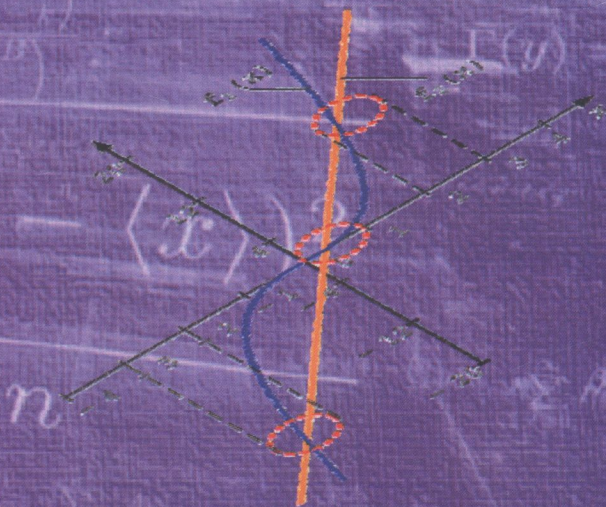


519.610.78.8)
МЗєр

МАТЕМАТИЧНІ МЕТОДИ МОДЕЛЮВАННЯ

Навчальний посібник



18

519.6(075.8)
У34

МАТЕМАТИЧНІ МЕТОДИ МОДЕЛЮВАННЯ

Навчальний посібник

НТБ ВНТУ



477574

519.6(075.8) М34 2016

Математичні методи моделювання

КРЕМЕНЧУК

Кременчук
ПП Щербатих О.В.
2016

УДК 621.313:517.91

ББК 31.26

М 34

Авторський колектив

О. П. Чорний, В. К. Титюк, Н. М. Істоміна, В. А. Власенко

Рецензенти: *Л. І. Мазуренко*, – доктор техн. наук, професор, завідувач відділу електромеханічних систем Інституту електродинаміки НАН України

В. П. Ляшенко, – доктор техн. наук, професор, завідувач кафедри інформатики і вищої математики Кременчуцького національного університету імені Михайла Остроградського

І. В. Шелевицький, – доктор техн. наук, професор, завідувач кафедри інформаційних технологій та моделювання Криворізького економічного інституту Київського національного університету імені Вадима Гетьмана

Рекомендовано до видання вченою радою

Кременчуцького національного університету імені Михайла Остроградського

Міністерства освіти і науки України

(протокол №10, від 24.05.2016).

М 34

Математичні методи моделювання : навчальний посібник /

О. П. Чорний, В. К. Титюк, Н. М. Істоміна та ін. ; заг. ред.

О. П. Чорний. – Кременчук : ПП Щербатих О. В., 2016. – 234 с. : іл.

ISBN 978-617-639-116-6

Посібник присвячений чисельним методами математичного аналізу які використовуються у сучасних програмних засобах розрахунку математичних моделей. Теоретичні положення супроводжуються графічним матеріалом і прикладами розрахунку із застосуванням програмного середовища Mathcad. Матеріал посібника охоплює більшість задач, що постають перед студентами, аспірантами, науковими та інженерними робітниками.

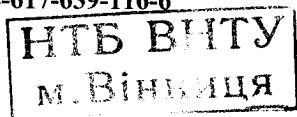
Іл. 33, табл. 7, бібліограф. 60 назв.

УДК 621.313:517.91

ББК 31.26

477544

ISBN 978-617-639-116-6



© О. П. Чорний, В. К. Титюк,
Н. М. Істоміна, В. А. Власенко, 2016

MATHEMATICAL METHODS FOR SIMULATION

THE TUTORIAL

Kremenchuk
PE Shcherbatykh O.V.
2016

UDC 621.313:517.91

BBK 31.26

M 34

COMPOSITE AUTHOR:

*Chorny Oleksii Petrovych,
Tytiuk Valerii Kostiantynovych,
Istomina Nataliia Mykolaivna,
Vlasenko Volodymyr Anatoliiovych.*

Reviewers:

Mazurenko Leonid Ivanovych,

*D.Sc. (Eng.), Professor, Head of Electromechanical Systems
Department, The Institute of Electrodynamics, The National
Academy of Sciences of Ukraine*

Liashenko Viktor Pavlovych,

*D.Sc. (Eng.), Professor, Head of Mathematics and Computer Science
Department, Kremenchuk Mykhailo Ostrohradskyi National
University*

Shelevytsky Ihor Volodymyrovych, *D.Sc. (Eng.), Professor, Head of
Information Technologies and Simulation Department, Kryvyi Rih
Institute of Economics, Kyiv National Economic University named
after Vadym Hetman*

*The tutorial is published by the decision of the Scientific Council of
Kremenchuk Mykhailo Ostrohradskyi National University
Ministry of Education and Science of Ukraine
(Record № 10 of 24 May, 2016).*

M 34 Chorny, O.P., Tytiuk, V.K., Istomina, N.M. and Vlasenko, V.A.
(2016), **Mathematical Methods for Simulation: the tutorial**, PE
Shcherbatykh O.V., Kremenchuk, Ukraine. [in Ukrainian]

ISBN 978-617-639-116-6

The tutorial elucidate numerical methods of mathematic analysis used at up-to-day mathematical models computation software. The drawings and computation examples with use of Mathcad software environment are given to an abstract theorems. The tutorial material envelope most work and research tasks of students, PhD students, scientists and engineers.

UDC 621.313:517.91

BBK 31.26

ISBN 978-617-639-116-6

© Chorny O.P., Tytiuk V.K.,
Istomina N.M., Vlasenko V.A., 2016

**«Мысль о пространстве рождает "ах",
оперу, взгляд в лорнет.
В цифрах есть нечто, чего в словах,
даже крикнув их, нет.»**

*И. А. Бродский
«Полдень в комнате», 1978*

«Все есть число»

Пифагор

ЗМІСТ

| | |
|---|----|
| ПЕРЕДМОВА | 9 |
| ЛЕКЦІЯ 1 ВСТУП В ПРЕДМЕТ | 11 |
| 1.1 Визначення поняття моделі. Співвідношення між моделлю та об'єктом | 11 |
| 1.2 Вимоги до моделі | 12 |
| 1.3 Класифікація моделей | 13 |
| Питання для самоконтролю | 16 |
| ЛЕКЦІЯ 2 ЕЛЕМЕНТИ ТЕОРІЇ ОБЧИСЛЮВАЛЬНИХ ПОХИБОК. 17 | |
| 2.1 Загальні характеристики чисельних методів | 17 |
| 2.2 Класифікація похибок обчислень | 20 |
| 2.3 Абсолютна і відносна похибки. Точні десятинні знаки. Вірні знаки числа | 21 |
| 2.4 Похибки вимірювань | 24 |
| 2.5 Абсолютна та відносна похибка обчислення елементарних арифметичних операцій | 25 |
| Питання для самоконтролю | 30 |
| Варіанти контрольних завдань до лекції 2 | 30 |
| ЛЕКЦІЯ 3 МЕТОДИ РОЗВ'ЯЗАННЯ СИСТЕМ ЛІНІЙНИХ АЛГЕБРАЇЧНИХ РІВНЯНЬ | 35 |
| 3.1 Прямі методи розв'язання СЛАР | 38 |
| 3.2 Ітераційні методи розв'язання СЛАР | 47 |
| 3.3 Розв'язання СЛАР в пакеті MathCAD | 52 |
| Питання для самоконтролю | 58 |
| Варіанти контрольних завдань до лекції 3 | 59 |
| ЛЕКЦІЯ 4 ЧИСЕЛЬНІ МЕТОДИ РОЗВ'ЯЗАННЯ НЕЛІНІЙНИХ РІВНЯНЬ ТА ЇХ СИСТЕМ | 62 |
| 4.1 Відділення коренів рівняння | 64 |
| 4.2 Уточнення коренів | 68 |
| 4.2.1 Метод поділу навпіл (бісекції, дихотомії) | 69 |
| 4.2.2 Метод хорд (метод січних) | 71 |
| 4.2.3 Метод Ньютона-Рафсона (метод дотичних) | 74 |
| 4.3 Метод простих ітерацій | 77 |
| 4.4 Розв'язання нелінійних рівнянь та їх систем у пакеті MathCAD. 79 | |
| Питання для самоконтролю | 88 |
| Завдання до лекції 4 | 89 |
| ЛЕКЦІЯ 5 НАБЛИЖЕННЯ ФУНКЦІЙ | 90 |
| 5.1 Постановка задачі інтерполяції | 91 |
| 5.2 Методи локальної інтерполяції | 93 |
| 5.2.1 Локальна кусково-постійна інтерполяція (метод найближчого сусіда) | 93 |

| | | |
|-------|--|------------|
| 5.2.2 | Локальна кусково-лінійна інтерполяція..... | 93 |
| 5.2.3 | Локальна кусково-квадратична інтерполяція..... | 96 |
| 5.3 | Методи глобальної інтерполяції..... | 96 |
| 5.3.1 | Інтерполяція степеневим багаточленом (поліномом) у канонічній формі..... | 97 |
| 5.3.2 | Інтерполяційний поліном Лагранжа..... | 98 |
| 5.3.3 | Інтерполяційний поліном Ньютона..... | 101 |
| 5.3.4 | Інтерполяція функцій у пакеті MathCAD..... | 103 |
| 5.4 | Постановка задачі апроксимації..... | 106 |
| 5.4.1 | Визначення параметрів емпіричної формули. Метод найменших квадратів (МНК)..... | 109 |
| 5.4.2 | Апроксимація тригонометричними поліномами..... | 112 |
| 5.4.3 | Апроксимація функцій у пакеті MathCAD..... | 114 |
| 5.5 | Похибка інтерполяції та апроксимації..... | 119 |
| 5.6 | Теорія множинності моделей..... | 124 |
| 5.7 | Згладжування експериментальних даних..... | 125 |
| | Питання для самоконтролю..... | 127 |
| | Завдання до самостійної роботи..... | 129 |
| | Варіанти контрольних завдань до лекції 5..... | 129 |
| | ЛЕКЦІЯ 6 РОЗВ'ЯЗАННЯ ОПТИМІЗАЦІЙНИХ ЗАДАЧ..... | 134 |
| 6.1 | Розв'язання одномірних задач оптимізації..... | 135 |
| 6.1.1 | Метод зворотного змінного кроку..... | 135 |
| 6.1.2 | Метод квадратичної апроксимації (метод Пауелла)..... | 136 |
| 6.2 | Розв'язання багатомірних задач оптимізації..... | 137 |
| 6.2.1 | Метод покоординатного спуску..... | 138 |
| 6.2.2 | Метод найшвидшого спуску (метод градієнтів)..... | 140 |
| 6.2.3 | Метод Нелдера-Міда (деформованого багатогранника)..... | 140 |
| 6.3 | Методи пошуку глобального екстремуму..... | 143 |
| | Питання для самоконтролю..... | 146 |
| | Завдання до самостійної роботи..... | 147 |
| | ЛЕКЦІЯ 7 ЧИСЕЛЬНЕ ДИФЕРЕНЦІЮВАННЯ ТА ІНТЕГРУВАННЯ..... | 148 |
| 7.1 | Чисельне диференціювання..... | 148 |
| 7.2 | Чисельне інтегрування..... | 153 |
| 7.2.1 | Методи прямокутників..... | 156 |
| 7.2.2 | Метод трапецій..... | 158 |
| 7.2.3 | Метод Сімпсона (метод парабол)..... | 159 |
| 7.2.4 | Похибка формул Ньютона-Котеса..... | 160 |
| 7.2.5 | T-інтегрування функцій в реальному масштабі часу..... | 161 |
| 7.3 | Спеціальні питання чисельного інтегрування функцій..... | 163 |
| 7.3.1 | Обчислення інтеграла з заданою точністю..... | 163 |

| | | |
|--|--|--|
| 7.3.2 | Інтегрування розривних функцій | 164 |
| 7.3.3 | Методи Монте-Карло (метод статистичних випробувань) | 164 |
| 7.3.4 | Обчислення визначених інтегралів засобами пакета MathCAD..... | 165 |
| | Завдання до самостійної роботи | 171 |
| | Питання для самоконтролю..... | 171 |
| | Варіанти контрольних завдань до лекції 7..... | 172 |
| ЛЕКЦІЯ 8 ЧИСЕЛЬНІ МЕТОДИ РОЗВ'ЯЗАННЯ ДИФЕРЕНЦІЙНИХ РІВНЯНЬ ТА ЇХ СИСТЕМ | | 174 |
| 8.1 | Постановка задачі..... | 174 |
| 8.2 | Методи розв'язання ЗДР..... | 177 |
| 8.2.1 | Метод Ейлера (метод ламаних) | 179 |
| 8.2.2 | Модифікований метод Ейлера | 180 |
| 8.2.3 | Методи Рунге-Кутти..... | 181 |
| 8.2.4 | Метод Адамса..... | 183 |
| 8.2.5 | Метод Мілна..... | 185 |
| 8.3 | Автоматичний вибір кроку інтегрування та контроль за точністю обчислень | 185 |
| 8.4 | Розв'язання диференціальних рівнянь та їх систем засобами пакету MathCAD..... | 187 |
| | Питання для самоконтролю..... | 196 |
| | Варіанти контрольних завдань до лекції 8..... | 198 |
| ЛЕКЦІЯ 9 ЧИСЕЛЬНЕ РОЗВ'ЯЗАННЯ СИСТЕМ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ В ЧАСТКОВИХ ПОХІДНИХ | | 201 |
| 9.1 | Рівняння математичної фізики..... | 202 |
| 9.1.1 | Рівняння параболічного типу..... | 205 |
| 9.1.2 | Рівняння гіперболічного типу..... | 206 |
| 9.1.3 | Рівняння еліптичного типу..... | 206 |
| 9.1.4 | Початкові та граничні умови | 207 |
| 9.2 | Методи розв'язання задач математичної фізики в електромеханічних системах..... | 209 |
| 9.2.1 | Метод кінцевих різниць | 210 |
| 9.2.2 | Метод кінцевих елементів..... | 216 |
| 9.3 | Програмні засоби для розв'язання рівнянь математичної фізики | 221 |
| | Питання для самоконтролю..... | 226 |
| | Предметний покажчик | 227 |
| | Література | 229 |
| | References | Ошибка! Закладка не определена. |

ПЕРЕДМОВА

«Цель расчетов – не числа, а понимание»

Р.В. Хемминг. Численные методы для научных работников и инженеров

Даний навчальний посібник призначено, в першу чергу, для студентів, але може бути корисним також для аспірантів і наукових працівників, які займаються розрахунком й аналізом моделей, складених на основі фундаментальних законів.

У даний час існує велика кількість літератури, присвяченої чисельним методам і математичному моделюванню, тому автори, готуючи матеріал до «Математичних методів моделювання», ні якою мірою не намагалися претендувати на оригінальність викладу матеріалу або його всеосяжність.

Сьогодні різні програмні системи, а також Internet, надають можливості моделювання, розрахунку та аналізу моделей із використанням чисельних методів. Для цього розроблено велику кількість різноманітних програмних засобів. Широко застосовуються такі стандартні пакети, як Microsoft Excel, Mathematica, MathCad, Maple, Matlab та ін. Майже щороку з'являються оновлені версії цих пакетів. Крім того, всі сучасні програмні засоби містять вбудовану систему прикладів розрахунку й навіть коротку теорію. Досить просто відкрити будь-який пакет, що реалізує чисельні методи, вибрати метод і виконати бажані обчислення. Проте слід зазначити, що зазвичай стандартний програмний пакет реалізує ті алгоритми, які вважали за потрібне застосувати самі розробники. Далеко не всі пакети дають користувачеві можливість вибирати конкретний чисельний метод і, особливо, оцінити достовірність і точність отриманого результату.

Таким чином, абсолютно не заперечуючи корисності існуючих програмних засобів, автори глибоко переконані, що користувач, застосовуючи той чи інший убудований у пакет метод, повинен чітко уявляти собі його суть і мати можливість контролювати точність реалізації обраного методу для кожного

конкретного завдання. В іншому випадку слід самому реалізовувати чисельні методи програмним шляхом. Такі дії можливо робити, лише озброївшись теоретичними знаннями з чисельних методів. І саме таку задачу автори ставили перед собою – допомогти читачам у вирішенні цього завдання.

Перш ніж перейти до викладення матеріалу, ми хотіли б зробити ще одне зауваження. Читач зверне увагу на те, що приклади розрахунку виконані із застосуванням практично лише одного програмного середовища – MathCad. Виникає питання, а чому не Maple або Matlab? Останній узагалі вважається найбільш поширеним серед дослідників через свою універсальність, наявність бібліотек, кількість застосовуваних методів розрахунку, аналізу результатів та їх візуалізації. Досвід та виконані дослідження показують, що розв’язок однієї й тієї ж задачі різними програмними середовищами, або навіть їх різними версіями, дає результати, які можуть дещо відрізнитися. Звісно, це не стосується задач лінійної алгебри, але спостерігається при чисельних розв’язках, наприклад, жорстких чи напівжорстких систем диференціальних рівнянь. Ці спостереження тільки наголошують на необхідності чіткого розуміння щодо вибору методу розв’язку, його реалізації й навіть виду та значення отриманого результату.

Автори висловлюють подяку рецензентам, критичні зауваження яких дозволили покращити викладення теорії чисельних методів.

Зауваження та пропозиції автори просять надсилати на адресу:

Кафедра систем автоматичного управління і електроприводу,
Кременчуцький національний університет імені Михайла Остроградського,
вул. Першотравнева, 20, м. Кременчук Полтавської обл., Україна. 39600.

Automatic Control Systems and Electric Drive Department
Kremenchuk Mykhailo Ostrohradskyi National University,
vul. Pershotravneva, 20, Kremenchuk, Poltava Region, Ukraine, 39600.

ЛЕКЦІЯ 1

ВСТУП В ПРЕДМЕТ

— Ты любишь математику больше, чем меня!
— Дорогая! Как ты могла подумать!
— Ну тогда докажи!
— Ладно. Пусть R —множество всех любимых объектов...

1.1 Визначення поняття моделі.

Співвідношення між моделлю та об'єктом

Модель – це представлення об'єкта, системи або поняття в деякій формі, відмінній від реального існування. Модель є засобом, що допомагає в поясненні, розумінні або удосконалюванні системи. Модель може бути точною копією об'єкта (хоча й в іншому масштабі і з іншого матеріалу) або відображати деякі характерні властивості об'єкта в абстрактній формі. Тому модель – це інструмент для прогнозування наслідків при дії вхідних сигналів на об'єкт, який підвищує ефективність суджень і інтуїції фахівців [59].

Всі моделі – спрощені уявлення реального світу або абстракції. Звичайно відкидають велику частину реальних характеристик досліджуваного об'єкта і вибирають ті його особливості, що ідеалізують варіант реальної події.

Подібність моделі з об'єктом характеризується ступенем ізоморфізму. Для того щоб бути цілком

ізоморфною, модель повинна задовольняти дві умови: по-перше, повинна існувати взаємно однозначна відповідність між елементами моделі й елементами, що представляють об'єкт; по-друге, повинні бути збережені точні співвідношення (взаємодії) між елементами.

Процес моделювання полягає в наступному: загальна задача дослідження системи розділяється на ряд більш простих; чітко формулюються цілі моделювання; підшукується аналогія; розглядається спеціальний числовий приклад, що відповідає даній задачі; вибираються певні позначення; записуються очевидні співвідношення. Якщо отримана модель піддається математичному опису, її розширюють, у противному випадку – спрощують.

1.2 Вимоги до моделі

Найбільш загальні вимоги до моделі можуть бути сформульовані таким чином: модель повинна бути простою і зрозумілою користувачу, цілеспрямованою, гарантованою від абсурдних результатів, зручною в керуванні і спілкуванні, повною з погляду розв'язання головних завдань, адаптивною, що дозволяє легко переходити до інших модифікацій або оновлювати дані, дозволяти поступові зміни, тобто, будучи спочатку простою, вона може у взаємодії з користувачем ставати усе складнішою.

Ідея уявлення системи за допомогою моделі носить настільки загальний характер, що дати повну класифікацію усіх функцій моделі важко.

Модель може служити для досягнення двох цілей: описової, якщо модель служить для пояснення і кращого розуміння об'єкта, і керівної, коли модель дозволяє передбачити або відтворити характеристики об'єкта, що визначають її поведінку.

1.3 Класифікація моделей

Моделі можна класифікувати різними способами, але жоден із них не є вичерпним. Тому наведемо визначення для двох найбільш поширених методів моделювання.

Фізичне моделювання (рос. *физическое моделирование*; англ. *physical simulation*, нім. *physikalische Modellierung*) полягає в наступному [50]:

- створення матеріальної моделі, що має таку саму фізичну природу (такий самий фізичний зміст), як і дійсне явище, що вивчається на основі критеріїв геометричного, кінематичного й динамічного моделювання;
- відтворення на моделі й дослідження процесів, що якісно однакові з процесами у реальному об'єкті. Під час фізичного моделювання процесу необхідно забезпечити геометричну, часову та фізичну подібності.

Фізичне моделювання – метод експериментального вивчення фізичних явищ, який базується на їх фізичній подібності.

Метод застосовується у випадках, коли:

- відсутня математична модель явища (машини, процесу тощо), або така модель дуже складна,

вимагає багато вихідних даних, одержання яких ускладнене.

- відтворення явища (машини, процесу) в реальних масштабах недоцільне.

Математичне моделювання (рос. *моделирование математическое*; англ. *mathematical simulation*, нім. *mathematische Modellierung*) – метод дослідження процесів або явищ, коли реальний об'єкт представляється у вигляді набору взаємопов'язаних математичних об'єктів; рівнянь, графіків, таблиць та інше [16, 17]:

При складанні математичного опису об'єкта слід пам'ятати, що математична модель – це математичне зображення реальності (фізичне представлення математичних понять) [2].

Згідно з принципами системного аналізу [48], необхідно визначити структурно-функціональні зв'язки у системі та відобразити їх за допомогою схем заміщення або графів.

За отриманою схемою заміщення чи графом складаються системи рівнянь, які описують процеси, що відбуваються у досліджуваній системі.

Згідно з [42], найчастіше побудова математичної моделі ґрунтується на застосуванні фундаментальних фізичних законів до конкретної ситуації:

1. Закон збереження енергії.
2. Закон збереження матерії.
3. Закон збереження імпульсу.

Для великої кількості об'єктів при створенні математичної моделі неможливо вказати фундаментальні закони або принципи, яким він підпорядковується. У

цьому випадку застосовують аналогії з відомими об'єктами. Наприклад, електромеханічні аналогії можуть замінювати опис процесів, що відбуваються у різноманітних фізичних, біологічних і соціальних об'єктах.

Формальна класифікація моделей ґрунтується на математичних засобах, що використовуються для розв'язання поставлених задач [16, 17].

Розрізняють моделі:

- лінійні або нелінійні моделі;
- зосереджені або розподілені системи;
- детерміновані або стохастичні;
- статичні або динамічні.

Процес математичного моделювання включає наступні основні етапи:

1. Розробка математичної моделі об'єкта.

Цей етап є найбільш складним, трудомістким і відповідальним. На основі теоретичних знань, емпіричних й інтуїтивних підходів складаються математичні рівняння, що враховують найбільш важливі та істотні, з точки зору дослідника, властивості об'єкта. При розробці математичної моделі необхідно уникати невиправданого ускладнення моделі, відкидаючи несуттєві взаємозв'язки між характеристиками об'єкта. Таким чином, при розробці математичної моделі необхідно прийняти допущення й обмеження на діапазоні застосування моделі.

2. Визначення числових значень параметрів математичної моделі.

3. Одержання розв'язків математичної моделі.

На цьому етапі, попередньо задаючись значенням параметрів рівнянь, які є математичною моделлю,

знаходять їхні розв'язки в аналітичному або числовому вигляді. Для складних моделей розв'язок відшуковують на ЕОМ.

4. Оцінка адекватності отриманих результатів.

Отримані з використанням моделі результати необхідно оцінити з точки зору їхньої відповідності основним фізичним законам і іншими обмеженнями.

Питання для самоконтролю

1. Дайте визначення поняття моделі.
2. Сформулюйте вимоги до моделі.
3. Перелічіть методи моделювання.
4. Поясніть термін «фізичне моделювання».
5. Для яких випадків застосовуються фізичні моделі?
6. Поясніть термін «математичне моделювання».
7. Для яких випадків застосовуються математичні моделі?
8. Які види математичних моделей ви знаєте?
9. Назвіть фундаментальні фізичні закони, на яких ґрунтуються математичні моделі.
10. Перелічіть основні етапи математичного моделювання.

ЛЕКЦІЯ 2

ЕЛЕМЕНТИ ТЕОРІЇ ОБЧИСЛЮВАЛЬНИХ ПОХИБОК

«Всякая точная наука основывается на приближенности.»

Бертран Рассел

Телефонный звонок математика:

- Алло, это квартира Сидорова Ивана Петровича?

- Нет, это квартира Каца Абрама Самуиловича.

- Извините, это 22-38-89?

- Нет, это 22-38-88.

- Надо же! В шестом знаке ошибка, а такой эффект!

2.1 Загальні характеристики чисельних методів

Чисельні методи – це математичний інструментарій, за допомогою якого математичне завдання формулюється у вигляді, зручному для розв'язання на комп'ютері. У такому разі говорять про перетворення математичної задачі в обчислювальну задачу. При цьому послідовність виконання необхідних арифметичних і логічних операцій визначається алгоритмом її розв'язання. Алгоритм повинен бути ітеративним, реентерабельним і складатися з відносно невеликих блоків, які багаторазово виконуються для різних вхідних даних.

Слід зазначити, що з появою швидких та потужних цифрових комп'ютерів роль чисельних методів для вирішення наукових та інженерних задач значно зросла. Хоча аналітичні методи розв'язання математичних задач, як і раніше, дуже важливі, чисельні методи істотно розширюють область вирішальності наукових та інженерних завдань, не дивлячись на те, що самі рівняння математичних моделей з ускладненням структури сучасних об'єктів стають погано обумовленими та жорсткими, що суттєвим чином ускладнює їх розв'язання. Взявши виконання рутинних обчислень на себе, комп'ютери звільняють час вченого або інженера для творчості: формулювання задач і генерування гіпотез, аналізу та інтерпретації результатів розрахунків і т.п.

Чисельні методи забезпечують системний формалізований підхід до вирішення математичних завдань. Проте за умов їх ефективного використання окрім уміння присутня і деяка частка мистецтва, яке залежить від здібностей користувача, оскільки для вирішення кожної математичної задачі існує декілька можливих чисельних методів і їх програмних реалізацій для різних типів комп'ютерів. На жаль, для вибору ефективного способу вирішення поставленого завдання лише інтуїції замало, потрібні глибокі знання та певні звички. Хоча існує безліч чисельних методів, всі вони (як і алгоритми, які їм відповідають) мають багато спільних властивостей і характеристик.

Чисельні методи у загальному випадку, [9]:

◆ передбачають проведення великої кількості рутинних арифметичних обчислень за допомогою

рекурсивних співвідношень, які використовуються для організації ітерацій, тобто повторюваних циклів обчислень зі зміненими початковими умовами для поліпшення результату;

- ◆ спрямовані на локальне спрощення задачі, коли, наприклад, використовувані нелінійні залежності лінеаризуються за допомогою своїх обчислених похідних або похідні замінюють різницевиими апроксимаціями;

- ◆ значно залежать від близькості початкового наближення (або декількох наближень), необхідного для початку обчислень;

- ◆ залежать від властивостей нелінійних функцій, які використовуються в математичних моделях, що накладає обмеження (для забезпечення єдиного розв'язку) на їх диференційованість, на швидкість зміни функцій та ін..

Чисельні методи **характеризуються**:

- ◆ різною швидкістю збіжності, тобто числом ітерацій, виконання яких необхідне для отримання заданої точності розв'язку;

- ◆ різною стійкістю, тобто збереженням достовірності розв'язку під час подальших ітерацій.

- ◆ різною точністю одержуваного розв'язку в разі виконання однакового числа ітерацій або циклів обчислень.

Чисельні методи **розрізняються**:

- ◆ по широті та легкості застосування, тобто за ступенем своєї універсальності й інваріантності для розв'язування різних математичних завдань;

- ◆ за складністю їх програмування;

◆ по можливостям використання в разі їх реалізації наявних бібліотек функцій і процедур, створених для підтримки різних алгоритмічних мов;

◆ за ступенем чутливості до погано обумовлених (або некоректних) математичних задач, коли малим змінам вхідних даних можуть відповідати великі зміни розв'язку.

2.2 Класифікація похибок обчислень

Під час чисельного розв'язання математичних і прикладних задач майже завжди на тому чи іншому етапі виникають похибки таких трьох типів, [10].

1. Похибка задачі. Вона пов'язана з наближеним характером вихідної моделі (зокрема, з неможливістю врахувати всі чинники в процесі вивчення модельованого явища). Крім того, параметрами математичного опису моделі є наближені числа (наприклад, через неможливість виконання абсолютно точних вимірювань). Для обчислювача похибка завдання являється *неусувною* (*безумовною*), хоча постановник завдання іноді може її змінити.

Іноді неусувну похибку поділяють на дві частини:

а) похибка, яка є наслідком невідповідності математичного опису задачі до реальності, яку називають *похибкою математичної моделі*.

б) похибка, яка є наслідком неточності завдання чисельних даних, які входять в математичний опис задачі, яку називають *похибкою параметрів*; ця складова похибки може бути зменшена за рахунок, наприклад, підвищення точності вимірювань.

2. Похибка методу. Ця похибка пов'язана з технікою розв'язування сформульованої математичної задачі. Вона з'являється унаслідок заміни початкової математичної моделі іншою або кінцевою послідовністю спрощених, наприклад, лінійних моделей. У разі створення чисельних методів є можливість відстеження таких похибок і зведення їх до як завгодно малого рівня. Тому похибка методу є *усуненою (умовною)*.

3. Похибка округлень (похибка операцій). Цей тип похибок обумовлений необхідністю виконувати арифметичні операції над числами, одержаними усіканням до кількості розрядів, які підтримуються обчислювальною технікою.

2.3 Абсолютна і відносна похибки.

Точні десятинні знаки. Вірні знаки числа

Кожне позитивне число A може бути зображено у вигляді кінцевої або нескінченної десяткового дробу, [11],

$$A = k_1 10^m + k_2 10^{m-1} + \dots + k_n 10^{m-n+1} + \dots,$$

де k_1, k_2, \dots, k_n – десятинні знаки (цифри) числа A , $k_1 \neq 0$.

Однак у випадку розв'язання завдань на комп'ютерах ми можемо використовувати лише числа з кінцевою і заздалегідь визначеною кількістю розрядів. Припустимо, що для зображення чисел використовується n розрядів. Тоді наближене число a , що представляє точне число A , матиме наступний вигляд:

$$a = k_1 10^m + k_2 10^{m-1} + \dots + k_n 10^{m-n+1},$$

де $k_1 \neq 0$, n – кількість десятинних знаків.

Значущими цифрами будемо вважати всі цифри в запису числа, починаючи з першої ненульової зліва.

Наприклад:

$$a_1 = 0,000105 \text{ – три значущих цифри,}$$

$$a_2 = 20 \text{ – дві значущих цифри,}$$

$$a_3 = 3000 \text{ – чотири значущих цифри,}$$

$$a_4 = 3,0 \cdot 10^3 \text{ чи } 30 \cdot 10^2 \text{ – дві значущих цифри.}$$

Якщо позначити через A – точне значення деякої величини, а через a – його наближене значення, то *абсолютною похибкою* наближеного числа називають величину:

$$\Delta = |A - a|.$$

Оскільки точне значення A невідомо, то користуються *граничною абсолютною похибкою*:

$$\Delta = |A - a| \leq \Delta_a. \quad (2.1)$$

Інтервал, в якому знаходиться точне значення A :

$$a - \Delta_a \leq a + \Delta_a.$$

Основною характеристикою точності наближеного числа є його *відносна похибка*:

$$\delta = \frac{\Delta}{|A|}.$$

Оскільки точне значення A невідомо, то користуються наступною формулою:

$$\delta = \frac{\Delta}{|a|}. \quad (2.2)$$

Відносна похибка числа пов'язана з його кількістю вірних знаків.

Значуща цифра називається *вірною*, якщо відносна похибка числа не перевищує $\frac{1}{2}$ одиниці розряду, що відповідає цій цифрі.

Пояснимо детальніше порядок визначення вірних значущих цифр:

- 1) обчислюємо відносну похибку наближеного числа;
- 2) округлюємо значення відносної похибки до розряду перед першою значущою цифрою;
- 3) кількість нулів після коми перед округленим значенням відносної похибки буде кількістю вірних цифр наближеного числа.

Приклад 2.1. Яка точність використання в обчисленнях числа $\pi = 3,14$?

Точне значення $\pi = 3,1415926\dots$

Тоді абсолютна похибка:

$$\Delta_{\pi} = |3,1415926 - 3,14| = 0,0015926 \approx 0,0016.$$

Відносна похибка:

$$\delta_{\pi} = 0,0016 / 3,14 = 0,00051.$$

Округлюємо відносну похибку до трьох знаків після коми $\delta_{\pi}^* = 0,001$. Відповідно наближене число 3,14 містить 2 вірних значущих цифри числа π .

Тепер розглянемо випадок використання числа 3,1416:

$$\Delta_{\pi} = |3,1415926 - 3,1416| = 7,4 \times 10^{-6},$$

$$\delta_{\pi} = 7,4 \times 10^{-6} / 3,1416 = 2,355 \times 10^{-6} = 0,02355 \times 10^{-4},$$

$$\delta_{\pi}^* = 0,02 \times 10^{-4} = 0,000002.$$

Відповідно наближене число 3,1416 містить 5 вірних значущих цифри числа π .

Будь-яке число можна записати у вигляді:

$$a \cdot 10^b,$$

така форма запису називається нормалізованою.

В нормалізованому науковому записі, порядок b вибирається такий, щоб абсолютна величина a залишалась не менше одиниці, але строго менше десяти.

Приклад 2.2. Наведіть нормалізований запис для чисел 1234 та 0,00057?

Розв'язок: $1,234 \times 10^3$ та $5,7 \times 10^{-4}$.

2.4 Похибки вимірювань

Результати вимірювань ніколи не бувають точними. Їх завжди отримують з деякою похибкою.

При вимірюванні слід врахувати три умови.

1. При прямих вимірюваннях приладом максимально можлива абсолютна похибка Δ – дорівнює ціні поділки шкали приладу, якщо значення фізичної величини збігається зі штрихом шкали; дорівнює подвійній ціні поділки шкали, якщо показчик не збігається зі штрихом шкали або коливається.

2. Записувати результати прямих вимірювань слід з указівкою похибок, при цьому:

1) похибку слід округлювати з надлишком до однієї значущої цифри;

2) числові значення результатів вимірювань беруться так, щоб їх остання цифра була в тому самому розряді, що й цифра похибки.

Приклад 2.3. При вимірюванні довжини прямокутника лінійкою отримали 10 см. Як правильно записати результат вимірювання?

Похибка вимірювання дорівнює ціні поділки лінійки $\Delta = 1$ мм. Тому результат вимірювань записуємо наступним чином:

$$100 \pm 1 \text{ мм.}$$

2.5 Абсолютна та відносна похибка обчислення елементарних арифметичних операцій

Важливою проблемою при проведенні обчислень з використанням наближених чисел є питання про вплив похибки вихідних даних на похибку отриманого результату.

Пряма задача теорії похибок полягає в наступному: відомі похибки деякої системи величин, вимагається визначити похибку даної функції від цих величин. На практиці важлива також *обернена задача теорії похибок*: які повинні бути абсолютні похибки аргументів функції, щоб абсолютна похибка функції не перевищувала заданої величини. Ця задача математично невизначена, оскільки задану граничну похибку функції можна забезпечити, встановлюючи по-різному граничні абсолютні похибки її аргументів.

Якщо для безперервної функції,

$$y = f(x_1, x_2 \dots x_n) = f(x_i),$$

відомі абсолютна та відносна похибка вхідних факторів – Δx_i та δx_i , тоді, згідно [8], абсолютна та відносна похибки функції обчислюються наступним чином:

$$\Delta y \approx \left| \sum_{i=1}^n \frac{\partial y}{\partial x_i} \Delta x_i \right|, \quad (2.3)$$

$$\delta y \approx \sum_{i=1}^n \left| \frac{\partial y}{\partial x_i} \right| \delta x_i = \frac{\Delta y}{|y|}. \quad (2.4)$$

Розглянемо отримання похибок для основних математичних операцій на прикладі функції двох змінних – $y = f(x_1, x_2)$.

Сума $y = x_1 + x_2$:

$$\begin{aligned}\Delta y &\approx \left| \sum_{i=1}^n \frac{\partial y}{\partial x_i} \Delta x_i \right| = \left| \frac{\partial y}{\partial x_1} \Delta x_1 + \frac{\partial y}{\partial x_2} \Delta x_2 \right| = \\ &= |1 \cdot \Delta x_1 + 1 \cdot \Delta x_2| = \Delta x_1 + \Delta x_2\end{aligned}$$

$$\delta y \approx \sum_{i=1}^n \left| \frac{\partial y}{\partial x_i} \right| \delta x_i = \left| \frac{\partial y}{\partial x_1} \right| \delta x_1 + \left| \frac{\partial y}{\partial x_2} \right| \delta x_2 = \delta x_1 + \delta x_2,$$

$$\delta y \approx \frac{\Delta y}{|y|} = \frac{\Delta x_1 + \Delta x_2}{|x_1 + x_2|}.$$

Добуток $y = x_1 x_2$:

$$\begin{aligned}\Delta y &\approx \left| \sum_{i=1}^n \frac{\partial y}{\partial x_i} \Delta x_i \right| = \left| \frac{\partial y}{\partial x_1} \Delta x_1 + \frac{\partial y}{\partial x_2} \Delta x_2 \right| = \\ &= |x_2 \Delta x_1 + x_1 \Delta x_2|,\end{aligned}$$

$$\begin{aligned}\delta y &\approx \sum_{i=1}^n \left| \frac{\partial y}{\partial x_i} \right| \delta x_i = \left| \frac{\partial y}{\partial x_1} \right| \delta x_1 + \left| \frac{\partial y}{\partial x_2} \right| \delta x_2 = \\ &= |x_2 \delta x_1 + x_1 \delta x_2|,\end{aligned}$$

$$\delta y \approx \frac{\Delta y}{|y|} = \frac{|x_2 \cdot \Delta x_1 + x_1 \cdot \Delta x_2|}{|x_1 x_2|} = \left| \frac{\Delta x_1}{x_1} + \frac{\Delta x_2}{x_2} \right|.$$

Ділення $y = \frac{x_1}{x_2}$:

$$\begin{aligned}\Delta y &\approx \left| \sum_{i=1}^n \frac{\partial y}{\partial x_i} \Delta x_i \right| = \left| \frac{\partial y}{\partial x_1} \Delta x_1 + \frac{\partial y}{\partial x_2} \Delta x_2 \right| = \left| \frac{1}{x_2} \Delta x_1 - \frac{x_1}{(x_2)^2} \Delta x_2 \right| = \\ &= \left| \frac{x_2 \Delta x_1 - x_1 \Delta x_2}{(x_2)^2} \right|,\end{aligned}$$

$$\begin{aligned} \delta y &\approx \sum_{i=1}^n \left| \frac{\partial y}{\partial x_i} \right| \delta x_i = \left| \frac{\partial y}{\partial x_1} \right| \delta x_1 + \left| \frac{\partial y}{\partial x_2} \right| \delta x_2 = \\ &= \left| \frac{1}{x_2} \delta x_1 - \frac{x_1}{(x_2)^2} \delta x_2 \right| = \left| \frac{x_2 \delta x_1 - x_1 \delta x_2}{(x_2)^2} \right|, \\ \delta y &\approx \frac{\Delta y}{|y|} = \left| \frac{x_2 \Delta x_1 + x_1 \Delta x_2}{(x_2)^2} \right| \left| \frac{x_2}{x_1} \right| = \left| \frac{x_2 \Delta x_1 + x_1 \Delta x_2}{x_1 x_2} \right|. \end{aligned}$$

Зведемо отримані результати до таблиць.

Таблиця 2.1 – Обчислення абсолютних похибок функцій двох змінних

| Функція | Абсолютна похибка |
|-----------------------|---|
| $y = x_1 + x_2$ | $\Delta y \approx \Delta x_1 + \Delta x_2$ |
| $y = x_1 x_2$ | $\Delta y \approx x_2 \Delta x_1 + x_1 \Delta x_2 $ |
| $y = \frac{x_1}{x_2}$ | $\Delta y \approx \left \frac{x_2 \Delta x_1 - x_1 \Delta x_2}{(x_2)^2} \right $ |

Таблиця 2.2 – Обчислення відносних похибок функцій двох змінних

| Функція | Відносна похибка | |
|-----------------------|---|---|
| $y = x_1 + x_2$ | $\delta y \approx \delta x_1 + \delta x_2$ | $\delta y \approx \frac{\Delta x_1 + \Delta x_2}{ x_1 + x_2 }$ |
| $y = x_1 x_2$ | $\delta y \approx x_2 \delta x_1 + x_1 \delta x_2 $ | $\delta y \approx \left \frac{\Delta x_1}{x_1} + \frac{\Delta x_2}{x_2} \right $ |
| $y = \frac{x_1}{x_2}$ | $\delta y \approx \left \frac{x_2 \delta x_1 - x_1 \delta x_2}{(x_2)^2} \right $ | $\delta y \approx \left \frac{x_2 \cdot \Delta x_1 + x_1 \cdot \Delta x_2}{x_1 x_2} \right $ |

Наведемо основні правила поширення абсолютної та відносної похибки, [31]:

1. Гранична абсолютна похибка як суми, так і різниці декількох наближених чисел дорівнює сумі граничних абсолютних похибок доданків.

2. Гранична відносна похибка добутку або відношення наближених чисел дорівнює сумі граничних відносних похибок співмножників.

3. Гранична абсолютна похибка обчислення функції дорівнює добутку абсолютної величини її похідної на граничну абсолютну похибку аргументу.

Розглянемо приклад обчислення похибок для функцій двох змінних.

Приклад 2.4. Прикладена напруга і струм через резистор склали 100 В і 10 А з точністю до 0,5 од. Оцінити похибку в обчисленні опору резистора за формулою:

$$R = U/I = 100/10 = 10 \text{ Ом.}$$

За умовами задачі $\Delta U = 0,5$ В и $\Delta I = 0,5$ А.

Знайдемо абсолютну та відносну похибки обчислення активного опору:

$$\Delta R = \frac{|I\Delta U - U\Delta I|}{(I)^2} = \frac{|10 \times 0,5 - 100 \times 0,5|}{10^2} = 0,45 ;$$

$$\delta R = \frac{|I\Delta U + U\Delta I|}{UI} = \frac{|10 \times 0,5 + 100 \times 0,5|}{100 \times 10} = 0,045 .$$

При обчисленнях над наближеними числами похибка результату залежить від порядку обчислень.

Приклад 2.5. Обчислити вираз $(1 \times 6) / 6$ з точністю до двох знаків після коми.

$$1,00 \times 6,00 = 6,00$$

$$6,00/6,00=1,00$$

Оскільки $(1 \times 6)/6 = (1/6) \times 6$, то

$$1,00/6,00=0,17;$$

$$0,17 \times 6,00=1,02$$

У деяких завданнях алгоритм обчислень може виявитися нестійким до погрешностей округлення.

Приклад 2.6. Задано систему рівнянь,

$$10,1x + 9,9y = 20,0$$

$$9,9x + 10,1y = 20,0$$

Ця система має розв'язок $(1; 1)$.

Припустимо, що вихідні дані отримані з похибкою 0,5 %, тоді система рівнянь матиме вигляд:

$$10,1x + 9,9y = 20,1$$

$$9,9x + 10,1y = 19,9$$

а її розв'язок дорівнюватиме $(1,5; 0,5)$, що відхиляється від точного розв'язку на 50%.

Щоб уникнути накопичення помилок, необхідно застосовувати деякі рекомендації з вибору кращої послідовності обчислень.

Наприклад:

- вибирати той алгоритм, який вимагає меншої кількості обчислень;
- вибирати чисельний метод, стійкий до похибок заокруглень;
- уникати віднімання та ділення на різницю близьких по розряду чисел, перетворюючи формулу обчислень, наприклад

$$(a+b)^2 - a^2 = 2ab + b^2 \quad \text{при } a \gg b,$$

$$\frac{1}{\sqrt{a^2 + b^2} - a} = \frac{\sqrt{a^2 + b^2} - a}{b^2}.$$

Питання для самоконтролю

1. Які основні джерела похибок?
2. Як за записом наближеного числа оцінити його абсолютну і відносну похибки?
3. Що таке значуща цифра числа?
4. Що називається вірною цифрою числа?
5. Як правильно записувати результат наближених обчислень або вимірювань?
6. Що становить пряму і обернену задачі теорії похибок?
7. Як визначається похибка наближеного значення функції багатьох змінних?
8. Як отримати формули для визначення похибок арифметичних операцій?
9. Що таке нормалізована форма запису числа?

Варіанти контрольних завдань до лекції 2

Таблиця 2.3 – Варіанти завдань

| № зп | Завдання 1 | Завдання 2 |
|------|--|--|
| 1 | Округляючи число 2,1514 до 3 значущих цифр визначити його абсолютну і відносну похибку. | Дано числа $a = 1,137$ і $b = 1,073$ з абсолютними похибками 0,011. Оцінити абсолютну та відносну похибки їх різниці. |
| 2 | Округляючи число 0,16152 до 4 значущих цифр визначити його абсолютну і відносну похибку. | Дано числа $a = 9,147$ і $b = 0,173$ з абсолютними похибками 0,01. Оцінити абсолютну та відносну похибки їх суми. |
| 3 | Округляючи число 0,01284 до 2 значущих цифр визначити його абсолютну і відносну похибку. | Виміряні значення напруги і струму рівні $U = 200$ В, $I = 25$ А. Похибка вимірювань відповідно 4 і 2%. Оцініть точність визначення потужності кола. |

| № зп | Завдання 1 | Завдання 2 |
|------|---|--|
| 4 | Округляючи число 1,225 до 3 значущих цифр визначити його абсолютну і відносну похибку. | Радіус кола склав 12 см з точністю 4 мм. Оцініть похибки обчислення площі круга. |
| 5 | Округляючи число 392,85 до 3 значущих цифр визначити його абсолютну і відносну похибку. | Виміряний значення струму на резисторі $R = 20$ Ом одно $I = 5$ А. Похибка вимірювання дорівнює 4%. Клас точності виготовлення резистора – 10%. Оцініть точність визначення потужності кола. |
| 6 | Округляючи число 0,15458 до 3 значущих цифр визначити його абсолютну і відносну похибку. | Виміряне значення напруги на резисторі $R = 10$ Ом $U = 75$ В. Похибка вимірювання дорівнює 2%. Клас точності виготовлення резистора – 5%. Оцініть точність визначення потужності кола. |
| 7 | Визначити абсолютну похибку наближеного числа 1,3267 за його відносною похибкою $\delta = 0,1$ %. | Сторони прямокутника дорівнюють 10,5 на 5,5 м з точністю 5 см. Оцініть точність визначення площі прямокутника. |
| 8 | Визначити абсолютну похибку наближеного числа 2,32 за його відносною похибкою $\delta = 0,7$ %. | Довжина ребра куба склала 8 см з точністю 4 мм. Оцініть точність визначення площі поверхні куба. |
| 9 | Визначити абсолютну похибку наближеного числа 35,72 за його відносною похибкою $\delta = 1$ %. | Довжина ребра куба склала 12 см з точністю 5 мм. Оцініть точність визначення площі обсягу куба. |
| 10 | Визначити абсолютну похибку наближеного числа 0,896 за його відносною похибкою $\delta = 5$ %. | Активний опір кола – 0,5 Ом з класом точності 5%. Індуктивність кола – 4,5 мГн з точністю 10%. Оцініть точність обчислення постійної часу кола $T = L / R$ |

| № зп | Завдання 1 | Завдання 2 |
|------|---|--|
| 11 | Визначити абсолютну похибку наближеного числа 232,44 за його відносною похибкою $\delta = 1\%$. | Радіус кола склав 14 см з точністю 5 мм. Оцініть похибка обчислення площі круга. |
| 12 | Визначити абсолютну похибку наближеного числа 94,525 за його відносною похибкою $\delta = 3\%$. | Виміряне значення струму на резисторі $R = 10 \text{ Ом}$, $I = 7 \text{ А}$. Похибка вимірювання дорівнює 2%. Клас точності виготовлення резистора – 6%. Оцініть точність визначення потужності кола. |
| 13 | Визначити кількість вірних знаків наближеного числа 1,3267 за його відносною похибки $\delta = 0,1\%$. | Виміряне значення напруги на резисторі $R = 18 \text{ Ом}$ $U = 45 \text{ В}$. Похибка вимірювання дорівнює 3%. Клас точності виготовлення резистора – 8%. Оцініть точність визначення потужності кола. |
| 14 | Визначити кількість вірних знаків наближеного числа 2,32 за його відносною похибкою $\delta = 0,7\%$. | Сторони прямокутника дорівнюють 5,5 на 8,5 м з точністю 2 см. Оцініть точність визначення площі прямокутника. |
| 15 | Визначити кількість вірних знаків наближеного числа 35,72 за його відносною похибкою $\delta = 1\%$. | Довжина ребра куба склала 12 см з точністю 5 мм. Оцініть точність визначення площі поверхні куба. |
| 16 | Визначити кількість вірних знаків наближеного числа 0,896 за його відносною похибкою $\delta = 5\%$. | Довжина ребра куба склала 8 см з точністю 2 мм. Оцініть точність визначення площі обсягу куба. |
| 17 | Визначити кількість вірних знаків наближеного числа 232,44 за його відносною похибкою $\delta = 1\%$. | Активний опір кола – 0.25 Ом з класом точності 10%. Індуктивність кола – 3,5 мГн з точністю 4%. Оцініть точність обчислення постійної часу кола $T = L / R$ |

| № зп | Завдання 1 | Завдання 2 |
|------|---|--|
| 18 | Визначити кількість вірних знаків наближеного числа 94,525 за його відносною похибкою $\delta = 3\%$. | Катети прямокутного трикутника рівні 5,5 на 8,5 м з точністю 2 см. Оцініть точність визначення площі трикутника. |
| 19 | Визначити кількість вірних знаків наближеного числа 0,3941 за його абсолютною похибкою $\Delta x = 0,0025$. | Радіус кола склав 11 см з точністю 3 мм. Оцініть похибка обчислення площі круга. |
| 20 | Визначити кількість вірних знаків наближеного числа 0,1132 за його абсолютною похибкою $\Delta x = 0,001$. | Виміряне значення струму на резисторі $R = 25 \text{ Ом}$, $I = 11 \text{ А}$. Похибка вимірювання дорівнює 3%. Клас точності виготовлення резистора – 2%. Оцініть точність визначення потужності кола. |
| 21 | Визначити кількість вірних знаків наближеного числа 38,2543 за його абсолютною похибкою $\Delta x = 0,0027$. | Виміряне значення напруги на резисторі $R = 28 \text{ Ом}$, $U = 85 \text{ В}$. Похибка вимірювання дорівнює 4%. Клас точності виготовлення резистора – 1%. Оцініть точність визначення потужності кола. |
| 22 | Визначити кількість вірних знаків наближеного числа 293,481 за його абсолютною похибкою $\Delta x = 0,2$. | Сторони прямокутника дорівнюють 1.5 на 2.5 м з точністю 1 см. Оцініть точність визначення площі прямокутника. |
| 23 | Визначити кількість вірних знаків наближеного числа 2,325 за його абсолютною похибкою $\Delta x = 0,01$. | Довжина ребра куба склала 9 см з точністю 3 мм. Оцініть точність визначення площі поверхні куба. |
| 24 | Визначити кількість вірних знаків наближеного числа 14,00231 за його абсолютною похибкою $\Delta x = 0,004$. | Довжина ребра куба склала 11 см з точністю 4 мм. Оцініть точність визначення площі обсягу куба. |

| № зп | Завдання 1 | Завдання 2 |
|------|---|--|
| 25 | Визначити кількість вірних знаків наближеного числа 0,0842 за його абсолютною похибкою $\Delta x = 0,0015$. | Активний опір кола – 0.55 Ом з класом точності 6 %. Індуктивність кола – 2,5 мГн з точністю 5%. Оцініть точність обчислення постійної часу кола $T = L / R$ |
| 26 | Визначити кількість вірних знаків наближеного числа 0,00381 за його абсолютною похибкою $\Delta x = 0,0001$. | Катети прямокутного трикутника дорівнюють 55 на 85 см з точністю 8 мм. Оцініть точність визначення площі трикутника. |
| 27 | Визначити кількість вірних знаків наближеного числа 32,285 за його абсолютною похибкою $\Delta x = 0,02$. | Висота підйому вантажу – 12м з точністю 40 см. Швидкість підйому вантажу 45 см/с з похибкою 10 %. Оцініть точність визначення часу підйому вантажу. |
| 28 | Визначити кількість вірних знаків наближеного числа 0,2113 за його абсолютною похибкою $\Delta x = 0,05$. | Вага кранового візка становить 3,5 тонни. Швидкість руху візка – 0,5 м/с. Вважаючи, що точність вимірювань дорівнює 5 %, визначте похибку обчислення кінетичної енергії візка. |
| 29 | Визначити кількість вірних знаків наближеного числа 22,351 за його абсолютною похибкою $\Delta x = 0,0025$. | Вимірний значення струму на резисторі $R = 45$ Ом одно $I = 9$ А. Похибка вимірювання дорівнює 6%. Клас точності виготовлення резистора - 8%. Оцініть точність визначення потужності кола. |
| 30 | Визначити кількість вірних знаків наближеного числа 293,481 за його абсолютною похибкою $\Delta x = 0,25$. | Вимірне значення напруги на резисторі $R = 18$ Ом, $U = 28$ В. Похибка вимірювання дорівнює 3 %. Клас точності виготовлення резистора – 8%. Оцініть точність визначення потужності кола. |

ЛЕКЦІЯ 3

МЕТОДИ РОЗВ'ЯЗАННЯ СИСТЕМ ЛІНІЙНИХ АЛГЕБРАЇЧНИХ РІВНЯНЬ

Разносчик продал одному покупателю 10 яблук, 5 груш и 3 лимона за 1 рубль 10 копеек; другому покупателю по той же цене он продал 10 яблук, 3 груши и 1 лимон за 78 копеек; третьему покупателю он продал по той же цене 2 груши и 1 лимон за 22 копейки. Почем он продавал в отдельности яблоко, грушу и лимон?

"Арифметика". XIX век.

Задача розв'язання системи лінійних алгебраїчних рівнянь (СЛАР) має важливе прикладне значення при розв'язанні різноманітних проблем науки і техніки.

Розрахунки сталих режимів в електричних схемах в кінцевому підсумку зводяться до складання і розв'язку СЛАР. До розв'язання СЛАР зводиться й задача розрахунку зусиль в опорах механізму при статичному навантаженні у механіці. Крім того, завдання лінійної алгебри є допоміжними при реалізації багатьох алгоритмів обчислювальної математики, обробки результатів експериментальних досліджень.

Задача розв'язання СЛАР виникає при знаходженні канонічних інтерполяційних поліномів, при вирішенні завдань апроксимації функцій за допомогою методу найменших квадратів.

Необхідність розв'язання СЛАР високих порядків виникає в процесі реалізації алгоритмів розв'язання рівнянь в часткових похідних методами кінцевих елементів.

Постановка задачі. Необхідно вирішити СЛАР виду

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

де x_k – невідомі величини, $k = 1..n$; a_{ij} – задані елементи розширеної матриці системи рівнянь, відомі числові коефіцієнти.

СЛАР може бути представлена у вигляді:

$$\sum_{i=1}^n a_{ij}x_k = b_i, \quad i, j = 1..n. \quad (3.2)$$

СЛАР часто записують у матричній формі:

$$A \cdot X = B, \quad (3.3)$$

де A – матриця коефіцієнтів системи розмірності $n \times n$; B – вектор-стовпець вільних членів довжиною n ; X – вектор-стовпець невідомих довжиною n :

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

Найчастіше в ході розв'язування СЛАР використовують матриці. Виділяють такі види матриць:

- *квадратна* (матриця, в якій співпадає кількість рядків і стовпців);

- *діагональна* (квадратна матриця, для якої всі недіагональні елементи рівні нулю);
- *одинична* (діагональна матриця, для якої всі діагональні елементи дорівнюють одиниці);
- *нульова* (всі елементи дорівнюють нулю);
- *матриці-стовпці та матриці-рядки*;
- *розширена* (отримана поєднанням матриць коефіцієнтів і вільних членів).

Перед тим, як приступати до розв'язування СЛАР, необхідно переконатись, що система взагалі має розв'язок.

Існують наступні ознаки *однорідності* системи:

1. *Правило Крамера*: якщо визначник системи відмінний від нуля, $\Delta \neq 0$, система має єдиний розв'язок.

2. *Теорема Кронекера-Капеллі*: якщо ранг матриці коефіцієнтів дорівнює рангу розширеної матриці $\text{rank}(A) = \text{rank}(D)$, система має єдиний розв'язок.

СЛАР, що має один і більше розв'язків називається *сумісною*. Сумісна система, яка має лише один розв'язок називається *однорідною*.

Методи розв'язування систем лінійних алгебраїчних рівнянь ділять на три класи: прямі, ітераційні, методи лінійного програмування (рис. 3.1).

Для розв'язання СЛАР застосовують в основному два класи методів: прямі й ітераційні. Прямі методи є універсальними і застосовуються для розв'язання систем порівняно невисокого порядку ($n \sim 200$). Ітераційні методи вигідно використовувати для СЛАР високого порядку зі слабо заповненими матрицями.

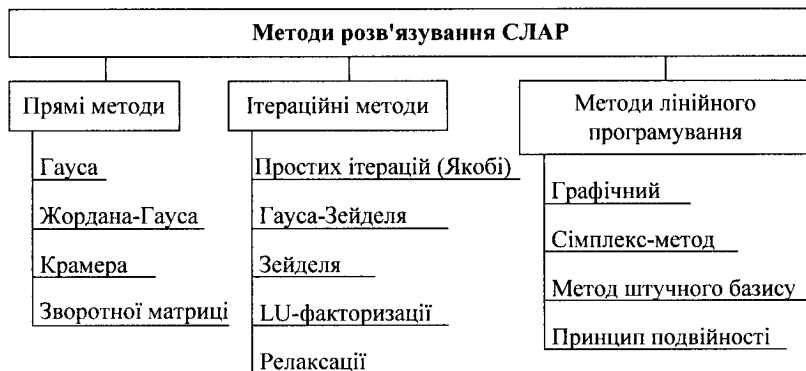


Рисунок 3.1 – Класифікація методів розв'язування СЛАР

3.1 Прямі методи розв'язання СЛАР

Прямі методи є аналітичними, а тому забезпечують точний розв'язок СЛАР.

До прямих методів належать метод зворотної матриці, метод Гауса та метод Крамера (метод визначників).

Розглянемо детальніше *метод зворотної матриці*.

Якщо СЛАР записати в матричній формі:

$$A \cdot X = B,$$

то, множачи обидві частини цієї рівності на зворотну матрицю A^{-1} , отримаємо розв'язок у вигляді:

$$X = A^{-1}B, \tag{3.4}$$

де A^{-1} – зворотна матриця,

$$A^{-1} = \frac{1}{|A|} C^T;$$

C^T – транспонована союзна матриця, елементи якої визначаються наступним чином,

Ще одним прикладом прямих методів є метод послідовного виключення, або *метод Гауса*.

Застосування методу Гауса полягає в наступному, [13]. Перше рівняння системи (3.1) ділимо на коефіцієнт a_{11} . Отримаємо наступне рівняння:

$$x_1 + c_{12}x_2 + \dots + c_{1n}x_n = y_1,$$

$$\text{де } c_{12} = \frac{a_{12}}{a_{11}}, y_1 = \frac{b_1}{a_{11}}.$$

З цього рівняння нескладно виділити змінну x_1 і підставити отриманий вираз в інші рівняння системи.

Цей процес можна послідовно застосовувати до решти $(n-1)$ рівнянь вихідної системи рівнянь.

У підсумку такого процесу, який називається *прямим ходом* метода Гауса, система рівнянь перетворюється до наступного вигляду:

$$\begin{cases} x_1 + c_{12}x_2 + \dots + c_{1n}x_n = y_1, \\ 0 + x_2 + \dots + c_{2n}x_n = y_2, \\ \dots \\ 0 + 0 + \dots + x_n = y_n. \end{cases} \quad (3.5)$$

Останнє рівняння дає нам розв'язок для змінної x_n .

Виконуючи *зворотний хід* по рівняннях системи, ми послідовно можемо обчислити всі змінні системи рівнянь.

Суть методу Гауса полягає в перетворенні вихідної матриці коефіцієнтів СЛАР в нижню трикутну матрицю, у якій всі елементи, розташовані нижче головної діагоналі, дорівнюють нулю.

Точність результатів буде визначатися точністю виконання арифметичних операцій при перетворенні елементів матриці. Для зменшення похибки при діленні на

діагональний елемент рекомендується здійснити таку перестановку рівнянь, щоб поставити на діагональ найбільший по модулю з усіх елементів розглянутого рівняння. Така процедура називається *вибором головного елемента стовпця*.

Кількість арифметичних операцій в методі Гауса пов'язана з розмірністю системи n і дорівнює $2/3 n^3$.

Приклад 3.2. Аналітичне розв'язання СЛАР другого порядку методом Гауса:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}.$$

З першого рівняння знаходимо:

$$x_1 = -\frac{a_{12}x_2}{a_{11}} + \frac{b_1}{a_{11}}.$$

І підставляємо в друге рівняння:

$$a_{21}\left(-\frac{a_{12}x_2}{a_{11}} + \frac{b_1}{a_{11}}\right) + a_{22}x_2 = b_2.$$

Розкриваємо дужки і групуємо:

$$x_2\left(a_{22} - \frac{a_{12}a_{21}}{a_{11}}\right) = b_2 - \frac{a_{21}b_1}{a_{11}}.$$

Звідки знаходимо x_2 :

$$x_2 = \left(b_2 - \frac{a_{21}b_1}{a_{11}}\right) / \left(a_{22} - \frac{a_{12}a_{21}}{a_{11}}\right),$$

$$x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}} \cdot \frac{a_{11}}{a_{11}a_{22} - a_{12}a_{21}},$$

$$x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}}.$$

Підставляємо отриманий вираз для x_2 у перше рівняння системи:

$$a_{11}x_1 + a_{12} \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}} = b_1.$$

Виразимо x_1 :

$$x_1 = \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}} \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}}.$$

Зведемо під загальний знаменник, розкриємо дужки та спростимо коефіцієнти:

$$x_1 = \frac{b_1(a_{11}a_{22} - a_{12}a_{21}) - a_{12}(a_{11}b_2 - a_{21}b_1)}{a_{11}(a_{11}a_{22} - a_{12}a_{21})},$$

$$x_1 = \frac{a_{11}a_{22}b_1 - a_{12}a_{21}b_1 - a_{11}a_{12}b_2 + a_{12}a_{21}b_1}{a_{11}(a_{11}a_{22} - a_{12}a_{21})},$$

$$x_1 = \frac{a_{11}a_{22}b_1 - a_{11}a_{12}b_2}{a_{11}(a_{11}a_{22} - a_{12}a_{21})},$$

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}.$$

Отримали наступні вирази для обчислення невідомих x_1 і x_2 через відомі коефіцієнти системи рівнянь:

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}},$$

$$x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}}.$$

Як видно з прикладу 3.2 аналітичний розв'язок навіть системи другого порядку вимагає докладання багатьох зусиль.

Приклад 3.3. Розв'язання СЛАР третього порядку методом Гауса. Дана наступна СЛАР:

$$\begin{cases} 10x_1 - 2x_2 + x_3 = 9 \\ x_1 + 5x_2 - x_3 = 8 \\ 4x_1 + 2x_2 + 8x_3 = 32 \end{cases}.$$

При чисельному розв'язанні СЛАР методом Гауса користуються методом множення та складання різних рядків матриці таким чином, щоб отримати трикутну матрицю. Запишемо СЛАР у вигляді розширеної матриці, при чому у третьому рядку можна поділити всі елементи на 2:

$$D = \left[\begin{array}{ccc|c} 10 & -2 & 1 & 9 \\ 1 & 5 & -1 & 8 \\ 2 & 1 & 4 & 16 \end{array} \right].$$

Помножимо 2-ий рядок на -10 , а третій на -5 :

$$D = \left[\begin{array}{ccc|c} 10 & -2 & 1 & 9 \\ -10 & -50 & 10 & -80 \\ -10 & -5 & -20 & -80 \end{array} \right].$$

Складемо 2-ий і 3-ій рядки з 1-им:

$$D = \left[\begin{array}{ccc|c} 10 & -2 & 1 & 9 \\ 0 & -52 & 11 & -71 \\ 0 & -7 & -19 & -71 \end{array} \right].$$

Помножимо 2-ий рядок на 7 , 3-ій на -52 , і складемо між собою:

$$D = \left[\begin{array}{ccc|c} 10 & -2 & 1 & 9 \\ 0 & -364 & 77 & -497 \\ 0 & -364 & 988 & 3692 \end{array} \right] = \left[\begin{array}{ccc|c} 10 & -2 & 1 & 9 \\ 0 & -364 & 77 & -497 \\ 0 & 0 & 1065 & 3195 \end{array} \right].$$

Визначимо невідомі:

$$x_3 = \frac{3195}{1065} = 3,$$

$$x_2 = \frac{-497 - 77 \times 3}{-364} = \frac{-728}{-364} = 2,$$

$$x_1 = \frac{9 + 2 \times 2 - 1 \times 3}{10} = 1.$$

Менш трудомістким є *метод Крамера*. Значення невідомих визначаються за формулою:

$$X_i = \frac{\det A_i}{\det A},$$

де A_i – додаткова матриця, яка утворюється підстановкою замість відповідного стовпця значень матриці вільних членів:

$$A_1 = \begin{bmatrix} b_1 & a_{12} & \dots & a_{1n} \\ b_2 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ b_n & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad A_2 = \begin{bmatrix} a_{11} & b_1 & \dots & a_{1n} \\ a_{21} & b_2 & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & b_n & \dots & a_{nn} \end{bmatrix},$$

$$\dots, \quad A_n = \begin{bmatrix} a_{11} & a_{12} & \dots & b_1 \\ a_{21} & a_{22} & \dots & b_2 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & b_n \end{bmatrix}.$$

Цей метод вимагає обчислення $n+1$ визначників порядку n .

Приклад 3.4. Визначення струмів у гілках схеми методом Крамера. Для заданої схеми визначити струми в окремих гілках методом контурних струмів.

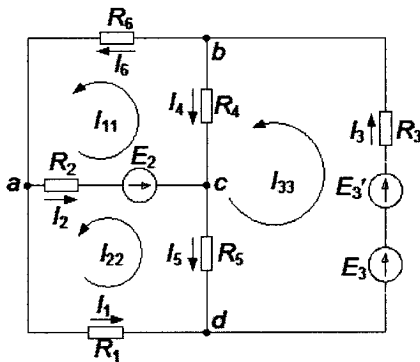


Рисунок 3.2 - Електрична схема

Параметри елементів схеми: $R_1=24$ Ом, $R_2=70$ Ом, $R_3=44$ Ом, $R_4=12$ Ом, $R_5=20$ Ом, $R_6=30$ Ом, $E_2=40$ В, $E_3=19,6$ В, $E_3'=4,4$ В.

Для цієї схеми система рівнянь для контурних струмів матиме вигляд:

$$\begin{cases} 112I_{11} - 70I_{22} - 12I_{33} = 40, \\ -70I_{11} + 114I_{22} - 20I_{33} = -40, \\ -12I_{11} - 20I_{22} + 76I_{33} = 24. \end{cases}$$

Знаходимо головний визначник системи:

$$\Delta = \begin{vmatrix} 112 & -70 & -12 \\ -70 & 114 & -20 \\ -12 & -20 & 76 \end{vmatrix} =$$

$$\begin{aligned} &= (112 \times 114 \times 76) + (-70) \times (-20) \times (-12) + (-70) \times (-20) \times (-12) - \\ &- (-12) \times 114 \times (-12) - 112 \times (-20) \times (-20) - (-70) \times (-70) \times 76 = \\ &= 503152. \end{aligned}$$

Допоміжні визначники:

$$\Delta_{11} = \begin{vmatrix} 40 & -70 & -12 \\ -40 & 114 & -20 \\ 24 & -20 & 76 \end{vmatrix} = 174592,$$

$$\Delta_{22} = \begin{vmatrix} 112 & 40 & -12 \\ -70 & -40 & -20 \\ -12 & 24 & 76 \end{vmatrix} = 38400,$$

$$\Delta_{33} = \begin{vmatrix} 112 & -70 & 40 \\ -70 & 114 & -40 \\ -12 & -20 & 24 \end{vmatrix} = 176352.$$

Тоді контурні струми дорівнюють:

$$I_{11} = \frac{\Delta_{11}}{\Delta} = \frac{174592}{503150} = 0,347 \text{ А};$$

$$I_{22} = \frac{\Delta_{22}}{\Delta} = \frac{38400}{503150} = 0,076 \text{ А};$$

$$I_{33} = \frac{\Delta_{33}}{\Delta} = \frac{176352}{503150} = 0,3505 \text{ А}.$$

Приклад 3.5. Розв'язання СЛАР методом Крамера в MS EXCEL. Візьмемо СЛАР з прикладу 3.1. Вихідні дані та необхідні формули:

| Коефіцієнти СЛАР | | | Вільні члени |
|------------------|------|------|--------------|
| 1,84 | 2,25 | 2,53 | -6,09 |
| 2,32 | 2,6 | 2,82 | -6,98 |
| 1,83 | 2,06 | 2,24 | -5,52 |

| Матриця обчислення D1 | | |
|-----------------------|------|------|
| -6,09 | 2,25 | 2,53 |
| -6,98 | 2,6 | 2,82 |
| -5,52 | 2,06 | 2,24 |

| Визначники | |
|------------|-------------------|
| Головний | =МОПРЕД (A2:C4) |
| D1 | =МОПРЕД (G2:I4) |
| D2 | =МОПРЕД (G7:I9) |
| D3 | =МОПРЕД (G12:I14) |

| Корені рівняння | |
|-----------------|---------|
| x1 | =B9/B8 |
| x2 | =B10/B8 |
| x3 | =B11/B8 |

| Матриця обчислення D2 | | |
|-----------------------|-------|------|
| 1,84 | -6,09 | 2,53 |
| 2,32 | -6,98 | 2,82 |
| 1,83 | -5,52 | 2,24 |

| Матриця обчислення D3 | | |
|-----------------------|------|-------|
| 1,84 | 2,25 | -6,09 |
| 2,32 | 2,6 | -6,98 |
| 1,83 | 2,06 | -5,52 |

Результати обчислень :

| Коефіцієнти СЛАР | | | Вільні члени |
|------------------|------|------|--------------|
| 1,84 | 2,25 | 2,53 | -6,09 |
| 2,32 | 2,6 | 2,82 | -6,98 |
| 1,83 | 2,06 | 2,24 | -5,52 |

| Матриця обчислення D1 | | |
|-----------------------|------|------|
| -6,09 | 2,25 | 2,53 |
| -6,98 | 2,6 | 2,82 |
| -5,52 | 2,06 | 2,24 |

| Визначники | |
|------------|---------|
| Головний | -0,0006 |
| D1 | -0,0031 |
| D2 | 0,01038 |
| D3 | -0,0055 |

| Корені рівнянь | |
|----------------|---------|
| x1 | 5,38832 |
| x2 | -17,828 |
| x3 | 9,52921 |

| Матриця обчислення D2 | | |
|-----------------------|-------|------|
| 1,84 | -6,09 | 2,53 |
| 2,32 | -6,98 | 2,82 |
| 1,83 | -5,52 | 2,24 |

| Матриця обчислення D3 | | |
|-----------------------|------|-------|
| 1,84 | 2,25 | -6,09 |
| 2,32 | 2,6 | -6,98 |
| 1,83 | 2,06 | -5,52 |

Розроблено багато інших прямих методів розв'язання СЛАР: метод LU-розкладання, метод Холецкого, метод обертань і багато інших, [54, 21, 58].

Основним недоліком прямих методів розв'язання СЛАР є дуже великий обсяг арифметичних операцій, виконуваних над дійсними числами. Так як обсяг обчислень практично експоненціально зростає з ростом порядку СЛАР, то помилки округлення можуть призводити до суттєвої неточності отриманих результатів. Для отримання задовільної точності рекомендується

уникати ситуацій, коли окремі коефіцієнти СЛАР значно (3-4 порядку) відрізняються один від одного.

Також обов'язковою процедурою є числова перевірка отриманого розв'язку.

При формальному підході розв'язання СЛАР не зустрічає утруднень: розв'язок системи можна знайти, розкривши визначники у формулі Крамера; для знаходження власних значень матриці досить виписати характеристичне рівняння і знайти його корні. Однак ці рекомендації зустрічають заперечення з багатьох сторін.

Так, при безпосередньому розкритті визначників розв'язок системи з n невідомими вимагає порядку $n(n+1)n!$ арифметичних операцій, [18]. Так, наприклад, для обчислення одного визначника при $n=100$ потрібно виконати 10^{157} операцій з дійсними числами. Таке число операцій за прийнятний час не в змозі виконати наймогутніші сучасні суперкомп'ютери, і навіть у далекій перспективі це буде неможливо.

Іншою причиною, по якій ці класичні способи незастосовні навіть при малих n , є сильний вплив на остаточний результат округлень при обчисленнях. Результат обчислень часто далекий від істинного значення через вплив обчислювальної похибки. Точно так само йде справа при знаходженні власних значень матриць з використанням явного виразу характеристичного багаточлена.

3.2 Ітераційні методи розв'язання СЛАР

Для отримання гарантованої точності одержуваного розв'язку, а також для розв'язання СЛАР високих (кілька

тисяч) розмірностей слід застосовувати ітераційні методи розв'язання СЛАР. Ітераційні методи відносяться до класу методів послідовного наближення.

Ітераційні методи – це методи послідовного наближення, які використовують повторюваний процес, доки різниця між результатами останніх кроків не стане меншою заданої точності ε ,

$$\left| X^{(k)} - X^{(k-1)} \right| \leq \varepsilon. \quad (3.6)$$

Ітераційні методи потребують додаткових умов для отримання збіжного процесу. Якщо в процесі розв'язку СЛАР ітераційним методом з кожним кроком відбувається наближення до коренів системи такий процес називається *збіжним*, в іншому випадку *розбіжним*,

$$\lim_{k \rightarrow \infty} X^{(k)} = X. \quad (3.7)$$

Ідея ітераційних методів розв'язання СЛАР полягає в наступному, [6].

Напишемо рівняння вихідної СЛАР у вигляді:

$$\begin{cases} x_1^{j+1} = \frac{1}{a_{11}} (a_{1,n+1} - a_{12}x_1^{j+1} - \dots - \dots^{j+1}) + x_1^j \\ x_2^{j+1} = \frac{1}{a_{22}} (a_{2,n+1} - a_{21}x_1^{j+1} - \dots - \dots^{j+1}) + x_2^j \\ \dots \\ x_n^{j+1} = \frac{1}{a_{nn}} (a_{n,n+1} - a_{n1}x_1^{j+1} - \dots - \dots^{j+1}) + x_n^j \end{cases}. \quad (3.8)$$

Задається стовпець початкових наближень $x_1^0, x_2^0, \dots, x_n^0$. Підставляючи ці значення в праві частини наведеної

вище системи рівнянь отримаємо нове наближення розв'язку $x_1^1, x_2^1, \dots, x_n^1$.

Повторюючи цей обчислювальний процес достатнє число разів, ми отримуємо на кожному кроці все більш точне значення розв'язку СЛАР.

При використанні ітераційних обчислювальних процесів будь-якого виду необхідно враховувати деякі особливості:

1. Ітераційний процес може бути розбіжним. Стійкість обчислювального процесу часто залежить від значень початкового наближення, вибраного для старту процесу. Необхідно контролювати стійкість і збіжність ітераційного процесу.

2. Ітераційний процес, навіть якщо він збігається, не обов'язково збігається до точного розв'язання вихідної системи. Тому необхідно обов'язково перевіряти отримані розв'язки шляхом прямої підстановки отриманих значень у вихідну СЛАР.

Умови збіжності ітераційного процесу:

1. Всі елементи перетвореної матриці коефіцієнтів за модулем менше 1.

2. Хоча б одна із норм матриці менша за 1.

Норма матриці – це функція, що узагальнює поняття довжини вектора та абсолютного значення числа.

Виділяють наступні норми матриці:

– *одинична норма*, або *L-норма* – максимальний елемент серед суми елементів стовпців матриці

$$\|A\|_1 = \max \left(\sum_j |a_{i,j}| \right);$$

– *нескінченна норма* – максимальний елемент серед суми рядків матриці,

$$\|A\|_{\infty} = \max \left(\sum_i |a_{i,j}| \right);$$

– *евклідова норма* – сума квадратів всіх елементів матриці,

$$\|A\|_E = \sqrt{\sum_{i,j=1}^n |a_{i,j}|^2}.$$

Приклад 3.6. Обчислимо норми для наведеної матриці,

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

Одинична норма: $\|A\|_1 = \max(5 \ 5) = 5$.

Нескінченна норма: $\|A\|_{\infty} = \max \begin{pmatrix} 7 \\ 3 \end{pmatrix} = 7$.

Евклідова норма: $\|A\|_E = \sqrt{4^2 + 3^2 + 1^2 + 2^2} = \sqrt{30} = 5,477$.

Метод Якоби (метод простих ітерацій).

Ітераційний процес описується наступним рівнянням:

$$X^{(k+1)} = HX^{(k)} + g, \quad (3.9)$$

де k – номер ітерації; ε – задана точність розв'язку; H , g

– перетворені матриці коефіцієнтів і вільних членів,

$$H = \begin{bmatrix} 0 & -a_{12}/a_{11} & \dots & -a_{1n}/a_{11} \\ -a_{21}/a_{22} & 0 & \dots & -a_{2n}/a_{22} \\ \vdots & \vdots & & \vdots \\ -a_{m1}/a_{mn} & -a_{m2}/a_{mn} & \dots & 0 \end{bmatrix}, \quad (3.10)$$

$$g = \begin{bmatrix} b_1/a_{11} \\ b_2/a_{22} \\ \vdots \\ b_m/a_{mn} \end{bmatrix}. \quad (3.11)$$

Приклад 3.7. Для наведеної СЛАР обчислити перше наближення за методом простих ітерацій,

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

Спочатку потрібно змінити порядок рівнянь, потім записати матриці коефіцієнтів і вільних членів:

$$\begin{cases} 4x_1 + 3x_2 = 1 \\ x_1 + 2x_2 = 4 \end{cases}, \quad A = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

Запишемо матриці H та g :

$$H = \begin{bmatrix} 0 & -\frac{3}{4} \\ -\frac{1}{2} & 0 \end{bmatrix}, \quad g = \begin{bmatrix} 1/4 \\ 2 \end{bmatrix}.$$

В якості початкових наближень приймається матриця $X^{(0)} = g$. Рівняння для обчислення першого наближення буде мати вигляд:

$$X^{(1)} = HX^{(0)} + g = \begin{bmatrix} 0 & -\frac{3}{4} \\ -\frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 1/4 \\ 2 \end{bmatrix} + \begin{bmatrix} 1/4 \\ 2 \end{bmatrix} = \begin{bmatrix} -5/4 \\ 15/8 \end{bmatrix} = \begin{bmatrix} -1,25 \\ 1,875 \end{bmatrix}.$$

Для підвищення збіжності та точності ітераційного процесу можна запропонувати наступний нескладний підхід. При обчисленні кожного ітераційного рівняння замість значень змінних на попередньому кроці розрахунків можна використовувати уточнені значення змінних, отримані з уже перетворених рівнянь поточного кроку розрахунків:

$$\begin{cases} x_1^{m+1} = f_1(x_1^m, x_2^m, \dots, x_{n-1}^m, x_n^m) \\ x_2^{m+1} = f_1(x_1^{m+1}, x_2^m, \dots, x_{n-1}^m, x_n^m) \\ x_3^{m+1} = f_1(x_1^{m+1}, x_2^{m+1}, \dots, x_{n-1}^m, x_n^m) \\ \dots \\ x_n^{m+1} = f_1(x_1^{m+1}, x_2^{m+1}, \dots, x_{n-1}^{m+1}, x_n^m) \end{cases} \quad (3.12)$$

Така зміна ітераційного процесу, відома як *метод Зейделя*, як правило призводить до прискорення збіжності.

3.3 Розв'язання СЛАР в пакеті MathCAD

Пакет MathCAD має достатній арсенал засобів роботи з матрицями.

Чисельний розв'язок СЛАР у пакеті MathCAD можна здійснити різноманітними засобами. Зокрема, СЛАР можна вирішити матричним методом за допомогою запису (3.4), а також вбудованими функціями: **Isolve(A,B)** та **rref(D)**, де параметри *A* і *B* відповідають параметрам рівняння (3.3). Функція **rref(D)** реалізує прямий та зворотний хід методу Гауса. Розглянемо приклади розв'язання СЛАР прямими методами.

Приклад 3.8. Завдання СЛАР у матричному виді та перевірка наявності розв'язку даної системи:

$$\begin{cases} 10x_1 - 70x_2 = 7, \\ -3x_1 + 2x_2 + 6x_3 = 4, \\ 5x_1 - 2x_2 + 5x_3 = 6. \end{cases}$$

Матрицю коефіцієнтів та вектор вільних членів створюємо через пункт меню Вставка/Матриця. Лістинг у пакеті MathCAD:

Задамо матрицю коефіцієнтів A та вільних членів B

$$A := \begin{pmatrix} 10 & -7 & 0 \\ -3 & 2 & 6 \\ 5 & -2 & 5 \end{pmatrix} \quad B := \begin{pmatrix} 7 \\ 4 \\ 6 \end{pmatrix}$$

Отримаємо розширену матрицю D

$$D := \text{augment}(A, B) = \begin{pmatrix} 10 & -7 & 0 & 7 \\ -3 & 2 & 6 & 4 \\ 5 & -2 & 5 & 6 \end{pmatrix}$$

Перевіримо СЛАР за правилом Крамера та Кронекера-Капеллі

$$\Delta_A := |A| = -95 \quad \text{rank}(A) = 3 \quad \text{rank}(D) = 3$$

Як бачимо, система має єдиний розв'язок.

Приклад 3.9. Розв'язання СЛАР методом зворотної матриці у програмі MathCAD.

За базу для розрахунку візьмемо лістинг з прикладу 3.8, до якого додаються такі рядки:

Обчислимо невідомі

за методом зворотної матриці Виконаємо перевірку

$$X := A^{-1} \cdot B = \begin{pmatrix} -0.442 \\ -1.632 \\ 0.989 \end{pmatrix} \quad A \cdot X - B = \begin{pmatrix} -1.776 \times 10^{-15} \\ 0 \\ 0 \end{pmatrix}$$

Приклад 3.10. Розв'язання СЛАР у програмі MathCAD убудованою функцією.

За базу для розрахунку візьмемо лістинг з прикладу 3.8, до якого додаються такі рядки:

$$X := \text{lsolve}(A, B) = \begin{pmatrix} -0.442 \\ -1.632 \\ 0.989 \end{pmatrix}$$

Приклад 3.11. Розв'язання СЛАР у програмі MathCAD методом Гауса.

За базу для розрахунку візьмемо лістинг з прикладу 3.8, до якого додаються такі рядки:

```
Розв'язання системи методом Гауса

E_d := rref(D) =  $\begin{pmatrix} 1 & 0 & 0 & -0.442 \\ 0 & 1 & 0 & -1.632 \\ 0 & 0 & 1 & 0.989 \end{pmatrix}$ 

Виділення вектору, що містить розв'язок

X := E_d  $\langle \text{rank}(D) \rangle = \begin{pmatrix} -0.442 \\ -1.632 \\ 0.989 \end{pmatrix}$ 
```

Крім того, в пакеті MathCAD можна отримати розв'язок СЛАР ітераційними методами, реалізуючи їх як засобами пакету, так і звичайним математичним записом. Однак, слід зазначити, що реалізація останнього засобу в пакеті MathCAD не є ефективною. Для реалізації першого способу (вбудованими можливостями) розв'язання СЛАР можна використовувати так званий "обчислювальний блок" з директивою **Given** та функціями **Find**, **Minerr**¹. Обчислювальний блок починається службовим словом-директивою **Given** і має таку структуру:

Given

Рівняння

Обмеження

Вирази з функціями **Find** або **Minerr**

¹ Функції **Find** і **Minerr** призначені для розв'язання нерівностей та нелінійних рівнянь. Однак їх можна використовувати й для розв'язання СЛАР.

У блоці використовується одна з двох функцій:

Find(v_1, v_2, \dots, v_n) – повертає значення однієї або кількох змінних для точного розв’язку;

Minerr(v_1, v_2, \dots, v_n) – повертає значення однієї або кількох змінних для наближеного розв’язку;

Між цими функціями існує принципова різниця. Перша функція намагається знайти точний розв’язок. Друга функція намагається знайти максимальне наближення до точного розв’язку (навіть до неіснуючого розв’язку) шляхом мінімізації середньоквадратичної похибки. Для обчислення та перевірки збіжності ітераційного процесу (тобто обчислення норми матриці) можна застосувати одну з передбачених функцій пакету, за порядком одинична, нескінченна та евклідова норми відповідно:

norm1(A), normi(A), norme(A).

Слід зауважити, що розв’язання СЛАР з використанням обчислювального блоку залежить від значення змінної TOL (точність, з якою пакет виконує обчислення). Приклади розв’язання СЛАР ітераційними методами наведено нижче.

Приклад 3.12. Приведення СЛАР до ітераційного виду, перевірка збіжності ітераційного процесу.

Виконаємо перетворення для двох СЛАР:

$$\begin{cases} 10x_1 - 70x_2 = 7, \\ -3x_1 + 2x_2 + 6x_3 = 4, \\ 5x_1 - 2x_2 + 5x_3 = 6, \end{cases} \quad \begin{cases} 8x_1 + 4x_2 + 3x_3 = 7, \\ -3x_1 + 10x_2 + x_3 = 11, \\ 1x_1 + 2x_2 + 8x_3 = 10. \end{cases}$$

Задамо матрицю коефіцієнтів A та вільних членів B

$$A := \begin{pmatrix} 10 & -7 & 0 \\ -3 & 2 & 6 \\ 5 & -2 & 5 \end{pmatrix} \quad B := \begin{pmatrix} 7 \\ 4 \\ 6 \end{pmatrix}$$

Приведемо матрицю до ітераційного виду

$$i := 0.. \text{rows}(A) - 1 \quad j := 0.. \text{cols}(A) - 1$$

$$a_{i,j} := \text{if} \left(i \neq j, -\frac{A_{i,j}}{A_{i,i}}, 0 \right) \quad b_i := \frac{B_i}{A_{i,i}}$$

$$a = \begin{pmatrix} 0 & 0.7 & 0 \\ 1.5 & 0 & -3 \\ -1 & 0.4 & 0 \end{pmatrix} \quad b = \begin{pmatrix} 0.7 \\ 2 \\ 1.2 \end{pmatrix}$$

Перевіримо збіжність ітераційного процесу

$$\text{norm1}(a) = 3 \quad \text{normi}(a) = 4.5 \quad \text{norme}(a) = 3.592$$

Задамо матрицю коефіцієнтів A та вільних членів B

$$A := \begin{pmatrix} 8 & 4 & 3 \\ -3 & 10 & 1 \\ 1 & 2 & 8 \end{pmatrix} \quad B := \begin{pmatrix} 7 \\ 11 \\ 10 \end{pmatrix}$$

Приведемо матрицю до ітераційного виду

$$i := 0.. \text{rows}(A) - 1 \quad j := 0.. \text{cols}(A) - 1$$

$$a_{i,j} := \text{if} \left(i \neq j, -\frac{A_{i,j}}{A_{i,i}}, 0 \right) \quad b_i := \frac{B_i}{A_{i,i}}$$

$$a = \begin{pmatrix} 0 & -0.5 & -0.375 \\ 0.3 & 0 & -0.1 \\ -0.125 & -0.25 & 0 \end{pmatrix} \quad b = \begin{pmatrix} 0.875 \\ 1.1 \\ 1.25 \end{pmatrix}$$

Перевіримо збіжність ітераційного процесу

$$\text{norm1}(a) = 0.75 \quad \text{normi}(a) = 0.875 \quad \text{norme}(a) = 0.754$$

За обчисленими нормами видно, що для першої СЛАР ітераційний процес буде розбіжним, а для другої - збіжним.

Приклад 3.13. Розв'язання СЛАР методом простих ітерацій.

Візьмемо СЛАР із прикладу 3.12, для якої ітераційний процес буде збіжним. Лістинг розрахунку продовжимо наступними рядками:

Виберемо початкове наближення та кількість ітерацій

$$X^{(0)} := b \quad N := 10 \quad k := 1.. N$$

Обчислимо ітераційний процес $X^{(k)} := a \cdot X^{(k-1)} + b$

Переглянемо отримані значення невідомих та зробимо числову перевірку

$$X^{(N)} = \begin{pmatrix} 1.728 \times 10^{-5} \\ 1 \\ 1 \end{pmatrix} \quad A \cdot X^{(N)} - B = \begin{pmatrix} 1.494 \times 10^{-4} \\ -8.558 \times 10^{-5} \\ 8.457 \times 10^{-5} \end{pmatrix}$$

Приклад 3.14. Розв'язання СЛАР за допомогою обчислювального блоку.

$$A := \begin{pmatrix} 10 & -7 & 0 \\ -3 & 2 & 6 \\ 5 & -2 & 5 \end{pmatrix} \quad B := \begin{pmatrix} 7 \\ 4 \\ 6 \end{pmatrix}$$

Визначимо нумерацію елементів матриці невідомих
 $i := 0.. \text{rows}(A) - 1$

Задамо початкове наближення $x_i := 0$

Отримаємо розв'язок системи використовуючи командні слова Find і Minerr:

$$\text{Given } A \cdot x - B = 0 \quad X_f := \text{Find}(x) = \begin{pmatrix} -0.442 \\ -1.632 \\ 0.989 \end{pmatrix}$$

$$\text{Given } A \cdot x - B = 0 \quad X_m := \text{Minerr}(x) = \begin{pmatrix} -0.442 \\ -1.632 \\ 0.989 \end{pmatrix}$$

Числова перевірка показує, що отримані однакові корені різними засобами:

$$A \cdot X_f - B = \begin{pmatrix} 0 \\ 0 \\ 1.776 \times 10^{-15} \end{pmatrix} \quad A \cdot X_m - B = \begin{pmatrix} 0 \\ 0 \\ 1.776 \times 10^{-15} \end{pmatrix}$$

Перед словом `given` обов'язково повинні бути задані початкові наближення. Сам обчислювальний блок складається із командного слова `given`, рівняння СЛАР у матричній формі, записаного через тотожне дорівнює, та розв'язку зі словом `Find` (або `Minerr`).

Питання для самоконтролю

1. На які групи поділяються на практиці методи, що використовуються для розв'язання СЛАР? У чому різниця прямих та ітераційних методів розв'язання СЛАР?
2. Як перевіряється однорідність СЛАР?
3. У чому полягає суть методу зворотної матриці?
4. У чому полягає суть методу Гауса? Поясніть поняття “прямий” та “зворотний” хід методу Гауса.
5. Навіщо у методі Гауса потрібний етап “вибір головного елемента”?
6. У чому полягає суть методу Крамера?
7. В чому переваги й недоліки ітераційних методів розв'язання СЛАР?
8. Як перевірити збіжність ітераційного процесу?
9. В чому полягає метод простих ітерацій?
10. Що є умовою закінчення ітераційного процесу?
11. Як засобами пакету `MathCAD` розв'язуються СЛАР?

Варіанти контрольних завдань до лекції 3

Таблиця 3.1 – СЛАР у матричному виді

| № вар. | Матриця коефіцієнтів <i>A</i> | | | Стовпець <i>B</i> |
|--------|-------------------------------|------|------|-------------------|
| 1 | 1,84 | 2,25 | 2,53 | -6,09 |
| | 2,32 | 2,60 | 2,82 | -6,98 |
| | 1,83 | 2,06 | 2,24 | -5,52 |
| 2 | 2,58 | 2,93 | 3,13 | -6,66 |
| | 1,32 | 1,55 | 1,58 | -3,58 |
| | 2,09 | 2,25 | 2,34 | -5,01 |
| 3 | 2,18 | 2,44 | 2,49 | -4,34 |
| | 2,17 | 2,31 | 2,49 | -3,91 |
| | 3,15 | 3,22 | 3,17 | -5,27 |
| 4 | 1,54 | 1,70 | 1,62 | -1,97 |
| | 3,69 | 3,73 | 3,59 | -3,74 |
| | 2,45 | 2,43 | 2,25 | -2,26 |
| 5 | 1,53 | 1,61 | 1,43 | -5,13 |
| | 2,35 | 2,31 | 2,07 | -3,69 |
| | 3,83 | 3,73 | 3,45 | -5,98 |
| 6 | 2,36 | 2,37 | 2,13 | 1,48 |
| | 2,51 | 2,40 | 2,10 | 1,92 |
| | 2,59 | 2,41 | 2,06 | 2,16 |
| 7 | 3,43 | 3,38 | 3,09 | 5,52 |
| | 4,17 | 4,00 | 3,65 | 6,93 |
| | 4,30 | 4,10 | 3,67 | 7,29 |
| 8 | 3,88 | 3,78 | 3,45 | 10,41 |
| | 3,00 | 2,79 | 2,39 | 8,36 |
| | 2,67 | 2,37 | 1,96 | 7,62 |
| 9 | 3,40 | 3,26 | 2,90 | 13,05 |
| | 2,64 | 2,39 | 1,96 | 10,30 |
| | 4,64 | 4,32 | 3,85 | 17,89 |
| 10 | 2,53 | 2,36 | 1,93 | 12,66 |
| | 3,95 | 4,11 | 3,66 | 21,97 |
| | 2,78 | 2,43 | 1,94 | 13,93 |

| № вар. | Матриця коефіцієнтів A | | | Стовпець B |
|--------|--------------------------|------|------|--------------|
| 11 | 2,16 | 1,96 | 1,56 | 13,16 |
| | 3,55 | 3,23 | 2,78 | 21,73 |
| | 4,85 | 4,47 | 3,97 | 29,75 |
| 12 | 2,69 | 2,47 | 2,07 | 19,37 |
| | 2,73 | 2,39 | 1,92 | 19,43 |
| | 2,93 | 2,52 | 2,02 | 20,80 |
| 13 | 3,72 | 3,47 | 3,06 | 30,74 |
| | 4,47 | 4,10 | 3,63 | 36,80 |
| | 4,96 | 4,53 | 4,01 | 40,79 |
| 14 | 4,35 | 4,39 | 3,67 | 40,15 |
| | 4,04 | 3,65 | 3,17 | 36,82 |
| | 3,14 | 2,69 | 2,17 | 28,10 |
| 15 | 4,07 | 3,79 | 3,37 | 40,77 |
| | 2,84 | 2,44 | 1,95 | 27,68 |
| | 4,99 | 4,50 | 3,97 | 49,37 |
| 16 | 3,19 | 2,89 | 2,47 | 33,91 |
| | 4,43 | 4,02 | 3,53 | 47,21 |
| | 3,40 | 2,92 | 2,40 | 32,92 |
| 17 | 2,57 | 2,26 | 1,84 | 28,66 |
| | 4,47 | 4,03 | 3,57 | 50,27 |
| | 4,89 | 4,40 | 3,87 | 55,03 |
| 18 | 2,83 | 2,50 | 2,08 | 33,28 |
| | 3,00 | 2,55 | 2,07 | 33,59 |
| | 3,72 | 3,21 | 2,68 | 43,43 |
| 19 | 3,78 | 3,44 | 3,02 | 46,81 |
| | 4,33 | 3,88 | 3,39 | 53,43 |
| | 4,76 | 4,24 | 3,71 | 58,73 |
| 20 | 4,59 | 4,24 | 3,82 | 59,54 |
| | 4,83 | 4,36 | 3,88 | 62,33 |
| | 4,06 | 3,53 | 3,01 | 52,11 |
| 21 | 4,56 | 4,20 | 3,78 | 61,86 |
| | 3,21 | 2,73 | 2,25 | 42,98 |
| | 4,58 | 4,04 | 3,52 | 61,67 |

| № вар. | Матриця коефіцієнтів <i>A</i> | | | Стовпець <i>B</i> |
|--------|-------------------------------|------|------|-------------------|
| 22 | 3,75 | 3,39 | 2,97 | 53,38 |
| | 4,18 | 3,70 | 3,22 | 59,28 |
| | 4,43 | 3,88 | 3,36 | 62,62 |
| 23 | 2,95 | 2,58 | 2,16 | 44,16 |
| | 5,11 | 4,62 | 4,14 | 46,68 |
| | 4,38 | 3,82 | 3,30 | 65,34 |
| 24 | 2,93 | 2,55 | 2,14 | 46,41 |
| | 3,47 | 2,98 | 2,50 | 54,78 |
| | 4,78 | 4,22 | 3,70 | 75,81 |
| 25 | 3,74 | 3,36 | 2,94 | 63,26 |
| | 4,02 | 3,51 | 3,04 | 67,51 |
| | 4,18 | 3,61 | 3,09 | 70,03 |
| 26 | 4,07 | 4,28 | 3,87 | 84,43 |
| | 5,30 | 4,79 | 4,32 | 95,45 |
| | 5,11 | 4,54 | 4,03 | 91,69 |
| 27 | 4,90 | 4,50 | 4,09 | 94,18 |
| | 3,79 | 3,27 | 2,81 | 71,57 |
| | 4,01 | 3,43 | 2,91 | 75,45 |
| 28 | 4,25 | 3,84 | 3,43 | 86,07 |
| | 3,86 | 3,34 | 2,87 | 77,12 |
| | 5,40 | 4,82 | 4,30 | 108,97 |
| 29 | 3,35 | 2,94 | 2,53 | 70,69 |
| | 5,41 | 4,88 | 4,41 | 115,38 |
| | 3,88 | 3,30 | 2,78 | 81,07 |
| 30 | 3,05 | 2,64 | 2,23 | 67,17 |
| | 4,14 | 3,61 | 3,14 | 91,43 |
| | 5,63 | 5,03 | 4,52 | 125,40 |

ЛЕКЦІЯ 4

ЧИСЕЛЬНІ МЕТОДИ РОЗВ'ЯЗАННЯ НЕЛІНІЙНИХ РІВНЯНЬ ТА ЇХ СИСТЕМ

Получив перерасчет электроэнергии за 3 месяца профессор математики передал его в качестве задания своим аспирантам. Уже на следующий день молодые специалисты установили, что данная система нелинейных уравнений не имеет решений в области действительных чисел.

У багатьох наукових та інженерних задачах виникає необхідність розв'язання рівнянь виду:

$$f(x, p_1, p_2, \dots, p_n) = 0, \quad (4.1)$$

де f – задана функція; x – невідома величина; p_1, p_2, \dots, p_n – параметри задачі.

Будь-яке число, що обертає функцію $f(x)$ в нуль, називається *коренем рівняння*.

Лінійними називаються рівняння, які не містять ступенів та експонент, поділу на змінну й добутку змінних, кореня будь-якого ступеня зі змінної.

Лінійне алгебраїчне рівняння однієї змінної – це рівняння прямої,

$$y(x) = ax + b. \quad (4.2)$$

Відповідно всі інші рівняння являються нелінійними. Залежно від виду функції $f(x)$ нелінійні рівняння поділяються на два класи – алгебраїчні та трансцендентні.

Рівняння (4.1) називається *алгебраїчним*, якщо функція $f(x)$ є алгебраїчною функцією, тобто поліномом n -го ступеня:

$$P_n(x) = a_0x^0 + a_1x^1 + a_2x^2 + \dots + a_{n-1}x^{n-1} + a_nx^n. \quad (4.3)$$

Якщо функція $f(x)$ містить тригонометричні, показникові, логарифмічні й інші функції, які не є алгебраїчними, то рівняння (4.1) називається *трансцендентним*.

Методи розв'язання нелінійних рівнянь діляться на:

- прямі (аналітичні, точні);
- ітераційні.

Прямі методи – дозволяють записати розв'язок у вигляді деякого співвідношення (формули). При цьому значення коренів можуть бути обчислені за цією формулою за кінцеве число арифметичних операцій. Подібні методи розвинені для розв'язання деяких класів тригонометричних, логарифмічних, показових, а також найпростіших алгебраїчних рівнянь.

Однак переважна більшість нелінійних рівнянь, що зустрічаються на практиці, не вдається розв'язати прямими методами. Навіть для алгебраїчного рівняння вище четвертого ступеня не вдається отримати аналітичного розв'язку у вигляді формули з кінцевим числом арифметичних дій. У всіх таких випадках доводиться звертатися до чисельних методів, що дозволяє отримати наближені значення коренів з будь-якою заданою точністю.

При чисельному підході завдання про розв'язання нелінійних рівнянь розбивається на два етапи:

1. **Відділення коренів** (локалізація), тобто знаходження таких відрізків на осі x , в межах яких міститься єдиний корінь, і
2. **Уточнення коренів**, тобто обчислення наближених значень коренів із заданою точністю.

Методи які використовуються при виконанні цих етапів наведені на рис. 4.1.

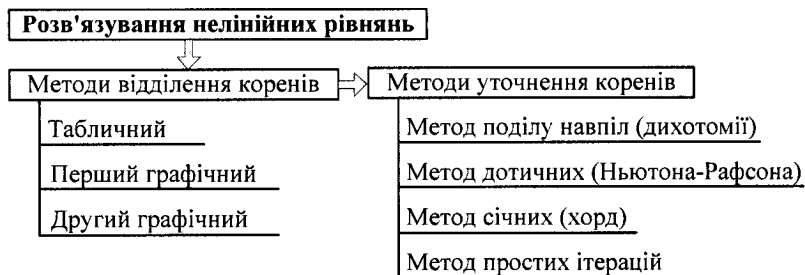


Рисунок 4.1 – Класифікація методів розв'язування нелінійних рівнянь

4.1 Відділення коренів рівняння

Процес відділення коренів рівняння $f(x) = 0$ ґрунтується на першій теоремі Больцано-Коші.

Перша теорема Больцано-Коші. Нехай функція визначена і неперервна в замкнутому проміжку $[a, b]$ і на кінцях цього проміжку приймає значення різних знаків,

$$f(a) \cdot f(b) < 0. \quad (4.4)$$

Тоді між a і b знайдеться точка c , в якій функція звертається в нуль:

$$f(c) = 0, \quad (a < c < b).$$

Зауважимо, що цей корінь буде єдиним, якщо похідна $f'(x)$ існує і зберігає постійний знак всередині інтервалу $[a, b]$.

Приклад 4.1. Перевірте наявність коренів у наведених рівнянь згідно з першою теоремою Больцано-Коші,

$$f(x) = x^3 - 6x + 2.$$

Розглянемо існування функції на інтервалі $[-4; 4]$. Знайдемо значення функції на кінцях проміжку:

$$f(-4) = (-4)^3 - 6(-4) + 2 = -38,$$

$$f(4) = (4)^3 - 6(4) + 2 = 42.$$

Функція на кінцях інтервалу міняє знак, значить вибраний інтервал містить корені рівняння $f(x)$. Для того, щоб з'ясувати скільки коренів містить вибраний інтервал, необхідно відділити корені.

Відділення коренів можна виконати *табличним способом*. Припустимо, що всі шукані корені рівняння $f(x)$ знаходяться на відрізку $[a, b]$. Вибір цього відрізка (інтервалу пошуку коренів) може бути зроблений, наприклад, на основі аналізу конкретної фізичної чи іншої задачі. Значення $f(x)$ обчислюються починаючи з точки a рухаючись вправо з деяким кроком h .

Як тільки виявляється пара сусідніх значень $f(x)$, які мають різні знаки, то відповідні значення аргументу x можна вважати межами відрізка, що містить корінь (рис. 4.2,а).

Надійність розглянутого підходу до відокремлення коренів рівняння залежить як від характеру функції $f(x)$ так і від вибраної величини кроку h .

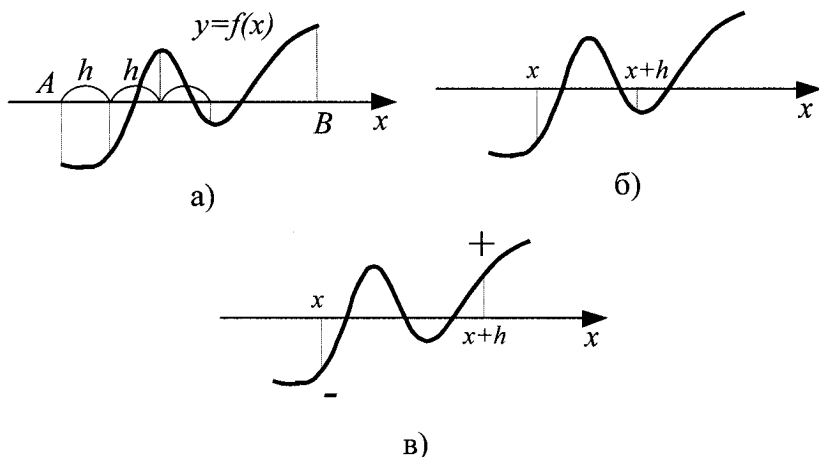


Рисунок 4.2 – Ілюстрація процесу відокремлення коренів

На рис. 4.2,б,в видно, що при неправильно вибраному кроку відокремлення, один інтервал може містити 2 і більше коренів. Передбачаючи подібні ситуації, слід вибрати досить малі значення x .

Приклад 4.2. Відокремити корені наведеної функції на інтервалі $[-4, 4]$ табличним способом,

$$f(x) = x^3 - 6x + 2.$$

Вибираємо крок $h=1$.

Складемо таблицю значень функції:

| | | | | | | | | | |
|--------|-----|----|----|----|---|----|----|----|----|
| x | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 |
| $f(x)$ | -38 | -7 | 6 | 7 | 2 | -3 | -2 | 11 | 42 |

Отже, рівняння має три дійсних кореня, що лежать в інтервалах:

$$[-3, -2], [0, 1] \text{ и } [2, 3].$$

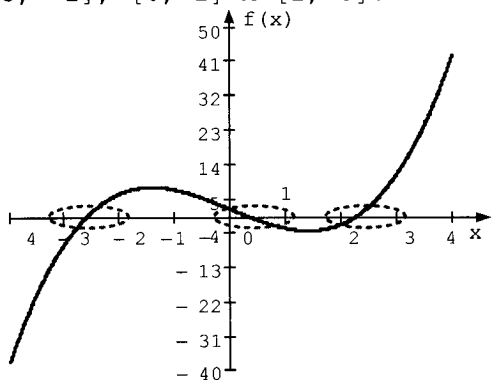
Графічні методи дозволяють візуально визначити малі інтервали, що містять корені рівняння. Тому всі графічні методи зручно застосовувати у математичних пакетах з автоматичною побудовою графіків.

Перший графічний спосіб: малі інтервали, що містять корені, визначаються перетинами функції з віссю аргументів Ox .

Другий графічний спосіб: початкова функція $f(x)$ розділяється на складові функції $f_1(x)$ та $f_2(x)$. Інтервали, на яких ці функції перетинаються одна з одною містять корені рівняння.

Приклад 4.3. Відокремити корені наведеної функції на інтервалі $[-4, 4]$ першим графічним способом, $f(x) = x^3 - 6x + 2$.

Побудуємо графік функції на заданому інтервалі, з кроком $h=1$. Візуально відділяємо три корені: $[-3, -2]$, $[0, 1]$ та $[2, 3]$.



Приклад 4.4. Відокремити корені наведеної функції на інтервалі $[-4, 4]$ другим графічним способом,

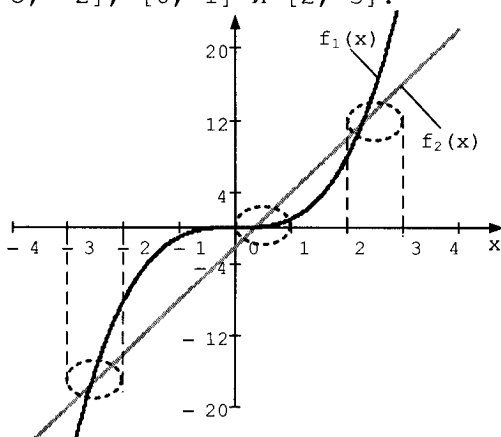
$$f(x) = x^3 - 6x + 2.$$

Прирівнюємо функцію до нуля, та частину функції переносимо вправо від знака дорівнює:

$$x^3 - 6x + 2 = 0, \quad x^3 = 6x - 2.$$

Визначаємо нові функції: $f_1(x) = x^3$, $f_2(x) = 6x - 2$.

Побудуємо графіки цих функцій в одній координатній площі. Візуально відділяємо три корені: $[-3, -2]$, $[0, 1]$ и $[2, 3]$.



4.2 Уточнення коренів

Завдання полягає в отриманні наближеного значення кореня, що належить відрізку $[a, b]$, із заданою точністю (погрішністю). Це означає, що обчислене значення кореня має відрізнятися від точної не більше ніж на величину ε :

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

Процедура чисельного визначення наближених значень коренів нелінійних рівнянь, як правило, полягає у виборі *початкового наближення* до кореня $x_0 \in [a, b]$ і обчисленні за деякою формулою наступних наближень $x_1, x_2, \dots, x_k, \dots$ (ітерацій).

Якщо ці значення з ростом k прагнуть до точного значення кореня x^* ,

$$\lim_{k \rightarrow \infty} \{x_k\} = x^*,$$

то ітераційний процес збігається.

Збіжність ітераційного процесу означає, що похибка кожного наступного наближення повинна бути менше похибки попереднього наближення, тобто похибка наближених значень з кожним кроком повинна зменшуватися:

$$|x^* - x_{k+1}| < |x^* - x_k|.$$

Постановка задачі.

При розгляді методів уточнення коренів будемо вважати, що функція $f(x)$ неперервна на відрізку $[a, b]$, на кінцях приймає значення різних знаків, $f(a) \cdot f(b) < 0$ і похідна $f'(x)$ зберігає на цьому відрізку знак. Потрібно знайти наближене значення кореня рівняння $f(x) = 0$, яке належить відрізку $[a, b]$ з точністю ε .

4.2.1 Метод поділу навпіл (бісекції, дихотомії)

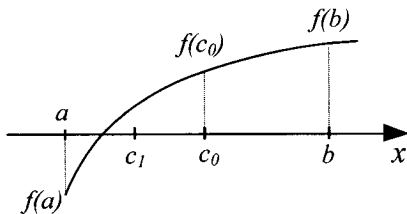


Рисунок 4.3 – Ілюстрація методу дихотомії

В якості початкового наближення кореня приймаємо середину відрізка $[a, b]$, [11]:

$$c_0 = \frac{a+b}{2}.$$

Потім досліджуємо значення функції $f(x)$ на кінцях відрізка $[a, c_0]$ і $[c_0, b]$. Той з відрізків, на кінцях якого $f(x)$ приймає значення різних знаків, містить шуканий корінь; тому його приймаємо в якості нового відрізка $[a_1, b_1]$ (на рис. 4.3 це відрізок $[a, c_0]$). Другу половину відрізка $[a, b]$, на якій $f(x)$ не змінює знак,

відкидаємо. В якості наступного наближення кореня приймаємо середину нового відрізка

$$c_1 = \frac{a_1 + b_1}{2}.$$

Таким чином, k -е наближення обчислюється як:

$$c_k = \frac{a_k + b_k}{2}. \quad (4.5)$$

Припинити ітераційний процес слід, коли буде досягнута задана точність, тобто при виконанні умови

$$|x^* - c_k| < \varepsilon, \quad |b_k - a_k| < 2\varepsilon.$$

Приклад 4.5. Методом поділу навпіл уточнити корінь наведеного рівняння, яке належить малому інтервалу $[0, 1]$, з точністю $\varepsilon=0,1$,

$$f(x) = x^4 + 2x^3 - x - 1.$$

Визначаємо значення функції на кінцях інтервалу:

$$f(0) = 0^4 + 2 \cdot 0^3 - 0 - 1 = -1, \quad f(1) = 1^4 + 2 \cdot 1^3 - 1 - 1 = 1.$$

Знаходимо значення середини інтервалу та значення функції у цій точці:

$$c_0 = \frac{a+b}{2} = \frac{0+1}{2} = 0,5;$$

$$f(0,5) = 0,5^4 + 2 \cdot 0,5^3 - 0,5 - 1 = -1,19.$$

На інтервалі $[a, c_0]$ функція знак не міняє, цей інтервал відкидаємо. А на інтервалі $[c_0, b]$ функція міняє знак з «-» на «+», тому приймаємо $a_1 = c_0 = 0,5$ і $b_1 = b = 1$.

Обчислимо середину нового інтервалу $[0,5; 1]$ та значення функцій у точці c_1 :

$$c_1 = \frac{a_1 + b_1}{2} = \frac{0,5 + 1}{2} = 0,75;$$

$$f(0,75) = 0,75^4 + 2 \cdot 0,75^3 - 0,75 - 1 = -0,59.$$

Відповідно вибираємо новий інтервал:

$$a_2 = c_1 = 0,75; \quad b_2 = b_1 = 1.$$

Обчислимо середину нового інтервалу $[0,75; 1]$ та значення функцій у точці c_2 :

$$c_2 = \frac{a_2 + b_2}{2} = \frac{0,75 + 1}{2} = 0,875;$$

$$f(0,75) = 0,875^4 + 2 \cdot 0,875^3 - 0,875 - 1 = 0,051.$$

Відповідно вибираємо новий інтервал:

$$a_3 = c_1 = 0,75;$$

$$b_3 = c_2 = 0,875.$$

Перевіримо умову припинення ітераційного процесу, приймаємо $x^* = c_2 = 0,875$:

$$|b_3 - a_3| = |0,875 - 0,75| = 0,125 < 2\varepsilon = 0,2,$$

умова виконується, припиняємо ітераційний процес, шуканий корінь знайдено із заданою точністю.

4.2.2 Метод хорд (метод січних)

В якості початкового наближення кореня, на відміну від методу поділу навпіл, беремо не середину відрізка, а точку, де перетинає вісь абсцис січна лінія (хорда), проведена через точки A і B , [11].

Запишемо тотожність для отримання рівняння прямої, що проходить через точки A і B :

$$\frac{y - f(a)}{f(b) - f(a)} = \frac{x - a}{b - a}.$$

Для точки перетину прямої з віссю абсцис ($x = x_0, y = 0$) отримаємо рівняння,

$$x_0 = a - \frac{b - a}{f(b) - f(a)} f(a).$$

В якості нового інтервалу для продовження ітераційного процесу вибираємо той з двох $[a, x_0]$ і $[x_0, b]$

на кінцях якого функція $f(x)$ приймає значення різних знаків.

Наступна ітерація полягає у визначенні нового наближення x_1 як точки перетину хорди AB_1 з віссю абсцис (рис. 4.4) і т.д.

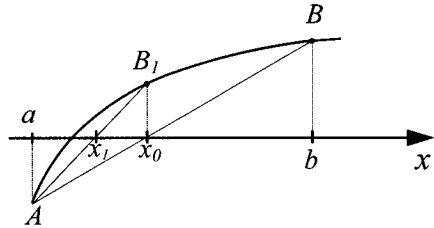


Рисунок 4.4 – Ілюстрація методу хорд

Закінчуємо процес уточнення кореня, коли відстань між черговими наближеннями стане менше заданої точності, тобто $|x_{k+1} - x_k| < \varepsilon$.

При виконанні ітераційного процесу важливим є вибір нерухомої точки послідовних наближень. Метод хорд використовується, коли на відрізку $[a, b]$ зберігає знак не тільки перша похідна $f'(x)$, а й друга – $f''(x)$, тоді хорда, яка з'єднає кінці дуги AB графіка функції, завжди розташована з боку увігнутості цієї дуги, і точка x_0 , що є нульовим наближенням кореня, завжди розташована ближче до того кінця відрізка $[a, b]$, у якому знак функції $f(x)$ протилежний $f''(x)$, тобто $f(x) \cdot f''(x) < 0$.

Можливі два випадки:

1) якщо $f(b) \cdot f''(b) < 0$, то за початкове наближення береться права межа відрізка $[a, b]$, тобто $x_0 = b$, і обчислення проводять за формулою

$$x_{i+1} = x_i - \frac{f(x_i)}{f(x_i) - f(a)}(x_i - a); \quad (4.6)$$

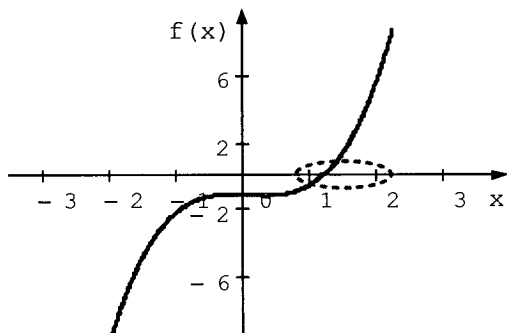
2) якщо $f(a) \cdot f''(a) < 0$, то за початкове наближення береться ліва межа відрізка $[a, b]$, тобто $x_0 = a$, і обчислення проводять за формулою

$$x_{i+1} = x_i - \frac{f(x_i)}{f(b) - f(x_i)}(b - x_i). \quad (4.7)$$

Приклад 4.6. Знайти корінь наведеного рівняння з точністю $\varepsilon = 0,01$ методом хорд,

$$f(x) = x^3 - 0,2x^2 - 0,2x - 1,2.$$

Виділимо корені даного рівняння першим графічним способом.



Корінь лежить в інтервалі $[1, 2]$.

Зменшуємо інтервал методом хорд.

Знайдемо другу похідну функції:

$$f'(x) = 3x^2 - 0,4x - 0,2; \quad f''(x) = 6x - 0,4.$$

Обчислимо значення функції та її другої похідної на кінцях інтервалу:

$$f(1) = 1^3 - 0,2 \cdot 1^2 - 0,2 \cdot 1 - 1,2 = -0,6;$$

$$f(2) = 2^3 - 0,2 \cdot 2^2 - 0,2 \cdot 2 - 1,2 = 5,6$$

$$f''(1) = 6 \cdot 1 - 0,4 = 5,6;$$

$$f''(2) = 6 \cdot 2 - 0,4 = 11,6.$$

Перевіримо умови

$$f(b) \cdot f''(b) < 0, \quad f(a) \cdot f''(a) < 0;$$

$$5,6 \cdot 11,6 > 0, \quad -0,6 \cdot 5,6 < 0.$$

Виконується умова для границі a , відповідно за початкове наближення приймаємо $x_0 = a$.

Для обчислення послідовного наближення до кореня використовуємо формулу (4.7):

$$x_1 = x_0 - \frac{f(x_0)}{f(b) - f(x_0)}(b - x_0) = 1 - \frac{-0,6}{5,6 + 0,6}(2 - 1) = 1,097;$$

$$f(x_1) = 1,097^3 - 0,2 \cdot 1,097^2 - 0,2 \cdot 1,097 - 1,2 = -0,341;$$

$$x_2 = x_1 - \frac{f(x_1)}{f(b) - f(x_1)}(b - x_1) = 1,097 - \frac{-0,341}{5,6 + 0,341}(2 - 1,097) = 1,149;$$

$$f(x_2) = 1,149^3 - 0,2 \cdot 1,149^2 - 0,2 \cdot 1,149 - 1,2 = -0,178;$$

$$x_3 = x_2 - \frac{f(x_2)}{f(b) - f(x_2)}(b - x_2) = 1,149 - \frac{-0,178}{5,6 + 0,178}(2 - 1,149) = 1,175;$$

$$f(x_3) = 1,175^3 - 0,2 \cdot 1,175^2 - 0,2 \cdot 1,175 - 1,2 = -0,089;$$

$$x_4 = x_3 - \frac{f(x_3)}{f(b) - f(x_3)}(b - x_3) = 1,175 - \frac{-0,089}{5,6 + 0,089}(2 - 1,175) = 1,188;$$

$$f(x_4) = 1,188^3 - 0,2 \cdot 1,188^2 - 0,2 \cdot 1,188 - 1,2 = -0,044;$$

$$x_5 = x_4 - \frac{f(x_4)}{f(b) - f(x_4)}(b - x_4) = 1,188 - \frac{-0,044}{5,6 + 0,044}(2 - 1,188) = 1,194;$$

$$f(x_5) = 1,194^3 - 0,2 \cdot 1,194^2 - 0,2 \cdot 1,194 - 1,2 = -0,021.$$

Перевіримо умову припинення ітераційного процесу $|x_{k+1} - x_k| < \varepsilon$:

$$|x_5 - x_4| = |1,194 - 1,188| = 0,006 < \varepsilon.$$

Знайдений корінь рівняння $x = 1,194$ (для перевірки точний розв'язок складає 1,2).

4.2.3 Метод Ньютона-Рафсона (метод дотичних)

Ідея *методу Ньютона-Рафсона* зводиться до заміни функції $f(x)$ на кожній ітерації дотичною до неї в точці x_k [15]. Нехай нам відоме початкове наближення до кореня x_0 . Проведемо в цій точці дотичну пряму до кривої $f(x)$. Ця дотична перетне вісь абсцис в точці x_1 , яку будемо

розглядати в якості наступного наближення (рис. 4.5). Значення x_1 легко знайти як точку перетину дотичної з віссю Ox :

$$y = kx + b,$$

де k – коефіцієнт дотичної, який визначається похідною функції у вибраній точці, $k = \operatorname{tg}(\alpha) = f'(x_0)$.

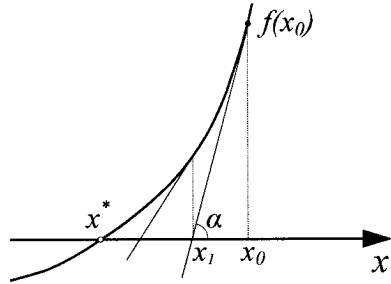


Рисунок 4.5 – Ілюстрація методу дотичних

Якщо записати рівняння дотичної через похідну функції та її значення у точці початкового наближення отримаємо:

$$y = f(x_0) + f'(x_0)(x - x_0). \quad (4.8)$$

Тоді перше наближення отримаємо як:

$$f(x_0) + f'(x_0)(x - x_0) = 0, \quad x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Аналогічно можуть бути знайдені і наступні наближення:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad (4.9)$$

де k – номер ітерації, $k = 1, 2, 3 \dots$

Закінчуємо процес уточнення кореня, коли відстань між черговими наближеннями стане менше заданої точності, тобто, $|x_{k+1} - x_k| < \varepsilon$.

У цьому методі за початкове наближення x_0 вибирається той з кінців відрізка $[a, b]$, у якому знаки $f(x_0)$ і $f'(x_0)$ збігаються, тобто $f(x) \cdot f'(x) > 0$.

Приклад 4.7. Знайти корінь наведеного рівняння з точністю $\varepsilon=0,01$ методом дотичних, на малому інтервалі $[1; 2]$: $f(x)=x^2-2$.

Знайдемо похідну функції:

$$f'(x) = \frac{d}{dx}(x^2-2) = 2x.$$

Перевіримо виконання умови $f(x) \cdot f'(x) > 0$, на обох кінцях інтервалу:

$$f(1) = 1^2 - 2 = -1, \quad f'(1) = 2 \cdot 1 = 2, \quad f(1) \cdot f'(1) = -1 \cdot 2 = -2 < 0;$$

$$f(2) = 2^2 - 2 = 2, \quad f'(2) = 2 \cdot 2 = 4, \quad f(2) \cdot f'(2) = 2 \cdot 4 = 8 > 0.$$

Тож за початкове наближення приймаємо праву межу малого інтервалу $x_0 = b = 2$.

Згідно формули методу дотичних визначимо перше наближення

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 2 - \frac{2}{4} = 1,5;$$

$$f(x_1) = f(1,5) = 1,5^2 - 2 = 0,25;$$

$$f'(x_1) = f'(1,5) = 2 \cdot 1,5 = 3.$$

Відповідно друге наближення:

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 1,5 - \frac{0,25}{3} = 1,417;$$

$$f(x_2) = f(1,417) = 1,417^2 - 2 = 0,008;$$

$$f'(x_2) = f'(1,417) = 2 \cdot 1,417 = 2,834.$$

Третє наближення:

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = 1,417 - \frac{0,008}{2,834} = 1,414;$$

$$f(x_3) = f(1,414) = 1,414^2 - 2 = -0,001;$$

$$f'(x_3) = f'(1,414) = 2 \cdot 1,414 = 2,828.$$

Перевіримо умову припинення ітераційного процесу:

$$|x_3 - x_2| < \varepsilon, \quad |1,414 - 1,417| < 0,01, \quad 0,003 < 0,01.$$

Шуканий корінь рівняння із заданою точністю $x = 1,414$.

Розглянемо застосування методу Ньютона на прикладі обчислення квадратного кореня, [20]. Якщо ми шукаємо розв'язок рівняння $x^2 = A$, необхідно до формули (4.9) підставити значення $f(x) = \sqrt{x}$. Після спрощення виразу отримаємо:

$$x_{k+1} = \frac{x_k + \frac{A}{x_k}}{2}. \quad (4.10)$$

Приклад 4.8. Обчислити квадратний корінь числа 23 за допомогою пакету MS Excel.

В комірці B5 задаємо формулу (4.10) і протягуємо її нижче, отримаємо послідовні ітерації до значення квадратного кореня з числа 23.

| B5 | | f' = (B4+SBS2/B4)/2 | |
|----|------------|---------------------|---|
| | A | B | C |
| 1 | A | корінь A | |
| 2 | 23 | 4,79583152 | |
| 3 | № Ітерація | | |
| 4 | 0 | 3 | |
| 5 | 1 | 5,33333333 | |
| 6 | 2 | 4,82291667 | |
| 7 | 3 | 4,79590758 | |
| 8 | 4 | 4,79583152 | |
| 9 | 5 | 4,79583152 | |

4.3 Метод простих ітерацій

Використовується для функцій для яких можна отримати перетворену функцію:

$$x = \phi(x). \quad (4.11)$$

Обчислення проводяться за формулою:

$$x_{k+1} = \phi(x_k). \quad (4.12)$$

Приклад 4.9. Знайти корінь наведеного рівняння з точністю $\varepsilon = 0,01$ методом простих ітерацій при початковому наближенні $x_0 = 0$: $f(x) = x^3 - 12x - 4$.

Отримаємо перетворену функцію:

$$x^3 - 12x - 4 = 0, \quad 12x = x^3 - 4, \\ x = \frac{x^3 - 4}{12}, \quad \text{відповідно } x_{k+1} = \frac{x_k^3 - 4}{12}.$$

Обчислимо послідовні наближення:

$$x_1 = \frac{0^3 - 4}{12} = -0,333, \\ x_2 = \frac{(-0,333)^3 - 4}{12} = -0,336, \\ x_3 = \frac{(-0,336)^3 - 4}{12} = -0,336.$$

Відповідно x_2 є коренем функції і для його обчислення достатньо 2-х ітерацій.

Розглянемо можливі варіації отримання $x = \phi(x)$ для наведеного рівняння:

$$x^3 - 12x - 4 = 0, \quad x^3 = 12x + 4, \\ \frac{x^3}{x^2} = \frac{12x + 4}{x^2}, \quad x = \frac{12x + 4}{x^2}.$$

В цьому випадку ітераційна формула має вигляд:

$$x_{k+1} = \frac{12x_k + 4}{x_k^2}.$$

При спробі обчислення значень з заданим початковим наближенням отримаємо операцію поділу навіл, що неприпустимо.

Ще одна варіація отримання $x = \phi(x)$:

$$x^3 - 12x - 4 = 0, \quad x^3 = 12x + 4, \quad x = \sqrt[3]{12x + 4}.$$

В цьому випадку ітераційна формула має вигляд:

$$x_{k+1} = \sqrt[3]{12x_k + 4}.$$

Обчислимо послідовні наближення:

$$x_1 = \sqrt[3]{12x_0 + 4} = \sqrt[3]{12 \cdot 0 + 4} = 1,587; \\ x_2 = \sqrt[3]{12x_1 + 4} = \sqrt[3]{12 \cdot 1,587 + 4} = 2,846;$$

Як бачимо в цьому випадку ітераційний процес має розбіжний характер. Тому в методі простих ітерацій вибір функції $x = \phi(x)$ має велике значення.

Метод простих ітерацій має суттєвий недолік: сфера його застосування обмежена. Якщо в межах кореня похідна $|f'(x)| < 1$ і початкове наближення достатньо близьке до кореня, то процедура ітерації збігається, в іншому випадку ітераційний процес є розбіжним. Збіжність методу суттєва якщо $|f'(x)| \ll 1$.

4.4 Розв'язання нелінійних рівнянь та їх систем у пакеті MathCAD

Етапи відділення та уточнення коренів нелінійного рівняння достатньо легко реалізувати як засобами пакету MathCAD² так і методами, описаними вище. Побудова графіків в пакеті автоматична, тому легко користуватись графічними способами при відділенні коренів.

Приклад 4.10. Відділити корені наведеного рівняння за допомогою пакету MATHCAD,

$$f(x) = \cos(x) - 0,2x.$$

Відділимо корені численно з представленням результату у таблиці:

```

< завдання функції
f(x) := cos(x) - 0.2·x

< завдання інтервалу існування функції
a := -10    b := 10

< завдання кроку та обчислення кількості
малих інтервалів
h := 1      n :=  $\frac{b-a}{h}$       n = 20

```

² Необхідно зазначити, що до розв'язків, які повертають вбудовані функції пакету при розв'язанні нелінійних рівнянь та їх систем, треба ставитися обережно, тобто перевіряти їх. Власне кажучи, це зауваження стосується й інших функцій пакету.


```
< завдання нумерації та обчислення  
значень аргументу
```

```
i := 0, 1.. n      z0 := a      zi+1 := zi + h
```

```
< завдання матриці для табличного  
перегляду значень аргументу та функції
```

```
ab := augment (z, f(z))T
```

| | | | | | | | |
|------|---|-------|-------|-------|-------|------|-----|
| | 0 | 1 | 2 | 3 | 4 | 5 | |
| ab = | 0 | -10 | -9 | -8 | -7 | -6 | -5 |
| | 1 | 1.161 | 0.889 | 1.454 | 2.154 | 2.16 | ... |

Продивляючись результати обчислення аргументу та функції у таблиці, отримаємо наступні малі інтервали, що містять корені даного рівняння:

$[-4; -3]$, $[-2; -1]$, $[1; 2]$.

Відділимо корені першим графічним способом.

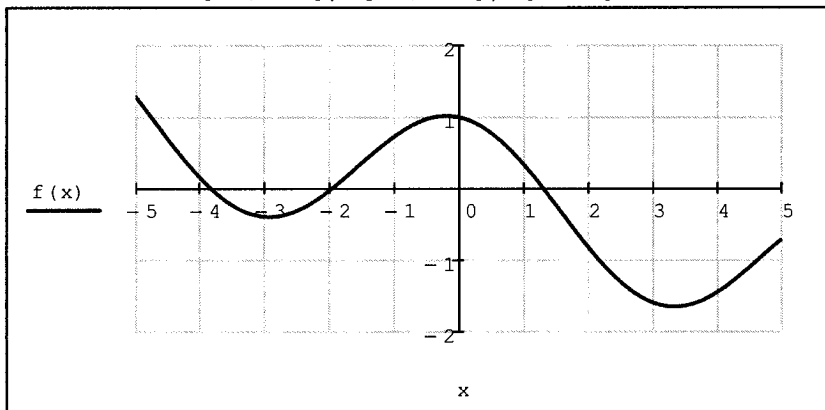
Для цього побудуємо графік у системі координат Декарта (X-Y Plot).

Візуально можна зменшити діапазон побудови до $[-5; 5]$.

При побудові графіку слід задати розбиття по осі $0x$ 10, щоб кожна клітинка сітки складала 1, та вказати тип графіку з перетином осей (crossed).

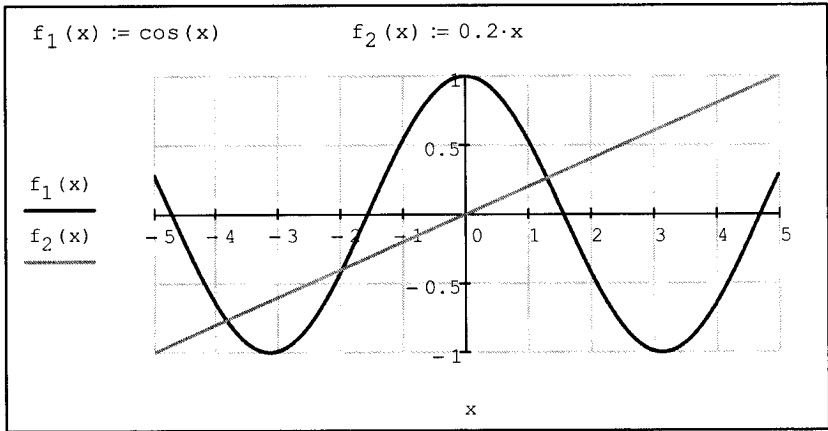
З наведеного нижче лістингу, видно, що відділяються такі інтервали з коренем:

$[-4; -3]$, $[-2; -1]$, $[1; 2]$.



Для другого графічного способу перед побудовою графіку визначимо функції $f_1(x)$, $f_2(x)$:

$$\cos(x) - 0,2x = 0, \cos(x) = 0,2x, f_1(x) = \cos(x), f_2(x) = 0,2x.$$



Слід зазначити, що відділення коренів у пакеті MathCAD необхідно проводити і численно і графічно, для запобігання помилок у визначенні малих інтервалів.

Для чисельного обчислення коренів рівнянь з однією змінною в пакеті призначені функції **polyroots** і **root**. Перша функція має наступний формат:

$$\mathbf{polyroots}(\text{“Вектор”}),$$

і дозволяє обчислити корені (як дійсні так і комплексні) звичайного полінома ступеня n . Аргумент “Вектор” - матриця-стовбець розміру $n+1$, елементами якого є коефіцієнти полінома. Функція повертає розв’язок у вигляді матриці-стовпця.

Приклад 4.11. Обчислити корені наведеного поліному за допомогою убудованої функції **polyroots**,

$$f(x) = x^3 + 3x^2 - 24x + 1.$$

Лістинг у пакеті MathCAD:

```

f(x) := x3 + 3·x2 - 24·x + 1 < завдання поліному
< завдання елементів матриці коефіцієнтів поліному
a3:= 1 a2:= 3 a1:= -24 a0:= 1
< обчислення коренів функцією "polyroots", перевірка
x:= polyroots(a) =  $\begin{pmatrix} -6.638 \\ 0.042 \\ 3.596 \end{pmatrix}$  f(x) =  $\begin{pmatrix} -8.527 \times 10^{-14} \\ 0 \\ 0 \end{pmatrix}$ 

```

Функція **root** обчислює один корінь (дійсний або комплексний) і має слідуєчий формат:

root("Вираз", "Змінна").

Аргумент "Вираз" визначає позначення функції, аргумент "Змінна" вказує ім'я ведучої змінної (початкове наближення), відносно, якої розв'язується рівняння. Перед використанням функції **root** значення ведучої змінної повинно бути визначено (наприклад з графіку функції).

Пакет використовує це значення (x_0) як початкову точку для організації послідовних наближень до шуканого кореня x^* , припиняючи обчислення при виконанні вимоги $|f(x^*)| < TOL$, тобто, наближене обчислення кореня проводиться з точністю, що визначається значенням зумовленої змінної пакета TOL (за замовчанням $TOL = 0,001$). При обчисленні дійсного кореня змінній x_0 надається дійсне значення, а при обчисленні комплексного – комплексне. При перевищенні наперед заданої кількості ітерацій без виконання даної вимоги пакет припиняє обчислення й ідентифікує функцію **root** діагностичним

повідомленням “Not converning” – немає збіжності.

Основними причинами появи даного повідомлення є:

- відсутність у функції $f(x)$ кореня даного типу;
- некоректне визначення першого наближення x_0 .

Приклад 4.12. Уточнити корені рівняння $f(x)=\cos(x)-0,2x$ для всіх малих інтервалів визначених у прикладі 4.10 за допомогою функції `root`.

На етапі відділені коренів були визначені такі інтервали: $[-4; -3]$, $[-2; -1]$, $[1; 2]$. Продовжуємо лістинг з прикладу 4.10. Задаючи у якості початкового наближення будь-яку з меж даних інтервалів отримає точне значення коренів:

```
z := -4    root ( f ( z ) , z ) = -3.837
z := -3    root ( f ( z ) , z ) = -3.837

z := -2    root ( f ( z ) , z ) = -1.977
z := -1    root ( f ( z ) , z ) = -1.977

z := 1     root ( f ( z ) , z ) = 1.306
z := 2     root ( f ( z ) , z ) = 1.306
```

Приклад 4.13. Уточнити корені рівняння $f(x)=\cos(x)-0,2x$ методами поділу навпіл, дотичних та січних з точністю $\epsilon=0,001$.

Для визначених на етапі відділення коренів інтервалів (приклад 4.10) використаємо такі методи уточнення коренів: для $[-4; -3]$ метод хорд; для $[-2; -1]$ метод дотичних; для $[1; 2]$ метод поділу навпіл.

Для виконання умови припинення ітераційного процесу скористаємось функцією «until», як показано у [52]. Реалізуємо у пакеті метод хорд:

```

< завдання функції та точності розв'язання
f(x) := cos(x) - 0.2·x   ε := 0.001
< завдання меж малого інтервалу
a := -4   b := -3
< визначення другої похідної функції
df(x) :=  $\frac{d^2}{dx^2}f(x) \rightarrow -\cos(x)$ 
< визначення початкового наближення
x0 := if (f(a) · df(a) > 0, a, b) = -4
< визначення нерухомої межі інтервалу
a := if (f(a) · df(a) < 0, a, b) = -3
< завдання рівняння методу хорд
mh(c) :=  $c - \frac{(a - c) \cdot f(c)}{f(a) - f(c)}$ 
< визначення першого наближення
x1 := mh(x0) = -3.727
< завдання надлишкової кількості ітерацій
k := 1.. 100
< обчислення кореня із заданою точністю
xk+1 := until (|xk - xk-1| - ε, mh(xk))
n := last(x) = 27   xn = -3.837

```

Отримали корінь $x = -3,837$ з точністю $\epsilon = 0,001$ протягом 27 ітерацій (точний розв'язок $x = -3,837$).

Реалізуємо у пакеті метод дотичних:

```

< завдання функції та точності розв'язання
f(x) := cos(x) - 0.2·x   ε := 0.001
< завдання меж малого інтервалу
a := -2   b := -1
< визначення похідної функції
df(x) :=  $\frac{d}{dx}f(x) \rightarrow -\sin(x) - 0.2$ 

```

```

< завдання функції та точності розв'язання
f(x) := cos(x) - 0.2·x    ε := 0.001
< завдання меж малого інтервалу
a := -2    b := -1
< визначення похідної функції
df(x) :=  $\frac{d}{dx} f(x) \rightarrow -\sin(x) - 0.2$ 
< визначення початкового наближення
x0 := if(f(a)·df(a) < 0, a, b) = -2
< завдання рівняння методу дотичних
md(c) :=  $c - \frac{f(c)}{df(c)}$ 
< визначення першого наближення
x1 := md(x0) = -1.977
< завдання надлишкової кількості ітерацій
k := 1.. 100
< обчислення кореня із заданою точністю
xk+1 := until(|xk - xk-1| - ε, md(xk))
n := last(x) = 2    xn = -1.977

```

Отримали корінь $x = -1,977$ з точністю $\varepsilon = 0,001$ протягом 2 ітерацій (точний розв'язок $x = -1,977$).

Реалізуємо у пакеті метод поділу навпіл:

```

< завдання функції та точності розв'язання
f(x) := cos(x) - 0.2·x    ε := 0.001
< завдання меж малого інтервалу,
як нульових елементів матриць
a0 := 1    b0 := 2
< завдання умови вибору інтервалу, на якому
функція міняє знак
R(d, c) := if(f(d)·f(c) > 0, c, d)
< завдання надлишкової кількості ітерацій
k := 0.. 100

```

```

< обчислення кореня із заданою точністю

$$\begin{pmatrix} a_{k+1} \\ b_{k+1} \end{pmatrix} := \text{until} \left[ \left| a_k - b_k \right| - \varepsilon, \begin{pmatrix} R \left( a_k, \frac{a_k + b_k}{2} \right) \\ R \left( b_k, \frac{a_k + b_k}{2} \right) \end{pmatrix} \right]$$

n := last (a) = 10      x :=  $\frac{a_n + b_n}{2} = 1.306$ 

```

Отримали корінь $x=1,306$ з точністю $\varepsilon=0,001$ протягом 10 ітерацій (точний розв'язок $x=1,306$).

За допомогою математичних пакетів задача, розв'язок якої в аналітичному вигляді практично не можливо отримати, стає доступною. Ми говоримо про розв'язання систем нелінійних рівнянь. У пакеті MathCAD ця задача може бути вирішена двома засобами:

- ❖ за допомогою обчислювального блоку (**Given**, **Find** та **Minerr**), застосування якого описане в пункті 3.3;
- ❖ методом Ньютона.

Приклад 4.14. Розв'язати систему нелінійних рівнянь за допомогою обчислювального блоку,

$$\begin{cases} x_1^2 + x_2^2 + \sin(x_3) = 41,5, \\ x_1 x_2 - x_2 \operatorname{tg}(x_3) = 20, \\ x_1 x_2 - \sin(x_3) = 10. \end{cases}$$

Лістинг у пакеті MathCAD:

```

< завдання системи нелінійних рівнянь
f1 (x1, x2, x3) := x12 + x22 + sin (x3) - 41.5
f2 (x1, x2, x3) := x1·x2 - x2·tan (x3) - 20
f3 (x1, x2, x3) := x1·x2 - sin (x3) - 10
< завдання початкових наближень
x1 := 4      x2 := 2      x3 := 0

```

```

< реалізація обчислювального блоку
Given
    f1 (x1, x2, x3) = 0
    f2 (x1, x2, x3) = 0
    f3 (x1, x2, x3) = 0
vx := Find (x1, x2, x3)    vx =  $\begin{pmatrix} 6.363 \\ 1.416 \\ -1.443 \end{pmatrix}$ 
< перевірка
x1 := vx0    x2 := vx1    x3 := vx2
    f1 (x1, x2, x3) = 0
    f2 (x1, x2, x3) = -7.105 x 10-15
    f3 (x1, x2, x3) = 0

```

Приклад 4.15. Розв'язати систему нелінійних рівнянь методом Ньютона¹ з точністю 10⁻³,

$$\begin{cases} x^2 - y^2 = 0, \\ xy - 4 = 0. \end{cases}$$

Лістинг у пакеті MathCAD:

```

< запис матриці Якобі
J(x, y) :=  $\begin{bmatrix} \frac{d}{dx} \left[ \text{tr} \left[ \left( f(x, y)^T \right)^{\langle 0 \rangle} \right] \right] & \frac{d}{dy} \left[ \text{tr} \left[ \left( f(x, y)^T \right)^{\langle 0 \rangle} \right] \right] \\ \frac{d}{dx} \left[ \text{tr} \left[ \left( f(x, y)^T \right)^{\langle 1 \rangle} \right] \right] & \frac{d}{dy} \left[ \text{tr} \left[ \left( f(x, y)^T \right)^{\langle 1 \rangle} \right] \right] \end{bmatrix}$ 
< завдання початкових наближень
x0 := 1.5    y0 := 2.5

```

¹ Приклади розв'язання систем нелінійних рівнянь методом Зейделя засобами пакету не наводяться.

< обчислення визначника матриці Якобі

$$|J(x_0, y_0)| = 17$$

< завдання умовної функції обчислення наступного наближення

$$\Delta a(c, d) := -J(c, d)^{-1} \cdot f(c, d)$$

< обчислення наближень, перевірка

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} := \begin{pmatrix} x_k + \Delta a(x_k, y_k)_0 \\ x_k + \Delta a(x_k, y_k)_0 \end{pmatrix} \quad \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

$$n := x_{\text{last}(x)} = 2 \quad f(x_n, y_n) = \begin{pmatrix} 0 \\ 5.615 \times 10^{-3} \end{pmatrix}$$

Пояснення щодо матриці Якобі $J(x, y)$: Задається вона функцією користувача, кожний елемент якої обчислюється так: з транспонованого (позначка "T") вектора $f(x, y)$ виділяється (позначка "<n>") стовбець з номером n залежно від номера стовбця матриці Якобі, за допомогою функції `tr` вибирається перший елемент n-го стовпця транспонованого вектора $f(x, y)$ для якого й обчислюється частинна похідна за вказаною змінною.

Питання для самоконтролю

1. Наведіть класифікацію нелінійних рівнянь. Що є розв'язком нелінійних рівнянь та їх систем?
2. Чим викликана необхідність застосування чисельних методів розв'язання нелінійних рівнянь та їх систем?
3. З яких етапів складається знаходження коренів нелінійних рівнянь?
4. В чому суть методів відділення коренів?

5. Як графічне відділення коренів уточнюється з допомогою обчислень? Які властивості функції однієї змінної при цьому використовуються?

6. Перелічить методи уточнення коренів нелінійного рівняння.

7. В чому суть методу поділу навпіл? Як оцінюється похибка кореня, знайденого методом поділу навпіл?

8. В чому суть методу хорд? Як вибирається початкове наближення?

9. У чому полягає відмінність методу хорд від методу поділу навпіл і що в них спільного?

10. В чому суть методу Ньютона для розв'язку нелінійних рівнянь та їх систем? Як вибирається початкове наближення? Які умови застосовності методу Ньютона?

11. В чому суть методу простої ітерації для розв'язку нелінійних рівнянь? Які умови збіжності ітераційної послідовності в методі простих ітерацій?

12. Як впливає значення змінної TOL на обчислення коренів рівнянь?

13. Як функціями пакету MathCAD обчислити корені алгебраїчних, трансцендентних рівнянь та їх систем.

Завдання до лекції 4

1. Відокремити корені рівняння $f(x) = \cos x - \sin x = 0$ на інтервалі $[0,5]$.

2. Знайти корінь рівняння $f(x) = x^3 - 2x^2 - 5 = 0$ методом поділу навпіл з точністю $\varepsilon = 0,001$ на інтервалі $[2, 3]$.

3. Знайти корінь рівняння $f(x) = x^3 + x + 3 = 0$ методом хорд з точністю $\varepsilon = 0,01$ на інтервалі $[-2, -1]$.

ЛЕКЦІЯ 5

НАБЛИЖЕННЯ ФУНКЦІЙ

Метод решения хорош, если с самого начала мы можем предвидеть – и далее подтвердить это, – что, следуя этому методу, мы достигнем цели.

Г. Лейбниц

Завдання наближення (апроксимації) полягає в заміні функціональної залежності, заданої на множині $X \subseteq R$ у вигляді таблиці, графіка, формули або не в явному вигляді, більш простою наближеною функцією.

Тобто якась функція $f(x)$ апроксимується на множині аргументів X іншою функцією $\varphi(x)$, яка зручніша для обчислень і близька в деякому сенсі до $f(x)$.

Якщо X складається з дискретної множини точок x_0, x_1, \dots, x_n то наближення називається *точковим*, а якщо це відрізок (a, b) , то наближення називають *безперервним* або *інтегральним*. При вирішенні задач обробки результатів експериментів, діагностики параметрів або ідентифікації стану об'єктів по виміряним параметрам стану його координат, в основному використовують точкові наближення.

Одним з видів точкової апроксимації є інтерполявання, при якому апроксимуюча функція приймає в заданих точка x_i , ті ж значення y_i що й функція $f(x)$, тобто

$$\phi(x_i) = y_i \quad i = 0, 1, \dots, n$$

У цьому випадку, близькість функції інтерполяції до заданої функції полягає в тому, що їх значення співпадають на заданій системі точок.

Для отримання точкового середньоквадратичного наближення таблично заданої функції $y=f(x)$, апроксимуючу функцію $\phi(x)$ будують з умови мінімуму величини,

$$S = \sum_{i=0}^n [y_i - \phi(x_i)]^2,$$

де y_i – значення функції $f(x)$ в точках x_i .

На рисунку 5.1 показані якісні графіки:

- інтерполяційної функції (суцільна лінія)
- результати середньоквадратичного наближення (пунктирна лінія).

Точками відзначені табличні значення функції $f(x)$.



Рисунок 5.1 – Ілюстрація суті інтерполяції

5.1 Постановка задачі інтерполяції.

Якщо значення функції $f(x)$ задані в $n + 1$ в точках x_0, x_1, \dots, x_n відрізка (a, b) і досить проста для обчислень наближуюча її функція $\phi(x)$ збігається з $f(x)$ в цих точках, а в інших точках (a, b) наближено представляє $f(x)$, то таке наближення називає *інтерполяцією*, [51].

Нехай на відрізку $[x_0, x_n]$ задані таблично точки x_0, x_1, \dots, x_n (вузли інтерполяції) і відомі значення функції $y = f(x)$ в цих точках (табл. 5.1): $y_0 = f(x_0), y_1 = f(x_1), \dots, y_n = f(x_n)$.

Таблиця 5.1 – Таблично задані значення аргументу та функції

| | | | | | |
|-----|-------|-------|-------|---------|-------|
| x | x_0 | x_1 | x_2 | \dots | x_n |
| y | y_0 | y_1 | y_2 | \dots | y_n |

Потрібно побудувати функцію $\varphi(x)$, яка збігається у вузлах інтерполяції зі значеннями функції $f(x)$.

Таблиця значень функції $f(x)$ може бути інтерпольовано нескінченною безліччю різних кривих, кожна з яких буде графіком функції, для якої виконані всі умови інтерполяції.

Для практики важливий випадок апроксимації функції алгебраїчними багаточленами

$$\varphi(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots + a_mx^m.$$

Всі методи інтерполяції можна розділити на локальні і глобальні.

Для випадку *локальної інтерполяції* на кожному інтервалі $[x_{i-1}, x_i]$ будується окремий поліном. Для випадку *глобальної інтерполяції* відшукується єдиний поліном на всьому інтервалі (a, b) , [13]. При цьому шуканий поліном називається **інтерполяційним поліномом**.

5.2 Методи локальної інтерполяції

Найвідоміші методи локальної інтерполяції:

- кусково-постійна інтерполяція;
- кусково-лінійна інтерполяція;
- сплайн-інтерполяція (квадратичні, кубічні сплайни

та ін.)

5.2.1 Локальна кусково-постійна інтерполяція (метод найближчого сусіда)

На кожному інтервалі $[x_{i-1}, x_i]$ інтерполяційний багаточлен дорівнює константі, а саме лівому або правому значенню функції, [14].

Для лівої кусково-постійної інтерполяції:

$$\varphi(x) = f(x_{i-1}), \quad (5.1)$$

якщо $x_{i-1} \leq x \leq x_i$.

Для правої кусково-постійної інтерполяції:

$$\varphi(x) = f(x_i), \quad \text{якщо } x_{i-1} \leq x \leq x_i. \quad (5.2)$$

Легко зрозуміти, що умови інтерполяція виконуються. Побудована функція є розривною, що обмежує її застосування.

5.2.2 Локальна кусково-лінійна інтерполяція

На кожному інтервалі $[x_{i-1}, x_i]$ інтерполяційний багаточлен є лінійною функцією $\phi_i(x) = k_i x + l_i$, [14].

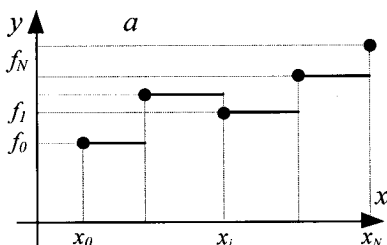


Рисунок 5.2 – Локальна кусково-постійна інтерполяція

Значення коефіцієнтів знаходяться від виконання умов інтерполяції в кінцях відрізка,

$$\varphi(x_{i-1}) = f(x_{i-1}), \varphi(x_i) = f(x_i).$$

Отримуємо систему рівнянь:

$$k_i x_{i-1} + l_i = f(x_{i-1})$$

$$k_i x_i + l_i = f(x_i)$$

Звідки знаходимо:

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

Отже, функцію $\varphi(x)$ можна записати у вигляді:

$$\varphi(x) = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} (x - x_{i-1}) + f(x_{i-1}), \quad (5.3)$$

$$x_{i-1} \leq x \leq x_i \quad i = 1, 2, \dots, n-1.$$

При лінійній інтерполяції спочатку потрібно визначити інтервал, в який потрапляє значення x , а потім підставити його в формулу.

Підсумкова функція буде безперервною, але похідна буде розривною в кожному вузлі інтерполяції. Похибка такої інтерполяції менше, ніж в разі кусково-постійної інтерполяції.

Приклад 5.1. Для заданої таблиці значень функції потрібно знайти значення функції при $x = 1$, $x = 3,2$ наступними методами: 1. Кусково-постійної інтерполяції. 2. Кусково-лінійної інтерполяції.

| | | | | |
|------|----|-----|-----|-----|
| x | 0 | 2 | 3 | 3,5 |
| f(x) | -1 | 0,2 | 0,5 | 0,8 |

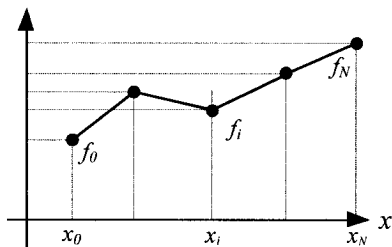


Рисунок 5.3 – Локальна кусково-лінійна інтерполяція

Розв'язок 1.

Точка $x = 1$ належить першому локальному відрізку $[0; 2]$, тобто $i = 1$ і отже:

- за формулами лівої кусково-постійної інтерполяції $\varphi(1) = f(0) = -1$ (рис. 5.4, а);
- за формулами правої кусково-постійної інтерполяції $\varphi(1) = f(2) = 0,2$ (рис. 5.4, б).

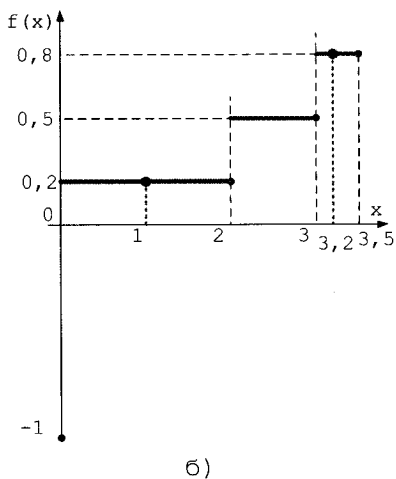
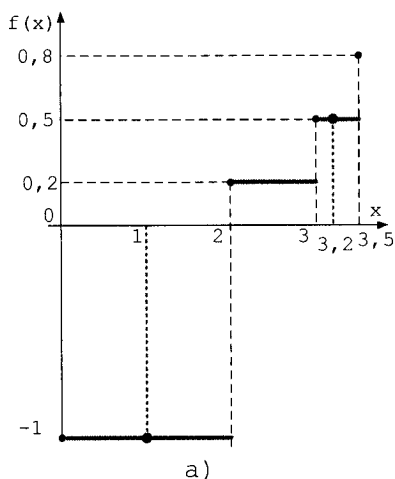


Рисунок 5.4 - Ліва та права кусково-постійні інтерполяції функції

Точка $x = 3,2$ належить третьому локальному відрізку $[3; 3,5]$, тобто $i = 3$ і отже:

- при лівій кусково-постійній інтерполяції $\varphi(3,2) = f(3) = 0,5$;
- при правій - $\varphi(3,2) = f(3,5) = 0,8$.

Розв'язок 2.

Скористаємося формулами кусково-лінійної інтерполяції (5.3) для відрізку $[0; 2]$:

$$\begin{aligned}\varphi(1) &= \frac{f(x_1) - f(x_0)}{x_1 - x_0} (x - x_0) + f(x_0) = \\ &= \frac{0,2 - (-1)}{2 - 0} (1 - 0) + (-1) = -0,4\end{aligned}$$

Скористаємося формулами кусково-лінійної інтерполяції для відрізку $[3; 3,5]$:

$$\begin{aligned}\phi(3,2) &= \frac{f(x_3) - f(x_2)}{x_3 - x_2} (x - x_3) + f(x_3) = \\ &= \frac{0,8 - 0,5}{3,5 - 3} (3,2 - 3,5) + 0,8 = 0,62.\end{aligned}$$

5.2.3 Локальна кусково-квадратична інтерполяція

Якщо розглянути інтервал, що містить три вузлових точки, наприклад x_{i-1} , x_i , x_{i+1} , то аналогічно можна побудувати інтерполяційний поліном другого порядку, тобто параболу, [14]. Для випадку рівновіддалених вузлів ($x_i - x_{i-1} = h = \text{const}$):

$$\begin{aligned}\phi(x) &= f(x_{i-1}) \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} (x - x_{i-1}) + \\ &+ \frac{f(x_{i+1}) - 2f(x_i) + f(x_{i-1}))}{x_i - x_{i-1}} (x - x_{i-1})(x - x_i)\end{aligned}, \quad (5.4)$$

або

$$\begin{aligned}\phi(x) &= f(x_{i-1}) \frac{f(x_i) - f(x_{i-1}))}{h} (x - x_{i-1}) + \\ &+ \frac{f(x_{i+1}) - 2f(x_i) + f(x_{i-1}))}{h} (x - x_{i-1})(x - x_i)\end{aligned}. \quad (5.5)$$

5.3 Методи глобальної інтерполяції

Найвідоміші методи глобальної інтерполяції:

- ступеневий багаточлен (поліном);
- поліном Лагранжа;
- поліном Ньютона;
- поліноми Ерміта;
- поліноми Чебишева;
- тригонометричні поліноми і ін.

5.3.1 Інтерполяція степеневим багаточленом (поліномом) у канонічній формі

Виберемо в якості апроксимуючої функції поліном ступеня n в канонічному вигляді:

$$\varphi(x) = P_n(x) = c_0 + c_1x + c_2x^2 + \dots + c_nx^n. \quad (5.6)$$

Коефіцієнти полінома c_i визначаються з умов Лагранжа $P_n(x_i) = y_i$, $i=0,1,\dots,n$, що дає систему рівнянь з $n + 1$ невідомих:

$$c_0 + c_1x_0 + c_2x_0^2 + \dots + c_nx_0^n = y_0,$$

$$c_0 + c_1x_1 + c_2x_1^2 + \dots + c_nx_1^n = y_1,$$

...

$$c_0 + c_1x_n + c_2x_n^2 + \dots + c_nx_n^n = y_n.$$

Систему рівнянь можна коротко записати в такій формі:

$$\sum_{k=0}^n c_k x_i^k = y_i, \quad i=0,1,\dots,n; \quad (5.7)$$

або в матричній формі,

$$A \cdot C = Y, \quad (5.8)$$

де C – вектор-стовпець, що містить невідомі коефіцієнти c_i , Y – вектор-стовпець, складений з табличних значень функції y_i , а матриця A має вид,

$$A = \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{bmatrix}.$$

Система лінійних алгебраїчних рівнянь щодо невідомих c_i , матиме розв'язок, якщо визначник системи

відмінний від нуля. Визначник матриці A відомий як визначник Вандермонда, має аналітичний вираз:

$$\det A = \sum_{\substack{i,j=0 \\ (i \neq j)}}^n (x_i - x_j) \neq 0$$

З цього виразу видно, що $\det A \neq 0$, якщо серед вузлів x_i немає співпадаючих.

Формально розв'язок може бути записаний в матричному вигляді:

$$C = A^{-1} \cdot Y.$$

5.3.2 Інтерполяційний поліном Лагранжа

На відміну від інтерполяційного полінома в канонічній формі для обчислення значень полінома Лагранжа не потрібно попередньо визначати коефіцієнти полінома шляхом розв'язання СЛАР.

Лагранж запропонував будувати інтерполяційний поліном у вигляді наступного розкладання, [60]:

$$L_n(x) = \sum_{i=0}^n y_i l_i(x) = y_0 l_0(x) + y_1 l_1(x) + \dots + y_n l_n(x), \quad (5.9)$$

де $l_i(x)$ – базисні поліноми n -го степеню,

$$l_i(x) = \frac{(x - x_0)(x - x_1) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_0)(x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)}. \quad (5.10)$$

Наприклад, легко бачити, що

$$l_0(x) = \frac{(x - x_1)(x - x_2) \dots (x - x_n)}{(x_0 - x_1)(x_0 - x_2) \dots (x_0 - x_n)},$$

$$l_1(x) = \frac{(x - x_0)(x - x_2) \dots (x - x_n)}{(x_1 - x_0)(x_1 - x_2) \dots (x_1 - x_n)}.$$

Розглянемо окремі випадки.

Нехай $n = 1$, тобто задані значення функції тільки в двох точках. Тоді базові поліноми мають вигляд:

$$l_0(x) = \frac{(x - x_1)}{(x_0 - x_1)}, \quad l_1(x) = \frac{(x - x_0)}{(x_1 - x_0)}.$$

Тоді інтерполяційний поліном запишемо як:

$$L_n(x) = \sum_{i=0}^1 y_i l_i(x) = y_0 l_0(x) + y_1 l_1(x) = y_0 \frac{(x - x_1)}{(x_0 - x_1)} + y_1 \frac{(x - x_0)}{(x_1 - x_0)},$$

або

$$L_n(x) = \frac{y_1 - y_0}{x_1 - x_0} x + y_1 - \frac{y_1 - y_0}{x_1 - x_0}.$$

Отриманий вираз відповідає формулі кусково-лінійної інтерполяції.

Якщо $n = 2$ тоді:

$$l_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)}, \quad l_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)},$$

$$l_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}.$$

Тоді інтерполяційний поліном запишемо як:

$$L_n(x) = \sum_{i=0}^2 y_i l_i(x) = y_0 l_0(x) + y_1 l_1(x) + y_2 l_2(x),$$

$$L_n(x) = y_0 \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} + y_1 \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} + y_2 \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}.$$

Отриманий вираз відповідає формулі так званої квадратичної або параболічної інтерполяції.

Приклад 5.2. Для заданої таблично функції знайти її значення при $x = 1$, використовуючи інтерполяційний поліном Лагранжа.

| | | | | |
|--------|----|-----|-----|-----|
| x | 0 | 2 | 3 | 3,5 |
| $f(x)$ | -1 | 0,2 | 0,5 | 0,8 |

Для цього випадку $n = 3$, тобто поліном Лагранжа має третій порядок. Обчислимо значення базисних поліномів при $x = 1$:

$$l_0(1) = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)} = \frac{(1-2)(1-3)(1-3,5)}{(0-2)(0-3)(0-3,5)} = 0,238;$$

$$l_1(1) = \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)} = \frac{(1-0)(1-3)(1-3,5)}{(2-0)(2-3)(2-3,5)} = 1,667;$$

$$l_2(1) = \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)} = \frac{(1-0)(1-2)(1-3,5)}{(3-0)(3-2)(3-3,5)} = -1,667;$$

$$l_3(1) = \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)} = \frac{(1-0)(1-2)(1-3)}{(3,5-0)(3,5-2)(3,5-3)} = 0,762;$$

$$L_n(1) = \sum_{i=0}^3 f(x_i) l_i(x) = (-1) \cdot 0,238 + 0,2 \cdot 1,667 + 0,5 \cdot (-1,667) + 0,8 \cdot 0,762 = -0,129.$$

На відміну від інтерполяційного полінома в канонічній формі для обчислення значень полінома Лагранжа не потрібно попередньо визначати коефіцієнти полінома шляхом розв'язання системи рівнянь.

Однак для кожного значення аргументу x поліном Лагранжа доводиться перераховувати знову, коефіцієнти ж канонічного полінома обчислюються тільки один раз.

Тому практичне застосування полінома Лагранжа виправдано тільки в тому випадку, коли інтерполяційна функція обчислюється в порівняно невеликій кількості точок x .

5.3.3 Інтерполяційний поліном Ньютона

Розглянемо ще одну форму запису інтерполяційного полінома, [60]:

$$N_n(x) = A_0 + A_1(x-x_0) + A_2(x-x_0)(x-x_1) + \dots + A_n(x-x_0)(x-x_1)\dots(x-x_{n-1}). \quad (5.11)$$

Вимоги збігу значень полінома з заданими значеннями функції у вузлових точках $N_n(x_i) = y_i$, $i=0,1,\dots,n$ призводить до системи лінійних рівнянь з трикутною матрицею для невідомих коефіцієнтів $\{A, i = 0,1,\dots,n\}$:

$$\begin{cases} A_0 = y_0, \\ A_0 + A_1(x_1 - x_0) = y_1, \\ A_0 + A_1(x_2 - x_0) + A_2(x_2 - x_0)(x_2 - x_1) = y_2, \\ \dots \end{cases}$$

розв'язання якої не складає труднощів.

Інтерполяційний поліном, записаний у формі $N_n(x)$, називають *інтерполяційним поліномом Ньютона*.

Розрізняють *першу* й *другу* інтерполяційні формули Ньютона.

Перша формула Ньютона використовується для інтерполяції в точках близьких до початку таблиці, у цьому випадку формула Ньютона записується у вигляді:

$$N_n(x) = y_0 + \frac{\Delta y_0}{1!h}(x-x_0) + \frac{\Delta^2 y_0}{2!h^2}(x-x_0)(x-x_1) + \dots + \frac{\Delta^n y_0}{n!h^n} \cdot (x-x_0) \cdot (x-x_1) \cdot \dots \cdot (x-x_{n-1}), \quad (5.12)$$

де h – крок інтерполяції, $h = x_{i+1} - x_i$; Δy_i – кінцева різниця 1-го порядку, $\Delta y_i = y_{i+1} - y_i$; $\Delta^2 y_i$ – кінцева різниця 2-го порядку, $\Delta^2 y_i = \Delta y_{i+1} - \Delta y_i$; $\Delta^n y_i$ – кінцева різниця n -го порядку, $\Delta^n y_i = \Delta^{n-1} y_{i+1} - \Delta^{n-1} y_i$.

Друга інтерполяційна формула Ньютона використовується для інтерполяції в точках, близьких до кінця таблиці для функцій із значеннями аргументу, які знаходяться на однаковій відстані. Вона відрізняється від першої лише порядком перебору вузлів:

$$N_n(x) = y_n + \frac{\Delta y_{n-1}}{1!h}(x-x_n) + \frac{\Delta^2 y_{n-2}}{2!h^2}(x-x_n)(x-x_{n-1}) + \dots$$

$$\dots + \frac{\Delta^n y_0}{n!h^n}(x-x_n)(x-x_{n-1})\dots(x-x_0).$$
(5.13)

При обчисленні кінцевих різниць зручно користуватися таблицею.

Приклад 5.3. Для заданої таблично функції побудувати поліном Ньютона.

| | | | | | |
|-------|----|----|---|---|---|
| x_i | 0 | 1 | 2 | 3 | 4 |
| y_i | 22 | 12 | 6 | 4 | 6 |

Оскільки кількість точок у наведеній таблиці мала, для інтерполяції будемо застосовувати першу формулу Ньютона. Обчислимо всі кінцеві різниці:

| | | | | | |
|----------------|-----|----|----|---|---|
| i | 0 | 1 | 2 | 3 | 4 |
| x_i | 0 | 1 | 2 | 3 | 4 |
| y_i | 22 | 12 | 6 | 4 | 6 |
| Δy_i | -10 | -6 | -2 | 2 | |
| $\Delta^2 y_i$ | 4 | 4 | 4 | | |

$$P_2(x) = 22 + \frac{-10}{1! \cdot 1}(x-0) + \frac{4}{2! \cdot 1^2}(x-0)(x-1) = 2x^2 - 12x + 22.$$

5.3.4 Інтерполяція функцій у пакеті MathCAD

Пакет містить групу вбудованих функцій інтерполяції:

interp(vx,vy,x) – повертає значення функції в точці x , що обчислюється методом лінійної інтерполяції на основі значень з векторів vx и vy .

lspline(vx,vy) – повертає вектор коефіцієнтів, який використовується функцією **interp** для побудови лінійного сплайна. При цьому побудований сплайн на межах інтервалу має рівні нулю другу та третю похідні.

pspline(vx,vy) – повертає вектор коефіцієнтів, який використовується функцією **interp** для побудови параболічного сплайна. При цьому побудований сплайн на межах інтервалу має рівну нулю третю похідну.

cspline(vx,vy) – повертає вектор коефіцієнтів, який використовується функцією **interp** для побудови кубічного сплайна. При цьому на поведінку сплайна на межах інтервалу ніяких обмежень не накладається.

interp(vs,vx,vy,x) – повертає інтерпольоване значення в точці x . Вектор vs є результатом, який повертає одна з функцій **cspline**, **pspline** або **lspline**.

У даних функціях використовуються наступні аргументи: vx , vy – вектори одного розміру, які містять значення аргументу та відповідні значення табличної функції. Елементи векторів повинні бути відсортовані в порядку збільшення значень аргументу; vs – вектор коефіцієнтів, який повертає одна з функцій **cspline**, **pspline** або **lspline**; x – значення змінної, в якій необхідно обчислити інтерпольоване значення y . Передбачається, що значення x лежить у межах зміни елементів vx .

Приклад 5.4. Обчислити значення функції, що задана таблицею, в точках $x_1^* = 1,5$, $x_2^* = 2,3$ такими методами:

1. Лінійна інтерполяція.
2. Квадратична сплайн-інтерполяція.
3. Кубічна сплайн-інтерполяція.
4. Поліном Лагранжа.
5. поліном Ньютона.

| | | | | | |
|---|-----|---|-----|---|------|
| x | 0 | 1 | 2 | 3 | 4 |
| y | 1,5 | 1 | 2,5 | 3 | -0,5 |

Для початку задамо табличну функцію у вигляді матриць-стовпців:

```

< завдання матриць-стовпців x і y,
  обчислення довжини заданих матриць

x :=  $\begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$       y :=  $\begin{pmatrix} 1.5 \\ 1 \\ 2.5 \\ 3 \\ -0.5 \end{pmatrix}$       d := length(x) = 5
                                     n := d - 1 = 4

< завдання аргументів шуканих точок
xc1 := 1.5      xc2 := 2.3
  
```

Для отримання лінійної інтерполяції скористаємось убудованою функцією **linterp**:

```

< завдання функції лінійної інтерполяції
yL(z) := linterp(x, y, z)

< перевірка на відомих значеннях
yL(0) = 1.5      yL(1) = 1      yL(2) = 2.5

< обчислення шуканих точок
yL(xc1) = 1.75      yL(xc2) = 2.65
  
```

Лістинг у пакеті MathCAD для отримання квадратичної сплайн-інтерполяції:

```

< завдання функції квадратичної інтерполяції
y2(z) := interp(cspline(x, y), x, y, z)

< перевірка на відомих значеннях
y2(0) = 1.5    y2(1) = 1    y2(2) = 2.5

< обчислення шуканих точок
y2(xc1) = 1.688    y2(xc2) = 2.891

```

Лістинг у пакеті MathCAD для отримання кубічної сплайн-інтерполяції:

```

< завдання функції кубічної інтерполяції
y3(z) := interp(pspline(x, y), x, y, z)

< перевірка на відомих значеннях
y3(0) = 1.5    y3(1) = 1    y3(2) = 2.5

< обчислення шуканих точок
y3(xc1) = 1.65    y3(xc2) = 2.919

```

Лістинг у пакеті MathCAD для отримання інтерполяції поліномом Лагранжа:

```

< завдання формули поліному Ларанжа

$$L(z) := \sum_{i=0}^n \left( y_i \cdot \prod_{j=0}^n \text{if} \left( i = j, 1, \frac{z - x_j}{x_i - x_j} \right) \right)$$


< перевірка поліному на відомих значеннях
L(0) = 1.5    L(1) = 1

< обчислення шуканих точок
L(xc1) = 1.688    L(xc2) = 2.891

```

Лістинг у пакеті MathCAD для отримання інтерполяції поліномом Ньютона:

```

< завдання матриці a, яка містить обчислення всіх
кінцевих різниць для заданої функції
a := y^T    i:=1.. n    j:=0.. n-1
ai,j := if[j ≥ (d - i), 0, ai-1,j+1 - ai-1,j]

```

$$a = \begin{pmatrix} 1.5 & 1 & 2.5 & 3 & -0.5 \\ -0.5 & 1.5 & 0.5 & -3.5 & 0 \\ 2 & -1 & -4 & 0 & 0 \\ -3 & -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

< виділення кінцевих різниць для y_0

$$a := a^{(0)} \quad a^T = (1.5 \quad -0.5 \quad 2 \quad -3 \quad 0)$$

< завдання нумерації елементів i
обчислення кроку інтерполяції

$$i := 0..n \quad h := x_2 - x_1 \quad h = 1$$

< завдання формули поліному Ньютона

$$N(z) := a_0 + \sum_{i=1}^n \left[\frac{a_i}{i! \cdot h^i} \cdot \prod_{k=0}^{i-1} (z - x_k) \right]$$

< перевірка на відомих значеннях

$$N(0) = 1.5 \quad N(1) = 1 \quad N(2) = 2.5$$

< обчислення шуканих точок

$$N(xc1) = 1.688 \quad N(xc2) = 2.891$$

5.4 Постановка задачі апроксимації

У деяких випадках побудова апроксимуючої функції методами інтерполяції виявляється абсолютно недоцільною. Наприклад, якщо мова йде про обробку експериментальних даних, отриманих в результаті спостережень або вимірювань. Експериментальні дані завжди містять в собі помилки різного роду:

- *Систематичні помилки*, як правило, дають відхилення в одну сторону від істинного значення вимірюваної величини. Вони можуть бути

постійними або закономірно змінюватися при повторенні досліду, і їх причини і характер відомі. Систематичні помилки можуть бути викликані умовами експерименту, дефектом вимірювального приладу, його незадовільним градуванням.

- *Випадкові помилки* визначаються великим числом факторів, які не можуть бути усунені або в повній мірі враховані при вимірюванні або при обробці результатів. Вони мають випадкових, несистематичний характер, дають відхилення від істинного значення в ту чи іншу сторони при повторенні вимірювань.
- *Грубі помилки* сильно спотворюю результат вимірювань і за величиною можуть значно перевищувати систематичні і випадкові помилки.

Експериментальні дані неминуче містять випадкові помилки. Теоретично, випадкова помилка може бути зменшена до як завгодно малої величини шляхом провидіння багаторазових вимірювань. Однак більш ефективним способом позбавлення від випадкових помилок є підхід, що дозволяє отримати витончені дані належної математичною обробкою наявних результатів вимірювань.

Один з поширених способів математичної обробки експериментальних даних складається з визначення виду і параметрів функціонального зв'язку між досліджуваними величинами на підставі результатів вимірювань.

Нехай, в ході експерименту з вивчення залежності між величинами y та x шляхом вимірювань була отримана таблиця значень (табл. 5.2).

Таблиця 5.2 – Експериментальні дані

| | | | | | |
|-----|-------|-------|-------|-----|-------|
| x | x_0 | x_1 | x_2 | ... | x_n |
| y | y_0 | y_1 | y_2 | ... | y_n |

Завдання полягає в тому, щоб знайти формулу

$$y = f(x),$$

що наближено виражає цю залежність. Наближена залежність отримана на підставі експериментальних даних, називається *емпіричною формулою*.

Взагалі для апроксимації використовуються поліноми, сплайни та відрізки ряду Фур'є [19].

Якщо при інтерполяції функцій ми використовували умову рівності значень інтерполяційного полінома і даної функції в вузлах інтерполяції то в даному випадку вимога точного збігу не потрібна, так як дані не отримані точно.

У цих випадках можна вимагати лише наближеного виконання умов інтерполяції

$$|\phi(x_i) - f(x_i)| < \varepsilon.$$

Ця умова означає, що функція апроксимації $\phi = f(x)$ проходить не точно через задані точки, а в деякій їх околиці (рис. 5.5).

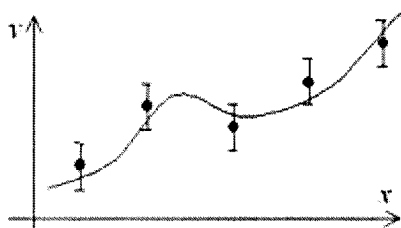


Рисунок 5.5 – Функція апроксимації та входні точки

Побудова емпіричної формули складається з двох етапів:

1. Вибір виду залежності $\phi = f(x, a_0, a_0, a_1, \dots, a_m)$, яка

містить невідомі параметри (a_0, a_1, \dots, a_m) .

2. Визначення найкращих, в деякому сенсі, параметрів вибраної залежності.

Вид формули апроксимації вибирається з фізичних або геометричних міркувань: експериментальні точки наносяться на графік і приблизно вибирається загальний вигляд залежності шляхом порівняння отриманої кривої з графіками відомих функцій.

При розв'язанні електротехнічних задач в основному використовуються наступні типи функцій:

- ступеневий поліном;
- експоненціальний поліном;
- кусково-лінійна функція;
- кусково-нелінійна функція;
- трансцендентні функції.

5.4.1 Визначення параметрів емпіричної формули.

Метод найменших квадратів (МНК)

Нехай для вихідних даних $x_i, f_i, i=1, 2, \dots, N$ вибраний вигляд емпіричної залежності:

$$y = \varphi(x, a_0, a_1, \dots, a_m),$$

з невідомими параметрами (a_0, a_1, \dots, a_m) .

Запишемо суму квадратів відхилень між обчисленими за емпіричною формулою і заданими дослідними даними:

$$S(a_0, a_1, \dots, a_m) = \sum_{i=1}^N (\varphi(x, a_0, a_1, \dots, a_m) - f_i)^2.$$

Параметри (a_0, a_1, \dots, a_m) будемо знаходити з умов мінімуму функції $S(a_0, a_1, \dots, a_m)$. У цьому полягає **метод найменших квадратів (МНК)**, [35].

Таким чином, невідомі параметри (a_0, a_1, \dots, a_m) визначають знаючи, що в точці мінімуму всі часткові похідні від S по (a_0, a_1, \dots, a_m) дорівнюють нулю:

$$\frac{dS}{da_0} = 0, \quad \frac{dS}{da_1} = 0, \dots, \quad \frac{dS}{da_m} = 0$$

Метод найменших квадратів не є єдиним способом мінімізації відхилень. Існують інші способи вирішення цього завдання, зокрема, метод вибраних точок і метод середніх, [8].

Розглянемо застосування МНК для окремого випадку, широко використовуваного на практиці. В якості емпіричної функції розглянемо поліном:

$$\phi(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots + a_mx^m. \quad (5.14)$$

Формула для визначення суми квадратів відхилень набуде вигляду:

$$S(a_0, a_1, \dots, a_m) = \sum_{i=1}^N (a_0 + a_1x_i + a_2x_i^2 + \dots + a_mx_i^m - f_i)^2. \quad (5.15)$$

Обчислимо похідні:

$$\frac{dS}{da_0} = 2 \sum_{i=1}^N (a_0 + a_1x_i + a_2x_i^2 + \dots + a_mx_i^m - f_i)^2,$$

$$\frac{dS}{da_1} = 2 \sum_{i=1}^N (a_0 + a_1x_i + a_2x_i^2 + \dots + a_mx_i^m - f_i)^2 x_i,$$

$$\frac{dS}{da_m} = 2 \sum_{i=1}^N (a_0 + a_1x_i + a_2x_i^2 + \dots + a_mx_i^m - f_i)^2 x_i^m.$$

Прирівнюючи коефіцієнти при невідомих (a_0, a_1, \dots, a_m) отримаємо наступну систему лінійних рівнянь:

$$\begin{aligned}
 Na_0 + a_1 \sum_{i=1}^N x_i + a_2 \sum_{i=1}^N x_i^2 + \dots + a_m \sum_{i=1}^N x_i^m &= \sum_{i=1}^N f_i, \\
 a_0 \sum_{i=1}^N x_i + a_1 \sum_{i=1}^N x_i^2 + a_2 \sum_{i=1}^N x_i^3 \dots + a_m \sum_{i=1}^N x_i^m &= \sum_{i=1}^N x_i f_i, \quad (5.16) \\
 a_0 \sum_{i=1}^N x_i^m + a_1 \sum_{i=1}^N x_i^{m+1} + a_2 \sum_{i=1}^N x_i^{m+2} \dots + a_m \sum_{i=1}^N x_i^{2m} &= \sum_{i=1}^N x_i^m f_i.
 \end{aligned}$$

Ця система рівнянь носить назву *нормальної*.

Вирішуючи цю систему лінійних рівнянь, отримуємо значення коефіцієнтів (a_0, a_1, \dots, a_m) .

Розглянемо приклади.

У випадку полінома першого порядку $m = 1$, апроксимуюча емпірична залежність має вигляд $\varphi(x) = a_0 + a_1 x$. Система нормальних рівнянь буде мати простий вигляд:

$$\begin{aligned}
 \sum_{i=1}^N (a_0 + a_1 x_i - y_i) &= 0, \quad Na_0 + a_1 \sum_{i=1}^N x_i = \sum_{i=1}^N f_i, \\
 \text{або } \sum_{i=1}^N (a_0 + a_1 x_i - y_i) &= 0, \quad a_0 \sum_{i=1}^N x_i + a_1 \sum_{i=1}^N x_i^2 = \sum_{i=1}^N x_i f_i.
 \end{aligned}$$

У випадку полінома другого порядку $m = 2$, $\varphi(x) = a_0 + a_1 x + a_2 x^2$, тоді:

$$\begin{aligned}
 Na_0 + a_1 \sum_{i=1}^N x_i + a_2 \sum_{i=1}^N x_i^2 &= \sum_{i=1}^N f_i, \\
 a_0 \sum_{i=1}^N x_i + a_1 \sum_{i=1}^N x_i^2 + a_2 \sum_{i=1}^N x_i^3 &= \sum_{i=1}^N x_i f_i,
 \end{aligned}$$

$$a_0 \sum_{i=1}^N x_i^2 + a_1 \sum_{i=1}^N x_i^3 + a_2 \sum_{i=1}^N x_i^4 = \sum_{i=1}^N x_i^2 f_i .$$

Приклад 5.5. Для заданих експериментальних даних знайти лінійну емпіричну залежність за методом МНК:

| | | | | | | |
|------|-----|------|-----|-----|------|-----|
| x | -5 | -3,5 | -2 | 1,5 | 3,25 | 5 |
| f(x) | 0,5 | 1,2 | 1,4 | 1,6 | 1,7 | 1,5 |

Система нормальних рівнянь для лінійної залежності:

$$Na_0 + a_1 \sum_{i=1}^N x_i = \sum_{i=1}^N f_i ,$$

$$a_0 \sum_{i=1}^N x_i + a_1 \sum_{i=1}^N x_i^2 = \sum_{i=1}^N x_i f_i .$$

Враховуючи, що $N=6$:

$$\sum_{i=1}^6 x_i = -0,75; \quad \sum_{i=1}^6 x_i^2 = 79,0625; \quad \sum_{i=1}^6 f_i = 7,9; \quad \sum_{i=1}^6 x_i f_i = 5,925;$$

отримаємо

$$6a_0 + a_1(-0,75) = 7,9;$$

$$a_0(-0,75) + a_1(79,0625) = 5,925 .$$

Розв'язуючи систему лінійних рівнянь, отримаємо:

$$a_0 = 1,328; \quad a_1 = 0,0875 .$$

Отже, лінійна залежність має вигляд:

$$\phi(x) = 1,328 + 0,0875x .$$

Обчислимо суму квадратів відхилень за формулою:

$$S = \sum_{i=1}^6 (\phi(x_i) - f_i), \quad S = 0,343 .$$

5.4.2 Апроксимація тригонометричними поліномами

У багатьох завданнях необхідно апроксимувати первинну функцію $f(x)$ на інтервалі $T = 2l$. Розв'язок зручно знаходити у вигляді тригонометричного полінома

n -го порядку в системі ортогональних функцій на відрізку $[0; 2l]$. При деяких обмеженнях $f(x)$ можна представити у вигляді суми тригонометричного ряду Фур'є:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos\left(\frac{k\pi x}{l}\right) + b_k \sin\left(\frac{k\pi x}{l}\right) \right). \quad (5.17)$$

Коефіцієнти ряду Фур'є знаходять по формулах Ейлера-Фур'є:

$$a_0 = \frac{1}{l} \int_0^{2l} f(x) dx, \quad a_k = \frac{1}{l} \int_0^{2l} f(x) \cos\left(\frac{k\pi x}{l}\right) dx,$$

$$b_k = \frac{1}{l} \int_0^{2l} f(x) \sin\left(\frac{k\pi x}{l}\right) dx,$$

де $k = 1, 2, 3, \dots$

Подання функції у вигляді ряду Фур'є називають гармонійним аналізом.

Нехай $f(x)$ задана таблично. Застосовуючи формулу трапецій для чисельного інтегрування одержимо формули Бесея для знаходження коефіцієнтів ряду Фур'є:

$$\begin{cases} a_0 = \frac{2}{m+1} \sum_{i=0}^m f(x_i) \\ a_k = \frac{2}{m+1} \sum_{i=0}^m f(x_i) \cos\left(\frac{2k\pi i}{m+1}\right) \\ b_k = \frac{2}{m+1} \sum_{i=0}^m f(x_i) \sin\left(\frac{2k\pi i}{m+1}\right) \end{cases}$$

Тригонометричний поліном з коефіцієнтами, обчисленими по цих формулах, буде найкращим середньо квадратичним наближенням для функції $f(x)$ серед інших, тому що забезпечує мінімальну суму квадратів

відхилень:

$$\delta_n^2 = \sum_{i=0}^m (\phi(x_i) - f(x_i))^2.$$

5.4.3 Апроксимація функцій у пакеті MathCAD

Для виконання лінійної апроксимації в пакеті передбачені дві функції: **intercept(vx,vy)** та **slope(vx,vy)**. Дані функції дозволяють обчислити коефіцієнти a_0 та a_1 апроксимуючого рівняння. Функції **intercept(vx,vy)** і **slope(vx,vy)** повертають значення відповідно a_0 – зміщення прямої по вісі ординат та a_1 – тангенс кута нахилу прямої, які найкращим чином наближають набір даних, що представлені в векторах vx та vy, в сенсі найменших квадратів. У пакеті реалізована можливість виконання лінійної апроксимації загального вигляду:

$$F(x, k_1, k_2, \dots)$$

що забезпечує найкращу апроксимацію лінійних комбінацій заданих функцій в сенсі найменших квадратів.

Функція **linfit(vx,vy,f)** повертає вектор коефіцієнтів лінійної апроксимації загального вигляду для лінійної комбінації функції F , яка дає найкращу апроксимацію даних з векторів vx и vy. Вектор f повинен містити функції $f_i(x)$ в символьному вигляді. Для виконання нелінійної апроксимації загального вигляду в пакеті передбачена функція **genfit(vx,vy,vz,f)**, яка повертає вектор коефіцієнтів функції F , що дають мінімальну середньоквадратичну похибку наближення функцією F за даними з векторів vx и vy. Параметр vz – вектор початкових наближень (оскільки

нелінійна задача може мати більше одного розв'язку), поблизу якого шукається розв'язок. Вектор f повинен містити функцію та її частинні похідні за всіма параметрами у символічному вигляді.

Функція **regress(vx,vy,k)** – повертає вектор, який потребує функція **interp**, щоб знайти поліном k -го порядку, коефіцієнти якого обчислюються за всією сукупністю заданих точок у векторах vx и vy .

Функція **loess(vx,vy,span)** – повертає вектор, який потребує функція **interp**, щоб знайти сукупність поліномів другого порядку, який найкращим чином наближає дані з векторів vx и vy . Аргумент **span**>0 вказує розмір локальної області наближених даних, тобто визначає, наскільки великі набори даних будуть наближатися окремими поліномами. Чим більше значення **span**, тим дужче згладжуються дані. При великому значенні **span** функція **loess** працює як **regress**.

Функція **interp(vs,vx,vy,x)** – повертає інтерпольоване значення y , яке відповідає значенню x . Вектор vs обчислюється функціями **regress** або **loess**.

Можливість представлення періодичних (а за визначених умов й неперіодичних) функцій рядом Фур'є в пакеті забезпечується групою функцій, в яких реалізовано так званий, алгоритм швидкого перетворення Фур'є (перехід від функції, що задана множиною точок (x_n, y_n) , $n = 0, 1, \dots, N$ до параметрів ряду Фур'є називається прямим перетворенням Фур'є, а зворотний перехід – зворотним перетворенням Фур'є).

Функції **fft(rx)** та **cfft(cx)** виконують пряме, так зване, "швидке перетворення Фур'є" (ШПФ) для даних, що

представлені відповідно функціям дійсними значеннями в векторі **gx** та комплексними – в векторі **cx**. Причому вектор **vx** повинен містити 2^n елементів, де n – ціле число.

Функція **ifft(cx)** виконує зворотне ШПФ для вектора **cx** з комплексними елементами (вектор **vx** повинен містити 2^{m+1} елементів). Функція повертає вектор з дійсними елементами.

Функція **icfft(cx)** виконує зворотне ПФ за повним алгоритмом, при якому як початковий вектор, так й вектор результату містять елементи з комплексними значеннями.

Крім описаних функцій, пакет містить функції альтернативного перетворення Фур'є: **FFT**, **CFFT**, **IFFT**, **ICFFT**¹, в яких використовуються інші нормуючі множники.

Приклад 5.6. Для заданої табличної функції вибрати й розрахувати апроксимовану функцію. Порівняти отримані результати.

| | | | | | |
|---|-----|-----|----|----|----|
| X | 2 | 4 | 6 | 8 | 10 |
| Y | 3,5 | 6,2 | 13 | 23 | 43 |

Для зручності подальшої апроксимації функцію задамо у вигляді матриць-стовпців.

¹ В посібнику наводяться приклади використання описаних типів побудови наближених функцій не застосовуючи функції пакета (крім **intercept** та **slope**), їх опис та приклади використання можна подивитись, наприклад, у літературі [52, **Ошибка! Источник ссылки не найден.**] та в довідковій базі MathCAD.

< таблицно задані аргумент і функція,
визначення довжини векторів та нумерації елементів

$$x := \begin{pmatrix} 2 \\ 4 \\ 6 \\ 8 \\ 10 \end{pmatrix} \quad y := \begin{pmatrix} 3.5 \\ 6.2 \\ 13 \\ 23 \\ 43 \end{pmatrix} \quad d := \text{length}(x) = 5$$

$$n := d - 1 \quad i := 0..n$$

< апроксимація вихідної функції та
обчислення середньоквадратичного відхилення

< апроксимація експоненціальною функцією

$$ae := \begin{pmatrix} \text{exp}(\text{intercept}(x, \ln(y))) \\ \text{slope}(x, \ln(y)) \end{pmatrix} = \begin{pmatrix} 1.839 \\ 0.316 \end{pmatrix}$$

$$fe(z) := ae_0 \cdot e^{ae_1 \cdot z} \quad Ee := \sum_{e=0}^n fe(x_e) - y_e)^2 = 0.916$$

< апроксимація лінійною функцією

$$al := \begin{pmatrix} \text{intercept}(x, y) \\ \text{slope}(x, y) \end{pmatrix} = \begin{pmatrix} -11 \\ 4.79 \end{pmatrix} \quad fl(z) := al_0 + al_1 \cdot z$$

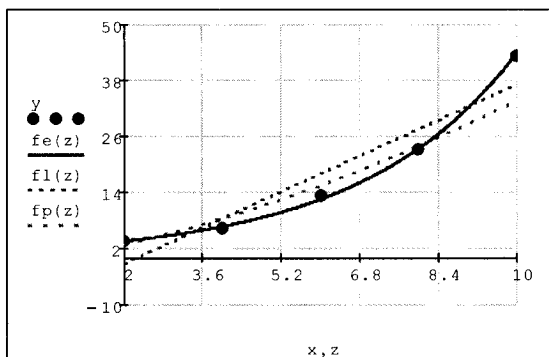
$$El := \sum_{e=0}^n fl(x_e) - y_e)^2 = 106.388$$

< апроксимація ступеневкою функцією

$$ap := \begin{pmatrix} \text{exp}(\text{intercept}(\ln(x), \ln(y))) \\ \text{slope}(\ln(x), \ln(y)) \end{pmatrix} = \begin{pmatrix} 0.977 \\ 1.533 \end{pmatrix}$$

$$fp(z) := ap_0 \cdot z^{ap_1} \quad Ep := \sum_{e=0}^n fp(x_e) - y_e)^2 = 102.898$$

Відповідно найбільш точна апроксимація при використанні експоненціальної функції. Побудуємо всі отримані функції в одній координатній площині.



У деяких випадках нелінійну апроксимацію зручно провести не через лінеаризацію, а через мінімізацію (наприклад, коли лінеаризація пов'язана з логарифмуванням негативних значень) за алгоритмом пошуку параметричних коефіцієнтів, що мінімізують цільову функцію – суму квадратів відхилень точок від кривої. Вказану можливість демонструє наступний приклад.

Приклад 5.7. Задану табличну функцію апроксимувати поліномом 2-го ступеня за допомогою функції мінімізації.

| | | | | | |
|---|-----|-----|----|----|----|
| X | 2 | 4 | 6 | 8 | 10 |
| Y | 3,5 | 6,2 | 13 | 23 | 43 |

Задаємо ступінь апроксимуючого поліному та початкові наближення його коефіцієнтів.

$$K := 2 \quad e := 0.. K \quad am_e := 0$$

Реалізуємо обчислення коефіцієнтів поліному апроксимації за допомогою обчислювального блоку і функції Minerr.

Given

$$\sum_{i=0}^n \left[y_i - \left[\sum_{e=0}^K am_e \cdot (x_i)^e \right] \right]^2 = 0$$

(7.875)

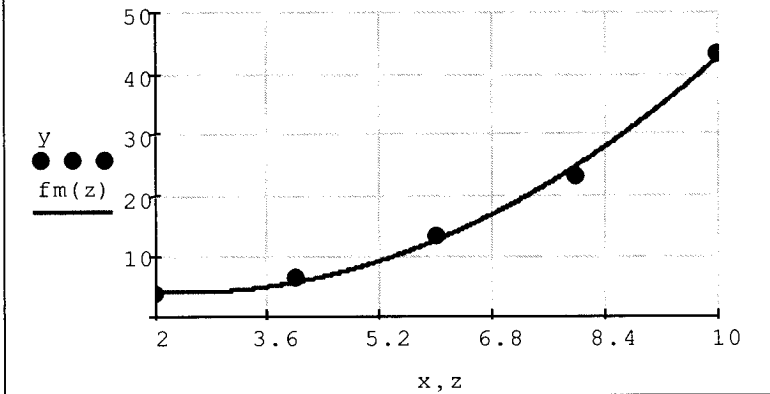
$$am := \text{Minerr}(am) = -3.3$$

(0.674)

Обчислимо середньоквадратичне відхилення та побудуємо графік.

$$fm(z) := \sum_{i=0}^K (am_i \cdot z^i)$$

$$Em := \sum_{e=0}^n (fm(x_e) - y_e)^2 = 4.328$$



5.5 Похибка інтерполяції та апроксимації

В загальному випадку похибка в кожній табличній точці буде:

$$\varepsilon_i = |P(x_i) - y_i|, \quad (5.18)$$

де $P(x_i)$ – значення апроксимуючого поліному в i -ій точці; y_i – значення табличної вихідної функції.

Тоді сума квадратів похибок визначається виразом:

$$E = \sum_{i=0}^n (P(x_i) - y_i)^2. \quad (5.19)$$

Похибка інтерполяції тим вище, чим ближче точка x лежить до кінців відрізка $[x_0, x_n]$. За межами відрізка інтерполяції (тобто при екстраполяції) похибка зросте суттєво. Оскільки у вузлах інтерполяції вихідна та інтерполююча функція повністю співпадають, то похибка інтерполяції дорівнює нулю. При значеннях $x \neq x_i$ різниця значень вихідної функції та функції інтерполяції,

$$R_n(x) = |y(x) - P_n(x)|, \quad (5.20)$$

у загальному випадку відмінна від нуля і являє собою похибку методу.

Залишкова середня квадратична похибка апроксимації оцінюється:

$$\Delta = \sqrt{\frac{E}{n-1}}. \quad (5.21)$$

Приклад 5.8. Для лінійної функції $y(x) = 3x + 2$ були виміряні з деякими помилками значення цієї функції:

| | | | | |
|---|---|---|----|----|
| x | 1 | 2 | 3 | 4 |
| y | 4 | 9 | 10 | 14 |

Для отриманої табличної функції виконати:

- апроксимацію лінійним поліномом,
- лінійну інтерполяцію між двома найближчими вузлами;
- інтерполяцію поліномом Лагранжа.

Для точки $x_0=1,5$ обчислити:

- точне значення функції за її аналітичного подання;
- наближене значення функції за рівнянням лінійної регресії;
- наближене значення функції за методом кусково-лінійної інтерполяції;
- наближене значення функції за поліномом Лагранжа.

Для отриманих значень функцій обчислити їх абсолютні та відносні похибки в порівнянні з точним значенням функції.

Апроксимація за МНК.

Функція апроксимації буде мати вигляд $y_m(x) = a_0 + a_1 x^2$. Невідомі коефіцієнти a_0, a_1 визначимо із системи рівнянь:

$$(n+1)a_0 + a_1 \sum_{i=0}^n x_i = \sum_{i=0}^n y_i$$

$$a_0 \sum_{i=0}^n x_i + a_1 \sum_{i=0}^n (x_i)^2 = \sum_{i=0}^n y_i x_i$$

Обчислимо коефіцієнти для цієї системи за допомогою наведеної нижче таблиці:

| i | 0 | 1 | 2 | n=3 | $\sum_{i=0}^n$ |
|-----------|---|----|----|-----|----------------|
| x_i | 1 | 2 | 3 | 4 | 10 |
| y_i | 4 | 9 | 10 | 14 | 37 |
| x_i^2 | 1 | 4 | 9 | 16 | 30 |
| $x_i y_i$ | 4 | 18 | 30 | 56 | 108 |

Запишемо систему рівнянь:

$$\begin{cases} 4a_0 + 10a_1 = 37, \\ 10a_0 + 30a_1 = 108. \end{cases}$$

Помножимо перше рівняння на -3 і складемо його з другим, отримаємо:

$$\begin{aligned} -12a_0 - 30a_1 &= -111, \\ 10a_0 + 30a_1 &= 108, \\ \hline -2a_0 &= -3. \end{aligned}$$

Звідки $a_0 = 1,5$. Підставимо це значення у перше рівняння:

$$4 \cdot 1,5 + 10a_1 = 37, \quad a_1 = \frac{37-6}{10} = 3,1.$$

Відповідно функція апроксимації за МНК:

$$y_m(x) = 1,5 + 3,1x^2.$$

Кусково-лінійна інтерполяція.

Значення функції з використанням лінійної інтерполяції між двома найближчими вузлами можна визначити з використанням знань аналітичної геометрії. Рівняння прямої що проходить через дві точки з координатами (x_1, y_1) и (x_2, y_2) має вигляд:

$$y(x) = y_1 + (x - x_1) \frac{(y_2 - y_1)}{(x_2 - x_1)}.$$

Оскільки $x_0 = 1,5$, підставляючи в це рівняння координати 1-го і 2-го вузлів отримуємо:

$$y_{k1}(x) = 4 + (x - 1) \frac{(9 - 4)}{(2 - 1)},$$

$$y_{k1}(x) = 5x - 1.$$

Інтерполяція поліномом Лагранжа.

Рівняння інтерполяційного полінома Лагранжа для таблиці, яка містить 4 точки даних, має вигляд:

$$\begin{aligned} L(x) = & y_0 \frac{(x - x_1)(x - x_2)(x - x_3)}{(x_0 - x_1)(x_0 - x_2)(x_0 - x_3)} + \\ & + y_1 \frac{(x - x_0)(x - x_2)(x - x_3)}{(x_1 - x_0)(x_1 - x_2)(x_1 - x_3)} + \\ & + y_2 \frac{(x - x_0)(x - x_1)(x - x_3)}{(x_2 - x_0)(x_2 - x_1)(x_2 - x_3)} + \\ & + y_3 \frac{(x - x_0)(x - x_1)(x - x_2)}{(x_3 - x_0)(x_3 - x_1)(x_3 - x_2)}. \end{aligned}$$

Підставимо координати вузлів інтерполяції та обчислимо поліном Лагранжа для невідомого x :

$$L(x) = 4 \frac{(x-2)(x-3)(x-4)}{(1-2)(1-3)(1-4)} + 9 \frac{(x-1)(x-3)(x-4)}{(2-1)(2-3)(2-4)}$$

$$+ 10 \frac{(x-1)(x-2)(x-4)}{(3-1)(3-2)(3-4)} + 14 \frac{(x-1)(x-2)(x-3)}{(4-1)(4-2)(4-3)},$$

$$L(x) = 4 \frac{x^3 - 9x^2 + 26x - 24}{-6} + 9 \frac{x^3 - 8x^2 + 19x - 12}{2}$$

$$+ 10 \frac{x^3 - 7x^2 + 14x - 8}{-2} + 14 \frac{x^3 - 6x^2 + 11x - 6}{6},$$

$$L(x) = \frac{7}{6}x^3 - 9x^2 + \frac{143}{6}x - 12.$$

Обчислення похибок інтерполяції та апроксимації.

Для обчислення скористаємось таблицею $y(1,5) = 6,5$:

| | $x_0=1,5$ | Δ | $\delta, \text{в.о.}$ | $\delta, \%$ |
|-------------|-----------|----------|-----------------------|--------------|
| $y_m(x)$ | 8,475 | 1,975 | 0,304 | 30,4 |
| $y_{kl}(x)$ | 6,5 | 0 | 0 | 0 |
| $L(x)$ | 7,438 | 0,938 | 0,144 | 14,4 |

Для 1-го відрізка заданої функції найточнішою виявилась лінійна інтерполяція.

5.6 Теорія множинності моделей

Оцінка методу найменших квадратів, заснована на «внутрішньому» критерії (коефіцієнт регресії), не придатна для знаходження моделі оптимальної складності через слабку завадостійкість внутрішніх критеріїв.

Використання коефіцієнта регресії як єдиного критерію якості моделі призводить до правила: «чим складніше модель, тим вона точніша».

Теорія множинності моделей стверджує, що за експериментальними даними принципово не можна знайти єдину модель. Наприклад, в якості полінома регресії можна взяти поліном будь-якого виду і будь-якого ступеня, і для кожного з них регресійний аналіз вкаже значення коефіцієнтів. Звідси випливає, що для кожного об'єкта, що розглядається як деякий «чорний ящик», можна скласти не одну єдину, а нескінченна безліч моделей, що мають однакові або майже однакові зовнішні прояви. Так, ясно, що не розкривши електричний вимикач, ми ніяким способом не можемо встановити, скільки розривів електричного кола він виконує: один, два або три.

У багатьох наукових роботах ще досі автори вказують, що їх теорія або формула дала криву, яка дуже близька пройшла через експериментальні точки. Такого роду «докази» правильності теорій не можна приймати серйозно.

Вирішення питання про вибір єдиного рівняння регресії оптимальної складності дає принцип зовнішнього доповнення. Тільки зовнішнє доповнення може служити арбітром при порівнянні суперечних теорій. Помилка, виміряна на всіх експериментальних точках, не є зовнішнім доповненням.

5.7 Згладжування експериментальних даних

Згладжування даних експерименту є спеціальною операцією усереднення за допомогою інтерполяційних поліномів, яка забезпечує отримання усередненого значення за заданим значенням і ряду прилеглих значень $(\dots, y_{i-2}, y_{i-1}, y_i, y_{i+1}, y_{i+2}, \dots)$, відомих з деякою похибкою. Наприклад, обчислювальні схеми лінійного згладжування за трьома точкам, [8]:

$$y_0^* = \frac{5y_0 + 2y_1 - y_2}{6},$$
$$y_i^* = \frac{y_{i-1} + y_i - y_{i+1}}{3},$$
$$y_n^* = \frac{5y_n + 2y_{n-1} - y_{n-2}}{6}.$$

Сутність лінійного згладжування при постійному кроці зводиться до наступного. Вибирають число точок згладжування N в групі, яке має бути непарним і зазвичай становить 3 або 5. Середній точці приписують індекс i , а

точкам, симетричним щодо нього, індекси $i \pm 1, i \pm 2$. Після згладжування i -ої точки, що входить в групу, її зміщують на одну і процес згладжування повторюють.

Лінійне згладжування по п'яти точках, [8]:

$$y_0^* = \frac{3y_0 + 2y_1 + y_2 - y_4}{5},$$

$$y_1^* = \frac{4y_0 + 3y_1 + 2y_2 + y_3}{10},$$

$$y_i^* = \frac{y_{i-2} + y_{i-1} + y_0 + y_{i+1} + y_{i+2}}{5},$$

$$y_{n-1}^* = \frac{y_{n-3} + 2y_{n-2} + 3y_{n-1} + 4y_n}{10},$$

$$y_{n-2}^* = \frac{3y_n + 2y_{n-1} + y_{n-2} - y_{n-4}}{5}.$$

Лінійне згладжування по семи точках, [8]:

$$y_0^* = \frac{39y_0 + 8y_1 - 4(y_2 + y_3 - y_4) + y_5 - 2y_6}{42},$$

$$y_1^* = \frac{8y_0 + 19y_1 + 16y_2 + 6y_3 - 4y_4 - 7y_5 + 4y_7}{42},$$

$$y_2^* = \frac{-4y_0 + 16y_1 + 19y_2 + 12y_3 + 2y_4 - 4y_5 + y_6}{42}.$$

$$y_i^* = \frac{7y_i + 6(y_{i+1} - y_{i-1}) + 3(y_{i+2} - y_{i-2}) - 2(y_{i+3} - y_{i-3})}{21},$$

$$y_{n-2}^* = \frac{y_{n-6} - 4y_{n-5} + 2y_{n-4} + 12y_{n-3} + 19y_{n-2} + 16y_{n-1} - 4y_n}{42}.$$

$$y_{n-1}^* = \frac{4y_{n-6} - 7y_{n-5} - 4y_{n-4} + 6y_{n-3} + 16y_{n-2} + 19y_{n-1} + 8y_n}{42},$$

$$y_n^* = \frac{-2y_{n-6} + 4y_{n-5} + y_{n-4} - 4y_{n-3} - 4y_{n-2} + 8y_{n-1} + 39y_n}{42}.$$

Питання для самоконтролю

1. У чому полягає завдання апроксимації функцій?
2. У чому полягає особливість апроксимації таблично заданої функції методом інтерполяції?
3. Як пов'язана ступінь інтерполяційного полінома з кількістю вузлів при глобальній інтерполяції?
4. В чому полягає ідея локальної інтерполяції?
5. У чому відмінність інтерполяції кубічними сплайнами від кусково-кубічної інтерполяції?
6. Як здійснюється підбір емпіричної формули для встановленої з досліду залежності?
7. Яким чином зводиться задача побудови нелінійних апроксимуючих функцій до випадку лінійної функції?
8. У чому полягає ідея методу найменших квадратів і яка основна область його застосування? Чим відрізняється цей метод від методу інтерполяції?
9. Як обґрунтовується метод найменших квадратів з ймовірнісної точки зору?
10. Сформулюйте задачу інтерполяції, назвіть галузі її використання?
11. Що називається вузлом інтерполяції?
12. Яку функцію називають інтерполяційною?
13. Як будується інтерполяційний поліном Лагранжа?
14. Що називається кінцевою різницею?
15. Як обчислюються кінцеві різниці першого, другого й n -порядку?
16. Перша та друга формули Ньютона. В чому їх різниця?

17. Проведіть порівняльний аналіз інтерполяційних поліномів Лагранжа та Ньютона?

18. Як оцінюється похибка метода інтерполяції?

19. Що називається сплайн-інтерполяцією? У чому її перевага?

20. Яким вимогам повинен задовольняти кубічний сплайн?

21. Які вбудовані функції інтерполяції ви знаєте? Як вирішується задача інтерполяції засобами пакету MathCAD?

22. Сформулюйте задачу апроксимації, назвіть області її застосування?

23. Що називають емпіричною функцією (або формулою)?

24. В чому різниця задачі апроксимації від задачі інтерполювання?

25. З яких етапів складається побудова емпіричної формули?

26. В чому сутність МНК?

27. Як обчислюється відхилення апроксимуючої функції від експериментальних значень?

28. Як отримати систему рівнянь для визначення коефіцієнтів при лінійному та квадратичному наближеннях за методом МНК?

29. Назвіть приклади апроксимуючих функцій та їх лінійні аналоги.

30. Назвіть особливості застосування ряду Фур'є.

31. Як вирішується задача апроксимації в пакеті MathCAD?

Завдання до самостійної роботи

1. Знайти самостійно параметри квадратичної і гіперболічної апроксимації і вибрати найкращу з трьох.

2. Розглянути самостійно і законспектувати тему «Метод вирівнювання (лінеаризація даних)»

3. Розглянути самостійно і законспектувати тему «Інтерполяція кубічними сплайнами», [31,32].

Варіанти контрольних завдань до лекції 5

У завданні наведено аналітичний вираз лінійної функції $y = ax + b$ і таблиця виміряних з деякими помилками значень цієї функції.

Необхідно виконати наступне:

1. По таблиці виміряних значень обчисліть коефіцієнти лінійної регресії за методом найменших квадратів.

2. Як аргумент x_0 довільно виберіть середину кожного з трьох відрізків між наявними вузлами даних.

3. Для вибраного аргументу x_0 обчисліть наступні величини: точне значення функції $y_0(x_0)$ за її аналітичного подання; наближене значення функції $y_{0A}(x_0)$ за рівнянням лінійної регресії; наближене значення функції $y_{0I}(x_0)$ за методом лінійної інтерполяції між двома найближчими вузлами; наближене значення функції $y_{0Л}(x_0)$ за інтерполяційним поліномом Лагранжа.

4. Для отриманих оцінок $y_{0A}(x_0)$, $y_{0I}(x_0)$, $y_{0Л}(x_0)$ обчисліть їх абсолютні та відносні похибки в порівнянні з точним значенням функції $y_0(x_0)$.

Варіант 1
Аналітична
функція

| a | b |
|---|----|
| 4 | -2 |
| x | y |
| 1 | 3 |
| 2 | 5 |
| 3 | 12 |
| 4 | 15 |

Варіант 2
Аналітична
функція

| a | b |
|----|-----|
| -4 | -2 |
| x | y |
| 1 | -5 |
| 2 | -11 |
| 3 | -13 |
| 4 | -16 |

Варіант 3
Аналітична
функція

| a | b |
|---|----|
| 5 | 3 |
| x | y |
| 1 | 10 |
| 2 | 14 |
| 3 | 19 |
| 4 | 21 |

Варіант 4
Аналітична
функція

| a | b |
|---|----|
| 6 | 4 |
| x | y |
| 1 | 9 |
| 2 | 14 |
| 3 | 23 |
| 4 | 30 |

Варіант 5
Аналітична
функція

| a | b |
|---|----|
| 3 | -8 |
| x | y |
| 1 | -7 |
| 2 | -1 |
| 3 | 2 |
| 4 | 3 |

Варіант 6
Аналітична
функція

| a | b |
|---|----|
| 7 | -8 |
| x | y |
| 1 | 0 |
| 2 | 7 |
| 3 | 15 |
| 4 | 19 |

Варіант 7
Аналітична
функція

| a | b |
|----|----|
| -3 | 12 |
| x | y |
| 1 | 8 |
| 2 | 7 |
| 3 | 2 |
| 4 | -1 |

Варіант 8
Аналітична
функція

| a | b |
|----|----|
| -4 | 16 |
| x | y |
| 1 | 11 |
| 2 | 10 |
| 3 | 6 |
| 4 | -1 |

Варіант 9
Аналітична
функція

| a | b |
|---|----|
| 6 | 9 |
| x | y |
| 1 | 12 |
| 2 | 23 |
| 3 | 25 |
| 4 | 36 |

Варіант 10
Аналітична
функція

| a | b |
|---|----|
| 7 | -5 |
| x | y |
| 1 | 4 |
| 2 | 6 |
| 3 | 15 |
| 4 | 27 |

Варіант 11
Аналітична
функція

| a | b |
|----|-----|
| -8 | -5 |
| x | y |
| 1 | -11 |
| 2 | -24 |
| 3 | -28 |
| 4 | -32 |

Варіант 12
Аналітична
функція

| a | b |
|----|-----|
| -6 | 5 |
| x | y |
| 1 | 0 |
| 2 | -5 |
| 3 | -15 |
| 4 | -22 |

Варіант 13
Аналітична
функція

| a | b |
|----|-----|
| -6 | 5 |
| x | y |
| 1 | -3 |
| 2 | -4 |
| 3 | -17 |
| 4 | -16 |

Варіант 14
Аналітична
функція

| a | b |
|---|----|
| 7 | 5 |
| x | y |
| 1 | 10 |
| 2 | 22 |
| 3 | 22 |
| 4 | 36 |

Варіант 15
Аналітична
функція

| a | b |
|---|----|
| 7 | -5 |
| x | y |
| 1 | 0 |
| 2 | 12 |
| 3 | 12 |
| 4 | 26 |

Варіант 16
Аналітична
функція

| a | b |
|-----|----|
| -11 | 45 |
| x | y |
| 1 | 32 |
| 2 | 26 |
| 3 | 8 |
| 4 | 4 |

Варіант 17
Аналітична
функція

| a | b |
|----|-----|
| -4 | -6 |
| x | y |
| 1 | -12 |
| 2 | -11 |
| 3 | -22 |
| 4 | -19 |

Варіант 18
Аналітична
функція

| a | b |
|----|-----|
| -8 | 3 |
| x | y |
| 1 | -7 |
| 2 | -10 |
| 3 | -25 |
| 4 | -26 |

Варіант 19
Аналітична
функція

| a | b |
|----|-----|
| -9 | 5 |
| x | y |
| 1 | -7 |
| 2 | -14 |
| 3 | -20 |
| 4 | -32 |

Варіант 20
Аналітична
функція

| a | b |
|----|----|
| 13 | 4 |
| x | y |
| 1 | 13 |
| 2 | 24 |
| 3 | 46 |
| 4 | 53 |

Варіант 21
Аналітична
функція

| a | b |
|-----|-----|
| -16 | 18 |
| x | y |
| 1 | -2 |
| 2 | -20 |
| 3 | -27 |
| 4 | -49 |

Варіант 22
Аналітична
функція

| a | b |
|----|-----|
| -8 | 1 |
| x | y |
| 1 | -9 |
| 2 | -19 |
| 3 | -20 |
| 4 | -34 |

Варіант 23
Аналітична
функція

| a | b |
|---|----|
| 9 | -5 |
| x | y |
| 1 | 1 |
| 2 | 12 |
| 3 | 24 |
| 4 | 30 |

Варіант 24
Аналітична
функція

| a | b |
|----|-----|
| -9 | -8 |
| x | y |
| 1 | -19 |
| 2 | -30 |
| 3 | -32 |
| 4 | -47 |

Варіант 25
Аналітична
функція

| a | b |
|----|-----|
| -6 | 3 |
| x | y |
| 1 | -6 |
| 2 | -5 |
| 3 | -17 |
| 4 | -18 |

Варіант 26
Аналітична
функція

| a | b |
|-----|-----|
| -11 | -3 |
| x | y |
| 1 | -17 |
| 2 | -21 |
| 3 | -38 |
| 4 | -44 |

Варіант 27
Аналітична
функція

| a | b |
|---|----|
| 6 | 11 |
| x | y |
| 1 | 14 |
| 2 | 27 |
| 3 | 27 |
| 4 | 38 |

Варіант 28
Аналітична
функція

| a | b |
|---|----|
| 7 | -3 |
| x | y |
| 1 | 1 |
| 2 | 15 |
| 3 | 16 |
| 4 | 28 |

Варіант 29
Аналітична
функція

| a | b |
|---|----|
| 8 | 8 |
| x | y |
| 1 | 13 |
| 2 | 28 |
| 3 | 30 |
| 4 | 43 |

Варіант 30
Аналітична
функція

| a | b |
|---|----|
| 4 | 12 |
| x | y |
| 1 | 14 |
| 2 | 22 |
| 3 | 22 |
| 4 | 30 |

Варіант 31
Аналітична
функція

| a | b |
|----|-----|
| -6 | -6 |
| x | y |
| 1 | -14 |
| 2 | -16 |
| 3 | -26 |
| 4 | -28 |

Варіант 32
Аналітична
функція

| a | b |
|---|----|
| 1 | 12 |
| x | y |
| 1 | 11 |
| 2 | 16 |
| 3 | 13 |
| 4 | 18 |

Варіант 33
Аналітична
функція

| a | b |
|----|-----|
| -7 | -7 |
| x | y |
| 1 | -17 |
| 2 | -17 |
| 3 | -30 |
| 4 | -32 |

Варіант 34
Аналітична
функція

| a | b |
|---|----|
| 2 | 9 |
| x | y |
| 1 | 9 |
| 2 | 15 |
| 3 | 13 |
| 4 | 19 |

Варіант 35
Аналітична
функція

| a | b |
|----|----|
| -4 | 12 |
| x | y |
| 1 | 6 |
| 2 | 6 |
| 3 | -2 |
| 4 | -2 |

Варіант 36
Аналітична
функція

| a | b |
|---|----|
| 4 | -3 |
| x | y |
| 1 | -1 |
| 2 | 7 |
| 3 | 7 |
| 4 | 15 |

ЛЕКЦІЯ 6

РОЗВ'ЯЗАННЯ ОПТИМІЗАЦІЙНИХ ЗАДАЧ

Процесс решения задачи представляет собой поиск выхода из затруднения или пути обхода препятствия, – это процесс достижения цели, которая первоначально не кажется доступной

Д. Поппа

Лекція присвячена вирішенню найбільш простого типу задач оптимізації – пошуку екстремуму функцій однієї змінної без обмежень. Така задача формулюється як знаходження такого значення вхідної змінної об'єкта, яке відповідає екстремальному (мінімальному або максимальному) значенню цільової функції.

Постановка задачі.

Для заданої цільової функції $f(x)$ у заданій області визначення $[x_1, x_2]$ необхідно знайти екстремальне (мінімальне або максимальне) значення цільової функції, а також відповідне значення аргументу x .

Задачі відшукування мінімального або максимального значення не утворюють окремих задач. З елементарних міркувань ясно, що пошук мінімуму функції $f(x)$ тотожний пошукові максимуму функції $g = -f(x)$. Тому надалі ми розглядатимемо задачу пошуку екстремуму як задачу відшукування саме максимуму цільової функції, [23].

6.1 Розв'язання одномірних задач оптимізації

6.1.1 Метод зворотного змінного кроку

Задаємо довільну точку x_0 – початкове наближення до локального максимуму, точність знаходження максимуму ε та початковий крок аргументу Δ .

Зробивши один крок від точки x_0 , отримаємо нове значення аргументу $x_1 = x_0 + \Delta$. Порівняємо значення функції, $y_0 = f(x_0)$ і $y_1 = f(x_1) = f(x_0 + \Delta)$.

Можливі два різних продовження в наближенні до точки максимуму x^* .

I. $y_1 > y_0$ – відбулося збільшення значення функції. Тоді приймемо як нове стартове значення $x_0(1) = x_1$, і зробимо крок Δ від цієї точки $x_0(1)$ до точки $x_1(1)$, тобто $x_1(1) = x_0(1) + \Delta$.

II. $y_1 < y_0$ – значення функції зменшилося. У цьому випадку початковою точкою обчислень стає точка $x_0 = x_1$, а крок аргументу Δ змінюється: $\Delta = -\beta \Delta$, де β – деяке позитивне число, $\beta < 1$.

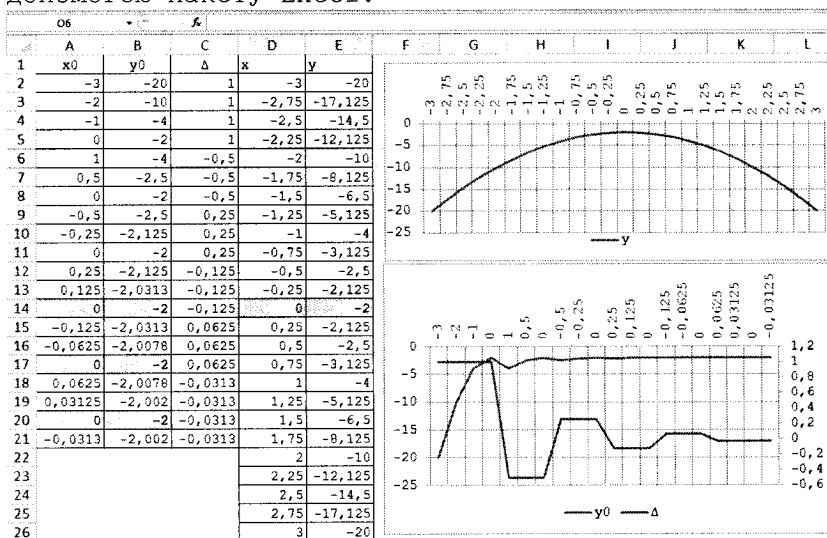
Далі повторюємо обчислення за схемою I або II, аж до досягнення заданої точності. У якості умови досягнення заданої точності обчислень виступає умова $|\Delta| < \varepsilon$.

Пошук максимуму функції зазначеним методом являє собою коливальний процес, що відбувається близько точки x^* локального мінімуму функції $f(x)$, [49].

Приклад 6.1. Знайти екстремум наведеної функції методом зворотного змінного кроку, при початковому наближенні $x_0 = -3$,

$$y(x) = -(2x^2 + 2).$$

Далі наведено розв'язання цієї задачі за допомогою пакету Excel.



6.1.2 Метод квадратичної апроксимації (метод Пауелла)

Заснований на апроксимації функції поліномом другого порядку в деякій околиці та розрахунку на його основі координати точки оптимуму, [27].

Нехай відомі значення функції в трьох точках x_0 , x_1 , x_2 складові відповідно y_0 , y_1 , y_2 . Тоді функцію $f(x)$ можна апроксимувати поліномом,

$$g(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1).$$

Оптимальне значення можна оцінити за формулою:

$$x^* = \frac{x_1 + x_0}{2} - \frac{a_1}{2a_2}. \quad (6.1)$$

де

$$a_1 = \frac{y_1 - y_0}{x_1 - x_0},$$

$$a_2 = \frac{1}{x_2 - x_1} \left(\frac{y_2 - y_0}{x_2 - x_0} - \frac{y_1 - y_0}{x_1 - x_0} \right).$$

Ми обчислюємо значення в новій, четвертій точці. Точку з найгіршим значенням функції можна відкинути. Тоді, послідовно застосовуючи цей алгоритм до трьох найкращих точок, ми будемо отримувати послідовність, яка сходиться до екстремуму x^* .

6.2 Розв'язання багатомірних задач оптимізації

Основні алгоритми будемо розглядати на прикладі функції двох змінних виду: $z = f(x, y)$.

Для наочного зображення таких функцій часто використовують лінії рівня. На рис. 6.1 наведено приклад побудови для параболоїду $z = x^2 + y^2$, на рис. 6.2 – для функції Хіммельблау.

Функція Хіммельблау – це мультимодальна функція двох змінних, використовувана для перевірки ефективності алгоритмів оптимізації, визначається формулою:

$$f(x, y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2. \quad (6.2)$$

Має локальний максимум у точках, $x = -0,270845$, $y = 0,923039$, із значенням функції $f(x, y) = 181,617$; і чотири рівнозначних локальних мінімуму у точках:

$$f(3; 2) = 0, \quad f(-2,805118; 3,131312) = 0,$$

$$f(-3,779310; -3,283186) = 0, \quad f(3,584428; -1,848126) = 0.$$

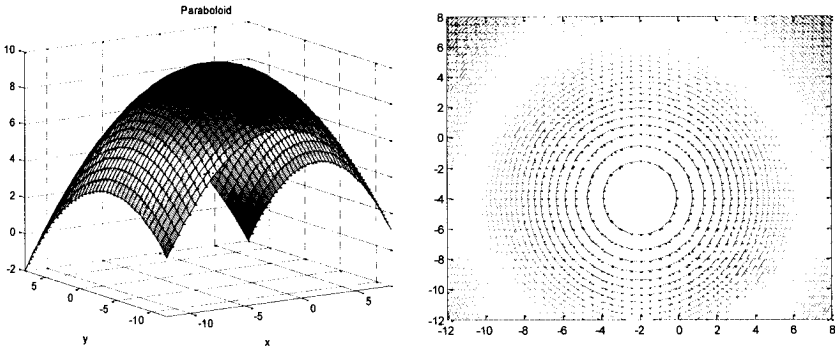


Рисунок 6.1 – Параболоїд

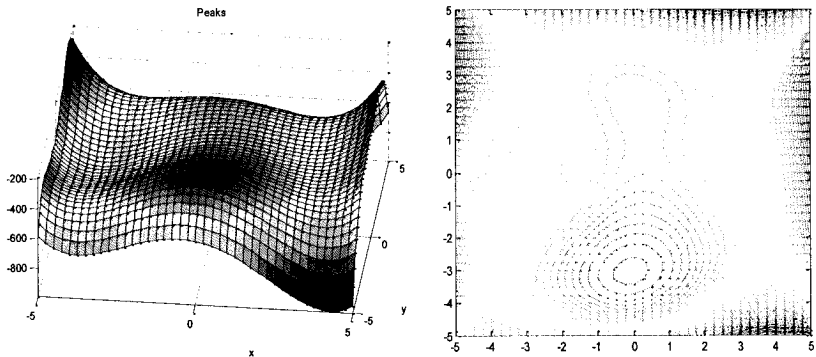


Рисунок 6.2 – Функція Хіммельблау

6.2.1 Метод покоординатного спуску

Алгоритм покоординатного спуску полягає у зведенні багатовимірної задачі до послідовності одновимірних задач, які розв'язуються методами пошуку екстремуму функції однієї змінної, [23].

Спочатку в прямокутній області зафіксуємо координату $y = y^0$, тоді функція $f(x, y^0)$ буде залежати тільки від однієї змінної x .

Знайдемо максимум x_1^0 , функції $f(x, y^0)$ змінюючи координату x . Потім зафіксуємо перший аргумент $x = x_1^0$ і знайдемо максимум y^1 функції $f(x_1^0, y)$ щодо другого аргументу y . Якщо в області пошуку функція $f(x, y)$ досить гладка, то процес спуску по координатам буде лінійно сходиться до мінімуму (рис. 6.3).

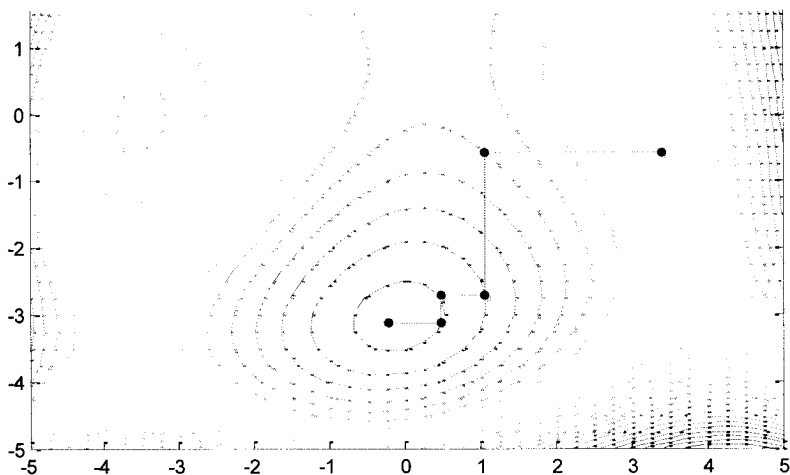


Рисунок 6.3 – Ілюстрація методу покоординатного спуску

У процесі, що сходиться, з наближенням до мінімуму функції $f(x, y)$ відстані між послідовними точками покоординатних мінімумів будуть прагнути до нуля. Тому в якості критеріїв закінчення ітераційного процесу координатного спуску вибирають умови:

$$|x^k - x^{k-1}| < \varepsilon_x, \quad |y^k - y^{k-1}| < \varepsilon_y. \quad (6.3)$$

6.2.2 Метод найшвидшого спуску (метод градієнтів)

Відомо, що градієнт функції $f(z_1, z_n, z_2)$ в кожній точці спрямований у бік найшвидшого локального зростання цієї функції. Отже, для пошуку мінімуму необхідно спускатися в протилежному напрямку.

Якщо мінімізується функція диференційована і обмежена знизу, то її градієнт задовольняє умови Ліпшиця, то ітераційний процес,

$$z_j^{k+1} = z_j^k - a \cdot \text{grad}_{z_j} f(z_1^k, z_2^k, \dots, z_n^k), \quad (6.4)$$

буде сходиться до мінімуму функції f з довільної початкової точки з координатами $z_1^0, z_2^0, \dots, z_n^0$, [41].

Параметр a у формулі визначає довжину кроку у напрямку спуску. Довжину кроку можна вибрати з умови мінімізації функції впродовж напрямку, протилежного градієнту. Такий варіант градієнтного методу називають методом найшвидшого спуску.

В іншому варіанті градієнтного спуску довжина спуску a вибирається методом дроблення. За допомогою градієнтного спуску мінімум гладких функцій знаходиться значно швидше, ніж при використанні координатного спуску. Поблизу точки мінімуму складові градієнта функції мають малі значення, що призводить до зростання чутливості ітераційного процесу до погрешностей обчислень і ускладнює пошук на заключному етапі.

6.2.3 Метод Нелдера-Міда (деформованого багатогранника)

Метод Нелдера-Міда, також відомий як метод деформованого багатогранника – метод безумовної

оптимізації функції від декількох змінних, що не використовує похідної (точніше – градієнтів) функції, а тому легко застосовний до негладких та / або зашумлених функцій.

Суть методу полягає в послідовному переміщенні і деформації симплекса навколо точки екстремуму. У двомірному просторі регулярним симплексом є правильний трикутник, а в тривимірному – правильний тетраедр.

Ідея методу полягає в порівнянні значень функції в $(n + 1)$ вершинах симплекса і переміщенні симплекса в напрямку оптимальної точки за допомогою ітераційної процедури.

У симплексному методі, запропонованому спочатку, регулярний симплекс використовувався на кожному етапі. Нелдер і Мід запропонували кілька модифікацій цього методу, що допускають, щоб симплекси були неправильними. В результаті вийшов дуже надійний метод прямого пошуку, який є одним з найефективніших при $N < 7$, [53].

Опишемо ідею цього стосовно функції двох змінних для пошуку максимуму функції $z = f(x, y)$.

На початку розрахунку задається початковий симплекс, для 2-мірного випадку, правильний трикутник, і обчислюються значення функції в вершинах трикутника.

Відділяється вершина з найменшим значенням функції z_{min} . Для решти двох вершин трикутника знаходимо центр ваги. Від знайденого центру ваги в знайденому напрямку робимо крок пошуку, знаходимо координати нової точки z_4 .

Якщо в точці z_4 знайдено нове максимальне значення, то робимо черговий крок у цьому ж напрямку.

Якщо черговий крок призводить до зменшення функції, то в тій точці проводиться відображення симплекса, вибирається новий напрямок руху (рис. 6.4).

Метод Нелдера-Міда схожий з покоординатно спуском, однак тут рух проводиться в напрямку потенційно максимального градієнта.

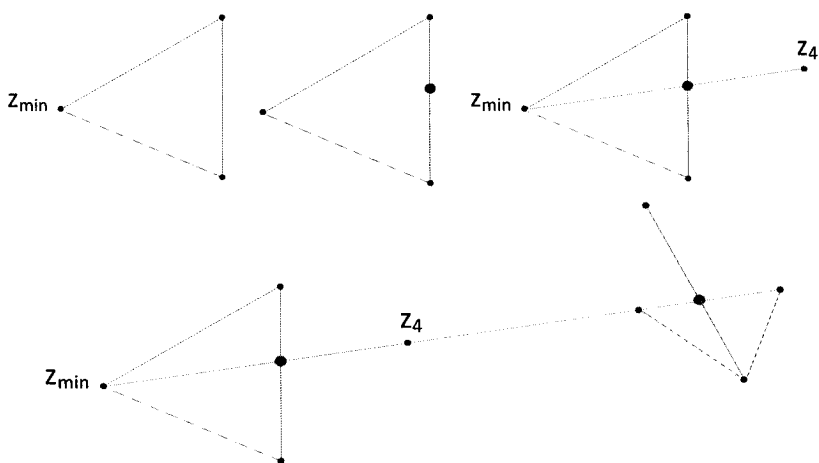


Рисунок 6.4 – Ідея методу Нелдера-Міда

Характеристики методу Нелдера-Міда:

- ❖ Метод Нелдера-Міда не накладає обмежень на гладкість функції.
- ❖ Даний метод є ефективним при низькій швидкості обчислення функції, що мінімізується.
- ❖ Відсутність теорії збіжності. Алгоритм може розходитися навіть на гладких функціях.

6.3 Методи пошуку глобального екстремуму

Всі описані вище методи забезпечують пошук локальних екстремумів функції багатьох змінних.

Для пошуку глобального екстремуму необхідно виконувати багаторазові розрахунки або застосовувати більш складні методи пошуку, які містять елементи імовірного характеру.

Для вирішення завдань такого роду був запропонований ряд алгоритмів випадкового пошуку (неспрямованого, спрямованого, спрямованого з самонавчанням), генетичні та еволюційні алгоритми, алгоритми імітації відпалу.

Однією з новітніх різновидів генетичних алгоритмів пошуку є алгоритм бджолиного рою (в англійських статтях так само зустрічаються назви Particle Swarm Optimization, Artificial Bee Colony Algorithm і Bees Algorithm), [53].

Використання методу бджолиного рою базується на ідеях моделювання багатоагентних систем, що використовуються для дослідження динаміки децентралізованих систем.

Метод бджолиного рою.

Кожна бджола в рої розглядається як частка або агент. Всі частинки рою діють індивідуально відповідно до одного керуючого принципу: рухатися в напрямку найкращої персональної та найкращої глобальної позиції, постійно перевіряючи значення поточної позиції.

Позиція бджоли є координатами в досліджуваному N -вимірному просторі.

Персональна найкраща позиція (ПНП) – позиція з найбільшим значенням цільової функції, виявлена бджолою. Кожна бджола має власну ПНП. У кожній точці вздовж шляху руху бджола порівнює значення цільової функції в поточній позиції зі значенням ПНП. Якщо поточна позиція має значення придатності вище, значення ПНП замінюється на значення поточної позиції.

Глобальна найкраща позиція (ГНП) визначається як позиція з найбільшим значенням цільової функції, виявлена всім роєм. Інформація про значення ГНП доступна кожній окремій бджолі. Якщо в процесі руху одна з бджіл виявить позицію з більш високим значенням цільової функції, ГНП замінюється поточною позицією цієї бджоли.

Алгоритм методу бджолиного рою викладено нижче.

Крок 1. Визначення області пошуку оптимальних значень, завдання параметрів алгоритму пошуку, наприклад розмір рою.

Крок 2. В області пошуку випадковим чином розміщуємо бджіл і задаємо початкові швидкості їх руху.

Крок 3. Для кожної бджоли в рої виконуємо переміщення в нове положення відповідно до її положення і швидкості. Перевіряємо вихід бджоли за границю досліджуваної області і виконуємо необхідні обмежуючі дії.

Крок 4. Для кожної бджоли обчислюємо значення цільової функції в її новій позиції. Порівнюємо це значення зі значенням ПНП бджоли, і при необхідності замінюємо ПНП поточною позицією. Порівнюємо це

значення зі значенням ГНП рою, і при необхідності замінюємо ГНП поточною позицією.

Крок 5. Для кожної бджоли обчислюємо нову швидкість переміщення відповідно до рівняння:

$$v_n^{i+1} = w \cdot v_n^i + c_1 \cdot \psi_1 \cdot (p_n - x_n) + c_2 \cdot \psi_2 \cdot (g_n - x_n), \quad (6.5)$$

де v_n^i – швидкість бджоли по вимірюванню n на i -ій ітерації; w – інерційна вага, це число (знаходиться в інтервалі $[0, 1]$) відображає, якою мірою частка зберігає свою первісну швидкість; p_n, g_n – значення координати n відповідно для ПНП бджоли і для ГНП всього рою; ψ_1, ψ_2 – випадкова величина в діапазоні $[-1, 1]$; c_1, c_2 – постійні вагові коефіцієнти, що визначають тяжіння до власної ПНП і до ГНП рою відповідно, c_1 визначає, який вплив на частинку надає її пам'ять про ПНП, а c_2 визначає, який вплив на частинку надають інші члени рою. Збільшення c_1 передбачає дослідження простору розв'язку шляхом руху кожної частинки у напрямку свого ПНП; збільшення c_2 передбачає дослідження передбачуваного глобального максимуму.

Ці коефіцієнти іноді розглядаються як пізнавальний і соціальний фактори.

Крок 6. Перевірка stop-критерію, якщо пошук не закінчений, перейти до кроку 3.

Рух кожної бджоли є компромісом між рухом у напрямку передбачуваного глобального максимуму і рухом у напрямку виявленого локального максимуму.

Приклад траєкторії однієї з бджіл у процесі пошуку максимуму функції Хіммельблау наведено на рис. 6.5.

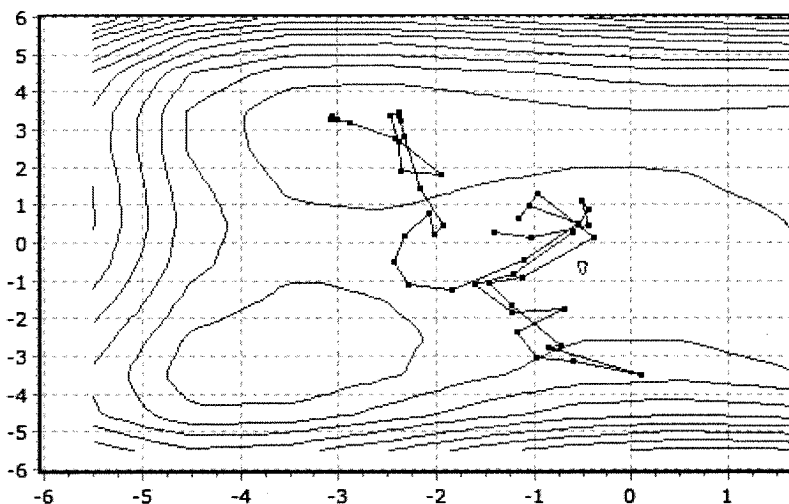


Рисунок 6.5 – Траєкторія руху частинки
(метод бджолиного рою)

Траєкторія є складною ламаною, з вигляду нагадує траєкторію частинки, що бере участь в броунівському русі.

Питання для самоконтролю

1. У чому полягає необхідна і достатня умова екстремуму одновимірної функції?
2. Як знайти екстремум функції?
3. Як поводить ся дотична до опуклої функції? Поведінка її в області екстремуму?
4. Можна вважати, що глобальний мінімум є локальним? А навпаки?
5. Поясніть суть методу Нелдера-Міда.
6. Поясніть суть методу найшвидшого спуску.
7. Поясніть суть методу покоординатного спуску.
8. Поясніть суть методу квадратичної апроксимації.

9. Поясніть суть методу зворотного змінного кроку.
10. Опишіть функцію Хіммельблау.
11. Які методи пошуку глобального екстремуму ви знаєте?
12. Що таке персональна і глобальна найкращі позиції у методі бджолиного рою?

Завдання до самостійної роботи

Використовуючи літературні джерела [23, 27, 41, 49, 53], законспекуйте відповіді на наступні питання:

1. У чому полягає умова унімодальності функції і яка умова використовується?
2. Поняття опуклої функції.
3. Як поводитьься похідна в області точки екстремуму?
4. Чи вірно твердження, що будь-яка опукла безперервна на відрізку функція є на цьому відрізку унімодальною?
5. Як поводитьься дотична до опуклої функції? Поведінка її в області екстремуму?
6. У чому відмінність між пасивним і послідовним пошуком?
7. Що називають інтервалом невизначеності в задачах одновимірної оптимізації?
8. Яким чином порівнюють ефективність методів прямого пошуку?

ЛЕКЦІЯ 7

ЧИСЕЛЬНЕ ДИФЕРЕНЦІЮВАННЯ ТА ІНТЕГРУВАННЯ

"...учащиеся старших групп должны уметь дифференцировать простые линейные формы..."

Из методички по проблемам обучения детей в детском саду: умение отличать круг от квадрата.

7.1 Чисельне диференціювання

Похідною функції $y = f(x)$ називається границя відношення приросту функції Δy до відповідного приросту аргументу Δx при зменшенні Δx до нуля:

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}, \quad \Delta y = f(x + \Delta x) - f(x). \quad (7.1)$$

У численних розрахунках похідну наближено знаходять вважаючи Δx рівним деякому кінцевому числу, отримуючи наближене рівняння для обчислення похідної:

$$f'(x) = \frac{\Delta y}{\Delta x}. \quad (7.2)$$

Це співвідношення називається апроксимацією (наближенням) похідної за допомогою кінцевих різниць. Це співвідношення може бути використано для наближеного обчислення похідної від функції, заданої як аналітичним виразом, так і таблично.

Якщо функція задана як аналітичний вираз, то вибір величини Δx довільний і визначається характером

поведінки функції. Для отримання хорошої точності величину Δx вибирають досить малою, такою щоб на інтервалі $[x, x + \Delta x]$ функція $f(x)$ була б монотонна і змінювалася несуттєво. Якщо функція задана таблично, величина Δx дорівнює різниці між сусідніми вузлами таблиці в околиці яких обчислюється похідна. При цьому, якщо кількість вузлів невелика і вузли розташовані на великій відстані один від одного, формула кінцевих різниць може давати істотну погрешність.

Загальний підхід до задачі чисельного диференціювання функції, заданої таблицею значень, ґрунтується на використанні інтерполяційних поліномів. Нехай в точках x_0, x_1, \dots, x_n відомі значення $f(x)$: y_0, y_1, \dots, y_n . За табличними даними апроксимуємо функцію $f(x)$ інтерполяційним поліном ступеня n : $f(x) \approx P_n(x)$. Тоді для k -ої похідної від функції $f(x)$ на відрізьку інтерполяції $[x_0, x_n]$ отримаємо наближену формулу:

$$\frac{d^k f(x)}{dx^k} \approx \frac{d^k P_n(x)}{dx^k}. \quad (7.3)$$

На практиці рідко вдаються до апроксимації функції інтерполяційним поліномом використовуючи методи глобальної інтерполяції, через велику похибку. Як правило, використовують локальну інтерполяцію. При цьому в околиці точки, для якій потрібно обчислити похідну, функцію інтерполують поліномом невисокого ступеня (наприклад, $n = 1, 2, 3$).

Розглянемо обчислення похідних функцій $f(x)$ в околиці табличній точки $x \sim x_i$. Для простоти будемо вважати, що табличні точки рівновіддалені одна від одної, тобто, $x_i - x_{i-1} = h = \text{const}$.

Локальна інтерполяція поліномом першого ступеня. Наблизимо в розглянутій околиці функцію $f(x)$ інтерполяційним поліномом першого ступеня $P_1(x)$, тобто прямою, що проходить через точки (x_{i-1}, y_{i-1}) і (x_i, y_i) :

$$f(x) \approx P_1(x) = y_{i-1} + \frac{y_i - y_{i-1}}{h} (x - x_{i-1}).$$

Тоді отримаємо ліве різницеве відношення:

$$\left. \frac{df}{dx} \right|_{x \sim x_i} \approx \frac{dP_1(x)}{dx} = \frac{y_i - y_{i-1}}{h}.$$

З іншого боку в даній околиці функцію $f(x)$ можна наблизити і так:

$$f(x) \approx \bar{P}_1(x) = y_i + \frac{y_{i+1} - y_i}{h} (x - x_i).$$

В цьому випадку отримаємо праве різницеве відношення:

$$\left. \frac{df}{dx} \right|_{x \sim x_i} \approx \frac{d\bar{P}_1(x)}{dx} = \frac{y_{i+1} - y_i}{h}.$$

Сенс цих назв добре зрозумілий з рис. 7.1.

Також видно, що для отримання наближених формул для другої і вищих похідних лінійного наближення функції недостатньо.

Локальна інтерполяція поліномом другого ступеня. Наблизимо в розглянутій околиці функцію $f(x)$ інтерполяційним поліномом другого ступеня $P_2(x)$, тобто параболою, що проходить значення функції в точках (x_{i-1}, y_{i-1}) , (x_i, y_i) і (x_{i+1}, y_{i+1}) :

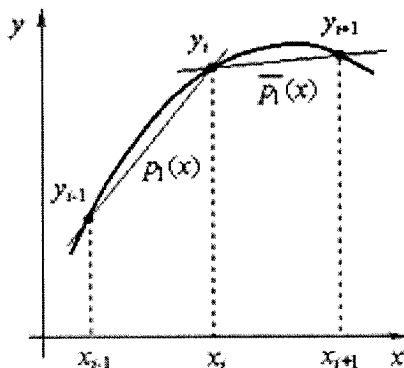


Рисунок 7.1 – Локальна заміна функції прямою

$$P_2(x) = y_i + \frac{y_{i+1} - y_{i-1}}{2h}(x - x_i) + \frac{y_{i+1} - 2y_i + y_{i-1}}{2h^2}(x - x_i)^2.$$

Диференціюючи цей вислів один раз, отримаємо нову наближену формулу для першої похідної:

$$\left. \frac{df}{dx} \right|_{x \sim x_i} \approx \frac{dP_2(x)}{dx} = \frac{y_{i+1} - y_{i-1}}{2h} + \frac{y_{i+1} - 2y_i + y_{i-1}}{h}(x - x_i).$$

Тут видно, що наближення залежить від x . Зокрема для $x = x_i$ маємо:

$$\left. \frac{df}{dx} \right|_{x=x_i} \approx \frac{y_{i+1} - y_{i-1}}{2h}.$$

Диференціюючи поліном $P_2(x)$ два рази, отримаємо наближену формулу для другої похідної:

$$\left. \frac{d^2 f}{dx^2} \right|_{x \sim x_i} \approx \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2}.$$

Аналогічним чином, залучаючи інтерполяцію поліномами більш високого ступеня, можна отримувати нові формули для першої і другої похідних, а також формули для похідних вищих порядків. Так, в разі інтерполяції функції $f(x)$ поліномом четвертого ступеня $P_4(x)$ можна отримати наступні формули (центральні різниці співвідношення) для першої і другої похідної:

$$\left. \frac{df}{dx} \right|_{x=x_i} \approx \frac{1}{12h}(y_{i-2} - 8y_{i-1} + 8y_{i+1} - y_{i+2}),$$

$$\left. \frac{d^2 f}{dx^2} \right|_{x=x_i} \approx \frac{1}{12h^2}(-y_{i-2} + 16y_{i-1} - 30y_i + 16y_{i+1} - y_{i+2}).$$

Джерелами похибки чисельного диференціювання є:

1. *Похибка апроксимації*, яка при зменшенні кроку h , як правило, зменшується.
2. *Похибки обчислення* значень функції y_i в вузлах, зростають зі зменшенням кроку h .

Наприклад, якщо при обчисленні значення функції $f(x)$ абсолютна похибка дорівнює ε , то при обчисленні дробів у формулах для лівого і правого різницевого співвідношень вона складе $2\varepsilon/h$. Тому сумарна похибка чисельного диференціювання може спадати при зменшенні кроку лише до деякого граничного значення, після чого подальше зменшення кроку не підвищить точності результату. Якщо функція задана в вузлах з абсолютною похибкою ε , то для конкретної формули чисельного диференціювання можна приблизно вказати оптимальний крок. Наприклад, для формули:

$$\left. \frac{df}{dx} \right|_{x=x_i} \approx \frac{y_{i+1} - y_{i-1}}{2h},$$

сумарна похибка (без урахування помилок) округлення приймає мінімальне значення при,

$$h = 2 \cdot \sqrt[3]{\frac{2\varepsilon}{M_3}}, \text{ де } M_3 = \min_{x_{i-1} \leq x \leq x_{i+1}} |f'''(x)|;$$

для формули:

$$\left. \frac{d^2 f}{dx^2} \right|_{x=x_j} \approx \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2},$$

вираз для оптимального кроку має вигляд:

$$h = 2 \cdot \sqrt[4]{\frac{2\varepsilon}{M_4}}, \text{ де } M_4 = \min_{x_{i-1} \leq x \leq x_{i+1}} |f''''(x)|.$$

7.2 Чисельне інтегрування

Поставимо задачу обчислити визначений інтеграл виду,

$$I = \int_a^b f(x)dx. \quad (7.4)$$

де $f(x)$ – безперервна функція на інтервалі (a, b) .

Обчислення визначеного інтеграла безпосередньо за допомогою формули Ньютона-Лейбніца:

$$\int_a^b f(x)dx = F(b) - F(a), \quad (7.5)$$

неможливо якщо:

- вид функції $f(x)$ не дозволяє аналітично виразити її первісну $F(x)$ через елементарні функції;
- підінтегральна функція $f(x)$ задана таблично.

У цих випадках використовують наближені методи інтегрування.

Сутність більшості наближених методів обчислення визначених інтегралів полягає в заміні підінтегральної функції $f(x)$ апроксимуючої функцією $\phi(x)$ для якої можна легко записати первісну в елементарних функціях, тобто,

$$\int_a^b f(x)dx = \int_a^b \phi(x)dx + R, \quad (7.6)$$

де R – похибка обчислення інтеграла.

З метою зменшення похибки, пов'язаної з апроксимацією підінтегральної функції, інтервал інтегрування (a, b) розбивають на n відрізків і на кожному частковому відрізку замінюють підінтегральну функцію апроксимуючої функцією $\phi(x)$. Тоді наближене значення інтеграла визначається сумою часткових інтегралів від функцій $\phi_i(x)$, взятих в межах від x_{i-1} до x_i :

$$\int_a^b f(x)dx \approx \sum_{i=1}^n \int_{x_{i-1}}^{x_i} \phi_i(x)dx . \quad (7.7)$$

Використовувані на практиці методи чисельного інтегрування можна згрупувати в залежності від способу апроксимації підінтегральної функції.

Методи Ньютона-Котеса засновані на поліноміальній апроксимації (інтерполяції) підінтегральної функції. Методи цього класу відрізняються один від одного ступенем використуваного полінома, від якого залежить кількість вузлів, де необхідно обчислювати функцію $f(x)$. Відрізок інтегрування розбивається, як правило, на відрізки рівної довжини, величина яких визначається як $h = (b - a) / n$ і називається кроком інтегрування.

Сплайнові методи базуються на апроксимації підінтегральної функції сплайнами. Методи розрізняються за типом вибраних сплайнів. Такі методи є сенс використовувати в задачах, де алгоритми сплайнової апроксимації застосовуються для обробки даних.

Методи найвищої алгебраїчної точності використовують нерівномірно розташовані вузли, розташовані так, щоб забезпечити мінімальну похибку інтегрування при заданій кількості вузлів. Методи розрізняються способом вибору вузлів. Найбільш широке застосування отримали методи Гаусса, в яких вузли інтегрування вибираються як корні поліномів Лежандра.

У методах Монте-Карло вузли вибираються за допомогою генератора випадкових чисел, результат в підсумку носить ймовірний характер. Методи виявляються особливо ефективні при обчисленні кратних інтегралів.

Похибка чисельного інтегрування, також як і при чисельному диференціюванні, має два основних джерела:

1. *Похибка апроксимації* – через наближену заміну підінтегральної функції апроксимуючої функцією. Зменшується зі збільшенням кількості відрізків розбиття інтервалу інтегрування n за рахунок більш точної апроксимації підінтегральної функції.

2. *Похибки неточності в обчисленні підінтегральної функції* у вузлових точках і помилки округлення зростають з ростом n і з деякого значення n^* починають переважати над похибкою апроксимації. Тому не слід вибирати надмірно велике значення n .

Отримання формул для обчислення наближеного значення інтеграла в методах Ньютона-Котеса полягає в апроксимації підінтегральної функції, на кожному частковому відрізку інтегрування (x_{i-1}, x_i) , інтервалу інтегрування (a, b) , інтерполяційним поліномом.

Точки розбивки $a = x_0, x_1, \dots, x_n = b$, будемо називати вузлами інтегрування, а відстані між вузлами $h_i = x_i - x_{i-1}$ – кроками інтегрування.

В окремому випадку крок інтегрування може бути постійним $h = (b - a) / n$.

До інтерполяційних методів Ньютона-Котеса можна віднести наступні методи:

- методи прямокутників (правих, лівих, середніх);
- метод трапецій;
- метод Сімпсона (метод парабол).

7.2.1 Методи прямокутників

Замінімо підінтегральну функцію $f(x)$ на відріжку інтегрування $[x_{i-1}, x_i]$ поліномом нульового степеню, тобто константою: $f(x) \approx \varphi_i(x) = c_i$. Підставляючи $\varphi_i(x)$ в

$\int_a^b f(x)dx \approx \sum_{i=1}^n \int_{x_{i-1}}^{x_i} \varphi_i(x)dx$ і інтегруючи, одержуємо:

$$\int_a^b f(x)dx \approx \sum_{i=1}^n c_i \int_{x_{i-1}}^{x_i} dx = \sum_{i=1}^n c_i x \Big|_{x_{i-1}}^{x_i} = \sum_{i=1}^n c_i h_i. \quad (7.8)$$

Таким чином, наближене значення інтеграла визначається сумою площ прямокутників, одна зі сторін яких є довжиною відрізка інтегрування h_i , а інша – апроксимуючою константою. Вибираючи константу c_i рівну значенню підінтегральної функції в лівій (рис. 7.2,а) або правій (рис. 7.2,б) границях відрізка $[x_{i-1}, x_i]$ приходимо до формул лівих і правих прямокутників, відповідно, вважаючи, $f_{i-1} = f(x_{i-1})$, $f_i = f(x_i)$:

$$\int_a^b f(x)dx \approx \sum_{i=1}^n h_i f_{i-1} = h_1 f_0 + h_2 f_1 + \dots + h_n f_{n-1},$$

$$\int_a^b f(x)dx \approx \sum_{i=1}^n h_i f_i = h_1 f_1 + h_2 f_2 + \dots + h_n f_n.$$

Для випадку постійного кроку формули інтегрування набувають вигляду:

- лівих прямокутників,

$$\int_a^b f(x)dx \approx \sum_{i=1}^n h_i f_{i-1} = h(f_0 + f_1 + \dots + f_{n-1}); \quad (7.9)$$

- правих прямокутників,

$$\int_a^b f(x)dx \approx \sum_{i=1}^n h_i f_i = h(f_1 + f_2 + \dots + f_n). \quad (7.10)$$

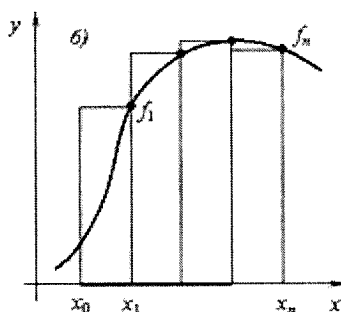
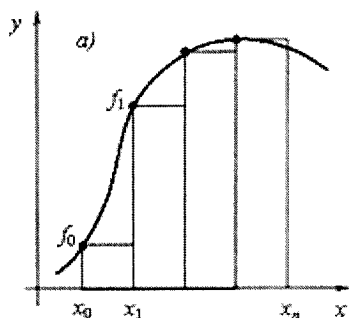


Рисунок 7.2 – Ілюстрація методів лівих (а) і правих (б) прямокутників

Найбільш широко на практиці використовується формула середніх прямокутників (рис. 7.3) в якій значення константи c_i вибирається рівною значенню підінтегральної функції в середній точці відрізка інтегрування:

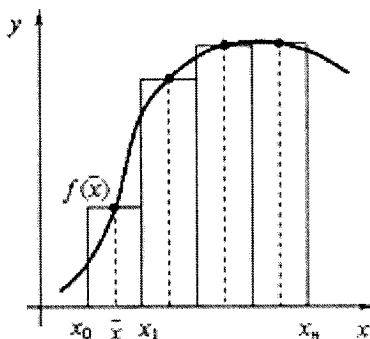


Рисунок 7.3 – Ілюстрація методу середніх прямокутників

$$\int_a^b f(x)dx \approx \sum_{i=1}^n h_i f(\bar{x}_i),$$

$$\bar{x}_i = x_{i-1} + h/2.$$

Для випадку постійного кроку інтегрування, коли $h_i = h$ формула середніх прямокутників матиме такий вигляд:

$$\int_a^b f(x)dx \approx h \sum_{i=1}^n f(\bar{x}_i). \quad (7.11)$$

З трьох розглянутих методів прямокутників метод середніх прямокутників є найбільш точним.

7.2.2 Метод трапецій

На частковому відрізку інтегрування $[x_{i-1}, x_i]$ замінимо підінтегральну функцію $f(x)$ інтерполяційним поліномом першого ступеня $f(x) \approx \phi(x)$, тобто прямої, що проходить через точки (x_{i-1}, f_{i-1}) , (x_i, f_i) :

$$\phi(x) = f_{i-1} + \frac{f_i - f_{i-1}}{x_i - x_{i-1}}(x - x_{i-1}).$$

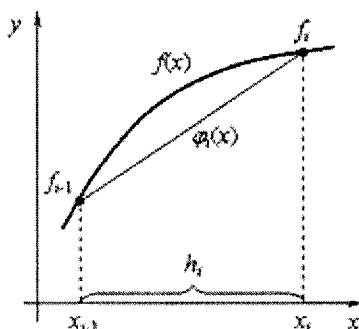


Рисунок 7.4 – Ілюстрація методу трапецій

Підставляючи цей вираз в формулу,

$$\int_a^b f(x)dx \approx \sum_{i=1}^n \int_{x_{i-1}}^{x_i} \phi_i(x)dx,$$

і виконуючи інтегрування по частковим відрізкам, приходимо до формули трапецій:

$$\int_a^b f(x)dx \approx \frac{1}{2} \sum_{i=1}^n \int_{x_{i-1}}^{x_i} h_i (f_{i-1} + f_i).$$

Для випадку постійного кроку інтегрування формула трапецій набирає вигляду:

$$\int_a^b f(x)dx \approx \frac{h}{2} [f_0 + 2f_1 + 2f_2 + \dots + 2f_{n-1} + f_n], \quad (7.12)$$

або

$$\int_a^b f(x)dx \approx \frac{h}{2} \left[\frac{f_0 + f_n}{2} + \sum_{i=1}^n f_i \right]. \quad (7.13)$$

7.2.3 Метод Сімпсона (метод парабол)

Розіб'ємо інтервал інтегрування (a, b) на парне число n рівних відрізків з кроками h . На кожному відрізку $[x_{i-1}, x_{i+1}]$, який містить три вузли, замінимо підінтегральну функцію інтерполяційним поліномом другого ступеня, тобто параболою:

$$f(x) \approx \varphi(x) = f_i + \frac{f_{i+1} - f_{i-1}}{2h}(x - x_i) + \frac{f_{i-1} - 2f_i + f_{i+1}}{2h^2}(x - x_i)^2.$$

Інтегруючи цей вираз на відрізку $[x_{i-1}, x_{i+1}]$, отримаємо:

$$\begin{aligned} \int_{x_{i-1}}^{x_{i+1}} \varphi(x) dx &= \int_{x_{i-1}}^{x_{i+1}} f_i + \frac{f_{i+1} - f_{i-1}}{2h}(x - x_i) + \\ &+ \frac{f_{i-1} - 2f_i + f_{i+1}}{2h^2}(x - x_i)^2 dx = \frac{h}{3}(f_{i-1} + 4f_i + f_{i+1}). \end{aligned}$$

Наближене значення інтеграла на інтервалі (a, b) отримаємо підсумовуванням часткових інтегралів по всіх відрізках $[x_{i-1}, x_{i+1}]$:

$$\begin{aligned} \int_a^b f(x) dx &\approx \frac{h}{3} \sum_{i=1}^{\frac{n}{2}} (f_{2i-2} + 4f_{2i-1} + f_{2i}) = \\ &= \frac{h}{3} (f_0 + 4f_1 + 2f_2 + 4f_3 + 2f_4 + \dots + 2f_{n-2} + 4f_{n-1} + f_n). \end{aligned} \quad (7.14)$$

Це співвідношення називається *формулою Сімпсона* або *формулою парабол*.

Якщо підінтегральну функцію $f(x)$ інтерполювати поліномом другого ступеня на відрізку $[x_{i-1}, x_{i+1}]$ з залученням додаткової точки $x_{i-1/2}$ – середини відрізка, то в цьому випадку число відрізків розбиття n може бути

довільним (не обов'язково парним), а формула Сімпсона в цьому випадку буде мати вигляд:

$$\int_a^b f(x)dx \approx \frac{h}{6} \sum_{i=1}^n \left(f_{i-1} + 4f_{i-\frac{1}{2}} + f_i \right) =$$

$$= \frac{h}{6} \left(f_0 + 4f_{\frac{1}{2}} + 2f_1 + 4f_{\frac{3}{2}} + 2f_2 + \dots + 2f_{n-1} + 4f_{n-\frac{1}{2}} + f_n \right).$$

При цьому $f_{i-\frac{1}{2}} = f(x_i - h/2)$, $f_i = f(x_i)$.

Отримана формула застосовна тільки в тих випадках, коли підінтегральна функція задана аналітично.

Формули, що використовуються для наближеного обчислення інтеграла, називаються *квадратурними формулами* і мають наступну структуру:

$$\int_a^b f(x)dx \approx \sum_{i=1}^n A_i f_i,$$

де f_i – значення підінтегральної функції в вузлових точках; A_i – вагові коефіцієнти, що залежать від функції $f(x)$.

7.2.4 Похибка формул Ньютона-Котеса

Методи лівих і правих прямокутників є методами першого порядку, для них похибка:

$$|R_0| \leq \frac{1}{2}(b-a)hM_1.$$

Методи середніх прямокутників і трапецій мають другий порядок точності, при цьому метод трапецій має вдвічі більшу за абсолютною величиною похибку в порівнянні з методом середніх прямокутників. Тому, якщо підінтегральна функція задана аналітично, то краще з

методів другого порядку застосовувати метод середніх прямокутників внаслідок його меншої похибки:

- для методу середніх прямокутників,

$$|R_0| \leq \frac{1}{24}(b-a)h^2 M_2;$$

- для методу трапецій,

$$|R_0| \leq \frac{1}{12}(b-a)h^2 M_2.$$

Метод Сімпсона має четвертий порядок точності з дуже малим числовим коефіцієнтом. Формула Сімпсона дозволяє отримати дуже високу точність, якщо четверта похідна підінтегральної функції не дуже велика. В іншому випадку, методи другого порядку точності можуть дати більшу точність, ніж метод Сімпсона,

$$|R_0| \leq \frac{1}{180}(b-a)h^2 M_4.$$

Наприклад, для функції $f(x) = -25x^4 + 45x^2 - 7$ формула трапецій при $n = 2$ для інтеграла в межах $(-1, 1)$ дає точний результат, рівний 4, тоді як за формулою Сімпсона отримаємо результат, неспівпадаючий навіть за знаком: $-8/3$.

7.2.5 Т-інтегрування функцій в реальному масштабі часу

Метод Т-інтегрування є чисельним методом інтегрування, розробленим Джоном Майклом Смітом, щоб спростити цифрове управління і моделювання літаків, космічних апаратів і комп'ютерне управління динамічними системами, [4]. Символ Т – це скорочення від "Tunable Numerical Integration" (адаптивне цифрове інтегрування).

Сам метод являє собою ітераційну формулу з постійним кроком інтегрування, робота якої налаштовується за допомогою коефіцієнтів фази і підсилення. Т-інтегратор передбачає налаштування параметрів фази і амплітуди.

Найпростіший варіант алгоритму Т-інтегратора виглядає наступним чином. Нехай $f(x)$ – підінтегральна функція, параметри фази і підсилення позначені символами P і G . Крім того, початкову точку інтегрування позначимо через x_0 а крок інтегрування позначимо через Δx . Т-інтегратор визначається наступною рекурентною формулою, $F_0 = 0$:

$$F_n = F_{n-1} + G\Delta x (Pf_n + (1-P)f_{n-1}). \quad (7.15)$$

Тут f_n означає $f(x_n)$. Величина F_n наближає шуканий інтеграл,

$$\int_{x_0}^{x_0 + n\Delta x} f(x) dx. \quad (7.16)$$

Якщо $G = 1$, то ця формула зводиться до наступних добре відомих чисельних методів інтегрування для заданих значень P : при $P = 0$ – формула лівих прямокутників; при $P = 1/2$ – формула трапецій; $P = 1$ – формула правих прямокутників; $P = 3/2$ – формула корекції Адамса-Башфорта та ін. Якщо G і/або P інші дійсні числа, то проводиться безліч нових інтеграторів першого порядку.

Правильний вибір параметрів G і/або P для конкретної динамічної системи може дати більш високу точність, ніж застосування стандартних схем інтегрування.

7.3 Спеціальні питання чисельного інтегрування функцій

7.3.1 Обчислення інтеграла з заданою точністю

Використовуючи наведені вище оцінки можна апріорі (до проведення розрахунку) визначити крок інтегрування h , при якому похибка обчисленого результату гарантовано не перевищить допустимий рівень похибки ε . Однак на практиці користуватися апріорними оцінками похибки не завжди зручно. Тоді контроль за точністю одержуваного результату можна організувати таким чином. На кожному частковому інтервалі обчислюють значення інтеграла з кроком інтегрування h і з кроком $h/2$. При необхідності обчислити результат з заданою точністю ε обчислення повторюють з послідовно зменшуваним (вдвічі) кроком до тих пір, поки не виконається умова:

$$|I^{(h/2)} - I^{(h)}| \leq \varepsilon.$$

При цьому довжина чергового інтервалу, за допомогою послідовного зменшення (або збільшення) початкової довжини вдвічі, встановлюється такою, щоб виконувалася нерівність:

$$|I^{(h/2)} - I^{(h)}| \leq \varepsilon \frac{h_i}{b-a}.$$

Спосіб обчислення інтеграла з автоматичним вибором кроку "приспосовується" до особливостей підінтегральної функції: в областях різкої зміни функції крок зменшується, а там, де функція змінюється слабо, – збільшується. Такого роду алгоритми називаються *адаптивними*.

7.3.2 Інтегрування розривних функцій

Якщо підінтегральна функція в деяких внутрішніх точках c_i , ($i = 1, 2, \dots$) інтервалу інтегрування терпить розрив першого роду (стрибок), то інтеграл обчислюють чисельно для кожної ділянки безперервності окремо і результат складають. Наприклад, в разі однієї точки розриву $x = c$ ($a < c < b$) маємо:

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx. \quad (7.17)$$

Для обчислення кожного з інтегралів у правій частині можна використовувати будь-який з розглянутих вище методів.

7.3.3 Методи Монте-Карло (метод статистичних випробувань)

На відрізку інтегрування (a, b) вибираємо n випадкових точок x_1, x_2, \dots, x_n , які є значеннями випадкової величини x з рівномірним розподілом на даному відрізку. Для кожної точки обчислимо площу прямокутника, одна сторона якого дорівнює $b-a$, а друга дорівнює значенню функції в цій точці $f(x_i)$. Тоді:

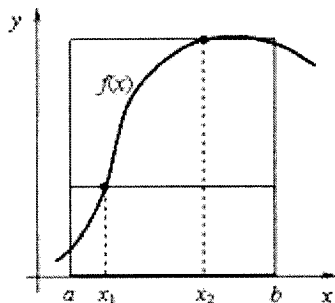


Рисунок 7.5 – Ілюстрація методу Монте-Карло

$$S_i = \frac{(b-a)}{f(x_i)}.$$

Внаслідок випадковості вузла x_i , значення площ S_i також буде носити випадковий характер. Як наближене значення інтеграла можна прийняти результат усереднення площ S_i :

$$\int_a^b f(x)dx \approx \frac{S_1 + S_2 + \dots + S_n}{n} = \frac{b-a}{n} \sum_{i=1}^n f(x_i). \quad (7.18)$$

Похибка обчислення інтеграла буде зменшуватися з ростом числа випробувань n по закону $\varepsilon \sim n^{-1/2}$.

Отримана формула формально збігається з формулою правих прямокутників, але відмінність полягає в тому, що розташування вузлів інтегрування не регулярне, а носить випадковий характер.

7.3.4 Обчислення визначених інтегралів засобами пакета MathCAD

Для обчислення визначеного інтегралу засобами пакета MathCAD необхідно скористатися кнопкою із зображенням визначеного інтеграла в математичній панелі інструментів (**Calculus Palette**). Після чого на екрані з'явиться макет інтеграла з чотирма вказівниками. Для визначення й обчислення інтеграла необхідно заповнити ці вказівники: вказати межі інтегрування, задати підінтегральний вираз й після знаку диференціювання вказати змінну, по якій проводиться інтегрування. Після цього можна обчислити значення інтеграла (якщо воно існує й може бути отримане пакетом), вказавши після нього знак '=' (обчислювальне дорівнює). Якщо процес отримання результату складний у обчислювальному відношенні, інтеграл ідентифікується зображенням

області, що містить обчислювальну конструкцію й словом WAIT у службовому рядку пакета.

Границями інтегрування можуть бути тільки дійсні константи, змінні та вирази (також символ нескінченності). Функція, що інтегрується, може бути як дійсною, так й комплексною. Усі константи й змінні (крім змінної інтегрування), що входять у підінтегральну функцію повинні бути попередньо визначені.

Приклад 7.1. Обчислити визначений інтеграл для функції $f(x)=x^2-6x+10$ на інтервалі [1; 5].

Для обчислення визначеного інтегралу необхідно задати функцію користувача та границі інтегрування:

```
< завдання підінтегральної функції та інтервалу
f(x) := x2 - 6x + 10      a:=1      b:=5
< обчислення інтегралу за формулою Ньютона-Лейбніца
SNL := ∫ab f(x) dx = 9.333
```

Приклад 7.2. Обчислити визначений інтеграл для функції,

$$f(x) = x^2 - 6x + 10,$$

на інтервалі [1; 5] методом лівих прямокутників.

Задамо кількість розбиття n , від неї залежать нумерація елементів, крок інтегрування, ранжирувана змінна x . Обчислимо визначений інтеграл за формулою лівих прямокутників.

```
n := 8      i := 0.. n      h := (b - a) / n = 0.5      xi := a + i · h
S := h · ∑i = 0n-1 f(xi) = 9.5
```

Оцінимо абсолютну та відносну похибки обчислення:

$$M(x) := \left| \frac{d}{dx} f(x) \right| \quad M := \max(M(a), M(b)) = 4$$

$$\Delta_S := M \cdot \frac{(b-a)^2}{2n} = 4$$

$$\delta_S := \frac{\Delta_S}{S_{NL}} = 0.429 \quad \delta_S \cdot 100 = 42.857$$

Висновок: визначений інтеграл для даної функції обчислений з похибкою у 42,9 %.

Визначимо необхідну кількість розбиття для забезпечення відносної похибки 5 %:

$$R := 0.05 \cdot S_{NL} = 0.467 \quad n := \left\lceil \frac{M \cdot (b-a)^2}{2 \cdot R} \right\rceil = 68.571$$

Пакет дозволяє обчислити інтеграли підінтегральна функція яких складна (наприклад, функції, які мають особливості на границях інтегрування). Однак у подібних випадках обчислення інтегралу може привести до повідомлення: ‘Not converning’.

Також на обчислення інтеграла може впливати маленьке значення змінної TOL.

Якщо особливості функції, що інтегрується відомі (наприклад, розриви, сингулярність тощо), то для обчислення інтегралу можна представити його в вигляді суми інтегралів з межами, що являють собою особливі точки, або спеціальним чином записати підінтегральну функцію, визначити й обчислити інтеграл.

Приклад 7.3. Обчислити інтеграл для наведеної функції:

$$\int_{-1}^1 |u(x)| dx, \text{ де } u(x) = \begin{cases} -2x^2, & x \in [-5, 0] \\ 8x^3, & x \in [0, 5] \end{cases}.$$

Функція, що інтегрується попередньо записується у вигляді трьох допоміжних функцій. Потім обчислюється інтеграл:

| | | | |
|---|---------------------|-----------|--------------------------------|
| $f(x) := 2 \cdot x^2$ | $y(x) := 8 \cdot x$ | $a := -5$ | $b := 5$ |
| $u(x) := \begin{cases} f(x) & \text{if } a \leq x \leq 0 \\ y(x) & \text{if } 0 < x \leq b \end{cases}$ | | | $\int_a^b u(x) dx = 183.333$ |

Можливості пакета щодо використання упорядкованих змінних як меж інтегрування, констант та змінних підінтегрального виразу дозволяють виводити результати обчислення інтеграла у вигляді таблиць та графіків первісних функцій. Наступний приклад ілюструє побудову графіка первісної для функції $f(x)$ на основі табуляції значень інтегралу за його верхньою межею інтегрування.

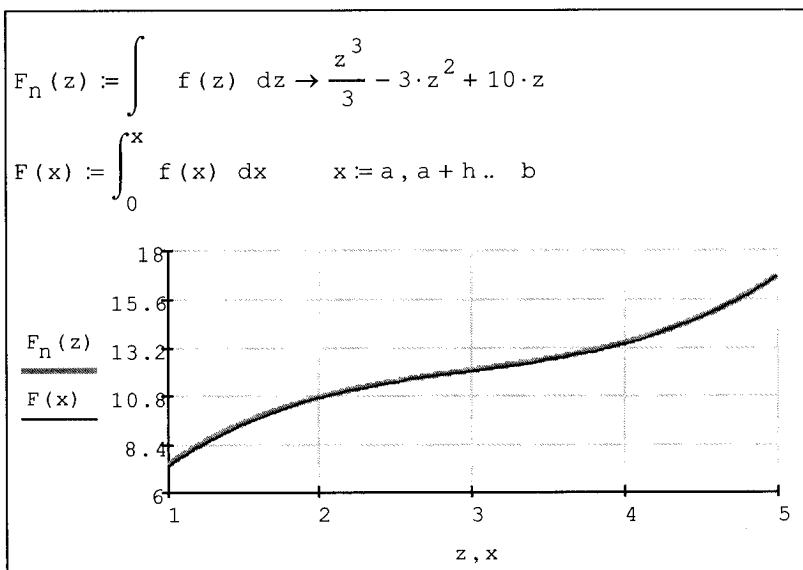
Приклад 7.4. Побудувати графік первісної для наведеної функції на інтервалі $[1; 5]$ шляхом табуляції значень визначеного інтегралу,

$$f(x) = x^2 - 6x + 10.$$

Задамо функцію, межі та крок інтегрування:

| | | | |
|-------------------------|----------|----------|------------|
| $f(x) := x^2 - 6x + 10$ | $a := 1$ | $b := 5$ | $h := 0.1$ |
|-------------------------|----------|----------|------------|

Знайдемо первісну функції в аналітичному вигляді, а також за допомогою табуляції значень визначеного інтегралу. Результати побудуємо в одній координатній площині.



Як бачимо аналітична та таблично знайдена первісні співпадають. Звісно на результат суттєво впливає вибраний крок інтегрування.

Для обчислення визначених інтегралів за методом Монте-Карло важливим питанням є отримання випадкових чисел.

У пакеті для отримання випадкових (псевдовипадкових) чисел, що рівномірно розподілені на інтервалі $[0, 1]$ передбачена функція **rnd**, яка має наступний формат:

rnd(x),

де x – права межа інтервалу $[0, x]$. Таким чином, для отримання випадкових чисел у межах $[0, 1]$ необхідно параметру x надати значення 1, тобто $x:=1$.

Приклад 7.5. Обчислити визначений інтеграл для функції $f(x)=x^2-6x+10$ на інтервалі $[1; 5]$ методом Монте-Карло¹.

Задамо функцію, межі інтегрування, кількість випадкових чисел N та випадкові точки x .

$$f(x) := x^2 - 6x + 10 \quad a := 1 \quad b := 5$$

$$N := 1000 \quad i := 1..N \quad x_i := a + (b - a) \cdot \text{rnd}(1)$$

Обчислимо значення визначеного інтегралу за методом Монте-Карло та оцінимо відносну похибку методу використовуючи значення інтегралу отримане за формулою Ньютона-Лейбніца:

$$S_{MC} := \frac{b - a}{N} \cdot \sum_{i=1}^N f(x_i) = 9.307327$$

$$S_{NL} := \int_a^b f(x) dx = 9.333333$$

$$\delta_{MC} := \frac{S_{NL} - S_{MC}}{S_{NL}} \cdot 100 = 0.278635$$

Згідно з розрахунком, методом Монте-Карло, при кількості випадкових чисел $N=1000$, визначений інтеграл знайдено з точністю **0,28 %**.

Слід зазначити, що якщо запустити опцію обчислення сторінки (Calculate Worksheet) отримаємо інше значення S_{MC} і, відповідно, інше значення відносної похибки. Тобто на відміну від формул Ньютона-Котеса результат обчислень неможливо повторити.

¹ В цьому прикладі при виведенні результатів на екран змінено значення опції Displayed Precision (точність виводу результату на екран) з меню Number Format з 3 на 6.

Завдання до самостійної роботи

Розглянути самостійно і законспектувати тему «Обчислення інтегралів від швидкоосцилюючих функцій».

Питання для самоконтролю

1. У чому особливість чисельного диференціювання з використанням інтерполяційних поліномів?

2. Які основні джерела похибки чисельного диференціювання?

3. Чому виникає необхідність використання чисельних методів інтегрування?

4. В чому полягає ідея методів Ньютона-Котеса для наближеного обчислення визначених інтегралів?

5. Як впливає на точність чисельного інтегрування величина кроку h ? Чи можна домогтися необмеженого зменшення похибки інтегрування шляхом послідовного зменшення кроку?

6. Пояснити, на чому заснований метод прямокутників. Яким чином можна обчислити площі часткових прямокутників?

7. У чому суть методу трапецій?

8. У чому суть методу Сімпсона?

9. У якому випадку формула Сімпсона дає точне значення інтеграла?

10. Як пов'язані точність інтегрування й крок інтегрування для методів прямокутників, трапецій, Сімпсона?

11. Який підхід використовується на практиці для обчислення інтеграла із заданою точністю?

12. Яка ідея обчислення визначеного інтеграла методом Монте-Карло?

13. Як оцінити похибку обчислення інтегралу?

14. В чому полягають особливості обчислення інтегралів у пакеті MathCAD?

Варіанти контрольних завдань до лекції 7

Для аналітично заданої у табл. 7.1 функції виконати чисельне інтегрування:

1. Методами прямокутників (правих, лівих, середніх).
2. Методом трапецій.
3. Методом Сімпсона.
4. Методом Монте-Карло.

Таблиця 7.1 – Функції до контрольного завдання

| Варіант | Підінтегральна функція $f(x)$ | Межі інтегрування | |
|---------|------------------------------------|-------------------|----------|
| | | a | b |
| 1 | $x \cdot \arctg(x)$ | 5 | 6,5 |
| 2 | $x^3 \cdot \ln(x)$ | 2 | 3,5 |
| 3 | $x^2 \cdot \cos(x)$ | 3π | $3,5\pi$ |
| 4 | $\frac{x^2}{8+x^3}$ | 0 | 2 |
| 5 | $\frac{1}{(16-x^2)^{\frac{3}{2}}}$ | 0,5 | 2 |
| 6 | $\frac{\sqrt{3+x}}{3-x}$ | 2 | 2,5 |
| 7 | $x\sqrt{1+x^2}$ | 0 | 1 |

| Варіант | Підінтегральна функція $f(x)$ | Межі інтегрування | |
|---------|--------------------------------|-------------------|----------|
| | | a | b |
| 8 | $e^{\cos(x)} \sin(x)$ | $3\pi/4$ | 2π |
| 9 | $x^5/(27x^3 + 1)$ | 2 | 5 |
| 10 | $\frac{\arctg(x)^{10}}{1+x^2}$ | 2 | 3 |
| 11 | $\frac{\sin(x)^4}{\cos(x)^6}$ | 0 | $\pi/4$ |
| 12 | e^{x^2} | 0 | 2 |
| 13 | $\sqrt{(\exp(x)-1)}$ | 0 | $\ln(2)$ |
| 14 | $\ln(x)$ | 1 | 2 |
| 15 | $1/(1+x^2)$ | 0 | 1 |
| 16 | $(x-1)$ | -1 | 2 |
| 17 | $x^2\sqrt{1-x^2}$ | 0 | 1 |
| 18 | $\arcsin(x)^2$ | 0,5 | 1 |
| 19 | $\sin(2x)\cos(x)$ | 0 | $\pi/4$ |
| 20 | $1/(1+\sin(x))$ | 0 | $\pi/4$ |
| 21 | $1/(\cos(x)^4)$ | $0,1\pi$ | $0,2\pi$ |
| 22 | $1/(2x^2 + 3x + 1)$ | 1 | 2 |
| 23 | $\sqrt{9-x^2}$ | 0 | 1 |
| 24 | $1/(x^2 - 3x + 2)$ | 3 | 4 |
| 25 | $1/(\sin(x)^3 \cos(x))$ | $0,1\pi$ | $0,3\pi$ |

ЛЕКЦІЯ 8

ЧИСЕЛЬНІ МЕТОДИ РОЗВ'ЯЗАННЯ ДИФЕРЕНЦІЙНИХ РІВНЯНЬ ТА ЇХ СИСТЕМ

*Уравнение было очень сложным,
но профессор с присущей ему скромностью
назвал его обыкновенным.*

*Все это называется одним словом:
дифференциальные уравнения*

8.1 Постановка задачі

Диференціальні рівняння широко використовуються для математичного моделювання процесів і явищ в найрізноманітніших галузях науки і техніки. Рух космічних об'єктів, кінетика хімічних реакцій, динаміка біологічних популяцій в природі, моделі економічного розвитку і т.п., всі ці явища досліджуються за допомогою диференціальних рівнянь.

Будь-який процес де відбувається зміна однієї змінної по відношенню до іншої змінної, описується диференціальним рівнянням.

Залежно від числа незалежних змінних диференціальні рівняння діляться на дві істотно різні категорії:

- звичайні диференціальні рівняння, що містять одну незалежну змінну,
- рівняння в часткових похідних, що містять кілька незалежних змінних.

Загалом звичайними диференціальними рівняннями (ЗДР) називаються такі рівняння, які містять одну або кілька похідних від шуканої функції. Їх можна записати у вигляді, [36]:

$$F(x, y, y', y'', \dots, y^{(n)}) = 0, \quad (8.1)$$

де x – незалежна змінна.

Розв'язком диференціального рівняння n -го порядку називається всяка n раз диференційована функція $y = \varphi(x)$, яка після її підстановки у вихідне рівняння звертає його в тотожність.

Графік функції $y = \varphi(x)$ називають *інтегральною кривою*.

Загальний розв'язок звичайного диференціального рівняння має вигляд:

$$y = \varphi(x, c_1, c_2, \dots, c_n), \quad (8.2)$$

і містить n довільних постійних c_1, c_2, \dots, c_n .

В залежності від значень c_1, c_2, \dots, c_n ЗДР може мати нескінчену множину розв'язків.

Єдині (часткові) розв'язки отримують за допомогою додаткових умов, яким повинні задовольняти шукані розв'язки. При цьому постійні c_1, c_2, \dots, c_n отримують конкретні значення, що визначаються додатковими умовами.

Залежно від способу завдання додаткових умов для отримання часткового розв'язку диференціального рівняння розглядають три типи задач, [36]:

- задача Коші,
- крайова задача,
- задача на власні значення.

В якості додаткових умов можуть задаватися значення шуканої функції й її похідних при деяких значеннях незалежної змінної, тобто в деяких точках. Кількість додаткових умов збігається з порядком рівняння.

Задача Коші. Якщо додаткові умови задаються в одній точці, то таке завдання називається задачею Коші. Додаткові умови у задачі Коші називаються початковими умовами, а точка $x = x_0$, в якій вони задаються, – початковою точкою.

Початкові умови можуть бути задані в такий спосіб:

$$y(x_0) = y_0, \quad y'(x_0) = y_{1,0}, \quad y^{(n-1)}(x_0) = y_{n-1,0}. \quad (8.3)$$

Крайова задача. Якщо додаткові умови задаються більш ніж в одній точці, тобто при різних значеннях незалежної змінної, то така задача називається крайовою. Самі додаткові умови називаються при цьому граничними (або крайовими) умовами.

На практиці, як правило граничні умови задаються в двох точках $x = a$ і $x = b$, які є межами відрізка, на якому розглядається розв'язання диференціального рівняння. Мінімальний порядок ЗДР, для якого може бути сформульована крайова задача, дорівнює двом.

Задача на власні значення. Третій тип задач для ЗДР – це задачі на власні значення. Такі задачі відрізняються

тим, що крім шуканих функцій $y(x)$ та їх похідних в рівняння входять додатково m невідомих параметрів $\lambda_1, \lambda_2, \dots, \lambda_m$, які називаються власними значеннями.

При цьому для знаходження єдиного (часткового) розв'язку на інтервалі (a, b) необхідно задати $n + m$ граничних умов.

Як приклад можна назвати задачу визначення власних частот, задачу знаходження коефіцієнтів загасання хвильових процесів і т.д.

8.2 Методи розв'язання ЗДР

Методи розв'язання звичайних диференціальних рівнянь можна розбити на наступні групи: аналітичні, наближені та чисельні.

Аналітичні методи дозволяють отримати розв'язок у вигляді формули шляхом аналітичних перетворень.

При цьому є можливість досліджувати аналітичним способом властивості загального розв'язку і отримувати з нього часткові розв'язки. Такі методи розвинені для ряду рівнянь першого порядку (з відокремлюваними змінними, однорідних, лінійних), а також для деякого типу рівнянь вищих порядків (наприклад, лінійних з постійними коефіцієнтами).

Наближені методи засновані на різних спрощеннях самих рівнянь шляхом обґрунтованого відкидання деяких членів, що містяться в них.

У деяких випадках спочатку знаходять точний розв'язок спрощеної задачі, а потім наближено обчислюють поправки, зумовлені малими членами, відкинутими на першому етапі. На цьому засновані

методи теорії збурень. Інший підхід пов'язаний з поданням розв'язку у вигляді розкладання по малому параметру, який міститься в задачі. До даної групи належать *асимптотичні методи*.

Чисельні методи розв'язання диференціальних рівнянь застосовують, коли не вдається отримати аналітичного розв'язку і застосування наближених методів також виявляється скрутним.

Наприклад, зовні просте рівняння $y' = x^2 + y^2$ не має елементарного аналітичного розв'язку і може бути розв'язано тільки у чисельному виді.

В даний час чисельні методи є основним інструментом при дослідженні більшості науково-технічних завдань.

Найбільш поширеним і універсальним підходом до чисельного розв'язання диференціальних рівнянь є *метод кінцевих різниць*.

Суть цього підходу полягає в наступному. Область безперервної зміни незалежної змінної (аргументу) замінюється дискретною безліччю точок, які називаються вузлами. Ці вузли утворюють розрахункову сітку. Шукана функція безперервного аргументу наближено замінюється функцією дискретного аргументу на заданій сітці. Ця функція називається сітковою. Початкове диференціальне рівняння замінюється різницеvim співвідношенням щодо сіткової функції. При цьому вхідні до рівняння похідні замінюються різницеvim співвідношеннями.

Така заміна диференціального рівняння різницеvim рівнянням називається його *апроксимацією на сітці*. Таким чином, розв'язання диференціального рівняння

зводиться до відшукування значень сіткової функції у вузлах сітки.

Методи, в яких наближений розв'язок в черговому i -му вузлі явно виражається через попередні значення y_{i-1} , y_{i-2} називаються **явними методами**.

На противагу цьому методи, в яких для обчислення наближеного розв'язку в черговому i -му вузлі необхідно додатково вирішувати деякі рівняння (лінійні або нелінійні) називаються **неявними методами**.

При цьому, якщо для обчислення потрібне одне значення, то метод називається **однокроковим**, а якщо кілька попередніх значень – **багатокроковим**.

Відомі такі методи розв'язання ЗДР: Ейлера. Гюна, Рунге-Кутти, Адамса, методи прогнозу та корекції.

8.2.1 Метод Ейлера (метод ламаних)

Розглянемо розв'язання диференціального рівняння:

$$\frac{dy}{dx} = f(x, y), \quad (8.4)$$

при $a \leq x \leq b$, $y(a) = y_a$.

Замінюючи в цьому рівнянні похідну в околиці кожного i -го вузла сітки правим різницеvim співвідношенням, приходимо до різницевої схеми Ейлера, [40]:

$$\frac{y_{i+1} - y_i}{h} = f(x_i, y_i), \quad (8.5)$$

де $i = 0, 1, \dots, n-1$, $y(0) = y_a$

Послідовні значення y_i обчислюються за формулою:

$$y_{i+1} = y_i + hf(x_i, y_i). \quad (8.6)$$

Таким чином, метод Ейлера є явним однокроковим методом.

Метод Ейлера має дуже просту геометричну інтерпретацію (рис. 8.1). Шукана інтегральна крива $y(x)$ на відрізку (a, b) наближається ламаною, нахил якої на кожній елементарній ділянці $[x_i, x_{i+1}]$ визначається нахилом інтегральної кривої рівняння точки (x_i, y_i) .

Метод Ейлера має перший порядок точності $\varepsilon \sim h$.

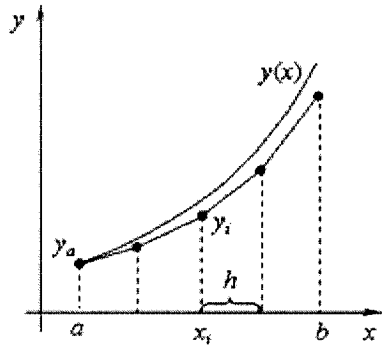


Рисунок 8.1 – Метод Ейлера

8.2.2 Модифікований метод Ейлера

Точність методу Ейлера можна підвищити, якщо скористатися для апроксимації інтеграла більш точною формулою інтегрування – формулою трапецій. Цей метод називають *методом Гюна* або *методом Ейлера з перерахунком, модифікованим методом Ейлера*.

Запишемо рівняння прямої, що виходить з точки (x_i, y_i) з нахилом, рівним нахилу інтегральної кривої середині відрізка $[x_i, x_{i+1}]$. Цій умові відповідає пряма, що проходить через точки $(x_i, y_i), (x_{i+1}, y_c)$ (рис. 8.2):

$$y = y_i + f\left(x_i + \frac{h}{2}, y_{i+1/2}\right)(x - x_i).$$

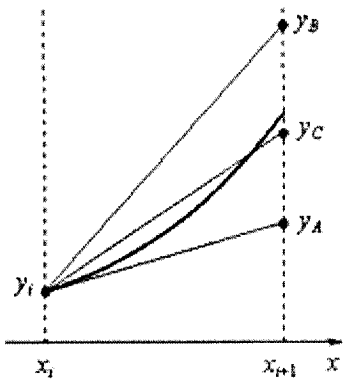


Рисунок 8.2 – Метод Гюна

Для точки $x = x_{i+1}$

отримаємо:

$$y_{i+1} = y_i + f(x_i + h/2, y_{i+1/2})h.$$

звідки витікає різницеве співвідношення:

$$\frac{y_{i+1} - y_i}{h} = f(x_i + h/2, y_{i+1/2}).$$

Якщо значення $y_{i+1/2}$, що відноситься до середини відрізка $[x_i, x_{i+1}]$, наближено обчислити за методом Ейлера:

$$y_{i+1/2} \approx y_i + \frac{h}{2} f(x_i, y_i),$$

то отримаємо двокрокову різницеву схему з другим порядком точності $\varepsilon \sim h^2$:

$$\frac{y_{i+1} - y_i}{h} = f\left(x_i + \frac{h}{2}, y_i + \frac{h}{2} f(x_i, y_i)\right), \quad (8.7)$$

де $i = 0, 1, \dots, n-1$, $y(0) = y_a$.

8.2.3 Методи Рунге-Кутти

Метод Рунге-Кутти дозволяє будувати схеми різного порядку точності. Вони найбільш застосовні в практичних обчисленнях.

Розглянемо метод другого порядку, який вимагає лише два обчислення функції $f(x, y)$ на кожному кроці, [38]:

– перше обчислення,

$$k_1 = h_i f(x_i, y_i) = h_i f_i; \quad (8.8)$$

- друге обчислення, здійснюється дробовий крок, який використовує k_1 ,

$$k_2 = h_i f(x_i + \alpha h_i, y_i + \beta k_1). \quad (8.9)$$

Для повного кроку береться комбінація значень функції:

$$y_{i+1} = y_i + \gamma_1 k_1 + \gamma_2 k_2. \quad (8.10)$$

Невідомі коефіцієнти визначаються розкладанням k_1 і k_2 в околиці точки (x_i, y_i) :

$$\gamma_1 + \gamma_2 = 1, \quad \gamma_2 \beta = \frac{1}{2}, \quad \gamma_2 \alpha = \frac{1}{2}. \quad (8.11)$$

Коефіцієнт α можна вибрати довільно поклавши його рівним 1 або 1/2. Якщо α прийняти 1/2, отримуємо обчислювальну схему:

$$y_{i+1} = y_i + h_i f\left(x_i + \frac{1}{2} h_i, y_i + \frac{1}{2} h_i f_i\right). \quad (8.12)$$

Якщо α прийняти 1:

$$y_{i+1} = y_i + \frac{h_i}{2} [f(x_i, y_i) + f(x_i + h_i, y_i + h_i f_i)]. \quad (8.13)$$

Аналогічно будуються обчислювальні схеми методу Рунге-Кутти більш високих порядків.

Найбільш широко використовується схема четвертого порядку точності, [12]:

$$y_{i+1} = y_i + \frac{h_i}{6} [k_1 + 2k_2 + 2k_3 + k_4], \quad (8.14)$$

$$k_1 = f(x_i, y_i);$$

$$k_2 = f\left(x_i + \frac{h_i}{2}, y_i + \frac{h_i}{2} k_1\right);$$

$$k_3 = f\left(x_i + \frac{h_i}{2}, y_i + \frac{h_i}{2}k_2\right);$$

$$k_4 = f(x_i + h_i, y_i + h_i k_3).$$

Метод Рунге-Кутти потребує суттєво більшого обсягу обчислень у порівнянні з методом Ейлера та його модифікаціями, проте це окупається підвищеною точністю, що дає можливість проводити обчислення з великим кроком. Іншими словами, для отримання результатів з однаковою точністю в методі Ейлера буде потрібно значно менший крок, ніж в методі Рунге-Кутти, [30].

Методи Рунге-Кутти легко програмуються, вони числово стійкі для широкого класу задач, величину кроку можна змінювати. Методи Рунге-Кутти легко переносяться на розв'язання систем диференціальних рівнянь.

8.2.4 Метод Адамса

У розглянутих однокрокових методах чисельного інтегрування диференціальних рівнянь використовують на кожному кроці обчислень інформацію про значення шуканого розв'язку тільки у попередній точці.

Обчислювальні схеми багатокрокових методів будуються так, що отримана раніше інформація про розв'язок використовується повторно на кількох кроках.

Іншими словами обчислюване значення розв'язку в поточному вузлі залежить від даних не тільки в одному попередньому вузлі, але і в ряді попередніх.

Серед багатокрокових методів різних порядків широко застосовуються *методи Адамса*.

Ідея методів Адамса полягає в тому, що якщо в декількох $k + 1$ початкових точках $x_i, x_{i-1}, \dots, x_{i-k}$ знайдено наближені значення $y(x)$: $y_i, y_{i-1}, \dots, y_{i-k}$, то по даним цих вузлів будується інтерполяційний поліном $P_k(x)$ для функції $f(x, y(x))$.

Вважаючи наближено, що $y'(x) = P_k(x)$ запишемо:

$$y_{i+1} \approx y_i + \int_{x_i}^{x_{i+1}} P_k(x) dx. \quad (8.15)$$

Якщо виразити $P_k(x)$ через кінцеві різниці до третього порядку використовуючи формулу Ньютона для інтерполяції назад, отримаємо екстраполяційну формулу Адамса-Башфорта, [46]:

$$y_{i+1} = y_i + \frac{h}{24} (55f_i - 59f_{i-1} + 37f_{i-2} - 9f_{i-3}). \quad (8.16)$$

Тут використовуються значення функції f_i в точках $x_i, x_{i-1}, x_{i-2}, x_{i-3}$ попередніх відрізка інтегрування.

Після обчислення y_{i+1} визначають $f_{i+1} = f(x_{i+1}, y_{i+1})$, а потім уточнюють y_{i+1} за такою інтерполяційною формулою Адамса-Мултона, [46]:

$$y_{i+1} = y_i + \frac{h}{24} (9f_{i+1} + 19f_i - 5f_{i-1} + f_{i-2}). \quad (8.17)$$

Формули методів Адамса мають четвертий порядок точності. Для початку обчислень потрібні, крім y_0 , ще три значення шуканого розв'язку y_1, y_2, y_3 . Їх можна визначити будь-яким однокроковим методом того ж порядку, наприклад, методом Рунге-Кутти. Величину кроку цього однокрокового методу необхідно вибрати меншою, ніж в обчисленні, для досягнення більшої точності в початкових точках відрізка інтегрування.

8.2.5 Метод Мілна

Метод Мілна є багатокроковим методом четвертого порядку точності типу "прогноз-корекція". Для його роботи необхідно знайти будь-яким однокроковим методом чотири значення шуканого розв'язку y_0, y_1, y_2, y_3 . Подальші обчислення проводяться за наступною схемою:

1. По чотирьом попереднім точкам передбачаємо наступні значення y_{i+1} :

$$y_{i+1}^{пред.} = y_{i-3} + \frac{4}{3}h(2f_i - f_{i-3} + 2f_{i-2}), \quad (8.18)$$

де $f_i = f(x_i, y_i); (i = 3, 4, 5, \dots)$.

2. Обчислюємо значення правої частини рівняння:

$$f_{i+1}^{пред.} = f(x_{i+1}, y_{i+1}^{пред.}). \quad (8.19)$$

3. Коригуємо значення y_{i+1} :

$$y_{i+1}^{кор.} = y_{i-1} + \frac{h}{3}(f_{i-1} + 4f_i + 2f_{i+1}^{пред.}). \quad (8.20)$$

де $(i = 3, 4, 5, \dots)$.

Гранична абсолютна похибка значення y_i в методі Мілна дорівнює:

$$\varepsilon = \frac{1}{29} |y_i^{пред.} - y_i^{кор.}|. \quad (8.21)$$

8.3 Автоматичний вибір кроку інтегрування та контроль за точністю обчислень

Питання вибору кроку інтегрування є дуже важливим при вирішенні задачі. Від величини кроку залежить точність одержуваного розв'язку і час, що витрачається на його отримання. Як правило вважають, що для підвищення точності слід брати менший крок. Однак:

- ❖ зменшення кроку інтегрування призводить до збільшення часу обчислень;
- ❖ занадто малі значення кроку можуть призвести не до підвищення точності, а, навпаки, до збільшення похибки в силу накопичення обчислювальної помилки;
- ❖ вибір занадто великого кроку інтегрування може привести не тільки до великої похибки, але і до отримання абсолютно невірною результату.

Тому вибір кроку це завжди певний компроміс між точністю і часом.

Критерієм, за яким можна судити про точність одержуваних результатів є порівняння наближеного результату з точним.

Однак, якщо точний результат невідомий, то критерієм точності може служити порівняння наближених результатів в кожному вузлі, отриманих при різних кроках інтегрування, наприклад h і $h/2$:

$$\Delta_y = \frac{y_{i+1}^{(h/2)} - y_{i+1}^{(h)}}{y_{i+1}^{(h/2)}}. \quad (8.22)$$

Якщо ця похибка порівнянна із заданою точністю обчислень, то крок можна збільшити; в іншому випадку, коли зазначена величина занадто велика, значення кроку слід зменшити. Використовуючи цю оцінку, можна побудувати методи з автоматичним вибором кроку і контролем точності протягом всього часу обчислень. Такі алгоритми називають **адаптивними**, тобто здатними підлаштовуватися під умови конкретного завдання, [44].

Що стосується вибору початкового (пробного) значення кроку, то тут, на жаль, не існує універсального

рецепта. І в кожному конкретному випадку крок вибирається виходячи з характерних властивостей розв'язуваної задачі.

8.4 Розв'язання диференціальних рівнянь та їх систем засобами пакету MathCAD

У пакеті реалізовано набір різноманітних методів у вигляді функцій для розв'язання диференціальних рівнянь та їх систем різного порядку. Назва функції говорить про реалізований в ній метод (алгоритм). Нижче наведені функції, що використовуються для розв'язання задачі Коши¹:

- ❖ **rkfixed**(y, x_n, x_k, n, D) – метод Рунге-Кутти 4-го порядку з фіксованим кроком.
- ❖ **rkadapt**(y, x_n, x_k, n, D) – метод Рунге-Кутти 4-го порядку з кроком, що змінюється.
- ❖ **Bulstoer**(y, x_n, x_k, n, D) – метод Булірш-Штера;
- ❖ **Adams**(y, x_n, x_k, n, D) – метод Адамса.

Аргументи функцій:

- y – вектор початкових значень;
- x_n, x_k – межі інтервалу, на якому повинен бути знайдений розв'язок;
- n – кількість кроків інтегрування;
- D – вектор-функція правих частин диференціального рівняння.

¹ Пакет має функції, що тут не розглядаються, які призначені для розв'язання диференціальних рівнянь та їх систем у випадках, коли необхідно отримати розв'язок тільки в кінцевій точці (більш ефективно й без зайвої праці).

Наведені функції повертають таблицю розв'язків одного рівняння або системи рівнянь, в якій нульовий стовбець – значення аргументу x (їх задає користувач через x_n, x_k, n), перший стовбець – значення ординат розв'язку (якщо диференціальне рівняння, що розв'язується, першого порядку); якщо розв'язується диференціальне рівняння n -го порядку або система з n диференціальних рівнянь, то розв'язок буде міститися в n стовпцях, починаючи з першого (нумерація стовбців наводиться для випадку коли значення змінної $ORIGIN=0$, якщо значення $ORIGIN=1$, то в першому стовбці буде міститися значення аргументу, а в наступних – розв'язку).

Функція **rkfixed** майже завжди справляється з поставленою задачею. Однак у деяких випадках краще використовувати більш складні методи. Для диференціальних рівнянь та їх систем, в яких функції поведуться плавно, для більш точного розв'язання краще застосовувати функцію **rkadapt**. Матриця результатів, яку повертає ця функція, буде представлена для точок, що знаходяться на однаковій відстані, яку задає користувач. Коли відомо, що розв'язок гладкий, краще використовувати функцію **Bulstoer**. У цьому випадку розв'язок буде точніше.

Слід зауважити, що розв'язання диференціального рівняння (це стосується й систем) за формулами Рунге-Кутти в пакеті MathCAD можна реалізувати (у математичному записі) кількома варіантами. Нижче наведено приклади розв'язання диференціального рівняння, використовуючи функції пакета та можливий варіант обчислення похибки результатів.

Приклад 8.1. Розв'язати рівняння $y' = y(1-x)$ на інтервалі $[0, 1]$ з початковою умовою $y(0)=1$ методом Рунге-Кутти 4-го порядку фіксованим кроком.

Дане рівняння може бути розв'язано аналітично. При заданих початкових умовах цей розв'язок має вигляд:

$$y = e^{x - \frac{1}{2}x^2}.$$

Спочатку розглянемо пряме використання формул метода Рунге-Кутти:

```

< функція для задачі Коші, інтервал
f(x, y) := y · (1 - x)      a := 0      b := 1

< кількість розбиття, крок розв'язання, нумерація
n := 10      h := (b - a) / n = 0.1      i := 1.. n

< формули Рунге-Кутти 4-го порядку
k1(x, y) := h · f(x, y)
k2(x, y) := h · f(x + h/2, y + k1(x, y)/2)
k3(x, y) := h · f(x + h/2, y + k2(x, y)/2)
k4(x, y) := h · f(x + h, y + k3(x, y))
K(x, y) := k1(x, y) + 2 · k2(x, y) + 2 · k3(x, y) + k4(x, y)

< початкові умови та
обчислення за методом Рунге-Кутти

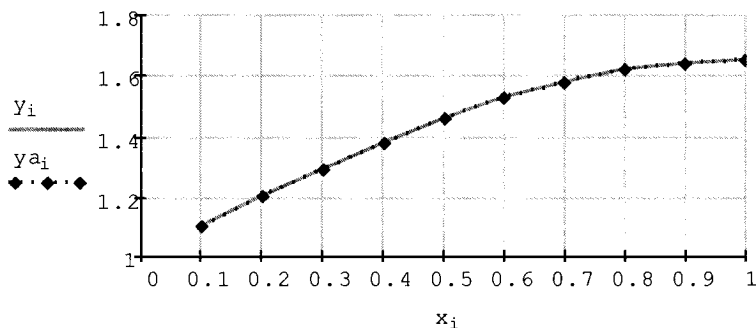
$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} := \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \begin{pmatrix} x_i \\ y_i \end{pmatrix} := \begin{pmatrix} x_{i-1} + h \\ y_{i-1} + \frac{K(x_{i-1}, y_{i-1})}{6} \end{pmatrix}$$


< функція точного (аналітичного) розв'язку

$$y_{a_i} := e^{\left[ x_i - \frac{1}{2} \cdot (x_i)^2 \right]}$$


```

< порівняння числового та аналітичного розв'язків



$$E_y := \sum_{i=1}^n (y_{a_i} - y_i)^2 = 4.683 \times 10^{-13}$$

Тепер розглянемо використання убудованих функцій пакету. Для цього задамо праву частину ЗДР, початкову умову, інтервал, кількість розбиття і функцію аналітичного розв'язку.

$D(x, y) := y_0 \cdot (1 - x) \quad y_0 := 1 \quad a := 0 \quad b := 1 \quad n := 10$

$$y_a(x) := e^{\left[x \frac{1}{2} \cdot (x)^2 \right]}$$

Реалізуємо убудовану функцію метода Рунге-Кутти з постійним кроком і перевіримо точність розв'язку у порівнянні з аналітичним.

$Z1 := \text{rkfixed}(y, a, b, n, D) \quad x := Z1^{(0)} \quad y1 := Z1^{(1)}$

$i := 0.. n \quad y_{a_i} := y_a(x_i)$

$$E_y := \sum (y_a - y1)^2 = 4.683 \times 10^{-13}$$

Також оцінимо точність згідно з (8.22). Для цього обчислимо ще одне значення y_2 з подвійним розбиттям (тобто $h/2$, або $2n$).

```
Z2 := rkfixed(y, a, b, 2·n, D)  x2 := Z2<0>  y2 := Z2<1>
```

З отриманих аргументу та функції виділимо вузли, які співпадають з вузлами x , y_1 . Обчислимо відхилення в кожному вузлі та середньоквадратичне відхилення.

```
Z2 := rkfixed(y, a, b, 2·n, D)  x2 := Z2<0>  y2 := Z2<1>
```

```
x2wi := x22·i      y2wi := y22·i
```

| | |
|----|------------------------|
| | 0 |
| 0 | 0 |
| 1 | 6.395·10 ⁻⁸ |
| 2 | 1.028·10 ⁻⁷ |
| 3 | 1.255·10 ⁻⁷ |
| 4 | 1.38·10 ⁻⁷ |
| 5 | 1.446·10 ⁻⁷ |
| 6 | 1.48·10 ⁻⁷ |
| 7 | 1.496·10 ⁻⁷ |
| 8 | 1.503·10 ⁻⁷ |
| 9 | 1.505·10 ⁻⁷ |
| 10 | 1.505·10 ⁻⁷ |

$$\Delta y := \frac{y_{2w} - y_1}{y_{2w}} =$$

$$E_y := \sum (y_{2w} - y_1)^2 = 4.145 \times 10^{-13}$$

Як бачимо середньоквадратичне відхилення обчислене для аналітичного розв'язку та для розв'язку з подвійним кроком мають однаковий порядок. В кожному вузлі відхилення становить 10^{-7} .

Реалізуємо убудовану функцію метода Рунге-Кутти зі змінним кроком і метод Булірш-Штера, та оцінимо їх відхилення від аналітичного розв'язку.

$$\begin{aligned}
Z3 &:= \text{Rkadapt}(y, a, b, n, D) & x3 &:= Z3^{(0)} & y3 &:= Z3^{(1)} \\
E3 &:= \sum a - y3^2 = 0 \\
Z4 &:= \text{Bulstoer}(y, a, b, n, D) & x4 &:= Z4^{(0)} & y4 &:= Z4^{(1)} \\
E4 &:= \sum a - y4^2 = 1.972 \times 10^{-15}
\end{aligned}$$

Приклад 8.2. Розв'язати наведену систему диференціальних рівнянь на інтервалі $[0, 1]$ з початковими умовами $y_0(0)=0, y_1(0)=0$, методами Ейлера та Рунге-Кутти з постійним кроком,

$$\begin{aligned}
\frac{dy_0}{dx} &= y_0 + y_1 - x^2 + x - 2, \\
\frac{dy_1}{dx} &= -2y_0 + 4y_1 + 2x^2 - 4x - 7.
\end{aligned}$$

Система рівнянь записується як функція користувача, де x - аргумент, y - функція, $y1$ - перша похідна функції.

```

< межі інтегрування, кількість розбиття, нумерація
a:=0    b:=1    n:=10    i:=1.. n
< обчислення кроку інтегрування

```

$$h := \frac{b - a}{n} = 0.1$$

```

< запис диференціальних рівнянь

```

$$f_1(x, y0, y1) := y0 + y1 - x^2 + x - 2$$

$$f_2(x, y0, y1) := -2 \cdot y0 + 4 \cdot y1 + 2 \cdot x^2 - 4 \cdot x - 7$$

```

< завдання вектору початкових наближень
та обчислення розв'язку за формулами Ейлера

```

$$\begin{pmatrix} x_0 \\ y0_0 \\ y1_0 \end{pmatrix} := \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \quad \begin{pmatrix} x_i \\ y0_i \\ y1_i \end{pmatrix} := \begin{pmatrix} x_{i-1} + h \\ y0_{i-1} + h \cdot f_1(x_{i-1}, y0_{i-1}, y1_{i-1}) \\ y1_{i-1} + h \cdot f_2(x_{i-1}, y0_{i-1}, y1_{i-1}) \end{pmatrix}$$

$$yE := y0$$

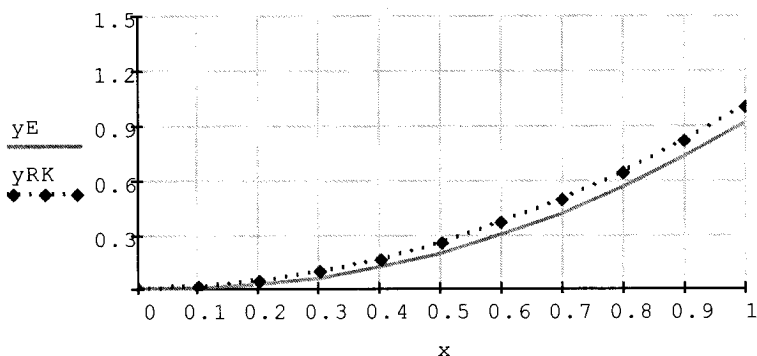
< завдання вектору початкових наближень
та матриці прaviх частин

$$y := \begin{pmatrix} 0 \\ 2 \end{pmatrix} \quad D(x, y) := \begin{pmatrix} y_0 + y_1 - x^2 + x - 2 \\ -2 \cdot y_0 + 4 \cdot y_1 + 2 \cdot x^2 - 4 \cdot x - 7 \end{pmatrix}$$

< обчислення розв'язку убудованим
методом Рунге-Кутти з постійним кроком

$$Z := \text{rkfixed}(y, a, b, n, D) \quad y_{\text{RK}} := Z \langle 1 \rangle$$

< побудова графіку



При розв'язанні система диференціальних рівнянь n -го порядку приводиться до системи диференціальних рівнянь першого порядку: вводяться позначення для шуканої функції та її $(n-1)$ похідних. Розглянемо такий випадок на прикладі.

Приклад 8.3. Розв'язати диференціальне рівняння $y'' = y(x^2 + 1)$ на інтервалі $[0, 1]$ з початковими умовами $y_0(0) = 1$, $y_1(0) = 0$.

Введемо позначення для шуканої функції та її похідної $y = y_0$, $y' = y_1$, тобто система диференціальних рівнянь запишеться так:

$$\begin{cases} y_0' = y_1, \\ y_1' = (x^2 + 1)y_0. \end{cases}$$

Аналітичний розв'язок даного рівняння:

$$y_a(x) = e^{\frac{1}{2}x^2}.$$

Розв'яжемо рівняння методом Булірш-Штера і порівняємо отриманий результат з аналітичним розв'язком:

< межі інтегрування, кількість розбиття, нумерація, обчислення кроку інтегрування

$$a := 0 \quad b := 1 \quad n := 20 \quad i := 1..n \quad h := \frac{b-a}{n} = 0.05$$

< завдання вектору початкових наближень та матриці правих частин

$$y := \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad D(x, y) := \begin{bmatrix} y_1 \\ (x^2 + 1) \cdot y_0 \end{bmatrix}$$

< обчислення розв'язку методом Булірш-Штера

$$Z := \text{Bulstoer}(y, a, b, n, D) \quad x := Z^{(0)} \quad y := Z^{(1)}$$

< завдання аналітичного розв'язку

$$y_a(x) := \exp\left(\frac{x^2}{2}\right) \quad y_{a_i} := y_a(x_i)$$

< обчислення середньоквадратичного відхилення

$$E4 := \sum_{i=1}^n (y_{a_i} - y_i)^2 = 0$$

Приклад 8.4. Провести моделювання запуску двигуна постійного струму незалежного збудження з наступними паспортними даними: потужність - 1300 кВт; частота обертання - 38 об/хв.; напруга - 600 В; струм - 2450 А; ККД - 88,5%; загальний активний опір силового кола - $R_{\Sigma} = 0,024$ Ом; індуктивність якірного кола двигуна - $L_{\Sigma} = 5,95$ мГн.

Система диференційних рівнянь ДПС незалежного збудження має вигляд:

$$\begin{cases} L_{\Sigma} \frac{dI}{dt} = U - k_E \cdot \omega - IR_{\Sigma}; \\ J \frac{d\omega}{dt} = k_M \cdot I - M_c. \end{cases}$$

Обчислимо параметри необхідні для вказаної системи ДУ:

< паспортні дані двигуна

$$P_n := 1300 \cdot 10^3 \quad n_n := 38 \quad U_n := 600 \quad I_n := 2450$$

$$\eta_n := 0.885 \quad R_\Sigma := 0.024 \quad L_\Sigma := 5.95 \cdot 10^{-3}$$

< обчислення коефіцієнтів потоку, номінального моменту двигуна та моменту інерції

$$\omega_n := \frac{2 \cdot \pi \cdot n_n}{60} = 3.979 \quad M_n := \frac{P_n}{\omega_n} = 3.267 \times 10^5$$

$$K_E := \frac{U_n - I_n \cdot R_\Sigma}{\omega_n} = 136.002 \quad K_M := \frac{M_n}{I_n} = 133.341$$

$$J := 1.5 \cdot 10^{-4} \cdot M_n^{1.5} = 2.801 \times 10^4$$

Для запису матриці правих частин перетворимо систему рівнянь таким чином, щоб зліва залишились самі похідні (розглянемо прямий пуск при номінальній напрузі без навантаження):

$$\begin{cases} \frac{dI}{dt} = \frac{U - k_E \cdot \omega - I R_E}{L_E}; \\ \frac{d\omega}{dt} = \frac{k_M \cdot I}{J}. \end{cases}$$

Введемо заміну: вектор невідомих позначимо Y , тоді струму буде відповідати нульовий елемент цієї матриці Y_0 , швидкості - перший Y_1 .

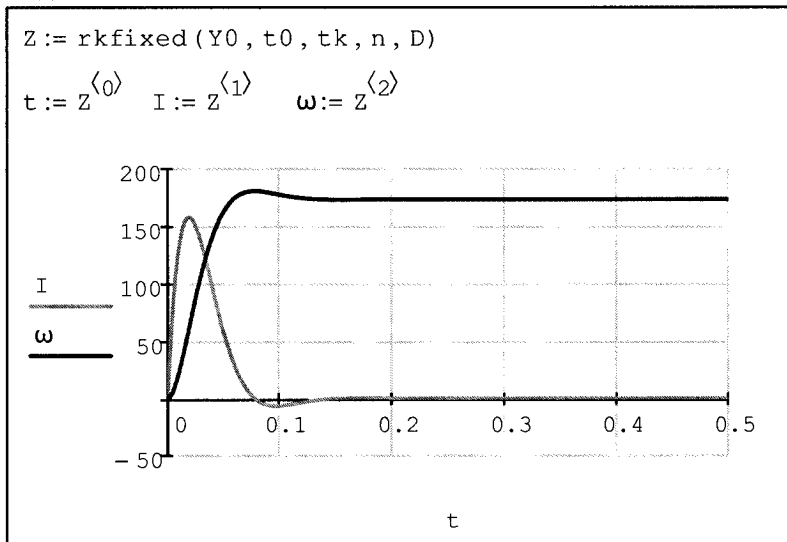
На момент пуску струм і швидкість двигуна дорівнюють нулю, відповідно початкові умови нульові. Задамо вектор початкових умов та матрицю правих частин:

$$Y_0 := \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad D(t, Y) := \begin{pmatrix} \frac{U_n - R_\Sigma \cdot Y_0 - K_E \cdot Y_1}{L_\Sigma} \\ \frac{K_M \cdot Y_0}{J} \end{pmatrix}$$

Задамо час пуску двигуна та кількість точок інтегрування:

```
t0:= 0   tk:= 0.5   n:= 1000
```

Знайдемо розв'язок методом Рунге-Кутти та побудуємо графіки перехідних процесів за струмом та швидкістю:



Питання для самоконтролю

1. У чому полягає постановка початкової і крайової задач для звичайних диференціальних рівнянь?
2. Як здійснити перехід від рівняння другого порядку до системи рівнянь першого порядку?
3. В чому полягає ідея чисельного розв'язання початкової задачі для звичайних диференціальних рівнянь?
4. Що називається порядком точності чисельного методу розв'язання ЗДР?
5. Чим відрізняються явні і неявні розрахункових схеми?

6. Чим відрізняються методи Рунге-Кутти і Адамса?
7. Що називається диференціальним рівнянням n -го порядку?
8. Сформулюйте задачу Коші.
9. Що є розв'язком диференціального рівняння?
10. Чому виникає необхідність застосовувати чисельні методи розв'язання диференціальних рівнянь?
11. В чому відміна чисельних методів розв'язання диференціальних рівнянь?
12. В чому суть методу Ейлера та його модифікацій?
13. Розробити й намалювати схему алгоритму, що реалізує метод Ейлера?
14. В чому суть методу Рунге-Кутти?
15. В яких випадках краще використовувати чисельні методи розв'язання диференціальних рівнянь з постійним та змінним кроком?
16. Розробити й намалювати схему алгоритму, що реалізує метод Рунге-Кутти.
17. Як оцінюється похибка різноманітних чисельних методів розв'язання диференціальних рівнянь?
18. В чому полягає ідея розв'язання системи диференціальних рівнянь?
19. В чому полягає ідея розв'язання диференціальних рівнянь n -го порядку?
20. В чому полягають особливості розв'язання диференціальних рівнянь та їх систем у пакеті MathCAD?

Варіанти контрольних завдань до лекції 8

Для заданої системи диференціальних рівнянь виконати чисельне розв'язання за допомогою довільно вибраного програмного засобу.

Таблиця 8.1 – Диференціальні рівняння та початкові умови до індивідуального завдання

| № | Диференціальні рівняння, межі інтегрування [a, b], початкові умови, $y(0)$ | № | Диференціальні рівняння, межі інтегрування [a, b], початкові умови, $y(0)$ |
|---|---|----|--|
| 1 | $y' = -yz + \frac{\sin(x)}{x}$ $z' = -z^2 + \frac{3,5x}{1+x^2}$ [0; 1], $y(0) = 0; z(0) = -0,4122$ | 2 | $y' = -2xy^2 + z^2 - x^2 - 1$ $z' = \frac{1}{z^2} - y - \frac{x}{z}$ [0; 0,5], $y(0) = 1; z(0) = 1$ |
| 3 | $y' = 1 - \frac{1}{z}$ $z' = \frac{1}{y-x}$ [0; 1], $y(0) = -1; z(0) = 1$ | 4 | $y' = -2xy^2 + z^2 - x - 1$ $z' = z^2 - y - x$ [0; 1], $y(0) = 1; z(0) = 1$ |
| 5 | $y' = -2z$ $z' = 2y$ [0; 5], $y(0) = 1; z(0) = 0$ | 6 | $y' = 0,5z$ $z' = -8y$ [0; 4π], $y(0) = 2; z(0) = 2$ |
| 7 | $y' = e^{-(y^2+z^2)} + 3x$ $z' = 4y^2 + z$ [0; 0,3], $y(0) = 0,5; z(0) = 1$ | 8 | $y' = 2y - 3z + 2\sin(2x)$ $z' = y - 2z - \cos(2x)$ [0; 2π], $y(0) = 1; z(0) = 1$ |
| 9 | $y' = z$ $z' = -y$ [0; 4π], $y(0) = -1; z(0) = 1$ | 10 | $y' = z$ $z' = y$ [0; 2], $y(0) = 1; z(0) = -1$ |

| № | Диференційні рівняння, межі інтегрування [a, b], початкові умови, $y(0)$ | № | Диференційні рівняння, межі інтегрування [a, b], початкові умови, $y(0)$ |
|----|---|----|--|
| 11 | $y' = e^{-(y^2+z^2)} + 2x,$ $z' = 2y^2 + z,$ <p>[0; 1], $y(0) = 0,5; z(0) = 1$</p> | 12 | $y' = \sin(y^2) + x + z,$ $z' = x + y + 1 - 3z^2,$ <p>[0; 1], $y(0) = 1; z(0) = 0,5$</p> |
| 13 | $y' = -2z$ $z' = 0,5y,$ <p>[0; 4π], $y(0) = 0; z(0) = 1$</p> | 14 | $y' = -z$ $z' = 1,01y - 0,2z,$ <p>[0; 15], $y(0) = 0; z(0) = -1$</p> |
| 15 | $y' = -2y^2 + z^2 - x^2$ $z' = \frac{1}{z^2} - y - \frac{x}{z},$ <p>[0; 1], $y(0) = 1; z(0) = 1$</p> | 16 | <p>Рівняння Енона</p> $y' = 1 + z - 1,4y^2,$ $z' = 0,3y,$ <p>[0; 20], $y(0) = 1; z(0) = 1$</p> |
| 17 | $y' = \frac{y}{x} + \sqrt{x^2 + y^2},$ $z' = \frac{y+z}{z^2 - x},$ <p>[1; 1,5], $y(0) = 0,25; z(0) = 0$</p> | 18 | <p>Нелінійний математичний маятник: $g=9,81 \text{ м/с}^2; l=1 \text{ м},$</p> $y' = z, \quad z' = \frac{g}{l} \sin(y)$ <p>[0; 1], $y(0) = -1; z(0) = 1$</p> |
| 19 | $y' = \ln\left(2x + \sqrt{2x^2 + z^2}\right),$ $z' = \sqrt{2x^2 + y^2},$ <p>[0; 1], $y(0) = 0,5; z(0) = 1$</p> | 20 | <p>Рівняння Дуффінга з двома потенційними ямами</p> $y' = z$ $z' = 0,15z + 0,5y(1 - y^2) +$ $+ 0,15\cos(0,8x)$ <p>[0; 1], $y(0) = -1; z(0) = 1$</p> |

| № | Диференційні рівняння, межі інтегрування [a, b], початкові умови, $y(0)$ | № | Диференційні рівняння, межі інтегрування [a, b], початкові умови, $y(0)$ |
|----|--|----|--|
| 21 | $y' = 2y + z$ $z' = y + 2z - e^{-2x}$ $[0; 1],$ $y(0) = 0; z(0) = 0$ | 22 | <p>Атрактор Уеди</p> $y' = z$ $z' = -0,1z - z^3 + 13,4 \cos(x)$ $[0; 1],$ $y(0) = -1; z(0) = 1$ |
| 23 | $y' = 8z$ $z' = -2y$ $[0; 4\pi],$ $y(0) = 4; z(0) = 0$ | 24 | <p>Рівняння Лотки-Вольтерра: a=4; b=2,5; c=2; d=1</p> $y' = (a - cz)y$ $z' = (-b + dy)z$ $[0; 50],$ $y(0) = 3; z(0) = 1$ |
| 25 | $y' = -z$ $z' = 10y - 7z$ $[0; 5],$ $y(0) = 0; z(0) = -7$ | 26 | $y' = -z$ $z' = 6y - z$ $[0; 5],$ $y(0) = 1; z(0) = 2$ |
| 27 | $y' = z$ $z' = -9y - 6z$ $[0; 3],$ $y(0) = -2; z(0) = -2$ | 28 | $y' = -z$ $z' = 13y + 4z$ $[0; 4],$ $y(0) = 0; z(0) = 3$ |
| 29 | $y' = 4y + z + 2x$ $z' = -2y + z$ $[0; 1],$ $y(0) = -1; z(0) = 1$ | 30 | $y' = 4y - 3z$ $z' = 9y - 7z$ $[0; 3],$ $y(0) = 3; z(0) = 2$ |

ЛЕКЦІЯ 9

ЧИСЕЛЬНЕ РОЗВ'ЯЗАННЯ СИСТЕМ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ В ЧАСТКОВИХ ПОХІДНИХ

"...Раз он в море закинул невод ..."

А.С. Пушкин. Сказка о рыбаке и рыбке

Рівняння з частковими похідними описують багато фізичних процесів в таких областях, як механіка суцільних середовищ, термодинаміка, квантова механіка, електродинаміка, теорія пружності та багато інших. Тому розділ математики, що вивчає властивості можливих розв'язків рівнянь з частковими похідними, називається математичною фізикою, а самі рівняння часто називають рівняннями математичної фізики.

Визначення точного аналітичного розв'язку, на жаль, можливо лише для дуже обмеженого кола одновимірних задач при використанні цілого ряду припущень, що негативно відбиваються на адекватності отриманих результатів. Для розв'язання задач математичної фізики в разі кількох вимірювань необхідно використовувати чисельні методи, що дозволяють перетворити диференціальні рівняння або їх системи в системи алгебраїчних рівнянь. Для розв'язання отриманих нелінійних систем алгебраїчних рівнянь або лінійних

систем великої розмірності використовують ітераційні методи. При цьому однією з найбільш складних проблем є забезпечення збіжності ітераційного процесу, що в значній мірі визначає час обчислень. Точність розв'язку визначається кроком координатної сітки, кількістю ітерацій і розрядної сіткою комп'ютера.

Математичними моделями багатьох фізичних процесів є лінійні диференціальні рівняння другого порядку. Тому основна увага в цьому розділі приділяється саме цим рівнянням.

9.1 Рівняння математичної фізики

Розробка і дослідження значної частини елементів сучасних електротехнічних систем пов'язана з розв'язання так званих задач *математичної фізики* (або скорочено – *матфізики*, до яких відносяться задачі теплопровідності, дифузії, електростатики та електродинаміки, завдання про перебіг рідини, про розподіл щільності електричного струму в провідному середовищі, завдання про деформації твердих тіл і багато інших.

Подібні задачі описуються диференціальними рівняннями в часткових похідних з додатковими рівняннями, що виражають граничні і початкові умови. У навчальному посібнику розглядаються рівняння в часткових похідних не вище другого порядку, оскільки ці рівняння охоплюють достатньо широкий діапазон фізичних явищ, покладених в основу функціонування електротехнічних систем.

У загальному випадку лінійне диференціальне рівняння в часткових похідних другого порядку з 2 незалежними змінними має вигляд, [3]:

$$\begin{aligned}
 & a(x, y) \frac{\partial^2 u}{\partial x^2} + b(x, y) \frac{\partial^2 u}{\partial x \partial y} + c(x, y) \frac{\partial^2 u}{\partial y^2} + \\
 & + e(x, y) \frac{\partial u}{\partial x} + f(x, y) \frac{\partial u}{\partial y} + g(x, y) u = f(x, y).
 \end{aligned}
 \tag{9.1}$$

Задача називається *стаціонарною*, якщо її розв'язок не залежить від часу, і *нестационарною* – якщо така залежність існує. Задача з однією просторовою змінною називаються *одновимірною*, з двома змінними – *двовимірною*, з трьома – *тривимірною*.

Коефіцієнти a, b, c – це функції, які двічі неперервно-диференційовані і не рівні нулю одночасно. У залежності від значень цих функцій розрізняють кілька типів квазілінійних диференціальних рівнянь з частковими похідними другого порядку. Щоб визначити тип рівняння в заданій точці (x_1, x_2) простору, обчислимо значення,

$$D = b^2 - 4ac.$$

Диференціальне рівняння є параболічним, якщо $D = 0$, гіперболічним – коли $D > 0$ і еліптичним – коли $D < 0$. Слід зазначити, що тип рівняння може змінюватися в залежності від значень (x_1, x_2) координат точки.

Для скорочення запису диференційних рівнянь у часткових похідних використовують спеціальні векторні оператори, [37].

Оператор «**набла** (оператор Гамільтона) – векторний диференційний оператор, компоненти якого є частковими похідними за координатами. Для тривимірного евклідового простору в прямокутній декартовій системі координат (ПДСК) оператор **набла** визначається наступним чином:

$$\nabla = \frac{\partial u}{\partial x} \bar{i} + \frac{\partial u}{\partial y} \bar{j} + \frac{\partial u}{\partial z} \bar{k}, \quad (9.2)$$

де \bar{i} , \bar{j} , \bar{k} – одиничні вектори по осях x , y , z відповідно.

Оператор Лапласа дорівнює скалярному добутку оператора набла самого на себе:

$$\Delta = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \quad (9.3)$$

За допомогою цих диференціальних операторів значно спрощується запис стандартних операцій над векторними полями, до яких належить і електромагнітне поле.

Ротор, або вихор - векторний диференціальний оператор над векторним полем. Ротор дорівнює векторному добутку оператора набла на векторне поле F :

$$\text{rot } F = \nabla \times F = \begin{bmatrix} \bar{i} & \bar{j} & \bar{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{bmatrix}. \quad (9.4)$$

Дивергенція (від лат. *Divergere* – виявляти розходження) – це лінійний диференціальний оператор на векторному полі, що характеризує потік даного поля через поверхню досить малої (в умовах конкретного завдання) околиці кожної внутрішньої точки області визначення поля. Оператор дивергенції, застосований до поля F , позначають як,

$$\text{div } F = \nabla \cdot F = \frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} + \frac{\partial u}{\partial z}. \quad (9.5)$$

Отож, дивергенція векторного поля F дорівнює скалярному добутку оператора набла на векторне поле F .

Градiєнт (від лат. *Gradiens* – крокуючий, зростаючий) – вектор, своїм напрямком вказує напрямком найбільшого зростання функції φ , значення якої змінюється від однієї точки простору в іншу (скалярного поля), а по величині (модулю) рівний швидкості росту цієї величини в цьому напрямку.

Наприклад, якщо взяти в якості φ висоту поверхні землі над рівнем моря, то її градієнт в кожній точці поверхні буде показувати «напрямок самого крутого підйому», і своєю величиною характеризувати крутизну схилу. Оператор градієнта позначають як,

$$\mathit{grad} \varphi = \left[\nabla \varphi \frac{\partial u}{\partial x} \bar{i}, \frac{\partial u}{\partial y} \bar{j}, \frac{\partial u}{\partial z} \bar{k} \right]. \quad (9.6)$$

9.1.1 Рівняння параболічного типу

Багато нестационарних, тобто таких, що змінюються в часі, фізичних процесів описуються рівняннями параболічного типу. Нестационарне рівняння теплопровідності у векторній формі має вигляд:

$$f(x, y, z) = \rho(x, y, z) C(x, y, z) \frac{\partial T}{\partial t} - \mathit{div}(k(x, y, z) \cdot \mathit{grad}(T)), \quad (9.7)$$

або у операторному вигляді,

$$f(x, y, z) = (x, y, z) C(x, y, z) \frac{\partial T}{\partial t} - \nabla(k(x, y, z) \cdot \nabla(T)).$$

Рівняння дифузії, рівняння безперервності для електронів і дірок також є рівняннями параболічного типу, і дуже важливі для моделювання процесів переносу заряду в напівпровідниках.

9.1.2 Рівняння гіперболічного типу

Багато фізичних процесів пов'язані з виникненням коливань в деякому середовищі. Наприклад, коливання струни, коливання мембрани, поширення звукових та електромагнітних коливань і ін. Вони описуються хвильовим рівнянням, яке належить до рівнянь гіперболічного типу:

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} + f(x, t) \quad (9.8)$$

при $0 < x < l, t \geq 0$

Це рівняння описує малі поздовжні коливання стрижня і поперечні коливання струни, де u – відхилення від положення рівноваги і a – швидкість поширення збурення. Хвильове рівняння описує процес поширення малих акустичних коливань.

9.1.3 Рівняння еліптичного типу

Більшість стаціонарних, тобто таких, що не змінюються в часі, фізичних процесів, описуються рівняннями еліптичного типу, в найпростішому випадку однорідного середовища і відсутності джерел.

Прикладом рівнянь еліптичного типу є рівняння Лапласа:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \Delta u = 0, \quad (9.9)$$

або його поширення – рівняння Пуассона,

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

Ці рівняння описують потік ідеальної рідини в стаціонарних потоках, стаціонарний розподіл температури або напруженості електричних або магнітних полів. Рівняння Лапласа описує ці процеси у разі відсутності джерел енергії або стоків, а рівняння Пуассона – ті ж процеси при наявності розподілених в області G джерел, що задаються правою частиною рівняння $-f(x, y, z)$.

9.1.4 Початкові та граничні умови

З курсу вищої математики відомо, що диференціальні рівняння, як правило, мають безліч розв'язків. Це пов'язано з появою в процесі інтегрування констант, при будь-яких значеннях яких розв'язок задовольняє вихідному рівнянню.

Розв'язання задач матфізики пов'язано з побудуванням залежностей від координат і часу певних фізичних величин, які, безумовно, повинні задовольняти вимогам однозначності, скінченності і безперервності [37]. Іншими словами, будь-яка задача матфізики передбачає пошук єдиного розв'язку (якщо він взагалі існує). Тому математичне формулювання фізичної задачі повинне крім основних рівнянь (диференціальних рівнянь у частинних похідних), що описують шукані функції всередині даної області, включати додаткові рівняння (диференціальні або алгебраїчні), що описують шукані функції на границях даної області в будь-який момент часу і в усіх внутрішніх точках області в початковий момент часу. Ці додаткові рівняння називають відповідно *граничними* і *початковими умовами задачі*.

Граничні умови. Припустимо, необхідно вирішити певну задачу, яка описується рівняннями математичної фізики, для деякої області Θ . Тоді для знаходження однозначної відповіді необхідно задати граничні умови (ГУ), тобто описати шукані змінні на границі Ω області Θ деякими рівняннями.

Якщо область Θ є певною областю в тривимірному просторі, то границя Ω буде являти собою замкнуту поверхню в цьому просторі, що обмежує задану область. Якщо область Θ є деякою поверхню в двовимірному просторі, то границя Ω буде являти собою замкнутий контур в цьому просторі, обмежуючий задану поверхню. І, нарешті, якщо область Θ є деякий відрізок в одновимірному просторі, то границя Ω буде являти собою дві точки на границях заданого відрізка.

По виду рівнянь, які задають ГУ, розрізняють граничні умови першого роду (умови Діріхле), другого роду (умови Неймана) і третього роду.

Граничні умови першого роду або крайова задача Діріхле мають вигляд:

$$u(X, t) = g(X, t), X \in \Omega, t \geq 0, \quad (9.10)$$

де $u(X, t)$ – шукана функція; $g(X, t)$ – деяка задана на границі Ω функція; X – координати граничної точки в просторі (наприклад, для тривимірного простору $X = (x, y, z)$); t – час.

Граничні умови другого роду або крайова задача Неймана мають вигляд:

$$\frac{\partial u(X, t)}{\partial n} = g(X, t), X \in \Omega, t \geq 0. \quad (9.11)$$

де n – внутрішня нормаль до границі області Ω .

Іншими словами, умови Неймана задають потік на границі, точніше, проекцію вектору потоку на внутрішню нормаль до границі.

Наприклад, в задачах теплопровідності ГУ другого роду задають тепловий потік, в завданні про розподіл електростатичного поля в непровідячому середовищі – проекцію вектору напруженості електричного поля на нормаль до границі і т.ін.

Граничні умови третього роду є більш загальним випадком крайових задач Діріхле та Неймана.

9.2 Методи розв’язання задач математичної фізики в електромеханічних системах

Методи розв’язання задач математичної фізики діляться на аналітичні і чисельні.

Для *аналітичних методів* розв’язком є алгебраїчні функції, в які підставляють значення параметрів, що визначають поле. Для *чисельних методів* розв’язок має вигляд сукупності числових значень, що описують поле для одного часткового поєднання заданих параметрів.

Перевага аналітичних методів полягає в можливості отримання спільного розв’язку, який дозволяє скласти повне уявлення про вплив різних параметрів на магнітне поле машини. На противагу цьому при використанні чисельних методів необхідно виконувати розрахунок для кожної сукупності значень параметрів, тому їх недолік полягає в тому, що загальна картина часто може бути отримана тільки ціною великої кількості розрахунків.

Серед аналітичних методів слід зазначити *метод конформних перетворень*. Метод полягає в такому

перетворенні області розрахунку, коли описи границь набувають просту форму, для якої розв'язок рівняння поля відомо. Відшукання функціональної залежності, яка правильно відображає заміну поля, є основною складністю цього методу. Конформне перетворення є, безумовно, найбільш дієвим методом аналітичного визначення лапласових полів, що дозволяє враховувати вплив границь набагато більш складної конфігурації, ніж інші аналітичні методи. Воно може бути виконано, наприклад, для визначення полів у повітряному зазорі електричної машини з урахуванням зубчатості сердечників. Найбільш підходящим в цьому випадку може виявитися конформне перетворення Шварца-Крістоффеля, [1].

Чисельні методи, які отримали в останні роки широке визнання, в значній мірі зобов'язані прогресу швидкодіючих цифрових обчислювальних машин, які дозволили дослідникам вирішувати з високим ступенем точності задачі по визначенню різноманітних фізичних полів. Без використання ЕОМ такі розрахунки були б надзвичайно трудомісткі або взагалі неможливі.

Серед чисельних методів найбільш поширеними слід визнати метод кінцевих різниць (МКР), метод кінцевих елементів (МКЕ) і метод граничних елементів (МГЕ).

9.2.1 Метод кінцевих різниць

Метод кінцевих різниць використовує заміну диференціальних рівнянь кінцево-різницевиими рівняннями, [3]. Основний недолік цього методу – складність точного опису границь і оптимального

накладення на область розрахунку кінцево-різницевої сітки.

Метод кінцевих різниць передбачає заміну безперервного розподілу скалярного або векторного магнітного потенціалу дискретним.

З цією метою область, де розраховується магнітне поле, покривається сіткою. Тому метод ще називають методом сіток.

Залежно від системи координат, в якій записані рівняння електромагнітного поля, форма осередків сітки може бути прямокутною, квадратною (прямокутна система координат) або у вигляді сегментів (полярна система координат). Система координат і відповідна їй форма осередків сітки вибираються такими, щоб найбільш точно апроксимувати границі розрахункової області. Точність апроксимації границь окремих ділянок може іноді вимагати використання сітки з різною формою осередків для однієї і тієї ж розрахункової області.

У методі кінцевих різниць розв'язання рівняння з частковими похідними, як правило, зводиться до розв'язання систем лінійних алгебраїчних рівнянь з розрідженими матрицями. Реалізація методу здійснюється в три етапи.

1. Область безперервного аргументу або аргументів замінюється дискретною безліччю вузлів, що називається різницевою сіткою. У ній виділяють внутрішні і граничні вузли. Функція дискретного аргументу, визначена на різницевій сітці, називається сітковою функцією.

2. Диференціальне рівняння, а також початкові і граничні умови заміняються (апроксимуються)

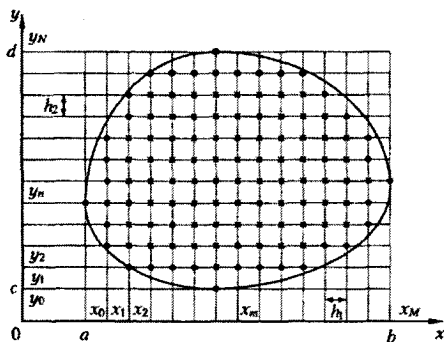
різницеви́ми аналогами. В результаті диференціальне рівняння зводиться до системи алгебраїчних (різницеви́х) рівнянь, яка називається різницево́ю схемою. Така система повинна мати єдиний розв'язок. Необхідно, щоб із збільшенням кількості вузлів сітки розв'язок різницево́ї схеми наближався (збігався) з розв'язком початкового диференціального рівняння.

3. Розв'язується система рівнянь (в більшості випадків матриця системи рівнянь має дуже велику розмірність і є розрідженою).

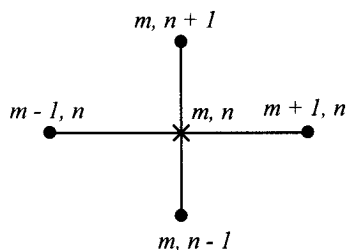
Основні ідеї методу сіток розглянемо на приклад задачі Діріхле для рівняння Пуассона з граничною умовою,

$$u(g) = \varphi(x, y); (X, y) \in g,$$

де g – границя області G (рис. 9.1, а), в якій шукається розв'язок і (x, y) , що задовольняє рівнянню Пуассона і граничним умовам.



а)



б)

Рисунок 9.1 – Область визначення розв'язку рівняння у часткових похідних: а – сіткова область; б – шаблон різницево́ї схеми

Спочатку область G безперервної зміни аргументів з границею g замінюють її сітковою областю G , з границею g_h . Для цього проводять лінії $x_m = \text{const}$ і $y_n = \text{const}$, так що $x_m = m \cdot h_1$, $m = 0, 1, \dots, M$ і $y_n = n \cdot h_2$, $n = 0, 1, \dots, N$. Величини h_1 і h_2 , які називаються кроками сітки, в загальному випадку можуть бути різними. Точки перетину ліній $x_m = \text{const}$ і $y_n = \text{const}$ називають вузлами сітки. Розрізняють два типи вузлів – внутрішні і граничні. Внутрішніми називають такі вузли, для яких чотири сусідніх вузла (по два в кожному напрямку) належать області $G + g$.

Замінімо диференційний оператор Лапласа різницеvim оператором. З цією метою виберемо шаблон різницевої схеми – система вузлів, які використовуються для заміни похідних кінцевими різницями. Шаблон, що містить p точок, називається p -точковим. Для апроксимації других похідних, що входять до оператора Лапласа, використовуємо п'ятиточковий шаблон, зображений на рис. 9.1, б.

Використовуючи розкладання функції в ряд Тейлора на п'ятиточковому шаблоні, отримаємо наближення других часткових похідних:

$$\begin{aligned} \frac{\partial^2 u_{m,n}}{\partial x^2} &= \frac{u_{m-1,n} - 2u_{m,n} + u_{m+1,n}}{h^2}, \\ \frac{\partial^2 u_{m,n}}{\partial y^2} &= \frac{u_{m,n-1} - 2u_{m,n} + u_{m,n+1}}{h^2}. \end{aligned} \quad (9.12)$$

Тоді різницеve рівняння, яке відповідає рівнянню Пуассона можна буде записати у вигляді:

$$\frac{u_{m-1,n} - 2u_{m,n} + u_{m+1,n}}{h^2} + \frac{u_{m,n-1} - 2u_{m,n} + u_{m,n+1}}{h^2} = f_{m,n}.$$

Тоді вузлове рівняння для внутрішнього вузла представимо в такому вигляді:

$$u_{m-1,n} + u_{m+1,n} - 4u_{m,n} + u_{m,n-1} + u_{m,n+1} = h^2 f_{m,n}. \quad (9.13)$$

Аналогічним чином можна отримати рівняння для граничних вузлів. Ці рівняння можуть залежати від форми границі g .

Отже, для знаходження невідомих значень $u_{m,n}$ у вузлах сітки отримаємо систему лінійних алгебраїчних рівнянь, в якій число рівнянь дорівнює кількості невідомих. Відзначимо, що кількість рівнянь може бути досить великою. Так, для вирішення завдання з високою точністю потрібно задати кількість вузлів M і N досить великою, тому кількість рівнянь може досягати декількох тисяч. Наприклад, кожне рівняння (9.13) містить всього п'ять невідомих, але в системі їх близько N^2 . Матриця цієї системи є сильно розрідженою. Існують спеціалізовані методи розв'язання СЛАР з розрідженими матрицями.

Приклад 9.1. Складіть та розв'яжіть рівняння роботи турбомеханізмів з урахуванням розподіленої комунікаційної мережі.

Рівняння поширення хвиль тиску і продуктивності у рідких і газоподібних середовищах у лінеаризованій постановці можуть бути отримані з рівнянь Нав'є-Стокса і матимуть вигляд:

$$\begin{cases} -\frac{\partial H}{\partial x} = \frac{\rho}{s} \left(\frac{\partial Q}{\partial t} + 2aQ \right); \\ -\frac{\partial H}{\partial t} = \rho \frac{c^2}{s} \times \frac{\partial Q}{\partial x}, \end{cases}$$

де H , Q - тиск і витрати в деякій частині перерізу трубопроводу; ρ - густина рідини або газу; S - площа поперечного перерізу; c - швидкість звуку в середовищі; a - середнє значення коефіцієнта розподілених втрат тиску на внутрішнє тертя; x - переміщення; t - час.

Розглянемо один з різновидів методу кінцевих різниць. Представимо частину комунікаційної мережі уздовж просторової координати x , розділену на деяке число інтервалів рівної довжини N (рис. 9.2). На кожному інтервалі тиск і продуктивність будуть сталими.

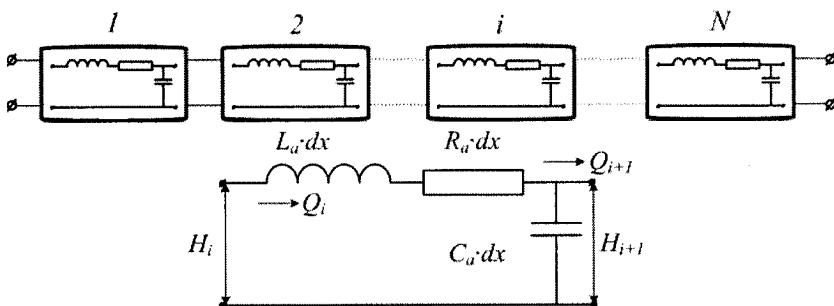


Рисунок 9.2 - До заміни довгою лінією послідовно включеними чотиріполусниками

Таке перетворення відповідає заміні довгої лінії послідовним з'єднанням N чотиріполусників (рис. 9.2), де R_0 , L_0 , c_0 - питомі акустичні параметри комунікаційної мережі; x - коефіцієнт втрат напору:

$$R_0 = \rho \frac{\lambda c}{2S}; \quad L_0 = \rho \frac{c}{S}; \quad c_0 = \frac{S}{\rho c^2}.$$

Тоді для i -го чотиріполусника рівняння тиску і продуктивності будуть мати вигляд:

$$\begin{cases} \frac{dQ_i}{dt} = \frac{H_i - H_{i-1} - R_0 Q_i}{L_0}; \\ \frac{dH_i}{dt} = \frac{Q_{i-1} - Q_i}{c_0}. \end{cases}$$

Таким чином, система рівнянь турбомеханізму, що реально складається з двох диференціальних рівнянь у

часткових похідних, наближено заміняється системою з $2N+4$ звичайних диференціальних рівнянь.

Результати розв'язання еквівалентної системи звичайних диференціальних рівнянь наведені на рис. 9.3.

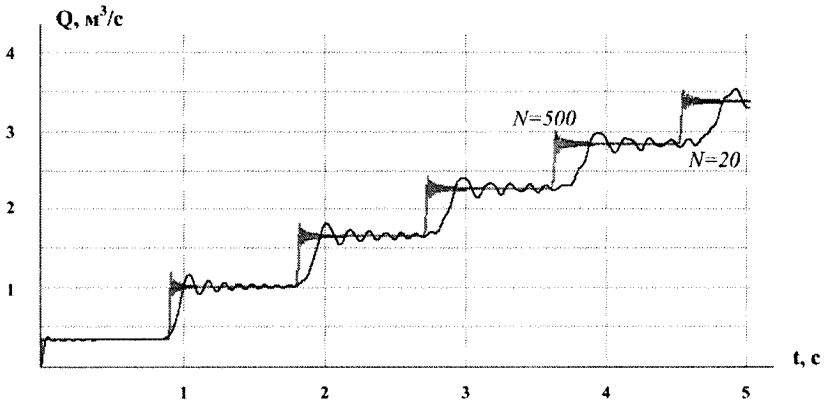


Рисунок 9.3 – Графіки переходних процесів продуктивності турбомеханізму при різних кількостях інтервалів розбиття в моделі комунікаційної мережі

При обмеженій кількості інтервалів розбиття трубопроводу на графіку витрат з'являються високочастотні коливання Гіббса. Для їх згладжування необхідно збільшувати кількість ділянок розбиття трубопроводу, що приводить до підвищення порядку системи та до значного збільшення часу розрахунку моделі.

9.2.2 Метод кінцевих елементів

Основна ідея МКЕ полягає в тому, що будь-яку безперервну функцію, таку, як векторний або скалярний магнітний потенціал, індукцію, температуру і т.п., можна апроксимувати дискретною моделлю, яка будується на безлічі кусково-неперервних функцій, визначених на кінцевому числі підобластей. Розв'язання рівнянь поля в МКЕ визначається виходячи з умови мінімуму

енергетичного функціоналу або ортогональності нев'язких рівнянь поля і інтерполяційних функцій кінцевих елементів [43, 47].

При побудові в МКЕ дискретну модель безперервної функції надходять у такий спосіб:

- ❖ в даній області фіксується скінченне число точок, які називаються вузловими;
- ❖ значення безперервної величини в кожній вузловій точці вважається змінною, яка повинна бути визначена;
- ❖ область визначення безперервної величини розбивається на кінцеве число підобластей, які називаються елементами; ці елементи мають загальні вузлові точки і в сукупності апроксимують форму області;
- ❖ безперервна величина апроксимується на кожному елементі поліномом, який визначається за допомогою вузлових значень цієї величини; для кожного елемента визначається свій поліном, який підбирають так, щоб зберігалася безперервність величини уздовж границь кожного елемента.

При використанні методу кінцевих елементів дискретизація простору виконується у вигляді елементарних трикутників. Так як необхідна дискретизація всієї області розрахунку, включаючи немагнітне середовище, застосування цього методу для моделювання поля електричної машини стикається з проблемою необхідності дуже дрібної дискретизації в області повітряного зазору. Це значно збільшує загальну кількість елементарних трикутників, ускладнюючи розрахунок.

Крім того розміри трикутників коливаються у широких межах. Дискретизація області розрахунку за допомогою елементарних трикутників зображена на рисунку 9.4.

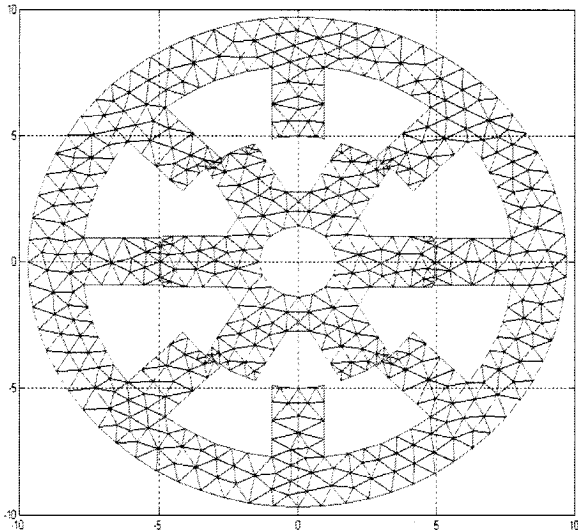


Рисунок 9.4 – Дискретизація магнітної системи вентиляно-індукторного двигуна методом триангуляції Делоне у пакеті Matlab PDE Toolbox

Виникнення МКЕ пов'язують з вирішенням завдань космічних досліджень. Вперше він був опублікований в 1956 р в статті Тернера, Клута, Мартіна і Топпа, [47]. Ця робота сприяла появі інших праць; був опублікований ряд статей із застосуваннями МКЕ до завдань будівельної механіки, механіки суцільних середовищ, інших областей техніки. Метод застосовувався до задач, описуваних рівняннями дифузії, Гельмгольца, Лапласа, Пуассона. У перших публікаціях за допомогою МКЕ вирішувалися

завдання поширення тепла. Потім цей метод був розширений і на завдання електромеханіки.

Сьогодні МКЕ має гарне теоретичне обґрунтування і з чисельної процедури розв'язання задач будівельної механіки перетворився в загальний метод чисельного розв'язання диференціальних рівнянь в часткових похідних. Цей прогрес був досягнутий завдяки вдосконаленню швидкодіючих цифрових обчислювальних машин, що дозволили значно прискорити проведення багатьох складних числових розрахунків.

Найбільш важливі переваги МКЕ в порівнянні з МКР, завдяки яким він широко використовується для розрахунку магнітних полів, полягають у наступному:

1) можливість точного опису криволінійних границь областей;

2) простота зміни дискретизації області в різних її ділянках для отримання більш високої точності розрахунків при найменшому числі вузлів розрахункової сітки;

3) можливість завдання граничних умов другого роду на границях будь-якої протяжності, а також змішаних граничних умов;

4) можливість накладення граничних умов з розривним поверхневим навантаженням.

Розрахунок магнітного поля в електричній машині чисельними методами дозволяє детально проаналізувати його розподіл в окремих елементах магнітного ланцюга. Складна межа зубцевої зони, великі розміри розрахункової області і наявність середовищ з різними магнітними проникностями вимагають великої кількості вузлів при

використанні МКР або МКЕ. Кількість вузлів, на яке розбивається область поля електричної машини, може досягати декількох десятків тисяч. Система рівнянь такого порядку може бути розв'язана тільки за допомогою досконалих обчислювальних програм і з використанням швидкодіючої обчислювальної техніки. Як правило, ці розрахунки досить трудомісткі, що ускладнює їх застосування в процедурах проектування.

Як показала практика, дуже часто розрахунок повного поля машини представляє значні складності, що пов'язано з надзвичайно великим обсягом обчислень. Крім того, визначальним фактором при виконанні розрахунку є правильне і досить ретельне накладення кінцево-елементної сітки на область розрахунку. В іншому випадку завдання може навіть не сходитися. Складність накладення сітки полягає в тому, що область розрахунку поля електричної машини має маленький повітряний зазор, що вимагає дуже високого ступеня дискретизації, в той же час на периферійних ділянках, де невеликий градієнт зміни поля, крок сітки бажано збільшити.

Таким чином, досвід розрахунку і наведені вище міркування говорять, що розрахунок поля у всьому поперечному перерізі машини з урахуванням насичення сердечників дуже часто не може забезпечити ні необхідної точності, ні прийнятних обчислювальних витрат. Тому доводиться виконувати розрахунки для обмеженої області, яка охоплює, наприклад, лише взаємодіючі зубці або головки зубців, зубцеві або полюсні поділи машини. Облік насичення решти областей може бути виконаний при розрахунку магнітного кола на основі методів теорії кіл.

9.3 Програмні засоби для розв'язання рівнянь математичної фізики

Найбільшого поширення для дослідження магнітних характеристик електричних машин у спеціалізованому програмному забезпеченні здобув метод кінцевих елементів. Для рішення зазначених задач з використанням МКЕ сьогодні широко використовуються універсальні програмні комплекси Ansys Maxwell, Comsol Multiphysics, які дозволяють вирішувати довільні комплекси міждисциплінарних технічних задач, що пов'язані з розв'язанням рівнянь різних типів (наприклад, пов'язані з розрахунки електромагнітних полів електричної машини та її нагрівання).

Існує ціла низка більш простих та вузько спеціалізованих програм для розв'язання рівнянь математичної фізики – ELCUT, JUMP, LOMAN, FEMM, FemLab та ін.

Розглянемо розрахунок магнітостатичних полів у пакеті Ansys Maxwell.

На рисунку 9.5 наведено приклад розрахунку поля простої електромагнітної системи, що включає магнітопровід із магнітом'якого матеріалу та котушки, що виконана з провідника. Моделювання виконано в програмному середовищі Ansys Maxwell. Картина поля зображена у вигляді векторів напруженості магнітостатичного поля.

За допомогою графічного редактора програма Ansys Maxwell дозволяє задати необхідні геометричні параметри об'єкту дослідження. В процесі моделювання задаються

основні параметри системи: матеріали, з яких складається система, крайові умови, набір джерел поля, та інші.

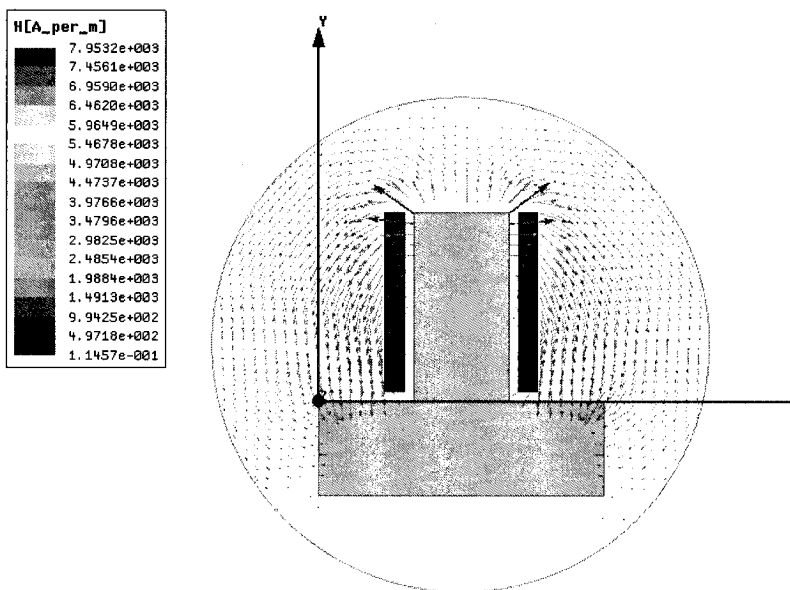


Рисунок 9.5 – Двовимірна модель електромагнітної системи в середовищі Ansys Maxwell

В результаті розрахунку визначається ряд параметрів поля: розподілення напруженості, індукції, векторного потенціалу, векторів намагніченості, а також розраховуються магнітний потік, потокозчеплення, індуктивність тощо.

Розрахунок поля можливо виконати також у тривимірній постановці задачі. На рисунку 9.6 зображена об'ємна картина поля тієї ж самої системи у вигляді векторів напруженості магнітного поля.

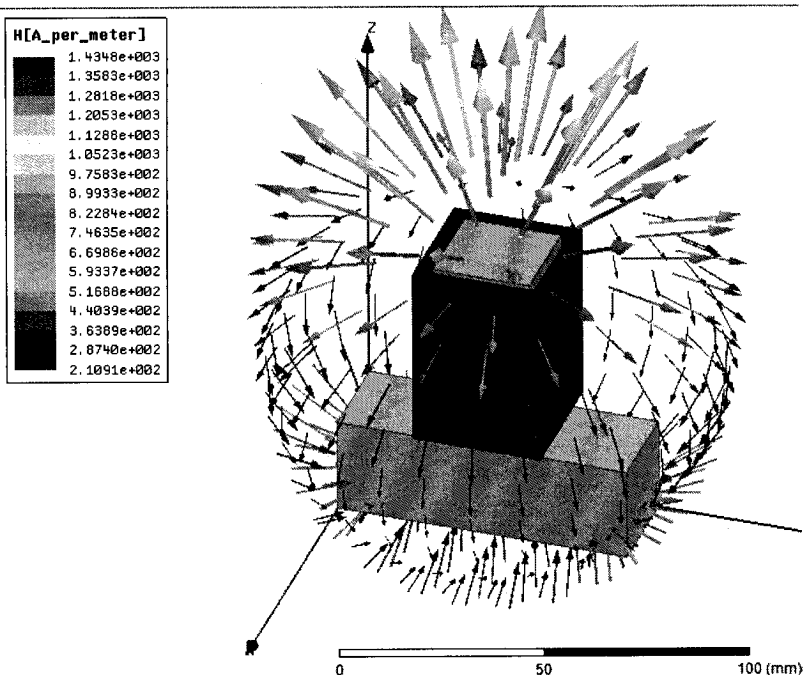


Рисунок 9.6 – Тривимірний розрахунок поля в середовищі Ansys Maxwell

Розглянемо аналіз двигуна постійного струму у програмі Ansys Maxwell. На основі технічних даних двигуна постійного струму з незалежним збудженням 2ПН-112-L, його геометричних розмірів та обмоточних даних, за допомогою програми RMXprt було розроблено геометричну модель двигуна постійного струму (рис. 9.7). На рис 9.8, 9.9 зображено результати розрахунку електромагнітних полів при роботі двигуна постійного струму.

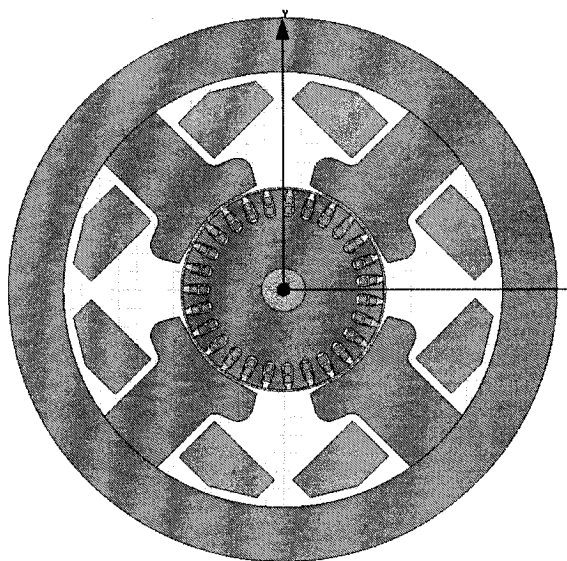


Рисунок 9.7 – Двовимірний модель двигуна 2ПН-112-L

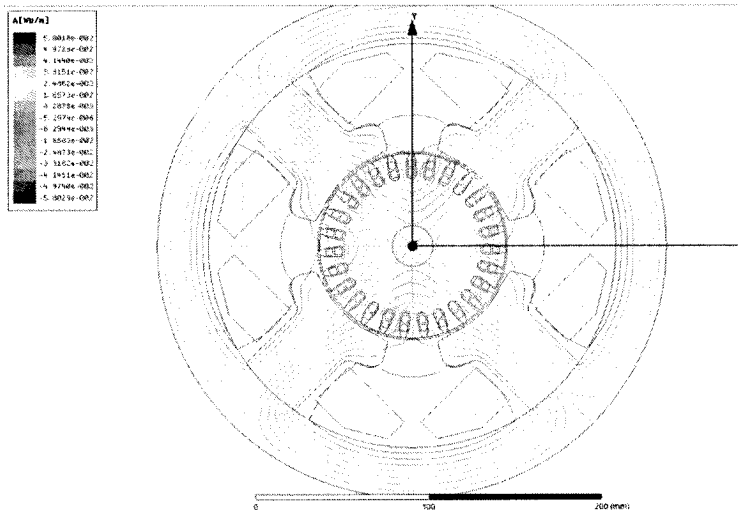


Рисунок 9.8 – Розподілення силових ліній магнітного поля в моделі двигуна

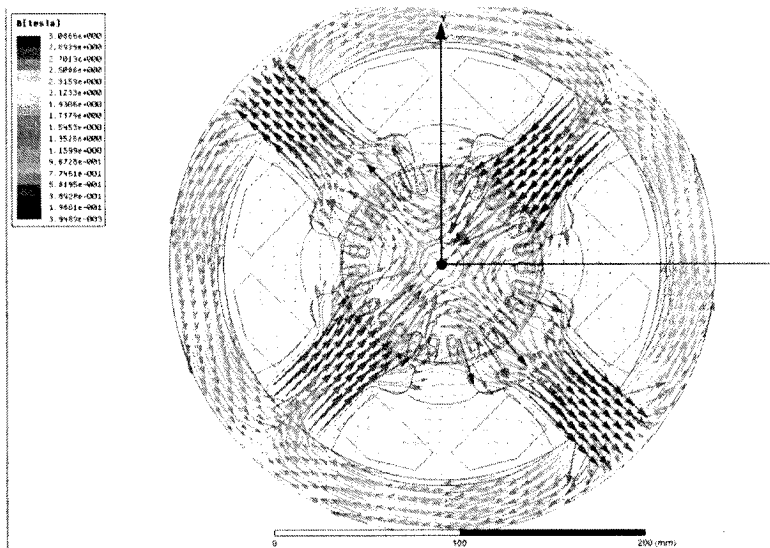


Рисунок 9.9 – Розподілення векторів магнітної індукції

На рис. 9.10 зображено тривимірну модель двигуна 2ПН-112-Л.

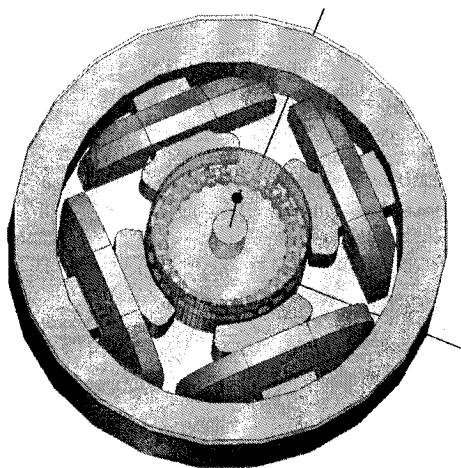


Рисунок 9.10 – Тривимірна модель двигуна 2ПН-112-Л

Питання для самоконтролю

1. Які задачі відносяться до математичної фізики?
2. Чим відрізняються стаціонарна та нестаціонарна задачі?
3. Дайте визначення векторних операторів набла, ротор, дивергенція та градієнт?
4. Що таке рівняння параболічного типу?
5. Які фізичні процеси описують рівняння гіперболічного типу?
6. Які фізичні процеси описують рівняння еліптичного типу?
7. Чим обумовлені початкові та граничні умови задачі?
8. Які існують різновиди граничних умов?
9. Сформулюйте крайову задачу (умови) Діріхле.
10. Сформулюйте крайову задачу (умови) Неймана.
11. Назвіть аналітичні та чисельні методи розв'язку рівнянь матфізики?
12. У чому полягає перевага чисельних методів над аналітичними?
13. У чому полягає особливість методу кінцевих різниць?
14. Опишіть алгоритм розрахунку при застосуванні методу кінцевих елементів?
15. Що таке тріангуляція та як вона використовується в чисельних методах?
16. Назвіть основні переваги методу кінцевих елементів?
17. Які програмні пакети базуються на методі кінцевих елементів?

Предметний покажчик

| | | | |
|---|----------|---|----------|
| абсолютна похибка | 22 | ітераційні методи розв'язання СЛАР48 | |
| адаптивні алгоритми | 162, 185 | квадратна матриця | 36 |
| аналітичні методи розв'язання ЗДР | 176 | квадратурні формули | 159 |
| апроксимація | 90 | конформне перетворення Шварца-Крістоффеля | 209 |
| апроксимація на сітці | 177 | корінь рівняння | 62 |
| асимптотичні методи | 177 | крайова задача | 175 |
| вибір головного елемента стовпця | 41 | крайові умови | 175 |
| випадкові помилки | 107 | крок сітки | 212 |
| відділення коренів | 64 | кроки інтегрування | 154 |
| відносна похибка | 22 | лінійні рівняння | 62 |
| вірна значуща цифра | 23 | локальна інтерполяція | 92 |
| власні значення | 176 | локальна кусково-квадратична інтерполяція | 96 |
| вузли інтегрування | 154 | локальна кусково-лінійна інтерполяція | 94 |
| вузли інтерполяції | 92 | локальна кусково-постійна інтерполяція | 93 |
| вузли сітки | 212 | метод бджолиного рою | 142 |
| глобальна інтерполяція | 92 | метод Гауса | 40 |
| глобальна найкраща позиція | 143 | метод градієнтів | 139 |
| градієнт | 204 | метод деформованого багатогранника | 139 |
| гранична абсолютна похибка | 22 | метод дихотомії | 69 |
| граничні умови задачі | 206 | метод дотичних | 75 |
| граничні умови третього роду | 208 | метод Ейлера | 179 |
| грубі помилки | 107 | метод зворотного змінного кроку | 134 |
| двовимірна задача | 202 | метод зворотної матриці | 38 |
| дивергенція | 203 | метод Зейделя | 52 |
| діагональна матриця | 37 | метод квадратичної апроксимації | 135 |
| евклідова норма матриці | 50 | метод кінцевих елементів | 215 |
| екстраполяційна формула Адамса-Башфорта | 183 | метод кінцевих різниць | 177, 209 |
| емпірична формула | 108 | метод конформних перетворень | 208 |
| задача Коші | 175 | метод Крамера | 43 |
| задача на власні значення | 175 | метод ламаних | 179 |
| збіжний ітераційний процес | 48 | метод Мілна | 184 |
| звичайні диференціальні рівняння | 174 | метод Монте-Карло | 163 |
| зворотна матриця | 38 | метод найближчого сусіда | 93 |
| зворотний хід метода Гауса | 40 | метод найменших квадратів | 109 |
| значуща цифра | 22 | метод найшвидшого спуску | 139 |
| інтегральна крива | 174 | метод Нелдера-Міда | 139 |
| інтегральне наближення | 90 | метод Ньютона-Рафсона | 74 |
| інтерполяційна формула Адамса-Мултона | 183 | метод Пауелла | 135 |
| інтерполяційний поліном | 92 | метод поділу напіл | 69 |
| інтерполяційний поліном Лагранжа | 98 | метод покоординатного спуску | 137 |
| інтерполяційний поліном Ньютона | 101 | | |
| інтерполяція | 91 | | |
| ітераційні методи | 48 | | |

| | | | |
|---|--------|--|-----|
| метод простих ітерацій..... | 50, 77 | похибка параметрів | 20 |
| метод Рунге-Кутти | 180 | початкові умови задачі..... | 206 |
| метод Сімпсона | 158 | правило Крамера..... | 37 |
| метод трапецій..... | 157 | пряма задача теорії похибок..... | 25 |
| метод хорд..... | 71 | прямий хід метода Гауса..... | 40 |
| методи Адамса..... | 182 | прямі методи розв'язання СЛАР | 38 |
| методи Монте-Карло | 153 | рівняння в часткових похідних | 174 |
| методи найвищої алгебраїчної точності | 153 | рівняння гіперболічного типу | 205 |
| методи Ньютона-Котеса..... | 153 | рівняння еліптичного типу | 205 |
| методи теорії збурень | 177 | рівняння математичної фізики | 201 |
| модифікований метод Ейлера..... | 179 | рівняння параболічного типу | 204 |
| наближені методи розв'язання ЗДР | 176 | різницева сітка | 210 |
| нелінійне алгебраїчне рівняння | 63 | різницева схема..... | 211 |
| нелінійне трансцендентне рівняння .. | 63 | розбіжний ітераційний процес..... | 48 |
| нелінійні рівняння..... | 62 | розширена матриця | 37 |
| нескінченна норма матриці | 50 | ротор..... | 203 |
| нестационарна задача | 202 | систематичні помилки..... | 106 |
| неусувна похибка | 20 | системи лінійних алгебраїчних рівнянь | 35 |
| неявні методи розв'язання ЗДР | 178 | сіткова функція | 210 |
| норма матриці | 49 | сплайнові методи | 153 |
| нормалізований запис числа | 23 | стаціонарна задача..... | 202 |
| нормальна система нелінійних рівнянь..... | 111 | сумісна СЛАР..... | 37 |
| нульова матриця..... | 37 | теорема Кронекера-Капеллі..... | 37 |
| обернена задача теорії похибок..... | 25 | Т-інтегрування | 160 |
| одинична матриця | 37 | точкове наближення | 90 |
| одинична норма матриці | 49 | транспонована союзна матриця | 38 |
| одновимірна задача..... | 202 | тривимірна задача..... | 202 |
| однорідна СЛАР..... | 37 | тригонометричний ряд Фур'є | 112 |
| оператор «набла» | 202 | умова Ліпшиця | 139 |
| оператор Лапласа | 203 | умови Діріхле | 207 |
| параболоїд..... | 136 | умови збіжності ітераційного процесу..... | 49 |
| персональна найкраща позиція | 143 | умови Неймана..... | 207 |
| перша теорема Больцано-Коші..... | 64 | уточнення коренів..... | 64 |
| похибка задачі | 20 | формула Ньютона-Лейбніца..... | 152 |
| похибка математичної моделі..... | 20 | формули Беселя | 113 |
| похибка методів лівих і правих прямокутників | 159 | формули Ейлера-Фур'є | 113 |
| похибка методу | 21 | функція Химмельблау..... | 136 |
| похибка методу середніх прямокутників | 160 | чисельне диференціювання | 147 |
| похибка методу Сімпсона | 160 | чисельні методи | 17 |
| похибка методу трапецій | 160 | чисельні методи розв'язання ЗДР .. | 177 |
| похибка округлень | 21 | швидке перетворення Фур'є..... | 115 |
| | | явні методи розв'язання ЗДР..... | 178 |

Література

1. Driscoll T. A. Schwarz-Christoffel mapping, Cambridge Monogr. Appl. Comput. Math. 8 / T. A. Driscoll, L. N. Trefethen. – Cambridge : Cambridge Univ. Press, 2002. – 132 p.
2. Mathematical model [Електронний ресурс] // Стаття з офіційного сайта Encyclopedia Britannica. – Режим доступу : <http://www.britannica.com>
3. Szabo B. Finite Element Analysis / B. Szabo, I. Babushka. – New York : John Wiley and Sons, 1991. – 368 p.
4. T-integration. [Електронний ресурс]. – Режим доступу: <http://en.wikipedia.org/w/index.php?title=T-integration&oldid=402346593>
5. Алберг Дж. Теория сплайнов и ее приложения / Дж. Алберг, Э. Нильсон, Дж. Уолш. – М. : Мир. 1972. – 318 с.
6. Амосов А. А. Вычислительные методы для инженеров : учебное пособие / А. А. Амосов, Ю. А. Дубинский, Н. В. Копченова. – М. : Высшая Школа, 1994. – 544 с.
7. Бахвалов Н. С. Численные методы / Н. С. Бахвалов, Н. П. Жидков, Г. М. Кобельков. – М. : Наука, 1987. – 600 с.
8. Бахвалов Н. С. Численные методы в задачах и упражнениях / Н. С. Бахвалов, А. В. Лапин, Е. В. Чижонков. – М. : Высш. шк., 2000. – 192 с.
9. Березин И. С. Методы вычислений / И. С. Березин, Н. П. Жидков. – М. : Наука, 1966. – Т. 1. – 632 с.
10. Бидасюк Ю. М. Mathsoft MathCAD 12. Самоучитель / Ю. М. Бидасюк. – М. : Вильямс, 2006. – 224 с.
11. Боглаев Ю. П. Вычислительная математика и программирование / Ю. П. Боглаев. – М. : Высшая Школа, 1990. – 546 с.
12. Вержбицкий В. М. Численные методы (математический анализ и обыкновенные дифференциальные уравнения) : учеб. пособие для вузов / В. М. Вержбицкий. – М. : ООО «Издательский дом «ОНИКС 21 век», 2005. – 400 с.
13. Вержбицкий В. М. Численные методы. Математический анализ и обыкновенные дифференциальные уравнения / В. М. Вержбицкий. – М. : Высш.шк., 2001. – 383 с.
14. Волков Е. А. Численные методы / Е. А. Волков. – СПб. : Лань, 2004. – 248 с.
15. Воробьева Г. Н. Практикум по вычислительной математике / Г. Н. Воробьева, А. Н. Данилова. – М. : Высшая Школа, 1990. – 207 с.
16. Глушков В. М. Энциклопедія кібернетики. Т1 / В. М. Глушков, М. М. Амосов, І. П. Артеменко, О. О. Бакаев и др. ; за

ред. В. М. Глушкова. – К. : Головна редакція УРЕ АН УРСР, 1973. – 596 с.

17. Глушков В. М. Энциклопедія кібернетики. Т2 / В. М. Глушков, М. М. Амосов, І. П. Артеменко, О. О. Бакаєв и др.; за ред. В. М. Глушкова. – К. : Головна редакція УРЕ АН УРСР, 1973. – 590 с.

18. Голуб Дж. Матричные вычисления / Дж. Голуб, Ч. Ван Лоун. – М. : Мир, 1999. – 548 с.

19. Голубинский А. Н. Методы аппроксимации экспериментальных данных и построения моделей [Электронный ресурс] / А. Н. Голубинский // Вестник Воронежского института МВД России. – Воронеж, 2007. – Вып. № 2. – Режим доступа: <http://cyberleninka.ru/article/n/metody-approksimatsii-eksperimentalnyh-dannyh-i-postroeniya-modeley#ixzz4NSSbAJCb>

20. Демидович Б. П. Основы вычислительной математики / Б. П. Демидович, И. А. Марон. – М. : Наука, 1970. – 664 с.

21. Деммель Дж. Вычислительная линейная алгебра / Дж. Деммель. – М. : Мир, 2001. – 430 с.

22. Дьяконов В. П. Справочник по алгоритмам и программам на языке бейсик для персональных ЭВМ / В. П. Дьяконов. – М. : Наука, 1987. – 240 с.

23. Жалдак М. І. Основи теорії і методів оптимізації : навчальний посібник / М. І. Жалдак, Ю. В. Триус. – Черкаси : Брама-Україна, 2005. – 608 с.

24. Заварыкин В. М. Численные методы : учеб. пособие для студентов физ.-мат. спец. пед. ин-тов / В. М. Заварыкин, В. Г. Житомирский, М. П. Лапчик. – М. : Просвещение, 1990. – 176 с.

25. Загірняк М. В. Електромагнітні розрахунки : підручник / М. В. Загірняк. – Харків: «Друкарня Мадрид», 2015. – 320 с.

26. Зверев В. С. Mathcad. Математическая компьютерная поддержка проектирования : учеб. пос. для инженерно-технических специальностей / В. С. Зверев. – Астрахань : Изд-во АГТУ, 1996. – 160 с.

27. Измаилов А. Ф. Численные методы оптимизации / А. Ф. Измаилов, М. В. Солодов. – М. : Физматлит, 2005. – 304 с.

28. Калиткин Н. Н. Численные методы / Н. Н. Калиткин. – М. : Наука, 1978. – 512 с.

29. Корнейчук Н. П. Экстремальные свойства полиномов и сплайнов / Н. П. Корнейчук, В. Ф. Бабенко, А. А. Лигун ; отв. ред. А. И. Степанец ; ред. С. Д. Кошис, О. Д. Мельник. – К. : Наукова думка, 1992. – 304 с.

30. Коробицын В. В. Исследование поведения явных методов Рунге-Кутты при решении систем обыкновенных дифференциальных уравнений с разрывной правой частью / В. В. Коробицын, В. Б. Маренич, Ю. В. Фролова // Математические структуры и моделирование. – 2007. – Вып. 17. – С. 19–25.
31. Крылов В. И. Вычислительные методы / В. И. Крылов, В. В. Бобков, П. И. Монастырный. – М. : Наука, 1976. – Т.1. – 302 с.
32. Мала гірнича енциклопедія. Т1 ; за ред. В. С. Білецького. – Донецьк : Донбас, 2004. – 640 с.
33. Мала гірнича енциклопедія. Т2 ; за ред. В. С. Білецького. – Донецьк : Донбас, 2007. – 652 с.
34. Мала гірнича енциклопедія. Т3 ; за ред. В. С. Білецького. – Донецьк : Східний видавничий дім, 2013. – 644 с.
35. Мудров А. Е. Численные методы для ПЭВМ на языках Бейсик, Фортран и Паскаль / А. Е. Мудров. – Томск : МП "РАСКО", 1991. – 272 с.
36. Мэтьюз Д. Численные методы. Использование MATLAB / Д. Мэтьюз, К. Финк. – М. : Вильямс, 2001. – 720 с.
37. Норри Д. Введение в метод конечных элементов / Д. Норри, Ж. де Фриз. – М. : Мир, 1981. – 304 с.
38. Ортега Дж. Введение в численные методы решения дифференциальных уравнений / Дж. Ортега, У. Пул. – М. : Наука. Гл. ред. физ.-мат. лит., 1986. – 288 с.
39. Охорзин В. А. Прикладная математика в системе MATHCAD : учебное пособие / В. А. Охорзин. – СПб. : Лань, 2009. – 352 с.
40. Плис А. И. Лабораторный практикум по высшей математике / А. И. Плис, Н. А. Сливина. – М. : Высш. шк., 1994. – 416 с.
41. Полак Э. Численные методы оптимизации. Единый поход / Э. Полак. – М. : Мир, 1994. – 374 с.
42. Самарский А. А. Математическое моделирование. Идеи. Методы. Примеры / А. А. Самарский, А. П. Михайлов. – М. : Физматлит, 2001. – 320 с.
43. Сегерлинд Л. Применение метода конечных элементов / Л. Сегерлинд. – М. : Мир, 1979. – 392 с.
44. Семененко М. Г. Математическое моделирование в MathCad / М. Г. Семененко. – М. : Альтекс-А, 2003. – 206 с.
45. Сквиорцов Л. М. Адаптивные методы численного интегрирования в задачах моделирования динамических систем / Л. М. Сквиорцов // Изв. РАН. Теория и системы управления. – 1999. – № 4. – С. 72–78.

46. Скворцов Л. М. Явные адаптивные методы численного решения жестких систем / Л. М. Скворцов // Математическое моделирование. – 2000. – № 2 (12). – С. 97–107.
47. Стренг Г. Теория метода конечных элементов / Г. Стренг, Д. Фикс. – М. : Мир, 1977. – 349 с.
48. Сурмин Ю. П. Теория систем и системный анализ : учебное пособие / Ю. П. Сурмин. – К. : МАУП, 2003. – 368 с.
49. Сухарев А. Г. Курс методов оптимизации : учебное пособие / А. Г. Сухарев, А. В. Тимохов, В. В. Федоров. – М. : Физматлит, 2005. – 367 с.
50. Тарг С. М. Моделирование физическое / С. М. Тарг, С. Л. Вишневецкий, В. А. Арутюнов // Большая советская энциклопедия.
51. Турчак Л. И. Основы численных методов / Л. И. Турчак. – М. : Наука, 1987. – 320 с.
52. Тюканов А. С. Основы численных методов : учебное пособие [Электронный ресурс] / А. С. Тюканов. – Российский государственный педагогический университет им. А. И. Герцена. – Режим доступа: http://physics.herzen.spb.ru/library/01/01/nm_labs/index.htm
53. Уайлд Д. Дж. Методы поиска оптимума / Д. Дж. Уайлд. – М. : Наука, 1997. – 268 с.
54. Форсайт Дж. Численное решение систем линейных алгебраических уравнений / Дж. Форсайт, К. Молер. – М. : Мир, 1969. – 167 с.
55. Фриск В. В. Mathcad. Расчеты и моделирование цепей на ПК / В. В. Фриск. – М. : Солон-Пресс, 2006. – 88 с.
56. Хайрер Э. Решение обыкновенных дифференциальных уравнений. Жесткие и дифференциально-алгебраические задачи / Э. Хайрер, Г. Ваннер. – М. : Мир, 1999. – 685 с.
57. Хайрер Э. Решение обыкновенных дифференциальных уравнений. Нежесткие задачи / Э. Хайрер, С. Нёрсетт, Г. Ваннер. – М. : Мир, 1990. – 512 с.
58. Ханова А. А. Численное решение уравнений и систем уравнений : учебное пособие / А. А. Ханова. – Астрахань : Астраханский государственный технический университет, 2001. – 45 с.
59. Чорний О. П. Моделювання електромеханічних систем : підручник / О. П. Чорний, А. В. Луговой, Д. Й. Родькін і др. – Кременчук, 2001. – 410 с.
60. Шуп Т. Е. Прикладные численные методы в физике и технике/ Терри Е. Шуп. – М. : Высшая Школа, 1990. – 254 с.