

ДЕРЖАВНА ПОДАТКОВА АДМІНІСТРАЦІЯ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ДЕРЖАВНОЇ ПОДАТКОВОЇ
СЛУЖБИ УКРАЇНИ

Кафедра інтелектуальних систем прийняття рішень

“До друку”

Проректор з навчальної та методичної
роботи

М.М. Касьяненко

« _____ » _____ 2009 р.

Навчально-методичний комплекс дисципліни

«Чисельні методи в інформатиці»

для підготовки бакалаврів денної форми навчання
за напрямом 0804 «Комп'ютерні науки»
спеціальності 6.080400 «Інтелектуальні системи прийняття рішень»
статус дисципліни: нормативна

Ірпінь 2009

Навчально-методичний комплекс дисципліни «Чисельні методи в інформатиці» включає: робочу навчальну програму, опорний конспект лекцій, методичні рекомендації до проведення практичних занять, методичні вказівки до проведення індивідуальних та самостійних робіт, контрольні питання з дисципліни, літературу.

Автор: П.Ф. Жук, д. ф.-м. наук, доцент, професор

Розглянуто і схвалено на засіданні кафедри інтелектуальних систем прийняття рішень, протокол № __ від “__” _____ 2009 р.

Завідувач кафедри _____ С.П. Ріппа, д.е.н., професор

Розглянуто і схвалено на засіданні вченої ради факультету економіки та оподаткування, протокол № __ від “__” _____ 2009 р.

Голова вченої ради факультету економіки та оподаткування _____ Г.М. Калач, к.е.н., доцент

Завідувач навчально-методичного відділу _____ О.О. Бойко

Реєстраційний № ____

Зміст

Передмова	4
Робоча навчальна програма	5
Передмова	6
Опис навчальної дисципліни «Чисельні методи в інформатиці»	8
Структура навчальної дисципліни «Чисельні методи в інформатиці»	9
Зміст навчальної дисципліни за модулями	11
Модуль 1	11
Змістовний модуль 1 (теми 1–7). Чисельні методи алгебри і теорії наближень	11
Модуль 2	16
Змістовний модуль 2 (теми 8–11). Чисельні методи аналізу і математичної фізики	16
Методи і форми проміжного та підсумкового контролю	22
Розподіл балів при рейтинговій системі	23
Опорний конспект лекцій з курсу	24
Методичні вказівки до проведення практичних занять з дисципліни	156
Карта практичних занять з навчальної дисципліни «Чисельні методи в інформатиці»	158
Методичні вказівки до організації самостійної та індивідуальної роботи студентів з дисципліни	213
Вступ	214
Теоретичні основи	216
Контрольні питання з дисципліни	240
Завдання до самостійної та індивідуальної роботи	246
Документація ПМК	253
Критерії оцінки знань	254
Модульні контрольні роботи з курсу	255
Комплексна контрольна робота для підсумкового контролю знань з курсу	258
Література	273

Передмова

На сучасному етапі широкого застосування комп'ютерних методів в науково-практичних дослідженнях одним з основних методів дослідження складних явищ є метод математичного моделювання. При побудові математичних моделей складних, зокрема соціально-економічних систем, часто виникають задачі, для яких знаходження аналітичного розв'язку або надто ускладнене, або неможливе.

Ефективність розв'язання таких моделей безпосередньо пов'язана з вибором того чи іншого алгоритму, який природно має свою особливу область застосування. Тому важливою задачею є вміння вибирати та застосовувати найбільш реальну стратегію чисельного розв'язання задач.

Галузь математики, яка займається розробкою спеціальних математичних методів та алгоритмів, що дозволяють одержати чисельний розв'язок та оцінити його точність, називається чисельні методи.

Навчальна дисципліна «Чисельні методи в інформатиці» – обов'язковий компонент загальної та професійної освіти. Метою викладання цієї дисципліни є підготовка майбутніх бакалаврів спеціальності «Інтелектуальні системи прийняття рішень» до розробки та використання комп'ютерно-орієнтованих обчислювальних алгоритмів розв'язання завдань, що виникають при математичному моделюванні реального світу.

При викладанні навчальної дисципліни «Чисельні методи в інформатиці» ставляться наступні завдання:

- навчити студентів основним чисельним методам алгебри, аналізу, диференціальних рівнянь і математичної фізики;
- дати студентам уявлення про теорію різницевих схем;
- прищепити студентам стійкі навички математичного моделювання з використанням комп'ютерів, надати досвіду проведення обчислювальних експериментів;
- прищепити студентам вміння самостійно вивчати навчальну і наукову літературу в галузі обчислювальної математики.

Теоретичним фундаментом дисципліни є вища та дискретна математика, інформатика, теорія ймовірностей, теорія програмування та алгоритмічні мови. Практичним засобом реалізації чисельних методів є сучасна комп'ютерна техніка та прикладне програмне забезпечення.

У результаті вивчення дисципліни студент повинен одержати фундаментальні теоретичні знання у галузі чисельних методів і закріпити їх на практичних заняттях.

Загальний обсяг навчальної дисципліни - 135 годин, з них лекції: 46 годин, практичні: 22 години, самостійна робота: 53 годин, індивідуальні заняття: 14 годин.

РОБОЧА НАВЧАЛЬНА ПРОГРАМА З КУРСУ
“ЧИСЕЛЬНІ МЕТОДИ В ІНФОРМАТИЦІ”

Передмова

При побудові математичних моделей складних, зокрема соціально-економічних систем, часто виникають задачі, для яких знаходження аналітичного розв'язку або надто ускладнене, або неможливе. Це обумовлює необхідність розробки спеціальних математичних методів та алгоритмів, що дозволяють одержати чисельний розв'язок та оцінити його точність. Ця галузь математики одержала назву "Чисельні методи".

Метою дисципліни є засвоєння теоретичних основ, формування у студентів практичних навичок щодо використання основних методів чисельного розв'язання складних математичних моделей.

Завдання курсу полягає у вивченні теоретичних відомостей та набуття студентами практичних навичок чисельного розв'язання задач великої розмірності, систем алгебраїчних та диференціальних рівнянь, нелінійних математичних моделей тощо, опануванні сучасними пакетами прикладних програм, що дозволяють здійснювати чисельні розрахунки.

Предмет навчальної дисципліни - математичний інструментарій, що дозволяє знаходити чисельні розв'язки складних математичних моделей.

Практичним засобом реалізації чисельних методів є сучасна комп'ютерна техніка та прикладне програмне забезпечення.

У результаті вивчення дисципліни студенти повинні:

знати:

- теоретичні основи чисельного розв'язання складних математичних моделей;
- одержати практичні навички побудови алгоритмів чисельного розв'язання прикладних задач та їх реалізації на персональних комп'ютерах;
- вміти:
- використовувати сучасне програмне забезпечення, що дозволяє здійснювати чисельні розрахунки.

Міжпредметні зв'язки навчальної дисципліни.

Вивчення навчальної дисципліни "Чисельні методи в інформатиці" проводиться після таких дисциплін: «Основи дискретної математики», «Вища математика», «Теорія ймовірностей, імовірнісні процеси і математична статистика», «Основи програмування та алгоритмічні мови».

Методи і форми викладання дисципліни.

Вивчення дисципліни передбачає лекційні та лабораторні заняття. Значна частина матеріалу дисципліни відведена під індивідуальні заняття студентів під керівництвом викладача та СРС.

Методики активізації процесу навчання.

Навчальні технології, що застосовуються для активізації навчально-пізнавальної діяльності студентів наступні:

Проблемні лекції – застосовуються при викладанні основного лекційного матеріалу.

Робота в малих групах – основний метод активізації роботи студентів.

Кейс-метод (метод аналізу конкретних ситуацій) – застосовується для наближення процесу навчання до реальної лабораторної діяльності

спеціалістів і передбачає розгляд конкретних виробничих, управлінських ситуацій у процесі вивчення і виконання навчальних завдань.

Презентації – виступи перед аудиторією – використовуються для представлення результатів роботи в малих групах на практичних заняттях, звітів про виконання індивідуальних завдань студентів.

Головна мета самостійної роботи — поглиблення знань і одержання практичних навичок з основних питань курсу “Чисельні методи в інформатиці” шляхом самостійної роботи з літературними джерелами та закріплення знань на практичних та лабораторних заняттях.

Студентам рекомендуються такі форми самостійної роботи:

- опрацювання лекційного матеріалу з використанням конспекту лекцій та рекомендованої літератури;
- самостійне вивчення окремо визначених тем та питань на основі навчальної літератури та методичних розробок кафедри;
- самостійне виконання індивідуальних завдань;
- підготовка та виконання контрольних робіт;
- самостійна підготовка до ПМК.

Індивідуальна робота здійснюється за графіком навчального процесу по темам у формі: індивідуальних занять, консультацій, перевірки індивідуальних завдань, перевірки та захисту завдань, що винесені на поточний контроль тощо.

Форми і засоби проміжного та підсумкового контролю: експрес-контроль рівня готовності студента до проведення практичних занять; перевірка виконання поза аудиторних завдань; оцінка роботи студента під час заняття (виступи, доповнення, участь у дискусії); виконання домашніх завдань; контрольні роботи в кінці залікового кредиту. Оцінка індивідуальних результатів здобуття знань студентами проводиться у формі заліку за кредитно-модульною методологією навчання, критерії якої визначаються у навчальній робочій програмі за стобальною системою, яка трансформується у стандартні залікові диференційовані оцінки відповідно до вимог Міністерства освіти та науки України.

Критерії оцінки успішності повинні відповідати навчальній програмі й найбільш важливим вимогам до знань студентів:

1. Знання фактів, явищ. Вірне, науково достовірне їх пояснення.
2. Оволодіння науковими термінами, поняттями, законами, методами, правилами; вміння користуватися ними при поясненні нових фактів, розв’язуванні різних питань і виконанні практичних завдань.
3. Максимальна ясність, точність думки, вміння відстоювати свої погляди, захищати їх.

Форма підсумкового контролю – ПМК.

Загальний обсяг навчальної дисципліни - 135 годин, з них лекції: 46 годин, практичні: 22 години, самостійна робота: 53 годин, індивідуальні заняття: 14 годин.

Опис навчальної дисципліни
“Чисельні методи в інформатиці”

за напрямом 0804 “Комп’ютерні науки”
спеціальності 6.080400 “Інтелектуальні системи прийняття рішень”

Предмет: математичний інструментарій, що дозволяє знаходити чисельні розв’язки складних математичних моделей.

Мета: засвоєння теоретичних основ, формування у студентів практичних навичок щодо використання основних методів чисельного розв’язання складних математичних моделей.

Змістово-модульна структура дисципліни Курс: 3 Семестр: 6	Напрямок, спеціальність, освітньо- кваліфікаційний рівень	Характеристика навчальної дисципліни
Кількість кредитів: Національних -2,5 ECTS -3,75 Модулів:2 Змістових модулів:2 Загальна кількість годин: 135 годин	Шифр та назва напряму 0804 "Комп’ютерні науки" Шифр та назва спеціальності: 6.080400 "Інтелектуальні системи прийняття рішень" Освітньо- кваліфікаційний рівень – бакалавр	Нормативна Рік підготовки: 3 Семестр: 6 Лекції: 46 годин Практичні: 22 години Самостійна робота: 53 годин Індивідуальні заняття: 14 годин Вид контролю: ПМК

Передумови вивчення: «Основи дискретної математики», «Вища математика», «Теорія ймовірностей, імовірнісні процеси і математична статистика», «Основи програмування та алгоритмічні мови».

Структура навчальної дисципліни
“Чисельні методи в інформатиці”

За напрямом підготовки 0804 “Комп’ютерні науки”

Спеціальності 6.080400 «Інтелектуальні системи прийняття рішень»

№ п/п	Змістові модулі	Кількість годин				СРС
		Всього	Лекції	Практичні	Індивідуальні заняття	
Модуль 1=1,75 ЗК						
Змістовний модуль 1. Чисельні методи алгебри і теорії наближень						
Т 1.	Вступ. Етапи розв’язання прикладних задач на ЕОМ	3	2			1
Т 2.	Сучасні пакети прикладних програм, призначених для чисельного розв’язання математичних моделей	6	2			4
Т 3.	Загальна характеристика чисельних методів	6	2			4
Т 4.	Чисельні методи розв’язування нелінійних рівнянь	10	4	2	2	2
Т 5.	Чисельні методи розв’язування систем лінійних алгебраїчних рівнянь	16	6	2	4	4
Т 6.	Інтерполювання функцій	12	4	4		4
Т 7.	Задача наближення функцій однієї змінної	10	4	2		4
Усього по заліковому кредиту		63	24	10	6	23
Форма контролю - контрольна робота						
Усього по модулю 1		63	24	10	6	23
Модуль 2= 2 ЗК						
Змістовний модуль 2. Чисельні методи аналізу і математичної фізики						
Т 8.	Чисельне диференціювання функцій	16	4	2	2	8
Т 9.	Чисельне інтегрування функції	20	6	2	2	10
Усього по заліковому кредиту		36	10	4	4	18
Форма контролю - контрольна робота						

Т 10.	Чисельне інтегрування звичайних диференціальних рівнянь	16	6	2	2	6
Т 11.	Чисельні методи математичної фізики	20	6	6	2	6
Усього по заліковому кредиту		36	12	8	4	12
Форма контролю-контрольна робота						
Усього по модулю - 2		72	22	12	8	30
Разом годин з курсу		135	46	22	14	53

Зміст навчальної дисципліни

Змістовний модуль 1. Чисельні методи алгебри і теорії наближень Тема 1. Вступ. Етапи розв'язання прикладних задач на ЕОМ

Вступ. Етапи розв'язання прикладних задач на ЕОМ. Поняття та приклади математичних моделей. Місце чисельних методів при розв'язуванні прикладних задач. Зв'язок чисельних методів з іншими дисциплінами інформаційного циклу.

Лекція 1. Вступ. Етапи розв'язання прикладних задач на ЕОМ (2 год.)

1. Вступ.
2. Етапи розв'язання прикладних задач на ЕОМ.
3. Поняття та приклади математичних моделей.
4. Місце чисельних методів при розв'язуванні прикладних задач.
5. Зв'язок чисельних методів з іншими дисциплінами інформаційного циклу.

Перелік питань до самостійної роботи

1. Застосування методів математичного моделювання в комп'ютерних науках.
2. Історія розвитку чисельних методів.
3. Основні етапи побудови математичної моделі. Кільце Самарського.

Література [3], [4], [5], [8], [10].

Тема 2. Сучасні пакети прикладних програм, призначені для чисельного розв'язання математичних моделей

Математичний пакет Maple та його основні можливості. Аналітичне та чисельне диференціювання та інтегрування в Maple. Розв'язання диференціальних рівнянь та побудова графіків розв'язків. Розв'язання задач лінійної алгебри за допомогою модуля LinearAlgebra.

Математичний пакет Mathematica та його можливості. Інтерфейс та основні модулі пакету. Виконання чисельних розрахунків в пакеті Mathematica.

Математичний пакет MATLAB. Інтерфейс та основні модулі пакету. Робота з матрицями в пакеті MatLab. Використання спеціальних функцій пакету MatLab для чисельного розв'язання алгебраїчних та диференціальних рівнянь.

Лекція 2. Сучасні пакети прикладних програм, призначених для чисельного розв'язання математичних моделей (2 год.)

1. Математичний пакет Maple та його основні можливості.
2. Аналітичне та чисельне диференціювання та інтегрування в Maple.
3. Математичний пакет Mathematica та його можливості. Виконання чисельних розрахунків в пакеті Mathematica.

4. Математичний пакет MATLAB.
5. Інтерфейс та основні модулі пакету.
6. Робота з матрицями в пакеті MatLab.
7. Використання спеціальних функцій пакету MatLab для чисельного розв'язання алгебраїчних та диференціальних рівнянь.

Перелік питань до самостійної роботи

1. Математичний пакет MATLAB.
2. Інтерфейс та основні модулі пакету.
3. Робота з матрицями в пакеті MatLab.
4. Використання спеціальних функцій пакету MatLab для чисельного розв'язання алгебраїчних та диференціальних рівнянь.

Література [9], [10], [14].

Тема 3. Загальна характеристика чисельних методів

Загальна характеристика чисельних методів. Поняття некоректно поставленої задачі. Стійкість чисельних алгоритмів. Типи похибок, які виникають при розв'язуванні прикладної задачі.

Лекція 3. Загальна характеристика чисельних методів(2 год.)

1. Загальна характеристика чисельних методів.
2. Поняття некоректно поставленої задачі.
3. Стійкість чисельних алгоритмів.
4. Типи похибок, які виникають при розв'язуванні прикладної задачі.

Перелік питань до самостійної роботи

1. Приклади некоректно поставлених задач.
2. Приклади нестійких чисельних алгоритмів.
3. Безпомилкові обчислення.

Література [4] - [8], [11].

Тема 4. Чисельні методи розв'язування нелінійних рівнянь

Чисельні методи розв'язування нелінійних рівнянь. Алгоритми ізолювання та уточнення коренів. Графічний метод та метод табулювання. Метод ділення навпіл. Блок схема, оцінка кількості ітерацій. Метод хорд та дотичних.

Лекція 4. Чисельні методи розв'язування нелінійних рівнянь (4 год.)

1. Чисельні методи розв'язування нелінійних рівнянь.
2. Алгоритми ізолювання та уточнення коренів.
3. Графічний метод та метод табулювання.
4. Метод ділення навпіл.
5. Блок схема, оцінка кількості ітерацій.
6. Метод хорд та дотичних.

Практичне заняття 1. Чисельні методи розв'язування нелінійних рівнянь (2 год.)

1. Алгоритми ізолювання коренів.
2. Особливості застосування графічного методу та методу табулювання.
3. Алгоритми уточнення коренів.
4. Приклади застосування методів ділення навпіл, хорд та дотичних, ітерацій.

Перелік питань до самостійної роботи

1. Розв'язування нелінійних рівнянь однієї змінної.
2. Відокремлення коренів за допомогою графічного методу та метод табулювання. Знаходження коренів за допомогою метода ділення навпіл.
3. Розв'язування нелінійних рівнянь однієї змінної за допомогою методу ітерацій, хорд та дотичних.

Індивідуальні заняття

1. Порівняння методів хорд та дотичних для розв'язування нелінійних алгебраїчних рівнянь. Суть комбінованого методу хорд та дотичних.
2. Суть методу простої ітерації для розв'язування нелінійних алгебраїчних рівнянь. Переваги та недоліки методу.

Література [3], [4], [6], [11], [13], [14].

Тема 5. Чисельні методи розв'язування систем лінійних алгебраїчних рівнянь

Чисельні методи розв'язування систем лінійних алгебраїчних рівнянь. Метод Гауса. Прямий та обернений хід методу Гауса. Вибір головного елемента. Оцінка кількості арифметичних дій методу Гауса. Блок-схема методу Гауса.

Метод ітерацій розв'язування систем лінійних алгебраїчних рівнянь. Достатня умова збіжності методу ітерацій. Метод Гауса-Зейделя. Блок-схема методу Гауса-Зейделя. Перетворення СЛАУ до вигляду, зручному для ітерування.

Лекція 5. Чисельні методи розв'язування систем лінійних алгебраїчних рівнянь (4 год.)

1. Чисельні методи розв'язування систем лінійних алгебраїчних рівнянь.
2. Метод Гауса.
3. Прямий та обернений хід методу Гауса.
4. Вибір головного елемента.
5. Оцінка кількості арифметичних дій методу Гауса.
6. Блок-схема методу Гауса.

Лекція 6. Метод ітерацій розв'язування систем лінійних алгебраїчних рівнянь (2 год.)

1. Метод ітерацій розв'язування систем лінійних алгебраїчних рівнянь.
2. Достатня умова збіжності методу ітерацій.
3. Метод Гауса-Зейделя.

4. Блок-схема методу Гауса-Зейделя.
5. Перетворення СЛАУ до вигляду, зручному для ітерування.

Практичне заняття 2. Чисельні методи розв'язування систем лінійних алгебраїчних рівнянь (2 год.)

1. Приклади математичних моделей, що містять системи лінійних алгебраїчних рівнянь.
2. Застосування методу Гауса для розв'язування модельних лінійних систем третього порядку. Приклади розв'язування систем лінійних алгебраїчних рівнянь методом ітерацій.
3. Розв'язування системи лінійних алгебраїчних рівнянь третього порядку за допомогою методу Гауса, методу оберненої матриці (знаходження оберненої матриці шляхом розбиття на клітини), методу простої ітерації та ітераційним методом Зейделя.

Перелік питань до самостійної роботи

1. Розв'язання систем лінійних алгебраїчних рівнянь з симетричними матрицями. Метод квадратних коренів.
2. Розв'язання систем нелінійних алгебраїчних рівнянь. Метод Ньютона. Метод найшвидшого спуску.
3. Чисельні методи розв'язання нелінійних рівнянь в моделях економічної динаміки.

Індивідуальні заняття

1. Поняття матриці системи лінійних алгебраїчних рівнянь та оберненої матриці.
 2. Порівняння точних та ітераційних методів (їх переваги та недоліки).
 3. Суть методу Гауса з виділенням головного члену.
 4. Порівняння методу простої ітерації та методу Гауса-Зейделя для розв'язування системи лінійних алгебраїчних рівнянь.
 5. Порівняння методів хорд та дотичних для розв'язування нелінійних алгебраїчних рівнянь. Суть комбінованого методу хорд та дотичних.
 6. Суть методу простої ітерації для розв'язування нелінійних алгебраїчних рівнянь. Переваги та недоліки методу.
 7. Використання формул методу розбиття на клітини для знаходження оберненої матриці.
 8. Використання методу Гауса з виділенням головного члену.
- Література [1], [4], [6], [7], [9], [11], [13].

Тема 6. Інтерполювання функцій

Інтерполювання функцій однієї змінної. Теорема про існування інтерполяційного многочлену. Многочлен Лагранжа. Інтерполювання за рівномірною сіткою вузлів. Многочлен Ньютона. Поняття скінчених різниць. Побудова многочлену Ньютона за таблицею скінчених різниць. Поняття сплайн-інтерполяції. Складання системи лінійних рівнянь для визначення

коефіцієнтів сплайн-інтерполяції. Алгоритм прогонки для знаходження коефіцієнтів сплайн-інтерполяції.

Лекція 7. Інтерполювання функцій (2 год.)

1. Інтерполювання функцій однієї змінної.
2. Теорема про існування інтерполяційного многочлену.
3. Многочлен Лагранжа.
4. Інтерполювання за рівномірною сіткою вузлів.
5. Многочлен Ньютона. Поняття скінчених різниць. Побудова многочлену Ньютона за таблицею скінчених різниць.

Лекція 8. Поняття сплайн-інтерполяції. (2 год.)

1. Поняття сплайн-інтерполяції.
2. Складання системи лінійних рівнянь для визначення коефіцієнтів сплайн-інтерполяції.
3. Алгоритм прогонки для знаходження коефіцієнтів сплайн-інтерполяції.

Практичне заняття 3. Інтерполювання функцій. (4 год.)

1. Приклади задач, які використовують інтерполювання функцій однієї змінної. Побудова лінійних та квадратичних многочленів Лагранжа та Ньютона. Складання таблиць скінчених різниць.

2. Інтерполяція функції однієї змінної. Інтерполяційні формули Ньютона (інтерполяція вперед і назад) для обчислення значень функції на підставі табличної залежності $y(x)$ при заданих значеннях аргументу.

3. Знаходження проміжних значень функції однієї змінної за допомогою многочлену Лагранжа. Знаходження проміжних значень функції однієї змінної за допомогою многочлену Ньютона. Побудова кубічного сплайну за інтерполяційною таблицею.

Перелік питань до самостійної роботи

1. Призначення інтерполяційних формул.
2. Інтерполяційні формули Ньютона (інтерполяція вперед і назад).
3. Визначення абсолютної та відносної похибки арифметичних обчислень.
4. Використання інтерполяційних формул для обчислення значень функції на підставі табличної залежності $y(x)$ при заданих значеннях аргументу.
5. Визначення залишку інтерполяційного полінома.

Література [2] - [7], [11] - [14].

Тема 7. Задача наближення функцій однієї змінної

Задача наближення функцій однієї змінної. Найкраще наближення функцій за допомогою многочленів. Метод найменших квадратів. Формули наближення функцій за допомогою лінійних та квадратичних многочленів.

Лекція 9. Задача наближення функцій однієї змінної (4 год.)

1. Задача наближення функцій однієї змінної.

2. Найкраще наближення функцій за допомогою многочленів.
3. Метод найменших квадратів.
4. Формули наближення функцій за допомогою лінійних та квадратичних многочленів.

Практичне заняття 4. Задача наближення функцій однієї змінної. (2 год.)

1. Побудова лінійних та квадратичних наближення для функцій однієї змінної за методом найменших квадратів.

Перелік питань до самостійної роботи

1. Стійкість методу найменших квадратів.
2. Обчислювальна схема наближення функцій
3. Похибки наближення.

Література [2] - [7], [11], [14].

Модуль 2

Змістовний модуль 2. Чисельні методи аналізу і математичної фізики

Тема 8. Чисельне диференціювання функцій

Чисельне диференціювання функцій. Некоректність задачі чисельного диференціювання. Формули для знаходження першої та другої похідної. Оцінки похибки чисельного диференціювання. Регуляризація задачі чисельного диференціювання.

Лекція 10. Чисельне диференціювання функцій (4 год.)

1. Чисельне диференціювання функцій.
2. Некоректність задачі чисельного диференціювання.
3. Формули для знаходження першої та другої похідної.
4. Оцінки похибки чисельного диференціювання.
5. Регуляризація задачі чисельного диференціювання.

Практичне заняття 5. Чисельне диференціювання функцій (2 год.)

1. Різницеві формули для визначення похідних на основі інтерполяційного полінома Ньютона.

2. Формули наближеного диференціювання на основі на інтерполяційної формули Стірлінга.

3. Чисельне диференціювання на основі інтерполяційних формул Гаусса і Бесселя. Різницеві формули для визначення похідної довільного порядку.

4. Метод невизначених коефіцієнтів (для нерівномірного розбиття). Знаходження першої та другої похідної табличної функції.

Перелік питань до самостійної роботи

1. Використання формули Рунге покращення апроксимації при чисельному диференціюванні.

2. Виведення формули чисельного диференціювання високого порядку апроксимації з рівномірним розбиттям області диференціювання.

3. Виведення формули чисельного диференціювання з нерівномірним розбиттям області диференціювання за заданим шаблоном.

Індивідуальні заняття

1. Регуляризація методів чисельного диференціювання
2. Формули чисельного диференціювання високого порядку точності

3. Алгоритми чисельного диференціювання

Література [3], [4], [5], [7], [8], [11] - [14].

Тема 9. Чисельне інтегрування функції однієї змінної

Чисельне інтегрування функції однієї змінної. Метод прямокутників та метод трапецій. Оцінка похибки. Блок-схеми методів. Вибір кроку сітки. Чисельне інтегрування функції однієї змінної. Метод Симпсона. Оцінка похибки. Блок-схема методу. Порівняльний аналіз методів прямокутників, трапецій та Симпсона.

Лекція 11. Чисельне інтегрування функції однієї змінної (4 год.)

1. Чисельне інтегрування функції однієї змінної.
2. Метод прямокутників та метод трапецій.
3. Оцінка похибки.
4. Блок-схеми методів.
5. Вибір кроку сітки.

Лекція 12. Метод Симпсона (2 год.)

1. Метод Симпсона.
2. Оцінка похибки.
3. Блок-схема методу.
4. Порівняльний аналіз методів прямокутників, трапецій та Симпсона.

Практичне заняття 6. Чисельне інтегрування функції (2 год.)

1. Приклади задач, які вимагають чисельного інтегрування функції однієї змінної.

2. Складання формул прямокутників та трапецій для двох вузлів. Складання формул Симпсона за трьома вузлами.

3. Організація обчислень при чисельному інтегруванні.
4. Чисельне інтегрування. Формули трапецій та Симпсона.
5. Формули Рунге покращення апроксимації, застосовані до формул трапецій та Симпсона.

6. Формула Гауса. Обчислення інтегралу табличної функції за методами прямокутників та метод трапецій. Обчислення інтегралу табличної функції за методом Симпсона.

Перелік питань до самостійної роботи

1. Формула Лапласа. Формула Грегорі.
2. Обчислення невластних інтегралів.
3. Метод Люстерника.
4. Метод Монте-Карло та його застосування.

Індивідуальні заняття

1. Порівняння формул прямокутників, трапецій та Сімпсона для чисельного інтегрування.
 2. Призначення методу Рунге.
 3. Призначення формул Ньютона-Котеса.
 4. Суть методів підвищеної точності.
 5. Використання формул прямокутників, трапецій та Сімпсона для чисельного інтегрування.
 6. Використання формули Рунге покращення апроксимації при чисельному інтегруванні.
 7. Використання формули Гауса підвищеної точності.
- Література [2], [3], [5], [7], [9], [11], [13], [14].

Тема 10. Чисельне інтегрування звичайних диференціальних рівнянь

Чисельні методи розв'язування задачі Коші для звичайних диференціальних рівнянь першого порядку. Метод Ейлера. Оцінка локальної та глобальної похибки методу Ейлера. Блок-схема методу Ейлера. Вибір кроку сітки. Чисельні методи розв'язування задачі Коші для звичайних диференціальних рівнянь першого порядку. Метод Рунге-Кутта. Оцінка локальної та глобальної похибки методу Рунге-Кутта. Блок-схема методу Рунге-Кутта. Вибір кроку сітки. Узагальнення методу Рунге-Кутта для систем звичайних диференціальних рівнянь першого порядку. Блок-схема алгоритму. Поняття про граничні задачі для звичайних диференціальних рівнянь другого порядку та методи їх розв'язування.

Лекція 13. Чисельне інтегрування звичайних диференціальних рівнянь (2 год.)

1. Чисельні методи розв'язування задачі Коші для звичайних диференціальних рівнянь першого порядку.
2. Метод Ейлера.
3. Оцінка локальної та глобальної похибки методу Ейлера.
4. Блок-схема методу Ейлера.
5. Вибір кроку сітки.

Лекція 14. Чисельне інтегрування звичайних диференціальних рівнянь (2 год.)

1. Метод Рунге-Кутта.
2. Оцінка локальної та глобальної похибки методу Рунге-Кутта.
3. Блок-схема методу Рунге-Кутта.
4. Вибір кроку сітки.

Лекція 15. Методи розв'язання крайових задач для звичайних диференціальних рівнянь (2 год.)

1. Постановки крайових задач.
2. Метод стрільби.
3. Метод пристрілки.

4. Метод лінеаризації.

5. Метод продовження за параметром.

Практичне заняття 7. Чисельне інтегрування звичайних диференційних рівнянь (2 год.)

1. Складання математичних моделей кінетичних та динамічних задач. Організація обчислень при наближеному розв'язуванні задачі Коші для звичайних диференційних рівнянь першого порядку за методом Ейлера. Застосування методу Рунге-Кутта.

2. Практичне застосування апостеріорних оцінок локальної та глобальної похибки методу Рунге-Кутта.

3. Метод Ейлера з уточненням.

Перелік питань до самостійної роботи

1. Метод Адамса з другими різницями.

2. Метод прогонки.

3. Розв'язування граничної задачі для звичайного диференційних рівнянь другого порядку.

4. Різницеві методи розв'язання диференційних рівнянь вищих порядків.

5. Оцінка похибки, збіжності та стійкості різницевих методів. Формула Адамса.

6. Метод Гальоркіна.

Індивідуальні заняття

1. Порівняння методів Рунге-Кутта та Адамса.

2. Суть методу скінчених різниць.

3. Алгоритм методу прогонки.

4. Використання методу прогонки.

5. Застосування інтерполяційної формули Ньютона для інтегрування диференційних рівнянь 1-го порядку.

Література: [1], [2] - [6], [8], [11], [14].

Тема 11. Чисельні методи математичної фізики

Поняття задач математичної фізики та їх класифікація. Чисельні методи розв'язування двовимірних рівнянь Лапласа та Пуассона. Метод послідовних ітерацій. Блок-схема алгоритму. Чисельні методи розв'язування одновимірних рівнянь теплопровідності. Поняття часового шару. Двошарові скінчено-різницеві схеми. Оцінка похибки апроксимації. Умова стійкості алгоритму. Блок-схема алгоритму. Чисельні методи розв'язування одновимірних хвильових рівнянь. Задача коливання струни та стержня. Трьохшарові скінчено-різницеві схеми. Оцінка похибки апроксимації. Умова стійкості алгоритму. Блок-схема алгоритму.

Лекція 16. Чисельні методи математичної фізики (2 год.)

1. Поняття задач математичної фізики та їх класифікація.

2. Чисельні методи розв'язування двовимірних рівнянь Лапласа та Пуассона.
3. Метод послідовних ітерацій. Блок-схема алгоритму.

Лекція 17. Чисельні методи розв'язування одновимірних рівнянь теплопровідності (2 год.)

1. Чисельні методи розв'язування одновимірних рівнянь теплопровідності.
2. Поняття часового шару.
3. Двошарові скінчено-різницеві схеми.
4. Оцінка похибки апроксимації.
5. Умова стійкості алгоритму.
6. Блок-схема алгоритму.

Лекція 18. Чисельні методи розв'язування одновимірних хвильових рівнянь (2 год.)

1. Чисельні методи розв'язування одновимірних хвильових рівнянь.
2. Задача коливання струни та стержня.
3. Трьохшарові скінчено-різницеві схеми.
4. Оцінка похибки апроксимації.
5. Умова стійкості алгоритму.
6. Блок-схема алгоритму.

Практичне заняття 8. Чисельні методи математичної фізики (6 год.)

1. Приклади математичних моделей, які зводяться до рівнянь Лапласа та Пуассона.
2. Організація обчислень при наближеному розв'язування двовимірних рівнянь Лапласа та Пуассона.
3. Застосування формул метод послідовних ітерацій. Приклади математичних моделей, які зводяться до рівнянь теплопровідності та одновимірних хвильових рівнянь. Організація обчислень при застосуванні двошарових та трьохшарових скінчено-різницевих схем. Практична перевірка умов стійкості скінчено-різницевої схеми.
4. Метод сіток розв'язання диференційних рівнянь еліптичного, гіперболічного та параболічного типів, а також змішаної задачі.
5. Розв'язування двовимірного рівняння Пуассона за методом послідовних ітерацій. Розв'язування одновимірного рівняння теплопровідності скінчено-різницевим методом. Розв'язування одновимірного хвильового рівняння скінчено-різницевим методом. Чисельне розв'язування задачі коливання струни.

Перелік питань до самостійної роботи

1. Метод прогонки розв'язання граничних задач для рівнянь в частинних похідних.
2. Метод прямих розв'язку граничних задач для диференційних рівнянь в частинних похідних.

3. Варіаційні методи розв'язання граничних задач рівнянь математичної фізики. Метод Рітца.

Індивідуальні заняття

1. Поняття апроксимації, збіжності та стійкості методу сіток.
2. Порівняння диференціальних рівнянь еліптичного, гіперболічного та параболічного типів.
3. Особливість розв'язання змішаної задачі.
4. Використання методу сіток для розв'язання диференціальних рівнянь еліптичного типу.
5. Використання методу сіток для розв'язання диференціальних рівнянь гіперболічного типу.
6. Використання методу сіток для розв'язання диференціальних рівнянь параболічного типу.
7. Використання методу сіток для розв'язання змішаної задачі для рівняння коливання струни.

Література [1] - [6], [9], [11], [13].

Методи і форми проміжного та підсумкового контролю

Контроль засвоєння студентами дисципліни здійснюється в кілька етапів:

1. контрольний захід після кожного залікового кредиту;
2. підсумкова атестація з кожного модульного контролю;
3. підсумкова атестація з формування сумарної модульної оцінки;
4. ПМК.

Програмний матеріал навчальної дисципліни складається з двох модулів, які охоплюють 11 тем. Оцінювання проводиться з урахуванням всіх видів навчального процесу:

- знань з теорії за відсутності практичного заняття по темі відбувається у вигляді письмового контрольного заходу на 10 - 15 хв. під час лекції;

- знань, умінь і практичних навичок студента за результатами захисту звіту про виконання практичного завдання;

- індивідуальної роботи студента;

- самостійної роботи студента.

Контрольний захід проводиться у вигляді контрольної роботи в аудиторії за розкладом у кожній академічній групі окремо. На контрольну роботу відводиться дві академічні години. При цьому завдання включають три запитання з теоретичного матеріалу і одне практичне завдання певного змістовного модуля.

У випадку неявки студента на контрольний захід з поважних причин, підтверджених документально, викладач проводить контрольне опитування студента під час чергової консультації.

Узагальнююче оцінювання знань, умінь і практичних навичок студента здійснюється за 100 бальною системою.

Умови переведення даних 100-бальної шкали оцінювання в 4-х бальну та шкалу ECTS

Переведення даних 100-бальної шкали оцінювання в 4-х бальну та шкалу за системою ECTS здійснюється в такому порядку:

Оцінка за шкалою ECTS	Оцінка за бальною шкалою, прийнятою в НУДПСУ	Оцінка за національною шкалою
A	84-100	5 (відмінно)
BC	67-83	4 (добре)
DE	50-66	3 (задовільно)
FX	33-49	2 (незадовільно) з можливістю повторного складання
F	0-32	2 (незадовільно) з обов'язковим повторним вивченням

Розподіл балів при рейтинговій системі
з дисципліни “Чисельні методи в інформатиці”

	Модуль I=43							Модуль II=57											
	ЗМ I=43							ЗМ II=57											
Теми	T1	T2	T3	T4	T5	T6	T7	Модульна контрольна робота 7 балів	Підсумкова атестація по I-му модулю MI=ЗМI = 43 балів	T8	T9	Контрольний захід - 5	T10	T11	Модульна контрольна робота - 10	Підсумкова атестація по II-му модулю MII=ЗМII = 57 балів	Підсумкова атестація M1+M2=43+57=100		
Виконання та захист практичних завдань	-	-	-	5	5	10	5			5	5		5	5				5	15
Виконання та захист інд. роботи	-	-	-	2	2	-	-			2	2		2	2				2	2
Виконання та захист СРС	1	1	1	1	1	1	1			1	1		1	1				1	1
Всього	1	1	1	8	8	11	6			8	8		8	8				8	8

ОПОРНИЙ КОНСПЕКТ ЛЕКЦІЙ З КУРСУ
«ЧИСЕЛЬНІ МЕТОДИ В ІНФОРМАТИЦІ»

Лекція 1. Вступ. Етапи розв'язання прикладних задач на ЕОМ

Мета лекції: ознайомити з місцем дисципліни «Чисельні методи в інформатиці» в системі підготовки студентів за напрямом «Комп'ютерні науки», з методикою застосування чисельних методів при розв'язуванні прикладних задач на ЕОМ.

План лекції:

1. Вступ.
2. Етапи розв'язання прикладних задач на ЕОМ.
3. Поняття та приклади математичних моделей.
4. Місце чисельних методів при розв'язуванні прикладних задач.
5. Зв'язок чисельних методів з іншими дисциплінами інформаційного циклу.

Неможливо уявити собі сучасну науку без широкого застосування математичного моделювання. Сутність цієї методології полягає в заміні початкового об'єкту математичною моделлю з метою подальшому вивченні моделі за допомогою реалізуємих на комп'ютерах обчислювально-логічних алгоритмів.

Цей метод пізнання, конструювання, проектування містить в собі багато переваг як теорії, так і експерименту. Робота не з самим об'єктом (явищем, процесом), а з його моделлю дає можливість безболісно, відносно швидко і без істотних витрат досліджувати його властивості і поведінку в будь-яких мислимих ситуаціях (переваги теорії). В той же час обчислювальні (комп'ютерні, симуляції, імітаційні) експерименти з моделями об'єктів дозволяють, спираючись на потужність сучасних обчислювальних методів і технічних інструментів інформатики, детально і глибоко вивчати об'єкти в достатній повноті, неприступній чисто теоретичним підходам (переваги експерименту). Недивно, що методологія математичного моделювання бурхливо розвивається, охоплюючи всі нові сфери - від розробки технічних систем і управління ними до аналізу найскладніших економічних і соціальних процесів.

Елементи математичного моделювання використовувалися з самого початку появи точних наук, і не випадково, що деякі методи обчислень носять імена таких корифеїв науки, як Ньютон і Ейлер, а слово «алгоритм» походить від імені арабського вченого Аль-Хорезмі. Друге «народження» цієї методології припало на кінець 40-х-початок 50-х років ХХ століття і було обумовлено принаймні двома причинами. Перша з них - поява ЕОМ (комп'ютерів), хоча і скромних за нинішніми мірками, але проте позбавили вчених від величезної за об'ємом рутинної обчислювальної роботи. Друга - безпрецедентний соціальний заказ - виконання національних програм СРСР і США по створенню ракетно-ядерного щита, які не могли бути реалізовані традиційними методами.

Математичне моделювання справилося з цим задачею: ядерні вибухи і польоти ракет і супутників були попередньо «здійснені» в надрах ЕОМ за допомогою математичних моделей і лише потім втілені на практиці. Цей успіх у цілому визначив подальші досягнення методології, без застосування

якої в розвинутих країнах жоден великомасштабний технологічний, екологічний або економічний проект тепер всерйоз не розглядається (казане справедливо і по відношенню до деяких соціально-політичних проектів).

Зараз математичне моделювання вступає в третій принципіальне важливий етап свого розвитку, «вбудовуючись» в структури так званого інформаційного суспільства. Вражаючий прогрес засобів переробки, передачі і зберігання інформації відповідає світовим тенденціям про ускладнення і взаємного проникнення різних сфер людської діяльності. Без володіння інформаційним «ресурсами» не можна і думати про вирішення складних і різноманітних проблем, що стоять перед світовим співтовариством.

Проте інформація як така часто мало що дає для аналізу і прогнозу, для ухвалення рішень і контролю за їх виконанням. Потрібні надійні способи переробки інформаційної «сировини» в готовий «продукт», тобто в точне знання. Історія методології математичного моделювання переконує: вона може і має бути інтелектуальним ядром інформаційних технологій всього процесу інформатизації суспільства.

Технічні, екологічні, економічні і інші системи, що вивчаються сучасною наукою, більше не піддаються дослідженню (у потрібній повноті і точності) звичайними теоретичними методами. Прямий натурний експеримент над ними – це довгий, дорогий, часто або небезпечний, або просто неможливий, оскільки багато з цих систем існують в «єдиному екземплярі». Ціна помилок і прорахунків у обігу з ними неприпустимо висока. Тому математичне (ширше - інформаційне) моделювання є обов'язковою складовою науково-технічного прогресу.

Сама постановка питання про математичне моделювання будь-якого об'єкту породжує чіткий план дій. Його можна умовно розбити на три етапи: *модель - алгоритм – програма*.

На першому етапі вибирається (або будується) «еквівалент» об'єкту, що відображає в математичній формі найважливіші його властивості - закони, яким він підкоряється, зв'язки, властиві складовим його частинам, і т.п. Математична модель (або її фрагменти) досліджується теоретичними методами, що дозволяє отримати важливі попередні знання про об'єкт.

Другий етап - вибір (або розробка) алгоритму для реалізації моделі на комп'ютері. Модель представляється у формі, зручній для застосування чисельних методів, визначається послідовність операцій, які потрібно провести, щоб знайти шукані величини із заданою точністю. Обчислювальні алгоритми повинні не спотворювати основні властивості моделі і початкового об'єкту, бути економічними і такими, що адаптуються до особливостей вирішуваних завдань і використовуваних комп'ютерів.

На третьому етапі створюються програми, що «перекладають» модель і алгоритм на доступну комп'ютеру мову. До них також пред'являються вимоги економічності і адаптивності. Їх можна назвати «електронним» еквівалентом об'єкту, що вивчається, вже придатним для безпосереднього випробування на «експериментальній установці» - комп'ютері.

Створивши тріаду «модель-алгоритм-програма», дослідник отримує в руки універсальний, гнучкий і недорогий інструмент, який спочатку підлаштовується, тестується в «пробних» обчислювальних експериментах. Після того, як адекватність тріади початковому об'єкту засвідчена, з моделлю проводяться різноманітні і докладні «досліди», що дають всі потрібні якісні і кількісні властивості і характеристики об'єкту.

Процес моделювання супроводжується поліпшенням і уточненням, по мірі необхідності, всіх ланок тріади. Будучи методологією, математичне моделювання не підміняє собою математику, фізику, біологію та інші наукові дисципліни, не конкурує з ними. Створення і застосування тріади неможливе без опори на самі різні методи і підходи - від якісного аналізу нелінійних моделей до сучасних мов програмування.

Воно дає нові додаткові стимули самим різним напрямкам науки. Розглядаючи питання ширше, нагадаємо, що моделювання присутнє майже у всіх видах творчої активності людей різних «спеціальностей» - дослідників і підприємців, політиків і воєначальників. Звичайно ж, математичне моделювання плідне лише при виконанні добре відомих професійних вимог: чітке формулювання основних понять і припущень, апостеріорний аналіз адекватності моделей, гарантована точність обчислювальних алгоритмів і т.п.

Приклади простих математичних моделей:

$$pV = \frac{mRT}{M} \text{ - рівняння стану ідеального газу;}$$

$$F = G \frac{m_1 \cdot m_2}{R^2} \text{ - закон всесвітнього тяжіння;}$$

$$\frac{mv^2}{2} = mgh \text{ - закон збереження енергії;}$$

$$F = k \frac{q_1 \cdot q_2}{R^2} \text{ - закон Кулона.}$$

Основні етапи математичного моделювання:

1. Розробка моделі - *формалізація*. Вивчається в прикладних і фундаментальних науках.
2. Розробка методу (алгоритму) розв'язання рівняння моделі - *алгоритмізація*. Вивчається в обчислювальній математиці.
3. Створення програми - *програмування*. Вивчається в інформатиці.
4. Розрахунки, аналіз результатів - *практичне використання*.

Предметом чисельних методів є алгоритми розв'язання математичних задач, що виникають при дослідженні реальних об'єктів методом математичного моделювання.

Наприклад, нехай потрібно знайти R або v із записаних вище моделей. Що загального в цих задачах? Те, що потрібно розв'язати рівняння вигляду:

$$x^2 = a \quad (1)$$

Чисельні методи не вивчають методи розв'язання конкретних задач, а пропонують алгоритми розв'язання задач у абстрактного вигляді.

З погляду звичайної математики точний розв'язок рівняння (1) такий:

$$x_{1,2}^* = \pm \sqrt{a},$$

причому якщо $a > 0$, то є два дійсних розв'язків;

якщо $a = 0$, то є тільки тривіальний розв'язок $x_1^* = x_2^* = 0$;

якщо $a < 0$, то дійсних розв'язків немає.

Але знак $\sqrt{\quad}$ не вирішує задачу, оскільки не дає практичного способу (алгоритму) обчислення значення x для конкретного значення a .

Чисельний метод пропонує наступний алгоритм обчислення x^* :

Вибрати початкове значення x , наприклад $x_0 = a$. Це початкове наближення розв'язку. Обчислювати нові наближення розв'язку x_i за формулою:

$$x_i = \frac{1}{2} \left(x_{i-1} + \frac{a}{x_{i-1}} \right) \quad (2)$$

до досягнення умови:

$$|x_i - x_{i-1}| \leq \varepsilon \quad (3)$$

Тут $i = 1, 2, \dots$ - номер обчислення - *ітерації*, ε - необхідна точність.

Приклад. Розв'язати рівняння $x^2 - 3 = 0$ з точністю $\varepsilon = 0,001$.

Задамо $x_0 = a = 3$. Обчислимо перше наближення:

$$x_1 = \frac{1}{2} \left(x_0 + \frac{a}{x_0} \right) = \frac{1}{2} \left(3 + \frac{3}{3} \right) = 2,$$

оцінимо точність $|x_1 - x_0| = |2 - 3| = 1 \geq \varepsilon$. Необхідна точність не досягнута, потрібно продовжити розрахунок.

Обчислимо друге наближення: $x_2 = \frac{1}{2} \left(x_1 + \frac{a}{x_1} \right) = \frac{1}{2} \left(2 + \frac{3}{2} \right) = 1,75 = \frac{7}{4}$,

оцінимо точність $|1,75 - 2| = 0,25 > \varepsilon$.

Обчислимо третє наближення: $x_3 = \frac{1}{2} \left(x_2 + \frac{a}{x_2} \right) = \frac{1}{2} \left(\frac{7}{4} + \frac{12}{7} \right) = 1,7321$,

оцінимо точність $|1,7321 - 1,75| = 0,0179 > \varepsilon$.

Обчислимо четверте наближення: $x_4 = 1,73205$, оцінимо точність $|1,73205 - 1,7321| = 0,00005 < \varepsilon$ - точність досягнута.

Відповідь: $x_{1,2} = +1,73205$.

Точне значення (до 8 значущих цифр): $\sqrt{3} = 1,7320508$.

Розглянутий приклад демонструє принципи, загальні для ітераційних методів розв'язання задач обчислювальної математики:

1. Вихідна задача (1) замінюється іншою задачею – обчислювальним алгоритмом по формулах (2), (3), де використовуються тільки арифметичні операції. Прийнято називати (2) *формулою ітераційного процесу (ітераційним процесом)*, (3) - *умовою завершення ітераційного процесу*.

1. Задача (2) містить новий параметр i - *номер ітерації*. Очевидно, що число ітерацій впливає на точність рішення. Якщо $\lim_{i \rightarrow \infty} x_i = x^* = \pm\sqrt{a}$, то ітераційний процес є *таким, що збігається* – він дозволяє отримати розв'язок початкової задачі (1).

2. Рішення, отримане ітераційним методом, завжди є наближенням, оскільки точне рішення отримати неможливе - потрібні нескінченні обчислення.

Важливо підкреслити, що формула (2) отримана з (1) шляхом тотожних перетворень:

$$x^2 - a = 0 \Leftrightarrow \underset{*(-1)}{a - x^2} = 0 \Leftrightarrow \underset{+2x^2}{a + x^2} = 2x^2 \Leftrightarrow \underset{:2x}{x} = \frac{1}{2} \left(x + \frac{a}{x} \right)$$

Але не всяке тотожне перетворення дозволяє отримати ітераційний процес, що збігається.

Наприклад:

а) $x^2 - a = 0 \Rightarrow x^2 = a \Rightarrow x = \frac{a}{x}$,

Виконаємо розрахунок при $a=3$:

$$x_0 = 3; x_1 = \frac{a}{x_0} = 1; x_2 = \frac{a}{x_1} = 3; x_3 = 1 \dots$$

Ітераційний процес *не збігається*; значення наближень *коливаються*.

б) $x^2 - a = 0 \Rightarrow x = x^2 + x - a$

$$x_0 = 3; x_1 = 3^2 + 3 - 3 = 9; x_2 = 9^2 + 9 - 3 = 87 \dots$$

Ітераційний процес *розбіжний*.

Розглянутий приклад ілюструє один з видів чисельних методів - *ітераційний*.

Види чисельних методів:

1. Прямі – точний розв'язок отримують за скінчене число арифметичних дій.

2. Ітераційні - точний розв'язок може бути отриманий теоретично у вигляді границі нескінченної послідовності, що збігається.

3. Імовірнісні - методи випадкового пошуку розв'язку (*вгадування*).

Всі види чисельних методів дозволяють отримати тільки наближений розв'язок задачі, тобто чисельний розв'язок завжди містить похибку.

Контрольні запитання:

1. Роль чисельних методів при розв'язанні складних задач, що виникають при моделюванні економіко-математичних систем.
2. Поняття тріади «модель-алгоритм-програма».
3. Приклади математичних моделей.
4. Предмет дисципліни «Чисельні методи».
5. Види чисельних методів.

Література [3], [4], [5], [8], [10].

Лекція 2. Сучасні пакети прикладних програм, призначені для чисельного розв'язання математичних моделей

Мета лекції: ознайомитися з сучасними пакетами прикладних програм – Maple, Mathematica, MATLAB та їх основними можливостями при чисельному розв'язанні математичних моделей.

План лекції:

1. Математичний пакет Maple та його основні можливості.
2. Аналітичне та чисельне диференціювання та інтегрування в Maple.
3. Математичний пакет Mathematica та його можливості. Виконання чисельних розрахунків в пакеті Mathematica.
4. Математичний пакет MATLAB.
5. Інтерфейс та основні модулі пакету.
6. Робота з матрицями в пакеті MatLab.
7. Використання спеціальних функцій пакету MatLab для чисельного розв'язання алгебраїчних та диференційних рівнянь.

Зовсім недавно користувач ЕОМ, вирішуючи навіть прості обчислювальні завдання, був вимушений освоювати основи програмування і готувати кустарні програми, навряд чи потрібні кому-небудь ще, окрім їх творця. Тим часом можливості комп'ютерів постійно зростали. Зараз персональний комп'ютер (ПК) набагато перевершує по своїх можливостях перші ЕОМ, що займали цілі кімнати і зали. А швидкість обчислень нинішніх ПК в сотні разів перевершує швидкість обчислень легендарних ІВМ РС ХТ і АТ (перших ПК) і впритул наближається до швидкості обчислень СУПЕРЕОМ недавнього минулого.

У зв'язку з цим почав змінюватися погляд на призначення комп'ютера. На перше місце вийшло застосування їх для роботи з текстовими процесорами (наприклад, Microsoft Word) і прикладними програмними системами для автоматизації офісної діяльності. На жаль, при цьому багато користувачів почали забувати про те, що ЕОМ спочатку створювалися для обчислень.

Проте часи змінюються і вічні цінності, до яких належать розум і освіта, знов повертаються. Останніми роками у всьому світі істотно зріс інтерес до серйозного застосування ПК, зокрема в області математичних розрахунків. Цьому у великій мірі сприяла розробка спеціальних

комп'ютерних математичних програмних систем, що різко понизили потребу в написанні власних програм при вирішенні математичних завдань.

Перше покоління таких систем було орієнтоване на операційну систему MS-DOS і з'явилося, здавалося б, зовсім недавно - на початку 90-х рр. Так або інакше, але комп'ютерний мир знов заговорив про «штучний інтелект», розуміючи під цим здатність електронної машини видавати нетривіальні рішення і навчатися вирішенню нових завдань. Інтерес до комп'ютерного моделювання в найширших областях помітно зріс після шахових баталій між суперкомп'ютером фірми IBM і колишнім чемпіоном світу по шахах Гарі Каспаровим. Як відомо, вони завершилися тріумфальною перемогою машини - або, точніше кажучи, колективного розуму тих, хто створив її і її програмне забезпечення.

Останніми роками показником інтелектуальної потужності комп'ютерів, у тому числі і персональних, стали вже не програми для гри в шахи, а новітні програмні системи символічної математики або комп'ютерної алгебри. Створені для проведення символічних перетворень математичних виразів, ці системи були доведені до рівня, що дозволяє різко полегшити, а часом і замінити, працю найшанобливішої наукової еліти миру - математиків: теоретиків і аналітиків. Вже з'явилися відкриття, зроблені за допомогою таких систем, - але не ними самими! Навряд чи є хоч один дійсно серйозний науковий проект, пов'язаний з математикою, де вони не застосовувалися.

Системи символічної математики довгий час були орієнтовані на великі комп'ютери, з появою ПК класу IBM PC і Macintosh і із зростанням їх можливостей ці системи були перероблені під них і доведені до рівня масових серійних програмних систем. Зараз системи символічної математики (або комп'ютерної алгебри) випускаються самого різного «калібру» - від розрахованої «на всіх» системи Mathcad, вражаюче компактної, швидкої і зручної для простих символічних обчислень системи Derive і до комп'ютерних монстрів Mathematica, MATLAB і Maple, що мають тисячі вбудованих і бібліотечних функцій і дивовижні можливості графічної візуалізації обчислень.

Всі ці системи працюють на персональних комп'ютерах, оснащених популярними операційними системами класу Windows 95/98/NT/2000. Але не тільки на них - є версії під операційні системи Linux, Unix, Mac і ін. Вони давно знайомі користувачам великих комп'ютерів і навіть суперкомп'ютерів.

До середнього рівня таких систем відносяться системи класу Mathcad, що інтенсивно розвиваються, мають (на додаток до прекрасних засобів числових обчислень) придбане за ліцензією у фірми Waterloo Maple Inc. (творця систем Maple) ядро символічних обчислень. Ядро системи Maple використовується і в іншій маститій системі - MATLAB, додаючи їй незвичайні для неї можливості символічної математики.

Система Maple була створена групою учених, що займаються символічними обчисленнями (The Symbolic Group), організованою Кейтом Геддом (Keith Geddes) і Гастоном Гоне (Gaston Gonnet) в 1980 році в університеті Waterloo, Канада. Спочатку вона була реалізована на великих

комп'ютерах і пройшла довгий шлях апробації, увібравши в своє ядро і бібліотеки велику частину математичних функцій і правил їх перетворень, вироблених математикою за сторіччя розвитку.

Системам класу Maple присвячені сотні книг. Достатньо повний список (близько 400 найменувань) книг по системах Maple можна знайти на сайті розробника цієї системи - компанії Waterloo Maple Software (www.maplesoft.com).

Навряд чи ця могутня математична система, що розділяє претензії на світове лідерство з системами Mathematica фірми Wolfram Research Inc., потрібна секретарці або навіть директорові невеликої комерційної фірми. Але, поза сумнівом, будь-яка серйозна наукова лабораторія або кафедра вузу повинні мати в своєму розпорядженні подібну систему, якщо вони серйозно зацікавлені в автоматизації виконання математичних розрахунків будь-якого ступеня складності.

Не дивлячись на свою спрямованість на серйозні математичні обчислення, системи класу Maple необхідні досить широкої категорії користувачів: студентам і викладачам вузів, інженерам, аспірантам, науковцям і що навіть вчиться математичних класів загальноосвітніх і спеціальних шкіл. Всі вони знайдуть в Maple численні гідні можливості для застосування.

Порівняння системи Maple з лідером серед систем комп'ютерної математики - системою Mathematica - непродуктивно. У кожній програмі є свої достоїнства і недоліки. А головне - у них є свої прихильники, яких марно переконувати, що інша система в чомусь краща. Це все одно, що порівнювати великих виконавців джазової і рок-музики Луї Амстронга і Бі Бі Кінга. Всі, хто серйозно застосовують системи комп'ютерної математики, повинно працювати з декількома системами, бо тільки це гарантує високий рівень надійності складних обчислень.

Треба відзначити, що інтерфейс Maple більш інтуїтивно зрозумілий, ніж у строгої Mathematica. Maple на перший погляд має декілька менш могутню графіку, але простота управління параметрами і легкість підготовки графічних процедур часто дозволяє візуалізувати вирішення математичних завдань з меншими зусиллями, чим при використанні системи Mathematica. Обидві системи в останніх реалізаціях зробили якісний стрибок у напрямі ефективності вирішення завдань в чисельному вигляді, зокрема за рахунок підвищення швидкості виконання матричних операцій.

Особливо ефективно використання Maple при навчанні математиці. Високий «інтелект» цієї системи символічної математики об'єднується в ній з прекрасними засобами математичного чисельного моделювання і просто приголомшливими можливостями графічної візуалізації рішень. Застосування таких систем, як Maple, можливо при викладанні і самоосвіті від самих основ до вершин математики.

Перерахуємо основні можливості системи Maple.

Інтерфейс:

- робота з багатьма вікнами;

- виведення графіків в окремих вікнах або у вікні документа;
- представлення вихідних і вхідних даних у вигляді природних математичних формул;
- завдання текстових коментарів різними шрифтами;
- можливість використання гіперпосилань і підготовки електронних документів;
- зручне управління за допомогою клавіатури через головне меню і інструментальну панель;
- управління за допомогою миші.

Символьні і чисельні обчислення:

- диференціювання функцій;
- чисельна і аналітична інтеграція;
- обчислення меж функцій;
- розкладання функцій в ряди;
- обчислення сум і творів;
- інтегральні перетворення Лапласа, Фур'є і др.;
- дискретні Z-преобразования;
- пряме і зворотне швидке перетворення Фур'є;
- робота з кусочно-заданими функціями.

Робота з рівняннями в чисельному і символьному вигляді:

- вирішення систем лінійних і нелінійних рівнянь;
- вирішення систем диференціальних рівнянь;
- символьне обчислення рядів;
- робота з рекурентними функціями;
- вирішення трансцендентних рівнянь;
- вирішення систем з нерівностями.

Робота з функціями:

- обчислення значень всіх елементарних функцій;
- обчислення значень більшості спеціальних математичних функцій;
- перерахунок координат крапок між різними координатними системами;
- завдання функцій користувача.

Лінійна алгебра:

- понад сто операцій з векторами і матрицями;
- вирішення систем лінійних рівнянь;
- формування спеціальних матриць і їх перетворення;
- обчислення власних значень і власних векторів матриць;
- підтримка швидких векторних і матричних алгоритмів пакету програм NAG.

Графічна візуалізація результатів обчислень:

- побудова графіків багатьох функцій;
- різні типи осей (з лінійним і логарифмічним масштабом);

- графіки функцій в декартовій і полярній системах координат;
- спеціальні види графіків (точки масивів, векторні графіки, діаграми рівнів і ін.);
- системи координат, визначувані користувачем;
- графіки, що представляють вирішення диференціальних рівнянь;
- графіки тривимірних поверхонь з функціональним зафарбовуванням;
- побудова пересічних в просторі об'єктів;
- завдання користувачем забарвлення графіків;
- імпорт графіків з інших пакетів і програмних систем;
- анімація графіків;
- створення і програвання анімаційних файлів.

Програмування:

- вбудована мова процедурного програмування;
- простий і типовий синтаксис мови програмування;
- обширний набір типів даних;
- типи даних, що задаються користувачем;
- засоби відладки програм;
- могутні бібліотеки функцій;
- завдання зовнішніх функцій і процедур;
- підтримка мов програмування 3 і Fortran;
- можливість запису формул у форматі Latex.

Контрольні запитання:

1. Необхідність сучасних пакетів прикладних програм Maple, Mathematica, MATLAB.
2. Порівняти можливості пакета Maple с можливостями Mathematica, MATLAB.
3. Особливості символьних і чисельних обчислень прикладних програм Maple, Mathematica, MATLAB.
4. Робота з рівняннями в чисельному і символьному вигляді у пакетах Maple, Mathematica, MATLAB.
5. Реалізація операцій лінійної алгебри у пакетах Maple, Mathematica, MATLAB.
6. Графічна візуалізація результатів обчислень у пакетах Maple, Mathematica, MATLAB. Роль чисельних методів при розв'язанні складних задач, що виникають при моделюванні економіко-математичних систем.

Література [9], [10].

Лекція 3. Загальна характеристика чисельних методів

Мета лекції: ознайомитися з основними поняттями теорії чисельних методів: похибка, коректність, стійкість.

План лекції:

1. Загальна характеристика чисельних методів.
2. Поняття коректно поставленої задачі.
3. Стійкість чисельних алгоритмів.
4. Типи похибок, які виникають при розв'язуванні прикладної задачі.

Нехай необхідно розв'язати рівняння

$$Au = f. \quad (1)$$

За рахунок неточно заданих вхідних даних насправді ми маємо рівняння

$$\tilde{A}u = \tilde{f}. \quad (2)$$

Назвемо $\delta_1 = u - \tilde{u}$ - неусувною похибкою.

Застосування методу розв'язання (2) приводить до рівняння

$$\tilde{A}_h \tilde{u}_h = \tilde{f}_h, \quad (3)$$

де $h > 0$ - малий параметр. Назвемо $\delta_2 = \tilde{u} - \tilde{u}_h$ - похибкою методу.

Реалізація методу на ЕОМ приводить до рівняння

$$\tilde{A}_h^* \tilde{u}_h^* = \tilde{f}_h^*. \quad (4)$$

Назвемо $\delta_3 = \tilde{u}_h - \tilde{u}_h^*$ - похибкою заокруглення.

Тоді повна похибка $\delta = u - \tilde{u}_h^* = \delta_1 + \delta_2 + \delta_3$.

Означення. Кажуть, що задача (1) коректна, якщо

$$1) \forall f \in F \quad \exists! u \in U ;$$

2) задача (1) стійка, тобто

$$\forall \varepsilon > 0 \exists \delta > 0: \|A - \tilde{A}\| < \delta, \|f - \tilde{f}\| < \delta \Rightarrow \|u - \tilde{u}\| < \varepsilon.$$

Якщо задача (1) некоректна, то або розв'язок її не існує, або він неєдиний, або він нестійкий, тобто

$$\exists \varepsilon > 0 \forall \delta > 0: \|A - \tilde{A}\| < \delta, \|f - \tilde{f}\| < \delta \Rightarrow \|u - \tilde{u}\| > \varepsilon.$$

Нехай число x задано своїм наближенням x^* .

Абсолютна похибка $\Delta x \leq |x - x^*|$.

Відносна похибка $\delta x \leq \frac{\Delta x}{|x|}$ або $\frac{\Delta x}{|x^*|}$.

Значущими цифрами називаються всі цифри, починаючи з першої ненульової зліва.

Вірна цифра – це значуща, якщо абсолютна похибка за рахунок відкидання всіх молодших розрядів не перевищує одиниці розряду цієї цифри. Тобто, якщо $x^* = \alpha_n \dots \alpha_0, \alpha_{-1} \dots \alpha_{-p} \dots$, то α_{-p} - вірна, якщо $\Delta x \leq 10^{-p}$ (інколи $\Delta x \leq w \cdot 10^{-p}$, $\frac{1}{2} \leq w < 1$, наприклад, $w = 0,55$).

Підрахунок похибок в ЕОМ

Підрахуємо відносно похибку заокруглення числа x на ЕОМ з плаваючою комою. В β -ічній системі числення число представляється у вигляді

$$x = \pm(\alpha_1\beta^{-1} + \alpha_2\beta^{-2} + \dots + \alpha_t\beta^{-t} + \dots)\beta^p, \quad (5)$$

де $0 \leq \alpha_k < \beta$, $\alpha_1 \neq 0$, $k = 1, 2, \dots$

Якщо в ЕОМ t розрядів, то при відкиданні молодших розрядів ми оперуємо з наближеним значенням

$$x^* = \pm(\alpha_1\beta^{-1} + \alpha_2\beta^{-2} + \dots + \alpha_t\beta^{-t})\beta^p$$

і відповідно похибка заокруглення $x - x^* = \pm\beta^p(\alpha_{t+1}\beta^{-t-1} + \dots)$. Тоді її можна оцінити так

$$|x - x^*| \leq \beta^{p-t-1}(\beta - 1)(1 + \beta^{-1} + \dots) \leq \beta^{p-t-1}(\beta - 1)\frac{1}{1 - \beta^{-1}} = \beta^{p-t}$$

Якщо в представленні (5) взяти $\alpha_1 = 1$, $\alpha_i = 0$, $i \geq 2$, то $|x| \geq \beta^p \beta^{-1}$. Звідси остаточно

$$\delta x \leq \frac{\beta^{p-t}}{\beta^{p-1}} = \beta^{-t+1}.$$

При більш точних способах заокруглення можна отримати оцінку $\delta x \leq \frac{1}{2}\beta^{-t+1} = \varepsilon$. Число ε називається "машинним іпсилон". Наприклад, для $\beta = 2$, $t = 24$, $\varepsilon = 2^{-24} \approx 10^{-7}$.

Підрахунок похибок обчислення значення функції

Нехай задана функція $y = f(x_1, \dots, x_n) \in C^1(\Omega)$. Необхідно обчислити її значення при наближеному значенні аргументів $\vec{x}^* = (x_1^*, \dots, x_n^*)$, де $|x_i - x_i^*| \leq \Delta x_i$ та оцінити похибку обчислення значення функції $y^* = f(x_1^*, \dots, x_n^*)$. Маємо

$$|y - y^*| = |f(\vec{x}) - f(\vec{x}^*)| = \left| \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\xi) (x_i - x_i^*) \right| \leq \sum_{i=1}^n B_i \cdot \Delta x_i, \text{ де } B_i = \max_{\vec{x} \in U} \left| \frac{\partial f}{\partial x_i}(\vec{x}) \right|.$$

Тут $U = \{\vec{x} = (x_1, \dots, x_n) : |x_i - x_i^*| \leq \Delta x_i\} \in \Omega$, $i = \overline{1, n}$. Отже з точністю до величин першого порядку малості по $\Delta x = \max_i \Delta x_i$

$$\Delta y = |y - y^*| \prec \sum_{i=1}^n b_i \cdot \Delta x_i,$$

де $b_i = \left| \frac{\partial f}{\partial x_i}(\vec{x}^*) \right|$ та " \prec " означає приблизно менше.

Розглянемо похибки арифметичних операцій.

1. Сума: $y = x_1 + x_2$, $x_1, x_2 > 0$,

$$\Delta y \leq \Delta x_1 + \Delta x_2, \quad \delta y \leq \frac{\Delta x_1 + \Delta x_2}{x_1 + x_2} \leq \max(\delta x_1, \delta x_2).$$

2. Різниця: $y = x_1 - x_2, \quad x_1 > x_2 > 0,$

$$\Delta y \leq \Delta x_1 + \Delta x_2, \quad \delta y \leq \frac{x_2 \delta x_1 + x_1 \delta x_2}{x_1 - x_2}.$$

При близьких x_1, x_2 зростає відносна похибка (за рахунок втрати вірних цифр).

3. Добуток: $y = x_1 \cdot x_2, \quad x_1, x_2 > 0,$

$$\Delta y < x_2 \Delta x_1 + x_1 \Delta x_2 \quad \delta y \leq \delta x_1 + \delta x_2.$$

4. Частка: $y = \frac{x_1}{x_2}, \quad x_1, x_2 > 0,$

$$\Delta y < \frac{x_2 \Delta x_1 + x_1 \Delta x_2}{(x_2)^2} \quad \delta y \leq \delta x_1 + \delta x_2.$$

При малих x_2 зростає абсолютна похибка (за рахунок зростання результату ділення).

Пряма задача аналізу похибок : обчислення $\Delta y, \delta y$ по заданих $\Delta x_i, i = \overline{1, n}$.

Обернена задача: знаходження $\Delta x_i, i = \overline{1, n}$ по заданих $\Delta y, \delta y$. Якщо $n > 1$, маємо одну умову $\sum_{i=1}^n b_i \Delta x_i < \varepsilon$ для багатьох невідомих Δx_i . Вибирають їх із однієї з умов

$$\forall i: b_i \Delta x_i < \frac{\varepsilon}{n} \quad \text{або} \quad \Delta x_i < \frac{\varepsilon}{\sum_{i=1}^n b_i}.$$

Контрольні запитання:

1. Об'єктивна обумовленість похибки чисельної оцінки показників.
2. Поняття про наближені числа.
3. Похибки в арифметичних обчисленнях. Джерела виникнення похибок.
4. Методи зменшення похибок обчислень.
5. Дії та операції над наближеними числами.
6. Статистичні методи оцінки похибок.
7. Середньоквадратичні похибки.
8. Поняття про стійкість, збіжність та коректність чисельних алгоритмів..

Література [4] - [8], [11].

Лекція 4. Чисельні методи розв'язування нелінійних рівнянь

Мета лекції: ознайомитися з найпростішими чисельними методами розв'язування нелінійних рівнянь (метод ділення навпіл, метод хорд та дотичних). Навчитися аналізувати складність цих методів.

План лекції:

1. Чисельні методи розв'язування нелінійних рівнянь.
2. Алгоритми ізолювання та уточнення коренів.
3. Графічний метод та метод табулювання.
4. Метод ділення навпіл.
5. Блок схема, оцінка кількості ітерацій.
6. Метод хорд та дотичних.

Постановка задачі. Нехай маємо рівняння $f(x) = 0$, \bar{x} - його розв'язок, тобто $f(\bar{x}) \equiv 0$.

Задача розв'язання цього рівняння розпадається на етапи:

1. Існування та кількість коренів.
2. Відділення коренів, тобто розбиття числової вісі на інтервали, де знаходиться один корінь.
3. Обчислення кореня із заданою точністю ε .

Для розв'язання перших двох задач використовуються методи математичного аналізу та алгебри, а також графічний метод. Далі розглядаються методи розв'язання третього етапу.

Метод ділення навпіл

Припустимо на $[a, b]$ знаходиться лише один корінь рівняння

$$f(x) = 0 \quad (1)$$

для $f(x) \in C[a, b]$, який необхідно визначити. Нехай $f(a)f(b) < 0$.

Припустимо, що $f(a) > 0, f(b) < 0$. Покладемо $x_1 = \frac{a+b}{2}$ і підрахуємо $f(x_1)$.

Якщо $f(x_1) < 0$, тоді шуканий корінь \bar{x} знаходиться на інтервалі (a, x_1) . Якщо ж $f(x_1) > 0$, то $\bar{x} \in (x_1, b)$. Далі з двох інтервалів (a, x_1) і (x_1, b) вибираємо той, на границях якого функція $f(x)$ має різні знаки, Позначимо його (a_1, b_1) . Знаходимо точку x_2 – середину вибраного інтервалу, підраховуємо $f(x_2)$ і повторюємо вказаний процес.

В результаті отримаємо послідовність інтервалів (a_k, b_k) , що містять шуканий корінь \bar{x} , причому довжина кожного послідуочого інтервалу вдвічі менше попереднього.

Цей процес продовжується доти, поки довжина отриманого інтервалу (a_n, b_n) не стане меншою за ε : $b_n - a_n < \varepsilon$. Тоді

$$|x_n - \bar{x}| < \varepsilon. \quad (2)$$

Ця умова для деякого n буде виконуватись за теоремою Больцано – Коші. Оскільки

$$|b_{k+1} - a_{k+1}| = \frac{1}{2}|b_k - a_k|,$$

то

$$|x_n - \bar{x}| \leq \frac{1}{2^n}(b - a) < \varepsilon. \quad (3)$$

Звідси отримаємо нерівність для обчислення кількості ітерацій n для виконання умови (2):

$$n = n(\varepsilon) \geq \left[\log_2 \left(\frac{b-a}{\varepsilon} \right) \right] + 1.$$

Степінь збіжності – лінійна, тобто геометричної прогресії з знаменником

$$q = \frac{1}{2}.$$

Переваги методу: простота, надійність. Недоліки методу: низька швидкість збіжності; метод не узагальнюється на системи.

Метод простої ітерації

Спочатку рівняння

$$f(x) = 0 \quad (1)$$

замінюється еквівалентним

$$x = \varphi(x). \quad (2)$$

Ітераційний процес має вигляд

$$x_{n+1} = \varphi(x_n), \quad n = 0, 1, \dots \quad (3)$$

Початкове наближення x_0 задається.

Для збіжності ітераційного процесу велике значення має вибір функції $\varphi(x)$. Перший спосіб заміни рівняння полягає в відділенні змінної з якогось члена рівняння. Більш продуктивним є перехід від рівняння (1) до (2) з функцією $\varphi(x) = x + \tau(x)f(x)$, де $\tau(x)$ – знакостала функція на тому відрізку, де шукаємо корінь.

Кажуть, що ітераційний метод *збігається*, якщо $\lim_{k \rightarrow \infty} x_k = \bar{x}$.

Далі $U_r = \{x : |x - x_0| \leq r\}$ відрізок довжини $2r$ з серединою в точці x_0 . З'ясуємо умови, при яких збігається метод простої ітерації.

Теорема 1

Якщо $\max_{x \in [a, b] = U_r} |\varphi'(x)| \leq q < 1$, то метод простої ітерації збігається і має місце оцінка

$$|x_n - \bar{x}| \leq \frac{q^n}{1-q} |x_0 - x_1| \leq \frac{q^n}{1-q} (b - a) \quad (3)$$

◁ Нехай $x_{k+1}, x_k \in U_r$. Тоді

$$|x_k - x_{k+1}| = |\varphi(x_k) - \varphi(x_{k-1})| = |\varphi'(\xi_k)(x_k - x_{k-1})| \leq$$

$$\begin{aligned}
\xi_k &= x_k + \theta_k (x_{k-1} - x_k), \quad 0 < \theta_k < 1 \\
&\leq |\varphi'(\xi_k)| \cdot |x_k - x_{k-1}| \leq q |x_k - x_{k-1}| = \dots = q^k |x_1 - x_0| \\
|x_{k+p} - x_k| &= |x_{k+p} - x_{k+p-1} + \dots + x_{k+1} - x_k| = |x_{k+p} - x_{k+p-1}| + \dots + |x_{k+1} - x_k| \leq \\
&\leq (q^{k+p-1} + q^{k+p-2} + \dots + q^k) |x_1 - x_0| = \frac{q^k - q^{k+p-1}}{1-q} |x_1 - x_0| \xrightarrow{k \rightarrow \infty} 0 \quad (4)
\end{aligned}$$

Таким чином $\{x_k\}$ – фундаментальна послідовність. Значить вона збіжна. При $p \rightarrow \infty$ в (4) отримуємо (3). \triangleright

Визначимо кількість ітерацій для досягнення точності ε . З оцінки (3) отримаємо

$$|x_n - \bar{x}| \leq \frac{q^n}{1-q} (b-a) < \varepsilon \Rightarrow n(\varepsilon) = n \geq \left\lceil \frac{\ln \frac{\varepsilon(1-q)}{b-a}}{\ln q} \right\rceil + 1.$$

Практично ітераційний процес зупиняємо при $|x_n - x_{n-1}| < \varepsilon$. Але ця умова не завжди гарантує, що $|x_n - \bar{x}| < \varepsilon$.

Зауваження Умова збіжності методу може бути замінена на умову Ліпшиця

$$|\varphi(x) - \varphi(y)| \leq q |x - y|, \quad 0 < q < 1.$$

Переваги методу: простота; при $q < \frac{1}{2}$ – швидше збігається ніж метод ділення навпіл; метод узагальнюється на системи. Недоліки методу: 1) при $q > \frac{1}{2}$ – збігається повільніше ніж метод ділення навпіл, 2) виникають труднощі при зведенні $f(x) = 0$ до $x = \varphi(x)$.

Метод релаксації

Якщо в методі простої ітерації для рівняння $x = x + \tau f(x) \equiv \varphi(x)$ вибрати $\tau(x) = \tau = const$, то ітераційний процес приймає вигляд

$$x_{n+1} = x_n + \tau f(x_n), \quad (1)$$

$k = 0, 1, 2, 3, \dots$ x_0 – задано. Метод можна записати у вигляді $\frac{x_{k+1} - x_k}{\tau} = f(x_k)$,

$k = 0, 1, \dots$. Оскільки $\varphi'(x) = 1 + \tau f'(x)$, то метод збігається при умові

$$|\varphi'(x)| = |1 + \tau f'(x)| \leq q < 1.$$

Нехай $f'(x) < 0$, тоді (3) запишеться у вигляді: $-q \leq 1 + \tau f'(x) \leq q < 1$. Звідси

$$\tau |f'(x_k)| \leq 1 + q < 2, \quad \text{і } 0 < \tau < \frac{2}{|f'(x)|}.$$

Поставимо задачу знаходження τ , для якого $q = q(\tau) \rightarrow \min$. Для того, щоб вибрати оптимальний параметр τ , розглянемо рівняння для похибки $z_k = x_k - \bar{x}$.

Підставивши $x_k = \bar{x} + z_k$ в (1), отримаємо

$$z_{k+1} = z_k + \tau f(\bar{x} + z_k).$$

В припущенні $f(x) \in C^1[a, b]$ з теореми про середнє маємо

$$f(\bar{x} + z_k) = f(\bar{x}) + z_k f'(\bar{x} + \theta z_k) = z_k f'(\bar{x} + \theta z_k) = z_k f'(\xi_k)$$

$$z_{k+1} = z_k + \tau f'(\xi_k) \cdot z_k$$

$$|z_{k+1}| \leq |1 + \tau f'(\xi_k)| \cdot |z_k| \leq \max_U |1 + \tau f'(\xi_k)| |z_k|$$

$$|z_{k+1}| \leq \max\{|1 - \tau M_1|, |1 - \tau m_1|\} |z_k|$$

$$m_1 = \min_{[a,b]} |f'(x)|, \quad M_1 = \max_{[a,b]} |f'(x)|$$

Таким чином, задача вибору оптимального параметра зводиться до знаходження τ , для якого функція

$$q(\tau) = \max\{|1 - \tau M_1|, |1 - \tau m_1|\}$$

приймає мінімальне значення: $q(\tau) \rightarrow \min$.

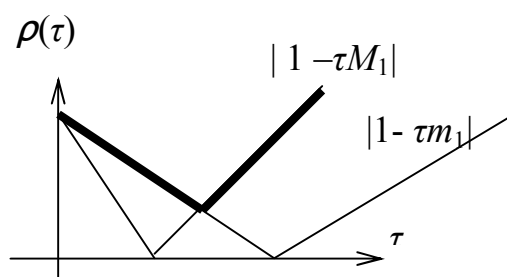


Рис. 1

З графіка видно, що точка мінімуму визначається умовою $|1 - \tau M_1| = |1 - \tau m_1|$. Тому

$$1 - \tau_0 m_1 = \tau_0 M_1 - 1 \Rightarrow \tau_0 = \frac{2}{M_1 + m_1} < \frac{2}{|f'(x)|}$$

При цьому значенні τ маємо

$$q(\tau_0) = \rho_0 = \frac{M_1 - m_1}{M_1 + m_1}.$$

Тоді для похибки вірна оцінка

$$|x_n - \bar{x}| \leq \frac{(\rho_0)^n}{1 - \rho_0} (b - a) < \varepsilon$$

Кількість ітерацій

$$n = n(\varepsilon) \geq \left[\frac{\ln \varepsilon (1 - \rho_0) / (b - a)}{\ln \rho_0} \right] + 1$$

Задача. Дати геометричну інтерпретацію методу простої ітерації для випадків:

$$0 < \varphi'(x) < 1; \quad -1 < \varphi'(x) < 0; \quad \varphi'(x) < -1; \quad \varphi'(x) > 1.$$

Задача. Знайти оптимальне $\tau = \tau_0$ для методу релаксації при $f'(x) > 0$.

Метод Ньютона (метод дотичних)

Припустимо, що рівняння $f(x) = 0$ має простий дійсний корінь \bar{x} , тобто $f(\bar{x}) = 0$, $f'(\bar{x}) \neq 0$. Нехай виконуються умови: $f(x) \in C^1[a, b]$, $f(a) \cdot f(b) < 0$. Тоді

$$0 = f(\bar{x}) = f(x_k + \bar{x} - x_k) = f(x_k) + f'(\xi_k)(\bar{x} - x_k),$$

де $\xi_k = x_k + \theta_k(\bar{x} - x_k)$, $0 < \theta_k < 1$, $\xi_k \approx x_k$. Тому наступне наближення виберемо з рівняння

$$f(x_k) + f'(x_k)(x_{k+1} - x_k) = 0.$$

Звідси маємо ітераційний процес

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, 2, \dots; \quad x_0 \text{- задане.}$$

Метод Ньютона ще називають методом лінеаризації або методом дотичних.

Задача. Дати геометричну інтерпретацію методу Ньютона.

Метод Ньютона можна інтерпретувати як метод простої ітерації з

$$\varphi(x) = x - \frac{f(x)}{f'(x)}, \quad \text{тобто } \tau(x) = -\frac{1}{f'(x)}.$$

Тому $\varphi'(x) = 1 - \frac{[f'(x)f'(x) - f(x)f''(x)]}{(f'(x))^2} = \frac{f(x)f''(x)}{(f'(x))^2}$. Якщо \bar{x} - корінь $f(x)$,

то $\varphi'(\bar{x}) = 0$. Тому знайдеться окіл кореня, де

$$|\varphi'(x)| = \left| \frac{f(x)f''(x)}{(f'(x))^2} \right| < 1.$$

Це означає, що збіжність методу Ньютона залежить від вибору x_0 .

Недолік методу Ньютона: необхідність обчислювати на кожній ітерації не тільки значення функції, а й похідної.

Модифікований метод Ньютона позбавлений цього недоліку і має вигляд:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_0)}, \quad k=0, 1, 2, \dots$$

Цей метод має лише лінійну збіжність: $|x_{k+1} - \bar{x}| = O(|x_k - \bar{x}|)$.

Задача. Дати геометричну інтерпретацію модифікованого методу Ньютона.

В методі Ньютона, для якого $f'(x_k)$ замінюється на $\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$

дає метод січних:

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k) \quad k = 1, 2, \dots; \quad x_0, x_1 \text{- задані}$$

Задача. Дати геометричну інтерпретацію методу січних.

Збіжність методу Ньютона

Теорема 1 Нехай $f(x) \in C^2[a, b]$; \bar{x} простий дійсний корінь рівняння

$$f(x) = 0 \quad (1)$$

і $f'(x) \neq 0$ при $x \in U_r = \{x : |x - \bar{x}| < r\}$. Якщо

$$q = \frac{M_2 |x_0 - \bar{x}|}{2m_1} < 1, \quad (2)$$

де $m_1 = \min_{U_r} |f'(x)|$, $M_2 = \max_{U_r} |f''(x)|$, то для $x_0 \in U_r$ метод Ньютона

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \quad (3)$$

збігається і має місце оцінка

$$|x_n - \bar{x}| \leq q^{2^n - 1} |x_0 - \bar{x}|. \quad (4)$$

з (3) маємо

$$x_{k+1} - \bar{x} = x_k - \frac{f(x_k)}{f'(x_k)} - \bar{x} = \frac{(x_k - \bar{x})f'(x_k) - f(x_k)}{f'(x_k)} = \frac{F(x_k)}{f'(x_k)}, \quad (5)$$

де $F(x) = (x - \bar{x})f'(x) - f(x)$, така, що

$$1) F(\bar{x}) = 0; \quad 2) F'(x) = (x - \bar{x})f''(x).$$

Тоді

$$F(x_k) = F(\bar{x}) + \int_{\bar{x}}^{x_k} F'(t) dt = \int_{\bar{x}}^{x_k} (t - \bar{x}) f''(t) dt.$$

Так як $(t - \bar{x})$ не міняє знак на відрізку інтегрування, то скористаємося теоремою про середнє значення:

$$F(x_k) = f''(\xi_k) \int_{\bar{x}}^{x_k} (t - \bar{x}) dt = \frac{(x_k - \bar{x})^2}{2} f''(\xi_k), \quad (6)$$

де $\xi_k = \bar{x} + \theta_k(x_k - \bar{x})$, $0 < \theta_k < 1$. З (5), (6) маємо

$$x_{k+1} - \bar{x} = \frac{(x_k - \bar{x})^2}{2f'(x_k)} f''(\xi_k) \quad (7)$$

Доведемо оцінку (3) за індукцією. Так як $x_0 \in U_r$, то

$$|\xi_0 - \bar{x}| = |\theta_0(x_0 - \bar{x})| < |\theta_0| |x_0 - \bar{x}| < r \Rightarrow \xi_0 \in U_r.$$

Тоді $|f''(\xi_0)| \leq M_2$, тому

$$|x_1 - \bar{x}| \leq \frac{(x_0 - \bar{x})^2 M_2}{2m_1} = \frac{M_2 |x_0 - \bar{x}|}{2m_1} |x_0 - \bar{x}| = q |x_0 - \bar{x}| < r, x_1 \in U_r,$$

Ми довели твердження (4) при $n = 1$. Нехай воно справджується при $n = k$

$$|x_k - \bar{x}| \leq q^{2^k - 1} |x_0 - \bar{x}| < r, |\xi_k - \bar{x}| = |\theta_k(x_k - \bar{x})| < r.$$

Тоді $x_k, \xi_k \in U_r$.

Доведемо (4) для $n = k + 1$. З (7) маємо

$$\begin{aligned}
|x_{k+1} - \bar{x}| &\leq \frac{|x_k - \bar{x}|^2 M_2}{2m_1} \leq (q^{2^{k-1}})^2 \frac{|x_0 - \bar{x}|^2 M_2}{2m_1} = \\
&= q^{2^{k+1}-2} \frac{|x_0 - \bar{x}| M_2}{2m_1} |x_0 - \bar{x}| = q^{2^{k+1}-1} |x_0 - \bar{x}|
\end{aligned}$$

Таким чином (4) справджується для $n = k + 1$. Значить (4) виконується і для довільного n . Таким чином $x_n \xrightarrow{n \rightarrow \infty} x$.

З (4) маємо оцінку кількості ітерацій для досягнення точності ε

$$n \geq \left\lceil \log_2 \left(1 + \frac{\ln \frac{\varepsilon}{b-a}}{\ln q} \right) \right\rceil + 1.$$

Кажуть, що ітераційний метод має степінь збіжності m , якщо

$$|x_{k+1} - \bar{x}| = O(|x_k - \bar{x}|^m).$$

Для методу Ньютона $|x_{k+1} - \bar{x}| = \frac{|x_k - \bar{x}|^2 |f''(\xi_k)|}{2|f'(x_k)|} \Rightarrow |x_{k+1} - \bar{x}| = O(|x_k - \bar{x}|^2)$.

Значить степінь збіжності методу Ньютона $m=2$. Для методу простої ітерації і ділення навпіл $m=1$.

Теорема 2 Нехай $f(x) \in C^2[a, b]$ та \bar{x} простий корінь рівняння $f(x) = 0$ ($f'(\bar{x}) \neq 0$) $\forall x \in [a, b]$. Якщо $f'(x)f''(x) > 0$ ($f'(x)f''(x) < 0$) то для методу Ньютона при $x_0 = b$ послідовність наближень $\{x_k\}$ монотонно спадає (монотонно зростає при $x_0 = a$).

∇ Нехай

$$f'(x) > 0, f''(x) > 0. \quad (8)$$

Доведемо твердження теореми методом мат. індукції. Нехай $x_* < x_k \leq b$. З формули (3)

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

За теоремою про середнє

$$x_k - x_{k+1} = \frac{f'(\xi_k)}{f'(x_k)} (x_k - x_*), \quad \xi_k \in (x_*, x_k).$$

В силу (8) $f'(x)$ зростаюча функція, тому $0 < \frac{f'(\xi_k)}{f'(x_k)} < 1$. Звідси

$0 < x_k - x_{k+1} < (x_k - x_*)$ і значить $x_k < x_{k+1} \leq x_*$. Аналогічно доводяться твердження для $f'(x) < 0, f''(x) < 0$ та $f'(x)f''(x) < 0$. Δ

Задача. Довести теорему 2 при а) $f'(x)f''(x) > 0$, б) $f'(x)f''(x) < 0$.

Задача. Знайти степінь збіжності методу січних.

Зауваження 1 Якщо \bar{x} r -кратний корінь тобто

$$f^{(m)}(\bar{x}) = 0, m = 0, 1, \dots, p-1; f^{(p)}(\bar{x}) \neq 0,$$

то в методі Ньютона необхідна наступна модифікація

$$x_{k+1} = x_k - p \frac{f(x_k)}{f'(x_k)} \text{ і } q = \frac{M_{p+1}|x_0 - \bar{x}|}{m_p p(p+1)} < 1.$$

Зауваження 2 Метод Ньютона можна застосовувати і для обчислення комплексного кореня $z_{k+1} = z_k - \frac{f(z_k)}{f'(z_k)}$. В теоремі про збіжність

$$q = \frac{|x_0 - \bar{x}| M_2}{2m_1}, \text{ де } m_1 = \min_{U_r} |f'(z)|, M_2 = \max_{U_r} |f''(z)|. \text{ Тут } |z| - \text{ модуль комплексного числа.}$$

Геометричний зміст методу Ньютона.

В методі рівняння $f(x) = 0$ заміняється лінійним

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

Точка перетину цієї прямої з віссю ox і дає значення x_{k+1} . а

Якщо $f(a)f''(a) > 0$ та $f''(x)$ не міняє знак, то потрібно вибирати $x_0 = a$; при цьому $\{x_k\} \uparrow \bar{x}$.

Якщо $f(b)f''(b) > 0$, то $x_0 = b$; маємо $\{x_k\} \downarrow \bar{x}$. Пояснення на рисунку 2.

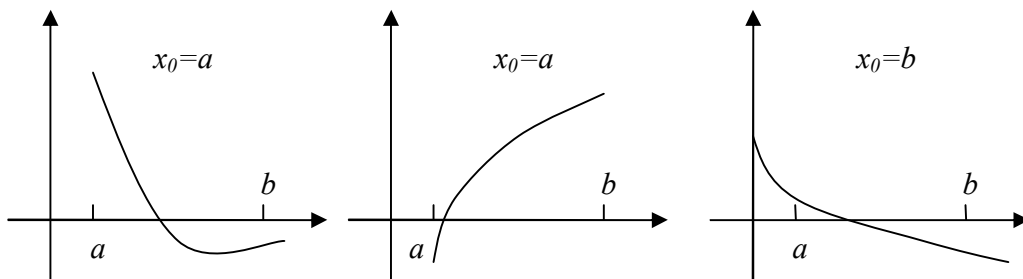


Рис. 2

Зауваження 1 Якщо \bar{x} p -кратний корінь тобто

$$f^{(m)}(\bar{x}) = 0, m = 0, 1, \dots, p-1; f^{(p)}(\bar{x}) \neq 0,$$

то в методі Ньютона необхідна наступна модифікація

$$x_{k+1} = x_k - p \frac{f(x_k)}{f'(x_k)} \text{ і } q = \frac{M_{p+1}|x_0 - \bar{x}|}{m_p p(p+1)} < 1.$$

Зауваження 2 Метод Ньютона можна застосовувати і для обчислення комплексного кореня $z_{k+1} = z_k - \frac{f(z_k)}{f'(z_k)}$. В теоремі про збіжність

$q = \frac{|x_0 - \bar{x}|M_2}{2m_1}$, де $m_1 = \min_{U_r} |f'(z)|$, $M_2 = \max_{U_r} |f''(z)|$. Тут $|z|$ - модуль комплексного числа.

Переваги методу Ньютона: 1) висока швидкість збіжності; 2) узагальнюється на системи рівнянь; 3) узагальнюється на комплексні корені.

Недоліки методу Ньютона: 1) на кожній ітерації обчислюється не тільки $f(x_k)$, а і похідна $f'(x_k)$; 2) збіжність залежить від початкового наближення x_0 , так як від нього залежить умова збіжності $q = \frac{M_2|x_0 - \bar{x}|}{2m_1} < 1$;

3) потрібно, щоб $f(x) \in C^2[a, b]$.

Контрольні запитання:

1. Етапи розв'язання нелінійного рівняння.
2. Швидкість збіжності метода ділення навпіл.
3. Метод простої ітерації.
4. Метод релаксації.
5. Метод Ньютона (метод дотичних).
6. Геометричний зміст методу Ньютона.

Література [3], [4], [6], [11].

Лекція 5. Чисельні методи розв'язування систем лінійних алгебраїчних рівнянь

Мета лекції: ознайомитися з чисельними методами розв'язування систем лінійних алгебраїчних рівнянь, навчитися аналізувати їх складність.

План лекції:

1. Чисельні методи розв'язування систем лінійних алгебраїчних рівнянь.
2. Метод Гауса.
3. Прямий та обернений хід методу Гауса.
4. Вибір головного елемента.
5. Оцінка кількості арифметичних дій методу Гауса.
6. Блок-схема методу Гауса.
7. Метод квадратних коренів.
8. Обумовленість систем лінійних алгебраїчних рівнянь.

Методи розв'язування СЛАР поділяються на прямі та ітераційні. При умові точного виконання обчислень прямі методи за скінчену кількість операцій в результаті дають точний розв'язок. Використовуються вони для невеликих та середніх СЛАР $n=10^2-10^4$. Ітераційні методи використовуються для великих СЛАР $n>10^5$, як правило розріджених. В результаті отримуємо послідовність наближень, яка збігається до розв'язку.

Метод Гаусса

Розглянемо задачу розв'язання СЛАР

$$A\vec{x} = \vec{b}, \quad (1)$$

причому $A = (a_{ij})_{i,j=1}^n$, $\det A \neq 0$, $\vec{x} = (x_i)_{i=1}^n$, $\vec{b} = (b_j)_{j=1}^n$. Метод Крамера з обчисленням визначників для такої системи має складність $Q = O(n!n)$.

Запишемо СЛАР у вигляді

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

Якщо $a_{11} \neq 0$, то ділимо перше рівняння на нього і виключаємо x_1 з інших рівнянь:

$$\begin{cases} x_1 + a_{12}^{(1)}x_2 + \dots + a_{1n}^{(1)}x_n = a_{1n+1}^{(1)} \\ a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n = a_{2n+1}^{(1)} \\ \dots \\ a_{n2}^{(1)}x_2 + \dots + a_{nn}^{(1)}x_n = a_{nn+1}^{(1)} \end{cases}$$

Процес повторюємо для x_2, \dots, x_n . В результаті отримуємо систему з трикутною матрицею

$$\begin{cases} x_1 + a_{12}^{(1)}x_2 + \dots + a_{1n}^{(1)}x_n = a_{1n+1}^{(1)} \\ x_2 + \dots + a_{2n}^{(2)}x_n = a_{2n+1}^{(2)} \\ \dots \\ x_n = a_{nn+1}^{(n)} \end{cases} \quad (2)$$

Це прямий хід методу Гаусса. Формули прямого ходу

$$\begin{cases} k = \overline{1, n-1}: \\ a_{kj}^{(k)} = \frac{a_{kj}^{(k-1)}}{a_{kk}^{(k-1)}}, j = \overline{k+1, n+1}; \\ a_{ij}^{(k)} = a_{ij}^{(k-1)} - a_{ik}^{(k-1)}a_{kj}^{(k)}, \\ i = \overline{k+1, n}; j = \overline{k+1, n+1}. \end{cases}$$

Звідси

$$x_n = a_{nn+1}^{(n)}, x_i = a_{in+1}^{(i)} - \sum_{j=i+1}^n a_{ij}^{(i)}x_j, i = \overline{n-1, 1}. \quad (3)$$

Це формули оберненого ходу.

Складність, тобто кількість операцій, яку необхідно виконати для реалізації методу, - $Q_{np} = \frac{2}{3}n^3 + O(n^2)$ для прямого ходу, $Q_{ob} = n^2 + O(n)$ для оберненого ходу.

Умова $a_{kk}^{(k-1)} \neq 0$ не суттєва, оскільки знайдеться m , для якого $|a_{mk}^{(k-1)}| = \max_i |a_{ik}^{(k-1)}| \neq 0$ (оскільки $\det A \neq 0$). Тоді міняємо місцями рядки номерів k і m . Елемент $a_{kk}^{(k-1)} \neq 0$ називається ведучим.

Введемо матриці

$$M_k = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ 0 & & m_{kk} & \\ & & \vdots & \ddots \\ 0 & & m_{nk} & 1 \end{pmatrix},$$

елементи якої обчислюється так: $m_{kk} = \frac{1}{a_{kk}^{(k-1)}}$, $m_{nk} = -\frac{a_{nk}^{(k-1)}}{a_{kk}^{(k-1)}}$.

Нехай на k -му кроці маємо систему $A_{k-1}\vec{x} = \vec{b}_{k-1}$. Множимо цю СЛАР зліва на M_k : $M_k A_{k-1}\vec{x} = M_k \vec{b}_{k-1}$. Позначимо $A_k = M_k A_{k-1}$; $A_0 = A$. Тоді прямий хід методу Гаусса можна записати у вигляді

$$M_n M_{n-1} \dots M_1 A \vec{x} = M_n M_{n-1} \dots M_1 \vec{b}.$$

Позначимо останню систему, яка співпадає з (2), так

$$U\vec{x} = \vec{c}, \quad U = (u_{ij})_{i,j=1}^n, \quad (3)$$

причому

$$\begin{cases} u_{ii} = 1, \\ u_{ij} = 0, i > j. \end{cases}$$

Таким чином $U = M_n M_{n-1} \dots M_1 A$. Введемо матриці

$$L_k = M_k^{-1} = \begin{pmatrix} 1 & & 0 & 0 \\ & \ddots & & \\ 0 & & a_{kk}^{(k-1)} & 0 \\ & & \vdots & \ddots \\ 0 & & a_{nk}^{(k-1)} & 1 \end{pmatrix}.$$

Тоді

$$A = L_1 \dots L_n U = LU; \quad L = L_1 \dots L_n$$

L - нижня трикутня матриця, U - верхня трикутня матриця. Таким чином метод Гаусса можна трактувати, як розклад матриці A в добуток двох трикутних матриць - (LU) – розклад

Введемо матрицю перестановок на k -му кроці (це матриця, отримана з одиничної матриці перестановкою k - того і m - того рядка). Тоді при множенні на неї матриці A_{k-1} робимо ведучим елементом максимальний за модулем.

$$P_k = \begin{pmatrix} 1 & \dots & 0 & & \\ \cdot & \cdot & \cdot & & \\ \cdot & & 0 & 1 & \\ \cdot & & 1 & 0 & \\ 0 & \cdot & \cdot & \cdot & 1 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow k \\ \leftarrow m \\ \\ \end{matrix}$$

$$\begin{matrix} \uparrow & \uparrow \\ k & m \end{matrix}$$

За допомогою цих матриць перехід до трикутної системи (3) тепер має вигляд:

$$M_n M_{n-1} P_{n-1} \dots M_1 P_1 A \vec{x} = M_n M_{n-1} P_{n-1} \dots M_1 P_1 \vec{b}.$$

Твердження

Знайдеться така матриця P - перестановок, що $PA = LU$ - розклад матриці на нижню трикутну з ненульовими діагональними елементами і верхню трикутну матрицю з одиницями на діагоналі.

Висновки про переваги трикутного розкладу:

1. Розділення прямого і оберненого ходів дає змогу економно розв'язувати декілька систем з одноковою матрицею та різними правими частинами.
2. Зберігання M , або L та U на місці A .
3. Обчислюючи l - кількість перестановок, можна встановити знак визначника.

Метод квадратних коренів

Цей метод призначений для розв'язання систем рівнянь із симетричною матрицею

$$A\vec{x} = \vec{b}, A^T = A. \quad (1)$$

Він оснований на розкладі матриці A в добуток:

$$A = S^T D S, \quad (2)$$

S – верхня трикутна матриця, S^T – нижня трикутна матриця, D – діагональна матриця.

Виникає питання: як обчислити S , D по матриці A ? Маємо

$$(DS)_{ij} = \begin{cases} d_{ii} s_{ij}, & i \leq j \\ 0, & i > j \end{cases}$$

$$(S^T DS)_{ij} = \sum_{l=1}^n s_{il}^T d_{ll} s_{lj} = \sum_{l=1}^{i-1} s_{li} s_{lj} d_{ll} + s_{ii} s_{ij} d_{ii} + \underbrace{s_{li}^T \sum_{l=i+1}^n s_{li} s_{lj} d_{ll}}_{=0} = a_{ij}, i, j = \overline{1, n} \quad (3)$$

Якщо $i = j$, то

$$|s_{ii}^2| d_{ii} = a_{ii} - \sum_{l=1}^{i-1} |s_{li}^2| d_{ll} \equiv p_i.$$

Тому

$$d_{ii} = \text{sign}(p_i) \quad , \quad s_{ii} = \sqrt{|p_i|} .$$

Якщо $i < j$, то

$$s_{ij} = \left(a_{ij} - \sum_{l=1}^{i-1} s_{li} d_{ll} s_{lj} \right) / (s_{ii} d_{ii}), \quad i = \overline{1, n}, j = \overline{i+1, n} .$$

Якщо $A > 0$ (тобто головні мінори матриці A додатні), то всі $d_{ii} = +1$.

Знайдемо розв'язок рівняння (1). Враховуючи (2), маємо:

$$S^T D \vec{y} = \vec{b} \quad , \quad (4)$$

$$S \vec{x} = \vec{y} . \quad (5)$$

Оскільки S – верхня трикутна матриця, а $S^T D$ – нижня трикутна матриця, то

$$y_i = \frac{b_i - \sum_{j=1}^{i-1} s_{ji} d_{jj} y_j}{s_{ii} d_{ii}}, \quad i = \overline{1, n} \quad (6)$$

$$x_n = \frac{y_n}{s_{nn}}, \quad x_i = \frac{y_i - \sum_{j=1}^{i-1} s_{ij} x_j}{s_{ii}}, \quad i = \overline{n-1, 1}. \quad (7)$$

Метод застосовується лише для симетричних матриць. Його складність -

$Q = \frac{1}{3} n^3 + O(n^2)$. Переваги цього методу:

1. він витрачає в 2 рази менше пам'яті ніж метод Гаусса для зберігання

$A^T = A$ (необхідний об'єм пам'яті $\frac{n(n+1)}{2} \sim \frac{n^2}{2}$);

2. метод однорідний, без перестановок;

3. якщо матриця A має багато нульових елементів, то і в матриця S також.

Остання властивість дає економію в пам'яті та кількості арифметичних операцій. Наприклад, якщо A має m ненульових стрічок по діагоналі, то

$Q = O(m^2 n)$.

Обчислення визначника та оберненої матриці

Кількість операцій обчислення детермінанту за означенням - $Q_{\det} = n!$. В методі Гаусса - $PA = LU$. Тому

$$\det P \det A = \det L \det U \Rightarrow \det A = (-1)^l \det L \det U = (-1)^l \prod_{k=1}^n a_{kk}^{(k)}, \quad (1)$$

де l - кількість перестановок. Ясно, що за методом Гаусса

$$Q_{\det} = \frac{2}{3} n^3 + O(n^2) .$$

В методі квадратного кореня $A = S^T D S$. Тому

$$\det A = \det S^T \det D \det S = \prod_{k=1}^n d_{kk} \prod_{k=1}^n s_{kk}^2. \quad (2)$$

Тепер $Q_{\det} = \frac{1}{3}n^3 + O(n^2)$.

За означенням

$$AA^{-1} = E, \quad (3)$$

де A^{-1} обернена до матриці A . Позначимо

$$A^{-1} = (\alpha_{ij})_{i,j=1}^n.$$

Тоді $\vec{\alpha}_j = (\alpha_{ij})_{i=1}^n$ - вектор-стовпчик оберненої матриці. З (3) маємо

$$A\vec{\alpha}_j = \vec{e}_j, j = \overline{1, n} \quad (4)$$

\vec{e}_j - стовпчики одиничної матриці: $\vec{e}_j = (\delta_{ij})_{i=1}^n, \delta_{ij} = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$. Для знаходження A^{-1} необхідно розв'язати n систем. Для знаходження A^{-1} методом Гаусса необхідна кількість операцій $Q = 2n^3 + O(n^2)$.

Метод прогонки

Це економний метод для розв'язання СЛАР з три діагональною матрицею:

$$\begin{cases} -c_0 y_0 + b_0 y_1 = -f_0, \\ a_i y_{i-1} - c_i y_i + b_i y_{i+1} = -f_i, \\ a_N y_{N-1} - c_N y_N = -f_N. \end{cases}$$

(1)

(2)

(3)

Матриця системи

$$A = \begin{pmatrix} -c_0 & b_1 & & 0 \\ a_0 & -c_1 & \ddots & \\ & \ddots & \ddots & b_N \\ 0 & & a_N & -c_N \end{pmatrix}$$

тридіагональна.

Розв'язок представимо у вигляді

$$y_i = \alpha_{i+1} y_{i+1} + \beta_{i+1}, i = \overline{0, N-1} \quad (4)$$

Замінімо в (4) $i \rightarrow i-1$ і підставимо в (2), тоді

$$(a_i \alpha_i - c_i) y_i + b_i y_{i+1} = -f_i - a_i \beta_i$$

Звідси

$$y_i = \frac{b_i}{c_i - a_i \alpha_i} y_{i+1} + \frac{f_i + a_i \beta_i}{c_i - a_i \alpha_i}.$$

Тому з (5)

$$\alpha_{i+1} = \frac{b_i}{c_i - a_i \alpha_i}, \beta_{i+1} = \frac{f_i + a_i \beta_i}{c_i - a_i \alpha_i}, i = \overline{1, N-1}.$$

Умова розв'язності (1) $c_i - a_i \alpha_i \neq 0$.

Щоб знайти всі α_i, β_i , треба задати перші значення. З (1):

$$\alpha_1 = \frac{b_0}{c_0}, \beta_1 = \frac{f_0}{c_0} \quad (5)$$

Після знаходження всіх α_i, β_i обчислюємо y_N з системи

$$\begin{cases} a_N y_N - c_N y_N = -f_N \\ y_{N-1} = \alpha_N y_N + \beta_N \end{cases}.$$

Звідси

$$y_N = \frac{f_N + a_N \beta_N}{c_N - a_N \alpha_N}. \quad (6)$$

Алгоритм

1. Покладемо $\alpha_1 = \frac{b_0}{c_0}, \beta_1 = \frac{f_0}{c_0}$.
2. Позначимо $z_i = c_i - a_i \alpha_i$, обчислимо $\alpha_{i+1} = \frac{b_i}{z_i}, \beta_{i+1} = \frac{f_i + a_i \beta_i}{z_i}$, для $i = \overline{1, N-1}$
3. Знайдемо $y_N = \frac{f_N + a_N \beta_N}{c_N - a_N \alpha_N}$.
4. Обчислюємо $y_i = \alpha_{i+1} y_{i+1} + \beta_{i+1}, i = \overline{N-1, 0}$.

Складність алгоритму $Q = 8N - 2$.

Метод можна застосовувати, коли $c_i - a_i \alpha_i \neq 0 \forall i, |\alpha_i| \leq 1$. Якщо $|\alpha_i| \geq q > 1$, то $\Delta y_0 \geq q^N \Delta y_N$ (тут Δy_i абсолютна похибка обчислення y_i), а це приводить до експоненціального накопичення похибок заокруглення, тобто нестійкості алгоритму прогонки.

Теорема (про достатні умови стійкості метода прогонки).

Нехай

$$a_i, b_i \neq 0, \text{ та } |c_i| \geq |a_i| + |b_i|, \forall i, (a_0 = b_N = 0)$$

та хоча би одна нерівність строга. Тоді

$$|\alpha_i| \leq 1, \text{ та } z_i = c_i - a_i \alpha_i \neq 0, i = \overline{1, N}.$$

Задача. Довести теорему про стійкість методу прогонки.

Обумовленість систем лінійних алгебраїчних рівнянь

Нехай задано СЛАР

$$A\vec{x} = \vec{b} \quad (1)$$

Припустимо, що матриця і права частина системи задані неточно і фактично розв'язуємо систему

$$B\bar{y} = \bar{h} \quad (2)$$

де $B = A + C$, $\bar{h} = \bar{b} + \bar{\eta}$, $\bar{y} = \bar{x} + \bar{z}$.

Малість детермінанту $|\det A| \ll 1$ не є необхідною умовою різкого збільшення похибки. Це ілюструє наступний приклад:

$$A = \text{diag}(\varepsilon), a_{ij} = \varepsilon \delta_{ij},$$

Тоді $\det A = \varepsilon^n \ll 1$, але $x_i = \frac{b_i}{\varepsilon}$. Тому $\Delta x_i = \frac{\Delta b_i}{\varepsilon}$.

Оцінимо похибку розв'язку (2). Підставивши значення B, \bar{h} та $\bar{z} = \bar{y} - \bar{x}$, отримуємо:

$$(A + C)(\bar{x} + \bar{z}) = \bar{b} + \bar{\eta}.$$

Віднімемо від цієї рівності (1) $A\bar{z} + C\bar{x} + C\bar{z} = \bar{\eta}$. Тоді

$$A\bar{z} = \bar{\eta} - C\bar{x} - C\bar{z}, \bar{z} = A^{-1}(\bar{\eta} - C\bar{x} - C\bar{z}).$$

Введемо норми векторів: $\|\bar{z}\|$:

$$\|\bar{z}\|_1 = \sum_{i=1}^n |z_i|, \quad \|\bar{z}\|_2 = \left(\sum_{i=1}^n |z_i|^2\right)^{\frac{1}{2}}, \quad \|\bar{z}\|_\infty = \max_i |z_i|.$$

Норми матриці, що відповідають нормам вектора, тобто

$$\|A\|_m = \sup_{\|\bar{x}\|_m \neq 0} \frac{\|A\bar{x}\|_m}{\|\bar{x}\|_m}, m = 1, 2, \infty,$$

такі: $\|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|$, $\|A\|_2 = \max_i \sqrt{\lambda_i(A^T A)}$, $\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$, де

$\lambda_i(B)$ – власні значення матриці B .

Позначимо $\delta(\bar{x}) = \frac{\|\bar{z}\|}{\|\bar{x}\|}$, $\delta(\bar{b}) = \frac{\|\bar{\eta}\|}{\|\bar{b}\|}$, $\delta(A) = \frac{\|C\|}{\|A\|}$ – відносні похибки \bar{x}, \bar{b}, A ,

де $\|\cdot\| = \|\cdot\|_k$ одна з введених вище норм.

Для характеристики зв'язку між похибками правої частини та розв'язку вводять поняття обумовленості матриці системи.

Число обумовленості матриці A - $\text{cond}(A) = \|A\| \|A^{-1}\|$.

Теорема Якщо $\exists A^{-1}$ та $\|A^{-1}\| \|C\| < 1$, то

$$\delta(\bar{x}) \leq \frac{\text{cond} A}{1 - \text{cond} A \delta(A)} (\delta(A) + \delta(\bar{b})), \quad (3)$$

де $\text{cond} A$ - число обумовленості.

$\triangleleft A\bar{z} = \bar{\eta} - C\bar{x} - C\bar{z}$, $\bar{z} = A^{-1}\bar{\eta} - A^{-1}C\bar{x} - A^{-1}C\bar{z}$;

$$\begin{aligned}\|\bar{z}\| &\leq \|A^{-1}\bar{\eta}\| + \|A^{-1}C\bar{x}\| + \|A^{-1}C\bar{z}\| \leq \|A^{-1}\|\bar{\eta}\| + \|A^{-1}\| \|C\|\bar{x}\| + \|A^{-1}\| \|C\|\bar{z}\| \\ \|\bar{z}\| &\leq \frac{\|A^{-1}\|(\bar{\eta} + \|C\|\bar{x}\|)}{1 - \|A^{-1}\| \|C\|}\end{aligned}$$

Оцінка похибки

$$\begin{aligned}\delta(\bar{x}) &\leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \|C\|} \left(\frac{\|\bar{\eta}\|}{\|\bar{x}\|} + \|C\| \right) = \frac{\|A^{-1}\| \|A\|}{1 - \|A^{-1}\| \|A\| \frac{\|C\|}{\|A\|}} \left(\frac{\|\bar{\eta}\|}{\|A\| \|\bar{x}\|} + \delta(A) \right) \leq \\ &\leq \frac{\text{cond}A}{1 - \text{cond}A \delta(A)} \left(\frac{\|\bar{\eta}\|}{\|\bar{x}\|} + \delta(A) \right). \triangleright\end{aligned}$$

Наслідок Якщо $C \equiv 0$, то $\delta(\bar{x}) \leq \text{cond}A \delta(\bar{b})$.

Властивості $\text{cond}A$:

1^o. $\text{cond}A \geq 1$;

2^o. $\text{cond}A \geq \frac{\max_i |\lambda_i(A)|}{\min_i |\lambda_i(A)|}$;

3^o. $\text{cond}(AB) \leq \text{cond}A * \text{cond}B$;

4^o. $A^T = A^{-1} \Rightarrow \text{cond}A = 1$.

Друга властивість має місце оскільки довільна норма матриці не менше її найбільшого за модулем власного значення. Значить $\|A\| \geq \max |\lambda_A|$.

Оскільки власні значення матриць A^{-1} та A взаємно обернені, то

$$\|A^{-1}\| \geq \max \frac{1}{|\lambda_A|} = \frac{1}{\min |\lambda_A|}.$$

Якщо $\text{cond}A \gg 1$, то система називається *погано обумовленою*.

Оцінка впливу похибок заокруглення при обчисленні розв'язку СЛАР така (Дж. Уілкінсон): $\delta(A) = O(n\beta^{-t})$, $\delta(\bar{b}) = O(\beta^{-t})$, де β - розрядність ЕОМ, t - кількість розрядів, що відводиться під мантису числа. З оцінки (3) витікає: $\delta(\bar{x}) = \text{cond}A \times O(n\beta^{-t})$. Висновок: найпростіший спосіб підвищити точність обчислення розв'язку погано обумовленої СЛАР – збільшити розрядність ЕОМ при обчисленнях. Інші способи пов'язані з розглядом цієї СЛАР як некоректної задачі із застосуванням відповідних методів її розв'язання.

Приклад погано обумовленої системи – системи з матрицею Гільберта

$$H_n = \left(\frac{1}{i+j-1} \right)_{i,j=1}^n, \text{ наприклад } \text{cond}H_8 \approx 10^9.$$

Контрольні запитання:

1. Необхідність чисельних методів розв'язування систем лінійних алгебраїчних рівнянь.
2. Алгоритм методу Гауса.
3. Оцінка кількості арифметичних дій методу Гауса.
4. Формули методу квадратних коренів.
5. Методи обчислення визначника та оберненої матриці.
6. Метод прогонки.
7. Поняття обумовленості систем лінійних алгебраїчних рівнянь.

Література [1], [4], [6], [11].

Лекція 6. Метод ітерацій розв'язування систем лінійних алгебраїчних рівнянь

Мета лекції: ознайомитися з основними ітераційними методами розв'язування систем лінійних алгебраїчних рівнянь. Навчитися аналізувати умови збіжності цих методів.

План лекції:

1. Метод ітерацій розв'язування систем лінійних алгебраїчних рівнянь.
2. Достатня умова збіжності методу ітерацій.
3. Метод Гауса-Зейделя.
4. Блок-схема методу Гауса-Зейделя.
5. Перетворення СЛАУ до вигляду, зручному для ітерування.

Ітераційні методи розв'язання СЛАУ

Систему

$$A\vec{x} = \vec{b} \quad (1)$$

зводимо до вигляду

$$\vec{x} = B\vec{x} + \vec{f} \quad (2)$$

Будь яка система

$$\vec{x} = \vec{x} - C(A\vec{x} - \vec{b}) \quad (3)$$

має вигляд (2) і при $\det C \neq 0$ еквівалентна системі (1). Наприклад, для $C = \tau E$

$$\vec{x} = \vec{x} - \tau(A\vec{x} - \vec{b}) \quad (3')$$

Метод простої ітерації (м.п.і.)

Цей метод застосовується до рівняння (2)

$$\vec{x}^{k+1} = B\vec{x}^k + \vec{f}, \quad (4)$$

\vec{x}^0 –початкове наближення задано. Ітераційний процес збігається, тобто $\|\vec{x}^k - \vec{x}\| \rightarrow 0, k \rightarrow \infty$, якщо

$$\|B\| \leq q < 1 \quad (5)$$

При цьому має місце оцінка

$$\|\bar{x}^n - \bar{x}\| \leq \frac{q^n}{1-q} \|\bar{x}^1 - \bar{x}^0\| \quad (6)$$

Метод Якобі

Припустимо $a_{ii} \neq 0, \forall i$. Зведемо систему (1) до вигляду

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

Ітераційний процес запишемо у вигляді

$$x_i^{k+1} = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j^k - \sum_{j=i+1}^n \frac{a_{ij}}{a_{ii}} x_j^k + \frac{b_i}{a_{ii}} \quad k = 0, 1, \dots, i = \overline{1, n} \quad (7)$$

Ітераційний процес збігається до розв'язку, якщо виконується умова

$$\sum_{\substack{j=1 \\ i \neq j}}^n |a_{ij}| \leq |a_{ii}|, \forall i = \overline{1, n}$$

Це умова діагональної переваги матриці A . Якщо ж

$$\sum_{\substack{j=1 \\ i \neq j}}^n |a_{ij}| \leq q \cdot |a_{ii}|, \forall i = \overline{1, n}, 0 \leq q < 1, \quad (8)$$

то має місце оцінка точності:

$$\|\bar{x}^n - \bar{x}\| \leq \frac{q^n}{1-q} \|\bar{x}^1 - \bar{x}^0\|.$$

Метод Зейделя

В компонентному вигляді ітераційний метод Зейделя записується так:

$$x_j^{k+1} = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j^{k+1} - \sum_{j=i+1}^n \frac{a_{ij}}{a_{ii}} x_j^k + \frac{b_i}{a_{ii}} \quad (9)$$

На відміну від методу Якобі на k -му-кроці попередні компоненти розв'язку беруться з $k+1$ -ої ітерації.

Достатня умова збіжності методу Зейделя - $A^T = A > 0$.

Матрична інтерпретація методів Якобі і Зейделя

Подамо матрицю A у вигляді

$$A = A_1 + D + A_2,$$

де A_1 - нижній трикутник матриці A , A_2 - верхній трикутник матриці A , D - її діагональ. Тоді систему (1) запишемо у вигляді

$$D\bar{x} = A_1\bar{x} + A_2\bar{x} + \vec{b}$$

або

$$\bar{x} = D^{-1}A_1\bar{x} + D^{-1}A_2\bar{x} + D^{-1}\vec{b}$$

Матричний запис методу Якобі:

$$\bar{x}^{k+1} = -D^{-1}(A_1 + A_2)\bar{x}^k + D^{-1}\bar{b};$$

методу Зейделя:

$$\bar{x}^{k+1} = -D^{-1}A_1\bar{x}^{k+1} - D^{-1}A_2\bar{x}^k + D^{-1}\bar{b} \quad \text{або} \quad (D + A_1)\bar{x}^{k+1} = -A_2\bar{x}^k + \bar{b}.$$

Необхідна і достатня умова збіжності методу Якобі:

- всі корені рівняння $\det(D + (A_1 + A_2)\lambda) = 0$ по модулю більше 1.

Необхідна і достатня умова збіжності методу Зейделя:

- всі корені рівняння $\det(A_1 + D + A_2\lambda) = 0$ по модулю більше 1.

Однокрокові (двошарові) ітераційні методи

Канонічною формою однокрокового ітераційного методу розв'язку СЛАР є його запис у вигляді

$$B_k \frac{\begin{matrix} \rightarrow k+1 & \rightarrow k \\ x & -x \end{matrix}}{\tau_{k+1}} + A\bar{x}^k = \bar{b} \quad (10)$$

Тут $\{B_k\}$ - послідовність матриць (пере обумовлюючі матриці), що задають ітераційний метод на кожному кроці; $\{\tau_{k+1}\}$ - ітераційні параметри.

Якщо $B_k = E$, то ітераційний процес називається *явним*

$$\bar{x}^{k+1} = \bar{x}^k - \tau_{k+1}(A\bar{x}^k + \bar{b}).$$

Якщо $B_k \neq E$, то ітераційний процес називається *неявним*

$$B_k\bar{x}^{k+1} = \bar{F}^k.$$

В цьому випадку на кожній ітерації необхідно розв'язувати СЛАР.

Якщо $\tau_{k+1} \equiv \tau$, $B_k \equiv B$, то ітераційний процес називається *стаціонарним*; інакше – *нестабілізованим*.

Методам, що розглянуті вище відповідають:

методу простої ітерації (3'):

$$B_k = E, \tau_{k+1} = \tau;$$

методу Якобі:

$$B_k = D, \tau_{k+1} = 1;$$

методу Зейделя:

$$B_k = D + A_1, \tau_{k+1} = 1.$$

Збіжності стаціонарних ітераційних процесів у випадку симетричних матриць

Розглянемо випадок симетричних матриць $A^T = A$ і стаціонарний ітераційний процес $B_k \equiv E, \tau_{k+1} \equiv \tau$

Нехай для A справедливі нерівності

$$\gamma_1 E \leq A \leq \gamma_2 E, \gamma_1, \gamma_2 > 0 \quad (11)$$

Тоді при виборі $\tau = \tau_0 = \frac{2}{\gamma_1 + \gamma_2}$ ітераційний процес збігається. Найбільш

точним значенням γ_1, γ_2 при яких виконуються обмеження (11) є

$$\gamma_1 = \min_i \lambda_i(A), \gamma_2 = \max_i \lambda_i(A). \quad \text{Тоді} \quad q = q_0 = \frac{\gamma_2 - \gamma_1}{\gamma_2 + \gamma_1} = \frac{1 - \xi}{1 + \xi}, \xi = \frac{\gamma_1}{\gamma_2}$$

(Зауважимо, що аналогічно обчислюється q і для методу релаксації розв'язання нелінійних рівнянь, де $\gamma_1 = m_1 = \min |f'(x)|$, $\gamma_2 = M_1 = \max |f'(x)|$) і справедлива оцінка

$$\|\bar{x}^n - \bar{x}\| \leq \frac{q^n}{1-q} \|\bar{x}^0 - \bar{x}\|.$$

Явний метод з багатьма параметрами $\{\tau_k\}$:

$$B \equiv E, \{\tau_k\}: \min_{\tau} q(\tau) \quad n = n(\varepsilon) \rightarrow \min,$$

які обчислюються за допомогою нулів багаточлена Чебишова, називаються ітераційним методом з чебишевським набором параметрів.

Метод верхньої релаксації

Узагальненням методу Зейделя є метод верхньої релаксації:

$$(D + \omega A_1) \frac{\bar{x}^{k+1} + \bar{x}^k}{\omega} + A\bar{x}^k = \bar{b},$$

де D - діагональна матриця з елементами a_{ii} по діагоналі. $\omega > 0$ - заданий числовий параметр.

Тепер $B = D + \omega A_1, \tau = \omega$. Якщо $A^T = A > 0$, то метод верхньої релаксації збігається при умові $0 < \omega < 2$. Параметр підбирається експериментально з умови мінімальної кількості ітерацій.

Методи варіаційного типу

До цих методів відносяться: метод мінімальних нев'язок, метод мінімальних поправок, метод найшвидшого спуску, метод спряжених градієнтів. Вони дозволяють обчислювати наближення без використання апріорної інформації про γ_1, γ_2 в (11).

Нехай $B = E$. Для методу мінімальних нев'язок параметри τ_{k+1} обчислюються з умови

$$\|\bar{r}^{k+1}\|^2 = \|\bar{r}^k\|^2 - 2\tau_{k+1}(\bar{r}^k, A\bar{r}^k) + \tau_{k+1}^2 \|A\bar{r}^k\|^2 \rightarrow \min.$$

Тому

$$\tau_{k+1} = \frac{(A\bar{r}^k, \bar{r}^k)}{\|A\bar{r}^k\|^2} \quad \text{де } \bar{r}^k = A\bar{x}^k - \bar{b} \text{ - нев'язка.}$$

Умова для завершення ітераційного процесу:

$$\|\bar{r}^n\| < \varepsilon.$$

Швидкість збіжності цього методу співпадає із швидкістю методу простої ітерації з одним оптимальним параметром $\tau_0 = \frac{2}{\gamma_1 + \gamma_2}$.

Аналогічно будуються методи з $B \neq E$. Матриця B називається переобумовлювачем і дозволяє підвищити швидкість збіжності ітераційного

процесу. Його вибирають з умов а) легко розв'язувати СЛАР $B\bar{x}^k = \bar{F}_k$ (діагональний, трикутний, добуток трикутних та інше); б) зменшення числа обумовленості матриці $B^{-1}A$ у порівнянні з A .

Контрольні запитання:

1. Метод простої ітерації.
2. Метод Якобі.
3. Метод Зейделя.
4. Однокрокові (двошарові) ітераційні методи.
5. Метод верхньої релаксації.
6. Методи варіаційного типу.

Література [1] - [6], [11].

Лекція 7. Інтерполювання функцій

Мета лекції: засвоїти поняття інтерполяційного многочлену. Навчитися будувати інтерполяційні многочлени Лагранжа і Ньютона.

План лекції:

1. Інтерполювання функцій однієї змінної.
2. Теорема про існування інтерполяційного многочлену.
3. Многочлен Лагранжа.
4. Інтерполювання за рівномірною сіткою вузлів.
5. Многочлен Ньютона. Поняття скінчених різниць. Побудова многочлену Ньютона за таблицею скінчених різниць.
6. Кусково - лінійна інтерполяція.

Постановка задачі інтерполювання

Нехай функція $f(x) \in C[a, b]$ задана своїми значеннями

$y_i = f(x_i), x_i \in [a, b], i = \overline{0, n}$, причому при $x_i \neq x_j$ для $i \neq j$.

Функція $\Phi(x)$ називається *інтерполюючою* для $f(x)$ на сітці $\{x_i\}_{i=0}^n$, якщо $\Phi(x_i) = y_i, i = \overline{0, n}$.

Задача інтерполювання функції має не єдиний розв'язок. Виберемо систему лінійно незалежних функцій $\{\varphi_k(x)\}_{k=0}^n, \varphi_k(x) \in C[a, b]$ і побудуємо лінійну комбінацію

$$\Phi(x) = \Phi_n(x) = \sum_{k=0}^n c_k \varphi_k(x), \quad (1)$$

яка називається *узагальненим багаточленом*. Умови інтерполювання дають СЛАР

$$\sum_{k=0}^n c_k \varphi_k(x_i) = y_i, i = \overline{0, n}, \quad (2)$$

розв'язком якої є $\vec{c} = (c_0, \dots, c_n)$. Якщо

$$D(x_0, \dots, x_n) = \begin{vmatrix} \varphi_0(x_0) & \dots & \varphi_n(x_0) \\ \dots & \dots & \dots \\ \varphi_0(x_n) & \dots & \varphi_n(x_n) \end{vmatrix} \neq 0,$$

то система (2) має єдиний розв'язок.

Система функцій $\{\varphi_k(x)\} k = \overline{0, n}$ називається *системою Чебишова*, якщо, $D(x_0, \dots, x_n) \neq 0, \forall \{x_i\}_{i=0}^n, x_i \in [a, b], i = \overline{0, n}, x_i \neq x_j \forall i \neq j$.

Приклади систем Чебишова.

1. $\varphi_k(x) = x^k$ - алгебраїчна система.

Визначник $D(x_0, \dots, x_n) \neq 0$ є визначником Вандермонда:

$$D(x_0, \dots, x_n) = \begin{vmatrix} 1 & x_0 & \dots & x_0^{n-1} \\ \dots & \dots & \dots & \dots \\ 1 & x_n & \dots & x_n^{n-1} \end{vmatrix} = \prod_{0 \leq k < m \leq n} (x_k - x_m) \neq 0.$$

2. $\varphi_k(x) = L_k(x)$ - ортогональні багаточлени Лежандра; $\varphi_k(x) = T_k(x)$ - ортогональні багаточлени Чебишова.

3. $\varphi_k(x)$: $1, \cos x, \sin x, \dots, \cos nx, \sin nx$.

Тоді $\Phi_n(x) = T_n(x) = a_0 + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx)$ - тригонометричний багаточлен.

Інтерполяційна формула Лагранжа

Якщо $\varphi_k(x) = x^k$, то $\Phi_n(x) = P_n(x) = \sum_{k=0}^n c_k x^k$. Задача інтерполявання функції $f(x)$ алгебраїчним, багаточленом полягає в знаходженні коефіцієнтів $c_k, k = \overline{0, n}$ для яких виконується умова $f(x_i) = \varphi(x_i) i = \overline{0, n}$. Представимо інтерполяційний багаточлен у вигляді

$$P_n(x) = L_n(x) = \sum_{k=0}^n f(x_k) \Phi_k^{(n)}(x) \quad (1)$$

Тут $L_n(x)$ - інтерполяційний поліном; $\Phi_k^{(n)}(x)$ - поліноми n -го степеня, які називають *множниками Лагранжа*. З умови $L_n(x_i) = f(x_i)$ випливає, що множник Лагранжа повинен задовольняти умови

$$\Phi_k^{(n)}(x_i) = \delta_{ik}. \quad (2)$$

Так як $\Phi_k^{(n)}(x)$ - багаточлен степеня n , то він має вигляд

$$\Phi_k^{(n)}(x) = A_k (x - x_0) \dots (x - x_{k-1}) (x - x_{k+1}) \dots (x - x_n),$$

де A_k - число. Знайдемо його з умови $\Phi_k^{(n)}(x_k) = 1$:

$$A_k = \frac{1}{(x_k - x_0) \dots (x_k - x_{k-1}) (x_k - x_{k+1}) \dots (x_k - x_n)}.$$

Таким чином багаточлени $\Phi_k^{(n)}(x)$ мають вигляд:

$$\Phi_k^{(n)}(x) = \frac{(x-x_0)\dots(x-x_{k-1})(x-x_{k+1})\dots(x-x_n)}{(x_k-x_0)\dots(x_k-x_{k-1})(x_k-x_{k+1})\dots(x_k-x_n)} \quad (3)$$

Позначивши $\omega_n(x) = \prod_{i=0}^n (x-x_i)$, маємо $\Phi_k^{(n)}(x) = \frac{\omega_n(x)}{(x-x_k)\omega_n'(x_k)}$. Остаточно формула Лагранжа має вигляд:

$$L_n(x) = \sum_{k=0}^n f(x_k) \frac{\omega_n(x)}{(x-x_k)\omega_n'(x_k)} \quad (4)$$

Залишковий член інтерполяційного полінома

В заданих точках (точки інтерполювання) значення функції та полінома співпадають, але в інших точках в загальному випадку не співпадають. Отже доцільно розглянути питання про похибку інтерполювання.

Замінюючи функцію $f(x)$ на $L_n(x)$ ми допускаємо похибку $r_n(x) = f(x) - L_n(x)$. Це *залишковий член* інтерполювання.

З означення випливає, що, $r_n(x_i) = 0$, $x_i \in [a, b]$, $i = \overline{0, n}$. Оцінимо похибку у довільній точці $x \in [a, b]$. Введемо допоміжну функцію:

$$g(t) = f(t) - L_n(t) - K\omega_n(t), \quad t \in [a, b], \quad g(x_i) = 0, \quad i = \overline{0, n}.$$

Знайдемо таке K , щоб $g(x) = 0$, в деякій точці $x \in [a, b]$, $x \neq x_i$, $i = \overline{0, n}$. Легко бачити, що

$$K = \frac{f(x) - L_n(x)}{\omega_n(x)}.$$

Припустимо що $f(x) \in C^{n+1}[a, b]$, тоді $g(t) \in C^{n+1}[a, b]$. Функція $g(t) = 0$ в $(n+2)$ -х точках, а саме: $t = x$, $t = x_i$, $i = \overline{0, n}$. З теореми Ролля випливає, що існує $(n+1)$ -а точка, де $g'(t_i) = 0$, $i = \overline{0, n}$. Продовжуючи цей процес, отримаємо, що існує хоча б одна точка $\xi \in [a, b]$ така, що $g^{(n+1)}(\xi) = 0$. Так як $g^{(n+1)}(t) = f^{(n+1)}(t) - 0 - K(n+1)!$, то

$$g^{(n+1)}(\xi) = f^{(n+1)}(\xi) - (n+1)! \frac{f(x) - L_n(x)}{\omega_n(x)} = 0.$$

Звідси

$$r_n(x) = f(x) - L_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_n(x) \quad (1)$$

Оскільки ξ невідомо, то використовують оцінку залишкового члена:

$$|r_n(x)| = |f(x) - L_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |\omega_n(x)|, \quad (2)$$

де $M_{n+1} = \sup_{x \in [a, b]} |f^{(n+1)}(x)|$.

Багаточлени Чебишова. Мінімізація залишкового члена інтерполяційного полінома

Як вибрати вузли інтерполяції, щоб похибка інтерполювання була мінімальною? Спочатку обґрунтуємо теоретичний апарат, завдяки якому будемо досліджувати це питання.

Багаточленом Чебишова (n-того степеня, 1-го роду) називається поліном, який задається такими рекурентними співвідношеннями:

$$T_{n+1}(x) - 2xT_n(x) + T_{n-1}(x) = 0, \quad (1)$$

$$T_0(x) = 1, T_1(x) = x. \quad (2)$$

Знайдемо явний вигляд багаточлена Чебишова. Будемо шукати розв'язок рівняння (1) у вигляді $T_n(x) = q^n$, де $q = q(x)$. З (1), отримуємо характеристичне рівняння $q^2 - 2xq + 1 = 0$. Тоді при $|x| \geq 1 \Rightarrow q_{1,2} = x \pm \sqrt{x^2 - 1}$, а при $|x| < 1 \Rightarrow q_{1,2} = \cos \varphi \pm i \sin \varphi$, $i^2 = -1$, $\varphi = \arccos x$.

Розглянемо обидва випадки детальніше.

а) При $|x| \leq 1$ $T_n(x) = A \cos(n\varphi) + B \sin(n\varphi)$. З (2) випливає, що $A=1, B=0$ і тому

$$T_n(x) = \cos(n \arccos x). \quad (3)$$

б) При $|x| > 1$

$$T_n(x) = \frac{1}{2} \left[\left(x + \sqrt{x^2 - 1} \right)^n + \left(x - \sqrt{x^2 - 1} \right)^n \right]. \quad (4)$$

Знайдемо нулі та екстремуми багаточлена Чебишова.

$$T_n(x) = 0, x \in [-1, 1], \cos(n \arccos x) = 0, \arccos x = \frac{2k+1}{2n} \pi, k = \overline{0, n-1}.$$

Отже нулі багаточлена Чебишова:

$$x_k = \cos \frac{2k+1}{2n} \pi \in [-1, 1], k = \overline{0, n-1}.$$

Локальні екстремуми багаточлена Чебишова на $x \in [-1, 1]$:

$$x'_k = \cos \frac{k\pi}{2n}, k = \overline{0, n}, T_n(x'_k) = (-1)^k.$$

Коефіцієнт при старшому члені багаточлена дорівнює 2^{n-1} . Введемо нормований багаточлен Чебишова $\overline{T}_n(x) = 2^{1-n} T_n(x) = x^n + \dots$. Тоді

$$\|\overline{T}_n(x)\|_{C[-1,1]} = \max_{x \in [-1,1]} |T_n(x)| = 2^{1-n}.$$

Відхиленням двох функцій $f(x)$ та $\Phi(x)$ називається величина

$$\Delta(f, \Phi) = \|f(x) - \Phi(x)\|_{C[a,b]}.$$

Теорема. (Чебишова) Серед усіх багаточленів n -го степеня з коефіцієнтом 1 при старшому степені $\overline{T}_n(x)$ найменше відхиляється від 0 на $[-1, 1]$, тобто

$$\|\overline{T}_n(x) - 0\|_{C[-1,1]} = \inf_{\overline{P}_n(x)} \|\overline{P}_n(x)\|_{C[-1,1]} = 2^{1-n}.$$

◁ Будемо доводити від супротивного: припустимо, що існує багаточлен, такий, що

$$\|\overline{Q}_n(x)\| < 2^{1-n}.$$

Тоді $Q_{n-1}(x) = \overline{T}_n(x) - \overline{Q}_n(x)$ - поліном степеня не вище $n-1$ і не рівний тотожно нулю. Дослідимо його знаки:

$$\text{sign}(Q_{n-1}(x'_k)) = \text{sign}(\overline{T}_n(x'_k) - \overline{Q}_n(x'_k)) = \text{sign}(\overline{T}_n(x'_k)) = \alpha(-1)^k, \alpha = \pm 1.$$

Значить $\exists z_k, k = \overline{0, n-1}$ таке, що $Q_{n-1}(z_k) = 0$. Це протиріччя, бо $Q_{n-1}(x)$ - поліном степеня $\leq n-1$. \triangleright

Тепер узагальнимо наш багаточлен Чебишова на довільний проміжок. Нагадаємо $T_n(t) = \cos(n \arccos t)$ для $-1 \leq t \leq 1$. Від змінної $t \in [-1, 1]$

перейдемо до $x \in [a, b]$. Запровадимо заміну: $t = \frac{2}{b-a}x - \frac{b+a}{b-a}$,

$x = \frac{b+a}{2} + \frac{b-a}{2}t$. Тоді

$$T_n^{[a,b]}(t) = \overline{T}_n\left(\frac{2}{b-a}x - \frac{b+a}{b-a}\right) = 2^{1-n} \cos\left(n \arccos\left(\frac{2}{b-a}x - \frac{b+a}{b-a}\right)\right).$$

Побудований нами багаточлен Чебишова на $[a, b]$ не є нормованим. Нормований багаточлен Чебишова на $[a, b]$:

$$\overline{T}_n^{[a,b]}(x) = \frac{(b-a)^n}{2^{2n-1}} \cos\left(n \arccos\left(\frac{2x - (b+a)}{b-a}\right)\right).$$

Відповідно його нулі $x_k = \frac{a+b}{2} - \frac{b-a}{2}t_k, t_k = \cos \frac{(2k+1)\pi}{2n}, k = \overline{0, n-1}$, а

точки екстремуму $x'_k = \frac{a+b}{2} - \frac{b-a}{2}t'_k, t'_k = \cos \frac{k\pi}{2n}, k = \overline{0, n}$. Теорема

Чебишова вірна і для проміжку $[a, b]$. Тепер $\|\overline{T}_n^{[a,b]}(t)\|_{C[a,b]} = \frac{(b-a)^n}{2^{2n-1}}$.

Перейдемо до питання мінімізації залишкового члена. Нагадаємо, що

$$|r_n(x)| = |f(x) - L_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |\omega_n(x)|, \quad (5)$$

де $M_{n+1} = \max_{x \in [a,b]} |f^{(n+1)}(x)|$, $\omega_n(x) = \prod_{i=0}^n (x - x_i) = x^{n+1} + \dots$. Поставимо задачу

$$\inf_{P_n(x)} \max_{x \in [a,b]} |\omega_n(x)|.$$

За теоремою Чебишова розв'язком цієї задачі є поліном Чебишова $\omega_n(x) = \overline{T}_{n+1}^{[a,b]}(x)$. Якщо співпадають поліноми, то співпадають їх нулі. Отже: x_k - вузли інтерполяції співпадають з нулями багаточлена Чебишова:

$x_k = \frac{a+b}{2} - \frac{b-a}{2}t_k, t_k = \cos \frac{(2k+1)\pi}{2(n+1)}, k = \overline{0, n}$. В цьому випадку

$$|r_n(x)| = |f(x) - L_n(x)| \leq \frac{M_{n+1}}{(n+1)!} \frac{(b-a)^{n+1}}{2^{2n+1}} \quad (6)$$

Цю оцінку не можна покращити! Так для $f(x) = \bar{P}_{n+1}(x) = x^{n+1} + \dots$ її $(n+1)$ похідна дорівнює $(n+1)!$, тому $M_{n+1} = (n+1)!$. Різниця $f(x) - L_n(x) = \bar{T}_{n+1}^{[a,b]}(x)$, отже $\max_{[a,b]} |f(x) - L_n(x)| = \frac{(b-a)^{n+1}}{2^{2n+1}}$.

Розділені різниці

Розділені різниці є аналогом похідної для функцій, що задана таблицею.

Розділеною різницею 1-го порядку для функції $f(x)$ називатимемо

$$f(x_i; x_j) = \frac{f(x_i) - f(x_j)}{x_i - x_j}.$$

Розділеною різницею 2-го порядку для функції $f(x)$ називатимемо

$$f(x_i; x_j; x_k) = \frac{f(x_i; x_j) - f(x_j; x_k)}{x_i - x_k}.$$

Аналогічно визначаються розділені різниці довільного порядку.

Наведемо деякі властивості розділених різниць:

1⁰. $f(x_0; \dots; x_n) = \sum_{i=0}^n \frac{f(x_i)}{\prod_{i \neq j} (x_i - x_j)}$.

2⁰. Розділена різниця – лінійний функціонал
 $(\alpha_1 f_1 + \alpha_2 f_2)(x_0; x_1) = \alpha_1 f_1(x_0; x_1) + \alpha_2 f_2(x_0; x_1)$

3⁰. Розділена різниця – симетричний функціонал
 $f(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = f(x_1, \dots, x_j, \dots, x_i, \dots, x_n)$

4⁰. Для $f(x) \in C^n([a, b])$, $\exists \xi \in (a, b) : f(x_0, x_1, \dots, x_n) = \frac{f^{(n)}(\xi)}{n!}$.

Задача. Довести властивість 1⁰ розділених різниць

$$f(x_0; \dots; x_n) = \sum_{i=0}^n \frac{f(x_i)}{\prod_{i \neq j} (x_i - x_j)}.$$

Таблиця розділених різниць має вигляд

x_i	f_i	p.p.1	p.p.2	p.p.n
x_0	$f(x_0)$				
x_1	$f(x_1)$	$f(x_0; x_1)$			
x_2	$f(x_2)$	$f(x_1; x_2)$	$f(x_0; x_1; x_2)$		
.....			$f(x_0; \dots; x_n)$	
x_n	$f(x_n)$	$f(x_{n-1}; x_n)$	$f(x_n; x_{n-1}; x_{n-2})$		

Інтерполяційна формула Ньютона

Запишемо формулу Лагранжа інтерполяційного багаточлена

$$L_n(x) = \sum_{i=0}^n f(x_i) \frac{\omega_n(x)}{(x-x_i)\omega'_n(x_i)}, \quad (1)$$

де $\omega_n(x) = \prod_{j=0}^n (x-x_j)$.

Позначимо $\Phi_j(x) = L_j(x) - L_{j-1}(x)$. Тоді, оскільки

$$\begin{aligned} L_n(x) &= L_0(x) + (L_1(x) - L_0(x)) + \dots + (L_n(x) - L_{n-1}(x)), \\ L_j(x_i) &= L_{j-1}(x_i) = f(x_i) \quad \forall i \in \overline{0, j-1}, \end{aligned}$$

то

$$\Phi_j(x_i) = A_j(x-x_0)\dots(x-x_{j-1}), \quad (2)$$

де A_j визначається з умови $\Phi_j(x_j) = L_j(x_j) - L_{j-1}(x_j) = f(x_j) - L_{j-1}(x_j)$. Звідси

$$\Phi_j(x) = \frac{f(x_j) - L_{j-1}(x_j)}{(x_j - x_0)\dots(x_j - x_{j-1})} (x-x_0)\dots(x-x_{j-1}).$$

Тоді

$$\begin{aligned} A_j &= \frac{f(x_j) - L_{j-1}(x_j)}{(x_j - x_0)\dots(x_j - x_{j-1})} = \frac{f(x_j)}{(x_j - x_0)\dots(x_j - x_{j-1})} - \\ &- \sum_{i=0}^{j-1} \frac{f(x_i)}{(x_i - x_0)\dots(x_i - x_{i-1})(x_i - x_{i+1})\dots(x_i - x_{j-1})(x_j - x_i)} = \\ &= \frac{f(x_j)}{(x_j - x_0)\dots(x_j - x_{j-1})} + \sum_{i=0}^{j-1} \frac{f(x_i)}{(x_i - x_0)\dots(x_i - x_{i-1})(x_i - x_{i+1})\dots(x_i - x_{j-1})(x_i - x_j)} = \\ &= \sum_{i=0}^j \frac{f(x_i)}{(x_i - x_0)\dots(x_i - x_{i-1})(x_i - x_{i+1})\dots(x_i - x_j)} = f(x_0, \dots, x_j). \end{aligned}$$

Звідси маємо інтерполяційну формулу Ньютона вперед ($x_0 \rightarrow x_n$):

$$L_n(x) = f(x_0) + f(x_0, x_1)(x-x_0) + \dots + f(x_0, \dots, x_n)(x-x_0)\dots(x-x_{n-1}). \quad (3)$$

Аналогічно, інтерполяційна формула Ньютона назад ($x_n \rightarrow x_0$):

$$L_n(x) = f(x_n) + f(x_n, x_{n-1})(x-x_n) + \dots + f(x_0, \dots, x_n)(x-x_n)\dots(x-x_1). \quad (4)$$

Маємо рекурсію за степенем багаточлена

$$L_n(x) = L_{n-1}(x) + f(x_0; \dots; x_n)(x-x_0)\dots(x-x_{n-1}).$$

Звідси $L_n(x) = f(x_0) + (x-x_0)\{f(x_0, x_1) + (x-x_1)\{\dots + (x-x_{n-1})f(x_0, x_1, \dots, x_n)\}\}$

і цю формулу розкриваємо починаючи з середини (це аналог формули Герона обчислення значення багаточлена).

Виведемо нову формулу для похибки інтерполювання. Для $x \neq x_i, i = \overline{0, n}$ розглянемо розділену різницю

$$f(x; x_0; \dots; x_n) = \frac{f(x)}{(x-x_0)\dots(x-x_n)} + \sum_{k=0}^n \frac{f(x_k)}{(x_k - x) \prod_{i \neq k} (x_k - x_i)}.$$

Звідси

$$f(x) = f(x_0) \frac{(x-x_1)\dots(x-x_n)}{(x_0-x_1)\dots(x_0-x_n)} + \dots + f(x_n) \frac{(x-x_0)\dots(x-x_{n-1})}{(x_n-x_0)\dots(x_n-x_{n-1})} + f(x; x_0; \dots; x_n)(x-x_0)\dots(x-x_n) = L_n(x) + f(x; x_0; \dots; x_n)\omega_n(x)$$

Тоді похибка має вигляд

$$r_n(x) = f(x) - L_n(x) = f(x; x_0; \dots; x_n)\omega_n(x) \quad (5)$$

Це нова форма для залишкового члена.

Порівнюючи з формулою залишкового члена в пункті 6.3, маємо

$$f(x, x_0, \dots, x_n) = \frac{f^{(n+1)}(\xi)}{(n+1)!},$$

що доводить властивість 4^0 розділених різниць

Нехай маємо сітку рівновіддалених вузлів: $x_i = a + ih, i = \overline{0, n}, h = \frac{b-a}{n}$.

Позначимо $\Delta f_0 = f_1 - f_0, \Delta^2 f_0 = \Delta f_1 - \Delta f_0 = f_2 - 2f_1 + f_0 \dots$ - скінченні різниці. Запишемо формули Ньютона у нових позначеннях:

$$L_n(x) = L_n(x_0 + th) = f_0 + t\Delta f_0 + \dots + \frac{t(t-1)\dots(t-n+1)}{n!} \Delta^n f_0, t = \frac{x-x_0}{h}.$$

Це інтерполяційна формула Ньютона вперед по рівновіддалених вузлах.

Задача. Побудувати інтерполяційну формулу Ньютона назад по рівновіддалених вузлах.

Інтерполювання з кратними вузлами

Нехай $f(x)$ задана таблицею значень $f^{(j)}(x_i) \quad i = \overline{0, n}, j = \overline{0, k_i - 1}; k_i$ - кратності відповідних вузлів. Побудуємо $H_m^{(i)}(x_j) = f^{(i)}(x_j)$ - інтерполяційний багаточлен Ерміта по кратних вузлах, причому $m = \sum_{i=1}^n k_i - 1$.

Якщо $k_i \equiv 1$, то $H_m(x) = L_n(x)$.

Для побудови $H_m(x)$ в загальному випадку для кожної точки x_i введемо k_i точок $x_{ij}^\varepsilon = x_i + j\varepsilon, \quad i = \overline{0, n}; j = \overline{0, k_i - 1}$. З умови $x_{ik_i-1}^\varepsilon = x_i + \varepsilon(k_i - 1) < x_{i+1} \quad \forall i$ можна вибрати ε .

Нехай $f(x) \in C^{m+1}[a, b]$. Запишемо інтерполяційну формулу Ньютона

$$L_m^\varepsilon(x) = f(x_{00}^\varepsilon) + f(x_{00}^\varepsilon, x_{01}^\varepsilon)(x-x_{00}^\varepsilon) + \dots + f(x_{00}, \dots, x_{nk_n-1})(x-x_{00})\dots(x-x_{nk_n-1})$$

При $\varepsilon \rightarrow 0$ маємо $x_{ij}^\varepsilon \rightarrow x_i$. Крім того $f(x_{i0}^\varepsilon; \dots; x_{ik_i-1}^\varepsilon) = f(x_i; \dots; x_i) = \frac{f^{k_i}(x_i)}{k_i!}$.

Тому $L_m^\varepsilon(x) \rightarrow H_m(x)$ та

$$R_m(x) = f(x) - H_m(x) = \frac{f^{(m+1)}(\xi)}{(m+1)!} \Omega_m(x),$$

де $\Omega_m(x) = (x-x_0)^{k_0} \dots (x-x_n)^{k_n}$.

Збіжність процесу інтерполювання

Виникає питання, чи буде прямувати до нуля похибка інтерполювання $f(x) - L_n(x)$, якщо число вузлів n збільшувати?

Введемо норму $\|f(x) - L_n(x)\|_{C[a,b]} = \max_{x \in [a,b]} |f(x) - L_n(x)|$. Тоді для

$f(x) \in C^{n+1}[a,b]$ справджується оцінка

$$\|f(x) - L_n(x)\|_{C[a,b]} \leq \frac{M_{n+1}}{(n+1)!} \|\omega_n(x)\|_{C[a,b]}, \quad (1)$$

де $M_{n+1} = \max_{x \in [a,b]} |f^{(n+1)}(x)|$, $\omega_n(x) = \prod_{i=0}^n (x - x_i)$. А яка оцінка буде для довільної неперервної функції?

Кажуть, що інтерполяційний процес для функції $f(x)$ збігається в точці $x \in [a,b]$, якщо

$$\lim_{n \rightarrow \infty} L_n(x) = f(x), \forall \{x_i\}_{i=1}^n : h = \max_{i=1, n} h_i \rightarrow 0, (h_i = x_i - x_{i-1}).$$

Якщо $\|f(x) - L_n(x)\|_{C[a,b]} \xrightarrow{n \rightarrow \infty} 0$, то інтерполяційний процес збігається рівномірно.

Розглянемо приклади поведінки інтерполяційних багаточленів при $n \rightarrow \infty$ для деяких функцій.

Приклад 1 Послідовність інтерполяційних багаточленів (сітка рівномірна), побудованих для неперервної функції $f(x) = |x|$, $-1 \leq x \leq 1$ (функція неперервна, але негладка), не збігається на $x \in [-1,1]$, крім точок $x = -1, 0, 1$. На рис. 1 дано графіки самої функції (штрихова лінія) та інтерполяційного багаточлена (суцільна лінія) на рівномірній сітці $x_i = -1 + ih$, $h = \frac{2}{n}$, $i = \overline{0, n}$ для $n=10$.

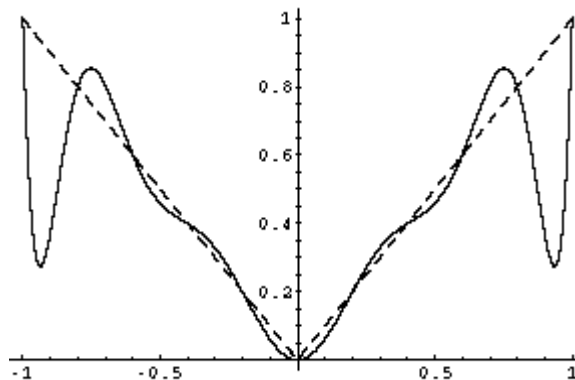


Рис. 3

Приклад 2 Функція Рунге $f(x) = \frac{1}{1 + 40x^2}$, $-1 \leq x \leq 1$ (функція аналітична!).

Для рівномірної сітки

$$x_i = -1 + ih, \quad h = \frac{2}{n}, \quad i = \overline{0, n}$$

маємо графіки: суцільна лінія – інтерполяційного багаточлена; пунктирна – самої функції; $n=10$.

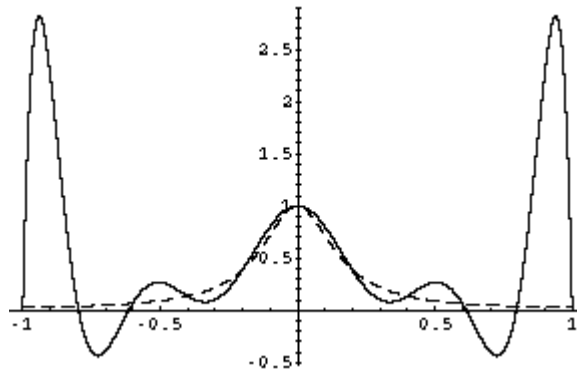


Рис. 4

Пояснити чому рівномірна сітка дає великі похибки інтерполювання біля кінців інтервалу інтерполювання допомагає рис. 3. На цьому рисунку суцільною лінією представлено графік функції $\omega_n(x) = \prod_{i=0}^n (x - x_i)$ ($n=8$) для рівномірної сітки. Як бачимо максимальні за модулем значення цієї функції припадають на кінці інтервалу.

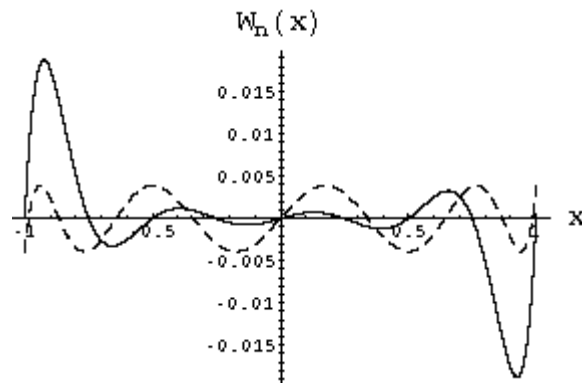


Рис. 5

Для порівняння на цьому ж рисунку (штрихова лінія) побудовано графік $\omega_n(x) = \prod_{i=0}^n (x - x_i)$, що відповідає чебишовським вузлам, які мінімізують похибку інтерполювання. Тепер відхилення $|\omega_n(x)|$ розподілено рівномірно по всьому проміжку інтерполювання.

Теорема 1 (Фабера) $\forall \{x_i\}_{i=0}^n$ існує $f(x) \in C[a, b]$, для якої інтерполяційний процес не збігається рівномірно.

Теорема 2 (Марцинкевича) $\forall f(x) \in C[a, b] \quad \exists \{x_i\}_{i=0}^n$ такі, що послідовність $\{L_n(x)\}$ збігається рівномірно до $f(x)$.

Розглянемо оператори $P_n : C[a, b] \rightarrow L_n$, тобто $P_n f(x) = L_n(x)$.

Теорема 3 Стала Лебега $\|P_n\| = \max_{j \in [a,b]} \sum_{j=0}^n |\varphi_j^{(n)}(x)|$, де $\varphi_j^{(n)}(x) = \frac{\omega_n(x)}{(x-x_j)\omega_n'(x_j)}$.

Теорема 4 Для $f(x) \in C[a, b]$

$$\|f(x) - L_n(x)\|_{C[a,b]} \leq (1 + \|P_n\|)E_n(f),$$

де $E_n(f) = \inf_{Q_n(x)} \|f(x) - Q_n(x)\|_{C[a,b]}$ - відхилення багаточлена n -го степеня найкращого рівномірного наближення від $f(x)$.

Теорема 5 Нехай P_n^E - оператор інтерполяції на рівномірній сітці, P_n^T - оператор інтерполяції на чебишовській сітці. Тоді на $[-1;1]$ маємо наближені оцінки:

$$\|P_n^E\| \approx C_1 2^n, \|P_n^T\| \approx C_2 \ln(n).$$

Останні оцінки пояснюють розбіжність процесу інтерполювання при великих n .

Кусково - лінійна інтерполяція

Інтерполяція багаточленом Лагранжа або Ньютона на відрізку $[a, b]$ з використанням великої кількості вузлів інтерполяції часто приводить до поганого наближення через розбіжність процесу інтерполювання. Для того щоб уникнути великої похибки, весь відрізок $[a, b]$ розбивають на частинні відрізки $[x_{i-1}, x_i]$ і на кожному з частинних відрізків замінюють функцію $f(x)$ багаточленом невисокого степеню. В цьому і полягає кусково-поліноміальна інтерполяція.

Розглянемо найпростішу таку інтерполяцію – кусково лінійну. Нехай задана $f(x)$ значеннями $f(x_i)$ $i = \overline{0, n}$. Побудуємо функцію $\Phi_1(x)$ – лінійну на $x \in [x_{i-1}, x_i]$, що інтерполює ці значення:

$$\Phi_1(x) = L_1^i(x) = f(x_{i-1}) \frac{x - x_{i-1}}{x_i - x_{i-1}} + f(x_i) \frac{x_i - x}{x_i - x_{i-1}}, \quad x \in [x_{i-1}, x_i] \quad (2)$$

Представимо її у вигляді

$$\Phi_1(x) = \sum_{i=0}^n f(x_i) \varphi_i(x). \quad (3)$$

З умов інтерполювання маємо

$$\Phi_1(x_j) = \sum_{i=0}^n f(x_i) \varphi_i(x_j) = f(x_j).$$

Звідси

$$\varphi_i(x_j) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}.$$

Значить

$$\varphi_i(x) = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}}, & x_{i-1} \leq x \leq x_i \\ \frac{x_{i+1} - x}{x_{i+1} - x_i}, & x_i \leq x \leq x_{i+1} \\ 0, & x \leq x_{i-1}, \quad x \geq x_{i+1} \end{cases}$$

Теорема 1 Для $f(x) \in C^2[a, b]$ вірна оцінка

$$\|f(x) - \Phi_1(x)\|_{C[a,b]} \leq \frac{M_2}{8} |h|^2, \quad (4)$$

де $\Phi_1(x)$ - кусково-лінійна функція побудована по значеннях $f(x_i), i = \overline{0, n}$,
 $|h| = \max_i h_i, h_i = x_i - x_{i-1}$.

▷ Маємо для $x \in [x_{i-1}, x_i]$

$$z(x) = f(x) - \Phi_1(x) = f(x) - L_1^i(x) = \frac{f''(\xi_i)}{2!} (x - x_{i-1})(x - x_i).$$

Звідси

$$|f(x) - \Phi_1(x)| \leq \frac{M_2^i}{2} |(x - x_{i-1})(x - x_i)| \leq \frac{M_2^i h_i^2}{8}, \quad (5)$$

де $M_2^i = \max_{x \in [x_{i-1}, x_i]} |f''(x)|$. Остання оцінка отримана з нерівності

$$\max_{[x_{i-1}, x_i]} |(x - x_{i-1})(x - x_i)| = \frac{h_i^2}{4}.$$

Тоді

$$\max_{i=1, n} \max_{x \in [x_{i-1}, x_i]} |z(x)| \leq \frac{1}{8} h^2 M_2, \quad (6)$$

де $M_2 = \max_{[a,b]} |f''(x)|$, $h = \max_i h_i$, що доводить (4) ◁

Задача. Довести оцінку $|f'(x) - \Phi_1'(x)| \leq hM_2$.

Отже маємо збіжність процесу інтерполювання за допомогою кусково-лінійної функції $\|f(x) - \Phi_1(x)\|_{C[a,b]} \xrightarrow{h \rightarrow 0 (n \rightarrow \infty)} 0$.

Розглянемо деякі простори:

1. $H_0 = L_2(a, b)$ - гільбертів простір, в якому скалярний добуток визначається

так: $(u, v) = \int_a^b u(x)v(x)dx$, а норма $\|u\|_0 = \sqrt{(u, u)}$.

2. $H_k = W_2^k(a, b)$. Тепер скалярний добуток $(u, v)_k = \sum_{m=0}^k \int_a^b u^{(m)}(x)v^{(m)}(x)dx$, а

норма $\|u\|_k = \sqrt{\|u\|_0^2 + \dots + \|u^{(k)}\|^2}$.

Теорема 2 Нехай $f(x) \in H_2 = W_2^2(a,b)$. Тоді $\|f^{(k)} - \Phi_1^{(k)}\|_0 \leq |h|^{2-k} \|f\|_2, k = 1,2$.

Зауважимо, що кусково-лінійна інтерполяція негладка, тому на практиці застосовують квадратичні, а найчастіше – кубічні поліноми на кожному під проміжку $x \in [x_{i-1}, x_i]$.

Контрольні запитання:

1. Постановка задачі інтерполювання.
2. Інтерполяційна формула Лагранжа.
3. Залишковий член інтерполяційного полінома.
4. Багаточлени Чебишова. Мінімізація залишкового члена інтерполяційного полінома.
5. Розділені різниці.
6. Інтерполяційна формула Ньютона.
7. Інтерполювання з кратними вузлами.
8. Збіжність процесу інтерполювання.
9. Кусково - лінійна інтерполяція.

Література [2], [3], [5], [6], [7], [11].

Лекція 8. Поняття сплайн-інтерполяції

Мета лекції: засвоїти поняття сплайн-функції, ознайомитися з алгоритмами побудови сплайн-інтерполяції в одно- та двомірних випадках.

План лекції:

1. Поняття сплайн-інтерполяції.
2. Складання системи лінійних рівнянь для визначення коефіцієнтів сплайн-інтерполяції.
3. Алгоритм прогонки для знаходження коефіцієнтів сплайн-інтерполяції.

Сплайн (spline) в перекладі означає рейка, якою користувалися креслярі при проведенні гладкої кривої, що з'єднувала задані точки на площині.

Функція $s(x)$ називається *сплайном* степеня m і дефекту k , якщо

- 1) $s(x) \in \pi_m$ (множина поліномів степеня m) для $x \in [x_{i-1}, x_i], i = \overline{1, n}$;
- 2) $s(x) \in C^{m-k}[a, b]$.

Приклади: 1) $\Phi_1(x): m = 1, k = 1$; 2) $\Phi_3(x): m = 3, k = 2$.

Зараз ми побудуємо сплайн, для якого $m = 3, k = 1$.

Функція $s_3(x) = s(x)$ називається *кубічним інтерполяційним природнім сплайном*, якщо

$$1) s(x) \in \pi_3 \text{ для } x \in [x_{i-1}, x_i], i = \overline{1, n}, \text{ (кубічний);} \quad (1)$$

$$2) s(x) \in C^2[a, b] \text{ (має дефект 1);} \quad (2)$$

$$3) s(x_i) = f(x_i), i = \overline{0, n}, \text{ (інтерполює } f(x)); \quad (3)$$

$$4) s''(a) = s''(b) = 0 \text{ (природній).} \quad (4)$$

Умови (4), так звані умови природності, необхідні, щоб разом було $4n$ умови для знаходження $4n$ коефіцієнтів сплайну. Замість них можуть бути такі умови:

$$s''(a) = A, \quad s''(b) = B \quad (4a)$$

$$s'(a) = A, \quad s'(b) = B \quad (4б)$$

$$s(a) = s(b), \quad s'(a) = s'(b), \quad s''(a) = s''(b) \quad (4в)$$

Умови (4в) – це так звані умови періодичності.

Побудуємо природній сплайн. З (1) та (2) маємо

$$s''(x) = m_{i-1} \frac{x_i - x}{h_i} + m_i \frac{x - x_{i-1}}{h_i}, \quad (5)$$

де $m_i = s''(x_i)$ і вони є невідомими коефіцієнтами; $h_i = x_i - x_{i-1}$.

Двічі інтегруючи (5), маємо:

$$s(x) = m_{i-1} \frac{(x_i - x)^3}{6h_i} + m_i \frac{(x - x_{i-1})^3}{6h_i} + A_i \frac{x_i - x}{h_i} + B_i \frac{x - x_{i-1}}{h_i}.$$

A_i, B_i знаходимо з умов $s(x_{i-1}) = f_{i-1}, s(x_i) = f_i$. Остаточню

$$s(x) = m_{i-1} \frac{(x_i - x)^3}{6h_i} + m_i \frac{(x - x_{i-1})^3}{6h_i} + \left(f_{i-1} - \frac{m_{i-1}h_i^2}{6} \right) \frac{x_i - x}{h_i} + \left(f_i - \frac{m_i h_i^2}{6} \right) \frac{x - x_{i-1}}{h_i}, \quad x \in [x_{i-1}, x_i]. \quad (6)$$

З (4) маємо $m_0 = m_n = 0$.

Враховуючи, що $s'(x_i - 0) = s'(x_i + 0)$ отримаємо СЛАР для знаходження всіх $m_i = s''(x_i)$:

$$\begin{cases} \frac{h_i}{6} m_{i-1} + \frac{h_i + h_{i+1}}{3} m_i + \frac{h_{i+1}}{6} m_{i+1} = \frac{f_{i+1} - f_i}{h_i} - \frac{f_i - f_{i-1}}{h_i}, & i = \overline{1, n-1} \\ m_0 = m_n = 0. \end{cases} \quad (7)$$

Це тридіагональна СЛАР; її можна розв'язати методом прогонки за $O = O(N)$ арифметичних операцій.

Задача. Написати СЛАР для кубічного інтерполяційного сплайну, якщо $s'(a) = A, s'(b) = B$ та розробити алгоритм її розв'язання (тобто написати формули методу прогонки).

Теорема Нехай $f(x) \in C^4[a, b]$, тоді має місце оцінка

$$\|f^{(k)}(x) - s^{(k)}(x)\|_{C[a,b]} \leq M_4 |h|^{4-k}, \quad k = 0, 1, 2,$$

де $M_4 = \max_{[a,b]} |f^{(4)}(x)|, |h| = \max_i h_i$.

Введемо клас функцій $U = \{u(x): u(x) \in W_2^2[a, b], u(x_i) = f_i, i = \overline{0, n}\}$ – це функції досить гладкі і приймають задані значення. Якщо ввести такий функціонал $\Phi(u) = \int_a^b (u''(x))^2 dx$, то

$$\Phi(s) = \inf_{u \in U} \Phi(u),$$

де $s(x)$ – кубічний природній інтерполяційний сплайн.

Оскільки кривизна графіка кривої $u(x)$ пропорційна $u''(x)$, то це фактично означає, що сплайн має в середньоквадратичному розумінні найменшу кривизну серед всіх функцій $u(x) \in W_2^2[a, b]$, що інтерполюють значення $f(x_i)$.

Для того, щоб не розв'язувати СЛАР (7) інколи будують наближення до сплайну $\tilde{s}(x)$, яке отримується заміною $m_i = s''(x_i)$ на

$$f_{\tilde{x}\tilde{x},i} \equiv \frac{1}{\tilde{h}_i} \left(\frac{f_{i+1} - f_i}{h_{i+1}} - \frac{f_i - f_{i-1}}{h_i} \right) \approx f''(x_i) \approx s''(x_i),$$

де $\tilde{h}_i = \frac{h_i + h_{i+1}}{2}$, причому $f''(x_i) - f_{\tilde{x}\tilde{x},i} = O(h_i^2)$. При цьому і

$\tilde{s}(x) - s(x) = O(h^4)$. Відмітимо, що $\tilde{s}(x)$ не є сплайном дефекту 1.

Зауваження 1 Складемо матрицю A розмірності $(n-1) \times (n-1)$:

$$A = \begin{pmatrix} \frac{h_1 + h_2}{3} & \frac{h_2}{6} & 0 & \dots & 0 \\ \frac{h_2}{6} & \frac{h_2 + h_3}{3} & \frac{h_3}{6} & \dots & 0 \\ 0 & \frac{h_3}{6} & \frac{h_3 + h_4}{3} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \frac{h_{n-1}}{6} & \frac{h_{n-1} + h_n}{3} \end{pmatrix}$$

і матрицю H розмірності $(n+1) \times (n-1)$:

$$H = \begin{pmatrix} \frac{1}{h_1} & -\left(\frac{1}{h_1} + \frac{1}{h_2}\right) & \frac{1}{h_2} & 0 & \dots \\ 0 & \frac{1}{h_2} & -\left(\frac{1}{h_2} + \frac{1}{h_3}\right) & \frac{1}{h_3} & \dots \\ 0 & 0 & \frac{1}{h_3} & -\left(\frac{1}{h_2} + \frac{1}{h_3}\right) & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \frac{1}{h_{n-1}} & -\left(\frac{1}{h_{n-1}} + \frac{1}{h_n}\right) \end{pmatrix}$$

Тоді можна записати СЛАР (7) відносно моментів $\vec{m} = (m_1, m_2, \dots, m_{n-1})$ у вигляді:

$$A\vec{m} = H\vec{f}, \text{ де } \vec{f} = (f_0, f_1, \dots, f_n)^T.$$

Зауваження 2 Нагадаємо формулу для інтерполяційного багаточлена

Лагранжа $L_n(x) = \sum_{i=0}^n f(x_i)\Phi_i^{(n)}(x)$, де $\Phi_i^{(n)}$ - множники Лагранжа. Це

представлення інтерполяційного багаточлена Лагранжа по системі функцій

$\{\Phi_i^{(n)}\}$. Для $\Phi_1(x) = \sum_{i=1}^n f(x_i)\varphi_i(x)$ маємо представлення по системі кусково-

лінійних функцій $\{\varphi_i(x)\}$. Для $\Phi_3(x) = \sum_{i=1}^n (f(x_i)\varphi_i^0(x) + f'(x_i)\varphi_i^1(x))$ -

представлення по системі $\{\varphi_i^0, \varphi_i^1\}$.

Аналогічно, якщо представити кубічний сплайн у вигляді

$$s_3(x) = \sum_{i=0}^n c_i B_3^i(x),$$

то відповідна система для кубічного сплайну буде $\{B_3^i(x)\}_{i=1}^n$. Тут $B_3^i(x)$ - так званий кубічний B_3 -сплайн (формула дається, а графік представлено на рис.

б):

$$B_3^i(z) = \frac{1}{6h} \begin{cases} \left(\frac{z-x_{i-2}}{h}\right)^3, & z \in [x_{i-2}, x_{i-1}] \\ -3\left(\frac{z-x_{i-1}}{h}\right)^3 + 3\left(\frac{z-x_{i-1}}{h}\right)^2 + 3\left(\frac{z-x_{i-1}}{h}\right) + 1, & z \in [x_{i-1}, x_i] \\ -3\left(\frac{x_{i+1}-z}{h}\right)^3 + 3\left(\frac{x_{i+1}-z}{h}\right)^2 + 3\left(\frac{x_{i+1}-z}{h}\right) + 1, & z \in [x_i, x_{i+1}] \\ \left(\frac{x_{i+2}-z}{h}\right)^3, & z \in [x_{i+1}, x_{i+2}] \\ 0, & z < x_{i-2}, z > x_{i+2}. \end{cases}$$

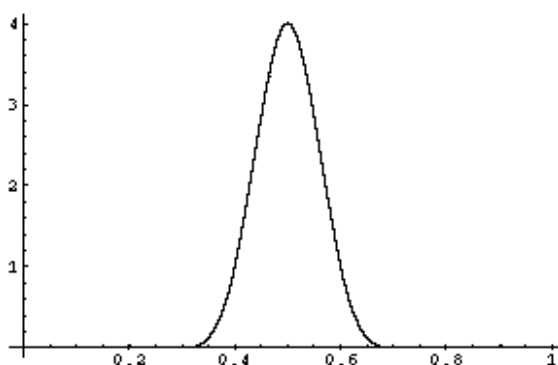


Рис. 6

Задача. Показати, що B_3^i - є кубічним сплайном дефекту 1.

Для знаходження коефіцієнтів c_i записується СЛАР з умов інтерполювання.

Параметричні сплайни

На практиці часто виникає задача побудови кривої по заданим точкам $(x_i, y_i), i = \overline{1, n}$. В цьому випадку використовують сплайни. Якщо відповідна функція $y = f(x)$ однозначна, то сплайн будується за алгоритмом, що розглянуто у пункті 6.12.

Окремо розглянемо випадок, коли точки $(x_i, y_i), i = \overline{1, n}$ в площині (x, y) розташовані у довільний спосіб:

В цьому випадку відповідна функція задається параметрично

$$x = x(t), y = y(t) \quad t \in [A, B]. \quad (1)$$

Для значень $x_i, i = \overline{1, n}$ побудуємо кубічний сплайн $s_x(t)$ такий, що $s_x(t_i) = x_i, i = \overline{1, n}$, а для $y_i, i = \overline{1, n}$ будуємо сплайн $s_y(t)$, для якого $s_y(t_i) = y_i, i = \overline{1, n}$. Тоді параметрична функція

$$(s_x(t), s_y(t)) \quad t \in [A, B] \quad (2)$$

називається параметричним сплайном для функції (1).

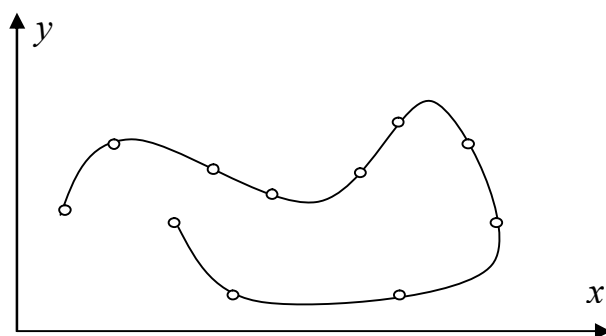


Рис. 7

Стає питання про вибір параметру t . Нехай $t_i = i, i = \overline{1, n}$, тобто для табличних даних $(x_i, y_i), i = \overline{1, n}$ параметром виступає номер точки в площині (x, y) . Тоді для параметричного сплайну неперервний параметр t змінюється на інтервалі $t \in [1, n]$.

Побудова сплайнів $s_x(t)$ та $s_y(t)$ здійснюється за алгоритмом наведеним в пункті 6.12 по значенням $f_i = x_i, i = \overline{1, n}$ та $f_i = y_i, i = \overline{1, n}$.

Розглянемо тепер побудову замкненої гладкої кривої.

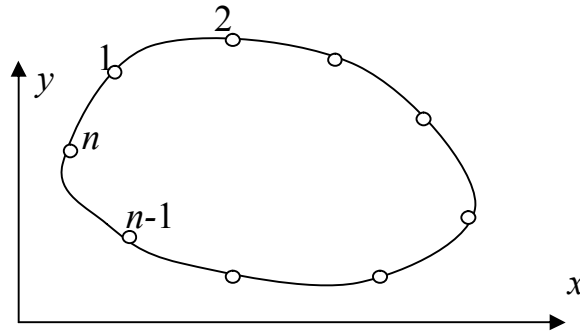


Рис. 8

Параметризуємо її як в попередньому випадку. Відмінність полягає в тому, що тепер функції $x = x(t)$ та $y = y(t)$ періодичні з періодом $T = n$, тобто

$$x(t) = x(t + n), y(t) = y(t + n) \quad \forall t.$$

Наприклад, для значень в точках маємо:

$$x_1 = x_{n+1}, y_1 = y_{n+1} \text{ і } x_0 = x_n, y_0 = y_n. \quad (3)$$

Побудуємо алгоритм реалізації періодичного параметричного кубічного сплайну. Як і для звичайного сплайну на інтервалі $t \in [t_i, t_{i+1}] \quad \forall i$ маємо:

$$s(t) = m_{i-1} \frac{(t_i - t)^3}{6h_i} + m_i \frac{(t - t_{i-1})^3}{6h_i} + \left(f_{i-1} - \frac{m_{i-1}h_i^2}{6} \right) \frac{t_i - t}{h_i} + \left(f_i - \frac{m_i h_i^2}{6} \right) \frac{t - t_{i-1}}{h_i},$$

де $s(t)$ одна з функцій $s_x(t)$ або $s_y(t)$; $f_i = x_i, i = \overline{1, n}$ або $f_i = y_i, i = \overline{1, n}$; $h_i = t_{i+1} - t_i = 1, \forall i$. Для знаходження коефіцієнтів сплайну $m_i = s''(t_i)$ з умови неперервності першої похідної сплайна маємо СЛАР:

$$\begin{cases} \frac{h_i}{6} m_{i-1} + \frac{h_i + h_{i+1}}{3} m_i + \frac{h_{i+1}}{6} m_{i+1} = \frac{f_{i+1} - f_i}{h_i} - \frac{f_i - f_{i-1}}{h_i}, & i = \overline{1, n}; \\ m_0 = m_n, m_{n+1} = m_1. \end{cases} \quad (4)$$

Додаткові умови на коефіцієнти m_i випливають з періодичності сплайну та його других похідних.

Системі (4) відповідає матриця розмірності $(n \times n)$:

$$A = \begin{pmatrix} \left\langle \frac{h_1 + h_2}{3} \right\rangle & \frac{h_2}{6} & 0 & \dots & \left\langle \frac{h_n}{6} \right\rangle \\ \frac{h_2}{6} & \frac{h_2 + h_3}{3} & \frac{h_3}{6} & \dots & 0 \\ 0 & \frac{h_3}{6} & \frac{h_3 + h_4}{3} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \left\langle \frac{h_1}{6} \right\rangle & \dots & 0 & \frac{h_n}{6} & \frac{h_n + h_1}{3} \end{pmatrix},$$

яка є майже тридіагональною: “заважають” два елементи матриці, що виділені кутовими дужками.

Для розв'язання таких систем застосовують метод циклічної прогонки. Розглянемо алгоритм цього методу для більш загальної системи:

$$\begin{cases} a_i y_{i-1} - c_i y_i + b_i y_{i+1} = -f_i, & i = \overline{1, n}; \\ y_0 = y_n, y_{n+1} = y_1. \end{cases} \quad (5)$$

Формули методу:

$$1^0. \alpha_2 = \frac{b_1}{c_1}; \beta_2 = \frac{f_1}{c_1}; \gamma_2 = \frac{a_1}{c_1};$$

$$2^0. z_i = c_i - a_i \alpha_i; \alpha_{i+1} = \frac{b_i}{z_i}; \beta_{i+1} = \frac{f_i + a_i \beta_i}{z_i}; \gamma_{i+1} = \frac{a_i \gamma_i}{z_i}, i = \overline{2, n};$$

$$3^0. p_{n-1} = \beta_n; q_{n-1} = \alpha_n + \gamma_n;$$

$$4^0. p_i = \alpha_{i+1} p_{i+1} + \beta_{i+1}; q_i = \alpha_{i+1} q_{i+1} + \gamma_{i+1}, i = \overline{n-2, 1};$$

$$5^0. y_n = \frac{f_n + b_n p_1 + a_n p_{n-1}}{c_n - b_n q_1 - a_n q_{n-1}};$$

$$6^0. y_i = p_i + y_n q_i; i = \overline{1, n-1}.$$

Метод стійкий ($|\alpha_i| < 1$), якщо $a_i, b_i > 0, |c_i| > b_i + a_i$. Для системи (4) ці умови виконані.

Метод економний, оскільки кількість арифметичних операцій, що витрачається на його реалізацію, $Q = O(n)$.

Розглянуті в цьому пункті параметричні сплайни мають хороші апроксимативні та екстремальні властивості, тому побудовані по ним криві добре відновлюють задані як при малій, так досить великій кількості точок інтерполювання.

Контрольні запитання:

1. Поняття сплайну степеня m і дефекту k .
2. Кубічний інтерполяційний природний сплайн.
3. Система лінійних рівнянь для визначення коефіцієнтів сплайн-інтерполяції.
4. Параметричні сплайни.
5. Алгоритм реалізації періодичного параметричного кубічного сплайну.

Література [2] - [7], [11].

Лекція 9. Задача наближення функцій однієї змінної

Мета лекції: засвоїти поняття найкращого наближення функцій, алгоритмів побудови найкращих наближень функцій за допомогою многочленів.

План лекції:

1. Задача наближення функцій однієї змінної.

2. Найкраще наближення функцій за допомогою многочленів.
3. Метод найменших квадратів.
4. Формули наближення функцій за допомогою лінійних та квадратичних многочленів.

Наближення функцій застосовують у випадках, якщо

- функція складна (трансцендентна або є розв'язком складної задачі) і її замінюють функцією, яка легко обчислюється (найчастіше, поліномом);
- необхідно побудувати функцію неперервного аргументу для функції, яка задана своїми значеннями (таблична);
- таблична функція наближається табличною ж функцією (згладжування).

Інтерполювання не кращий спосіб наближення функцій через розбіжність цього процесу для поліномів. Тим більше доцільність застосування інтерполювання сумнівна, якщо функція таблична, а її значення неточні. Потрібно будувати апроксимуючу функцію з інших міркувань.

Найбільш загальний принцип: наблизити $f(x)$ функцією $\Phi(x)$ так, щоб досягалася деяка задана точність ε :

$$\|f(x) - \Phi(x)\| < \varepsilon.$$

Але розв'язок в такій постановці може не існувати або бути не єдиним.

Загальна постановка задачі наближення така. Нехай маємо елемент f лінійного нормованого простору R . Побудуємо підпростір M_n , в якому елементи є лінійною комбінацією

$$\Phi = \sum_{i=0}^n c_i \varphi_i \in M_n \subset R \quad (1)$$

по елементах лінійно незалежної системи

$$\{\varphi_i\}_{i=0}^{\infty}, \varphi_i \in R \quad (2)$$

Відхилення $\Phi \in M_n$ від $f \in R$ є число

$$\Delta(f, \Phi) = \|f - \Phi\|.$$

Позначимо

$$\inf_{\Phi \in M_n} \|f - \Phi\| = \Delta(f).$$

Елемент Φ_0 такий, що

$$\Delta(f, \Phi_0) = \|f - \Phi_0\| = \inf_{\Phi \in M_n} \|f - \Phi\| = \Delta(f), \quad (3)$$

називається *елементом найкращого наближення* (ЕНН).

Ясно, що умову точності треба перевіряти на цьому елементі. У випадку її невиконання треба збільшувати кількість елементів n в (1).

Теорема 1 Для будь-якого лінійного нормованого простору R існує елемент найкращого наближення $\Phi_0 \in M_n$.

◁ Введемо $F(\vec{c}) = F(c_0, c_1, \dots, c_n) = \|f - \Phi\| = \left\| f - \sum_{i=0}^n c_i \varphi_i \right\|$. Це неперервна функція аргументів $\vec{c} = (c_0, c_1, \dots, c_n)$. Для елементів, які задовольняють умові

$$\|\Phi\| > 2\|f\| \quad f \in R_1 \quad \Phi \in M_n, \quad (4)$$

маємо

$$F(\bar{c}) = \|f - \Phi\| \geq \|\Phi\| - \|f\| > 2\|f\| - \|f\| = \|f\| > \Delta(f).$$

Значить ЕНН $\Phi_0 \in \{\Phi : \|\Phi\| \leq 2\|f\|\} = \bar{U} \subset M_n$. За теоремою Кантора $\exists \Phi_0$, де $F(\bar{c})$ досягає мінімуму. Причому $\|f - \Phi_0\| \leq \|f - \Phi\|$. \triangleright

Елементів найкращого наближення в лінійному нормованому просторі може бути і декілька.

Простір R називається *строго нормованим*, якщо з умови

$$\|f + g\| = \|f\| + \|g\|, \quad \|f\| \neq 0, \quad \|g\| \neq 0$$

впливає, що $\exists \lambda \neq 0$ таке, що

$$g = \lambda f \quad (5)$$

Теорема 2 Якщо простір R строго нормований, то елемент найкращого наближення Φ_0 єдиний.

\triangleleft Доведення від супротивного. Нехай існують $\Phi_0^{(1)} \neq \Phi_0^{(2)}$ – два елементи найкращого наближення. Візьмемо $\alpha \in [0, 1]$, тоді

$$\begin{aligned} \Delta(f) &\leq \|f - \alpha \Phi_0^{(1)} - (1 - \alpha) \Phi_0^{(2)}\| = \|\alpha(f - \Phi_0^{(1)}) + (1 - \alpha)(f - \Phi_0^{(2)})\| \leq \\ &\leq \alpha \|f - \Phi_0^{(1)}\| + (1 - \alpha) \|f - \Phi_0^{(2)}\| = \alpha \Delta(f) + (1 - \alpha) \Delta(f) = \Delta(f) \end{aligned}$$

Тобто всі “ \leq ” можна замінити на “ $=$ ”. Отримаємо

$$\|\alpha(f - \Phi_0^{(1)}) + (1 - \alpha)(f - \Phi_0^{(2)})\| = \alpha \|f - \Phi_0^{(1)}\| + (1 - \alpha) \|f - \Phi_0^{(2)}\|$$

За припущенням $\exists \lambda$ таке, що $\alpha(f - \Phi_0^{(1)}) = \lambda(1 - \alpha)(f - \Phi_0^{(2)})$. Виберемо

$\alpha = \frac{1}{2}$. Тоді $(f - \Phi_0^{(1)}) = \lambda(f - \Phi_0^{(2)})$. Оскільки $\|f - \Phi_0^{(1)}\| = \|f - \Phi_0^{(2)}\| = \Delta(f)$, то

остання рівність має місце тільки для $\lambda = 1$. Звідси

$$f - \Phi_0^{(1)} = f - \Phi_0^{(2)} \Rightarrow \Phi_0^{(1)} = \Phi_0^{(2)}.$$

Отже, ми отримали протиріччя з припущенням, що і доводить існування єдиного елемента найкращого наближення. \triangleright

Теорема 3 Гільбертів простір H – строго нормований.

\triangleleft ($H : (u, v) \forall u, v \in H \quad \|u\| = \sqrt{(u, u)}$). Нехай

$$\|f + g\| = \|f\| + \|g\|, \quad x, y \in H \quad (6)$$

$$\|f + g\|^2 = \|f\|^2 + 2\|f\|\|g\| + \|g\|^2.$$

З іншого боку

$$\|f + g\|^2 = (f + g, f + g) = \|f\|^2 + 2(f, g) + \|g\|^2$$

Звідси $\|f\|\|g\| = (f, g)$. Для довільного гільбертового простору $(f, g) \leq \|f\|\|g\|$.

Таким чином на елементах (6) нерівність Коші – Буняковського перетворюється в рівність. Розглянемо

$$\|f - \lambda g\|^2 = \|f\|^2 - 2\lambda(f, g) + \lambda^2 \|g\|^2 = \|f\|^2 - \lambda\|f\|\|g\| + \lambda^2 \|g\|^2 = (\|f\| - \lambda\|g\|)^2,$$

Тоді для $\lambda = \frac{\|f\|}{\|g\|}$ маємо $\|f - \lambda g\| = 0$. Звідси $\exists \lambda : f = \lambda g$, тобто H – строго нормований. \triangleright

Наслідок $R = H \Rightarrow \exists! \Phi_0 \in M_n$.

Приклади строго нормованих просторів:

1) $L_2(a, b)$ з нормою $\|u\| = \sqrt{\int_a^b u^2 dx}$.

2) $L_p(a, b)$ з нормою $\|u\|_p = \left(\int_a^b u^p dx \right)^{\frac{1}{p}}$, $p > 1$.

Простір $C[a, b]$ не є строго нормованим, але в ньому існує єдиний елемент найкращого наближення (про цей факт в наступному пункті).

Найкраще рівномірне наближення

Найкраще рівномірне наближення – це наближення в просторі $R = C[a, b]$, де $\|f\|_{C[a, b]} = \max_{x \in [a, b]} |f(x)|$ – рівномірна метрика.

Теорема 1 (Хаара) Для того, щоб $\forall f \in C[a, b]$ існував єдиний елемент найкращого рівномірного наближення необхідно і достатньо, щоб система $\{\varphi_i\}_{i=0}^{\infty}$ була системою Чебишова.

Система $\{\varphi_i\}_{i=0}^{\infty}$ називається системою Чебишова, якщо елемент $\Phi_n(x) = \sum_{i=0}^n c_i \varphi_i(x)$ має не більше n нулів, причому $\sum_{i=1}^n c_i^2 \neq 0$. Наприклад, системою Чебишова є поліноміальна система $\{x^i\}_{i=0}^{\infty}$.

Позначимо $Q_n^0(x)$ – багаточлен найкращого рівномірного наближення (далі – БНРН). Його відхилення від f : $\Delta(f) = \|Q_n^0(x) - f(x)\|_C = \inf_{Q_n(x)} \|Q_n(x) - f(x)\|$.

Теорема 2 (Чебишова) $Q_n^0(x)$ – БНРН неперервної функції $f(x)$ тоді та тільки тоді, якщо на відрізку $[a, b]$ існує хоча б $(n+2)$ -а точки $a \leq x_0 \dots \leq x_m \leq b$, $m \geq n+1$ такі, що

$$f(x_i) - Q_n^0(x_i) = \alpha (-1)^i \Delta(f), \quad (1)$$

де $i = \overline{0, m}$, $\alpha = \pm 1$.

Точки $\{x_i\}_{i=0}^m$, які задовольняють умовам теореми Чебишова, називають точками чебишовського альтернансу.

Теорема 3 $Q_n^0(x)$ – БНРН для неперервної функції єдиний.

\triangleleft Припустимо, існують два БНРН степеня n : $Q_n^{(1)}(x) \neq Q_n^{(2)}(x)$:

$$\Delta(f) = \|f - Q_n^{(1)}\|_C = \|f - Q_n^{(2)}\|_C.$$

Звідси випливає, що

$$\left\| f - \frac{Q_n^{(1)} + Q_n^{(2)}}{2} \right\| \leq \left\| \frac{f - Q_n^{(1)}}{2} \right\| + \left\| \frac{f - Q_n^{(2)}}{2} \right\| = \Delta(f),$$

тобто багаточлен $\frac{Q_n^{(1)}(x) + Q_n^{(2)}(x)}{2}$ також є БНРН. Нехай x_0, x_1, \dots, x_m – відповідні йому точки чебишовського альтернансу.

Це означає, що

$$\left| \frac{Q_n^{(1)}(x_i) + Q_n^{(2)}(x_i)}{2} - f(x_i) \right| = \Delta(f),$$

або

$$\left[Q_n^{(1)}(x_i) - f(x_i) \right] + \left[Q_n^{(2)}(x_i) - f(x_i) \right] = 2\Delta(f) \quad (2)$$

Так як $\left| Q_n^{(k)}(x_i) - f(x_i) \right| \leq \Delta(f)$, $k = 1, 2$, то (2) можливе лише у тому випадку, коли

$$Q_n^{(1)}(x_i) - f(x_i) = Q_n^{(2)}(x_i) - f(x_i), \quad i = \overline{0, n+1}.$$

Звідки випливає, що $Q_n^{(1)}(x) = Q_n^{(2)}(x)$, а це суперечить початковому припущенню. \triangleright

Приклади побудови БНРН

Скінченного алгоритму побудови БНРН для довільної функції не існує. Є ітераційний. Але в деяких випадках можна побудувати БНРН за теоремою Чебишова.

1⁰. Потрібно наблизити $f(x)$ багаточленом нульового степеня.

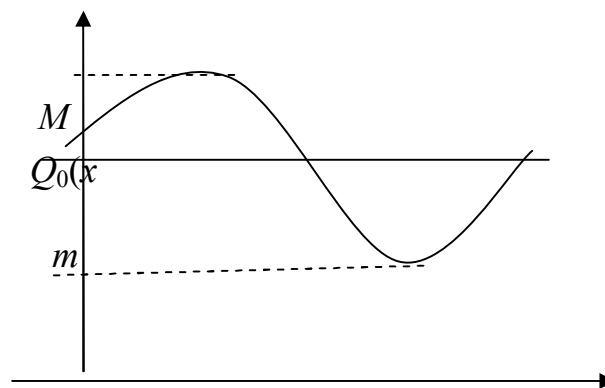


Рис. 11

Нехай $M = \max_{[a,b]} f(x) = f(x_0)$, $m = \min_{[a,b]} f(x) = f(x_1)$. тоді $Q_0(x)$ – БНРН має вигляд (див. рис. 11):

$$Q_0(x) = \frac{M + m}{2},$$

де $f(x_0) = \frac{M+m}{2} = \frac{M-m}{2}$, $f(x_1) = \frac{M+m}{2} = m - \frac{M+m}{2} = -\frac{M+m}{2}$,
 $\Delta(f) = \frac{M+m}{2}$, а x_0, x_1 – точки чебишовського альтернансу.

2⁰. Опукла функція $f(x) \in C[a, b]$ наближається багаточленом першого степеня

$$Q_1(x) = c_0 + c_1x.$$

Оскільки $f(x)$ опукла, то різниця $f(x) - (c_0 + c_1x)$ може мати лише одну внутрішню точку екстремуму. Тому точки a, b є точками чебишовського альтернансу. Нехай ξ третя – точка чебишовського альтернансу. Згідно з теоремою Чебишова, маємо систему:

$$\begin{cases} f(a) - c_0 - c_1a = \alpha\Delta(f) \\ f(\xi) - c_0 - c_1\xi = -\alpha\Delta(f) \\ f(b) - c_0 - c_1b = \alpha\Delta(f) \end{cases}$$

Звідси $f(b) - f(a) = c_1(b - a)$ та $c_1 = \frac{f(b) - f(a)}{b - a}$.

Цю систему треба замкнути, використавши ще одне рівняння з умови: точка ξ є точкою екстремуму різниці $f(x) - (c_0 + c_1x)$. Тому для диференційованої функції $f(x)$ для визначення ξ маємо рівняння (дотична і січна паралельні):

$$f'(\xi) = c_1 = \frac{f(b) - f(a)}{b - a}.$$

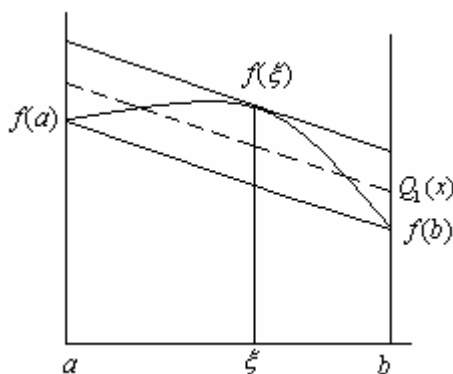


Рис. 12

Геометрично ця процедура виглядає наступним чином (див. рис. 12). Проводимо січну через точки $(a, f(a)), (b, f(b))$. Для неї тангенс кута дорівнює c_1 . Проводимо паралельну їй дотичну до кривої $y = f(x)$, а потім пряму, рівновіддалену від січної та дотичної, яка і буде графіком $Q_1(x)$. При цьому $x_0 = a$, $x_1 = \xi$, $x_2 = b$.

3⁰. Потрібно наблизити $f(x) = x^{n+1}$, $x \in [-1, 1]$ багаточленом степеня n $Q_n^0(x)$. Введемо

$$\begin{aligned}\bar{P}_{n+1}(x) &= x^{n+1} - Q_n(x) = x^{n+1} - a_1 x^n - \dots \\ \Delta(f) &= \inf_{Q_n(x)} \|x^{n+1} - Q_n^0(x)\|_C = \inf_{P_{n+1}} \|\bar{P}_{n+1} - 0\|_C = \|\bar{T}_{n+1}(x)\| \Rightarrow \\ x^{n+1} - Q_n^0(x) &= \bar{T}_{n+1}(x) \Rightarrow Q_n^0(x) = x^{n+1} - \bar{T}_{n+1}(x).\end{aligned}$$

Задача. Для прикладу 3^0 вказати точки чебишовського альтернансу $\{x_i\}, i = \overline{0, n+1}$.

4^0 . Потрібно наблизити $f(x) = P_{n+1}(x) = a_0 + \dots + a_{n+1}x^{n+1}$, $a_{n+1} \neq 0$, $x \in [a, b]$ БНРН степеня n . Запишімо його у вигляді:

$$Q_n^0(x) = P_{n+1}(x) - a_{n+1}\bar{T}_{n+1}^{[a,b]},$$

де $\bar{T}_{n+1}^{[a,b]}(x)$ - нормований багаточлен Чебишова на проміжку $x \in [a, b]$.

Дійсно це БНРН: вираз у правій частині є багаточленом степеня n , оскільки коефіцієнт при x^{n+1} дорівнює нулю, а його нулі $x_k = \frac{b+a}{2} + \frac{b-a}{2}t_k, t_k = \cos \frac{2k+1}{2(n+1)}\pi, k = \overline{0, n}$ є точками чебишевського альтернансу для $Q_n^0(x)$.

Задача. Показати, що для $f(x)$ парної (непарної) функції БНРН це багаточлен по парних (непарних) степенях x .

5^0 . Телескопічний метод. Дуже часто БНРН точно знайти не вдається. В таких випадках шукається багаточлен, близький до нього. Бажано щоб цей багаточлен був невисокого степеня (менше арифметичних операцій на його обчислення) Спочатку будують такий багаточлен $P_n(x) = \sum_{j=0}^n a_j x^j$, щоб

відхилення від $f(x)$ була достатньо малою. (наприклад меншою за $\frac{\varepsilon}{2}$).

Можна це зробити, наприклад, за формулою Тейлора. Потім наближають багаточлен $P_n(x)$ багаточленом найкращого рівномірного наближення $P_{n-1}(x)$ (за алгоритмом п. 4; для простоти $x \in [-1, 1]$):

$$P_{n-1}(x) = P_n(x) - a_n T_n(x) 2^{1-n}.$$

Оскільки $|T_n(x)| \leq 1$ на відріжку $[-1, 1]$, то

$$|P_{n-1}(x) - P_n(x)| \leq |a_n| 2^{1-n}.$$

Далі наближають багаточлен $P_{n-1}(x)$ багаточленом найкращого рівномірного наближення $P_{n-2}(x)$ і т. д. Пониження степеня продовжується до тих пір, поки сумарна похибка від таких послідовних апроксимацій залишається меншою за задане мале число ε .

Найкраще середньоквадратичне наближення

Наблизимо функцію $f(x) \in H$ з гільбертового простору H функціями з скінченно-вимірного підпростору M_n простору H . Тут H гільбертів

простір із скалярним добутком (u, v) , норма і відстань для якого визначаються формулами:

$$\|u\| = \sqrt{(u, u)}, \Delta(u, v) = \|u - v\|.$$

Побудуємо

$$u = \sum_{i=0}^n c_i \varphi_i \in M_n \subset H, \quad (1)$$

де $\{\varphi_i\}_{i=0}^{\infty}$ - лінійно незалежна система елементів з H .

Елемент найкращого середньоквадратичного наближення (в подальшому ЕНСКН) Φ_0 такий, що

$$\|f - \Phi_0\| = \sqrt{(f - \Phi_0, f - \Phi_0)} = \inf_{\Phi \in M_n} \|f - \Phi\|.$$

Теорема 1 Нехай $f \in H$, $\Phi_0 \in M_n$ - елемент найкращого середньоквадратичного наближення

$$\|f - \Phi_0\| = \inf_{\Phi \in M_n} \|f - \Phi\|.$$

Тоді

$$(f - \Phi_0, \Phi) = 0, \forall \Phi \in M_n \quad (2)$$

◁ Нехай (2) не виконується, тобто $\exists \Phi_1$:

$$(f - \Phi_0, \Phi_1) = \alpha \neq 0, \Phi_1 \in M_n, \|\Phi_1\| = 1.$$

Побудуємо $\Phi_2 = \Phi_0 + \alpha\Phi_1$,

$$\|f - \Phi_2\|^2 = (f - \Phi_2, f - \Phi_2) = \|f - \Phi_0\|^2 - \alpha^2 < \|f - \Phi_0\|^2.$$

Отже, елемент Φ_2 кращий за елемент найкращого середньоквадратичного наближення Φ_0 . А це суперечність. ▷

Наслідок

$f = \Phi_0 + v$, де $\Phi_0 \in M_n$, а $v \perp M_n$ (поправка v з ортогонального доповнення до M_n).

Знайти ЕНСКН

$$\Phi_0 = \sum_{i=0}^n c_i \varphi_i \quad (3)$$

означає знайти коефіцієнти c_i .

Для виконання (2) достатньо, щоб

$$(f - \Phi_0, \varphi_k) = 0, k = \overline{0, n}. \quad (4)$$

Підставимо (3) у формулу (4):

$$\left(f - \sum_{i=0}^n c_i \varphi_i, \varphi_k \right) = 0.$$

Таким чином маємо СЛАР для c_i :

$$\sum_{i=0}^n c_i (\varphi_i, \varphi_k) = (f, \varphi_k), k = \overline{0, n} \quad (5)$$

З теореми 1 витікає лише достатність умов (5) для знаходження коефіцієнтів c_i . Розглянемо задачу $\|f - \Phi_0\| = \inf_{\varphi \in M_n} \|f - \Phi\|$, як задачу мінімізації функції багатьох змінних:

$$F(a_0, \dots, a_n) \equiv \|f - \Phi\|^2 = \left\| f - \sum_{i=0}^n a_i \varphi_i \right\|^2.$$

Умови мінімуму цієї функції приводять до (5).

Задача. Показати, що для коефіцієнтів c_i елемента найкращого середньо квадратичного наближення умови (5) є необхідними та достатніми.

Матриця СЛАР (5) складається з елементів $g_{ik} = (\varphi_i, \varphi_k)$, тобто це матриця Грамма: $G = (g_{ik})_{i,k=0}^n$. Оскільки це матриця Грамма лінійно незалежної системи, то $\det G \neq 0$, що ще раз доводить існування та єдиність ЕНСКН. Оскільки $G^T = G$, то для розв'язку цієї системи використовують метод квадратних коренів.

Якщо взяти $x \in [0, 1]$ та $\varphi_i = x^i$, $i = \overline{0, n}$, $H = L_2(0, 1)$, то

$$g_{ik} = \int_0^1 x^i x^k dx = \frac{1}{i+k+1}, \quad i, k = \overline{0, n}.$$

Це матриця Гілберта, яка є погано обумовленою: $\text{cond} G \cong 10^7$, $n = 6$. Праві частини

$$f_k = (f, \varphi_k) = \int_0^1 f(x) x^k dx,$$

як правило, обчислюються наближено, тому похибки обчислення c_i можуть бути великими.

Що робити? Якщо вибирати систему $\{\varphi_i\}_{i=0}^{\infty}$ ортонормованою, тобто

$$(\varphi_i, \varphi_k) = \delta_{ik} = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases},$$

то система (5) має явний розв'язок:

$$\Phi_0 = \sum_{i=0}^n (f, \varphi_i) \varphi_i, \quad (6)$$

Якщо $\{\varphi_i\}$ – повна ортонормована система, то довільну функцію можна представити у вигляді ряду Фур'є:

$$f = \sum_{i=0}^{\infty} (f, \varphi_i) \varphi_i, \quad (7)$$

і $f - \Phi_0 = \sum_{i=n+1}^{\infty} c_i \varphi_i = v$ - залишок (похибка). Таким чином ЕНСКН є відрізком ряду Фур'є. Далі

$$\begin{aligned} \|f - \Phi_0\|^2 &= (f - \Phi_0, f - \Phi_0) = \|f\|^2 - 2(f, \Phi_0) + \|\Phi_0\|^2 = \\ &= \|f\|^2 - 2\|\Phi_0\|^2 - \underbrace{2(v, \Phi_0)}_{=0 \text{ за теоремою 1}} + \|\Phi_0\|^2 = \\ &= \|f\|^2 - \|\Phi_0\|^2 = \sum_{i=0}^{\infty} c_i^2 - \sum_{i=0}^n c_i^2 = \underbrace{\sum_{i=n+1}^{\infty} c_i^2}_{\text{квадрат похибки}} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Останнє витікає з відповідної теореми математичного аналізу. Таким чином, якщо $\{\varphi_i\}_{i=0}^{\infty}$ - повна ортонормована система, то

$$\sum_{i=n+1}^{\infty} c_i \xrightarrow{n \rightarrow \infty} 0 \text{ та } \Phi_0 \xrightarrow{n \rightarrow \infty} f.$$

Значить вірна

Теорема 2 В гільбертовому просторі H послідовність ЕНСКН $\{\Phi_0^{(n)}\}$ по повній ортонормованій системі $\{\varphi_i\}_{i=0}^{\infty}$ збігається до f .

Зауваження 1. Відхилення можна обчислити за формулою:

$$\Delta^2(f) = \|f - \Phi_0\|^2 = \|f\|^2 - 2(f, \Phi_0) + \|\Phi_0\|^2 = \|f\|^2 - \|\Phi_0\|^2 = \|f\|^2 - \sum_{i=0}^n c_i^2.$$

Якщо $\{\varphi_i\}_{i=0}^{\infty}$ - ортогональна система, але не нормована, тобто $(\varphi_i, \varphi_k) = \delta_{ik} \|\varphi_i\|^2$, то

$$c_i = \frac{(f, \varphi_i)}{\|\varphi_i\|^2}, \quad \Phi_0 = \sum_{i=0}^n \frac{(f, \varphi_i) \varphi_i}{\|\varphi_i\|^2}, \quad \|f - \Phi_0\|^2 = \|f\|^2 - \sum_{i=0}^n \frac{c_i^2}{\|\varphi_i\|^2}.$$

Для функції $f(x)$, щоб побудувати ЕНСКН покладемо $H = L_{2,\alpha}(a, b)$, в якому скалярний добуток виберемо наступним чином

$$(u, v) = \int_a^b u(x)v(x)d\alpha(x) \text{ (інтеграл Стільт'єса),}$$

де $\alpha(x)$ - зростаюча функція. Можливі випадки:

1. $\alpha(x) \in C^1[a, b]$, тоді $\alpha'(x) = \rho(x) > 0$ та $(u, v) = \int_a^b \rho(x)u(x)v(x)dx$;
2. $\alpha(x)$ - функція стрибків, $\alpha(x) = \alpha(x_k - 0)$, де $x_{k-1} \leq x \leq x_k$, $k = \overline{1, N}$. Якщо ввести $\rho_k = \alpha(x_k + 0) - \alpha(x_k - 0)$, то $(u, v) = \sum_{i=1}^n \rho_i u(x_i)v(x_i)$.

Перший вибір $\alpha(x)$ використовується при апроксимації функцій неперервного аргументу, а другий - для табличних функцій.

Системи ортогональних функцій

Як вибрати ортонормальну або ортогональну систему функцій $\{\varphi_i\}_{i=0}^{\infty}$? Розглянемо деякі з найбільш вживаних таких систем.

1⁰. Якщо $H = L_2(-1,1)$; $\rho \equiv 1$ (ваговий множник), то $\varphi_i(x) = L_i(x)$ - система багаточленів Лежандра, які мають вигляд

$$L_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n.$$

Використовують також рекурентні формули

$$(n+1)L_{n+1}(x) = (2n+1)xL_n(x) - nL_{n-1}(x),$$

до яких додаємо умови

$$L_0(x) = 1, L_1(x) = x.$$

Це ортогональна система в тому сенсі, що

$$(L_i, L_k) = \int_{-1}^1 L_i(x)L_k(x)dx = \delta_{ik} \|L_i(x)\|^2, \quad \|L_i(x)\|^2 = \frac{2}{2i+1}$$

і тому $c_i = \frac{(f, L_i)}{\|L_i\|^2} = \frac{2i+1}{2} (f, L_i)$.

Зауваження Якщо потрібно побудувати наближення на довільному проміжку (a,b) , то бажано перейти до проміжку $(-1,1)$, тобто по $f(x)$ на $[a,b]$ побудувати $\bar{f}(t)$ з $t \in [-1,1]$ заміною $x = At + B$, $t = \alpha x + \beta$ та для побудови багаточлена НСКН для $\bar{f}(t)$ використати багаточлени Лежандра $L_i(t)$.

Можна робити навпаки - систему багаточленів перевести з $[a,b]$ на $[-1,1]$, але це вимагає більше обчислень і процес побудови ЕНСН складніше.

2⁰. Якщо $H = L_{2,\rho}(-1,1)$, $\rho(x) = \frac{1}{\sqrt{1-x^2}}$, скалярний добуток

$(u, v) = \int_{-1}^1 \frac{u(x)v(x)}{\sqrt{1-x^2}} dx$ (це невластні інтеграли другого роду), то $\varphi_i(x) = T_i(x)$, де $\{T_i(x)\}$ система ортогональних багаточленів Чебишова 1-го роду, які мають вигляд

$$T_n(x) = \cos(n \arccos(x)).$$

Рекурентна формула для цих багаточленів:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad T_0 = 1 \quad T_1 = x,$$

$$\|T_n\|^2 = \begin{cases} \pi, & n=0 \\ \frac{\pi}{2}, & n=1,2,\dots \end{cases}$$

3⁰. H гільбертів простір з ваговим множником $\rho(x) = (1-x)^\alpha (1+x)^\beta$.

Система $\varphi_i(x) = P_n^{(\alpha,\beta)}(x)$ - багаточленів Якобі, $\alpha, \beta > -1$ (α, β - числові параметри) ортогональна в сенсі скалярного добутку

$(u, v) = \int_{-1}^1 (1-x)^\alpha (1+x)^\beta u(x)v(x)dx$. Ця система є узагальненням випадків 1⁰ та

2⁰. Диференціальна формула для багаточленів:

$$P_n^{(\alpha, \beta)}(x) = \frac{(-1)^n}{2^n n!} (1-x)^{-\alpha} (1+x)^{-\beta} \frac{d^n}{dx^n} [(1-x)^{n+\alpha} (1+x)^{n+\beta}];$$

Рекурентна формула:

$$\begin{aligned} & 2(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta)P_{n+1}^{(\alpha, \beta)}(x) = \\ & = (2n+\alpha+\beta+1)[(2n+\alpha+\beta)(2n+\alpha+\beta+2)x + \alpha^2 - \beta^2]P_n^{(\alpha, \beta)}(x) - \\ & - 2(n+\alpha)(n+\beta)(2n+\alpha+\beta+2)P_{n-1}^{(\alpha, \beta)}(x), \quad n=1, 2, \dots \end{aligned}$$

де

$$P_0^{(\alpha, \beta)} = 1, \quad P_{-1}^{(\alpha, \beta)} = 0, \quad \|P_n^{(\alpha, \beta)}\|^2 = \frac{2^{\alpha+\beta+1} \Gamma(\alpha+n+1) \Gamma(\beta+n+1)}{n!(\alpha+\beta+2n+1) \Gamma(\alpha+\beta+n+1)},$$

та

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt, \quad \Gamma(z+1) = z\Gamma(z), \quad \Gamma(n+1) = n!$$

Коли $\alpha = \beta = 0$: $P_n^{(0,0)}(x) = L_n(x)$, а для $\alpha = \beta = -\frac{1}{2}$: $P_n^{(-\frac{1}{2}, -\frac{1}{2})}(x) = T_n(x)$.

4⁰. $H = L_{2,\rho}[0, \infty)$, $\rho(x) = x^\alpha e^{-x}$, $\alpha > -1$. Цьому ваговому множнику відповідає система багаточленів Лагерра $\varphi_i(x) = L_i^\alpha(x)$, які задаються диференціальною формулою:

$$L_n^\alpha(x) = (-1)^n x^{-\alpha} e^x \frac{d^n}{dx^n} [x^{\alpha+n} e^{-x}],$$

або в рекурентній формі

$$(n+1)L_{n+1}^\alpha = (2n+\alpha+1-x)L_n^\alpha - (n+\alpha)L_{n-1}^\alpha,$$

де $L_0^\alpha = 1$, $L_{-1}^\alpha = 0$ та з нормою $\|L_n^\alpha\|^2 = n! \Gamma(\alpha+n+1)$.

5⁰. $H = L_{2,\alpha}(-\infty, \infty)$, $\rho(x) = e^{-x^2}$. Систему ортогональних функцій вибираємо як систему багаточленів Ерміта $\varphi_i(x) = H_i(x)$, які задаються диференціальною формулою:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2})$$

або в рекурентній формі

$$H_{n+1} = 2xH_n - 2nH_{n-1},$$

де $H_0 = 1$, $H_{-1} = 0$ та $\|H_n\|^2 = 2^n n! \sqrt{\pi}$.

6⁰. $H = L_2(0, 2\pi)$, $\rho(x) \equiv 1$, $f(x) = f(x+2\pi)$. $f(x)$ – 2π -періодичні функції. За систему ортогональних функцій вибираємо тригонометричну систему

$$\varphi_0 = \frac{1}{\sqrt{2\pi}}, \quad \varphi_{2k-1}(x) = \frac{1}{\sqrt{\pi}} \cos(kx), \quad \varphi_{2k}(x) = \frac{1}{\sqrt{\pi}} \sin(kx),$$

де

$$\|\varphi_m\|^2 = 1.$$

Елемент найкращого середньоквадратичного наближення представляє собою *тригонометричний багаточлен*

$$\Phi_0(x) = T_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx)),$$

формули для обчислення цих коефіцієнтів наведені в наступному пункті.

7⁰. Якщо потрібно апроксимувати табличну функцію, то

$$H = l_2, \quad x_i \equiv i, \quad i = \overline{0, N}, \quad (u, v) = \frac{1}{N+1} \sum_{i=0}^N u_i v_i$$

і за систему ортогональних функцій вибираємо наступну систему багаточленів $\varphi_k(x) = p_k^{(N)}(x)$, $k = \overline{0, m}$ ($m \leq N$) – систему багаточленів Чебишова дискретного аргументу, які задається формулою

$$p_k^{(N)}(x) = \sum_{j=0}^k \frac{(-1)^j C_k^j C_{k+j}^j}{N^{(j)}} x^{(j)},$$

$x^{(j)} = x(x-1)\dots(x-j+1)$ - факторіальний багаточлен; C_k^j - число сполук.

Рекурентна формула

$$\frac{(m+1)(N-m)}{2(2m+1)} p_{m+1}^{(N)} = \left(\frac{N}{2} - x\right) p_m^{(N)} - \frac{m(N+m+1)}{2(2m+1)} p_{m-1}^{(N)}, \quad p_0^{(N)} = 1, \quad p_{-1}^{(N)} = 0.$$

Наприклад $p_1^{(N)} = 1 - \frac{2x}{N}$, $p_2^{(N)} = 1 - \frac{6x}{N} + \frac{6x^2}{N(N-1)}$.

У випадку, якщо задані вузли $t_i = t_0 + ih$, $i = \overline{0, N}$, то робимо заміну $x_i = \frac{t_i - t_0}{h} = i$.

Середньоквадратичне наближення періодичних функцій

Нехай маємо періодичну функцію $f(x)$ неперервного аргументу, з періодом $T = 2\pi$, тобто $f(x + 2\pi) = f(x)$. В просторі $H_2 = L_2[0, 2\pi]$ визначений скалярний добуток

$$(u, v) = \int_0^{2\pi} u(x)v(x)dx.$$

В якості системи лінійно-незалежних функцій $\{\varphi_i\}$ виберемо тригонометричну систему функцій:

$$\varphi_0(x) = 1; \quad \varphi_{2k-1}(x) = \cos kx; \quad \varphi_{2k}(x) = \sin kx, \quad k = \pm 1, \pm 2, \dots;$$

яка є повною нормованою системою в $L_2[0, 2\pi]$.

Будемо шукати $\Phi(x)$ у вигляді тригонометричного багаточлена

$$\Phi(x) \equiv T_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx). \quad (1)$$

За теорією найкращого середньоквадратичного наближення коефіцієнти обчислюємо за формулами:

$$\begin{cases} a_0 = (f, \varphi_0) = \frac{1}{2\pi} \int_0^{2\pi} f(x) dx, \\ a_k = (f, \varphi_k^c) = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx dx, \\ b_k = (f, \varphi_k^s) = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx dx. \end{cases} \quad (2)$$

Відхилення:

$$\Delta^2(f) = \|f\|^2 - \left(2\pi a_0^2 + \sum_{k=1}^n \pi(a_k^2 + b_k^2) \right).$$

Тепер нехай функція $f(x)$ задана таблично:

$$f_i = f(x_i), i = \overline{1, N}.$$

Тригонометрична система $\varphi_0(x), \varphi_{2k-1}(x), \varphi_{2k}(x)$ ортогональна в $H = L_2(\omega)$ для $\omega = \left\{ x_i = i \frac{\pi}{N}, i = \overline{1, N} \right\}$ в сенсі скалярного добутку

$$(u, v) = \frac{1}{N} \sum_{i=1}^N u_i v_i, u_i = u(x_i).$$

Тоді

$$\begin{cases} a_0 = \frac{1}{N} \sum_{i=1}^N f_i, \\ a_k = \frac{2}{N} \sum_{i=1}^N f_i \cos kx_i, \\ b_k = \frac{2}{N} \sum_{i=1}^N f_i \sin kx_i. \end{cases} \quad (3)$$

Це формули Бесселя. В формулі (1) $\Phi(x) \equiv T_n(x)$ (тобто багаточлен той же), але коефіцієнти визначаємо за формулою (3).

Зауваження Як правило кількість даних значень $N \gg 2n + 1$. Але якщо

$N = 2n + 1$, то $n = \frac{N-1}{2}$ і N -непарне. При цьому $T_{\frac{N-1}{2}}(x)$ – БНСКН і звідси

$$\Delta^2(f) = \left\| f(x) - T_{\frac{N-1}{2}}(x) \right\|^2 = \frac{1}{N} \sum_{i=1}^N \left[f(x_i) - T_{\frac{N-1}{2}}(x_i) \right]^2 \rightarrow \inf_{a_k, b_k}$$

Оскільки найменше значення відхилення $\Delta^2(f) = 0$, то тригонометричний багаточлен найкращого середньоквадратичного наближення співпадає з інтерполяційним тригонометричним багаточленом і

$$T_{\frac{N-1}{2}}(x_i) = f(x_i).$$

Для визначення коефіцієнтів a_i, b_i за формулою Бесселя (3) необхідна кількість операцій $Q = O(N^2)$. Існують алгоритми, які дозволяють обчислити за $Q = O(N \log N)$ операцій. Це так званий алгоритм швидкого перетворення Фур'є. Якщо в (3) існує група доданків, які рівні між собою, тобто число N можна представити як $N = p_1 p_2$, то можна так вибрати сітку, що $Q = O(N \max(p_1, p_2))$. Якщо ж $N = n^m$, то $Q = O(Nm) = O(N \log_2 N)$.

Метод найменших квадратів (МНК)

Нехай в результаті вимірювань функції $f(x)$ маємо таблицю значень:

$$y_i \approx f(x_i), \quad i = \overline{1, N}, \quad x_i \in [a, b]. \quad (1)$$

За даними цієї таблиці треба побудувати аналітичну формулу $\Phi(x; a_0, a_2, \dots, a_n)$ таку, що

$$\Phi(x_i; a_0, a_2, \dots, a_n) \approx y_i, \quad i = \overline{1, N}. \quad (2)$$

Виконувати це інтерполюванням тобто задавати

$$\Phi(x_i; a_0, a_2, \dots, a_n) = y_i, \quad i = \overline{1, N} \quad (3)$$

нераціонально, бо $N \gg n$ і система перевизначена; її розв'язки як правило не існують. Вигляд функції $\Phi(x; a_0, a_2, \dots, a_n)$ і число параметрів a_i у деяких випадках відомі. В інших випадках вони визначаються за графіком, побудованим за відомими значеннями $f(x_i)$ так, щоб залежність (2) була досить простою і добре відображала результати спостережень. Але такі міркування не дають змогу побудувати єдиний елемент та й ще найкращого наближення.

Тому визначають параметри a_0, \dots, a_n так, щоб у деякому розумінні всі рівняння системи (2) одночасно задовольнялись з найменшою похибкою, наприклад, щоб виконувалося

$$I(a_0, \dots, a_n) = \sum_{i=1}^N [y_i - \Phi(x_i; a_0, \dots, a_n)]^2 \rightarrow \min. \quad (4)$$

Такий метод розв'язання системи (2) і називають методом найменших квадратів, оскільки мінімізується сума квадратів відхилення $\Phi(x; a_0, a_2, \dots, a_n)$ від значень $f(x_i)$.

Для реалізації мінімуму необхідно та достатньо виконання умов:

$$\frac{\partial I}{\partial a_i} = 0, \quad i = \overline{0, n}. \quad (5)$$

Якщо $\Phi(x_i; a_0, \dots, a_n)$ лінійно залежить від параметрів a_0, \dots, a_n , тобто

$$\Phi(x_i; a_0, \dots, a_n) = \sum_{k=0}^n a_k \varphi_k(x), \quad (6)$$

то з (3) маємо СЛАР

$$\sum_{k=0}^n a_k \varphi_k(x_i) = y_i, \quad i = \overline{1, N}, \quad (7)$$

яку називають *системою умовних рівнянь*. Позначивши

$$C = (\varphi_k(x_i))_{\substack{i=\overline{1, N} \\ k=\overline{0, n}}}, \quad \vec{a} = (a_0, \dots, a_n)^T, \quad \vec{y} = (y_1, \dots, y_N)^T,$$

маємо матричний запис СЛАР (7)

$$C\vec{a} = \vec{y} \quad (8)$$

Помноживши систему умовних рівнянь (8) зліва на транспоновану до C матрицю C^T отримаємо систему *нормальних рівнянь*

$$C^T C \vec{a} = C^T \vec{y} \quad (9)$$

$$G = A = C^T C, \quad \dim G = n + 1, \quad G = [g_{ik}]_{i,k=0}^n,$$

$$g_{ik} = \sum_{j=1}^N c_{ij}^T c_{jk} = \sum_{j=1}^N c_{ji} c_{jk} = \sum_{j=1}^N \varphi_k(x_j) \varphi_j(x_i), \quad C^T \vec{y} = \left(\sum_{i=1}^N c_{ik} y_i \right)_{k=0}^n,$$

з якої власно і обчислюють невідомі коефіцієнти.

Покажемо, що МНК є методом знаходження ЕНСКН, якщо визначити скалярний добуток

$$(u, v) = \sum_{i=1}^N u(x_i) v(x_i).$$

Поставимо задачу знаходження ЕНСКН:

$$\Delta(f, \Phi) = \|f - \Phi\|^2 = (f - \Phi, f - \Phi) = \sum_{i=1}^N (y_i - \Phi(x_i, \vec{a}))^2 \rightarrow \inf.$$

За теорією середньоквадратичного наближення для цього необхідно, щоб коефіцієнти a_0, \dots, a_n знаходилися з системи:

$$\sum_{j=0}^n a_k (\varphi_k, \varphi_j) = (\varphi_k, f), \quad k = \overline{0, n},$$

а це співпадає з (9).

Якщо відома інформація про обчислювальну похибку для значень $f(x_i)$

$$|f(x_i) - y_i| < \varepsilon_i,$$

то вибирають такий скалярний добуток $(u, v) = \sum_{i=1}^N \rho_i u(x_i) v(x_i)$, де $\rho_i = \frac{1}{\varepsilon_i^2}$.

Нехай тепер $\Phi(x, a_0, \dots, a_n)$ - нелінійна функція параметрів $\vec{a} = (a_0, \dots, a_n)$, наприклад:

$$\Phi = a_0 e^{a_1 x} + a_2 e^{a_3 x} + \dots,$$

або

$$\Phi = a_0 \cos a_1 x + a_2 \sin a_3 x + \dots$$

Складемо функціонал:

$$S(a_0, \dots, a_n) = \sum_{i=1}^N \rho_i [y_i - \Phi(x, \vec{a})]^2 \rightarrow \inf_{\vec{a}} \quad (10)$$

Оскільки тепер $\Phi(x, a_0, \dots, a_n)$ нелінійна, то застосуємо метод лінеаризації.

Нехай відомі наближені значення $\vec{a}^0 = (a_0^0, \dots, a_n^0)$. Розкладемо $\Phi(x, \vec{a})$ в околі \vec{a}^0 . Тоді отримаємо лінійне наближення до $\Phi(x, \vec{a})$:

$$\Phi(x, \vec{a}) \approx \Phi(x, \vec{a}^0) + \sum_{k=0}^n \frac{\partial \Phi}{\partial a_k}(x, \vec{a}^0) (a_k - a_k^0).$$

Якщо ввести позначення

$$\vec{z} = \vec{a} - \vec{a}^0 \quad y_i^* = y_i - \Phi(x, \vec{a}^0), \quad c_{ik} = \Phi'_{a_k}(x_i, \vec{a}^0),$$

то отримаємо систему умовних рівнянь відносно поправок до \vec{a}^0 :

$$C\vec{z} = \vec{y}^* \quad (11)$$

Замінімо її на систему нормальних рівнянь

$$C^T C\vec{z} = C^T \vec{y}^* \quad (12)$$

Знайшовши \vec{z} , обчислюємо наступне наближення: $\vec{a}^1 = \vec{a}^0 + \vec{z}$. Цей процес можна продовжувати: на кожній ітерації знаходимо $\vec{z}^m, m = 0, 1, \dots$ і уточнюємо наближення до \vec{a} : $\vec{a}^m = \vec{a}^{m-1} + \vec{z}^{m-1}$.

Умова припинення ітерацій

$$\|\vec{z}^m\| = \left(\sum_{k=0}^n (z_k^m)^2 \right)^{\frac{1}{2}} < \varepsilon.$$

Важливим є вибір початкового наближення \vec{a}^0 . З системи умовних рівнянь (нелінійної) виберемо деякі $n+1$. Розв'язок цієї системи і дасть початкове наближення.

Для деяких простих нелінійних залежностей від невеликої кількості параметрів задачу можна лінеаризувати аналітично. Наприклад, розглянемо наближення даних алометричним законом

$$y_i \approx f(x_i), \quad \Phi(x, A, \alpha) = Ax^\alpha.$$

Система умовних рівнянь має вигляд:

$$\Phi(x_i) = Ax_i^\alpha = y_i, \quad i = \overline{1, N}.$$

Прологарифмуємо її:

$$\psi(x_i) = \ln \Phi(x_i) = \ln A + \alpha \ln x_i = \ln y_i, \quad i = \overline{1, N}$$

Введемо $a = \ln A$. Тепер функція $\psi(x, a, \alpha)$ лінійна. Система умовних рівнянь відносно параметрів a та α має вигляд.

$$C\vec{z} = \vec{b}, \quad \vec{z} = (a, \alpha), \quad \vec{b} = (\ln y_i)_{i=1}^N,$$

$$C = \begin{pmatrix} 1 & \ln x_1 \\ \dots & \dots \\ 1 & \ln x_N \end{pmatrix}.$$

Запишемо систему нормальних рівнянь для методу найменших квадратів

$$C^T C \bar{z} = C^T \bar{b}, \quad (13)$$

$$G = C^T C = \begin{pmatrix} N & \sum_{i=1}^N \ln x_i \\ \sum_{i=1}^N \ln x_i & \sum_{i=1}^N (\ln x_i)^2 \end{pmatrix}, \quad C^T \bar{b} = \begin{pmatrix} \sum_{i=1}^N \ln y_i \\ \sum_{i=1}^N \ln x_i \ln y_i \end{pmatrix}.$$

Розв'язавши систему (13), знаходимо і α , та $A = \exp(a)$.

Згладжуючі сплайни

Якщо значення в точках x_i неточно $\tilde{f}_i = f_i + \varepsilon_i$, то застосовують згладжування. Для цього треба побудувати нову таблицю із згладженими значеннями \bar{f}_i .

Наведемо деякі прості формули згладжування:

$$m = 1: \quad \bar{f}_i = \frac{1}{3} [\tilde{f}_{i-1} + \tilde{f}_i + \tilde{f}_{i+1}], \quad N = 3.$$

$$\bar{f}_i = \frac{1}{5} [\tilde{f}_{i-2} + \dots + \tilde{f}_{i+2}], \quad N = 5.$$

$$\bar{f}_i = \frac{1}{N} \left[\tilde{f}_{i-\frac{N}{2}} + \dots + \tilde{f}_{i+\frac{N}{2}} \right], \quad N - \text{парне.}$$

$$m = 3: \quad \bar{f}_i = \frac{1}{3 \cdot 5} [-3\tilde{f}_{i-2} + 12\tilde{f}_{i-1} + 17\tilde{f}_i + 12\tilde{f}_{i+1} - 3\tilde{f}_{i+2}], \quad N = 5.$$

Їх отримуємо в такий спосіб: до \tilde{f}_i застосовуємо апроксимацію, будуємо багаточлен НСКН

$$Q_m(x) = \sum_{k=0}^m c_k p_k^N(x),$$

де p_k^N – система багаточленів Чебишова дискретного аргументу. Беремо значення

$$\bar{f}_i = Q_m(x_i),$$

які приводять до наведених вище формул.

Але ці формули не дають гарантію, що в результаті ми отримаємо функцію, яка задовольняє умові:

$$|\bar{f}_i - f_i| < \varepsilon_i$$

Згладжуючі сплайни дають можливість побудувати наближення з заданою точністю. Нагадаємо деякі відомості про сплайни. Явний вигляд кубічного сплайна:

$$s(x) = m_{i-1} \frac{(x_i - x)^3}{6h_i} + m_i \frac{(x - x_{i-1})^3}{6h_i} + \left(f_{i-1} - \frac{m_{i-1}h_i^2}{6} \right) \frac{x_i - x}{h_i} + \left(f_i - \frac{m_i h_i^2}{6} \right) \frac{x - x_{i-1}}{h_i}, \quad x \in [x_{i-1}, x_i], \quad h_i = x_i - x_{i-1}. \quad (1)$$

Тут $s(x_i) = f_i, i = \overline{0, n}$, а $m_i = s''(x_i)$ задовольняють систему:

$$\begin{cases} \frac{h_i}{6} m_{i-1} + \frac{h_i + h_{i+1}}{3} m_i + \frac{h_{i+1}}{6} m_{i+1} = \frac{f_{i+1} - f_i}{h_{i+1}} - \frac{f_i - f_{i-1}}{h_i}, & i = \overline{1, n-1}, \\ m_0 = m_n = 0. \end{cases} \quad (2)$$

В матричній формі ця система має вигляд:

$$A\vec{m} = H\vec{f}. \quad (3)$$

Тут

$$\vec{m} = (m_1, \dots, m_{n-1})^T, \quad \vec{f} = (f_0, \dots, f_n)^T,$$

$$A = \underbrace{\begin{pmatrix} h_1 + h_2 & h_2 & 0 \\ 3 & 6 & \\ h_2 & h_2 + h_3 & \dots \\ 6 & 3 & \dots \\ 0 & \dots & \dots \end{pmatrix}}_{n-1}, \quad H = \underbrace{\begin{pmatrix} \frac{1}{h_1} & -\left(\frac{1}{h_1} + \frac{1}{h_2}\right) & \frac{1}{h_2} & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots \end{pmatrix}}_{n+1}.$$

Кубічний інтерполяційний сплайн мінімізує функціонал:

$$\Phi(u) = \int_a^b (u'')^2 dx:$$

$$\Phi(s) = \inf_{u \in U} \Phi(u) \quad U = \left\{ u(x) : u(x_i) = f_i, i = \overline{0, n}, u(x) \in W_2^2(a, b) \right\} \quad (5)$$

Введемо функціонал

$$\Phi_1(u) = \Phi(u) + \sum_{i=0}^n \rho_i [\tilde{f}_i - u(x_i)]^2.$$

Згладжуючим сплайном назвемо функцію g , яка є розв'язком задачі:

$$\Phi_1(g) = \inf_{u \in W_2^2(a, b)} \Phi_1(u), \quad (5)$$

Перший доданок в $\Phi_1(u)$ дає мінімум „згину”, другий – середньоквадратичне наближення до значень \tilde{f}_i . Покажемо, що g є сплайном.

Нехай існує функція $g(x)$. Побудуємо кубічний сплайн такий, що $s(x_i) = g(x_i)$. З того, що $g(x)$ є розв'язком задачі (5), маємо

$$\Phi_1(s) \geq \Phi_1(g) \Rightarrow \int_a^b (s'')^2 dx + \sum \rho_i [\tilde{f}_i - s(x_i)]^2 \geq \int_a^b (g'')^2 dx + \sum \rho_i [\tilde{f}_i - g(x_i)]^2$$

Звідси

$$\Phi(s) \geq \Phi(g).$$

Так як кубічний інтерполяційний сплайн $s(x)$ мінімізує функціонал (4), то $\Phi(s) \leq \Phi(g)$. Тому $\Phi(s) = \Phi(g)$. Звідки $s = g$.

Позначимо

$$\mu_i = g(x_i), i = \overline{0, n} \quad (6)$$

Якщо б ми знали μ_i , то для побудови g достатньо було б розв'язати систему

$$A\vec{m} = H\vec{\mu} \quad (7)$$

Підставимо (1) та (6) в $\Phi_1(g)$:

$$\Phi_1(g) = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} \left(m_{i-1} \frac{x_i - x}{h_i} + m_i \frac{x - x_{i-1}}{h_i} \right)^2 dx + \sum_{i=0}^n \rho_i (\tilde{f}_i - \mu_i)^2 = \inf \Phi_1(u). \quad (8)$$

Після перетворень маємо

$$\begin{aligned} \Phi_1(g) &= \sum_{i=1}^n \int_{x_{i-1}}^{x_i} \left(m_{i-1} \frac{x_i - x}{h_i} + m_i \frac{x - x_{i-1}}{h_i} \right)^2 dx + \sum_{i=0}^n \rho_i (\tilde{f}_i - \mu_i)^2 = \\ &= \sum_{i=1}^n m_i \left(m_{i-1} \frac{h_i}{6} + \frac{h_i + h_{i+1}}{3} m_i + \frac{h_{i+1}}{6} m_{i+1} \right) + \sum_{i=0}^n \rho_i (\tilde{f}_i - \mu_i)^2 = \\ &= (A\vec{m}, \vec{m}) + \sum_{i=0}^n \rho_i (\tilde{f}_i - \mu_i)^2 \end{aligned}$$

Задача. Показати, що для кубічного згладжуючого сплайну g має місце формула:

$$\begin{aligned} \Phi_1(g) &= \sum_{i=1}^n \int_{x_{i-1}}^{x_i} \left(m_{i-1} \frac{x_i - x}{h_i} + m_i \frac{x - x_{i-1}}{h_i} \right)^2 dx + \sum_{i=0}^n \rho_i (\tilde{f}_i - \mu_i)^2 = \\ &= \sum_{i=1}^n m_i \left(m_{i-1} \frac{h_i}{6} + \frac{h_i + h_{i+1}}{3} m_i + \frac{h_{i+1}}{6} m_{i+1} \right) + \sum_{i=0}^n \rho_i (\tilde{f}_i - \mu_i)^2 = \\ &= (A\vec{m}, \vec{m}) + \sum_{i=0}^n \rho_i (\tilde{f}_i - \mu_i)^2. \end{aligned}$$

Так як $\Phi_1(g)$ представляє собою квадратичну функція відносно $\vec{m} = (m_0, \dots, m_N)$, то необхідною і достатньою умовою мінімуму є

$$\frac{\partial \Phi_1}{\partial \mu_j} = 0, j = \overline{0, n}.$$

Знаходимо:

$$\begin{aligned}
\frac{\partial \Phi_1}{\partial \mu_j} &= \frac{\partial}{\partial \mu_j} (A\vec{m}, \vec{m}) + 2\rho_j(\mu_j - \tilde{f}_j) = \\
&= 2 \left[\frac{\partial}{\partial \mu_j} (A\vec{m}, \vec{m}) \right] + 2\rho_j(\mu_j - \tilde{f}_j) = 2 \left(\frac{\partial}{\partial \mu_j} (H\vec{\mu}), \vec{m} \right) + \\
&+ 2\rho_j(\mu_j - \tilde{f}_j) = 2 \left(\frac{\partial \vec{m}}{\partial \mu_j}, H^T \vec{m} \right) + 2\rho_j(\mu_j - \tilde{f}_j) = \\
&= 2(H^T \vec{m})_j + 2\rho_j(\mu_j - \tilde{f}_j) = 0
\end{aligned}$$

Отже, з умови мінімізації функціоналу

$$\Phi(u) = \int_a^b (u'')^2 dx + \sum_{i=0}^n \rho_i (f_i - u_i)^2$$

ми отримали таку систему рівнянь :

$$2(H^T \vec{m})_i + 2\rho_i(\mu_i - f_i) = 0 \quad (9)$$

де, як і раніше μ_i – це невідомі значення згладжуючого сплайну :

$$\mu_i = s(x_i), \quad m_i = s''(x_i).$$

Можна записати (9) матричному вигляді, якщо ввести матрицю $R = \text{diag } \rho_i$:

$$H^T \vec{m} + R\vec{\mu} = R\vec{f}. \quad (10)$$

Тут \vec{f} – вектор заданих значень функції.

Таким чином маємо для \vec{m} та $\vec{\mu}$ дві системи (7) і (10). Виключаючи $\vec{\mu}$ отримаємо таку систему лінійних рівнянь

$$(A + HR^{-1}H^T)\vec{m} = H\vec{f} \quad (11)$$

Розв'язавши її, можемо обчислити

$$\vec{\mu} = \vec{f} - R^{-1}H^T \vec{m} \quad (12)$$

і підставити знайдені значення μ_i та m_i в формулу для сплайну

$$g(x) = m_{i-1} \frac{(x_i - x)^3}{6h_i} + m_i \frac{(x - x_{i-1})^3}{6h_i} + \left(\mu_{i-1} - \frac{m_{i-1}h_i^2}{6} \right) \frac{x_i - x}{h_i} + \left(\mu_i - \frac{m_i h_i^2}{6} \right) \frac{x - x_{i-1}}{h_i}. \quad (13)$$

Тепер звернемо увагу на матрицю системи (10)

$$A' = A + HR^{-1}H^T.$$

Оскільки матриці H , H^T – трьохдіагональні, то матриця $HR^{-1}H^T$ буде п'ятидіагональною, а тому п'ятидіагональною буде й A' . Розв'язують зазвичай системи з такими матрицями наступним чином :

1. або методом квадратних коренів; для матриць із такою структурою цей метод має складність $Q = O(nm) = O(2n) = O(n)$, так як в нашому випадку півширина діагональної смуги $m = 2$.

2. або методом п'ятидіагональної прогонки, що також має складність $O(n)$.

Зауваження: ρ_i вибирають так $\rho_i = \frac{1}{\varepsilon_i^2}$.

Контрольні запитання:

1. Загальна постановка задачі наближення.
2. Поняття елемента найкращого наближення.
3. Найкраще рівномірне наближення.
4. Найкраще середньоквадратичне наближення.
5. Поняття системи ортогональних функцій.
6. Середньоквадратичне наближення періодичних функцій.
7. Метод найменших квадратів.
8. Згладжуючі сплайни.

Література [2] - [7], [11].

Лекція 10. Чисельне диференціювання функцій

Мета лекції: ознайомитися з поняттям чисельного диференціювання функцій, навчитися проводити регуляризацію задач чисельного диференціювання.

План лекції:

1. Чисельне диференціювання функцій.
2. Некоректність задачі чисельного диференціювання.
3. Формули для знаходження першої та другої похідної.
4. Оцінки похибки чисельного диференціювання.
5. Регуляризація задачі чисельного диференціювання.

Задача чисельного диференціювання виникає у випадку коли необхідно обчислити похідну функції, значення якої задані таблицею. Нехай задано

$$f_i = f(x_i), \quad i = \overline{0, n}, \quad x_i \in [a, b].$$

Проінтерполюємо ці значення. Тоді

$$f(x) = L_n(x) + r_n(x), \quad (1)$$

де залишковий член $r_n(x)$ у формі Ньютона має вигляд:

$$r_n(x) = f(x; x_0, \dots, x_n) \omega_n(x), \quad \omega_n(x) = \prod_{i=0}^n (x - x_i).$$

Звідси

$$f^{(k)}(x) = L_n^{(k)}(x) + r_n^{(k)}(x). \quad (2)$$

За наближене значення похідної в точці x беремо $f^{(k)}(x) \approx L_n^{(k)}(x), x \in [a, b]$.

Оцінимо похибку наближення - $r_n^{(k)}(x)$. За формулою Лейбніца:

$$r_n^{(k)}(x) = \sum_{j=0}^k C_k^j f^{(j)}(x; x_0, \dots, x_n) \omega_n^{(k-j)}(x). \quad (3)$$

З властивості розділених різниць маємо для $f(x) \in C^{n+k+1}[a, b]$:

$$f^{(j)}(x; \underbrace{x_0, \dots, x_n}_{j+1}) = j! f(x; \dots, x_0, \dots, x_n) = \frac{j!}{(n+j+1)!} f^{(n+j+1)}(\xi_j), \quad \xi_j \in [a, b].$$

Остаточню вираз для похибки наближення похідної має вигляд:

$$r_n^{(k)}(x) = \sum_{j=0}^k \frac{k!}{(k-j)!(n+j+1)!} f^{(n+j+1)}(\xi_j) \omega_n^{(k-j)}(x), \quad (4)$$

Оцінка похибки матиме вигляд:

$$\left| f^{(k)}(x) - L_n^{(k)}(x) \right| \leq M \sum_{j=0}^k \frac{k!}{(k-j)!(n+j+1)!} \left| \omega_n^{(k-j)}(x) \right|,$$

де $M = \max_{0 \leq j \leq k} \max_{x \in [a,b]} \left| f^{(n+j+1)}(x) \right|$.

Нагадаємо, що процес інтерполювання розбіжний. Крім того, якщо $k > n$, то $L_n^{(k)}(x) \equiv 0$. Тому не можна брати великими значення n та k . Як правило $k = 1, 2$, іноді $k = 3, 4$. Відповідно, $n = k$, або $n = k + 1$, або $n = k + 2$.

Подивимося як залежить порядок збіжності процесу чисельного диференціювання від кроку. Нехай $x_i = x_0 + ih$, $h > 0$ – крок. Тоді за умови $x_n - x_0 = O(h)$:

$$\omega_n(x) = (x - x_0) \dots (x - x_n) = O(h^{n+1}), \quad x \in [x_0, x_n].$$

Перша похідна від $\omega_n(x)$ має порядок на одиницю менше, тобто

$$\omega_n'(x) = O(h^n).$$

Далі

$$r_n^{(k)}(x) = O(h^{n+1-k}), \quad \text{тому } f^{(k)}(x) - L_n^{(k)}(x) = O(h^{n+1-k}).$$

При умові $n \geq k$ останній вираз збігається до нуля, тобто

$$f^{(k)}(x) - L_n^{(k)}(x) \xrightarrow{h \rightarrow 0} 0. \quad (5)$$

Далі

$$r_n^{(k)}(x) = \underbrace{f(x; x_0; \dots; x_n) \omega_n^{(k)}(x)}_{O(h^{n+1-k})} + \underbrace{\sum_{j=1}^k C_k^j f^{(j)}(x; x_0; \dots; x_n) \omega_n^{(k-j)}(x)}_{O(h^{n+2-k})}.$$

Якщо

$$\omega_n^{(k)}(\bar{x}) = 0, \quad \text{то } r_n^{(k)}(\bar{x}) = O(h^{n+2-k}). \quad (6)$$

Точки $x = \bar{x}$ називаються *точками підвищеної точності формул чисельного диференціювання*.

Приклад 1 Виведемо формули чисельного диференціювання для $k = 1, n = 1$.

Виберемо точки $x_0, x_1 = x_0 + h$ і інтерполяційний багаточлен має вигляд:

$$L_1(x) = f_0 + (x - x_0) \frac{f_1 - f_0}{h}.$$

Для похідної отримаємо вираз:

$$f'(x) \approx L_1'(x) = \frac{f_1 - f_0}{h}, \quad x \in [x_0, x_1].$$

Розписавши за формулою Тейлора, отримаємо вираз для похибки:

$$r_1'(x) = \frac{f^{(3)}(\xi_1)}{3!}(x-x_0)(x-x_1) + \frac{f^{(2)}(\xi_0)}{2!}(2x-x_1-x_0) = O(h).$$

Якщо $2\bar{x} - x_1 - x_0 = 0$, то $r_1'(\bar{x}) = O(h^2)$. Тобто $\bar{x} = \frac{x_1 + x_0}{2}$ – точка підвищеної точності. Більш точно (див. приклад 3)

$$|r_1'(\bar{x})| \leq \frac{h^2}{24} M_3, \text{ де } M_3 = \max_{x \in [a, b]} |f^{(3)}(x)|.$$

Приклад 2 Аналогічно виведемо формули чисельного диференціювання для $k=1, n=2$. Виберемо точки $x_0, x_1 = x_0 + h, x_2 = x_0 + 2h$. Інтерполяційний поліном має вигляд:

$$L_2(x) = f_0 + (x-x_0)\frac{f_1-f_0}{h} + (x-x_0)(x-x_1)\frac{f_2-2f_1+f_0}{2h^2}.$$

Тоді замінимо $f'(x) \approx L_2'(x) = \frac{f_1-f_0}{h} + (2x-x_0-x_1)\frac{f_2-2f_1+f_0}{2h^2}$, $x \in [x_0, x_2]$.

Якщо сюди підставити $x = x_0$, то отримаємо $f'(x_0) \approx \frac{-f_2+4f_1-3f_0}{2h}$. Для

точки $x = x_1$ маємо $f'(x_1) \approx \frac{f_2-f_0}{2h} = f_{x,1}$. Для точки $x = x_2$ маємо

$f'(x_2) \approx \frac{f_0-4f_1+3f_2}{2h}$, $x \in [x_0, x_2]$. Для похибки маємо оцінку $r_2'(x) = O(h^2)$.

Позначимо для $x \in [x_i, x_{i+1}]$ $f_{x,i} = \frac{f_{i+1}-f_i}{h} \approx f'(x)$ (різницева похідна вперед);

для $x \in [x_{i-1}, x_i]$ - $f_{\bar{x},i} = \frac{f_i-f_{i-1}}{h} \approx f'(x)$ (різницева похідна назад);

для $x \in [x_{i-1}, x_{i+1}]$ - $f_{x,i} = \frac{f_{i+1}-f_{i-1}}{2h} \approx f'(x)$ (центральна різницева похідна).

Замість $f'(x_i)$ можна взяти будь-яке із значень: $f_{x,i}$, $f_{\bar{x},i}$ або $f_{x,i}$.

Задача. Знайти точки підвищеної точності формул чисельного диференціювання для $k=1, n=2$ і оцінити похибку в цих точках.

Приклад 3 При $n=1, k=1$ оцінимо точність формул чисельного диференціювання за формулою Тейлора.

а) Нехай $f(x) \in C^2[a, b]$. Тоді

$$f'(x_0) - \frac{f(x_0+h) - f(x_0)}{h} = f'(x_0) - \frac{1}{h} \left[f_0 + hf'_0 + \frac{h^2}{2} f''(\xi) - f_0 \right] = -\frac{h}{2} f''(\xi),$$

$$\left| f'(x_0) - \frac{f_1 - f_0}{h} \right| \leq \frac{M_2 h}{2}, \quad M_2 = \max_{[x_0, x_1]} |f''(\xi)|.$$

б) Нехай $f(x) \in C^3[a, b]$. Тоді, розписавши розклад по формулі Тейлора до третьої похідної, маємо оцінку:

$$f'(\bar{x}) - \frac{f(x_0 + h) - f(x_0)}{h} = f'(\bar{x}) - \frac{1}{h} \left[f(\bar{x}) + \frac{h}{2} f'(\bar{x}) + \frac{h^2}{8} f''(\bar{x}) + \frac{h^3}{48} f'''(\xi) - \right. \\ \left. - f(\bar{x}) + \frac{h}{2} f'(\bar{x}) - \frac{h^2}{8} f''(\bar{x}) + \frac{h^3}{48} f'''(\eta) \right] = -\frac{h^2}{24} f'''(\zeta). \\ \left| f'(\bar{x}) - \frac{f_1 - f_0}{h} \right| \leq \frac{h^2 M_3}{24}, \bar{x} = \frac{x_1 + x_0}{2}.$$

Задача. Показати, що якщо $f(x) \in C^3[a, b]$, то $\left| f'(x_1) - \frac{f_2 - f_0}{2h} \right| \leq \frac{M_3 h^2}{6}$.

Приклад 4 При $n = 2, k = 2$ маємо:

$$L_2(x) = f_{i-1} + \frac{f_i - f_{i-1}}{h} (x - x_{i-1}) + \frac{f_{i+1} - 2f_i + f_{i-1}}{h^2} (x - x_{i-1})(x - x_i); \\ L_2''(x) = \frac{f_{i+1} - 2f_i + f_{i-1}}{h^2}.$$

Для $f(x) \in C^4[a, b]$ оцінимо точність формул чисельного диференціювання за формулою Тейлора:

$$f''(x_1) - \frac{f_2 - 2f_1 + f_0}{h^2} = f''(x_1) - \frac{f(x_1 + h) - 2f(x_1) + f(x_1 - h)}{h^2} = \\ = f''(x_1) - \frac{1}{h^2} \left[f_1 + hf_1'' + \frac{h^2}{2} f_1'' + \frac{h^3}{6} f_1''' + \frac{h^4}{24} f^{(4)}(\xi_1) - 2f_1 + \right. \\ \left. + f_1 - hf_1'' + \frac{h^2}{2} f_1'' - \frac{h^3}{6} f_1''' + \frac{h^4}{24} f^{(4)}(\xi_2) \right] = \\ = \frac{h^2}{12} f^{(4)}(\xi) \quad \xi_1, \xi_2, \xi \in [x_0, x_2]$$

Отже, $\left| f_1'' - \frac{f_2 - 2f_1 + f_0}{h^2} \right| \leq \frac{M_4 h^2}{12}$.

Задача. Побудувати формулу чисельного диференціювання $k = 2, n = 2$ у випадку нерівновіддалених вузлів: $x_0, x_1 = x_0 + h_1, x_2 = x_1 + h_2$. Оцінити точність формули. Знайти точки підвищеної точності оцінити похибку.

Крім інтерполяційних формул для чисельного диференціювання можна застосовувати сплайни. Нехай $f_i = f(x_i)$. Побудуємо інтерполяційний сплайн першого степеня $s_1(x)$, для якого має місце оцінка $|f^{(k)}(x) - s_1^{(k)}(x)| = O(h^{2-k}), k = 0, 1$. Звідси при $k = 1$ маємо $|f'(x) - s_1'(x)| = O(h)$.

Для кубічного інтерполяційного сплайну $s_3(x)$ маємо для першої та другої похідних

$$|f^{(k)}(x) - s_3^{(k)}(x)| = O(h^{4-k}), \quad k = 1, 2.$$

Про обчислювальну похибку чисельного диференціювання

Нехай значення функції обчислені з деякою похибкою. Постає питання про вплив цих похибок на значення похідних обчислених за формулами чисельного диференціювання.

Перед цим зробимо зауваження про вплив збурення функції на значення звичайних похідних.

Нехай $f(x) \in C^1[a, b]$ і її збурення має вигляд:

$$\tilde{f}(x) = f(x) + \frac{1}{n} \sin(\omega x).$$

При $n \rightarrow \infty$ маємо $\|f(x) - \tilde{f}(x)\|_{C[a, b]} = \frac{1}{n} \rightarrow 0$ звідси $\tilde{f}(x) \xrightarrow{n \rightarrow \infty} f(x)$. Таким

чином це малі збурення. Маємо $\tilde{f}'(x) = f'(x) + \frac{\omega}{n} \cos(\omega x)$. Нехай $\omega = n^2$, тоді

$$\|\tilde{f}' - f'\|_{C[a, b]} = \frac{|\omega|}{n} = n \xrightarrow{n \rightarrow \infty} \infty.$$

Цей приклад ілюструє нестійкість оператора диференціювання. Є сподівання, що ця нестійкість має місце і для чисельного диференціювання.

Нехай $\tilde{f}_i = f_i + \delta_i$, $f_i = f(x_i)$, $i = \overline{0, n}$ $|\delta_i| \leq \delta$. Розглянемо вплив похибок δ_i на конкретних формулах чисельного диференціювання.

Приклад 1 Оцінімо вплив збурень на похибку обчислення першої похідної $n = 1, k = 1$.

$$f'_i - \frac{\tilde{f}_i - \tilde{f}_{i-1}}{h} = f'_i - \frac{f_i - f_{i-1}}{h} - \frac{\delta_i - \delta_{i-1}}{h},$$
$$\left| f'_i - \frac{\tilde{f}_i - \tilde{f}_{i-1}}{h} \right| \leq \left| f'_i - \frac{f_i - f_{i-1}}{h} \right| + \left| \frac{\delta_i - \delta_{i-1}}{h} \right| \leq \frac{M_2 h}{2} + \frac{2\delta}{h} \xrightarrow{h \rightarrow 0} \infty, \quad (1)$$

Таким чином, як і для аналітичного диференціювання, маємо некоректність: при малих збуреннях $|\delta_i| \leq \delta$ можуть бути як завгодно великі похибки, якщо $\frac{\delta}{h} \rightarrow \infty$ при $h \rightarrow 0$.

Мінімізуємо вплив цих збурень. Позначимо

$$\varphi(h) = \frac{M_2 h}{2} + \frac{2\delta}{h}.$$

Тоді мінімум цієї функції досягається для таких h :

$$\varphi'(h) = \frac{M_2}{2} - \frac{2\delta}{h^2} = 0, \quad h_0 = 2\sqrt{\frac{\delta}{M_2}}.$$

При такому значенні h оцінка похибки (1) така:

$$\varphi(h_0) = 2\sqrt{M_2\delta} = O\left(\delta^{\frac{1}{2}}\right)_{\delta \rightarrow 0} \rightarrow 0.$$

Приклад 2 Подивимося на вплив збурень на похибку обчислення першої похідної при використанні центральної різницевої похідної:

$$\left|f'_i - \frac{\tilde{f}_{i+1} - \tilde{f}_{i-1}}{2h}\right| \leq \left|f'_i - \frac{f_{i+1} - f_{i-1}}{2h}\right| + \left|\frac{\delta_{i+1} - \delta_{i-1}}{2h}\right| \leq \frac{M_3 h^2}{6} + \frac{\delta}{h} = \varphi(h).$$

З рівняння $\varphi'(h) = \frac{M_3 h}{3} - \frac{\delta}{h^2} = 0$ маємо: $h_0^3 = \frac{3\delta}{M_3}$, $h_0 = \sqrt[3]{\frac{3\delta}{M_3}}$. Отже,

$$\varphi(h_0) = \frac{M_3}{6} \sqrt[3]{\frac{9\delta^2}{M_3^2}} + \frac{\delta}{\sqrt[3]{\frac{3\delta}{M_3}}} = \frac{1}{2} \sqrt[3]{\frac{M_3\delta^2}{3}} + \sqrt[3]{\frac{M_3\delta^2}{3}} = \frac{3}{2} \sqrt[3]{\frac{M_3\delta^2}{3}} = O\left(\delta^{\frac{2}{3}}\right).$$

Таким чином швидкість збіжності при $\delta \rightarrow 0$ похибки формули чисельного диференціювання центральною похідною вища ніж для формули з прикладу 1 (похідна вперед або назад).

Задача. Дослідити похибку чисельного диференціювання для $n=2, k=2$, вибрати оптимальний крок h_0 , дати оцінку $\varphi(h_0)$.

Контрольні запитання:

1. Постановка задачі чисельного диференціювання функцій.
2. Формула для похибки наближення похідної.
3. Обчислювальна похибка чисельного диференціювання.
4. Означення точок підвищеної точності формул чисельного диференціювання.
5. Вплив збурень на похибку обчислення похідної.

Література [3], [4], [5], [7], [8], [11]

Лекція 11. Чисельне інтегрування функції

Мета лекції: ознайомитися із задачею чисельного інтегрування функції, засвоїти метод прямокутників та метод трапецій, стратегії вибору кроку сітки інтегрування.

План лекції:

1. Чисельне інтегрування функції однієї змінної.
2. Метод прямокутників та метод трапецій.
3. Оцінка похибки.
4. Блок-схеми методів.
5. Вибір кроку сітки.

Постановка задачі чисельного інтегрування

Нехай потрібно знайти

$$I(f) = \int_a^b \rho(x)f(x)dx \quad (1)$$

де f – задана функція, $\rho(x) > 0$ – деякий ваговий множник. Ця задача часто вимагає чисельного вирішення, оскільки

- значна кількість інтегралів типу (1) не можуть бути обчислені аналітично;
- інформація про функцію f може бути задана у вигляді таблиці.

Нагадаємо, що за означенням

$$I(f) = \lim_{\Delta \rightarrow 0} \sum_{i=1}^n \rho(\xi_i)f(\xi_i)\Delta x_i,$$

де $\Delta x_i = x_i - x_{i-1}$, а $\{x_i\}_{i=0}^n$ – розбиття проміжку $[a, b]$, $x_i \in [a, b]$, $\xi_i \in [x_{i-1}, x_i]$. Тому візьмо як наближення таку суму, яка називається *квадратурною формулою*:

$$I_n(f) = \sum_{k=0}^n c_k f(x_k), \quad (2)$$

де x_k – вузли квадратурної формули (2), а c_k – її вагові множники. Задача полягає в тим, щоб вибрати $\{x_k, c_k\}_{k=0}^n$, так щоб похибка була найменша:

$$R_n(f) = I(f) - I_n(f) \rightarrow \min$$

Квадратурну формулу (2) називають *квадратурною формулою замкнутого типу*, якщо $x_0 = a$ та $x_n = b$, і *відкритого типу*, якщо $x_0 > a$ та $x_n < b$.

Кажуть, що квадратурна формула (2) має *m -ий степінь алгебраїчної точності*, якщо

$$R_n(f) = 0, \forall f \in \pi_m, \quad (3)$$

π_m – множина поліномів m -го степеня, і $\exists P_{m+1}(x) \in \pi_{m+1}$, такий що $R_n(P_{m+1}) \neq 0$.

Цю умову можна замінити умовою

$$R_n(x^\alpha) = 0, \alpha = \overline{0, m}, R_n(x^{m+1}) \neq 0 \quad (4)$$

(вона більш зручна для перевірки).

Розглянемо деякі підходи до побудови квадратурних формул.

1) *Інтерполяційний*. Він приводить до квадратурних формул інтерполяційного типу. В інтегралі (1) покладають $f(x) \approx L_n(x)$ по деяких вузлах $\{x_k\}_{k=0}^n$ (вузли фіксовані). Тоді:

$$I_n(f) = I(L_n(x)) = \int_a^b \rho(x) \sum_{k=0}^n \frac{f(x_k)\omega_n(x)}{(x-x_k)\omega_n'(x_k)} dx = \sum_{k=0}^n f(x_k) \int_a^b \rho(x) \frac{\omega_n(x)}{(x-x_k)\omega_n'(x_k)} dx$$

Отже вузлами цієї квадратурної формули є вузли інтерполяційного багаточлена, а вагові множники

$$c_k = \int_a^b \rho(x) \frac{\omega_n(x)}{(x-x_k)\omega'_n(x_k)} dx.$$

2) *Найвищого алгебраїчного степеня точності.* Вибираємо одночасно x_k і c_k з умови $R_n(x^\alpha) = 0, \alpha = \overline{0, m}$, щоб m було максимальним. Отримуємо систему нелінійних алгебраїчних рівнянь, розв'язавши яку отримуємо квадратурні формули найвищого алгебраїчного степеня точності.

3) *Складені квадратурні формули.* Проміжок $[a, b]$ розбиваємо на підпроміжки (наприклад однокової довжини), а потім на кожному проміжку використовуємо, з невеликим ступенем, формули з пункту 1 або 2. Наприклад, для формул інтерполяційного типу:

$$I(f) = \sum_{i=1}^N \int_{x_{i-1}}^{x_i} \rho(x) f(x) dx \approx \sum_{i=1}^N \sum_{k=0}^n c_k^i f(x_k^i) = I_h(f).$$

Кажуть, що квадратурна формула складеного типу I_h має *порядок (ступінь) точності p по кроку h* , якщо $R_h(f) = I(f) - I_h(f) = O(h^p)$.

4). *Квадратурні формули оптимальні на класі функцій.* Вибираємо $\{x_k, c_k\}$ так, щоб досягався $\inf_{\{x_k, c_k\}} \sup_{f \in F} R_n(f)$. Це ми можемо робити, коли знаємо з

яким класом функцій маємо справу.

Зауваження 1 (про квадратурні формули інтерполяційного типу)

При підвищенні степеня інтерполяції погіршується якість наближення функції внаслідок розбіжності процесу інтерполяції: $\|f - L_n\|_c \xrightarrow{n \rightarrow \infty} \neq 0$. Але

$R_n(f) \xrightarrow{n \rightarrow \infty} 0$, наприклад, для $f \in C[a, b]$. І все ж таки розбіжність процесу інтерполювання дає ознаки: при $n \rightarrow \infty \max_k |c_k| \rightarrow \infty$ і це приводить

до поганих наслідків чисельного інтегрування. Дійсно, розглянемо випадок, коли функція задана неточними значеннями:

$$\tilde{f}(x_k) = f(x_k) + \delta_k, \quad |\delta_k| < \delta.$$

Тоді

$$\delta I_n(f) = I_n(\tilde{f}) - I_n(f) = \sum_{k=0}^n c_k \delta_k.$$

Якщо всі $c_k > 0$, то

$$|\delta I| = \sum_{k=0}^n c_k |\delta_k| \leq \delta \sum_{k=0}^n c_k = \delta(b-a).$$

При $\rho \equiv 1$, якщо підставити $f \equiv 1$, то отримаємо $b - a = \int_a^b dx = \sum_{k=0}^n c_k$. При

$\rho \neq 1$ $\sum_{k=0}^n c_k = \int_a^b \rho(x) dx$, бо хоча б нульовий степінь точності будь-яка квадратурна формула повинна мати.

Нагадаємо, що при $n \rightarrow \infty$ $\max_k |c_k| \rightarrow \infty$, а оскільки $\sum c_k > 0$, то $\exists c_k > 0$ і $\exists c_k < 0$, тому з ростом n зростає $|c_k|$, а відповідно і вплив похибки на результат. Тому не можна використовувати великі степені і використовують складені квадратурні формули.

Зауваження 2 Ясно, що квадратурні формули інтерполяційного типу мають алгебраїчний степінь точності принаймні $m = n$, бо ми заміняємо $f \rightarrow L_n$, а якщо $f \in \pi_n$, то $f \equiv L_n$. Але виявляється, що для парних n та симетричному розташуванні вузлів інтегрування, $m = n + 1$, тобто алгебраїчний степінь точності на одиницю вищий степеня інтерполяції.

Квадратурні формули прямокутників

Припустимо, що $\rho \equiv 1$. Тоді можна побудувати такі квадратурна формули інтерполяційного типу при $n = 0$:

а) лівих прямокутників: $x_0 = a$: $I_0^{лів} = (b - a)f(a)$;

б) правих прямокутників: $I_0^{прав} = (b - a)f(b)$ $x_0 = b$;

в) середніх прямокутників:

$$I_0 = (b - a)f(x_0), x_0 = \frac{a + b}{2}. \quad (1)$$

Знайдемо тепер алгебраїчну степінь точності цих квадратурних формул. Для лівих прямокутників:

$$I_0^{лів}(1) = b - a = I(1), I_0^{лів}(x) = a(b - a) \neq I(x) = \int_a^b x dx = \frac{x^2}{2} \Big|_a^b = \frac{b^2 - a^2}{2}.$$

Отже степінь точності $m = 0$. Така ж вона буде і для $I_0^{прав}$. А для середніх прямокутників

$$I_0(x) = (b - a) \frac{a + b}{2} = I(x), I_0(x^2) \neq I(x^2),$$

тому $m = 1$. Отож нею і будемо користуватися.

Оцінимо для неї похибку. Взагалі для формули інтерполяційного типу:

$$R_n(f) = I(f) - I_n(f) = I(f) - I(L_n) = I(f - L_n) = I(r_n) = \int_a^b r_n(x) dx,$$

де $r_n(x)$ – залишковий член інтерполяції. Далі

$$|R_n(f)| \leq (b-a) \max_x |r_n(x)| \leq (b-a) \frac{M_{n+1}}{n+1} \max_x |\omega_n(x)|.$$

Для I_0 :

$$|R_0(f)| = \left| \int_a^b r_0(x) dx \right| \leq \int_a^b |r_0(x)| dx \leq \int_a^b \frac{M_1}{1!} |x - x_0| dx = M_1 \int_a^b |x - x_0| dx \leq M_1 \frac{b^2 - a^2}{4},$$

Але це погана оцінка, вона не використовує той факт, що квадратурна формула має степінь точності на одиницю вищу. Отримаємо кращу оцінку.

Маємо при $f \in C^2[a, b]$:

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2} f''(\xi), \quad x_0 \equiv \frac{a+b}{2}$$

де $\xi \in [a, b]$. Тоді

$$\begin{aligned} R_0(f) &= \int_a^b f(x) dx - \int_a^b L_0(x) dx = \int_a^b [f(x) - f(x_0)] dx = \\ &= \int_a^b \left[(x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2} f''(\xi) \right] dx = \int_a^b \frac{(x - x_0)^2}{2} f''(\xi) dx = \\ &= f''(\eta) \int_a^b \frac{(x - x_0)^2}{2} dx = \frac{f''(\eta)}{24} (b - a)^3. \end{aligned}$$

Таким чином

$$|R_0(f)| \leq \frac{M_2}{24} (b - a)^3 \quad (2)$$

Але тут у нас немає впливу на точність (величину похибки). Тому використовують формулу складеного типу. Якщо сітка рівномірна, то складена квадратурна формула прямокутників має вигляд

$$I(f) = \sum_{i=1}^N \int_{x_{i-1}}^{x_i} f(x) dx \approx \sum_{i=1}^N h f(\bar{x}_i) = I_h^{np}(f), \quad (3)$$

де $\bar{x}_i = x_{i-\frac{1}{2}} = x_i - \frac{h}{2}$.

Оцінимо похибку цієї квадратурної формули:

$$R_h(f) = I(f) - I_h(f) = \sum_{i=1}^N \int_{x_{i-1}}^{x_i} [f(x) - f(\bar{x}_i)] dx = \sum_{i=1}^N f''(\eta_i) \frac{h^3}{24},$$

$$|R_h(f)| \leq \frac{M_2}{24} \sum_{i=1}^N h^3 = \frac{M_2 h^2}{24} \sum_{i=1}^N h = \frac{M_2 h^2 (b-a)}{24} \quad (4)$$

Тобто ця формула має степінь точності $p = 2$ по кроку h . (Не слід плутати з алгебраїчним степенем точності $m = 1$ для цієї формули).

Якщо $f(x) \in C^4[a, b]$, то

$$f(x) - f(\bar{x}_i) = f(\bar{x}_i) + (x - \bar{x}_i)f'(\bar{x}_i) + \frac{(x - \bar{x}_i)^2}{2} f''(\bar{x}_i) + \frac{(x - \bar{x}_i)^3}{6} f'''(\bar{x}_i) + \frac{(x - \bar{x}_i)^4}{24} f^{(4)}(\xi_i) - f(\bar{x}_i).$$

При непарних степенях інтеграли пропадуть і тому:

$$R_h(f) = \sum_{i=1}^N \int_{x_{i-1}}^{x_i} \frac{(x - \bar{x}_i)^2}{2} f''(\bar{x}_i) dx + \sum_{i=1}^N \int_{x_{i-1}}^{x_i} \frac{(x - \bar{x}_i)^4}{24} f^{(4)}(\xi_i) dx = \frac{h^2}{24} \sum_{i=1}^N h f''(\bar{x}_i) + \sum_{i=1}^N \frac{h^5 f^{(4)}(\eta_i)}{1920}.$$

Оскільки $\sum_{i=1}^N h f''(\bar{x}_i)$ - це квадратурна формула середніх прямокутників для $f''(x)$ з похибкою $O(h^2)$, то

$$R_h(f) = \frac{h^2}{24} \int_a^b f''(x) dx + O(h^4) = O(h^4).$$

$$R_h^{np}(f) = R_h(f) + \alpha(h), \text{ де } R_h(f) = \frac{h^2}{24} \int_a^b f''(x) dx, \alpha(h) = O(h^4). \quad (5)$$

Вона використовується для побудови програм, що автоматично вибирають крок інтегрування.

Формула трапеції

Нехай $x_0 = a$, $x_1 = b$, $L_1(x) = f(x)$. Тоді отримаємо формулу:

$$I_1(f) = \frac{b-a}{2} [f(a) + f(b)] \quad (1)$$

Формула має алгебраїчний степінь точності $m = 1$, оскільки $I(x^2) \neq I_1(x^2)$.

Це формула замкненого типу. Залишковий член:

$$R_1(f) = \int_a^b \frac{f''(\xi)(x-a)(x-b)}{2} dx = -\frac{(b-a)^3}{12} f''(\xi) \quad (2)$$

Оцінка залишкового члена:

$$|R_1(f)| \leq M_2 \frac{(b-a)^3}{12}. \quad (3)$$

З геометричної точки зору замінюється площа криволінійної трапеції площею звичайної трапеції.

Складена квадратурна формула трапецій:

$$I_h(f) = \sum_{i=1}^N \frac{h}{2} [f(x_{i-1}) + f(x_i)] = \frac{h}{2} f(a) + \sum_{i=1}^{N-1} hf(x_i) + \frac{h}{2} f(b), \quad (4)$$

де $x_i = a + ih$, $h = \frac{b-a}{N}$, $i = \overline{0, N}$ та

$$|R_h(f)| \leq M_2 \frac{(b-a)}{12} h^2, \quad f \in C^2[a, b], \quad (5)$$

Якщо $f \in C^4[a, b]$, то

$$R_h(f) = R_h^0(f) + \alpha(h), \quad \text{де } R_h^0(f) = -\frac{h^2}{12} \int_a^b f''(x) dx, \quad \alpha(h) = O(h^4). \quad (6)$$

Задача. Використовуючи явний вигляд головних членів похибки складених квадратурних формул прямокутників та трапецій, побудувати лінійною комбінацією цих двох формул квадратурну формулу четвертого степеня точності за кроком h .

Контрольні запитання:

1. Постановка задачі чисельного інтегрування.
2. Квадратурні формули прямокутників.
3. Формула трапеції.
4. Інтерполяційний підхід до побудови квадратурних формул.
5. Оцінка похибку квадратурної формули.

Література [2], [3], [5], [7], [9], [11].

Лекція 12. Метод Симпсона

Мета лекції: засвоїти теоретичні основи методу Симпсона, одержати практичні навички побудови алгоритмів чисельного інтегрування та їх реалізації на персональних комп'ютерах, ознайомитися з сучасним програмним забезпеченням чисельного інтегрування.

План лекції:

1. Метод Симпсона.
2. Оцінка похибки.
3. Блок-схема методу.
4. Порівняльний аналіз методів прямокутників, трапецій та Симпсона.

Квадратурна формула Симпсона

Нехай $x_0 = a$, $x_1 = \frac{a+b}{2}$, $x_2 = b$. Замість f використовуємо $L_2(x)$. Тоді отримаємо квадратурну формулу:

$$I_2(f) = \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right]. \quad (1)$$

Це квадратурна формула Симпсона.

Задача. Довести, що алгебраїчна степінь точності квадратурної формули Симпсона $m = 3$.

Задача. Довести, що для $f \in C^4[a, b]$ залишковий член квадратурної формули Симпсона має місце представлення:

$$R_2(f) = \frac{1}{24} \int_a^b (x-a) \left(x - \frac{a+b}{2}\right)^2 (x-b) f^{(4)}(\xi) dx = \frac{f^{(4)}(\xi)}{2880} (b-a)^5, \quad (2)$$

та вірна оцінка:

$$|R_2(f)| \leq \frac{M_4}{2880} (b-a)^5. \quad (3)$$

Складена квадратурна формула Симпсона має вигляд:

$$\begin{aligned} I_h(f) &= \sum_{i=1}^N \frac{h}{6} \left[f(x_{i-1}) + 4f\left(x_{i-\frac{1}{2}}\right) + f(x_i) \right] = \\ &= \frac{h}{6} \left[f(x_0) + 4f\left(x_{\frac{1}{2}}\right) + 2f(x_1) + \dots + 2f(x_{N-1}) + 4f\left(x_{N-\frac{1}{2}}\right) + f(x_N) \right] \end{aligned} \quad (4)$$

Якщо $f \in C^4[a, b]$, то має місце оцінка:

$$|R_h(f)| \leq \frac{M_4}{2880} (b-a)h^4, \quad p = 4. \quad (5)$$

Якщо $f \in C^6[a, b]$, то

$$R_h(f) = R_h^0(f) + \alpha(h), \text{ де } R_h^0(f) = \frac{h^4}{2880} \int_a^b f^{(4)}(x) dx, \alpha(h) = O(h^6). \quad (6)$$

Задача. Побудувати інтерполяційну квадратурну формулу для $n = 3$, $x_0 = a$, $x_1 = \frac{3a+b}{4}$, $x_2 = \frac{a+3b}{4}$, $x_3 = b$. Який алгебраїчний степінь точності вона має?

Принцип Рунге

Нехай задана деяка величина I (сіткова функція, інтеграл, неперервна функція). Нехай $I_h \approx I$ та $I_h \rightarrow I$ при $h \rightarrow 0$. Нехай похибка послідовності I_h представляється у вигляді

$$R_h = I - I_h = R_h^0 + \alpha(h) \quad (1)$$

де $R_h^0 = C \cdot h^m$ - головний член похибки, C не залежить від h , $\alpha(h) = o(h^m)$. Обчислимо $I_{h/2}$. З (1) слідує, що

$$I = I_h + Ch^m + \alpha(h), \quad I = I_{h/2} + C \frac{h^m}{2^m} + \alpha(h).$$

Звідси

$$I_{h/2} - I_h = \frac{Ch^m}{2^m} (2^m - 1) + \alpha(h).$$

З (1)

$$R_{h/2}^0 = \frac{Ch^m}{2^m} = \frac{I_{h/2} - I_h}{2^m - 1} \quad (2)$$

та $R_h^0 = \frac{2^m}{2^m - 1} (I_{h/2} - I_h)$. Формула (2) носить назву *апостеріорної оцінки* похибки обчислення I за допомогою наближення $I_{h/2}$. (*Апріорні оцінки* це оцінки отримані до обчислення величини I_h , *апостеріорні оцінки* – під час її обчислення).

З формули (2) витікає такий алгоритм обчислення інтегралу із заданою точністю ε :

- обчислюємо $I_h, I_{h/2}, R_{h/2}^0$; перевіряємо чи $|R_{h/2}^0| < \varepsilon$. Якщо так, то $I \approx I_{h/2}$.
Якщо ж ні, то
- обчислюємо $I_{h/2}, I_{h/4}, R_{h/4}^0$; перевіряємо $|R_{h/4}^0| < \varepsilon$ і т.д.
- Процес продовжуємо поки не буде виконана умова $|R_{h/2^k}^0| < \varepsilon$ $k = 1, 2, \dots$

Зауваження Ми даємо оцінку не похибки, а її головного члена з точністю $\alpha(h)$, тому такий метод може давати збої, якщо не виконана умова $|\alpha(h)| \ll |R_h^0|$.

За допомогою головного члена похибки можна отримати краще значення для I :

$$\tilde{I}_{h/2} = I_{h/2}^{(1)} = I_{h/2} + R_{h/2}^0 = \frac{2^m}{2^m - 1} I_{h/2} - \frac{1}{2^m - 1} I_h. \quad (3)$$

Це екстраполяційна формула Річардсона: $I_h - \tilde{I}_{h/2} = \alpha(h)$.

Для квадратурної формули трапецій $p = 2$ і

$$I - I_h = Ch^2 + O(h^4), \quad R_{h/2}^0 = \frac{I_{h/2} - I_h}{3}.$$

Маємо

$$R_h = -\frac{h^2}{12} \int_a^b f''(x) dx + O(h^4) = O(h^2).$$

Отже, якщо застосовувати екстраполяційну формулу Річардсона, то

$$\tilde{I}_{h/2} = \frac{4}{3} I_{h/2} - \frac{1}{3} I_h \quad (4)$$

і $I_h - \tilde{I}_{h/2} = O(h^4)$.

Задача. Написати явний вигляд квадратурної формули, яка отримується екстраполяцією Річардсона з квадратурної формули трапецій.

Якщо невідомо m , то можна використати принцип Рунге для його знаходження. Для цього використаємо $I_h, I_{h/2}, I_{h/4}$:

$$I_{h/2} - I_h = \frac{Ch^m}{2^m} (2^m - 1) + \alpha(h), \quad I_{h/4} - I_{h/2} = \frac{Ch^m}{2^{2m}} (2^m - 1) + \alpha(h).$$

З точністю $\alpha(h)$ маємо

$$2^m = \frac{I_{h/2} - I_h}{I_{h/4} - I_{h/2}}.$$

Звідки

$$m = \log_2 \frac{I_{h/2} - I_h}{I_{h/4} - I_{h/2}}.$$

Оцінка $|R_{h/4}^0| < \varepsilon$ - найбільш точна, тому $I \approx I_{h/4}$.

Покажемо чому формулу прямокутників рідко використовують з принципом Рунге. Нехай точки, в яких обчислюється значення функції позначаються: в $I_h \leftrightarrow \circ$, в $I_{h/2} \leftrightarrow *$.

Для формули трапецій використовуються такі точки:

$$\begin{array}{l} \circ - \quad \circ - \quad \circ - \quad \circ - \quad \circ - \quad \circ - \quad \circ - \quad \circ - \quad \circ - \quad \circ - \quad I_h \\ \circ - * - \circ - * - \circ - * - \circ - * - \circ - * - \circ - * - \circ - * - \circ - * - \circ - * - \circ - * - \circ - * - \circ - * - \circ - * - \circ - I_{h/2} \end{array}$$

Для формули прямокутників:

$$\begin{array}{l} - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad - \circ - \quad I_h \\ - * - I_{h/2} \end{array}$$

Як бачимо для формули трапецій необхідно підраховувати нові значення в N точках, а для формули прямокутників в $2N$.

Для економного використання обчислених значень функції на сітці з кроком h для формули трапецій запишемо

$$I_h^{mp} = \frac{h}{2} \left(f(a) + 2 \sum_{i=1}^{N-1} f(x_i) + f(b) \right),$$

$$I_{h/2}^{mp} = \frac{h}{4} \left(f(a) + 2 \sum_{i=1}^{N-1} f(x_i) + 2 \sum_{i=1}^{N-1} f(x_{i-1/2}) + f(b) \right) = \frac{1}{2} I_h^{mp} + \frac{h}{2} \sum_{i=1}^{N-1} f(x_{i-1/2}). \quad (5)$$

Отже, на одному кроці принципу Рунге кількість обчислень $Q^{mp} = O(N)$, а для $Q^{np} = O(2N)$.

Цей принцип застосовується і для формули Сімпсона $m = 4$. Головна частина залишкового члена для цієї формули:

$$R_{h/2}^0 = \frac{I_{h/2} - I_h}{15}.$$

$$\tilde{I}_{h/2} = \frac{16}{15} I_{h/2} - \frac{1}{15} I_h, \quad I_h - \tilde{I}_{h/2} = O(h^6)$$

Розглянемо використання так званих адаптивних квадратурних формул, в яких змінний крок вибирається за принципом Рунге. Для цього запишемо формулу трапецій із змінним кроком:

$$I_h^{mp}(f) = \sum_{i=1}^N \frac{h_i}{2} [f(x_{i-1}) + f(x_i)], \quad \text{де } h_i = x_i - x_{i-1}.$$

Оцінимо похибку на кожному інтервалі:

$$R_{h_i} = I_i - I_{h_i} = \int_{x_{i-1}}^{x_i} f(x) dx - \frac{h_i}{2} [f(x_{i-1}) + f(x_i)] = -\frac{h_i^3}{6} f''(x_{i-1/2}) + O(h_i^5).$$

Таким чином $p = 3$ і головний член похибки:

$$R_{h_{i/2}}^0 = \frac{(I_{h_{i/2}} - I_{h_i})}{7}.$$

Умова припинення ділення навпіл проміжку $[x_{i-1}, x_i]$:

$$\left| R_{h_{i/2}}^0 \right| \leq \frac{\varepsilon \cdot h_i}{b - a}.$$

Це забезпечує точність на всьому інтервалі

$$\left| R_{h/2} \right| = \left| \sum_{i=1}^N R_{h_i/2} \right| \leq \sum_{i=1}^N \frac{\varepsilon h_i}{b-a} = \varepsilon \frac{b-a}{b-a} = \varepsilon.$$

Ще одне застосування принципу Рунге – високоточне обчислення інтегралу від достатньо гладкої функції за допомогою таблиці Ромберга. Для побудови цієї таблиці обчислимо за допомоги складеної квадратурної формули трапецій із сталим кроком h послідовність значень $I_h = I_h^{(0)}, I_{h/2} = I_{h/2}^{(0)}, I_{h/4} = I_{h/4}^{(0)}, I_{h/8} = I_{h/8}^{(0)}, \dots$, які мають похибку $O(h^2)$. За допомогою екстраполяції Ричардсона (3) з коефіцієнтами лінійної комбінації $\left(\frac{4}{3}, -\frac{1}{3}\right)$ уточнимо ці значення (див. також формулу (4)). Отримаємо $I_{h/2}^{(1)}, I_{h/4}^{(1)}, I_{h/8}^{(1)}, \dots$. Вони мають похибку $O(h^4)$. Знову використовуємо екстраполяцію Ричардсона з коефіцієнтами лінійної комбінації $\left(\frac{16}{15}, -\frac{1}{15}\right)$. Отримаємо $I_{h/4}^{(2)}, I_{h/8}^{(2)}, \dots$, які мають точність $O(h^6)$ і т.д.. Отримані значення можна розмістити в такій таблиці Ромберга:

$$\begin{array}{cccc} I_h^{(0)} & & & \\ I_{h/2}^{(0)} & I_{h/2}^{(1)} & & \\ I_{h/4}^{(0)} & I_{h/4}^{(1)} & I_{h/4}^{(2)} & \\ I_{h/8}^{(0)} & I_{h/8}^{(1)} & I_{h/8}^{(2)} & I_{h/8}^{(3)} \end{array}$$

Всі значення крім останнього $I_{h/8}^{(3)}$ можна оцінити за принципом Рунге (див. формулу (2)). Використання формули (5) для обчислення $I_{h2^{-k}}^{(0)}$ та лінійні комбінації (2) дають простий та економічний алгоритм обчислення I . Початкове значення h можна брати рівним $b-a$, або $\frac{b-a}{n}$, де n ціле.

Квадратурні формули найвищого алгебраїчного степеня точності

Розглянемо інтеграл

$$I(f) = \int_a^b \rho(x) f(x) dx, \quad (1)$$

де

$$\rho(x) > 0 \quad x \in [a, b] \quad \left| \int_a^b \rho(x) x^i dx \right| < \infty.$$

Розглянемо задачу побудови квадратурної формули

$$I_n(f) = \sum_{k=1}^n c_k f(x_k), \quad (2)$$

яка при заданому n була б точною для алгебраїчного багаточлена можливо більшого степеня. Такі квадратурні формули існують, вони називаються *квадратурні формули найвищого алгебраїчного степеня точності* або формули Гаусса (або Гаусса – Кристофеля).

В (2) невідомими є $c_k, x_k, k = \overline{1, n}$. Їх обирають з умови, що (2) точна для будь-якого багаточлена степеня p , а це еквівалентно умові, щоб формула була точною для функції $f(x) = x^\alpha, \alpha = 0, 1, \dots, p$. Звідси отримуємо умови:

$$I_n(x^\alpha) = \int_a^b \rho(x)x^\alpha dx = \sum_{k=1}^n c_k x_k^\alpha, \quad a = \overline{0, p}. \quad (3)$$

Ми хочемо отримати формули для $m \rightarrow \max$. Щоб кількість рівнянь була рівною кількості невідомих нам потрібно, щоб $p + 1 = 2n$.

Задача. Побудувати квадратурну формулу найвищого степеня точності (розв'язати систему рівнянь (3)) для $a = -1, b = 1, \rho(x) = 1$.

Теорема Гаусса Квадратурна формула (2) буде точною для будь-якого багаточлена степеня $p = 2n - 1$, тобто $\forall f(x) \in \pi_{2n-1}$ тоді і тільки тоді, коли виконуються умови:

1) поліном $\omega(x) = (x - x_1)(x - x_2) \dots (x - x_n)$ ортогональний з вагою $\rho(x)$ до будь-якого багаточлена степеня менше n Q_{n-1} :

$$\int_a^b \omega(x) Q_{n-1}(x) \rho(x) dx = 0; \quad (4)$$

2) формула (2) є квадратурною формулою інтерполяційного типу, тобто коефіцієнти обчислюються за формулою

$$c_k = \int_a^b \rho(x) \frac{\omega(x)}{(x - x_k) \omega'(x_k)} dx \quad (5)$$

◁ Необхідність. Нехай формула (2) точна для багаточлена степеня $p = 2n - 1$, тобто $I(f) = I_n(f) \quad \forall f(x) \in \pi_{2n-1}$. Тоді

$$I(f) = \int_a^b \rho(x) \omega(x) Q_{n-1}(x) dx = \sum_{k=1}^n c_k \omega(x_k) Q_{n-1}(x_k) = 0.$$

Тобто виконується (4). Тепер покладемо

$$f(x) = \frac{\omega(x)}{(x - x_j) \omega'(x_j)} \in \pi_{n-1} \subset \pi_{2n-1}.$$

Отримаємо

$$\int_a^b \rho(x) f(x) dx = \int_a^b \rho(x) \frac{\omega(x)}{(x - x_j) \omega'(x_j)} dx = \sum_{k=1}^n c_k \frac{\omega(x_k)}{(x_k - x_j) \omega'(x_j)} = \sum_{k=1}^n c_k \delta_{kj} = c_j,$$

тобто виконується і умова (5).

Достатність. Нехай виконується (4) і (5). Подамо $\forall f(x) \in \pi_{2n-1}$ у вигляді

$$f(x) = \omega(x) Q_{n-1}(x) + R_{n-1}(x).$$

Розглянемо

$$I(f) = \int_a^b \rho(x) f(x) dx = \int_a^b \rho(x) (\omega(x) Q_{n-1}(x) + R_{n-1}(x)) dx =$$

$$= \sum_{k=1}^n c_k \omega(x_k) Q_{n-1}(x_k) + \sum_{k=1}^n c_k R_{n-1}(x_k)$$

Так як $R_{n-1}(x_k) = f(x_k) - \omega(x_k) Q_{n-1}(x_k) = f(x_k)$, то

$$I(f) = \sum_{k=1}^n c_k f(x_k) = I_n(f).$$

Тобто формула (2) є точною для будь-якого багаточлена степеня $2n-1$. \triangleright

Отже, з точністю до сталого множника багаточлени $\omega(x)$ співпадають з багаточленами n -того степеня ортогональної системи багаточленів. Ця система ортогональна на проміжку $[a, b]$ з вагою $\rho(x)$.

Вивчимо деякі властивості квадратурних формул Гаусса.

1) Покажемо, що c_k, x_k визначаються однозначно.

Представимо багаточлен $\omega(x)$ у вигляді $\omega(x) = a_0 + a_1 x + \dots + a_{n-1} x^{n-1} + x^n$.

Умови ортогональності (4) приймуть вигляд

$$\int_a^b \rho(x) \omega(x) x^\alpha dx = \int_a^b \rho(x) (a_0 + a_1 x + \dots + a_{n-1} x^{n-1} + x^n) x^\alpha dx = 0, \quad \alpha = \overline{0, n-1}$$

або

$$\int_a^b \rho(x) (a_0 + a_1 x + \dots + a_{n-1} x^{n-1}) x^\alpha dx = - \int_a^b \rho(x) x^n x^\alpha dx$$

Покажемо, що відповідна однорідна система рівнянь

$$\int_a^b \rho(x) (a_0 + a_1 x + \dots + a_{n-1} x^{n-1}) x^\alpha dx = 0, \quad \alpha = \overline{0, n-1} \quad (6)$$

має єдиний розв'язок $a_0 = a_1 = \dots = a_{n-1} = 0$.

Помножимо систему (6) на a_α і просумуємо по всіх $\alpha = \overline{0, n-1}$:

$$\sum_{\alpha=0}^{n-1} a_\alpha \int_a^b \rho(x) (a_0 + a_1 x + \dots + a_{n-1} x^{n-1}) x^\alpha dx =$$

$$= \int_a^b \rho(x) \sum_{\alpha=0}^{n-1} \sum_{j=0}^{n-1} a_j a_\alpha x^\alpha x^j dx = \int_a^b \rho(x) \left(\sum_{j=0}^{n-1} a_j x^j \right)^2 dx = 0.$$

Звідси і з умови $\rho(x) > 0$ випливає, що $a_0 = a_1 = \dots = a_{n-1} = 0$. Тому і відповідна неоднорідна система має єдиний розв'язок. Отже існує єдиний багаточлен $\omega(x)$ степеня n , який ортогональний з вагою $\rho(x)$ до будь-якого багаточлена степеня $n-1$.

2) Покажемо, що найвищий степінь точності формули Гаусса $p = 2n-1$. З теореми випливає, що $p \geq 2n-1$. Покажемо, що існує багаточлен степеня $2n$, для якого формула не є точною. Для цього введемо функцію $f(x) = \omega^2(x) \in \pi_{2n}$. Маємо

$$I(f) = \int_a^b \rho(x) \omega^2(x) dx > 0,$$

але

$$I_n(f) = \sum_{k=1}^n c_k \omega^2(x_k) = 0.$$

Отже, $I(f) \neq I_n(f)$. Звідси $p \leq 2n - 1$, тобто $p = 2n - 1$.

3) Коефіцієнти формул Гаусса додатні, тобто $c_k > 0$. Розглянемо багаточлени

$$\varphi_j = \left[\frac{\omega(x)}{(x - x_j) \omega'(x_j)} \right]^2,$$

які мають степінь $2n - 2$ і властивості:

$$\varphi_i(x_k) = \delta_{ik}, \quad I(\varphi_j) = \int_a^b \rho(x) \varphi_j(x) dx > 0.$$

Так як для цих багаточленів справедлива теорема Гаусса, то

$$I(\varphi_j) = I_n(\varphi_j) = \sum_{k=1}^n c_k \varphi_j(x_k) = \sum_{k=1}^n c_k \left[\frac{\omega(x_k)}{(x_k - x_j) \omega'(x_j)} \right]^2 = \sum_{k=1}^n c_k \delta_{jk}^2 = c_j.$$

Звідси випливає, що $c_j > 0$, $j = \overline{1, n}$.

4) Теорема Нехай вагова функція $\rho(x) > 0$ $x \in [a, b]$ $f(x) \in C^{2n}[a, b]$. Тоді існує точка $\xi \in [a, b]$ така, що

$$R_n(f) = \frac{f^{(2n)}(\xi)}{(2n)!} \int_a^b \rho(x) \omega^2(x) dx. \quad (7)$$

◁ Розглянемо інтерполяційний багаточлен Ерміта з двократними вузлами

$$H_{2n-1}(x): f(x_i) = H_{2n-1}(x_i), \quad f'(x_i) = H'_{2n-1}(x_i), \quad i = \overline{1, m}.$$

Для нього

$$r_{2n-1}(x) = f(x) - H_{2n-1}(x) = \frac{f^{(2n)}(\xi)}{(2n)!} \omega^2(x).$$

Звідси

$$R_{2n-1}(x) = \int_a^b \rho(x) r_{2n-1}(x) dx = \frac{f^{(2n)}(\xi)}{(2n)!} \int_a^b \rho(x) \omega^2(x) dx. \triangleright$$

Частинні випадки квадратурної формули Гаусса

1) Розглянемо відрізок $[-1, 1]$ і ваговий множник $\rho(x) = 1$, тобто виведемо формули Гаусса для обчислення інтегралу

$$\int_{-1}^1 f(x) dx.$$

Щоб знайти вузли квадратурна формули розглянемо багаточлени Лежандра

$$(n+1)L_{n+1}(x) = (2n+1)xL_n(x) - nL_{n-1}(x), \quad L_0 = 1, \quad L_1(x) = x.$$

Багаточлени Лежандра задовольняють умовам теореми 1 (пункт 1), тому $\omega(x) = L_n(x)$

і вузлами квадратурної формули є корені цього багаточлена. Вагові множники цієї формули обчислюються за формулою

$$c_k = \int_{-1}^1 \frac{\omega(x)}{(x-x_k)\omega'(x_k)} dx,$$

а залишковий член

$$R_n(f) = \frac{2^{2n+1}}{(2n+1)!(2n)!(2n)!} f^{(2n)}(\xi).$$

Приклад. Побудуємо квадратурну формулу

$$\int_0^1 f(x) dx \approx \sum_{k=1}^n c_k f(x_k).$$

При $n=2$ потрібно знайти c_0, c_1, x_0, x_1 . Заміною $t = 2x - 1$ переведемо $x \in [0,1]$ на проміжок $t \in [-1,1]$. Запишемо $L_2(t) = \frac{3}{2}t^2 - \frac{1}{2}$. Звідси

$$L_2(x) = \frac{3(2x-1)-1}{2} = \frac{12x^2 - 12x + 2}{2} = 6x^2 - 6x + 1 = 0.$$

Звідси $x_1 = \frac{3-\sqrt{3}}{6}, x_2 = \frac{3+\sqrt{3}}{6}$. За формулою (5) знайдемо

$$c_1 = \int_0^1 \frac{x-x_2}{x_1-x_2} dx = \frac{1}{2}, \quad c_2 = \int_0^1 \frac{x-x_1}{x_2-x_1} dx = \frac{1}{2}.$$

Тобто квадратурна формула має вигляд

$$\int_0^1 f(x) dx = \frac{1}{2} \left(f\left(\frac{3-\sqrt{3}}{6}\right) + f\left(\frac{3+\sqrt{3}}{6}\right) \right).$$

2) Розглянемо відрізок $[-1,1]$ і вага $\rho(x) = (1-x^2)^{-\frac{1}{2}}$, тобто виведемо формули Гаусса для обчислення інтегралу

$$I(f) = \int_{-1}^1 \frac{f(x) dx}{\sqrt{1-x^2}}.$$

Багаточлени Чебишова задовольняють умовам теореми 1 (п.1), тому

$$\omega(x) = \bar{T}_n(x) = \frac{1}{2^{n-1}} \cos(n \arccos x).$$

Вузлами квадратурної формули є корені цього багаточлена Чебишова, тобто корені рівняння $\cos(n \arccos x) = 0$. Звідси

$$x_k = \cos \frac{(2k-1)\pi}{2n}, \quad k = \overline{1, n}.$$

Відповідні коефіцієнти обчислюються за формулами

$$c_k = \int_{-1}^1 \frac{T_n(x) dx}{\sqrt{1-x^2} T_n'(x_k)(x-x_k)} = \frac{\pi}{n}, \quad k = \overline{1, n}.$$

Отже, квадратурні формули найвищого степеня точності (ці формули називають формулами Ерміта) мають вигляд

$$I_n(f) = \frac{\pi}{n} \sum_{k=1}^n f(x_k), \quad (1)$$

де x_k – корені багаточлена Чебишова.

Залишковий член має вигляд

$$R_n(f) = \frac{\pi}{2^{2n-1}(2n)!} f^{(2n)}(\xi).$$

3) Розглянемо проміжок $(-\infty, \infty)$ і вагу $\rho(x) = e^{-x^2}$, тобто виведемо формули Гаусса для обчислення інтегралу

$$\int_{-\infty}^{\infty} e^{-x^2} f(x) dx.$$

За теорією

$$\omega(x) = H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^2} e^{-x^2},$$

де $H_n(x)$ багаточлени Ерміта. Багаточлени Ерміта обчислюються також за рекурентними формулами

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x), \quad H_{-1} = 0, \quad H_0 = 1.$$

Коефіцієнти квадратурної формули обчислюються за формулами

$$c_k = \frac{2^{n+1} n! \sqrt{\pi}}{[H_n'(x_k)]^2}.$$

Залишковий член

$$R_n(f) = \frac{n! \sqrt{\pi}}{2^n (2n)!} f^{(2n)}(\xi).$$

Наприклад, при $n = 1$ $H_1(x) = 2x$. Корінь $x_0 = 0$,

$$c_0 = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

Квадратурна формула має вигляд:

$$I_1(x) = \sqrt{\pi} f(0).$$

4) Розглянемо відрізок $[0, \infty]$ і ваговий множник $\rho(x) = x^\alpha e^{-x}$, тобто виведемо формули Гаусса для обчислення інтегралу

$$\int_0^{\infty} x^\alpha e^{-x} f(x) dx.$$

За теорією

$$\omega(x) = L_n^\alpha(x) = (-1)^n x^{-\alpha} e^x \frac{d^n}{dx^n} (x^{\alpha+n} e^{-x}),$$

де $L_n^\alpha(x)$ багаточлени Лагера. Коефіцієнти обчислюються за формулами

$$c_k = \frac{P(n+1)P(n+\alpha+1)}{x_k [L_n^\alpha(x_k)]^2}.$$

Залишковий член при $\alpha = 0$ рівний

$$R_n(f) = \frac{(n!)^2}{(2n)!} f^{(2n)}(\xi).$$

7) Інтегрування швидко осцилюючих функцій.

Розглянемо інтеграл

$$I(f) = \int_a^b f(x) e^{j\omega x} dx, \quad j^2 = -1.$$

При великих ω застосування складених квадратурних формул інтерполяційного типу приводить до великої похибки і при малих кроках h .

Розглянемо $e^{j\omega x}$ як ваговий коефіцієнт, тобто $\rho(x) = e^{j\omega x}$. Замінімо $[a, b]$ на

$[-1, 1]$: $x_i = \frac{a+b}{2} + \frac{b-a}{2} d_i$, $d_i \in [-1, 1]$, $i = \overline{1, n}$ (вузли можуть бути не

рівновіддалені, якщо рівновіддалені, то $d_i = -1 + i \frac{2}{n}$, $i = \overline{1, n}$).

Замінімо $f(x)$ на інтерполяційний багаточлен Лагранжа $L_{n-1}(x)$ з вузлами x_i і отримаємо формулу

$$I_n(f) = \int_a^b L_{n-1}(x) e^{j\omega x} dx, \quad (2)$$

яка буде точною для всіх багаточленів не вище $n-1$ степеня. Тобто, якщо в (2) підставити багаточлен Лагранжа, то можна обчислити інтеграл і отримати квадратурну формулу

$$S_n^\omega(f) = \frac{b-a}{2} \exp\left\{j\omega \frac{a+b}{2}\right\} \sum_{i=1}^n D_i\left(\omega \frac{b-a}{2}\right) f(x_i), \quad D_i(p) = \int_{-1}^1 \left(\prod_{\substack{k=1 \\ k \neq i}}^n \frac{\xi - d_k}{d_i - d_k} \right) \exp(jp\xi) d\xi$$

При $n = 3, d_1 = -1, d_2 = 0, d_3 = 1$ – це формула Філона. Можна брати і більше

точок, наприклад $n = 5, d_1 = -1, d_2 = -\frac{1}{2}, d_3 = 0, d_4 = \frac{1}{2}, d_5 = 1$. Ці формули не

потрібно застосовувати, коли немає швидко осцилюючого множника.

Обчислення невластних інтегралів

Розглянемо обчислення інтегралів з такими особливостями :

а) інтеграли другого роду, тобто

$$I = \int_a^b F(x)dx, \quad F(x) \xrightarrow{x \rightarrow a \vee x \rightarrow b} \infty;$$

б) інтеграли першого роду

$$I = \int_a^{\infty} F(x)dx.$$

Розглянемо спочатку інтеграли другого роду, тобто

$$I = \int_a^b F(x)dx \quad F(x) \xrightarrow{x \rightarrow a \vee x \rightarrow b} \infty.$$

1) *Мультиплікативний спосіб*. Представимо підінтегральну функцію у вигляді $F(x) = \rho(x)f(x)$, причому $\rho(x)$ – особлива, а $f(x)$ – гладка. Далі для обчислення інтегралу

$$I = \tilde{I}(f) = \int_a^b \rho(x)f(x)dx$$

використовуємо відповідні квадратурні формули Гаусса.

Приклад 1 Потрібно обчислити інтеграл

$$I = \int_{-1}^1 \frac{dx}{\sqrt{1-x^4}}.$$

Точки $x = \pm 1$ є особливими.

Представимо підінтегральну функцію у вигляді:

$$F(x) = \underbrace{\frac{1}{\sqrt{1-x^2}}}_{\rho(x)} \underbrace{\frac{1}{\sqrt{1+x^2}}}_{f(x)},$$

отримаємо інтеграл вигляду

$$I = \int_0^1 \frac{1}{\sqrt{1+x^2}} \frac{dx}{\sqrt{1-x^2}},$$

$$\text{де } \rho(x) = \frac{1}{\sqrt{1-x^2}}.$$

Далі використовуємо квадратурну формулу Ерміта (1) з попереднього пункту і обчислюємо інтеграл.

Приклад 2 Обчислити інтеграл

$$I = \int_0^{\pi} \ln(\sin x) dx.$$

Особливі точки $x = 0, x = \pi$.

Зведемо цю особливість до степеневі:

$$\rho(x) = \frac{1}{\sqrt{x(\pi-x)}},$$

тоді

$$f(x) = \sqrt{x(\pi-x)} \ln(\sin x) \xrightarrow{x \rightarrow 0, \pi} 0.$$

Для знаходження інтегралу з таким $\rho(x)$ застосовуємо квадратурні формули Чебишова. Неприємності виникають, оскільки $f'(x) \xrightarrow{x \rightarrow 0, \pi} \infty$ (хоча квадратурні формули даватимуть наближене значення). Тому застосовують другий спосіб розв'язання проблеми:

2) Адитивний спосіб. Представимо підінтегральну функцію у вигляді

$$F(x) = f(x) + \psi(x),$$

причому $\psi(x)$ – особлива, $f(x)$ – гладка. Розбиваємо інтеграл на два:

$$I = I_1 + I_2.$$

$I_1 = \int_a^b f(x) dx$ – обчислюють чисельно (наприклад, формули Сімпсона чи трапецій),

$I_2 = \int_a^b \psi(x) dx$ – пробують обчислити аналітично (можливо апроксимувати функцію $\psi(x)$, наприклад, рядом).

Приклад 3 Обчислити інтеграл з прикладу 2:

$$I = \int_0^{\pi} \ln(\sin x) dx.$$

$$I = 2 \int_0^{\pi/2} \ln(\sin x) dx.$$

Представимо

$$\ln(\sin x) = \ln \frac{\sin x}{x} + \ln x.$$

Отримаємо два інтеграли:

$$I_1 = \int_a^b \ln \frac{\sin x}{x} dx \text{ обчислюємо чисельно,}$$

$$\text{а } I_2 = \int_0^{\pi/2} \ln x dx = (x \ln x - x) \Big|_0^{\pi/2} = \frac{\pi}{2} \left(\ln \frac{\pi}{2} - 1 \right).$$

Розглянемо тепер інтеграли першого роду

$$I = \int_a^{\infty} F(x) dx$$

Нехай $a > 0$. Зробимо заміну $t = \frac{x-a}{x}$, $x = \frac{a}{1-t}$. Тоді

$$I = a \int_0^1 F\left(\frac{a}{1-t}\right) \frac{dt}{(1-t)^2},$$

а це інтеграл другого роду.

Якщо $a=0$, то робимо заміну $t = e^{-x}$, $x = -\ln t$, тоді

$$I = \int_0^1 F(-\ln t) \frac{dt}{t}.$$

Знову отримуємо інтеграл другого роду.

Якщо $a < 0$ (не можна зробити заміну $t = \frac{x-a}{x}$, тому що виникає особливість в точці $x=0$), розбиваємо інтеграл на два:

$$I = \int_a^0 F(x) dx + \int_0^{\infty} F(x) dx$$

і обчислюємо за допомогою попередніх пунктів.

Мультиплікативний спосіб обчислення інтегралів першого роду ґрунтується на представленні підінтегральної функції у вигляді

$$F(x) = \rho(x)f(x),$$

де, наприклад,

$$\rho(x) = x^\alpha e^{-x}, \quad x \in [0, \infty).$$

Такий ваговий коефіцієнт відповідає багаточленам Лагера. При $x \in (-\infty, \infty)$, $\rho(x) = e^{-x^2}$ приходимо до багаточленів Ерміта.

Ще один спосіб ґрунтується на обрізанні верхньої границі. Для цього інтеграл запишемо у вигляді

$$I = \int_a^{\infty} F(x)dx = \int_a^b F(x)dx + \int_b^{\infty} F(x)dx.$$

Верхня границя b обчислюють з умови

$$\left| \int_b^{\infty} F(x)dx \right| < \frac{\varepsilon}{2},$$

де ε – задана точність. Для обчислення $\int_a^b F(x)dx$ використовують квадратурні формули складеного типу.

Обчислення кратних інтегралів

Розглянемо інтеграл

$$I = \int_a^b \int_c^d f(x, y) dx dy \quad (1)$$

Цей інтеграл зводиться до повторного, якщо ввести

$$F(x) = \int_c^d f(x, y) dy \quad (2)$$

Тоді

$$I = \int_a^b F(x) dx. \quad (3)$$

До однократних інтегралів можна застосувати квадратурну формулу середніх прямокутників. Тоді

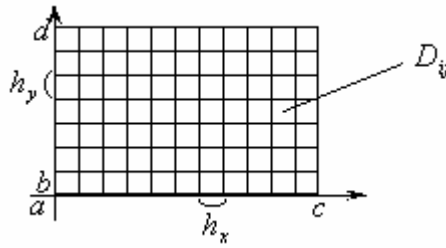
$$I \approx I_0 = (b-a)F(\bar{x}) = (b-a) \int_c^d f(\bar{x}, y) dy \approx (b-a)(d-c)f(\bar{x}, \bar{y}), \quad (4)$$

$$\bar{x} = \frac{a+b}{2}, \quad \bar{y} = \frac{c+d}{2}$$

Кубатурна формула (це формула наближеного обчислення інтегралу (1)), якщо використовується формула трапеції має вигляд:

$$I \approx I_1 = \frac{(b-a)(d-c)}{4} [f(a, c) + f(b, c) + f(a, d) + f(b, d)].$$

Точність залежить від поведінки функції та від довжини інтервалів $[a, b]$, $[c, d]$. Аналог формул складеного типу для (1) будується таким чином: розбиваємо D на комірки (див. малюнок):



$$D = \bigcup_{i,j} D_{ij} \quad D_{ij} = \{x_{i-1} \leq x \leq x_i, y_{j-1} \leq y \leq y_j\}, x_i = a + ih_x, i = \overline{0, N_x},$$

$$y_j = b + jh_y, j = \overline{0, N_y}, h_x = \frac{b-a}{N_x}, h_y = \frac{d-c}{N_y}$$

Тоді для кожного інтегралу по комірці застосовуємо кубатурну формулу прямокутників (4):

$$I = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \iint_{D_{ij}} f(x, y) dx dy \approx I_{0,h} = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} h_x h_y f(\bar{x}_i, \bar{y}_j)$$

$$\bar{x}_i = x_i - \frac{h_x}{2} \quad \bar{y}_j = y_j - \frac{h_y}{2}.$$

Якщо $f(x, y) \in C^2(\bar{D})$, то $I - I_{0,h} = O(h_x^2 + h_y^2)$. Степінь точності по крокам сітки – 2. Складність методу пропорційна кількості комірок

$$Q = O(N_x, N_y) = O(N^2),$$

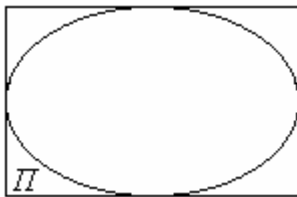
$N \approx N_x \approx N_y$. В 3-вимірному просторі $f = f(x, y, z)$ складність - $Q = O(N^3)$.

Якщо D не прямокутник, то замість f введемо

$$\bar{f}(x, y) = \begin{cases} f(x, y), & x \in D, \\ 0, & x \in \Pi \setminus D. \end{cases}$$

де Π – найменший охоплюючий D прямокутник $D \subset \Pi$

Тоді $I = \iint_{\Pi} \bar{f}(x, y) dx dy$, що розглядався вище.



Недолік такого підходу: $\bar{f}(x, y)$ може бути розривною функцією і через це низька точність обчислення інтегралу.

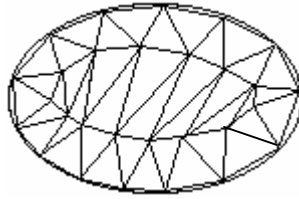
Наступний підхід базується на відповідній заміні змінних

$$x = x(\xi, \eta) \quad y = y(\xi, \eta) \quad D \rightarrow \Pi,$$

$$I = \iint_{\Pi} f(x(\xi, \eta), y(\xi, \eta)) J(\xi, \eta) d\xi d\eta,$$

де Π - прямокутник в площині (ξ, η) ; $J(\xi, \eta)$ – Якобіан переходу. Якщо границя області D гладка, то якобіан буде мати особливості, що також знижує швидкість збіжності.

Ще один підхід в обчисленні інтегралу по довільній області D базується на триангулюванні області. Якщо область довільного вигляду, то її можна розбити на трикутники таким чином $D = \bigcup_{k=1}^N D_k$:



Тоді $I = \sum_{k=1}^N I_k$, $I_k = \iint_{D_k} f(x, y) dx dy$. Застосуємо кубатурні формули до

кожного $I_k \approx I_k^h$. Для цього замінимо функцію поліномом першого степеня

$$f(x, y) \cong L_1(x, y) = A + Bx + Cy.$$

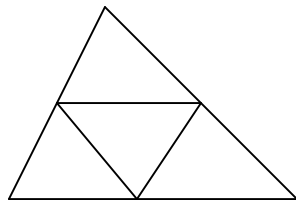
Задача. Побудувати явний вигляд кубатурної формули, яка дозволяє наближено обчислити I_k по трикутнику D_k , якщо замінити $f(x, y) \approx L_1(x, y)$ на інтерполяційний багаточлен 1-го степеня.

Точність такого підходу

$$I - I^h = I - \sum_{k=1}^N I_k^h = O(h^2),$$

де $h = \max_k \text{diam} D_k$.

Складність обчислення інтеграла: $Q = O(h^{-2})$ для $D \subset R_2$; $Q = O(h^{-3})$ для



$D \subset R_3$.

Можна згустити сітку, поділивши один трикутник D_k на чотири (див. малюнок):

І, нарешті, розглянемо ідею методу статистичних випробувань (метод Монте-Карло) для обчислення інтегралу $I = \iint_D f(x, y) dx dy$. Замінимо

наближено

$$I = \iint_{\Pi} \bar{f}(x, y) dx dy \approx \frac{1}{N} \sum_{i=1}^N \bar{f}(\xi_i, \eta_i) \cdot \text{mes} \Pi,$$

де Π – найменший охоплюючий D прямокутник $D \subset \Pi$; $\bar{f}(x, y)$ – продовження функції (4); ξ_i, η_i – незалежні реалізації рівномірно розподілених на $[a, b]$ та $[c, d]$ випадкових величин ξ та η . Складність цього методу

$Q = O(N)$. Оцінка точності $I - I_N = O(N^{-\frac{1}{2}})$ носить ймовірністний характер. Позитивна сторона методу – незалежна від розмірності складність; негативна – низька точність.

Контрольні запитання:

1. Квадратурна формула Сімпсона.
2. Принцип Рунге оцінки похибки.
3. Екстраполяційна формула Річардсона.
4. Поняття адаптивних квадратурних формул.
5. Квадратурні формули найвищого алгебраїчного степеня точності.
6. Методи обчислення невластних інтегралів.
7. Методи обчислення кратних інтегралів.

Література [3] - [6], [10], [11].

Лекція 13. Чисельне інтегрування звичайних диференціальних рівнянь

Мета лекції: засвоїти теоретичні основи методу Ейлера, сформулювати уявлення про алгоритми чисельного інтегрування звичайних диференціальних рівнянь та їх реалізацій на персональних комп'ютерах.

План лекції:

1. Чисельні методи розв'язування задачі Коші для звичайних диференціальних рівнянь першого порядку.
2. Метод Ейлера.
3. Оцінка локальної та глобальної похибки методу Ейлера.
4. Блок-схема методу Ейлера.
5. Вибір кроку сітки.

Постановка задачі: нехай потрібно знайти розв'язок диференціального рівняння з початковими умовами

$$\frac{du}{dt} = f(t, u), \quad t > t_0, \quad u(t_0) = u_0, \quad (1)$$

$$u = (u_1, \dots, u_m)^T, \quad f = (f_1, \dots, f_m)^T, \quad u_k = u_k(t), \quad f_k = f_k(t, u_1, \dots, u_m).$$

Якщо довільна функція f_k неперервна по кожній своїй змінній та по u_j вони Ліпшиць-неперервні, тобто

$$|f_k(t, \dots, u_j, \dots) - f_k(t, \dots, v_j, \dots)| \leq L_j |u_j - v_j|, \quad \forall j, k,$$

то розв'язок задачі (1) існує і єдиний.

Нехай задано рівняння m -ого порядку та початкові умови:

$$\begin{cases} v^{(m)}(t) = F(t, v, v', \dots, v^{(m-1)}), \\ t = t_0 : v = v_1, v' = v_2, \dots, v^{(m-1)} = v_m. \end{cases} \quad (2)$$

Введемо компоненти вектора \vec{u} : $u_k(t) = v^{(k-1)}(t)$. Тоді задача (2) записується у вигляді системи (1):

$$\begin{cases} \frac{du_1}{dt} = u_2, & u_1(t_0) = v_1, \\ \dots \\ \frac{du_{m-1}}{dt} = u_m, & u_{m-1}(t_0) = v_{m-1}, \\ \frac{du_m}{dt} = F(t, u_1, \dots, u_m), & u_m(t_0) = v_m. \end{cases}$$

Тому далі, в основному, розглядаються методи розв'язання задачі (1).

Наближені аналітичні методи

1) Метод послідовних наближень (метод Пікара)

Потрібно розв'язати диференціальне рівняння з відповідними початковими умовами:

$$\frac{du}{dt} = f(t, u), \quad u(t_0) = u_0. \quad (1)$$

Проінтегруємо (1)

$$u(t) = u(t_0) + \int_{t_0}^t f(\xi, u(\xi)) d\xi \quad (2)$$

Задаємо $u^{(0)}(t)$ і запишемо ітераційний процес:

$$u^{(k+1)}(t) = u_0 + \int_{t_0}^t f(\xi, u^{(k)}(\xi)) d\xi, \quad k = 0, 1, \dots \quad (3)$$

Існує $T = T(u_0, L)$ ($L = \max_j L_j$ - стала Лібшиця) така, що

$$u^{(n)}(t) \xrightarrow{n \rightarrow \infty} u(t), \quad t \in [t_0, T].$$

Тому $u(t) \approx u^{(n)}(t)$.

Недолік методу: необхідно проведення аналітичного інтегрування

2) Метод рядів Тейлора

Нехай розв'язок задачі (1) можна представити у вигляді ряду

$$u(t) = \sum_{k=0}^{\infty} \frac{u^{(k)}(t_0)}{k!} (t - t_0)^k.$$

Будемо шукати наближення у вигляді скінченної суми

$$u(t) \approx u^N(t) = \sum_{k=0}^N \frac{u^{(k)}(t_0)}{k!} (t - t_0)^k, \quad t \in [t_0, t_1]. \quad (4)$$

Для визначення $u^{(k)}(t_0)$ диференціюємо рівняння (1) по t :

$$\begin{aligned} u^{(0)}(t_0) &= u_0, \quad u^{(1)}(t_0) = f(t_0, u_0), \\ u^{(2)}(t_0) &= f_{t,0} + f_{u,0} f_0, \quad u^{(3)}(t_0) = f_{tt,0} + 2f_{tu,0} f_0 + f_{uu,0} f_0^2, \dots \end{aligned}$$

Якщо $\tau = t_1 - t_0$ малий параметр, то

$$|u(t) - u^N(t)| = O(\tau^{N+1}).$$

Недоліки методу:

- зростання кількості доданків при обчисленні $u^{(n)}(t_0)$;
- необхідно аналітичного диференціювання.

Методи типу Ейлера

Розглянемо задачу Коші:

$$\frac{du}{dt} = f(t, u), u(t_0) = u_0. \quad (1)$$

Використаємо перше наближення за допомогою рядів Тейлора на проміжку $t_0 \leq t \leq t_1$:

$$u(t) \approx u^1(t) = u(t_0) + (t - t_0)f(t_0, u_0).$$

Обчислимо наближене значення в точці t_1 :

$$u(t_1) = u_1 \approx u^1(t_1) = u_0 + \tau f(t_0, u_0),$$

де $\tau = t_1 - t_0$ - деякий крок. Якщо позначити $y_1 = u(t_1)$, $y_0 = u(t_0)$, то маємо формулу

$$y_1 = y_0 + \tau f(t_0, y_0).$$

Застосовуючи такий підхід для $t_n \leq t \leq t_{n+1}$, отримаємо рекурентну формулу

$$y_{n+1} = y_n + \tau f(t_n, y_n), n = 0, 1, \dots, y_0 = u_0, \quad (2)$$

Це формула методу Ейлера. Крок інтегрування може змінюватися:

$$\tau = \tau_n = t_{n+1} - t_n.$$

Геометрична інтерпретація методу Ейлера представлена на рис.1. Його друга назва – метод ламаних.

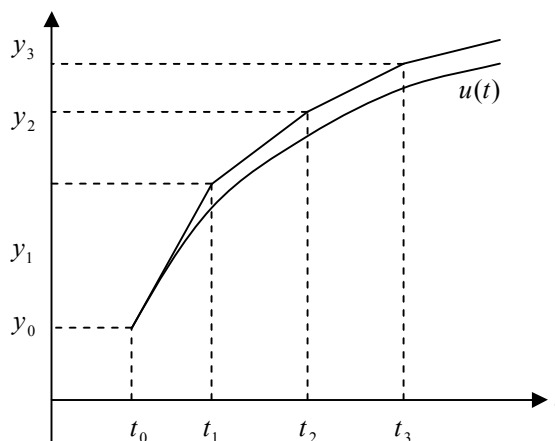


Рис. 1

Цей метод відноситься до однокрокових, тобто розв'язок на наступному кроці обчислюється тільки по одному значенню на попередньому кроці. Загальна формула однокрокових методів

$$y_{n+1} = y_n + \tilde{\Phi}(t_n, y_n, f) \quad (3)$$

Для методу Ейлера $\tilde{\Phi} = \tau f(t_n, y_n)$.

Величина $R(\tau) = y(t_{n+1}) - u(t_{n+1})$, де y_{n+1} обчислюється за формулою (3), причому $y_n \equiv u(t_n)$ називається *похибка методу (3) на одному кроці*.

Загальна похибка на $(n+1)$ -му кроці складається з похибок на всіх попередніх кроках.

Величина $z_{n+1} = y_{n+1} - u(t_{n+1})$, де y_{n+1} і всі попередні y_k , $k = 1, 2, \dots$ також наближені, називається *похибкою методу* (3).

Якщо виразити $y_{n+1} = u(t_{n+1}) + z_{n+1}$ і підставити в (3), то

$$u(t_{n+1}) + z_{n+1} = u(t_n) + z_n + \tau \tilde{\Phi}(t_n, u_n + z_n).$$

Тоді можна записати рівняння для z_n :

$$z_{n+1} = z_n + \tau \left[-\frac{u_{n+1} - u_n}{\tau} + \tilde{\Phi}(t_n, u_n) \right] + \tau [\tilde{\Phi}(t_n, u_n + z_n) - \tilde{\Phi}(t_n, u_n)].$$

Величина $\psi_n = -\frac{u_{n+1} - u_n}{\tau} + \tilde{\Phi}(t_n, u_n)$ називається *похибкою апроксимації методу* (3).

Для методу Ейлера

$$\begin{aligned} \psi_n &= -\frac{u_{n+1} - u_n}{\tau} + f(t_n, u_n) = -\frac{1}{\tau} \left[u_n + \tau \frac{du}{dt}(t_n) + O(\tau^2) - u_n \right] + f(t_n, u_n) = \\ &= -\frac{du}{dt}(t_n) + f(t_n, u_n) + O(\tau) = O(\tau) \xrightarrow{\tau \rightarrow 0} 0. \end{aligned}$$

Похибка апроксимації це нев'язка, коли замість y в різницеве рівняння (3) підставляємо точний розв'язок задачі Коші u .

Метод (3) має *похибку на одному кроці степеня m* , якщо $|R_n(\tau)| = O(\tau^{m+1})$.

Кажуть, що чисельний метод має *похибку апроксимації степеня p* , якщо $|\psi_n| = O(\tau^p)$.

Для методу Ейлера похибка на одному кроці

$$\begin{aligned} R_n(\tau) &= y_{n+1} - u(t_{n+1}) = u(t_n) + \tau f(t_n, u(t_n)) - u(t_n + \tau) = u(t_n) + \tau f(t_n, u(t_n)) - \\ &- u(t_n) - \tau u'(t_n) + O(\tau^2) = O(\tau^2). \end{aligned}$$

Отже $m = 1$. Маємо зв'язок похибок: $\psi(\tau) = \frac{1}{\tau} R(\tau)$ і тому $p = 1$.

Кажуть, що метод (3) має *ступінь точності m* , якщо

$$\forall n \quad z_n = y_n - u(t_n) = O(\tau^m).$$

Теорема Нехай $f(t, u) \in C^1(\bar{D}_T)$, та $|f_u(t, u)| \leq L$, де $\bar{D}_T = \{t_0 < t \leq T, |u(t)| \leq K\}$. Тоді метод Ейлера (3) має ступінь точності $m = 1$ і для нього має місце оцінка

$$|z_n| \leq M \max_j |\psi_j| = O(\tau),$$

де $M = M(T, L)$.

▷ Для z маємо задачу

$$z_{n+1} = z_n + \tau \psi_n + \tau [f(t_n, u_n + z_n) - f(t_n, u_n)].$$

Оцінимо z_{n+1} :

$$|z_{n+1}| \leq |z_n| + \tau |\psi_n| + \tau L |z_n| \leq (1 + \tau L) |z_n| + \tau |\psi_n| \leq$$

$$\begin{aligned} & \dots \leq (1 + \tau L)^n |z_0| + \sum_{j=0}^n \tau (1 + \tau L)^{n-j} |\psi_j| \leq \\ & \leq (1 + \tau L)^n \max_j |\psi_j| \cdot \sum_{j=0}^n \tau \leq Te^{L\tau} \max_j |\psi_j| \leq M \max_j |\psi_j| = O(\tau), \end{aligned}$$

так як $(1 + \tau L)^n = (1 + \tau L)^{\frac{1}{\tau} n\tau} = (1 + \tau L)^{\frac{1}{\tau} t_n} \leq (1 + \tau L)^{\frac{1}{\tau} T} \leq e^{LT}$. Позначимо $Te^{LT} = M$ і отримуємо бажану оцінку. \triangleleft

Таким чином порядок точності методу Ейлера $m = 1$.

Метод Ейлера можна вивести із таких міркувань. Інтегруємо (1) по t :
 $t_n < t < t_{n+1}$

$$u(t_{n+1}) = u(t_n) + \int_{t_n}^{t_{n+1}} f(t, u(t)) dt. \quad (4)$$

Застосуємо формулу лівих прямокутників для інтегралу

$$\int_{t_n}^{t_{n+1}} f(t, u(t)) dt \approx \tau f(t_n, u(t_n))$$

і підставимо в (4). Отримаємо формулу для явного методу Ейлера

$$y_{n+1} = y_n + \tau f(t_n, y_n).$$

Застосуємо формулу правих прямокутників для інтегрування

$$\int_{t_n}^{t_{n+1}} f(t, u(t)) dt \approx \tau f(t_{n+1}, u(t_{n+1}))$$

і отримаємо з (4) формулу для неявного методу Ейлера

$$y_{n+1} = y_n + \tau f(t_{n+1}, y_{n+1}). \quad (5)$$

Ці формули 1-го степеня точності по кроку τ . Заміняючи інтеграл за квадратурною формулою трапеції

$$\int_{t_n}^{t_{n+1}} f(t, u(t)) dt \approx \frac{\tau}{2} [f(t_n, u(t_n)) + f(t_{n+1}, u(t_{n+1}))],$$

ми отримаємо формулу *методу трапеції* інтегрування задачі Коші

$$y_{n+1} = y_n + \frac{\tau}{2} [f(t_n, u(t_n)) + f(t_{n+1}, u(t_{n+1}))]. \quad (6)$$

Це неявний метод.

Обчислимо похибку апроксимації цього методу

$$\begin{aligned} \psi_n &= -\frac{u_{n+1} - u_n}{\tau} + \frac{1}{2} [f(t_n, u_n) + f(t_{n+1}, u_{n+1})] = \\ &= -\frac{1}{\tau} [u_{n+\frac{1}{2}} + \frac{\tau}{2} \dot{u}_{n+\frac{1}{2}} + \frac{1}{2} \left(\frac{\tau}{2}\right)^2 \ddot{u}_{n+\frac{1}{2}} + O(\tau^3)] - \\ &- u_{n+\frac{1}{2}} + \frac{\tau}{2} \dot{u}_{n+\frac{1}{2}} - \frac{1}{2} \left(\frac{\tau}{2}\right)^2 \ddot{u}_{n+\frac{1}{2}} + O(\tau^3) + \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{2} \left[f_{n+\frac{1}{2}} + \frac{\tau}{2} \dot{f}_{n+\frac{1}{2}} + O(\tau^2) + f_{n+\frac{1}{2}} - \frac{\tau}{2} \dot{f}_{n+\frac{1}{2}} + O(\tau^2) \right] = \\
& = -\dot{u}_{n+\frac{1}{2}} + O(\tau^2) + f_{n+\frac{1}{2}} = O(\tau^2).
\end{aligned}$$

Таким чином метод трапецій має другий порядок апроксимації.

Задача. Показати, що похибка на одному кроці методу трапецій є величина порядку $O(\tau^3)$.

Таким чином ми отримали більш точний метод. Недолік методу трапецій – неявність (треба розв'язувати нелінійне рівняння). Для розв'язання проблеми неявності застосуємо метод предиктор-корректор (предиктор - попередник, коректор - уточнювач).

Обчислимо попередньо за явним методом Ейлера

$$\bar{y}_{n+1} = y_n + \tau f(t_n, y_n). \quad (7)$$

Уточнимо за методом трапецій

$$y_{n+1} = y_n + \frac{\tau}{2} [f(t_n, y_n) + f(t_{n+1}, \bar{y}_{n+1})]. \quad (8)$$

Формули (7) та (8) утворюють *метод Ейлера-Коші*.

Оцінимо похибку на одному кроці:

$$R(\tau) = y_1 - u(t_0 + \tau) = u_0 + \frac{\tau}{2} [f(t_0, y_0) + f(t_0 + \tau, u_0 + \tau f(t_0, y_0))] - u(t_0 + \tau),$$

Звідси $R(0) = 0$. Далі

$$\begin{aligned}
R'(\tau) &= \frac{1}{2} [f(t_0, y_0) + f(t_0 + \tau, u_0 + \tau f_0)] - \frac{du}{dt}(t_0 + \tau) + \frac{\tau}{2} f_t(t_0 + \tau, u_0 + \tau f(t_0, u_0)) + \\
&+ \frac{\tau}{2} f_u(t_0 + \tau, u_0 + \tau f(t_0, u_0)) \cdot f(t_0, u_0), \quad R'(0) = \frac{1}{2} (f_0 + f_0) - \frac{du}{dt}(t_0) = 0.
\end{aligned}$$

Наступна похідна

$$R''(\tau) = f_t(t_0 + \tau, u_0 + \tau f_0) + f_u(t_0 + \tau, u_0 + \tau f_0) \cdot f_0 + \tau(\dots) - u''(t_0),$$

$$R''(0) = f_{t,0} + f_{u,0} \cdot f_0 - \frac{d^2u}{dt^2}(t_0) = 0,$$

так як $\frac{d^2u}{dt^2} = \frac{d}{dt} f(t, u) = f_t(\dots) + f_u(\dots) f_0$.

І, нарешті, $R'''(0) = \frac{3}{2} (f_{u,0} + 2f_{uu,0} + f_{uu} \cdot f_0^2) - u'''(t_0) \neq 0$. Таким чином похибка на одному кроці має порядок $p = 2$.

Ще один метод цього типу:

$$y_{n+\frac{1}{2}} = y_n + \frac{\tau}{2} f(t_n, y_n), \quad (9)$$

$$y_{n+1} = y_n + \tau f\left(t_{n+\frac{1}{2}}, y_{n+\frac{1}{2}}\right). \quad (10)$$

Це модифікований метод Ейлера. На етапі коректор (11) використовуємо формулу середніх прямокутників, а предиктор – це метод Ейлера на інтервалі (t_n, t_{n+1}) .

Задача. Показати, що для модифікованого методу Ейлера $R(\tau) = O(\tau^3)$, тобто $m = 2$.

Контрольні запитання:

1. Постановка задачі чисельного інтегрування звичайних диференціальних рівнянь.
2. Метод послідовних наближень Пікара.
3. Метод рядів Тейлора.
4. Методи типу Ейлера.
5. Похибка методу Ейлера.
6. Формула методу трапеції інтегрування задачі Коші.
7. Метод Ейлера-Коші.
8. Модифікований метод Ейлера.

Література [3], [4], [6], [9], [11].

Лекція 14. Чисельне інтегрування звичайних диференціальних рівнянь

Мета лекції: засвоїти теоретичні основи методу Рунге-Кутта, одержати уявлення про можливості його практичного застосування.

План лекції:

1. Метод Рунге-Кутта.
2. Оцінка локальної та глобальної похибки методу Рунге-Кутта.
3. Блок-схема методу Рунге-Кутта.
4. Вибір кроку сітки.

Методи типу Рунге-Кутта

Методи типу Ейлера мають низьку точність ($m = 1,2$). Рунге запропонував метод третього, а Кутта розвинув його ідею та отримав методи четвертого порядку точності.

Розглянемо задачу Коші

$$\frac{du}{dt} = f(t, u), u(t_0) = u_0. \quad (1)$$

Явний m - етапний (стадійний) метод Рунге-Кутта полягає в наступному. Нехай розв'язок $y_n = y(t_n)$ вже відомий. Задаються числові коефіцієнти α_i , β_{ij} , $i = 2, m$, $j = 1, i - 1$ та p_i , $i = 1, m$ і послідовно обчислюються за формулами:

$$\begin{aligned} k_1 &= \tau f(t_n, y_n), \\ k_2 &= \tau f(t_n + \alpha_2 \tau, y_n + \beta_{21} \tau k_1), \\ k_3 &= \tau f(t_n + \alpha_3 \tau, y_n + \beta_{31} \tau k_1 + \beta_{32} \tau k_2), \dots, \\ k_m &= \tau f(t_n + \alpha_m \tau, y_n + \beta_{m1} \tau k_1 + \beta_{m2} \tau k_2 + \dots + \beta_{m,m-1} \tau k_{m-1}). \end{aligned}$$

Потім з формули

$$y_{n+1} = y_n + \sum_{i=1}^m p_i k_i \quad (2)$$

знаходимо нове значення $y_{n+1} = y(t_{n+1}) \approx u(t_{n+1})$.

Коефіцієнти α_i , β_{ij} , p_i вибираємо з міркувань точності. Наприклад, для того, щоб рівняння (2) апроксимувало рівняння (1), необхідно, щоб $\sum_{i=1}^m p_i = 1$.

Інформація про m -стадійний метод записується в таблиці Батчера:

$$\begin{array}{c|ccc} \alpha_1 & & & \\ \alpha_2 & \beta_{21} & & \\ \vdots & \vdots & & \\ \alpha_m & \beta_{m1} & \cdots & \beta_{mm-1} \\ \hline & p_1 & \cdots & p_{m-1} \quad p_m \end{array}$$

Похибка на одному кроці

$$\begin{aligned} R(\tau) &= y_1 - u(t_0 + \tau) = u_0 + \sum_{i=1}^m p_i k_i(\tau) - u(t_0 + \tau) = \\ &= R(0) + R'(0)\tau + \dots + \frac{\tau^p}{p!} R^{(p)}(0) + O(\tau^{p+1}) \end{aligned}$$

Для того, щоб $R(\tau) = O(\tau^{p+1})$, тобто метод мав p -й степінь точності необхідно і достатньо, щоб $R(0) = R'(0) = \dots = R^{(p)}(0) = 0$. Загального розв'язку цієї нелінійної системи немає, тому розглянемо частинні випадки.

1. Методи першого порядку $m = 1$. Невідоме p_1 .

$$\begin{aligned} k_1(\tau) &= \tau f(t_0, y_0), \quad R(\tau) = u_0 + p_1 k_1(\tau) - u(t_0 + \tau), \quad R(0) = u_0 - u(t_0) = 0, \\ R'(\tau) &= p_1 k'_1(\tau) - \dot{u}(t_0 + \tau) = p_1 f(t_0, y_0) - \dot{u}(t_0 + \tau), \end{aligned}$$

$$R'(0) = (p_1 - 1)f_0 = 0 \Rightarrow p_1 = 1.$$

Ясно, що $R''(0) \neq 0$. Тому $R(\tau) = O(\tau^2)$. Отримуємо явний метод Ейлера

$$y_{n+1} = y_n + \tau f(t_n, y_n).$$

2. Методи другого порядку $m = 2$. Тут отримуємо сімейство методів. Невідомі $p_1, p_2, \alpha_2, \beta_{21}$. Вони вибираються з умови: $R(\tau) = O(\tau^3)$, $p = 2$. Маємо

$$p_1 + p_2 = 1, \quad 2\alpha_2 p_2 = 1, \quad 2\beta_{21} p_2 = 1$$

і один параметр залишається вільним. Далі один параметр p_1 фіксуємо і отримуємо конкретний метод.

а) $p_1 = 0, p_2 = 1, \alpha_2 = \beta_{21} = \frac{1}{2}$ – модифікований метод Ейлера

$$y_{n+1} = y_n + \tau f\left(t_{n+\frac{1}{2}}, y_n + \frac{\tau}{2} f_n\right).$$

б) $p_1 = \frac{1}{2}, p_2 = \frac{1}{2}, \alpha_2 = \beta_{21} = 1$ – метод Ейлера-Кوشі

$$y_{n+1} = y_n + \frac{\tau}{2} [f(t_n, y_n) + f(t_{n+1}, y_n + \tau f_n)].$$

в) $p_1 = \frac{1}{3}, p_2 = \frac{2}{3}, \alpha_2 = \beta_{21} = \frac{3}{4}$ – ще один метод другого порядку точності:

$$y_{n+1} = y_n + \frac{\tau}{3} [f(t_n, y_n) + 2f(t_{n+\frac{3}{4}}, y_n + \frac{3}{4} f_n)].$$

3. Методи третього порядку $m = 3$. $R(\tau) = O(\tau^4)$. Тому $p = 3$.

Запишемо результат вибору параметрів у вигляді таблиці Батчера (один з частинних випадків):

$$\begin{array}{c|cc} 0 & & \\ \hline \frac{1}{3} & \frac{1}{3} & \\ \frac{2}{3} & 0 & \frac{2}{3} \\ \hline & \frac{1}{4} & 0 & \frac{3}{4} \end{array}$$

Задача. Довести, що метод, якому відповідає ця таблиця Батчера, має третій степінь точності.

4. Методи четвертого порядку $m = 4$, $R(\tau) = O(\tau^5)$, $p = 4$. Найбільш поширені методи:

$$\begin{array}{c|ccc} 0 & & & \\ \hline \frac{1}{2} & \frac{1}{2} & & \\ \frac{1}{2} & 0 & \frac{1}{2} & \\ 1 & \frac{1}{2} & 0 & 1 \\ \hline & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{array} \quad \begin{array}{c|ccc} 0 & & & \\ \hline \frac{1}{4} & \frac{1}{4} & & \\ \frac{1}{2} & 0 & \frac{1}{2} & \\ 1 & 1 & -2 & 2 \\ \hline & \frac{1}{6} & 0 & \frac{2}{3} & \frac{1}{6} \end{array}$$

Розглянемо як пов'язаний порядок методу p і кількість стадій m :

кількість стадій m	1	2	3	4	5	6
точність методу p	1	2	3	4	4	5

Ясно чому методи Рунге-Кутта при $m > 5$ рідко застосовуються: збільшення кількості стадій (а значить збільшення об'єму обчислювальної роботи) не дає виграшу в точності.

Теорема Нехай m - стадійний метод Рунге-Кутта має p -й степінь точності на кроці, а $f(t, u)$ задовольняє умову Ліпшиця. Тоді метод має p -й степінь точності і для похибки методів Рунге-Кутта має місце така оцінка:

$$|z_n| = |y_n - u(t_n)| \leq T \cdot e^{\alpha_0 T} \cdot \max_j |\psi_j| = O(\tau^p),$$

де $\psi_j = \frac{1}{\tau} R_j(\tau)$ – похибка апроксимації метода, $\alpha_0 = pLm(1 + BL\tau_0)^{m-1}$, L – стала Ліпшиця, $B = \max_{i,j} |f_{ij}|$, $T = \max_n |t_n|$.

Методи з контролем точності на кроці

Часто в ході розрахунків необхідно змінювати крок інтегрування, контролюючи величину похибки методу на кроці. При практичній оцінці цієї величини можна, наприклад, поступати так.

Перший підхід використовує принцип Рунге. Головний член похибки на одному кроці (t_n, t_{n+1}) інтегрування має вигляд

$$\frac{\varphi^{(p+1)}(0)\tau^{p+1}}{(p+1)!}.$$

В результаті двох кроків інтегрування однокроковим методом, наприклад, методом Рунге – Кутта p -го степеня точності, буде отримано наближення $y^\tau(t_n + 2\tau)$ до значення $u(t_n + 2\tau)$ таке, що

$$y^\tau(t_n + 2\tau) - u(t_n + 2\tau) \approx 2 \frac{\varphi^{(p+1)}(0)\tau^{p+1}}{(p+1)!} = 2 \overset{0}{R}(\tau).$$

Якщо тепер застосувати метод Рунге – Кутта p -го степеня з одним кроком довжини 2τ на інтервалі (t_n, t_{n+2}) , то отримаємо наближене значення $y^{2\tau}(t_n + 2\tau)$, для якого

$$y^{2\tau}(t_n + 2\tau) - u(t_n + 2\tau) \approx \frac{\varphi^{(p+1)}(0)(2\tau)^{p+1}}{(p+1)!} = \overset{0}{R}(2\tau) = 2^{p+1} \overset{0}{R}(\tau).$$

З цих співвідношень випливає представлення головного члена похибки на кроці

$$y^\tau(t_n + 2\tau) - u(t_n + 2\tau) \approx 2 \overset{0}{R}(\tau) = \frac{y_{n+1}^{2\tau} - y_{n+1}^\tau}{2^p - 1}.$$

При необхідності можна уточнити отримане наближене значення, додавши

до нього величину головного члена похибки, тобто покласти

$$u(t_n + 2\tau) \approx y^\tau + \frac{y^\tau - y^{2\tau}}{2^p - 1}.$$

Позначимо $g(\tau) = \frac{y_{n+1}^{2\tau} - y_{n+1}^\tau}{2^p - 1}$. Якщо $|g(\tau)| \leq \varepsilon$, де ε – деяка задана мала величина (похибка на одному кроці), то τ – успішний крок і $y_{n+1}^\tau \approx u(t_n + 2\tau)$.

Якщо $|g(\tau)| > \varepsilon$, то зменшуємо крок $\tau := \frac{\tau}{2}$. Треба ще передбачити умову збільшення кроку. Задаємо деяке $\delta \ll \varepsilon$ і якщо $|g(\tau)| \leq \delta$, то $y_{n+1}^\tau \approx u(t_n + 2\tau)$ і далі беремо $\tau := 2\tau$. Параметр δ вибирають, наприклад, так: $\delta = \alpha \cdot \varepsilon \cdot 2^{-p}$, $0 < \alpha < 1$, де p – порядок точності методу.

Інший підхід вибору кроку інтегрування заключається в використанні методів різного степеня точності. Отже, якщо в нас є два методи степеня точності на кроці p та $p+1$:

$$y_{n+1}^{(1)} - u(t_n) = O(\tau^{p+1}),$$

$$y_{n+1}^{(2)} - u(t_n) = O(\tau^{p+2}),$$

то головний член похибки першого методу

$$g(\tau) = y_{n+1}^{(1)} - y_{n+1}^{(2)} = O(\tau^{p+1}).$$

Далі з головним членом похибки $g(\tau)$ методу степені p оперуємо так як і в першому підході.

Багато мати можливість здійснювати крок інтегрування і оцінювати похибку при меншій кількості обчислення значень правих частин. Виграш досягається, якщо використовують методи, які називаються *вкладеними*. Таблиця Батчера для них має вигляд:

α_2	β_{21}		
\vdots	\vdots		
α_m	β_{m1}	\dots	β_{mm-1}
	p_1	p_{m-1}	p_m
	\bar{p}_1	\bar{p}_{m-1}	\bar{p}_m

Метод з параметрами p_1, \dots, p_{m-1}, p_m має порядок точності p , а з параметрами $\bar{p}_1, \dots, \bar{p}_{m-1}, \bar{p}_m$ - $p+1$. Коефіцієнти α_i, β_{ij} у обох методів однакові.

Найпростіший приклад вкладених методів для $m = 2$ має таблицю Батчера:

0	
1	1
	1 0
	0 1

Перший метод, якому відповідають коефіцієнти $p_1 = 1, p_2 = 0$, це метод Ейлера, $p = 1$. Другий - $p_1 = 0, p_2 = 1$ - метод Ейлера - Коші, $p = 2$.

Іншим прикладом може служити сукупність формул шестістадійного $m = 6$ методу Рунге - Кутта - Фельберга:

$$k_1 = \tau f(t_n, y_n), \quad k_2 = \tau f\left(t_n + \frac{\tau}{2}, y_n + \frac{k_1}{2}\right), \quad k_3 = \tau f\left(t_n + \frac{\tau}{2}, y_n + \frac{1}{4}(k_1 + k_2)\right),$$

$$k_4 = \tau f\left(t_n + \tau, y_n - k_2 + 2k_3\right), \quad k_5 = \tau f\left(t_n + \frac{2\tau}{3}, y_n + \frac{1}{27}(7k_1 + 10k_2 + k_4)\right),$$

$$k_6 = \tau f\left(t_n + \frac{\tau}{5}, y_n + \frac{1}{625}(28k_1 - 125k_2 + 546k_3 + 54k_4 - 378k_5)\right),$$

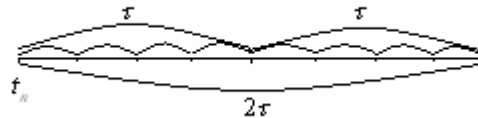
$$y_{n+1} = y_n + \frac{1}{6}(k_1 + 4k_3 + k_4),$$

з головним членом похибки

$$y_n(t_n + \tau) - u(t_n + \tau) = g(\tau) + O(\tau^5),$$

$$g(\tau) = -\frac{1}{366}(42k_1 + 224k_3 + 21k_4 - 162k_5 - 125k_6).$$

Методи Рунге – Кутта – Фельберга мають четвертий та п'ятий степінь точності. Порівняємо кількість обчислень правих частин в методах Рунге-Кутта та Рунге – Кутта – Фельберга ($p = 4$ для обох). Згідно схеми кроків по змінній t та по стадіях $i = \overline{1,4}$ для методу Рунге – Кутта необхідно обчислити для оцінки похибки 11 значень функції, а для методу Рунге – Кутта – Фельберга 6 значень функції.



Основний недолік методів Рунге-Кутта: щоб отримати досить високий степінь точності потрібно багато раз обчислювати значення функції.

Достоїнстю методів Рунге – Кутта є можливість зміни кроку інтегрування $\tau = \tau_n$ і за рахунок цього задовольняти умову точності на кроці.

Контрольні запитання:

1. Поняття методів типу Рунге-Кутта.
2. Методи інтегрування з контролем точності на кроці.
3. Методи Рунге – Кутта – Фельберга.
4. Явний m - етапний метод Рунге-Кутта.
5. Застосування принципу Рунге до методів типу Рунге-Кутта.

Література [3], [4], [6], [11].

Лекція 15. Методи розв'язання крайових задач для звичайних диференціальних рівнянь

Мета лекції: засвоїти теоретичні основи методів розв'язання крайових задач для звичайних диференціальних рівнянь (методу стрільби, пристрілки, лінеаризації, продовження за параметром), одержати уявлення про побудову їх алгоритмів.

План лекції:

1. Постановки крайових задач.
2. Метод стрільби.
3. Метод пристрілки.
4. Метод лінеаризації.
5. Метод продовження за параметром.

Почнемо з постановки крайових задач.

1) Нелінійна двоточкова крайова задача

$$\frac{d\vec{U}}{dx} = \vec{F}(x, \vec{U}) \quad a < x < b \quad (1)$$

$$\vec{\varphi}(\vec{U}(a), \vec{U}(b)) = \vec{d} \quad (2)$$

Тут

$$\vec{U} = (u_1, \dots, u_m)^T, \quad u_k = u_k(x), \quad \vec{F} = (f_1, \dots, f_m)^T, \quad f_k = f_k(x, \vec{U}), \quad \vec{\varphi} = (\varphi_1, \dots, \varphi_m)^T, \\ \varphi_k = \varphi_k(\vec{U}(a), \vec{U}(b)), \quad \vec{d} = (d_1, \dots, d_m), \quad d_k - \text{числа.}$$

2) Лінійна двоточкова крайова задача

$$\frac{d\vec{U}}{dx} = A(x)\vec{U}(x) + \vec{F}(x) \quad (3)$$

$$B_1\vec{U}(a) + B_2\vec{U}(b) = \vec{d} \quad (4)$$

$A(x) = (a_{ij}(x))_{i,j=1}^m$, $\vec{F} = (f_1, \dots, f_m)$, $f_k = f_k(x)$, B_1, B_2 - числові матриці $m \times m$; \vec{d} - вектор.

Крайові умови (2) і (4) називаються *нероздільними*. Часто зустрічаються *розділені крайові умови* (наприклад, для лінійної задачі):

$$C_1\vec{U}(a) = \vec{d}_1, \quad C_2\vec{U}(b) = \vec{d}_2, \quad (4')$$

де $C_1 - (m-k) \times m$ - матриця, $C_2 - k \times m$ - матриця; $\text{rang} C_1 = m-k$, $\text{rang} C_2 = k$; $\vec{d}_1 - (m-k)$ - вектор, $\vec{d}_2 - k$ - вектор.

До (3),(4') зводиться крайова задача для рівнянь вищих порядків. Нехай задана крайова задача:

$$\begin{cases} u^{(m)}(x) = p_1(x)u^{(m-1)}(x) + \dots + p_m(x)u(x) + f(x) \\ \alpha_{i1}u^{(m-1)}(a) + \dots + \alpha_{im}u(a) = \mu_i, \quad i = \overline{1, m-k} \\ \beta_{i1}u^{(m-1)}(b) + \dots + \beta_{ik}u(b) = \nu_i, \quad i = \overline{1, k} \end{cases}$$

Вона зводиться до задачі (3), (4) з

$$A(x) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ p_m & p_{m-1} & p_{m-2} & \dots & p_1 \end{pmatrix},$$

$$C_1 = (\alpha_{ij})_{i=1, m-k}^{j=\overline{i, m}}, \quad C_2 = (\beta_{ij})_{i=1, k}^{j=\overline{1, m}},$$

$$\vec{d}_1 = (\mu_1, \dots, \mu_{m-k})^T, \quad \vec{d}_2 = (\nu_1, \dots, \nu_k)^T.$$

Вважаємо, що всі задачі мають єдині розв'язки. Розглянемо методи розв'язання цих задач.

Метод стрільби

Розглянемо лінійну крайову задачу з нерозділеними крайовими умовами:

$$\frac{d\vec{U}}{dx} = A(x)\vec{U}(x) + \vec{F}(x) \quad (1)$$

$$B_1\vec{U}(a) + B_2\vec{U}(b) = \vec{d} \quad (2)$$

Метод стрільби зводить крайову задачу до послідовності задач Коші. Для цього розв'яжемо $(m+1)$ задачу Коші:

$$\frac{d\vec{Y}_0}{dx} = A(x)\vec{Y}_0 + \vec{F}(x), \quad \vec{Y}_0(a) = 0 \quad (3)$$

$$\frac{d\vec{Y}_i}{dx} = A(x)\vec{Y}_i(x), \quad \vec{Y}_i(a) = \vec{\delta}_i, \quad \vec{\delta}_i = (\delta_{ij})_{j=1}^m, \quad i = \overline{1, m} \quad (4)$$

Матриця $Y(x) = (\vec{Y}_i(x))_{i=\overline{1, m}}$ називається *фундаментальною матрицею* однорідної системи (1). Розв'язок (1) шукаємо у вигляді:

$$\vec{Y}(x) = \vec{Y}_0(x) + \sum_{i=1}^m c_i \vec{Y}_i(x) \quad (5)$$

Він задовольняє (1) $\forall c_i$. Самі c_i знаходяться з (2):

$$B_1 \left(\vec{Y}_0(a) + \sum_i c_i \vec{Y}_i(a) \right) + B_2 \left(\vec{Y}_0(b) + \sum_i c_i \vec{Y}_i(b) \right) = \vec{d} \quad \text{або} \\ [B_1 + B_2 Y(b)] \vec{c} = \vec{d} - B_2 \vec{Y}_0(b) \quad (6)$$

Розв'язуючи СЛАР (6) знаходимо c_i . За єдиністю $\vec{Y}(x) = \vec{U}(x)$.

Алгоритм А1.

1) Розв'язуємо задачу Коші (3), знаходимо $\vec{Y}_0(b)$.

2) Розв'язуємо m задач Коші (4), знаходимо $Y(b)$.

3) Розв'язуємо СЛАР (6), що дає нам $c_i, i = \overline{1, m}$.

4) $\vec{Y}(x) = U(x)$ знаходимо з (5).

Складність цього алгоритму співпадає зі складністю розв'язання $(m+1)$ задач Коші.

Якщо крайові умови розділені

$$C_1 \vec{Y}(a) = \vec{d}_1, \quad C_2 \vec{Y}(b) = \vec{d}_2,$$

то можна зменшити кількість задач Коші, які необхідно розв'язати. Для цього побудуємо вектор \vec{V}_0 такий, що

$$C_1 \vec{V}^0 = \vec{d}_1 \quad (7)$$

Це завжди можна зробити, оскільки кількість рівнянь менше кількості невідомих. Далі будуємо $\vec{V}_i, i = \overline{1, k}$ такі, що

$$C_1 \vec{V}^i = 0, \quad i = \overline{1, k} \quad (8)$$

Знову це можна здійснити бо $\text{rang} C_1 = m - k$.

Розв'яжемо задачі Коші:

$$\frac{d\vec{Y}^0}{dx} = A\vec{Y}^0 + \vec{F}, \quad \vec{Y}^0(a) = \vec{V}^0 \quad (9)$$

$$\frac{d\vec{Y}^i}{dx} = A\vec{Y}^i, \quad \vec{Y}^i(a) = \vec{V}^i, \quad i = \overline{1, k}. \quad (10)$$

Сталі c_i знаходимо з умови виконання другої крайової умови.

Алгоритм А2.

- 1). Розв'язуємо СЛАР (7) – (8).
- 2). Розв'язуємо задачу Коші (9).
- 3). Розв'язуємо k задач Коші (10).
- 4). Розв'язуємо СЛАР:

$$B_2 \bar{Y}(b) \equiv C_2 \left[\bar{Y}^0(b) + \sum_{i=1}^k c_i \bar{Y}^i(b) \right] = \bar{d}_2. \quad (11)$$

- 5). Розв'язок

$$\bar{Y}(x) = \bar{Y}^0(x) + \sum_{i=1}^k c_i \bar{Y}^i(x). \quad (12)$$

Оскільки для А1 та А2 розв'язок задачі Коші шукається чисельно, то маємо фактично значення

$$\bar{Y}^i(x_n), \quad n = \overline{0, N}, \quad x_n \in [a, b].$$

Їх треба запам'ятовувати для (12). Для запобігання цього у випадку розділеної крайової задачі запишемо алгоритм А3.

Алгоритм А3.

- 1). Розв'язуємо (8) – (9).
- 2). Розв'язуємо задачу Коші (9).
- 3). Розв'язуємо задачі Коші (10) і не запам'ятовуємо $\bar{Y}^i(x_n)$, а знаходимо тільки $\bar{Y}^i(x_N) = \bar{Y}^i(b)$.
- 4). Розв'язуємо СЛАР (11).
- 5). Розв'язуємо ще одну задачу Коші:

$$\frac{d\bar{Y}}{dx} = A\bar{Y} + F, \quad \bar{Y}(a) = \bar{V}^0 + \sum_{i=1}^k c_i \bar{V}^i.$$

Тоді за формулою (12) $\bar{Y}(x) = \bar{U}(x)$.

Зрозуміло, що “стріляти”, тобто починати розв'язувати задачу Коші, треба з того боку, де задано більше крайових умов.

Суттєвий недолік алгоритмів!. Серед власних значень $A(x)$, як правило, є такі, що $\text{Re } \lambda_i(x) > 0$. Тоді лінійно незалежні розв'язки задачі Коші нарастають експоненціально. Це призводить до наростання похибок заокруглень та погано обумовленої матриці системи (6) або (11) (розв'язки $\bar{Y}^i(x)$ стають майже лінійно залежні).

Тому $[a, b]$ розбивають на проміжки $[x_{p-1}, x_p]$, $p = \overline{1, M}$, і розв'язують задачу Коші на підпроміжках, а в кінці $x = x_p$ ортогоналізують отримані розв'язки. Зрозуміло, що для $x = b$ отримують не $\bar{Y}^i(b)$, а деякі $\bar{W}^i(b)$, які залежать від $\bar{Y}^i(b)$ та відповідних перетворень ортогоналізації. З їх

допомогою по $\vec{W}^i(b)$ обчислюють $\vec{Y}^i(b)$ та “прогоняють” ці умови для всіх значень $\vec{Y}(a) = \vec{Y}^0(a) + \sum c_i \vec{Y}^i(a)$. Така ідея методу ортогональної прогонки Годунова, що широко застосовується на практиці.

Метод пристрілки

Це метод для розв’язання крайової задачі для нелінійних рівнянь аналогічний методу стрільби.

Розглянемо крайову задачу з розділеними крайовими умовами:

$$\frac{d\vec{U}}{dx} = \vec{F}(x, \vec{U}), a < x < b, \quad (1)$$

$$u_i(a) = c_i, i = \overline{k+1, m}, \quad (2)$$

$$\varphi_i(\vec{U}(b)) = d_i, i = \overline{1, k}. \quad (3)$$

При $x = a$ невідомі k початкових умов $u_i(a), i = \overline{1, k}$. Будемо їх шукати.

Розв’яжемо задачу Коші:

$$\begin{cases} \frac{d\vec{Y}}{dx} = \vec{F}(x, \vec{Y}), a < x < b; \\ \vec{Y}(a) = \vec{C} = (c_i)_{i=1}^m. \end{cases},$$

де $c_i, i = \overline{1, k}$ невідомі. Їх шукаємо з крайової умови (3):

$$f_i(c_1, \dots, c_k) \equiv \varphi_i(\vec{\varphi}(b; c_1, \dots, c_k)) - d_i = 0, i = \overline{1, k}.$$

Це система нелінійних рівнянь. Задаємо початкові значення $c_i^{(0)}, i = \overline{1, k}$. За якимось ітераційним методом знаходимо її розв’язок. Найзручніше використовувати метод січних.

Метод пристрілки найбільш прозоро виглядає для $k = 1$. В цьому випадку нам необхідно знайти тільки c_1 . Використаємо метод ділення навпіл.

Знайдемо $c_1^{(0)}$ таке, що

$$\varphi_1(\vec{y}(b; c_1^{(0)})) - d_1 > 0,$$

та $c_1^{(1)}$ таке, що

$$\varphi_1(\vec{y}(b; c_1^{(1)})) - d_1 < 0.$$

Тоді вибираємо

$$c_1^{(2)} = \frac{1}{2}(c_1^{(0)} + c_1^{(1)}).$$

З трьох $c_1^{(0)}, c_1^{(1)}, c_1^{(2)}$ вибираємо таке, що $\varphi_1(\vec{y}(b; c_1)) - d_1$ міняє знак. Процес продовжуємо до виконання умови

$$|\varphi_1(\vec{y}(b; c_1^{(k)})) - d_1| < \varepsilon,$$

де $\varepsilon > 0$ - задана точність.

Метод лінеаризації

Розглянемо задачу:

$$\frac{d\vec{U}}{dx} = \vec{F}(x, \vec{U}) \quad a < x < b, \quad (1)$$

$$\vec{\varphi}(\vec{U}(a), \vec{U}(b)) = \vec{d}. \quad (2)$$

Метод лінеаризації для задачі (1) це аналог методу Ньютона для систем нелінійних рівнянь. Нехай $\vec{Y}^0(x)$ – деяке наближення. Побудуємо його уточнення $\vec{Z}^0(x)$ до точного розв'язку $\vec{U}(x)$:

$$\vec{U}(x) = \vec{Y}^0(x) + \vec{Z}^0(x).$$

З (1) маємо $\frac{d\vec{Z}^0}{dx} = \Phi_F(x, \vec{V})\vec{Z}^0(x) + \vec{F}(x, \vec{Y}^0) - \frac{d\vec{Y}^0}{dx}$. Замінюючи середнє значення $\vec{V}(x)$ на $\vec{Y}^0(x)$ отримаємо лінійне рівняння:

$$\frac{d\vec{Z}^0}{dx} = \Phi_F(x, \vec{Y}^0)\vec{Z}^0 + \vec{F}(x, \vec{Y}^0) - \frac{d\vec{Y}^0}{dx} \quad (3)$$

Аналогічно:

$$\Phi_a(\vec{Y}^0(a), \vec{Y}^0(b))\vec{Z}^0(a) + \Phi_b(\vec{Y}^0(a), \vec{Y}^0(b))\vec{Z}^0(b) = \vec{d} - \vec{\varphi}(\vec{Y}^0(a), \vec{Y}^0(b)) \quad (4)$$

Тут $\Phi_F = \left(\frac{\partial F_i}{\partial u_j} \right)_{i,j=1,\overline{m}}$ – матриця Якобі правої частини $\vec{F}(x, \vec{U})$;

$$\Phi_a = \left(\frac{\partial \varphi_i}{\partial u_j(a)} \right)_{i,j=1,\overline{m}}, \quad \Phi_b = \left(\frac{\partial \varphi_i}{\partial u_j(b)} \right)_{i,j=1,\overline{m}} \quad \text{- матриці Якобі для } \vec{\varphi}(\vec{U}(a), \vec{U}(b))$$

по крайовим умовам в точках $x = a$ та $x = b$ відповідно. Задача (3)-(4) лінійна і розв'язується методом стрільби (з ортогоналізацією). Розв'язавши цю задачу, маємо наступне наближення $\vec{Y}^1(x) = \vec{Y}^0(x) + \vec{Z}^0(x)$. Цей процес продовжуємо до виконання умови точності $\|\vec{Z}^k(x)\| < \varepsilon$.

Недоліки методу:

1) Наявність похідної $\frac{d\vec{Y}^0}{dx}$ в правій частині. Оскільки розв'язок задач

Коші чисельний, то для її обчислення треба застосовувати формули чисельного диференціювання. Це може привести до великих похибок за рахунок нестійкості задачі чисельного диференціювання.

2) Збіжність залежить від вибору \vec{Y}^0 .

Метод продовження за параметром

Суттєвим недоліком методу лінеаризації є необхідність задавати хороше початкове наближення та чисельне диференціювання попереднього наближення. Розглянемо метод, який позбавлений цих недоліків.

Розглянемо задачу знаходження вектора $\vec{U}(x) = (u_i)_{i=1,\overline{n}}$, що задовольняє умовам:

$$\frac{d\vec{U}}{dx} = \vec{F}(x, \vec{U}) \quad a < x < b, \quad (1)$$

$$\vec{\varphi}(\vec{U}(a), \vec{U}(b)) = \vec{d}. \quad (2)$$

Нехай розв'язок цієї задачі існує та єдиний.

Розв'яжемо задачу Коші

$$\frac{d\vec{Y}}{dx} = \vec{F}(x, \vec{Y}), \quad \vec{Y}(a) = \vec{Y}^0 \quad (3)$$

Вибір \vec{Y}^0 здійснимо так, щоб було задовольнялося як можна більша кількість з крайових умов (2). Наприклад, якщо $\varphi_i(\vec{U}(a), \vec{U}(b)) \equiv u_i(a)$, то вибираємо

$$y_i^0 = d_i.$$

Обчислимо $\vec{d}^0 = \vec{\varphi}(\vec{Y}(a), \vec{Y}(b))$. Якщо $\vec{d}^0 \equiv \vec{d}$, то $\vec{Y} \equiv \vec{U}$. Але, як правило, $\vec{d}^0 \neq \vec{d}$ і тому необхідно уточнювати початкове наближення. Розглянемо параметричну крайову задачу

$$\frac{d\vec{V}}{dx} = \vec{F}(x, \vec{V}), \quad a < x < b, \quad (4)$$

$$\vec{\varphi}(\vec{V}(a), \vec{V}(b)) = \lambda \vec{d} + (1 - \lambda) \vec{d}^0, \quad (5)$$

яка залежить від параметра λ : $\vec{V} = \vec{V}(x, \lambda)$. Ясно, що $\vec{V}(x, 0) = \vec{Y}(x)$, а $\vec{V}(x, 1) = \vec{U}$.

Спробуємо продовжити розв'язок задачі (4), (5) від відомого $\vec{Y}(x)$ до шуканого $\vec{U}(x)$. Для цього продиференціюємо (4), (5) по λ :

$$\frac{d}{dx} \frac{\partial V_i}{\partial \lambda} = \sum_{j=1}^n \frac{\partial F_i}{\partial u_j} \frac{\partial V_j}{\partial \lambda},$$

$$\sum_{j=1}^n \frac{\partial \varphi_i}{\partial u_j(a)} \frac{\partial V_j(a)}{\partial \lambda} + \sum_{j=1}^n \frac{\partial \varphi_i}{\partial u_j(b)} \frac{\partial V_j(b)}{\partial \lambda} = d_i - d_i^0.$$

Позначимо $\vec{Z} = \frac{\partial \vec{V}}{\partial \lambda}$. Тоді останню систему можна записати у вигляді:

$$\frac{d\vec{Z}}{dx} = \Phi_F(x, \vec{V}) \vec{Z}, \quad a < x < b, \quad (4)$$

$$\Phi_a(\vec{V}(a), \vec{V}(b)) \vec{Z}(a) + \Phi_b(\vec{V}(a), \vec{V}(b)) \vec{Z}(b) = \vec{d} - \vec{d}^0, \quad (5)$$

$$\frac{\partial \vec{V}}{\partial \lambda} = \vec{Z}, \quad \vec{V}(x, 0) = \vec{Y}^0. \quad (6)$$

де $\Phi_F = \left(\frac{\partial F_i}{\partial u_j} \right)_{i,j=1,\overline{n}}$ - матриця Якобі правої частини рівняння (1) $\vec{F}(x, \vec{U})$;

$\Phi_a = \left(\frac{\partial \varphi_i}{\partial u_j(a)} \right)_{i,j=1,\overline{n}}$ - матриця Якобі лівої частини $\vec{\varphi}(\vec{U}(a), \vec{U}(b))$ крайової

умови (2) по першому аргументу $\vec{U}(a)$; $\Phi_b = \left(\frac{\partial \phi_i}{\partial u_j(b)} \right)_{i,j=1,\overline{n}}$ - матриця Якобі

лівої частини $\vec{\phi}(\vec{U}(a), \vec{U}(b))$ крайової умови (2) по другому аргументу $\vec{U}(b)$.

Задача (4)-(6) не простіше ніж вихідна задача (1)-(2), а ще й складніша за неї. Спростимо її, застосувавши до задачі Коші (6) чисельний метод, наприклад, метод Ейлера:

$$\vec{V}^{k+1}(x) = \vec{V}^k(x) + \Delta\lambda \vec{Z}^k(x), \quad \vec{V}^0(x) = \vec{Y}(x).$$

Тут $\vec{V}^k(x) = \vec{V}(x, \lambda_k)$, $\Delta\lambda = \lambda_{k+1} - \lambda_k$, $\lambda_0 = 0$, $\lambda_K = 1$, $k = \overline{1, K}$.

Знайдене наближення $\vec{V}^{k+1}(x)$ використовується для знаходження наступного наближення $\vec{Z}^{k+1}(x)$ лінійної крайової задачі (4), (5).

Повністю алгоритм розв'язання крайової задачі (1), (2) цим методом такий:

1. Розв'язуємо задачу Коші (3). Задаємо початкові значення $\vec{V}^0(x) = \vec{Y}(x)$.
2. Для $k = \overline{1, K}$ розв'язуємо лінійні крайові задачі:

$$\frac{d\vec{Z}^k}{dx} = \Phi_F(x, \vec{V}^k) \vec{Z}^k, \quad a < x < b,$$

$$\Phi_a(\vec{V}^k(a), \vec{V}^k(b)) \vec{Z}^k(a) + \Phi_b(\vec{V}^k(a), \vec{V}^k(b)) \vec{Z}^k(b) = \vec{d} - \vec{d}^0;$$

3. Продовжуємо розв'язок по параметру λ :

$$\vec{V}^{k+1}(x) = \vec{V}^k(x) + \Delta\lambda \vec{Z}^k(x).$$

4. Шуканий розв'язок $\vec{U}(x) \approx \vec{V}^K(x)$.

Лінійні крайові задачі пункту 2 розв'язуються, наприклад, методом стрільби. Для розв'язання задачі Коші (6) можна застосовувати більш точні методи ніж метод Ейлера.

Контрольні запитання:

1. Постановки крайових задач.
2. Алгоритми методу стрільби.
3. Алгоритм методу пристрілки.
4. Формули методу лінеаризації.
5. Метод продовження за параметром.

Література: [2], [3], [4], [6], [11].

Лекція 16. Чисельні методи математичної фізики

Мета лекції: ознайомитися з поняттям задач математичної фізики та їх класифікацією, засвоїти теоретичні основи методу сіток, одержати уявлення про алгоритми чисельного інтегрування рівнянь Лапласа та Пуассона.

План лекції:

1. Поняття задач математичної фізики та їх класифікація.

2. Чисельні методи розв'язування двовимірних рівнянь Лапласа та Пуассона.
3. Метод послідовних ітерацій. Блок-схема алгоритму.

У реальних фізичних процесах шукана функція залежить від декількох змінних, а це приводить до рівнянь в частинних похідних від шуканої функції. Як і у випадку звичайних диференціальних рівнянь (ЗДУ), в цьому випадку для вибору одного конкретного рішення, що задовольняє рівнянню в частинних похідних, необхідно задавати додаткові умови (тобто краєві умови). Найчастіше такі задачі на практиці не мають аналітичного розв'язку і доводиться використовувати чисельні методи їх рішення, зокрема метод сіток, метод скінчених різниць і так далі.

Лінійні рівняння в частинних похідних в загальному вигляді записуються у вигляді

$$A(x,y)\frac{\partial^2 u}{\partial x^2} + B(x,y)\frac{\partial^2 u}{\partial x \partial y} + C(x,y)\frac{\partial^2 u}{\partial y^2} + a(x,y)\frac{\partial u}{\partial x} + b(x,y)\frac{\partial u}{\partial y} + c(x,y)u = F(x,y), \quad (1)$$

де: A, B, C, a, b, c - задані неперервні функції двох змінних, що мають неперервні частинні похідні, u - шукана функція. Для скорочення запису введемо позначення

$$u_{xx} = \frac{\partial u}{\partial x^2}; \quad u_{xy} = \frac{\partial^2 u}{\partial x \partial y}; \quad u_{yy} = \frac{\partial^2 u}{\partial y^2}; \quad u_x = \frac{\partial u}{\partial x}; \quad u_y = \frac{\partial u}{\partial y}.$$

Розглядатимемо спрощену форму запису (1) вигляду

$$A(x,y)u_{xx} + B(x,y)u_{xy} + C(x,y)u_{yy} + a(x,y)u_x + b(x,y)u_y + c(x,y)u = F(x,y) \quad (2)$$

і розглянемо окремий випадок (7.2), коли $a = b = c = F \equiv 0$, тобто

$$A(x,y)u_{xx} + B(x,y)u_{xy} + C(x,y)u_{yy} = 0. \quad (3)$$

Приклади рівнянь з частинними похідними:

- $u_t = u_{xx}$ (одновимірне рівняння теплопровідності);
- $u_t = u_{xx} + u_{yy}$ (двовимірне рівняння теплопровідності);
- $u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\theta\theta} = 0$ (рівняння Лапласа в полярних координатах);
- $u_{tt} = u_{xx} + u_{yy} + u_{zz}$ (тривимірне хвильове рівняння);
- $u_{tt} = u_{xx} + \alpha u_t + \beta u$ (телеграфне рівняння).

Відзначимо, що в приведених раніше прикладах невідома функція u залежить більш ніж від однієї змінної. Змінні t, x , по яких проводиться диференціювання, називаються незалежними змінними.

Методи розв'язання рівнянь в частинних похідних:

- метод розділення змінних
- метод інтегральних перетворень
- метод перетворення координат
- чисельні методи
- метод теорії обурень
- метод функцій Гріна
- метод інтегральних рівнянь
- варіаційні методи
- метод розкладання по власних функціях

Для кожного класу рівнянь існує своя загальна теорія і методи розв'язання. Рівняння в частинних похідних можна класифікувати по багатьом ознакам:

1. По порядку рівняння (*порядком рівняння* називають найвищий порядок частинних похідних, що входять в рівняння):

- $u_t = u_x$ (рівняння першого порядку);
- $u_t = u_{xx}$ (рівняння другого порядку);
- $u_t = uu_{xxx} + \sin x$ (рівняння третього порядку).

2. По числу змінних (*числом змінних* називають число незалежних змінних).

- $u_t = u_{xx}$ (рівняння з двома змінними)
- $u_t = u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\theta\theta}$ (рівняння з трьома змінними r, θ, t)

3. По критерію лінійне/нелінійне.

Лінійним рівнянням другого порядку з двома незалежними змінними називається рівняння вигляду

$$Au_{xx} + Bu_{xy} + Cu_{yy} + Du_x + Eu_y + Fu = G, \quad (3)$$

де A, B, C, D, E, F і G константи або задані функції змінних x і y .

- $u_{xx} + uu_{yy} = 0, u_t = e^{-t}u_{xx} + \sin t$ (лінійні рівняння);
- $uu_{xx} + u_t = 0, xu_x + uy_y + u^2 = 0$ (нелінійні рівняння).

4. По критерію однорідне/неоднорідне.

Рівняння (1) називається однорідним, якщо права частина $G(x,y)$ тотожно дорівнює нулю для всіх x і y . Якщо $G(x,y)$ не дорівнює нулю тотожно, то рівняння називається неоднорідним.

5. По вигляду коефіцієнтів.

Якщо коефіцієнти A, B, C, D, E, F і G рівняння (1) - константи, то рівняння (1) називається рівнянням з постійними коефіцієнтами, інакше рівнянням із змінними коефіцієнтами.

Існує декілька типів лінійних рівнянь.

Параболічний тип. Рівняння параболічного типу описують процеси теплопровідності і дифузії і визначаються умовою

$$B^2 - 4AC = 0$$

Гіперболічний тип. Рівняння гіперболічного типу описують коливальні системи і хвилеві рухи і визначаються умовою

$$B^2 - 4AC > 0$$

Еліптичний тип. Рівняння еліптичного типу описують сталі процеси і визначаються умовою

$$B^2 - 4AC < 0$$

Приклади лінійних рівнянь різних типів:

$$u_t = u_{xx}, \quad B^2 - 4AC = 0 \quad (\text{параболічне});$$

$$u_{tt} = u_{xx}, \quad B^2 - 4AC = 4 \quad (\text{гіперболічне});$$

$$u_{\xi\eta} = 0, \quad B^2 - 4AC = 1 \quad (\text{гіперболічне});$$

$$u_{xx} + u_{yy} = 0, \quad B^2 - 4AC = -4 \quad (\text{еліптичне});$$

$yu_{xx} + u_{yy} = 0, \quad B^2 - 4AC = -4y$ (еліптичне при $y > 0$, параболічне при $y = 0$, гіперболічне при $y < 0$).

Розглянемо рівняння Лапласа

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \quad (4)$$

Рівняння (4) описує розповсюдження електромагнітних хвиль (полів). Розглядатимемо рівняння Лапласа в прямокутній області $\Omega = \{(x, y), 0 \leq x \leq a, 0 \leq y \leq b\}$ з краєвими умовами

$$u(0, y) = f_1(y); \quad u(a, y) = f_2(y); \quad u(x, 0) = f_3(x); \quad u(x, b) = f_4(x),$$

де f_1, f_2, f_3, f_4 - задані функції. Відмітимо, що найчастіше область буває не прямокутною.

Введемо позначення $u_{ij} = u(x_i, y_j)$. Накладаємо на прямокутну область сітку $x_i = h \cdot i; i=0, 1, \dots, n; y_j = l \cdot j; j=0, 1, \dots, m$. Тоді $x_n = h \cdot n, y_m = l \cdot m$.

Частинні похідні апроксимуємо по формулах

$$\frac{\partial^2 u}{\partial x^2} = \frac{u_{i+1,j} - 2 \cdot u_{i,j} + u_{i-1,j}}{h^2} + O(h^2);$$

$$\frac{\partial^2 u}{\partial y^2} = \frac{u_{i,j+1} - 2 \cdot u_{i,j} + u_{i,j-1}}{l^2} + O(l^2),$$

і замінимо рівняння Лапласа конечнорозностним рівнянням

$$\frac{u_{i+1,j} - 2 \cdot u_{i,j} + u_{i-1,j}}{h^2} + \frac{u_{i,j+1} - 2 \cdot u_{i,j} + u_{i,j-1}}{l^2} = 0, \quad (5)$$

де: $i=1, \dots, n-1, j=1, \dots, m-1$ (тобто для внутрішніх вузлів).

Похибка заміни диференціального рівняння різницеvim складає величину $O(h^2 + l^2)$. Виразимо $u_{i,j}$ при $h = l$, і отримуємо систему

$$\begin{aligned} u_{ij} &= (u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}) / 4; \\ u_{i0} &= f_3(x_i); \\ u_{im} &= f_4(x_i); \\ u_{0j} &= f_1(y_j); \\ u_{nj} &= f_2(y_j). \end{aligned} \quad (6)$$

Систему (6) лінійних алгебраїчних рівнянь можна розв'язати будь-яким ітераційним методом (Зейделя, простих ітерацій і так далі). При цій побудові системи використовувалася схема типу “хрест” (рис. 1).

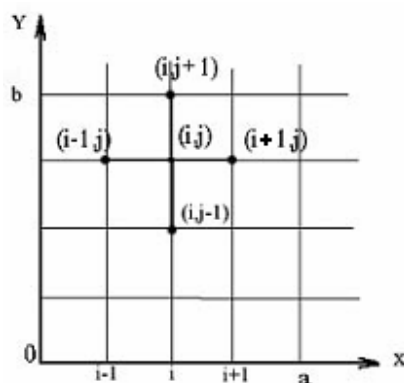


Рис. 1. Схема “Хрест”

Будуємо послідовність ітерацій за методом Гауса-Зейделя

$$u_{ij}^{(s+1)} = \frac{1}{4} (u_{i-1,j}^{(s+1)} + u_{i+1,j}^{(s)} + u_{i,j+1}^{(s)} + u_{i,j-1}^{(s+1)}),$$

де s -поточна ітерація.

Умова закінчення ітераційного процесу

$$\max_{i,j} |u_{ij}^{(s+1)} - u_{ij}^{(s)}| < \varepsilon. \quad (7)$$

Умова (7) ненадійно і на практиці використовують інший критерій

$$\max_{i,j} |u_{ij}^{(s+1)} - u_{ij}^{(s)}| \leq \varepsilon(1 - \nu),$$

$$\text{де } \nu = \frac{\max_{i,j} |u_{ij}^{(s+1)} - u_{ij}^{(s)}|}{\max_{i,j} |u_{ij}^{(s)} - u_{ij}^{(s-1)}|}.$$

Схема “крест” - явная устойчивая схема (малое изменение входных данных ведет к малому изменению выходных данных).

Контрольні запитання:

1. Поняття лінійних рівнянь в частинних похідних.

2. Класифікація рівнянь з частинними похідними.
3. Перелік методів розв'язання рівнянь в частинних похідних.
4. Рівняння Лапласа.
5. Метод сіток для рівняння Лапласа.

Література [1], [2], [3], [6], [7].

Лекція 17. Чисельні методи розв'язування одновимірних рівнянь теплопровідності

Мета лекції: ознайомитися з чисельними методами розв'язування одновимірних рівнянь теплопровідності, засвоїти теоретичні основи застосування методу сіток до параболічних рівнянь, одержати уявлення про алгоритми чисельного інтегрування рівнянь теплопровідності.

План лекції:

1. Чисельні методи розв'язування одновимірних рівнянь теплопровідності.
2. Поняття часового шару.
3. Двошарові скінчено-різницеві схеми.
4. Оцінка похибки апроксимації.
5. Умова стійкості алгоритму.
6. Блок-схема алгоритму.

Як приклад рівняння параболічного типу зупинимося на рівнянні теплопровідності для однорідного стрижня завдовжки $0 \leq x \leq l$:

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} \quad (1)$$

де $u = u(x, t)$ - температура і t - час. Припустимо, що $a = 1$. Тобто від рівняння (1) перейдемо до рівняння

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad (2)$$

Нехай заданий розподіл температури $u(x, 0) = \varphi_1(x)$ у початковий момент часу $t = 0$ і закони зміни температури залежно від часу на кінцях стрижня $x = 0$ і $x = l$: $u(0, t) = \varphi_2(t)$; $u(l, t) = \varphi_3(t)$. Потрібно знайти розподіл температури $u = u(x, t)$ уздовж стрижня завдовжки $0 \leq x \leq l$ у будь-який момент часу t . Функція $u(x, t)$ має бути неперервна і двічі неперервно диференціюєма по своїх змінних в області $\Omega = \{t, x : t \in [0, T], x \in [0, l]\}$.

Область Ω замінимо сітковою, розбиваючи її за допомогою кроку h по x і за допомогою кроку k по t . В результаті заміни неперервної області Ω дискретною множиною вузлових точок $\Omega_h = \{x_i, t_j\}, i = \overline{0, n}; j = \overline{0, m}$, вихідна задача змінюється. Тепер шукатимемо рішення $u(x, t)$ тільки на

дискретній множині Ω . Тобто $u_{ij} = u(x_i, t_j)$ - двовимірна таблиця значень шуканої функції у вузлових точках.

Представимо рівняння (51) в скінчено-різницевої формі, замінюючи $\frac{\partial u}{\partial t}$ і $\frac{\partial^2 u}{\partial x^2}$ скінчено-різницевим аналогом у вузлових точках x_i, t_j :

$$\frac{\partial u(x_i, t_j)}{\partial t} = \frac{u(x_i, t_j + k) - u(x_i, t_j - k)}{2k} = \frac{u_{ij+1} - u_{ij-1}}{2k}$$

$$\frac{\partial^2 u(x_i, t_j)}{\partial x^2} = \frac{u(x_i + h, t_j) - 2u(x_i, t_j) + u(x_i - h, t_j)}{h^2} = \frac{u_{i+1j} - 2u_{ij} + u_{i-1j}}{h^2}$$

Отримаємо скінчено-різницевий аналог початкового завдання: потрібно знайти значення функції $u(x_i, t_j)$, що задовольняє скінчено-різницевому рівнянню вигляду:

$$\frac{u_{ij+1} - u_{ij-1}}{2k} = \frac{u_{i+1j} - 2u_{ij} + u_{i-1j}}{h^2}, \quad i = \overline{1, n-1}; j = \overline{1, m-1} \quad (3).$$

і додатковим умовам:

$$u(x_i, 0) = \varphi_1(x_i), i = \overline{0, n}, u(0, t_j) = \varphi_2(t_j), j = \overline{0, m}, u(l, t_j) = \varphi_3(t_j), j = \overline{0, m}$$

Отримаємо систему лінійних рівнянь алгебри, яку можна вирішити будь-яким відомим методом. Дослідження показали, що значення h і k мають бути зв'язані між собою таким чином: $k = \delta h^2$, де $\delta = Const$, $0 < \delta < 1/2$. Апроксимуємо рівняння (2) скінчено-різницевим

$$\frac{u_{ij+1} - u_{ij-1}}{2\delta h^2} = \frac{2\delta u_{i+1j} - 4\delta u_{ij} + 2\delta u_{i-1j}}{2\delta h^2} \quad (4)$$

Вирішуючи систему (4) з урахуванням додаткових умов, отримаємо $u(x_i, t_j)$ - шукану функцію в точках $x_i, t_j, i = \overline{0, n}; j = \overline{0, m}$.

Другий варіант скінчено-різницевого аналога диференціального рівняння (2), т.з. *явна схема* (рис. 1), виходить за рахунок того, що перші похідні $\frac{\partial u}{\partial t}$ в вузлових точках x_i, t_j представлені в вигляді:

$$\frac{\partial u(x_i, t_j)}{\partial t} = \frac{u(x_i, t_j + h) - u(x_i, t_j)}{k},$$

а друга похідна залишається попередньою. Отримаємо двошарову схему в скінчено-різницевої формі:

$$\frac{u_{ij+1} - u_{ij}}{k} = \frac{u_{i+1j} - 2u_{ij} + u_{i-1j}}{h^2}.$$

Покладаючи $k = \delta h^2$, отримуємо $u_{ij+1} - u_{ij} = \delta u_{i+1j} - 2\delta u_{ij} + \delta u_{i-1j}$, або $u_{ij+1} = \delta u_{i+1j} + (1 - 2\delta)u_{ij} + \delta u_{i-1j}$, $i = \overline{1, n-1}$; $j = \overline{0, m-1}$. По цій формулі для кожного значення u_{ij+1} для шара $(j+1)$ по осі t використовуються три значення $u_{ij}, u_{i-1j}, u_{i+1j}$ на попередньому шарі з номером j . Для початку обчислень використовуємо додаткові умови.

В результаті розв'язання задачі в скінчено-різницевої формі ми набуваємо значень шуканої функції в точках x_i, t_j , які є наближеним розв'язком вихідної задачі. На практиці вважають $\delta = 1/6$, тоді розрахункова формула спрощується і набуває наступного вигляду:

$$u_{ij+1} = \frac{1}{6}(u_{i+1j} + 4u_{ij} + u_{i-1j}), \quad i = \overline{1, n-1}; \quad j = \overline{0, m}.$$

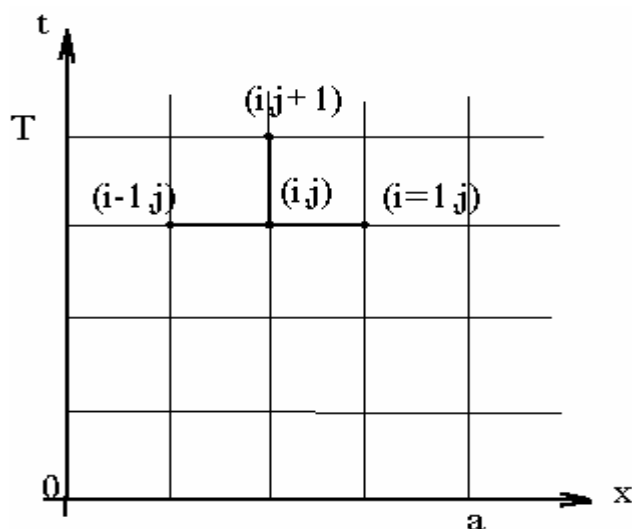


Рис. 1. Явна схема

Дана розрахункова формула дає якнайкраще наближення до шуканого рішення, забезпечуючи стійкість скінчено-різницевої схеми і якнайкращу апроксимацію диференціального рівняння скінчено-різницевим.

У разі неявної схеми (рис. 2) використовується інший вид апроксимації і нове співвідношення між кроками h і k у вигляді $h^2 = kS$.

Вихідне диференціальне рівняння (2) апроксимується скінчено-різницевим рівнянням вигляду

$$\frac{u_{ij+1} - u_{ij}}{h^2/S} = \frac{u_{i+1j+1} - 2u_{ij+1} - u_{i-1j+1}}{h^2} \quad (5)$$

Початкові і граничні умови залишаються тими ж, що у попередньому випадку. Для розв'язання системи лінійних рівнянь (5) застосовується метод прогону.

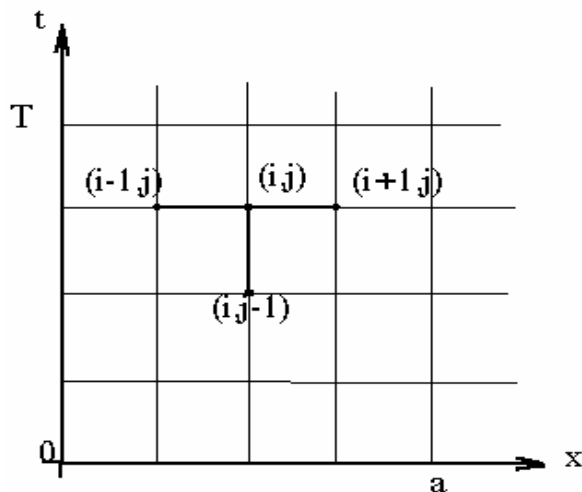


Рис. 2. Неявна схема

Суть методу прогону полягає в тому, що спочатку обчислюються значення $u_{i0} = \varphi_{1i}$, вибирається значення S з метою отримання необхідної швидкості просування осі t .

Позначимо $u_{ij} = u(x_i, t_j)$, $\varphi_1(x_i) = \varphi_{1i}$, $\varphi_2(t_j) = \varphi_{2j}$, $\varphi_3(t_j) = \varphi_{3j}$. У прямому ході на черговому $(j+1)$ тимчасовому шарі обчислюються допоміжні функції:

$$a_{1j+1} = \frac{1}{2+S}, \quad b_{1j+1} = \varphi_{2j+1} + Su_{1j}.$$

У зворотному ході обчислюються значення шуканої функції на $(j+1)$ шарі по формулі $u_{ij+1} = a_{ij+1}(b_{ij+1} + u_{i+1j+1})$. Величина $u_{nj+1} = \varphi_{3j+1}$ є значенням шуканої функції в точці (x_n, t_{j+1}) , а $u_{0j+1} = \varphi_{2j+1}$ - в точці (x_0, t_{j+1}) . Похибка методу $O(h^2 + k)$. З аналізу стійкості неявної схеми витікає, що слід призначати $S > 0$.

Контрольні запитання:

1. Рівняння теплопровідності для однорідного стрижня.
2. Метод сіток для рівняння теплопровідності.
3. Двошарова схема для рівняння теплопровідності.
4. Неявна схема для рівняння теплопровідності.
5. Умова стійкості різницевих схем для рівняння теплопровідності.

Література [2], [3], [7], [10].

Лекція 18. Чисельні методи розв'язування одновимірних хвильових рівнянь

Мета лекції: ознайомитися з чисельними методами розв'язування одновимірних хвильових рівнянь, засвоїти теоретичні основи застосування

методу сіток до гіперболічних рівнянь, одержати уявлення про алгоритми чисельного інтегрування гіперболічних рівнянь.

План лекції:

1. Чисельні методи розв'язування одновимірних хвильових рівнянь.
2. Задача коливання струни та стержня.
3. Трьохшарові скінчено-різницеві схеми.
4. Оцінка похибки апроксимації.
5. Умова стійкості алгоритму.
6. Блок-схема алгоритму.

Розглянемо рівняння коливання однорідної і обмеженої струни. Завдання полягає у відшуканні функції $u(x,t)$ при $t > 0$, що задовольняє рівнянню гіперболічного типу

$$\frac{\partial^2 u}{\partial t^2} = C \cdot \frac{\partial^2 u}{\partial x^2}, \quad (1)$$

де: $0 < x < a$; $0 < t < T$, початковим умовам

$$\begin{aligned} u(x, 0) &= f(x); \\ \frac{\partial u}{\partial t}(x, 0) &= g(x); \\ 0 &\leq x \leq a \end{aligned} \quad (2)$$

і краєвим умовам

$$\begin{aligned} u(0, t) &= \mu_1(t); \\ u(a, t) &= \mu_2(t); \\ 0 &\leq t \leq T. \end{aligned} \quad (3)$$

Замінімо t на $\sqrt{C}t$ і отримаємо рівняння

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2},$$

надалі вважатимемо, що $C = 1$.

Для побудови різницевої схеми рішення задачі (1) - (3) побудуємо в області $D = \{(x,t), 0 \leq x \leq a, 0 \leq t \leq T\}$ сетку $x_i = ih$; $i=0,1,\dots,n$; $a = h \cdot n$; $t_j = j\tau$; $j=0,1,\dots,m$; $m\tau = T$.

Апроксимуємо рівняння (1) різницевими похідними другого порядку точності по крокам h, τ :

$$\frac{u_{ij-1} - 2 \cdot u_{ij} + u_{ij+1}}{\tau^2} = \frac{u_{i+1,j} - 2 \cdot u_{i,j} + u_{i-1,j}}{h^2}. \quad (4)$$

Вважаючи $\lambda = \tau/h$ перепишемо (4), виразив $U_{i,j+1}$. Таким чином отримаємо трьохшарову різницеву схему

$$u_{i,j+1} = \lambda^2 (u_{i+1,j} + u_{i-1,j}) + 2 \cdot (1 - \lambda^2) \cdot u_{i,j} - u_{i,j-1}, \quad (5)$$

де: $i=1, \dots, n; j=1, \dots, m$. Задаємо нульові граничні умови $\mu_1(t)=0, \mu_2(t)=0$. Тоді в (5) маємо $u_{0j} = 0, u_{nj} = 0$ для всіх j .

Схема (5) називається трьохшаровою, оскільки вона зв'язує значення шуканої функції на трьох тимчасових шарах $j-1, j, j+1$.

Чисельне рішення задачі полягає в обчисленні наближених значень u_{ij} розв'язку $u(x,t)$ в вузлах (x_i, t_j) при $i=1, \dots, n; j=1, \dots, m$. Алгоритм розв'язання заснований на тому, що розв'язок на кожному наступному шарі ($j=2, 3, \dots, n$) можна отримати за допомогою перерахунку розв'язків з двох попередніх шарів ($j=0, 1, \dots, n-1$) по формулі (5). При $j=0$ розв'язок відомий з початкової умови $u_{i0} = f(x_i)$. Для обчислення розв'язку на першому шарі ($j=1$) покладемо

$$\frac{\partial u}{\partial t}(x,0) \approx \frac{u(x,\tau) - u(x,0)}{\tau}, \quad (6)$$

тоді $u_{i1} = u_{i0} - \tau g(x_i)$, $i=1, 2, \dots, n$. Тепер для обчислення розв'язків на наступних шарах можна використовувати формулу (5).

Описана схема апроксимує задачу (1) - (3) з точністю $O(\tau + h^2)$. Невисокий порядок апроксимації по τ пояснюється грубістю апроксимації по формулі (6).

Схема буде стійкою, якщо виконана умова $\tau < h$.

Контрольні запитання:

1. Поняття гіперболічного рівняння.
2. Рівняння коливання однорідної і обмеженої струни.
3. Трьохшарова різницєва схема для рівняння коливання струни.
4. Алгоритм розв'язання трьохшарової різницєвої схеми.
5. Умова стійкості трьохшарової різницєвої схеми.

Література [2], [3], [7].

МЕТОДИЧНІ ВКАЗІВКИ
до проведення практичних занять з дисципліни
“ЧИСЕЛЬНІ МЕТОДИ В ІНФОРМАТИЦІ”

ПЕРЕДМОВА

Курс “Чисельні методи в інформатиці” – обов'язковий компонент загальної та професійної освіти. Значення курсу у загальноосвітній підготовці визначається насамперед тим, при побудові математичних моделей складних, зокрема соціально-економічних систем, часто виникають задачі, для яких знаходження аналітичного розв'язку або надто ускладнене, або неможливе. Це обумовлює необхідність розробки спеціальних математичних методів та алгоритмів, що дозволяють одержати чисельний розв'язок та оцінити його точність.

Дисципліна “Чисельні методи в інформатиці” призначена для оволодіння майбутніми бакалаврами спеціальності «Інтелектуальні системи прийняття рішень» основами знань у галузі комп'ютерних методів дослідження математичних моделей, на базі яких провадиться подальше вивчення спеціальних дисциплін, пов'язаних з фаховою діяльністю.

Метою дисципліни є засвоєння теоретичних основ, формування у студентів практичних навичок щодо використання основних методів чисельного розв'язання складних математичних моделей.

Завдання курсу полягає у вивченні теоретичних відомостей та набуття студентами практичних навичок чисельного розв'язання задач великої розмірності, систем алгебраїчних та диференціальних рівнянь, нелінійних математичних моделей тощо, опануванні сучасними пакетами прикладних програм, що дозволяють здійснювати чисельні розрахунки.

Предмет навчальної дисципліни - математичний інструментарій, що дозволяє знаходити чисельні розв'язки складних математичних моделей.

Теоретичним фундаментом дисципліни є вища та дискретна математика, інформатика, теорія ймовірностей, теорія програмування та алгоритмічні мови. Практичним засобом реалізації чисельних методів є сучасна комп'ютерна техніка та прикладне програмне забезпечення.

У результаті вивчення дисципліни студенти повинні:

- знати теоретичні основи чисельного розв'язання складних математичних моделей;
- одержати практичні навички побудови алгоритмів чисельного розв'язання прикладних задач та їх реалізації на персональних комп'ютерах;
- опанувати сучасне програмне забезпечення, що дозволяє здійснювати чисельні розрахунки.

Критерії оцінки успішності повинні відповідати навчальній програмі й найбільш важливим вимогам до знань студентів:

1. Знання фактів, явищ. Вірне, науково достовірне їх пояснення.
2. Оволодіння науковими термінами, поняттями, законами, методами, правилами; вміння користуватися ними при поясненні нових фактів, розв'язуванні різних питань і виконанні практичних завдань.
3. Максимальна ясність, точність думки, вміння відстоювати свої погляди, захищати їх.

КАРТА ПРАКТИЧНИХ ЗАНЯТЬ З НАВЧАЛЬНОЇ ДИСЦИПЛІНИ
“ЧИСЕЛЬНІ МЕТОДИ В ІНФОРМАТИЦІ”

Тема		Години	Форми контролю та звітності	Максим. кількість балів
Т. 4.	Чисельні методи розв’язування нелінійних рівнянь.	2		
Практичне заняття №1. Чисельні методи розв’язування нелінійних рівнянь			Поточний контроль	
Т. 5.	Чисельні методи розв’язування систем лінійних алгебраїчних рівнянь.	2		
Практичне заняття №2. Чисельні методи розв’язування систем лінійних алгебраїчних рівнянь.			Поточний контроль	
Т. 6.	Інтерполювання функцій	4		
Практичне заняття №3. Інтерполювання функцій			Поточний контроль	
Т. 7.	Задача наближення функцій однієї змінної	2		
Практичне заняття №4. Задача наближення функцій однієї змінної.			Поточний контроль	
Т. 8.	Чисельне диференціювання функцій	2		
Практичне заняття №5. Чисельне диференціювання функцій.			Поточний контроль	
Т. 9.	Чисельне інтегрування функції	2		
Практичне заняття №6. Чисельне інтегрування функції.			Поточний контроль	
Т. 10.	Чисельне інтегрування звичайних диференційних рівнянь	2		
Практичне заняття №7. Чисельне інтегрування звичайних диференційних рівнянь.			Поточний контроль	
Т. 11.	Чисельні методи математичної фізики	6		
Практичне заняття №8. Чисельні методи математичної фізики			Поточний контроль	
Разом годин з курсу		22		

ПОЯСНЕННЯ ЗА РОЗДІЛАМИ ТА ТЕМАМИ

ТЕМА 4. Чисельні методи розв'язування нелінійних рівнянь

Чисельні методи розв'язування нелінійних рівнянь. Алгоритми ізолювання та уточнення коренів. Графічний метод та метод табулювання. Метод ділення навпіл. Метод простої ітерації. Метод релаксації. Метод хорд та дотичних. Оцінка кількості ітерацій.

Практичне заняття №1. Чисельні методи розв'язування нелінійних рівнянь

Мета: засвоїти основні методи чисельного розв'язування нелінійних рівнянь від однієї змінної.

План заняття

1. Алгоритми ізолювання коренів.
2. Особливості застосування графічного методу та методу табулювання.
3. Алгоритми уточнення коренів.
4. Приклади застосування методів ділення навпіл, хорд та дотичних, ітерацій.

Теоретичні відомості, що необхідні для виконання даної роботи.

Розглянемо задачу знаходження коренів рівняння

$$f(x) = 0, \quad (1)$$

де $f(x)$ – задана функція дійсного змінного.

Розв'язування даної задачі можна розкласти на декілька етапів:

- а) дослідження розташування коренів (в загальному випадку на комплексній площині) та їх кратність;
- б) відділення коренів, тобто виділення областей, що містять тільки один корінь;
- в) обчислення кореня з заданою точністю за допомогою одного з ітераційних алгоритмів.

Далі розглядаються ітераційні процеси, що дають можливість побудувати числову послідовність x_n , яка збігається до шуканого кореня x_* рівняння (1).

1. Метод ділення проміжку навпіл (метод дихотомії)

Нехай $f \in C[a, b]$, $f(a)f(b) < 0$ і відомо, що рівняння (1) має єдиний корінь $x_* \in [a, b]$. Покладемо $a_0 = a$, $b_0 = b$, $x_0 = (a_0 + b_0)/2$. Якщо $f(x_0) = 0$, то $x_* = x_0$. Якщо $f(x_0) \neq 0$, то покладемо

$$a_{n+1} = \begin{cases} x_n, & \text{якщо } \text{sign } f(a_n) = \text{sign } f(x_n), \\ a_n, & \text{якщо } \text{sign } f(a_n) \neq \text{sign } f(x_n), \end{cases} \quad (2)$$

$$b_{n+1} = \begin{cases} x_n, & \text{якщо } \text{sign } f(b_n) = \text{sign } f(x_n), \\ b_n, & \text{якщо } \text{sign } f(b_n) \neq \text{sign } f(x_n), \end{cases} \quad (3)$$

$$x_{n+1} = \frac{a_{k+1} + b_{k+1}}{2}, \quad n = 0, 1, 2, \dots, \quad (4)$$

і обчислимо $f(x_{n+1})$. Якщо $f(x_{n+1}) = 0$, то ітераційний процес зупинимо і будемо вважати, що $x_* \approx x_{n+1}$. Якщо $f(x_{n+1}) \neq 0$, то повторюємо розрахунки за формулами (2)-(4).

З формул (2), (3) видно, що $\text{sign } f(a_{n+1}) = \text{sign } f(a_n)$ і $\text{sign } f(b_{n+1}) = \text{sign } f(b_n)$. Тому $f(a_{n+1})f(b_{n+1}) < 0$, а отже шуканий корінь x_* знаходиться на проміжку $[a_{n+1}, b_{n+1}]$. При цьому має місце оцінка збіжності

$$|x_n - x_*| \leq \frac{b - a}{2^{n+1}}. \quad (5)$$

Звідси випливає, що кількість ітерацій, які необхідно провести для знаходження наближеного кореня рівняння (1) з заданою точністю ε задовольняє співвідношенню

$$n \geq \left\lceil \log_2 \frac{b - a}{\varepsilon} \right\rceil. \quad (6)$$

де $[c]$ – ціла частина числа c .

Серед переваг даного методу слід відзначити простоту реалізації та надійність. Послідовність $\{x_n\}$ збігається до кореня x_* для довільних неперервних функцій $f(x)$. До недоліків можна віднести невисоку швидкість збіжності методу та неможливість безпосереднього узагальнення систем нелінійних рівнянь.

2. Метод простої ітерації

Метод простої ітерації застосовується до розв'язування нелінійного рівняння виду

$$x = \varphi(x). \quad (7)$$

Перейти від рівняння (1) до рівняння(7) можна багатьма способами, наприклад, вибравши

$$\varphi(x) = x + \psi(x) f(x), \quad (8)$$

де $\psi(x)$ – довільна знакостала неперервна функція.

Вибравши нульове наближення x_0 , наступні наближення знаходяться за формулою

$$x_{n+1} = \varphi(x_n), \quad n = 0, 1, 2, \dots \quad (9)$$

Наведемо достатні умови збіжності методу простої ітерації.

Теорема 1. Нехай для вибраного початкового наближення x_0 на проміжку

$$S = \{x : |x - x_0| \leq \delta\} \quad (10)$$

функція $\varphi(x)$ задовольняє умові Ліпшиця

$$|\varphi(x') - \varphi(x'')| \leq q|x' - x''|, \quad x', x'' \in S \quad (11)$$

де $0 < q < 1$, і виконується нерівність

$$|\varphi(x_0) - x_0| \leq (1 - q)\delta. \quad (12)$$

Тоді рівняння (7) має на проміжку S єдиний корінь x_* , до якого збігається послідовність (9), причому швидкість збіжності визначається нерівністю

$$|x_n - x_*| \leq \frac{q^n}{1-q} |\varphi(x_0) - x_0|. \quad (13)$$

Зауваження: якщо функція $\varphi(x)$ має на проміжку S неперервну похідну $\varphi'(x)$, яка задовольняє умові

$$|\varphi'(x)| \leq q < 1, \quad (14)$$

то функція $\varphi(x)$ буде задовольняти умові (11) теореми 1.

З (13) можна отримати оцінку кількості ітерацій, які потрібно провести для знаходження розв'язку задачі (7) з наперед заданою точністю ε :

$$n \geq \left\lceil \frac{\ln \frac{|\varphi(x_0) - x_0|}{(1-q) \cdot \varepsilon}}{\ln(1/q)} \right\rceil + 1. \quad (15)$$

Наведемо ще одну оцінку, що характеризує збіжність методу простої ітерації:

$$|x_n - x_*| \leq \frac{q}{1-q} |x_n - x_{n-1}|. \quad (16)$$

3. Метод релаксації

Для збіжності ітераційного процесу (9) суттєве значення має вибір функції $\varphi(x)$. Зокрема, якщо в (8) вибрати $\psi(x) = \tau = \text{const}$, то отримаємо метод релаксації.

$$x_{n+1} = x_n + \tau f(x_n), \quad n = 0, 1, 2, \dots, \quad (17)$$

який збігається при

$$-2 < \tau f'(x) < 0. \quad (18)$$

Якщо в деякому околі кореня виконуються умови

$$f'(x) < 0, \quad 0 < m_1 < |f'(x)| < M_1, \quad (19)$$

то метод релаксації збігаються при $\tau \in (0, 2/M_1)$. Збіжність буде найкращою при

$$\tau = \tau_{\text{opt}} = 2/(m_1 + M_1). \quad (20)$$

При такому виборі τ для похибки $z_n = x_n - x_*$ буде мати місце оцінка

$$|z_n| \leq q^n |z_0|, \quad n = 0, 1, 2, \dots, \quad (21)$$

де $q = (M_1 - m_1)/(M_1 + m_1)$.

Кількість ітерацій, які потрібно провести для знаходження розв'язку з точністю ε визначається нерівністю

$$n \geq \left\lceil \frac{\ln(|z_0|/\varepsilon)}{\ln(1/q)} \right\rceil + 1. \quad (22)$$

Зауваження: якщо виконується умова $f'(x) > 0$, то ітераційний метод (17) потрібно записати у вигляді

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

4. Метод Ньютона

Метод Ньютона застосовується до розв'язування задачі (1), де $f(x)$ є неперервно-диференційованою функцією. На початку обчислень вибирається початкове наближення x_0 . Наступні наближення обчислюються за формулою

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, \dots, \quad f'(x_n) \neq 0. \quad (23)$$

З геометричної точки зору x_{n+1} є значенням абсциси точки перетину дотичної до кривої $y=f(x)$ в точці $(x_n, f(x_n))$ з віссю абсцис. Тому метод Ньютона називають також методом дотичних.

Теорема 2. Якщо $f(x) \in C^2[a, b]$, $f(a)f(b) < 0$, а $f''(x)$ не змінює знака на $[a, b]$, то виходячи з початкового наближення $x_0 \in [a, b]$, що задовольняє умові $f(x_0)f''(x_0) > 0$, можна обчислити методом Ньютона єдиний корінь x_* рівняння (1) з будь-якою степінню точності.

Теорема 3. Нехай x_* – простий дійсний корінь рівняння (1) і $f(x) \in C^2(S)$, де $S = \{x : |x - x_*| \leq \delta\}$,

$$0 < m_1 = \min_{x \in S} |f'(x)|, \quad M_2 = \max_{x \in S} |f''(x)|, \quad (24)$$

причому

$$q = \frac{M_2 |x_0 - x_*|}{2m_1} < 1. \quad (25)$$

Тоді для $x_0 \in S$ метод Ньютона збігається, причому для похибки справедлива оцінка

$$|x_n - x_*| \leq q^{2^n - 1} |x_0 - x_*|. \quad (26)$$

З оцінки (26) видно, що метод Ньютона має квадратичну збіжність, тобто похибка на $(n+1)$ -й ітерації пропорційна квадрату похибки на n -й ітерації.

Модифікований метод Ньютона

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)}, \quad n = 0, 1, 2, \dots \quad (27)$$

дозволяє не обчислювати похідну $f'(x)$ на кожній ітерації, а отже і позбутися можливого ділення на нуль. Однак цей алгоритм має тільки лінійну збіжність.

Кількість ітерацій, які потрібно провести для знаходження розв'язку задачі (1) з точністю ε задовольняє нерівності

$$n \geq \left\lceil \log_2 \left(\frac{\ln(|x_0 - x_*|/\varepsilon)}{\ln(1/q)} \right) + 1 \right\rceil + 1. \quad (28)$$

Приклади розв'язання типових задач

Приклад 1. Розв'язати рівняння

$$x + \sin x - 1 = 0 \quad (29)$$

методом ділення проміжку навпіл з точністю $\varepsilon=10^{-4}$.

Розв'язання. Спочатку знайдемо проміжок, де рівняння має єдиний корінь. Оскільки похідна функції $f(x) = x + \sin x - 1$ не змінює знак, то корінь

у рівнянні (29) буде один. Легко бачити, що $f(0)=-1 < 0$, а $f\left(\frac{\pi}{2}\right) = \frac{\pi}{2} > 0$. Отже

корінь належить проміжку $\left[0, \frac{\pi}{2}\right]$. Виберемо $a_0 = 0$, $b_0 = \frac{\pi}{2}$. Згідно з

формулою (6), отримаємо, що для знаходження кореня з точністю 10^{-4} необхідно провести 13 інтеграцій. Відповідні значення x_n наведені в табл. 1.

Табл.1

	x_n	$f(x_n)$
	0785398E+00	0492505E+00
	0392699E+00	0224617E+00
	0589049E+00	0144619E+00
	0490874E+00	0377294E-01
	0539961E+00	0540639E-01
	0515418E+00	0831580E-02
	0503146E+00	0146705E-01
	0509282E+00	0316819E-02
	0512350E+00	0257611E-02
	0510816E+00	0295467E-03
0	0511583E+00	0114046E-02
1	0511199E+00	0422535E-03
2	0511007E+00	0635430E-04
3	0510911E+00	0116016E-03

Приклад 2. Знайти додатні корені рівняння

$$x^3 - x - 1 = 0 \quad (30)$$

методом простої ітерації з точністю $\varepsilon=10^{-4}$.

Розв'язання. Графічне дослідження рівняння (30) показує, що існує єдиний дійсний додатній корінь цього рівняння і він належить проміжку $[1,2]$. Оскільки на цьому проміжку $x \neq 0$, то рівняння (30) можна подати у вигляді

$$x = \sqrt{\frac{1}{x} + 1}. \quad (31)$$

Позначимо $\varphi(x) = \left(\frac{1}{x} + 1\right)^{1/2}$. Перевіримо виконання умов теореми про збіжність методу простої ітерації. Виберемо $x_0 = 1,5$, тоді $\delta = 0,5$. Розглянемо

$$\varphi'(x) = -\frac{1}{2\sqrt{x^3 + x^4}}; \quad \max_{1 \leq x \leq 2} |\varphi'(x)| = \frac{1}{2\sqrt{2}},$$

тобто $q = \frac{1}{2\sqrt{2}}$.

тоді

$$|\varphi(x_0) - x_0| = \left| \sqrt{\frac{2}{3} + 1} - 1,5 \right| = 0,205, \quad (1 - q)\delta = 0,5 \cdot \left(1 - \frac{1}{2\sqrt{2}}\right) \approx 0,3232,$$

а отже умова (12) виконується. З формули (15) маємо, що кількість ітерацій, які необхідно провести для знаходження кореня з точністю $\varepsilon = 10^{-4}$ повинна задовольняти умові $n \geq 8$. Відповідні значення x_n та $x_n - \varphi(x_n)$ наведені в табл.2.

Табл.2

	x_n	$x_n - \varphi(x_n)$
	0150000E+01	0209006E+00
	0129099E+01	0411454E-01
	0133214E+01	0901020E-02
	0132313E+01	0193024E-02
	0132506E+01	0415444E-03
	0132464E+01	0892878E-04
	0132473E+01	0191927E-04
	0132471E+01	0417233E-05
	0132472E+01	0953674E-06

Виходячи з нерівності (16) і отриманих результатів видно, що для досягнення заданої точності достатньо було провести 5 ітерацій ($n=5$). Взагалі слід відзначити, що апостеріорна оцінка (16) є більш точною і її використання може заощадити деяку кількість обчислень.

Приклад 3. Методом релаксації знайти найменший за модулем від'ємний корінь рівняння

$$x^3 - 3x^2 - 1 = 0 \quad (32)$$

з точністю $\varepsilon = 10^{-4}$.

Розв'язання. Спочатку виділимо корені рівняння (32) користуючись наступною таблицею

Табл.3

x	4	3	2	1				
$\text{sign}f(x)$								

З даної таблиці видно, що рівняння має три корені розташовані на проміжках $[-3;-2]$, $[-1;0]$, $[0;1]$. Будемо знаходити корінь на проміжку $[-1;0]$. Обчисливши значення $f(-0,5)=-0,375$ можна уточнити проміжок існування кореня $[-1;-0,5]$.

Позначимо $f(x)=x^3-3x^2-1$. Тоді $f'(x)=3x^2+6x < 0$, $x \in [-1;-0,5]$ і є монотонно зростаючою функцією на $[-1;-0,5]$ (оскільки $f''(x)=6x+6 \geq 0$).

Тому
$$m_1 = \min_{x \in [-1;-0,5]} |f'(x)| = |f'(-0,5)| = 2,25,$$

$$M_1 = \max_{x \in [-1;-0,5]} |f'(x)| = |f'(-1)| = 3.$$

Тоді, відповідно до формул (20) і (21), будемо мати вигляд

$$x_{n+1} = x_n + \tau_{\text{опт}} (x_n^3 + 3x_n^2 - 1). \quad (33)$$

Вибравши за початкове наближення точку $x_0=-0,5$ будемо мати оцінку $|z_0| \leq 0,5$, а кількість ітерацій, які потрібно провести для знаходження розв'язку з точністю $\varepsilon=10^{-4}$ буде дорівнювати 5 (див. (22)). В табл. 4 наведені відповідні дані ітераційної послідовності:

Табл.4

	x_n	$f(x_n)$
	0500000E+00	0142857E+00
	0642857E+00	0985700E-02
	0652714E+00	0105500E-04
	0652704E+00	0596046E-07
	0652704E+00	0000000E+00
	0652704E+00	0000000E+00

Із наведених даних видно, що необхідна точність досягається раніше 5-ї ітерації. Це досить характерно для апріорних оцінок типу (22).

Приклад 4. Методом Ньютона знайти найменший додатний корінь рівняння

$$x^3+3x^2-1=0 \quad (34)$$

з точністю $\varepsilon=10^{-4}$.

Розв'язання. З табл. 3 видно, що рівняння (34) має єдиний додатний корінь, що належить проміжку $[0;1]$. обчислимо $f(0,5)=-0,125$. Тепер будемо шукати корінь на проміжку $[0,5;1]$. Нехай $f(x)=x^3+3x^2-1$. Тоді $f'(x)=3x^2+6x > 0$, $f''(x)=6x+6 > 0$, $x \in [0,5;1]$.

$$m_1 = \min_{x \in [0,5;1]} |f'(x)| = |f'(0,5)| = 3,75,$$

$$M_2 = \max_{x \in [-1;-0,5]} |f''(x)| = |f''(1)| = 12.$$

Виберемо $x_0=1$, тоді $|x_0 - x_*| \leq 0,5$. З формули (25) маємо

$$q = \frac{12 \cdot 0,5}{2 \cdot 3,75} = 0,8 < 1.$$

Тобто всі умови теореми про збіжність методу Ньютона виконані. З формули (28) маємо, що для досягнення заданої точності достатньо провести 7 ітерацій. Відповідні обчислення наведені в табл. 5.

Табл.5

	x_n	$f(x_n)$
	01000000E+01	03000000E+01
	06666667E+00	06296297E+00
	05486111E+00	06804019E-01
	05323902E+00	01218202E-02
	05320890E+00	04395228E-06
	05320889E+00	04230802E-07
	05320889E+00	04230802E-07
	05320889E+00	04230802E-07

Задачі для самоконтролю

Знайти одним з ітераційних методів дійсні корені рівнянь з точністю ε (наприклад $\varepsilon=10^{-4}$).

- 1) $x^3 - 5x^2 + 4x + 0,092 = 0$
- 2) $x^3 - 4x^2 - 7x + 13 = 0$
- 3) $x^4 + x^3 - 6x^2 + 20x - 16 = 0$
- 4) $x^3 + \sin x - 12x + 1 = 0$
- 5) $x^3 - 10x^2 + 44x + 29 = 0$
- 6) $x + \sin x - 12x = 0,25$
- 7) $3x + \cos x + 1 = 0$
- 8) $x^3 - 3x^2 - 17x + 22 = 0$
- 9) $x^4 - 2x^3 - 3,74x^3 + 8,18x - 3,48 = 0$
- 10) $x^2 + 4\sin x - 1 = 0$
- 11) $x^3 + 4\sin x = 0$
- 12) $x^4 - 10x^3 + 48,16x^2 + 108,08x + 70,76 = 0$
- 13) $x^4 - 3x^3 + 20x^2 + 44x + 54 = 0$
- 14) $x^3 - 3x^2 - 14x - 8 = 0$
- 15) $x^3 - x - 1 = 0$
- 16) $3x - \cos x - 1 = 0$
- 17) $3x^2 - \cos^2 \pi x = 0$

- 18) $x^2 + 4\sin x = 0$
- 19) $(x-1)^3 + 0,5e^x = 0$
- 20) $x^3 + 4x - 6 = 0$
- 21) $x^3 - 2x^2 + x + 1 = 0$
- 22) $x^2 \lg x - 1 = 0$
- 23) $x^3 + 6x^2 + 9x + 2 = 0$
- 24) $\operatorname{sh}x - 12\operatorname{th}x - 0,311 = 0$
- 25) $e^x - 2(x-1)^2 = 0$
- 26) $e^{-x} + x^2 - 2 = 0$
- 27) $x^4 + 4x - 2 = 0$
- 28) $x^4 + 2x - 1 = 0$
- 29) $x^3 - x^2 + x - 3 = 0$
- 30) $x^5 + x - 3 = 0$
- 31) $x^7 + x + 4 = 0$
- 32) $2^x + x^2 - 1,15 = 0$
- 33) $3^{-x} - x^2 + 1 = 0$
- 34) $x^4 - 2x^3 + x^2 - 2x + 1 = 0$
- 35) $x^5 - 5x + 2 = 0$
- 36) $x^7 + 6x - 5 = 0$
- 37) $x^4 + 2x - 2 = 0$
- 38) $(x-1)^2 - \sin 2x = 0$
- 39) $x^4 + 2x^2 - 6x + 2 = 0$
- 40) $x^5 - 3x^2 + 1 = 0$
- 41) $5x^3 + 2x^2 - 15x - 6 = 0$
- 42) $x^6 - 3x^2 + x - 1 = 0$
- 43) $(x-1)^2 - 0,5e^x = 0$
- 44) $3x^4 + 4x^3 - 12x^2 - 5 = 0$
- 45) $x^2 \cos 2x = 1$
- 46) $x^2 - 3 + 0,5^x = 0$
- 47) $x^2 - 10\sin x = 0$

ТЕМА 5. Чисельні методи розв'язування систем лінійних алгебраїчних рівнянь

Чисельні методи розв'язування систем лінійних алгебраїчних рівнянь. Метод Гауса. Прямий та обернений хід методу Гауса. Вибір головного елемента. Оцінка кількості арифметичних дій методу Гауса. Метод ітерацій розв'язування систем лінійних алгебраїчних рівнянь. Достатня умова

збіжності методу ітерацій. Метод Гауса-Зейделя. Перетворення СЛАУ до вигляду, зручному для ітерування.

Практичне заняття 2. Чисельні методи розв'язування систем лінійних алгебраїчних рівнянь

Мета: засвоїти основні методи чисельного розв'язування систем лінійних алгебраїчних рівнянь.

План заняття

1. Приклади математичних моделей, що містять системи лінійних алгебраїчних рівнянь.
2. Розв'язування системи лінійних алгебраїчних рівнянь, обчислення визначника, знаходження оберненої матриці методом Гауса.
3. Розв'язування системи лінійних алгебраїчних рівнянь методом квадратних коренів.
4. Розв'язування системи лінійних алгебраїчних рівнянь методом простої ітерації та ітераційним методом Зейделя.

Теоретичні відомості, що необхідні для виконання даної роботи, містяться в конспекті лекцій за темою 5.

Приклади розв'язання типових задач

Приклад 1. Здійснюючи трикутний розклад методом Гауса з вибором ведучого елемента в стовпчику, знайти розв'язок системи $Ax = b$, де

$$A = \begin{pmatrix} 2 & -4 & -1 \\ -2 & 3 & -2 \\ 4 & -11 & -13 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \\ 10 \end{pmatrix}.$$

Розв'язання. Введемо розширену матрицю системи

$$\bar{A} = \left(\begin{array}{ccc|c} 2 & -4 & -1 & 1 \\ -2 & 3 & -2 & 1 \\ 4 & -11 & -13 & 10 \end{array} \right).$$

Розв'язування системи запишемо у вигляді перетворення матриці \bar{A} множенням на матриці перестановок рядків P_k та на матриці виключення невідомих M_k . Маємо

$$P_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}; P_1 \bar{A} = \left(\begin{array}{ccc|c} 4 & -11 & -13 & 10 \\ -2 & 3 & -2 & 1 \\ 2 & -4 & -1 & 1 \end{array} \right);$$
$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0,5 & 1 & 0 \\ -0,5 & 0 & 1 \end{pmatrix}; \bar{A}_1 = M_1 P_1 \bar{A} = \left(\begin{array}{ccc|c} 4 & -11 & -13 & 10 \\ 0 & -2,5 & -8,5 & 6 \\ 0 & 1,5 & 5,5 & -4 \end{array} \right);$$

$$M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0,6 & 1 \end{pmatrix}; \overline{A_2} = M_2 P_2 \overline{A_1} = \left(\begin{array}{ccc|c} 4 & -11 & -13 & 10 \\ 0 & -2,5 & -8,5 & 6 \\ 0 & 0 & 0,4 & -0,4 \end{array} \right).$$

З останньої трикутної системи дістаємо розв'язок системи

$$x_3 = \frac{-0,4}{0,4} = -1, \quad x_2 = \frac{6 - 8,5}{-2,5} = 1, \quad x_1 = \frac{10 - 13 + 11}{4} = 2.$$

Приклад 2. Обчислити визначник матриці A попередньої задачі.

Розв'язання. Виходячи з трикутного розкладу матриці A , маємо

$$\det A = (-1)^l \cdot u_{11} \cdot u_{22} \cdot u_{33} = -4 \cdot (-2,5) \cdot 0,4 = 4,$$

де l - кількість перестановок рядків при перетворенні матриці A .

Приклад 3. Знайти матрицю обернену до матриці A задачі 1.

Розв'язання. Шукана матриця $C = A^{-1}$ задовольняє матричне рівняння

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

Із правими частинами цієї системи виконаємо ті самі перетворення, що і з матрицею A в задачі 1:

$$P_1 B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad M_1 P_1 B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0,5 \\ 0 & 0 & -0,5 \end{pmatrix},$$

$$B_2 = M_2 P_2 M_1 P_1 B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0,5 \\ 0 & 0,6 & -0,2 \end{pmatrix}.$$

Розв'язуючи матричне рівняння $A_2 C = B_2$, тобто три системи з трикутною матрицею A_2 , знаходимо

$$C = \begin{pmatrix} -15,25 & -10,25 & -2,75 \\ -8,5 & -5,5 & -1,5 \\ 2,5 & 1,5 & -0,5 \end{pmatrix}.$$

Приклад 4. Методом квадратних коренів знайти розв'язок системи $Ax = b$, де

$$A = A^* = \begin{pmatrix} 4 & 2 & 2 & 2 \\ 2 & -3 & -1 & -1 \\ 2 & -1 & -4 & -2 \\ 2 & -1 & -2 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} 8 \\ -6 \\ -9 \\ 3 \end{pmatrix}.$$

Розв'язання. Згідно формул методу знаходимо розклад матриці $A = S^*DS$, де

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

S^* - матриця спряжена (транспонована) до S .

Розв'язуючи систему з нижньою трикутною матрицею

$$S^*Dy = b,$$

знаходимо $y_1=4$, $y_2=5$, $y_3=4$, $y_4=4$.

Розв'язуючи систему з верхньою трикутною матрицею $Sx = y$, знаходимо $x_4=2$, $x_3=1$, $x_2=1$, $x_1=0$.

Приклад 5. Обчислити визначник матриці A попередньої задачі.

Розв'язання. Згідно з розкладом матриці A задачі 4 маємо

$$\det A = \prod_{i=1}^n d_{ii} \cdot \prod_{i=1}^n s_{ii}^2 = 256.$$

Приклад 6. Знайти матрицю обернену до матриці A задачі 4.

Відповідь.

$$A^{-1} = \begin{pmatrix} \frac{45}{256} & \frac{13}{128} & \frac{5}{64} & -\frac{1}{32} \\ \frac{13}{128} & -\frac{19}{64} & \frac{5}{32} & -\frac{1}{16} \\ \frac{5}{64} & \frac{5}{32} & -\frac{3}{15} & -\frac{1}{8} \\ -\frac{1}{32} & -\frac{1}{16} & -\frac{1}{8} & \frac{1}{4} \end{pmatrix}.$$

Приклад 7. Знайти розв'язок системи $Ax = b$, де

$$A = \begin{pmatrix} 3,00 & 1,00 & -1,00 \\ 6,00 & 2,01 & 0,00 \\ 3,00 & 1,02 & 3,01 \end{pmatrix}; \quad b = \begin{pmatrix} -2,00 \\ -2,01 \\ 1,99 \end{pmatrix}$$

методом Гауса, виконуючи обчислення з чотирма значущими цифрами.

Розв'язання. Запишемо послідовність перетворень розширеної матриці A системи методом Гауса з вибором ведучого елемента в стовпчику:

$$P_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad P_1 \bar{A} = \left(\begin{array}{ccc|c} 6,00 & 2,01 & 0,00 & -2,01 \\ 3,00 & 1,00 & -1,00 & -2,00 \\ 3,00 & 1,02 & 3,01 & 1,99 \end{array} \right);$$

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ -0,5 & 1 & 0 \\ -0,5 & 0 & 1 \end{pmatrix}; \bar{A}_1 = M_1 \bar{A} P_1 = \left(\begin{array}{ccc|c} 6 & 2,01 & 0 & -2,01 \\ 0 & -5 \cdot 10^{-3} & -1 & -0,995 \\ 0 & 1,5 \cdot 10^{-2} & 3,01 & 2,995 \end{array} \right);$$

$$P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}; P_2 \bar{A} = \left(\begin{array}{ccc|c} 6 & 2,01 & 0 & -2,01 \\ 0 & 1,5 \cdot 10^{-2} & 3,01 & 2,995 \\ 0 & -5 \cdot 10^{-3} & -1 & -0,995 \end{array} \right);$$

$$M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0,3333 & 1 \end{pmatrix}; \bar{A}_2 = M_2 \bar{A}_1 P_2 = \left(\begin{array}{ccc|c} 6 & 2,01 & 0 & -2,01 \\ 0 & 1,5 \cdot 10^{-2} & 3,01 & 2,995 \\ 0 & 0 & 3 \cdot 10^{-3} & 3,2 \cdot 10^{-3} \end{array} \right).$$

Звідси $x_3=1,067$, $x_2=-14,47$, $x_1=4,511$, що сильно відрізняється від точного розв'язку $x=(0,-1,1)$.

Приклад 8. Розв'язати задачу 7, виконуючи обчислення з п'ятьма значущими цифрами.

Розв'язання. На відміну від попередньої задачі обчислення на останньому кроці дають

$$M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0,33333 & 1 \end{pmatrix}; \bar{A}_2 = \left(\begin{array}{ccc|c} 6 & 2,01 & 0 & -2,01 \\ 0 & 1,5 \cdot 10^{-2} & 3,01 & 2,995 \\ 0 & 0 & 3,3 \cdot 10^{-3} & 3,3 \cdot 10^{-3} \end{array} \right).$$

Звідси $x_3=1$, $x_2=-1$, $x_1=0$, що збігається з точним розв'язком системи.

Приклад 9. Визначити число обумовленості матриці A із задачі 7.

Розв'язання. Використовуючи знайдені в задачі 7 матриці P_k, M_k

$$\text{маємо } A^{-1} = \begin{pmatrix} 2,240 \cdot 10^3 & -1,492 \cdot 10^3 & 7,447 \cdot 10^3 \\ -6,687 \cdot 10^3 & 4,455 \cdot 10^3 & -2,223 \cdot 10^3 \\ 3,33 \cdot 10^2 & -2,2 \cdot 10^2 & 1,111 \cdot 10^2 \end{pmatrix}.$$

Звідси $\|A^{-1}\|_{\infty} = 1,121 \cdot 10^4$, $\|A\|_{\infty} = 8,01$, $\text{cond}_{\infty}(A) = 8,979 \cdot 10^4$. Таким великим числом обумовленості пояснюється відмінність розв'язку задачі 7 від точного:

$$c_k = \frac{(k-1)!!}{k!!} \delta_k, \quad \delta_k = \begin{cases} \frac{\pi}{2}, & k - \text{парне}, \\ 1, & k - \text{непарне} \end{cases}.$$

Приклад 10. Знайти розв'язок системи $Ax = b$, де

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 6 & 2,1 & 0 \\ 3 & 1,15 & 1,9999 \end{pmatrix}, \quad b = \begin{pmatrix} +22,01 \\ +2,15 \cdot 10^{-2} \\ -0,8499 \end{pmatrix}.$$

виконуючи обчислення з чотирма значущими цифрами. Пояснити результат.

Розв'язання. В силу припущень, замість системи, що задана, розв'язується система з наближеною матрицею

$$\tilde{A} = \begin{pmatrix} 3 & 1 & -1 \\ 6 & 2,1 & 0 \\ 3 & 1,15 & 2,000 \end{pmatrix}.$$

Але вона вироджена, оскільки $\det \tilde{A} = 0$. В той час як $\det A = -3 \cdot 10^{-5}$.

Виродженість матриці \tilde{A} обумовлена похибкою заокруглення елемента $a_{33} = 1,9999 \approx 2,000$. Для того щоб одержати задовільний результат, треба збільшити кількість значущих цифр при обчисленнях. Розв'яжіть цей приклад з п'ятьма значущими цифрами.

Точний розв'язок $x = (0,1, -1)$.

Приклад 11. Яку похибку в розв'язку системи лінійних алгебраїчних рівнянь

$$Ax = b, \text{ де } A = \begin{pmatrix} 1 & a & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & a \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

спричинює похибка ε в компоненті b_n вектора b ($b_n = b_n + \varepsilon$)? Які властивості матриці A обумовлюють цю похибку?

Розв'язання. Оскільки розв'язок системи можна знайти за рекурентними формулами $x_n = b_n$, $\tilde{x}_{n-1} = b_{n-1} - ab_n, \dots$, то неважко знайти, що похибки в x_i є:

$$\begin{aligned} \tilde{x}_{n-1} - x_{n-1} &= b_{n-1} - ab_n - b_{n-1} + ab_n = -a\varepsilon, \\ \tilde{x}_{n-2} - x_{n-2} &= a^2\varepsilon, \dots, \tilde{x}_1 - x_1 = (-1)^{n-1} a^{n-1} \varepsilon. \end{aligned}$$

Якщо $a > 1$, то ця похибка швидко зростає з ростом n . Матриця A при цьому є погано обумовленою, що і спричинює велику похибку в розв'язку.

Приклад 12. Знайти оцінку швидкості збіжності методу простої ітерації

$$y_{k+1} = By_k + b$$

для системи лінійних алгебраїчних рівнянь

$$x = Bx + b,$$

де B - матриця розмірності n , причому відомо, що для деякої норми $\|B\| = \rho < 1$.

Розв'язання. Якщо x^* - розв'язок системи, то $x^* = Bx^* + b$. Тому маємо $x^* - y_k = B(x^* - y_{k-1})$. Звідси

$$\|x^* - y_k\| \leq \rho \|x^* - y_{k-1}\| \leq \dots \leq \rho^k \|x^* - y_0\|.$$

Оскільки $x^* - y_{k-1} = y_k - y_{k-1} + B(x^* - y_{k-1})$, то

$$\|x^* - y_{k-1}\| \leq \|y_k - y_{k-1}\| + \|B\| \|x^* - y_{k-1}\|,$$

$$\|x^* - y_{k-1}\| \leq \frac{1}{1-\rho} \|y_k - y_{k-1}\|.$$

Остаточно маємо

$$\|x^* - y_k\| \leq \frac{\rho^k}{1-\rho} \|y_1 - y_0\|.$$

Оскільки $\rho < 1$, то для $k \rightarrow \infty$ $\|x^* - y_k\| \rightarrow 0$, що означає збіжність методу.

Приклад 13. Дослідити збіжність ітераційного процесу

$$y_{k+1} = By_k + b$$

з матрицями

$$\text{а) } B = \begin{pmatrix} -3/5 & 3/5 \\ 2/5 & 1/5 \end{pmatrix}; \quad \text{б) } B = \begin{pmatrix} 1/10 & 4/5 \\ 3/5 & -1/5 \end{pmatrix}.$$

Розв'язання. а) $\|B\|_\infty = 6/5 > 1$, $\|B\|_E = \left(\sum_{i,j} b_{ij}^2 \right)^{1/2} = \sqrt{23}/5 < 1$ що означає

збіжність ітераційного процесу в нормі

$$\|x\| = \left(\sum_{i=1}^n x_i^2 \right)^{1/2},$$

яка узгоджена з матричною нормою Шура $\|B\|_E$.

б) $\|B\|_\infty = 9/11 < 1$, що означає збіжність ітераційного процесу в нормі $\|x\|_\infty$. В той же час для цієї матриці $\|B\|_E = \sqrt{21/20} > 1$.

Матрична норма Шура називається також нормою Ерхарда-Шмідта або нормою Фробеніуса.

Приклад 14. Нехай матриця системи лінійних алгебраїчних рівнянь

$$Ax = b$$

має діагональну перевагу, тобто

$$\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \leq q |a_{ii}|, \quad i = \overline{1, n}, \quad 0 < q < 1,$$

а також є симетричною і додатно визначеною. Дати оцінку швидкості збіжності методу Зейделя.

Розв'язання. Для похибки $z_k = y_k - x = \left(z_k^j \right)_{j=1}^n$ маємо

$$a_{ii} z_{k+1}^{(i)} = - \sum_{j < i} a_{ij} z_{k+1}^{(j)} - \sum_{j > i} a_{ij} z_k^{(j)},$$

$$|a_{ii}| \|z_{k+1}^{(i)}\| = \sum_{j<i} |a_{ij}| \|z_{k+1}^{(j)}\| + \sum_{j>i} |a_{ij}| \|z_k^{(j)}\|.$$

Нехай

$$\max_i |z_k^{(i)}| = |z_k^{(i_0)}| = \|z_k\|_C.$$

Тоді

$$|a_{i_0 i_0}| \|z_{k+1}\|_C \leq \sum_{j<i_0} |a_{i_0 j}| \|z_{k+1}\|_C + \sum_{j>i_0} |a_{i_0 j}| \|z_k\|_C,$$

$$\|z_{k+1}\|_C \leq \left[\sum_{j<i_0} |a_{i_0 j}| / |a_{i_0 i_0}| - \sum_{j>i_0} |a_{i_0 j}| \right] \|z_k\|_C.$$

Користуючись умовою діагональної переваги, маємо

$$\sum_{j>i_0} |a_{i_0 j}| \leq q |a_{i_0 i_0}| - \sum_{j<i_0} |a_{i_0 j}| < q \left(|a_{i_0 i_0}| - \sum_{j<i_0} |a_{i_0 j}| \right)$$

тобто

$$\|z_{k+1}\|_C \leq q \|z_k\|_C \leq \dots \leq q^{k+1} \|z_0\|_C.$$

Приклад 15. Дослідити збіжність методу простої ітерації та методу Зейделя для системи

$$Ax = b,$$

$$\text{де } A = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, |\rho| < 1.$$

Розв'язання. Для застосування методу простої ітерації зведемо систему до вигляду

$$x = Bx + b,$$

де

$$B = I - A = \begin{pmatrix} 0 & \rho \\ \rho & 0 \end{pmatrix}.$$

Власні значення матриці B визначаються з рівняння

$$|B - \lambda I| = \begin{vmatrix} -\lambda & \rho \\ \rho & -\lambda \end{vmatrix} = \lambda^2 - \rho^2 = 0.$$

Звідси $\lambda = \pm \rho, |\lambda| = \rho < 1$. Останнє означає збіжність методу.

Із критерію Сільвестра маємо $A^* = A > 0$, бо $\det(a_{ii}) = 1 > 0$, $\det A = 1 - \rho^2 > 0$. Тому метод Зейделя також збіжний.

Задачі для самоконтролю

Задача 1. Нехай векторна норма $\|x\|_p$ визначається формулою

$$\|x\|_p = \left\{ \sum_{k=1}^n |x_k|^p \right\}^{1/p}, \quad p \geq 1, \quad \|x\|_\infty = \max_{1 \leq k \leq n} |x_k|.$$

Довести, що для узгоджених норм матриці $A = (a_{ik})_{i,k=1}^n$

$$\|A\|_p = \sup_{\|x\|_p \neq 0} \frac{\|Ax\|_p}{\|x\|_p}$$

мають місце такі формули:

$$\text{а) } \|A\|_1 = \max_k \sum_{i=1}^n |a_{ik}|;$$

$$\text{б) } \|A\|_\infty = \max_i \sum_{k=1}^n |a_{ik}|;$$

$$\text{в) } \|A\|_2 = \sqrt{\max \lambda(A^*A)}, \text{ де } \lambda(A^*A) \text{ - власне значення матриці } A^*A.$$

Задача 2. Довести такі нерівності:

$$\text{а) } \frac{1}{\sqrt{n}} \|A\|_2 \leq \|A\|_1 \leq \sqrt{n} \|A\|_2;$$

$$\text{б) } \frac{1}{\sqrt{n}} \|A\|_2 \leq \|A\|_\infty \leq \sqrt{n} \|A\|_2;$$

$$\text{в) } n^{-1} \leq \frac{\text{cond}_\infty A}{\text{cond}_2 A} \leq n;$$

$$\text{г) } n^{-1} \leq \frac{\text{cond}_1 A}{\text{cond}_2 A} \leq n,$$

де $\text{cond}_k A = \|A\|_k \|A^{-1}\|_k$ – число обумовленості матриці A .

Задача 3. Довести, що матриця обернена до

$$A = \begin{pmatrix} 1 & a & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & a \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

має вигляд

$$A^{-1} = \begin{pmatrix} 1 & -a & a^2 & -a^3 & \dots & (-a)^{n-1} \\ 0 & 0 & -a & a^2 & \dots & (-a)^{n-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & -a \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

і матриця A має число обумовленості

$$\text{cond}_\infty A = (1+a) \frac{a^{n-1} - 1}{a - 1}.$$

Підрахувати число обумовленості для $n=20$, $a=5$.

Задача 4. Дати відповіді на запитання:

а) Як зв'язані між собою число обумовленості матриці та її детермінант?

б) Чи обов'язково погано обумовлена матриця має малі власні значення?

Задача 5. Довести, що детермінант матриці Коші

$$K_n = \left(\frac{1}{a_i + b_j} \right)_{i,j=1}^n$$

дорівнює

$$\begin{aligned} \det K_n = D_n &= \prod_{p=1}^{n-1} \left[\prod_{q=p+1}^n (a_p - a_q) \right] \cdot \prod_{r=1}^{n-1} \left[\prod_{s=r+1}^n (b_r - b_s) \right] / \prod_{i=1}^n \left[\prod_{j=1}^n (a_i + b_j) \right] = \\ &= \prod_{1 \leq i < j \leq n} (a_i - a_j)(b_i - b_j) \left(\prod_{1 \leq i < j \leq n} (a_i + b_j) \right)^{-1}. \end{aligned}$$

Задача 6. Знайти елементи матриці $K_n^{-1} = (b_{ij})_{i,j=1}^n$, оберненої до матриці Коші $K_n = \left(\frac{1}{a_i + b_j} \right)_{i,j=1}^n$.

Задача 7. Знайти елементи матриці $H_n^{-1} = (a_{ij})_{i,j=1}^n$, оберненої до матриці Гільберта

$$H_n = \left(\frac{1}{j+i-1} \right)_{i,j=1}^n$$

і показати, що a_{ij} – цілі числа.

Задача 8. Довести, що (див. задачу 7)

$$\max |a_{ij}| = \frac{1}{4\sqrt{2}\pi^2 n} (\sqrt{2} + 1)^{4n} (1 + O(n^{-1})),$$

$$\|H_n^{-1}\|_\infty = \frac{1}{(2\pi)^{3/2} 2^{7/4} n^{1/2}} (\sqrt{2} + 1)^{4n} (1 + O(n^{-1}))$$

і дати оцінку числа обумовленості матриці Гільберта H_n .

Задача 9. Привести приклади систем $Ax=b$, для яких розбігається метод простої ітерації, але збігається метод Зейделя і навпаки.

Задача 10. Методом Зейделя розв'язати СЛАР, привівши її до вигляду, зручного для ітерацій

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

Задача 11. Використовуючи схему Гауса, розв'язати систему рівнянь з точністю до 0,0001:

$$\begin{cases} 4,21x_1 - 3,25x_2 + 1,13x_3 = 3,06; \\ 5,09x_1 + 2,17x_2 - 4,23x_3 = 2,75; \\ 1,43x_1 - 0,4x_2 - 0,82x_3 = -1,05. \end{cases}$$

Задача 12. Методом ітерацій розв'язати СЛАР з точністю до 0,001:

$$\begin{cases} x_1 = 0,32x_1 - 0,05x_2 + 0,31x_3 - 0,18x_4 + 1,24; \\ x_2 = 0,47x_1 - 0,23x_2 + 0,06x_3 - 0,9; \\ x_3 = 0,28x_1 + 0,44x_2 - 0,21x_3 + 1,42; \\ x_4 = 0,08x_1 - 0,16x_2 + 0,54x_3 - 0,32x_4 - 2,17. \end{cases}$$

ТЕМА 6. Інтерполювання функцій

Інтерполювання функцій однієї змінної. Многочлен Лагранжа. Інтерполювання за рівномірною сіткою вузлів. Многочлен Ньютона. Поняття скінчених різниць. Побудова многочлену Ньютона за таблицею скінчених різниць.

Практичне заняття 3. Інтерполювання функцій

Мета: засвоїти основні методи інтерполювання функцій однієї змінної.

План заняття

1. Приклади задач, які використовують інтерполювання функцій однієї змінної.
2. Побудова многочленів Лагранжа. Знаходження проміжних значень функції однієї змінної за допомогою многочлену Лагранжа.
3. Інтерполяційні формули Ньютона (інтерполяція вперед і назад). Складання таблиць скінчених різниць.

Теоретичні відомості, що необхідні для виконання даної роботи, містяться в конспекті лекцій за темою 6.

Приклади розв'язання типових задач

Приклад 1. За допомогою формули Лагранжа побудувати інтерполяційний многочлен для функції, що задана наступною таблицею

x_i	-2	-1	0	1	2
f_i	3	-1	-1	3	-13

Таб. 1

Обчислити значення в точці $x=0,5$.

Розв'язання. За формулою Лагранжа

$$L_4(x) = 3 \frac{(x+1)x(x-1)(x-2)}{(-1)(-2)(-3)(-4)} - 1 \frac{(x+2)x(x-1)(x-2)}{1(-1)(-2)(-3)} -$$

$$- 1 \frac{(x+2)(x+1)(x-1)(x-2)}{2 \cdot 1(-1)(-2)} + 3 \frac{(x+2)(x+1)x(x-2)}{3 \cdot 2 \cdot 1(-1)} -$$

$$- 13 \frac{(x+2)(x+1)x(x-1)}{4 \cdot 3 \cdot 2 \cdot 1}.$$

Далі

$$L_4(0,5) = 3 \frac{1,5 \cdot 0,5 \cdot (-0,5)(-1,5)}{(-1)(-2)(-3)(-4)} - 1 \frac{2,5 \cdot 0,5(-0,5)(-1,5)}{1(-1)(-2)(-3)} -$$

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

$$- 13 \frac{2,5 \cdot 1,5 \cdot 0,5(-0,5)}{4 \cdot 3 \cdot 2 \cdot 1} = \frac{23}{16} = 1,438.$$

Приклад 2. Побудувати інтерполяційний многочлен за формулою Ньютона для функції, що задана таблицею 1, та обчислити значення в точці $x=0,5$.

Розв'язання. Таблиця розділених різниць для цих даних має вигляд

x_i	f_i				
-2	3				
-1	-1	-4			
0	-1	0	2	0	-1
1	3	4	-10	-4	
2	-13	-16			

Тоді за формулою Ньютона

$$L_4(x) = 3 - 4(x+2) + 2(x+2) + 0(x+2)(x+1)x - 1(x+2)(x+1)x(x-1).$$

Далі

$$L_4(0,5) = 3 - 4 \cdot 2,5 + 2 \cdot 2,5 \cdot 1,5 - 2,5 \cdot 1,5 \cdot 0,5(-0,5) = \frac{23}{16} = 1,438.$$

Приклад 3. За допомогою інтерполяції обчислити $e^{0,15}$, якщо

x_i	0.	0.1	0.2
e^{x_i}	1	1,10517	1,22140

Оцінити похибку.

Розв'язання. Позначимо $f(x) = e^x$. Тоді таблиця розділених різниць для даних в таблиці має вигляд

x_i	f_i		
0	1,0		
0,1	1,105187	1,0517	0,553
0,2	1,22140	1,1623	

Тому

$$L_2(x) = 1 + 1,0517x + 0,553x(x - 0,1)$$

та

$$L_2(0,15) = 1 + 0,10517 + 0,00415 = 1,16190.$$

Оцінка для похибки буде мати вигляд

$$|R_2(x)| \leq \frac{M_3}{3!} |x(x - 0,1)(x - 0,2)|,$$

де $M_3 = \max_{0 \leq x \leq 0,2} |f^{(3)}(x)| = e^{0,2}$. Тому

$$|R_2(0,15)| \leq \frac{1,2214}{6} 0,15 \cdot 0,05 \cdot 0,05 = 7,63 \cdot 10^{-5}.$$

Приклад 4. За значеннями $\cos x$ в точках $x = 0, \pi/4, \pi/2$ знайти $\cos(\pi/5)$. Оцінити похибку.

Розв'язання. Таблиця розділених рівнянь має вигляд

x_i	f_i		
0	1,0		
$\pi/4$	$\sqrt{2}/2$	$(\sqrt{2}/2 - 1)\pi/4$	$(1 - \sqrt{2})/(\pi/4)/(\pi/2)$
$\pi/2$	0	$-\sqrt{2}/2/(\pi/4)$	

Приклад 5. а) Побудувати на проміжку $[-1, 1]$ многочлени Чебишева четвертого та п'ятого степеня з коефіцієнтом 1 при старшому степені. Обчислити їх відхилення від 0.

Розв'язання. За рекурентною формулою

$$T_{n+1}(x) = 2xT_n - T_{n-1}(x), \quad T_0(x) = 1, \quad T_1(x) = x$$

знаходимо

$$T_2(x) = 2x^2 - 1, \quad T_3(x) = 4x^3 - 3x, \quad T_4(x) = 8x^4 - 8x^2 + 1,$$

$$T_5(x) = 16x^5 - 20x^3 + 5x.$$

Шукані многочлени Чебишева мають вигляд

$$\bar{T}_4(x) = x^4 - x^2 + \frac{1}{8}, \quad T_5(x) = x^5 - \frac{5}{4}x^3 + \frac{5}{16}x.$$

За формулою відхилення від нуля такі

$$\|\bar{T}_4(x)\| = \frac{1}{8}, \quad \|T_5(x)\| = \frac{1}{16}.$$

б) Побудувати на проміжку $[0,1]$ многочлен Чебишева четвертого степеня з коефіцієнтом 1 при старшому степені. Обчислити її відхилення від 0.

Розв'язання. У многочлені Чебишева

$$T_4(\xi) = 8\xi^4 - 8\xi^2 + 1, \quad -1 \leq \xi \leq 1$$

(див. попередню задачу) зробимо заміну змінної

$$\xi = \frac{2}{b-a} \left(x - \frac{b+a}{2} \right) = 2x - 1, \quad a \leq x \leq b.$$

При $a=0, b=1$ маємо $\xi=2x-1$. Тоді многочлен Чебишева для проміжку $[0,1]$ має вигляд

$$T_4^{[0,1]}(x) = 128x^4 - 256x^3 + 160x^2 - 32x + 1.$$

Шуканий многочлен такий

$$\bar{T}_4^{[0,1]}(x) = x^4 - 2x^3 + \frac{5}{4}x^2 - \frac{1}{4}x + \frac{1}{128}.$$

Його відхилення від 0 дорівнює

$$\|\bar{T}_4^{[0,1]}(x)\| = \frac{1}{128}.$$

Приклад 6. Оцінити похибку інтерполяції функції e^x на $[0,1]$ многочленом четвертого степеня побудованим за вузлами.

Розв'язання. З оцінки залишкового члена інтерполяційної формули маємо

$$|R_4(x)| \leq \frac{M_5}{5!} |\omega_4(x)| = \frac{M_5}{5!} |T_5^{[0,1]}(x)| \leq \frac{e}{120} \frac{1}{2^9} = 4,4 \cdot 10^{-5}.$$

Приклад 7. Скільки чебишевських вузлів інтерполяції необхідно вибрати, щоб похибка інтерполяції для функції $f(x) = \ln(1+x)$, $x \in [0,1]$ не перевищувала $\varepsilon = 10^{-4}$.

Розв'язування. З оцінки залишкового члена інтерполяційної формули маємо

$$|R_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |\omega_n(x)| = \frac{M_{n+1}}{(n+1)!} |T_{n+1}^{[0,1]}(x)| \leq \frac{M_{n+1}}{(n+1)!} \frac{1}{2^{2n+1}}.$$

Далі

$$f^{(n+1)}(x) = (-1)^{n+1} \frac{n!}{(1+x)^{n+1}}.$$

Тому n шукаємо з умови

$$|R_n(x)| \leq \frac{1}{(n+1)2^{2n+1}} < 10^{-4}.$$

Підстановкою знаходимо, що

$$|R_5(x)| \leq \frac{1}{6 \cdot 2^{11}} \approx 8 \cdot 10^{-5} < 10^{-4}.$$

Таким чином треба вибрати 6 чебишевських вузлів.

Задачі для самоконтролю

1. Функція задана таблично:

x_i	-1	1	2	3	5
y_i	2	-2	-4	-2	2

Інтерполюванням обчислити значення функції в точці $x = 0$ та дати оцінку похибки, якщо відомо $M_5 = 0.1$.

2. Функція $f(x) = x \exp(-x)$ задана значеннями

$$f(0.25) = 0.1947, f(0.5) = 0.3033$$

Обчислити за допомогою прямої інтерполяції цих значень корінь рівняння $f(x) = 0.25$. Оцінити похибку.

3. Функція задана таблично:

x_i	-1	1	2	3	5
y_i	2	2	4	2	-2

Інтерполюванням обчислити значення функції в точці $x = 0$ та дати оцінку похибки, якщо відомо $M_5 = 0.05$.

4. Функція $f(x) = x \exp(-x)$ задана значеннями

$$f(0.25) = 0.1947, f(0.5) = 0.3033$$

Обчислити за допомогою оберненої інтерполяції цих значень корінь рівняння $f(x) = 0.3$. Оцінити похибку.

5. Вибрати крок таблиці для функції $f(x) = \arctg x$, $-1 \leq x \leq 1$ так, щоб похибка лінійного інтерполювання по двох сусідніх точках не перевищувала 10^{-4} .

6. Функція $f(x) = x + e^x$ задана значеннями

$$f(0) = 1, f(0.5) = 2.1487.$$

Обчислити за допомогою прямої інтерполяції цих значень корінь рівняння $f(x) = 2$. Оцінити похибку.

7. Функція $f(x) = e^x + x$ задана значеннями

$$f(0) = 1, f(0.5) = 2.1487$$

Обчислити за допомогою оберненого інтерполювання цих значень корінь рівняння $f(x) = 2$. Оцінити похибку.

8. Вибрати крок чотирьохзначної таблиці для функції $f(x) = e^{-x^2}$, $x \in [0,1]$ так, щоб похибка лінійного інтерполювання не перевищувала 10^{-4} .

9. Побудувати багаточлен, що інтерполіює дані

x	-2	-1	0	1	3	4
y	112	6	0	4	162	256

Обчислити значення функції в точці $x = 2$.

10. Вибрати крок таблиці для функції $f(x) = \ln(1+x)$, $x \in [0,1]$ так, щоб похибка квадратичного інтерполювання не перевищувала 10^{-4} .

11. Побудувати багаточлен, що інтерполіює дані

x	-2	0	1	2	3	5
y	-1	3	5	1	3	7

Обчислити значення для $x = -1$. Оцінити похибку, якщо $M_6 = 3$.

12. Функція $f(x) = x + e^x$ задана значеннями

$$f(0.5) = 2.1487, f(1) = 3.8871.$$

Обчислити за допомогою оберненої інтерполяції цих значень корінь рівняння $f(x) = 3$. Оцінити похибку.

13. Побудувати багаточлен, що інтерполіює дані

$$x_0 = -1: y_0 = 2; y'_0 = -10; y''_0 = 32;$$

$$x_1 = 0: y_1 = -1; y'_1 = -1;$$

$$x_2 = 2: y_2 = -19.$$

14. Вибрати крок таблиці для функції $f(x) = \arctg x$, $-1 \leq x \leq 1$ так, щоб похибка лінійного інтерполювання по двох сусідніх точках не перевищувала 10^{-4} .

15. Побудувати багаточлен, що інтерполіює дані

X	-1	0	2	3	5
Y	0	2	4	1	-3

Обчислити значення для $x = 1$. Оцінити похибку, якщо $M_5 = 3$.

16. Вибрати крок таблиці для функції $f(x) = \operatorname{tg} x$, $-1 \leq x \leq 1$ так, щоб похибка лінійного інтерполювання не перевищувала 10^{-4} .

17. Побудувати багаточлен, що інтерполіює дані

$$x_0 = -1: y_0 = 2; y'_0 = -2; y''_0 = 4;$$

$$x_1 = 0: y_1 = -1; y'_1 = -1;$$

18. Вибрати крок таблиці для функції $f(x) = \arccos x$, $-1 \leq x \leq 1$ так, щоб похибка лінійного інтерполювання не перевищувала 10^{-4} .

19. Побудувати багаточлен, що інтерполіює дані

x	0	0.1	0.2	0.3	0.4
y	-0.22	0.02	0.14	0.21	0.35

Обчислити значення для $x = 0.15$. Оцінити похибку, якщо $M_5 = 1$.

20. Дана таблиця синусів для $0 \leq \alpha \leq 45^\circ$ з кроком 1° . Якою є похибка лінійної інтерполяції між двома сусідніми точками?

21. Побудувати багаточлен, що інтерполіює дані

x	0	0.1	0.2	0.3	0.4
y	-0.22	0.02	0.14	0.21	0.35

Обчислити значення для $x = 0.15$. Оцінити похибку, якщо $M_5 = 1$.

22. Вибрати крок таблиці для функції $f(x) = chx$, $-1 \leq x \leq 1$ так, щоб похибка лінійного інтерполювання не перевищувала 10^{-4} .

23. Побудувати багаточлен, що інтерполює дані

x	0	0.2	0.4	0.6	0.8
y	0.32	0.12	0.14	0.27	0.35

Обчислити значення для $x = 0.1$. Оцінити похибку, якщо $M_5 = 1$.

24. Яка похибка інтерполювання функції $f(x) = e^x + x$, $0 \leq x \leq 1$ поліномом 4-го степеня по чебишовських вузлах?

25. Використовуючи таблицю значень функції $y = \sin x$, знайти $\sin 37^\circ$ та похибку результату.

x	30°	35°	40°	45°
y	0,5000	0,5736	0,6428	0,7071

26. Функція $f(x) = \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$, $x \in [0,1]$ задана значеннями

$$f(0.4) = 0.4284, f(0.5) = 0.5204, f(0.6) = 0.6039.$$

Обчислити за допомогою інтерполювання цих значень корінь рівняння $f(x) = 0.5$. Оцінити похибку.

ТЕМА 7. Задача наближення функцій однієї змінної

Найкраще наближення функцій за допомогою многочленів. Метод найменших квадратів. Формули наближення функцій за допомогою лінійних та квадратичних многочленів.

Практичне заняття 4. Задача наближення функцій однієї змінної

Мета: засвоїти поширені методи побудови наближень для функцій однієї змінної.

План заняття

1. Побудова лінійних та квадратичних наближення для функцій однієї змінної за методом найменших квадратів.

Теоретичні відомості, що необхідні для виконання даної роботи, містяться в конспекті лекцій за темою 7.

Метод найменших квадратів відноситься до класу методів апроксимацій. Ідея методу полягає в тому, щоб за даними експерименту побудувати приблизно функцію у вигляді многочлена з тим розрахунком, щоб сума квадратів відхилень побудованої функції від експериментальної у вузлових точках була мінімальна. Будуватимемо функцію у вигляді многочлена

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

Використовуємо для побудови результату експерименту, задані в таблиці

x_i	x_0	x_1	\dots	x_m
y_i	y_0	y_1	\dots	y_m

Побудувати многочлен, значить, визначити його коефіцієнти $a_i, i = \overline{0, n}$.

Для цього введемо функцію $S = \sum_{i=0}^m \delta_i^2$ і будемо вимогати, щоб

$\min S = \min \sum_{i=0}^m \delta_i^2$, де $\delta_i = f(x_i) - y_i$ - відхилення функції від

експериментальних даних у вузлах $x_i, i = \overline{0, m}$.

Використовуючи вигляд $f(x)$, маємо:

$$\min_{a_0, a_1, \dots, a_n} S = \min_{a_0, \dots, a_n} \left\{ \sum_{i=0}^m (a_n x_i^n + a_{n-1} x_i^{n-1} + \dots + a_0 - y_i)^2 \right\}.$$

Необхідними умовами екстремуму функції S є рівність нулю її першої похідній по всіх змінних $\frac{\partial S}{\partial a_0} = 0; \frac{\partial S}{\partial a_1} = 0; \dots \frac{\partial S}{\partial a_n} = 0$. Розписавши ці умови,

отримаємо СЛАР вигляду:

$$\begin{cases} \frac{\partial S}{\partial a_0} = 2 \sum_{i=0}^m (a_n x_i^n + a_{n-1} x_i^{n-1} + \dots + a_0 - y_i) = 0 \\ \frac{\partial S}{\partial a_1} = 2 \sum_{i=0}^m (a_n x_i^n + a_{n-1} x_i^{n-1} + \dots + a_0 - y_i) x_i = 0 \\ \dots \\ \frac{\partial S}{\partial a_n} = 2 \sum_{i=0}^m (a_n x_i^n + a_{n-1} x_i^{n-1} + \dots + a_0 - y_i) x_i^n = 0 \end{cases}$$

Запишемо систему для визначення a_0, \dots, a_n у нормальній формі:

$$\begin{cases} m a_0 + \sum_{i=0}^m x_i a_1 + \sum_{i=0}^m x_i^2 a_2 + \dots + \sum_{i=0}^m x_i^n a_n = \sum_{i=0}^m y_i \\ \sum_{i=0}^m x_i a_0 + \sum_{i=0}^m x_i^2 a_1 + \dots + \sum_{i=0}^m x_i^{n+1} a_n = \sum_{i=0}^m x_i y_i \\ \dots \\ \sum_{i=0}^m x_i^n a_0 + \sum_{i=0}^m x_i^{n+1} a_1 + \dots + \sum_{i=0}^m x_i^{2n} a_n = \sum_{i=0}^m x_i^n y_i \end{cases}$$

Вирішуємо систему одним з відомих методів і знаходимо a_0, \dots, a_n , які потім підставляємо в шуканий многочлен.

Запишемо алгоритм методу найменших квадратів.

1. Вводимо таблицю чисел $x_i, y_i, i = \overline{0, m}$.

2. Обчислюємо $\sum_{i=0}^m x_i, \sum_{i=0}^m x_i^2, \sum_{i=0}^m x_i^3, \dots, \sum_{i=0}^m x_i^{2n}, \sum_{i=0}^m y_i, \sum_{i=0}^m x_i y_i, \dots, \sum_{i=0}^m x_i^n y_i$.

3. Вирішуючи будь-яким відомим методом отриману систему лінійних рівнянь алгебри, знаходимо a_0, \dots, a_n - коефіцієнти шуканого багаточлена.

Приклад розв'язання типових задач

x_i	$\frac{\pi}{6} = 0.524$	$\frac{\pi}{4} = 0.785$	$\frac{\pi}{2} = 1.571$
y_i	0.5	0.707	1.0

Таб.1

Приклад 1. Для таблиці 1 вузлових точок, приведених вище, побудувати апроксимаційний багаточлен $P_2(x) = a_2 x^2 + a_1 x + a_0$ другого порядку за методом найменших квадратів.

Розв'язання. Для цього необхідно обчислити наступні суми:

$$\sum_{i=0}^2 x_i = 2.88; \quad \sum_{i=0}^2 x_i^2 = 3.358; \quad \sum_{i=0}^2 y_i = 2.207; \quad \sum_{i=0}^2 x_i y_i = 2.388$$

$$\sum_{i=0}^2 x_i^3 = 4.504; \quad \sum_{i=0}^2 x_i^4 = 6.544; \quad \sum_{i=0}^2 x_i^2 y_i = 3.041$$

і вирішити СЛАР щодо невідомих коефіцієнтів a_0, a_1, a_2 :

$$\begin{cases} 4a_0 + 2.88a_1 + 3.358a_2 = 2.207; \\ 2.88a_0 + 3.358a_1 + 4.504a_2 = 2.388; \\ 3.358a_0 + 4.504a_1 + 6.544a_2 = 3.041. \end{cases}$$

Значення невідомих коефіцієнтів рівні:

$$a_0 = -0.004; \quad a_1 = 1.146; \quad a_2 = -0.323.$$

Тоді шуканий багаточлен другого порядку матиме вигляд:

$$P_2(x) = -0.323x^2 + 1.146x - 0.004.$$

Неважко помітити, що у вузлових точках значення многочлена і табличної функції не збігаються. Похибка обчислень за даною формулою в контрольній точці, в порівнянні з дійсним значенням, складає

$$\left| \sin\left(\frac{\pi}{3}\right) - P_2\left(\frac{\pi}{3}\right) \right| = |0.866 - 0.844| = 0.022$$

Задачі для самоконтролю

Задача. Для вузлових точок $x_i = i \times 0,1; \quad i=1,2,\dots,20$, побудувати багаточлен найкращого середньоквадратичного наближення другого порядку за методом найменших квадратів (30 варіантів експериментальних даних наведено в таб. 2)

Таб. 2

	Значення $Y_i=y(X_i)$					
	Варіант 1	Варіант 2	Варіант 3	Варіант 4	Варіант 5	Варіант 6
1	2	3	4	5	6	7
1	2,05	2,09	2,02	1,99	2,23	2,07
2	1,94	2,05	1,98	2,03	2,29	2,17
3	1,92	2,19	1,67	2,20	2,27	2,21
4	1,87	2,18	1,65	2,39	2,62	2,31
5	1,77	2,17	1,57	2,19	2,72	2,10
6	1,88	2,27	1,42	2,61	2,82	2,09
7	1,71	2,58	1,37	2,35	3,13	2,12
8	1,60	2,73	1,07	2,60	3,49	1,63
9	1,56	2,82	0,85	2,55	3,82	1,78
10	1,40	3,04	0,48	2,49	3,95	1,52
11	1,50	3,03	0,35	2,50	4,22	1,16
12	1,26	3,45	-0,30	2,52	4,48	1,07
13	0,99	3,62	-0,61	2,44	5,06	0,85
14	0,97	3,85	-1,20	2,35	5,50	0,56
15	0,91	4,19	-1,39	2,26	5,68	0,10
16	0,71	4,45	-1,76	2,19	6,19	-0,25
17	0,43	4,89	-2,28	2,24	6,42	-0,65
18	0,54	5,06	-2,81	2,34	7,04	-1,06
19	0,19	5,63	-3,57	1,96	7,57	-1,66
20	0,01	5,91	-4,06	2,19	8,10	-2,01

	Значення $Y_i=y(X_i)$					
	Варіант 7	Варіант 8	Варіант 9	Варіант 10	Варіант 11	Варіант 12
1	2,18	-0,10	-0,16	2,09	2,15	0,10
2	2,43	-0,21	0,01	2,31	2,41	-0,01
3	2,40	0,01	0,10	2,72	2,58	-0,19
4	2,43	0,05	0,16	2,77	2,84	-0,11
5	2,65	-0,13	0,05	2,78	3,28	-0,31
6	2,75	-0,23	0,35	2,97	3,46	-0,78
7	2,67	-0,21	0,19	3,00	4,02	-0,64
8	2,66	-0,43	0,50	3,51	4,11	-0,85
9	2,63	-0,57	0,74	3,43	4,61	-1,18
10	2,75	-0,44	1,03	3,58	5,03	-1,39
11	2,41	-0,44	1,06	3,58	5,34	-1,79
12	2,24	-0,83	1,49	3,54	5,86	-2,02
13	2,12	-0,78	1,79	3,82	6,33	-2,48
14	1,74	-0,81	2,03	3,90	6,81	-2,93
15	1,57	-1,06	2,22	3,77	7,21	-3,26
16	1,17	-1,41	2,50	3,81	7,67	-3,91
17	0,96	-1,40	2,88	4,00	8,23	-4,41
18	0,63	-1,70	3,21	3,97	8,68	-4,91

19	0,25	-1,96	3,63	4,08	9,35	-5,30
20	-0,01	-1,91	3,90	4,08	9,93	-6,00
	Значення $Y_i=y(X_i)$					
	Варіант 13	Варіант 14	Варіант 15	Варіант 16	Варіант 17	Варіант 18
1	0,17	0,80	0,04	0,08	-0,02	0,14
2	0,07	0,29	0,47	0,14	0,44	0,23
3	0,17	0,52	0,78	0,37	0,51	0,44
4	0,05	0,77	1,01	0,36	0,67	0,54
5	0,12	0,93	1,19	0,44	0,69	0,72
6	0,00	1,20	1,60	0,48	1,04	3,76
7	0,01	1,20	1,93	0,27	1,1	0,37
8	-0,05	1,35	2,22	0,39	1,3	0,64
9	-0,21	1,39	2,50	0,50	1,7	357
10	-0,50	1,48	3,01	0,48	2,0	3,44
11	-0,50	1,52	3,22	0,69	2,1	3,41
12	-0,86	1,71	3,71	0,50	2,4	0,30
13	-1,24	1,72	4,23	0,31	2,90	-0,0
14	-1,47	1,87	4,78	0,37	3,50	-0,03
15	-1,79	1,86	5,27	0,43	3,99	-0,47
16	-2,25	1,89	5,75	0,33	4,06	-0,68
17	-2,55	2,04	6,16	0,31	4,54	-0,93
18	-3,18	1,73	6,76	0,09	4,99	-1,28
19	-3,60	2,04	7,30	0,08	5,36	-1,53
20	-3,93	2,03	8,00	0,03	5,99	-1,93
	Значення $Y_i=y(X_i)$					
	Варіант 19	Варіант 20	Варіант 21	Варіант 22	Варіант 23	Варіант 24
1	-1,86	-1,65	-1,89	-1,84	-1,92	-1,90
2	-1,95	-2,00	-2,07	-1,98	-1,60	-1,80
3	-2,12	-1,87	-2,30	-1,72	-1,57	-1,82
4	-2,06	-1,89	-2,26	-1,58	-1,41	-1,86
5	-2,15	-1,75	-2,34	-1,69	-1,36	-1,83
6	-2,00	-1,59	-2,66	-1,59	-0,97	-2,00
7	-2,12	-1,44	-2,88	-1,58	-0,59	-2,01
8	-2,31	-1,61	-2,85	-1,64	-0,71	-2,05
9	-2,29	-1,00	-3,16	-1,55	-0,15	-2,46
10	-2,57	-1,17	-3,49	-1,35	0,01	-2,68
11	-2,56	-0,87	-3,88	-1,33	0,22	-2,85
12	-2,86	-0,47	-4,22	-1,47	0,63	-2,98
13	-2,85	-0,33	-4,45	-1,50	1,07	3,30
14	-3,03	-0,00	-4,99	-2,65	1,42	-3,40
15	-3,25	0,34	-5,36	-1,65	1,68	-3,90
16	-3,08	0,49	-5,71	-1,87	2,49	-4,37
17	-3,29	0,81	-6,51	-1,61	2,57	-4,65
18	-3,67	1,37	-6,76	-1,86	3,09	-5,00

19	-3,70	1,72	-7,35	- 1,84	3,40	-5,42
20	-3,85	2,03	-8,02	- 1,91	4,00	-6,13
	Значення $Y_i=y(X_i)$					
	Варіант 25	Варіант26	Варіант27	Варіант 28	Варіант 29	Варіант 30
1	-1,80	-1,65	-1,88	- ,01	-4,13	-3,97
2	-1,66	-1,64	-1,69	- ,06	-4,11	-4,07
3	-1,36	-1,41	-1,52	- ,88	-3,87	-4,04
4	-1,41	-0,91	-1,55	-3,98	-3,74	-4,30
5	-1,13	-0,63	-1,16	-4,36	-3,85	-4,27
6	-0,82	-0,34	-1,27	-4,18	-3,71	-4,54
7	-0,74	-0,12	-1,23	-4,16	-3,53	-4,79
8	-0,76	0,25	-1,36	-4,51	-3,56	-5,07
9	-0,64	0,64	-1,26	-4,53	-3,19	-5,30
10	-0,46	0,96	-1,47	-4,38	-3,04	-5,51
11	-0,30	1,50	-1,72	-4,76	-2,83	-5,83
12	-0,27	1,77	-1,76	-4,66	-2,54	-6,06
13	-0,22	2,24	-2,00	-4,82	-2,41	-6,40
14	-0,11	2,93	-2,03	-4,77	-1,97	-6,83
15	-0,02	3,17	-2,35	-5,12	-1,78	-7,54
16	0,11	3,77	-2,46	-5,23	-1,53	-7,68
17	0,11	4,42	-2,88	-5,40	-1,04	-8,36
18	-0,02	4,79	-3,27	-5,84	-0,86	-8,91
19	0,03	5,50	-3,68	-5,86	-0,48	-9,39
20	0,01	6,01	-3,98	-6,01	0,09	-9,98

ТЕМА 8. Чисельне диференціювання функцій

Формули для знаходження першої та другої похідної. Оцінки похибки чисельного диференціювання. Некоректність задачі чисельного диференціювання. Регуляризація задачі чисельного диференціювання.

Практичне заняття 5. Чисельне диференціювання функцій

Мета: засвоїти основні формули чисельного диференціювання функцій однієї змінної.

План заняття

1. Різницеві формули для визначення похідних на основі інтерполяційного полінома Ньютона.
2. Формули наближеного диференціювання на основі на інтерполяційної формули Стірлінга.
3. Чисельне диференціювання на основі інтерполяційних формул Гаусса і Бесселя.

Теоретичні відомості, що необхідні для виконання даної роботи.

При розв'язуванні практичних задач часто потрібно знайти похідну певного порядку від функції $y = f(x)$, заданої таблично. Також можливо, що, в силу складності аналітичного виразу функції $f(x)$, неопосередковане

диференціювання її викликає труднощі. В цих випадках звичайно звертаються до чисельного диференціювання.

Для виведення формул чисельного диференціювання замінюють дану функцію $f(x)$ на певному відрізку $[a;b]$ інтерполяційною функцією $P(x)$ (найчастіше - поліномом), а потім покладають: $f'(x) = P'(x)$ при $a \leq x \leq b$. Аналогічно роблять при знаходженні похідних вищих порядків функції $f(x)$.

Якщо для інтерполяційної функції $P(x)$ відома похибка $R(x) = f(x) - P(x)$, то похибка похідної $P'(x)$ виражається формулою

$$r(x) = f'(x) - P'(x) = R'(x),$$

тобто похибка похідної інтерполяційної функції дорівнює похідній від похибки цієї функції. Це також вірно і для похідних вищих порядків.

Слід відмітити, що чисельне диференціювання є менш точною операцією, ніж інтерполювання. Дійсно близькість ординат двох кривих $y = f(x)$ і $Y = P(x)$ на відрізку $[a;b]$ ще не гарантує близькість похідних $f'(x)$ і $P'(x)$, тобто малого розходження кутових коефіцієнтів дотичних до даних кривих при однакових значеннях аргументу (рис. 1).

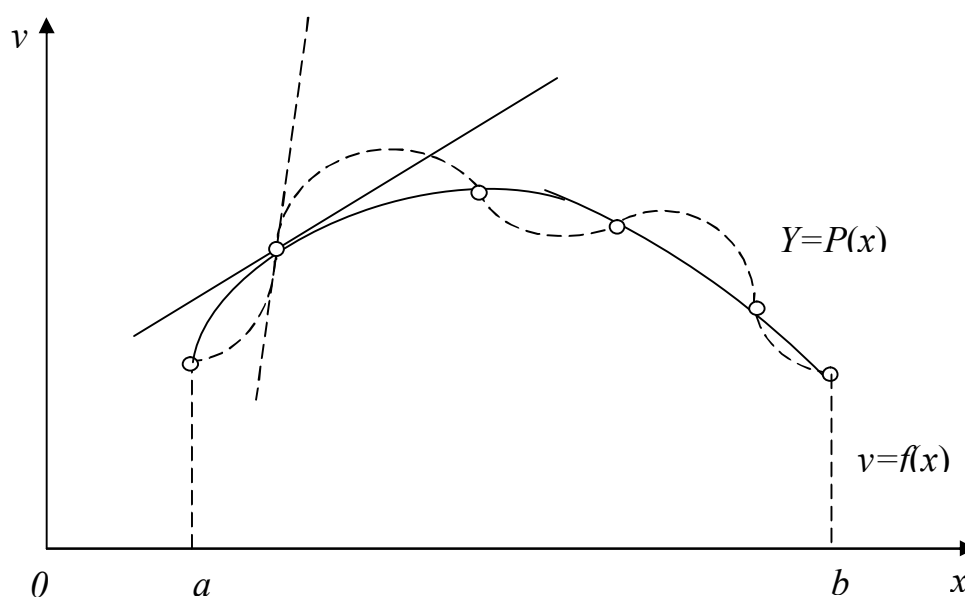


рис. 1

Нехай маємо табличну функцію $y(x)$, задану в рівновіддалених точках x_i ($i=0,1,2,\dots,n$) відрізка $[a;b]$ за допомогою значень $y_i = f(x_i)$. Для знаходження на $[a;b]$ похідних $y' = f'(x)$, $y'' = f''(x)$ і т.д. функцію y наближено замінимо інтерполяційним поліномом Ньютона, побудованим для системи вузлів x_0, x_1, \dots, x_k ($k \leq n$). Маємо:

$$y(x) = y_0 + q\Delta y_0 + \frac{q(q-1)}{2!} \Delta^2 y_0 + \frac{q(q-1)(q-2)}{3!} \Delta^3 y_0 + \frac{q(q-1)(q-2)(q-3)}{4!} \Delta^4 y_0 + \dots,$$

де $q = \frac{x - x_0}{h}$ і $h = x_{i+1} - x_i$ ($i = 0, 1, \dots$). Перемноживши біноми, отримаємо:

$$y(x) = y_0 + q\Delta y_0 + \frac{q^2 - q}{2}\Delta^2 y_0 + \\ + \frac{q^3 - 3q^2 + 2q}{6}\Delta^3 y_0 + \frac{q^4 - 6q^3 + 11q^2 - 6q}{24}\Delta^4 y_0 + \dots$$

Оскільки $\frac{dy}{dx} = \frac{dy}{dq} \cdot \frac{dq}{dx} = \frac{1}{h} \cdot \frac{dy}{dq}$, то

$$y'(x) \approx \frac{1}{h} \left(\Delta y_0 + \frac{2q-1}{2}\Delta^2 y_0 + \frac{3q^2-6q+2}{6}\Delta^3 y_0 + \frac{2q^3-9q^2+11q-3}{12}\Delta^4 y_0 + \right. \\ \left. + \frac{5q^4-40q^3+105q^2-100q+24}{120}\Delta^5 y_0 + \dots \right). \quad (1)$$

Аналогічно, оскільки $y'' = \frac{d(y')}{dx} = \frac{d(y')}{dq} \cdot \frac{dq}{dx}$, то

$$y''(x) \approx \frac{1}{h^2} \left(\Delta^2 y_0 + (q-1)\Delta^3 y_0 + \frac{6q^2-18q+11}{12}\Delta^4 y_0 + \frac{2q^3-12q^2+21q-10}{12}\Delta^5 y_0 \dots \right) \quad (2)$$

Аналогічно можна знайти похідні функції будь-якого порядку.

Зауважимо, що при знаходженні похідних $y'(x)$, $y''(x)$, ... у точці x в якості x_0 потрібно вибирати найближче табличне значення аргументу.

Іноді необхідно знайти похідні функції y в вузлах таблиці x_i . В цьому випадку формули чисельного диференціювання спрощуються. Оскільки кожне табличне значення можна вважати за початкове, то покладемо $x = x_0$, $q = 0$; тоді матимемо:

$$y'(x_0) \approx \frac{1}{h} \left(\Delta y_0 - \frac{\Delta^2 y_0}{2} + \frac{\Delta^3 y_0}{3} - \frac{\Delta^4 y_0}{4} + \frac{\Delta^5 y_0}{5} - \dots \right) \quad (3)$$

і

$$y''(x_0) \approx \frac{1}{h^2} \left(\Delta^2 y_0 - \Delta^3 y_0 + \frac{11}{12}\Delta^4 y_0 - \frac{5}{6}\Delta^5 y_0 + \dots \right). \quad (4)$$

Якщо $P_k(x)$ – інтерполяційний поліном Ньютона, який має скінченні різниці $\Delta y_0, \Delta^2 y_0, \dots, \Delta^k y_0$, і $R_k(x) = y(x) - P_k(x)$ – похибка, то похибка похідної дорівнює $R_k'(x) = y'(x) - P_k'(x)$. Як відомо,

$$R_k(x) = \frac{(x-x_0)(x-x_1)\dots(x-x_k)}{(k+1)!} y^{(k+1)}(\xi) = h^{k+1} \frac{q(q-1)\dots(q-k)}{(k+1)!} y^{(k+1)}(\xi),$$

де ξ – деяке проміжне число між значеннями x_0, x_1, \dots, x_k, x . Тому, припустивши, що $y(x) \in C^{(k+2)}$, отримаємо:

$$R_k'(x) = \frac{dR_k}{dq} \cdot \frac{dq}{dx} = \frac{h^k}{(k+1)!} \left\{ y^{(k+1)}(\xi) \frac{d}{dq} (q(q-1)\dots(q-k)) + q(q-1)\dots(q-k) \frac{d}{dq} (y^{(k+1)}(\xi)) \right\}.$$

Звідси при $x = x_0$ (отже, при $q = 0$), використовуючи співвідношення $\frac{d}{dq} (q(q-1)\dots(q-k))_{q=0} = (-1)^k k!$, матимемо:

$$R_k'(x_0) = (-1)^k \frac{h^k}{k+1} y^{(k+1)}(\xi).$$

Оскільки $y^{(k+1)}(\xi)$ в багатьох випадках важко оцінити, то при h малим наближено покладають:

$$y^{(k+1)}(\xi) \approx \frac{\Delta^{k+1} y_0}{h^{k+1}}$$

і, отже,

$$R_k'(x_0) \approx \frac{(-1)^k}{h} \frac{\Delta^{k+1} y_0}{k+1}. \quad (5)$$

Аналогічно можна знайти похибку $R_k''(x_0)$ для похідної $y''(x_0)$.

Наведені вище формули численного диференціювання для функції y в точці $x = x_0$ використовують лише односторонні значення функції при $x > x_0$. Більшу точність мають симетричні формули диференціювання, що враховують значення функції y як при $x > x_0$, так і при $x < x_0$. Ці формули зазвичай називаються *центральними формулами диференціювання*. Виведемо одну з них, узявши за основу інтерполяційну формулу Стірлінга.

Нехай $\dots, x_{-3}, x_{-2}, x_{-1}, x_0, x_1, x_2, x_3, \dots$ – система рівновіддалених вузлів з кроком $x_{i+1} - x_i = h$ і $y_i = f(x_i)$ – значення функції $y = f(x)$ у вузлах. Покладаючи $q = \frac{x - x_0}{h}$ і замінюючи наближено функцію y інтерполяційним поліномом Стірлінга, маємо:

$$y(x) = y_0 + q\Delta y_{-\frac{1}{2}} + \frac{q^2}{2!} \Delta^2 y_{-1} + \frac{q(q^2-1)}{3!} \Delta^3 y_{-\frac{3}{2}} + \frac{q^2(q^2-1)}{4!} \Delta^4 y_{-2} + \\ + \frac{q(q^2-1)(q^2-2^2)}{5!} \Delta^5 y_{-\frac{5}{2}} + \frac{q^2(q^2-1)(q^2-2^2)}{6!} \Delta^6 y_{-3} + \dots,$$

де використані позначення

$$\Delta y_{-\frac{1}{2}} = \frac{\Delta y_{-1} + \Delta y_0}{2}, \quad \Delta^3 y_{-\frac{3}{2}} = \frac{\Delta^3 y_{-2} + \Delta^3 y_{-1}}{2}, \quad \Delta^5 y_{-\frac{5}{2}} = \frac{\Delta^5 y_{-3} + \Delta^5 y_{-2}}{2}.$$

Тоді, враховуючи, що $\frac{dq}{dx} = \frac{1}{h}$, отримуємо:

$$y'(x) \approx \frac{1}{h} \left(\Delta y_{-\frac{1}{2}} + q\Delta^2 y_{-1} + \frac{3q^2 - 1}{6} \Delta^3 y_{-\frac{3}{2}} + \frac{2q^3 - q}{12} \Delta^4 y_{-2} + \frac{5q^4 - 15q^2 + 4}{120} \Delta^5 y_{-\frac{5}{2}} + \frac{3q^5 - 10q^3 + 4q}{360} \Delta^6 y_{-3} + \dots \right),$$

$$y''(x) \approx \frac{1}{h^2} \left(\Delta^2 y_{-1} + q\Delta^3 y_{-\frac{3}{2}} + \frac{6q^2 - 1}{12} \Delta^4 y_{-2} + \frac{2q^3 - 3q}{12} \Delta^4 y_{-\frac{5}{2}} + \frac{15q^4 - 30q^2 + 4}{360} \Delta^6 y_{-3} + \dots \right).$$

Зокрема, покладаючи $q = 0$, матимемо формули Стірлінга:

$$y'(x_0) \approx \frac{1}{h} \left(\Delta y_{-\frac{1}{2}} - \frac{1}{6} \Delta^3 y_{-\frac{3}{2}} + \frac{1}{30} \Delta^5 y_{-\frac{5}{2}} + \dots \right), \quad (6)$$

і

$$y''(x_0) \approx \frac{1}{h^2} \left(\Delta^2 y_{-1} - \frac{1}{12} \Delta^4 y_{-2} + \frac{1}{90} \Delta^6 y_{-3} + \dots \right). \quad (7)$$

Формули наближеного диференціювання, засновані на першій інтерполяційній формулі Гаусса, такі:

$$y'(x) \approx \frac{1}{h} \left(\Delta y_0 + \frac{2q-1}{2!} \Delta^2 y_{-1} + \frac{3q^2-1}{3!} \Delta^3 y_{-1} + \frac{2q^3-3q^2-q+1}{12} \Delta^4 y_{-2} + \dots \right), \quad (8)$$

$$y''(x) \approx \frac{1}{h^2} \left(\Delta^2 y_{-1} + q\Delta^3 y_{-1} + \frac{6q^2-6q+1}{12} \Delta^4 y_{-2} + \dots \right). \quad (9)$$

Формули наближеного диференціювання, засновані на другій інтерполяційній формулі Гаусса, такі:

$$y'(x) \approx \frac{1}{h} \left(\Delta y_{-1} + \frac{2q+1}{2} \Delta^2 y_{-1} + \frac{3q^2-1}{6} \Delta^3 y_{-1} + \frac{2q^3+3q^2-q-1}{12} \Delta^4 y_{-2} + \dots \right), \quad (10)$$

$$y''(x) \approx \frac{1}{h^2} \left(\Delta^2 y_{-1} + q\Delta^3 y_{-2} + \frac{6q^2-6q-1}{12} \Delta^4 y_{-2} + \dots \right). \quad (11)$$

Формули наближеного диференціювання, засновані на інтерполяційній формулі Бесселя, такі:

$$y'(x) \approx \frac{1}{h} \left(\Delta y_0 + \frac{2q-1}{2} \cdot \frac{\Delta^2 y_{-1} + \Delta^2 y_0}{2} + \frac{3q^2-3q+\frac{1}{2}}{6} \Delta^3 y_{-1} + \frac{2q^3-3q^2-q+1}{12} \cdot \frac{\Delta^4 y_{-2} + \Delta^4 y_{-1}}{2} + \dots \right), \quad (12)$$

$$y''(x) \approx \frac{1}{h^2} \left(\frac{\Delta^2 y_{-1} + \Delta^2 y_0}{2} + \frac{2q-1}{2} \Delta^3 y_{-1} + \frac{6q^2-6q-1}{12} \cdot \frac{\Delta^4 y_{-2} + \Delta^4 y_{-1}}{2} + \dots \right). \quad (13)$$

Приклад розв'язання типових задач

Задача. Знайти значення першої та другої похідних табличної функції

x	$y(x)$	x	$y(x)$
0,8	2,857	2,4	6,503
1,2	3,946	2,8	7,010
1,6	4,938	3,2	7,288
2,0	5,801	3,6	7,301

при: 1) $x_1 = 1,2$; 2) $x_2 = 2,23$; 3) $x_3 = 2,76$; 4) $x_4 = 3,1$.

Розв'язування. Складемо діагональну таблицю скінчених різниць даної функції:

x	$y(x)$	Δy_i	$\Delta^2 y_i$	$\Delta^3 y_i$
0,8	2,857		0,097	
1,2	3,946	1,089		0,032
1,6	4,938	0,992	0,129	
2,0	5,801	0,863		0,032
2,4	6,503	0,702	0,161	
2,8	7,010	0,507		0,034
3,2	7,288	0,278	0,195	
3,6	7,301	0,013		0,034
			0,265	

1) Нехай $x_1 = 1,2$; тоді $t = \frac{x - x_0}{h} = \frac{1,2 - 1,2}{0,4} = 0$. Скористаємося для розрахунків формулами (3), (4),

$$y'(x_0) \approx \frac{1}{h} \left(\Delta y_0 - \frac{\Delta^2 y_0}{2} + \frac{\Delta^3 y_0}{3} - \frac{\Delta^4 y_0}{4} + \frac{\Delta^5 y_0}{5} - \dots \right),$$

$$y''(x_0) \approx \frac{1}{h^2} \left(\Delta^2 y_0 - \Delta^3 y_0 + \frac{11}{12} \Delta^4 y_0 - \frac{5}{6} \Delta^5 y_0 + \dots \right),$$

які отримані з першої інтерполяційної формули Ньютона. Знаходимо

$$y'(1,2) \approx \frac{1}{0,4} \cdot \left(0,992 + \frac{1}{2} \cdot 0,129 - \frac{1}{3} \cdot 0,032 \right) = 2,5 \cdot (0,992 + 0,0645 - 0,0107) = 2,614;$$

$$y''(1,2) \approx \frac{1}{0,4^2} \cdot (-0,129 + 0,032) = 0,606.$$

2) Нехай $x_0 = 2,0$; тоді $t = \frac{2,23 - 2,0}{0,4} = 0,575$. Скористаємося для розрахунків формулами (12), (13),

$$y'(x) \approx \frac{1}{h} \left(\Delta y_0 + \frac{2t-1}{2} \cdot \frac{\Delta^2 y_{-1} + \Delta^2 y_0}{2} + \frac{3t^2 - 3t + \frac{1}{2}}{6} \Delta^3 y_{-1} + \dots \right),$$

$$y''(x) \approx \frac{1}{h^2} \left(\frac{\Delta^2 y_{-1} + \Delta^2 y_0}{2} + \frac{2t-1}{2} \Delta^3 y_{-1} + \frac{6t^2 - 6t - 1}{12} \cdot \frac{\Delta^4 y_{-2} + \Delta^4 y_{-1}}{2} + \dots \right),$$

які отримані із формули Бесселя. Знаходимо

$$y'(2,23) \approx \frac{1}{0,4} \cdot \left(0,702 + \frac{1,15-1}{2} \cdot \frac{-0,161-0,195}{2} + \frac{0,992-1,725+0,5}{6} \cdot (-0,034) \right) = 2,5(0,702 - 0,0134 + 0,0013) = 1,725.$$

$$y''(2,23) \approx \frac{1}{0,4^2} \left(\frac{-0,161-0,195}{2} + \frac{1,15-1}{2} \cdot (-0,034) \right) = 6,25(-0,178 - 0,0026) = -1,129.$$

3) Нехай $x_0 = 2,8$; тоді $t = \frac{2,76 - 2,8}{0,4} = -0,1$. Скористуємося формулами

$$y'(x) \approx \frac{1}{h} \left(\frac{\Delta y_{-1} + \Delta y_0}{2} + t \cdot \Delta^2 y_{-1} + \frac{3t^2 - 1}{6} \cdot \frac{\Delta^3 y_{-2} + \Delta^3 y_{-1}}{2} + \dots \right),$$

$$y''(x) \approx \frac{1}{h^2} \left(\Delta^2 y_{-1} + t \frac{\Delta^3 y_{-2} + \Delta^3 y_{-1}}{2} + \dots \right),$$

які отримують із формул Стірлінга (6), (7). Знаходимо

$$y'(2,76) \approx \frac{1}{0,4} \cdot \left(\frac{0,507 + 0,207}{6} + 0,1 \cdot 0,229 + \frac{0,03-1}{6} \cdot \frac{-0,034-0,036}{2} \right) = 2,5 \cdot (0,3925 + 0,0229 + 0,0057) = 1,053,$$

$$y''(2,76) \approx \frac{1}{0,4^2} \cdot \left(-0,229 - 0,1 \cdot \frac{-0,034-0,036}{2} \right) = 6,25(-0,229 + 0,0035) = -1,409.$$

4) Нехай $x_0 = 2,8$, тоді $t = \frac{3,1 - 2,8}{0,4} = 0,75$. Скористаємося формулами

$$y'(x) \approx \frac{1}{h} \left(\Delta y_0 + \frac{2t-1}{2!} \Delta^2 y_{-1} + \frac{3t^2-1}{3!} \Delta^3 y_{-1} + \frac{2t^3-3t^2-t+1}{12} \Delta^4 y_{-2} + \dots \right),$$

$$y''(x) \approx \frac{1}{h^2} \left(\Delta^2 y_{-1} + t \Delta^3 y_{-1} + \frac{6t^2-6t+1}{12} \Delta^4 y_{-2} + \dots \right)$$

які отримують із першої формули Гаусса (див. (8), (9)). Знаходимо

$$y'(3,1) \approx \frac{1}{0,4} \cdot \left(0,278 + \frac{1,5-1}{2} \cdot (-0,229) + \frac{1,6875-1}{6} \cdot (-0,036) \right) = 2,5 \cdot (0,278 - 0,0572 - 0,0041) = 0,542;$$

$$y''(3,1) \approx \frac{1}{0,4^2} \cdot (-0,229 + 0,75 \cdot (-0,036)) = -1,600.$$

Задачі для самоконтролю

Задача. За допомогою інтерполяційних формул Ньютона, Гаусса, Стірлінга, Бесселя знайти значення першої та другої похідних при даних значеннях аргумента для функції, заданої таблично.

Таблиця 1

x	$y(x)$	x	$y(x)$
2,4	3,526	3,6	4,222
2,6	3,782	3,8	4,331
2,8	3,945	4,0	4,507
3,0	4,043	4,2	4,775
3,2	4,104	4,4	5,159
3,4	4,155	4,6	5,683

- 1) $x = 2,4 + 0,05n$;
- 2) $x = 3,12 + 0,03n$;
- 3) $x = 4,5 - 0,06n$;
- 4) $x = 4,04 - 0,04n$ ($n = 1, 3, 5, 7, \dots, 29$).

Таблиця 2

x	$y(x)$	x	$y(x)$
1,5	10,517	4,5	8,442
2,0	10,193	5,0	8,482
2,5	9,807	5,5	8,862
3,0	9,387	6,0	9,701
3,5	8,977	6,5	11,132
4,0	8,637	7,0	13,302

- 1) $x = 1,6 + 0,08n$;
- 2) $x = 3,27 + 0,11n$;
- 3) $x = 6,3 - 0,12n$;
- 4) $x = 5,85 - 0,09n$ ($n = 2, 4, 6, 8, \dots, 30$).

ТЕМА 9. Чисельне інтегрування функції однієї змінної

Метод прямокутників та метод трапецій. Вибір кроку сітки. Метод Симпсона. Оцінка похибки. Блок-схема методу. Порівняльний аналіз методів прямокутників, трапецій та Симпсона.

Практичне заняття 6. Чисельне інтегрування функції

Мета: засвоїти поширені методи чисельного інтегрування функцій однієї змінної.

План заняття

1. Чисельне інтегрування. Формули трапецій та Сімсона.

2. Формули Рунге покращення апроксимації, застосовані до формул трапецій та Сімпсона.
3. Обчислення інтегралу табличної функції за методами прямокутників та метод трапецій. Обчислення інтегралу табличної функції за методом Сімпсона

Теоретичні відомості, що необхідні для виконання даної роботи, містяться в конспекті лекцій за темою 9.

Приклади розв'язання типових задач

Приклад 1. Обчислити інтеграл $\int_{0,7}^{1,3} \frac{dx}{\sqrt{2x^2 + 0,3}}$ за формулою трапецій з

трьома десятковими знаками.

Розв'язання. З метою досягнення заданої точності необхідно визначити значення n таким, що

$$\frac{(b-a)^3}{12n^2} \cdot M_2 < 0,0005.$$

У нашому випадку $a=0,7$; $b=1,3$; $M_2 \geq \max|f''(x)|$ на $[0,7; 1,3]$, де

$$f(x) = 1/\sqrt{2x^2 + 0,3}.$$

Знаходимо $f'(x) = -2x/\sqrt{(2x^2 + 0,3)^3}$, $f''(x) = (8x^2 - 0,6)/\sqrt{(2x^2 + 0,3)^5}$;

$$M_2 = (8 \cdot 1,3^2 - 0,6)/\sqrt{(2 \cdot 0,7^2 + 0,3)^5} \approx 6,98.$$

Покладемо $M_2 = 7$, тоді n визначаємо з нерівності

$$\frac{(1,3 - 0,7)^3 \cdot 7}{12n^2} < 0,0005.$$

Отримаємо в результаті $n > 16$. Візьмемо $n = 20$.

Обчислення інтегралу виконуємо за формулою $I \approx h \cdot ((y_0 + y_{20})/2 + y_1 + y_2 + \dots + y_{19})$, де $h = (b-a)/n = 0,6/20 = 0,003$, $y_i = y(x_i) = 1/\sqrt{2x_i^2 + 0,3}$, $x_i = 0,7 + i \cdot h$ ($i=0, 1, 2, \dots, 20$).

Приклад 2. Обчислити інтеграл $\int_2^6 \lg(x^2 + 2) dx$ за формулою Сімпсона

при $n_1 = 2$, $n_2 = 4$ та знайти його уточнене значення.

Розв'язання. При $n_1=2$ маємо $h_1 = (6-2)/2 = 2$, при $n_2=4$ - $h_2 = (6-2)/4 = 1$.

Таблиця значень підінтегральної функції з кроком $h_2 = 1$ для $x_i = 2 + i \cdot h$ ($i=0, 1, 2, 3, 4$) така:

i	x_i	$\lg(x_i^2 + 2)$
0	2	0,875061
1	3	1,096910
2	4	1,290035

3	5	1,454845
4	6	1,596597

Використовуючи формулу Сімпсона отримаємо:

при $n_1=2$

$$I_1 \approx (y_0 + 4y_2 + y_4)h_1/3 = (0,8751 + 4 \cdot 1,29 + 1,5966) \cdot 2/3 = 5,0878,$$

при $n_2=4$

$$I_2 \approx (y_0 + 4y_1 + 2y_2 + 4y_3 + y_4)h_2/3 =$$

$$= (0,8751 + 4 \cdot 1,096910 + 2 \cdot 1,29 + 4 \cdot 1,4548 + 1,5966) \cdot 1/3 = 5,0862.$$

Уточнене значення інтегралу знайдемо за формулою Рунге:

$$I_{1,2} = I_2 + \frac{n_1^m}{n_2^m - n_1^m} (I_2 - I_1),$$

де для формули трапецій $m = 2$, а для формули Сімпсона $m = 4$. Отже,

$$I_{1,2} = 5,0862 + \frac{2^4}{4^4 - 2^4} (5,0862 - 5,0878) = 5,0861$$

Задачі для самоконтролю

Задача 1. Обчислити інтеграл, наведений у таблиці 1 згідно номера варіанта, за формулою трапецій з трьома десятковими знаками.

Задача 2. Обчислити інтеграл, наведений у таблиці 1 згідно номера варіанта, за формулою Сімпсона при $n_1 = 2$, $n_2 = 4$ та знайти його уточнене значення.

Таб. 1

№ варіанта	$F_1(x, y_1, y_2)$	$F_2(x, y_1, y_2)$	$y_1(a)$	$y_2(a)$	a	b
1	2	3	4	5	6	7
1	$\arctg\left(\frac{1}{1+y_1^2+y_2^2}\right)$	$\sin(y_1 y_2)$	1	0	-1	1
2	$\arctg(x^2+y_2^2)$	$\sin(x+y_1)$	0,5	1,5	0	2
3	$\frac{y_1+x}{y_1^2+y_2^2}$	$\cos(y_1+xy_2)$	-1	1	0	4
4	$e^{y_1 y_2}$	$xy_1 y_2$	1	0	0	5
5	$\frac{x}{\sqrt{1+x^2+y_2^2}}$	$\frac{y_2}{\sqrt{1+x^2+y_1^2}}$	0,2	0	-1	1
6	$\sin(x^2+y_2^2)$	$\cos(xy_1)$	0	0	0	4
7	$\sin y_2$	$\cos y_1$	0,5	-0,5	1	3
8	$x \cos(y_1+y_2)$	$\sin(y_1-y_2)$	-0,6	2	2	5
9	$\sin y_1 \cos^3 y_2$	$\cos y_1 \cos y_2$	0	0	-1	3
10	$\sin(y_1 y_2)$	$\cos(xy_1 y_2)$	0,5	1,2	0	2
11	$\arctg\left(\frac{1}{1+y_1^2+y_2^2}\right)$	$\sin(y_1 y_2)$	1	1	1	4

12	$\frac{y_2}{\sqrt{1+x^2+y_1^2}}$	$y_2 \cos x - \sin 2x$	0,8	3,5	2	3
13	$\frac{y_1 + x}{y_1^2 + y_2^2}$	$\cos(y_1 + xy_2)$	1	-1	2	4
14	$\sin(xy_2)$	$\cos(xy_1y_2)$	0	-3	2	5
15	$\cos(y_1y_2)$	$\sin(y_1 + y_2)$	0	0	0	2
16	$y_2 \ln x$	$\sin(x - y_2)$	-2	-1	1	4
17	$\cos y_1 \cos y_2$	$\frac{y_2}{\sqrt{1+x^2+y_1^2}}$	0	1	-1	1
18	$\arctg(\frac{1}{1+y_1^2+y_2^2})$	$\sin y_1$	0	0	-2	1
19	$y_1 + y_2$	$\cos(xy_1)$	0	0	0	4
20	$x \cos(y_1 + y_2)$	$Y_1 y_2$	-1	1	0	4
21	$\cos(y_1y_2)$	$y_2 \cos x - \sin 2x$	-1	5	2	4
22	$\frac{y_2}{\sqrt{1+x^2+y_1^2}}$	$\sin(y_1 - y_2)$	1	1	0	3
23	$\cos(x + y_2)$	$\sin(x - y_2)$	0,7	-0,5	0	4
24	$x \cos(y_1 + y_2)$	$\frac{y_2}{\sqrt{1+x^2+y_1^2}}$	0	0	0	2
25	$\frac{x}{\sqrt{1+x^2+y_2^2}}$	$\frac{y_2}{\sqrt{1+x^2+y_1^2}}$	0,2	0	0	3
26	$\sin y_2$	$\cos(y_1y_2)$	1	-1	0	1
27	$\frac{y_2}{\sqrt{1+x^2+y_1^2}}$	$\sin(y_1y_2)$	-2	3	-1	1
28	$\sin(xy_2)$	$x + \cos(xy_1)$	0	0	0	2
29	$y_2 \ln x$	$y_1 y_2$	0	0	-5	0
30	$\sin(y_1y_2)$	$\cos(xy_1y_2)$	-1	2	0	2

ТЕМА 10. Чисельне інтегрування звичайних диференціальних рівнянь

Методи Ейлера і Рунге-Кутта. Оцінка локальної та глобальної похибки методу. Вибір кроку сітки. Узагальнення методу Рунге-Кутта для систем звичайних диференціальних рівнянь першого порядку.

Практичне заняття 7. Чисельне інтегрування звичайних диференціальних рівнянь

Мета: засвоїти найпростіші методи чисельного інтегрування звичайних диференціальних рівнянь.

План заняття

1. Організація обчислень при наближеному розв'язуванні задачі Коші для звичайних диференціальних рівнянь першого порядку за методами Ейлера та Рунге-Кутта.
2. Метод Ейлера з уточненням.

3. Метод Рунге-Кутта для системи диференціальних рівнянь першого порядку та диференціальних рівнянь вищих порядків.

Теоретичні відомості, що необхідні для виконання даної роботи, містяться в конспекті лекцій за темою 10.

Приклади розв'язання типових задач

Приклад 1. Використовуючи метод Ейлера з уточненням, скласти таблицю наближених значень розв'язку диференційного рівняння $y' = x^2 + y^2$, що задовольняє початковій умові $y(0) = 0,1$ на відрізку $[0; 0,5]$, $h = 0,1$. Всі обчислення вести з точністю до чотирьох десяткових знаків.

Розв'язання. Метод Ейлера з уточненням полягає у тому, що кожне значення $y_{k+1} = y(x_{k+1})$, де $y(x)$ – шукана функція, а $x_{k+1} = x_0 + h(k+1)$, $k=0, 1, 2, \dots$, визначається наступним чином: за початкове наближення вибирається

$$y_{k+1}^{(0)} = y_k^{(0)} + h \cdot (x_k^2 + y_k^2);$$

потім знайдене значення уточнюється за формулою

$$y_{k+1}^{(i+1)} = y_k + \frac{h}{2} \left((x_k^2 + y_k^2) + (x_k^2 + (y_{k+1}^{(i)})^2) \right)$$

Уточнення продовжують доти, доки у межах заданої точності два послідовних наближення не будуть співпадати.

Обчислення занесено у таблицю:

k+1	X_{k+1}	Y_k	i	$Y_{k+1}^{(i)}$
1	0,1	0,1	0	0,102
			1	0,10352
			2	0,103536
2	0,2	0,1035	0	0,10857
			1	0,111125
			2	0,11115
			3	0,11115
3	0,3	0,11115	0	0,121385
			1	0,125004
			2	0,125005
4	0,4	0,125	0	0,14256
			1	0,1473
			2	0,147366
			3	0,147367
5	0,5	0,1474	0	0,174573
			1	0,18051
			2	0,180616
			3	0,180617

Відповідь:

$y(0,1) = 0,1035$; $y(0,2) = 0,11115$; $y(0,3) = 0,125$; $y(0,4) = 0,1474$; $y(0,5) = 0,1806$.

Приклад 2. Розв'язати задачу Коші для звичайного диференціального рівняння методом Рунге-Кутта 4 порядку.

$$y' - 2y + x^2 = 1, \quad x \in [0;1], \quad y(0) = 1.$$

Розв'язання. Покладаємо $n = 10$, $h = (1 - 0)/10 = 0,1$.

Початкова точка $x_0 = 0$, $y_0 = 1$.

Обчислення для першої точки.

З початку обчислюємо значення K_0, K_1, K_2, K_3 :

$$K_0 = f(x_0; y_0) = f(0; 1) = 1 + 2 \cdot 1 - 0^2 = 3$$

$$K_1 = f\left(x_0 + \frac{h}{2}; y_0 + h \cdot \frac{K_0}{2}\right) = f(0,05; 1,15) = 1 + 2 \cdot 1,15 - 0,05^2 = 3,2975$$

$$K_2 = f\left(x_0 + \frac{h}{2}; y_0 + h \cdot \frac{K_1}{2}\right) = f(0,05; 1,164875) = 1 + 2 \cdot 1,164875 - 0,05^2 = 3,32725$$

$$K_3 = f(x_0 + h; y_0 + h \cdot K_2) = f(0,1; 1,332725) = 1 + 2 \cdot 1,332725 - 0,1^2 = 3,65545$$

Обчислюємо значення y_1 :

$$y_1 = y_0 + \frac{h}{6} \cdot (K_0 + 2 \cdot K_1 + 2 \cdot K_2 + K_3) = 1 + \frac{0,1}{6} \cdot (3 + 2 \cdot 3,2975 + 2 \cdot 3,32725 + 3,65545) = 1,3317491667$$

Аналогічно можна вычислити значення функції во 2, 3, ..., 10 точках.

Приклад 3. Для системи диференціальних рівнянь першого порядку

$$\begin{cases} y_1' = f_1(x, y_1, y_2, \dots, y_n) \\ y_2' = f_2(x, y_1, y_2, \dots, y_n) \\ \dots \\ y_n' = f_n(x, y_1, y_2, \dots, y_n) \end{cases}$$

з початковими умовами $y_1(x_0) = y_{1,0}$, $y_2(x_0) = y_{2,0}$, ..., $y_n(x_0) = y_{n,0}$, формули Рунге-Кутта 4-го порядку такі:

$$y_{1,i+1} = y_{1,i} + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

$$y_{2,i+1} = y_{2,i} + \frac{h}{6}(l_1 + 2l_2 + 2l_3 + l_4)$$

.....

$$y_{n,i+1} = y_{n,i} + \frac{h}{6}(m_1 + 2m_2 + 2m_3 + m_4)$$

де

$$k_1 = f_1(x_i, y_{1,i}, y_{2,i}, \dots, y_{n,i})$$

$$l_1 = f_2(x_i, y_{1,i}, y_{2,i}, \dots, y_{n,i})$$

.....

$$m_1 = f_n(x_i, y_{1,i}, y_{2,i}, \dots, y_{n,i})$$

$$k_2 = f_1\left(x_i + \frac{h}{2}, y_{1,i} + \frac{hk_1}{2}, y_{2,i} + \frac{hl_1}{2}, \dots, y_{n,i} + \frac{hm_1}{2}\right)$$

.....

$$m_2 = f_n\left(x_i + \frac{h}{2}, y_{1,i} + \frac{hk_1}{2}, y_{2,i} + \frac{hl_1}{2}, \dots, y_{n,i} + \frac{hm_1}{2}\right)$$

$$k_3 = f_1\left(x_i + \frac{h}{2}, y_{1,i} + \frac{hk_2}{2}, y_{2,i} + \frac{hl_2}{2}, \dots, y_{n,i} + \frac{hm_2}{2}\right)$$

.....

$$m_3 = f_n\left(x_i + \frac{h}{2}, y_{1,i} + \frac{hk_2}{2}, y_{2,i} + \frac{hl_2}{2}, \dots, y_{n,i} + \frac{hm_2}{2}\right)$$

$$k_4 = f_1(x_i + h, y_{1,i} + hk_3, y_{2,i} + hl_3, \dots, y_{n,i} + hm_3)$$

.....

$$m_4 = f_n\left(x_i + \frac{h}{2}, y_{1,i} + \frac{hk_1}{2}, y_{2,i} + \frac{hl_1}{2}, \dots, y_{n,i} + \frac{hm_1}{2}\right)$$

Наприклад, рівняння $y' - 2y'' + y = 0$ перетворюємо до системи першого порядку:

$$\begin{cases} y_1' = y_2 \\ y_2' = 2y_2 - y_1 \end{cases}$$

Тоді формули Рунге-Кутта 4-го порядку такі:

$$k_1 = y_{2,i}$$

$$l_1 = 2y_{2,i} - y_{1,i}$$

$$k_2 = y_{2,i} + \frac{hl_1}{2}$$

$$l_2 = 2\left(y_{2,i} + \frac{hl_1}{2}\right) - \left(y_{1,i} + \frac{hk_1}{2}\right)$$

$$k_3 = y_{2,i} + \frac{hl_2}{2}$$

$$l_3 = 2\left(y_{2,i} + \frac{hl_2}{2}\right) - \left(y_{1,i} + \frac{hk_2}{2}\right)$$

$$k_4 = y_{2,i} + hl_3$$

$$l_4 = 2(y_{2,i} + hl_3) - (y_{1,i} + hl_3)$$

$$y_{1,i+1} = y_{1,i} + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

$$y_{2,i+1} = y_{2,i} + \frac{h}{6}(l_1 + 2l_2 + 2l_3 + l_4)$$

На першому кроці з $(x_0, y_{1,0}, y_{2,0})$ отримуємо $(x_1, y_{1,1}, y_{2,1})$, де в даному прикладі $y_{2,1}$ - похідна y_1 в даній точці.

Задачі для самоконтролю

Задача 1. Розв'язати диференціальне рівняння і систему диференціальних рівнянь методами Ейлера і Рунге-Кутта, згідно варіанту, наведеному в табл. 1.

Табл. 1.

Номер варіанта	Номер завдання	Завдання для виконання роботи
1	1	$y' = \frac{x+y}{x-y} + e^x, \quad y_0 = -2$
	2	$y_1' = 3y_1 - y_2 + y_3 + e^x,$ $y_2' = y_1 + y_2 + y_3 + x,$ $y_3' = 4y_1 - y_2 + 4y_3,$ $y_1(0) = 0.3, \quad y_2(0) = -0.2, \quad y_3(0) = 0.$
	3	$y'''' - x \sin(x) = 0, \quad y(0) = 1,$ $y'(0) = 0, \quad y''(0) = 2, \quad y'''(0) = 3.$
2	1	$y' = \frac{3x + 0.55y^2 + 3}{x+1}, \quad y_0 = 1$
	2	$y_1' = 5y_1 + 2y_2 - 3y_3,$ $y_2' = 4y_1 + 5y_2 - 4y_3,$ $y_3' = 6y_1 + 4y_2 - 4y_3,$ $y_1(0) = 0.3, \quad y_2(0) = -1.2, \quad y_3(0) = 3.$

	3	$3y''' + 2y - 4 = 0,$ $y(0) = 2, \quad y'(0) = 1, \quad y''(0) = 3$
3	1	$y' = \frac{y}{x+2} + x^2 + 2x, \quad y_0 = 2$
	2	$y_1' = y_1 + 3y_2 - y_3 + e^{-x},$ $y_2' = y_1 + 4y_2 + 5y_3 - x^3,$ $y_3' = y_1 + y_2 - 5.5y_3 - x^2,$ $y_1(0) = 2, \quad y_2(0) = -5, \quad y_3(0) = 3.$
	3	$y'''' - 10y'''' + 9y' = 0, \quad y(0) = 1,$ $y'(0) = 1, \quad y''''(0) = 2, \quad y''''(0) = -3.$

Задача 2. Чисельно розв'язати задачу Коші для звичайного диференціального рівняння другого порядку, наведену в таблиці 2, відповідно до варіанту. Визначити близькість отриманого розв'язку до точного значення з допомогою оцінок: $\beta_1 = \max|y_{Ti} - y_{Mi}|$ - лінійна оцінка,

$\beta_2 = \frac{\sqrt{\sum_{i=1}^n (y_{Ti} - y_{Mi})^2}}{\sqrt{\sum_{i=1}^n y_{Ti}^2}}$ - інтегральна оцінка, де y_{Ti} - точний розв'язок, y_{Mi} - отриманий наближений розв'язок.

Таб. 2

Варіант	Диференціальне рівняння	Початкова умова	[a,b]	Крок	Точний розв'язок
1	$y'' - 2y' + y = 0$	$y(2)=1$ $y'(2) = 2$	[2;4]	0.2	$y = (1 - 3x)e^{x-2}$
2	$y'' - 3y' + 2y - 2x + 3 = 0$	$y(0)=1$ $y'(0) = 2$	[0;2]	0.2	$y = e^x + x$
3	$y'' + y = 4e^x$	$y(0)=4$ $y'(0) = -3$	[0;1]	0.1	$y = 2 \cos(x) - 5 \sin(x) + 2e^x$
4	$x^2 y'' + xy' = 0$	$y(1)=5$ $y'(1) = -1$	[1;1.5]	0.05	$y = 5 - \ln(x)$
5	$y'' - 2y' = 2e^x$	$y(1)=-1$ $y'(1) = 0$	[1;2]	0.1	$y = e^{2x-1} - 2e^x + e - 1$
6	$y'' + 4y = \cos(3x)$	$y(0)=0.8$ $y'(0) = 2$	[0;1]	0.1	$y = \cos(2x) + \sin(2x) - 0.2 \cos(3x)$
7	$y'' + 2y' + 2y = xe^x$	$y(0)=0$ $y'(0) = 0$	[0;1.5]	0.1	$y = e^{-x}(x - \sin(x))$
8	$(1 + x^2)y'' + (y')^2 + 1 = 0$	$y(0)=1$ $y'(0) = 1$	[0;0.5]	0.05	$y = 1 - x + 2 \ln(-1 + x)$
9	$y'' + 4y' + 4y = 0$	$y(0)=1$ $y'(0) = -1$	[0;1]	0.1	$y = (1 + x)e^{-2x}$
10	$y'' - 3y' = e$	$y(0)=2.2$ $y'(0) = 0.8$	[0;0.2]	0.02	$y = 2 + 0.1(e^{3x} + e^{5x})$
11	$x^2 y'' - 2y = 0$	$y(1)=0.83$ $y'(1) = 0.66$	[1;2]	0.1	$y = 0.5x^2 + \frac{1}{3x}$
12	$y'' - 5y' + 6y = e^x$	$y(0)=0$ $y'(0) = 0$	[0;0.2]	0.02	$y = -e^{2x} + 0.5(e^{2x} + x^x)$

13	$y'' + y = 1 + e^x$	$y(0)=2.5$ $y'(0) = 1.5$	[0;1]	0.1	$y = \cos(x) + \sin(x) + \frac{e^x}{2} + 1$
14	$x^2 y'' + 2.5y'x - y = 0$	$y(1)=2$ $y'(1) = 3.5$	[1;2]	0.1	$y = 3\sqrt{x} - x^{-2}$
15	$y'' + y' = x^2 - x + 2$	$y(0)=1$ $y'(0) = 0$	[0;1]	0.1	$y = \cos(x) + \sin(x) + x^2 - x$

ТЕМА 11. Чисельні методи математичної фізики

Чисельні методи розв'язування двовимірних рівнянь Лапласа та Пуассона. Метод послідовних ітерацій. Чисельні методи розв'язування одновимірних рівнянь теплопровідності та хвильових рівнянь. Двошарові та трьохшарові скінчено-різницеві схеми. Оцінки похибки апроксимації. Умови стійкості алгоритму.

Практичне заняття 8. Чисельні методи математичної фізики

Мета: засвоїти найпростіші методи чисельного інтегрування еліптичних, параболічних та гіперболічних рівнянь.

План заняття

1. Приклади математичних моделей, які зводяться до рівнянь теплопровідності та одновимірних хвильових рівнянь.
2. Побудова сіток і скінчено-різницевої апроксимації для диференціальних рівнянь еліптичного, гіперболічного та параболічного типів.
3. Організація обчислень при застосуванні двошарових та трьохшарових скінчено-різницевої схем. Практична перевірка умов стійкості скінчено-різницевої схеми.
4. Чисельне розв'язування двовимірного рівняння Лапласа скінчено-різницевим методом.
5. Чисельне розв'язування одновимірного рівняння теплопровідності скінчено-різницевим методом.
6. Чисельне розв'язування одновимірного хвильового рівняння скінчено-різницевим методом.

Теоретичні відомості, що необхідні для виконання даної роботи, містяться в конспекті лекцій за темою 11.

Приклад 1. Використовуючи метод сіток, скласти таблицю наближених значень розв'язку одновимірного хвильового рівняння $\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}$, що задовольняє початковим умовам на відрізку $[0, 1]$: $u(x,0) = x \sin(\pi x)$, $u_t(x,0) = (x+1)^2$, та крайовим умовам $u(0,t) = 2t$, $u(1,t) = 0$, $0 \leq t \leq 0,5$.

Розв'язання. Побудуємо на прямокутнику $0 \leq x \leq 0,5$, $0 \leq t \leq 0,5$ рівномірну сітку з кроком $h=0,1$ по x і кроком $\tau=0,1$ по t . Тоді, після апроксимації, отримуємо скінченно-різницеву схему:

$$u_{i,j+1} = u_{i+1,j} + u_{i-1,j} - u_{i,j-1},$$

де $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$, $n = (1 - 0)/h = 10$, $m = (0,5 - 0)/\tau = 5$.

Обчислення проводимо у такому порядку:

- обчислюємо значення $u_{i0} = x_i \sin(\pi x_i)$ при $x_i = ih$ (вони відповідають значенню $t_0 = 0$).
- обчислюємо значення $u_{0j} = 2t_j$ при $t_j = j\tau$ (вони відповідають значенню $x_0 = 0$).
- обчислюємо значення $u_{nj} = 0$ при $t_j = j\tau$ (вони відповідають значенню $x_n = 1$).
- обчислюємо значення $u_{i1} = u_{i0} + h(x_i + 1)^2$ при $x_i = ih$ (вони відповідають значенню $t_1 = 0,1$).
- обчислюємо значення u_{ij} для усіх інших значень x, t за формулою $u_{i,j+1} = u_{i+1,j} + u_{i-1,j} - u_{i,j-1}$, використовуючи вже відомі значення u_{ij} .

Одержанні розв'язки заносимо у таблицю 1:

Таб. 1

t_j	x_i										
	0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
0	0,0000	0,0309	0,1176	0,2427	0,3804	0,5000	0,5706	0,5663	0,4702	0,2781	0,0000
0,1	0,2000	0,1519	0,2616	0,4117	0,5764	0,7250	0,8266	0,8553	0,7942	0,6391	0,0000
0,2	0,4000	0,4307	0,4460	0,5953	0,7563	0,9031	1,0097	1,0546	1,0242	0,5161	0,0000
0,3	0,6000	0,6941	0,7644	0,7906	0,9219	1,0410	1,1310	1,1786	0,7764	0,3851	0,0000
0,4	0,8000	0,9337	1,0387	1,0910	1,0753	1,1498	1,2098	0,8529	0,5394	0,2603	0,0000
0,5	1,0000	1,1446	1,2604	1,3234	1,3189	1,2442	0,8717	0,5707	0,3367	0,1544	0,0000

У таблиці 1 для $x_i = 0,4$; $t_j = 0,2$ маємо $u_{4,2} = 0,7563$. Це значення обчислюється за формулою $u_{4,2} = u_{5,1} + u_{3,1} - u_{4,0} = 0,7250 + 0,4117 - 0,3804 = 0,7563$.

Приклад 2. Використовуючи метод сіток, розв'язати чисельно двовимірне рівняння Лапласа $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$ в прямокутній області $\Omega = \{(x, y), 0 \leq x \leq a, 0 \leq y \leq b\}$ з краєвими умовами $u(0, y) = f_1(y)$; $u(a, y) = f_2(y)$; $u(x, 0) = f_3(x)$; $u(x, b) = f_4(x)$.

Розв'язання. Введемо позначення $u_{ij} = u(x_i, y_j)$. Накладаємо на прямокутну область сітку $x_i = h \cdot i$; $i = 0, 1, \dots, n$; $y_j = l \cdot j$; $j = 0, 1, \dots, m$. Тоді $x_n = h \cdot n$, $y_m = l \cdot m$. Частинні похідні апроксимуємо по формулах

$$\frac{\partial^2 u}{\partial x^2} = \frac{u_{i+1,j} - 2 \cdot u_{i,j} + u_{i-1,j}}{h^2} + O(h^2);$$

$$\frac{\partial^2 u}{\partial y^2} = \frac{u_{i,j+1} - 2 \cdot u_{i,j} + u_{i,j-1}}{l^2} + O(l^2),$$

і замінимо рівняння Лапласа конечноразностним рівнянням

$$\frac{u_{i+1,j} - 2 \cdot u_{i,j} + u_{i-1,j}}{h^2} + \frac{u_{i,j+1} - 2 \cdot u_{i,j} + u_{i,j-1}}{l^2} = 0,$$

де: $i=1, \dots, n-1, j=1, \dots, m-1$ (тобто для внутрішніх вузлів). Виразимо $u_{i,j}$ при $h = l$, і отримуємо систему

$$\begin{aligned} u_{ij} &= (u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}) / 4; \\ u_{i0} &= f_3(x_i); \\ u_{im} &= f_4(x_i); \\ u_{0j} &= f_1(y_j); \\ u_{nj} &= f_2(y_j). \end{aligned}$$

Для знаходження наближеного розв'язку цієї системи будемо послідовність ітерацій за методом Гауса-Зейделя

$$u_{ij}^{(s+1)} = \frac{1}{4} (u_{i-1,j}^{(s+1)} + u_{i+1,j}^{(s)} + u_{i,j+1}^{(s)} + u_{i,j-1}^{(s+1)}),$$

де s -поточна ітерація. Умова закінчення ітераційного процесу

$$\max_{i,j} |u_{ij}^{(s+1)} - u_{ij}^{(s)}| < \varepsilon.$$

Приклад 3. Нехай заданий розподіл температури однорідного стрижня $u(x,0) = \varphi_1(x)$ у початковий момент часу $t=0$ і закони зміни температури залежно від часу на його кінцях $x=0$ і $x=l$: $u(0,t) = \varphi_2(t)$; $u(l,t) = \varphi_3(t)$. Потрібно знайти розподіл температури $u = u(x,t)$ уздовж стрижня у будь-який момент часу t .

Розв'язання. Область $\Omega = \{t, x : t \in [0, T], x \in [0, l]\}$ замінимо сітковою, розбиваючи її за допомогою кроку h по x і за допомогою кроку k по t . В результаті неперервна область Ω замінюється дискретною множиною вузлових точок $\Omega_h = \{x_i, t_j\}, i = \overline{0, n}; j = \overline{0, m}$. Тепер шукатимемо функцію $u(x,t)$ тільки на дискретній множині Ω , тобто $u_{ij} = u(x_i, t_j)$ - двовимірна таблиця значень шуканої функції у вузлових точках.

Скінчено-різницевий аналог початкового завдання такий: потрібно знайти значення функції $u(x_i, t_j)$, що задовольняють скінчено-різницевому рівнянню:

$$\frac{u_{ij+1} - u_{ij-1}}{2k} = \frac{u_{i+1j} - 2u_{ij} + u_{i-1j}}{h^2}, \quad i = \overline{1, n-1}; j = \overline{1, m-1}$$

і додатковим умовам:

$$u(x_i, 0) = \varphi_1(x_i), i = \overline{0, n}, u(0, t_j) = \varphi_2(t_j), j = \overline{0, m}, u(l, t_j) = \varphi_3(t_j), j = \overline{0, m}$$

Розрахункова формула при $k = h^2 / 6$ набирає наступного вигляду:

$$u_{ij+1} = \frac{1}{6} (u_{i+1j} + 4u_{ij} + u_{i-1j}), i = \overline{1, n-1}; j = \overline{0, m}.$$

Задачі для самоконтролю

Задача 1. Побудувати графік залежності температури вздовж стержня за часом. Довжина стержня $L=5$, на лівому кінці стержня $T = 100^0\text{C}$, правий кінець стержня знаходиться у ізоляційному матеріалі, на початку стержинь мав температуру $T_0=15^0\text{C}$, коефіцієнт температуропровідності $a=0,01$.

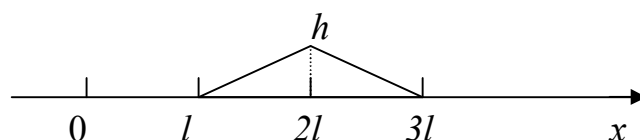
Задача 2. Знайти чисельний розв'язок задачі відповідно до варіанту:

Варіант	Задача
1	$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, 1])$ $u(0, t) = u(1, t) = 0, \quad u(x, 0) = \sin \pi x, \quad u_t(x, 0) = 0.$
2	$\frac{\partial^2 u}{\partial t^2} = 4 \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, 1]),$ $u(0, t) = u(1, t) = 0, \quad u(x, 0) = 1 - \cos 2\pi x, \quad u_t(x, 0) = 0.$
3	$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, 1]),$ $u(0, t) = u(1, t) = 0, \quad u(x, 0) = \sin^2 \pi x, \quad u_t(x, 0) = 1.$
4	$\frac{\partial^2 u}{\partial t^2} = 4 \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, 1]),$ $u(0, t) = u(1, t) = 0, \quad u(x, 0) = 1 - \cos 4\pi x, \quad u_t(x, 0) = 1.$
5	$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, \pi]),$ $u(0, t) = u(\pi, t) = 0, \quad u(x, 0) = \sin \pi x, \quad u_t(x, 0) = 0.$
6	$\frac{\partial^2 u}{\partial t^2} = 4 \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, 1]),$ $u(0, t) = u(1, t) = 0, \quad u(x, 0) = x(1-x), \quad u_t(x, 0) = 0.$
7	$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, \pi]),$ $u(0, t) = u(\pi, t) = 0, \quad u(x, 0) = x(\pi-x), \quad u_t(x, 0) = 0.$
8	$\frac{\partial^2 u}{\partial t^2} = 4 \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, \pi]),$ $u(0, t) = u(\pi, t) = 0, \quad u(x, 0) = 1 - \cos 2x, \quad u_t(x, 0) = 1$
9	$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, 1]),$ $u(0, t) = u(1, t) = 0, \quad u(x, 0) = \sin \pi x$
10	$\frac{\partial u}{\partial t} = 4 \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, 1]),$ $u(0, t) = u(1, t) = 0, \quad u(x, 0) = 1 - \cos 2\pi x.$

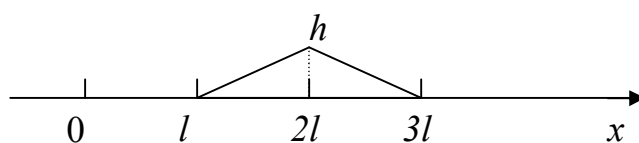
11	$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, \pi]),$ $u(0, t) = u(\pi, t) = 0, \quad u(x, 0) = \sin x.$
12	$\frac{\partial u}{\partial t} = 4 \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, \pi]),$ $u(0, t) = u(\pi, t) = 0, \quad u(x, 0) = 1 - \cos 2x.$
13	$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, 2]),$ $u(0, t) = u(2, t) = 0, \quad u(x, 0) = \sin \pi x + \cos \pi x - 1.$
14	$\frac{\partial u}{\partial t} = 4 \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, 1]),$ $u(0, t) = u(1, t) = 0, \quad u(x, 0) = \sin 2\pi x - \cos 2\pi x + 1.$
15	$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, 1]),$ $u(0, t) = u(1, t) = 0, \quad u(x, 0) = x(1 - x).$
16	$\frac{\partial u}{\partial t} = 4 \frac{\partial^2 u}{\partial x^2} \quad (x \in [0, \pi]),$ $u(0, t) = u(\pi, t) = 0, \quad u(x, 0) = x(\pi - x).$
17	$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad 0 < x < 1, \quad 0 < y < 1,$ $u(x, 0) = u(x, 1) = 0, \quad u(0, y) = y(1 - y), \quad u(1, y) = \sin \pi y$
18	$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad 0 < x < 1, \quad 0 < y < 1,$ $u(0, y) = y(1 - y), \quad u(1, y) = u_y(x, 0) = u_y(x, 1) = 0.$
19	$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad 0 < x < 1, \quad 0 < y < 1,$ $u(x, 0) = 0, \quad u(x, 1) = 0, \quad u(0, y) = u(1, y) = y.$
20	$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad 0 < x < 1, \quad 0 < y < 1,$ $u(x, 0) = \sin \pi x, \quad u(x, 1) = 0, \quad u(0, y) = u(1, y) = 0.$

Задача 3. Методом сіток побудувати графіки положення напівобмеженої струни $0 < x < \infty$ в моменти часу $t_k = \frac{kl}{2a}, k = 1, 2, 3, 4$, відповідно до варіанту:

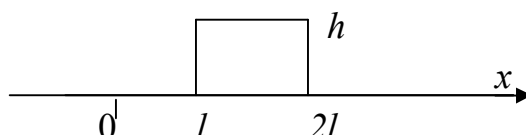
1. Струна закріплена в точці $x = 0$ та збуджена початковим відхиленням $\varphi(x)$ (див. рис.). Початкова швидкість нульова.



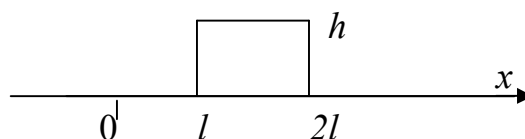
2. Струна вільна в точці $x = 0$ та збуджена початковим відхиленням $\varphi(x)$ (див. рис.). Початкова швидкість нульова.



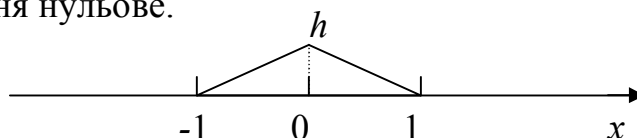
3. Струна закріплена в точці $x = 0$ та збуджена початковою швидкістю $\psi(x)$ (див. рис.). Початкове відхилення нульове.



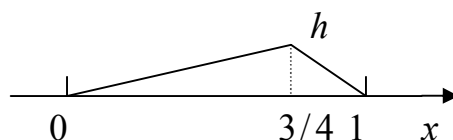
4. Струна вільна в точці $x = 0$ та збуджена початковою швидкістю $\psi(x)$ (див. рис.). Початкове відхилення нульове.



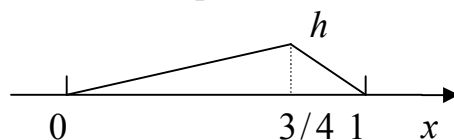
5. Струна збуджена початковою швидкістю $\psi(x)$ (див. рис.). Початкове відхилення нульове.



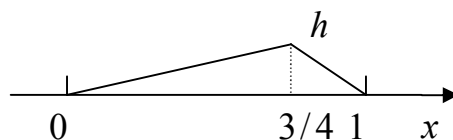
6. Струна закріплена в точках $x = 0, x = 1$ та збуджена початковим відхиленням $\varphi(x)$ (див. рис.). Початкова швидкість нульова.



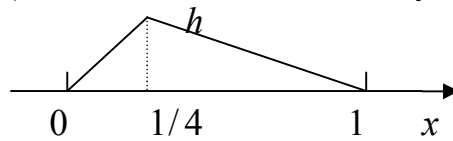
7. Струна закріплена в точці $x = 0$, вільна в точці $x = 1$ та збуджена початковим відхиленням $\varphi(x)$ (див. рис.). Початкова швидкість нульова.



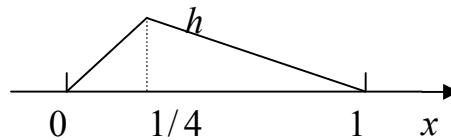
8. Струна вільна в точці $x = 0$, закріплена в точці $x = 1$ та збуджена початковим відхиленням $\varphi(x)$ (див. рис.). Початкова швидкість нульова.



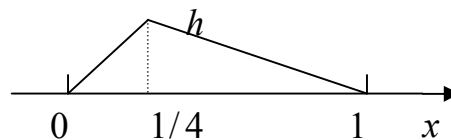
9. Струна закріплена в точках $x=0$, $x=1$ та збуджена початковим відхиленням $\varphi(x)$ (див. рис.). Початкова швидкість нульова.



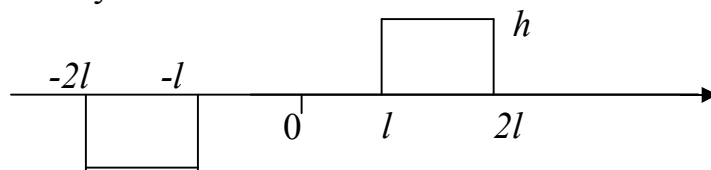
10. Струна закріплена в точці $x=0$, вільна в точці $x=1$ та збуджена початковим відхиленням $\varphi(x)$ (див. рис.). Початкова швидкість нульова.



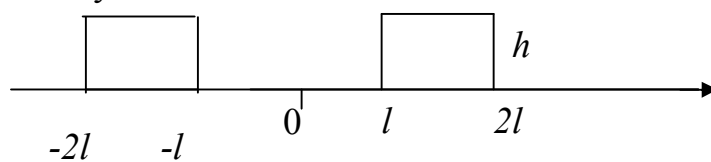
11. Струна вільна в точці $x=0$, закріплена в точці $x=1$ та збуджена початковим відхиленням $\varphi(x)$ (див. рис.). Початкова швидкість нульова.



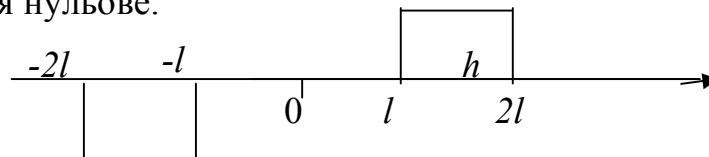
12. Струна збуджена початковою швидкістю $\psi(x)$ (див. рис.). Початкове відхилення нульове.



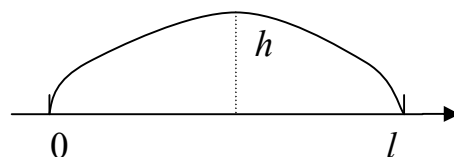
13. Струна збуджена початковою швидкістю $\psi(x)$ (див. рис.). Початкове відхилення нульове.



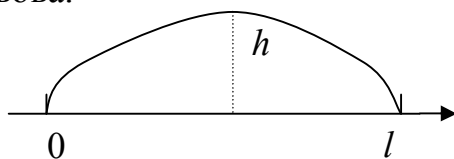
14. Струна збуджена початковою швидкістю $\psi(x)$ (див. рис.). Початкове відхилення нульове.



15. Струна закріплена в точках $x=0$, $x=1$, та збуджена початковим відхиленням $\varphi(x)$ у вигляді параболи (див. рис.). Початкова швидкість нульова.



16. Струна закріплена в точці $x=0$ та вільна в точці $x=l$. Струна збуджена початковим відхиленням $\varphi(x)$ у вигляді параболи (див. рис.). Початкова швидкість нульова.



Задача 4. Поставити і чисельно розв'язати граничну задачу про визначення температури стержня довжини L , якщо початкова температура є довільною функцією координати x вздовж вісі стержня, на боковій поверхні стержня здійснюється теплообмін з оточуючим середовищем заданої температури, а кінці стержня підтримуються при заданій температурі.

Задача 5. Поставити і чисельно розв'язати граничну задачу про визначення температури стержня довжини L , якщо початкова температура є довільною функцією координати x вздовж вісі стержня, на боковій поверхні стержня здійснюється теплообмін за законом Ньютона з оточуючим середовищем заданої температури, а на кінцях стержня подається ззовні заданий тепловий потік.

Задача 6. Поставити і чисельно розв'язати задачу про визначення температури стержня довжини L , якщо початкова температура є довільною функцією координати x вздовж вісі стержня, на боковій поверхні стержня здійснюється теплообмін з оточуючим середовищем заданої температури, на кінцях стержня здійснюється теплообмін за законом Ньютона з середовищем, температура якого задана.

Задача 7. Поставити і чисельно розв'язати задачу про визначення температури на внутрішній поверхні стінової панелі житлового будинку (тобто в кімнаті), яка ззовні закрита шаром пенопласту. Товщина пенопласту – b_1 , товщина панелі – b_2 . На внутрішній та зовнішній поверхні відбувається теплообмін з оточуючим середовищем за законом Ньютона. Температури повітря ззовні, в кімнаті, а також початковий розподіл в стіні та пенопласті задано як функції часу.

Задача 8. Поставити і чисельно розв'язати задачу про дифузю частинок у розчині, припускаючи, що концентрація частинок у розчині залежить лише від однієї координати z (висота) і часу t . Поставити граничні умови, що відповідають непроникливій перегородці.

Задача 9. Поставити і чисельно розв'язати задачу про поперечні коливання однорідної струни, закріпленої на одному кінці та вільною на іншому. В початковий момент струна знаходилася в стані рівноваги та була виведена з нього шляхом відхилення $u_0(x)$.

Задача 10. Поставити і чисельно розв'язати задачу про поперечні коливання однорідної струни, закріпленої на правому кінці та вільною на

лівому. В початковий момент струна знаходилася в стані рівноваги та була виведена з нього шляхом надання швидкості $u_1(x)$.

Задача 11. Поставити і чисельно розв'язати задачу про подовжні коливання однорідного стержня, закріпленого на лівому кінці та збудженого силою $P(t)$ на правому кінці. В початковий момент стержень знаходився в стані рівноваги.

Задача 12. Поставити і чисельно розв'язати задачу про подовжні коливання однорідного стержня, закріпленого на обох кінцях та збудженого силою теплового розширення, яке викликано змінною температурою $T(x,t)$. В початковий момент стержень знаходився в стані рівноваги. Вказівка: теплове розширення враховуються модифікацією закону Гука $\sigma = E \frac{\partial u}{\partial x} - \alpha T$. Тут E - модуль Юнга, α - коефіцієнт теплового розширення.

Задача 13. Поставити і чисельно розв'язати задачу про поперечні коливання однорідного стержня прямокутного перетину, жорстко закріпленого на правому кінці та вільного на лівому кінці. В початковий момент задані зміщення та швидкість стержня.

Задача 14. Поставити і чисельно розв'язати задачу про поперечні коливання однорідного стержня шарнірно опертого на кінцях, що навантажений вздовж довжини силою $F(x,t)$. В початковий момент стержень знаходився в стані рівноваги.

Задача 15. Поставити і чисельно розв'язати задачу про поперечні коливання однорідного стержня шарнірно опертого на лівому кінці та навантаженого на правому кінці силою $P(t)$. В початковий момент стержень знаходився в стані рівноваги.

Задача 16. Поставити і чисельно розв'язати задачу знаходження зміщення u при малих подовжніх коливаннях газу в циліндричній трубі, лівий кінець якої закритий перегородкою, а правий відкритий.

Задача 17. Поставити і чисельно розв'язати задачу про поперечні коливання прямокутної мембрани закріпленої на трьох сторонах та вільною на четвертій стороні $x_1 = l_1, 0 \leq x_2 \leq l_2$. В початковий момент задані зміщення та їх швидкість.

Задача 18. Поставити і чисельно розв'язати задачу про поперечні коливання прямокутної мембрани закріпленої на стороні $x_1 = l_1, 0 \leq x_2 \leq l_2$, вільною на сторонах $0 \leq x_1 \leq l_1, x_2 = 0$, $0 \leq x_1 \leq l_1, x_2 = l_2$ та навантаженою на стороні $x_1 = 0, 0 \leq x_2 \leq l_2$ силою $P(t)$. В початковий момент мембрана знаходилася в стані рівноваги.

МЕТОДИЧНІ ВКАЗІВКИ

до організації самостійної та індивідуальної роботи студентів з дисципліни

“ЧИСЕЛЬНІ МЕТОДИ В ІНФОРМАТИЦІ”

ВСТУП

Курс “Чисельні методи в інформатиці” – обов'язковий компонент загальної та професійної освіти. Значення курсу у загальноосвітній підготовці визначається насамперед тим, при побудові математичних моделей складних, зокрема соціально-економічних систем, часто виникають задачі, для яких знаходження аналітичного розв'язку або надто ускладнене, або неможливе. Це обумовлює необхідність розробки спеціальних математичних методів та алгоритмів, що дозволяють одержати чисельний розв'язок та оцінити його точність.

Дисципліна “Чисельні методи в інформатиці” призначена для оволодіння майбутніми бакалаврами спеціальності «Інтелектуальні системи прийняття рішень» основами знань у галузі комп'ютерних методів дослідження математичних моделей, на базі яких провадиться подальше вивчення спеціальних дисциплін, пов'язаних з фаховою діяльністю.

У результаті вивчення дисципліни студент повинен одержати фундаментальні теоретичні знання у галузі чисельних методів і закріпи їх на практичних заняттях.

Програма курсу “Чисельні методи в інформатиці” охоплює достатній обсяг матеріалу, який дозволяє підготувати бакалаврів належного рівня. Вона складена згідно з вимогами «Положення про програму дисципліни» і є нормативним документом, який визначає мету і завдання курсу, місце курсу серед дисциплін професійної підготовки.

Типова програма відповідає діючим підручникам, що використовується у навчальному процесі.

Метою дисципліни є засвоєння теоретичних основ, формування у студентів практичних навичок щодо використання основних методів чисельного розв'язання складних математичних моделей. Вона полягає в тому щоб забезпечити загальний розвиток світогляду студентів, їх ознайомлення з методологією чисельних методів, на основі яких додатково розглянути сучасні уявлення про методи аналізу та синтезу систем. Чисельні методи спрямовані на розв'язання складних проблем. Метою застосування чисельних методів до конкретної проблеми є підвищення ступеня обґрунтованості рішення, що приймається.

Досягнення цієї загальної мети у практиці викладення курсу можна здійснювати різними шляхами:

- конкретизацією теорій, явищ і процесів під час вивчення курсу та закріпленні знань, використовуючи навчальний матеріал предметів профциклу;
- показом практичного використання в даній професійній діяльності знань, отриманих під час вивчення курсу;
- складанням задач з професійно спрямованим змістом, виконанням при їх рішенні розрахунків, пов'язаних з майбутньою професійною діяльністю студентів;

- проведенням практичних робіт із курсу, інтегрованих з деякими практичними роботами профциклу;
- використанням діафільмів, кіно- і відеофільмів із загальноосвітніх дисциплін профциклу з ілюстрацією в них наступності та взаємозв'язку основ системного аналізу і професійних знань.

Завдання курсу полягає у вивченні теоретичних відомостей та набуття студентами практичних навичок чисельного розв'язання задач великої розмірності, систем алгебраїчних та диференційних рівнянь, нелінійних математичних моделей тощо, опануванні сучасними пакетами прикладних програм, що дозволяють здійснювати чисельні розрахунки.

Критерії оцінки успішності повинні відповідати навчальній програмі й найбільш важливим вимогам до знань студентів:

1. Знання фактів, явищ. Вірне, науково достовірне їх пояснення.
2. Оволодіння науковими термінами, поняттями, законами, методами, правилами; вміння користуватися ними при поясненні нових фактів, розв'язуванні різних питань і виконанні практичних завдань.
3. Максимальна ясність, точність думки, вміння відстоювати свої погляди, захищати їх.

Методи і форми викладання дисципліни

Вивчення дисципліни передбачає лекційні та практичні заняття. Значна частина матеріалу дисципліни відведена під індивідуальні заняття студентів під керівництвом та СРС. СРС використовується студентами для проробки лекційного матеріалу, навчально-методичної літератури, при підготовці до практичних та лабораторних занять.

Навчальна дисципліна “Чисельні методи в інформатиці” розглядається як елемент загальної освіти і відіграє роль апарату вивчення та засвоєння закономірностей навколишнього світу та фундаменту професійних знань галузевої техніки чи технології. Навчання слід спрямувати на розв'язання задач, формування умінь використовувати наукові знання для розв'язання практичних завдань у відповідних галузях професійної діяльності, прийняття обґрунтованих конструкторсько-технологічних рішень в конкретних виробничих завданнях, застосування фізичних методів і теорій у поясненні суті технічних і технологічних процесів.

Форми і засоби проміжного та підсумкового контролю: експрес-контроль рівня готовності студента до проведення та практичних робіт; перевірка виконання позааудиторних завдань; оцінка роботи студента під час заняття (виступи, доповнення, участь у дискусії); виконання домашніх завдань; контрольні роботи в кінці залікового кредиту. Оцінка індивідуальних результатів здобуття знань студентами проводиться у формі заліку за кредитно-модульною методологією навчання, критерії якої визначаються у навчальній робочій програмі за стобальною системою, яка трансформується у стандартні залікові диференційовані оцінки відповідно до вимог Міністерства освіти та науки України.

ТЕОРЕТИЧНІ ОСНОВИ

1. Елементи теорії похибок

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

- 1) неусувна похибка
- 2) похибка методу
- 3) похибка обчислень

Неусувна похибка є наслідком:

а) неточності вхідних даних, що входять до математичного описання задачі,

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

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

Похибка обчислень виникає при вводиті-виводі даних до ПЕОМ та при виконанні математичних операцій.

Основна задача теорії похибок – знаходження області невизначеності результату.

Розглянемо процес заокруглення чисел. Якщо число $x=4,167493$ і його потрібно заокруглити до п'яти десяткових знаків після коми, то будемо мати $x^*=4,16749$. Тобто, якщо старший розряд, що відкидається менше 5, то попередня цифра не змінюється. Якщо $x=4,167493$ потрібно заокруглити до чотирьох знаків після коми, то $x^*=4,1675$. Тобто, якщо старший розряд, що відкидається дорівнює, або більше 5, то попередня цифра в числі збільшується на 1.

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

Розглянемо приклади заокруглення чисел:

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

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

Абсолютна та відносна похибки

Нехай x – точне значення деякої величини, а x^* – її відоме наближене значення.

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

$$|x^* - x| \leq \Delta(x^*).$$

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

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

Відзначимо, що точність результату краще характеризує відносна похибка. Інформацію про абсолютну та відносну похибки можна використати для наступного представлення числа x :

$$x = x^* \pm \Delta(x^*),$$

$$x = x(1 \pm \delta(x^*)).$$

Значущими цифрами числа називаються всі цифри в його запису, починаючи з першої ненульової зліва.

Наприклад:

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

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

Приклад 1. Нехай $x^*=14,537$ і відомо, що $\Delta(x^*)=0,04$. Скільки вірних значущих цифр має число x^* ?

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

Приклад 2. Нехай $x^*=8,677142$ і $\Delta(x^*)=3 \cdot 10^{-4}$. Скільки вірних значущих цифр має число x^* ?

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

Приклад 3. Нехай $x^*=0,046725$ і $\Delta(x^*)=0,008$. Скільки вірних значущих цифр має число x^* ?

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

Пряма задача теорії похибок

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

Використовуючи формулу Лагранжа, будемо мати

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

де

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

При практичних розрахунках окрім оцінки (3) використовують оцінку

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

яку називають лінійною оцінкою похибки. Знайдемо відносну похибку:

$$\delta(y^*) \leq \sum_{j=1}^n \left| \frac{\frac{\partial f(x_1^*, x_2^*, \dots, x_n^*)}{\partial x_j}}{f(x_1^*, x_2^*, \dots, x_n^*)} \right|.$$

Використовуючи ці формули, визначимо похибки результатів математичних операцій.

1. Похибка суми

$$y = f(x_1, x_2) = x_1 + x_2, \quad x_1, x_2 > 0.$$

Оскільки $f'_{x_j}(x^*) = 1$, то з (4) будемо мати

$$\Delta(y^*) = \Delta(x_1^*) + \Delta(x_2^*),$$

і відповідно

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

Аналогічно знаходимо похибки для інших математичних операцій.

2. Похибка різниці.

$$y = f(x_1, x_2) = x_1 - x_2, \quad x_1 > x_2 > 0.$$

$$\Delta(y^*) = \Delta(x_1^*) + \Delta(x_2^*),$$

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

3. Похибка множення.

$$y = f(x_1, x_2) = x_1 \cdot x_2, \quad x_1, x_2 > 0.$$

$$\Delta(y^*) = |x_2^*| \Delta(x_1^*) + |x_1^*| \Delta(x_2^*),$$

$$\delta(y^*) = \delta(x_1^*) + \delta(x_2^*).$$

4. Похибка ділення.

$$y = f(x_1, x_2) = x_1 / x_2, \quad x_1, x_2 > 0.$$

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

$$\delta(y^*) = \delta(x_1^*) + \delta(x_2^*).$$

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

Розглянемо деякі приклади.

Приклад 4. Заокруглюючи наступні числа до трьох значущих цифр, визначити абсолютну та відносну похибки отриманих наближених чисел:

1) 0,1545; 2) 1,343; 3) -372,75.

Розв'язання.

1) $x = 0,1545$. Заокруглення до трьох значущих цифр дає $x^* = 0,155$, тоді $\Delta(x^*) = 0,0005 = 5 \cdot 10^{-4}$, а відносна похибка

$$\delta(x^*) = 5 \cdot 10^{-4} / 0,155 \approx 0,32 \cdot 10^{-4}.$$

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

$$\delta(x^*) = 3 \cdot 10^{-3} / 1,34 = 2,2 \cdot 10^{-3}.$$

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

$$\delta(x^*) = 0,25 / 373 = 6,7 \cdot 10^{-4}.$$

Приклад 5. Визначити кількість вірних цифр в числі x^* , якщо відома його відносна похибка:

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

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

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

Розв'язання.

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

2) Обчислимо абсолютну похибку $\Delta(x^*) = x^* \delta(x^*) = 9,4698 \cdot 0,1 \cdot 10^{-2} = 0,0094698$. Тоді в числі x^* будуть вірними дві цифри 9 та 4.

3) Абсолютна похибка буде дорівнювати $\Delta(x^*)=47361 \cdot 0,01=473,61$. Отже в числі x^* будуть вірними дві цифри 4 та 7.

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

Приклад 6. На гіпотетичній ЕОМ з мантисою довжини чотири знайти суму $S=0,2764+0,3944+1,475+26,46+1364$

а) сумуючи від меншого доданку до більшого;

б) сумуючи від більшого доданку до меншого.

Розв'язання.

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

$$S_4=S_3+26,46=2,15+226,46=28,61,$$

$$S=S_5=S_4+1364=29+1393.$$

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

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

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

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

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

Приклад 7. Нехай числа $\sqrt{2,01}=1,417744688$ та $\sqrt{2}=1,414213562$ задані з десятьма вірними значущими цифрами. Скільки вірних значущих цифр матиме число $x^* = \sqrt{2,01} - \sqrt{2}$?

Розв'язання. Віднімаючи, отримаємо $x^*=0,003531126$. Позначимо $x_1^*=1,417744688$, $x_2^*=1,414213562$. Тоді абсолютні похибки $\Delta(x_1^*) = \Delta(x_2^*) = 0,5 \cdot 10^{-9}$. Абсолютна похибка різниці $x^* = x_1^* - x_2^*$ буде дорівнювати $\Delta(x^*) = \Delta(x_1^*) + \Delta(x_2^*) = 10^{-9}$. Оскільки $10^{-9} < 0,5 \cdot 10^{-8}$, то робимо висновок, що число x^* має шість вірних значущих цифр 3,5,3,1,1,2.

Відзначимо, що те ж саме значення можна отримати, подавши x^* у вигляді

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

причому для цього достатньо взяти величини x_1^* і x_2^* достатньо взяти з сімома вірними значущими цифрами.

Приклад 8. Оцінити похибку обчислення функції

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

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

Розв'язання. Згідно з формулою (4), для абсолютної похибки результату отримаємо

$$\begin{aligned} \Delta(f^*) &= \left| \frac{2x^* z^*}{(y^*)^3} \right| \Delta(x^*) + \left| \frac{3(x^*)^2 z^*}{(y^*)^4} \right| \Delta(y^*) + \left| \frac{(x^*)^2}{(y^*)^3} \right| \Delta(z^*) = \frac{2 \cdot 0,15 \cdot 1,14}{(2,13)^3} \cdot 0,005 + \\ &+ \frac{3 \cdot (0,15)^2 \cdot 1,14}{(2,13)^4} \cdot 0,01 + \frac{(0,15)^2}{(2,13)^3} \cdot 0,007 = 0,00017695 + 0,00003738 + \\ &+ 0,000016298 \approx 0,00023 = 2,3 \cdot 10^{-4}. \end{aligned}$$

$$\text{Знайдемо } f(x^*, y^*, z^*) = \frac{(0,15)^2 \cdot 1,14}{(2,13)^3} = 0,0022265429.$$

$$\text{Тоді } \delta(f^*) = \frac{2,3 \cdot 10^{-4}}{0,00265429} = 0,08665.$$

Приклад 9. Висота та радіус основи циліндра виміряні з точністю до 0,5%. Яка відносна похибка при обчисленні об'єму циліндра, якщо $\pi^* = 3,14$?

Розв'язання. $V = \pi R^2 h$. Більш точне значення $\pi = 3,14159265$, отже $\Delta(\pi^*) = 0,16 \cdot 10^{-2}$, а $\delta(\pi^*) = 0,16 \cdot 10^{-2} / 3,14 = 0,0005 = 0,05\%$. Тоді, згідно до формули про відносну похибку добутку будемо мати

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

Приклад 10. Ребро куба виміряне з точністю до 0,02 см. дорівнює 8 см. Знайти абсолютну та відносну похибки при обчисленні об'єму куба.

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

Приклад 11. Визначити відносну похибку числа, що записане в ЕОМ з числення β та довжиною мантиси t .

Розв'язання. Число x^* можна записати в ЕОМ у вигляді

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

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

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

Тоді

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

Отже $\delta(x^*) \leq \beta^{1-t}$. Якщо ж числа вводяться за правилами заокруглення, то $d_{t+1} \leq 0,5\beta$ і тоді будемо мати, що

$$\delta(x^*) \leq \frac{1}{2}\beta^{1-t}.$$

Обернена задача теорії похибок

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

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

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

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

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

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

б) вважаємо всі похибки рівними, причому максимально можливими,

тобто покладемо

$$\Delta(x_1^*) = \Delta(x_2^*) = \dots = \Delta(x_n^*) = \delta,$$

де

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

Приклад 12. Сторона квадрату дорівнює 2м. З якою точністю її потрібно виміряти, щоб похибка знаходження площі не перевищувала 1см^2 ?

Розв'язання. Позначимо сторону квадрату через x ; $S=x^2$, $S'=2x$. Тоді за формулою отримаємо

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

Приклад 13. З якою кількістю вірних значущих цифр потрібно взяти вільний член квадратного рішення

$$x^2 - 2x + \lg 2 = 0,$$

щоб отримати корені рівняння з чотирма вірними значущими цифрами?

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

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

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

Аналогічно, розглядаючи функцію $f(z) = 1 - \sqrt{1 - z}$ отримаємо, що для знаходження кореня x_2 з точністю $0,5 \cdot 10^{-4}$ потрібно обчислити $\lg 2$ з чотирма вірними значущими цифрами після коми, тобто $\lg 2 = 0,3010$.

Приклад 14. В п'ятизначних логарифмічних таблицях дано значення десяткових логарифмів з точністю до $\varepsilon = 0,5 \cdot 10^{-6}$. Оцінити величину можливої похибки при знаходженні числа за його логарифмом, якщо саме число знаходиться між 300 та 400.

Розв'язання. Позначимо $y = \lg x$, $x \in [300; 400]$. За умовою задачі $\Delta(y^*) \leq 0,5 \cdot 10^{-6}$ і потрібно знайти $\Delta(x^*)$. Маємо $y' = \frac{1}{x \ln 10}$. Тоді будемо мати $\Delta(x^*) = x(\ln 10) \cdot \Delta(y^*) \leq 400 \cdot 2,30 \cdot 0,5 \cdot 10^{-6} = 0,00046$, отже x можна знайти принаймні з трьома вірними значущими цифрами після коми.

2. Методи розв'язання нелінійних систем

Розглянемо систему рівнянь:

$$\begin{cases} f_1(x_1, \dots, x_n) = 0 \\ \dots \quad \dots \quad \dots \\ f_n(x_1, \dots, x_n) = 0 \end{cases}$$

Перепишемо її у векторному вигляді:

$$\vec{f}(\vec{x}) = 0 \quad (1)$$

Метод простої ітерації

В цьому методі рівняння (1) зводиться до еквівалентного вигляду

$$\vec{x} = \vec{\Phi}(\vec{x}) \quad (2)$$

Ітераційний процес представимо у вигляді:

$$\vec{x}^{k+1} = \vec{\Phi}(\vec{x}^k), \quad (3)$$

початкове наближення \vec{x}^0 – задано.

Нехай оператор Φ визначений на множині H . За теоремою про стискуючі відображення ітераційний процес (3) сходиться, якщо виконується умова

$$\|\vec{\Phi}(\vec{x}) - \vec{\Phi}(\vec{y})\| \leq q \|\vec{x} - \vec{y}\|, \quad 0 < q < 1, \quad (4)$$

або

$$\|\Phi'(\vec{x})\| \leq q < 1 \quad (5)$$

$\vec{x} \in U_r$, $\Phi'(x) = \left(\frac{\partial \varphi_i}{\partial x_j} \right)_{i,j=1}^n$. Для похибки справедлива оцінка

$$\|\vec{x}^m - \vec{x}\| \leq \frac{q^n}{1-q} \|\vec{x}^1 - \vec{x}^0\|.$$

Частинним випадком методу простої ітерації є метод релаксації для рівняння (1)

$$\vec{x}^{k+1} = \vec{x}^k - \tau \vec{F}(\vec{x}^k),$$

де $\tau < \frac{2}{\|F'(\vec{x})\|}$.

Метод Ньютона

Розглянемо рівняння

$$\vec{F}(\vec{x}) = 0.$$

Представимо його у вигляді

$$\vec{F}(\vec{x}^k) + F'(\vec{\xi}^k)(\vec{x} - \vec{x}^k) = 0, \quad (6)$$

де $\vec{\xi}^k = \vec{x}^k + \theta_k(\vec{x}^k - \vec{x}^k)$, $0 < \theta_k < 1$ Тут $F(x) = \left(\frac{\partial f_i}{\partial x_j} \right)_{i,j=1}^n$ - матриця Якобі

для $\vec{F}(\vec{x})$. Можемо наближено вважати $\vec{\xi}^k \approx \vec{x}^k$. Тоді з (6) матимемо:

$$\vec{F}(\vec{x}^k) + F'(\vec{x}^k)(\vec{x}^{k+1} - \vec{x}^k) = 0 \quad (7)$$

Ітераційний процес представимо у вигляді:

$$\vec{x}^{k+1} = \vec{x}^k - [F'(\vec{x}^k)]^{-1} \vec{F}(\vec{x}^k). \quad (8)$$

Для реалізації методу Ньютона потрібно, щоб існувала обернена матриця $[F'(\vec{x}^k)]^{-1}$.

Можна не шукати обернену матрицю, а розв'язувати на кожній ітерації СЛАР

$$\begin{cases} A_k \bar{z}^k = \bar{F}(\bar{x}^k) \\ \bar{x}^{k+1} = \bar{x}^k - \bar{z}^k \end{cases}, k = 0, 1, 2, \dots \quad (9)$$

де \bar{x}^0 — задано, а матриця $A_k = F'(\bar{x}^k)$.

Метод має квадратичну збіжність, якщо добре вибрано початкове наближення. Складність методу (при умові використання методу Гаусса розв'язання СЛАР (9)) на кожній ітерації

$$Q_n = \frac{2}{3}n^3 + O(n^2),$$

де n — розмірність системи (1).

Модифікований метод Ньютона

Ітераційний процес має вигляд :

$$\bar{x}^{k+1} = \bar{x}^k - [F'(\bar{x}^0)]^{-1} F(\bar{x}^k).$$

Тепер обернена матриця обчислюється тільки на нульовій ітерації. На інших обчислення нового наближення зводиться до множення матриці $A_0 = [F'(\bar{x}^0)]^{-1}$ на вектор $\bar{F}(\bar{x}^k)$ та додавання до \bar{x}^k .

Запишемо метод у вигляді системи лінійних рівнянь (аналог (9))

$$\begin{cases} A_0 \bar{z}^k = \bar{F}(\bar{x}^k) \\ \bar{x}^{k+1} = \bar{x}^k - \bar{z}^k \end{cases} \quad (10)$$

Оскільки матриця A_0 розкладається на трикутні (або обертається) один раз, то складність цього методу на одній ітерації (окрім нульової) $Q_n = O(n^2)$. Але цей метод має лінійну швидкість збіжності.

Можливе циклічне застосування модифікованого методу Ньютона, тобто коли обернену матрицю похідних шукаємо та обертаємо через певне число кроків ітераційного процесу.

3. Алгебраїчна проблема власних значень.

Нехай задано матрицю $A: (n \times n)$. Тоді задача на власні значення ставиться так: знайти число λ та вектор $\bar{x} \neq 0$, що задовольняють рівнянню

$$A\bar{x} = \lambda\bar{x}. \quad (1)$$

λ називається власним значенням A , а \bar{x} — власним вектором. З (1) маємо

$$\det(A - \lambda E) \equiv P_n(\lambda) \equiv (-1)^n \lambda^n + a_n \lambda^{n-1} + \dots + a_0 = 0.$$

Тут $P_n(\lambda)$ — характеристичний багаточлен.

Для розв'язання багатьох задач механіки, фізики, хімії потрібне знаходження всіх власних значень $\lambda_i, i = \overline{1, n}$, а іноді й всіх власних векторів \bar{x}_i , що відповідають λ_i . Цю задачу називають повною проблемою власних значень.

В багатьох випадках потрібно знайти лише максимальне або мінімальне за модулем власне значення матриці. При дослідженні стійкості коливальних процесів іноді потрібно знайти два максимальних за модулем власних значення матриці. Останні дві задачі (2) - (4) називають частковими проблемами власних значень.

Степеневий метод

1) Знаходження λ_{\max} : $|\lambda_1| \equiv \lambda_{\max} > |\lambda_2| \geq |\lambda_3| \geq \dots$

Нехай \bar{x}^0 - заданий вектор, будемо послідовно обчислювати вектори

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

Тоді $\bar{x}^k = A^k \bar{x}^0$. Нехай $\{\bar{e}_i\}_{i=1}^n$ - система власних векторів. Представимо \bar{x}^0 у вигляді:

$$\bar{x}^0 = \sum_{i=1}^n c_i \bar{e}_i.$$

Оскільки $A\bar{e}_i = \lambda_i \bar{e}_i$, то $\bar{x}^k = \sum_{i=1}^n c_i \lambda_i^k \bar{e}_i$. При великих k $\bar{x}^k \approx c_1 \lambda_1^k \bar{e}_1$. Тому

$$\mu_1^{(k)} = \frac{x_m^{k+1}}{x_m^k} = \lambda_1 + O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right).$$

Значить $\mu_1^{(k)} \xrightarrow{k \rightarrow \infty} \lambda_1$.

Якщо матриця $A = A^T$ симетрична, то існує ортонормована система векторів $(\bar{e}_i, \bar{e}_j) = \delta_{ij}$. Тому

$$\begin{aligned} \mu_1^{(k)} &= \frac{(\bar{x}^{k+1}, \bar{x}^k)}{(x^k, x^k)} = \frac{\left(\sum_i c_i \lambda_i^{k+1} \bar{e}_i, \sum_j c_j \lambda_j^k \bar{e}_j\right)}{\left(\sum_i c_i \lambda_i^k \bar{e}_i, \sum_j c_j \lambda_j^k \bar{e}_j\right)} = \frac{\sum_i c_i^2 \lambda_i^{2k+1}}{\sum_i c_i^2 \lambda_i^{2k}} = \\ &= \frac{c_1^2 \lambda_1^{2k+1} + c_2^2 \lambda_2^{2k+1} + \dots}{c_1^2 \lambda_1^{2k} + c_2^2 \lambda_2^{2k} + \dots} = \lambda_1 + O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right) \xrightarrow{k \rightarrow \infty} \lambda_1 \end{aligned}$$

Це означає збіжність до максимального за модулем власного значення з квадратичною швидкістю.

Якщо $|\lambda_1| > 1$, то при проведенні ітерацій відбувається зріст компонент вектора \bar{x}^k , що приводить до "переповнення" (overflow). Якщо ж $|\lambda_1| < 1$, то це приводить до зменшення компонент (underflow). Позбутися негативу такого явища можна нормуючи вектори \bar{x}^k на кожній ітерації.

Алгоритм степеневого методу знаходження максимального за модулем власного значення з точністю ε виглядає так:

1. $\bar{x}^0 \rightarrow \bar{e}_0 = \frac{\bar{x}^0}{\|\bar{x}^0\|}$;
2. $\bar{x}^{k+1} = A\bar{e}^k, \mu_1^{(k)} = (\bar{x}^{k+1}, \bar{e}^k), \bar{e}^{k+1} = \frac{\bar{x}^{k+1}}{\|\bar{x}^{k+1}\|}, k = 0, 1, \dots$;
3. $|\mu_1^{(k+1)} - \mu_1^{(k)}| \geq \varepsilon$ goto 2;
4. $\lambda_1 \approx \mu_1^{(k+1)}$.

За цим алгоритмом для симетричної матриці $A^T = A$ швидкість прямування $\mu_1^{(k)}$ до λ_{\max} - квадратична.

2. Знаходження $\lambda_2: |\lambda_1| > |\lambda_2| > |\lambda_3| \geq \dots$. Нехай λ_1, \bar{e}_1 відомі, то

$$\mu_2^{(k)} = \frac{x_m^{k+1} - \lambda_1 x_m^k}{x_m^k - \lambda_1 x_m^{k-1}} \xrightarrow{k \rightarrow \infty} \lambda_2,$$

де $\bar{x}^{k+1} = A\bar{x}^k$, x_m^k - m -та компонента \bar{x}^k . Алгоритм обчислення λ_2, \bar{e}_2 , використовує нормування векторів та скалярні добутки для обчислення $\mu_2^{(k)}$.

3. Знаходження мінімального власного числа $\lambda_{\min}(A) = \min_i |\lambda_i(A)|$.

Припустимо, що $\lambda_i(A) > 0$ та відоме λ_{\max} . Розглянемо матрицю $B = \lambda_{\max} E - A$. Маємо

$$\lambda_i(B) = \lambda_{\max} - \lambda_i(A), \forall i,$$

Тому $\max_i \lambda_i(B) = \lambda_{\max} - \min_i \lambda_i(A)$. Звідси $\lambda_{\min}(A) = \lambda_{\max}(A) - \lambda_{\max}(B)$.

Якщо $\exists i: \lambda_i(A) < 0$, то будемо матрицю $\bar{A} = \sigma E + A, \sigma > 0, \bar{A} > 0$ і для неї попередній розгляд дає необхідний результат. Замість λ_{\max} в матриці B можна використовувати $\|A\|$.

Ще один спосіб обчислення мінімального власного значення полягає в використанні обернених ітерацій:

$$A\bar{x}^{k+1} = \bar{x}^k, k = 0, 1, \dots \quad (3)$$

Але цей метод вимагає більшої кількості арифметичних операцій: складність методу на основі формули (2) $Q = O(n^2)$, а на основі (3) - $Q = O(n^3)$, оскільки треба розв'язувати СЛАР, але збігається метод (3) швидче.

Ітераційний метод обертання

Цей метод розв'язання повної проблеми власних значень для симетричних матриць $A^T = A$. Існує матриця U , що приводить матрицю A до діагонального виду:

$$A = U\Lambda U^T, \quad (1)$$

де Λ - діагональна матриця, по діагоналі якої стоять власні значення λ_i ; U - унітарна матриця, тобто: $U^{-1} = U^T$.

З (1) маємо

$$\Lambda = U^T A U \quad (2)$$

Нехай $\exists \tilde{U}$ - матриця, така що $\tilde{\Lambda} = \tilde{U}^T A \tilde{U}$ і $\tilde{\Lambda} = (\tilde{\lambda}_{ij})_{i,j}^n$ $|\tilde{\lambda}_{ij}| < \delta \ll 1, i \neq j$.

Тоді діагональні елементи мало відрізняються від власних значень

$$|\tilde{\lambda}_{ii} - \lambda_i(A)| < \varepsilon = \varepsilon(\delta).$$

Введемо $t(A) = \sum_{\substack{i,j=1 \\ i \neq j}}^n (a_{ij})^2$. З малості величини $t(A)$ витікає, що діагональні

елементи малі. Покладаємо $A = A_0$ і за допомогою матриць обертання U_k

$$U_k = \begin{pmatrix} 1 & 0 & 0 & 0 \cdots 0 & 0 & 0 & 0 \\ 0 & \ddots & \vdots & \vdots \cdots \vdots & \vdots & \vdots & 0 \\ 0 & 0 \cdots & \cos \varphi & 0 \cdots 0 & -\sin \varphi & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 \cdots & \sin \varphi & 0 \cdots 0 & \cos \varphi & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \cdots \vdots & & \ddots & 0 \\ 0 & 0 & 0 & 0 \cdots 0 & 0 & & 1 \end{pmatrix} \begin{matrix} \leftarrow i \\ \\ \\ \leftarrow j \\ \\ \end{matrix}$$

$$\begin{matrix} \uparrow i & \uparrow j \end{matrix}$$

що повертають систему векторів на кут φ , побудуємо послідовність $\{A_k\}$ таку, що $A_k \rightarrow \Lambda$ при $k \rightarrow \infty$. Матриця обертання U_k є унітарною: $U_k^{-1} = U_k^T$.

Послідовно будуємо:

$$A_{k+1} = U_k^T A_k U_k, \quad (3)$$

Процес (3) називається *монотонним*, якщо: $t(A_{k+1}) < t(A_k)$. Для матриці (3) виконується:

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} \cos 2\varphi + \frac{1}{2} (a_{jj}^{(k)} - a_{ii}^{(k)}) \sin 2\varphi. \quad (4)$$

Оскільки $t(A_{k+1}) = t(A_k) - 2[a_{ij}^{(k)}]^2$, якщо вибрати φ з умови $a_{ij}^{(k+1)} = 0$, то

$$\varphi = \varphi_k = \frac{1}{2} \arctan p^{(k)} \quad p^{(k)} = \frac{2a_{ij}^{(k)}}{a_{ii}^{(k)} - a_{jj}^{(k)}}, \text{ де } |a_{ij}^{(k)}| = \max_{\substack{m,l \\ m \neq l}} |a_{ml}^{(k)}|.$$

Тоді $t(A_k) \rightarrow 0, k \rightarrow \infty$. Чим більше n , тим більше ітерацій необхідно для зведення A до Λ . Якщо матриця несиметрична, то застосовують *QR*, *QL* методи.

4. Багатокрокові методи розв'язання задачі Коші. Методи Адамса
Для розв'язання задачі Коші

$$\frac{du}{dt} = f(t, u), \quad t > 0, \quad u(0) = u_0 \quad (1)$$

введемо сітку $\omega_\tau = \{t_n = n\tau, n = 0, 1, \dots\}$ з постійним кроком $\tau > 0$. Позначимо через $y_n = y(t_n)$, $f_n = f(t_n, y_n)$ функції, визначені на сітці ω_τ .

Лінійним m -кроковим різницевим методом називається рівняння:

$$\frac{a_0 y_n + a_1 y_{n-1} + \dots + a_m y_{n-m}}{\tau} = b_0 f_n + b_1 f_{n-1} + \dots + b_m f_{n-m}, \quad n = m, m+1, \dots, \quad (2)$$

де a_k, b_k – числові коефіцієнти, які не залежать від n , причому $a_0 \neq 0$.

Рівняння (2) слід розглядати як рекурентне співвідношення, яке виражає нове значення $y_n = y(t_n)$ через знайдені раніше значення $y_{n-1}, y_{n-2}, \dots, y_{n-m}$. Розрахунок починається з $n = m$, тобто з рівняння

$$\frac{a_0 y_m + a_1 y_{m-1} + \dots + a_m y_0}{\tau} = b_0 f_m + b_1 f_{m-1} + \dots + b_m f_0.$$

Як бачимо з рівняння, для початку розрахунків необхідно задати m початкових значень. Значення y_0 визначається початковою умовою задачі (1), а саме покладають $y_0 = u_0$. Величини y_1, \dots, y_{m-1} можна обчислити, наприклад, за методом Рунге-Кутта, або за методом рядів Тейлора. В подальшому будемо вважати, що початкові значення y_0, y_1, \dots, y_{m-1} задані.

З рівняння (2) видно, що на відміну від методу Рунге-Кутта багатокрокові різницеві методи допускають обчислення правих частин тільки в точках основної сітки ω_τ .

Метод (2) називається *явним*, якщо $b_0 = 0$, і відповідно, шукане значення y_n виражається явно через y_{n-1}, \dots, y_{n-m} . Якщо $b_0 \neq 0$, то метод (2) називається *неявним*.

В неявному методі для пошуку y_n потрібно розв'язувати нелінійне рівняння:

$$\frac{a_0}{\tau} y_n - b_0 f(t_n, y_n) = F[y_{n-1}, y_{n-2}, \dots, y_{n-m}],$$

де

$$F[y_{n-1}, y_{n-2}, \dots, y_{n-m}] = \sum_{k=1}^m \left(b_k f_{n-k} - \frac{a_0}{\tau} y_{n-k} \right).$$

Як правило це рівняння розв'язують методом Ньютона, обираючи, наприклад, початкове наближення $y_n^{(0)}$ рівним y_{n-1} .

Коефіцієнти рівняння (2) визначені з точністю до множника. Для уникнення цієї неоднозначності, будемо вважати, що виконується умова:

$$\sum_{k=0}^m b_k = 1, \quad (3)$$

яка означає, що права частина різницевого рівняння (2) апроксимує праву частину диференціального рівняння (1).

В практичному використанні найбільш поширені методи Адамса, які являють собою частинний випадок багатокрокових методів (2), коли похідна $u'(t)$ апроксимується тільки по двох точках, t_n і t_{n-1} , тобто

$$a_0 = -a_1 = 1, \quad a_k = 0, \quad k = \overline{2, m}.$$

Таким чином, *методи Адамса* мають вигляд:

$$\frac{y_n - y_{n-1}}{\tau} = \sum_{k=0}^m b_k f_{n-k}.$$

У випадку, коли $b_0 = 0$, методи Адамса називають *явними*, А при $b_0 \neq 0$, методи Адамса називають *неявними*.

Розглянемо детальніше процедуру побудови методів Адамса. Інтегруємо рівняння (1) по $t \in [t_{n-1}, t_n]$:

$$\frac{u(t_{n-1}) - u(t_n)}{\tau} = \frac{1}{\tau} \int_{t_{n-1}}^{t_n} f(t, u(t)) dt. \quad (4)$$

Замінімо $f(t, u(t))$ на інтерполяційний поліном. Виберемо вузлами інтерполявання точки $t_n, t_{n-1}, \dots, t_{n-m}$. Використаємо багаточлен степеня m за формулою Ньютона по рівновіддалених вузлах, тобто

$$\begin{aligned} f(t, u(t)) &\approx L_m(t) = \\ &= f_n + s \Delta f_{n-1} + \dots + \frac{s(s+1) \cdot \dots \cdot (s+m-1)}{m!} \Delta^m f_{n-m}, \end{aligned}$$

де $s = \frac{t - t_n}{\tau}$, крок τ сталий. В результаті підстановки в (4) $f(t, u(t)) \approx L_m(t)$ отримуємо метод

$$\frac{y_n - y_{n-1}}{\tau} = f_n + \beta_1 \Delta f_{n-1} + \dots + \beta_m \Delta^m f_{n-m}, \quad (4a)$$

де

$$\beta_k = \int_{-1}^0 \frac{s(s+1) \cdot \dots \cdot (s+k-1)}{k!} ds, \quad k = \overline{1, m}$$

$$\Delta f_{n-1} = f_n - f_{n-1}, \quad \Delta^2 f_{n-2} = f_n - 2f_{n-1} + f_{n-2}.$$

Похибка методу на одному кроці

$$R(\tau) = \int_{t_{n-1}}^{t_n} r_m(t) dt = O(\tau^{m+2}),$$

де

$$r_m(t) = f(t) - L_m(t),$$

звідки степінь точності на одному кроці $p = m + 1$.

Формула (4а) називається *формулою Адамса-Мултона* (неявний метод Адамса, інтерполяційний метод Адамса).

Виберемо вузли інтерполювання точки t_{n-1}, \dots, t_{n-m} . Отримаємо багаточлен степеня $m - 1$:

$$f(t, u(t)) \approx L_{m-1}(t) = f_{n-1} + v \Delta f_{n-2} + \dots + \frac{v(v+1) \cdot \dots \cdot (v+m-2)}{(m-1)!} \Delta^{m-1} f_{n-m},$$

де $v = \frac{t - t_{n-1}}{\tau}$. Підставляючи в (4), маємо метод

$$\frac{y_n - y_{n-1}}{\tau} = f_{n-1} + \gamma_1 \Delta f_{n-1} + \dots + \gamma_{m-1} \Delta^{m-1} f_{n-m}, \quad (5)$$

де

$$\gamma_k = \int_0^1 \frac{v(v+1) \cdot \dots \cdot (v+k-1)}{k!} dv, \quad k = \overline{1, m-1}.$$

Формула (5) називається *формулою Адамса-Баушфорта* (явний метод Адамса, екстраполяційний метод Адамса). Похибка цього методу на одному кроці $R(\tau) = O(\tau^{m+1})$. Степінь точності на одному кроці $p = m$.

5. Метод невизначених коефіцієнтів побудови багатокрокових методів

Розглянемо задачу Коші

$$\frac{du}{dt} = f(t, u), \quad t > 0, \quad u(t_0) = u_0. \quad (1)$$

і багатокроковий метод

$$\sum_{k=0}^m \frac{a_k}{\tau} y_{n-k} = \sum_{k=0}^m b_k f_{n-k} \quad (2)$$

Підберемо коефіцієнти a_k, b_k так, щоб досягти найвищої точності методу (2).

Введемо похибку апроксимації $\psi(\tau)$ для формули (2). *Похибкою апроксимації* на розв'язку або нев'язкою різницевого методу (2) називається функція

$$\psi(\tau) = -\sum_{k=0}^m \frac{a_k}{\tau} u_{n-k} + \sum_{k=0}^m b_k f(t_{n-k}, u_{n-k}), \quad (3)$$

де $u_{n-k} = u(t_n - k\tau)$. Її отримують підстановкою точного розв'язку $u(t)$ задачі (1) в різницеве рівняння (2). Розглянемо питання про порядок похибки апроксимації при $\tau \rightarrow 0$ в залежності від вибору коефіцієнтів $a_k, b_k, k = \overline{0, m}$.

Розкладемо функції $u_{n-k} = u(t_n - k\tau)$ в точці $t = t_n$ по формулі Тейлора:

$$u(t_n - \tau k) = \sum_{j=0}^p u^{(j)}(t_n) \cdot \frac{(-k\tau)^j}{j!} + O(\tau^{p+1}). \quad (4)$$

Тут $u^{(j)}$ – j -та похідна. Далі

$$f(t_n - k\tau, u(t_n - k\tau)) = \frac{du}{dt}(t_n - k\tau) = \sum_{j=0}^{p-1} u^{(j+1)}(t_n) \frac{(-k\tau)^j}{j!} + O(\tau^p), \quad k = \overline{1, m}. \quad (5)$$

Підставляючи ці вирази в формулу (3), отримаємо:

$$\begin{aligned} \psi(\tau) &= -\sum_{k=0}^m \frac{a_k}{\tau} \left(\sum_{j=0}^p \frac{(-k\tau)^j u^{(j)}(t_n)}{j!} \right) + \sum_{k=0}^m b_k \left(\sum_{j=0}^{p-1} \frac{(-k\tau)^j u^{(j+1)}(t_n)}{j!} \right) + O(\tau^p) = \\ &= -\sum_{j=0}^p \left(\frac{a_k}{\tau} \cdot \sum_{k=0}^m \frac{(-k\tau)^j u^{(j)}(t_n)}{j!} \right) + \sum_{j=1}^p \left(\sum_{k=0}^m b_k \cdot \frac{(-k\tau)^{j-1} u^{(j)}(t_n)}{(j-1)!} \right) + O(\tau^p). \end{aligned}$$

Після перетворень, приходимо до розкладу:

$$\psi(\tau) = -\left(\sum_{k=0}^m \frac{a_k}{\tau} \right) u(t_n) + \sum_{j=1}^p \left(\sum_{k=0}^m (-k\tau)^{j-1} \left(a_k \cdot \frac{k}{j} + b_k \right) \right) \frac{u^{(j)}(t_n)}{(j-1)!} + O(\tau^p).$$

Звідки видно, що похибка апроксимації має порядок p , якщо виконуються умови:

$$E_0 = \sum_{k=0}^m a_k = 0, \quad E_l = \sum_{k=0}^m \frac{k^{l-1}}{l!} (ka_k + lb_k) = 0, \quad l = \overline{1, p} \quad (6)$$

Разом з умовою нормування $\sum_{k=0}^m b_k = 1$, рівняння (6) утворюють систему з $p+2$ лінійних алгебраїчних рівнянь відносно $2(m+1)$ невідомих $a_0, a_1, \dots, a_m, b_0, b_1, \dots, b_m$.

Умови нормування запишуться у вигляді:

$$\lim_{\tau \rightarrow 0} \sum_{k=0}^m \frac{a_k}{\tau} u_{n-k} = \frac{du}{dt}(t_n), \quad \lim_{\tau \rightarrow 0} \sum_{k=0}^m b_k f(t_{n-k}, u_{n-k}) = f(t_n, u_n).$$

З урахуванням (4), (5) маємо

$$\sum_{k=0}^m b_k = 1, \quad \sum_{k=0}^m ka_k = 1.$$

З рівняння (6) при $l=0$ отримуємо

$$\sum_{k=0}^m ka_k - \sum_{k=0}^m b_k = 0.$$

Тому умови нормування додатково до (6) дають тільки рівняння

$$\sum_{k=0}^m b_k = 1.$$

Для того, щоб система (5) не була перевизначеною, необхідно вимагати, щоб $p \leq 2m$. Ця вимога означає, що порядок апроксимації лінійних m - крокових різницевих методів не може перевищувати $2m$ (неявні методи). Найвищий порядок апроксимації явних методів – $2m - 1$.

Для методів Адамса умови p -го порядку апроксимації мають вигляд:

$$\sum_{k=1}^m k^{l-1} b_k = 1, \quad l = 2, 3, \dots, p, \quad b_0 = 1 - \sum_{k=1}^m b_k.$$

Звідки бачимо, що найвищий порядок апроксимації неявного m - крокового методу Адамса дорівнює $m + 1$, а найвищий порядок апроксимації явного методу Адамса ($b_0 = 0$) дорівнює m .

6. Питання реалізації багатокрокових методів

1). Перша проблема, яка виникає при застосування багатокрокових методів це вибір додаткових початкових умов. Треба знайти $(m - 1)$ додаткове початкове значення y_1, \dots, y_{m-1} . Шляхи вирішення проблеми такі.

а) Можна використовувати *методи Рунге-Кутта* для пошуку цих початкових значень. Але треба, щоб ці методи мали або точність апроксимації p , або точність похибки на кроці p . Недолік такого способу: сама процедура обчислення за методами Рунге-Кутта займає великий об'єм, та й багатокрокові методи об'ємні.

б) Можна використовувати *метод рядів Тейлора*. Знову ж таки потрібно узгоджувати точність p . Функцію розкладають в ряд:

$$u^{(p-1)}(t) = \sum_{k=0}^{p-1} \frac{u^{(k)}(t_0)}{k!} (t - t_0)^k,$$

де індекс $p - 1$ над u вказує кількість членів ряду. Похибка

$$u - u^{(p-1)} = O(\tau^p).$$

2). Друга проблема: реалізація неявних методів

$$y_n = \tau \frac{b_0}{a_0} f(t_n, y_n) + \Phi(y_{n-1}, \dots, y_{n-m}) = \varphi(y_n).$$

а) Можна використовувати для знаходження розв'язку нелінійного рівняння метод простої ітерації:

$$y_n^{(k+1)} = \varphi(y_n^{(k)}),$$

де індекс k означає k -ту ітерацію. Умова збіжності методу простої ітерації $\varphi'(u) \leq q < 1$. Тоді маємо:

$$|\varphi'(u)| = \left| \tau \frac{b_0}{a_0} f'_u(t, u) \right| \leq q < 1 \Rightarrow \tau < \frac{a_0}{b_0 \cdot L}, \quad \text{де } |f'_u(t, u)| \leq L,$$

L – максимум похідної чи стала Ліпшиця. Тобто для збіжності методу необхідно, щоб крок τ був досить малим.

б) Можна використовувати метод Ньютона. Як відомо, умови збіжності методу Ньютона залежать від вдалого вибору початкових умов. Хороше початкове наближення таке: або $y_n^{(0)} = y_{n-1}$, або $y_n^{(0)}$ обчислюється за явною m - кроковою формулою. В результаті умови збіжності методу Ньютона менш жорсткі, ніж у методу простої ітерації.

в) можна використовувати *метод предиктор-коректор*. Запишемо формули неявного методу (C – коректор, P – предиктор):

$$C: y_n = \tau \frac{b_0}{a_0} f(t_n, y_n) + \Phi(y_{n-1}, \dots, y_{n-m}) - \text{неявний } m\text{- кроковий метод, } p = 2m.$$

$$P: \bar{y}_n = \bar{\Phi}(y_{n-1}, \dots, f_{n-1}, \dots) - \text{явний } m\text{- кроковий метод, } p = 2m - 1.$$

Далі виконується така процедура:

$$P: \bar{y}_n = \bar{\Phi}(y_{n-1}, \dots, f_{n-1}, \dots);$$

$$E: \bar{f}_n = f(t_n, \bar{y}_n);$$

$$C: y_n = \tau \frac{b_0}{a_0} f(t_n, \bar{y}_n) + \Phi(y_{n-1}, \dots);$$

$$E: f_n = f(t_n, y_n),$$

де E – підрахування правої частини рівняння (Equation). Схема $PECE$ – це метод *предиктор-коректор*.

Іноді, щоб підвищити точність, використовують схему $PE(CE)^j$. Ця схема аналогічна методу простої ітерації, де j – це максимальна кількість ітерацій.

7. Стійкість методів розв'язання задачі Коші

Важливими є питання стійкості чисельних методів для задач Коші, тобто неперервна залежність розв'язків від збурень початкових даних, похибок заокруглення, тощо.

Чисельний метод розв'язання задачі Коші

$$\frac{du}{dt} = f(t, u), u(t_0) = u_0 \quad (1)$$

називається *стійкий за початковими даними*, якщо для довільної задачі Коші (тобто $\forall f(t, u)$)

$$\forall \varepsilon > 0 \exists \sigma = \sigma(\varepsilon) > 0 \text{ при } |u - \bar{u}_0| < \sigma, \text{ маємо } |y_n - \bar{y}_n| < \varepsilon, \quad \forall n.$$

Чисельний метод називається *нестійким*, якщо

$$\exists \varepsilon > 0, \forall \sigma = \sigma(\varepsilon) > 0, \exists f(t, u), \text{ що при } |u - \bar{u}_0| < \sigma, \text{ маємо}$$

$$|y_n - \bar{y}_n| > \varepsilon, \quad n \geq N(\varepsilon).$$

Перше означення складніше перевіряти ніж друге. Перевіркою умови нестійкості можна вибракувати погані методи на таких класах правих частин:

$$a) f \equiv 0; \text{ б) } f = \lambda u, \lambda < 0; \text{ в) } f = \lambda u, \operatorname{Re} \lambda < 0.$$

Розглянемо кожен випадок.

а) $f \equiv 0$. Якщо застосувати метод Рунге-Кутта, то $y_{n+1} = y_n = \dots = y_0$, а це означає, що малі збурення y_0 приводять до таких самих збурень в y_{n+1} : $\delta = \varepsilon$.

Розглянемо тепер багатокроковий метод, що приводить до різницевого рівняння n -го порядку

$$\sum_{k=0}^m \frac{a_k}{\tau} y_{n-k} = 0 \quad (2)$$

Часткові розв'язки шукаємо у вигляді $y_n = q^n$, $y_{n-k} = q^{n-k}$. Підставимо їх і отримаємо характеристичне рівняння

$$\sum_{k=0}^m a_k q^{m-k} = 0 \quad (3)$$

з коренями q_1, \dots, q_m .

Розв'язок рівняння (2) знаходимо у вигляді:

якщо q – дійсні різні, то $y^n = \sum_{k=0}^m c_k q_k^n$;

якщо q – дійсні кратні, то $y^n = \dots + c_j q_j^n + \tilde{c}_j n q_j^n + \dots$;

якщо q – комплексні, то $y^n = \dots + c_j \rho_j^n \cos(n \operatorname{Im} q_j) + \tilde{c}_j \rho_j^n \sin(n \operatorname{Im} q_j) + \dots$,

де $\rho_j = \operatorname{mod} q_j$, якщо $\exists \rho_j > 1 \Rightarrow n \rightarrow \infty, y_n \rightarrow \infty$, c_k знаходять з початкових умов.

При $|q_k| > 1$ багатокроковий метод – нестійкий.

Кажуть, що багатокроковий метод розв'язання задач задовольняє *кореневій умові*, якщо всі корені характеристичного рівняння (3) лежать в одиничному крузі, а на колі немає кратних коренів, тобто $|q_j| \leq 1$, для $q_j = q_m$, $|q_j| = |q_m| < 1$.

Щоб перевірити, як працює коренева умова розглянемо задачу. Побудуємо явний двокроковий метод найвищого степеня точності. $m = 2$, $p = 3$:

$$\frac{y_n + 4y_{n-1} - 5y_{n-2}}{6\tau} = \frac{2}{3} f_{n-1} + \frac{1}{3} f_{n-2}.$$

Його характеристичне рівняння

$$q^n + 4q^{n-1} - 5q^{n-2} = 0,$$

$$q^2 + 4q - 5 = 0, \quad q_1 = 1, \quad q_2 = -5.$$

Явний двокроковий метод найвищого степеня – нестійкий метод оскільки $q_2 = -5$.

Теорема. Нехай багатокроковий метод задовольняє кореневій умові та права частина задовольняє умову Ліпшиця, $|f(t, u) - f(t, v)| \leq L|u - v|$, тоді має місце оцінка точності

$$|z_n| = |y_n - u(t_n)| \leq M \left(\max_{k=0, m-1} |z_k| + \max_{k=0, n-1} |\psi_k| \right), \quad (4)$$

де ψ_k – похибка апроксимації.

б) $f = \lambda u$, $\lambda < 0$. В цьому випадку розв'язок задачі Коші (1) $u(t) = u_0 e^{\lambda(t-t_0)} \rightarrow 0$ і він є асимптотично стійким відносно 0. Слід очікувати, що чисельний метод має ту ж властивість. Тобто

$$|y_n| < |y_{n-1}| < \dots < |y_0| = |u_0| \quad (5)$$

Для явного методу Ейлера

$$y_{n+1} = y_n + \tau f(F_n, y_n) = y_n + \tau \lambda y_n = (1 + \tau \lambda) y_n = q y_n, \quad q = 1 + \tau \lambda.$$

Тому (5) має місце тільки якщо $|q| < 1$ або $-1 < q < 1$. Права нерівність виконується для довільного τ , а ліва тільки для $\tau < \frac{2}{|\lambda|}$.

Метод розв'язання задачі Коші називається *умовно стійким*, якщо він стійкий при $\tau \leq \tau_0$. Якщо ж він стійкий для $\forall \tau > 0$, то такий метод називається *абсолютно стійким*.

Таким чином явний метод Ейлера – умовно стійкий. Для неявного методу Ейлера:

$$y_{n+1} = y_n + \tau f(t_{n+1}, y_{n+1}) = y_n + \tau \lambda y_{n+1}.$$

Звідси запишемо $y_{n+1} = \frac{1}{1 - \tau \lambda} y_n$, $0 < q = \frac{1}{1 - \tau \lambda} < 1, \lambda < 0$. Таким чином неявний метод Ейлера абсолютно стійкий на тестовому рівнянні.

в) $f = \lambda u$, $\text{Re } \lambda < 0$. Розглянемо систему

$$\frac{d\vec{u}}{dt} = A\vec{u},$$

де A – матриця простої структури, тобто $\exists H: A = H\Lambda H^{-1} = \text{diag}(\lambda_i)$, $\lambda_i = \lambda_i(A)$. Попередня система зводиться до такої:

$$\frac{d\vec{v}}{dt} = \Lambda\vec{v}, \text{ або } \frac{dv_i}{dt} = \lambda_i v_i, \quad i = \overline{1, n}.$$

Тому $v_i = v_i^0 e^{\lambda_i t} = v_i^0 e^{\operatorname{Re} \lambda_i t} (\cos \varphi(t) + j \sin \varphi(t))$, де φ – кут, для якого $\operatorname{tg} \varphi = \frac{\operatorname{Im} \lambda_i}{\operatorname{Re} \lambda_i}$. Дослідимо явний та неявний метод Ейлера на такому тестовому прикладі.

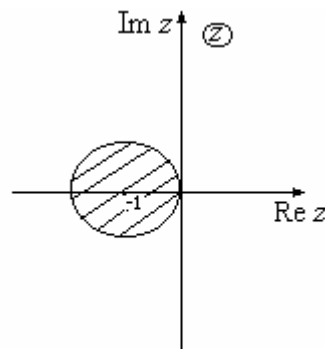


Рис. 1

Явний метод Ейлера: $y_{n+1} = q y_n$, $q = 1 + \tau \lambda = 1 + z$, $z = \tau \lambda = \operatorname{Re} z + j \operatorname{Im} z$. Умова $|q| \leq 1$ виконується для внутрішності кола радіуса 1 з центром в точці $(-1, 0)$ (рис. 1):

$$|q_i| = |1 + \tau \lambda_i| < 1, \forall i.$$

Звідси

$$\tau \leq \frac{2 \operatorname{Re} \lambda_i}{|\lambda_i|^2} = \tau_0,$$

Таким чином явний метод Ейлера умовно стійкий і для тестового рівняння з $\operatorname{Re} \lambda < 0$.

Для неявного методу Ейлера $q = \frac{1}{1 - z}$.

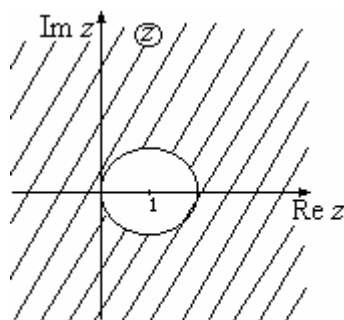


Рис. 2

Умова $|q| \leq 1$ виконується для зовнішності кола радіуса $R = 1$, з центром в точці $z_0 = (1, 0)$ (рис. 2), в той час, як точки $z_i = \tau \lambda_i$ лежать в лівій півплощині \textcircled{Z} . Значить умова стійкості задовольняється для довільних τ . Таким чином

неявний метод Ейлера є абсолютно стійким. Недолік неявного методу Ейлера - низька точність: $p = 1$.

Кажуть, що метод розв'язання задачі Коші є A -стійким, якщо область його стійкості для тестового рівняння $f = \lambda u$, $\operatorname{Re} \lambda < 0$ включає всю півплощину $\operatorname{Re} z < 0$, $z = \tau \lambda$.

Теорема. (Далквіст). Явний лінійний багатокроковий метод не може бути A -стійким. Порядок неявного A -стійкого методу не може бути вищим за 2. Найбільш точний серед цих методів є метод трапецій.

Задача Коші (1) називається *жорсткою* для $t \in [t_0, T]$, якщо $\operatorname{Re} \lambda_i < 0$ та

$$S(t) = \frac{\max_i |\operatorname{Re} \lambda_i(t)|}{\min_i |\operatorname{Re} \lambda_i(t)|} \gg 1,$$

де $\lambda_i = \lambda_i(t)$ власні значення матриці Якобі правої частини рівняння

$$A_n = \left(\frac{\partial f_i}{\partial u_j}(t_n) \right)_{i,j=1}^m.$$

Характерна особливість розв'язків жорсткої задачі Коші є наявність компонент $e^{\lambda_i t}$, які повільно змінюються (коли $|\operatorname{Re} \lambda_i| \approx \min_i |\operatorname{Re} \lambda_i|$) та швидко спадають (коли $|\operatorname{Re} \lambda_i| \approx \max_i |\operatorname{Re} \lambda_i|$). Стійкість явних методів (Рунге-Кутта, явного методу Ейлера) визначається компонентами, які швидко змінюються. В той час як в точному розв'язку вони швидко прямують до нуля і дають малий вклад. Сам розв'язок змінюється повільно разом з повільними компонентами. Але умова стійкості явних методів на всьому проміжку $[t_0, T]$ повинна бути орієнтована на крок $\tau < \frac{2}{\max_i |\lambda_i|}$. Це дуже жорстка умова. Тому

такі системи називаються жорсткими.

Застосування A -стійких методів знімає проблему стійкості: крок τ можна міняти тільки з умови виконання малої похибки на одному кроці. Але точність A -стійких методів не більше 2. Що робити? Послаблюють означення A -стійкості.

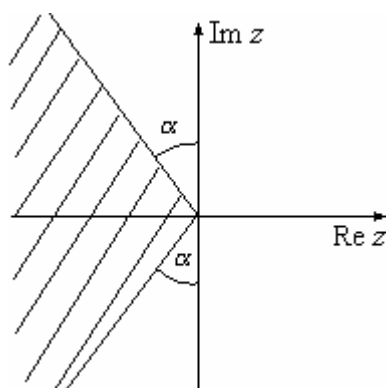


Рис. 3

Метод розв'язання задачі Коші називають $A(\alpha)$ -стійким $0 \leq \alpha \leq \frac{\pi}{2}$, якщо область стійкості методу для тестового рівняння з $f = \lambda u$, $\text{Re} \lambda < 0$, включає частину півплощини $\text{Re} z$ ($z = \tau \lambda$), що лежить в середині між променями, які утворюють з віссю $\text{Im} z$ кути α (рис. 3). Зауважимо, що A -стійкість = $A(0)$ -стійкість !

Найбільш поширеними для розв'язування жорстких систем є методи Куртіса-Хершфільта:

$$\sum_{k=0}^m \frac{a_k}{\tau} y_{n-k} = f(t_n, y_n).$$

Області їх стійкості для різних m приведена на рис. 4:

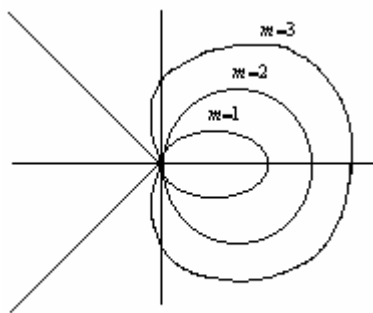


Рис. 4

Це неявні m - крокові методи з $b_0 = 1$, $b_k = 0$ $k = \overline{1, m}$ (інші назви: методи диференціювання назад, чисто неявні методи).

Для жорстких задач треба використовувати реалізацію неявних методів за допомогою ітераційного методу Ньютона з вибором початкового наближення за явною схемою.

Для жорстких задач також застосовують неявні методи Рунге-Кутта. Але для їх реалізації треба розв'язувати систему з $m \times N$ нелінійних рівнянь, m – кількість стадій, N – розмірність системи. Таблиця Батчера цих методів має вигляд

$$\left(\begin{array}{c|ccc} \alpha_1 & \beta_{11} & \cdots & \beta_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_m & \beta_{m1} & \cdots & \beta_{mm} \\ \hline & p_1 & \cdots & p_m \end{array} \right).$$

КОНТРОЛЬНІ ПИТАННЯ З ДИСЦИПЛІНИ

Перелік питань до залікових кредитів

Заліковий кредит 1

1. Етапи розв'язання прикладних задач на ЕОМ.
2. Поняття та приклади математичних моделей.
3. Місце чисельних методів при розв'язуванні прикладних задач.
4. Зв'язок чисельних методів з іншими дисциплінами інформаційного циклу.
5. Математичний пакет MATLAB.
6. Інтерфейс та основні модулі пакету.
7. Робота з матрицями в пакеті MatLab.
8. Використання спеціальних функцій пакету MatLab для чисельного розв'язання алгебраїчних та диференціальних рівнянь.
9. Загальна характеристика чисельних методів.
10. Поняття некоректно поставленої задачі.
11. Стійкість чисельних алгоритмів.
12. Типи похибок, які виникають при розв'язуванні прикладної задачі.
13. Чисельні методи розв'язування нелінійних рівнянь.
14. Алгоритми ізолювання та уточнення коренів.
15. Графічний метод та метод табулювання.
16. Метод ділення навпіл.
17. Блок-схема, оцінка кількості ітерацій.
18. Метод хорд та дотичних.
19. Чисельні методи розв'язування систем лінійних алгебраїчних рівнянь. Метод Гауса.
20. Прямий та обернений хід методу Гауса.
21. Вибір головного елемента.
22. Оцінка кількості арифметичних дій методу Гауса.
23. Блок-схема методу Гауса.
24. Метод ітерацій розв'язування систем лінійних алгебраїчних рівнянь. Достатня умова збіжності методу ітерацій.
25. Метод Гауса-Зейделя. Блок-схема методу Гауса-Зейделя.
26. Перетворення СЛАУ до вигляду, зручному для ітерування.
27. Інтерполювання функцій однієї змінної.
28. Теорема про існування інтерполяційного многочлену.
29. Многочлен Лагранжа.
30. Інтерполювання за рівномірною сіткою вузлів.
31. Многочлен Ньютона. Поняття скінчених різниць.
32. Побудова многочлену Ньютона за таблицею скінчених різниць.
33. Поняття сплайн-інтерполяції.
34. Складання системи лінійних рівнянь для визначення коефіцієнтів сплайн-інтерполяції.
35. Алгоритм прогонки для знаходження коефіцієнтів сплайн-інтерполяції.

36. Задача наближення функцій однієї змінної.
37. Найкраще наближення функцій за допомогою многочленів.
38. Метод найменших квадратів.
39. Формули наближення функцій за допомогою лінійних та квадратичних многочленів.

Заліковий кредит 2

1. Постановка задачі чисельного диференціювання функцій.
2. Некоректність задачі чисельного диференціювання.
3. Формули для знаходження першої та другої похідної.
4. Формули чисельного диференціювання, що не використовують різниці.
5. Регуляризація задачі чисельного диференціювання
6. Точки підвищеної точності формул чисельного диференціювання.
7. Обчислювальна похибка чисельного диференціювання.
8. Вплив збурень на похибку обчислення похідної.
9. Формули Ньютона-Котеса.
10. Постановка задачі чисельного інтегрування.
11. Квадратурні формули прямокутників.
12. Формула трапеції.
13. Інтерполяційний підхід до побудови квадратурних формул.
14. Оцінка похибку квадратурної формули
15. Квадратурна формула Сімпсона.
16. Принцип Рунге оцінки похибки.
17. Екстраполяційна формула Річардсона.
18. Поняття адаптивних квадратурних формул.
19. Квадратурні формули найвищого алгебраїчного степеня точності.
20. Методи обчислення невластних інтегралів.
21. Методи обчислення кратних інтегралів.
22. Метод прямокутників та метод трапецій. Оцінка похибки. Блок-схеми методів. Вибір кроку сітки.
23. Чисельне інтегрування функції однієї змінної. Метод Симпсона. Оцінка похибки. Блок-схема методу.
24. Порівняльний аналіз методів прямокутників, трапецій та Симпсона.

Заліковий кредит 3

1. Чисельні методи розв'язування задачі Коші для звичайних диференційних рівнянь першого порядку. Метод Ейлера.
2. Оцінка локальної та глобальної похибки методу Ейлера.
3. Блок-схема методу Ейлера. Вибір кроку сітки.
4. Чисельні методи розв'язування задачі Коші для звичайних диференційних рівнянь першого порядку.
5. Метод Рунге-Кутта.
6. Оцінка локальної та глобальної похибки методу Рунге-Кутта.
7. Блок-схема методу Рунге-Кутта. Вибір кроку сітки.

8. Узагальнення методу Рунге-Кутта для систем звичайних диференціальних рівнянь першого порядку. Блок-схема алгоритму.
9. Поняття про граничні задачі для звичайних диференціальних рівнянь другого порядку та методи їх розв'язування.
10. Чисельні методи розв'язування двовимірних рівнянь Лапласа та Пуассона.
11. Метод послідовних ітерацій. Блок-схема алгоритму.
12. Чисельні методи розв'язування одновимірних рівнянь теплопровідності.
13. Поняття часового шару. Двошарові скінчено-різницеві схеми. Оцінка похибки апроксимації. Умова стійкості алгоритму. Блок-схема алгоритму.
14. Чисельні методи розв'язування одновимірних хвильових рівнянь.
15. Задача коливання струни та стержня. Трьохшарові скінчено-різницеві схеми.
16. Оцінка похибки апроксимації. Умова стійкості алгоритму. Блок-схема алгоритму.

Перелік питань з курсу

1. Роль чисельних методів при розв'язанні складних задач, що виникають при моделюванні економіко-математичних систем.
2. Поняття тріади «модель-алгоритм-програма».
3. Приклади математичних моделей.
4. Види чисельних методів.
5. Об'єктивна обумовленість похибки чисельної оцінки показників.
6. Поняття про наближені числа.
7. Похибки в арифметичних обчисленнях. Джерела виникнення похибок.
8. Дії та операції над наближеними числами.
9. Поняття середньоквадратичної похибки.
10. Поняття про стійкість, збіжність та коректність чисельних алгоритмів.
11. Етапи розв'язання нелінійного рівняння.
12. Швидкість збіжності метода ділення навпіл.
13. Метод простої ітерації.
14. Метод Ньютона (метод дотичних).
15. Геометричний зміст методу Ньютона.
16. Ітераційні методи розв'язання нелінійних рівнянь.
17. Розв'язання систем рівнянь. Метод Ньютона.
18. Необхідність чисельних методів розв'язування систем лінійних алгебраїчних рівнянь.
19. Алгоритм методу Гауса.
20. Оцінка кількості арифметичних дій методу Гауса.
21. Формули методу квадратних коренів.
22. Методи обчислення визначника та оберненої матриці.
23. Метод прогонки.
24. Поняття обумовленості систем лінійних алгебраїчних рівнянь.

25. Метод простої ітерації.
26. Метод Зейделя розв'язування систем лінійних алгебраїчних рівнянь.
27. Однокрокові (двошарові) ітераційні методи.
28. Метод верхньої релаксації.
29. Методи варіаційного типу.
30. Алгебраїчна проблема власних значень.
31. Постановка задачі інтерполювання.
32. Інтерполяційна формула Лагранжа.
33. Залишковий член інтерполяційного полінома.
34. Багаточлени Чебишова. Мінімізація залишкового члена інтерполяційного полінома.
35. Розділені різниці.
36. Інтерполяційна формула Ньютона.
37. Інтерполювання з кратними вузлами.
38. Збіжність процесу інтерполювання.
39. Кусково - лінійна інтерполяція.
40. Інтерполяційні формули, що використовують центральні різниці.
41. Алгебраїчні поліноми найкращого рівномірного наближення.
42. Тригонометричні поліноми найкращого рівномірного наближення.
43. Середньоквадратичне наближення.
44. Загальна постановка задачі наближення.
45. Поняття елемента найкращого наближення.
46. Найкраще рівномірне наближення.
47. Найкраще середньоквадратичне наближення.
48. Поняття системи ортогональних функцій.
49. Середньоквадратичне наближення періодичних функцій.
50. Метод найменших квадратів.
51. Згладжуючі сплайни.
52. Ортонормовані системи.
53. Інтерполяційні формули Гауса, Бесселя, Стірлінга.
54. Інтерполяція періодичних функцій. Поліном Ерміта.
55. Інтерполяція функцій багатьох змінних.
56. Поняття сплайну степеня m і дефекту k .
57. Кубічний інтерполяційний природній сплайн.
58. Система лінійних рівнянь для визначення коефіцієнтів сплайн-інтерполяції.
59. Параметричні сплайни.
60. Алгоритм реалізації періодичного параметричного кубічного сплайну
61. Ортогональні системи поліномів. Поліноми Лежандра, Якобі.
62. Поліноми Чебишева 1-го та 2-го порядку. Поліноми Лаггера та Ерміта.
63. Некоректність задачі чисельного диференціювання.
64. Формули для знаходження першої та другої похідної.
65. Регуляризація задачі чисельного диференціювання
66. Формули чисельного диференціювання для нерівновіддалених вузлів.

67. Метод невизначених коефіцієнтів в задачі чисельного диференціювання.
68. Постановка задачі чисельного диференціювання функцій.
69. Обчислювальна похибка чисельного диференціювання.
70. Точки підвищеної точності формул чисельного диференціювання.
71. Вплив збурень на похибку обчислення похідної.
72. Вираження різниць через похідні.
73. Формули чисельного диференціювання, що не використовують різниці.
74. Формули Ньютона-Котеса.
75. Постановка задачі чисельного інтегрування.
76. Квадратурні формули прямокутників.
77. Формула трапеції.
78. Інтерполяційний підхід до побудови квадратурних формул.
79. Оцінка похибку квадратурної формули
80. Квадратурна формула Сімпсона.
81. Принцип Рунге оцінки похибки.
82. Екстраполяційна формула Річардсона.
83. Поняття адаптивних квадратурних формул.
84. Квадратурні формули найвищого алгебраїчного степеня точності.
85. Методи обчислення невластних інтегралів.
86. Методи обчислення кратних інтегралів.
87. Постановка задачі чисельного інтегрування звичайних диференціальних рівнянь.
88. Стійкість методів розв'язання задачі Коші.
89. Метод послідовних наближень Пікара.
90. Метод рядів Тейлора.
91. Методи типу Ейлера.
92. Похибка методу Ейлера.
93. Формула методу трапеції інтегрування задачі Коші.
94. Метод Ейлера-Коші.
95. Модифікований метод Ейлера.
96. Поняття методів типу Рунге-Кутта.
97. Методи інтегрування з контролем точності на кроці.
98. Методи Рунге – Кутта – Фельберга.
99. Явний m - етапний метод Рунге-Кутта.
100. Застосування принципу Рунге до методів типу Рунге-Кутта.
101. Постановки крайових задач.
102. Алгоритми методу стрільби.
103. Алгоритм методу пристрілки.
104. Багатокрокові методи розв'язання задачі Коші. Методи Адамса.
105. Формули чисельного інтегрування Чебишева.
106. Апроксимація частинних похідних.
107. Метод Рунге знаходження похибки інтегрування.
108. Обчислення невластних та кратних інтегралів.
109. Метод сіток розв'язання диференціальних рівнянь еліптичного типу.

110. Метод сіток розв'язання диференційних рівнянь гіперболічного типу.
111. Метод сіток розв'язання диференційних рівнянь параболічного типу.
112. Поняття апроксимації, стійкості та збіжності різницевих схем.
113. Різницева апроксимація диференційного рівняння та граничних умов. Поняття коректності та стійкості різницевої схеми.
114. Метод прогонки розв'язання граничних задач для рівнянь в частинних похідних.
115. Математичний пакет Maple та його основні можливості.
116. Бібліотека алгоритмів NAG (Numerical Algorithms Group). Чисельне диференціювання та інтегрування в Maple.
117. Математичний пакет Mathematica та його можливості. Інтерфейс та основні модулі пакету.
118. Виконання чисельних розрахунків в пакеті Mathematica.
119. Математичний пакет MATLAB. Інтерфейс та основні модулі пакету.
120. Виконання чисельних розрахунків в пакеті Mathematica.

ЗАВДАННЯ ДО САМОСТІЙНОЇ ТА ІНДИВІДУАЛЬНОЇ РОБОТИ

1. Завдання до самостійної роботи

Задача 1. Заокруглюючи наступні числа до трьох значущих цифр, визначити абсолютну та відносну похибки наближених чисел:

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

Задача 2. Визначити кількість вірних цифр в числі x , якщо його відносна похибка

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

Задача 3. Визначити, яка рівність точніша:

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

Задача 4. Якою буде відносна похибка, якщо число π наблизити числом 3,14?

Задача 5. Записати число π з п'ятьма вірними значущими цифрами та визначити відносну похибку отриманого наближення.

Задача 6. Знайти $\sqrt{3,02} - \sqrt{3}$ з трьома вірними значущими цифрами.

Задача 7. При вимірі радіуса кола з точністю до 0,5 см, отримали число 14 см. Знайти абсолютну та відносну похибки при обчисленні площі кола.

Задача 8. Кожне ребро куба, виміряне з точністю 0,02 см виявилось рівним 15 см. Знайти абсолютну та відносну похибки при обчисленні площі куба.

Задача 9. Визначити відносну похибку обчислення повної поверхні зрізаного конуса, якщо радіуси його основ R і r та твірна ℓ , виміряні з точністю до 0,01 см, рівні: $R=23,64$ см, $r=17,31$ см, $\ell=10,21$ см.

Задача 10. Обчислити значення функції f . Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних вірними:

$$f=x_1, x_2,$$

де

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

Задача 11. Обчислити значення функції f . Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних вірними:

$$f=x_1, x_2, x_3,$$

де

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

Задача 12. Обчислити значення функції f . Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних вірними:

$$f=x_1 x_2 + x_2 x_3,$$

де

$$x_1=2,104$$
 , $x_2=1,935$, $x_3=0,845$.

Задача 13. Обчислити значення функції f . Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних вірними:

$$f=x_1/x_2$$

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

Задача 14. Обчислити значення функції f . Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних вірними:

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

де $x_1=0,93$, $x_2=1,123$.

Задача 15. Обчислити значення функції f . Знайти абсолютну та відносну похибки результату, вважаючи всі значущі цифри вхідних даних вірними:

$$f = \frac{x_1 + x_2^2}{x_3},$$

де $x_1=3,15$, $x_2=0,831$, $x_3=1,123$.

Задача 16. Оцінити абсолютну та відносну похибки обчислення функції:

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

при $x=2,34 \pm 0,01$, $y=1,25 \pm 0,02$, $z=3,05 \pm 0,02$;

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

при $x=0,757 \pm 0,001$, $y=21,7 \pm 0,05$, $z=1,84 \pm 0,05$;

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

при $x=4 \pm 0,1$, $y=3 \pm 0,05$, $z=1 \pm 0,08$;

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

при $x=1,02 \pm 0,01$, $y=2,35 \pm 0,02$, $z=3,04 \pm 0,01$;

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

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

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

при $x=27,51 \pm 0,001$, $y=21,78 \pm 0,003$, $z=32,5 \pm 0,06$;

$$7) f(x, y, z) = \frac{1}{64} \pi \sqrt{x^4 - y^4},$$

при $x=36,5 \pm 0,01$, $y=26,35 \pm 0,005$, $\pi=3,14$.

Задача 17. Знайти межі абсолютної та відносної похибки аргументів, які дозволяють обчислити з чотирма вірними знаками функції

$$f = \frac{x_1 + x_2^2}{3},$$

де $x_1=2,10415$, $x_2=1,93521$, $x_3=0,84542$.

Задача 18. Оцінити похибку в визначені кута $x=60^\circ$ за п'ятизначною таблицею синусів.

Задача 19. З якою кількістю вірних значущих цифр потрібно взяти значення аргументу x , щоб обчислити значення функції $f(x) = x^3 \sin x$ з точністю до $0,1 \cdot 10^{-5}$?

Задача 20. З якою точністю потрібно обчислити $\sin \frac{\pi}{8}$, щоб відносна похибка обчислення коренів рівняння

$$x^2 - 2x + \sin \frac{\pi}{8} = 0$$

не перевищувала 10^{-3} ?

Задача 21. З якою відносною похибкою треба виміряти висоту $h = 0.5$ м та радіус основи $r = 10$ для того, щоб відносна похибка обчислення об'єма конуса не перевищувала 0,1%?

Задача 22. Обчислити та визначити похибки результату: $x = \frac{ab}{\sqrt[5]{c}}$, якщо $a = 3,55 (\pm 0,01)$; $b = 2,0435 (\pm 0,0004)$; $c = 28,7 (\pm 0,1)$.

Задача 23. Обчислити з точністю до 0,001 $\sqrt{10}$, $\ln(1,1)$, $\sin(0,1)$.

Задача 24. Знайти наближено суму ряду $\sum_{n=1}^{\infty} \frac{1}{n^4}$ з точністю до 0,001.

Задача 25. Відділити корені аналітично та уточнити один з них методом половинного ділення з точністю до 0,01: $3x^4 + 4x^3 - 12x^2 - 5 = 0$.

Задача 26. Відділити корені графічно та уточнити один з них методом хорд з точністю до 0,01: $x - \sin x = 0,25$.

Задача 27. Відділити корені графічно та уточнити один з них методом дотичних з точністю до 0,001: $\operatorname{tg}(0,55x + 0,1) = x^2$.

Задача 28. Відділити корені аналітично та уточнити один з них методом ітерацій: $x^3 - 2x^2 + 7x + 3 = 0$.

Задача 29. Використовуючи схему Гауса, розв'язати систему рівнянь з точністю до 0,0001:

$$\begin{cases} 3,21x_1 - 4,25x_2 + 2,13x_3 = 5,06; \\ 7,09x_1 + 1,17x_2 - 2,23x_3 = 4,75; \\ 0,43x_1 - 1,4x_2 - 0,62x_3 = -1,05. \end{cases}$$

Задача 30. Методом ітерацій розв'язати СЛАР з точністю до 0,001:

$$\begin{cases} x_1 = 0,23x_1 - 0,04x_2 + 0,21x_3 - 0,18x_4 + 1,24; \\ x_2 = 0,45x_1 - 0,23x_2 + 0,06x_3 - 0,8; \\ x_3 = 0,26x_1 + 0,34x_2 - 0,11x_3 + 0,52; \\ x_4 = 0,05x_1 - 0,26x_2 + 0,34x_3 - 0,12x_4 - 1,17. \end{cases}$$

Задача 31. Знайти наближене значення функції при даному значенні аргументу за допомогою інтерполяційного полінома Лагранжа:

а)	х	у	Значення аргументу	б)	х	у	Значення аргументу
	0,43	1,6356	0,702		0,115	8,6573	0,1264
	0,48	1,7323	0,512		0,120	8,2933	0,1315
	0,55	1,8728	0,645		0,125	7,9583	0,1232
	0,62	2,0335	0,736		0,130	7,6489	0,1334
	0,70	2,2285	0,608		0,135	7,3624	0,1285
	0,75	2,3600			0,140	7,09613	

Задача 31. Використовуючи першу чи другу інтерполяційну формулу Ньютона, обчислити значення функції при даних значеннях аргументу.

х	у	Значення аргументу			
		х ₁	х ₂	х ₃	х ₄
1,415	0,88855				
1,420	0,89000	1,4161	1,4625	1,4135	1,470
1,425	0,89064	1,4263	1,4575	1,410	1,4662
1,430	0,89167	1,4302	1,4632	1,423	1,4732
1,435	0,89269				
1,440	0,89370				
1,445	0,89470				
1,450	0,89570				
1,455	0,89668				
1,460	0,89765				

Задача 32. За допомогою інтерполяційної формули Ньютона знайти значення першої а другої похідних при даних значеннях аргументу для функції заданої табличною

х	у	х	у	Значення аргументу
2,4	3,526	3,4	5,100	2,45
2,6	3,782	3,6	4,222	3,15
2,8	3,945	3,8	4,331	4,44
3,0	4,043	4,0	4,507	4,00
3,2	4,104	4,2	4,775	

Задача 33. Обчислити інтеграл $\int_{0,8}^{1,6} \frac{dx}{\sqrt{2x^2+1}}$ за формулою трапецій з трьома десятковими знаками.

Задача 34. Обчислити інтеграл $\int_{1,2}^2 \frac{\ln(x+2)}{x} dx$ за формулою Сімпсона при $n = 8$.

Задача 35. За допомогою рядів обчислити інтеграл $\int_0^1 \frac{dx}{\sqrt{1+x^4}}$ з точністю 0,001.

Задача 36. Обчислити визначений інтеграл $\int_0^{0,1} \frac{\ln(1+x)}{x} dx$ розклавши підінтегральну функцію в ряд, взявши три перші члени ряду. Оцінити похибку.

Задача 37. Використовуючи метод Ейлера, скласти таблицю наближених значень розв'язку диференційного рівняння $y' = 2x^2 - y$, що задовольняє початковій умові $y(0) = 0$ на відрізку $[0;1]$, $h = 0,1$. Всі обчислення вести з точністю до чотирьох десяткових знаків.

Задача 38. Використовуючи метод Рунге-Кута, скласти таблицю наближених значень розв'язку диференційного рівняння $y' = \sin(x - y) + 5(x + y)$, що задовольняє початковій умові $y(0) = 0$ на відрізку $[0; 1]$, $h = 0,1$.

Задача 39. Використовуючи метод сіток, скласти таблицю наближених значень розв'язку диференційного рівняння $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$, що задовольняє початковим умовам $u(x,0) = x(1-x)$ на відрізку $x \in [0,1]$ та крайовим умовам $u(0,t) = u(1,t) = 0$. Розв'язок одержати з кроком $h = 0,1$, визначаючи значення функції $u(x,t)$ з трьома десятковими знаками.

Задача 40. Довести, що алгебраїчна степінь точності квадратурної формули Сімпсона $m = 3$.

Задача 41. Довести, що для $f \in C^4[a, b]$ залишковий член квадратурної формули Сімпсона має представлення:

$$R_2(f) = \frac{1}{24} \int_a^b (x-a) \left(x - \frac{a+b}{2}\right)^2 (x-b) f^{(4)}(\xi) dx = \frac{f^{(4)}(\xi)}{2880} (b-a)^5$$

та вірна оцінка:

$$|R_2(f)| \leq \frac{M_4}{2880} (b-a)^5.$$

Задача 42. Побудувати явний вигляд кубатурної формули, яка дозволяє наближено обчислити I_k по трикутнику D_k , якщо замінити $f(x, y) \approx L_1(x, y)$ на інтерполяційний багаточлен 1-го степеня.

Задача 43. Показати, що для модифікованого методу Ейлера $R(\tau) = O(\tau^3)$, тобто $m = 2$.

Задача 44. Побудувати багаточлен найкращого середньоквадратичного наближення другого степеня для $f(x) = 2x^3 - 1$, $-1 \leq x \leq 1$. Знайти відхилення.

2. Індивідуальні завдання

1. Метод прогонки розв'язання граничних задач для рівнянь в частинних похідних.
2. Метод прямих розв'язання граничних задач для диференціальних рівнянь в часткових похідних.
3. Варіаційні методи розв'язання граничних задач рівнянь математичної фізики. Метод Рітца.
4. Наближені методи розв'язання інтегральних рівнянь.
5. Варіаційно-проекційні методи розв'язання граничних задач рівнянь математичної фізики. Метод Гальоркіна.
6. Чисельні методи розв'язання нелінійних рівнянь в моделях економічної динаміки.
7. Методи покоординатного та найшвидшого градієнтного спуску.
8. Метод Монте-Карло в обчислювальній математиці.
9. Пакет прикладних програм Maple.
10. Пакет прикладних програм Mathematica.
11. Пакет прикладних програм MATLAB.
12. Пакет прикладних програм MathCAD.
13. Оптимізації алгоритмів чисельного розв'язання звичайних диференціальних рівнянь. Управління вибором кроку інтегрування.
14. Варіаційні і проекційні методи апроксимацій.
15. Жорсткі задачі для звичайних диференціальних рівнянь.
16. Чисельні методи розв'язування крайових задач математичної фізики.
17. Методи обчислення кратних інтегралів.
18. Методи обчислення невластних інтегралів.
19. Дослідження збіжності процесу інтерполювання.
20. Написати систему лінійних рівнянь для кубічного інтерполяційного сплайну, якщо $s'(a) = A$, $s'(b) = B$ та розробити алгоритм її розв'язання.
21. Кубічний B_3 -сплайн.
22. Методи побудови багаточленів найкращого рівномірного наближення.
23. Системи ортогональних функцій. Система багаточленів Лежандра.
24. Системи ортогональних функцій. Система багаточленів Чебишова.
25. Системи ортогональних функцій. Система багаточленів Якобі.
26. Системи ортогональних функцій. Система багаточленів Лагерра.
27. Системи ортогональних функцій. Система багаточленів Ерміта.
28. Обчислення визначених інтегралів за формулою Гауса.
29. Методи обчислення визначників.
30. Обчислення елементів оберненої матриці.
31. Визначення власних значень і векторів матриці методом Крилова.
32. Інтерполяція функції за формулою Гауса.
33. Інтерполяція функції за формулою Стірлінга.
34. Інтерполяція функції за формулою Беселя.

Документація ПМК
з курсу “ЧИСЕЛЬНІ МЕТОДИ В ІНФОРМАТИЦІ”

ПОЯСНЮВАЛЬНА ЗАПИСКА

Пакет контрольних завдань з курсу «Чисельні методи в інформатиці», складений згідно програми курсу і освітньо-кваліфікаційних вимог до підготовки бакалаврів за спеціальністю 6.080400 «Інтелектуальні системи прийняття рішень».

Пакет завдань включає в себе 2 частини:

- 1- модульні контрольні роботи для контролю окремих тем;
- 2- комплексну контрольну роботу для підсумкового контролю знань.

Вимоги до виконання завдань та оцінка

Оцінка підсумкового контролю знань студентів – чотирибальна /«відмінно», «добре», «задовільно», «незадовільно»/.

Критерії оцінки знань

Критерії оцінки підсумкового контролю знань студентів базуються на навчальній програмі, робочому плані та найбільш важливих умовах до знань студентів:

1. знання фактів, явищ і вірне, науководостовірне їх пояснення;
2. оволодіння науковими термінами, поняттями, законами, методами, правилами; вміння користуватися ними при вирішенні різних питань і виконанні практичних завдань;
3. максимальна ясність, точність думки;
4. знання повинні мати практичну значимість; студенти повинні вміти безпосередньо застосувати їх на комп'ютері.

Відповіді на теоретичні питання повинні бути повними, логічними, доведеними. Практичні завдання студентів повинні бути виконані з точним дотриманням вказівок викладача.

На оцінку «відмінно» відповідь студента повинна відповідати пунктам 1-4, на «добре» – 1, 2, 4, на «задовільно» 1,4.

МОДУЛЬНІ КОНТРОЛЬНІ РОБОТИ З КУРСУ
«ЧИСЕЛЬНІ МЕТОДИ В ІНФОРМАТИЦІ»

Модульна контрольна робота №1

1. Обчислити та визначити похибки результату: $x = \frac{a^2 b}{\sqrt[5]{c}}$, якщо $a = 3,53$ ($\pm 0,01$); $b = 2,0432$ ($\pm 0,0004$); $c = 28,5$ ($\pm 0,1$).

2. Знайти наближено суму ряду $\sum_{n=1}^{\infty} \frac{1}{n^3}$ з точністю до 0,001.

3. Відділити корені аналітично та уточнити один з них методом половинного ділення з точністю до 0,01: $2x^4 + 5x^3 - 18x^2 - 9 = 0$.

4. Відділити корені графічно та уточнити один з них методом хорд з точністю до 0,01: $x - \sin x = 0,21$.

5. Відділити корені графічно та уточнити один з них методом дотичних з точністю до 0,001: $\text{tg}(0,67x + 0,12) = x^2$.

6. Відділити корені аналітично та уточнити один з них методом ітерацій: $x^3 - 3x^2 + 5x + 2 = 0$.

7. Використовуючи схему Гауса, розв'язати систему рівнянь з точністю

до 0,0001:
$$\begin{cases} 2,31x_1 - 3,15x_2 + 1,15x_3 = 2,26; \\ 8,19x_1 + 2,14x_2 - 1,43x_3 = 3,56; \\ 0,73x_1 - 2,4x_2 - 1,64x_3 = -2,15. \end{cases}$$

8. Методом ітерацій розв'язати СЛАР з точністю до 0,001:

$$\begin{cases} x_1 = 0,323x_1 - 0,05x_2 + 0,25x_3 - 0,28x_4 + 4,14; \\ x_2 = 0,35x_1 - 0,13x_2 + 0,08x_3 - 2,9; \\ x_3 = 0,29x_1 + 0,14x_2 - 0,31x_3 + 0,92; \\ x_4 = 0,15x_1 - 0,36x_2 + 0,24x_3 - 0,23x_4 - 2,17. \end{cases}$$

9. Знайти наближене значення функції при даному значенні аргументу за допомогою інтерполяційного полінома Лагранжа:

	х	у	Значення Аргументу
а)	0,43	1,6356	0,702
	0,48	1,7323	0,512
	0,55	1,8728	0,645
	0,62	2,0335	0,736
	0,70	2,2285	0,608
	0,75	2,3600	
	х	у	Значення аргументу
б)	0,115	8,6573	0,1264
	0,120	8,2933	0,1315
	0,125	7,9583	0,1232
	0,130	7,6489	0,1334
	0,135	7,3624	0,1285
	0,140	7,09613	

10. Використовуючи першу чи другу інтерполяційну формулу Ньютона, обчислити значення функції при даних значеннях аргументу.

x	y	Значення аргументу			
		x ₁	x ₂	x ₃	x ₄
1,415	0,45655				
1,420	0,48500	1,4161	1,4625	1,4135	1,470
1,425	0,489264	1,4263	1,4575	1,410	1,4662
1,430	0,49157	1,4302	1,4632	1,423	1,4732
1,435	0,489249				
1,440	0,489570				
1,445	0,489770				
1,450	0,489575				
1,455	0,489768				
1,460	0,489865				

Модульна контрольна робота №2

1. За допомогою інтерполяційної формули Ньютона знайти значення першої та другої похідних при даних значеннях аргументу для функції, заданої табличною

x	y	x	y	Значення аргументу
2,4	4,523	3,4	6,105	2,45
2,6	4,781	3,6	5,227	3,15
2,8	4,943	3,8	5,338	4,44
3,0	5,045	4,0	5,509	4,00
3,2	5,102	4,2	5,773	

2. За складеною квадратурною формулою прямокутників обчислити

$$\int_0^{\pi/2} \frac{dx}{1 + \cos x}$$

та оцінити точність за допомогою принципу Рунге ($h = \frac{\pi}{4}$).

3. За допомогою формули Сімпсона обчислити

$$\int_0^1 \ln(1+x) dx$$

та оцінити точність.

4. За допомогою рядів обчислити інтеграл $\int_0^1 \frac{dx}{\sqrt{1+x^4}}$ з точністю 0,001.

5. Використовуючи метод Ейлера, скласти таблицю наближених значень розв'язку диференційного рівняння $y' = 3x^2 - y^2$, що задовольняє початковій умові $y(0) = 1$ на відрізку $[0;1]$, $h = 0,1$. Всі обчислення вести з точністю до чотирьох десяткових знаків.

6. Використовуючи метод Рунге-Кута, скласти таблицю наближених значень розв'язку диференційного рівняння $y' = \cos(2x - y) + 1,5(x + y)$, що задовольняє початковій умові $y(0) = -1$ на відрізку $[0;1]$, $h = 0,1$.

7. Розв'язати чисельно задачу математичної фізики

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad (x \in [0,1]), \quad u(0,t) = u_x(1,t) = 0, \quad u(x,0) = \sin \frac{\pi}{2} x.$$

Перелік рекомендованої літератури

Основна література

1. Форсайт Дж., Молер К. Численное решение систем линейных алгебраических уравнений. – М.: Мир, 1969. – 168с.
2. Ляшко И.И., Макаров В.Л., Скоробагатько А.А. Методы вычислений. – Киев: Вища школа, 1977. – 406с.
3. Калиткин Н.Н. Численные методы, – М.: Наука, 1978. – 512с.
4. Форсайт Дж., Малькольм М., Моулер К. Машинные методы математических вычислений. – М.: Мир, 1980. – 280с.
5. Бабенко К.И., Основы численного анализа. - М: Наука, 1986 г.
6. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы: Уч. пособие. - М.: Наука, 1987.
7. Гаврілюк І.П., Макаров В.Л. Методи обчислень. – Київ, Вища школа, в 2-х частинах, 1995.
8. Лященко М.Я., Головань М.С., Чисельні методи. - К: Либідь, 1996 р.
9. Каханер Д., Моулер К., Нэш С. Численные методы и программное обеспечение. - М.: Мир, 1998.
10. Фельдман Л. П., Петренко А. І., Дмитрієва О. А. Чисельні методи в інформатиці. ДДТУ, 2006.
11. Москальков М. М. Чисельні методи. КНУ, 2006.
12. Копченлова Н.В., Марон И.А. Вычислительная математика в примерах и задачах. – М.: Наука. 1972, – 368с.
13. Данилович В. П. Чисельні методи в задачах і вправах: Навч. посібник для студ. спец. "Прикладна математика" - Львів: Державний ун-т "Львівська політехніка", 1995.
14. Фельдман Л. П., Пасько В. П. Чисельні методи та математичні пакети. Лабораторный практикум. ДДТУ, 2006

Додаткова література

1. Березин И.С., Жидков Н.П. Методы вычислений. В 2-х томах. - М.: Государственное из-во физико-математической литературы, 1959.
2. Крылов В.И., Бобков В.В., Монастырский П.Н. Вычислительные методы высшей математики, в 2 томах. – Минск: Вышэйшая школа, 1972.
3. Бахвалов Н.С. Численные методы. – М.:Наука, 1973.–632с.
4. Волков Е.А. Численные методы,– М.: Наука, 1982. – 256с.
5. Жалдак М.І., Рамський Ю.С., Чисельні методи математики - К: Радянська школа, 1984 р.
6. Турчак Л.И. Основы численных методов. - М.: Наука, 1986.
7. Заварыкин В.М., Житомирский В.Г., Лапчик М.П., Вычислительная математика - М.: Просвещение, 1990.
8. Шарапов О.Д., Семьонов Д.Є., Дербенцев В.Д. Дискретний аналіз: Навч.-метод. посібник для самост. вивч. дисц. – К.: КНЕУ, 2002. – 114 с.