 + 0.98 + 2.5 = 3.7 < 5.$$

Одержали, що в кожному рядку сума модулів недіагональних елементів менша за модуль діагонального елемента. Отже, метод Якобі буде збіжним за будь-якого наближення  $X^{(0)}$ .

Будуємо ітераційний процес згідно з (9.12), результат занесемо до таблиці (9.1). У верхній рядок будемо вносити значення невідомих з попе-

Таблиця 9.1– Ітераційний процес згідно з (9.12)

$X^{(k+1)} \backslash X^{(k)}$	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$	$P$
$x_1^{(k+1)}$	0	0.125	-0.375	-0.25	0.4175
$x_2^{(k+1)}$	-0.092	0	-0.4	-0.217	0.5256
$x_3^{(k+1)}$	-0.11	0.173	0	-0.07	-0.0697
$x_4^{(k+1)}$	0.044	-0.196	0.5	0	1.1087

редньої ітерації, в перший стовпець – розраховані значення невідомих у поточній ітерації.

За нульове наближення візьмемо вектор  $P$  (таке наближення можна одержати з міркувань, що перед нульовою ітерацією вектор розв'язків наближався вектором  $\Theta$ ).

Ітерація  $k = 1$  подана в таблиці(9.2).

Одержали вектор розв'язків першої ітерації

$$X^{(1)} = (0.2322; 0.275; -0.1021; 0.9892)^T.$$

Визначаємо відхилення розв'язків нульової та першої ітерацій:

$$\Delta^{(1)}X = (|P - X_i^{(1)}|) = (0.1853; 0.2506; 0.0324; 0.1195).$$

Таблиця 9.2– Ітерація  $k = 1$ 

$X^{(1)} \backslash X^{(0)}$	0.4175	0.5256	-0.0697	1.1087	$P$
0.2322	0	0.125	-0.375	-0.25	0.4175
0.275	-0.0917	0	-0.4	-0.217	0.5256
-0.1021	-0.11	0.173	0	-0.07	-0.0697
0.9892	0.044	-0.196	0.5	0	1.1087

Таблиця 9.3– Ітерація  $k = 2$ 

$X^{(2)} \backslash X^{(1)}$	0.2322	0.275	-0.1021	0.9892	$P$
0.2428	0	0.125	-0.375	-0.25	0.4175
0.3308	-0.0917	0	-0.4	-0.217	0.5256
-0.1168	-0.11	0.173	0	-0.07	-0.0697
1.0139	0.044	-0.196	0.5	0	1.1087

Перевіряємо точність наближення:

$$\|X^{(1)} - X^{(0)}\| = \max (0.1853; 0.2506; 0.0324; 0.1195) = 0.2606 > 10^{-4}.$$

Задана точність не задовольняється.

Переходимо до наступної ітерації.

Ітерація  $k = 2$  подана в таблиці (9.3).

Одержали вектор розв'язків другої ітерації:

$$X^{(2)} = (0.2428; 0.3308; -0.1168; 1.0139)^T.$$

Визначаємо відхилення розв'язків нульової та першої ітерацій:

$$\Delta^{(2)} X = (|P - X_i^{(2)}|) = (0.0107; 0.0558; 0.0147; 0.0247).$$

Перевіряємо точність наближення:

$$\|X^{(2)} - X^{(1)}\| = \max (0.2428; 0.3308; -0.1168; 1.0139) = 0.0558 > 10^{-4}.$$

Задана точність не задовольняється.

Виконуючи обчислення за такою схемою, одержимо, що

$$X^{(6)} = \begin{pmatrix} 0.24993 \\ 0.32996 \\ -0.11005 \\ 1.00008 \end{pmatrix}; X^{(7)} = \begin{pmatrix} 0.2499 \\ 0.33000 \\ -0.1100 \\ 0.99998 \end{pmatrix}; \Delta^{(7)} X = \begin{pmatrix} 6.33E - 05 \\ 4.66E - 05 \\ 4.1E - 05 \\ 9.29E - 05 \end{pmatrix};$$



$$\max(\Delta x_i) = 9 \cdot 10^{-5} < 10^{-4}.$$

тобто для заданої точності  $\varepsilon = 10^{-4}$  ітераційний процес збігся за 7 ітераціями. ■

Із розглянутого приклада видно, що процес розв'язання системи методом Якобі збігається не так вже й швидко.

Однією з різновидів методу простих ітерацій, який у багатьох випадках (але НЕ ЗАВЖДИ!!!) прискорює ітераційний процес, є метод Зейделя.

### 9.3 Метод Зейделя розв'язання СЛАР

Метод Зейделя так само, як і метод Якобі є стаціонарний ітераційний метод. Головна ідея методу Зейделя полягає в тому, що під час чергової ітерації  $k+1$  в ітераційному процесі (9.10), розраховуючи компоненти вектора невідомих  $X^{(k+1)}$  для обчислення  $i$ -ї компоненти вектора беруть вектор  $X = (x_1^{k+1}; x_2^{k+1}; \dots; x_{i-1}^{k+1}; x_i^k; x_{i+1}^k; \dots; x_n^k)$ , тобто

$$\begin{cases} x_1^{(k+1)} = -\frac{a_{12}}{a_{11}}x_2^k - \dots - \frac{a_{1n}}{a_{11}}x_n^k + \frac{b_1}{a_{11}} \\ x_2^{(k+1)} = -\frac{a_{21}}{a_{22}}x_1^{(k+1)} - \frac{a_{23}}{a_{22}}x_3^k \dots - \frac{a_{2n}}{a_{22}}x_n^k + \frac{b_2}{a_{22}} \\ \dots\dots\dots \\ x_n^{(k+1)} = -\frac{a_{n1}}{a_{nn}}x_1^{(k+1)} - \frac{a_{n2}}{a_{nn}}x_2^{k+1} - \dots - \frac{a_{nn-1}}{a_{nn}}x_{n-1}^{k+1} + \frac{b_n}{a_{nn}}. \end{cases} \quad (9.16)$$

Критерії збіжності і зупинення методу Зейделя збігається з критерієм методу простих ітерацій (9.10) і (9.8) відповідно.

Зазвичай метод Зейделя збігається швидше, ніж метод простих ітерацій і навіть інколи збігається для систем, для яких метод простої ітерації є розбіжним. Але не завжди. Існують приклади і протилежних ситуацій, коли для однієї СЛАР розбіжний ітераційний процес за методом Зейделя є збіжним за методом простих ітерацій.

Збіжність або розбіжність методу Зейделя й методу простих ітерацій залежить лише від вигляду матриці коефіцієнтів.

#### ПРИКЛАД 2

Розв'язати систему з прикладу 4 за методом Зейделя зі зберіганням точності наближення  $\varepsilon = 10^{-4}$ .

**Розв'язування**

Обґрунтування правомірності застосування методу Зейделя для розв'язання цієї СЛАР беремо з попереднього прикладу.

Як і в попередньому прикладі ітераційний процес (9.16) має складові

$X^{(k+1)} \backslash X^{(k)}$	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$	$P$
$x_1^{(k+1)}$	0	0.125	-0.375	-0.25	0.4175
$x_2^{(k+1)}$	-0.092	0	-0.4	-0.217	0.5256
$x_3^{(k+1)}$	-0.11	0.173	0	-0.07	-0.0697
$x_4^{(k+1)}$	0.044	-0.196	0.5	0	1.1087

За нульове наближення знову обираємо

$$X^{(0)} = P = (0.4175; 0.5256; -0.0697; 1.1087).$$

Ітерація  $k = 1$  подана в таблиці (9.5). Одержали вектор наближення до розв'язку

$$X^{(1)} = (0.2322; 0.2920; -0.1222; 1.0006).$$

Похибка

$$\Delta^{(1)} = (0.1853; 0.2336; 0.0525; 0.1081); \quad \max \Delta^{(1)} = 0.2336 > 10^{-4}.$$

Точність не задовольняється.

Ітерація  $k = 2$  подана в таблиці (9.5).

Одержали вектор наближення до розв'язку

$$X^{(2)} = (0.2497; 0.3348; -0.1092; 0.9996).$$

Похибка

$$\Delta^{(2)} = (0.0175; 0.0428; 0.0131; 0.0011); \quad \max \Delta^{(2)} = 0.0428 > 10^{-4}.$$

Точність не задовольняється. Похибка зменшилася, але точність не задовольняється.

Переходимо до наступної ітерації.

Виконуючи обчислення за такою схемою, одержимо, що

$$X^{(4)} = \begin{pmatrix} 0.24997 \\ 0.3300 \\ -0.1100 \\ 0.99997 \end{pmatrix}; X^{(5)} = \begin{pmatrix} 0.25002 \\ 0.33000 \\ -0.1100 \\ 1.0000 \end{pmatrix}; \Delta X = \begin{pmatrix} 2.4E - 05 \\ 1.4E - 05 \\ 2.0E - 06 \\ 2.99E - 06 \end{pmatrix};$$

Таблиця 9.4- Ітерація  $k = 1$ 

$X^{(1)} \backslash X^{(0)}$	0.4175	0.5256	-0.0697	1.1087	$P$
$x_1^{(1)}=0.2322$	0	0.125	-0.375	-0.25	0.4175
$X$	<b>0.2322</b>	0.5256	-0.0697	1.1087	
$x_2^{(1)}=0.2919$	-0.0917	0	-0.4	-0.217	0.5256
$X$	<b>0.2322</b>	<b>0.29197</b>	-0.0697	1.1087	
$x_3^{(1)}=-0.12224$	-0.11	0.1733	0	-0.07	-0.0697
$X$	<b>0.2322</b>	<b>0.29197</b>	<b>-0.12224</b>	1.1087	
$x_4^{(1)}=1.0005$	0.044	-0.196	0.5	0	1.1087

Таблиця 9.5- Ітерація  $k = 2$ 

$X^{(2)} \backslash X^{(1)}$	0.2322	0.2920	-0.1222	1.0006	$P$
$x_1^{(2)}=0.2497$	0	0.125	-0.375	-0.25	0.4175
$X$	<b>0.2497</b>	0.5256	-0.0697	1.1087	
$x_2^{(2)}=0.3348$	-0.0917	0	-0.4	-0.217	0.5256
$X$	<b>0.2497</b>	<b>0.3348</b>	-0.0697	1.1087	
$x_3^{(2)}=-0.1092$	-0.11	0.1733	0	-0.07	-0.0697
$X$	<b>0.2497</b>	<b>0.2248</b>	<b>-0.1092</b>	1.1087	
$x_4^{(2)}=0.9996$	0.044	-0.196	0.5	0	1.1087

$$\max_{i=1,4}(\Delta x_i) = 2.4 \cdot 10^{-5} < 10^{-4},$$

тобто метод Зейделя для заданої точності  $\varepsilon = 10^{-4}$  ітераційний збігся за 5 ітерацій на відміну від 7 методу Якобі. ■

## Лекція 10

### Чисельні методи розв'язання систем нелінійних рівнянь

#### 10.1 Поставлення задачі

Системи лінійних алгебраїчних рівнянь – це лише окремий випадок систем рівнянь. На практиці розв'язування задач моделювання приводить переважно до систем нелінійних рівнянь. Системи  $n$  лінійних рівнянь з  $n$  невідомими  $x_1, x_2, \dots, x_n$  у загальному випадку прийнято записувати в такий спосіб:

$$\begin{aligned} f_1(x_1, x_2, \dots, x_n) &= 0 \\ f_2(x_1, x_2, \dots, x_n) &= 0 \\ &\dots\dots\dots \\ f_n(x_1, x_2, \dots, x_n) &= 0, \end{aligned} \tag{10.1}$$

де  $f_1, f_2, \dots, f_n$  – будь-які функції незалежних змінних, зокрема й нелінійні щодо невідомих.

Якщо

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$$

вектор-стовпець розмірності  $n$  з елементами  $x_i$  ( $x \in R^n$ ),

$$F(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \dots \\ f_n(x) \end{pmatrix}$$

вектор-функція розмірності  $n$ , елементами якої є функції

$$f_i(x) = f_i(x_1, x_2, \dots, x_n), i = \overline{1, n}, (F : R^n \Rightarrow R^n),$$

то систему (10.1) можна записати у векторному вигляді

$$F(x) = 0. \tag{10.2}$$

Розв'язати систему (10.2) – означає знайти таке  $x^* \in R^n$ , для якого  $F(x^* = 0)$ , тобто  $f_i(x_1^*, x_2^*, \dots, x_n^*) = 0 \forall i = (\overline{1, n})$ .

Більшість математичних моделей різних процесів і явищ записуються в загальному випадку у вигляді (10.2), тому ця задача має величезне практичне значення.

Система нелінійних рівнянь може не мати розв'язків, мати єдиний розв'язок, скінченну чи нескінченну кількість розв'язків. Питання щодо кількості розв'язків повинно ухвалюватися для кожної конкретної задачі окремо.

Одним із найбільш простих алгоритмів її розв'язку є метод Ньютона.

## 10.2 Метод Ньютона

Математичним підґрунтям методу є лінеаризація функцій  $f_1, f_2, \dots, f_n$  шляхом розкладання в ряд Тейлора в околі точки початкового наближення до розв'язку системи рівнянь і нехтування всіма членами ряду, окрім лінійних щодо приростів змінних.

Для однієї змінної ряд Тейлора в околі певної точки  $x = x_0$  виглядає як :

$$f(x) = f(x_0) + \frac{1}{1!}(x - x_0)f'(x_0) + \frac{1}{2!}(x - x_0)^2 f''(x_0) + \dots + \frac{1}{n!}(x - x_0)^n f^{(n)}(x_0). \quad (10.3)$$

Для функцій  $f_1, f_2, \dots, f_n$  системи рівнянь (10.1) візьмемо лише лінійну частину (до другої похідної) розкладання в ряд Тейлора в околі точки  $x^{(0)} = \{x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}\}$  :

$$f_i(x_1, x_2, \dots, x_n) = f_i(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}) + (x_1 - x_1^{(0)}) \frac{\partial}{\partial x_1} f_i(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}) + (x_2 - x_2^{(0)}) \frac{\partial}{\partial x_2} f_i(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}) + \dots + (x_n - x_n^{(0)}) \frac{\partial}{\partial x_n} f_i(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$$

Уведемо позначення для змінних:

$\Delta x_i^{(0)} = (x_i - x_i^{(0)})$  – приріст  $i$ -тої змінної,  $f_i$  – значення  $i$ -тої функції,  $F_{ij} \partial / \partial x_0 f_i(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$  – значення першої частинної похідної функції  $f_i$  за змінною  $x_j$ .

Після перетворення дістанемо систему лінійних рівнянь порядку  $n$  щодо приросту змінних  $\Delta x_j$  :

$$\begin{aligned} F'_{11}\Delta x_1 + F'_{12}\Delta x_2 + \dots + F'_{1n}\Delta x_n &= -f_1; \\ F'_{21}\Delta x_1 + F'_{22}\Delta x_2 + \dots + F'_{2n}\Delta x_n &= -f_2; \\ &\dots\dots\dots \\ F'_{n1}\Delta x_1 + F'_{n2}\Delta x_2 + \dots + F'_{nn}\Delta x_n &= -f_n; \end{aligned} \quad (10.4)$$

або в матричній формі

$$\begin{pmatrix} F'_{11} & F'_{12} & \dots & F'_{1n} \\ F'_{21} & F'_{22} & \dots & F'_{2n} \\ \dots & \dots & \dots & \dots \\ F'_{n1} & F'_{n2} & \dots & F'_{nn} \end{pmatrix} \cdot \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \dots \\ \Delta x_n \end{pmatrix} = \begin{pmatrix} -f_1 \\ -f_2 \\ \dots \\ -f_n \end{pmatrix}. \quad (10.5)$$

У скороченому вигляді можна записати  $(F')(\Delta x) = -(f)$ , де матриця значень частинних похідних  $(F')$  називається матрицею Якобі, чи якобіаном системи рівнянь.

Розв'язок цієї системи (за умови  $\det(F') \neq 0$ ) надає вектор відхилень до початкового наближення  $\Delta x = -(F') \cdot (-1) \cdot (f)$ . Додавання його до вектора початкового наближення надає нові, уточнені значення змінних:

$$x^{(1)} = x^{(0)} + \Delta x^{(0)}. \quad (10.6)$$

Продовжуючи ітераційний процес, дістанемо нові наближення розв'язків системи лінійних рівнянь за скороченою формулою

$$x^{(k+1)} = x^{(k)} + \Delta x^{(k)} \quad (10.7)$$

або у загальному вигляді

$$x^{(k+1)} = x^{(k)} - F^{-1}(x^{(k)}) \cdot f(x^{(k)}); \quad (10.8)$$

де  $F^{-1}(x^{(k)})$  – обернена матриця Якобі  $F'$  для наближення  $x^{(k)} = \{x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}, \dots\}$ ,  $k = 0, 1, 2, \dots$ . Отже, метод Ньютона буде ітераційну послідовність  $\{x^{(k)}\}$ ,  $(x^{(k)} \in R^n)$ ,  $k = 0, 1, 2, \dots$ , наближень розв'язку  $x^*$  (початкове наближення  $x_0$  задається) за такою ітераційною формулою:

$$x^{(k+1)} = x^{(k)} - [F'(x^{(k)})]^{-1} \cdot F(x^{(k)}); \quad (10.9)$$

$$F'(x) = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_n(x)}{\partial x_1} & \cdots & \frac{\partial f_n(x)}{\partial x_n} \end{pmatrix} \quad (10.10)$$

матриця Якобі, тобто  $F'(x(k))$  – матриця розмірності  $n \times n$ .

Процес (10.7–10.9) триває доти, поки не виконається умова

$$\|F(x^{(k)})\| < \varepsilon,$$

де  $\varepsilon$  – задана точність розв'язку задачі (10.2).

Ідея метода Ньютона полягає в тому, що на  $k$ -й ітерації ( $x(k)$  – поточне наближення розв'язку) таке наближення розв'язку  $x(k+1)$  розміщене, як розв'язок системи лінійних рівнянь:

$$F_k(x) = 0, \quad (10.11)$$

де  $F_k(x) \equiv F(x^{(k)}) + F'(x^{(k)})(x - x^{(k)})$  (перші два члени розклада-  
ння в ряд Тейлора функції  $F(x)$  в околі точки  $x(k)$ , тобто система  
(10.11) є лінеаризацією (лінійним наближенням) системи (10.2).

Оскільки система (10.11) лінійна відносно  $x$ , то її розв'язок може  
бути знайдений аналітично:

$$F_k(x) = F(x^{(k)}) + F'(x^{(k)})(x - x^{(k)}) = 0 \quad (10.12)$$

або

$$F'(x^{(k)})(x - x^{(k)}) = -F(x^{(k)}). \quad (10.13)$$

Із застосуванням оберненої матриці

$$(x - x^{(k)}) = -[F'(x^{(k)})]^{-1} \cdot F(x^{(k)}), \quad (10.14)$$

звідси і одержуємо розв'язок системи (10.4)

$$x = x^{(k)} - [F'(x^{(k)})]^{-1} \cdot F(x^{(k)}). \quad (10.15)$$

Якщо послідовність  $\{x(k)\}$ ,  $k = 0, 1, 2, \dots$ , побудована згідно з (10.9),  
збігається, то за достатньо загальних умов швидкість її збіжності  
буде квадратичною, тобто

$$\|x^{(k+1)} - x^*\| \leq M \cdot \|x^{(k)} - x^*\|^2,$$

починаючи з деякого  $k$ , де  $M$  – деяка додатна константа.

Частинні похідні, потрібні для розрахунку матриці Якобі, можна обчислити аналітично або ж, якщо це неможливо чи то важко, діставати за формулами наближеного диференціювання.

Умови збіжності методу Ньютона для систем нелінійних рівнянь досліджували відомі вчені: Канторович, Островський, Віллерс, Стенін. Узагальнюючи їхні дослідження, можна вважати за достатні умови збіжності розв'язків систем нелінійних рівнянь методу Ньютона такі:

- 1 Матриця Якобі для початкового наближення  $F'(x^{(0)})$  повинна мати обернену матрицю  $F^{-1}$  з нормою, меншою за певну величину  $A$ , тобто

$$\|F^{-1}(x^{(0)})\| \leq A. \quad (10.16)$$

- 2 Норма добутку оберненої матриці Якобі на вектор заданих функцій  $f(x)$  повинна мати значення, менше за певну величину :

$$\|F^{-1}(x^{(0)}) \cdot f(x^{(0)})\| \leq B. \quad (10.17)$$

- 3 Значення матриці Якобі для частинних похідних другого порядку мають задовольняти умові:

$$\sum_{k=1}^n \left| \frac{\partial^2 f_i(\bar{x})}{\partial x_j \partial x_k} \right| \leq C, \quad (10.18)$$

де  $i, j = 1, 2, \dots, n$  – певні значення наближень розв'язків системи рівнянь в околі точки  $(0)$ .

- 4 Сталі величини  $A, B$  і  $C$  мають задовольняти умові

$$2 \cdot n \cdot A \cdot B \cdot C \leq 1. \quad (10.19)$$

Основними недоліками метода Ньютона є:

- збіжність лише для достатньо близьких до розв'язку початкових наближень  $x^{(0)}$ ;
- висока трудомісткість методу, оскільки на кожній ітерації необхідно обчислювати матриці  $F'(x^{(k)})$  і  $[F'(x^{(k)})]^{-1}$ .



Основною перевагою методу Ньютона є висока швидкість збіжності.

Необхідно зазначити, що в методі Ньютона формулу (10.9) можна записати у вигляді

$$x^{(k+1)} = x^{(k)} + h^{(k)},$$

де  $h^{(k)} = -[F'(x^{(k)})]^{-1} \cdot F(x^{(k)})$ . Проте під час практичної реалізації методу вектор  $h^{(k)}$  ефективніше обчислювати як розв'язок системи лінійних рівнянь вигляду

$$F'(x^{(k)}) \cdot h = -F(x^{(k)}).$$

Алгоритм розв'язування системи нелінійних рівнянь за методом Ньютона виглядає так:

- 1) обираємо початкове наближення  $(0) = \{x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}\}$ ;
- 2) обчислюємо матрицю Якобі ( $F'$ ) – значення частинних похідних  $F'_{ji}$  для обраного наближення  $(k)$  ( $k$  – номер кроку ітерації);
- 3) розв'язуємо систему лінійних рівнянь (10.6) щодо приростів змінних  $\Delta x(k) = -(F')^{-1} \cdot (x(k)) \cdot f(x(k))$ ;
- 4) до вектора наближення  $(k)$  додаємо вектор приростів змінних  $\Delta x(k)$  та дістаємо нове наближення  $(k+1) = (k) + \Delta x(k)$ ;
- 5) перевіряємо умову завершення процесу розв'язування системи рівнянь. Якщо умови не досягнуто, то значення  $k$  збільшуємо на одиницю і повторюємо процедуру з п. 2, інакше процес ітерації зупиняємо.



## ПРИКЛАД 1

---

Розв'язати систему нелінійних рівнянь з точністю  $\varepsilon = 0.001$ . Відокремити корені графічно. Знайти розв'язок системи рівнянь на інтервалах:  $x \in [-1.2; -1.0]$ ;  $y \in [-0.2; 0.0]$

$$\begin{cases} x + x^2 + y^2 = 0.1; \\ y + 2xy = 0.1. \end{cases}$$

### Розв'язування

Для локалізації кореня системи рівнянь на заданому інтервалі побудуємо графіки функцій  $x + x^2 + y^2 - 0.1 = 0$  та  $y + 2xy - 0.1 = 0$ , як подано на рисунку 10.1.

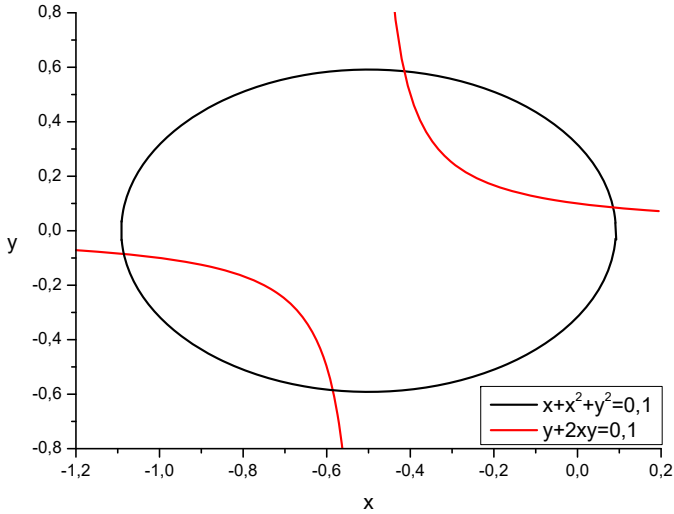


Рисунок 10.1– Графічний розв'язок системи рівнянь

Перепишемо систему у вигляді  $f(x) = 0$  та знайдемо якобіан

$$\begin{cases} x + x^2 + y^2 - 0.1 = 0; \\ y + 2xy - 0.1 = 0. \end{cases}$$

Тоді

$$f(x) = \begin{pmatrix} x + x^2 + y^2 - 0.1 \\ y + 2xy - 0.1 \end{pmatrix}; \quad W(x) = \begin{pmatrix} 1 + 2x & 2y \\ 2x & 1 + 2x \end{pmatrix}.$$

За методом Ньютона кожне наступне наближення вектора коренів шукають так:

$$x^{(k+1)} = x^{(k)} + \Delta x^{(k)}; \quad \Delta x^{(k)} = -[W(x^{(k)})]^{-1} f(x^{(k)})$$

з умовою виходу  $\|\Delta x^{(k)}\| < \varepsilon$ .

Початкове наближення для заданого інтервалу  $x^{(0)} = (-1; -0.1)$ .

Умови збіжності методу Ньютона:

$$2nABC \leq 1,$$

де

$$\begin{aligned}\|W^{-1}(x^{(0)})\| &\leq A, \\ \|W^{-1}(x^{(0)})f(x^{(0)})\| &\leq B, \\ \sum_{k=1}^n \left| \frac{\partial^2 f_i(x^{(0)})}{\partial x^j \partial x_k} \right| &\leq C.\end{aligned}$$

Для нашого випадку  $A = 2.8$ ;  $B = 0.03$ ;  $C = 2$ ;  $n = 2$ .  
Отже, умова збіжності методу виконується, а саме

$$2nABC = 2 \cdot 2 \cdot 2.8 \cdot 0.03 \cdot 2 = 0.67 < 1,$$

тому метод сходиться.

На 9-й ітерації за методом Ньютона одержали такі результати:

$$x \approx -1.0854; y \approx -0.0854. \quad \blacksquare$$

## Список літератури

1. Абакумова О. О. Чисельні методи. Лабораторний практикум : навчальний посібник. Київ, 2020. 74 с.
2. Данилевський М. П., Якунін А. В., Кузнецова Г. А. Елементи обчислювальної математики : конспект лекцій. Харків, 2012. 157с.
3. Чисельні методи : навчальний посібник / Л. О. Волонтир, Л. В. Зелінська, Н. А. Потапова, І. А. Чіков. Вінниця, 2020. 322 с.
4. Матюшенко М. В., Сівак Є. М. Методи наближених обчислень : навчальний посібник. Харків, 2021. 92 с.
5. Задачин В. М., Коношенко І. Г. Чисельні методи : навчальний посібник. Харків, 2014. 180 с.
6. Дичка І. А., Онаї М. В., Гадиняк Р. А. Чисельні методи розв'язання задач лінійної алгебри та нелінійних рівнянь : лабораторний практикум. Київ, 2018. 95 с.
7. Колесницький О. К., Арсенюк І. Р., Месюра В. І. Чисельні методи : навчальний посібник. Вінниця, 2017. 131 с.

Електронне навчальне видання

Дворниченко Аліна Василівна

## ЧИСЕЛЬНІ МЕТОДИ ОБРОБКИ ДАНИХ

Конспект лекцій

для здобувачів спеціальності  
*113 «Прикладна математика»*  
очної форми здобуття вищої освіти

У двох частинах  
Частина I

Відповідальний за випуск І. В. Коплик  
Редакторка Н. М. Мажуга  
Комп'ютерне верстання А. В. Дворниченко

Формат 60x84/16. Ум. друк. арк. 6,74. Обл.-вид. арк. 6,28.

Видавець і виготовлювач  
Сумський державний університет,  
вул. Римського-Корсакова, 2, м. Суми, 40007  
Свідоцтво суб'єкта видавничої справи ДК № 3062 від 17.12.2007.