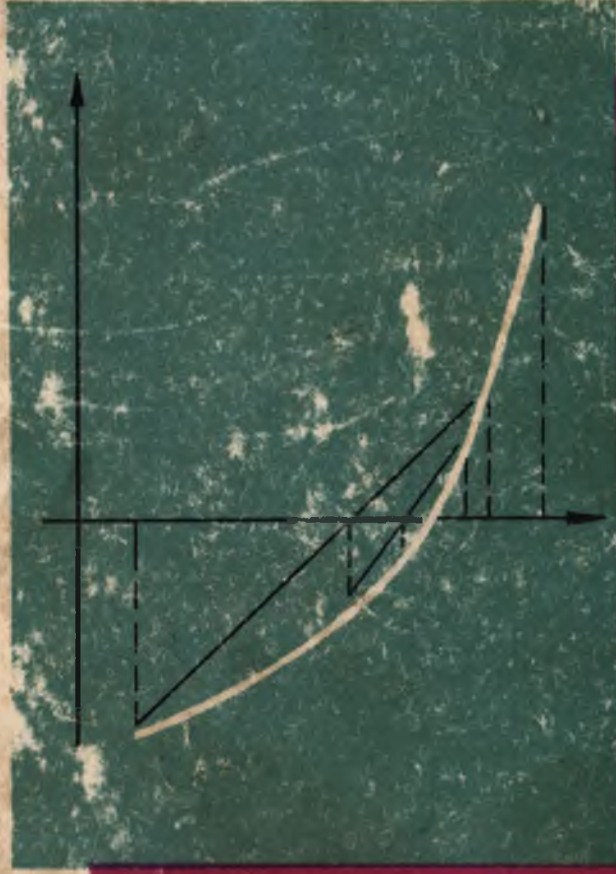


М.Я. ЛЯЩЕНКО, М.С. ГОЛОВАНЬ

ЧИСЕЛЬНІ МЕТОДИ



«Либідь»

М.Я. ЛЯЩЕНКО, М.С. ГОЛОВАНЬ

ЧИСЕЛЬНІ МЕТОДИ

*Затверджено Міністерством освіти України
як підручник для студентів педагогічних
навчальних закладів*

НБ ПНУС



601116

КИЇВ
"ЛИБІДЬ"
1996

*Розповсюдження та тиражування
без офіційного дозволу видавництва заборонено*

Рецензенти: І. Ф. Следзінський, Ю. С. Рамський, професори
Редактор В. М. Кириченко

Лященко М. Я., Головань М. С.
Л99 Чисельні методи: Підручник.—К.: Либідь, 1996.—
288 с.
ISBN 5-325-00652-5.

Відповідно до програми курсу розглянуто чисельні методи розв'язування нелінійних рівнянь і систем лінійних алгебраїчних рівнянь, лінійне програмування, чисельне диференціювання й інтегрування функцій та звичайних диференціальних рівнянь, методи обробки експериментальних даних. Теоретичні положення інтерпретуються геометрично, ілюструються прикладами. Для більшості чисельних методів побудовано алгоритми, які реалізуються на ЕОМ. Алгоритми записано алгоритмічною мовою Бейсік. Лабораторні роботи дадуть можливість закріпити здобуті знання.

Для студентів математичних і фізичних спеціальностей вищих педагогічних навчальних закладів. Може бути корисним учителям математики та інформатики середніх навчальних закладів.

Обчислювальну техніку останніми роками широко застосовують у всіх сферах діяльності людини. Вона стала каталізатором науково-технічного прогресу. Нагальною є потреба в оволодінні знаннями і навичками використання цієї техніки. Саме тому до програм навчальних закладів введено курс "Основи інформатики та обчислювальної техніки".

Обчислювальні машини можна використовувати ефективно лише за умови глибокого знання чисельних методів математики. Проте вищі педагогічні навчальні заклади, які готують учителів середньої школи, й досі не мають підручника з цього предмета. Пропонована книжка є першою спробою заповнити цю прогалину.

У підручнику розглянуто чисельні методи алгебри і математичного аналізу, лінійного програмування і статистичної обробки результатів експериментів. Значну увагу приділено не лише питанням теоретичного характеру (побудові відповідних формул, доведенню збіжності методів, апріорним оцінкам похибок, геометричній інтерпретації методів), а й питанням практичної реалізації їх (вибір обчислювальних схем, ілюстрація численними докладно проаналізованими прикладами, реалізація методів на сучасних ЕОМ, апостеріорні оцінки похибок наближених розв'язків задач). Двосторонні методи чисельного інтегрування функцій і звичайних диференціальних рівнянь, які й нині є предметом математичних досліджень, можна використати для написання курсових і дипломних робіт.

Лабораторні роботи (у 30-ти варіантах кожна) складено так, щоб студенти могли використати подані в підручнику програми, а також програми, побудовані виконавцями самостійно. Поряд з ЕОМ можна скористатися й мікрокалькуляторами "Електроніка" МК-61, МК-52 і МК-72.

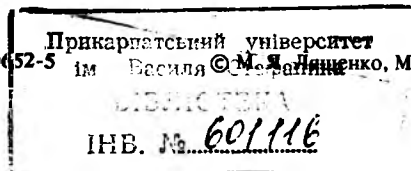
Параграф 1.1 автори написали разом.

Розділи 1,2,4,8 і §§ 3.5, 3.6 та 5.8 написав М. С. Головань, решту — проф. М. Я. Лященко.

Л 1602120000-011
224-96 Без оголошення

ББК 22.193я73

ISBN 5-325-00652-5
Прикарпатський університет
ім. Василя Стефанишина © М. Я. Лященко, М. С. Головань, 1996



Розділ 1. МАТЕМАТИЧНІ МОДЕЛІ І ЧИСЕЛЬНІ МЕТОДИ

§ 1.1. РОЛЬ МАТЕМАТИЧНОГО МОДЕЛЮВАННЯ В РОЗВ'ЯЗУВАННІ ЗАДАЧ НАВКОЛИШНЬОГО СВІТУ

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

На роль математики в суспільстві впливає рівень розвитку математичного апарату науки і ступінь досконалості знань про досліджуваний об'єкт, можливість описати його найсуттєвіші риси і властивості мовою математичних понять, тобто можливість побудувати математичну модель досліджуваного об'єкта.

Математична модель будується на основі деяких спрощень та ідеалізації об'єкта, а тому завжди є наближенням його описом. Але завдяки заміні реального об'єкта його математичною моделлю виникає можливість сформулювати задачу його вивчення як математичну, для розв'язування якої застосовують математичний апарат, що не залежить від природи досліджуваного об'єкта. Математичні моделі — це, як правило, різноманітні рівняння, що є записами законів природи, які управляють досліджуваним нами об'єктом чи явищем.

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

Одна й та сама диференціальна модель може описувати процеси чи явища різної природи. Так, модель, що задається рівнянням виду $y' = ky$, описує і радіоактивний розпад речовини, і зміну атмосферного тиску зі зміною висоти над землею поверхнею, і охолодження тіла

внаслідок конвективного теплообміну з наколишнім середовищем, і зміну струму в електричному колі з індуктивністю при розмиканні, і зростання колонії живих організмів, що перебувають у сприятливих умовах тощо. А модель, що задається рівнянням $y'' + \omega^2 y = 0$, може описувати і коливання кулі, підвішеної до пружини, і коливання математичного маятника, і зміну заряду конденсатора, пластини якого замкнені на індуктивність, і крутильні коливання тощо. Коливання різні (механічні, електромагнітні), але фундаментальна сутність їх однакова, тому вони описуються тією самою математичною моделлю. Отже, одна й та сама математична модель може мати багато реальних прообразів. Проте для описування різних сторін одного явища іноді виникає потреба у використанні іншої математичної моделі. Такі моделі здебільшого взаємно доповнюють одна одну.

Історія науки знає багато прикладів, коли в межах вдало побудованої математичної моделі за допомогою обчислень і міркувань, як кажуть, "на кінчику пера", вдавалося передбачити існування нових фізичних явищ та об'єктів. Так, спираючись на диференціальні моделі, після виконання складних обчислень астрономи Дж. Адамс (Англія) у вересні 1845 р. і У. Левер'є (Франція) у вересні 1846 р. незалежно один від одного дійшли висновку про існування невідомої доти планети і вказали її розміщення. За розрахунками Левер'є астроном І. Галле (Німеччина) знайшов цю планету. Її назвали Нептуном. Це відкриття доводило справедливість геліоцентричної теорії М. Коперника (1473–1543), яка до цього протягом трьохсот років залишалася гіпотезою, значною мірою ймовірною, але все-таки гіпотезою. Аналогічно на "кінчику пера" П. Ловелл (США) передбачив існування дев'ятої планети Сонячної системи, яку за його розрахунками в 1930 р. відкрив К. Томбо (США) і яку назвали Плутоном.

Англійський фізик Дж. Максвелл (1831–1879) у праці "Динамічна теорія електромагнітного поля" (1864 р.) вивів систему диференціальних рівнянь, які кількісно зв'язували між собою напруженості електричного і магнітного полів. З цих рівнянь випливав висновок про існування електромагнітних хвиль і електромагнітна природа світла. Г. Герц (Німеччина) у 1883 р. експериментально добув електромагнітні хвилі, вивчив їх природу, а в 1895 р. О. С. Попов (Росія) використав ці хвилі для передавання і приймання сигналів без проводів на відстань, поклавши початок розвитку техніки радіозв'язку. Це відкриття є основою розвитку радіо, телебачення, радіолокації.

Англійський фізик П. А. М. Дірак у 1928 р. побудував квантовомеханічне рівняння електрона. З його розв'язку випливало існування елементарної частинки з позитивним зарядом, маса якої дорівнює масі електрона. Таку частинку відкрив у 1932 р. фізик К. Д. Андерсон (США) у складі космічних променів і назвав її позитроном. Це була перша античастинка.

Метод математичного моделювання відіграв велику роль у корабле-і авіабудуванні. Розроблені російським ученим, академіком

О. М. Криловим (1863–1945) математичні моделі конструкцій кораблів стали основою вітчизняного кораблебудування.

Літакобудівники давно зіткнулися з негативним, а за певних умов і небезпечним впливом коливань окремих частин літальних апаратів. Коли літальний апарат досягає певної швидкості, яку називають критичною, у ньому виникають самозбудні коливання крил і оперення. Ці коливання, які дістали назву флатера, часто були причиною катастроф літаків у повітрі. Математичну теорію флатера літака розробив у 30-х роках нашого століття на той час ще молодий математик М. В. Келдиш (1911–1978), який пізніше став академіком. Він не лише розробив методи чисельного розрахунку флатера і його моделювання в аеродинамічних трубах, а й запропонував ефективні практичні методи боротьби з ним. Тоді ж М. В. Келдиш вивчив природу іншого явища, яке часто було причиною аварій літаків. Йдеться про явище шими — самозбудні коливання носового колеса триколісного шасі літака під час руління, розбігу і пробігу.

Таких прикладів, коли певна диференціальна модель давала змогу передбачити існування в природі певних явищ чи об'єктів, у науці відомо багато. Всі вони свідчать про прогнозуючу роль математики в науці. Математика стала знаряддям пізнання матеріального світу.

§ 1.2. ОБЧИСЛЮВАЛЬНИЙ ЕКСПЕРИМЕНТ ТА ЙОГО ТЕХНОЛОГІЧНІ ЕТАПИ

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

Тепер виділяють п'ять етапів технологічного циклу обчислювального експерименту: побудова математичної моделі задачі, розробка методу розв'язання математичної моделі, програмування, розрахунки на ЕОМ, аналіз результатів розрахунків і застосування. Схематично технологічний цикл обчислювального експерименту зображено на рис. 1.1.

• На першому етапі вивчають суть досліджуваного явища чи процесу, з'ясовують його склад, закономірності і взаємозв'язки окремих його частин. Виділяють у явищі основні і другорядні фактори. Останніми на початковому етапі досліджень нехтують. Основні фактори і закономірності досліджуваного явища записують у вигляді математичних формул.

Найчастіше — це диференціальні або інтегральні рівняння. Одночасно встановлюють межі застосування побудованої моделі, оскільки жодна математична модель не адекватна реальності. Природа значно багатша і різноманітніша у своїх проявах, ніж будь-які математичні моделі, що їх описують. У прикладних задачах побудова математичної моделі — один з найскладніших і найвідповідальніших етапів обчислювального експерименту. Складність полягає в тому, що він вимагає поєднання математичних і спеціальних знань. Тому над створенням математичної моделі працюють спільно математики і спеціалісти тієї галузі, до якої належить досліджуваний об'єкт.

Коли математичну модель побудовано, приступають до її теоретичного аналізу: досліджують, чи коректно поставлено задачу, чи має вона розв'язок, єдиний він чи ні тощо. Часто розв'язувати складні прикладні задачі починають і без вичерпного аналізу їх математичних властивостей, оскільки для такого аналізу можна затратити час, який перевищує відведений для розв'язування прикладної задачі строк.

• Другий етап обчислювального експерименту пов'язаний з пошуками методу розв'язування математичної моделі. Як правило, прикладні задачі описуються моделями, розв'язання яких не можна знайти у вигляді аналітичних формул. Тоді з відомих чисельних методів підбирають найбільш ефективний для пошуку наближеного розв'язку моделі або ж розробляють новий метод. Метод розв'язування моделі записують у вигляді алгоритму. Для однієї математичної задачі можна скласти, як правило, кілька алгоритмів. Визначення критеріїв для оцінки якості обчислювальних алгоритмів є предметом теорії чисельних методів. Основна її мета — побудова ефективних чисельних методів, які давали б можливість знаходити розв'язок поставленої математичної задачі з наперед заданою точністю з якомога меншими затратами машинного часу. Оцінюючи ефективність чисельного методу, враховують такі його якості, як універсальність, простоту організації обчислювального процесу і контролю його точності, швидкість збіжності, стійкість обчислювального процесу.

Під час обчислювального експерименту важливо використовувати ефективні чисельні методи, оскільки для однієї задачі часто доводиться виконувати велику кількість розрахунків, змінюючи лише деякі параметри.

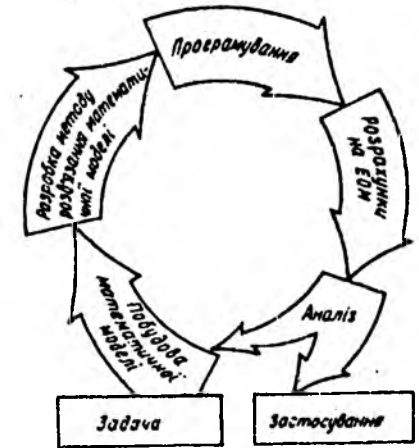


Рис. 1.1

В обчислювальній математиці для обґрунтування пропонуєваних чисельних методів часто використовують найновіші досягнення фундаментальної математики. Але є і зворотний вплив. Результати багатьох досліджень, проведених у рамках обчислювального експерименту, привели до появи нових ідей, понять і навіть математичних теорій.

На третьому етапі обчислювального експерименту складають програму для реалізації розробленого алгоритму на ЕОМ. З розвитком ЕОМ розвивалось і програмування, яке тепер є самостійною наукою із своїми фундаментальними принципами, методами і підходами. Сучасні ЕОМ оснащені різноманітним програмним забезпеченням, яке дає змогу до мінімуму спростити діалог користувача з ЕОМ. Серед програмного забезпечення є значна кількість так званих пакетів прикладних програм, призначених для розв'язування певного класу прикладних задач. Багате й різноманітне програмне забезпечення ЕОМ часто значно спрощує процес програмування алгоритму розв'язування конкретної прикладної задачі.

Четвертий етап обчислювального експерименту полягає в проведенні розрахунків на ЕОМ. Спочатку проводять тестові розрахунки, які потрібні для налагодження програми і перевірки її відповідності реальному процесу математичної моделі. Для тестової задачі підбирають такі значення вхідних даних, щоб можна було надійно судити про достовірність розв'язку цієї задачі. Зіставлення знайденого за допомогою ЕОМ розв'язку з контрольним дає можливість уточнити математичну модель, або переконатися в тому, що вона добре описує досліджуваний об'єкт чи явище. Після цього приступають до повномасштабних, виробничих розрахунків за розробленою програмою. Проведення обчислень на ЕОМ — це копітка, трудомістка і дослідницька робота, яка виконується в діалоговому режимі з ЕОМ. На цьому етапі обчислювального експерименту за допомогою математичної моделі передбачується поведінка досліджуваного об'єкта чи явища в умовах, де натурні експерименти ще не проводились або де вони зовсім неможливі.

Після проведення розрахунків на машині настає п'ятий етап обчислювального експерименту — всебічний аналіз результатів обчислень, на основі якого роблять висновки. Якщо результати надійні, то їх передають замовнику для подальшого використання; якщо результати розрахунків викликають сумнів щодо їх достовірності, тоді математичну модель модифікують, як правило, ускладнюють, і починається новий цикл обчислювального експерименту.

§ 1.3. ДЕЯКІ ЗАСТОСУВАННЯ ОБЧИСЛЮВАЛЬНОГО ЕКСПЕРИМЕНТУ

Математичне моделювання і проведення на його основі обчислювального експерименту — могутній і економічно вигідний засіб для проведення не тільки наукових досліджень, а й виконання найрізноманітніших експериментальних і конструкторських робіт. Так, проведення

обчислювальних експериментів при проектуванні літаків і кораблів економічно вигідніше від створення експериментальних зразків і проведення натурних експериментів, які можуть бути небезпечними для випробувачів. Крім того, обчислювальний експеримент можна провести для дослідження таких ситуацій, які не можна відтворити в реальних умовах. Саме з допомогою обчислювального експерименту вивчають еволюцію Всесвіту, еволюцію життя на Землі або еволюцію яких-небудь популяцій, тобто явища, які ми не здатні спостерігати протягом людського життя.

Зауважимо, що обчислювальний експеримент не може повністю замінити реальний, оскільки він має справу не з самим явищем, а лише з його математичною моделлю. Але обчислювальний експеримент може привести до відкриття нових реальних явищ. Підтвердженням цього є відкриття фізичного ефекту T -шару групою вчених Інституту прикладної математики АН колишнього СРСР під керівництвом А.М.Тихонова і О.А.Самарського (1968 р.). Суть цього ефекту полягає в тому, що в плазмі, яка взаємодіє з магнітним полем, за певних умов можуть виникати зони відносно високої температури. Ці зони називають тепловими шарами або T -шарами. В них зосереджується електричний струм, який розігріває плазму і підтримує високу температуру. В натурному експерименті ефект T -шару вдалося виявити лише через кілька років після його відкриття за допомогою обчислювального експерименту.

Однією з типових задач для обчислювального експерименту є проблема лазерного термоядерного синтезу — невичерпного джерела одержання енергії. Взяти під контроль процес термоядерного синтезу (злиття ядер дейтерію або тритію) і навчитися ним керувати — означає розв'язати енергетичну проблему. У 1962 р. радянські вчені М.Г.Басов і О.М.Крохін висунули ідею запалення термоядерної реакції за допомогою лазерів. Вона полягає в тому, щоб за допомогою лазерного променя за дуже короткий час (10^{-9} с) зарядити високою енергією невелику тверду кульку із замороженої суміші дейтерію і тритію (термоядерну мішень). Енергія лазерного імпульсу створює у термоядерній мішені високу температуру, яка є передумовою для початку термоядерної реакції. Як при цьому протікають процеси в мішені, чи досягаються умови для виникнення термоядерної реакції, як ці умови залежать від параметрів мішені, потужності лазерного випромінювання, тривалості імпульсу? Вивчити ці та інші питання сьогодні в натурному експерименті неможливо хоча б тому, що немає лазерів з відповідними характеристиками. Тому єдиний шлях дослідження — обчислювальний експеримент, за допомогою якого вдалося встановити такі подробиці лазерного термоядерного синтезу, які навряд чи будуть встановлені і виміряні навіть тоді, коли з'являться відповідні лазери.

Завдяки обчислювальному експерименту вдалося розв'язати не тільки багато важливих прикладних задач, а й перевірити деякі гіпотези класичної математики. Відомою топологічною задачею є проблема чо-

тирьох фарб, сформульована ще в 1852 р. лондонським студентом Гутрі. Він виявив, що за допомогою чотирьох фарб можна розфарбувати карту Англії так, що кожні два сусідні графства будуть зафарбовані в різні кольори, і висунув гіпотезу про те, що для розфарбовування будь-якої карти на площині досить чотирьох фарб. Ця гіпотеза була підтверджена тільки в 1976 р. американськими математиками К.Аппелем і В.Хакеном за допомогою ЕОМ.

У сучасних науці і техніці засобами обчислювального експерименту розв'язують задачі, пов'язані з космонавтикою, технологічними процесами, прогнозуванням і управлінням екологічними системами, вивченням гео- і астрофізичних явищ, вивченням фундаментальних проблем в галузі генетики, біотехнології, хімії, фізики та економіки.

§ 1.4. СТРУКТУРА ПОХИБКИ РОЗВ'ЯЗКУ ЗАДАЧІ

Побудувавши математичну модель, намагаються знайти її розв'язок. Для складних прикладних задач, як правило, не існує точного розв'язку у вигляді явних формул або скінченної послідовності арифметичних операцій, кожна з яких виконується точно. Тоді вдаються до чисельних методів — могутнього математичного засобу розв'язування задач. Найпростіші чисельні методи виникли і широко використовувалися задовго до появи ЕОМ. Але є багато прикладних задач, для яких знайти розв'язок без застосування ЕОМ практично неможливо. Сучасні швидкодіючі ЕОМ стали стимулом для розробки нових чисельних методів.

Застосовувати чисельні методи для розв'язування прикладних задач на базі ЕОМ треба обережно, оскільки точність знайденого розв'язку залежить від багатьох факторів. При цьому слід уміти оцінити похибку обчисленого розв'язку.

Похибка розв'язку задачі складається з похибки математичної моделі, неусувної похибки, похибки методу і обчислювальної похибки.

Похибка математичної моделі пов'язана з тим, що модель описує явище наближено, з припущеннями і спрощеннями. Тому треба мати уявлення про точність кінцевого результату, щоб спростити побудову математичної моделі.

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

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

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

Знайти чисельно розв'язок $y(x)$ в усіх точках відрізка $[a; b]$ немож-

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

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

$$y_{i+1} - y_i = hf(x_i, y_i), \quad i = 0, 1, 2, \dots, n-1. \quad (1.2)$$

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

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

Похибку, яку дістають від заміни неперервної моделі дискретною, називають *похибкою дискретизації* (або *похибкою апроксимації*).

Крім похибки дискретизації, існує інший тип похибки чисельних методів. В основі багатьох методів лежить ідея ітераційного процесу, в ході якого будується за певним правилом послідовність наближень до розв'язку задачі. Якщо ця послідовність має границю, коли кількість членів послідовності прямує до нескінченності, тоді ця границя буде розв'язком даної задачі. Але на ЕОМ можна обчислити тільки скінченну кількість членів послідовності. Похибку, спричинену обривом ітераційного процесу, називають *похибкою збіжності*. Похибку методу намагаються звести до величини, яка в кілька разів менша від похибки вхідних даних.

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

Обчислювальні похибки пов'язані з похибками округлення чисел. Обчислення, як ручні, так і на ЕОМ, виконують з певною кількістю значущих цифр. Це вносить у результат похибку округлення, яка нагромаджується в ході обчислень. Похибки округлення можуть по-різному впливати на кінцевий результат. У результаті виконання мільйонів операцій, кожна з яких вносить невелику похибку, сумарна похибка округлень може значно перевищити шуканий результат обчислень. Але в окремих операціях похибки округлень можуть мати різні знаки і частково компенсувати одна одну. Тому, якщо немає систематичних причин, випадкове нагромадження похибок округлення незначне.

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

Обчислювальні похибки виникають і під час перетворення чисел з

однієї системи числення в іншу, якщо основа однієї системи числення не є степенем основи іншої. Це може призвести до того, що в новій системі числення число стане ірраціональним.

Втрата точності може статися і при додаванні до великого числа дуже малих чисел. Для зменшення похибки додавати числа варто в порядку їх зростання. У машинній арифметиці комутативний і дистрибутивний закони алгебри не завжди виконуються. Обчислювальний алгоритм треба будувати так, щоб похибка округлень була значно меншою від усіх інших похибок.

§1.5. ПОНЯТТЯ СТІЙКОСТІ ТА КОРЕКТНОСТІ

Похибки у вхідних даних задачі — неусувні. Обчислювач не може їх зменшити, але мусить знати, як вони впливають на точність кінцевого результату. Одні задачі мають похибку результату такого самого порядку, як і порядок похибки вхідних даних, в інших задачах похибка результату може на кілька порядків перевищувати похибку вхідних даних. Чутливість задачі до неточностей у вхідних даних характеризується поняттям *стійкості*.

Задача називається *стійкою* за вхідними даними, якщо її розв'язок неперервно залежить від вхідних даних, тобто малому приросту Δx вхідної величини відповідає малий приріст Δy шуканого розв'язку. Іншими словами, малі похибки вхідних даних спричинюють малі похибки розв'язку задачі. Якщо ця умова не виконується, то задача вважається *нестійкою* за вхідними даними. Це означає, що навіть незначні похибки вхідних даних можуть привести до як завгодно великих похибок розв'язку, тобто розв'язок може бути зовсім спотворений. Тому застосовувати безпосередньо до таких задач чисельні методи не можна, оскільки похибки округлень при застосуванні методу будуть катастрофічно нагромаджуватись у ході обчислень. Наведемо приклад нестійкої задачі, який належить Уїлкінсону.

Коренями многочлена

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

є числа $x_1 = 1, x_2 = 2, \dots, x_{20} = 20$.

Нехай один з коренів многочлена (1.3) обчислено з незначною похибкою. Наприклад, коефіцієнт -210 при x^{19} замінено на коефіцієнт $-210 + 2^{-23}$ ($2^{-23} \approx 10^{-7}$). Тоді дещо змінений многочлен $\tilde{P}(x) = P(x) - 2^{-23}x^{19}$ має такі округлені до двох десяткових (після коми) знаків корені:

$$x_1 = 1,00; x_2 = 2,00; \dots; x_8 = 8,01; x_9 = 8,92;$$

$$x_{10,11} = 10,10 \pm 0,64i; x_{12,13} = 11,79 \pm 1,65i; \dots;$$

$$x_{18,19} = 19,50 \pm 1,94i; x_{20} = 20,85.$$

Незначна похибка в коефіцієнті -210 даного многочлена викликала суттєво інші значення коренів (десять з них стали комплексними).

Причиною цього є нестійкість задачі, оскільки корені обчислювали з точністю до 11 значущих цифр і похибка округлень незначна.

Іноді під час розв'язування стійкої за вхідними даними задачі нестійким може бути метод її розв'язування.

Нехай треба обчислити інтеграли

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

Інтегруючи за частинами, маємо

$$\begin{aligned} \int_0^1 x^n e^{x-1} dx &= x^n e^{x-1} \Big|_0^1 - \int_0^1 nx^{n-1} e^{x-1} dx = \\ &= 1 - n \int_0^1 x^{n-1} e^{x-1} dx. \end{aligned}$$

Звідси дістанемо

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

Використавши рекурентне співвідношення, обчислимо перші дев'ять інтегралів

$$I_1 = 0,367879, \quad I_2 = 0,264242, \quad I_3 = 0,207274,$$

$$I_4 = 0,170904, \quad I_5 = 0,145480, \quad I_6 = 0,127120,$$

$$I_7 = 0,110160, \quad I_8 = 0,118720, \quad I_9 = -0,0684800.$$

Значення інтеграла I_9 помилкове, оскільки підінтегральна функція $x^9 e^{x-1}$ в усіх точках відрізка $[0; 1]$ невід'ємна. Помилка зумовлена похибкою округлення значення I_1 до шести значущих цифр. Ця похибка наближено дорівнює $4,4 \cdot 10^{-7}$. При обчисленні I_2 вона множить на -3 і т.д. Похибка в I_9 дорівнює $9! \cdot 4,4 \cdot 10^{-7} \approx 0,1601$. Вона спотворила істинне значення I_9 , яке з трьома значущими цифрами дорівнює $0,0916$.

Введемо тепер поняття *коректності* задачі.

Задача називається *коректно поставленою*, якщо для будь-яких вхідних даних з деякого класу існує єдиний і стійкий за вхідними даними її розв'язок. Наведена вище задача обчислення коренів многочлена (1.3) є некоректно поставленою, а обчислення інтегралів (1.4) — коректно поставленою задачею. Прикладом некоректно поставленої задачі є також задача чисельного диференціювання функцій (див. § 5.9).

Для розв'язування некоректно поставлених задач застосовувати класичні чисельні методи не варто, оскільки похибки округлень при розрахунках можуть катастрофічно зростати і призвести до результату, далекого від шуканого розв'язку. Для розв'язування некоректно поставлених задач використовують так звані методи регуляризації, які заміняють дану задачу коректно поставленою.

Розділ 2. ЧИСЕЛЬНІ МЕТОДИ РОЗВ'ЯЗУВАННЯ РІВНЯНЬ З ОДНІЄЮ ЗМІННОЮ

§ 2.1. ПОСТАНОВКА ЗАДАЧІ

Нехай задано рівняння з однією змінною

$$f(x) = 0, \quad (2.1)$$

де функція $f(x)$ визначена і неперервна на деякому проміжку $\langle a; b \rangle$.

Розв'язати рівняння означає знайти множину його коренів, тобто таких значень $x \in \langle a; b \rangle$, при яких рівняння (2.1) перетвориться в тотожність. Корінь рівняння (2.1) називають ще нулем функції $f(x)$. Якщо функція $f(x)$ — алгебраїчний многочлен, то рівняння (2.1) називається алгебраїчним. Якщо функція $f(x)$ містить тригонометричні, показникові або логарифмічні функції, тоді рівняння (2.1) називають трансцендентним.

Знайти точні значення коренів заданого рівняння можна лише для найпростіших функцій $f(x)$: алгебраїчних многочленів не вище четвертого степеня, деяких многочленів степеня $n \geq 5$ і деяких трансцендентних функцій.

Універсальних методів для знаходження точних значень коренів алгебраїчних рівнянь степеня $n \geq 5$ і трансцендентних рівнянь не існує. Крім того, розв'язуючи практичні задачі, часто дістають рівняння з коефіцієнтами, які є наближеними числами. Тоді постановка задачі знаходження точних коренів не має смислу. Тому важливого значення набувають наближені методи знаходження коренів рівняння з достатньою для практики точністю. Задача знаходження коренів рівняння (2.1) вважається розв'язаною, якщо корені обчислені із наперед заданою точністю.

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

Знаходження наближених коренів рівняння (2.1) складається з двох етапів:

1) відокремлення коренів, тобто знаходження досить малих відрізків, на кожному з яких міститься один і тільки один корінь рівняння;

2) обчислення коренів з наперед заданою точністю.

Перший етап називають ще задачею визначення відрізків ізоляції коренів, а другий — уточненням наближених коренів. Перший етап складніший за другий, оскільки для загального випадку немає досить ефективних методів відокремлення коренів. Для знаходження коренів з наперед заданою точністю застосовують методи, які дають можливість уточнювати знайдені наближення коренів.

Зазначимо, що корені рівняння (2.1) можуть бути дійсними і комплексними. Далі розглянуто наближені методи обчислення тільки дійсних коренів рівняння (2.1).

§ 2.2. ВІДОКРЕМЛЕННЯ КОРЕНІВ.

ТЕОРЕМА ПРО ОЦІНКУ ПОХИВКИ НАВЛИЖЕНОГО ЗНАЧЕННЯ КОРЕНЯ

Корінь x^* рівняння (2.1) вважається відокремленим на відрізку $[a; b]$, якщо $x^* \in [a; b]$ і на цьому відрізку дане рівняння не має інших коренів. Щоб відокремити корені рівняння (2.1), треба розбити область визначення даного рівняння на проміжки, на кожному з яких міститься один і тільки один корінь або немає жодного кореня. Відокремлюють корені графічним і аналітичним методами, а також методом послідовного перебору.

Для відокремлення коренів графічним методом будують графік функції $y = f(x)$ і знаходять точки перетину графіка з віссю абсцис та кінці відрізків ізоляції коренів. Часто рівняння (2.1) записують у вигляді $\varphi(x) = g(x)$ і будують графіки функцій $y_1 = \varphi(x)$ і $y_2 = g(x)$, потім знаходять межі, в яких містяться абсциси точок перетину графіків функцій y_1 і y_2 .

Приклад 1. Відокремити корені рівняння $f(x) = 2\cos x - \frac{4}{3\pi}x + 2 = 0$.

Розв'язання. Будемо

будувати графіки функцій $y_1 = 2\cos x$ і $y_2 = \frac{4}{3\pi}x - 2$ (рис. 2.1).

З графіка видно, що дане рівняння має три корені, причому $x_1 = \left[\frac{\pi}{2}; \pi\right]$, $x_2 = \left[\frac{3}{2}\pi; 2\pi\right]$, $x_3 \in \left[2\pi; \frac{5}{2}\pi\right]$. Оскільки $|2\cos x| < 2$ для будь-яких x , а $y_2(x) > 2$ для $x > 3\pi$ і $y_2(x) < -2$ для $x < 0$, то інших коренів дане рівняння не має.

Аналітичний метод відокремлення коренів ґрунтується на теоремах з курсу математичного аналізу. Сформулюємо їх.

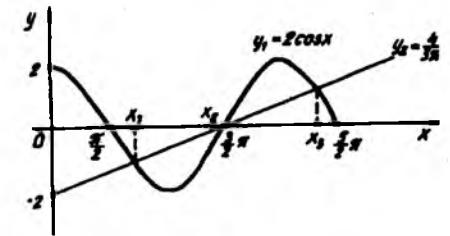


Рис. 2.1

Теорема 1 (теорема існування кореня). Якщо функція неперервна на $[a; b]$ і набуває на кінцях цього відрізка значень протилежних знаків, тобто $f(a)f(b) < 0$, то всередині відрізка $[a; b]$ існує хоча б один корінь рівняння $f(x) = 0$.

Зазначимо, що теорема не дає відповіді на питання про кількість коренів рівняння (2.1), які належать $[a; b]$. При виконанні умов теореми рівняння може мати й кілька коренів. На рис. 2.2 зображено графік функції $y = f(x)$, яка задовольняє усі вимоги теореми 1 і має на $[a; b]$ чотири нулі. У досить малому околі точки x_3 теорему існування кореня застосувати не можна, бо при переході зліва направо через точку x_3 знак функції $f(x)$ не змінюється. Точка x_3 — кратний корінь

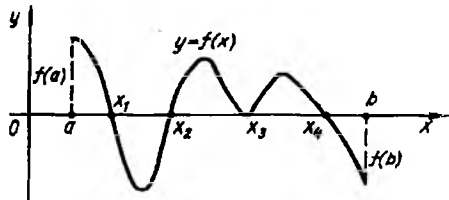


Рис. 2.2

ривняння (2.1) і його не можна відокремити, користуючись теоремою 1. Тому далі вважатимемо, що $f'(x) \neq 0$ для всіх $x \in [a; b]$.

Теорема 2 (теорема існування і єдиності кореня). Якщо функція $f(x)$, неперервна і диференційовна на $[a; b]$, набуває на кінцях цього відрізка значень різних знаків, а похідна $f'(x)$ зберігає сталий знак всередині відрізка $[a; b]$, то рівняння $f(x) = 0$ на цьому відрізку має корінь, причому єдиний.

У відповідності з теоремами 1 і 2 алгоритм відокремлення коренів рівняння (2.1) можна сформулювати так:

1. Знайти область визначення рівняння.
2. Знайти критичні точки функції $f(x)$.
3. Записати інтервали монотонності функції $f(x)$.
4. Визначити знак функції $f(x)$ на кінцях інтервалів монотонності.
5. Визначити відрізки, на кінцях яких функція $f(x)$ набуває значень протилежних знаків.
6. Знайдені відрізки ізоляції коренів при необхідності звузити.

П р и к л а д 2. Відокремити корені рівняння

$$f(x) = x^3 - 2x - 5 = 0.$$

Р о з в ' я з а н н я.

1. Область визначення $X = (-\infty; +\infty)$;
2. $f'(x) = 3x^2 - 2 = 0$. Звідси маємо критичні точки $x_1 = -\frac{\sqrt{6}}{3}$;
 $x_2 = \frac{\sqrt{6}}{3}$;
3. Запишемо інтервали монотонності
 $(-\infty; -\frac{\sqrt{6}}{3})$, $(-\frac{\sqrt{6}}{3}; \frac{\sqrt{6}}{3})$, $(\frac{\sqrt{6}}{3}; +\infty)$;

4. Визначимо знаки функції $f(x)$ на кінцях інтервалів монотонності

$$\lim_{x \rightarrow -\infty} f(x) = -\infty; f\left(-\frac{\sqrt{6}}{3}\right) = \frac{4\sqrt{6} - 45}{9} < 0,$$

$$f\left(\frac{\sqrt{6}}{3}\right) = \frac{-4\sqrt{6} - 45}{9} < 0, \lim_{x \rightarrow +\infty} f(x) = \infty;$$

5. Відрізком ізоляції кореня є проміжок $\left[\frac{\sqrt{6}}{3}; +\infty\right)$.

6. Методом проб звузимо знайдений проміжок ізоляції кореня до одиничної довжини. Оскільки значення $\frac{\sqrt{6}}{3}$ близьке до одиниці, то обчислимо $f(1) = -6 < 0$; $f(2) = -1 < 0$, $f(3) = 16 > 0$. Отже, корінь даного рівняння належить відрізку $[2; 3]$.

Нехай на $[a; b]$ функція $f(x)$ рівняння (2.1) задовольняє умови теореми 1, тоді, застосовуючи ЕОМ, можна відокремити всі корені рівняння (2.1) (крім кратних) методом послідовного перебору коренів. Для цього беруть початкове значення $x = a$, фіксований крок $\Delta x = h$ і обчислюють значення функції $f(x)$ у точках $x_i = a + ih$ ($i = 0, 1, 2, \dots$). На кінцях кожного з відрізків $[a + ih, a + (i + 1)h]$ ($i = 0, 1, 2, \dots$) визначають знак функції $f(x)$. Якщо знаки однакові, тобто $f(a + ih)f(a + (i + 1)h) > 0$, то на відрізку $[a + ih; a + (i + 1)h]$ рівняння (2.1) не має кореня; якщо знаки функції протилежні, то на даному відрізку є корінь рівняння, значення якого $\bar{x} = a + ih + \frac{1}{2}h$ є наближеним значенням

кореня з точністю $\varepsilon = \frac{1}{2}h$, оскільки $|x^* - \bar{x}| \leq \frac{1}{2}h$. Після цього переходять до наступного відрізка. Такий процес продовжують доти, поки правий кінець розглядуваного відрізка не досягне точки b . Цей алгоритм реалізований у програмі 2.1.

Програма 2.1

```

10 ' ----- Метод послідовного -----
20 '           перебору коренів
30 ' -----
40 PRINT "Ввести [a;b], на якому визначена f(x)"
50 INPUT "A="; A : INPUT "B="; B
60 INPUT "Ввести крок h"; H
70 DEF FNF(X)=           : ' Після знака рівності
80 X=A: U=FNF(X)         : ' записують вираз функції f(x).
90 X=X+H                 : ' U, V — значення функції
100 IF X>B THEN 160      : ' на лівому і правому кінцях
110 V=FNF(X)             : ' розглядуваного відрізка.

```

```

120 IF SGN(U)=SGN(V) THEN 140
130 PRINT "На ["X-H";"X"] є корінь"
140 U=V: X=X+H
150 GOTO 100
160 END

```

Зазначимо, що метод послідовного перебору коренів пов'язаний з порівняно великим обсягом обчислень і не гарантує втрат коренів рівняння, особливо тоді, коли корені дуже близькі один до одного, а крок h недостатньо малий. У таких випадках треба провести новий підрахунок з дрібнішим кроком.

Нехай відомо наближене значення \bar{x} кореня рівняння $f(x) = 0$, де функція $f(x)$ — визначена і неперервна на $[a; b]$. Оцінимо похибку наближеного кореня цього рівняння.

Теорема 3. Нехай $x^* \in [a; b]$ — точне значення кореня рівняння $f(x) = 0$, а $\bar{x} \in [a; b]$ — його наближене значення. Якщо існує таке число $m_1 > 0$, що $|f'(x)| \geq m_1$ для будь-яких $x \in [a; b]$, то справедлива оцінка

$$|\bar{x} - x^*| \leq \frac{|f(\bar{x})|}{m_1}. \quad (2.2)$$

Д о в е д е н н я. Застосувавши теорему Лагранжа, маємо

$$f(\bar{x}) - f(x^*) = f'(c)(\bar{x} - x^*),$$

де c лежить між точками \bar{x} і x^* .

Оскільки $f(x^*) = 0$ і $|f'(c)| \geq m_1$, то

$$|f(\bar{x}) - f(x^*)| = |f'(c)(\bar{x} - x^*)| \geq m_1 |\bar{x} - x^*|.$$

Звідси

$$|\bar{x} - x^*| \leq \frac{|f(\bar{x})|}{m_1}.$$

Теорему доведено.

Зазначимо, що за число m_1 у формулі (2.2) можна взяти найменше значення функції $|f'(x)|$ на відрізку $[a; b]$, якщо $f'(x)$ неперервна на $[a; b]$.

Формула (2.2) може давати занадто грубу оцінку точності. На практиці іноді точність знайденого кореня \bar{x} оцінюють близькістю до нуля значення $f(\bar{x})$. При цьому вважають, що коли число

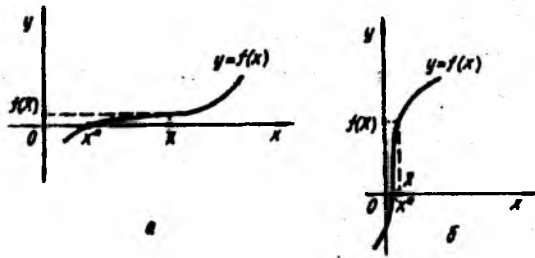


Рис. 2.3

$|f(\bar{x})|$ мале, то \bar{x} — добре наближення x^* , якщо $|f(\bar{x})|$ велике, то \bar{x} — грубе наближення кореня. Такий підхід не завжди правильний. Це видно з рис. 2.3.

П р и к л а д 3. Оцінити абсолютну похибку наближеного кореня $\bar{x} = 2,09$ рівняння $f(x) = x^3 - 2x - 5 = 0$.

Р о з в ' я з а н н я. Обчислимо

$$f(\bar{x}) = f(2,09) = -0,050671.$$

Оскільки $f(2,2) = 1,248$, то $x^* \in [2,09; 2,2]$. Похідна $f'(x) = 3x^2 - 2$ монотонно зростає на відрізку $[2,09; 2,2]$, тому за m_1 можна взяти $f'(2,09) = 11,1043$. Отже,

$$|\bar{x} - x^*| \leq \frac{0,050671}{11,1043} = 0,0046.$$

§ 2.3. УТОЧНЕННЯ КОРЕНЯ МЕТОДОМ ПОДІЛУ ВІДРІЗКА ПОПОЛАМ

Метод поділу відрізка пополам (або метод дихотомії) застосовний для уточнення кореня рівняння $f(x) = 0$ з наперед заданою точністю, допустимою для даної ЕОМ, якщо функція $f(x)$ задовольняє умови теореми 2.

Позначимо через x^* — точне значення кореня рівняння (2.1) на відрізку $[a; b]$, а ε — його граничну абсолютну похибку. Суть методу в тому, що відрізок $[a; b]$ ділять пополам точкою $c = 0,5(a + b)$ і обчислюють $f(c)$. Якщо $f(c) = 0$, то $x = c$ є точним значенням кореня. Якщо $f(c) \neq 0$, але $b - a \leq 2\varepsilon$, то $|x^* - c| \leq \varepsilon$ і значення $x = c$ буде шуканим наближеним коренем. Якщо $f(c) \neq 0$ і $b - a > 2\varepsilon$, тоді розглядають той з двох відрізків $[a; c]$ і $[c; b]$, на кінцях якого функція $f(x)$ набуває значень протилежних знаків. Позначимо цей відрізок $[a_1; b_1]$. На відрізку $[a_1; b_1]$ функція $f(x)$ задовольняє умови теореми 2. Далі відрізок $[a_1; b_1]$ точкою $c_1 = 0,5(a_1 + b_1)$ ділять пополам і міркують так само, як і для відрізка $[a; b]$. В результаті процесу ділення відрізків пополам дістають послідовність вкладених відрізків $[a; b]$, $[a_1; b_1]$, $[a_2; b_2]$, ..., $[a_n; b_n]$, ..., кожен з яких містить точне значення кореня x^* . Довжина відрізка $[a_n; b_n]$ дорівнює $(b - a)/2^n$, тому $\lim_{n \rightarrow \infty} (b_n - a_n) = 0$. Звідси випливає, що для деякого n буде справедлива нерівність $b_n - a_n \leq 2\varepsilon$. Тоді $c_n = \frac{a_n + b_n}{2}$ буде наближеним значенням кореня x^* з точністю ε , тобто $|x^* - c_n| \leq \varepsilon$. При цьому абсолютна похибка знайденого кореня не перевищує ε . Метод дихотомії легко реалізується на ЕОМ, але потребує значного обсягу обчислень, щоб досягти високої точності наближеного кореня.

Програма методу має циклічний характер. У програмі значення a_n, b_n, c_n ($n = 1, 2, \dots$) трактуватимемо як послідовні значення змінних a, b, c . При кожному проходженні циклу виконується серія команд

$c = (a+b)/2$
 якщо $f(a)f(c) < 0$
 то $b := c$
 інакше $a := c$
 все

Повторення команд циклу продовжують доти, поки не виконається одна з умов $f(c) = 0$ або $b - a \leq 2\epsilon$. Програма 2.2 містить алгоритм методу поділу відрізка пополам для рівняння $x^3 - 2x - 5 = 0$.

Програма 2.2

```
10 ' -----Метод поділу відрізка-----
20 ' ----- пополам-----
30 INPUT "Ввести А і В"; A,B
40 INPUT "Ввести точність EPS"; EPS
50 DEF FNF(X)=X^3-2*X-5
60 C=0.5*(A+B) : Y=FNF(C)
70 IF B-A<=2*EPS OR Y=0 THEN 110
80 C=0.5*(A+B) : Y=FNF(C)
90 IF Y*FNF(A)<0 THEN B=C ELSE A=C
100 GOTO 70
110 X=0.5*(A+B)
120 PRINT "Шуканий корінь X="; X
130 END
```

Таблиця 2.1

Точність кореня	Наближений корінь	Кількість ітерацій
10^{-3}	2,0957	9
10^{-6}	2,094559	16
10^{-9}	2,0945514832	29
10^{-12}	2,094551481542	39

У рядку 50 програми записують вигляд лівої частини рівняння (2.1).

У табл. 2.1 подано результати уточнення кореня $x^* \in [2;3]$ рівняння $x^3 - 2x - 5 = 0$ з наперед заданою точністю ϵ методом дихотомії. З таблиці видно,

що збільшення точності шуканого кореня веде до збільшення кількості ітерацій методу для досягнення заданої точності.

§ 2.4. МЕТРИЧНІ ПРОСТОРИ І ПРИНЦИП СТИСКУЮЧИХ ВІДБРАЖЕНЬ

О з н а ч е н н я 1. Непорожня множина X елементів довільної природи називається *метричним простором*, якщо будь-яким двом елементам $x, y \in X$ поставлено у відповідність невід'ємне число $\rho(x, y)$, яке задовольняє такі умови:

- 1) $\rho(x, y) = 0$ тоді і тільки тоді, коли $x = y$;
- 2) $\rho(x, y) = \rho(y, x)$ (аксіома симетрії);
- 3) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ для будь-яких $x, y, z \in X$ (аксіома трикутника).

Елементи множини X називають *точками метричного простору*, а дійсне число $\rho(x, y)$ — *відстанню* між елементами (точками) x і y .

Метричний простір позначають $R = (X, \rho)$. Властивості 1–3 відстані ρ називають *аксіомами метричного простору*.

Відстань ρ між елементами множини X можна означати по-різному. При цьому дістають різні простори. Наведемо кілька прикладів метричних просторів.

1. Множина дійсних чисел з відстанню

$$\rho(x, y) = |x - y| \quad (2.3)$$

є метричним простором, який позначають R або R_1 .

2. Множина n -вимірних векторів з дійсними координатами і відстанню

$$\rho_0(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}, \quad (2.4)$$

де $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n)$, є метричним простором, який називається *n -вимірним евклідовим простором*; його позначають R_n .

Перші дві властивості безпосередньо випливають з формули (2.4). Для доведення третьої властивості скористаємось нерівністю Коші

$$\left(\sum_{i=1}^n a_i b_i \right)^2 \leq \left(\sum_{i=1}^n a_i^2 \right) \left(\sum_{i=1}^n b_i^2 \right),$$

яка правильна для будь-яких дійсних чисел a_i і b_i ($i = 1, 2, \dots, n$).

Справді, поклавши в нерівності Коші $a_i = x_i - y_i$, $b_i = y_i - z_i$, знайдемо

$$\begin{aligned} \rho_0^2(x, z) &= \sum_{i=1}^n (x_i - z_i)^2 = \sum_{i=1}^n ((x_i - y_i) + (y_i - z_i))^2 = \\ &= \sum_{i=1}^n (x_i - y_i)^2 + 2 \sum_{i=1}^n (x_i - y_i)(y_i - z_i) + \sum_{i=1}^n (y_i - z_i)^2 \leq \end{aligned}$$

$$\begin{aligned} &\leq \sum_{i=1}^n (x_i - y_i)^2 + 2 \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \sqrt{\sum_{i=1}^n (y_i - z_i)^2} + \\ &+ \sum_{i=1}^n (y_i - z_i)^2 = \left(\sqrt{\sum_{i=1}^n (x_i - y_i)^2} + \sqrt{\sum_{i=1}^n (y_i - z_i)^2} \right)^2 = \\ &= (\rho_0(x, y) + \rho_0(y, z))^2. \end{aligned}$$

Звідси

$$\rho_0(x, z) \leq \rho_0(x, y) + \rho_0(y, z).$$

3. Множина n -вимірних векторів з відстанню

$$\rho_1(x, y) = \max_{1 \leq i \leq n} |x_i - y_i| \quad (2.5)$$

є метричним простором, який позначають R_n^0 .

Справедливість перших двох аксіом метричного простору очевидна. Доведемо третю аксіому.

Справді, з нерівності

$$|x_i - z_i| \leq |x_i - y_i| + |y_i - z_i| \leq \max_{1 \leq i \leq n} |x_i - y_i| + \max_{1 \leq i \leq n} |y_i - z_i|$$

випливає нерівність

$$\max_{1 \leq i \leq n} |x_i - z_i| \leq \max_{1 \leq i \leq n} |x_i - y_i| + \max_{1 \leq i \leq n} |y_i - z_i|.$$

Це означає, що відстань, яку визначає формула (2.5), задовольняє й третю аксіому метричного простору, тому простір R_n^0 — метричний.

4. Множина n -вимірних векторів з відстанню

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

є метричним простором, який позначають R_n^1 .

Останні три приклади показують, що одна й та сама множина елементів може бути по-різному метризована. В результаті дістають різні метричні простори.

Наявність відстані між кожними двома елементами метричного простору дає можливість ввести поняття границі послідовності метричного простору.

Означення 2. Точку x^* довільного метричного простору R називають *границею послідовності* $x^{(1)}, x^{(2)}, \dots, x^{(k)}, \dots$ точок цього метричного простору і записують $x^{(k)} \rightarrow x^*$ ($k \rightarrow \infty$) або $\lim_{k \rightarrow \infty} x^{(k)} = x^*$, якщо

$$\lim_{k \rightarrow \infty} \rho(x^{(k)}, x^*) = 0.$$

Послідовність $\{x^{(k)}\}$ точок, що має границю, називається *збіжною*.

Відомо, що збіжність послідовності точок $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$ ($k = 1, 2, \dots$) до точки $x = (x_1, x_2, \dots, x_n)$ у просторах R_n, R_n^0, R_n^1 , означає *збіжність за координатами*.

Доведено, що збіжна послідовність точок метричного простору може мати *тільки одну границю*. Відстань $\rho(x, y)$ є неперервною функцією своїх аргументів, тобто якщо $x^{(k)} \rightarrow x_0, y^{(k)} \rightarrow y_0$ коли $k \rightarrow \infty$, то $\rho(x^{(k)}, y^{(k)}) \rightarrow \rho(x_0, y_0)$.

Означення 3. Послідовність точок $\{x^{(k)}\}$ ($k = 1, 2, \dots$) метричного простору R називається *фундаментальною*, якщо для будь-якого числа $\epsilon > 0$ існує таке натуральне число $N(\epsilon)$, що $\rho(x^{(k)}, x^{(m)}) < \epsilon$ для всіх k і $m \geq N(\epsilon)$, тобто

$$\lim_{k, m \rightarrow \infty} \rho(x^{(k)}, x^{(m)}) = 0.$$

Відомо, що кожна фундаментальна послідовність збіжна. Для простору R_1 правильне й обернене твердження. Отже, для цього простору справедливий критерій Коші збіжності послідовності: щоб послідовність дійсних чисел $\{x_n\}$ була збіжною, необхідно і достатньо, щоб вона була фундаментальною.

Критерій Коші справедливий і для будь-якого n -вимірного евклідового простору R_n . Але не в усякому метричному просторі кожна фундаментальна послідовність має границю.

Наприклад, відкритий інтервал $X = (0; 1)$ з відстанню $\rho(x, y) = |x - y|$ між двома точками x і $y \in X$ є метричним простором. Послідовність $\frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots$ будучи фундаментальною в цьому просторі, не має в ньому границі.

Означення 4. Метричний простір R називається *повним*, якщо в ньому кожна фундаментальна послідовність має границю.

Це означає, що метричні простори, в яких справджується критерій Коші, є повними. Незавжди впевнитись, що всі метричні простори, наведені у прикладах 1–4, повні. Доведемо, наприклад, повноту простору R_n .

Нехай послідовність векторів $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}) \in R_n$ ($k = 1, 2, \dots$) — фундаментальна. Тоді для будь-якого $\epsilon > 0$ існує натуральне число $N(\epsilon)$ таке, що

$$\rho(x^{(k)}, x^{(m)}) = \sqrt{\sum_{i=1}^n (x_i^{(k)} - x_i^{(m)})^2} < \epsilon$$

для всіх k і $m > N(\epsilon)$. Звідси для кожного $i = 1, 2, \dots, n$ маємо

$$|x_i^{(k)} - x_i^{(m)}| \leq \rho(x^{(k)}, x^{(m)}) < \epsilon,$$

для усіх $k > N(\epsilon)$ і $m > N(\epsilon)$.

Згідно з критерієм Коші збіжності послідовності дійсних чисел роби-

мо висновок, що для кожного фіксованого $l = 1, 2, \dots$ числа послідовність $\{x_i^{(k)}\}$ збіжна і має деяку границю x_l : $\lim_{k \rightarrow \infty} x_i^{(k)} = x_l$. Числа $x_1, x_2, \dots, x_n \in R_n$. Отже, простір R_n повний.

Нехай R_1 і R_2 — метричні простори; X — певна множина точок простору R_1 . Якщо кожній точці $x \in X$ за певним правилом поставлено у відповідність точку $y \in R_2$, то кажуть, що на множині X задано оператор (відображення) U і записують $y = Ux$. Множину X називають областю визначення оператора U , точку x — аргументом оператора U , y — значенням оператора U , або образом точки x . Нехай оператор U здійснює відображення множини X у простір R_2 . Якщо множина значень оператора збігається з усім простором R_2 , то кажуть, що оператор U відображає множину X на простір R_2 .

Якщо оператор U визначений на множині X точок метричного простору R і відображає множину X у той самий метричний простір, то це записують так: $U: R \rightarrow R$. Окремим випадком такого оператора, заданого на множині дійсних чисел, є функція однієї змінної.

Нехай оператор U відображає метричний простір R у себе.

О з н а ч е н н я 5. Якщо існує таке додатне число $\alpha < 1$, що для будь-яких $x', x'' \in X$ виконується нерівність

$$\rho(Ux', Ux'') \leq \alpha \rho(x', x''),$$

то оператор U називається оператором стиску.

О з н а ч е н н я 6. Точка $x \in X$ називається нерухомою точкою оператора U , якщо

$$Ux = x.$$

Для знаходження нерухомих точок оператора U застосовують метод послідовних наближень. При цьому беруть довільну точку $x_0 \in X$ і послідовно знаходять

$$x_1 = Ux_0, \quad x_2 = Ux_1, \dots, \quad x_k = Ux_{k-1}, \dots$$

В результаті дістають послідовність $x_0, x_1, x_2, \dots, x_k, \dots$, яку називають послідовністю наближень, або ітераційною послідовністю.

Зазначимо, що нерухомі точки оператора U є розв'язками рівняння

$$x = Ux. \quad (2.6)$$

Для операторів стиску справедлива така теорема Банаха про нерухому точку.

Теорема 4 (принцип стискуючих відображень). Якщо оператор стиску U відображає повний метричний простір R у себе, то він має єдину нерухому точку, яку можна дістати методом послідовних наближень при будь-якій початковій точці $x_0 \in R$.

Д о в е д е н н я. Нехай x_0 — довільна фіксована точка простору R

$$x_k = Ux_{k-1} \quad (k = 1, 2, \dots). \quad (2.7)$$

Послідовність $\{x_k\}$ — фундаментальна. Справді,

$$\begin{aligned} \rho(x_{k+1}, x_k) &= \rho(Ux_k, Ux_{k-1}) \leq \alpha \rho(x_k, x_{k-1}) = \\ &= \alpha \rho(Ux_{k-1}, Ux_{k-2}) \leq \alpha^2 \rho(x_{k-1}, x_{k-2}) \leq \dots \leq \\ &\leq \alpha^k \rho(x_1, x_0). \end{aligned}$$

Якщо $m > k$, то, скориставшись аксіомою трикутника, маємо:

$$\begin{aligned} \rho(x_k, x_m) &\leq \rho(x_k, x_{k+1}) + \rho(x_{k+1}, x_m) \leq \\ &\leq \rho(x_k, x_{k+1}) + \rho(x_{k+1}, x_{k+2}) + \rho(x_{k+2}, x_m) \leq \dots \leq \\ &\leq \rho(x_k, x_{k+1}) + \rho(x_{k+1}, x_{k+2}) + \dots + \rho(x_{m-1}, x_m) \leq \\ &\leq (\alpha^k + \alpha^{k+1} + \dots + \alpha^{m-1}) \rho(x_1, x_0) = \\ &= \frac{\alpha^k(1 - \alpha^{m-k})}{1 - \alpha} \rho(x_1, x_0) \leq \frac{\alpha^k}{1 - \alpha} \rho(x_1, x_0). \end{aligned}$$

Для числа $\epsilon > 0$ існує натуральне число $N(\epsilon)$ таке, що для m і $k \geq N(\epsilon)$ виконуватиметься нерівність

$$\rho(x_k, x_m) \leq \frac{\alpha^k}{1 - \alpha} \rho(x_1, x_0) < \epsilon. \quad (2.8)$$

Це означає, що послідовність $\{x_k\}$ фундаментальна.

Оскільки метричний простір R — повний, то існує границя $\lim_{k \rightarrow \infty} x_k = x^*$

і x^* належить R .

Доведемо, що x^* є нерухомою точкою оператора U , тобто $Ux^* = x^*$. Справді, якщо $k \rightarrow \infty$,

$$\rho(Ux^*, x_k) = \rho(Ux^*, Ux_{k-1}) \leq \alpha \rho(x^*, x_{k-1}) \rightarrow 0,$$

тобто Ux^* є границею послідовності $\{x_k\}$ при $k \rightarrow \infty$. В силу єдиності границі маємо $Ux^* = x^*$.

Доведемо, що x^* — єдина нерухома точка.

Припустимо, що x^* і y^* дві різні нерухомі точки оператора U . Тоді

$$\rho(x^*, y^*) = \rho(Ux^*, Uy^*) \leq \alpha \rho(x^*, y^*).$$

Звідси і з нерівності $\alpha < 1$ дістанемо $\rho(x^*, y^*) = 0$. Отже, $x^* = y^*$. Цим доведено єдиність нерухомої точки. Теорему доведено.

Якщо виконуються умови теореми 4, то точки x_k ($n = 1, 2, \dots$), обчислені за формулами $x_k = Ux_{k-1}$, ($k = 1, 2, \dots$), утворюють послідовність $\{x_k\}$ точок простору R , яка збігається до розв'язку x^* рівняння $x = Ux$ при будь-якому початковому $x_0 \in R$.

Перейшовши в нерівності (2.8) до границі при $m \rightarrow \infty$ і фіксованому k , дістанемо

$$\rho(x^*, x_k) \leq \frac{\alpha^k}{1-\alpha} \rho(x_1, x_0). \quad (2.9)$$

Ця нерівність дає оцінку похибки k -го наближення x_k точного розв'язку рівняння (2.6). З неї випливає, що послідовні наближення збігаються до точного розв'язку з швидкістю геометричної прогресії із знаменником α . Нерівність (2.9) дає можливість також для конкретних задач попередньо оцінити кількість ітерацій, необхідних для обчислення x^* з наперед заданою точністю.

Оцінимо відхилення x_k від x^* через відстань між x_k і x_{k-1} . Якщо $m > k$, то

$$\begin{aligned} \rho(x_k, x_m) &\leq \rho(x_k, x_{k+1}) + \rho(x_{k+1}, x_{k+2}) + \dots + \\ &+ \rho(x_{m-1}, x_m) \leq \alpha \rho(x_{k-1}, x_k) + \alpha^2 \rho(x_{k-1}, x_k) + \dots + \\ &+ \alpha^{m-k} \rho(x_{k-1}, x_k) = \frac{\alpha(1-\alpha^{m-k})}{1-\alpha} \rho(x_{k-1}, x_k), \end{aligned}$$

тобто

$$\rho(x_k, x_m) \leq \frac{\alpha(1-\alpha^{m-k})}{1-\alpha} \rho(x_{k-1}, x_k).$$

Перейшовши в цій нерівності до границі при $m \rightarrow \infty$ і фіксованому k , дістанемо

$$\rho(x_k, x^*) \leq \frac{\alpha}{1-\alpha} \rho(x_{k-1}, x_k), \quad (2.10)$$

де $x^* = \lim_{m \rightarrow \infty} x_m$ — нерухома точка.

Нехай треба знайти нерухома точку x^* оператора U з наперед заданою точністю ε . Тоді обчислення x_k за формулами (2.7) (за умови, що послідовність $\{x_k\}$ збіжна) треба продовжувати доти, поки не стане справедливою нерівність

$$\frac{\alpha}{1-\alpha} \rho(x_{k-1}, x_k) \leq \varepsilon,$$

тобто

$$\rho(x_{k-1}, x_k) \leq \frac{1-\alpha}{\alpha} \varepsilon. \quad (2.11)$$

§ 2.5. МЕТОД ІТЕРАЦІЇ

Нехай задано рівняння $f(x) = 0$, де $f(x)$ — неперервна функція. Щоб знайти дійсні корені цього рівняння, замінимо його рівносильним:

$$x = \varphi(x). \quad (2.12)$$

Тепер, щоб розв'язати рівняння (2.12), застосовують *метод послідовних наближень (метод ітерацій)*. Вибирають деяке початкове наближення x_0 і послідовно обчислюють наступні наближення:

$$x_k = \varphi(x_{k-1}), \quad k = 1, 2, \dots \quad (2.13)$$

Якщо послідовність $\{x_k\}$ має границю x^* , тобто $\lim_{k \rightarrow \infty} x_k = x^*$, то ця границя і буде коренем рівняння (2.12). Справді, якщо функція $\varphi(x)$ неперервна, то

$$\lim_{k \rightarrow \infty} x_k = \lim_{k \rightarrow \infty} \varphi(x_{k-1}) = \varphi(\lim_{k \rightarrow \infty} x_{k-1}),$$

звідси $x^* = \varphi(x^*)$.

Збіжність послідовності $\{x_k\}$ забезпечується відповідним вибором функції $\varphi(x)$ і початкового наближення x_0 . Вибираючи по-різному

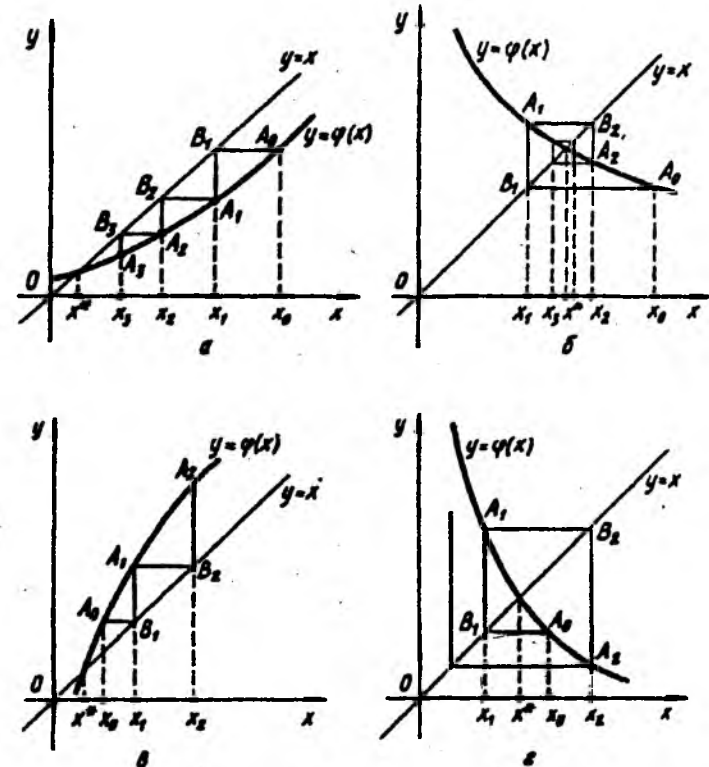


Рис. 2.4

функцію $\varphi(x)$, можна дістати різні ітераційні методи розв'язування рівняння (2.12).

Метод ітерації має простий геометричний зміст. Побудуємо графіки функцій $y = x$ і $y = \varphi(x)$. Абсциса точки перетину графіків цих функцій є коренем рівняння (2.12) (рис. 2.4).

На відріжку $[a; b]$ довільно вибираємо точку x_0 і проводимо через неї пряму, паралельну осі ординат до перетину з кривою $y = \varphi(x)$ в точці $A_0(x_0, \varphi(x_0))$. З точки A_0 проведемо пряму, паралельну осі абсцис, до перетину з прямою $y = x$. В результаті дістанемо точку B_1 з ординатою $\varphi(x_0)$. Спроектувавши точку B_1 на вісь Ox , знаходимо абсцису $x_1 = \varphi(x_0)$. Аналогічно через x_1 проводимо пряму, паралельну осі ординат, до перетину з кривою $y = \varphi(x)$ в точці $A_1(x_1, \varphi(x_1))$. З точки A_1 проводимо пряму, паралельну осі абсцис, до перетину з прямою $y = x$ в точці $B_2(x_2, \varphi(x_1))$, абсциса якої $x_2 = \varphi(x_1)$ і т.д. В результаті спільні абсциси точок A_1 і B_1 , A_2 і B_2 , ... є послідовними наближеннями x_1 , x_2 , x_3 , ... кореня x^* .

На рис. 2.4 зображено випадки: а) $0 < \varphi'(x) < 1$; б) $-1 < \varphi'(x) < 0$; в) $\varphi'(x) > 1$; г) $\varphi'(x) < -1$.

Якщо $|\varphi'(x)| < 1$, то послідовні наближення x_k збігаються до кореня x^* , якщо $|\varphi'(x)| > 1$, то послідовні наближення віддаляються від нього.

Достатні умови збіжності методу ітерацій дає теорема 5.

Теорема 5. Нехай рівняння $x = \varphi(x)$ має корінь x^* і в деякому околі R ($R = \{x: |x - x^*| \leq r\}$) цього кореня функція $\varphi(x)$ задовольняє умову Ліпшиця $|\varphi(x) - \varphi(x')| \leq q|x - x'|$, де $0 < q < 1$. Тоді для будь-якого $x_0 \in R$ послідовність $\{x_k\}$, обчислена за формулою (2.13), збігається до кореня x^* , причому швидкість збіжності характеризується нерівністю

$$|x_k - x^*| \leq q^k |x_0 - x^*|. \quad (2.14)$$

Д о в е д е н н я. Множина точок відрізка R з відстанню $\rho(x, y) = |x - y|$ є повним метричним простором. Функція $\varphi(x)$ відображає відрізок R у самого себе. Справді, якщо $x \in R$, то $y = \varphi(x)$ також належатиме R , оскільки

$$\rho(y, x^*) = |\varphi(x) - \varphi(x^*)| \leq q|x - x^*| < r.$$

Доведемо, що відображення φ — стискуюче.

Справді, для $x \in R$ і $x' \in R$

$$\rho(\varphi(x), \varphi(x')) = |\varphi(x) - \varphi(x')| \leq q|x - x'| = q\rho(x, x'),$$

де $0 < q < 1$.

З принципу стискуючих відображень випливає, що в околі R існує одна і тільки одна нерухома точка $x = x^*$, така, що $x^* = \varphi(x^*)$. Цю точку

можна знайти як границю послідовності $x_k = \varphi(x_{k-1})$, де $k = 1, 2, \dots$, $x_0 \in R$.

З умови Ліпшиця маємо:

$$|x_k - x^*| = |\varphi(x_{k-1}) - \varphi(x^*)| \leq q|x_{k-1} - x^*|,$$

тобто

$$|x_k - x^*| \leq q|x_{k-1} - x^*| \leq q^2|x_{k-2} - x^*| \leq$$

$$\leq \dots \leq q^k|x_0 - x^*|.$$

Теорему доведено.

Отже, послідовність $\{x_k\}$ збігається до кореня x^* із швидкістю геометричної прогресії із знаменником q . Чим менше число q , тим швидше збігається послідовність $\{x_k\}$, тобто тим менше число кроків треба виконати для того, щоб дістати наближене значення кореня із наперед заданою граничною абсолютною похибкою ε . Швидкість збіжності залежить також від вибору початкового наближення x_0 . Чим ближче до кореня x^* вибрано x_0 , тим швидше буде знайдено результат.

Зазначимо, що умова Ліпшиця з константою $0 < q < 1$ буде виконана для функції $\varphi(x)$ на відріжку R , якщо на R існує $\varphi'(x)$, причому

$$|\varphi'(x)| \leq q < 1, \quad x \in R. \quad (2.15)$$

При виконанні умов теореми 5 метод ітерацій є самокорегуючим методом, тобто він виправляє помилки, допущені при обчисленні наближень x_k (якщо вони не виходять за межі відрізка R), оскільки помилкові значення можна розглядати як нові початкові наближення. Тому при ручних обчисленнях не треба весь час виконувати ітерації з великою кількістю значущих цифр. Точність обчислень треба поступово збільшувати з наближенням до кореня.

Поклавши $\alpha = q$ в оцінках (2.9) і (2.10) методу послідовних наближень, знаходимо оцінки методу ітерацій:

$$|x_k - x^*| \leq \frac{q^k}{1 - q} |x_1 - x_0|, \quad (2.16)$$

$$|x_k - x^*| \leq \frac{q}{1 - q} |x_k - x_{k-1}|. \quad (2.17)$$

З оцінки (2.17) випливає, що для знаходження кореня рівняння (2.12) з точністю ε процес ітерації слід продовжувати доти, поки для двох послідовних наближень x_{k-1} і x_k не стане справедливою нерівність

$$|x_k - x_{k-1}| \leq \frac{1 - q}{q} \varepsilon. \quad (2.18)$$

Якщо $0 < q \leq 0,5$ або $-1 < \varphi'(x) < 0$ в околі кореня x^* , то виконання нерівності

$$|x_k - x_{k-1}| < \varepsilon$$

означає, що наближення x_k знайдено з точністю ε , і процес ітерацій можна закінчити.

Алгоритм методу ітерацій легко реалізується на ЕОМ. Програма 2.3 містить алгоритм методу ітерацій для рівняння $x = \sqrt[3]{2x + 5}$.

Програма 2.3

```

10 '----- Метод ітерацій -----
20 '      для рівняння X = F(X)
30 '-----
40 DEF FNF(X)=(2*X+5)^(1/3)
50 INPUT "Ввести точність наближеного значення кореня"; EPS
60 INPUT "Ввести сталу Q"; Q
80 INPUT "Ввести початкове наближення X0"; X0
90 EPS=EPS*(1-Q)/Q
100  X1=FNF(X0)
110  IF ABS(X1-X0) <= EPS THEN 130
120  X0=X1: GOTO 100
130 PRINT "Шуканий корінь X="; X1
140 END

```

Рівняння (2.1) до рівносильного рівняння (2.12) перетворюють так, щоб в околі кореня x^* виконувалась нерівність (2.15). Розглянемо деякі загальні прийоми такого перетворення.

Нехай корінь x^* рівняння (2.1) лежить на відрізку $[a; b]$. Замінімо рівняння (2.1) рівносильним йому

$$x = x - \lambda f(x) = \varphi(x), \quad \lambda \neq 0. \quad (2.21)$$

Підберемо сталу λ так, щоб в околі $[a; b]$ кореня x^* виконувалась нерівність

$$|\varphi'(x)| = |1 - \lambda f'(x)| \leq q < 1.$$

Звідси маємо

$$-1 < 1 - \lambda f'(x) < 1,$$

або

$$0 < \lambda f'(x) < 2.$$

Якщо знак $f'(x)$ на $[a; b]$ не змінюється, то стала λ повинна мати однаковий знак з $f'(x)$ і задовольняти умову

$$\lambda \leq \frac{2}{\max_{x \in [a; b]} |f'(x)|}. \quad (2.22)$$

При такому виборі λ нерівність (2.15) виконуватиметься. Іноді рівняння (2.1) замінюють рівносильним йому

$$x = x - \frac{1}{\mu} f(x), \quad \mu \neq 0. \quad (2.23)$$

Сталу μ вибирають так, щоб в околі $[a; b]$ кореня x^* справджувалась нерівність

$$|\varphi'(x)| = |1 - \frac{1}{\mu} f'(x)| < 1,$$

тобто виконувались умови:

$$-1 < 1 - \frac{1}{\mu} f'(x) < 1,$$

або

$$0 < \frac{1}{\mu} f'(x) < 2.$$

Остання подвійна нерівність виконуватиметься, якщо

$$\frac{f'(x)}{\mu} > 0 \quad \text{і} \quad |\mu| \geq \frac{1}{2} \max_{x \in [a; b]} |f'(x)|. \quad (2.24)$$

Таким чином, якщо на відрізку $[a; b]$ функція $f'(x)$ зберігає знак і обмежена, то завжди можна вказати число μ того самого знака, що й $f'(x)$, яке задовольнятиме нерівності (2.24) і цим самим забезпечуватиме виконання нерівності (2.15) для рівняння (2.23).

Якщо рівняння $x = \varphi(x)$ таке, що в околі $[a; b]$ кореня x^* має місце нерівність

$$|\varphi'(x)| \geq \nu > 1,$$

то ітераційний процес буде розбіжний. Тоді таке рівняння замінюють рівносильним йому $x = \psi(x)$, де $\psi(x)$ — функція, обернена до $\varphi(x)$. Для останнього рівняння метод ітерацій буде збіжним, оскільки

$$|\psi'(x)| = \left| \frac{1}{\varphi'(\psi(x))} \right| \leq \frac{1}{\nu} = q < 1.$$

П р и к л а д 4. Перетворити рівняння $f(x) = x^3 - 2x - 5 = 0$ до вигляду $x = \varphi(x)$, якщо його корінь $x^* \in [2; 3]$.

Р о з в ' я з а н н я. Застосуємо послідовно всі три способи.

1-й спосіб. Оскільки $f'(x) = 3x^2 - 2$ неперервна і монотонна на відрізку $[2; 3]$, то

$$\max_{x \in [2;3]} |f'(x)| = f'(3) = 25.$$

У відповідності з (2.22) знаходимо $\lambda < \frac{2}{25} = 0,08$. Тоді можна покласти, наприклад $\lambda = 0,078$, і шукане рівняння матиме вигляд

$$x = -0,078x^3 + 1,156x + 0,39.$$

При цьому для усіх $x \in [2;3]$ справджуватиметься нерівність

$$|\varphi'(x)| \leq 0,95.$$

2-й спосіб. Оскільки $f'(x) > 0$ на $[2;3]$, то у відповідності з (2.24) маємо

$$\mu > 0, \mu > \frac{25}{2}.$$

Наприклад, можна взяти $\mu = 13$. Тоді дане рівняння набере вигляду:

$$x = -\frac{1}{13}x^3 + \frac{15}{13}x + \frac{5}{13}.$$

При цьому

$$\varphi'(x) = -\frac{3}{13}x^2 + \frac{15}{13}.$$

Легко впевнитися, що на відрізку $[2;3]$

$$|\varphi'(x)| \leq 12/13 < 1.$$

3-й спосіб. Якщо дане рівняння записати у вигляді

$$x = \frac{1}{2}x^3 - \frac{5}{2},$$

то

$$\varphi(x) = \frac{1}{2}x^3 - \frac{5}{2}.$$

На відрізку $[2;3]$ похідна

$$\varphi'(x) = \frac{3}{2}x^2 > 1$$

і умова збіжності методу ітерацій не виконується. Для функції $\varphi(x)$ оберненою буде функція

$$\psi(x) = \sqrt[3]{2x + 5}$$

і заміна рівняння $x = \varphi(x)$ рівносильним йому рівнянням $x = \psi(x)$ приведе до рівняння

$$x = \sqrt[3]{2x + 5}.$$

Для нього

$$\psi'(x) = 2/(3\sqrt[3]{(2x + 5)^2})$$

і на відрізку $[2;3]$

$$\psi'(x) \leq \frac{2}{3\sqrt[3]{81}} < 0,154.$$

З розглянутих вище трьох варіантів рівнянь $x = \varphi(x)$ третій найбільш доцільний, оскільки має найменше значення q .

У табл. 2.2 подано результати розв'язування методом ітерацій кожного з трьох рівнянь, до якого було перетворене рівняння з прикладу 4.

Таблиця 2.2

Рівняння	Точність	x_0	Кількість ітерацій	Наближений корінь
$x = -0,078x^3 + 1,156x + 0,39;$ $q = 0,95$	10^{-6}	2	8	2,094551
	10^{-9}	2	11	2,094551481
	10^{-6}	3	10	2,094551
	10^{-9}	3	13	2,094551481
$x = -\frac{1}{13}x^3 + \frac{15}{13}x + \frac{5}{13};$ $q = 0,93$	10^{-6}	2	8	2,094551
	10^{-9}	2	11	2,094551481
	10^{-6}	3	10	2,094551
	10^{-9}	3	13	2,094551481
$x = \sqrt[3]{2x + 5};$ $q = 0,154$	10^{-6}	2	6	2,094551
	10^{-9}	2	9	2,094551481
	10^{-6}	3	7	2,094551
	10^{-9}	3	10	2,094551481

Порівнявши результати обчислень, побачимо, що кількість ітерацій методу залежить від наперед заданої точності наближеного кореня, параметра q , який характеризує величину модуля похідної $\varphi'(x)$ в околі кореня рівняння, а також вибору початкового наближення. Результати чисельного експерименту повністю узгоджуються з викладеними вище теоретичними висновками.

Вибір функції $\varphi(x)$ рівняння (2.12) повинен забезпечувати не тільки збіжність ітераційного процесу, а й його ефективність. При цьому критерієм ефективності може бути мінімальне число арифметичних операцій, необхідних для знаходження розв'язку, мінімальний час розв'язування задачі на ЕОМ тощо. Найчастіше за критерієм ефектив-

ності ітераційних методів розв'язування нелінійних рівнянь беруть порядок збіжності ітераційного процесу.

Нехай для деякого ітераційного процесу справедлива оцінка

$$|x_{k+1} - x^*| \leq C |x_k - x^*|^\alpha, \quad (2.25)$$

де $C = C(q)$ — деяка функція від q , обмежена зверху, а число α — порядок (ступінь, показник) збіжності ітераційного процесу.

З формули (2.25) випливає, що чим більший порядок збіжності, тим швидше збігатиметься ітераційна послідовність, якщо значення x_k досить близькі до кореня x^* . Якщо $\alpha = 1$, то вважають, що ітераційний метод збігається лінійно, якщо $\alpha = 2$, то ітераційний метод збігається квадратично і т.д.

Для методу ітерації маємо:

$$|x_{k+1} - x^*| = |\varphi(x_k) - \varphi(x^*)| \leq q |x_k - x^*|.$$

Звідси видно, що метод ітерації має лінійну збіжність. У § 2.6 показано, як побудувати функцію $\varphi(x)$, щоб підвищити порядок збіжності ітераційного процесу.

§ 2.6. МЕТОД НЬЮТОНА І ЙОГО МОДИФІКАЦІЇ

Нехай рівняння $f(x) = 0$ на відрізку $[a; b]$ має ізолюваний корінь x^* , тобто $f(a)f(b) < 0$, а функції $f(x)$ і $f'(x)$ неперервні і зберігають знак на $[a; b]$.

Нехай x_k — k -е наближення кореня. Розкладемо $f(x)$ в ряд Тейлора в околі точки x_k

$$f(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2}f''(x_k)(x - x_k)^2 + \dots$$

Замість рівняння $f(x) = 0$ розглядатимемо рівняння

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

яке враховує тільки лінійну відносно $x - x_k$ частину ряду Тейлора. Розв'язавши його відносно x , дістанемо

$$x = x_k - \frac{f(x_k)}{f'(x_k)}.$$

Взявши знайдене значення x за наступне наближення, матимемо

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, 2, \dots \quad (2.26)$$

Формула (2.26) визначає *метод Ньютона*. Він має просту геометричну інтерпретацію. Значення x_{k+1} є абсцисою точки перетину дотичної $y - f(x_k) = f'(x_k)(x - x_k)$ до кривої $y = f(x)$ в точці $(x_k, f(x_k))$

(рис. 2.5). Тому метод Ньютона називають ще *методом дотичних*. З рис. 2.5 видно, що послідовні наближення збігаються до кореня x^* монотонно.

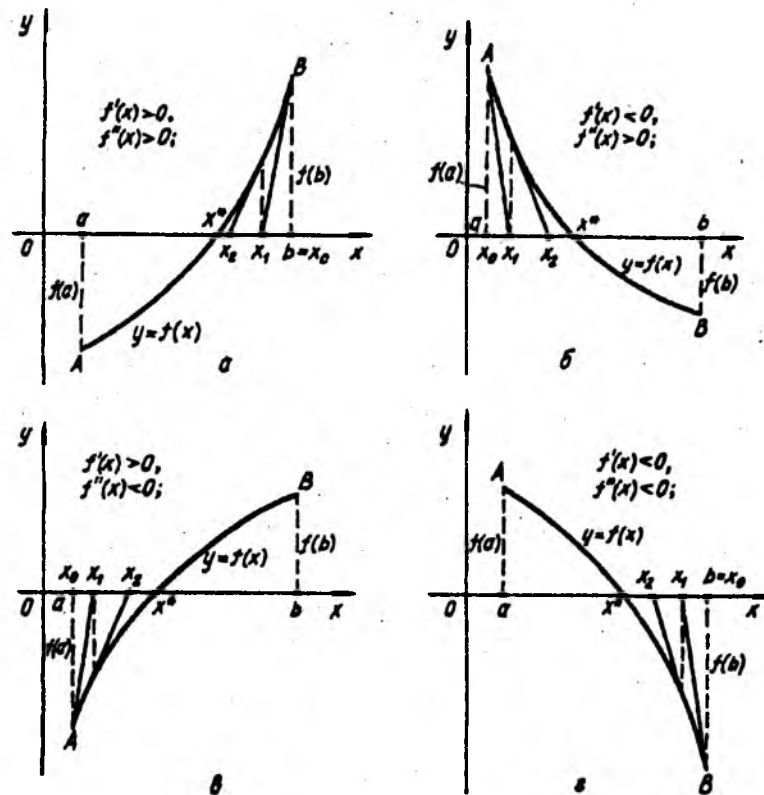


Рис. 2.5

За початкове наближення у методі Ньютона слід брати точку $x_0 \in [a; b]$, в якій $f(x_0)f'(x_0) > 0$.

Метод Ньютона є методом послідовних наближень $x_{k+1} = \varphi(x_k)$, де функція

$$\varphi(x) = x - \frac{f(x)}{f'(x)}. \quad (2.27)$$

Достатні умови збіжності методу Ньютона дає така теорема.
Теорема 6. Нехай на відрізку $[a; b]$ функція $f(x)$ має неперервні із сталими знаками похідні $f'(x) \neq 0$, $f''(x) \neq 0$ і $f(a)f(b) < 0$. Тоді існує

такий окіл $R \subset [a; b]$ кореня x^* рівняння $f(x) = 0$, що для будь-якого $x_0 \in R$ послідовність $\{x_k\}$, обчислена за формулою (2.26), збігається до кореня x^* .

Д о в е д е н н я. Для доведення збіжності послідовності $\{x_k\}$ до кореня x^* досить показати, що похідна $\varphi'(x)$ функції (2.27) задовольняє умови $0 < |\varphi'(x)| \leq q < 1$ для будь-якого $x \in R$, і взяти x_0 з околу R кореня x^* .

Для похідної $\varphi'(x)$ маємо вираз

$$\varphi'(x) = 1 - \frac{(f'(x))^2 - f(x)f''(x)}{(f'(x))^2} = \frac{f(x)f''(x)}{(f'(x))^2}.$$

Оскільки $f(x)$ неперервна і відмінна від нуля на $[a; b]$, то існують такі додатні m_1 і M_2 , що

$$|f'(x)| \geq m_1, \quad |f''(x)| \leq M_2$$

для будь-яких $x \in [a; b]$.

Тоді

$$|\varphi'(x)| \leq \frac{M_2}{m_1^2} |f(x)|.$$

З неперервності функції $f(x)$ випливає, що існує окіл R кореня x^* , в якому функція $f(x)$ задовольняє нерівність

$$|f(x)| \leq \frac{m_1^2}{M_2} q, \quad 0 < q < 1.$$

Внаслідок чого $|\varphi'(x)| \leq q < 1$ для всіх $x \in R$.

Таким чином, функція $\varphi(x)$ задовольняє умови теореми 5, з якої випливає, що послідовність $\{x_k\}$, обчислена за формулами (2.26), збігається до кореня x^* , якщо початкове наближення $x_0 \in R$.

Теорему доведено.

Оцінимо швидкість збіжності методу Ньютона. Для цього запишемо функцію $f(x)$ в околі точки $x_k \in R$ за формулою Тейлора із залишковим членом у формі Лагранжа

$$f(x) = f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2}f''(\eta)(x - x_k)^2,$$

де η лежить між точками x і x_k . Поклавши в ній $x = x^*$ і врахувавши, що $f(x^*) = 0$, дістанемо:

$$\frac{f(x_k)}{f'(x_k)} - x_k + x^* + \frac{1}{2} \frac{f''(\eta)}{f'(x_k)} (x^* - x_k)^2 = 0.$$

Звідси і з (2.26) маємо

$$x_{k+1} - x^* = \frac{1}{2} \frac{f''(\eta)}{f'(x_k)} (x^* - x_k)^2.$$

Поклавши

$$M_2 = \max_{x \in [a; b]} |f''(x)|, \quad m_1 = \min_{x \in [a; b]} |f'(x)|, \quad (2.28)$$

знаходимо

$$|x_{k+1} - x^*| \leq \frac{M_2}{2m_1} |x^* - x_k|^2. \quad (2.29)$$

З оцінки (2.29) випливає, що метод Ньютона збігатиметься до кореня x^* , якщо початкове наближення x_0 таке, що

$$\frac{M_2}{2m_1} |x^* - x_0| < 1,$$

причому в цьому випадку збіжність є квадратичною. Це означає, що похибка кожного наступного наближення пропорційна квадрату похибки попереднього наближення.

Зокрема, якщо $\frac{M_2}{2m_1} \leq 1$ і $|x^* - x_k| < 10^{-m}$, то з (2.29) маємо:

$$|x_{k+1} - x^*| < 10^{-2m}.$$

Отже, коли наближення x_k має m правильних десяткових знаків, то наступне наближення x_{k+1} матиме щонайменше $2m$ правильних десяткових знаків.

Для оцінки k -го наближення методу Ньютона можна скористатися формулою (2.2), в якій взяти $m_1 = \min_{x \in [a; b]} |f'(x)|$. Тоді ітераційний процес

можна закінчити, якщо стане правильною нерівність $\left| \frac{f(x_k)}{m_1} \right| < \varepsilon$, де ε — наперед задана точність наближення x_k .

Можна довести, що для методу Ньютона справедлива оцінка

$$|x_k - x^*| \leq \frac{M_2}{2m_1} |x_k - x_{k-1}|^2,$$

де m_1 і M_2 визначаються за формулами (2.28).

З цієї оцінки видно, що для досягнення заданої точності ε ітераційний процес треба продовжувати доти, поки для двох послідовних наближень x_k і x_{k-1} не виконуватиметься нерівність

$$|x_k - x_{k-1}| \leq \sqrt{\frac{2m_1}{M_2}} \varepsilon.$$

Якщо на відрізку $[a; b]$ справедлива нерівність $M_2 < 2m_1$, то ітераційний процес можна закінчити, коли виконується умова $|x_k - x_{k-1}| < \varepsilon$.

Програма 2.4 містить алгоритм методу Ньютона.

Програма 2.4

```

10 ' -----Метод Ньютона-----
20 ' -----для рівняння f(x) = 0-----
30 PRINT "Ввести кінці a і b відрізка ізоляції кореня"
40 INPUT "A="; A : INPUT "B="; B
50 INPUT "Ввести точність EPS"; EPS
60 INPUT "Ввести min f (x) на [a;b]"; M1
70 INPUT "Ввести max f'(x) на [a;b]"; M2
80 DEF FNF(X)=          : ' — вираз функції f(x)
90 DEF FNY(X)=          : ' — вираз похідної f'(x)
100 DEF FNZ(X)=         : ' — вираз другої похідної f''(x)
110 U=FNF(A): V=FNZ(A)  : ' Визначення початкового
120 IF U*V > 0 THEN X0=A ELSE X0=B : ' наближення X0
130 EPS= EPS*SQR(2*M1/M2)
140 P=FNF(X0)/FNY(X0)
150 X=X0-P
160 IF ABS(X-X0) < = EPS THEN 180
170 X0=X: GOTO 140
180 PRINT "Шуканий корінь X="; X : END
    
```

Перевага методу Ньютона перед методом ітерацій у тому, що він має вищу швидкість збіжності. Так, корінь $x^* \in [2;3]$ рівняння $x^3 - 2x - 5 = 0$ з точністю $\epsilon = 10^{-6}$ і $\epsilon = 10^{-9}$ методом Ньютона був обчислений за п'ять і шість ітерацій відповідно, тоді як методом ітерації він був обчислений не менш ніж за шість і десять ітерацій відповідно.

З формули (2.27) видно, що чим більше значення $|f'(x)|$ в околі кореня, тим менша поправка додається до попереднього наближення. Тому метод Ньютона зручно застосовувати тоді, коли в околі кореня графік функції $y = f(x)$ має значну крутість. Крім того, методом Ньютона можна знаходити не тільки дійсні корені рівнянь, а й комплексні. Метод Ньютона легко поширюється і на розв'язування систем нелінійних рівнянь з багатьма невідомими.

Недоліком методу Ньютона є те, що на кожній ітерації треба обчислювати не тільки значення функції $f(x)$, а й значення її похідної $f'(x)$. Обчислення похідної $f'(x)$ може бути значно складнішим від обчислення $f(x)$. У таких випадках похідну $f'(x_k)$ замінюють її наближеним значенням

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

Підставивши значення $f'(x_k)$ у формулу (2.26), знайдемо

$$x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}, \quad k = 1, 2, \dots \quad (2.30)$$

Метод, що визначається формулою (2.30), називають методом січних. Геометрично наближення x_{k+1} дістають як абсцису точки перетину осі Ox і січної, що проходить через точки $(x_{k-1}, f(x_{k-1}))$ і $(x_k, f(x_k))$ кривої $y = f(x)$ (рис. 2.6, а).

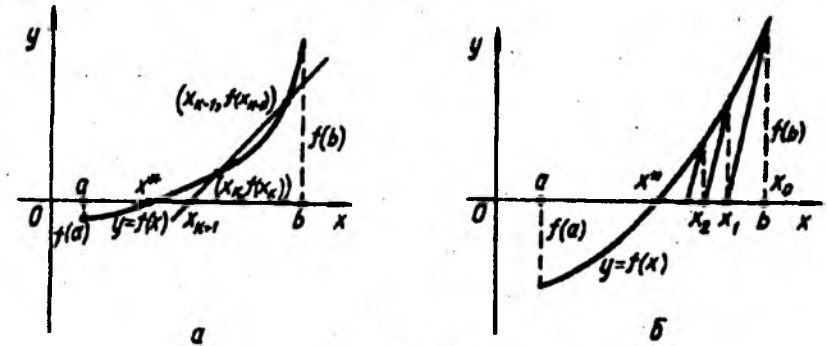


Рис. 2.6

Алгоритм методу січних починається із задання двох початкових наближень x_0 і x_1 , які вибирають на відрізку ізоляції кореня x^* . Тому складність цього методу полягає в знаходженні таких x_0 і x_1 , досить близьких до x^* , щоб метод був збіжним.

Якщо похідна $f'(x)$ мало змінюється на $[a; b]$, то кількість обчислень у методі Ньютона можна зменшити, коли значення похідних $f'(x_k)$ у точках $x_k, k = 1, 2, \dots$ замінити значенням $f'(x_0)$ у точці x_0 . Тоді формула (2.26) набере вигляду

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_0)}, \quad k = 0, 1, 2, \dots \quad (2.31)$$

В результаті дістали спрощений метод Ньютона. Геометрично він означає, що дотичні в точках $(x_k, f(x_k))$ замінюються прямими, паралельними дотичній, проведеній до кривої $y = f(x)$ у точці $(x_0, f(x_0))$ (рис. 2.6, б).

§ 2.7. МЕТОД ХОРД

Метод хорд — один з поширених ітераційних методів. Його ще називають методом лінійного інтерполювання, методом пропорційних частин, або методом хибного положення.

Нехай задано рівняння $f(x) = 0$, де $f(x)$ на відрізку $[a; b]$ має непервні похідні першого й другого порядків, які зберігають сталі знаки на цьому відрізку, і $f(a)f(b) < 0$, тобто корінь x^* рівняння відокремлений на $[a; b]$.

Ідея методу хорд в тому, що на досить малому відрізку дуга кривої $y = f(x)$ замінюється хордою і абсциса точки перетину хорди з віссю Ox є наближеним значенням кореня.

Нехай для визначеності $f'(x) > 0, f''(x) > 0, f(a) < 0, f(b) > 0$ (рис. 2.7, а). Візьмемо за початкове наближення шуканого кореня x^* значення $x_0 = a$. Через точки A_0 і B проведемо хорду і за перше

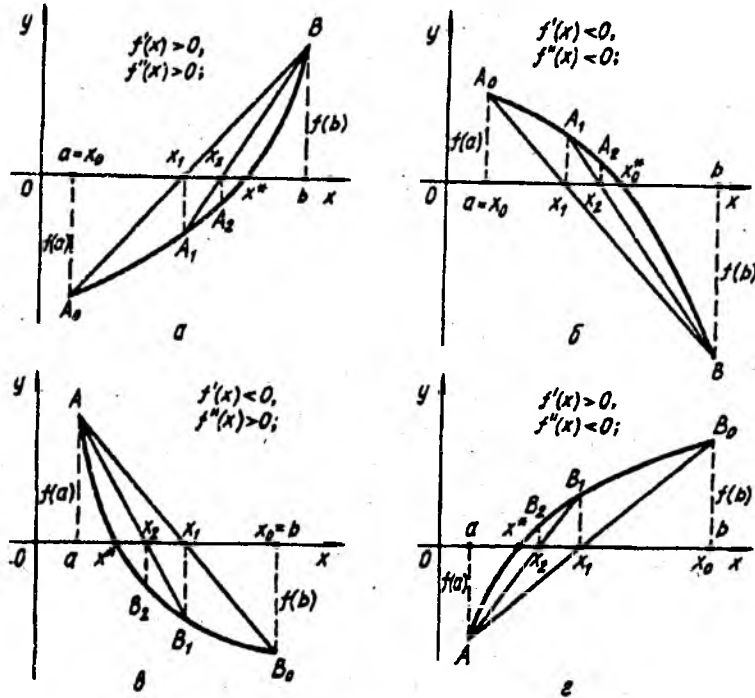


Рис. 2.7

наближення кореня x^* візьмемо абсцису x_1 точки перетину хорди з віссю Ox . Тепер наближене значення x_1 кореня можна уточнити, якщо застосувати метод хорд до відрізка $[x_1; b]$. Абсциса x_2 точки перетину хорди A_1B буде другим наближенням кореня. Продовжуючи цей процес необмежено, дістанемо послідовність $x_0, x_1, x_2, \dots, x_k, \dots$ наближених значень кореня x^* даного рівняння.

Для виведення формули методу хорд запишемо рівняння прямої, що проходить через точки $A_k(x_k, f(x_k))$ і $B(b, f(b))$:

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

Поклавши $y = 0$, знайдемо абсцису точки перетину хорди A_kB з віссю Ox :

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

Значення x можна взяти за наступне наближення, тобто

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

У цьому разі і тоді, коли $f(a) > 0, f(b) < 0, f'(x) < 0, f''(x) < 0$ (рис. 2.7, б) кінець b відрізка $[a; b]$ є нерухомим.

Якщо $f(a) > 0, f(b) < 0, f'(x) < 0, f''(x) > 0$ (рис. 2.7, в), або $f(a) < 0, f(b) > 0, f'(x) > 0, f''(x) < 0$ (рис. 2.7, г), аналогічно можна записати формулу

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

У цьому випадку точка a є нерухомим кінцем відрізка $[a; b]$.

У загальному випадку нерухомим буде той кінець відрізка ізоляції кореня, в якому знак функції $f(x)$ збігається із знаком другої похідної, а за початкове наближення x_0 можна взяти точку відрізка $[a; b]$, в якій $f(x_0)f''(x_0) < 0$.

Отже, метод хорд можна записати так:

$$x_{k+1} = x_k - \frac{f(x_k)}{f(x_k) - f(c)} (x_k - c), \quad k = 0, 1, 2, \dots, \quad (2.32)$$

де

$$c = \begin{cases} a, & \text{якщо } f(a)f''(a) > 0, \\ b, & \text{якщо } f(b)f''(b) > 0. \end{cases}$$

З формули (2.32) видно, що метод хорд є методом ітерацій $x_{k+1} = \varphi(x_k)$, в якому

$$\varphi(x) = x - \frac{f(x)}{f(x) - f(c)} (x - c). \quad (2.33)$$

Зауважимо, що рівняння

$$x = x - \frac{f(x)}{f(x) - f(c)} (x - c)$$

на відрізку $[a; b]$ рівносильне рівнянню $f(x) = 0$.

Достатні умови збіжності методу хорд дає така теорема.

Теорема 7. Нехай на відрізку $[a; b]$ функція $f(x)$ неперервна разом із своїми похідними до другого порядку включно, причому $f(a)f(b) < 0$, а похідні $f'(x)$ і $f''(x)$ зберігають сталі знаки на $[a; b]$, тоді існує такий окіл кореня x^* рівняння $f(x) = 0$, що для будь-якого початкового наближення x_0 з цього околу послідовність $\{x_k\}$, обчислена за формулою (2.32), збігатиметься до кореня x^* .

Д о в е д е н н я. Для доведення теореми досить показати, що в деякому околі $R \subset [a; b]$ кореня x^* похідна $\varphi'(x)$ функції (2.33) задовольняє умову $|\varphi'(x)| \leq q < 1$ для будь-яких $x \in R$.

Обчислимо

$$\varphi'(x) = 1 - \frac{f'(x)(x - x_0) + f(x)}{f(x) - f(x_0)} + \frac{f(x)f'(x)(x - x_0)}{(f(x) - f(x_0))^2}.$$

Поклавши $x = x^*$ і врахувавши, що $f(x^*) = 0$, маємо

$$\varphi'(x^*) = \frac{f(x_0) + f'(x^*)(x^* - x_0)}{f(x_0)}. \quad (2.34)$$

Запишемо для $f(x)$ в околі точки x^* формулу Тейлора із залишковим членом у формі Лагранжа:

$$f(x) = f(x^*) + f'(x^*)(x - x^*) + \frac{1}{2}f''(\eta)(x - x^*)^2,$$

де η лежить між x і x^* .

Поклавши в ній $x = x_0$, дістанемо

$$f(x_0) + f'(x^*)(x^* - x_0) = \frac{1}{2}f''(\eta)(x_0 - x^*)^2. \quad (2.35)$$

Із формули (2.34), враховуючи (2.35), знаходимо

$$\varphi'(x^*) = \frac{(x_0 - x^*)^2}{2f(x_0)} f''(\eta).$$

Оскільки $f(x)$ і $f''(x)$ — неперервні на $[a; b]$, то і $\varphi'(x)$ буде неперервною на $[a; b]$ функцією, тому

$$\lim_{x_0 \rightarrow x^*} \varphi'(x^*) = \lim_{x_0 \rightarrow x^*} \frac{(x_0 - x^*)^2}{2f(x_0)} f''(\eta) = 0.$$

Звідси і з неперервності $\varphi'(x)$ випливає, що на відрізку $[a; b]$ існує окіл R точки x^* такий, що $|\varphi'(x)| \leq q < 1$ для будь-якого $x \in R$. Тоді з теореми 5 випливає, що послідовність $\{x_k\}$, обчислена за формулою (2.32), збігається до кореня x^* , якщо початкове наближення $x_0 \in R$. Теорему доведено.

Для оцінки похибки, як і в методі дотичних, можна скористатися формулою (2.2).

Виведемо ще одну формулу, яка дає можливість оцінити абсолютну похибку наближення x_k через два послідовні наближення x_{k-1} і x_k .

Нехай $f'(x)$ — неперервна і зберігає на $[a; b]$ сталий знак, причому

$$0 < m_1 \leq |f'(x)| \leq M_1 < \infty,$$

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

З формули

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

дістаємо

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

Звідси, враховуючи, що $f(x^*) = 0$, маємо

$$f(x^*) - f(x_{k-1}) = \frac{f(x_{k-1}) - f(c)}{x_{k-1} - c} (x_k - x_{k-1}).$$

Застосувавши теорему Лагранжа, дістанемо

$$(x^* - x_{k-1})f'(\eta) = (x_k - x_{k-1})f'(\xi),$$

де η лежить між точками x^* і x_{k-1} , а ξ — між x_{k-1} і c . Далі запишемо:

$$(x^* - x_k + x_k - x_{k-1})f'(\eta) = (x_k - x_{k-1})f'(\xi)$$

або

$$x^* - x_k = \frac{f'(\xi) - f'(\eta)}{f'(\eta)} (x_k - x_{k-1}).$$

Оскільки $f'(x)$ зберігає на $[a; b]$ сталий знак, то

$$|f'(\xi) - f'(\eta)| \leq M_1 - m_1.$$

Тому

$$|x^* - x_k| \leq \frac{M_1 - m_1}{m_1} |x_k - x_{k-1}|. \quad (2.36)$$

Якщо на відрізку $[a; b]$ справедлива нерівність $M_1 \leq 2m_1$, то із (2.36) випливає оцінка:

$$|x^* - x_k| \leq |x_k - x_{k-1}|.$$

Отже, корінь x^* рівняння $f(x) = 0$ буде знайдено методом хорд із наперед заданою точністю ϵ , якщо для двох послідовних наближень x_{k-1}

і x_k справджуватиметься нерівність $|x_k - x_{k-1}| \leq \frac{m_1}{M_1 - m_1} \epsilon$.

§ 2.8. КОМБІНОВАНИЙ МЕТОД ДОТИЧНИХ І ХОРД

Характерна особливість методів дотичних і хорд та, що послідовності їх наближень монотонні. Причому, якщо для даного рівняння послідовність наближень методу хорд монотонно спадає, то послідовність наближень методу дотичних — монотонно зростає, і навпаки. Одночасне застосування цих методів дає змогу наближатися до кореня рівняння з двох боків, дістаючи наближення з недостаткою і надлишком.

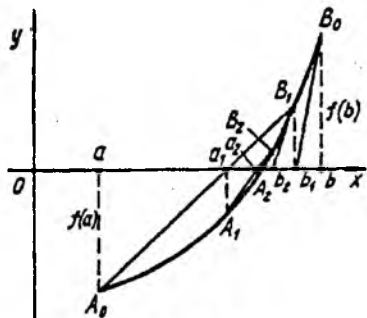


Рис. 2.8

Розглянемо рівняння $f(x) = 0$, корінь якого $x^* \in [a; b]$. Нехай, наприклад, $f'(x) > 0$, $f''(x) > 0$, $f(a) < 0$, $f(b) > 0$ (рис. 2.8).

У даному випадку за початкове наближення в методі хорд вибирають точку $x = a$, а в методі дотичних — точку b . На відрізку $[a; b]$ застосовують метод дотичних і хорд (спочатку дотичних, а потім хорд). У результаті дістають нові наближення a_1, b_1 , і початковий відрізок ізоляції кореня звужився. Для знаходження нових наближень застосовують метод дотичних і хорд уже на відрізку

$[a_1; b_1]$. У результаті дістають наближення a_2 і b_2 відповідно, причому $[a_2; b_2] \subset [a_1; b_1] \subset [a; b]$. Такий процес продовжують доти, поки довжина відрізка $[a_k; b_k]$ стане меншою або дорівнюватиме величині 2ϵ , де ϵ — наперед задана точність кореня.

За шукане значення кореня \bar{x} беруть півсуму наближень a_k і b_k , тобто $\bar{x} = 0,5(a_k + b_k)$, а модуль їх піврізниці дасть граничну абсолютну похибку наближеного кореня, тобто

$$|x^* - \bar{x}| \leq 0,5|a_k - b_k|.$$

Значимо, що на кожному кроці комбінованого методу за нерухомий кінець c у формулі методу хорд треба брати наближення, обчислене на цьому самому кроці за формулою дотичних.

Формули комбінованого методу дотичних і хорд мають вигляд

$$b_{k+1} = b_k - \frac{f(b_k)}{f'(b_k)}, \quad k = 0, 1, 2, \dots \quad (2.37)$$

$$a_{k+1} = a_k - \frac{f(a_k)(a_k - b_{k+1})}{f(a_k) - f(b_{k+1})}, \quad k = 0, 1, 2, \dots \quad (2.38)$$

За початкове наближення b_0 у формулі (2.37) методу дотичних беруть той з кінців відрізка $[a; b]$, в якому значення функції $f(x)$ і її другої

похідної мають однакові знаки, тоді протилежний кінець відрізка $[a; b]$ беруть за початкове наближення a_0 у формулі (2.38) методу хорд.

Завдяки своєрідній комбінації методів дотичних і хорд комбінований метод має вищу швидкість збіжності, ніж методи хорд і дотичних окремо взяті. Так, для рівняння $x^3 - 2x - 5 = 0$ його корінь $x^* \in [2; 3]$ з точністю до 10^{-6} і 10^{-9} комбінованим методом обчислено за дві і три ітерації відповідно, тоді як для методу дотичних потрібно було п'ять і шість ітерацій відповідно, а для методу хорд — дванадцять і дев'ятнадцять ітерацій відповідно.

ЛАБОРАТОРНА РОБОТА № 1

Тема. Чисельні методи розв'язування рівнянь з однією змінною.

Завдання. 1. Відокремити корені даного рівняння $f(x) = 0$ одним з методів: графічним, аналітичним або методом послідовного перебору.

2. Уточнити один з відокремлених коренів рівняння двома, вказаними в табл. 2.3, методами з точністю $\epsilon = 0,5 \cdot 10^{-3}$ і $\epsilon = 0,5 \cdot 10^{-6}$.

3. Порівняти використані методи між собою за кількістю ітерацій, потрібних для знаходження кореня із заданою точністю.

Пояснення до виконання роботи. Виконуючи п.1 завдання, відрізок ізоляції кореня бажано звужити до довжини, яка не перевищує одиниці.

При уточненні ізолюваного кореня методом ітерацій подане в табл. 2.3 рівняння слід звести до вигляду $x = \varphi(x)$, де функція $\varphi(x)$ на відрізку ізоляції кореня повинна задовольняти умову $|\varphi'(x)| < 1$.

Перед тим як користуватися методами хорд, дотичних і комбінованим, потрібно перевірити виконання достатніх умов збіжності відповідного методу.

Для чисельного розв'язування рівнянь методами дихотомії, ітерації і Ньютона можна скористатися поданими в підручнику відповідними програмами, передбачивши в них підрахунок кількості ітерацій, потрібних для досягнення заданої точності наближеного кореня.

Для розв'язування рівняння методами хорд і комбінованим відповідні програми скласти самостійно.

Таблиця 2.3

Варіант	Рівняння	Метод
1	$2^x + 5x - 3 = 0$	дихотомії, хорд
2	$\ln(1,5x) - 1,7x + 3 = 0$	ітерації, комбінований
3	$x^3 - 3x^2 + 2,5 = 0$	хорд, дотичних
4	$x^2 + 4\sin x = 0$	дихотомії, хорд
5	$\sin x - x - \ln(1 + x) + 1 = 0$	хорд, Ньютона
6	$x^3 - 3x^2 - 3,5 = 0$	дихотомії, комбінований
7	$x^2 - 20\sin x = 0$	хорд, дихотомії
8	$2 - x - \ln x = 0$	ітерації, хорд
9	$3x - \cos x - 1 = 0$	дихотомії, комбінований

Варіант	Рівняння	Метод
10	$x^2 - \ln(1+x) - 3 = 0$	Ітерації, комбінований
11	$2x^3 + 9x^2 - 4 = 0$	хорд, дотичних
12	$2e^x + 2x - 3 = 0$	хорд, Ньютона
13	$x\sqrt{x+1} - 1 = 0$	Ітерації, комбінований
14	$x^3 - 3x^2 + 6x + 3 = 0$	хорд, дотичних
15	$\operatorname{ctg} x - 0,5x = 0$	дихотомії, хорд
16	$2\sin(x-0,6) + x - 1,5 = 0$	Ітерації, комбінований
17	$2^x(x-2)^2 - 1 = 0$	дихотомії, хорд
18	$x^3 - 3x^2 + 9x + 2 = 0$	Ньютона, комбінований
19	$x^3 - \sin x = 0$	Ітерації, комбінований
20	$x^3 - 0,1x^2 + 0,4x - 1,5 = 0$	хорд, дотичних
21	$3\sin\sqrt{x} + 0,35x - 3,8 = 0$	Ітерації, хорд
22	$x^2 - \cos x = 0$	дихотомії, комбінований
23	$2x^3 - 3x^2 - 12x + 12 = 0$	Ньютона, комбінований
24	$3^x + 5x - 2 = 0$	хорд, дотичних
25	$(x-3)\cos x - 1 = 0$	дихотомії, хорд
26	$x - (3 + \sin 3,6x)^{-1} = 0$	Ітерації, комбінований
27	$x^3 + 4x - 6 = 0$	хорд, дотичних
28	$x^3 - 3x^2 - 24x - 5 = 0$	дихотомії, комбінований
29	$x - \sqrt{\ln(x+2)} = 0$	Ітерації, комбінований
30	$2x\sin x - \cos x = 0$	Ньютона, комбінований

Розділ 3. РОЗВ'ЯЗУВАННЯ СИСТЕМ ЛІНІЙНИХ АЛГЕБРАІЧНИХ РІВНЯНЬ

§ 3.1. ПОСТАНОВКА ЗАДАЧІ

Нехай дано систему n лінійних алгебраїчних рівнянь з n змінними

$$\sum_{j=1}^n a_{ij}x_j = b_i \quad (i = 1, 2, \dots, n). \quad (3.1)$$

Систему (3.1) можна записати у вигляді одного матричного рівняння

$$AX = B, \quad (3.2)$$

де

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

— матриця коефіцієнтів a_{ij} (індекс i вказує рівняння, якому належить коефіцієнт, а індекс j — змінну, при якій він стоїть),

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

— відповідно стовпець вільних членів і стовпець змінних.

Упорядкована сукупність n чисел c_1, c_2, \dots, c_n , яка, будучи підставленою в систему (3.1) замість x_1, x_2, \dots, x_n , перетворює всі рівняння в правильні числові рівності, називається *розв'язком системи* (3.1).

Якщо визначник системи (3.1)

$$\Delta = \det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \neq 0,$$

то вона має єдиний розв'язок. Його можна обчислити за формулою Крамера

$$x_k = \frac{\det A_k}{\det A}, \quad (k = 1, 2, \dots, n),$$

де матрицю A_k дістають з матриці A , замінивши її k -й стовпець стовпцем вільних членів.

Формули Крамера мають велике теоретичне значення, але через великий обсяг обчислювальної роботи мало ефективні при чисельному розв'язуванні лінійних систем. За цими формулами треба обчислити значення $(n+1)$ -го визначників порядку n , для чого виконати значну кількість арифметичних операцій.

Методи розв'язування систем лінійних рівнянь можна поділити на дві групи: *точні й ітераційні*.

Точними називають такі методи, які дають змогу знайти точний розв'язок системи (3.1) за допомогою виконання скінченної кількості арифметичних операцій у припущенні, що всі обчислення виконуються точно (без округлень), а коефіцієнти системи і вільні члени — точні числа. Але на практиці всі обчислення виконуються з обмеженою кількістю десяткових розрядів, а ірраціональні коефіцієнти і вільні члени, якщо такі є, замінюються раціональними числами. Тому в процесі обчислень вдаються до округлень, а це означає, що розв'язки, які обчислюються за точними методами, фактично є наближеними числами з певними похибками (похибками округлень). До точних належать *метод Гаусса, метод квадратних коренів, правило Крамера* тощо.

Ітераційними називають такі методи, які дають змогу знайти наближений розв'язок системи (3.1) із задалегідь указаною точністю шляхом виконання скінченної кількості арифметичних операцій, хоч самі обчислення можуть проводитись і без округлень, а коефіцієнти і вільні члени системи бути точними числами. Точний розв'язок системи (3.1) за допомогою ітераційних методів можна знайти тільки теоретично як границю збіжного нескінченного процесу. Розв'язуючи системи рівнянь ітераційними методами, крім похибок округлення, треба враховувати також похибку методу. До ітераційних належать *метод ітерації, метод Зейделя* тощо.

У процесі вивчення різних питань економіки, природознавства, техніки тощо доводиться розв'язувати системи лінійних алгебраїчних рівнянь. Зокрема, до таких систем зводиться чисельне розв'язування лінійних диференціальних та інтегральних рівнянь. У таких системах коефіцієнти і вільні члени рівнянь — числа наближені. А це веде до появи додаткових (так званих неусувних) похибок, які треба враховувати як у процесі обчислень, так і в остаточному округленні результату.

Коефіцієнти лінійних систем, які виникають під час обробки результатів експерименту, містять помилки спостережень. Якщо коефіцієнти системи задано формулами, то обчислення їх призводить до похибок округлення. Якщо систему рівнянь у пам'ять машини записати навить точно, то в процесі її розв'язування ЕОМ обов'язково виникнуть похибки округлень, які не можуть не вплинути на точність розв'язку. Проте,

якщо матриця A системи (3.2) майже вироджена, то можна сподіватись, що малі зміни в коефіцієнтах і (або) вільних членах також призведуть до значних змін у її розв'язку.

Якщо малі збурення коефіцієнтів і (або) вільних членів системи (3.1) дуже збурюють її розв'язок, то таку систему рівнянь називають *погано обумовленою*. Навпаки, якщо малі збурення коефіцієнтів і (або) вільних членів системи (3.1) мало збурюють її розв'язок, то таку систему називають *добре обумовленою*. Прикладом погано обумовленої є, наприклад, система вигляду

$$\begin{cases} 6,1x_1 + 3,4x_2 = 6,1, \\ 14,7x_1 + 8,2x_2 = 14,7, \end{cases} \quad (3.3)$$

розв'язком якої є пара (1; 0). Якщо число 6,1 у правій частині першого рівняння системи (3.3) змінити на 0,02, то система

$$\begin{cases} 6,1x_1 + 3,4x_2 = 6,12, \\ 14,7x_1 + 8,2x_2 = 14,7 \end{cases}$$

матиме розв'язком пару (5,1; -7,35). Отже, мале збурення (менше 0,33 %) одного з вільних членів системи (3.3) зовсім змінило розв'язок системи.

Якщо коефіцієнт 6,1 у лівій частині першого рівняння системи (3.3) змінити на 0,01, то розв'язком системи

$$\begin{cases} 6,11x_1 + 3,4x_2 = 6,1, \\ 14,7x_1 + 8,2x_2 = 14,7 \end{cases}$$

буде пара (0,36730972; 1,204918). І в цьому випадку мале збурення (менше 0,17 %) одного коефіцієнта системи (3.3) значно змінило її розв'язок.

Ознакою поганої обумовленості системи лінійних рівнянь є її майже виродженість. Іншими словами, якщо значення визначника системи досить мале порівняно з її коефіцієнтами, то така система близька до виродженої (особливої). Так, визначник системи (3.3) $\Delta = 0,04$ і становить близько 0,27 % значення найбільшого з коефіцієнтів системи ($a_{21} = 14,7$).

Далі розглянуто лише такі системи лінійних рівнянь, які не є виродженими (особливими), тобто їх визначник не дорівнює нулю, є добре обумовленими, тобто визначник системи є величиною не нижчого порядку, ніж її коефіцієнти, а коефіцієнти і вільні члени системи — точні числа.

§ 3.2. МЕТОД ПОСЛІДОВНОГО ВИКЛЮЧЕННЯ ЗМІННИХ (МЕТОД ГАУССА)

Найпростішим методом розв'язування систем лінійних алгебраїчних рівнянь є *метод послідовного виключення змінних, або метод Гаусса*. Є кілька модифікацій цього методу. Розглянемо *схему єдиного ділення*,

за якою систему розв'язують в два етапи. На першому етапі вихідну систему рівнянь зводять до рівносильної їй системи трикутної форми. Цей процес перетворення називають *прямим ходом*. На другому етапі, який називають *зворотним ходом*, знаходять розв'язок лінійної системи рівнянь трикутної форми.

Обмежимося розглядом системи трьох рівнянь з трьома змінними

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = a_{14}, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = a_{24}, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = a_{34}, \end{cases} \quad (3.4)$$

визначник Δ якої не дорівнює нулю.

Нехай $a_{11} \neq 0$ (цього завжди можна досягти перестановкою рівнянь системи, бо $\Delta \neq 0$, і тому завжди є таке рівняння, в якому $a_{11} \neq 0$). Поділимо коефіцієнти першого рівняння системи (3.4), включаючи й вільний член, на коефіцієнт a_{11} . Дістанемо нове рівняння

$$x_1 + a_{12}^{(1)}x_2 + a_{13}^{(1)}x_3 = a_{14}^{(1)}, \quad (3.5)$$

де

$$a_{ij}^{(1)} = a_{ij} / a_{11}, \quad (j = 2, 3, 4). \quad (3.6)$$

Виключимо тепер змінну x_1 з другого і третього рівнянь системи (3.4). Для цього рівняння (3.5) помножимо послідовно спочатку на коефіцієнт a_{21} і віднімемо його від другого рівняння системи (3.4), а потім на a_{31} і віднімемо від третього рівняння системи (3.4). Дістанемо систему двох рівнянь з двома змінними x_2 і x_3

$$\begin{cases} a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 = a_{24}^{(1)}, \\ a_{32}^{(1)}x_2 + a_{33}^{(1)}x_3 = a_{34}^{(1)}, \end{cases} \quad (3.7)$$

де коефіцієнти $a_{ij}^{(1)}$ обчислюють за формулами

$$a_{ij}^{(1)} = a_{ij} - a_{11}a_{1j}^{(1)} \quad (i = 2, 3; j = 2, 3, 4). \quad (3.8)$$

Верхній індекс (1) вказує на те, що над коефіцієнтами системи (3.4) виконано перше перетворення.

Далі поділимо коефіцієнти першого рівняння системи (3.7) на $a_{22}^{(1)}$ (якщо $a_{22}^{(1)} = 0$, то переставляємо рівняння місцями). Дістанемо рівняння

$$x_2 + a_{23}^{(2)}x_3 = a_{24}^{(2)}, \quad (3.9)$$

де

$$a_{2j}^{(2)} = a_{2j}^{(1)} / a_{22}^{(1)}, \quad (j = 3, 4). \quad (3.10)$$

Із системи (3.7) виключимо змінну x_2 так само, як і x_1 з системи (3.4). Дістанемо рівняння

$$a_{33}^{(2)}x_3 = a_{34}^{(2)}, \quad (3.11)$$

де

$$a_{3j}^{(2)} = a_{3j}^{(1)} - a_{32}^{(1)}a_{2j}^{(2)} \quad (j = 3, 4). \quad (3.12)$$

З рівняння (3.11) маємо

$$x_3 = a_{34}^{(3)}, \quad (3.13)$$

де

$$a_{34}^{(3)} = a_{34}^{(2)} / a_{33}^{(2)}. \quad (3.14)$$

Після трьох кроків перетворення дістанемо систему рівнянь трикутної форми

$$\begin{cases} x_1 + a_{12}^{(1)}x_2 + a_{13}^{(1)}x_3 = a_{14}^{(1)}, \\ x_2 + a_{23}^{(2)}x_3 = a_{24}^{(2)}, \\ x_3 = a_{34}^{(3)}, \end{cases} \quad (3.15)$$

яка еквівалентна системі (3.4).

На цьому прямий хід методу Гаусса завершено. Описаний процес перетворень системи (3.4) до рівносильної їй системи (3.15) можна здійснити, якщо виконуються умови $a_{11} \neq 0$, $a_{22}^{(1)} \neq 0$, $a_{33}^{(2)} \neq 0$. Можна довести, що коли $a_{11} \neq 0$, то для цього достатньо, щоб поряд з умовою $\Delta \neq 0$ виконувалась нерівність

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0.$$

Близькість діагональних коефіцієнтів системи (3.4) до нуля може призвести до значної втрати точності.

Оскільки системи (3.4) і (3.15) еквівалентні, то розв'язком системи (3.4) буде розв'язок системи (3.15), який можна записати формулами

$$\begin{cases} x_3 = a_{34}^{(3)}, \\ x_2 = a_{24}^{(2)} - a_{23}^{(2)}x_3, \\ x_1 = a_{14}^{(1)} - a_{13}^{(1)}x_3 - a_{12}^{(1)}x_2. \end{cases} \quad (3.16)$$

Цим завершено зворотний хід методу Гаусса.

Оскільки перетворення систем рівнянь є фактично перетвореннями їх коефіцієнтів при змінних і вільних членів, то для виконання перетворень немає потреби виписувати системи. Досить виписати лише матрицю коефіцієнтів і вільні члени і над ними виконати зазначені вище

перетворення. Всі записи доцільно розмішувати в окремій таблиці, яка складається з кількох частин, що відповідають певним крокам перетворення вихідної системи рівнянь (табл. 3.1).

Таблиця 3.1

Кроки перетворення	Рядок	Коефіцієнт при змінних			Вільний член	Контроль	
		x_1	x_2	x_3		контрольна сума Σ	рядкова сума σ
i	2	3	4	5	6	7	8
1	1	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	$\sum_{j=1}^4 a_{1j}$
	2	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	$\sum_{j=1}^4 a_{2j}$
	3	a_{31}	a_{32}	a_{33}	a_{34}	a_{35}	$\sum_{j=1}^4 a_{3j}$
2	4	1	$a_{12}^{(1)}$	$a_{13}^{(1)}$	$a_{14}^{(1)}$	$a_{15}^{(1)}$	$\sum_{j=2}^4 a_{1j}^{(1)} + 1$
	5		$a_{22}^{(1)}$	$a_{23}^{(1)}$	$a_{24}^{(1)}$	$a_{25}^{(1)}$	$\sum_{j=2}^4 a_{2j}^{(1)}$
	6		$a_{32}^{(1)}$	$a_{33}^{(1)}$	$a_{34}^{(1)}$	$a_{35}^{(1)}$	$\sum_{j=2}^4 a_{3j}^{(1)}$
3	7		1	$a_{23}^{(2)}$	$a_{24}^{(2)}$	$a_{25}^{(2)}$	$\sum_{j=3}^4 a_{2j}^{(2)} + 1$
	8			$a_{33}^{(2)}$	$a_{34}^{(2)}$	$a_{35}^{(2)}$	$\sum_{j=3}^4 a_{3j}^{(2)}$
4	9			1	x_3	\bar{x}_3	$1 + x_3$
	10		1		x_2	\bar{x}_2	$1 + x_2$
	11	1			x_1	\bar{x}_1	$1 + x_1$

У процесі розв'язування системи рівнянь треба організувати контроль правильності обчислень. Щоб вчасно виявити (і виправити) випадкові обчислювальні помилки, доцільно забезпечити контроль правильності обчислень у кожному рядку таблиці (так званий поточний контроль). Для цього до схеми обчислень введено два додаткових стовпці: 7-й — контрольна сума Σ і 8-й — рядкова сума σ .

Контрольна сума a_{i5} ($i = 1, 2, 3$) — це сума коефіцієнтів при змінних і вільного члена для кожного рівняння системи (3.4)

$$a_{i5} = \sum_{j=1}^4 a_{ij} \quad (i = 1, 2, 3). \quad (3.17)$$

У процесі розв'язування системи (3.4) за схемою єдиного ділення над контрольними сумами виконують ті самі перетворення, що й над відповідними елементами рядка, а знайдені в результаті цих перетворень значення нової контрольної суми записують у тому рядку, що й нові елементи відповідного рядка.

Рядкові суми — це суми всіх елементів відповідного рядка, що містяться в стовпцях з 3-го по 6-й включно. Поточний контроль обчислень за схемою єдиного ділення полягає в тому, що в кожному рядку обчислюють контрольну і рядкову суми та порівнюють їх між собою. Якщо вони збігаються або відрізняються на 1-2 одиниці нижчого розряду, що обумовлено впливом похибок округлення, то обчислення виконано правильно і можна переходити до обробки наступного рядка. Якщо ці суми в будь-якому рядку значно відрізняються між собою, то обчислення треба припинити і з'ясувати причину розбіжності. Значна розбіжність між контрольною і рядковою сумами може свідчити про випадкові помилки в обчисленнях або про нестійкість алгоритму обчислень відносно похибок округлень. Якщо причиною розбіжності є випадкова помилка, то її виправляють і продовжують обчислення. А якщо причиною розбіжності сум є нестійкість алгоритму обчислень відносно похибок округлення, то від нього слід відмовитися.

Зазначений спосіб організації контролю за обчисленнями по суті означає, що одночасно розв'язується дві системи лінійних алгебраїчних рівнянь з однаковою матрицею коефіцієнтів при змінних, але з різними вільними членами. Вільними членами вихідної системи (3.4) є числа a_{14} , a_{24} , a_{34} , а допоміжної системи — числа a_{15} , a_{25} , a_{35} .

Розв'язок x_1 , x_2 , x_3 системи (3.4) зв'язаний з розв'язком \bar{x}_1 , \bar{x}_2 , \bar{x}_3 допоміжної системи

$$\sum_{j=1}^3 a_{ij} \bar{x}_j = a_{i5} \quad (i = 1, 2, 3) \quad (3.18)$$

простою залежністю

$$\bar{x}_j = x_j + 1 \quad (j = 1, 2, 3). \quad (3.19)$$

Справді, підставивши (3.19) в систему (3.18) і врахувавши формули (3.17), дістанемо тотожність

$$\sum_{j=1}^3 a_{ij} x_j + \sum_{j=1}^3 a_{ij} = a_{i4} + \sum_{j=1}^3 a_{ij} = a_{i5} \quad (i = 1, 2, 3).$$

Під час виконання зворотного ходу методу Гаусса одночасно обчис-

люють як розв'язок x_1, x_2, x_3 системи (3.4), так і розв'язок $\bar{x}_1, \bar{x}_2, \bar{x}_3$ допоміжної системи (3.18). Ці розв'язки (в межах заданої точності) повинні задовольняти рівність (3.19). Саме в цьому й полягає суть заключного контролю.

Другою формою заключного контролю є безпосередня перевірка знайденого розв'язку підстановкою його в систему (3.4). Якби всі обчислення виконувалися точно, то в результаті підстановки дістали б правильну числову рівність. Але в процесі обчислень виконувалися округлення, тому значення лівих частин рівнянь системи (3.4), взагалі кажучи, не збігатимуться із значеннями їх правих частин. Значення різниць між вільними членами вихідної системи лінійних рівнянь і результатами підстановки в ці рівняння обчислених значень змінних називають *нев'язками*. Якщо невіздки досить малі, то можна стверджувати, що розв'язок системи (3.4) знайдено з малими похибками. Якщо невіздки досить значні, то це означає, що значення шуканих змінних x_1, x_2, x_3 обчислено з недостатньою точністю і їх треба уточнити. Це буває здебільшого тоді, коли проміжні обчислення виконують з недостатньою точністю. Зменшити невіздки можна так: розв'язують систему повторно, залишаючи в проміжних результатах більшу кількість десяткових розрядів, ніж при попередньому розв'язуванні, або обчислюють значення поправок до знайденого раніше розв'язку системи. Перший шлях досить громіздкий, потребує тим більших затрат обчислювальної роботи, чим вищий порядок системи. Тому перевагу доцільно надати другому шляху.

Нехай $x_1^{(0)}, x_2^{(0)}, x_3^{(0)}$ — наближений розв'язок системи (3.4). Підставивши його в систему (3.4), дістанемо невіздки

$$\begin{aligned} \delta_1 &= a_{14} - (a_{11}x_1^{(0)} + a_{12}x_2^{(0)} + a_{13}x_3^{(0)}), \\ \delta_2 &= a_{24} - (a_{21}x_1^{(0)} + a_{22}x_2^{(0)} + a_{23}x_3^{(0)}), \\ \delta_3 &= a_{34} - (a_{31}x_1^{(0)} + a_{32}x_2^{(0)} + a_{33}x_3^{(0)}). \end{aligned} \quad (3.20)$$

Шукатимемо тепер новий розв'язок системи (3.4) в такому вигляді

$$\begin{aligned} x_1 &= x_1^{(0)} + \varepsilon_1, \\ x_2 &= x_2^{(0)} + \varepsilon_2, \\ x_3 &= x_3^{(0)} + \varepsilon_3, \end{aligned} \quad (3.21)$$

де $\varepsilon_1, \varepsilon_2, \varepsilon_3$ — шукані похибки розв'язку. Підставивши (3.21) в систему (3.4) і взявши до уваги невіздки (3.20), для знаходження поправок ε_j ($j = 1, 2, 3$) дістаємо систему рівнянь

$$\sum_{j=1}^3 a_{ij}\varepsilon_j = \delta_i \quad (i = 1, 2, 3). \quad (3.22)$$

Система (3.22) відрізняється від системи (3.4) лише вільними членами; її розв'язок можна знайти за схемою єдиного ділення. Для цього

в табл. 3.1 досить додати новий стовпець 9, елементами якого будуть невіздки (3.20), тобто стовпець вільних членів системи (3.22). Значення поправок $\varepsilon_1, \varepsilon_2, \varepsilon_3$ будуть знайдені, якщо над невізтками $\delta_1, \delta_2, \delta_3$ виконати ті самі перетворення, що й над вільними членами системи (3.4). Обсяг обчислювальної роботи при цьому значно зменшується, бо перетворення виконуються лише над елементами одного (9-го) стовпця табл. 3.1.

П р и к л а д 1. Методом Гаусса розв'язати систему рівнянь

$$\begin{cases} 2,50x_1 + 0,94x_2 + 0,36x_3 = 6,804, \\ 0,87x_1 + 2,30x_2 + 0,76x_3 = 8,415, \\ 0,26x_1 + 0,97x_2 + 2,15x_3 = 8,877, \end{cases} \quad (3.23)$$

коефіцієнти і вільні члени якої є точними числами.

Розв'язання. Переконаємось спочатку, що система не вироджена і добре обумовлена. Для цього підраховуємо визначник системи: $\Delta = -9,035498$. Значення визначника системи становить 361 % значення найбільшого коефіцієнта системи ($a_{11} = 2,50$).

За схемою єдиного ділення систему розв'язуємо в такий спосіб.

1. Коефіцієнти і вільні члени системи (3.23) записуємо в перші три рядки (стовпці 3-6) табл. 3.2. Підраховуємо контрольні і рядкові суми, які збігаються між собою, і записуємо їх у 7-му і 8-му стовпцях.

Таблиця 3.2

Крок перетворення	Рядок	Коефіцієнти при змінних x_1, x_2, x_3			Вільний член	Контроль	
		3	4	5		6	7
1	1	2,50	0,94	0,36	6,804	10,604	10,604
1	2	0,87	2,30	0,76	8,415	12,345	12,345
	3	0,26	0,97	2,15	8,877	12,257	12,257
2	4	1	0,376	0,144	2,7216	4,2416	4,2416
	5		1,9729	0,6347	6,0472	8,6548	8,6548
	6		0,8722	2,1126	8,1694	11,1542	11,1542
3	7		1	0,3217	3,0651	4,3868	4,3868
	8			1,8320	5,4960	7,3280	7,3280
4	9			1	3,0000	4,0000	4,0000
	10		1		2,1000	3,1000	3,1000
	11		1		1,5000	2,5000	2,5000

2. Усі числа першого рядка, крім рядкової суми (стовпець 8), ділимо на $a_{11} = 2,50$. Результати ділення записуємо в четвертий рядок. Усі проміжні обчислення виконуємо з двома запасними десятковими розрядами. Рядкова сума $1 + 0,376 + 0,144 + 2,7216 = 4,2416$ збігається з контрольною сумою $10,604 : 2,50 = 4,2416$, а це означає, що випадкових помилок немає.

3. Обчислюємо коефіцієнти $a_{ij}^{(1)}$ ($i = 2,3; j = 2,3,4$) системи (3.7) за формулами (3.8) і записуємо їх у 5-й і 6-й рядки. Контрольні суми

$$12,345 - 0,87 \cdot 4,2416 = 12,345 - 3,690 = 8,6548,$$

$$12,257 - 0,26 \cdot 4,2416 = 12,257 - 1,103 = 11,1542$$

збігаються з відповідними рядковими сумами

$$1,9729 + 0,6347 + 6,0472 = 8,6548,$$

$$0,8722 + 2,1126 + 8,1694 = 11,1542,$$

а це означає, що обчислення виконано правильно.

4. Усі числа рядка 5, крім числа, що стоїть у 8-му стовпці, ділимо на $a_{22}^{(1)} = 1,9729$, і результати записуємо в 7-й рядок. Контрольна сума дорівнює $8,6548 : 1,9729 = 4,3868$, і вона збігається з рядковою сумою $1 + 0,3217 + 3,0651 = 4,3868$.

5. Коефіцієнт $a_{ij}^{(2)}$ ($i = 3; j = 3, 4$) рівняння (3.11) обчислюємо за формулами (3.12) і записуємо у 8-й рядок. Контрольна сума $11,1542 - 4,3868 \cdot 0,8722 = 7,3280$ збігається з рядковою $1,8320 + 5,4960 = 7,3280$.

6. Зворотний хід виконуємо за формулами (3.16). При цьому використовуються 8-й, 7-й і 4-й рядки табл. 3.1. З 8-го рядка дістаємо:

$$x_3 = 5,4960 : 1,8320 = 3,0000,$$

$$\bar{x}_3 = 7,3280 : 1,8320 = 4,0000.$$

З 7-го рядка маємо:

$$x_2 = 3,0651 - 0,3217 \cdot 3 = 2,1000,$$

$$\bar{x}_2 = 4,3868 - 0,3217 \cdot 4 = 3,1000.$$

Нарешті, з 4-го рядка знаходимо:

$$x_1 = 2,7216 - 0,144 \cdot 3 - 0,376 \cdot 2,1 = 1,5000,$$

$$\bar{x}_1 = 4,2416 - 0,144 \cdot 4 - 0,376 \cdot 3,1 = 2,5000.$$

Обчислені значення \bar{x}_j і x_j зв'язані між собою співвідношенням (3.19), що свідчить про відсутність випадкових обчислювальних помилок.

Безпосередньо підстановкою x_1 , x_2 і x_3 у систему (3.23) переконуємося, що її розв'язано точно: всі нев'язки дорівнюють нулю. Результати розв'язання системи за схемою єдиного ділення зведено в табл. 3.2.

Корені лінійних систем алгебраїчних рівнянь за методом Гаусса на сучасних ЕОМ обчислюють за спеціальними стандартними програмами. Такі програми записують різними мовами програмування. Нижче подано програму 3.1 методу Гаусса, записану мовою Бейсік для ЕОМ "Ямаха".

```

10 REM ----- МЕТОД ГАУССА-----
20 INPUT "Ввести розмірність квадратної матриці"; N
30 DIM A(N,N),B(N),X(N),C(N,N),G(N)
40 PRINT "Ввести елементи матриці по одному за рядками"
50 FOR I=1 TO N
60   FOR J=1 TO N
70     INPUT A(I,J)
80   NEXT J
90 NEXT I
100 PRINT "Ввести вільні члени по одному"
110 FOR I=1 TO N
120   INPUT B(I)
130 NEXT I
140 ' П Р Я М И Й Х І Д
150 N1=N-1
160 FOR K=1 TO N1
170   IF ABS(A(K,K))>0 THEN 300
180   K1=K+1
190   FOR M=K1 TO N
200     IF ABS(A(M,K))>0 THEN 270
210     FOR L=1 TO N
220       V=A(K,L)
230       A(K,L)=A(M,L)
240       A(M,L)=V
250     NEXT L
260     V=B(K)
270     B(K)=B(M)
280     B(M)=V
290   NEXT M
297 IF ABS(A(K,K))>0 THEN 300 ELSE 396
300   G(K)=B(K)/A(K,K)
310   K1=K+1
320   FOR I=K1 TO N
330     B(I)=B(I)-A(I,K)*G(K)
340   FOR J1=K TO N
350     J=N-J1+K
360     C(K,J)=A(K,J)/A(K,K)
370     A(I,J)=A(I,J)-A(I,K)*C(K,J)
380   NEXT J1
390 NEXT I
395 NEXT K
396   FOR I=1 TO N
397     FOR J=1 TO N
398       PRINT A(I,J);

```

Продовження програми.3.1

```

399 NEXT J
400 PRINT B(I):PRINT
401 NEXT I
409 ' Кінець прямого ходу
410 ' Зворотний хід
420 M=N: PR=0
425 IF A(M,M)<>0 THEN 430
426 IF B(M)<>0 THEN PRINT "Система несумісна": GOTO 500
427 IF PR=1 THEN PRINT "Покладемо x("M")=0":X(M)=1:
      M=M-1: GOTO 425
428 PRINT "Система має множину розв'язків":X(M)=0:M=M-1
429 PR=1: GOTO 425
430 X(M)=B(M)/A(M,M)
440 M=M-1
450 S=0
460 FOR L=M TO N1
470 S=S+C(M,L+1)*X(L+1)
475 NEXT L
480 X(M)=G(M)-S
490 IF M>1 THEN 440
492 FOR I=1 TO N
493 PRINT "X("I")=";X(I)
494 NEXT I
500 END

```

§ 3.3. ОБЧИСЛЕННЯ ОБЕРНЕНОЇ МАТРИЦІ МЕТОДОМ ГАУССА

Якщо треба розв'язати кілька систем лінійних рівнянь, що мають однаково невироджену матрицю коефіцієнтів, то це доцільно зробити за схемою єдиного ділення, розміщуючи вільні члени цих систем у сусідніх стовпцях праворуч від матриці коефіцієнтів. Тоді контрольна сума кожного рядка схеми буде сумою всіх елементів відповідного рядка розширеної матриці. Поточний контроль у процесі обчислень здійснюється (як і при розв'язуванні однієї системи) порівнянням між собою контрольних і рядкових сум. Над контрольною сумою кожного рядка слід виконувати ті самі перетворення, що й над іншими його елементами.

До одночасного розв'язування кількох систем лінійних рівнянь з тією самою невиродженою матрицею коефіцієнтів $A = (a_{ij})$ веде задача обчислення елементів матриці $A^{-1} = (x_{ij})$, оберненої до A . Оскільки

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

де E — одинична матриця, то, перемноживши матриці A і A^{-1} і прирівнявши кожну суму добутоків до відповідного елемента одиничної матриці E , дістанемо n систем рівнянь з n змінними x_{ij} ($i, j = 1, 2, \dots, n$) кожна.

Якщо, наприклад, помножити кожний рядок матриці A на перший стовпець матриці A^{-1} і прирівняти суму цих добутоків до відповідного елемента першого стовпця матриці E , то дістанемо систему вигляду

$$\begin{cases} a_{11}x_{11} + a_{12}x_{21} + \dots + a_{1n}x_{n1} = 1, \\ a_{21}x_{11} + a_{22}x_{21} + \dots + a_{2n}x_{n1} = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_{11} + a_{n2}x_{21} + \dots + a_{nn}x_{n1} = 0. \end{cases}$$

Аналогічно, помноживши кожний рядок матриці A на другий стовпець матриці A^{-1} і прирівнявши кожну з цих сум добутоків до відповідного елемента другого стовпця матриці E , дістанемо систему

$$\begin{cases} a_{11}x_{12} + a_{12}x_{22} + a_{13}x_{32} + \dots + a_{1n}x_{n2} = 0, \\ a_{21}x_{12} + a_{22}x_{22} + a_{23}x_{32} + \dots + a_{2n}x_{n2} = 1, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}x_{12} + a_{n2}x_{22} + a_{n3}x_{32} + \dots + a_{nn}x_{n2} = 0 \end{cases}$$

і т.д. Ці n систем рівнянь можна записати компактно в такому вигляді:

$$\sum_{j=1}^n a_{ij}x_{jk} = \delta_{ik} \quad (i, k = 1, 2, \dots, n),$$

де

$$\delta_{ik} = \begin{cases} 1, & \text{якщо } i = k, \\ 0, & \text{якщо } i \neq k. \end{cases}$$

Отже, системи мають одну й ту саму матрицю коефіцієнтів, але відрізняються одна від одної вільними членами. Тому їх можна розв'язувати одночасно в одній схемі, вводючи до неї n різних стовпців вільних членів. На конкретному прикладі покажемо, як це безпосередньо здійснюється.

П р и к л а д 2. Знайти обернену матрицю A^{-1} до матриці

$$A = \begin{pmatrix} 3,10 & 1,20 & 0,62 \\ 1,12 & 2,60 & 0,85 \\ 0,82 & 1,20 & 2,54 \end{pmatrix},$$

якщо елементи матриці A — точні числа.

Р о з в ' я з а н н я. Елементи оберненої матриці обчислимо методом Гаусса за схемою єдиного ділення. Обчислення зведемо в табл. 3.3, яка має три стовпці вільних членів (з 6-го по 8-й включно). Елементи оберненої матриці A^{-1} розміщено в тих самих стовпцях трьох останніх рядків, але в зворотному порядку: перший рядок оберненої матриці записано в 11-му рядку таблиці, другий — у 10-му і третій — у 9-му.

Таблиця 3.3

1	2	3	4	5	6	7	8	9	10
1	1	3,10	1,20	0,62	1	0	0	5,92	5,92
	2	1,12	2,60	0,85	0	1	0	5,57	5,57
	3	0,82	1,20	2,54	0	0	1	5,56	5,56
2	4	1	0,387097	0,2	0,322581	0	0	1,909677	1,909678
	5		2,166451	0,626	-0,361291	1	0	3,431162	3,431160
	6		0,882580	2,376	-0,264516	0	1	3,994065	3,994064
3	7		1	0,288952	-0,166766	0,461584	0	1,583771	1,583770
	8			2,120977	-0,117332	-0,407385	1	2,596260	2,596260
4	9			1	-0,055320	-0,192074	0,471481	1,224087	1,224087
	10		1		-0,150781	0,517084	-0,136235	1,230069	1,230068
	11	1			0,392012	-0,161747	-0,0415601	1,188704	1,188705

Отже, обернена матриця має такий вигляд:

$$A^{-1} = \begin{pmatrix} 0,392012 & -0,161747 & -0,0415601 \\ -0,150781 & 0,517084 & -0,136235 \\ -0,055320 & -0,192074 & 0,471481 \end{pmatrix}$$

Щоб дістати уявлення про те, з якою точністю обчислено обернену матрицю, треба провести заключний контроль. Для цього досить обчислити добуток цих матриць. Маємо

$$A \cdot A^{-1} = \begin{pmatrix} 1,0000016 & -0,0000007 & 0,0000002 \\ 0,0000008 & 0,9999989 & 0,0000006 \\ -0,0000002 & 0,0000004 & 1,0000005 \end{pmatrix}$$

Звідси видно, що внаслідок округлення елементи оберненої матриці обчислено не точно. В обчислювальній математиці використовують алгоритми для виправлення (уточнення) елементів наближеної оберненої матриці.

§ 3.4. МЕТОД КВАДРАТНИХ КОРЕНІВ

Цей метод використовують для знаходження розв'язку лінійної системи рівнянь

$$Ax = b, \quad (3.24)$$

в якій матриця $A = (a_{ij})$ симетрична, тобто елементи, симетричні відносно головної діагоналі, рівні між собою: $a_{ij} = a_{ji}$ ($i, j = 1, 2, \dots, n$). Відомо, що симетричну матрицю A завжди можна подати у вигляді добутку двох взаємно транспонованих трикутних матриць

$$A = T'T, \quad (3.25)$$

де

$$T = \begin{pmatrix} t_{11} & t_{12} & t_{13} & \dots & t_{1n} \\ 0 & t_{22} & t_{23} & \dots & t_{2n} \\ 0 & 0 & t_{33} & \dots & t_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & t_{nn} \end{pmatrix};$$

$$T' = \begin{pmatrix} t_{11} & 0 & 0 & \dots & 0 \\ t_{12} & t_{22} & 0 & \dots & 0 \\ t_{13} & t_{23} & t_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ t_{1n} & t_{2n} & t_{3n} & \dots & t_{nn} \end{pmatrix}.$$

Якщо тепер перемножити матриці T' і T , а потім прирівняти відповідні елементи матриць у рівності (3.25), то для знаходження $\frac{n(n+1)}{2}$ елементів t_{ij} ($i = 1, 2, \dots, n; j = i, i+1, \dots, n$) матриці T дістанемо систему $\frac{n(n+1)}{2}$ рівнянь

$$\begin{cases} t_{11}^2 + t_{21}^2 + \dots + t_{i1}^2 = a_{11} & (i = 1, 2, \dots, n) \\ t_{1i}t_{1j} + t_{2i}t_{2j} + \dots + t_{ij}t_{ij} = a_{ij} & i = 1, 2, \dots, n; j = i+1, i+2, \dots, n; i < j. \end{cases}$$

З цієї системи знаходимо послідовно елементи матриці T (і T'). Маємо

$$\begin{cases} t_{11} = \sqrt{a_{11}}, & t_{1j} = \frac{a_{1j}}{t_{11}} \quad (j = 2, 3, \dots, n), \\ t_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} t_{ki}^2} & (1 < i \leq n), \\ t_{ij} = (a_{ij} - \sum_{k=1}^{i-1} t_{ki}t_{kj}) / t_{ii} & (i < j), \\ t_{ij} = 0 & (i > j). \end{cases} \quad (3.26)$$

З рівності (3.25) випливає, що система (3.24) рівносильна двом системам рівнянь з трикутними матрицями $T'y = b$ і $Tx = y$.

Розв'язавши систему $T'y = b$ з нижньою трикутною матрицею T' , знайдемо

$$y_1 = \frac{b_1}{t_{11}}, \quad y_i = \frac{b_i - \sum_{k=1}^{i-1} t_{ki} y_k}{t_{ii}} \quad (1 < i \leq n). \quad (3.27)$$

Розв'язавши потім систему $Tx = y$ з верхньою трикутною матрицею T , знайдемо шуканий розв'язок системи (3.24)

$$x_n = \frac{y_n}{t_{nn}}, \quad x_i = \frac{y_i - \sum_{k=i+1}^n t_{ik}x_k}{t_{ii}} \quad (1 \leq i < n). \quad (3.28)$$

Всі обчислення за формулами (3.26)–(3.28) доцільно виконувати за спеціальною схемою (табл. 3.4), в якій забезпечується проміжний і заключний контроль введенням контрольних і рядкових сум. У методі квадратних коренів, як і в методі Гаусса, поряд з системою (3.24) одночасно розв'язують допоміжну систему

$$A\bar{x} = s. \quad (3.29)$$

Таблиця 3.4

Крок перетворення	Рядок	Коефіцієнти при змінних				Вільний член	Контроль	
		x_1	x_2	...	x_n		контрольна сума	рядкова сума
1	2	3	4	...	$n+2$	$n+3$	$n+4$	$n+5$
1	1	a_{11}	a_{12}		a_{1n}	b_1	s_1	
	2		a_{22}		a_{2n}	b_2	s_2	
	
	n				a_{nn}	b_n	s_n	
2	$n+1$	t_{11}	t_{12}	...	t_{1n}	y_1	s_1	σ_1
	$n+2$		t_{22}	...	t_{2n}	y_2	s_2	σ_2

	$2n$				t_{nn}	y_n	s_n	σ_n
3	$2n+1$				1	x_n	\bar{x}_n	$1+x_n$

	$3n-1$		1		...	x_2	\bar{x}_2	$1+x_2$
	$3n$	1			...	x_1	\bar{x}_1	$1+x_1$

Системи (3.24) і (3.29) мають однакову матрицю коефіцієнтів A , але різні вільні члени: у системі (3.24) — це числа b_i ($i = 1, 2, \dots, n$), а в системі (3.29) — числа

$$s_i = b_i + \sum_{k=1}^n a_{ik} \quad (i = 1, 2, \dots, n) \quad (3.30)$$

Розв'язки цих систем зв'язані співвідношенням:

$$\bar{x}_i = x_i + 1 \quad (i = 1, 2, \dots, n). \quad (3.31)$$

Оскільки система (3.29) рівносильна двом системам з трикутними матрицями $T'z = s$ і $T\bar{x} = z$, то елементи вектора z обчислюють за формулами

$$z_i = \frac{s_i}{t_{ii}}, \quad z_i = \frac{s_i - \sum_{k=1}^{i-1} t_{ik}z_k}{t_{ii}} \quad (1 < i \leq n), \quad (3.32)$$

а елементи вектора \bar{x} — за формулами

$$\bar{x}_n = \frac{z_n}{t_{nn}}, \quad \bar{x}_i = \frac{z_i - \sum_{k=i+1}^n t_{ik}\bar{x}_k}{t_{ii}} \quad (1 \leq i < n). \quad (3.33)$$

Поточний контроль здійснюють порівнянням контрольних сум z_i (стовпець $n+4$) з рядковими сумами σ_i (стовпець $n+5$), які обчислюють за формулами

$$\sigma_i = y_i + \sum_{k=i}^n t_{ik} \quad (i = 1, 2, 3, \dots, n). \quad (3.34)$$

Якщо обчислення виконано правильно, то суми збігаються, або внаслідок округлення проміжних обчислень відрізняються між собою на 1–2 одиниці нижчого розряду. Всі проміжні обчислення доцільно виконувати з 1–2 запасними цифрами.

Заключний контроль, як і в методі Гаусса, можна здійснити двоюко. Або перевірити виконання рівностей (3.31), або (і) обчислити нев'язки, підставивши знайдений розв'язок у систему (3.24).

Розрахункова таблиця 3.4 складається з трьох частин. У першій частині записано коефіцієнти і вільні члени системи (3.24), а також обчислені за формулами (3.30) контрольні суми s_i ($i = 1, 2, \dots, n$) (рядкові суми можна не записувати, бо вони збігаються з контрольними); у другій знайдені за формулами (3.26) коефіцієнти t_{ij} ($i = 1, 2, \dots, n; j = i+1, i+2, \dots, n$) матриці T , обчислені за формулами (3.27) і (3.32) вектори y і z , а також обчислені за формулами (3.34) рядкові суми σ_i ($i = 1, 2, \dots, n$); у третій — обчислені за формулами (3.28) і (3.33) вектори x і \bar{x} . Дві перші частини — це прямий хід, а третя — зворотний.

П р и к л а д 3. Методом квадратних коренів розв'язати систему рівнянь

$$\begin{cases} 3,45x_1 + 0,78x_2 - 0,97x_3 = 3,229, \\ 0,78x_1 + 2,63x_2 - 0,89x_3 = 4,026, \\ -0,97x_1 - 0,89x_2 + 2,41x_3 = 5,030, \end{cases} \quad (3.35)$$

коефіцієнти і вільні члени якої точні числа.

Р о з в ' я з а н н я. Всі проміжні обчислення виконуємо з шістьма

десятковими знаками, а остаточний результат подамо з п'ятьма десятковими знаками.

Прямий хід. Спочатку в перші три рядки табл. 3.5 заносимо коефіцієнти системи (3.35), а в 7-й стовпець заносимо суми коефіцієнтів

Таблиця 3.5

1	2	3	4	5	6	7	8
1	1	3,45	0,78	-0,97	3,229	6,489	
	2		2,63	-0,89	4,026	6,546	
	3			2,41	5,030	5,580	
2	4	1,857418	0,419938	-0,522230	1,738435	3,493559	3,493561
	5		1,566414	-0,428173	2,104147	3,242388	3,242388
	6			1,397835	4,892424	6,290259	6,290259
3	7			1	3,500001	4,500001	4,500001
	8		1		2,300000	3,300000	3,300000
	9	1			1,399999	2,399998	2,399999

і вільних членів кожного рядка (контрольні суми), обчислені за формулами (3.30). Далі переходимо до обчислення за формулами (3.26) елементів t_{ik} матриці T . Для $n=3$ ці формули набирають вигляду

$$t_{11} = \sqrt{a_{11}}, \quad t_{12} = \frac{a_{11}}{t_{11}}, \quad t_{13} = \frac{a_{13}}{t_{11}},$$

$$t_{22} = \sqrt{a_{22} - t_{12}^2}, \quad t_{23} = \frac{a_{23} - t_{12}t_{13}}{t_{22}}, \quad t_{33} = \sqrt{a_{33} - t_{13}^2 - t_{23}^2}.$$

За цими формулами знаходимо:

$$t_{11} = 1,857418, \quad t_{12} = 0,419938, \quad t_{13} = -0,522230,$$

$$t_{22} = 1,566414, \quad t_{23} = -0,428173, \quad t_{33} = 1,397835.$$

Якщо $n=3$, то з формул (3.27) і (3.32) для обчислення елементів векторів y і z дістаємо

$$y_1 = \frac{b_1}{t_{11}}, \quad y_2 = \frac{b_2 - t_{12}y_1}{t_{22}}, \quad y_3 = \frac{b_3 - t_{13}y_1 - t_{23}y_2}{t_{33}},$$

$$z_1 = \frac{s_1}{t_{11}}, \quad z_2 = \frac{s_2 - t_{12}z_1}{t_{22}}, \quad z_3 = \frac{s_3 - t_{13}z_1 - t_{23}z_2}{t_{33}}.$$

Звідси маємо

$$y_1 = 1,738435, \quad y_2 = 2,104147, \quad y_3 = 4,892424;$$

$$z_1 = 3,493559, \quad z_2 = 3,242388, \quad z_3 = 6,290259.$$

Поточний контроль здійснюємо, порівнюючи значення контрольних (стовпець 7) і рядкових (стовпець 8) сум. Рядкові суми обчислюють за формулами (3.34), які для $n=3$ набирають вигляду

$$t_{11} + t_{12} + t_{13} + y_1 = \sigma_1,$$

$$t_{22} + t_{23} + y_2 = \sigma_2,$$

$$t_{33} + y_3 = \sigma_3.$$

Найбільша розбіжність між контрольною і рядковою сумами дорівнює двом одиницям шостою десяткового розряду, що є результатом округлення проміжних результатів і цілком допустимо.

Зворотний хід. З формул (3.28) і (3.33) для обчислення елементів векторів x і \bar{x} для $n=3$ дістаємо

$$x_3 = \frac{y_3}{t_{33}}, \quad x_2 = \frac{y_2 - t_{23}x_3}{t_{22}}, \quad x_1 = \frac{y_1 - t_{13}x_3 - t_{12}x_2}{t_{11}},$$

$$\bar{x}_3 = \frac{z_3}{t_{33}}, \quad \bar{x}_2 = \frac{z_2 - t_{23}\bar{x}_3}{t_{22}}, \quad \bar{x}_1 = \frac{z_1 - t_{13}\bar{x}_3 - t_{12}\bar{x}_2}{t_{11}}.$$

Користуючись цими формулами, знаходимо

$$x_3 = 3,500001, \quad x_2 = 2,300000, \quad x_1 = 1,399999;$$

$$\bar{x}_3 = 4,500001, \quad \bar{x}_2 = 3,300000, \quad \bar{x}_1 = 2,399998.$$

Контроль обчислень здійснюємо перевіркою співвідношень (3.30). Всі обчислення зведемо в табл. 3.5.

З табл. 3.5 видно, що з точністю до $0,5 \cdot 10^{-5}$ $x_1 = 1,40000$, $x_2 = -2,30000$, $x_3 = 3,50000$. Підставивши ці значення в систему (3.35) впевнюємось, що вони задовольняють систему точно, адже відповідні нев'язки дорівнюють нулю.

§ 3.5. МЕТОД ПРОСТОЇ ІТЕРАЦІЇ

Застосування методу Гаусса для розв'язування системи лінійних рівнянь з великою кількістю невідомих досить громіздке. Крім того, кількість невідомих може бути така велика, що коефіцієнти системи не завжди можна розмістити в оперативній пам'яті ЕОМ. Тоді застосувати для її розв'язування метод Гаусса взагалі не можна. У цих випадках розв'язують систему ітераційними методами.

Розглянемо метод простої ітерації.

Нехай задано систему лінійних рівнянь

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n, \end{cases} \quad (3.36)$$

або в матричному вигляді

$$Ax = b, \quad (3.37)$$

де

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

Нехай діагональні елементи a_{ii} ($i = 1, 2, \dots, n$) матриці A відмінні від нуля. Тоді, розв'язавши перше рівняння системи (3.36) відносно x_1 , а друге — відносно x_2 і т.д., дістанемо систему

$$\begin{cases} x_1 = \alpha_{12}x_2 + \alpha_{13}x_3 + \dots + \alpha_{1n}x_n + \beta_1, \\ x_2 = \alpha_{21}x_1 + \alpha_{23}x_3 + \dots + \alpha_{2n}x_n + \beta_2, \\ \dots \\ x_n = \alpha_{n1}x_1 + \alpha_{n2}x_2 + \dots + \alpha_{n,n-1}x_{n-1} + \beta_n, \end{cases} \quad (3.38)$$

де

$$\alpha_{ij} = \begin{cases} -\frac{a_{ij}}{a_{ii}}, & i \neq j, \\ 0, & i = j; \end{cases} \quad \beta_i = \frac{b_i}{a_{ii}}.$$

Ввівши до розгляду матриці

$$\alpha = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix},$$

систему (3.38) запишемо у вигляді

$$x = \alpha x + \beta. \quad (3.39)$$

Систему (3.39) називають системою *нормального виду*.

Розв'яжемо її методом послідовних наближень. За початкове наближення візьмемо, наприклад, стовпець вільних членів, тобто $x^{(0)} = \beta$.

Тоді послідовно знаходимо

$$x^{(k+1)} = \alpha x^{(k)} + \beta \quad (k = 0, 1, 2, \dots) \quad (3.40)$$

або в розгорнутому вигляді

$$\begin{aligned} x_i^{(0)} &= \beta_i, \\ x_i^{(k+1)} &= \sum_{j=1}^n \alpha_{ij} x_j^{(k)} + \beta_i, \quad i = 1, 2, \dots, n; \quad k = 0, 1, 2, \dots \end{aligned} \quad (3.41)$$

Якщо послідовність наближень $x^{(0)}, x^{(1)}, \dots, x^{(k)}, \dots$ має границю $x^* = \lim_{k \rightarrow \infty} x^{(k)}$, то ця границя і буде розв'язком системи (3.39).

Справді, перейшовши до границі, коли $k \rightarrow \infty$, у рівності (3.40), маємо

$$\lim_{k \rightarrow \infty} x^{(k+1)} = \lim_{k \rightarrow \infty} (\alpha x^{(k)} + \beta) = \alpha \lim_{k \rightarrow \infty} x^{(k)} + \lim_{k \rightarrow \infty} \beta,$$

або

$$x^* = \alpha x^* + \beta.$$

Таким чином, вектор

$$x^* = \begin{pmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{pmatrix},$$

є розв'язком системи (3.39), а отже, і системи (3.36).

Метод послідовних наближень, який визначається формулами (3.40) або (3.41), називається методом *прості ітерації* або просто *методом ітерації*.

Праву частину рівності (3.39) можна розглядати як задання деякого оператора φ , який відображає вектор $x = (x_1, x_2, \dots, x_n)$ n -вимірного векторного простору у вектор у цього самого простору. При цьому

$$y_i = \sum_{j=1}^n \alpha_{ij} x_j + \beta_i, \quad i = 1, 2, \dots, n.$$

Тому знаходження розв'язку системи (3.39) зводиться до відшукування нерухомої точки відображення φ , яке визначається формулою

$$\varphi(x) = \alpha x + \beta. \quad (3.42)$$

Якщо відображення φ є стискующим, то розв'язок рівняння $x = \varphi(x)$, тобто рівняння (3.39), можна знайти методом послідовних наближень за формулами (3.40), починаючи з довільного вектора $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$.

За яких умов відображення φ буде стискующим? Відповідь на це питання залежить не тільки від самого відображення φ , а й від вибору метрики n -вимірного векторного простору.

Розглянемо три варіанти метрики.

1. Нехай відстань між векторами $x = (x_1, x_2, \dots, x_n)$ і $x' = (x_1', x_2', \dots, x_n')$ визначається формулою

$$\rho_0(x, x') = \sqrt{\sum_{i=1}^n (x_i - x_i')^2} \quad (3.43)$$

Множина усіх n -вимірних векторів з метрикою (3.43) є повним метричним простором R_n . Розглянемо $\rho(\varphi(x), \varphi(x'))$:

$$\rho_0(\varphi(x), \varphi(x')) = \sqrt{\sum_{i=1}^n \left(\sum_{j=1}^n \alpha_{ij} x_j + \beta_i - \sum_{j=1}^n \alpha_{ij} x_j' - \beta_i \right)^2} =$$

$$= \sqrt{\sum_{i=1}^n \left(\sum_{j=1}^n \alpha_{ij} (x_j - x'_j) \right)^2}.$$

Враховавши нерівність Коші, маємо

$$\begin{aligned} \sqrt{\sum_{i=1}^n \left(\sum_{j=1}^n \alpha_{ij} (x_j - x'_j) \right)^2} &\leq \sqrt{\sum_{i=1}^n \sum_{j=1}^n \alpha_{ij}^2 \sum_{j=1}^n (x_j - x'_j)^2} = \\ &= \sqrt{\sum_{i=1}^n \sum_{j=1}^n \alpha_{ij}^2} \sqrt{\sum_{j=1}^n (x_j - x'_j)^2} = \\ &= \rho_0(x, x') \sqrt{\sum_{i=1}^n \sum_{j=1}^n \alpha_{ij}^2} \end{aligned}$$

Якщо елементи матриці α задовольняють умову

$$l_0^2 = \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij}^2 < 1, \quad (3.44)$$

то відображення φ буде стискующим, і з принципу стискуючих відображень випливає існування єдиного розв'язку системи (3.39), який є границею послідовності, побудованої за формулами (3.40), починаючи з довільного початкового вектора $x^{(0)}$.

2. Нехай метрику введено згідно з формулою:

$$\rho_1(x, x') = \max_{1 \leq i \leq n} |x_i - x'_i|, \quad (3.45)$$

де $x = (x_1, x_2, \dots, x_n)$, $x' = (x'_1, x'_2, \dots, x'_n)$.

Відомо, що n -вимірний векторний простір з метрикою (3.45) є повним метричним простором. Розглянемо відстань $\rho(\varphi(x), \varphi(x'))$:

$$\begin{aligned} \rho(\varphi(x), \varphi(x')) &= \max_{1 \leq i \leq n} \left| \sum_{j=1}^n \alpha_{ij} x_j - \sum_{j=1}^n \alpha_{ij} x'_j \right| = \\ &= \max_{1 \leq i \leq n} \left| \sum_{j=1}^n \alpha_{ij} (x_j - x'_j) \right| \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |\alpha_{ij}| |x_j - x'_j| \leq \\ &\leq \max_{1 \leq i \leq n} \sum_{j=1}^n |\alpha_{ij}| \max_{1 \leq j \leq n} |x_j - x'_j| = \rho_1(x, x') \max_{1 \leq i \leq n} \sum_{j=1}^n |\alpha_{ij}|. \end{aligned}$$

Якщо елементи матриці α задовольняють умову

$$l_1 = \max_{1 \leq i \leq n} \sum_{j=1}^n |\alpha_{ij}| < 1, \quad (3.46)$$

то $\rho_1(\varphi(x), \varphi(x')) \leq l_1 \rho_1(x, x')$, $0 < l_1 < 1$), тобто відображення (3.42) є стискующим у розумінні метрики (3.45).

Тоді з принципу стискуючих відображень випливає існування і єдиність розв'язку $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ системи (3.39), а також збіжність до x^* послідовності $\{x^{(k)}\}$ наближень, побудованої за формулою (3.40).

3. Визначимо тепер відстань між векторами x і x' за формулою

$$\rho_2(x, x') = \sum_{i=1}^n |x_i - x'_i| \quad (3.47)$$

і розглянемо

$$\begin{aligned} \rho_2(\varphi(x), \varphi(x')) &= \sum_{i=1}^n \left| \sum_{j=1}^n \alpha_{ij} x_j - \sum_{j=1}^n \alpha_{ij} x'_j \right| = \\ &= \sum_{i=1}^n \left| \sum_{j=1}^n \alpha_{ij} (x_j - x'_j) \right| \leq \sum_{i=1}^n \sum_{j=1}^n |\alpha_{ij}| |x_j - x'_j| = \\ &= \sum_{j=1}^n \left(\sum_{i=1}^n |\alpha_{ij}| \right) |x_j - x'_j| \leq \\ &\leq \max_{1 \leq j \leq n} \left(\sum_{i=1}^n |\alpha_{ij}| \right) \sum_{j=1}^n |x_j - x'_j| = \\ &= \rho_2(x, x') \max_{1 \leq j \leq n} \sum_{i=1}^n |\alpha_{ij}|. \end{aligned}$$

Якщо елементи матриці α задовольняють умову

$$l_2 = \max_{1 \leq j \leq n} \sum_{i=1}^n |\alpha_{ij}| < 1, \quad (3.48)$$

то відображення (3.42) n -вимірного векторного простору в себе — стискує. Оскільки n -вимірний векторний простір з метрикою (3.47) є повним метричним простором, то з принципу стискуючих відображень випливає, що відображення (3.42) має єдину нерухому точку, яка є границею послідовності, побудованої за формулою (3.40).

Таким чином, доведено теорему.

Теорема 1. Якщо матриця α системи рівнянь (3.39) задовольняє одну з умов (3.44), (3.46) або (3.48), то система (3.39) має єдиний розв'язок $x^* = (x_1^*, x_2^*, \dots, x_n^*)$, який можна дістати як границю послідовності $\{x^{(k)}\}$, побудованої за формулою (3.40), починаючи з довільного початкового вектора $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$.

Із загальних оцінок для похибки n -го наближення методу послідовних наближень випливають аналогічні оцінки для наближень, знайде-

них методом простої ітерації. Для цього треба підставити у формули загальних оцінок вирази для метрики конкретних просторів.

Наприклад, з формули (2.10) для простору R_n^0 випливає оцінка

$$\max_{1 \leq i \leq n} |x_i^{(k)} - x_i^*| \leq \frac{l_1^k}{1 - l_1} \max_{1 \leq i \leq n} |x_i^{(1)} - x_i^{(0)}|,$$

де l_1 визначається формулою (3.46).

З формули (2.11) випливає, що метод простої ітерації слід закінчити, якщо стане справедливою нерівність

$$\max_{1 \leq i \leq n} |x_i^{(k)} - x_i^{(k-1)}| \leq \frac{1 - l_1}{l_1} \varepsilon, \quad (3.49)$$

де ε — наперед задана точність наближень.

Аналогічні оцінки можна знайти і для просторів R_n і R_n^1 .

З теореми 1 легко дістати достатні умови збіжності методу простої ітерації для системи (3.37).

Теорема 2. Якщо елементи матриці A задовольняють одну з умов

$$\sum_{i=1}^n \frac{1}{a_{ii}} \sum_{j=1}^n a_{ij}^2 < 1, \quad \sum_{j=1}^n |a_{ij}| < |a_{ii}|, \quad i = 1, 2, \dots, n;$$

або

$$\max_{1 \leq j \leq n} \sum_{i=1}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1,$$

то система рівнянь (3.37) має єдиний розв'язок $x^* = (x_1^*, x_2^*, \dots, x_n^*)$, який можна дістати як границю послідовності $\{x^{(k)}\}$, побудованої за формулою $x_i^{(k)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^n a_{ij} x_j^{(k-1)} \right)$, $i = 1, 2, \dots, n$; $k = 1, 2, \dots$, починаючи з довільного початкового наближення $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$.

Оцінка швидкості збіжності для простору R_n^0 має вигляд

$$\max_{1 \leq i \leq n} |x_i^{(k)} - x_i^*| \leq \frac{q_1^k}{1 - q_1} \max_{1 \leq i \leq n} |x_i^{(1)} - x_i^{(0)}|,$$

де

$$q_1 = \max_{1 \leq i \leq n} \frac{1}{|a_{ii}|} \sum_{j=1}^n |a_{ij}|.$$

Аналогічні оцінки легко дістати і для просторів R_n і R_n^1 .

Зазначимо, що теореми 1 і 2 дають достатні умови збіжності і невиконання умов цих теорем ще не означає розбіжності методу простої ітерації.

Невироджену систему лінійних рівнянь тотожними перетвореннями завжди можна звести до системи, для якої виконуватимуться достатні умови збіжності методу ітерацій.

Наприклад, нехай маємо систему

$$\begin{cases} 5x_1 - x_2 + 2x_3 = 8, \\ 3x_1 - 2x_2 + 9x_3 = 18, \\ 6x_1 - 5x_2 + 2x_3 = 10, \end{cases}$$

для якої не виконуються умови теореми 2.

Перетворимо цю систему до вигляду, в якого модулі елементів головної діагоналі були б більшими за суму модулів інших елементів відповідних рядків. Побудуємо нову систему, в якій першим і третім рівняннями будуть відповідно перше і друге рівняння даної системи, коефіцієнти яких задовольняють умову

$$\sum_{j=1}^n |a_{ij}| < |a_{ii}|, \quad i = 1, 3.$$

Друге рівняння нової системи дістанемо, якщо від третього рівняння даної системи віднімемо перше. В результаті матимемо систему

$$\begin{cases} 5x_1 - x_2 + 2x_3 = 8, \\ x_1 - 4x_2 = 2, \\ 3x_1 - 2x_2 + 9x_3 = 18, \end{cases}$$

коефіцієнти якої задовольняють умови теореми 2.

Цю систему легко звести і до нормального виду

$$\begin{cases} x_1 = \frac{1}{5} x_2 - \frac{2}{5} x_3 + \frac{8}{5}, \\ x_2 = \frac{1}{4} x_1 - \frac{1}{2}, \\ x_3 = -\frac{1}{3} x_1 + \frac{2}{9} x_2 + 2, \end{cases} \quad (3.50)$$

для якої виконуються достатні умови збіжності методу ітерацій.

Метод ітерацій легко реалізується на ЕОМ. Алгоритм розв'язування системи виду (3.38) передбачає:

1. Обчислення величини $l = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$.

2. Перевірку умови $l < 1$. Якщо ця умова не виконується, то процес обчислень закінчується і видається повідомлення про те, що метод застосувати не можна.

3. Обчислення допустимої похибки $\varepsilon_1 = \frac{1-l}{l} \varepsilon$.
 4. Вибір початкового наближення $x_i^{(0)} = \beta_i$ ($i = 1, 2, \dots, n$).
 5. Обчислення наступного наближення y_i ($i = 1, 2, \dots, n$) через попередні x_i ($i = 1, 2, \dots, n$).
 6. Перевірку умови $\max_{1 \leq i \leq n} |y_i - x_i| < \varepsilon_1$. Якщо ця умова виконується, то процес ітерацій завершується, інакше переходять до виконання п.5.
- Програма 3.2 містить алгоритм методу простої ітерації.

Програма 3.2

```

10 ' ----- Метод простих -----
20 ' ----- ітерацій для системи x=Ax+b -----
30 INPUT "Ввести розмірність N системи"; N
40 DIM A(N,N), B(N), X(N), Y(N)
50 PRINT "Ввести по одному елементи матриці A"
60 L=0
70 FOR I=1 TO N
80   S=0
90   FOR J=1 TO N
100    PRINT "A("I"; "J")=";
110    INPUT A(I,J)
120    S=S+ABS(A(I,J))
130   NEXT J
140   IF S<1 THEN 170
150   PRINT "Метод незастосовний"
160   GOTO 460
170   IF L<S THEN L=S
180 NEXT I
190 INPUT "Ввести точність EPS"; EPS
200 PRINT "Ввести по одному вільні члени системи"
210 FOR I=1 TO N
220   PRINT "B("I")=";
230   INPUT B(I)
240   X(I)=B(I)
250 NEXT I
260 EPS=ABS((1-L)*EPS/L)
270 M=0
280 FOR I=1 TO N
290   S=0
300   FOR J=1 TO N
310    S=S+A(I,J)*X(J)
320   NEXT J
330   Y(I)=S+B(I)
340   C=ABS(Y(I)-X(I))

```

```

350 IF M<C THEN M=C
360 NEXT I
370 IF M<=EPS THEN 420
380 FOR I=1 TO N
390   X(I)=Y(I)
400 NEXT I
410 GOTO 270
420 PRINT "Шуканий розв'язок"
430 FOR I=1 TO N
440   PRINT "X("I")="; Y(I)
450 NEXT I
460 END

```

§ 3.6. МЕТОД ЗЕЙДЕЛЯ

Метод Зейделя — деяка модифікація методу простої ітерації. У методі простої ітерації при обчисленні компонентів $x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_n^{(k+1)}$ вектора-стовпця $x^{(k+1)}$ на $(k+1)$ -му кроці використовуються значення $x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}$ вектора-стовпця $x^{(k)}$, обчисленого в попередньому кроці. Метод Зейделя відрізняється від методу простої ітерації тільки тим, що при обчисленні $(k+1)$ -го наближення компоненти x_i враховуються значення x_1, x_2, \dots, x_{i-1} , обчислені на цьому ж кроці.

Формули для знаходження послідовних наближень мають вигляд

$$x_1^{(k+1)} = \sum_{j=1}^n a_{1j} x_j^{(k)} + \beta_1,$$

$$x_i^{(k+1)} = \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} + \sum_{j=i}^n a_{ij} x_j^{(k)} + \beta_i,$$

$$x_n^{(k+1)} = \sum_{j=1}^{n-1} a_{nj} x_j^{(k+1)} + a_{nn} x_n^{(k)} + \beta_n, \quad k = 0, 1, 2, \dots$$

Зазначимо, що достатні умови збіжності для методу простої ітерації справедливі й для методу Зейделя.

Програма методу Зейделя відрізняється від програми методу простої ітерації тільки фрагментом обчислення наступних наближень. У програмі методу простої ітерації необхідно одночасно зберігати усі попередні $x_i^{(k)}$ й наступні $x_i^{(k+1)}$ наближення, оскільки найбільшу різницю $|x_i^{(k+1)} - x_i^{(k)}|$ можна знайти тільки після закінчення кроку ітерації. Користуючись методом Зейделя, немає потреби зберігати всі знайдені наближення $x_i^{(k)}$ ($i = 1, 2, \dots, n$), оскільки вони відразу використовуються для знаходження $x_i^{(k+1)}$. Тому у програмі методу Зейделя змінну $x_i^{(k+1)}$

позначимо змінною u , яка зберігає обчислене значення $x_i^{(k+1)}$ до того часу, поки воно не присвоїться змінній $x_i^{(k)}$ (у програмі вона позначена $X(I)$).

Фрагмент програми 3.3 методу Зейделя починається з рядка 270, оскільки рядки 30–260 такі самі, як і в програмі методу простої ітерації.

Програма 3.3

```

10 ' ----- Фрагмент програми методу -----
20 ' ----- Зейделя для системи  $x=Ax+b$  -----
270 M=0
280 FOR I=1 TO N
290   S=0
300   FOR J=1 TO N
310     S=S+A(I,J)*X(J)
320   NEXT J
330   Y=S+B(I)
340   C=ABS(Y-X(I))
350   IF M<C THEN M=C: X(I)=Y
360 NEXT I
370 IF M<=EPS THEN 390
380 GOTO 270
390 PRINT "Шуканий розв'язок:"
400 FOR I=1 TO N
410   PRINT "x("I")="; X(I)
420 NEXT I
430 END

```

:' Друкування
:' шуканого
:' розв'язку

Метод Зейделя часто збігається швидше, ніж метод простої ітерації. Наприклад, розв'язок $x_1 = 0,052980$, $x_2 = -0,486755$, $x_3 = 2,874172$ системи (3.50) з точністю $\varepsilon = 10^{-6}$ методом простої ітерації знайдено за 20 кроків, тоді як методом Зейделя його обчислено за 15 кроків. Але так буває не завжди. Є системи лінійних рівнянь, для яких метод простої ітерації збігається швидше, ніж метод Зейделя. Наприклад, для системи

$$\begin{cases} x_1 = -0,5x_1 + 0,6x_2 + 1, \\ x_2 = -0,6x_1 - 0,5x_2 + 2 \end{cases} \quad (3.51)$$

метод ітерації збігається і її розв'язок $x_1 = 1,034483$, $x_2 = 0,919540$ з точністю $\varepsilon = 10^{-6}$ знайдено за 68 кроків, а метод Зейделя розбігається; для системи

$$\begin{cases} x_1 = 0,5x_1 + x_2 + 1, \\ x_2 = -x_1 + 0,5x_2 + 2, \end{cases} \quad (3.52)$$

метод ітерацій розбігається, а метод Зейделя збігається, і її розв'язок $x_1 = 2$, $x_2 = 0$ обчислено за 23 ітерації.

Перевага ітераційних методів перед точним методом Гаусса в тому, що машинний час, потрібний для обчислень методом Гаусса, пропорційний n^3 , а ітераційними методами він пропорційний n^2 на одну ітерацію. Тому, якщо для розв'язування лінійної системи n рівнянь з n змінними ітераційними методами потрібно менше n ітерацій, то ці методи мають перевагу перед методом Гаусса.

Практичні задачі часто ведуть до систем лінійних рівнянь, які містять багато нульових коефіцієнтів. У таких випадках ітераційні методи дають велику економію машинного часу, бо трикутна матриця, яку дістають методом Гаусса, вже не матиме нульових елементів.

До того ж похибки округлень при використанні методу Гаусса можуть призвести до хибного результату, тоді як незначні похибки, допущені при обчисленнях ітераційними методами, не впливають на кінцевий результат.

Слід зазначити, що ітераційні методи дають можливість значно скоротити обсяг пам'яті ЕОМ, потрібної для зберігання коефіцієнтів системи, оскільки для обчислення наближення x_i використовуються коефіцієнти тільки i -го рівняння. Ця обставина особливо важлива для розв'язування систем рівнянь, коефіцієнти яких не вміщуються в оперативній пам'яті ЕОМ.

ЛАБОРАТОРНА РОБОТА № 2

Тема. Чисельне розв'язування систем лінійних алгебраїчних рівнянь.
Завдання. 1. Знайти розв'язок системи рівнянь виду

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3. \end{cases}$$

використавши три методи: Гаусса, простої ітерації і Зейделя з точністю $\varepsilon = 0,5 \cdot 10^{-6}$.
Значення коефіцієнтів системи і вільні члени взяті з табл. 3.6.

Таблиця 3.6

Варіант	i	a_{i1}	a_{i2}	a_{i3}	b_i	Варіант	i	a_{i1}	a_{i2}	a_{i3}	b_i
1	1	3,90	1,25	-0,98	4,905	4	1	2,78	0,38	-0,43	3,261
	2	0,74	3,45	-0,84	6,031		2	-0,78	3,14	-0,81	3,295
	3	-0,65	1,18	2,38	10,134		3	-0,45	-0,86	2,48	6,072
2	1	2,68	-0,68	0,48	3,868	5	1	3,96	-0,78	-0,35	2,525
	2	-0,73	2,92	-0,39	4,329		2	1,18	3,78	-0,87	7,301
	3	-0,58	-1,12	3,12	7,532		3	-0,96	-1,02	3,68	9,190
3	1	2,50	-0,91	-0,32	0,287	6	1	3,48	1,12	-0,94	4,158
	2	-0,91	3,64	-0,48	5,418		2	1,08	3,67	-0,87	6,908
	3	0,48	-0,98	2,14	5,908		3	-1,21	-1,43	4,14	9,507

Варіант	i	a_{i1}	a_{i2}	a_{i3}	b_i	Варіант	i	a_{i1}	a_{i2}	a_{i3}	b_i
7	1	2,75	1,12	-0,6	3,066	19	1	2,70	1,15	0,48	7,806
	2	1,06	2,98	-0,86	5,328		2	0,86	2,60	0,32	8,382
	3	-1,18	-1,36	3,02	5,790		3	1,05	0,74	2,10	9,861
8	1	3,45	0,78	-0,97	3,229	20	1	2,80	1,02	0,32	7,112
	2	0,78	2,63	-0,89	4,026		2	0,96	2,40	0,46	8,480
	3	-0,97	-0,89	2,41	5,030		3	0,76	0,98	2,02	9,804
9	1	3,21	0,81	-0,93	3,102	21	1	2,90	1,08	0,43	7,738
	2	0,81	2,49	-0,94	3,571		2	0,82	2,50	0,64	9,114
	3	-0,93	-0,94	2,53	5,391		3	0,38	0,96	1,80	8,558
10	1	3,67	0,68	-1,21	2,467	22	1	3,10	1,20	0,62	8,894
	2	0,68	2,71	-0,96	3,825		2	1,12	2,60	0,85	10,416
	3	-1,21	-0,96	2,69	5,513		3	0,82	1,20	2,54	12,074
11	1	3,78	0,67	-0,83	3,928	23	1	3,75	1,20	1,07	11,355
	2	0,67	2,76	-0,69	4,871		2	0,89	3,50	1,52	13,245
	3	-0,83	-0,69	2,39	5,616		3	0,79	1,71	3,20	14,376
12	1	4,05	-0,93	-0,41	2,096	24	1	4,20	1,50	0,92	12,210
	2	-0,93	3,76	0,25	8,221		2	1,32	4,50	1,20	15,030
	3	-0,41	0,25	3,2	11,201		3	0,98	1,45	3,50	15,015
13	1	3,74	1,12	-1,03	4,207	25	1	3,50	1,20	0,96	10,650
	2	1,12	2,43	-1,07	3,412		2	2,10	4,30	1,02	15,240
	3	-1,03	-1,07	2,7	5,547		3	0,87	1,70	3,20	14,538
14	1	3,91	0,88	-1,13	3,543	26	1	2,70	0,81	0,56	7,431
	2	0,88	2,77	-0,98	4,173		2	0,91	2,92	0,98	10,437
	3	-1,13	-0,98	2,41	4,599		3	0,64	0,93	1,96	8,793
15	1	3,80	1,10	0,98	10,716	27	1	3,02	1,20	0,47	8,460
	2	0,75	2,96	0,92	11,023		2	1,12	2,70	0,90	10,050
	3	0,60	1,20	3,20	13,900		3	0,29	0,99	1,88	8,432
16	1	2,40	1,10	0,60	7,680	28	1	4,50	1,20	0,64	11,190
	2	0,98	2,60	1,20	11,354		2	1,40	3,70	0,92	12,630
	3	0,56	1,10	2,70	12,008		3	0,56	1,23	2,00	9,423
17	1	2,50	1,05	0,75	8,170	29	1	2,50	0,94	0,36	6,804
	2	0,95	2,60	0,85	10,195		2	0,87	2,30	0,76	8,415
	3	0,68	1,05	2,15	10,284		3	0,26	0,97	2,15	8,877
18	1	2,60	1,10	0,70	8,260	30	1	3,75	1,05	0,63	9,720
	2	0,92	2,70	0,65	9,756		2	0,28	1,47	0,35	4,557
	3	0,48	0,88	1,98	9,072		3	1,12	0,75	3,15	12,705

2. Порівняти між собою розв'язки системи, знайдені згаданими методами. Визначити кількість ітерацій, потрібних для досягнення заданої точності розв'язку в методах простої ітерації і Зейделя.

Розділ 4. РОЗВ'ЯЗУВАННЯ ЗАДАЧ ЛІНІЙНОГО ПРОГРАМУВАННЯ

§ 4.1. ПРИКЛАДИ ЗАДАЧ ЛІНІЙНОГО ПРОГРАМУВАННЯ

Розглянемо деякі приклади задач лінійного програмування.

1. *Задача про використання сировини.* Для виготовлення двох видів продукції P_1 і P_2 використовують три види сировини S_1 , S_2 і S_3 . Запаси сировини першого, другого і третього видів відповідно дорівнюють 30, 24 та 20 одиниць. На виготовлення одиниці продукції P_1 потрібно п'ять одиниць сировини першого виду, дві одиниці сировини другого виду і чотири одиниці сировини третього виду. На виготовлення одиниці продукції P_2 треба витратити по шість одиниць сировини S_1 та S_2 і дві одиниці сировини S_3 . Прибуток від реалізації одиниці продукції P_1 становить 40 у.г.о.*, а від реалізації одиниці продукції P_2 — 30 у.г.о. Скласти такий план випуску продукції, щоб від її реалізації одержати максимальний прибуток.

Умову задачі запишемо у вигляді табл. 4.1.

Таблиця 4.1

Сировина	Запас сировини	Кількість одиниць сировини, потрібних для виготовлення одиниці продукції	
		P_1	P_2
S_1	30	5	6
S_2	24	2	6
S_3	20	4	2
Прибуток від реалізації одиниці продукції, у.г.о.		40	30

Побудуємо математичну модель цієї задачі.

Позначимо через x_1 і x_2 кількість одиниць продукції P_1 і P_2 відповідно, яку треба виготовити. Тоді на виготовлення продукції P_1 і P_2 буде витрачено $5x_1 + 6x_2$ одиниць сировини S_1 , $2x_1 + 6x_2$ одиниць сировини S_2 і $4x_1 + 2x_2$ одиниць сировини S_3 . Кількість сировини, витраченої на виготовлення продукції, не може перевищувати запасу сировини, тобто

* У.г.о. — умовна грошова одиниця.

$$\begin{cases} 5x_1 + 6x_2 \leq 30, \\ 2x_1 + 6x_2 \leq 24, \\ 4x_1 + 2x_2 \leq 20. \end{cases}$$

Якщо продукція P_1 не виробляється, то $x_1 = 0$, інакше $x_1 \geq 0$. Це саме стосується й продукції P_2 . Отже, змінні x_1 і x_2 повинні бути невід'ємними, тобто $x_1 \geq 0$, $x_2 \geq 0$.

Таким чином, треба знайти такі значення змінних x_1 і x_2 , за яких задовольняється система нерівностей

$$\begin{cases} 5x_1 + 6x_2 \leq 30, \\ 2x_1 + 6x_2 \leq 24, \\ 4x_1 + 2x_2 \leq 20; \\ x_1 \geq 0, x_2 \geq 0, \end{cases}$$

і лінійна функція $F = 40x_1 + 30x_2$ набуває найбільшого значення.

Сформулюємо задачу використання сировини в загальному вигляді.

Для виробництва n видів продукції P_j ($j = 1, 2, \dots, n$) використовують m видів сировини S_i ($i = 1, 2, \dots, m$). Запас сировини, кількість одиниць сировини, потрібної для виробництва одиниці продукції, а також прибуток від реалізації одиниці продукції подано в табл. 4.2. Треба скласти такий план випуску продукції, щоб від її реалізації одержати максимальний прибуток.

Таблиця 4.2

Сировина	Запас сировини	Кількість одиниць i -ої сировини, потрібної для виробництва одиниці j -ої продукції			
		P_1	P_2	...	P_n
S_1	b_1	a_{11}	a_{12}	...	a_{1n}
S_2	b_2	a_{21}	a_{22}	...	a_{2n}
...
S_m	b_m	a_{m1}	a_{m2}	...	a_{mn}
Прибуток від реалізації одиниці продукції		c_1	c_2	...	c_n

Позначимо через x_j — кількість одиниць j -ої продукції, яку треба виготовити. Тоді математичну модель задачі можна сформулювати у вигляді: знайти максимальне значення лінійної функції

$$F = c_1x_1 + c_2x_2 + \dots + c_nx_n$$

за обмежень

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b_2, \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m, \end{cases}$$

$$x_j \geq 0 \quad (j = 1, 2, \dots, n).$$

Лінійна функція виражає кінцеву мету оптимального планування, у цьому разі — одержання найбільшого прибутку. Тому цю функцію називають *функцією мети*, або *цільовою функцією*. Змінні x_1, x_2, \dots, x_n входять до функції мети і до системи обмежень у першому степені, тому сформульована задача є типовою задачею лінійного програмування. Функцію мети в таких задачах називають ще *лінійною формою*.

Зауважимо, що до розглядуваного типу задач належать також задачі оптимального використання ресурсів і добрив.

2. *Задача про дієту*. Денна дієта повинна містити не менше восьми одиниць вітаміну А, дев'яти одиниць вітаміну В і семи одиниць вітаміну С. Для дієти використовують два різних продукти. В одиниці першого продукту міститься дві одиниці вітаміну А, одна одиниця вітаміну В і одна одиниця вітаміну С. В одиниці другого продукту міститься одна одиниця вітаміну А, три одиниці вітаміну В і одна одиниця вітаміну С. Скласти дієту так, щоб вона задовольняла мінімальну денну потребу у вітамінах при найменшій її вартості, якщо ціна одиниці першого продукту 0,2 у.г.о., а другого — 0,4 у.г.о.

Побудуємо математичну модель задачі.

Через x_1 і x_2 позначимо відповідно кількість одиниць першого і другого продуктів у дієті. Тоді денна дієта міститиме $2x_1 + x_2$ одиниць вітаміну А, $x_1 + 3x_2$ одиниць вітаміну В і $x_1 + x_2$ одиниць вітаміну С. Потреба у вітамінах буде задоволена, якщо їх кількість буде не меншою від передбачених, тобто

$$\begin{cases} 2x_1 + x_2 \geq 8, \\ x_1 + 3x_2 \geq 9, \\ x_1 + x_2 \geq 7. \end{cases}$$

Зазначимо, що $x_1 \geq 0$, $x_2 \geq 0$. Вартість дієти дорівнює $0,2x_1 + 0,4x_2$ (у.г.о.). Задача полягає в тому, щоб, задовольнивши мінімальну потребу у вітамінах, досягти найменшої вартості даної дієти.

Таким чином, серед множини розв'язків системи нерівностей

$$\begin{cases} 2x_1 + x_2 \geq 8, \\ x_1 + 3x_2 \geq 9, \quad x_1 \geq 0, \\ x_1 + x_2 \geq 7, \quad x_2 \geq 0 \end{cases}$$

треба знайти такий, при якому лінійна функція $F = 0,2x_1 + 0,4x_2$ набуває мінімального значення.

Задачу про дієту можна узагальнити, якщо передбачити в дієті m різних поживних речовин у кількості, не меншій b_i ($i = 1, 2, \dots, m$)

одиниць. Денна дієта складається з n різних продуктів у кількості d_j ($j = 1, 2, \dots, n$) одиниць. Нехай a_{ij} — кількість одиниць i -ї поживної речовини, що міститься в одиниці j -го продукту. Визначити, які продукти i в якій кількості треба включити в дієту, щоб вона задовольняла мінімальну денну потребу в кожній поживній речовині при найменшій вартості використаних продуктів, якщо ціна одиниці j -го продукту дорівнює c_j ($j = 1, 2, \dots, n$) у.г.о.

Для побудови математичної моделі задачі позначимо через x_j ($j = 1, 2, \dots, n$) кількість одиниць j -го продукту в денній дієті. Задача полягає в тому, щоб знайти мінімальне значення лінійної функції

$$F = \sum_{j=1}^n c_j x_j$$

при обмеженнях

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \geq b_1, \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \geq b_m, \end{cases} \quad 0 \leq x_j \leq d_j, \quad j = 1, 2, \dots, n.$$

До цього типу задач належать також задачі на складання денного раціону відгодівлі тварин, сумішей з різних речовин і деякі задачі планування виробництва.

3. Транспортна задача. Нехай з трьох пунктів відправлення P_1, P_2, P_3 треба перевезти однорідний вантаж обсягом 30 т до трьох пунктів призначення M_1, M_2, M_3 , в тому числі з пункту P_1 — 12 т, з пункту P_2 — 8 т, з пункту P_3 — 10 т. Вантаж повинен надійти за призначенням у пункт M_1 — 6 т, пункт M_2 — 9 т, пункт M_3 — 15 т. Вартість перевезення з пункту P_i ($i = 1, 2, 3$) в пункт призначення M_j ($j = 1, 2, 3$) дорівнює c_{ij} . Задача полягає в тому, щоб весь вантаж був вивезений з пунктів відправлення, повністю був задоволений попит пунктів призначення і загальна вартість перевезень була мінімальною.

Позначимо через x_{ij} кількість одиниць вантажу, запланованих для перевезення з пункту P_i в пункт призначення M_j . Тоді умову задачі можна записати у вигляді табл. 4.3. У правому верхньому куті клітки — перетину рядка P_i та стовпця M_j , записано значення вартостей c_{ij} .

Побудуємо математичну модель задачі.

Весь вантаж буде вивезено з пунктів відправлення, якщо

$$\begin{cases} x_{11} + x_{12} + x_{13} = 12, \\ x_{21} + x_{22} + x_{23} = 8, \\ x_{31} + x_{32} + x_{33} = 10. \end{cases} \quad (1')$$

Попит пунктів призначення буде задоволений повністю, якщо

$$\begin{cases} x_{11} + x_{21} + x_{31} = 6, \\ x_{12} + x_{22} + x_{32} = 9, \\ x_{13} + x_{23} + x_{33} = 15. \end{cases} \quad (2)$$

Таблиця 4

Пункт відправлення	Пункт призначення			Запас
	M_1	M_2	M_3	
P_1	1 x_{11}	3 x_{12}	4 x_{13}	12
P_2	2 x_{21}	5 x_{22}	3 x_{23}	8
P_3	6 x_{31}	7 x_{32}	4 x_{33}	10
Попит	6	9	15	30

Вартість перевезення вантажу дорівнює

$$F = x_{11} + 3x_{12} + 4x_{13} + 2x_{21} + 5x_{22} + 3x_{23} + 6x_{31} + 7x_{32} + 4x_{33}. \quad (3')$$

Отже, серед невід'ємних розв'язків систем (1') і (2') треба знайти такий, за якого лінійна функція (3') набуває найменшого значення.

У загальному випадку транспортну задачу сформулюємо в такий спосіб.

Нехай однорідний вантаж, що знаходиться у m постачальників A_i у кількості a_i ($i = 1, \dots, m$) одиниць відповідно, треба перевезти n споживачам у кількості b_j ($j = 1, \dots, n$) одиниць. Вартість перевезення одиниці вантажу від i -го постачальника до j -го споживача становить c_{ij} одиниць. Спланувати перевезення вантажу так, щоб весь вантаж був вивезений від постачальників, повністю був задоволений попит споживачів і вартість перевезень була б мінімальною.

Позначимо через x_{ij} кількість одиниць вантажу, запланованого для перевезення від i -го постачальника до j -го споживача. Тоді умову задачі зручно подати у вигляді табл. 4.4, яку називають *матрицею планування*.

Математична модель транспортної задачі має такий вигляд. Знайти найменше значення лінійної функції

$$F = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}$$

за обмежень

$$\begin{cases} \sum_{j=1}^n x_{ij} = a_i, & i = 1, 2, \dots, m; \\ \sum_{i=1}^m x_{ij} = b_j, & j = 1, 2, \dots, n; \\ x_{ij} \geq 0, & i = 1, 2, \dots, m; j = 1, 2, \dots, n. \end{cases}$$

У розглядуваній моделі передбачається, що сумарні запаси вантажу дорівнюють сумарному попиту, тобто

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j.$$

Така модель називається *закритою*.

Таблиця 4.4

Поста- тальник	Споживач				Запас
	b_1	b_2	...	b_n	
A_1	c_{11} x_{11}	c_{12} x_{12}	...	c_{1n} x_{1n}	a_1
A_2	c_{21} x_{21}	c_{22} x_{22}	...	c_{2n} x_{2n}	a_2
...
A_m	c_{m1} x_{m1}	c_{m2} x_{m2}	...	c_{mn} x_{mn}	a_m
Попит	b_1	b_2	...	b_n	$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$

§ 4.2. ЗАГАЛЬНА ЗАДАЧА ЛІНІЙНОГО ПРОГРАМУВАННЯ

Серед розглянутих у § 4.1 задач лінійного програмування різних сюжетів неважко помітити дещо спільне. З математичного погляду в кожній задачі треба знайти такі невід'ємні значення змінних, які задовольняють певну систему лінійних рівнянь або нерівностей і надають функції мети найменшого або найбільшого значення. Математично загальну задачу лінійного програмування формують так.

Нехай задано систему лінійних рівнянь

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases} \quad (4.1)$$

де

$$x_j \geq 0, \quad (j = 1, 2, \dots, n), \quad (4.2)$$

і лінійну функцію

$$F = c_1x_1 + c_2x_2 + \dots + c_nx_n, \quad (4.3)$$

де a_{ij}, b_i, c_j — задані сталі величини. Вважатимемо, що в системі рівнянь (4.1) усі b_i ($i = 1, 2, \dots, m$) невід'ємні.

Загальна задача лінійного програмування полягає в тому, щоб знайти такі невід'ємні значення x_1, x_2, \dots, x_n , які задовольняють систему рівнянь (4.1) і надають лінійній функції (4.3) мінімальне значення. Загальну задачу лінійного програмування записують за допомогою знаків підсумовування або у векторній чи матричній формах.

Запис за допомогою знаків підсумовування. Знайти мінімальне значення лінійної функції $F = \sum_{j=1}^n c_j x_j$ за обмежень

$$\sum_{j=1}^n a_{ij} x_j = b_i, \quad i = 1, 2, \dots, m; \quad x_j \geq 0, \quad j = 1, \dots, n.$$

Векторна форма запису. Знайти найменше значення лінійної функції $F = CX$ за обмежень

$$A_1x_1 + A_2x_2 + \dots + A_nx_n = A_0, \quad X \geq 0, \quad (4.4)$$

де $C = (c_1, c_2, \dots, c_n)$, $X = (x_1, x_2, \dots, x_n)$ — вектори-рядки, CX — скалярний добуток векторів C і X ; вектори-стовпці A_j мають вигляд

$$A_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad A_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \dots, \quad A_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}, \quad A_0 = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

і складаються відповідно з коефіцієнтів при змінних і вільних членів системи (4.1).

Матрична форма запису. Мінімізувати лінійну функцію $F = CX$ за обмежень $AX = A_0, X \geq 0$, де $C = (c_1, c_2, \dots, c_n)$ — вектор коефіцієнтів лінійної функції, а

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, A_0 = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

— відповідно вектор-стовпець змінних, вектор-стовпець вільних членів і матриця коефіцієнтів при змінних системи (4.1). Матрицю A називають *матрицею умови задачі*. Її елементи — це техніко-економічні або технологічні показники. Вектор A_0 називають *вектором обмежень*. Його компоненти — невід'ємні числа. Задачу лінійного програмування, обмеження якої задано у вигляді системи (4.1), (4.2), називають *канонічною*.

У переважній більшості задач обмеження задають не системою рівнянь, а системою нерівностей

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b_2, \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m, \end{cases} \quad (4.5)$$

або

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \geq b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \geq b_2, \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \geq b_m. \end{cases} \quad (4.6)$$

Система обмежень може бути й мішаною: частина обмежень задана у вигляді рівнянь, частина у вигляді нерівностей (4.5), а частина — у вигляді нерівностей (4.6). Але будь-яку систему обмежень можна звести до системи рівнянь виду (4.1). Для цього досить до лівої частини кожної нерівності виду (4.5) додати (або відняти, якщо нерівність виду (4.6)) таку невід'ємну змінну, щоб відповідна нерівність перетворилася в рівняння.

Якщо в нерівності

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1$$

до лівої частини додати таку змінну $x_{n+1} \geq 0$, що

$$x_{n+1} = b_1 - (a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n),$$

то нерівність перетвориться в рівняння

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + x_{n+1} = b_1.$$

У лінійну функцію кожна додаткова змінна входить з нульовим коефіцієнтом, тобто

$$F = c_1x_1 + c_2x_2 + \dots + c_nx_n + 0 \cdot x_{n+1} + 0 \cdot x_{n+r},$$

де r — кількість введених у систему обмежень додаткових змінних.

П р и к л а д 1. Звести до канонічної задачі: знайти мінімум лінійної функції $F = 2x_1 + 3x_2 + x_3$ за обмежень

$$\begin{cases} 3x_1 - 2x_3 \leq 4, \\ 4x_1 + 2x_2 + x_3 = 5, \\ x_1 + x_2 - 3x_3 \geq 2 \quad x_j \geq 0 \quad (j = 1, 2, 3). \end{cases}$$

Р о з в ' я з а н н я. Задачу легко звести до канонічної, ввівши в першу нерівність системи змінну $x_4 \geq 0$, а в третю — $x_5 \geq 0$. Тоді система обмежень матиме вигляд

$$\begin{cases} 3x_1 - 2x_3 + x_4 = 4, \\ 4x_1 + 2x_2 + x_3 = 5, \\ x_1 + x_2 - 3x_3 - x_5 = 2, \\ x_j \geq 0 \quad (j = 1, 2, \dots, 5). \end{cases}$$

Лінійна функція при цьому не зміниться.

Якщо в системі обмежень (4.1), (4.2) не всі змінні підпорядковані умові невід'ємності, то така система не буде канонічною. Її легко звести до канонічної, якщо кожному змінну, не підпорядковану умові невід'ємності, замінити різницею двох невід'ємних змінних.

П р и к л а д 2. Звести до канонічної задачі: знайти мінімум лінійної функції $F = 3x_1 - 2x_2 + x_3 + 4x_4$ за обмежень

$$\begin{cases} x_1 + 3x_2 - x_3 + 2x_4 = 2, \\ 4x_1 + x_2 + 5x_3 + x_4 = 4, \\ x_1 + x_2 - 3x_3 - x_4 = 3, \\ x_1 \geq 0, \quad x_3 \geq 0. \end{cases}$$

Р о з в ' я з а н н я. У цій системі обмежень змінні x_2 і x_4 не підпорядковані умові невід'ємності. Тому замінимо ці змінні різницями $x_2 = x_2' - x_2''$, $x_4 = x_4' - x_4''$, де $x_2' \geq 0$, $x_2'' \geq 0$, $x_4' \geq 0$, $x_4'' \geq 0$. Підставивши значення x_2 і x_4 у лінійну функцію і систему обмежень, дістанемо канонічну задачу

$$F = 3x_1 - 2x_2' + 2x_2'' + x_3 + 4x_4' - 4x_4'';$$

$$\begin{cases} x_1 + 3x_2' - 3x_2'' - x_3 + 2x_4' - 2x_4'' = 2, \\ 4x_1 + x_2' - x_2'' + 5x_3 + x_4' - x_4'' = 4, \\ x_1 + x_2' - x_2'' - 3x_3 - x_4' + x_4'' = 3, \\ x_1 \geq 0, \quad x_2' \geq 0, \quad x_2'' \geq 0, \quad x_3 \geq 0, \quad x_4' \geq 0, \quad x_4'' \geq 0. \end{cases}$$

Часто в задачах лінійного програмування треба знайти не мінімум, а максимум лінійної функції. Оскільки $\max F = -\min(-F)$, то там, де треба знайти максимальне значення лінійної функції, досить змінити знак лінійної функції на протилежний і знайти її мінімум. Взявши

знайдений мінімум з протилежним знаком, дістанемо максимальне значення заданої лінійної функції.

Отже, в якому б вигляді не було поставлено задачу лінійного програмування, її завжди можна звести до канонічної.

Система (4.1) містить m рівнянь з n змінними. Тому вона може бути сумісною або несумісною, а рівняння системи — як лінійно залежними, так і лінійно незалежними. Якщо система (4.1) несумісна, то задача лінійного програмування не має розв'язку. Несумісність системи обмежень означає, що виробниче завдання не збалансоване, тому не може бути виконане. Отже, спеціалісти повинні його відповідно скоригувати.

Нехай система (4.1) сумісна і всі m її рівнянь лінійно незалежні (у протилежному разі з системи можна виключити частину рівнянь так, щоб решта рівнянь системи стали лінійно незалежними). Це означає, що серед n векторів A_1, A_2, \dots, A_n системи обмежень лише m лінійно незалежні, які утворюють базис m -вимірного простору. З курсу алгебри відомо, що сумісна система m лінійних рівнянь з n змінними ($m < n$) має безліч розв'язків, а кількості її базисних розв'язків не перевищує C_n^m . Базисний розв'язок — це розв'язок, що містить m базисних (основних) і $n-m$ вільних (неосновних) змінних. Вільні змінні в базисному розв'язку дорівнюють нулю. Якщо хоча б одна базисна змінна набуває нульового значення, то базисний розв'язок вироджений. Сформулюємо деякі основні означення.

Означення 1. Планом або допустимим розв'язком задачі лінійного програмування називають вектор $X = (x_1, x_2, \dots, x_n)$, який задовольняє умови (4.1), (4.2).

Означення 2. План $X = (x_1, x_2, \dots, x_n)$ називають опорним, якщо в рівності (4.4) вектори A_i ($i = 1, 2, \dots, m$) при додатних коефіцієнтах x_i лінійно незалежні.

Опорний план — це допустимий базисний розв'язок системи обмежень загальної задачі лінійного програмування. З означення опорного плану випливає, що він не може містити більш як m додатних компонент.

Означення 3. Опорний план називають невиродженим, якщо він містить рівно m додатних координат, у протилежному разі план називають виродженим.

Це означає, що невиродженим опорним планом буде допустимий базисний розв'язок системи (4.1) з m додатними базисними змінними. Якщо хоч одна базисна змінна допустимого базисного розв'язку системи (4.1) набуває нульового значення, то опорний план буде виродженим. Наприклад, розглянемо систему обмежень

$$\begin{cases} 2x_1 + x_2 + 3x_3 = 2, \\ x_1 - 2x_3 + x_4 = 1, \\ x_j \geq 0 \quad (j=1,2,3,4), \end{cases}$$

яку запишемо у вигляді

$$x_1A_1 + x_2A_2 + x_3A_3 + x_4A_4 = A_0.$$

В ній

$$A_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad A_0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Легко впевнитись, що вектори $X_1 = (0; 2; 0; 1)$, $X_2 = (1; 0; 0; 0)$, $X_3 = (0,5; 1; 0; 0,5)$ є допустимими розв'язками заданої системи. Допустимому розв'язку X_1 відповідають вектори A_2 і A_4 . Вони — лінійно незалежні і стоять у системі обмежень при додатних коефіцієнтах $x_2 = 2$, $x_4 = 1$. Отже, вектор X_1 — опорний план. Розв'язку X_2 відповідає вектор A_1 , який стоїть у системі обмежень при єдиній додатній координаті $x_1 = 1$. Вектор A_1 утворює лінійно незалежну систему, тому X_2 також опорний план. План X_3 не є опорним, бо має три додатні координати, а будь-які три двовимірні вектори лінійно залежні. План X_1 — невироджений, а план X_2 — вироджений.

Означення 4. Оптимальним планом або оптимальним розв'язком задачі лінійного програмування називають такий план, при якому лінійна функція набуває найменшого (або найбільшого) значення.

§ 4.3. ГЕОМЕТРИЧНИЙ ЗМІСТ ЗАДАЧ ЛІНІЙНОГО ПРОГРАМУВАННЯ

Розглянемо таку задачу лінійного програмування. Знайти мінімум лінійної функції (4.3) за обмежень (4.2), (4.5), заданих у вигляді нерівностей.

Покладемо в (4.2), (4.3) і (4.5) $n = 2$. Кожна з нерівностей системи (4.5) геометрично визначає півплощину, обмежену прямою $a_{i1}x_1 + a_{i2}x_2 = b_i$ ($i = 1, 2, \dots, m$). Умови невід'ємності $x_1 \geq 0$, $x_2 \geq 0$ визначають півплощини відповідно з граничними прямими $x_1 = 0$, $x_2 = 0$. Якщо система обмежень сумісна, то перетин півплощин, як опуклих множин, утворить опуклу множину точок, координати якої є розв'язком системи (4.5) (рис. 4.1). Сукупність цих точок є множиною планів або допустимих розв'язків задачі лінійного програмування, його називають *многокутником розв'язків*. Він може бути точкою, відрізком, променем, опуклим многокутником або необмеженою многокутною областю. Лінійну функцію на площині можна інтерпретувати як сім'ю деяких паралельних прямих, кожна з яких визначається значенням параметра F . Кожна така окрема пряма є прямою рівня лінійної функції. Нас цікавлять лише значення функції мети в точках, що є планами задачі, тобто ті прямі рівня цільової функції, що мають спільні точки з многокутником розв'язків.

Задача лінійного програмування на площині полягає в тому, щоб

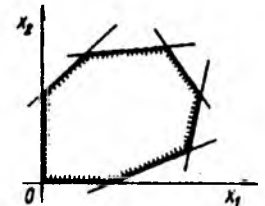


Рис. 4.1

знайти таку точку многокутника розв'язків, в якій лінійна функція набуває мінімального значення. Це означає, що серед прямих рівня треба вибрати ту, якій відповідає оптимальне значення F лінійної функції. Розглянемо геометричний зміст значення F . Якщо $F = 0$, пряма $c_1x_1 + c_2x_2 = 0$ проходить через початок координат. Саме ж значення F прямої $c_1x_1 + c_2x_2 = F$ є відстанню (відхиленням) даної прямої від початку координат, помноженою на довжину вектора C . Це відхилення вважається додатним для додатних значень F і від'ємним для від'ємних F .

Отже, серед прямих сім'ї F , що мають спільні точки з многокутником розв'язків, треба знайти найменш (або найбільш) віддалену від початку координат пряму, причому слід знайти не тільки абсолютну величину відстані, а й її знак. Ця шукана пряма повинна дотикатися до многокутника розв'язків у його граничних точках — вершині або стороні.

Якщо в системі обмежень (4.5) $n = 3$, то кожна нерівність геометрично означає півпростір тривимірного простору з граничною площиною $a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 = b_i$, ($i = 1, 2, \dots, m$), а умови $x_j \geq 0$ ($j = 1, 2, 3$) — півпростори, обмежені площинами $x_j = 0$ ($j = 1, 2, 3$). Якщо система обмежень сумісна, то перетин півпросторів утворює в тривимірному просторі *многогранник розв'язків*. Він може бути точкою, відрізком, променем, многокутником, многогранником або необмеженою многогранною областю. Лінійну функцію можна інтерпретувати як сім'ю паралельних площин.

В n -вимірному ($n > 3$) просторі кожна нерівність системи (4.5) визначає півпростір n -вимірного простору, обмеженого гіперплощинами $a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i$ ($i = 1, 2, \dots, m$), а умови невід'ємності — півпростори, обмежені гіперплощинами $x_j = 0$ ($j = 1, 2, \dots, n$). Якщо система обмежень сумісна, то перетин півпросторів утворює частину n -вимірного простору, що називається многогранником розв'язків, оскільки координати його точок задовольняють систему (4.5). Лінійна функція в n -вимірному просторі є сім'єю деяких паралельних гіперплощин, положення кожної з яких визначається значенням параметра F . Кожну таку гіперплощину можна інтерпретувати як гіперплощину рівня цільової функції F . Значення F гіперплощини

$$c_1x_1 + c_2x_2 + \dots + c_nx_n = F$$

можна глумачити, як відстань (відхилення) даної площини від початку координат, помножену на довжину вектора C . Відхилення є додатним для додатних F і від'ємним для від'ємних F .

Таким чином, в n -вимірному просторі геометричний зміст задачі лінійного програмування полягає в тому, щоб серед гіперплощин сім'ї F , які мають спільні точки з многогранником розв'язків, знайти точку найменш (або найбільш) віддалену від початку координат. При цьому слід враховувати й знак відстані, а не лише її абсолютну величину. Шукані гіперплощини повинні дотикатися до границі многогранника розв'язків — точки, ребра, грані.

Геометрична інтерпретація задачі лінійного програмування дає графічний метод її розв'язання, який застосовують в основному для розв'язування задач двовимірного і найпростіших задач тривимірного простору. Розглянемо деякі приклади.

П р и к л а д 3. Задача про використання сировини (§ 4.1). Знайти найбільше значення функції $F = 40x_1 + 30x_2$ за обмежень

$$\begin{cases} 5x_1 + 6x_2 \leq 30, \\ 2x_1 + 6x_2 \leq 24, \\ 4x_1 + 2x_2 \leq 20, \\ x_1 \geq 0, \quad x_2 \geq 0. \end{cases}$$

Р о з в ' я з а н н я. На рис. 4.2 зображено область допустимих розв'язків даної системи нерівностей. Це опуклий п'ятикутник $OABCD$, сторонами якого є відрізки прямих

$$\begin{aligned} 5x_1 + 6x_2 &= 30, \quad (l_1) \\ 2x_1 + 6x_2 &= 24, \quad (l_2) \\ 4x_1 + 2x_2 &= 20, \quad (l_3) \\ x_1 &= 0, \quad (Ox_2) \\ x_2 &= 0, \quad (Ox_1). \end{aligned}$$

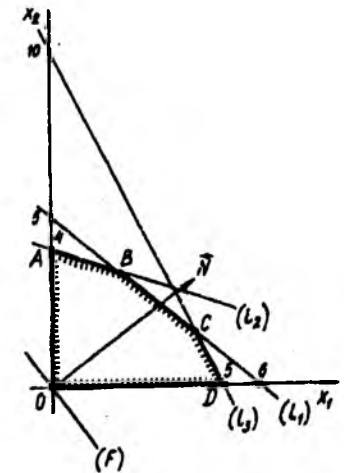


Рис. 4.2.

Треба знайти таку точку побудованого п'ятикутника, в якій лінійна функція набуває найбільшого значення. При фіксованих значеннях F лінійна функція на площині зображується прямою лінією $40x_1 + 30x_2 = F$. Змінюючи величину константи, дістанемо сім'ю паралельних ліній. Побудуємо, наприклад, лінію $40x_1 + 30x_2 = 0$. Якщо побудовану пряму рухати праворуч, то значення F зростатиме. Напрямок руху прямої $40x_1 + 30x_2 = \text{const}$ показує вектор $\vec{N} = (40; 30) = 10(4; 3)$, координатами якого є коефіцієнти при змінних лінійної функції, або пропорційні їм числа. При переміщенні побудованої прямої в напрямі вектора \vec{N} значення функції F зростає, а при переміщенні прямої в протилежному напрямі — спадає. З рис. 4.2 видно, що коли пряма $40x_1 + 30x_2 = 0$ рухається в напрямі вектора \vec{N} , лінійна функція F на многокутнику розв'язків набуває найбільшого значення в точці C . Щоб визначити її координати, розв'яжемо систему рівнянь

$$\begin{cases} 5x_1 + 6x_2 = 30, \\ 4x_1 + 2x_2 = 20. \end{cases}$$

Звідси маємо $x_1 = \frac{30}{7}$, $x_2 = \frac{10}{7}$. Підставивши ці значення в лінійну функцію, дістанемо $F_{\max} = 1500/7 \approx 214,28$.

Таким чином, щоб одержати максимальний прибуток у розмірі 214,28 у.г.о., слід запланувати 4,29 одиниці продукції першого виду і 1,43 одиниці продукції другого виду.

Дамо геометричну інтерпретацію опорного плану. Для цього задану систему обмежень за допомогою введення додаткових змінних $x_3 \geq 0$, $x_4 \geq 0$, $x_5 \geq 0$, зведемо до системи рівнянь

$$\begin{cases} 5x_1 + 6x_2 + x_3 = 30, \\ 2x_1 + 6x_2 + x_4 = 24, \\ 4x_1 + 2x_2 + x_5 = 20. \end{cases} \quad (4.7)$$

Система трьох рівнянь з п'ятьма змінними може мати не більше десяти ($C_3^5 = 10$) базисних розв'язків. Знайдемо деякі з них. Візьмемо змінні x_3, x_4, x_5 за базисні. Це можна зробити, бо вектори

$$A_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, A_4 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, A_5 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

лінійно незалежні. Нехай вільні змінні x_1 і x_2 дорівнюють нулю. Тоді з системи (4.7) дістанемо базисний розв'язок $X_1 = (0; 0; 30; 24; 20)$. Перші дві його компоненти є координатами точки площини. Знайдений розв'язок X_1 є опорним планом і відповідає вершині O п'ятикутника $OABCD$ допустимого розв'язку задачі лінійного програмування.

Тепер за базисні візьмемо змінні x_2, x_3, x_5 , оскільки вектори

$$A_2 = \begin{pmatrix} 6 \\ 6 \\ 2 \end{pmatrix}, A_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, A_5 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

лінійно незалежні. Тоді вільними будуть змінні x_1 і x_4 . Поклавши $x_1 = x_4 = 0$, з системи (4.7) знаходимо $x_2 = 4$, $x_3 = 6$, $x_5 = 12$. Отже, базисний розв'язок $X_2 = (0; 4; 6; 0; 12)$ є опорним планом задачі лінійного програмування і відповідає точці $A(0; 4)$ многокутника розв'язків.

Якщо за базисні змінні взяти x_1, x_3, x_4 , то поклавши $x_2 = x_5 = 0$, дістанемо базисний розв'язок $X_3 = (5; 0; 5; 14; 0)$. Цей розв'язок є опорним планом і відповідає вершині D многокутника розв'язків. Аналогічно можна показати, що базисні розв'язки $X_4 = (2; +\frac{10}{3}; 0; 0; \frac{16}{3})$

та $X_5 = (\frac{30}{7}; \frac{10}{7}; 0; \frac{48}{7}; 0)$ є опорними планами, які відповідають вершинам B і C многокутника розв'язків. Інші базисні розв'язки системи (4.7) не є допустимими.

Інші базисні розв'язки системи (4.7) не є допустимими.

Отже, кожному опорному плану відповідає вершина многокутника розв'язків задачі лінійного програмування, і навпаки, кожна вершина многокутника розв'язків відповідає опорному плану задачі лінійного програмування.

П р и к л а д 4. Знайти максимум лінійної функції $F = x_1 + 4x_2$ за обмежень

$$\begin{cases} -2x_1 + x_2 \leq 4, \\ x_1 + x_2 \leq 10, \quad x_1 \geq 0; \\ x_1 - 3x_2 \leq 6, \quad 0 \leq x_2 \leq 8. \end{cases} \quad (4.8)$$

Р о з в ' я з а н н я. Множину допустимих розв'язків даної задачі зображено на рис. 4.3. Рухаючи пряму $x_1 + 4x_2 = 0$ паралельно їй самій у напрямі вектора $\vec{N} = (1; 4)$, бачимо, що лінійна функція F набуває найбільшого значення в точці A . Її координати є розв'язком системи

$$\begin{cases} x_2 = 8, \\ -2x_1 + x_2 = 4, \\ x_1 + x_2 = 10, \end{cases}$$

тобто $x_1 = 2$, $x_2 = 8$. Значення лінійної функції $F_{\max} = 2 + 4 \cdot 8 = 34$.

З іншого боку, вершина A многокутника розв'язків є опорним планом даної задачі лінійного програмування. Знайдемо його. Для цього систему обмежень зведемо до канонічного вигляду

$$\begin{cases} x_2 + x_3 = 8, \\ -2x_1 + x_2 + x_4 = 4, \\ x_1 + x_2 + x_5 = 10, \\ x_1 - 3x_2 + x_6 = 6, \\ x_1 + 2x_2 - x_7 = 3, \\ x_j \geq 0 \quad (j = 1, 2, \dots, 7). \end{cases}$$

Вектори

$$A_1 = \begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \\ 1 \end{pmatrix}, A_2 = \begin{pmatrix} 1 \\ 1 \\ -3 \\ 2 \end{pmatrix}, A_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, A_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, A_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}$$

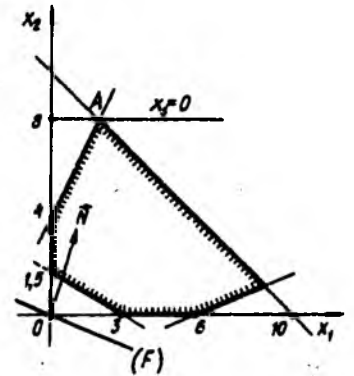


Рис. 4.3

лінійно незалежні, тобто утворюють базис 5-вимірного векторного простору, тому змінні x_1, x_2, x_3, x_6, x_7 можна взяти за базисні, а x_4 і x_5 за вільні змінні. Поклавши в останній системі $x_4 = x_5 = 0$, знаходимо $x_1 = -2, x_2 = 8, x_3 = 0, x_6 = 28, x_7 = 15$. Отже, вектор $X = (2; 8; 0; 0; 0; 28; 15)$ є опорним планом і відповідає вершині A многокутника розв'язків системи обмежень (4.8). Знайдений опорний план має лише чотири додатні координати (а не п'ять), тому він вироджений. Геометрично це означає, що у вершині A перетинаються не дві, а три межові прямі системи обмежень (4.8), на яких змінні x_3, x_4, x_5 відповідно дорівнюють нулю (див. рис. 4.3).

П р и к л а д 5. Задача про дієту (§ 4.1). Знайти мінімум функції $F = 0,2x_1 + 0,4x_2$ за обмежень

$$\begin{cases} 2x_1 + x_2 \geq 8, \\ x_1 + 3x_2 \geq 9, \quad x_1 \geq 0, \\ x_1 + x_2 \geq 7, \quad x_2 \geq 0. \end{cases} \quad (4.9)$$

Р о з в ' я з а н н я. На рис. 4.4 зображено необмежену багатокутну область розв'язків заданої системи обмежень, пряму $F = 2x_1 + 4x_2 = 0$ і вектор $\vec{N} (1; 2)$. Якщо рухати побудовану пряму F у напрямі вектора \vec{N} , то вона вперше дотикнеться до многокутника розв'язків у вершині C . Якщо пряму F і далі рухати в напрямі вектора \vec{N} , то значення лінійної функції на многокутнику розв'язків необмежено зростатиме. Отже, в точці C лінійна функція набуває найменшого значення. Знайдемо координати точки C , розв'язавши систему

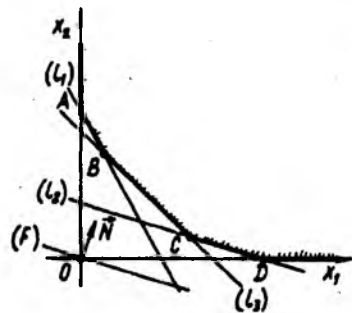


Рис. 4. 4.

Звідси $x_1 = 6, x_2 = 1$. Значення лінійної функції $F = 0,2 \cdot 6 + 0,4 \cdot 1 = 1,6$.

Таким чином, щоб забезпечити найменші витрати на денну дієту (1,6 у.г.о.), необхідно до неї включити шість одиниць маси першого продукту і одну одиницю маси другого продукту.

Зазначимо, що коли в цій задачі знаходити не мінімум, а максимум лінійної функції, то задача не матиме оптимального розв'язку, бо коли пряму F рухати в напрямі вектора \vec{N} , то значення лінійної функції необмежено зростає. А лінійна функція $F = x_1 - x_2$ на многокутнику розв'язків системи (4.9) необмежена як знизу, так і зверху, бо коли пряма $x_1 - x_2 = 0$ рухається в напрямі вектора $\vec{N} = (1; -1)$ або в протилежному напрямі, вона весь час перетинає многокутник розв'язків

і ніколи не залишає його. При цьому лінійна функція в напрямі вектора $\vec{N} = (1; -1)$ необмежено зростатиме, а в протилежному — необмежено спадатиме.

П р и к л а д 6. Знайти максимум лінійної функції $F = 3x_1 - x_2$ за обмежень

$$\begin{cases} x_1 + x_2 \leq 6, \\ 3x_1 - x_2 \leq 9, \\ -x_1 + 2x_2 \leq 8, \\ x_1 \geq 0, \quad x_2 \geq 0. \end{cases}$$

Р о з в ' я з а н н я. Область розв'язків системи обмежень є п'ятикутник $OABCD$ (рис. 4. 5).

Побудуємо пряму $3x_1 - x_2 = 0$ і вектор $\vec{N} = (3; -1)$. Рухаючи пряму F паралельно початковому її положенню в напрямі вектора \vec{N} , побачимо, що вона при виході з многокутника розв'язків матиме з ним не одну спільну точку, а безліч.

Пряма F збігається з прямою CD . Отже, в усіх точках відрізка CD лінійна функція F матиме одне і те саме значення, яке і буде її найбільшим значенням $F_{\max} = 9$. Таким чином, маємо не один, а безліч оптимальних розв'язків, що збігаються з точками відрізка CD і, зокрема, з двома вершинами C і D многокутника розв'язків.

П р и к л а д 7. Знайти мінімум лінійної функції $F = 3x_1 - x_2$ за обмежень

$$\begin{cases} x_1 + x_2 \leq 6, \\ 3x_1 - x_2 \geq 9, \\ -x_1 + 2x_2 \geq 8, \\ x_1 \geq 0, \quad x_2 \geq 0. \end{cases}$$

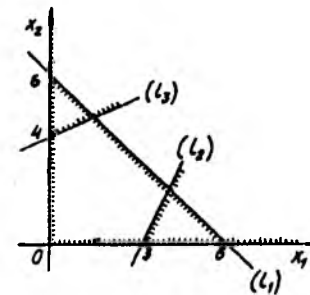


Рис. 4.6

Р о з в ' я з а н н я. Зображена на рис. 4. 6 область не містить жодної точки, яка задовольняла б усі нерівності системи обмежень. Це означає, що система обмежень несумісна і дана задача лінійного програмування оптимального розв'язку не має.

Геометрично найзручніше розв'язувати задачу лінійного програмування з обмеженнями у вигляді нерівностей з двома змінними. Але графічним методом можна розв'язувати й задачі, системи обмежень яких містять m лінійно незалежних рівнянь і n змінних, причому m і n повинні задо-

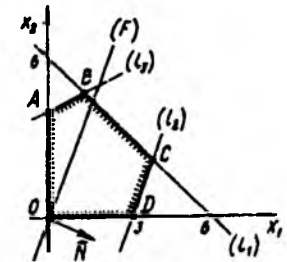


Рис. 4. 5

вільняти умову $n - m = 2$. У цьому разі за допомогою методу послідовних виключень Гаусса з системи виключають m змінних, які стануть базисними. Вільними залишаться дві змінні. Якщо від кожного рівняння перетвореної системи відняти відповідну невід'ємну змінну, то система рівнянь перетвориться в однорідну систему нерівностей відносно двох вільних змінних. За допомогою рівнянь перетвореної системи виражають також лінійну функцію через вільні змінні. Таким чином, перетворена задача містить лише дві змінні і її можна розв'язати графічним методом. Проілюструємо сказане на прикладі.

П р и к л а д 8. Знайти максимальне значення лінійної функції $F = -12x_1 - 9x_2 + 2x_3 + x_4 + 6x_5 + 2$ за обмежень

$$\begin{cases} -3x_1 - 7x_2 + x_3 + x_4 + 7x_5 = 13, \\ 2x_1 - 14x_2 + x_3 + 2x_4 + 13x_5 = 20, \\ 6x_1 - 23x_2 + x_3 + 3x_4 + 20x_5 = 19, \\ x_j \geq 0, \quad (j = 1, 2, \dots, 5). \end{cases}$$

Р о з в ' я з а н н я. Зведемо канонічну форму задачі лінійного програмування до задачі з однорідними обмеженнями у вигляді нерівностей. Для цього виділимо в системі обмежень базисні змінні, наприклад x_3, x_4, x_5 , методом повного виключення Гаусса. Одночасно перетворимо лінійну функцію, розглядаючи її як рівняння

$$12x_1 + 9x_2 - 2x_3 - x_4 - 6x_5 + F = 2,$$

розв'язане відносно базисної змінної F .

Виконуючи перетворення методом послідовного виключення, не слід вибирати за ведучий той рядок, в якому записано останнє рівняння. Тоді воно залишатиметься розв'язаним відносно базисної змінної F . Перетворення системи рівнянь подано в таблиці 4. 5. Після третьої ітерації систему рівнянь можна розв'язати відносно змінних x_3, x_4, x_5 :

$$\begin{cases} -7x_1 + 2x_2 + x_3 = 14, \\ 11x_1 + 5x_2 + x_4 = 55, \\ -x_1 - 2x_2 + x_5 = -8. \end{cases}$$

Змінні x_3, x_4, x_5 одночасно виключені з виразу лінійної функції: $F = -3x_1 - 6x_2 + 37$.

Враховавши, що $x_3 \geq 0, x_4 \geq 0, x_5 \geq 0$, можна кожне з цих рівнянь замінити відповідною нерівністю. Так, від першого рівняння $-7x_1 + 2x_2 + x_3 = 14$, віднявши невід'ємну змінну $x_3 \geq 0$, дістанемо нерівність $-7x_1 + 2x_2 \leq 14$.

Аналогічно перетворимо друге й третє рівняння останньої системи. В результаті прийдемо до задачі: знайти максимум лінійної функції $F = -3x_1 - 6x_2 + 37$ за обмежень

$$\begin{cases} -7x_1 + 2x_2 \leq 14, \\ 11x_1 + 5x_2 \leq 55, \\ -x_1 - 2x_2 \leq -8, \\ x_1 \geq 0, \quad x_2 \geq 0. \end{cases}$$

Таблиця 4.5

Ітерація	A_0	A_1	A_2	A_3	A_4	A_5	A_6
1	13	-3	-7	1	1	7	0
	20	2	-14	1	2	13	0
	19	6	-23	1	3	20	0
	2	12	9	-2	-1	-6	1
2	13	-3	-7	1	1	7	0
	7	5	-7	0	1	6	0
	6	9	-16	0	2	13	0
	28	6	-5	0	1	8	1
3	6	-8	0	1	0	1	0
	7	5	-7	0	1	6	0
	-8	-1	-2	0	0	1	0
	21	1	2	0	0	2	1
4	14	-7	2	1	0	0	0
	55	11	5	0	1	0	0
	-8	-1	-2	0	0	1	0
	37	3	6	0	0	0	1

На рис. 4.7 показано графічне розв'язання останньої задачі. Максимального значення лінійна функція F набуває в точці $A \left(\frac{70}{17}, \frac{33}{17} \right)$. На прямих AD і AB друга й третя нерівності системи обмежень перетворюються в рівності, тому на цих прямих x_4 і x_5 відповідно дорівнюють нулю. Отже, в точці A маємо $x_4 = 0, x_5 = 0$. З рівняння $-7x_1 + 2x_2 + x_3 = 14$ знаходимо $x_3 = 662/17$.

Таким чином, оптимальним розв'язком задачі буде вектор $X = \left(\frac{70}{17}, \frac{33}{17}, \frac{662}{17}, 0, 0 \right)$, а максимальне значення лінійної функції $F_{\max} = 13$.

Узагальнивши розглянуті приклади, доходимо висновку: загальна задача лінійного програмування може мати єдиний, безліч або зовсім

не мати оптимальних розв'язків. Якщо система обмежень несумісна, то задача лінійного програмування не має оптимального розв'язку; якщо система обмежень має єдиний допустимий розв'язок, то він і буде оптимальним. Якщо система обмежень має множину допустимих розв'язків і функція мети обмежена на ній, то задача лінійного програмування має оптимальний план (єдиний або безліч); якщо функція мети не обмежена знизу на множині планів, то загальна задача лінійного програмування не має оптимального розв'язку.

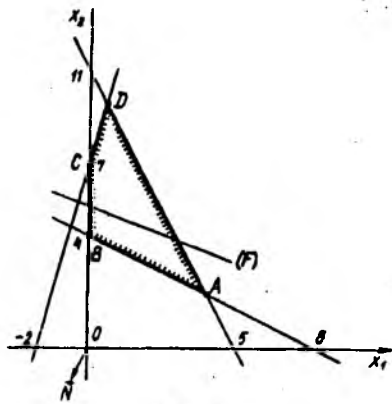


Рис. 4. 7.

Оптимальний план, якщо він існує, лежить не всередині многогранника розв'язків, а в якійсь з його границь — вершині, ребрі, грані, де не менш як $k = n - m$ компонент опорного плану дорівнюють нулю. Щоб знайти оптимальний розв'язок, треба переходити від одного опорного плану до іншого в напрямі зменшення або збільшення цільової функції залежно від того, що треба знайти — максимум чи мінімум лінійної функції.

Оптимальний план, якщо він існує, лежить не всередині многогранника розв'язків, а в якійсь з його границь — вершині, ребрі, грані, де не менш як $k = n - m$ компонент опорного плану дорівнюють нулю. Щоб знайти оптимальний розв'язок, треба переходити від одного опорного плану до іншого в напрямі зменшення або збільшення цільової функції залежно від того, що треба знайти — максимум чи мінімум лінійної функції.

§ 4.4. ЗАГАЛЬНІ ВІДОМОСТІ ПРО СИМПЛЕКС-МЕТОД

Нехай задача лінійного програмування має оптимальний розв'язок. З геометричного погляду це означає, що існує вершина многогранника розв'язків, в якій лінійна функція досягає свого оптимального значення. Кожній вершині многогранника розв'язків відповідає опорний план. А кожний опорний план визначається системою m лінійно незалежних векторів, що містяться серед n векторів A_1, A_2, \dots, A_n даної системи обмежень. Щоб знайти оптимальний план, досить розглянути лише опорні плани. Їх кількість не перевищує C_n^m . При великих значеннях m і n знайти серед них оптимальний простим перебором дуже важко. Тому необхідно мати такий аналітичний метод, який дає можливість цілеспрямовано розглядати опорні плани. Таким методом є *симплексний метод*. Виходячи з одного опорного плану, симплексний метод забезпечує побудову нового опорного плану, який надає лінійній функції значення меншого, ніж давав попередній опорний план. Цей процес продовжується доти, поки не буде знайдено оптимальний план. Оскільки кількість опорних планів задачі лінійного програмування обмежена, то обмежена й кількість ітерацій симплексного методу. Якщо задача лінійного програмування розв'язків не має, то симплекс-метод

встановлює цей факт у ході розв'язування задачі. Це означає, що під час обчислень можна встановити, чи є система обмежень сумісною і чи є лінійна функція обмеженою на множині планів задачі лінійного програмування.

Отже, симплексний метод дає спосіб обчислення опорного плану, перевіряє побудований опорний план на оптимальність і визначає спосіб побудови наступного опорного плану, ближчого до оптимального, ніж попередній. Як бачимо, симплексний метод полягає в послідовному поліпшенні плану задачі лінійного програмування, тому його називають ще *методом послідовного поліпшення плану*.

Розв'язування задачі симплексним методом складається з двох етапів: знаходження початкового опорного плану і оптимального плану. Алгоритм симплексного методу застосовний лише до канонічної форми задачі лінійного програмування. Тому при розв'язуванні задачі лінійного програмування симплексним методом систему обмежень спочатку зводять до канонічної форми.

Розглянемо докладніше ідею симплексного методу. Нехай треба знайти мінімум лінійної функції (4.3) за обмежень (4.1), (4.2).

Припустимо, що поставлена задача має плани, кожний її опорний план не вироджений, і система обмежень містить m одиничних векторів. Не порушуючи загальності, можна вважати, що такими векторами є перші m векторів. Запишемо систему обмежень у векторному вигляді

$$A_1 x_1 + A_2 x_2 + \dots + A_m x_m + A_{m+1} x_{m+1} + \dots + A_n x_n = A_0, \quad X \geq 0,$$

де

$$A_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad A_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \quad A_{m+1} = \begin{pmatrix} a_{1,m+1} \\ a_{2,m+1} \\ \vdots \\ a_{m,m+1} \end{pmatrix},$$

$$\dots, \quad A_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}, \quad A_0 = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Вектори A_1, A_2, \dots, A_m — лінійно незалежні одиничні вектори m -вимірного простору, тому вони утворюють базис цього простору. За базисні змінні в системі обмежень виберемо x_1, x_2, \dots, x_m , а вільні змінні x_{m+1}, \dots, x_n прирівняємо до нуля. Врахувавши, що $b_i \geq 0$ ($i = 1, \dots, m$), знаходимо початковий план

$$X_0 = (x_1 = b_1; x_2 = b_2; \dots; x_m = b_m; x_{m+1} = 0; \dots; x_n = 0). \quad (4.9)$$

Йому відповідає лінійна комбінація

$$x_1 A_1 + x_2 A_2 + \dots + x_m A_m = A_0, \quad (4.10)$$

в якій вектори A_1, A_2, \dots, A_m лінійно незалежні. Отже, побудова-

ний початковий план (4.9) є опорним. Він надає лінійній функції значення

$$F(X_0) = c_1x_1 + c_2x_2 + \dots + c_mx_m = f_0, \quad (4.11)$$

яке залежить лише від перших m компонент вектора $C = (c_1, c_2, \dots, c_n)$. Оскільки оптимальний план належить до опорних, то і оптимальне значення лінійної функції залежатиме від відповідних компонент вектора C .

Розглянемо, як, виходячи з початкового опорного плану (4.9), можна побудувати інший опорний план, який надасть лінійній функції значення меншого, ніж попередній план. Початковий опорний план визначається базисними векторами A_1, A_2, \dots, A_m . Побудувати наступний опорний план, виходячи з початкового, можна завдяки переходу від базису A_1, A_2, \dots, A_m до нового базису. Такий перехід здійснюється за допомогою вибору вектора, який вводять у базис, і вибору вектора, який виводять з нього.

Знайдемо критерій вибору таких векторів. Оскільки вектори A_1, A_2, \dots, A_m утворюють базис m -вимірного простору, то будь-який вектор системи A_1, A_2, \dots, A_n однозначно лінійно визначається через базисні

$$A_j = x_{1j}A_1 + x_{2j}A_2 + \dots + x_{mj}A_m, \quad j = 1, 2, \dots, n. \quad (4.12)$$

Зауважимо, що коефіцієнти x_{ij} ($i = 1, 2, \dots, m$) у виразі (4.12) є компонентами вектора A_j в базисі A_1, A_2, \dots, A_m . Виразу (4.12) вектора A_j відповідає єдине значення лінійної функції

$$f_j = c_1x_{1j} + c_2x_{2j} + \dots + c_mx_{mj}, \quad j = 1, 2, \dots, n, \quad (4.13)$$

де c_i ($i = 1, 2, \dots, m$) — коефіцієнти лінійної функції, які відповідають базисним векторам A_i ($i = 1, 2, \dots, m$). Значення f_j знайдемо, якщо у вираз лінійної функції замість змінних підставимо відповідні коефіцієнти подання j -го вектора через базисні вектори.

Нехай для деякого вектора, що не входить у базис, наприклад A_k ($k > m$), хоча б один з коефіцієнтів x_{ik} у виразі

$$x_{1k}A_1 + x_{2k}A_2 + \dots + x_{mk}A_m = A_k \quad (4.14)$$

додатний. Виразу (4.14) відповідає значення лінійної функції

$$c_1x_{1k} + c_2x_{2k} + \dots + c_mx_{mk} = f_k. \quad (4.15)$$

Задамо деяке, поки що невідоме, додатне число θ . Помножимо на нього обидві частини рівностей (4.14) і (4.15) і результати віднімемо від (4.10) і (4.11) відповідно:

$$\begin{aligned} (x_1 - \theta x_{1k})A_1 + (x_2 - \theta x_{2k})A_2 + \dots + \\ + (x_m - \theta x_{mk})A_m + \theta A_k = A_0, \end{aligned} \quad (4.16)$$

$$\begin{aligned} (x_1 - \theta x_{1k})c_1 + (x_2 - \theta x_{2k})c_2 + \dots + \\ + (x_m - \theta x_{mk})c_m = f_0 - \theta f_k. \end{aligned}$$

В останній рівності до обох частин додамо величину θc_k :

$$\begin{aligned} (x_1 - \theta x_{1k})c_1 + (x_2 - \theta x_{2k})c_2 + \dots + \\ + (x_m - \theta x_{mk})c_m + \theta c_k = f_0 - \theta(f_k - c_k). \end{aligned} \quad (4.17)$$

Лінійній комбінації (4.16) відповідає новий план

$$X_1 = (x_1 - \theta x_{1k}, x_2 - \theta x_{2k}, \dots, x_m - \theta x_{mk}, 0, \dots, \theta, \dots, 0),$$

якщо його коефіцієнти невід'ємні. Ті компоненти вектора X_1 , до яких входять недодатні x_{ik} , будуть невід'ємними, бо $\theta > 0$. Компоненти вектора X_1 з додатними x_{ik} ($i = 1, 2, \dots, m$) будуть невід'ємними, якщо

$$\theta \leq \min_i \frac{x_i}{x_{ik}}, \quad (4.18)$$

де мінімум береться серед тих i , для яких $x_{ik} > 0$. Отже, при виконанні умови (4.18) вектор X_1 буде планом. Опорний план не може мати $m+1$ додатних компонент. Тому в плані X_1 треба хоча б одну з компонент перетворити в нуль. Це можна зробити, якщо θ вибрати так:

$$\theta = \theta_0 = \min_i \frac{x_i}{x_{ik}}. \quad (4.19)$$

Оскільки розглядається не вироджена задача, тобто всі її опорні плани містять рівно m додатних компонент, то мінімум в (4.19) досягатиметься для єдиного значення i . Нехай $i = l$ ($1 \leq l \leq m$). Тоді для $\theta_0 = x_l / x_{lk}$ відповідні коефіцієнти виразу (4.16) і компоненти плану X_1 перетворяться в нуль. Підставивши значення θ_0 в (4.16) і план X_1 , дістанемо лінійну комбінацію

$$x'_1A_1 + \dots + x'_{l-1}A_{l-1} + x'_{l+1}A_{l+1} + \dots + x'_mA_m + x'_kA_k = A_0,$$

і відповідний їй опорний план

$$X_1 = (x'_1, \dots, x'_{l-1}, 0, x'_{l+1}, \dots, x'_m, 0, \dots, x'_k, \dots, 0),$$

де

$$\begin{cases} x'_i = x_i - \frac{x_l}{x_{lk}} x_{ik}, & i \neq l, \\ x'_k = \frac{x_l}{x_{lk}}, & i = l. \end{cases} \quad (4.20)$$

План X_1 надає лінійній функції F значення $f_0 - \theta_0(f_k - c_k)$.

Отже, дістали новий опорний план, базис якого складається з векторів $A_1, A_2, \dots, A_{l-1}, A_{l+1}, \dots, A_m, A_k$. Новий план X_1 надає лінійній функції F значення $F(X_1) = f_0 - \theta_0(f_k - c_k)$, яке буде менше від $F(X_0)$, якщо $f_k - c_k > 0$. Величина різниці $f_k - c_k$ називається оцінкою плану. Якщо для деякого вектора A_k оцінка плану $f_k - c_k > 0$, то план X_0 не є оптимальним і можна побудувати новий план X_1 такий, що $F(X_1) < F(X_0)$. План X_1 побудовано завдяки заміні вектора з початкового базису новим вектором. Критерій заміни векторів можна сформулювати так:

- 1) у новий базис вводять вектор A_k , для якого оцінка плану додатна;
- 2) із старого базису виводять той вектор A_l , для якого відношення

$\frac{x_l}{x_{lk}}$ з додатними x_{lk} набуває найменшого значення.

Процес заміни векторів продовжують доти, поки всі оцінки плану не стануть від'ємними або нулями, або для деякої оцінки $f_j - c_j > 0$ усі x_{ij} стануть недодатними. Недодатність оцінок плану означає, що побудований план — оптимальний.

Якщо на деякому кроці заміни векторів для якого-небудь j оцінка $f_j - c_j > 0$ й усі $x_{ij} \leq 0$, то це означає, що лінійна функція не обмежена знизу на многограннику розв'язків і її значення може бути як завгодно малим.

Значимо, що коли є кілька додатних оцінок плану, то вибирають найбільшу з них і той вектор, який відповідає цій найбільшій оцінці, вводять у новий базис. Якщо є кілька однакових найбільших оцінок, то можна вводити в новий базис будь-який з векторів, які відповідають цим найбільшим оцінкам. Якщо для деякого вектора A_j з додатною оцінкою плану є кілька однакових найменших відношень x_i / x_{ij} , то з базису можна вивести будь-який з відповідних векторів.

Нагадаємо, що компоненти нового плану можна обчислити за формулами (4.20). Для перевірки побудованого плану на оптимальність треба знати компоненти усіх векторів A_j ($j = 1, 2, \dots, n$) у новому базисі. Їх можна обчислити за формулами

$$a_{ij}' = a_{ij} - \frac{a_{ik}}{a_{lk}} a_{lj}, \quad i \neq l, \quad (4.21)$$

$$a_{ij}' = a_{ij} / a_{lk}, \quad i = l. \quad (4.22)$$

Таким чином, маючи компоненти всіх векторів системи обмежень у новому базисі і обчисливши оцінки плану, можна перевірити побудований план X_1 на оптимальність. Проілюструємо сказане прикладом.

П р и к л а д 9. Знайти мінімум лінійної функції $F = 2x_1 + 3x_2 + x_3 + 4x_4$ за обмежень

$$\begin{cases} 4x_1 + 11x_2 + x_3 = 70, \\ 2,5x_1 + 8x_2 + x_4 = 45, \\ x_j \geq 0 \quad (j=1,2,3,4). \end{cases}$$

Р о з в ' я з а н н я. Відповідно до загальних позначень маємо

$$X = (x_1, x_2, x_3, x_4), \quad C = (2; 3; 1; 4),$$

$$A_1 = \begin{pmatrix} 4 \\ 2,5 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 11 \\ 8 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad A_0 = \begin{pmatrix} 70 \\ 45 \end{pmatrix}.$$

Система обмежень містить два одиничні лінійно незалежні вектори A_3 і A_4 , які утворюють базис двовимірного векторного простору. Тоді вектори A_j ($j = 0, 1, \dots, 4$) виражаються через базисні

$$A_0 = x_3 A_3 + x_4 A_4 = 70 A_3 + 45 A_4,$$

$$A_1 = x_{31} A_3 + x_{41} A_4 = 4 A_3 + 2,5 A_4,$$

$$A_2 = x_{32} A_3 + x_{42} A_4 = 11 A_3 + 8 A_4,$$

$$A_3 = x_{33} A_3 + x_{43} A_4 = 1 \cdot A_3 + 0 \cdot A_4,$$

$$A_4 = x_{34} A_3 + x_{44} A_4 = 0 \cdot A_3 + 1 \cdot A_4.$$

Поклавши в системі обмежень $x_1 = x_2 = 0$, дістанемо початковий опорний план $X_0 = (0, 0, 70, 45)$. Значення цільової функції F дорівнює $F(X_0) = 2 \cdot 0 + 3 \cdot 0 + 1 \cdot 70 + 4 \cdot 45 = 250$. Знайдемо оцінки початкового опорного плану. Для цього обчислимо значення f_j ($j = 1, \dots, 4$) за формулою (4.13) як суму добутків компонент вектора A_j на компоненти вектора C , які відповідають базисним векторам

$$f_1 = 1 \cdot 4 + 4 \cdot 2,5 = 14; \quad f_2 = 1 \cdot 11 + 4 \cdot 8 = 43;$$

$$f_3 = 1 \cdot 1 + 4 \cdot 0 = 1; \quad f_4 = 1 \cdot 0 + 4 \cdot 1 = 4.$$

Тоді оцінки плану дорівнюють

$$f_1 - c_1 = 12, \quad f_2 - c_2 = 40, \quad f_3 - c_3 = 0, \quad f_4 - c_4 = 0.$$

Серед них є дві додатні. Більшою є оцінка $f_2 - c_2$, що відповідає вектору A_2 . Отже, його вводимо в базис. Щоб визначити, який з векторів A_3, A_4 треба вивести з базису, обчислимо відношення x_i / x_{i2} ($i = 3, 4$) для додатних x_{i2} , тобто поділимо компоненти вектора A_0 на відповідні додатні компоненти вектора A_2 :

$$A_0 : \frac{x_3}{x_{32}} = \frac{70}{11}, \quad \frac{x_4}{x_{42}} = \frac{45}{8}.$$

Меншим з обчислених відношень є відношення x_4 / x_{42} , яке відповідає вектору A_4 . Таким чином, замість вектора A_4 в базис вводимо вектор A_2 . Новий базис утворюють вектори A_2 і A_3 . У новому базисі вектори A_j ($j = 0, 1, \dots, 4$) мають такі вирази:

$$A_0 = x_2' A_2 + x_3' A_3,$$

$$A_1 = x_2' A_2 + x_3' A_3,$$

$$A_2 = x'_{22}A_2 + x'_{23}A_3 = A_2 + 0 \cdot A_3,$$

$$A_3 = x'_{32}A_2 + x'_{33}A_3 = A_3 + 0 \cdot A_2,$$

$$A_4 = x'_{42}A_2 + x'_{43}A_3,$$

де x' обчислюють за формулами (4.20), а x'_j — згідно з формулами (4.21) і (4.22). Новий опорний план X_1 , який визначається базисними векторами A_2 і A_3 , має вигляд $X_1 = (0, x'_2, x'_3, 0)$. Обчислимо його компоненти

$$x'_2 = \frac{x_4}{x_{42}} = \frac{45}{8}, \quad x'_3 = x_3 - \frac{x_4}{x_{42}} x_{32} = 70 - \frac{45}{8} \cdot 11 = \frac{65}{8}.$$

Отже, $X_1 = (0; \frac{45}{8}; \frac{65}{8}; 0)$. Значення лінійної функції F дорівнює

$$F(X_1) = c_2 x'_2 + c_3 x'_3 = \frac{3 \cdot 45}{8} + \frac{65}{8} = 25.$$

Обчислимо тепер компоненти векторів A_1 і A_4 в новому базисі:

$$x'_{21} = \frac{x_{41}}{x_{42}} = \frac{5}{16}, \quad x'_{31} = x_{31} - \frac{x_{41}}{x_{42}} x_{32} = 4 - \frac{5}{2} \cdot \frac{1}{8} \cdot 11 = \frac{1}{16},$$

$$x'_{41} = \frac{x_{44}}{x_{42}} = \frac{1}{8}, \quad x'_{34} = x_{34} - \frac{x_{44}}{x_{42}} x_{32} = 0 - \frac{1}{8} \cdot 11 = -\frac{11}{8}.$$

Тепер знаходимо значення f_j ($j = 1, 2, 3, 4$):

$$f_1 = \frac{3 \cdot 5}{16} + \frac{1}{16} = 1, \quad f_2 = 3 \cdot 1 + 1 \cdot 0 = 3,$$

$$f_3 = 1 \cdot 1 + 3 \cdot 0 = 1, \quad f_4 = -\frac{11}{8} + \frac{3}{8} = -1.$$

Нарешті, визначимо оцінки плану X_1 :

$$f_1 - c_1 = 1 - 2 = -1, \quad f_2 - c_2 = 0, \quad f_3 - c_3 = 0, \quad f_4 - c_4 = -5.$$

Усі оцінки опорного плану X_1 недодатні. Це означає, що побудований план оптимальний. Мінімум лінійної функції $F_{\min} = 25$ досягається при плані $X = (0; \frac{45}{8}; \frac{65}{8}; 0)$.

§ 4.5. ЗНАХОДЖЕННЯ ПОЧАТКОВОГО ОПОРНОГО ПЛАНУ

Щоб застосувати симплекс-метод для знаходження оптимального розв'язку задачі лінійного програмування, треба знати відправну точку — початковий опорний план. Якщо обмеження задачі лінійного програмування задано в канонічному вигляді

$$AX = A_0, \quad A_0 \geq 0, \quad X \geq 0 \quad (4.23)$$

і серед векторів A_1, A_2, \dots, A_n є одиничний базис, то початковим опорним планом буде вектор $X = (b_1, b_2, \dots, b_m, 0, \dots, 0)$. У деяких випадках одиничний базис у системі обмежень (4.23) легко виділити. Наприклад, нехай маємо систему обмежень

$$\begin{cases} 3x_1 - x_4 - 2x_6 = 6, \\ 2x_2 + x_4 + 3x_5 + x_6 = 8, \\ x_3 - 2x_4 + 5x_5 + 6x_6 = 5, \\ x_j \geq 0, \quad (j=1, 2, \dots, 6). \end{cases}$$

Якщо перше рівняння системи поділити на 3, а друге — на 2, то дістанемо систему

$$\begin{cases} x_1 - \frac{1}{3}x_4 - \frac{2}{3}x_6 = 2, \\ x_2 + \frac{1}{2}x_4 + \frac{3}{2}x_5 + \frac{1}{2}x_6 = 4, \\ x_3 - 2x_4 + 5x_5 + 6x_6 = 5, \\ x_j \geq 0, \quad (j=1, 2, \dots, 6). \end{cases}$$

У ній вектори A_1, A_2 і A_3 утворюють одиничний базис і всі вільні члени додатні. Поклавши $x_4 = x_5 = x_6 = 0$, знайдемо опорний план $X = (2, 4, 5, 0, 0, 0)$.

Якщо система (4.23) не містить у явному вигляді одиничного базису, то його в деяких випадках можна виділити методом повного виключення Гаусса.

П р и к л а д 10. Знайти опорний план системи обмежень

$$\begin{cases} x_1 + 2x_2 + 3x_4 - x_5 = 6, \\ x_1 + 3x_2 + x_3 + 4x_4 - x_5 = 12, \\ 2x_1 - x_3 + 12x_4 - x_5 = 14, \\ x_j \geq 0 \quad (j=1, 2, \dots, 5). \end{cases}$$

Р о з в ' я з а н н я. Застосуємо метод повного виключення Гаусса. Результати обчислень подано в табл. 4. 6, де ведучий елемент взято в рамку.

Після третьої ітерації перетворень системи методом повного виключення дістали одиничний базис: A_1, A_2, A_5 , йому відповідає план $X = (20, 6, 0, 0, 26)$.

Якщо обмеження задачі лінійного програмування задано у вигляді нерівностей

$$A_1x_1 + A_2x_2 + \dots + A_nx_n \leq A_0, \quad X \geq 0, \quad A_0 \geq 0, \quad (4.24)$$

то, додавши до лівої частини кожної нерівності системи (4. 24) по невід'ємній змінній $x_{n+1}, x_{n+2}, \dots, x_{n+m}$, дістанемо розширену систему лінійних обмежень

$$A_1x_1 + \dots + A_nx_n + A_{n+1}x_{n+1} + \dots + A_{n+m}x_{n+m} = A_0 \quad (4.25)$$

$$X \geq 0, A_0 \geq 0.$$

Таблиця 4.6

Ітерація	A_1	A_2	A_3	A_4	A_5	A_0	Σ
1	1	2	0	3	-1	6	11
	1	3	1	4	-1	12	20
	2	0	-1	12	-1	14	26
2	1	2	0	3	-1	6	11
	0	1	1	1	0	6	9
	0	-4	-1	6	1	2	4
3	1	0	-2	1	-1	-6	-7
	0	1	1	1	0	6	9
	0	0	3	10	1	26	40
4	1	0	1	11	0	20	33
	0	1	1	1	0	6	9
	0	0	3	10	1	26	40

Змінні $x_{n+1}, x_{n+2}, \dots, x_{n+m}$ називають *додатковими змінними*. В лінійну функцію вони входять з нульовими коефіцієнтами. Отже, початкова задача лінійного програмування перетворилась у розширену: знайти оптимальне значення лінійної функції

$$F = c_1x_1 + c_2x_2 + \dots + c_nx_n + 0 \cdot x_{n+1} + \dots + 0 \cdot x_{n+m}$$

за обмежень (4.25). Додаткові вектори $A_{n+1}, A_{n+2}, \dots, A_{n+m}$ утворюють одиничний базис m -вимірного векторного простору, а вектор $X = (x_1 = 0; \dots; x_m = 0; x_{m+1} = b_1; \dots; x_{n+m} = b_{n+m})$ є опорним планом розширеної задачі.

Зазначимо, що розв'язок розширеної задачі (якщо він існує) буде також розв'язком початкової задачі. Якщо розширена задача розв'язку не має, то не має розв'язку і початкова задача.

Застосовуючи симплексний метод до розширеної задачі, поступово замінюють у системі базисних векторів додаткові вектори $A_{n+1}, A_{n+2}, \dots, A_{n+m}$ векторами початкової системи обмежень. Якщо лінійна функція досягла свого оптимального значення, а в системі базисних векторів є хоча б один додатковий вектор, наприклад A_{n+i} , то це означає, що i -те

обмеження в початковій задачі має смисл строгої нерівності. Отже, початкова задача матиме оптимальний розв'язок (якщо його має розширена задача) і тоді, коли не всі додаткові вектори виведено з базису.

Часто, виділивши в системі обмежень одиничний базис і базисний розв'язок, який йому відповідає, можна впевнитися, що знайдений розв'язок не є опорним планом системи обмежень, бо серед базисних змінних є й від'ємні. З метою цілеспрямованого пошуку опорного плану треба від знайденого одиничного базису перейти до іншого. Для цього застосовують метод *симплексного перетворення*. Оскільки симплексні перетворення виконують над векторами $A_1, A_2, \dots, A_n, A_0$ системи обмежень, то ці вектори зведемо в табл. 4.7.

Таблиця 4.7

N рядка	A_0	A_1	A_2	...	A_j	...	A_n	Σ	θ
1	b_1	a_{11}	a_{12}	...	a_{1j}	...	a_{1n}		
2	b_2	a_{21}	a_{22}	...	a_{2j}	...	a_{2n}		
...
i	b_i	a_{i1}	a_{i2}	...	a_{ij}	...	a_{in}		
...
m	b_m	a_{m1}	a_{m2}	...	a_{mj}	...	a_{mn}		

Нехай вектор $A_0 \geq 0$, тобто всі $b_i \geq 0$ ($i = 1, 2, \dots, m$). Цього можна досягти, помноживши рівняння з від'ємними членами на -1 . Тепер задача полягає в тому, щоб у табл. 4.7 виділити одиничний базис, не порушивши невід'ємність компонент вектора A_0 . Перетворюють таблицю за таким алгоритмом.

1. З неединичних векторів A_1, A_2, \dots, A_n взяти той, у якого є хоча б один додатний елемент. Нехай таким вектором буде A_j . Якщо такого вектора в таблиці немає, то це означає, що система обмежень несумісна і процес симплексного перетворення завершено.

2. Знайти відношення θ_i елементів вектора A_0 до відповідних додатних елементів вектора A_j , записати їх у відповідному рядку стовпця θ і взяти з цих відношень найменше. Нехай таким буде відношення $\theta = b_i / a_{ij}$. Тоді елемент a_{ij} — ведучий, а рядок i і стовпець таблиці, на перетині яких лежить a_{ij} , відповідно ведучі рядок і стовпець.

3. Коефіцієнти ведучого рядка (крім θ_i) таблиці поділити на ведучий елемент і записати у відповідному рядку нової таблиці.

4. Елементи кожного наступного рядка нової таблиці дістають шляхом додавання відповідного рядка вихідної таблиці і рядка, записаного в п. 3, помноженого на таке число, щоб у ведучому стовпці при додаванні дістали нулі. На цьому заповнення нової таблиці завершується і перетворення нової таблиці починають з п. 1.

Такі перетворення продовжують доти, поки в таблиці не буде виді-

лено одиничний базис, якому відповідає опорний план, або ж не буде встановлено, що система обмежень несумісна. Столпець Σ в таблиці призначений для контролю обчислень, його елементи обчислюють так, як і за методом повного виключення Гаусса.

Зазначимо, що при виборі вектора A_j треба стежити за тим, щоб не повернутися до попереднього базису. Ця умова додатково обмежує вибір ведучого елемента. Розглянемо приклади.

Приклад 11. Знайти початковий опорний план системи

$$\begin{cases} x_1 - x_2 + x_3 - x_4 + x_5 = 6, \\ 2x_2 - x_3 - 3x_4 = 4, \\ x_1 + 2x_3 + x_4 = 2. \end{cases}$$

Розв'язання. Записуємо коефіцієнти системи в табл. 4.8.

Таблиця 4.8

Ітерація	Рядок	A_0	A_1	A_2	A_3	A_4	A_5	Σ	θ
1	1	6	1	-1	1	-1	1	7	6
	2	4	0	2	-1	-3	0	2	
	3	2	1	0	2	1	0	6	2
2	1	4	0	-1	-1	2	1	1	
	2	4	0	2	-1	-3	0	2	2
	3	2	1	0	2	1	0	6	
3	1	6	0	0	-3/2	-7/2	1	2	
	2	2	0	1	-1/2	-3/2	0	1	
	3	2	1	0	2	1	0	6	

За ведучий візьмемо, наприклад перший стовпець, оскільки вектор A_1 має додатні компоненти. Знайдемо відношення $\theta_1 = \frac{6}{1} = 6$ і $\theta_3 = 2$.

Менше з них лежить у третьому рядку. Отже, ведучий елемент $a_{31} = 1$. Над елементами вихідної системи виконаємо симплексні перетворення. Після першої ітерації перетворень одиничними є вектори A_1 і A_5 . Намагаючись знайти базис з векторів A_2, A_3, A_4 , візьмемо A_2 , що має додатний елемент $a_{22} = 2$, який і буде ведучим. Виконавши симплексні перетворення, після другої ітерації маємо одиничний базис A_1, A_2, A_5 і опорний план $X = (2; 2; 0; 0; 6)$, який йому відповідає.

Часто метод симплексних перетворень застосовують після повного виключення Гаусса. Останнім знаходять початковий одиничний базис і базисний розв'язок системи обмежень, який йому відповідає. Якщо знайдений базисний розв'язок не є опорним планом, тоді пошук одиничного базису продовжують методом симплексних перетворень.

Приклад 12. Знайти початковий опорний план системи

$$\begin{cases} 2x_1 + 3x_2 + 4x_3 + x_4 = 2, \\ x_1 + x_2 + 7x_3 + x_4 = 6, \\ 3x_1 + 2x_2 + x_3 + 5x_4 = 8. \end{cases}$$

Розв'язання. Спочатку виділимо одиничний базис методом повного виключення Гаусса. Після третьої ітерації (табл. 4.9) маємо одиничний базис A_1, A_2, A_4 і розв'язок $X_0 = (56; -30; 0; -20)$, який не є опорним. Тепер до нової таблиці застосуємо метод симплексних перетворень. Для цього її перший і третій рядки помножимо на -1 . Після шостої ітерації маємо таблицю, в якій тільки один відмінний від одиничного вектор A_1 з додатними елементами. Якщо взяти елемент $a_{21} = 30/77$ за ведучий і виконати один крок симплексного перетворення, то дістанемо базис A_1, A_3 . Цей базис уже мали після п'ятої ітерації і він не давав опорного плану. Очевидно, що процес симплексних перетворень завершився, бо неможливо вибрати ведучий стовпець за наведеними вище правилами. Отже, в цьому разі задана система обмежень не має жодного опорного плану.

Таблиця 4.9

Ітерація	Рядок	A_0	A_1	A_2	A_3	A_4	Σ	θ
1	1	2	2	3	4	1	12	
	2	6	1	1	7	1	16	
	3	8	3	2	1	5	19	
2	1	-10	0	1	-10	-1	-20	
	2	6	1	1	7	1	16	
	3	-10	0	-1	-20	2	-29	
3	1	-10	0	1	-10	-1	-20	
	2	16	1	0	17	2	36	
	3	-20	0	0	-30	1	-49	
4	1	-30	0	1	-40	0	-69	
	2	56	1	0	77	0	134	
	3	-20	0	0	-30	1	-49	
5	1	30	0	-1	40	0	69	3/4
	2	56	1	0	77	0	134	56/77
	3	20	0	0	30	-1	49	2/3

6	1	10/3	0	-1	0	4/3	11/3	10/3
	2	14/3	1	0	0	77/30	247/30	140/77
	3	2/3	0	0	1	-1/30	49/30	
7	1	70/77	-40/77	-1	0	0	-47/77	
	2	140/77	30/77	0	0	1	247/77	14/3
	3	56/77	1/77	0	1	0	134/77	56

§ 4.6. ВИКОРИСТАННЯ СИМПЛЕКС-ТАБЛИЦЬ

Розглянемо задачу лінійного програмування на знаходження мінімуму лінійної функції (4.3) за обмежень

$$\begin{cases} x_1 + a_{1,m+1}x_{m+1} + \dots + a_{1n}x_n = b_1, \\ x_2 + a_{2,m+1}x_{m+1} + \dots + a_{2n}x_n = b_2, \\ \dots \\ x_m + a_{m,m+1}x_{m+1} + \dots + a_{mn}x_n = b_m, \\ x_j \geq 0 \quad (j=1,2,\dots,n). \end{cases} \quad (4.26)$$

Систему обмежень подамо у векторному вигляді

$$x_1A_1 + x_2A_2 + \dots + x_mA_m + \dots + x_nA_n = A_0, \quad (4.27)$$

де

$$A_1 = \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 \\ 1 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \quad \dots, \quad A_m = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{pmatrix},$$

$$A_{m+1} = \begin{pmatrix} a_{1,m+1} \\ a_{2,m+1} \\ \cdot \\ \cdot \\ a_{m,m+1} \end{pmatrix}, \quad A_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdot \\ \cdot \\ a_{mn} \end{pmatrix}, \quad A_0 = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_m \end{pmatrix}.$$

Початковий опорний план $X_0 = (x_1 = b_1; \dots; x_m = b_m; 0; \dots; 0)$ задачі (4.3), (4.26) визначається системою m -вимірних одиничних векторів $A_1,$

A_2, \dots, A_m . Щоб дослідити цей опорний план на оптимальність, треба вектори A_j ($j = 1, 2, \dots, n$) системи (4.27) подати лінійною комбінацією базисних векторів і обчислити значення оцінок $f_j - c_j$. Оскільки базис системи (4.26) — одиничний, то коефіцієнтами у виразі вектора A_j через базисні будуть його компоненти, тобто $x_{ij} = a_{ij}$ ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$). Обчислення наступного опорного плану та перевірку його оптимальності зручно виконати, записавши умову задачі і дані, знайдені після побудови початкового опорного плану, в симплексну табл. 4.10.

У ній у першому стовпці записано номери рядків таблиці. Стовпець базису містить базисні вектори. У стовпці С базису записано коефіцієнти лінійної функції (4.3), які відповідають базисним векторам. У стовпці A_0 — початковий опорний план. У цьому самому стовпці в результаті обчислень знаходять оптимальний план. Стовпці A_j ($j = 1, 2, \dots, n$) заповнено коефіцієнтами вектора A_j через базисні вектори. В $(m+1)$ -му рядку стовпця A_0 записано значення лінійної функції $F(X_0)$, якого вона набуває при початковому опорному плані, а в стовпцях A_j — значення оцінок $f_j - c_j$. Оцінки можна обчислити, якщо від добутків елементів стовпця A_j на відповідні елементи стовпця С базису відняти відповідні коефіцієнти c_j . Далі обчислення симплекс-методом слід провести за таким алгоритмом.

1. Розглянути оцінки плану $(m+1)$ -го рядка. Якщо всі оцінки плану недодатні, то опорний план X_0 оптимальний і мінімум лінійної функції дорівнює $F(X_0)$. Якщо серед оцінок плану є хоча б одна додатна, то слід виконати п.2.

2. Якщо хоч для однієї додатної оцінки $f_j - c_j$ усі елементи стовпця A_j недодатні, то це означає, що лінійна функція на многограннику розв'язків не обмежена знизу і задача не має оптимального розв'язку. Якщо в кожному стовпці A_j з додатними оцінками плану є хоча б один додатний коефіцієнт, то слід перейти до п.3.

Коментар 1. Якщо в кожному стовпці A_j , які відповідають додатним оцінкам плану, є додатні коефіцієнти, то план X_0 не є оптимальним і можна побудувати новий опорний план, який надасть лінійній функції значення, меншого від попереднього. Задача полягає в тому, щоб один з базисних векторів замінити новим. Вектор, який вводять у базис, вибирають в п.3, а вектор, який виводять з базису, — в п.5.

3. Вибрати вектор з найбільшою додатною оцінкою. Нехай це буде вектор A_k . Перейти до п.4.

4. Обчислити відношення елементів стовпця A_0 до відповідних додатних елементів стовпця A_k , результати записати у відповідних рядках стовпця θ і перейти до п.5.

5. Серед векторів базису вибрати той, який відповідає найменшому з відношень, обчислених у п.4. Нехай це буде вектор A_i . Перейти до п.6.

Коментар 2. Елемент x_k називається ведучим, рядок і стовпець, на перетині яких він міститься, називаються ведучими. Після визначення ведучого елемента заповнюють нову симплексну таблицю. Перші три стовпці нової таблиці відрізнятимуться від старої тільки

Рядок	Базис	C базис- c_j	A_0	c_1	c_2	c_j	c_m	c_{m+1}	c_j	c_k	c_n
1	A_1	c_1	x_1	1	0	0	0	$x_{1,m+1}$	x_{1j}	x_{1k}	x_{1n}
2	A_2	c_2	x_2	0	1	0	0	$x_{2,m+1}$	x_{2j}	x_{2k}	x_{2n}
...
l	A_l	c_l	x_l	0	0	1	0	$x_{l,m+1}$	x_{lj}	x_{lk}	x_{ln}
...
m	A_m	c_m	x_m	0	0	0	1	$x_{m,m+1}$	x_{mj}	x_{mk}	x_{mn}
$m+1$		$f_j - c_j$	f_0	0	0	0	0	$f_{m+1} - c_{m+1}$	$f_j - c_j$	$f_k - c_k$	$f_n - c_n$

рядком з номером l . У ньому замість елементів l, A_l, c_l будуть записані елементи l, A_k, c_k . Інші клітини нової таблиці заповнюють згідно пп.6 і 7.

6. Поділити елементи l -го рядка, які відповідають векторам A_j ($j = 0, 1, \dots, n$), на ведучий елемент і результати записати у відповідні стовпці l -го рядка нової таблиці. Перейти до п.7.

7. Для знаходження елементів l -го ($l = 1, 2, \dots, l-1, l+1, \dots, m+1$) рядка нової таблиці треба від елементів l -го рядка старої таблиці відняти відповідні елементи l -го рядка нової таблиці, помножені на x_{lk} ($l \neq l$). Після заповнення нової таблиці перейти до п.1.

Зазначимо, що коли в симплексній таблиці є кілька додатних оцінок плану, то при ручних обчисленнях у базис вводять вектор, якому відповідає $\max \{\theta_{0j}(f_j - c_j)\}$, де максимум береться серед тих j , для яких $f_j - c_j > 0$, і $\theta_{0j} = \min (x_i / x_{ij})$, ($x_{ij} \geq 0$) обчислюється для кожного j .

Завдяки такому вибору вектора вдається досягти оптимального плану за мінімум ітерацій. Якщо ж є кілька рівних найбільших значень $\theta_{0j}(f_j - c_j)$, то з відповідних їм векторів вибирають той, якому відповідає $\min c_j$.

Якщо серед оцінок оптимального плану нульовими є тільки оцінки, які відповідають базисним векторам, то це означає, що оптимальний план єдиний. Якщо нульовій оцінці плану відповідає вектор, який не входить у базис, то оптимальний план не є єдиним.

Задачу знаходження максимального значення лінійної функції можна розв'язати, не зводячи її до задачі мінімізації. План задачі знаходження максимуму лінійної функції буде оптимальним тоді і тільки тоді, коли його оцінки будуть невід'ємними. Якщо умова оптимальності не виконується, то при ручних обчисленнях у базис вводять вектор, якому відповідає $\min \{\theta_{0j}(f_j - c_j)\}$, де мінімум береться серед тих j , для яких $f_j - c_j < 0$. Якщо мінімальних оцінок кілька, то в базис вводять вектор, якому відповідає $\max c_j$.

Приклад 13. Симплексним методом знайти мінімальне значення функції $F = x_1 + x_2 + x_3$ за обмежень

$$\begin{cases} x_1 - x_4 - 2x_6 = 5, \\ x_2 + 2x_4 - 3x_5 + x_6 = 3, \\ x_3 + 2x_4 - 5x_5 + 6x_6 = 5, \\ x_j \geq 0 \quad (j = 1, 2, \dots, 6). \end{cases}$$

Розв'язання. Запишемо систему обмежень у векторному вигляді

$$x_1 A_1 + x_2 A_2 + \dots + x_6 A_6 = A_0,$$

$$A_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, A_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, A_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, A_4 = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix},$$

$$A_5 = \begin{pmatrix} 0 \\ -3 \\ -5 \end{pmatrix}, A_6 = \begin{pmatrix} -2 \\ 1 \\ 6 \end{pmatrix}, A_0 = \begin{pmatrix} 5 \\ 3 \\ 5 \end{pmatrix}.$$

Дана система обмежень в явному вигляді містить одиничний базис A_1, A_2, A_3 . Тому змінні x_1, x_2, x_3 — базисні, а x_4, x_5, x_6 — вільні змінні. Поклавши $x_4 = x_5 = x_6 = 0$, знайдемо початковий опорний план $X_0 = (x_1 = 5; x_2 = 3; x_3 = 5; x_4 = 0; x_5 = 0; x_6 = 0)$, якому відповідає лінійна комбінація

$$x_1 A_1 + x_2 A_2 + x_3 A_3 = A_0$$

і значення лінійної функції $F(X_0) = 1 \cdot 5 + 1 \cdot 3 + 1 \cdot 5 = 13$.

Для перевірки плану X_0 на оптимальність заповнимо першу симплексну табл. 4.11 і обчислимо значення оцінок

$$f_4 - c_4 = 1 \cdot (-1) + 1 \cdot 2 + 1 \cdot 2 - 0 = 3,$$

$$f_5 - c_5 = 1 \cdot 0 + 1 \cdot (-3) + 1 \cdot (-5) - 0 = -8,$$

$$f_6 - c_6 = 1 \cdot (-2) + 1 \cdot 1 + 1 \cdot 6 - 0 = 5.$$

Таблиця 4.11

Ітерація	Рядок	Базис	С базису	$c_j - c_j$						Σ	θ	
				$c_1 - 1$	$c_2 - 1$	$c_3 - 1$	$c_4 - 0$	$c_5 - 0$	$c_6 - 0$			
1	1	A_1	1	5	1	0	0	-1	0	-2	3	
	2	A_2	1	3	0	1	0	2	-3	1	4	$\frac{3}{1}$
	3	A_3	1	5	0	0	1	2	-5	6	9	$\frac{5}{6}$
	4	$f_j - c_j$		13	0	0	0	3	-8	5	13	
2	1	A_1	1	$\frac{20}{3}$	1	0	$\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{5}{3}$	0	6	
	2	A_2	1	$\frac{13}{6}$	0	1	$-\frac{1}{6}$	$\frac{5}{3}$	$-\frac{13}{6}$	0	$\frac{5}{2}$	$\frac{13}{10}$
	3	A_6	0	$\frac{5}{6}$	0	0	$\frac{1}{6}$	$\frac{1}{3}$	$-\frac{5}{6}$	1	$\frac{3}{2}$	$\frac{5}{2}$
	4	$f_j - c_j$		$\frac{53}{6}$	0	0	$-\frac{5}{6}$	$\frac{4}{3}$	$-\frac{23}{6}$	0	$\frac{11}{2}$	

3	1	A_1	1	$\frac{71}{10}$	1	$\frac{1}{5}$	$\frac{3}{10}$	0	$-\frac{21}{10}$	0	$\frac{13}{2}$	
	2	A_4	0	$\frac{13}{10}$	0	$\frac{3}{5}$	$-\frac{1}{10}$	1	$-\frac{13}{10}$	0	$\frac{3}{2}$	
	3	A_6	0	$\frac{2}{5}$	0	$-\frac{1}{5}$	$\frac{1}{5}$	0	$-\frac{2}{5}$	1	1	
	4	$f_j - c_j$		$\frac{71}{10}$	0	$-\frac{4}{5}$	$-\frac{7}{10}$	0	$-\frac{21}{10}$	0	$\frac{7}{2}$	

Для базисних векторів значення оцінок нульові. Серед обчислених оцінок є дві додатні. Більша з них $f_6 - c_6 = 5$ відповідає вектору A_6 , який і введемо в базис. Для визначення вектора, який треба вивести з базису, обчислимо $\theta_{06} = \min\left(\frac{3}{1}; \frac{5}{6}\right) = \frac{5}{6}$. Мінімальне відношення відповідає вектору A_3 базису. Отже, його треба замінити вектором A_6 . Ведучим елементом у першій симплексній таблиці є число 6, яке стоїть на перетині третього рядка і стовпця A_6 . Виконаємо одну ітерацію симплексних перетворень і заповнимо другу симплекс-таблицю. Нові елементи ведучого рядка знайдемо, поділивши елементи ведучого рядка початкової симплекс-таблиці на ведучий елемент (число 6). За допомогою обчисленого рядка знайдемо елементи інших рядків другої таблиці методом симплексних перетворень, тобто помножимо його елементи на 2 і додамо до відповідних елементів першого рядка початкової симплекс-таблиці; віднімемо його елементи від відповідних елементів другого рядка; помножимо його на -5 і додамо до четвертого рядка. В результаті дістанемо заповнену другу таблицю. Правильність проведених обчислень контролюємо за допомогою стовпця Σ , кожний елемент якого обчислюється двома способами: за допомогою симплексних перетворень та додаванням елементів відповідних рядків. У другій симплексній таблиці маємо новий опорний план $X_1 = \left(\frac{20}{3}; \frac{13}{6}; 0; 0; 0; \frac{5}{6}\right)$, якому відповідає значення лінійної функції $F(X_1) = \frac{53}{6} < F(X_0) = 13$.

У четвертому рядку цієї таблиці маємо додатну оцінку $f_4 - c_4 = \frac{4}{3}$. Отже, вектор A_4 вводимо до базису. Знайшовши $\theta_{04} = \min\left(\frac{13}{10}; \frac{5}{2}\right) = \frac{13}{10}$, побачимо, що вектор A_2 треба вивести з базису. Ведучий елемент стоїть на перетині другого рядка і стовпця A_4 . Виконаємо один крок симплексних перетворень. У результаті дістанемо опорний план $X_2 =$

$= \left(\frac{71}{10}; 0; 0; \frac{13}{10}; \frac{2}{5}\right)$. У четвертому рядку останньої симплекс-таблиці усі оцінки недодатні. Отже, план X_2 — оптимальний. Він надав лінійній функції значення $F(X_2) = \frac{71}{10}$. На основі оцінок цього плану робимо

висновок, що оптимальний план для даної задачі єдиний, оскільки нульові оцінки відповідають тільки базисним векторам.

П р и к л а д 14. Знайти мінімум лінійної функції $F = x_1 - x_2$ за обмежень

$$\begin{cases} x_1 - x_3 + x_5 = 1, \\ -3x_1 + x_2 - x_3 = 8, \\ 5x_1 - x_3 + x_4 = 15, \\ x_j \geq 0 \quad (j = 1, 2, \dots, 5). \end{cases}$$

Р о з в ' я з а н н я. Задана система обмежень містить одиничний базис A_3, A_4, A_5 . Поклавши $x_1 = x_2 = 0$, дістанемо $x_3 = 8, x_4 = 15, x_5 = -1$. Початковий опорний план $X_0 = (0; 8; 0; 15; 1)$ надає лінійній функції F значення $F(X_0) = -8$. Для базисних векторів значення оцінок плану дорівнюють нулю. Обчислимо інші значення оцінок:

$$f_1 - c_1 = 0 \cdot 1 + (-1) \cdot (-3) + 0 \cdot 5 - 1 = 2,$$

$$f_3 - c_3 = 0 \cdot (-1) + (-1) \cdot (-1) + 0 \cdot (-3) - 0 = 1.$$

Заповнимо початкову симплексну табл. 4. 12.

Таблиця 4. 12

Рядок	Базис	С базису	c_j					
			A_0	A_1	A_2	A_3	A_4	A_5
1	A_5	0	1	1	0	-1	0	1
2	A_2	-1	8	-3	1	-1	0	0
3	A_4	0	15	5	0	-3	1	0
4	$f_j - c_j$		-8	2	0	1	0	0

Проаналізувавши четвертий рядок таблиці, робимо висновок, що план X_0 не є оптимальним і його можна поліпшити за рахунок введення в базис вектора A_1 або A_3 . Але всі коефіцієнти стовпця A_3 від'ємні. Це означає, що лінійна функція даної задачі не обмежена на множині планів і задача не має оптимального розв'язку.

П р и к л а д 15. Задача про використання сировини (§ 4.1). Знайти максимальне значення лінійної функції $F = 40x_1 + 30x_2$ за обмежень

$$\begin{cases} 5x_1 + 6x_2 \leq 30, \\ 2x_1 + 6x_2 \leq 24, \\ 4x_1 + 2x_2 \leq 20, \\ x_1 \geq 0, \quad x_2 \geq 0. \end{cases}$$

Р о з в ' я з а н н я. Запишемо задачу в канонічному вигляді, додавши до лівої частини кожної з нерівностей системи обмежень по одній додатковій невід'ємній змінній. В результаті дістанемо

$$\begin{cases} 5x_1 + 6x_2 + x_3 = 30, \\ 2x_1 + 6x_2 + x_4 = 24, \\ 4x_1 + 2x_2 + x_5 = 20, \\ x_j \geq 0 \quad (j = 1, 2, \dots, 5). \end{cases}$$

Додаткові змінні мають конкретний економічний зміст. Вони визначають невикористану при певному плані кількість сировини того чи іншого виду. Наприклад, змінна x_3 визначає кількість невикористаної сировини першого виду.

Запишемо систему у векторній формі

$$x_1A_1 + x_2A_2 + x_3A_3 + x_4A_4 + x_5A_5 = A_0,$$

де

$$A_1 = \begin{pmatrix} 5 \\ 2 \\ 4 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 6 \\ 6 \\ 2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad A_5 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad A_0 = \begin{pmatrix} 30 \\ 24 \\ 20 \end{pmatrix}.$$

Одиничні вектори A_3, A_4, A_5 утворюють базис тривимірного простору. Вони задають початковий опорний план $X_0 = (0; 0; 30; 24; 20)$. Значення лінійної функції $F(X_0) = 40 \cdot 0 + 30 \cdot 0 = 0$. Для перевірки плану X_0 на оптимальність складаємо симплекс-таблицю (табл. 4. 13). Обчислимо значення оцінок плану

$$f_1 = 0 \cdot 5 + 0 \cdot 2 + 0 \cdot 4 = 0, \quad f_2 = 0 \cdot 6 + 0 \cdot 6 + 0 \cdot 2 = 0, \\ f_1 - c_1 = -40, \quad f_2 - c_2 = -30.$$

Для векторів базису оцінки дорівнюють нулю. Серед обчислених оцінок дві від'ємні: $f_1 - c_1 = -40, f_2 - c_2 = -30$. Отже, опорний план X_0 не є оптимальним. Найменша оцінка відповідає вектору A_1 , який введемо в базис. Визначимо вектор, виключений з базису. Для цього обчислимо

$$\theta_{01} = \min\left(\frac{30}{5}, \frac{24}{2}, \frac{20}{4}\right).$$

Найменше з відношень відповідає вектору A_5 , тому його замінюємо вектором A_1 . Ведучим елементом є число 4, яке стоїть на перетині третього рядка і стовпця A_1 . Складаємо другу симплекс-таблицю. Обчислимо нові значення ведучого рядка. Для цього елементи ведучого рядка поділимо на 4 і запишемо в третьому рядку другої таблиці. Тепер за допомогою цього рядка виконаємо одну ітера-

Таблиця 4. 13

Ітерація	Рядок	Базис	С базису						Σ	θ	
				$c_1=40$	$c_2=30$	$c_3=0$	$c_4=0$	$c_5=0$			
1	1	A_3	0	30	5	6	1	0	0	42	6
	2	A_4	0	24	2	6	0	1	0	33	12
	3	A_5	0	20	4	2	0	0	1	27	5
	4	$f_j - c_j$		0	-40	-30	0	0	0	-70	
2	1	A_3	0	5	0	$\frac{7}{2}$	1	0	$\frac{5}{4}$	$\frac{33}{4}$	$\frac{10}{7}$
	2	A_4	0	14	0	5	0	1	$\frac{1}{2}$	$\frac{39}{2}$	$\frac{14}{5}$
	3	A_1	40	5	1	$\frac{1}{2}$	0	0	$\frac{1}{4}$	$\frac{27}{4}$	10
	4	$f_j - c_j$		200	0	-10	0	0	10	200	
3	1	A_2	30	$\frac{10}{7}$	0	1	$\frac{2}{7}$	0	$\frac{5}{14}$	$\frac{33}{14}$	
	2	A_4	0	$\frac{48}{7}$	0	0	$\frac{10}{7}$	1	$\frac{9}{7}$	$\frac{54}{7}$	
	3	A_1	40	$\frac{30}{7}$	1	0	$\frac{1}{7}$	0	$\frac{3}{7}$	$\frac{39}{7}$	
	4	$f_j - c_j$		$\frac{1500}{7}$	0	0	$\frac{20}{7}$	0	$\frac{45}{7}$	$\frac{1565}{7}$	

цію симплексних перетворень, тобто помножимо його елементи на 5 і віднімемо від першого рядка першої ітерації, потім помножимо його елементи на 2 і віднімемо від другого рядка першої ітерації, помножимо його на 40 і додамо до четвертого рядка першої ітерації. Правильність проведених обчислень перевіряємо за допомогою стовпця Σ, елементи якого обчислюють двічі: методом симплексних перетворень і підсумовуванням елементів рядка.

Після другої ітерації дістали план $X_1 = (5; 0; 5; 14; 0)$, якому відповідає значення лінійної функції $F(X_1) = 200$. У четвертому рядку $f_2 - c_2 = -10 < 0$, тому план X_1 не є оптимальним і вектор A_2 слід ввести

в базис. Обчислимо $\theta_{02} = \min \left(\frac{10}{7}, \frac{14}{5}, 10 \right) = \frac{10}{7}$. Число $7/2$, яке стоїть

на перетині першого рядка і стовпця A_2 , є ведучим. Вектор A_3 вилучаємо з базису. Виконавши третю ітерацію симплексних перетворень, дістаємо нову таблицю. Оскільки в четвертому рядку всі оцінки плану

$X_2 = (30/7; 10/7; 0; 48/7; 0)$ невід'ємні, то опорний план X_2 — оптимальний, йому відповідає значення лінійної функції $F(X_2) = 1500/7$.

Порівнявши розв'язання даної задачі симплекс-методом з її графічним розв'язанням (див. приклад 1, § 4.3) побачимо, що опорний план X_0 відповідає вершині O многокутника розв'язків, опорний план X_1 — вершині D , а опорний план X_2 — вершині C . Тому симплексний метод можна інтерпретувати як переміщення через сусідні вершини многогранника розв'язків (пов'язаних із зменшенням або збільшенням значення лінійної функції) у напрямі вершини, в якій лінійна функція набуде найменшого або найбільшого значення.

Зазначимо, що кожне число, записане в симплексній таблиці, має не тільки математичний, а й певний економічний зміст. Так, у четвертому рядку першої ітерації симплексної табл. 4.13 є дві від'ємні оцінки: $f_1 - c_1 = -40$, $f_2 - c_2 = -30$. Число -40 означає, що від включення до плану виробництва одиниці продукції P_1 прибуток збільшиться на 40 у.г.о. Якщо включити до плану виробництва одиницю продукції P_2 , то прибуток збільшиться на 30 у.г.о. Тому з економічного погляду найдоцільніше включати до плану продукцію P_1 . Цей висновок цілком узгоджується і з формальним критерієм оптимальності опорного плану симплексного методу, оскільки найменша оцінка відповідає вектору A_1 . Число $\theta_{01} = 5$ означає, що максимальна кількість продукції, яку можна включити до плану виробництва, дорівнює п'яти одиницям. При цьому сировину третього виду буде використано повністю.

Після завершення другої ітерації в стовпці A_0 дістали опорний план $X_1 = (5; 0; 5; 14; 0)$. Згідно з цим планом виробляється 5 одиниць продукції P_1 і залишається невикористаною 5 одиниць сировини S_1 і 14 одиниць сировини S_2 . Прибуток від реалізації виробленої продукції становить 200 у.г.о. В результаті симплексних перетворень змінився не тільки стовпець A_0 , а й інші стовпці таблиці. Їх економічний зміст став ще складнішим. Розглянемо, наприклад, стовпець A_2 . Число $1/2$ в третьому рядку показує, на скільки одиниць треба зменшити виробництво продукції P_1 , якщо запланувати випуск одиниці продукції P_2 . Числа $7/2$ і 5 у першому і другому рядках стовпця A_2 означають, скільки треба витратити одиниць сировини S_1 і S_2 відповідно, якщо запланувати виготовлення одиниці продукції P_2 . Число -10 у четвертому рядку означає, що прибуток, який буде одержано від реалізації одиниці продукції P_2 , становитиме 10 у.г.о. Інакше кажучи, включення до плану виробництва одиниці продукції P_2 зумовлює зменшення випуску продукції P_1 на $1/2$ одиниці і додаткові затрати $7/2$ одиниць сировини S_1 і 5 одиниць сировини S_2 , а загальний прибуток від реалізації продукції відповідно до нового плану зросте на 10 у.г.о. Дещо інший економічний зміст мають числа, записані в стовпці A_5 . Число $1/4$ в третьому рядку цього стовпця означає, що збільшення обсягу сировини S_3 на 1 одиницю дасть змогу збільшити випуск продукції P_2 на $1/4$ одиниці. Одночасно треба буде додатково взяти $5/4$ одиниці сировини S_1 і $1/2$ одиниці сировини S_2 . Збільшення випуску продукції P_2 на $1/4$ одиниці забезпечить збільшен-

ня прибутку на 10 у.г.о. З економічного аналізу даних симплексної таблиці після другої ітерації випливає, що план X_1 не є оптимальним. Це видно і з четвертого рядка симплекс-таблиці, оскільки в стовпці A_2 — від'ємна оцінка плану.

Після третьої ітерації знайшли оптимальний план $X_2 = (30/7; 10/7; 0; 48/7; 0)$ випуску продукції, який включає виробництво 30/7 одиниць продукції P_1 і 10/7 одиниць продукції P_2 . При такому плані випуску продукції буде повністю використано сировину першого і третього видів і залишиться невикористаною 48/7 одиниць сировини другого виду. Прибуток від реалізації запланованої продукції становитиме 1500/7 у.г.о.

Щоб розв'язати задачу лінійного програмування на ЕОМ, використовують пакети стандартних програм з математичного забезпечення ЕОМ. Програми симплекс-методу порівняно складні. Це пов'язано з необхідністю програмно досліджувати систему обмежень на сумісність, автоматично будувати початковий опорний план та аналізувати побудований опорний план на оптимальність на кожному кроці симплексних перетворень. Подамо програму 4.1, яка розв'язує канонічну задачу лінійного програмування (4.1) — (4.3). Вона складається з головної програми і десяти підпрограм. Основні відомості про кожен з підпрограм наведено в коментарях до них.

Програма 4. 1

```

5 '          Розв'язання
10 '          задачі лінійного програмування
20 '          симплекс-методом
30 '
40 ' У програмі використовуються такі
50 ' змінні
60 ' A(m,n) — матриця коефіцієнтів при
70 ' змінних системи обмежень,
80 ' доповнена двома
90 ' рядками
110 ' (m-1)-й — коефіцієнти
120 ' лінійної функції і m-й —
130 ' сума стовпців коефіцієнтів
140 ' матриці A від i=1,...,m-2;
150 ' U(m,m) — матриця, обернена до базисної
160 ' NB(m) — масив номерів базисних компонент
170 ' вектора розв'язків;
180 ' X(m) — вектор-розв'язок;
190 ' K — номер вектора, що вводиться в базис;
200 ' L — номер вектора, що виводиться із базису;
210 ' XK(m) — коефіцієнти розкладу k-го
220 ' вектора за векторами базису;

```

```

230 ' TETA=min(X(i)/XK(i)), XK(i)>0, i=1,...,m-2;
240 ' FJCJ — оцінка плану j-го стовпця;
250 ' P — ознака: P=1 — задачу поставлено
260 ' на максимум, P=0 — задачу поставлено
270 ' на мінімум.
280 CLS
290 INPUT "Ввести кількість змінних"; N
300 PRINT "Ввести кількість рівнянь у"
310 PRINT "системі обмежень"
320 INPUT MM
330 M=MM+2: EPS=1E-10
340 DIM A(M,N), X(MM+N), U(M,M), XK(M), NB(M)
350 PRINT "Ввести по одному коефіцієнти при"
360 PRINT "змінних у системі обмежень по стовпцях"
370 FOR J=1 TO N
380   FOR I=1 TO MM
390     INPUT A(I,J)
400   NEXT I
410 NEXT J
420 PRINT "Ввести по одному вільні члени"
430 PRINT "системі обмежень"
440 FOR I=1 TO MM
450   INPUT X(I)
460 NEXT I
470 PRINT "Ввести по одному коефіцієнти при"
480 PRINT "змінних цільової функції"
490 FOR J=1 TO N
500   INPUT A(M-1,J)
510 NEXT J
520 PRINT "Ввести ознаку розв'язку задачі:"
530 PRINT "0 — якщо знаходите мінімум функції мети"
540 PRINT "1 — якщо знаходите максимум функції мети"
550 INPUT IP
560 IF IP=0 THEN 590
570 IF IP=1 THEN 590
580 PRINT "Ви помилилися при введенні ознаки": GOTO 520
590 X(M-1)=0:X(M)=0
600 FOR J=1 TO N
610   A(M,J)=0
620   FOR I=1 TO MM
630     A(M,J)=A(M,J)-A(I,J)
640   NEXT I
650 NEXT J
660 FOR I=1 TO MM

```

Продовження програми 4.1

```

670 X(M)=X(M)-X(I)
680 NEXT I
690 ' Кінець заповнення початкової симплексної таблиці
700 PRINT
710 PRINT:PRINT "Чекайте результатів."
720 PRINT "Комп'ютер обробляє введені дані".
730 IF IP=0 THEN 770
740 FOR J=1 TO N
750 A(M-1,J)=A(M-1,J)
760 NEXT J
770 GOSUB 1210: ' Заповнення матриці U як одиничної
780 FOR I=1 TO MM
790 NB(I)=N+1
800 NEXT I
810 NI=0:NE=1
820 GOSUB 1390: ' Вибір вектора, що вводиться в базис
830 ' за оцінками m-го рядка
840 GOTO 900
850 GOSUB 2070: ' Вибір вектора, що вводиться в базис
860 ' за оцінками (m-1)-го рядка
870 GOTO 900
880 GOSUB 2190: ' Вибір вектора, що вводиться в базис
890 ' у випадку виродження
900 IF FMIN<-EPS THEN 980
910 ON NE GOTO 920, 1050, 1050
920 IF X(M)<-EPS THEN 1090
930 NE=2
940 FOR I=1 TO M
950 IF NB(I)>N THEN NE=3
960 NEXT I
970 ON NE GOTO 850, 850, 880
980 GOSUB 1550
990 GOSUB 1650
1000 IF TETA>=100000000000! THEN 1110
1010 GOSUB 1780
1020 GOSUB 1860
1030 NB(L)=K: NI=NI+1
1040 ON NE GOTO 820, 850, 880
1050 IF IP=1 THEN 1070
1060 X(M-1)=X(M-1)
1070 PR$="Задача має оптимальний розв'язок": IP=1
1080 GOTO 1120
1090 IP=2:PR$="Задача не має розв'язку"
1100 GOTO 1120

```

Продовження програми 4.1

```

1110 IP=3: PR$="Лінійна функція не обмежена"
1120 PRINT : PRINT PR$
1130 IF IP <> THEN 1200
1140 PRINT "Оптимальний план має ненульові компоненти:"
1150 FOR I=1 TO MM
1160 PRINT NB(I):"x("NB(I)")="X(I)
1170 NEXT I
1180 PRINT "Значення цільової функції дорівнює";X(M-1)
1190 PRINT "Задача розв'язана за "NI;"ітерацій"
1200 END
1210 ' Підпрограма формування
1220 ' одиничної матриці
1230 FOR I=1 TO M
1240 FOR J=1 TO M
1250 U(I,J)=0
1260 IF I=J THEN U(I,J)=1
1270 NEXT J
1280 NEXT I
1290 RETURN
1300 ' Підпрограма обчислення оцінки
1310 ' FJCJ, як добутку m-го рядка
1320 ' матриці U(m,m) на j-й
1330 ' стовпець матриці A(m,n)
1340 FJCJ=0
1350 FOR I=1 TO M
1360 FJCJ=FJCJ+U(M,I)*A(I,J)
1370 NEXT I
1380 RETURN
1390 ' Підпрограма вибору найменшої
1400 ' оцінки стовпців матриці A.
1410 ' Результат запам'ятовується через
1420 ' змінну FMIN. Номер стовпця, для
1430 ' якого знайдена мінімальна
1440 ' оцінка, запам'ятовується через змінну K
1450 FMIN=0: M1=M-2
1460 FOR J=1 TO N
1470 FOR II=1 TO M1
1480 IF J=NB(II) THEN 1530
1490 NEXT II
1500 GOSUB 1300: ' Обчислення оцінки плану
1510 ' для j-го стовпця
1520 IF FJCJ<FMIN THEN FMIN=FJCJ:K=J
1530 NEXT J
1540 RETURN

```

```

1550 ' Підпрограма множення матриці
1560 ' U(m,m) на k-й стовпець матриці
1570 ' A(m,n). Результат — вектор XK(m)
1580 FOR I=1 TO M
1590   XK(I)=0
1600   FOR J=1 TO M
1610     XK(I)=XK(I)+U(I,J)*A(J,K)
1620   NEXT J
1630 NEXT I
1640 RETURN
1650 ' Підпрограма знаходження мінімального
1655 ' відношення x(i)/xk(i),
1660 ' xk(i)>eps, i=1,...,m-2. Мінімальне
1670 ' відношення позначається
1680 ' ТЕТА, а номер рядка, для якого
1690 ' досягається мінімум, запам'ятовується
1700 ' через змінну L.
1710 TETA=1000000000000!: M1=M-2
1720 FOR I=1 TO M1
1730   IF XK(I)<EPS THEN 1760
1740   R=X(I)/XK(I)
1750   IF R<=TETA THEN TETA=R:L=I
1760 NEXT I
1770 RETURN
1780 ' Підпрограма обчислення нового
1790 ' опорного плану й значення
1800 ' цільової функції
1810 FOR I=1 TO M
1820 IF I=L THEN X(I)=TETA: GOTO 1840
1830   X(I)=X(I)-TETA*XK(I)
1840 NEXT I
1850 RETURN
1860 ' Підпрограма обчислення нових
1870 ' елементів матриці U(m,m)
1880 M1=M-2
1890 FOR J=1 TO M1
1900   U(L,J)=U(L,J)/XK(L)
1910 NEXT J
1920 FOR I=1 TO M
1930   FOR J=1 TO M1
1940     IF I<>L THEN U(I,J)=U(I,J)-U(L,J)*XK(I)
1950   NEXT J
1960 NEXT I
1970 RETURN

```

```

1980 ' Підпрограма множення (m-1)-го
1990 ' рядка матриці U(m,m) на j-й
2000 ' стовпець матриці A(m,n).
2010 ' Результат множення — FJCJ.
2020 FJCJ=0
2030 FOR I=1 TO M
2040   FJCJ=FJCJ+U(M-1,I)*A(I,J)
2050 NEXT I
2060 RETURN
2070 ' Підпрограма вибору мінімальної
2080 ' оцінки серед оцінок, обчислених
2090 ' за підпрограмою 1980-2060
2100 FMIN=0: M1=M-2
2110 FOR J=1 TO N
2120   FOR II=1 TO M1
2130     IF J=NB(II) THEN 2170
2140   NEXT II
2150   GOSUB 1980
2160   IF FJCJ<FMIN THEN FMIN=FJCJ: K=J
2170 NEXT J
2180 RETURN
2190 ' Підпрограма вибору мінімального
2200 ' FJCJ у випадку виродження, тобто
2210 ' коли в базисі опорного плану
2220 ' присутні штучні змінні
2230 ' з нульовими значеннями.
2240 ' Для кожного небазисного стовпця
2250 ' обчислюються як FJCJ (за п/п
2260 ' 1300-1380), так і F1C1 (за п/п
2270 ' 1980-2060).
2280 ' Мінімальне FJCJ вибирається
2290 ' тільки для тих стовпців, які
2300 ' не входять до базису і для яких
2310 ' ABS(F1C1)<EPS і FJCJ<0.
2320 '   Якщо таке FJCJ знайдено, воно
2330 '   запам'ятовується через змінну FMIN,
2340 '   а номер відповідного стовпця
2350 '   через змінну K.
2360 FMIN=0: M1=M-2
2370 FOR J=1 TO N
2380   FOR II=1 TO M1
2390     IF I=NB(II) THEN 2450
2400   NEXT II
2410   GOSUB 1980: F1C1=FJCJ

```

```

2420 GOSUB 1300
2430 IF ABS(F1C1)>EPS THEN 2450
2440 IF F1C1<FMIN THEN FMIN=F1C1: K=J
2450 NEXT J
2460 RETURN
    
```

ЛАБОРАТОРНА РОБОТА № 3

Тема. Симплексний метод розв'язування задачі лінійного програмування.

Завдання. 1. Серед невід'ємних розв'язків системи обмежень

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

знайти розв'язок, який надасть найменшого (або найбільшого) значення лінійній функції $F = c_1x_1 + c_2x_2 + \dots + c_nx_n$.

Завдання виконати двома способами:

- а) перетворенням симплекс-таблиць за допомогою електронних таблиць;
- б) за допомогою програми 4.1 або програми, взятої з бібліотеки стандартних програм наявної ЕОМ.

2. Знайдені результати порівняти між собою. Значення коефіцієнтів і вільні члени системи подано в табл. 4. 14, а вид екстремуму і значення коефіцієнтів лінійної функції — в табл. 4. 15.

Таблиця 4. 14

Варіант	i	a _{i1}	a _{i2}	a _{i3}	a _{i4}	a _{i5}	b _i
1	2	3	4	5	6	7	8
1	1	0	-3,78	3,37	-4,88	3,81	0,43
	2	-4,35	3,71	4,48	0,36	-4,33	0
	3	4,32	-3,38	4,29	3,87	2,73	4,87
2	1	2,57	-12,61	0	0,2	2,83	3,18
	2	0	3,19	0	0,81	1,27	1,86
	3	0	-0,02	3,87	0	-18,25	1,32
3	1	5,87	-2,37	0	2,28	23,64	0,56
	2	0	12,03	-28,73	33,17	6,28	27,43
	3	6,73	24,17	1,89	-1,74	0	30,07
4	1	47,25	0	-0,2	-7,63	24,21	5,84
	2	28,17	0	19,23	0	19,47	4,67
	3	-1,87	5,73	0	0	38,74	3,73

1	2	3	4	5	6	7	8
5	1	16,23	-38,02	0	6,45	22,87	0,31
	2	6,02	-32,74	0	-0,78	27,22	0,14
	3	-28,12	21,75	-2,16	2,17	2,76	3,22
6	1	-17,28	0	19,28	25,73	0	6,38
	2	41,64	0	0,1	26,82	6,28	32,17
	3	38,22	0,52	0	-14,17	0,12	44,22
7	1	12,34	22,56	-33,74	0	1,41	30,37
	2	0	0,73	3,21	32,29	1,01	21,74
	3	-27,38	16,21	0,68	-27,61	1,43	0
8	1	0,1	-7,45	0	18,43	11,41	16,78
	2	0	19,97	47,64	0	17,29	19,24
	3	-2,74	2,02	0	0	44,47	2,23
9	1	0	0	38,73	-11,65	24,74	2,49
	2	34,04	0	10,08	0	17,59	33,41
	3	0	1,73	24,47	0,02	6,23	34,56
10	1	-2,67	0	1,17	15,83	0	6,78
	2	8,74	36,2	0	-5,87	0	31,24
	3	11,73	0	0	12,8	18,83	19,47
11	1	4,05	-2,82	0	10,85	10,16	21,28
	2	-2,37	0	16,27	1,18	10,78	3,94
	3	12,28	18,02	-11,37	12,34	0	32,73
12	1	24,37	-2,28	49,93	0	2,87	17,28
	2	-2,34	16,83	0	29,3	17,21	22,27
	3	0	6,84	2,87	-6,19	24,45	23,17
13	1	0	6,08	-5,54	16,22	33,17	0
	2	25,07	0	3,17	44,47	19,83	17,24
	3	-8,39	0	38,24	16,51	32,17	3,07
14	1	10,08	-2,25	1,65	24,01	26,21	0,1
	2	16,27	0	30,23	11,37	3,28	2,29
	3	0	17,38	2,91	-2,83	18,87	2,92

Продовження табл. 4.14

1	2	3	4	5	6	7	8
15	1	-16,02	28,24	0	32,29	0	10,06
	2	0,27	-30,74	0	1,07	3,87	10,53
	3	4,47	-32,29	27,01	-3,84	-3,26	21,66
16	1	2,34	2,21	0,73	-3,85	0	2,74
	2	0	0,97	11,03	17,87	0	16,24
	3	-1,37	3,29	0	1,12	3,56	2,68
17	1	0,53	1,24	-0,12	3,11	0	1,74
	2	0	-3,78	1,17	0	3,24	4,37
	3	-2,04	0	1,25	2,64	20,83	3,28
18	1	0,63	32,16	-16,28	0	22,57	-2,28
	2	-17,12	0	3,24	21,86	2,16	1,05
	3	0	2,79	-2,28	3,37	1,77	1,03
19	1	6,74	12,17	-2,26	0	21,51	10,54
	2	0	-4,13	7,28	1,55	10,37	21,32
	3	-6,26	6,57	0	10,58	3,42	12,26
20	1	8,23	14,11	22,14	0	0,93	20,02
	2	0	0,37	2,13	21,38	0,74	13,47
	3	-17,54	11,23	0,63	-16,87	0,98	0
21	1	31,24	0	0,1	-4,78	14,35	3,87
	2	18,43	0	13,2	0	12,87	3,12
	3	-1,25	3,92	0	0	37,24	2,51
22	1	11,32	-24,83	0	2,71	13,84	10,4
	2	4,02	-22,34	0	-0,47	18,47	0,18
	3	19,24	14,37	-1,13	1,48	1,73	2,21
23	1	4,68	10,46	2,84	-12,39	0	10,96
	2	0	3,74	42,37	70,24	0	64,83
	3	-4,87	12,35	0	4,67	14,23	10,69
24	1	-8,01	14,12	0	16,3	0	5,03
	2	0,54	-61,28	0	2,14	6,92	21,35
	3	9,08	-64,25	54,37	-6,48	-6,53	42,87

Продовження табл. 4.14

1	2	3	4	5	6	7	8
25	1	4,08	4,28	1,63	-6,93	0	5,27
	2	0	1,85	23,06	34,58	0	32,18
	3	-2,18	6,48	0	2,25	6,83	5,36
26	1	8,12	-19,23	0	3,22	11,65	0,66
	2	3,07	-16,36	0	0,35	13,52	0,07
	3	-14,05	10,67	-1,15	1,08	1,39	1,66
27	1	-8,56	0	9,58	12,36	0	3,25
	2	20,76	0	0,2	13,32	3,21	16,21
	3	19,11	0,25	0	-7,13	0,09	22,11
28	1	0,2	-14,35	0	39,25	22,83	33,14
	2	0	38,35	60,28	0	34,17	38,24
	3	-4,47	4,12	0	0	48,32	4,87
29	1	23,12	0	-0,1	-3,83	12,11	2,93
	2	14,08	0	9,56	0	9,74	2,38
	3	-0,93	2,87	0	0	19,87	1,93
30	1	2,03	-1,41	0	5,93	5,18	11,35
	2	-1,19	0	8,14	0,7	5,37	1,86
	3	6,14	9,37	-5,83	6,74	0	16,36

Таблиця 4.15

Варіант	Екстремум	c_1	c_2	c_3	c_4	c_5
1	максимум	1	2	-3	4	-3
2	мінімум	2	-1	1	4	-3
3	мінімум	3	1	-2	-1	4
4	максимум	4	-2	4	-1	7
5	мінімум	2	1	-3	5	-1
6	максимум	3	-2	5	-1	4
7	максимум	2	3	-1	4	-5
8	мінімум	-1	3	4	-1	2
9	мінімум	2	3	-3	1	-4

Варіант	Екстремум	c_1	c_2	c_3	c_4	c_5
10	максимум	3	-1	-2	3	4
11	мінімум	1	1	-3	2	-4
12	мінімум	3	-2	1	-5	1
13	максимум	-2	1	-1	3	5
14	максимум	-1	3	-2	1	4
15	мінімум	1	-3	4	-5	2
16	максимум	3	2	-1	4	-7
17	мінімум	2	-3	4	-3	2
18	максимум	1	-1	3	4	-2
19	мінімум	3	4	-5	-1	3
20	мінімум	5	7	-2	1	-1
21	максимум	7	-1	3	-4	-5
22	мінімум	3	4	-1	2	-4
23	максимум	2	3	-1	-2	1
24	максимум	1	5	6	-3	-1
25	мінімум	4	-3	2	1	-2
26	максимум	-2	5	-1	4	-3
27	мінімум	3	-1	5	2	-4
28	максимум	5	-3	2	-1	-4
29	мінімум	-2	4	6	-4	5
30	мінімум	1	-3	2	4	-6

Пояснення до виконання роботи. Виконання першого завдання треба починати із знаходження початкового опорного плану. Його можна знайти, застосувавши метод симплексних перетворень (§ 4. 5). Після цього перетворіть симплекс-таблицю згідно з алгоритмом, описаним у § 4. 6. Таблицю перетворюйте доти, поки не знайдете оптимальний план або не встановите, що поставлена задача лінійного програмування не має розв'язку.

Ознакою оптимальності опорного плану при знаходженні мінімуму лінійної функції є відсутність додатних оцінок плану, а при знаходженні максимуму — відсутність від'ємних оцінок плану.

Задача не має оптимального плану, якщо хоч для однієї додатної (від'ємної) оцінки плану всі елементи таблиці, що стоять над цією оцінкою, недодатні.

Для виконання другого завдання роботи треба підготувати дані умови задачі згідно з вимогами програми, за допомогою якої розв'язують поставлену задачу.

Знайдені результати машинного розв'язання порівняйте з результатами обчислень, виконаних за допомогою електронних таблиць.

Розділ 5. ІНТЕРПОЛЮВАННЯ ФУНКЦІЙ. ЧИСЕЛЬНЕ ДИФЕРЕНЦІВАННЯ

§5.1. ПОСТАНОВКА ЗАДАЧІ

Нехай на відрізку $[a; b]$ визначено певний клас функцій $\{P(x)\}$, наприклад клас алгебраїчних многочленів, а в точках x_0, x_1, \dots, x_n цього проміжку задано значення деякої функції $y = f(x)$: $y_0 = f(x_0), y_1 = f(x_1), \dots, y_n = f(x_n)$. Наближену заміну функції f на відрізку $[a; b]$ однією з функцій $P(x)$ цього класу так, щоб функція $P(x)$ в точках x_0, x_1, \dots, x_n набувала тих самих значень, що й функція f , тобто щоб $P(x_i) = y_i$ ($i = 0, 1, \dots, n$), називають *інтерполюванням*, або *інтерполяцією*. Точки x_0, x_1, \dots, x_n називають *вузлами інтерполювання*, функцію $P(x)$ — *інтерполюючою функцією*, а формулу $f(x) \approx P(x)$, за допомогою якої обчислюють значення функції f у проміжку $[a; b]$, — *інтерполяційною формулою*.

З геометричного погляду задача інтерполювання полягає в знаходженні кривої $y = P(x)$ певного класу, яка проходить через точки площини з координатами (x_i, y_i) ($i = 0, 1, \dots, n$) (рис. 5.1).

Якщо функція $P(x)$ належить класу алгебраїчних многочленів, то інтерполювання називається *параболічним*. Параболічне інтерполювання найзручніше, оскільки многочлени, які прості за формою і не мають особливих точок, можуть набувати довільних значень, їх легко обчислювати, диференціювати й інтегрувати.

У деяких випадках доцільніше використовувати інші класи інтерполюючих функцій. Якщо, наприклад, функція f періодична, то функцію $P(x)$ природно вибирати з класу тригонометричних многочленів, а якщо функція f перетворюється в нескінченність у заданих точках або поблизу них, то функцію $P(x)$ доцільно вибирати з класу раціональних функцій.

Розглядатимемо лише задачу параболічного інтерполювання, яку сформулюємо так: в $n+1$ різних точках x_0, x_1, \dots, x_n задано значення

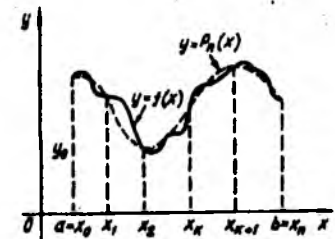


Рис. 5.1

$$\omega'_{n+1}(x) = \sum_{i=0}^n (x-x_0)(x-x_1)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n).$$

Поклавши тут $x = x_i$ ($i = 0, 1, \dots, n$), матимемо

$$\omega'_{n+1}(x_i) = (x_i - x_0)(x_i - x_1)\dots(x_i - x_{i-1})(x_i - x_{i+1})\dots(x_i - x_n). \quad (5.8)$$

Підставивши (5.7) і (5.8) в (5.5), знайдемо

$$L_n(x) = \omega_{n+1}(x) \sum_{i=0}^n \frac{y_i}{(x-x_i)\omega'_{n+1}(x_i)}. \quad (5.9)$$

Вирази $\frac{\omega_{n+1}(x)}{(x-x_i)\omega'_{n+1}(x_i)}$, ($i = 0, 1, \dots, n$), що є коефіцієнтами при y_i ($i = 0, 1, \dots, n$) у многочлені Лагранжа, називають *коефіцієнтами Лагранжа*.

Розглянемо два окремих випадки інтерполяційної формули Лагранжа (5.6).

1. Нехай $n=1$, тобто значення функції f задано в двох вузлах x_0 і x_1 . Позначимо ці значення y_0 і y_1 . Тоді з формули (5.6) дістанемо

$$f(x) \approx \frac{x-x_1}{x_0-x_1} y_0 + \frac{x-x_0}{x_1-x_0} y_1. \quad (5.10)$$

Формулу (5.10) називають *формулою лінійного інтерполювання*. При лінійному інтерполюванні дуга кривої $y = f(x)$ на відрізку $[x_0; x_1]$ замінюється відрізком прямої (5.10), що лежить між точками $(x_0; y_0)$ і $(x_1; y_1)$.

2. Нехай $n=2$. Функцію f задано в трьох вузлах x_i ($i = 0, 1, 2$) значеннями y_i ($i = 0, 1, 2$). У цьому разі формула (5.6) набирає вигляду

$$f(x) \approx \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} y_0 + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} y_1 + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} y_2. \quad (5.11)$$

Формулу (5.11) називають *формулою квадратичного інтерполювання*. При квадратичному інтерполюванні дуга кривої $y = f(x)$ на відрізку $[x_0; x_2]$ замінюється дугою параболи, що проходить через точки $(x_i; y_i)$ ($i = 0, 1, 2$).

Оцінка похибки інтерполяційної формули Лагранжа. Якщо функція f на відрізку $[a; b]$ є многочленом степеня, що менший або дорівнює n , то з єдиності інтерполяційного многочлена випливає, що інтерполяційний многочлен $L_n(x)$ тотожно дорівнює f , тобто $f(x) - L_n(x) \equiv 0$, $x \in [a; b]$.

Якщо f на відрізку $[a; b]$, який містить вузли інтерполяції x_i ($i = 0, 1, \dots, n$), не є многочленом степеня, що менший або дорівнює n , то різниця

$$R_n(f, x) = f(x) - L_n(x) \quad (5.12)$$

дорівнюватиме нулю лише у вузлах інтерполяції x_i ($i = 0, 1, \dots, n$), а в інших точках відрізка $[a; b]$ вона відмінна від тотожного нуля. Функцію $R_n(f, x)$, яка характеризує точність наближення функції f інтерполяційним многочленом $L_n(x)$, називають *залишковим членом* інтерполяційної формули Лагранжа (5.6), або *похибкою інтерполювання*. Якщо відомий аналітичний вираз функції f , то можна оцінити $R_n(f, x)$.

Справедлива така теорема.

Теорема. Якщо вузли інтерполювання x_i ($i = 0, 1, \dots, n$) різні і належать відрізку $[a; b]$, а функція f диференційовна $n+1$ раз на відрізку $[a; b]$, то для будь-якої точки $x \in [a; b]$ існує така точка $\xi \in (a; b)$, що для похибки інтерполювання справедлива рівність

$$R_n(f, x) = \frac{\omega_{n+1}(x)}{(n+1)!} f^{(n+1)}(\xi), \quad (5.13)$$

де $\omega_{n+1}(x) = (x-x_0)(x-x_1)\dots(x-x_n)$.

Д о в е д е н н я. Розглянемо допоміжну функцію

$$u(x) = f(x) - L_n(x) - k\omega_{n+1}(x), \quad (5.14)$$

де k — стала, яку доберемо пізніше.

Вузли інтерполювання x_i ($i = 0, 1, \dots, n$) є коренями функції u . Доберемо тепер сталу k так, щоб функція u на відрізку $[a; b]$ мала $(n+2)$ -й корінь у довільній, але фіксованій точці $\bar{x} \neq x_i$ ($i = 0, \dots, n$). Тоді $u(\bar{x}) = 0$, або $f(\bar{x}) - L_n(\bar{x}) - k\omega_{n+1}(\bar{x}) = 0$. Оскільки $\omega_{n+1}(\bar{x}) \neq 0$, то досить покласти

$$k = \frac{f(\bar{x}) - L_n(\bar{x})}{\omega_{n+1}(\bar{x})}. \quad (5.15)$$

При такому значенні k функція u на відрізку $[a; b]$ має $n+2$ коренів і дорівнює нулю на кінцях кожного з відрізків $[x_0; x_1]$, $[x_1; x_2]$, ..., $[x_i; \bar{x}]$, $[\bar{x}; x_{i+1}]$, ..., $[x_{n-1}; x_n]$.

За теоремою Ролля на кожному з цих відрізків є принаймні одна точка, в якій похідна $u'(x)$ дорівнює нулю. Тому на $(a; b)$ похідна $u'(x)$ дорівнює нулю не менш як $n+1$ раз, похідна другого порядку $u''(x)$ — не менш як n раз і т.д., нарешті похідна $u^{(n+1)}(x)$ — щонайменше один раз у точці, що лежить між найбільшим і найменшим з чисел x_0, x_1, \dots, x_n і \bar{x} .

Нехай ξ — корінь похідної $u^{(n+1)}(x)$, тобто $u^{(n+1)}(\xi) = 0$. Продиференціювавши по x функцію $u(x)$ (5.14) $n+1$ раз і врахувавши, що $L_n^{(n+1)}(x) = 0$, а $\omega_{n+1}^{(n+1)}(x) = (n+1)!$, дістанемо

$$u^{(n+1)}(x) = f^{(n+1)}(x) - k(n+1)!$$

Звідси, якщо $x = \xi$, маємо

$$f^{(n+1)}(\xi) - k(n+1)! = 0, \text{ або } k = \frac{f^{(n+1)}(\xi)}{(n+1)!}.$$

Порівнявши це значення k з (5.15), знайдемо

$$\frac{f(\bar{x}) - L_n(\bar{x})}{\omega_{n+1}(\bar{x})} = \frac{f^{(n+1)}(\xi)}{(n+1)!},$$

тобто

$$f(\bar{x}) - L_n(\bar{x}) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(\bar{x}).$$

Але точка \bar{x} — довільна, тому останню формулу можна записати ще й так:

$$R_n(f, x) = f(x) - L_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(x),$$

де точка ξ залежить від x і $\xi \in (a; b)$. Ця формула справедлива для всіх точок відрізка $[a; b]$, включаючи й вузли інтерполювання. У вузлах інтерполювання x_i ($i = 0, 1, \dots, n$) похибка інтерполювання $R(f, x_i) = 0$ ($i = 0, 1, \dots, n$). Теорему доведено.

Якщо $M_{n+1} = \max_{x \in [a; b]} |f^{(n+1)}(x)|$, то для абсолютної похибки інтерполяційної формули Лагранжа дістаємо таку оцінку:

$$|R_n(f, x)| = |f(x) - L_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |\omega_{n+1}(x)|. \quad (5.16)$$

З формули (5.13) дістаємо, зокрема, що залишковий член формули лінійного інтерполювання (5.10) дорівнює

$$R_1(f, x) = \frac{\omega_2(x)}{2!} f''(\xi),$$

де $\omega_2(x) = (x - x_0)(x - x_1)$, $\xi \in (x_0, x_1)$,

а залишковий член формули квадратичного інтерполювання (5.11)

$$R_2(f, x) = \frac{\omega_3(x)}{3!} f'''(\xi),$$

де $\omega_3(x) = (x - x_0)(x - x_1)(x - x_2)$, $\xi \in (x_0, x_2)$.

З формули (5.16) видно, що абсолютна похибка інтерполяційної формули Лагранжа пропорційна добутку двох множників M_{n+1} і $|\omega_{n+1}(x)|$, з яких M_{n+1} залежить лише від функції f , а величина другого,

$|\omega_{n+1}(x)|$, визначається виключно вибором вузлів інтерполювання. Зменшити величину абсолютної похибки інтерполяційної формули Лагранжа можна таким вибором вузлів інтерполювання, за якого множник $|\omega_{n+1}(x)|$ набуває найменшого максимального значення на відрізку $[a; b]$.

П. Л. Чебишов довів, що величина $\max_{[a; b]} |\omega_{n+1}(x)|$ має найменше

значення, якщо вузлами інтерполювання є числа $x_k = \frac{b+a}{2} + \frac{b-a}{2} \xi_k$, де $\xi_k = \cos \frac{(2k+1)\pi}{2(n+1)}$ ($k = 0, 1, \dots, n$) — нулі многочлена Чебишова $T_{n+1}(t) = \cos(n+1) \arccos t$. Вони дійсні, різні, належать $(-1; 1)$ і згущуються біля кінців інтервалу.

Доведено, що за такого вибору вузлів інтерполювання

$$|\omega_{n+1}(x)| \leq 2 \left(\frac{b-a}{4} \right)^{n+1}.$$

Навіть і тепер не можна гарантувати, що абсолютна похибка інтерполяційної формули Лагранжа буде як завгодно мала при досить великому n .

Приклад 1. Побудувати інтерполяційний многочлен Лагранжа третього степеня для функції f , заданої табл. 5.1, і знайти наближене значення функції в точці $x = 2$.

Розв'язання. Користуючись формулою (5.5), дістанемо

$$\begin{aligned} L_3(x) &= 10 \frac{(x-3)(x-4)(x-6)}{(1-3)(1-4)(1-6)} + 6 \frac{(x-1)(x-4)(x-6)}{(3-1)(3-4)(3-6)} + \\ &+ 8 \frac{(x-1)(x-3)(x-6)}{(4-1)(4-3)(4-6)} + 5 \frac{(x-1)(x-3)(x-4)}{(6-1)(6-3)(6-4)} = \\ &= \frac{1}{6} (-3x^3 + 32x^2 - 101x + 132), \end{aligned}$$

а $f(2) = L_3(2) = 5,66666666$.

Отже, щоб побудувати інтерполяційний многочлен Лагранжа, треба виконати значну обчислювальну роботу. Обсяг її дуже зростає тоді, коли треба підвищувати порядок многочлена: якщо до заданої системи вузлів інтерполювання x_i ($i = 0, 1, \dots, n$) додати ще хоч один вузол x_{n+1} , то для нової системи вузлів x_i ($i = 0, 1, \dots, n+1$) многочлен Лагранжа треба будувати заново.

Таблиця 5.1

i	0	1	2	3
x_i	1	3	4	6
y_i	10	6	8	5

Організація обчислень за інтерполяційною формулою Лагранжа. Якщо значення інтерполяційного многочлена Лагранжа треба обчислити лише в одній точці $x \neq x_i$ ($i = 0, 1, \dots, n$), де x_i ($i = 0, 1, \dots, n$) — вузли інтерполювання, то немає потреби явно будувати самий многочлен. Досить формулу (5.5) подати у вигляді

$$L_n(x) = \omega_{n+1}(x) \sum_{k=0}^n \frac{y_k}{P_k},$$

де $P_k = (x - x_k) \omega'_{n+1}(x_k)$ ($k = 0, 1, \dots, n$), і всі обчислення виконувати за обчислювальною схемою табл. 5.2.

За цією схемою обчислюють:

- 1) різниці $x_i - x_j$ ($i \neq j$) і $x - x_i$ ($i, j = 0, 1, \dots, n$);
- 2) добутки множників по рядках P_i ($i = 0, 1, \dots, n$);
- 3) частки $\frac{y_i}{P_i}$ ($i = 0, 1, \dots, n$);
- 4) суму $S = \sum_{i=0}^n \frac{y_i}{P_i}$;
- 5) добуток діагональних елементів (підкреслених у табл. 5.2 рискою), які дають значення многочлена $\omega_{n+1}(x)$ у точці x ;
- 6) значення многочлена Лагранжа в точці x .

Таблиця 5.2

i	x_i	Різниці $x_i - x_j$ ($i \neq j$)						P_i	y_i	$\frac{y_i}{P_i}$
		$x - x_0$	$x_0 - x_1$	$x_0 - x_2$...	$x_0 - x_{n-1}$	$x_0 - x_n$			
0	x_0	$x - x_0$	$x_0 - x_1$	$x_0 - x_2$...	$x_0 - x_{n-1}$	$x_0 - x_n$	P_0	y_0	y_0/P_0
1	x_1	$x_1 - x_0$	$x - x_1$	$x_1 - x_2$...	$x_1 - x_{n-1}$	$x_1 - x_n$	P_1	y_1	y_1/P_1
...
n	x_n	$x_n - x_0$	$x_n - x_1$	$x_n - x_2$...	$x_n - x_{n-1}$	$x - x_n$	P_n	y_n	y_n/P_n
$\omega_{n+1}(x) = (x - x_0)(x - x_1) \dots (x - x_n)$								$S = \sum_{i=0}^n \frac{y_i}{P_i}$		
$L_n(x) = \omega_{n+1}(x) \cdot S$										

Для кожного нового значення аргументу x або при підвищенні порядку многочлена (а це веде до залучення нових вузлів) всі обчислення виконують заново. А це значно збільшує обсяг обчислювальної роботи. Значно спростити і зменшити її можна тоді, коли скористатись інтер-

поляційною схемою Ейткіна. У цій схемі в процесі обчислення поступово залучаються все нові й нові вузли x_i доти, поки самі обчислення не покажуть, що необхідної точності досягнуто.

П р и к л а д 2. Значення функції e^x задано табл. 5.3.

Таблиця 5.3

i	0	1	2	3	4	5	6
x_i	0	0,5	1,0	1,3	1,5	1,7	1,9
y_i	1	1,6487	2,7183	3,6693	4,4817	5,4739	6,6859

За допомогою інтерполяційного многочлена Лагранжа 6-го степеня обчислити значення функції в точці $x = 1,1$, а також похибку такої апроксимації.

Р о з в ' я з а н н я. Оскільки треба обчислити значення інтерполяційного многочлена Лагранжа лише в одній точці, то немає потреби будувати многочлен в явному вигляді. Тому виконаємо обчислення за обчислювальною схемою табл. 5.2. В результаті дістанемо табл. 5.4.

Таблиця 5.4

i	x_j	Різниці $x_i - x_j$ ($i \neq j$)							P_i	y_i	$\frac{y_i}{P_i}$
		1,1	-0,5	-1,0	-1,3	-1,5	-1,7	-1,9			
0	0	1,1	-0,5	-1,0	-1,3	-1,5	-1,7	-1,9	3,464175	1	0,288669
1	0,5	0,5	0,6	-0,5	-0,8	-1,0	-1,2	-1,4	-0,2016	1,6487	-8,1780753
2	1,0	1,0	0,5	0,1	-0,3	-0,5	-0,7	-0,9	0,004725	2,7183	575,30158
3	1,3	1,3	0,8	0,3	-0,2	-0,2	-0,4	-0,6	0,0029952	3,6693	1225,06
4	1,5	1,5	1,0	0,5	0,2	-0,4	-0,2	-0,4	-0,0048	4,4817	-933,6875
5	1,7	1,7	1,2	0,7	0,4	0,2	-0,6	-0,2	0,0137088	5,4739	399,29826
6	1,9	1,9	1,4	0,9	0,6	0,4	0,2	-0,8	-0,0919296	6,6859	-72,728479
$\omega_7(1,1) = 0,0025344$									$S = 1185,3544$		
$e^{1,1} = L_6(1,1) = 3,0041625$											

Округливши результат, дістанемо $e^{1,1} = 3,0042$, що збігається з табличним значенням п'ятизначних математичних таблиць.

Для оцінки похибки інтерполяції скористаємось нерівністю (5.16), яка набирає вигляду

$$|R_6(1,1)| \leq \frac{M_7}{7!} |\omega_7(1,1)|.$$

Похідна сьомого порядку від функції e^x дорівнює e^x . На відрізку $[0; 1,9]$ функція e^x зростає, тому $1 \leq e^x \leq 6,6859$, $M_7 = 6,6859$, $|\omega_7(1,1)| = 0,0025344$,

$$|R_6(1,1)| \leq \frac{6,6859}{5040} \cdot 0,0025344 = 0,0000033 < 0,000005.$$

Отже, всі п'ять десяткових знаків наближеного числа правильні. Таку високу точність наближення маємо, по-перше через високий порядок інтерполяційного многочлена Лагранжа і, по-друге, тому, що значення $x = 1,1$ належить серединному відрізку $[1,0; 1,3]$ (а для серединних відрізків точність наближення значно вища, ніж для крайніх!) і, крім того, x лежить близько до вузла інтерполювання $x_2 = 1,0$.

Алгоритм інтерполювання за інтерполяційною формулою Лагранжа. Блок-схему алгоритму обчислення значення функції f в деякій точці x , за інтерполяційною формулою Лагранжа (5.6) подано на рис. 5.2, а його реалізацію на ЕОМ — програмою 5.1. Вхідними даними алгоритму є: степінь n многочлена Лагранжа, два масиви чисел x_i та y_i ($i = 0, 1, \dots, n$) і значення аргументу x , для якого треба обчислити $L_n(x)$. Внутрішній цикл алгоритму зі змінною j (параметр циклу) забезпечує обчислення кожного доданка суми (5.5), а зовнішній цикл зі змінною i (параметр цик-

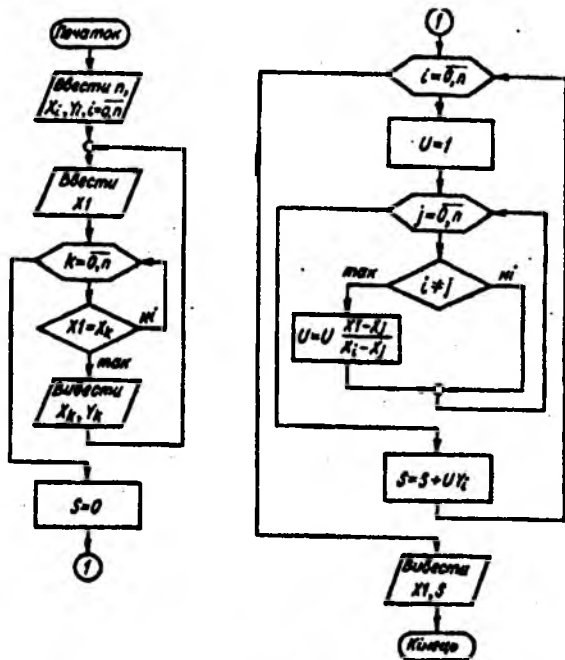


Рис. 5.2

лу) — обчислення всіх доданків суми (5.5), тобто значення функції в точці x . Алгоритмом передбачено також перевірку умови, чи не збігається значення аргументу x з яким-небудь з вузлів x_i ($i = 0, 1, \dots, n$) інтерполяційної формули (5.6): якщо збігається, то функції f надається відповідне значення y_i ($i = 0, 1, \dots, n$). Значення аргументу x і многочлена $L_n(x)$ в алгоритмі (і програмі 5.1) позначено відповідно $X1$ і S .

Програма 5.1

```

10 CLS
20 PRINT "Інтерполяція за Лагранжем"
30 PRINT "для N+1 вузлів"
40 PRINT
50 INPUT "Ввести N="; N: DIM X(N), Y(N)
60 PRINT "Ввести вузли і значення функції в них"
70 FOR I=0 TO N
80 PRINT "x("I")="; : INPUT X(I)
90 PRINT "y("I")="; : INPUT Y(I)
100 NEXT I
110 INPUT "Ввести x="; X1
120 FOR K=0 TO N
130 IF X1=X(K) THEN PRINT "y("X1")="; Y(K): GOTO 110
140 NEXT K
150 S=0
160 FOR I=0 TO N
170 U=1
180 FOR J=0 TO N
190 IF I=J THEN 210
200 U=U*(X1-X(J))/(X(I)-X(J))
210 NEXT J
220 S=S+U*Y(I)
230 NEXT I
240 PRINT "y("X")="; S : GOTO 110
250 END

```

§ 5.3. ПАРАБОЛІЧНЕ ІНТЕРПОЛЮВАННЯ ЗА СХЕМОЮ ЕЙТКІНА

Нехай функцію f , яка в точках x_i набуває значень $y_i = f(x_i)$ ($i = 0, 1, \dots, n$), задано таблично. Треба обчислити її значення в точці $x \in [x_0, x_n]$, яка не збігається з вузлами інтерполювання x_i ($i = 0, 1, \dots, n$). У цьому разі немає потреби знати загальний вираз многочлена Лагранжа явно. Треба лише вміти обчислити його значення в точці x . Це можна зробити, використавши так звану *інтерполяційну схему Ейткіна*, особливістю якої є однотипність обчислень.

Якщо функцію f задано в двох точках x_0 і x_1 (її значення відповідно дорівнюють y_0 і y_1), то її значення в точці $x \in (x_0, x_1)$ можна обчислити за формулою лінійного інтерполювання (5.10). Якщо значення функції f в точці x позначити через $P_{0,1}(x)$, то формулу лінійного інтерполювання (5.10) можна подати в рівносильному вигляді

$$P_{0,1}(x) = \frac{1}{x_1 - x_0} \begin{vmatrix} y_0 & x_0 - x \\ y_1 & x_1 - x \end{vmatrix},$$

права частина якої містить визначник другого порядку. Легко впевнитись, що

$$P_{0,1}(x_0) = \frac{1}{x_1 - x_0} \begin{vmatrix} y_0 & x_0 - x_0 \\ y_1 & x_1 - x_0 \end{vmatrix} = y_0, \quad P_{0,1}(x_1) = \frac{1}{x_1 - x_0} \begin{vmatrix} y_0 & x_0 - x_1 \\ y_1 & x_1 - x_1 \end{vmatrix} = y_1.$$

Нехай тепер функцію f задано в трьох точках x_0, x_1, x_2 (відповідні значення y_0, y_1, y_2) і треба обчислити її значення в точці $x \in (x_0, x_2)$, $x \neq x_1$. У цьому разі за схемою Ейткіна в точці x спочатку обчислюють значення двох лінійних многочленів

$$P_{0,1}(x_0) = \frac{1}{x_1 - x_0} \begin{vmatrix} y_0 & x_0 - x \\ y_1 & x_1 - x \end{vmatrix} \quad \text{і} \quad P_{1,2}(x) = \frac{1}{x_2 - x_1} \begin{vmatrix} y_1 & x_1 - x \\ y_2 & x_2 - x \end{vmatrix},$$

а потім значення квадратичного тричлена вигляду

$$P_{0,1,2}(x) = \frac{1}{x_2 - x_0} \begin{vmatrix} P_{0,1}(x) & x_0 - x \\ P_{1,2}(x) & x_2 - x \end{vmatrix}. \quad (5.17)$$

Безпосередня перевірка переконує, що $P_{0,1}(x_0) = y_0$, $P_{0,1}(x_1) = y_1$, $P_{1,2}(x_1) = y_1$, $P_{1,2}(x_2) = y_2$, $P_{0,1,2}(x_0) = y_0$, $P_{0,1,2}(x_1) = y_1$, $P_{0,1,2}(x_2) = y_2$.

Доведемо, що $P_{0,1,2}(x)$ збігається з інтерполяційним многочленом Лагранжа другого порядку. Справді, оскільки

$$P_{0,1}(x) = \frac{x - x_1}{x_0 - x_1} y_0 + \frac{x - x_0}{x_1 - x_0} y_1,$$

$$P_{1,2}(x) = \frac{x - x_2}{x_1 - x_2} y_1 + \frac{x - x_1}{x_2 - x_1} y_2,$$

то, розкривши визначник другого порядку формули (5.17), дістанемо

$$\begin{aligned} P_{0,1,2}(x) &= \frac{1}{x_2 - x_0} \left(\left(\frac{x - x_1}{x_0 - x_1} y_0 + \frac{x - x_0}{x_1 - x_0} y_1 \right) (x_2 - x) - \right. \\ &\quad \left. - \left(\frac{x - x_2}{x_1 - x_2} y_1 + \frac{x - x_1}{x_2 - x_1} y_2 \right) (x_0 - x) \right) = \\ &= \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} y_0 + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} y_1 + \\ &\quad + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} y_2. \end{aligned}$$

Ця схема узагальнюється на інтерполяційні многочлени вищих степенів. Якщо функцію f задано в чотирьох вузлах, то кубічне інтерполювання здійснюється за формулою

$$P_{0,1,2,3}(x) = \frac{1}{x_3 - x_0} \begin{vmatrix} P_{0,1,2}(x) & x_0 - x \\ P_{1,2,3}(x) & x_3 - x \end{vmatrix},$$

де $P_{0,1,2}(x)$ і $P_{1,2,3}(x)$ — значення квадратичних тричленів у точці $x \in (x_0, x_3)$, $x \neq x_1$, $x \neq x_2$. Тут значення многочлена $P_{0,1,2}(x)$ обчислюється за формулою (5.17), а значення $P_{1,2,3}(x)$ — за формулою:

$$P_{1,2,3}(x) = \frac{1}{x_3 - x_1} \begin{vmatrix} P_{1,2}(x) & x_1 - x \\ P_{2,3}(x) & x_3 - x \end{vmatrix}, \quad P_{2,3}(x) = \frac{1}{x_3 - x_2} \begin{vmatrix} y_2 & x_2 - x \\ y_3 & x_3 - x \end{vmatrix}.$$

Безпосередньою перевіркою впевнюємось, що $P_{1,2,3}(x_1) = y_1$, $P_{1,2,3}(x_2) = y_2$, $P_{1,2,3}(x_3) = y_3$, $P_{0,1,2,3}(x_i) = y_i$, ($i = 0, 1, 2, 3$) і $P_{0,1,2,3}(x)$ збігається з кубічним інтерполяційним многочленом Лагранжа.

Взагалі, якщо в $(n+1)$ -му вузлах інтерполювання x_i ($i = 0, 1, \dots, n$) функція f набуває значень y_i ($i = 0, 1, \dots, n$), то значення інтерполяційного многочлена степеня n в точці $x \in (x_0, x_n)$, що не збігається з вузлами інтерполювання, можна обчислити за формулою

$$P_{0,1,\dots,n}(x) = \frac{1}{x_n - x_0} \begin{vmatrix} P_{0,1,\dots,n-1}(x) & x_0 - x \\ P_{1,2,\dots,n}(x) & x_n - x \end{vmatrix},$$

де $P_{0,1,\dots,n-1}(x)$ і $P_{1,2,\dots,n}(x)$ — значення інтерполяційних многочленів $(n-1)$ -го степеня, обчислених у точці x на попередньому кроці обчислень. Легко впевнитись, що $P_{0,1,\dots,n}(x_i) = y_i$ ($i = 0, 1, \dots, n$) і $P_{0,1,\dots,n}(x)$ збігається з інтерполяційним многочленом Лагранжа n -го степеня.

Отже, щоб обчислити в точці x значення інтерполяційного многочлена n -го степеня за схемою Ейткіна, треба в цій точці обчислити значення n лінійних, $n-1$ квадратичних, $n-2$ кубічних многочленів і т. д., два многочлени $(n-1)$ -го степеня і, нарешті, один многочлен n -го

степеня. Всі ці многочлени виражають через визначник 2-го порядку, а це робить обчислення однотипними, циклічними.

П р и к л а д 3. За схемою Ейткіна обчислити з табличною точністю значення функції $\cos 0,28$, якщо функцію задано табл. 5.5.

Таблиця 5.5

i	0	1	2	3	4	5
x_i	0,2	0,3	0,4	0,5	0,6	0,7
y_i	0,98007	0,95534	0,92106	0,87758	0,82534	0,76484

Обчислення припинити, коли значення двох послідовних многочленів збігатимуться.

Р о з в ' я з а н н я. Виконавши обчислення з однією запасною цифрою, маємо

$$P_{0,1}(0,28) = \frac{1}{0,3 - 0,2} \begin{vmatrix} 0,98007 & 0,2 - 0,28 \\ 0,95534 & 0,3 - 0,28 \end{vmatrix} = \frac{1}{0,1} \begin{vmatrix} 0,98007 & -0,08 \\ 0,95534 & 0,02 \end{vmatrix} = 0,960286,$$

$$P_{1,2}(0,28) = \frac{1}{0,1} \begin{vmatrix} 0,95534 & 0,3 - 0,28 \\ 0,92106 & 0,4 - 0,28 \end{vmatrix} = \frac{1}{0,1} \begin{vmatrix} 0,95534 & 0,02 \\ 0,92106 & 0,12 \end{vmatrix} = 0,962196,$$

$$P_{0,1,2}(0,28) = \frac{1}{0,4 - 0,2} \begin{vmatrix} 0,960286 & -0,08 \\ 0,962196 & 0,12 \end{vmatrix} = 0,961050.$$

Оскільки $P_{0,1}(0,28) \neq P_{0,1,2}(0,28)$, то залуцаємо вузол $x_3 = 0,5$ і обчислюємо додатково

$$P_{2,3}(0,28) = \frac{1}{0,1} \begin{vmatrix} 0,92106 & 0,12 \\ 0,87758 & 0,22 \end{vmatrix} = 0,973236,$$

$$P_{1,2,3}(0,28) = \frac{1}{0,2} \begin{vmatrix} 0,962196 & 0,02 \\ 0,973236 & 0,22 \end{vmatrix} = 0,961092,$$

$$P_{0,1,2,3}(0,28) = \frac{1}{0,3} \begin{vmatrix} 0,961050 & -0,08 \\ 0,961092 & 0,22 \end{vmatrix} = 0,961061.$$

Але $P_{0,1,2,3}(0,28) \neq P_{0,1,2}(0,28)$, тому залуцаємо наступний вузол $x_4 = 0,6$ і обчислюємо додатково

$$P_{3,4}(0,28) = \frac{1}{0,1} \begin{vmatrix} 0,87758 & 0,22 \\ 0,82534 & 0,32 \end{vmatrix} = 0,992508,$$

$$P_{2,3,4}(0,28) = \frac{1}{0,2} \begin{vmatrix} 0,973236 & 0,12 \\ 0,992508 & 0,32 \end{vmatrix} = 0,961673,$$

$$P_{1,2,3,4}(0,28) = \frac{1}{0,3} \begin{vmatrix} 0,961092 & 0,02 \\ 0,961673 & 0,32 \end{vmatrix} = 0,961053,$$

$$P_{0,1,2,3,4}(0,28) = \frac{1}{0,4} \begin{vmatrix} 0,961061 & -0,08 \\ 0,961053 & 0,32 \end{vmatrix} = 0,961059.$$

Округливши значення $P_{0,1,2,3}(0,28)$ і $P_{0,1,2,3,4}(0,28)$ до п'ятого десяткового розряду, дістанемо: $P_{0,1,2,3}(0,28) = P_{0,1,2,3,4}(0,28) = 0,96106$, що відповідає табличному значенню.

Значення інтерполяційного многочлена в точці x доцільно обчислювати за схемою табл. 5.6.

Таблиця 5.6

i	x_i	y_i	$x_i - x$	$P_{i,i+1,i+2,\dots,i+j}$					
				$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$
0	x_0	y_0	$x_0 - x$	$P_{0,1}(x)$	$P_{0,1,2}(x)$	$P_{0,1,2,3}(x)$	$P_{0,1,2,3,4}(x)$	$P_{0,1,2,3,4,5}(x)$	$P_{0,1,2,3,4,5,6}(x)$
1	x_1	y_1	$x_1 - x$	$P_{1,2}(x)$	$P_{1,2,3}(x)$	$P_{1,2,3,4}(x)$	$P_{1,2,3,4,5}(x)$	$P_{1,2,3,4,5,6}(x)$	
2	x_2	y_2	$x_2 - x$	$P_{2,3}(x)$	$P_{2,3,4}(x)$	$P_{2,3,4,5}(x)$	$P_{2,3,4,5,6}(x)$		
3	x_3	y_3	$x_3 - x$	$P_{3,4}(x)$	$P_{3,4,5}(x)$	$P_{3,4,5,6}(x)$			
4	x_4	y_4	$x_4 - x$	$P_{4,5}(x)$	$P_{4,5,6}(x)$				
5	x_5	y_5	$x_5 - x$	$P_{5,6}(x)$					
6	x_6	y_6	$x_6 - x$						

В обчисленнях за цією схемою нові вузли x_i (що відповідає переходу до інтерполяційних многочленів вищих степенів) залуцають доти, поки самі обчислення не покажуть, що необхідної точності вже досягнуто.

Обчислення значення функції $\cos 0,28$ за схемою Ейткіна компактно подано в табл. 5.7.

Таблиця 5.7

i	x_i	y_i	$x_i - x$	$P_{i,i+1,\dots,i+j}$			
				$j=1$	$j=2$	$j=3$	$j=4$
0	0,2	0,98007	-0,08	0,960286	0,961050	0,961061	0,961059
1	0,3	0,95534	0,02	0,962196	0,961092	0,961053	
2	0,4	0,92106	0,12	0,973236	0,961673		
3	0,5	0,87758	0,22	0,992508			
4	0,6	0,82534	0,32				
5	0,7	0,76484	0,42				
6							

Інтерполяційну схему Ейткіна застосовують і тоді, коли вузли інтерполювання не є рівновіддаленими, а також при оберненому інтерполюванні та екстраполюванні.

Алгоритм і програма обчислення за схемою Ейткіна. Обчислення значення інтерполяційного многочлена за схемою Ейткіна має циклічний характер. Блок-схему відповідного алгоритму подано на рис. 5. 3, а його реалізацію на ЕОМ (мовою Бейсик) програмою 5. 2.

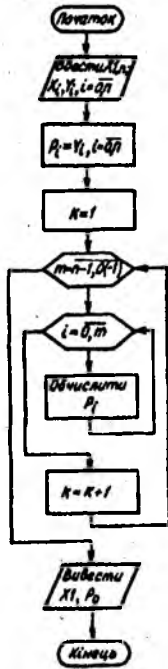


Рис. 5. 3

Програма 5. 2

```

10 PRINT "---- Інтерполяція за схемою Ейткіна -- "
20 PRINT "----- для N+1 вузлів -----"
30 INPUT "Ввести N="; N
40 DIM X (N), Y (N), P (N)
50 PRINT : PRINT "Ввести X (I) і Y (I)"
60 FOR I=0 TO N
70   PRINT "x ("I") = "; INPUT X (I)
80   PRINT "y ("I") = "; INPUT Y (I)
90 NEXT I
100 INPUT "Ввести аргумент X"; X1
110 FOR I=0 TO N
120   P(I) = Y(I)
130 NEXT I
140 K=1
150 FOR M=N-1 TO 0 STEP - 1
160   FOR I=0 TO M
170     P(I)=(P(I)*(X(I+K) -X1) -
              P(I+1)*(X(I)-(X1)))/(X(I+K)-X(I))
180   NEXT I
190   K=K+1
200 NEXT M
210 PRINT "Y("X1") = "; P(0) : GOTO 100
220 END
  
```

Для програми вхідними даними є: степінь многочлена n , масиви x_i та y_i ($i = 0, 1, \dots, n$) і значення аргументу x , який у програмі позначено $X1$. Внутрішній цикл зі змінною i (параметр циклу) обчислює значення визначника другого порядку в точці x за формулою

$$P_i(x) = (P_i(x)(x_{i+k} - x) - P_{i+1}(x)(x_i - x)) / (x_{i+k} - x_i), \quad (5.18)$$

де x_i, x_{i+k} — вузли інтерполювання, $P_i(x)$ — елементи масиву P . Спочатку до масиву P надсилаються елементи масиву Y . При першому виконанні внутрішнього циклу ($k=1$) за формулою (5. 18) використо-

вуються елементи масиву Y , але значення лінійних многочленів $P_{i,i+1}(x)$ ($i = 0, 1, \dots, n-1$) присвоюються відповідним елементам масиву P . При другому виконанні внутрішнього циклу ($k = 2$) обчислюються значення квадратних многочленів $P_{i,i+1,i+2}(x)$ ($i = 0, 1, \dots, n-2$) і їх значення присвоюються відповідним елементам масиву P і т. д., як це показано в табл. 5. 8.

Таблиця 5.8

Масиви		Значення многочленів k -го степеня після k -го виконання внутрішнього циклу				
$Y(i)$	$P(i)$	$k=1$	$k=2$...	$k=n-1$	$k=n$
y_0	$P(0)$	$P_{0,1}$	$P_{0,1,2}$...	$P_{0,1,\dots,n-1}$	$P_{0,1,2,\dots,n}$
y_1	$P(1)$	$P_{1,2}$	$P_{1,2,3}$		$P_{1,2,\dots,n}$	
y_2	$P(2)$	$P_{2,3}$	$P_{2,3,4}$			
y_3	$P(3)$	$P_{3,4}$	$P_{3,4,5}$			
\vdots	\vdots	\vdots	\vdots			
y_{n-2}	$P(N-2)$	$P_{n-2,n-1}$	$P_{n-2,n-1,n}$			
y_{n-1}	$P(N-1)$	$P_{n-1,n}$				
y_n	$P(N)$					

Отже, при кожному виконанні внутрішнього циклу дістаємо новий масив P , довжина якого на одиницю менша за довжину масиву, одержаного при попередньому виконанні цього циклу. В останньому ($k=n$) виконанні внутрішнього циклу дістанемо масив P лише з одним елементом $P(0) = P_{0,1,\dots,n}(x)$. Це і є шукане значення інтерполяційного многочлена Лагранжа n -го степеня в точці x .

Зовнішній цикл алгоритму зі змінною M (параметр циклу) слідкує за степенями інтерполяційних многочленів і виконується n раз. Після друкування результату програма переходить до виконання оператора з 100-го рядка. Ввівши нове значення аргументу, можна обчислити за програмою 5. 2 відповідне значення інтерполяційного многочлена.

§ 5.4. СКІНЧЕННІ РІЗНИЦІ ТА ЇХ ВЛАСТИВОСТІ

Нехай y_i ($i = 0, 1, \dots, n$) — значення функції f , обчислені для рівновіддалених значень аргументу x_i ($i = 0, 1, \dots, n$), $x_{i+1} - x_i = h$ ($i = 0, 1, \dots, n-1$), де h — крок таблиці.

Скінченними, або табличними, різницями першого порядку, або першими різницями, називають числа, які дорівнюють приростам значень функцій:

$$y_1 - y_0 = f(x_0 + h) - f(x_0) = \Delta y_0 = \Delta f(x_0),$$

$$y_2 - y_1 = f(x_0 + 2h) - f(x_0 + h) = \Delta y_1 = \Delta f(x_0 + h),$$

$$y_n - y_{n-1} = f(x_0 + nh) - f(x_0 + (n-1)h) = \Delta y_{n-1} = \Delta f(x_0 + (n-1)h).$$

Прирости різниць першого порядку називають *різницями другого порядку*, або *другими різницями*:

$$\Delta^2 y_0 = \Delta y_1 - \Delta y_0, \Delta^2 y_1 = \Delta y_2 - \Delta y_1, \dots, \Delta^2 y_{n-2} = \Delta y_{n-1} - \Delta y_{n-2}.$$

Взагалі, різниці будь-якого порядку k , або k -і різниці утворюються з різниць $(k-1)$ -го порядку за формулами

$$\Delta^k y_0 = \Delta^{k-1} y_1 - \Delta^{k-1} y_0,$$

$$\Delta^{k-1} y_1 = \Delta^{k-1} y_2 - \Delta^{k-1} y_1,$$

$$\Delta^k y_m = \Delta^{k-1} y_{m+1} - \Delta^{k-1} y_m.$$

Вважають, що за означенням $\Delta^0 y_k = y_k$.

Послідовні різниці зручно записувати у вигляді таблиць. Використовують дві форми таблиць: діагональну (табл. 5. 9) і горизонтальну (табл. 5. 10).

Таблиця 5. 9

x_i	y_i	Δy_i	$\Delta^2 y_i$	$\Delta^3 y_i$	$\Delta^4 y_i$	$\Delta^5 y_i$
x_0	y_0					
		Δy_0				
x_1	y_1		$\Delta^2 y_0$			
		Δy_1		$\Delta^3 y_0$		
x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$	
		Δy_2		$\Delta^3 y_1$		$\Delta^5 y_0$
x_3	y_3		$\Delta^2 y_2$		$\Delta^4 y_1$	
		Δy_3		$\Delta^3 y_2$		
x_4	y_4		$\Delta^2 y_3$			
		Δy_4				
x_5	y_5					

x_i	y_i	Δy_i	$\Delta^2 y_i$	$\Delta^3 y_i$	$\Delta^4 y_i$	$\Delta^5 y_i$
x_0	y_0	Δy_0	$\Delta^2 y_0$	$\Delta^3 y_0$	$\Delta^4 y_0$	$\Delta^5 y_0$
x_1	y_1	Δy_1	$\Delta^2 y_1$	$\Delta^3 y_1$	$\Delta^4 y_1$	
x_2	y_2	Δy_2	$\Delta^2 y_2$	$\Delta^3 y_2$		
x_3	y_3	Δy_3	$\Delta^2 y_3$			
x_4	y_4	Δy_4				
x_5	y_5					
Σ		$\sum_{i=0}^4 \Delta y_i$	$\sum_{i=0}^3 \Delta^2 y_i$	$\sum_{i=0}^2 \Delta^3 y_i$	$\sum_{i=0}^1 \Delta^4 y_i$	
S	$y_5 - y_0$	$\Delta y_4 - \Delta y_0$	$\Delta^2 y_3 - \Delta^2 y_0$	$\Delta^3 y_2 - \Delta^3 y_0$	$\Delta^4 y_1 - \Delta^4 y_0$	

Діагональну таблицю різниць дістають тоді, коли табличні різниці в кожному стовпці записують між відповідними значеннями зменшувального і від'ємника. В горизонтальній таблиці табличні різниці записують в одному рядку з від'ємником.

Як у діагональній, так і в горизонтальній таблицях скінченні різниці всіх порядків умовимось записувати в одиницях нижчого розряду значень функцій у вузлах інтерполювання. Далі дотримуватимемося цього правила.

Щоб проконтролювати правильність обчислення табличних різниць, таблицю доповнюють двома додатковими рядками Σ і S . У рядок Σ заносять суми різниць кожного із стовпців, а в рядок S — різниці між крайніми (останньою і першою) різницями кожного стовпця. Легко впевнитись, що сума різниць будь-якого стовпця таблиці дорівнює різниці між крайніми різницями попереднього стовпця, тобто

$$\sum_{i=0}^{n-1} \Delta y_i = y_n - y_0; \sum_{i=0}^{n-2} \Delta^2 y_i = \Delta y_{n-1} - \Delta y_0; \sum_{i=0}^{n-3} \Delta^3 y_i = \Delta^2 y_{n-2} - \Delta^2 y_0 \text{ і т. д.}$$

Приклад 4. Побудувати діагональну таблицю різниць для функції $y = \frac{1}{x}$ на відрізку $[1; 1,7]$ з кроком $h = 0,1$.

Розв'язання. У табл. 5. 11 подано відповідні значення різниць.

Таблиця 5.11

x_i	y_i	Δy_i	$\Delta^2 y_i$	$\Delta^3 y_i$	$\Delta^4 y_i$	$\Delta^5 y_i$	$\Delta^6 y_i$	$\Delta^7 y_i$
1	1,00000							
		-9091						
1,1	0,90909		1515					
		-7576		-349				
1,2	0,83333		1166		99			
		-6410		-250		-33		
1,3	0,76923		916		66		14	
		-5496		-184		-19		-9
1,4	0,71429		732		47		5	
		-4762		-137		-14		
1,5	0,66667		595		33			
		-4167		-104				
1,6	0,62500		491					
		-3676						
1,7	0,58824							
Σ		-41176	5415	-1024	245	-66	19	-9
5	-0,41176	5415	-1024	245	-66	19	-9	

Зв'язок між скінченними різницями і значеннями функцій. Скінченні різниці будь-якого порядку можна виразити через значення функції у вузлах інтерполювання. За означенням маємо

$$\Delta y_k = y_{k+1} - y_k.$$

Для різниць другого порядку

$$\Delta^2 y_k = \Delta y_{k+1} - \Delta y_k = (y_{k+2} - y_{k+1}) - (y_{k+1} - y_k) = y_{k+2} - 2y_{k+1} + y_k.$$

Аналогічно для різниць третього порядку

$$\begin{aligned} \Delta^3 y_k &= \Delta^2 y_{k+1} - \Delta^2 y_k = (y_{k+3} - 2y_{k+2} + y_{k+1}) - (y_{k+2} - 2y_{k+1} + y_k) = \\ &= y_{k+3} - 3y_{k+2} + 3y_{k+1} - y_k. \end{aligned}$$

Методом математичної індукції можна довести, що

$$\Delta^m y_k = y_{k+m} - m y_{k+m-1} + \frac{m(m-1)}{2!} y_{k+m-2} - \dots \quad (5.19)$$

$$\dots + (-1)^m y_k = \sum_{i=0}^m (-1)^i C_m^i y_{k+i}.$$

Справді, для $m = 1, 2, 3$ формулу доведено. Нехай тепер формула справедлива для $m \leq l$, тобто

$$\Delta^l y_k = y_{k+l} - C_l^1 y_{k+l-1} + C_l^2 y_{k+l-2} - \dots + (-1)^j C_l^j y_{k+l-j} + \dots + (-1)^l y_k.$$

і доведемо її справедливість для $m = l + 1$. Маємо

$$\begin{aligned} \Delta^{l+1} y_k &= \Delta^l y_{k+1} - \Delta^l y_k = (y_{k+l+1} - C_l^1 y_{k+l} + C_l^2 y_{k+l-1} - \dots \\ &\dots + (-1)^j C_l^j y_{k+l-j} + \dots + (-1)^l y_{k+1}) - (y_{k+l} - \\ &- C_l^1 y_{k+l-1} + C_l^2 y_{k+l-2} - \dots + (-1)^j C_l^j y_{k+l-j} + \dots + (-1)^l y_k) = \\ &= y_{k+l+1} - (C_l^0 + C_l^1) y_{k+l} + (C_l^1 + C_l^2) y_{k+l-1} - \dots + \\ &+ (-1)^j (C_l^{j-1} + C_l^j) y_{k+l-j} + \dots + (-1)^{l+1} y_k. \end{aligned}$$

Але

$$C_l^j + C_l^{j+1} = \frac{l!}{j!(l-j)!} + \frac{l!}{(j+1)!(l-j-1)!} = \frac{(l+1)!}{(j+1)!(l-j)!} = C_{l+1}^{j+1}.$$

Тому

$$\begin{aligned} \Delta^{l+1} y_k &= y_{k+l+1} - C_{l+1}^1 y_{k+l} + C_{l+1}^2 y_{k+l-1} - \dots + \\ &+ (-1)^j C_{l+1}^j y_{k+l-j} + \dots + (-1)^{l+1} y_k, \end{aligned}$$

що й треба було довести.

Формулу (5.19) легко запам'ятати: вона нагадує собою розклад виразу $(y_k - 1)^m$ за формулою бінома Ньютона, якщо в такому розкладі замість y_k^m взяти y_{k+m} , замість $y_k^{m-1} - y_{k+m-1}$ і т.д., і, нарешті, в останньому доданку замість $1 = y_k^0$ записати $y_{k+0} = y_k$.

Часто користуються формулами, які виражають значення функції в певному вузлі інтерполювання через скінченні різниці. Так, з формули $\Delta y_k = y_{k+1} - y_k$ випливає, що

$$y_{k+1} = y_k + \Delta y_k. \quad (5.20)$$

Аналогічно

$$\begin{aligned} y_{k+2} &= y_{k+1} + \Delta y_{k+1} = y_k + \Delta y_k + (\Delta y_{k+1} - \\ &- \Delta y_k) + \Delta y_k = y_k + 2\Delta y_k + \Delta^2 y_k. \end{aligned} \quad (5.21)$$

Для будь-якого l ($l = 0, 1, \dots, n$) справедлива формула

$$y_{k+1} = y_k + l\Delta y_k + \frac{l(l-1)}{2!} \Delta^2 y_k + \dots + \Delta^l y_k. \quad (5.22)$$

За цією формулою можна обчислити значення y_{k+1} для будь-якого l через значення y_k і скінченні різниці до l -го порядку включно. Формулу (5.22) символічно записують у вигляді

$$y_{k+1} = (1 + \Delta)^l y_k, \quad (5.23)$$

де вираз $(1 + \Delta)^l$ (l — ціле) розкривається за формулою бінома Ньютона. В цьому розкладі "добутки" вигляду $\Delta^l y_k$ означають скінченні різниці відповідних порядків.

Зв'язок між похідними функції і скінченними різницями. Нехай функція f має на відрізку $[a; b]$ неперервні похідні до порядку n включно. З означення похідної випливає, що

$$f'(x) = \lim_{h \rightarrow 0} \frac{\Delta f(x)}{h},$$

де $\Delta f(x) = f(x+h) - f(x)$, $x, x+h \in [a; b]$. Для малих значень h звідси випливає наближена формула

$$f'(x) \approx \frac{\Delta f(x)}{h}. \quad (5.24)$$

Знайдемо тепер

$$\lim_{h \rightarrow 0} \frac{\Delta^2 f(x)}{h^2} = \lim_{h \rightarrow 0} \frac{f(x+2h) - 2f(x+h) + f(x)}{h^2},$$

де $x, x+h, x+2h \in [a; b]$. При сталому значенні x і $h \rightarrow 0$ чисельник і знаменник цього дробу прямують до нуля. Для розкриття невизначеності типу $\frac{0}{0}$ застосуємо двічі правило Лопітала. Дістанемо

$$\lim_{h \rightarrow 0} \frac{\Delta^2 f(x)}{h^2} = \lim_{h \rightarrow 0} \frac{2f'(x+2h) - 2f'(x+h)}{2h} = \lim_{h \rightarrow 0} \frac{2f''(x+2h) - f''(x+h)}{1} = f''(x).$$

Звідси для досить малих h маємо наближену формулу

$$f''(x) \approx \Delta^2 f(x) / h^2. \quad (5.25)$$

Аналогічними міркуваннями для будь-якого натурального n можна дістати наближену формулу

$$f^{(n)}(x) \approx \Delta^n f(x) / h^n. \quad (5.26)$$

За формулами (5.24) — (5.26) можна обчислити наближене значення похідних, але їх точність (для $n > 1$) досить низька. Тому в обчислювальній практиці використовують точніші формули (§ 5.9).

З формул (5.24) — (5.26) випливає, що із зменшенням кроку таблиці h у λ раз перші різниці зменшуються приблизно в λ раз, другі — в λ^2 , треті — в λ^3 раз і т.д. Якщо, крім того, із зростанням порядку похідної їх модулі зростають повільно (не надто швидко), а крок таблиці досить малий, то модулі різниць із зростанням їх порядку зменшуватимуться. Тому при деякому l різниці $\Delta^l y$ стануть меншими за половину одиниці нижчого розряду табличних значень функцій, і тому з прийнятою точністю слід вважати, що вони дорівнюють нулю, а різниці $(l-1)$ -го порядку — сталі.

Але значення функції в таблиці подають здебільшого наближено, тому при обчисленні скінченних різниць похибка зростає. Якщо, наприклад, похибка округлення табличних значень функції дорівнює половині одиниці нижчого розряду, то похибка перших різниць дорівнюватиме вже одній одиниці нижчого розряду (при виконанні операції алгебраїчного додавання похибки додаються), похибка других різниць — 2 одиницям нижчого розряду, третіх — 4 одиницям і т.д., а похибки скінченних різниць порядку l дорівнюють 2^{l-1} одиниць нижчого розряду табличних значень функції.

З цього випливає, що, коли, наприклад, четверті різниці точних значень функції відрізняються одна від одної менше, ніж на половину одиниці нижчого розряду, то ці самі різниці, але обчислені для наближеного значення функції (абсолютна похибка яких дорівнює половині одиниці нижчого розряду), можуть відрізнитися одна від одної вже на $2^4 = 16$ одиниць нижчого розряду табличних значень функції.

Тому є підстави дати таке означення практично сталих різниць: якщо на деякій частині таблиці всі скінченні різниці l -го порядку відрізняються одна від одної не більш як на 2^l одиниць нижчого розряду табличних значень функції, то ці різниці називаються *практично сталими*.

Тоді різниці $(l+1)$ -го порядку не треба обчислювати, бо вони складатимуться лише з сумнівних цифр, а тому в межах даної точності вважають, що вони дорівнюють нулю і не беруть їх до уваги. Так, у табл. 5.11 різниці 5-го порядку відрізняються одна від одної менш як на $2^5 = 32$ одиниці розряду 10^{-5} , тому їх треба вважати практично сталими, а різниці 6-го порядку такими, що дорівнюють нулю. Тому на цій частині таблиці, тобто на відрізку $[1; 1,7]$, функція поводить себе приблизно як многочлен 5-го степеня.

Властивості скінченних різниць:

1. Скінченні різниці сталої C дорівнюють нулю, тобто $\Delta^k C = 0$.
2. Сталий множник C можна виносити за знак скінченної різниці

$$\Delta^k (Cf(x)) = C\Delta^k f(x).$$

3. Скінченні різниці алгебраїчної суми функцій дорівнюють алгебраїчній сумі скінченних різниць цих функцій, тобто

$$\Delta^k(f(x) \pm g(x)) = \Delta^k f(x) \pm \Delta^k g(x).$$

Ці властивості випливають з означення скінченних різниць.

4. Скінченна різниця функції $f(x)=x^n$ (n — натуральне)

$$\Delta(x^n) = nhx^{n-1} + \frac{n(n-1)}{2!} h^2 x^{n-2} + \dots + nh^{n-1}x + h^n. \quad (5.27)$$

Справді,

$$\Delta(x^n) = (x+h)^n - x^n.$$

Розклавши за формулою бінома Ньютона вираз $(x+h)^n$, дістанемо формулу (5.27).

За допомогою цих чотирьох властивостей легко дістати послідовні скінченні різниці многочлена

$$y = f(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n.$$

За властивостями 1—3 маємо

$$\Delta y = a_0 \Delta(x^n) + a_1 \Delta(x^{n-1}) + \dots + a_{n-1} \Delta(x).$$

Використавши властивість 4, дістанемо

$$\begin{aligned} \Delta y = & a_0 (nhx^{n-1} + \frac{n(n-1)}{2!} h^2 x^{n-2} + \dots + nh^{n-1}x + h^n) + \\ & + a_1 ((n-1)hx^{n-2} + \frac{(n-1)(n-2)}{2!} h^2 x^{n-3} + \dots + (n-1)h^{n-2}x + \\ & + h^{n-1}) + \dots + a_{n-1} h. \end{aligned}$$

Звідси

$$\Delta y = a_0 nhx^{n-1} + (a_0 \frac{n(n-1)}{2!} h^2 + a_1(n-1)h) x^{n-2} + \dots + a_{n-1} h.$$

Отже, перша різниця многочлена n -го степеня з старшим членом $a_0 x^n$ є многочлен $(n-1)$ -го степеня з старшим членом $a_0 nhx^{n-1}$.

Обчисливши аналогічно послідовні різниці многочлена будь-якого порядку, впевнюємось у справедливості такої теореми:

Теорема 1. Якщо $f(x)$ — многочлен n -го степеня відносно x із старшим членом $a_0 x^n$, то різниця $\Delta^k f(x)$ (коли h — стале) при $k < n$ є многочленом $(n-k)$ -го степеня із старшим членом $a_0 n(n-1) \dots (n-k+1) h^k x^{n-k}$, коли $k=n$, $\Delta^n f(x) = a_0 n! h^n$ і при $k > n$ різниця $\Delta^k f(x) \equiv 0$.

З цієї теореми проте не випливає, що коли при деякому h n -ї різниці функції f стали, то ця функція є многочленом n -го степеня. Справедлива така обернена теорема (яку подаємо без доведення).

Теорема 2. Якщо n -ї різниці функції f стали для будь-якого кроку h , то ця функція є многочленом n -го степеня.

Теорему 2 широко застосовують в обчислювальній практиці. Завдяки їй для знаходження проміжних значень функції за допомогою таблиці різниць із сталим кроком h обмежуються обчисленням значень інтер-

поляційних многочленів, степінь яких дорівнює порядку практично сталих різниць (див. приклад 5 § 5.5).

Сформульовані вище властивості скінченних різниць справедливі лише для точних різниць функції. Але в процесі обчислень значення округлюються, тому порядок практично сталих різниць істотно залежить від точності, з якою обчислюють значення функції, і від величини кроку таблиці.

Якщо значення функції $y = \frac{1}{x}$ на відрізку $[1; 1,7]$ обчислюють з точністю до третього десяткового знака, то (див.табл. 5.12) при тому самому кроці $h = 0,1$ практично сталими є вже треті різниці і функція себе поводить на цьому відрізку як многочлен третього степеня. Що ж до різниць вищих порядків, то вони значно зростають за рахунок похибок округлення. Отже, чим з меншою точністю обчислюється значення функції в таблиці, тим нижчого степеня многочленами треба користуватися при інтерполюванні.

Таблиця 5.12

x_i	y_i	Δy_i	$\Delta^2 y_i$	$\Delta^3 y_i$	$\Delta^4 y_i$	$\Delta^5 y_i$	$\Delta^6 y_i$
1	1,000						
		-91					
1,1	0,909		15				
		-76		-3			
1,2	0,833		12		0		
		-64		-3		2	
1,3	0,769		9		2		-6
		-55		-1		-4	
1,4	0,714		8		-2		9
		-47		-3		5	
1,5	0,667		5		3		
		-42		0			
1,6	0,625		5				
		-37					
1,7	0,588						
Σ		-412	54	-10	3	3	3
S	-0,412	54	-10	3	3	3	

Як величина кроку таблиці впливає на табличні різниці, можна певнитись з табл. 5.13, де функцію $y = \frac{1}{x}$ на відрізку $[1; 1,06]$ затабульовано з кроком $h = 0,01$. При такому кроці навіть для таблиць з п'ятьма знаками практично сталими є другі різниці. Тому для інтерполювання можна використати квадратичні многочлени.

Таблиця 5.13

x_i	y_i	Δy_i	$\Delta^2 y_i$	$\Delta^3 y_i$
1,00	1,00000			
		-990		
1,01	0,99010		19	
		-971		0
1,02	0,98039		19	
		-952		0
1,03	0,97087		19	
		-933		-2
1,04	0,96154		17	
		-916		1
1,05	0,95238		18	
		-898		
Σ		-5660	92	-1
S	-0,05660	92	-1	1

Табличні різниці аналітично заданої на відрізку $[a; b]$ функції f змінюються більш-менш повільно. Тому порушення повільної зміни табличних різниць може свідчити або про наявність у таблиці помилкових значень функції, або ж про наявність на даній ділянці особливостей у функції (наприклад, різко вираженого екстремуму). В табл. 5.14 показано, як помилка величиною в ϵ в значенні функції y_k впливає на різниці різних порядків. З таблиці видно, що, по-перше, із зростанням порядку різниці зростає величина помилки і, по-друге, найбільша помилка в значеннях різниць парного порядку $\Delta^{2i} y$ ($i = 1, 2, \dots$) для випадку діагональної таблиці різниць міститься в тому самому рядку, що й помилка в значенні функції y_k . Що ж до найбільшої помилки

різниць непарного порядку $\Delta^{2i-1} y$ ($i = 1, 2, \dots$), то її мають два послідовні значення різниць, які містяться відповідно над і під тим рядком, у якому міститься помилкове значення функції.

Таблиця 5.14

x_i	y_i	Δy_i	$\Delta^2 y_i$	$\Delta^3 y_i$	$\Delta^4 y_i$	$\Delta^5 y_i$	$\Delta^6 y_i$
...		$\Delta^5 y_{k-5+\epsilon}$	
x_{k-2}	y_{k-2}				$\Delta^4 y_{k-4+\epsilon}$		$\Delta^6 y_{k-5-6\epsilon}$
		Δy_{k-2}		$\Delta^3 y_{k-3+\epsilon}$		$\Delta^5 y_{k-4-5\epsilon}$	
x_{k-1}	y_{k-1}		$\Delta^2 y_{k-2+\epsilon}$		$\Delta^4 y_{k-3-4\epsilon}$		$\Delta^6 y_{k-4+15\epsilon}$
		$\Delta y_{k-1+\epsilon}$		$\Delta^3 y_{k-2-3\epsilon}$		$\Delta^5 y_{k-3+10\epsilon}$	
x_k	$y_{k+\epsilon}$		$\Delta^2 y_{k-1-2\epsilon}$		$\Delta^4 y_{k-2+6\epsilon}$		$\Delta^6 y_{k-3-20\epsilon}$
		$\Delta y_{k-\epsilon}$		$\Delta^3 y_{k-1+3\epsilon}$		$\Delta^5 y_{k-2-10\epsilon}$	
x_{k+1}	y_{k+1}		$\Delta^2 y_{k+\epsilon}$		$\Delta^4 y_{k-1-4\epsilon}$		$\Delta^6 y_{k-2+15\epsilon}$
		Δy_{k+1}		$\Delta^3 y_{k-\epsilon}$		$\Delta^5 y_{k-1+5\epsilon}$	
x_{k+2}	y_{k+2}				$\Delta^4 y_{k+\epsilon}$		$\Delta^6 y_{k-1-6\epsilon}$
...		$\Delta^5 y_{k-\epsilon}$	

Якщо в таблиці, наприклад, шості різниці практично стали всюди, крім виділеної в табл. 5.14 ділянки, то за виправлене значення шостої різниці $\Delta^6 y_{k-3}$ можна взяти середнє арифметичне сімох значень різниць

шостого порядку, бо саме ця сума не міститиме в собі помилки. Таким чином, виправлене значення шостої різниці

$$\begin{aligned} \overline{\Delta^6 y_{k-3}} &= \frac{1}{7} ((\Delta^6 y_{k-6} + \epsilon) + (\Delta^6 y_{k-5} - 6\epsilon) + (\Delta^6 y_{k-4} + 15\epsilon) + \\ &+ (\Delta^6 y_{k-3} - 20\epsilon) + (\Delta^6 y_{k-2} + 15\epsilon) + (\Delta^6 y_{k-1} - 6\epsilon) + \\ &+ (\Delta^6 y_k + \epsilon)) = \frac{1}{7} (\Delta^6 y_{k-6} + \Delta^6 y_{k-5} + \Delta^6 y_{k-4} + \\ &+ \Delta^6 y_{k-3} + \Delta^6 y_{k-2} + \Delta^6 y_{k-1} + \Delta^6 y_k). \end{aligned}$$

Тоді помилку в значенні y_k знаходять за формулою

$$\epsilon = \frac{1}{20} (\Delta^6 y_{k-3} - \overline{\Delta^6 y_{k-3}})$$

і виражають в цілих одиницях нижчого розряду табличних значень функції.

Якщо в таблиці практично сталими є четверті різниці, то за виправлене значення четвертої різниці $\Delta^4 y_{k-2}$ слід взяти середнє арифметичне п'ятьох значень різниць четвертого порядку

$$\begin{aligned} \overline{\Delta^4 y_{k-2}} &= \frac{1}{5} ((\Delta^4 y_{k-4} + \epsilon) + (\Delta^4 y_{k-3} - 4\epsilon) + (\Delta^4 y_{k-2} + 6\epsilon) + \\ &+ (\Delta^4 y_{k-1} - 4\epsilon) + (\Delta^4 y_k + \epsilon)), \end{aligned}$$

яка не матиме помилки. Тоді помилку ϵ в табличному значенні функції y_k знаходять за формулою

$$\epsilon = \frac{1}{6} (\Delta^4 y_{k-2} - \overline{\Delta^4 y_{k-2}}).$$

Цю помилку також виражають в цілих одиницях нижчого розряду табличних значень функції.

Після виправлення значення функції y_k знову обчислюють значення всіх різниць. У такий спосіб знаходять ті помилки, які допущено в процесі обчислення значення функції в "ізолюваних" вузлах (якщо в сусідніх вузлах значення функції обчислено правильно).

Якщо буде встановлено, що помилкових значень функції в таблиці немає, то при складанні таблиці різниць спиняються на різницях, які мало змінюються, а різницями вищих порядків нехтують.

§5.5. ПЕРШИЙ ІНТЕРПОЛЯЦІЙНИЙ МНОГОЧЛЕН НЬЮТОНА

Нехай значення функції f задано для рівновіддалених значень аргументу $x_0, x_1 = x_0 + h, x_2 = x_0 + 2h, \dots, x_n = x_0 + nh$. Позначимо ці значення функції відповідно через $y_0, y_1, y_2, \dots, y_n$.

Побудуємо інтерполяційний многочлен n -го степеня вигляду:

$$\begin{aligned} P_n(x) &= a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \\ &+ a_3(x - x_0)(x - x_1)(x - x_2) + \dots + \\ &+ a_n(x - x_0)(x - x_1)\dots(x - x_{n-1}) \end{aligned} \quad (5.28)$$

так, щоб у вузлах інтерполювання x_i ($i = 0, 1, \dots, n$) він набував значень y_i ($i = 0, 1, \dots, n$), тобто задовольняв умови

$$P_n(x_i) = y_i \quad (i = 0, 1, \dots, n) \quad (5.29)$$

У §5.1 було доведено, що існує єдиний інтерполяційний многочлен степеня n , який задовольняє умови (5.29). Ним є інтерполяційний многочлен Лагранжа. Проте, як показує формула (5.28), інтерполяційний многочлен n -го степеня можна записати у вигляді, відмінному від інтерполяційного многочлена Лагранжа. Якщо в многочлені Лагранжа кожний з доданків є многочленом n -го степеня і всі доданки між собою рівноправні, то в многочлені (5.28) доданками є многочлени, порядки яких з віддаленням від початку підвищуються на одиницю.

Користуючись умовами (5.29), визначимо коефіцієнти a_0, a_1, \dots, a_n многочлена (5.28).

Покладемо в (5.28) $x = x_0$. Тоді з умов (5.29) дістанемо

$$y_0 = P_n(x_0) = a_0, \quad a_0 = y_0.$$

Якщо $x = x_1$, маємо

$$y_1 = P_n(x_1) = a_0 + a_1(x_1 - x_0),$$

або, оскільки $x_1 - x_0 = h$ і $a_0 = y_0$

$$y_1 = y_0 + a_1 h,$$

звідки

$$a_1 = \frac{y_1 - y_0}{h} = \frac{\Delta y_0}{h}.$$

Аналогічно, якщо $x = x_2$, дістанемо

$$y_2 = P_n(x_2) = a_0 + a_1(x_2 - x_0) + a_2(x_2 - x_0)(x_2 - x_1).$$

Підставивши в цю формулу знайдені раніше значення коефіцієнтів a_0 та a_1 , і врахувавши, що $x_2 - x_0 = 2h$, $x_2 - x_1 = h$, дістанемо

$$y_2 = 2y_1 - y_0 + a_2 2! h^2.$$

Звідси

$$a_2 = \frac{y_2 - 2y_1 + y_0}{2! h^2} = \frac{\Delta^2 y_0}{2! h^2}.$$

Аналогічно, якщо $x = x_3$,

$$a_3 = \frac{y_3 - 3y_2 + 3y_1 - y_0}{3! h^3} = \frac{\Delta^3 y_0}{3! h^3}.$$

Взагалі, для будь-якого k

$$a_k = \frac{\Delta^k y_0}{k! h^k} \quad (k = 0, 1, 2, \dots, n) \quad (\Delta^0 y_0 = y_0).$$

Підставивши знайдені значення коефіцієнтів a_k ($k = 0, 1, \dots, n$) в (5.28), дістанемо

$$\begin{aligned} P_n(x) = & y_0 + \frac{\Delta y_0}{h} (x - x_0) + \frac{\Delta^2 y_0}{2! h^2} (x - x_0)(x - x_1) + \\ & + \frac{\Delta^3 y_0}{3! h^3} (x - x_0)(x - x_1)(x - x_2) + \dots + \\ & + \frac{\Delta^k y_0}{k! h^k} (x - x_0)(x - x_1) \dots (x - x_{k-1}) + \dots + \\ & + \frac{\Delta^n y_0}{n! h^n} (x - x_0)(x - x_1) \dots (x - x_{n-1}). \end{aligned} \quad (5.30)$$

Цей многочлен називають *першим інтерполяційним многочленом Ньютона*.

Замінивши функцію f відповідним їй інтерполяційним многочленом Ньютона, дістанемо наближену рівність

$$f(x) \approx P_n(x). \quad (5.31)$$

Цю формулу називають *першою інтерполяційною формулою Ньютона*.

На вигляд многочлен (і формула) Ньютона відрізняється від многочлена (і формули) Лагранжа. Але якщо ці многочлени побудовано для тієї самої функції f і однієї й тієї самої системи вузлів x_i ($i = 0, 1, \dots, n$), то за теоремою про єдиність розв'язку інтерполяційної задачі вони тотожно дорівнюватимуть одна одній.

Якщо до заданої системи рівновіддалених вузлів інтерполювання додати ще один, то відповідний інтерполяційний многочлен Лагранжа треба будувати заново, а в інтерполяційному многочлені Ньютона додається лише один новий доданок, а вже обчислені доданки зберігаються без змін.

З формули (5.31) для $n = 1$ дістанемо формулу лінійного інтерполювання

$$f(x) \approx y_0 + \frac{\Delta y_0}{h} (x - x_0), \quad (5.32)$$

а для $n = 2$ — формулу квадратичного інтерполювання

$$f(x) \approx y_0 + \frac{\Delta y_0}{h} (x - x_0) + \frac{1}{2!} \frac{\Delta^2 y_0}{h^2} (x - x_0)(x - x_1). \quad (5.33)$$

В обчислювальній практиці зручніше користуватися іншою формою запису многочлена Ньютона (5.30). Якщо покласти

$$\frac{x - x_0}{h} = t \quad \text{то} \quad \frac{x - x_1}{h} = t - 1, \quad \frac{x - x_2}{h} = t - 2, \quad \dots, \quad \frac{x - x_{n-1}}{h} = t - n + 1,$$

і многочлен (5.30) матиме вигляд

$$P_n(x_0 + th) = y_0 + t\Delta y_0 + \frac{t(t-1)}{2!} \Delta^2 y_0 + \dots + \frac{t(t-1)(t-2)}{3!} \Delta^3 y_0 + \dots + \frac{t(t-1)(t-2)\dots(t-n+1)}{n!} \Delta^n y_0, \quad (5.34)$$

а перша інтерполяційна формула Ньютона (5.31)

$$f(x) \approx P_n(x_0 + th). \quad (5.35)$$

Різницю

$$f(x) - P_n(x_0 + th) = R_n(x, f) \quad (5.36)$$

називають *залишковим членом першої інтерполяційної формули Ньютона*.

Оскільки для даної функції f і даної системи рівновіддалених вузлів x_i ($i = 0, 1, \dots, n$) існує єдиний інтерполяційний многочлен степеня n , то інтерполяційні многочлени Лагранжа і Ньютона збігатимуться між собою, тобто $L_n(x) = P_n(x)$. А тому й залишковий член інтерполяційної формули Ньютона (5.31) збігатиметься із залишковим членом формули Лагранжа. Отже,

$$R_n(x, f) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)(x - x_1) \dots (x - x_n).$$

З введенням змінної t залишковий член $R_n(x, f)$ набуває вигляду

$$R_n(x, f) = \frac{h^{n+1}}{(n+1)!} f^{(n+1)}(\xi) t(t-1)(t-2) \dots (t-n). \quad (5.37)$$

Якщо $M_{n+1} = \max_{[x_0, x_n]} |f^{(n+1)}(\xi)|$, то для абсолютної похибки інтерполяційної формули (5.35) дістаємо оцінку

$$|R_n(x, f)| \leq \frac{h^{n+1}}{(n+1)!} M_{n+1} |t(t-1) \dots (t-n)|. \quad (5.38)$$

Користуючись формулами (5.26), що встановлюють зв'язок між по-

хідними і скінченними різницями, залишковий член першої інтерполяційної формули Ньютона (5.37) можна записати так:

$$R_n(x; f) \approx \frac{\Delta^{n+1} y_0}{(n+1)!} t(t-1)\dots(t-n).$$

Інтерполяційний многочлен Ньютона містить різниці $\Delta y_0, \Delta^2 y_0, \dots, \Delta^n y_0$. Тому інтерполяційною формулою (5.31) або (5.35) зручно користуватися на початку таблиці, тобто для всіх $x \in (x_0; x_1)$. Для цих значень x змінна $t < 1$. Якщо $x \in [x_1; x_2]$, то користуватися інтерполяційною формулою (5.35) недоцільно, бо t буде більшим за 1. У цьому разі за x_0 треба взяти вузол x_1 і в інтерполяційному многочлені використовувати різниці $\Delta y_1, \Delta^2 y_1, \dots, \Delta^n y_1$. Тому формулу (5.31) або (5.35) називають також *інтерполяційною формулою Ньютона для інтерполювання вперед*.

Важливими окремими випадками є лінійне і квадратичне інтерполювання. Якщо у формулах (5.34), (5.35) і (5.38) покласти $n=1$, то дістанемо формулу лінійного інтерполювання і оцінку її залишкового члена

$$f(x) \approx y_0 + t\Delta y_0, \quad t = \frac{x-x_0}{h}, \quad |R_1(x, f)| \leq \frac{M_2 h^2}{2} |t(t-1)|.$$

Якщо $n=2$, з тих самих формул дістаємо формулу квадратичного інтерполювання і оцінку її залишкового члена

$$f(x) \approx y_0 + t\Delta y_0 + \frac{t(t-1)}{2!} \Delta^2 y_0, \quad |R_2(x, f)| \leq \frac{M_3 h^3}{3!} |t(t-1)(t-2)|.$$

Якщо перші різниці функції практично стали, то користуються формулою лінійного інтерполювання, а якщо другі різниці функції практично стали, то користуються формулою квадратичного інтерполювання.

П р и к л а д 5. Функцію $y = \frac{1}{x}$ задано табл. 5. 11. Знайти її значення в точці $x = 1,05$ і оцінити похибку інтерполювання.

Р о з в ' я з а н н я. Тут не треба будувати інтерполяційний многочлен у явному вигляді. Треба лише обчислити його значення в точці $x = 1,05 \in (1; 1,1)$. Це обчислення зручно виконати за схемою Горнера. Як було вже зазначено в табл. 5. 11, практично сталими треба вважати різниці п'ятого порядку, а різниці шостого порядку такими, що дорівнюють нулю. Тому многочлен Ньютона 5-го степеня можна подати у вигляді

$$P_5(x_0 + ht) = y_0 + t(\Delta y_0 + \frac{t-1}{2} (\Delta^2 y_0 + \frac{t-2}{3} (\Delta^3 y_0 + \frac{t-3}{4} (\Delta^4 y_0 + \frac{t-4}{5} \Delta^5 y_0))). \quad (5.39)$$

Щоб обчислити значення $P_5(x_0 + ht)$, треба знайти послідовно

$$S_0 = \Delta^5 y_0, \quad S_1 = \Delta^4 y_0 + \frac{t-4}{5} S_0, \quad S_2 = \Delta^3 y_0 + \frac{t-3}{4} S_1,$$

$$S_3 = \Delta^2 y_0 + \frac{t-2}{3} S_2, \quad S_4 = \Delta y_0 + \frac{t-1}{2} S_3, \quad S_5 = y_0 + t S_4,$$

прячому $S_5 = P_5(x_0 + ht)$.

У цьому прикладі $x_0 = 1, y_0 = 1, \Delta y_0 = -0,09091, \Delta^2 y_0 = 0,01515, \Delta^3 y_0 = -0,00349; \Delta^4 y_0 = 0,00099, \Delta^5 y_0 = -0,00033$. Визначимо $t = (x-x_0)/h$. Оскільки $h = 0,1, x = 1,05$ і $x_0 = 1$, то $t = 0,5$. Тому, виконавши обчислення з однією запасною цифрою, дістанемо

$$S_0 = -0,00033, \quad S_1 = 0,00099 - \frac{3,5}{5} (-0,00033) = 0,001221,$$

$$S_2 = -0,00349 + \frac{-2,5}{4} \cdot 0,001221 = -0,004253,$$

$$S_3 = 0,01515 - \frac{1,5}{3} (-0,004253) = 0,017277,$$

$$S_4 = -0,09091 - \frac{0,5}{2} \cdot 0,017277 = -0,095229,$$

$$S_5 = 1,00000 + 0,5(-0,095229) = 0,9523854.$$

Отже, округливши результат до 5-го десяткового знака, маємо $P_5(1,05) = 0,95239$. Це значення більше за табличне на одну одиницю нижчого розряду. Це підтверджується також тим, що залишковий член інтерполяційної формули від'ємний. Для оцінки похибки інтерполювання знаходимо

$$y' = -\frac{1}{x^2}, \quad y'' = \frac{2}{x^3}, \quad y''' = -\frac{3!}{x^4}, \quad y^{(4)} = \frac{4!}{x^5}, \quad y^{(5)} = \frac{-5!}{x^6}, \quad y^{(6)} = \frac{6!}{x^7}.$$

Звідси маємо

$$M_6 = \max_{[1; 1,05]} |y^{(6)}| = 6!, \quad h = 0,1,$$

$$|R_5(x, \frac{1}{x})| \leq \frac{6!}{6!} |0,5(-0,5)(-1,5)(-2,5)(-3,5)(-4,5)| \cdot 0,1^6 = 0,000014756 < 0,00002.$$

Отже, обчислене за інтерполяційною формулою значення $f(1,05) \approx 0,95239$ має принаймні чотири правильні значущі цифри.

П р и к л а д 6. Побудувати інтерполяційний многочлен Ньютона з цілочисловими вузлами інтерполювання, який наближав би значення функції $f(x) = \ln x$ на відрізку $[20; 21]$ з точністю до 0,00005.

Розв'язання. Оскільки

$$f(x) = \frac{1}{x}, \quad f'(x) = -\frac{1}{x^2}, \quad f''(x) = \frac{2!}{x^3}, \quad f^{(n)}(x) = -\frac{3!}{x^4}, \quad f^{(5)}(x) = \frac{4!}{x^5},$$

то величина $M_{n+1} = \max |f^{(n+1)}(x)|$ для всіх n найбільшого значення набудатиме, якщо $x = 20$.

Перевіримо послідовно, чи забезпечуватиме наперед задану точність лінійне інтерполювання ($n = 1$). Для цього оцінимо залишковий член інтерполяційної формули $R_1(x; \ln x)$. Маємо

$$M_2 = \frac{1}{20^2} = 0,0025; \quad \omega_2(x) = (x - 20)(x - 21) = x^2 - 41x + 420 = (x - 20,5)^2 - 0,25.$$

На відрізку $[20; 21]$ $|\omega_2(x)| \leq |\omega_2(20,5)| = 0,25$.

Тому

$$|R_1(x; \ln x)| \leq \frac{0,0025}{2!} \cdot 0,25 = 0,0003125.$$

Отже, лінійне інтерполювання не забезпечує наперед заданої точності. Тепер перевіримо, чи забезпечуватиме обчислення значень функції $f(x) = \ln x$ на $[20; 21]$ з точністю 0,00005 квадратичне інтерполювання ($n = 2$). За вузли інтерполювання візьмемо $x_0 = 20$, $x_1 = 21$, $x_2 = 22$. Тоді

$$\omega_3(x) = (x - 20)(x - 21)(x - 22) = x^3 - 63x^2 + 1322x - 9240.$$

Цей многочлен на кінцях відрізка $[20; 22]$, як і в точці $x = 21$, дорівнює нулю; у точці $x = 21 + \frac{1}{\sqrt{3}} \approx 21,57735$ має локальний мінімум, причому

$\omega_3(21 + \frac{1}{\sqrt{3}}) \approx -0,3876$, а в точці $x = 21 - \frac{1}{\sqrt{3}} \approx 20,42265$ — локальний максимум, причому $\omega_3(21 - \frac{1}{\sqrt{3}}) \approx 0,3852$. Отже, $|\omega_3(x)| \leq$

$$\leq |\omega_3(21 + \frac{1}{\sqrt{3}})| \leq 0,3876. \quad \text{Але } M_3 = \frac{2}{20^3} = 0,00025,$$

тому

$$|R_2(x, \ln x)| \leq \frac{0,00025}{3!} \cdot 0,3846 = 0,000016.$$

Оскільки $|R_2(x, \ln x)| \leq 0,000016 \leq 0,00005$, то квадратичне інтерполювання забезпечує обчислення значень функції $\ln x$ на $[20; 21]$ з точністю 0,00005.

З цього прикладу видно, що похибку інтерполяційної формули можна оцінити до побудови інтерполяційного многочлена. Це дає змогу заздалегідь підібрати систему вузлів, а отже і степінь інтерполяційного многочлена, який забезпечить наперед задану точність.

Для побудови інтерполяційного многочлена Ньютона складемо таблицю різниць функції $\ln x$ (табл. 5. 15).

Таблиця 5. 15

i	x_i	y_i	Δy_i	$\Delta^2 y_i$
0	20	2,995732	48790	-2270
1	21	3,044522	46520	
2	22	3,091042		

Оскільки $h = 1$, $n = 2$, то

$$P_2(x) = 2,995732 + 0,048790(x - 20) - \frac{0,002270}{2}(x - 20)(x - 21) = 1,543232 + 0,095325x - 0,001135x^2.$$

§ 5.6. ДРУГИЙ ІНТЕРПОЛЯЦІЙНИЙ МНОГОЧЛЕН НЬЮТОНА

Якщо значення аргументу лежить ближче до кінця відрізка інтерполювання, то застосовувати першу інтерполяційну формулу Ньютона невигідно, бо не вистачатиме скінченних різниць функції вищих порядків. Якщо, наприклад, треба обчислити значення функції в точці $x = 1,55$ (див. табл. 5. 11), то для першого інтерполяційного многочлена Ньютона є лише перша і друга різниці, а для значення аргументу 1,65 — лише перша різниця функції.

Тому в кінці відрізка інтерполювання користуються многочленом вигляду

$$P_n(x) = b_0 + b_1(x - x_n) + b_2(x - x_n)(x - x_{n-1}) + \dots + b_n(x - x_n)(x - x_{n-1}) \dots (x - x_1). \quad (5.40)$$

Коефіцієнти $b_0, b_1, b_2, \dots, b_n$ многочлена (5.40) визначають так, щоб його значення в рівновіддалених вузлах інтерполювання збігалося із значенням відповідної функції, тобто щоб виконувались умови $P_n(x_i) = y_i$ ($i = 0, 1, \dots, n$).

Підставивши послідовно в формулу (5.40) замість x значення x_n, x_{n-1}, \dots, x_0 , знаходимо коефіцієнти b_0, b_1, \dots, b_n . Якщо в (5.40) покласти $x = x_n$, то дістанемо $b_0 = y_n$. Аналогічно, поклавши $x = x_{n-1}$, маємо

$$y_{n-1} = b_0 + b_1(x_{n-1} - x_n).$$

Але $x_{n-1} - x_n = -h$ і $b_0 = y_n$, тому

$$b_1 = \frac{y_n - y_{n-1}}{h} = \frac{\Delta y_{n-1}}{h}.$$

Поклавши в (5.40) $x = x_{n-2}$ і замінивши коефіцієнти b_0, b_1 їх значеннями, дістанемо

$$b_2 = \frac{y_n - 2y_{n-1} + y_{n-2}}{2!h^2} = \frac{\Delta^2 y_{n-2}}{2!h^2}.$$

Взагалі, коли $x = x_{n-k}$ із (5.40) знаходимо

$$b_k = \frac{\Delta^k y_{n-k}}{k!h^k} \quad (k = 1, 2, \dots, n).$$

Підставивши значення цих коефіцієнтів в (5.40), маємо

$$P_n(x) = y_n + \frac{\Delta y_{n-1}}{h}(x - x_n) + \frac{\Delta^2 y_{n-2}}{2!h^2}(x - x_n)(x - x_{n-1}) + \dots + \frac{\Delta^n y_0}{n!h^n}(x - x_n)(x - x_{n-1})\dots(x - x_1). \quad (5.41)$$

Цей многочлен і називають *другим інтерполяційним многочленом Ньютона*.

Замінивши функцію $f(x)$ другим інтерполяційним многочленом Ньютона, дістанемо наближену рівність

$$f(x) \approx P_n(x). \quad (5.42)$$

Рівність (5.42) називають *другою інтерполяційною формулою Ньютона*, а різницю $f(x) - P_n(x) = R_n(x;f)$ — *залишковим членом* цієї формули.

З формул (5.42) і (5.41) для $n = 1$ дістанемо формулу лінійного інтерполювання

$$f(x) \approx y_n + \frac{\Delta y_{n-1}}{h}(x - x_n),$$

а для $n = 2$ — квадратичного інтерполювання

$$f(x) \approx y_n + \frac{\Delta y_{n-1}}{h}(x - x_n) + \frac{1}{2!} \frac{\Delta^2 y_{n-2}}{h^2}(x - x_n)(x - x_{n-1}).$$

У практичних обчисленнях зручніше користуватися іншим записом многочлена (5.41). Якщо ввести таке позначення $(x - x_n)/h = t$ або $x = x_n + th$, то інші множники многочлена (5.40) через t виражатимуться так:

$$\frac{x - x_{n-1}}{h} = \frac{x - x_n + x_n - x_{n-1}}{h} = t + 1,$$

$$\frac{x - x_{n-2}}{h} = \frac{x - x_n + x_n - x_{n-2}}{h} = t + 2,$$

.....

$$\frac{x - x_1}{h} = \frac{x - x_n + x_n - x_1}{h} = t + n - 1.$$

А сам многочлен (5.41) тоді набирає вигляду

$$P_n(x_n + th) = y_n + t\Delta y_{n-1} + \frac{t(t+1)}{2!} \Delta^2 y_{n-2} + \dots + \frac{t(t+1)(t+2)}{3!} \Delta^3 y_{n-3} + \dots + \frac{t(t+1)\dots(t+n-1)}{n!} \Delta^n y_0. \quad (5.43)$$

До виразів многочлена (5.41) і (5.43) входять різниці $\Delta y_{n-1}, \Delta^2 y_{n-2}, \dots, \Delta^n y_0$, які розміщуються в діагональній таблиці різниць по діагоналі знизу вгору, тому формулу (5.42) використовують для інтерполювання в кінці таблиці. Якщо треба обчислити значення функції в точці x , то за x_n беруть найближче, але більше за x значення аргументу з таблиці, так, щоб $x \in (x_{n-1}; x_n)$ і $|t| < 1$. Тому формулу (5.42) називають такою *інтерполяційною формулою Ньютона для інтерполювання назад*.

Оскільки інтерполяційні многочлени Лагранжа і Ньютона є різними формами запису одного й того самого інтерполяційного многочлена, то оцінка залишкового члена формули Ньютона буде такою самою, як і для формули Лагранжа, побудованої для тієї самої функції й тієї самої системи вузлів. Тому для абсолютної похибки інтерполяційної формули (5.42) справедлива оцінка

$$|R_n(x;f)| \leq M_{n+1} \frac{h^{n+1}}{(n+1)!} |t(t+1)(t+2)\dots(t+n)|, \quad (5.44)$$

де $M_{n+1} = \max_{[x_0; x_n]} |f^{(n+1)}(\xi)|$.

Якщо похідна функції невідома, але є табличні різниці функції до $(n+1)$ -го порядку, то залишковий член $R_n(x;f)$ інтерполяційної формули (5.42) можна подати наближено формулою

$$R_n(x;f) \approx \frac{\Delta^{n+1} y}{(n+1)!} t(t+1)(t+2)\dots(t+n).$$

П р и к л а д 7. Користуючись табл. 5. 11, обчислити значення функції $y = \frac{1}{x}$ у точці $x = 1,64$ й оцінити похибку інтерполяції.

Р о з в ' я з а н н я. Значення $x = 1,64 \in (1,6; 1,7)$ знаходиться в кінці таблиці, тому використовуватимемо другу інтерполяційну формулу Ньютона (5.41), обмежившись многочленом 5-го степеня. Сам многочлен запишемо так:

$$P_5(x_n + th) = y_n + t(\Delta y_{n-1} + \frac{t+1}{2} (\Delta^2 y_{n-2} +$$

$$+ \frac{t+2}{3} \left(\Delta^3 y_{n-3} + \frac{t+3}{4} \left(\Delta^4 y_{n-4} + \frac{t+4}{5} \Delta^5 y_{n-5} \right) \right) \quad (5.45)$$

Тому значення многочлена $P_5(x_n + th)$ можна обчислити

$$S_0 = \Delta^5 y_{n-5}; \quad S_1 = \Delta^4 y_{n-4} + \frac{t+4}{5} S_0; \quad S_2 = \Delta^3 y_{n-3} + \frac{t+3}{4} S_1;$$

$$S_3 = \Delta^2 y_{n-2} + \frac{t+2}{3} S_2; \quad S_4 = \Delta y_{n-1} + \frac{t+1}{2} S_3;$$

$$S_5 = y_n + t S_4.$$

Отже, $P_5(x_n + th) = S_5$.

У цьому випадку: $x_n = 1,7$; $h = 0,1$; $y_n = 0,58824$; $\Delta y_{n-1} = -0,03676$; $\Delta^2 y_{n-2} = 0,00491$; $\Delta^3 y_{n-3} = -0,00104$; $\Delta^4 y_{n-4} = 0,00033$; $\Delta^5 y_{n-5} = -0,00014$.

Значення $t = (x - x_n) / h = \frac{1,64 - 1,7}{0,1} = -0,6$. Виконавши проміжні обчислення з однією запасною цифрою, матимемо

$$S_0 = -0,00014; \quad S_1 = 0,00033 + \frac{3,4}{5}(-0,00014) = 0,000235;$$

$$S_2 = -0,00104 + \frac{2,4}{4} \cdot 0,000235 = -0,000899;$$

$$S_3 = 0,00491 + \frac{1,4}{3}(-0,000899) = 0,004490;$$

$$S_4 = -0,03676 + \frac{0,4}{2} \cdot 0,004490 = -0,035862;$$

$$S_5 = 0,58824 + (-0,6)(-0,035862) = 0,609757.$$

Округливши результат до 5-го десяткового розряду, дістанемо $P_5(1,64) = 0,60976$, що збігається з відповідним значенням у п'ятизначних математичних таблицях. Обчислене на мікрокалькуляторі значення $y(1,64) = 0,609756$.

Похибку інтерполяції оцінюватимемо за формулою (5.44). Оскільки

$$M_6 = \max_{[1; 1,7]} |f^{(6)}(x)| = \max_{[1; 1,7]} \left| \frac{6!}{x^7} \right| = 6! \quad (\text{див. приклад 5}), \quad \text{то}$$

$$\left| R_5 \left(x; \frac{1}{x} \right) \right| \leq 6! \frac{0,1^6}{6!} \left| (-0,6) \cdot 0,4 \cdot 1,4 \cdot 2,4 \cdot 3,4 \cdot 4,4 \right| = 0,000012 < 0,00002.$$

Отже, оцінка залишкового члена гарантує, що чотири десяткових знаки результату, обчисленого за формулою (5.45), правильні. Насправді правильні й усі п'ять десяткових знаків.

§ 5.7. ЕКСТРАПОЛЮВАННЯ Й ОБЕРНЕНЕ ІНТЕРПОЛЮВАННЯ

Екстраполювання. Нехай функцію f задано своїми значеннями $y_i = f(x_i)$ ($i = 0, 1, \dots, n$) у рівновіддалених точках x_i ($i = 0, 1, \dots, n$). Обчислення значень функції f для значень аргументу, що не належать відрізьку $[x_0; x_n]$, тобто для $x < x_0$ і $x > x_n$, називають *екстраполюванням*, або *екстраполяцією*.

Для обчислення значення функції f в точці $x < x_0$ доцільно застосувати першу інтерполяційну формулу Ньютона. У цьому разі $t = \frac{x - x_0}{h} < 0$, тому кажуть, що першу інтерполяційну формулу Ньютона використовують для екстраполювання назад. Якщо $x > x_n$, то $t = \frac{x - x_n}{h} > 0$. У цьому разі доцільно використовувати другу інтерполяційну формулу Ньютона. Тоді кажуть, що її використовують для екстраполювання вперед.

Отже, першу інтерполяційну формулу Ньютона застосовують для інтерполювання вперед ($t > 0$) і екстраполювання назад ($t < 0$), а другу — для інтерполювання назад ($t < 0$) і екстраполювання вперед ($t > 0$). Для екстраполювання можна використовувати також інтерполяційну формулу Лагранжа, якщо, наприклад, функцію задано не в рівновіддалених вузлах. Для обчислення значення функції в деякій точці $x \notin [x_0; x_n]$ можна скористатись також інтерполяційною схемою Ейткіна.

Слід зазначити, що екстраполювання є менш точною операцією, ніж інтерполювання. Тому до екстраполювання можна вдатися тоді, коли відстань точки x від точок x_0 або x_n менша за крок таблиці h . При більшій відстані слід чекати значних похибок. Тому межі застосування екстраполяції обмежені.

П р и к л а д 8. Для функції $y = \frac{1}{x}$, заданої табл. 5.11, обчислити значення функції, якщо $x = 0,95$ і $x = 1,76$.

Р о з в ' я з а н н я. Нехай $x = 0,95$. Тоді $x_0 = 1$, $y_0 = 1$, $\Delta y_0 = -0,09091$, $\Delta^2 y_0 = 0,01515$, $\Delta^3 y_0 = -0,00349$, $\Delta^4 y_0 = 0,00099$, $\Delta^5 y_0 = -0,00033$, а $t = (x - x_0) / h = (0,95 - 1) / 0,1 = -0,5$. Записавши тепер перший інтерполяційний многочлен Ньютона 5-го степеня у вигляді (5.39) і виконавши всі проміжні обчислення з однією запасною цифрою, дістанемо

$$S_0 = \Delta^5 y_0 = -0,00033;$$

$$S_1 = \Delta^4 y_0 + S_0(t - 4) / 5 = 0,00099 + \frac{(-4,5)}{5}(-0,00033) = 0,001287;$$

$$S_2 = \Delta^3 y_0 + S_1(t - 3) / 4 = -0,00349 + \frac{(-3,5)}{4} \cdot 0,001287 = -0,004616;$$

$$S_3 = \Delta^2 y_0 + S_2(t - 2) / 3 = 0,01515 + \frac{(-2,5)}{3}(-0,004616) = 0,018997;$$

$$S_4 = \Delta y_0 + S_3(t-1)/2 = -0,09091 + \frac{(-1,5)}{2} 0,018997 = -0,105158;$$

$$S_5 = y_0 + tS_4 = 1 + (-0,5)(-0,105158) = 1,052579;$$

$$P_5(0,95) = S_5 = 1,052579.$$

Округливши результат до п'ятого десяткового розряду, дістанемо $P_5(0,95) \approx 1,05258$. Це значення відрізняється від табличного (1,05263) на п'ять одиниць нижчого розряду.

Нехай тепер $x = 1,76$. Оскільки $1,76 > 1,7 = x_n$, то для обчислення значення функції використаємо другий інтерполяційний многочлен Ньютона 5-го степеня (5.45), поклавши $x_n = 1,7$; $y_n = 0,58824$; $\Delta y_{n-1} = -0,03676$; $\Delta^2 y_{n-2} = 0,00491$; $\Delta^3 y_{n-3} = -0,00104$; $\Delta^4 y_{n-4} = 0,00033$; $\Delta y_{n-5} = -0,00014$. Значення $t = \frac{x - x_n}{h} = (1,76 - 1,7) / 0,1 = 0,6$. Виконавши всі проміжні обчислення з однією запасною цифрою, маємо

$$S_0 = \Delta^5 y_{n-5} = -0,00014;$$

$$S_1 = \Delta^4 y_{n-4} + S_0(t+4)/5 = 0,00033 + \frac{4,6}{5} \cdot (-0,00014) = 0,000201;$$

$$S_2 = \Delta^3 y_{n-3} + S_1(t+3)/4 = -0,00104 + \frac{3,6}{4} \cdot 0,000201 = -0,000859;$$

$$S_3 = \Delta^2 y_{n-2} + S_2(t+2)/3 = 0,00491 + \frac{2,6}{3} \cdot (-0,000859) = 0,004166;$$

$$S_4 = \Delta y_{n-1} + S_3(t+1)/2 = -0,03676 + \frac{1,6}{2} \cdot 0,004166 = -0,033427;$$

$$S_5 = y_n + tS_4 = 0,58824 + 0,6 \cdot (-0,033427) = 0,5681838;$$

$$P_5(1,76) = 0,5681838.$$

Округлюючи результат до 5-го десяткового розряду, дістанемо: $P_5(1,76) \approx 0,56818$, що збігається з табличним значенням.

Обернене інтерполювання. Раніше розглядалися задачі на знаходження значень функції для значень аргументів, яких немає в таблиці. Проте на практиці часто стикаємось з оберненою задачею: за таблицею функції знайти значення аргументу, яке відповідає заданому значенню функції, якого в таблиці немає. Цю задачу називають *задачею оберненого інтерполювання*.

Якщо функція f строго монотонна (зростаюча або спадна) на заданій ділянці таблиці, то для неї існує обернена монотонна (зростаюча або спадна) функція $x = \varphi(y)$. У цьому разі обернене інтерполювання зводиться до звичайного інтерполювання для оберненої функції $x = \varphi(y)$. При цьому значення $y_i = f(x_i)$ вважають значеннями аргументу, а значення x_i — значеннями функції $x_i = \varphi(y_i)$. Але табличні різниці Δy функції f не зберігають сталих значень (значення y_i не рівновіддалені),

тому для оберненого інтерполювання зручно використовувати інтерполяційний многочлен Лагранжа або інтерполяційну схему Ейткіна.

Інтерполяційний многочлен Лагранжа в цьому разі набуває такого вигляду:

$$x = \sum_{k=0}^n \frac{(y - y_0)(y - y_1)\dots(y - y_{k-1})(y - y_{k+1})\dots(y - y_n)}{(y_k - y_0)(y_k - y_1)\dots(y_k - y_{k-1})(y_k - y_{k+1})\dots(y_k - y_n)} x_k.$$

Якщо функція f не монотонна, то для неї записують інтерполяційний многочлен Ньютона (5.30) або Лагранжа (5.5), а потім розв'язують рівняння $P_n(x) = y$ або $L_n(x) = y$ відносно x при заданому y . Це рівняння розв'язують здебільшого методом ітерацій, який збігатиметься при виконанні, наприклад, таких умов: $|\Delta^2 y_0| \ll |\Delta y_0|$, а різниці вищих порядків $\Delta^3 y_0, \Delta^4 y_0, \dots$ малі порівняно з Δy_0 і $\Delta^2 y_0$. Спосіб оберненого інтерполювання, в якому розв'язується рівняння $P_n(x) = y$ або $L_n(x) = y$ відносно x при заданому y , можна застосовувати також і тоді, коли функція f строго монотонна.

Приклад 9. Для функції $y = \frac{1}{x}$, заданої табл. 5.11, знайти значення x , якщо $y = \frac{1}{x} = 0,99305$.

Розв'язання. Обчислення за інтерполяційною схемою Ейткіна з двома запасними цифрами в проміжних результатах зведено в табл. 5.16.

Таблиця 5.16

i	y_i	x_i	$y_i - y$	$P_{ij+1, \dots, j+1}$			
				j=1	j=2	j=3	j=4
0	1,00000	1	0,00695	1,007020	1,006991	1,006989	1,006989
1	0,99010	1,01	-0,00295	1,006962	1,007005	1,007003	
2	0,98039	1,02	-0,01266	1,006702	1,007019		
3	0,97087	1,03	-0,02218	1,006227			
4	0,96154	1,04	-0,03151				

Оскільки $P_{0,1,2,3}(0,99305) = P_{0,1,2,3,4}(0,99305) = 1,006989$, то округливши до 4-го десяткового розряду, дістанемо шукане значення $x = 1,0070$, що відповідає табличному значенню.

§ 5.8. ІНТЕРПОЛЮВАННЯ ФУНКЦІЙ ЗА ДОПОМОГОЮ СПЛАЙНІВ

При інтерполюванні функцій з великою кількістю вузлів інтерполяційний поліном має високий степінь, що спричиняє коливання полінома на проміжках між вузлами інтерполювання. Щоб зменшити степінь інтерполяційного полінома вузли інтерполювання можна розбити на групи і будувати інтерполяційні поліноми з меншою кількістю вузлів. Але в цьому разі на стиках між вузлами порушуються аналітичні властивості інтерполяційного полінома, з'являються точки розриву похідних. Позбутися цих недоліків при інтерполюванні можна за допомогою *сплайнів*. Сплайн на проміжку між вузлами інтерполювання є поліномом невисокого степеня. На всьому відрізку інтерполювання сплайн — це функція, склеєна з різних частин поліномів заданого степеня. Наочне уявлення про сплайни дають криві, побудовані за допомогою лекал, а також трамвайні та залізничні колії. Найпростіший приклад сплайнів — ламані.

Нехай маємо розбиття відрізка $[a; b]$ осі дійсних чисел сіткою вузлів

$$\Delta_n: a = x_1 < x_2 < \dots < x_i < \dots < x_n = b, \quad (5.46)$$

де n — натуральне число.

Означення. Функцію $S_m(x) = S_{m,k}(x, \Delta_n)$ називають *поліноміальним сплайном степеня m дефекта k* ($1 \leq k \leq m$) з вузлами (5.46), якщо $S_m(x)$ на кожному з відрізків $[x_i, x_{i+1}]$ ($i = 1, 2, \dots, n-1$) є поліномом степеня не вище m ; $S_m(x)$ є неперервною на $[a; b]$ функцією разом із своїми похідними до $(m-k)$ -го порядку включно.

Точки x_i ($i = 1, 2, \dots, n$) називають *вузлами сплайна*. Похідна $(m-k+1)$ -го порядку функції $S_m(x)$ може мати розриви у вузлах x_i ($i = 1, 2, \dots, n-1$).

З означення видно, що дефект характеризує гладкість сплайна. Чим більший дефект сплайна, тим менш гладкою функцією є сплайн, і навпаки. Так, сплайн $S_{1,1}(x)$ є поліномом першого степеня дефекту 1 і є ламаною лінією. Сплайн $S_{2,1}(x)$ є поліномом другого степеня дефекту 1. Це є неперервна на відрізку $[a; b]$ функція разом із своєю похідною першого порядку. Цю функцію називають *параболічним сплайном*. Параболічний сплайн — це функція, склеєна з парабол. Так, параболічний сплайн $S_{2,2}(x)$ є неперервною на відрізку $[a; b]$ функцією. Похідні цієї функції у вузлах сплайна можуть мати розриви. Найбільше застосування мають сплайни, склеєні з поліномів третього степеня. Такі сплайни називають *кубічними*.

Розглянемо задачу інтерполювання функції $y = f(x)$ на відрізку $[a; b]$ за допомогою кубічних сплайнів дефекту 1. Нехай на відрізку $[a; b]$ дійсної осі задано сітку (5.46), у вузлах якої відомі значення y_i ($i = 1, 2, \dots, n$) функції f , визначеної на $[a; b]$. Тоді задачу інтерполювання функції кубічним сплайном можна сформулювати так. На відрізку $[a; b]$ знайти функцію $S(x)$, яка задовольняє вимоги:

1) на кожному з відрізків $[x_i; x_{i+1}]$ ($i = 1, 2, \dots, n-1$) $S(x)$ є кубічним поліномом:

$$S(x) = S_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3. \quad (5.47)$$

2) $S(x)$ неперервна на $[a; b]$ функція разом із своїми похідними до другого порядку включно, тобто

$$S^{(p)}(x_k) = S^{(p)}_i(x_k), \quad (k = 2, 3, \dots, n-1; p = 0, 1, 2); \quad (5.48)$$

3) у вузлах x_i виконуються рівності

$$S(x_i) = y_i \quad (i = 1, 2, \dots, n). \quad (5.49)$$

Функцію $S(x)$, що задовольняє умови 1 — 3, називають *інтерполяційним кубічним сплайном*.

Для знаходження $4n-4$ невідомих коефіцієнтів сплайна маємо $4n-6$ рівнянь виду (5.48), (5.49). Не вистачає двох рівнянь, які знаходять з крайових умов:

$$S''(x_1) = S''(x_n) = 0. \quad (5.50)$$

Умови (5.50) означають, що на кінцях відрізка інтерполювання сплайн $S(x)$ має нульову кривизну.

Позначимо $S_i = S(x_i)$, $S'_i = S'(x_i)$, $h_i = x_{i+1} - x_i$ і виразимо коефіцієнти сплайна $S(x)$ через значення $S(x)$ і $S'(x)$ на кінцях відрізка $[x_i; x_{i+1}]$. Розв'язавши відносно a_i, b_i, c_i, d_i систему рівнянь

$$\begin{cases} S_i = a_i, \\ S'_i = 2c_i, \\ S_{i+1} = a_i + b_i h_i + c_i h_i^2 + d_i h_i^3, \\ S'_{i+1} = 2c_i + 6d_i h_i, \end{cases}$$

маємо

$$a_i = S_i, \quad c_i = \frac{1}{2} S'_i, \quad d_i = \frac{S'_{i+1} - S'_i}{6h_i},$$

$$b_i = \frac{S_{i+1} - S_i}{h_i} - \frac{1}{2} S'_i h_i - \frac{h_i}{6} (S'_{i+1} - S'_i).$$

Підставивши знайдені значення коефіцієнтів сплайна у формулу (5.47) та виконавши тотожні перетворення, дістанемо сплайн на відрізку $[x_i, x_{i+1}]$ у вигляді

$$\begin{aligned} S(x) = & \frac{x_{i+1} - x}{h_i} S_i + \frac{(x_{i+1} - x)((x_{i+1} - x)^2 - h_i^2)}{6h_i} S'_i + \\ & + \frac{x - x_i}{h_i} S_{i+1} + \frac{(x - x_i)((x - x_i)^2 - h_i^2)}{6h_i} S'_{i+1} \end{aligned} \quad (5.51)$$

Звідси знаходимо

$$S'(x) = \frac{S_{i+1} - S_i}{h_i} + \frac{h_i^2 - 3(x_{i+1} - x)^2}{6h_i} S_i'' + \frac{3(x - x_i)^2 - h_i^2}{6h_i} S_{i+1}''.$$

Обчислимо односторонні границі похідної $S'(x)$ в точках x_i ($i = 1, 2, \dots, n-2$):

$$S'(x_{i+1} - 0) = \frac{S_{i+1} - S_i}{h_i} + \frac{h_i}{6} S_i'' + \frac{h_i}{3} S_{i+1}'',$$

$$S'(x_{i+1} + 0) = \frac{S_{i+2} - S_{i+1}}{h_{i+1}} - \frac{h_{i+1}}{3} S_{i+1}'' - \frac{h_{i+1}}{6} S_{i+2}''.$$

Згідно з умовами (5.48) функції $S'(x)$ і $S''(x)$ повинні бути неперервними на відрізку $[a; b]$. Легко впевнитись у тому, що друга похідна сплайна (5.51) на відрізку $[a; b]$ неперервна. З умови неперервності $S'(x)$ в точках x_{i+1} ($i = 1, \dots, n-2$) і врахувавши рівності (5.49), дістанемо $n-2$ рівняння

$$\frac{1}{h_i} (y_{i+1} - y_i) + \frac{h_i}{6} S_i'' + \frac{h_i}{3} S_{i+1}'' =$$

$$= \frac{1}{h_{i+1}} (y_{i+2} - y_{i+1}) - \frac{h_{i+1}}{3} S_{i+1}'' - \frac{h_{i+1}}{6} S_{i+2}'', \quad i = 1, 2, \dots, n-2;$$

або

$$h_i S_i'' + 2(h_i + h_{i+1}) S_{i+1}'' + h_{i+1} S_{i+2}'' = 6 \left(\frac{y_{i+2} - y_{i+1}}{h_{i+1}} - \frac{y_{i+1} - y_i}{h_i} \right),$$

$$i = 1, 2, \dots, n-2.$$

Позначимо $\alpha_{i+1} = \frac{h_{i+1}}{h_i + h_{i+1}}$, $r_{i+1} = 1 - \alpha_{i+1}$, $\beta_{i+1} = \frac{6}{h_i + h_{i+1}} \times$

$$\times \left(\frac{y_{i+2} - y_{i+1}}{h_{i+1}} - \frac{y_{i+1} - y_i}{h_i} \right), \text{ тоді остання система набере вигляду}$$

$$r_{i+1} S_i'' + 2S_{i+1}'' + \alpha_{i+1} S_{i+2}'' = \beta_{i+1}, \quad i = 1, 2, \dots, n-2. \quad (5.52)$$

Доповнивши систему (5.52) рівняннями (5.50), дістанемо систему лінійних алгебраїчних рівнянь для визначення невідомих параметрів сплайн-функції (5.51). Систему (5.50) і (5.52) легко розв'язати методом Гауса. В результаті прямого ходу система набере вигляду

$$\begin{cases} S_i'' = 0, \\ S_{i+1}'' = F_{i+1} - h_{i+1} S_{i+2}'', \quad i = 1, 2, \dots, n-3; \\ S_{n-1}'' = F_{n-1}, \\ S_n'' = 0, \end{cases} \quad (5.53)$$

де

$$h_1 = 0, \quad F_1 = 0, \quad h_i = \frac{\alpha_i}{2 - h_{i-1} r_{i-1}}, \quad F_i = \frac{\beta_i - F_{i-1} r_i}{2 - h_{i-1} r_{i-1}}, \quad i = 2, \dots, n-1.$$

З останньої системи послідовно знаходимо

$$S_n'' = 0, \quad S_{n-1}'' = F_{n-1},$$

$$S_i'' = F_i - h_i S_{i+1}'', \quad i = n-2, n-3, \dots, 2; \quad (5.54)$$

$$S_1'' = 0.$$

Підставивши значення $S_1'', S_2'', \dots, S_n''$ в (5.51), дістанемо вигляд кубічного полінома на кожному з відрізків $[x_i; x_{i+1}]$, тобто в явному вигляді матимемо на відрізку $[a; b]$ сплайн $S(x)$, який інтерполюватиме задану таблично функцію $y = f(x)$.

Зазначимо, що система рівнянь (5.53) завжди має розв'язок, тобто сплайн (5.51) існує.

Доведено, що кубічна сплайн-функція — це єдина функція з мінімальною кривиною серед усіх функцій, які інтерполюють задану функцію, і має квадратично інтегровну другу похідну. В цьому розумінні кубічний сплайн з крайовими умовами (5.50) є найкращою з функцій, які інтерполюють задану функцію.

Метод обчислення параметрів S_i'' кубічного інтерполяційного сплайна реалізований в програмі 5.3. За допомогою програми обчислюється значення сплайна в будь-якій точці відрізка $[a; b]$ за формулою (5.51). Якщо $x < a$, то виконується лінійна екстраполяція за формулою

$$S(x) = y_1 - (x - a) \left(\frac{x_2 - x_1}{6} S_2'' + \frac{y_2 - y_1}{x_2 - x_1} \right),$$

якщо $x > b$, то здійснюється екстраполяція за формулою

$$S(x) = y_n - (x - x_n) \left(\frac{x_n - x_{n-1}}{6} S_{n-1}'' + \frac{y_n - y_{n-1}}{x_n - x_{n-1}} \right).$$

П р и к л а д 10. Побудувати інтерполяційний кубічний сплайн для функції $y = 1 / (1 + 25x^2)$ на відрізку $[-1; 1]$ з вузлами $x_{i+1} = x_i + h$, $i = 1, 2, \dots, 20$, $x_1 = -1$, $h = 0,1$.

Р о з в ' я з а н н я. Нехай значення функції $y(x)$ у заданих вузлах дорівнюють $y_i = y(x_i)$, $i = 1, 2, \dots, 21$. Сплайн шукатимемо у вигляді (5.51). Врахуємо, що відстані $h_i = h$ між вузлами сталі, тоді система рівнянь для визначення невідомих параметрів має вигляд

$$\begin{cases} 4S'_1 + S'_3 = \frac{6}{h^2}(y_3 - 2y_2 + y_1), \\ S'_i + 4S'_{i+1} + S'_{i+2} = \frac{6}{h^2}(y_{i+2} - 2y_{i+1} + y_i), \quad i = 2, 3, \dots, 18; \\ S'_{19} + 4S'_{20} = \frac{6}{h^2}(y_{21} - 2y_{20} + y_{19}). \end{cases}$$

Зауважимо, що $S'_1 = 0$ і $S'_{21} = 0$.

З останньої системи послідовно знаходимо

$$S'_{20} = F_{20}, \quad S'_i = F_i - l_i S'_{i+1}, \quad i = 19, 18, \dots, 2;$$

де $l_1 = 0$, $F_1 = 0$,

$$l_i = \frac{1}{l_{i-1} + 4}, \quad F_i = \frac{6}{h^2}(y_{i+1} - 2y_i + y_{i-1}), \quad i = 2, 3, \dots, 20.$$

Обчисливши значення S'_i ($i = 19, 18, \dots, 1$), можна на кожному з відрізків $[x_1; x_2]$, $[x_2; x_3]$, ..., $[x_{20}; x_{21}]$ записати вигляд сплайна $S(x)$. У табл. 5.17 внесено результати інтерполювання даної функції побудованим інтерполяційним кубічним сплайном, інтерполяційним поліномом Лагранжа двадцятого степеня з вузлами, що збігаються з вузлами сплайна, і інтерполяційним поліномом Лагранжа третього степеня з вузлами $\{-1, -1/3, 1/3, 1\}$. З таблиці видно, що інтерполяційний кубічний сплайн має найменші відхилення від точних значень заданої функції у точках між вузлами інтерполювання. Поблизу точок $x = -1$ та $x = 1$ найбільші похибки має інтерполяційний поліном Лагранжа двадцятого степеня.

Таблиця 5.17

x	$y(x)$	$S(x)$	$L_{20}(x)$	$L_3(x)$
-0.975	0.040378	0.040469	-29.06340	0.051029
-0.875	0.049651	0.049624	2.177484	0.098116
-0.575	0.107926	0.107922	0.088064	0.208834
-0.275	0.345946	0.345536	0.348475	0.273738
0.025	0.984615	0.983324	0.985872	0.292828
0.225	0.441379	0.440515	0.446107	0.280101
0.525	0.126733	0.126740	0.081529	0.222833
0.725	0.070718	0.070721	0.857682	0.159202
0.825	0.055507	0.055492	8.648327	0.119750
0.925	0.044661	0.044711	-197.9204	0.075208

```

10 CLS: PRINT
20 PRINT"          СПЛАЙН-АПРОКСИМАЦІЯ, "
30 PRINT"          ІНТЕРПОЛЯЦІЯ, ЕКСТРАПОЛЯЦІЯ"
40 PRINT
50 INPUT "Задайте число вузлів N=";N
60 DIM X(N), Y(N), L(N), S2P(N)
70 PRINT "Ввести значення:"
80 FOR I=1 TO N
90   PRINT"х("I"), ";;"у("I")";: INPUT X(I), Y(I)
100 NEXT I
110 H1=X(2)-X(1)          : ' Розв'язування
115 EE=(Y(2)-Y(1))/H1    : '
120 F(1)=0: L(1)=0       : ' системи
130 FOR I=2 TO N-1      : '
140   H1=H1: H1=X(I+1)-X(I) : ' рівнянь
150 E=EE: EE=(Y(I+1) - Y(I))/H1 : ' Рядки
160 FF=6*(EE-E)/(H1+H1) : ' 110-210 —
170 ALL=H1/(H1+H1): R=1-ALL : ' прямий
180 ZN=2-L(I-1)*R       : ' хід, а
190 L(I)=ALL/ZN         : ' рядки
200 F(I) = (FF-F(I-1)*R)/ZN : ' 210-260 —
210 NEXT I              : ' зворотний
220 S2P(1)=0: S2P(N)=0 : ' хід
230 S2P(N-1)=F(N-1)
240 FOR I=N-2 TO 2
250   S2P(I)=F(I)-L(I)*S2P(I+1)
260 NEXT I
270 FOR I=1 TO N
280   PRINT "S2P("I") ="; S2P(I)
290 NEXT I
300 PRINT "Чи будете обчислювати значення сплайна? (т/н)"
310 INPUT B$
320 IF INSTR ("ТtТт", B$) = 0 THEN 540
325 IF INSTR ("NnНн", B$)=0 THEN 540
330 INPUT "Ввести аргумент X=";XX
340 I=1
350 IF XX<=X(N) THEN 400
360 H=X(N)-X(N-1)      : ' Екстра-
370 Y=H*S2P (N-1)/6+(Y(N)-Y(N-1))/H : ' поляція
380 Y=Y(N)-Y*(XX-X(N)) : ' функції.
390 GOTO 520          : ' Рядки
400 IF XX>X(1) THEN 450 : ' 360-380,
410 H=X(2)-X(1)      : ' для х < x(n),
420 Y=H*S2P(2)/6-(Y(2)-Y(1))/H : ' а рядки

```

```

430 Y=Y(1)+Y*(XX-X(1))      : ' 410-430,
440 GOTO 520                  : для x > x(n)
450 IF XX>X(1) THEN I=I+1: GOTO 450
460 J=I-1: HH=X(I)-XX: H=XX-X(J)
470 HI=X(I)-X(J): ZN=6*HI     : ' Інтер-
480 C1=HH/HI: C3=H/HI        : ' поляція
490 C2=HH*(HH*HH-HI*HI)/ZN   : ' функції
500 C4=H*(H*H-HI*HI)/ZN     : ' рядки
510 Y=C1*Y(J)+C2*S2P(J)+C3*Y(I)+C4*S2P(I)
520 PRINT "Значення Y("XX")=";Y: PRINT
530 GOTO 300
540 END

```

§ 5.9. ЧИСЕЛЬНЕ ДИФЕРЕНЦІЮВАННЯ ФУНКЦІЙ

До чисельного диференціювання функцій вдаються тоді, коли функцію задано таблично, а тому методи диференціального числення просто незастосовні, або коли функцію задано досить складним аналітичним виразом і тому обчислення похідних пов'язане із значними труднощами.

Щоб побудувати формули чисельного диференціювання, задану на відрізку $[a; b]$ функцію f замінюють відповідним інтерполяційним многочленом $P(x)$. Тоді

$$f(x) = P(x) + R(x; f), \quad (5.55)$$

де $R(x; f)$ — залишковий член інтерполяційної формули.

Якщо функція f на $[a; b]$ має похідні до k -го порядку включно, то, диференціюючи тотожність (5.55) по x , знаходять

$$\begin{aligned}
 f'(x) &= P'(x) + R'(x; f), \\
 f''(x) &= P''(x) + R''(x; f), \\
 &\dots\dots\dots \\
 f^{(k)}(x) &= P^{(k)}(x) + R^{(k)}(x; f).
 \end{aligned}$$

За наближені значення похідних від функції f беруть перші доданки правих частин цих рівностей:

$$f'(x) \approx P'(x), f''(x) \approx P''(x), \dots, f^{(k)}(x) \approx P^{(k)}(x), \quad x \in [a, b]. \quad (5.56)$$

Тоді залишкові члени $r_i(x)$ ($i=1, 2, \dots, k$) формул чисельного диференціювання (5.56) дорівнюватимуть похідним від залишкового члена інтерполяційної формули (5.55), тобто

$$r_i(x) \equiv f^{(i)}(x) - P^{(i)}(x). \quad (5.57)$$

Слід зазначити, що з малості залишкового члена інтерполяційної

формули $R(x; f)$ зовсім не впливає малість залишкових членів похідних (похибки чисельного диференціювання) $r_i(x)$, бо похідні від малих функцій можуть бути досить великими. Наприклад, функції

$y_1(x) = f(x)$ і $y_2(x) = f(x) + \frac{\cos n^3 x}{n^2}$ для великих значень n можуть відрізнятися між собою, як завгодно мало

$$|y_1 - y_2| \leq \frac{|\cos n^3 x|}{n^2} \leq \frac{1}{n^2}.$$

Але похідні від них для деяких значень x і великих значень n можуть значно відрізнятися між собою:

$$|y_1' - y_2'| \leq \frac{n^3 |\sin n^3 x|}{n^2} = n |\sin n^3 x| \leq n,$$

$$|y_1'' - y_2''| \leq n^4 |\cos n^3 x| \leq n^4.$$

Звідси видно, що неперервної залежності значень похідної від значень функції не існує. Тому задача чисельного диференціювання, взагалі кажучи, є менш точною операцією порівняно з інтерполяванням, є некоректною задачею. На рис. 5.4 в точці x_1 ординати функції f і многочлена P однакові, проте кутові коефіцієнти дотичних значно відрізняються.

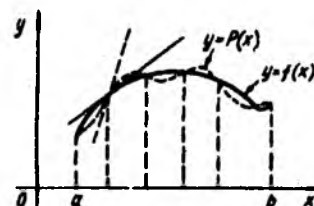


Рис. 5.4

Якщо інтерполяційний многочлен P на певній ділянці з достатньою точністю наближає функцію f , а сама функція f досить гладка і змінюється плавно на цій ділянці, то можна сподіватись, що при досить малому кроці інтерполявання похідні інтерполяційного многочлена також мало відрізнятимуться від похідних функції f . Проте не слід забувати, що із зростанням порядку похідної точність чисельного диференціювання здебільшого різко спадає. Тому на практиці формули чисельного диференціювання для похідних, вищих від другого порядку, застосовують досить рідко.

Формули чисельного диференціювання, побудовані за першою інтерполяційною формулою Ньютона. Нехай функцію f задано в рівновіддалених точках x_i ($i=0, 1, \dots, n$) відрізка $[a; b]$ значеннями $y_i = f(x_i)$ ($i=0, 1, \dots, n$). Щоб знайти похідні f' і f'' в точках x , близьких до x_0 , функцію f наближено замінюють першим інтерполяційним многочленом Ньютона (5.34), побудованим за вузлами x_0, x_1, \dots, x_n .

Кожний доданок формули (5.34) подамо за степенями t . Оскільки

$$t(t-1)(t-2)\dots(t-k) = t^{k+1} - S_k^{(1)}t^k + S_k^{(2)}t^{k-1} - \dots + (-1)^k S_k^{(k)}t,$$

де $S_k^{(i)}$ — сума всіх добутоків натуральних чисел від 1 до k по i множник:

(наприклад, для $n = 4$ маємо: $t(t-1)(t-2)(t-3)(t-4) = t^5 - (1+2+3+4)t^4 + (1\cdot2+1\cdot3+1\cdot4+2\cdot3+2\cdot4+3\cdot4)t^3 - (1\cdot2\cdot3+1\cdot2\cdot4+1\cdot3\cdot4+2\cdot3\cdot4)t^2 + 1\cdot2\cdot3\cdot4\cdot t - 1\cdot2\cdot3\cdot4 = t^5 - 10t^4 + 35t^3 - 50t^2 + 24t$), то формулу (5.34) можна записати так:

$$f(x) \approx P_n(x) = y_0 + t\Delta y_0 + \frac{t^2-t}{2!} \Delta^2 y_0 + \frac{t^3-3t^2+2t}{3!} \Delta^3 y_0 + \frac{t^4-6t^3+11t^2-6t}{4!} \Delta^4 y_0 + \frac{t^5-10t^4+35t^3-50t^2+24t}{5!} \Delta^5 y_0 + \dots + \frac{t^n - S_{n-1}^{(1)}t^{n-1} + S_{n-1}^{(2)}t^{n-2} - \dots + (-1)^{n-1}S_{n-1}^{(n-1)}t}{n!} \Delta^n y_0, \quad (5.58)$$

де $t = \frac{x-x_0}{h}$, $h = x_{i+1} - x_i$ ($i = 0, 1, \dots, n-1$).

Врахувавши, що похідні

$$\frac{df}{dx} = \frac{df(x)}{dt} \cdot \frac{dt}{dx} = \frac{1}{h} \frac{df(x)}{dt},$$

$$\frac{d^2f}{dx^2} = \frac{d}{dx} \left(\frac{1}{h} \frac{df(x)}{dt} \right) = \frac{1}{h} \frac{d}{dx} \left(\frac{df(x)}{dt} \right) = \frac{1}{h} \frac{d^2f(x)}{dt^2} \cdot \frac{dt}{dx} = \frac{1}{h^2} \frac{d^2f}{dt^2},$$

і продиференціювавши двічі по t наближену рівність (5.58), дістанемо

$$f'(x) \approx P_n'(x) = \frac{1}{h} (\Delta y_0 + \frac{2t-1}{2!} \Delta^2 y_0 + \frac{3t^2-6t+2}{3!} \Delta^3 y_0 + \frac{4t^3-18t^2+22t-6}{4!} \Delta^4 y_0 + \frac{5t^4-40t^3+105t^2-100t+24}{5!} \Delta^5 y_0 + \dots + \frac{nt^{n-1} - (n-1)S_{n-1}^{(1)}t^{n-2} + (n-2)S_{n-1}^{(2)}t^{n-3} - \dots + (-1)^{n-1}S_{n-1}^{(n-1)}}{n!} \Delta^n y_0), \quad (5.59)$$

$$f''(x) \approx P_n''(x) = \frac{1}{h^2} (\Delta^2 y_0 + \frac{6t-6}{3!} \Delta^3 y_0 + \frac{12t^2-36t+22}{4!} \Delta^4 y_0 + \frac{20t^3-120t^2+210t-100}{5!} \Delta^5 y_0 + \dots + (n(n-1)t^{n-2} - (n-1) \times$$

$$\times (n-2)S_{n-1}^{(1)}t^{n-3} + (n-2)(n-3)S_{n-1}^{(2)}t^{n-4} + \dots + (-1)^n 2S_{n-1}^{(n-2)}) \Delta^n y_0 / n!).$$

Аналогічно можна обчислити похідні вищих порядків.

Якщо похідні обчислюють за формулами (5.59), (5.60) в точці x , то за точку x_0 вибирають найближче табличне значення аргументу, яке менше за x .

Формули чисельного диференціювання (5.59), (5.60) значно спрощуються, якщо значення похідних обчислювати у вузлах інтерполявання. Оскільки кожне табличне значення можна взяти за x_0 , то, поклавши у формулах (5.59), (5.60) $t = 0$, дістанемо

$$f'(x_0) \approx P_n'(x_0) = \frac{1}{h} (\Delta y_0 - \frac{1}{2} \Delta^2 y_0 + \frac{1}{3} \Delta^3 y_0 - \frac{1}{4} \Delta^4 y_0 + \frac{1}{5} \Delta^5 y_0 - \frac{1}{6} \Delta^6 y_0 + \dots + \frac{(-1)^{n-1}}{n} \Delta^n y_0), \quad (5.61)$$

$$f''(x_0) \approx P_n''(x_0) = \frac{1}{h^2} (\Delta^2 y_0 - \Delta^3 y_0 + \frac{11}{12} \Delta^4 y_0 - \frac{5}{6} \Delta^5 y_0 + \frac{137}{180} \Delta^6 y_0 + \dots + (-1)^n \frac{2S_{n-1}^{(n-2)}}{n!} \Delta^n y_0). \quad (5.62)$$

Користуючись рівністю (5.57), можна знайти похибки формул чисельного диференціювання. Оскільки залишковий член першої інтерполяційної формули Ньютона

$$R_n(x; f) = \frac{h^{n+1}}{(n+1)!} f^{(n+1)}(\xi) t(t-1)(t-2) \dots (t-n),$$

то, припустивши, що функція $f(x)$ на $[a; b]$ має неперервну похідну $(n+2)$ -го порядку, дістанемо

$$r_1(x) = R_n'(x; f) = \frac{h^{n+1}}{(n+1)!} \frac{df^{(n+1)}(\xi)}{dx} t(t-1)(t-2) \dots (t-n) + \frac{h^n}{(n+1)!} f^{(n+1)}(\xi) \frac{d}{dt} (t(t-1)(t-2) \dots (t-n)).$$

Звідси, якщо $t = 0$ ($x = x_0$), для залишкового члена формули чисельного диференціювання (5.61) маємо

$$r_1(x_0) = R_n'(x_0) = (-1)^n \frac{h^n}{n+1} f^{(n+1)}(\xi),$$

бо

$$\frac{d}{dt} (t(t-1)(t-2) \dots (t-n))_{t=0} = (-1)^n n!$$

Але похідну $f^{(n+1)}(\xi)$ в багатьох випадках оцінити важко. Тому при досить малих h припускають, що

$$f^{(n+1)}(\xi) \approx \frac{\Delta^{n+1} y_0}{h^{n+1}}.$$

Тоді залишковий член для похідної в точці $x = x_0$ дорівнює

$$r_1(x_0) = \frac{(-1)^n \Delta^{n+1} y_0}{h^{n+1}}$$

Аналогічно можна знайти залишковий член $r_2(x) = R_n''(x; f)$ для другої похідної $f''(x)$.

П р и к л а д 12. У точках $x = 1$, $x = 1,1$ знайти першу і другу похідні від функції $y = \frac{1}{x}$, заданої табл. 5.11.

Р о з в ' я з а н н я. Оскільки п'ять різниць практично стали, то покладемо $h = 0,1$. Точка $x = 1$ розміщена на початку таблиці, тому першу й другу похідні обчислюватимемо відповідно за формулами (5.61) і (5.62). У цьому разі $h = 0,1$; $x_0 = 1$; $\Delta y_0 = -0,09091$; $\Delta^2 y_0 = 0,01515$; $\Delta^3 y_0 = -0,00349$; $\Delta^4 y_0 = 0,00099$; $\Delta^5 y_0 = -0,00033$.

Виконавши обчислення, знайдемо

$$f(1) \approx 10 \cdot (-0,09091 - \frac{1}{2} \cdot 0,01515 - \frac{1}{3} \cdot 0,00349 - \frac{1}{4} \cdot 0,00099 - \frac{1}{5} \cdot 0,00033) = -10(0,09091 + 0,007575 + 0,0011633 + 0,0002475 + 0,000066) = -10 \cdot 0,0999618 = -0,99962 \approx -1,00;$$

$$f'(1) \approx 100 \cdot (0,01515 - (-0,00349) + \frac{11}{12} \cdot 0,00099 - \frac{5}{9} \cdot (-0,00033)) = 100 \cdot 0,0198225 = 1,98225 \approx 1,98.$$

Точні значення похідних: $f'(1) = -1$, $f''(1) = 2$. Тому похибка в обчисленні першої похідної дорівнює $-0,00038$, другої $-0,01775$.

Щоб проілюструвати, як впливає на точність чисельного диференціювання кількість узятих у формулах (5.61) і (5.62) доданків, наведемо їх значення, обчислені з використанням відповідно одного, двох, трьох (а для $f'(1)$ і чотирьох) доданків. Маємо $f'(1) = -0,9091$; $f'(1) = -0,98485$; $f'(1) = -0,996483$; $f'(1) = -0,998958$; $f''(1) = 1,515$; $f''(1) = 1,864$; $f''(1) = 1,95475$.

Обчислюючи значення похідних у точці $x = 1,1$, вважаємо, що $h = 0,1$; $x_0 = 1,1$; $\Delta y_0 = -0,07576$; $\Delta^2 y_0 = 0,01166$; $\Delta^3 y_0 = -0,00250$; $\Delta^4 y_0 = -0,00066$; $\Delta^5 y_0 = -0,00019$.

Виконавши обчислення, матимемо

$$f(1,1) \approx 10 \cdot (-0,07576 - \frac{1}{2} \cdot 0,01166 + \frac{1}{3} \cdot (-0,00250) - \frac{1}{4} \cdot 0,00066 + \frac{1}{5} \cdot (-0,00019)) = -10 \cdot (0,07576 + 0,00583 + 0,0008333 + 0,000165 + 0,000038) = -10 \cdot 0,0826263 = -0,826263;$$

$$f'(1,1) \approx 100 \cdot (0,01166 - (-0,00250) + \frac{11}{12} \cdot 0,00066 -$$

$$- \frac{5}{6} \cdot (-0,00019)) = 100 \cdot (0,01166 + 0,00250 + 0,000605 + 0,0001583) = 100 \cdot 0,0149233 = 1,49233.$$

Значення похідних у точці $x = 1,1$, обчислені на ЕКОМ, дорівнюють: $f(1,1) = -0,82644626$, $f'(1,1) = 1,5026296$. Тому похибка в обчисленні першої похідної дорівнює $0,0001862$, другої $-0,0102996$.

Вплив на точність чисельного диференціювання в точці $x_0 = 1,1$ за формулами (5.61) і (5.62) від кількості врахованих у правій частині цих формул доданків можна простежити за такими даними:

$$f(1,1) = -0,7576; f'(1,1) = -0,8159; f''(1,1) = -0,824233; f'(1,1) = -0,825883; f''(1,1) = 1,166; f'''(1,1) = 1,416; f''(1,1) = 1,4765.$$

Щоб збільшити точність чисельного диференціювання, треба табличні значення функції брати з більшою кількістю цифр, а у формулах (5.61) і (5.62) — використовувати більшу кількість доданків.

Формули чисельного диференціювання, побудовані за другою інтерполяційною формулою Ньютона. Щоб знайти похідні від функції, заданої таблицію в рівновіддалених вузлах x_0, x_1, \dots, x_n , в точці x , що міститься в кінці таблиці, використовують другу інтерполяційну формулу Ньютона (5.42). Записавши кожний доданок цієї формули за степенями t , формулу (5.42) подамо в такому вигляді:

$$f(x) \approx P_n(x) = y_n + t \Delta y_{n-1} + \frac{t^2 + t}{2!} \Delta^2 y_{n-2} + \frac{t^3 + 3t^2 + 2t}{3!} \Delta^3 y_{n-3} + \frac{t^4 + 6t^3 + 11t^2 + 6t}{4!} \Delta^4 y_{n-4} + \frac{t^5 + 10t^4 + 35t^3 + 50t^2 + 24t}{5!} \Delta^5 y_{n-5} + \dots + \frac{t^n + S_{n-1}^{(1)} t^{n-1} + S_{n-1}^{(2)} t^{n-2} + S_{n-1}^{(3)} t^{n-3} + \dots + S_{n-1}^{(n-1)} t}{n!} \Delta^n y_0, \quad (5.63)$$

$$\text{де } t = \frac{x - x_n}{h}, \quad h = x_{i+1} - x_i \quad (i=0, 1, \dots, n-1).$$

Оскільки

$$\frac{df(x)}{dx} = \frac{df(x)}{dt} \cdot \frac{dt}{dx} = \frac{1}{h} \cdot \frac{df(x)}{dt},$$

$$\frac{d^2 f(x)}{dx^2} = \frac{d}{dx} \left(\frac{1}{h} \cdot \frac{df(x)}{dt} \right) = \frac{1}{h} \frac{d}{dx} \left(\frac{df(x)}{dt} \right) = \frac{1}{h} \frac{d^2 f(x)}{dt^2} \cdot \frac{dt}{dx} = \frac{1}{h^2} \frac{d^2 f(x)}{dt^2},$$

то, диференціюючи по t наближену рівність (5.63), дістанемо

$$f'(x) \approx P_n'(x) = \frac{1}{h} \left(\Delta y_{n-1} + \frac{2t+1}{2!} \Delta^2 y_{n-2} + \frac{3t^2+6t+2}{3!} \Delta^3 y_{n-3} + \frac{4t^3+18t^2+22t+6}{4!} \Delta^4 y_{n-4} + \frac{5t^4+40t^3+105t^2+100t+24}{5!} \Delta^5 y_{n-5} + \dots \right) \quad (5.64)$$

$$\begin{aligned}
& + \dots + \frac{n^{n-1} + (n-1)S_{n-1}^{(1)}t^{n-2} + (n-2)S_{n-1}^{(2)}t^{n-3} + \dots + S_{n-1}^{(n-1)}}{n!} \Delta^n y_0, \\
f'(x) \approx P'_n(x) &= \frac{1}{h^2} (\Delta^2 y_{n-2} + (t+1)\Delta^3 y_{n-3} + \\
& + \frac{6t^2 + 18t + 11}{12} \Delta^4 y_{n-4} + \frac{2t^3 + 12t^2 + 21t + 10}{12} \Delta^5 y_{n-5} + \quad (5.65) \\
& + \dots + \frac{n(n-1)t^{n-2} + (n-1)(n-2)S_{n-1}^{(1)}t^{n-3} + \dots + 2S_{n-1}^{(n-2)}}{n!} \Delta^n y_0).
\end{aligned}$$

Аналогічно знаходять і похідні вищих порядків. Якщо в формулах (5.64) і (5.65) покласти $t = 0$ ($x = x_n$), то дістанемо формули чисельного диференціювання функції f для табличних значень аргументу

$$\begin{aligned}
f'(x_n) \approx P'_n(x_n) &= \frac{1}{h} (\Delta y_{n-1} + \frac{1}{2} \Delta^2 y_{n-2} + \frac{1}{3} \Delta^3 y_{n-3} + \\
& + \frac{1}{4} \Delta^4 y_{n-4} + \frac{1}{5} \Delta^5 y_{n-5} + \dots + \frac{1}{n} \Delta^n y_0), \quad (5.66)
\end{aligned}$$

$$\begin{aligned}
f''(x_n) \approx P''_n(x_n) &= \frac{1}{h^2} (\Delta^2 y_{n-2} + \Delta^3 y_{n-3} + \frac{11}{12} \Delta^4 y_{n-4} + \\
& + \frac{5}{6} \Delta^5 y_{n-5} + \frac{137}{180} \Delta^6 y_{n-6} + \dots + \frac{2S_{n-1}^{(n-2)}}{n!} \Delta^n y_0). \quad (5.67)
\end{aligned}$$

Залишкові члени $r(x) = R^{(i)}(x; f)$ для похідних $f^{(i)}(x)$ ($i = 1, 2, \dots, k$), побудованих за другою інтерполяційною формулою Ньютона, можна знайти аналогічно до того, як знаходили їх для похідних, побудованих за першою інтерполяційною формулою Ньютона. Але тут залишковий член другої інтерполяційної формули Ньютона

$$R(x; f) = \frac{f^{(n+1)}(\xi)}{(n+1)!} t(t+1)(t+2)\dots(t+n).$$

П р и к л а д 13. У точці $x = 1,6$ знайти першу і другу похідні від функції $y = \frac{1}{x}$, заданої табл. 5. 11.

Р о з в ' я з а н н я. Похідні обчислюватимемо за формулами (5.66) і (5.67), поклавши $h = 0,1$; $x_n = 1,6$; $\Delta y_4 = -0,04167$; $\Delta^2 y_3 = 0,00595$; $\Delta^3 y_2 = -0,00137$; $\Delta^4 y_1 = 0,00047$; $\Delta^5 y_0 = -0,00019$.

Виконавши обчислення, дістанемо

$$\begin{aligned}
f'(1,6) &= 10(-0,04167 + \frac{1}{2} \cdot 0,00595 + \frac{1}{3}(-0,00137) + \frac{1}{4} \cdot 0,00047 + \\
& + \frac{1}{5}(-0,00019) = -10 \cdot (0,04167 - 0,002975 + 0,0004567 - \\
& - 0,0000925 + 0,00004) = -10 \cdot 0,0390992 = -0,390992;
\end{aligned}$$

$$\begin{aligned}
f''(1,6) &= 100(0,00595 + (-0,00137) + \frac{11}{12} \cdot 0,00047 + \frac{5}{6}(-0,00019) = \\
& = 100 \cdot (0,00595 - 0,00137 + 0,0003392 - 0,0001583) = \\
& = 100 \cdot 0,0047609 = 0,47609.
\end{aligned}$$

Оскільки значення різниць дано з 2-3 значущими цифрами, то наближені значення похідних $f'(1,6) \approx 0,39$, $f''(1,6) \approx 0,48$. Для порівняння подамо точніші значення цих похідних, обчислених на ЕКОМ: $f'(1,6) \approx 0,390625$; $f''(1,6) \approx 0,48828125$.

Покажемо, як впливає на значення похідних функції f у точці $x_n = -1,6$ кількість доданків у правих частинах наближених рівностей (5.66) і (5.67). Якщо їх один, два, три (а для f' і чотири), то дістаємо відповідно такі значення:

$$\begin{aligned}
f'(1,6) &= -0,4167, \quad f''(1,6) = -0,38695, \quad f'''(1,6) = -0,391517, \\
f'(1,6) &= -0,390592; \quad f''(1,6) = 0,595, \quad f'''(1,6) = 0,458, \\
f'(1,6) &= 0,49192.
\end{aligned}$$

Щоб обчислити значення похідних точніше, треба табличні значення функції обчислити з більшою кількістю цифр, а в формулах (5.66) і (5.67) взяти більшу кількість доданків.

Формули чисельного диференціювання, побудовані за інтерполяційною формулою Лагранжа. Якщо для функції f , заданої таблично в $(n+1)$ -му вузлі: $f(x_i) = f_i$ ($i = 0, 1, \dots, n$), побудовано інтерполяційний многочлен Лагранжа $L_n(x)$, то наближено

$$f'(x) = L'_n(x), \quad f''(x) = L''_n(x), \dots, \quad f^{(k)}(x) = L_n^{(k)}(x).$$

Нехай тепер функцію f задано таблично в рівновіддалених вузлах: $f(x_i) = f_i$ ($i = 0, 1, \dots, n$) з кроком $h = x_{i+1} - x_i$ ($i = 0, 1, \dots, n-1$). За цією системою точок (x_i, f_i) ($i = 0, 1, \dots, n$) побудуємо інтерполяційний многочлен Лагранжа. Тоді

$$f(x) \approx L_n(x) = \sum_{i=0}^n \frac{\omega_{n+1}(x)}{(x-x_i)\omega'_{n+1}(x_i)} f_i, \quad (5.68)$$

де $\omega_{n+1}(x) = (x-x_0)(x-x_1)\dots(x-x_n)$.

Перейдемо в многочлені Лагранжа до нової змінної t , поклавши $\frac{x-x_0}{h} = t$. Тоді

$$\begin{aligned}
x - x_i &= x - x_0 + x_0 - x_i = x - x_0 - (x_i - x_0) = ht - ih = h(t - i), \\
\omega_{n+1}(x) &= h^{n+1}t(t-1)\dots(t-n), \\
\omega'_{n+1}(x_i) &= (x_i - x_0)(x_i - x_1)\dots(x_i - x_{i-1})(x_i - x_{i+1})\dots(x_i - x_n) = \\
&= h^n i(t-1)(t-2)\dots 2 \cdot 1 \cdot (-1)(-2)\dots(-(n-i)) = \\
&= (-1)^{n-i} h^n i! (n-i)!, \quad (5.69)
\end{aligned}$$

а інтерполяційна формула Лагранжа (5.68) набере вигляду

$$f(x) \approx L_n(x) = \sum_{i=0}^n \frac{(-1)^{n-i}}{i!(n-i)!} \frac{i(t-1)(t-2)\dots(t-n)}{t-i} f_i \quad (5.70)$$

Звідси, врахувавши, що

$$\frac{dL_n(x)}{dx} = \frac{dL_n(x)}{dt} \cdot \frac{dt}{dx} = \frac{1}{h} \frac{dL_n(x)}{dt}$$

послідовно з (5.70) дістанемо

$$f'(x) \approx L'_n(x) = \frac{1}{h} \sum_{i=0}^n \frac{(-1)^{n-i}}{i!(n-i)!} \frac{d}{dt} \left(\frac{i(t-1)\dots(t-n)}{t-i} \right) f_i$$

$$f''(x) \approx L''_n(x) = \frac{1}{h^2} \sum_{i=0}^n \frac{(-1)^{n-i}}{i!(n-i)!} \frac{d^2}{dt^2} \left(\frac{i(t-1)\dots(t-n)}{t-i} \right) f_i$$

$$f^{(k)}(x) \approx L_n^{(k)}(x) = \frac{1}{h^k} \sum_{i=0}^n \frac{(-1)^{n-i}}{i!(n-i)!} \frac{d^k}{dt^k} \left(\frac{i(t-1)\dots(t-n)}{t-i} \right) f_i$$

Для оцінки похибки чисельного диференціювання

$$r_k(x) = f^{(k)}(x) - L_n^{(k)}(x)$$

скористаємося формулою (5.13) залишкового члена інтерполяційної формули Лагранжа

$$R(x; f) = \frac{\omega_{n+1}(x)}{(n+1)!} f^{(n+1)}(\xi)$$

де $\xi = \xi(x)$ лежить між точками x_0, x_1, \dots, x_n і x .

Коли припустити, що функція f має неперервні похідні порядку $(n+1+k)$, де $k \leq n$, то для залишкового члена наближеної рівності $f'(x) \approx L'_n(x)$, наприклад, знаходимо

$$r_1(x) = f'(x) - L'_n(x) = R'(x; f) = \frac{1}{(n+1)!} \left(\omega'_{n+1}(x) f^{(n+1)}(\xi) + \omega_{n+1}(x) \frac{d}{dx} f^{(n+1)}(\xi) \right)$$

Звідси, врахувавши формули (5.69), для похибки похідної у вузлах інтерполювання дістанемо

$$r_1(x) = R'(x; f) = \frac{(-1)^{n-i}}{(n+1)!} i!(n-i)! h^2 f^{(n+1)}(\xi)$$

де ξ — проміжне значення між x_0, x_1, \dots, x_n

Тема. Інтерполювання функцій.

Завдання. 1. Для таблично заданої функції (табл. 5.18) побудувати інтерполяційний многочлен Лагранжа і обчислити його значення в точках x_1, x_2, x_3

2. Оцінити похибку побудованого многочлена в одній з точок x_1, x_2, x_3 .

3. Обчислити точне значення заданої функції у вузлах x_1, x_2, x_3 і порівняти фактичну похибку з теоретичною.

4. Обчислити значення функції в точках x_1, x_2, x_3 за допомогою кубічних сплайнів та порівняти фактичні похибки інтерполювання функцій многочленом Лагранжа та кубічним сплайном.

Для інтерполювання в табл. 5.18 взято такі функції:

- sin x — для варіантів 1, 6, 11, 16, 21, 26;
- cos x — " " 2, 7, 12, 17, 22, 27;
- exp x — " " 3, 8, 13, 18, 23, 28;
- exp(-x) — " " 4, 9, 14, 19, 24, 29;
- sh x — " " 5, 10, 15, 20, 25, 30.

Таблиця 5.18

i	Варіант x(i)	1	2	3	4	5
0	.1	.0998334	.995004	1.10517	.904837	.100167
1	.5	.479426	.877583	1.64872	.606531	.521095
2	.8	.717356	.696707	2.22554	.449329	.888106
3	1.3	.963558	.267499	3.6693	.272532	1.69838
4	1.8	.973848	-.227202	6.04965	.165299	2.94217
5	2.6	.515501	-.856889	13.4637	.0742736	6.69473
x1 = .2, x2 = 1.4, x3 = 2.5						
i	Варіант x(i)	6	7	8	9	10
0	.1	.0998334	.995004	1.10517	.904837	.100167
1	.6	.564642	.825336	1.82212	.548812	.636654
2	1.2	.932039	.362358	3.32012	.301194	1.50946
3	1.8	.973848	-.227202	6.04965	.165299	2.94217
4	2.6	.515501	-.856889	13.4637	.0742736	6.69473
5	3	.14112	-.989992	20.0855	.0497871	10.0179
x1 = .3, x2 = 1.5, x3 = 2.8						
i	Варіант x(i)	11	12	13	14	15
0	.3	.29552	.955336	1.34986	.740818	.30452
1	.9	.783327	.62161	2.4596	.40657	1.02652
2	1.5	.997495	.0707372	4.48169	.22313	2.12928
3	2	.909297	-.416147	7.38906	.135335	3.62686
4	2.5	.598472	-.801144	12.1825	.082085	6.0502
5	3.1	.0415807	-.999135	22.198	.0450492	11.0765
x1 = .5, x2 = 1.7, x3 = 2.9						

i	Вариант		16	17	18	19	20
	x(i)	f(x)					
0	1		.841471	.540302	2.71828	.367879	1.1752
1	1.5		.997495	.0707372	4.48169	.22313	2.12928
2	2.3		.745705	-.666276	9.97418	.100259	4.93696
3	3		.14112	-.989992	20.0855	.0497871	10.0179
4	3.6		-.44252	-.896758	36.5982	.0273237	18.2855
5	4.5		-.97753	-.210796	90.0171	.011109	45.003
x1 = 1.2, x2 = 2.7, x3 = 4.3							
i	Вариант		21	22	23	24	25
	x(i)	f(x)					
0	1		.841471	.540302	2.71828	.367879	1.1752
1	1.6		.999574	-.0291995	4.95303	.201897	2.37557
2	2.5		.598472	-.801144	12.1825	.082085	6.0502
3	3.1		.0415807	-.999135	22.198	.0450492	11.0765
4	3.8		-.611858	-.790968	44.7012	.0223708	22.3394
5	4.5		-.97753	-.210796	90.0171	.011109	45.003
x1 = 1.3, x2 = 2.7, x3 = 4.3							
i	Вариант		26	27	28	29	30
	x(i)	f(x)					
0	1		.841471	.540302	2.71828	.367879	1.1752
1	1.5		.997495	.0707372	4.48169	.22313	2.12928
2	2.4		.675463	-.737394	11.0232	.090718	5.46623
3	3		.14112	-.989992	20.0855	.0497871	10.0179
4	3.9		-.687766	-.725932	49.4024	.0202419	24.6911
5	4.5		-.97753	-.210796	90.0171	.011109	45.003
x1 = 1.2, x2 = 2.8, x3 = 4.3							

Розділ 6. ЧИСЕЛЬНЕ ІНТЕГРУВАННЯ ФУНКЦІЙ

§ 6.1. ПОСТАНОВКА ЗАДАЧІ

Якщо функція f неперервна на відрізку $[a; b]$ і відома її первісна $F (F'(x) = f(x))$, то справедлива формула Ньютона—Лейбніца

$$\int_a^b f(x) dx = F(b) - F(a). \quad (6.1)$$

Проте цією формулою важко і навіть практично неможливо скористатися тоді, коли первісну F не можна виразити в елементарних функціях, як, наприклад, у таких інтегралах $\int_0^1 e^{-x^2} dx$, $\int_0^1 \frac{\sin x}{x} dx$,

$\int_0^1 \frac{dx}{\sqrt{1+x^3}}$, $\int_0^\pi \sqrt{1 - \varepsilon \sin^2 t} dt$ ($0 \leq \varepsilon < 1$); підінтегральну функцію f задано таблично або графічно і її аналітичний вираз невідомий; аналітичний вираз первісної F досить складний і незручний для обчислень.

У цих випадках треба будувати формули для наближеного обчислення визначених інтегралів. Особливо важливе значення мають методи чисельного інтегрування функцій, в яких для знаходження наближеного значення визначеного інтеграла використовуються значення підінтегральної функції та її похідних у скінченній кількості точок, що належать переважно проміжку інтегрування. Такі формули обчислення наближеного значення визначених інтегралів називають *формулами механічних квадратур*, або *квадратурними формулами*.

Найчастіше застосовують квадратурні формули, в яких використовуються значення підінтегральної функції f в окремих точках відрізка інтегрування, тобто формули вигляду

$$\int_a^b f(x) dx \approx \sum_{k=1}^n A_k f(x_k). \quad (6.2)$$

Суму в правій частині (6.2) називають *квадратурною сумою*, дійсні числа x_k і A_k відповідно *вузлами* і *коефіцієнтами* квадратурної формули. Вважатимемо, що вузли квадратурної формули (6.2) пронумеровано в порядку зростання $x_1 < x_2 < \dots < x_n$.

Рівність (6.2) наближена. Різницю між визначеним інтегралом і квадратурною сумою

$$R_n(f) = \int_a^b f(x) dx - \sum_{k=1}^n A_k f(x_k)$$

називають *залишковим членом*, або *похибкою* квадратурної формули (6.2).

Крім похибки, яка виникає від заміни інтеграла квадратурною сумою (*похибка методу* $|R_n(f)|$), є похибка, яка зумовлена виконанням арифметичних дій над наближеними числами — значеннями $f(x_k)$. Якщо абсолютні похибки значень $f(x_k)$ дорівнюють Δ_f , то абсолютна похибка квадратурної суми $\sum_{k=1}^n A_k f(x_k)$ дорівнюватиме

$$\tilde{R} = \Delta_f \sum_{k=1}^n |A_k|.$$

Це так звана *неусувна похибка*, яка зумовлена наближеними значеннями $f(x_k)$. У процесі обчислень виникає ще похибка за рахунок округлення проміжних результатів. Цю похибку можна зробити значно меншою порівняно з неусувною, якщо проміжні обчислення виконувати із запасними цифрами, які відкидають в остаточному результаті. Оцінюючи похибки чисельного інтегрування, треба враховувати також і *похибку остаточного округлення* Δ_0 . Отже, повна похибка чисельного інтегрування Δ_1 дорівнює сумі названих вище трьох похибок, тобто

$$\Delta_1 = |R_n(f)| + \tilde{R} + \Delta_0 = |R(f)| + \Delta_f \sum_{k=1}^n |A_k| + \Delta_0. \quad (6.3)$$

Для побудови квадратурних формул виду (6.2) часто вдаються до параболічного інтерполювання підінтегральної функції f . Для цього на проміжку $[a; b]$ вибирають скінченну послідовність точок $x_0, x_1, x_2, \dots, x_n$ і будують інтерполяційний многочлен Лагранжа

$$L_n(x) = \sum_{k=0}^n \frac{\omega_{n+1}(x)}{(x - x_k)\omega'_{n+1}(x_k)} f(x_k),$$

де $\omega_{n+1}(x) = (x - x_0)(x - x_1)\dots(x - x_n)$.
Тоді

$$f(x) = \sum_{k=0}^n \frac{\omega_{n+1}(x)}{(x - x_k)\omega'_{n+1}(x_k)} f(x_k) + R_n(f, x),$$

де $R_n(f, x)$ — залишковий член (похибка) інтерполювання.

Проінтегрувавши останню рівність по x у межах від a до b , дістанемо

$$\int_a^b f(x) dx = \sum_{k=0}^n A_k f(x_k) + R_n(f), \quad (6.4)$$

де

$$A_k = \int_a^b \frac{\omega_{n+1}(x)}{(x - x_k)\omega'_{n+1}(x_k)} dx \quad (k = 0, 1, \dots, n), \quad (6.5)$$

$$R_n(f) = \int_a^b R_n(f; x) dx. \quad (6.6)$$

Якщо залишковий член $R_n(f, x)$ інтерполювання функції f досить малий на всьому проміжку $[a; b]$, то у формулі (6.4), яка є точною рівністю, доданком $R_n(f)$ можна знехтувати. Тоді дістанемо наближену рівність

$$\int_a^b f(x) dx \approx \sum_{k=0}^n A_k f(x_k). \quad (6.7)$$

Квадратурну формулу (6.7), коефіцієнти якої обчислюють за формулами (6.5), називають *інтерполяційною*. Коефіцієнти інтерполяційних квадратурних формул залежать лише від вибору вузлів інтерполяції, але не від вигляду підінтегральної функції.

Справедлива теорема, яку подамо без доведення.

Теорема. Щоб квадратурна формула (6.7) з $n+1$ вузлом була інтерполяційною, необхідно й достатньо, щоб вона була точною, коли f є многочленом степеня, не вищого за n .

Якщо межі інтегрування a і b є вузлами інтерполювання, то квадратурну формулу (6.7) називають формулою *замкнутого типу*, в протилежному разі — *відкритого типу*.

З формули (6.5) для обчислення коефіцієнтів A_k квадратурної формули (6.7) видно, що значення коефіцієнтів A_k не залежать від вибору підінтегральної функції f , а залежать лише від вибору вузлів x_k ($k = 0, 1, \dots, n$). Обчисливши значення коефіцієнтів A_k ($k = 0, 1, \dots, n$) один раз, формулою (6.7) можна користуватися для обчислення наближених значень визначених інтегралів різних функцій f .

Формула (6.7) точна для многочлена n -го степеня, бо тоді $f(x) \equiv L_n(x)$ (похибка інтерполювання $R_n(f, x) = 0$). Зокрема, формула (6.7) точна, коли $f(x) = x^k$ ($k = 0, 1, \dots, n$). Проте, $R_n(x^{n+1}) \neq 0$.

Кажуть, що квадратурна формула має *алгебраїчний степінь точності* n , якщо вона точна для $f(x) = x^k$ ($k = 0, 1, \dots, n$) (або, що те саме, для будь-якого многочлена степеня, не вищого за n), і не дає точного результату для $f(x) = x^{n+1}$. Отже, інтерполяційні квадратурні формули з $n+1$ вузлом мають алгебраїчний степінь точності, не менший за n .

§ 6.2. КВАДРАТУРНІ ФОРМУЛИ НЬЮТОНА—КОТЕСА

До інтерполяційних квадратурних формул належать формули Ньютона—Котеса. Вузли цих формул рівновіддалені. Для їх побудови відрізок $[a; b]$, на якому визначено функцію f , ділять на n рівних частин завдовжки $h = \frac{b-a}{n}$ точками $x_0 = a$, $x_k = x_0 + kh$, $x_n = b$, обчислюють значення функції f у цих точках і будують інтерполяційний многочлен Лагранжа $L_n(x)$, значення якого в точках x_k ($k = 0, 1, \dots, n$) дорівнюють значенням функції f у цих точках. Замінивши в інтегралі $\int_a^b f(x) dx$ функцію f многочленом $L_n(x)$, дістають квадратурну формулу (6.7), коефіцієнти A_k якої обчислюють за формулою (6.5).

В інтегралі (6.5) переходять до нової змінної q за формулою $x = x_0 + qh$. Тоді $dx = hdq$ і нові межі інтегрування будуть: якщо $x = a$, то $q = 0$, а якщо $x = b = x_0 + hn$, то $q = n$. Виразимо тепер через q $x - x_k$, $\omega'_{n+1}(x_k)$ і $\omega_{n+1}(x)$. Маємо:

$$x - x_k = x - x_0 + x_0 - x_k = hq - kh = h(q - k);$$

$$\begin{aligned} \omega'_{n+1}(x_k) &= (x_k - x_0)(x_k - x_1)\dots(x_k - x_{k-1})(x_k - x_{k+1})\dots(x_k - x_n) = \\ &= kh(k-1)h\dots h(-h)\dots(-h(n-k)) = (-1)^{n-k} h^n k!(n-k)!; \end{aligned}$$

$$\begin{aligned} \omega_{n+1}(x) &= (x - x_0)(x - x_1)(x - x_2)\dots(x - x_n) = \\ &= (x - x_0)(x - x_0 + x_0 - x_1)(x - x_0 + x_0 - x_2)\dots(x - x_0 + x_0 - x_n) = \\ &= hq h(q-1)h(q-2)\dots h(q-n) = h^{n+1} q(q-1)(q-2)\dots(q-n). \end{aligned}$$

Підставивши ці значення в інтеграл (6.5), знайдемо

$$A_k = h \frac{(-1)^{n-k}}{k!(n-k)!} \int_0^n \frac{q(q-1)(q-2)\dots(q-n)}{q-k} dq \quad (k = 0, 1, \dots, n).$$

Але $h = \frac{b-a}{n}$, тому

$$A_k = (b-a) \frac{(-1)^{n-k}}{n \cdot k!(n-k)!} \int_0^n q(q-1)\dots(q-k+1)(q-k-1)\dots(q-n) dq \quad (k = 0, 1, \dots, n).$$

Коефіцієнти

$$H_k = \frac{(-1)^{n-k}}{n \cdot k!(n-k)!} \int_0^n q(q-1)\dots(q-k+1)(q-k-1)\dots(q-n) dq \quad (k = 0, 1, \dots, n)$$

називаються коефіцієнтами Котеса, який вперше підрахував їх значення для n від 1 до 10.

Отже, квадратурні формули Ньютона—Котеса набирають вигляду

$$\int_a^b f(x) dx \approx (b-a) \sum_{k=0}^n H_k f(x_k), \quad (6.8)$$

причому рівновіддалені від кінців відрізка інтегрування коефіцієнти Котеса рівні між собою, тобто $H_k = H_{n-k}$, і $\sum_{k=0}^n H_k = 1$. Доведено, що для

$n \geq 10$ (n — число вузлів) серед коефіцієнтів Котеса завжди є від'ємні. Тоді малі похибки в значеннях функції $f(x_k)$ можуть призвести до великих похибок у квадратурній сумі. Саме тому формули Ньютона—Котеса з великою кількістю вузлів мало придатні для обчислень і їх на практиці не використовують. Проте формули з 2-3 вузлами широко застосовують.

У наступних параграфах розглянуто окремі випадки застосування формул Ньютона—Котеса (6.8). Через важливість цих формул їх коефіцієнти і залишкові члени обчислимо незалежно від формули (6.8).

§ 6.3. ФОРМУЛА ПРЯМОКУТНИКІВ

Наближене значення інтеграла $\int_{x_0}^{x_0+h} f(x) dx$, де f неперервна на $[x_0; x_0+h]$, можна знайти, якщо функцію f замінити інтерполяційним многочленом нульового степеня, тобто для всіх $x \in [x_0, x_0+h]$ покласти $f(x) = f(c)$, де $c \in [x_0, x_0+h]$. Тоді дістанемо наближену рівність

$$\int_{x_0}^{x_0+h} f(x) dx \approx h f(c). \quad (6.9)$$

Якщо $f(x) \geq 0$ і неперервна на $[x_0; x_0+h]$, то наближену рівність (6.9) можна тлумачити геометрично так: за наближене значення площі криволінійної трапеції $ABCD$ (рис. 6.1), обмеженої знизу віссю абсцис, зверху графіком функції f , а з боків прямими $x = x_0$ і $x = x_0+h$, береться значення площі прямокутника $MNCD$. Тому формула (6.9) дістала назву *формули прямокутників*.

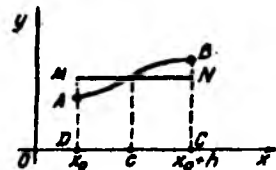


Рис. 6.1

Якщо $c = x_0$ або $c = x_0 + h$, або $c = x_0 + \frac{1}{2}h$, то (6.9) називають відповідно *формулою лівих* або *правих*, або *середніх прямокутників*. Знайдемо залишкові члени формул прямокутників, припустивши,

що підінтегральна функція f на $[x_0; x_0+h]$ має неперервну похідну першого порядку (для випадку формул лівих і правих прямокутників) і неперервну похідну другого порядку (для випадку формули середніх прямокутників).

Проінтегрувавши обидві частини формули Лагранжа

$$f(x) - f(x_0) = (x - x_0) f'(\xi_1), \quad x_0 \leq \xi_1 \leq x \leq x_0 + h$$

по x в межах від x_0 до $x_0 + h$, знайдемо

$$\int_{x_0}^{x_0+h} f(x) dx - f(x_0)h = \int_{x_0}^{x_0+h} (x - x_0) f'(\xi_1) dx.$$

Звідси залишковий член формули лівих прямокутників

$$R(f) = \int_{x_0}^{x_0+h} (x - x_0) f'(\xi_1) dx.$$

Але на відрізку $[x_0; x_0+h]$ множник $(x - x_0)$ зберігає знак, а функція $f'(\xi_1)$ неперервна, тому за узагальненою теоремою про середнє маємо

$$R_1(f) = \frac{h^2}{2} f'(\eta_1), \quad x_0 \leq \eta_1 \leq x_0 + h. \quad (6.10)$$

Аналогічно, проінтегрувавши по x у межах від x_0 до x_0+h обидві частини формули Лагранжа

$$f(x) - f(x_0 + h) = (x - x_0 - h) f'(\xi_2), \quad x_0 < x < \xi_2 \leq x_0 + h$$

і застосувавши до інтеграла $\int_{x_0}^{x_0+h} (x - x_0 - h) f'(\xi_2) dx$ узагальнену формулу

про середнє, для залишкового члена формули правих прямокутників знайдемо

$$R_1(f) = -\frac{h^2}{2} f'(\eta_2), \quad x_0 \leq \eta_2 \leq x_0 + h. \quad (6.11)$$

Якщо формулу Тейлора

$$f(x) = f\left(x_0 + \frac{1}{2}h\right) + \left(x - x_0 - \frac{h}{2}\right) f'\left(x_0 + \frac{h}{2}\right) + \frac{1}{2}\left(x - x_0 - \frac{h}{2}\right)^2 f''(\xi_3),$$

де ξ_3 — деяка проміжна точка, що лежить між x_0 і $x_0 + \frac{1}{2}h$ і залежить від x , проінтегрувати по x у межах від x_0 до x_0+h , то дістанемо

* Узагальнена теорема про середнє. Якщо функції f і g неперервні на $[a; b]$, а g зберігає знак на $[a; b]$ ($g(x) > 0$ або $g(x) < 0$), то існує $\xi \in [a; b]$ таке, що $\int_a^b f(x)g(x) dx =$

$$f(\xi) \int_a^b g(x) dx.$$

$$\int_{x_0}^{x_0+h} f(x) dx = hf\left(x_0 + \frac{h}{2}\right) + \frac{1}{2} \int_{x_0}^{x_0+h} \left(x - x_0 - \frac{h}{2}\right)^2 f''(\xi_3) dx.$$

Отже, залишковий член формули середніх прямокутників

$$R_2(f) = \frac{1}{2} \int_{x_0}^{x_0+h} \left(x - x_0 - \frac{h}{2}\right)^2 f''(\xi_3) dx.$$

Застосувавши до цього інтеграла узагальнену формулу про середнє, маємо

$$R_2(f) = \frac{h^3}{24} f''(\eta), \quad x_0 \leq \eta \leq x_0 + h. \quad (6.12)$$

Узагальнені формули прямокутників. Нехай тепер неперервну функцію f задано на великому проміжку $[a; b]$, а $\int_a^b f(x) dx$ бажано обчислити з більшою точністю. Для цього відрізок $[a; b]$ ділять на n рівних відрізків завдовжки $h = (b - a)/n$ точками

$$a = x_0 < x_1 < x_2 < \dots < x_k < x_{k+1} < \dots < x_n = b$$

і до кожного з відрізків $[x_k; x_{k+1}]$ ($k = 0, 1, \dots, n-1$) застосовують формулу (6.9). Дістають

$$\int_a^b f(x) dx = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(x) dx \approx h \sum_{k=0}^{n-1} f(c + kh),$$

де c — довільна точка з проміжку $[x_0; x_1]$.

Наближену рівність

$$\int_a^b f(x) dx \approx h \sum_{k=0}^{n-1} f(c + kh), \quad (6.13)$$

де $c \in [a; a+h]$, $h = (b-a)/n$ — крок інтегрування, називають узагальненою формулою прямокутників. Поклавши в (6.13) $c = a$, $c = a + \frac{1}{2}h$, можна вивести узагальнені формули відповідно лівих, правих і середніх прямокутників.

Обчислимо тепер залишкові члени узагальнених формул прямокутників, які дорівнюватимуть сумі залишкових членів формул прямокутників для кожного з проміжків $[x_k; x_{k+1}]$ ($k = 0, 1, \dots, n-1$).

Для залишкового члена узагальненої формули лівих прямокутників відповідно до формули (6.10) знаходимо

$$R_1(f) = \frac{h^2}{2} \sum_{k=0}^{n-1} f'(\eta_k) = \frac{h}{2} \cdot \frac{b-a}{n} \sum_{k=0}^{n-1} f'(\eta_k), \quad x_k < \eta_k < x_{k+1}.$$

Але $f(x)$ неперервна на $[a; b]$, тому існує точка $\xi_1 \in [a; b]$ така, що середнє арифметичне

$$\frac{1}{n} \sum_{k=0}^{n-1} f(\eta_k) = f(\xi_1).$$

Тому залишковий член узагальненої формули лівих прямокутників остаточно набирає вигляду

$$R_1(f) = \frac{b-a}{2} h f'(\xi_1), \quad \xi_1 \in [a; b]. \quad (6.14)$$

Аналогічно залишковий член узагальненої формули правих прямокутників

$$R_1(f) = -\frac{b-a}{2} h f'(\xi_2), \quad \xi_2 \in [a; b], \quad (6.15)$$

а залишковий член узагальненої формули середніх прямокутників

$$R_2(f) = \frac{b-a}{24} h^2 f''(\xi_3), \quad \xi_3 \in [a; b]. \quad (6.16)$$

Оскільки похідна від сталої і друга похідна від лінійної функції дорівнюють нулю, то з формул (6.14)–(6.16) випливає, що узагальнені формули лівих і правих прямокутників точні для сталої функції, а узагальнена формула середніх прямокутників точна для лінійної функції.

Обчислити значення залишкових членів (6.14)–(6.16) не можна, бо точки ξ_1, ξ_2, ξ_3 невідомі. Але для них справедливі такі оцінки

$$|R_1(f)| \leq h \frac{b-a}{2} M_1 = \frac{(b-a)^2}{2n} M_1, \quad (6.17)$$

$$|R_2(f)| \leq h^2 \frac{b-a}{24} M_2 = \frac{(b-a)^3}{24n^2} M_2, \quad (6.18)$$

де

$$M_1 = \max_{[a;b]} |f'(x)|, \quad M_2 = \max_{[a;b]} |f''(x)|.$$

П р и к л а д 1. Обчислити інтеграл

$$I = \int_0^1 x \cos x dx = \sin 1 + \cos 1 - 1 = 0,38177334 \quad (6.19)$$

за формулами прямокутників, поділивши відрізок інтегрування на $n = 10$ рівних частин, і оцінити похибку обчислень.

Р о з в ' я з а н н я. Крок інтегрування $h = \frac{1-0}{10} = 0,1$. Складаємо таблиці значень підінтегральної функції (табл. 6.1 і 6.2).

Таблиця 6.1

k	x_k	$f_k = f(x_k)$	k	x_k	$f_k = f(x_k)$
0	0	0	6	0,6	0,4952014
1	0,1	0,0995004	7	0,7	0,5353896
2	0,2	0,1960133	8	0,8	0,5573654
3	0,3	0,2866010	9	0,9	0,5594490
4	0,4	0,3684244	10	1,0	0,5403023
5	0,5	0,4387913			

Тоді за узагальненою формулою лівих прямокутників (6.13) з табл.6.1 маємо:

$$I_L = h \sum_{k=0}^9 f_k = 0,1 \cdot 3,5367358 = 0,35367358 \approx 0,353674$$

і за узагальненою формулою правих прямокутників:

$$I_R = h \sum_{k=1}^{10} f_k = 0,1 \cdot 4,0770381 = 0,40770381 \approx 0,407704.$$

Таблиця 6.2

k	$x_{k+\frac{1}{2}}$	$f_{k+\frac{1}{2}} = f\left(x_{k+\frac{1}{2}}\right)$	k	$x_{k+\frac{1}{2}}$	$f_{k+\frac{1}{2}} = f\left(x_{k+\frac{1}{2}}\right)$
0	0,05	0,0499375	5	0,55	0,4688885
1	0,15	0,1483157	6	0,65	0,5174545
2	0,25	0,2422281	7	0,75	0,5487667
3	0,35	0,3287805	8	0,85	0,5609857
4	0,45	0,4052012	9	0,95	0,5525989

З табл. 6.2 за формулою середніх прямокутників дістанемо:

$$I_{cp} = h \sum_{k=0}^9 f_{k+\frac{1}{2}} = 0,1 \cdot 3,8231573 = 0,38231573 \approx 0,382316.$$

Знайдемо тепер повну похибку чисельного інтегрування Δ , за цими квадратурними формулами, користуючись формулою (6.3).

Похибки методу $R(f)$ оцінюємо за формулами (6.17) і (6.18). Для цього знайдемо похідні

$$f'(x) = \cos x - x \sin x, \quad f''(x) = -2 \sin x - x \cos x.$$

Оскільки для всіх $x \in [0; 1]$ $f'(x) < 0$, то $f(x)$ спадає на $[0; 1]$, а тому найбільше значення f' має в точці $x = 0$. Отже,

$$M_1 = \max_{[0;1]} |f'(x)| = f(0) = 1,$$

$$M_2 = \max_{[0;1]} |f''(x)| \leq 2 |\sin x| + |x| |\cos x| = 2,3.$$

Неусувна похибка для всіх трьох формул прямокутників дорівнює

$$\tilde{R} = \Delta_f \sum_{k=1}^n |A_k| = \Delta_f \sum_{k=1}^n h = \Delta_f n h = \Delta_f (b - a).$$

Значення підінтегральної функції f у вузлах обчислювали з точністю $0,5 \cdot 10^{-7}$. Тому гранична абсолютна похибка $\Delta_f = 0,5 \cdot 10^{-7}$.

Похибка остаточного округлення Δ_0 дорівнює для лівих прямокутників $0,42 \cdot 10^{-6}$, для правих $0,19 \cdot 10^{-6}$ і середніх $0,27 \cdot 10^{-6}$.

З повної абсолютної похибки чисельного інтегрування Δ_I , яку подано в табл. 6. 3, видно, що наближені значення інтеграла (6.19), які обчислено за формулами лівих і правих прямокутників, мають одну правильну значущу цифру, а середніх прямокутників — дві. В останньому стовпці наведено значення модуля різниці між точним і наближеним значеннями інтеграла, яке вдвічі менше за повну абсолютну похибку Δ_I . З поданих у табл. 6. 3 даних видно, що основною у повній

Таблиця 6. 3

Узагальнена формула прямокутників	$R(f)$	\tilde{R}	Δ_0	Δ_I	$ I - I_{\text{нобл}} $
лівих	0,05	$0,5 \cdot 10^{-7}$	$0,42 \cdot 10^{-6}$	0,05000047	0,0280997
правих	0,05	$0,5 \cdot 10^{-7}$	$0,19 \cdot 10^{-6}$	0,05000024	0,0259305
середніх	0,00095833	$0,5 \cdot 10^{-7}$	$0,27 \cdot 10^{-6}$	0,00095865	0,0005425

абсолютній похибці чисельного інтегрування є похибка методу $R(f)$. Тому надалі при визначенні кроку інтегрування, який гарантував би обчислення наближеного значення інтеграла із наперед заданою точністю, досить виходити з оцінки залишкового члена $R(f)$, і всі проміжні обчислення виконувати з більшою, ніж ϵ , точністю, тобто з однією, а краще з двома запасними цифрами.

П р и м і т к а. Залишкові члени формул (6.14) і (6.15) лівих і правих прямокутників мають протилежні знаки. Отже, ці формули наближають інтеграл $I = \int_a^b f(x) dx$ з недостачею і надлишком. Тому за наближене значення інтеграла I можна взяти півсуму цих двосторонніх

наближень, поклавши $I \approx I^* = \frac{I_n + I_n}{2}$. Тоді для абсолютної похибки наближення I^* дістанемо $|I - I^*| \leq |I_n - I_n| / 2$. Так для інтеграла (6.19) маємо: $I^* = 0,5 \cdot (0,407704 + 0,353674) = 0,380689$, $|I - I^*| \leq 0,5 \cdot |I_n - I_n| = 0,027015$. Насправді $|I - I^*| = |0,381773 - 0,380689| = 0,001084$, тобто I^* має дві правильні значущі цифри, хоча I_n і I_n — лише одну.

Якщо наближене значення $\int_a^b f(x) dx$ треба обчислити з наперед заданою точністю $\epsilon > 0$, то відповідно до формул (6.17) і (6.18) відрізок $[a; b]$ треба поділити на n рівних відрізків так, щоб виконувалися нерівності

$$\frac{M_1(b-a)^2}{2n} < \epsilon \Leftrightarrow n > \frac{M_1(b-a)^2}{2\epsilon}, \quad (6.20)$$

$$\frac{M_2(b-a)^3}{24n^2} < \epsilon \Leftrightarrow n > \sqrt{\frac{M_2(b-a)^3}{24\epsilon}}. \quad (6.21)$$

Наприклад, щоб обчислити значення інтеграла (6.19) за узагальненою формулою середніх прямокутників з точністю $\epsilon = 0,5 \cdot 10^{-4}$, тобто з чотирма правильними десятковими знаками, треба відрізок $[0; 1]$ поділити не менш як на 44 рівних частини. Справді, з формули (6.21) маємо

$$n > \sqrt{\frac{M_2(b-a)^3}{24\epsilon}} = \sqrt{\frac{2,3 \cdot 10^4}{24 \cdot 0,5}} = 10^2 \sqrt{\frac{2,3}{12}} \approx 10^2 \cdot 0,438 \approx 44.$$

Якщо $n = 50$, то $h = 0,02$ і наближене значення інтеграла $I_{50} = 0,38179504$, чотири десяткових знаки якого збігаються з відповідними знаками точного розв'язку.

Але формулами (6.20) і (6.21) часто важко або й неможливо скористатися тоді, коли функцію f задано графічно або таблично і її аналітичний вираз невідомий, а також і тоді, коли функцію f задано таким складним виразом, що важко знайти значення M_1 і M_2 .

У цьому разі інтеграл обчислюють двічі: з кроком h і $h/2$, тобто відрізок $[a; b]$ ділять на n і $2n$ рівні частини. При цьому вважають, що точному значенню інтеграла належать ті десяткові знаки, які є однаковими в I_n і $I_{n/2}$. Наприклад, обчисливши інтеграл (6.19) за формулою середніх прямокутників, коли $n = 25, 50, 100$ і 200 , дістанемо $I_{25} = 0,38186606$; $I_{50} = 0,38179504$; $I_{100} = 0,38177874$; $I_{200} = 0,38177472$. Отже, I_{25} має три правильних значущих десяткових знаки, I_{50} — чотири, I_{100} — п'ять, бо ці знаки збігаються в двох послідовних наближеннях I_n і I_{2n} .

Ще один зручний для практики спосіб обчислення похибки точного інтегрування можна застосувати, використавши правило Рунге. У цьому

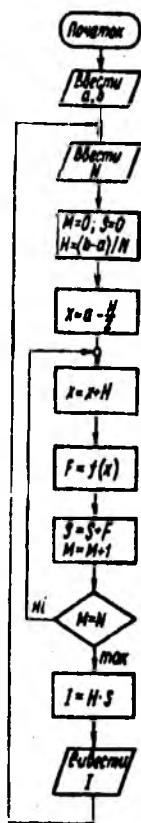


Рис. 6.2

разі відрізок $[a; b]$ ділять на n і $2n$ рівні частини і обчислюють I_1 та I_2 . Докладніше про це йдеться у § 6.6. На рис. 6.2 подано схему алгоритму чисельного інтегрування функцій за узагальненою формулою середніх прямокутників. Відповідною програмою для інтеграла (6.19) мовою Бейсік є програма 6.1. Ця програма дає змогу обчислити значення інтеграла з кроками h і $h/2$, що відповідають поділу відрізка інтегрування $[a; b]$ на n і $2n$ рівні частини. Обчислення за програмою припиняють, якщо відповідної точності $\epsilon > 0$ досягнуто, тобто, коли в двох послідовних наближеннях збігаються десяткові знаки, що відповідають заданій точності $\epsilon > 0$.

Програма 6.1

```

10 PRINT "---- Чисельне інтегрування методом"----
20 PRINT "---- середніх прямокутників"-----
30 INPUT "Ввести нижню межу інтегрування A="; A
40 INPUT "Ввести верхню межу інтегрування B="; B
50 DEF FNY(X)=X+COS(X)
60 INPUT "Ввести число інтервалів інтегрування
   N="; N
70 M=0: S=0: H=(B-A)/N: X=A-H/2
80 X=X+H
90 F=FNY(X)
100 S=S+F: M=M+1
110 IF M<>N THEN 80
120 I=S*H
130 PRINT "Для A="A; "B="B; "N="N
140 PRINT "Значення інтеграла I="I
150 GOTO 60
160 END

```

§ 6.4. ФОРМУЛА ТРАПЕЦІЙ

Для обчислення наближеного значення $I = \int_{x_0}^{x_0+h} f(x) dx$ підінтегральну функцію f , яка двічі неперервно диференційовна на відрізку $[x_0; x_0+h]$ ($h > 0$), замінюють інтерполяційним многочленом Лагранжа, що проходить через точки $(x_0; f(x_0))$ і $(x_0+h; f(x_0+h))$:

$$f(x) = \frac{x - x_0 - h}{x_0 - (x_0 + h)} f(x_0) + \frac{x - x_0}{x_0 + h - x_0} f(x_0 + h) +$$

$$+ \frac{1}{2} f''(\xi)(x - x_0)(x - x_0 - h), \quad x_0 < \xi < x_0 + h.$$

Проінтегрувавши цю рівність по x у межах від x_0 до $x_0 + h$, дістанемо

$$\int_{x_0}^{x_0+h} f(x) dx = \frac{h}{2} (f(x_0) + f(x_0 + h)) + \frac{1}{2!} \int_{x_0}^{x_0+h} f''(\xi)(x - x_0)(x - x_0 - h) dx.$$

Відкинувши в цій рівності залишковий член

$$R_2(f) = \frac{1}{2!} \int_{x_0}^{x_0+h} f''(\xi)(x - x_0)(x - x_0 - h) dx, \quad (6.22)$$

дістанемо квадратурну формулу

$$\int_{x_0}^{x_0+h} f(x) dx \approx \frac{h}{2} (f(x_0) + f(x_0 + h)). \quad (6.23)$$

Якщо $f(x)$ неперервна невід'ємна функція на відрізку $[x_0; x_0 + h]$, то наближену рівність (6.23) можна геометрично тлумачити так: за наближене значення криволінійної трапеції $ABCD$ (рис. 6.3) береться площа заштрихованої трапеції $ABCD$. Тому формула дістала назву *формули трапецій*.



Рис. 6.3

Якщо $f''(x)$ неперервна на відрізку $[x_0; x_0+h]$, то, застосувавши узагальнену теорему про середнє до інтеграла (6.22), знайдемо залишковий член формули трапецій у вигляді

$$R_2(f) = -\frac{h^3}{12} f''(\eta), \quad x_0 \leq \eta \leq x_0 + h. \quad (6.24)$$

Звідси випливає така оцінка для абсолютної похибки чисельного інтегрування за формулою трапецій

$$|R_2(f)| \leq \frac{h^3}{12} M_2, \quad M_2 = \max_{[x_0; x_0+h]} |f''(x)|. \quad (6.25)$$

Щоб обчислити наближене значення інтеграла $\int_a^b f(x) dx$, де f неперервна разом з похідними першого і другого порядків на $[a; b]$, з достатньою точністю, відрізок $[a; b]$ ділять на n рівних відрізків завдовжки $h = (b - a) / n$ і до кожного з відрізків $[x_k, x_{k+1}]$, ($k = 0, 1, \dots, n-1$) застосовують формулу трапецій із залишковим членом (6.24). Тоді

$$\int_a^b f(x) dx = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(x) dx = \sum_{k=0}^{n-1} \frac{h}{2} (f(x_k) + f(x_{k+1})) - \frac{h^3}{12} \sum_{k=0}^{n-1} f''(\eta_k), \quad x_k \leq \eta_k \leq x_{k+1} \quad (k = 0, 1, \dots, n-1).$$

Відкинувши в цій рівності залишковий член $R_2(f) = -\frac{h^3}{12} \sum_{k=0}^{n-1} f''(\eta_k)$,

дістають квадратурну формулу

$$\int_a^b f(x) dx \approx \frac{h}{2} (f(a) + 2 \sum_{k=1}^{n-1} f(a + kh) + f(b)), \quad h = (b - a) / n, \quad (6.26)$$

яку називають *узагальненою формулою трапецій*.

Оскільки $f'(x)$ неперервна на відрізку $[a; b]$, то існує точка $\eta \in [a; b]$ така, що $\frac{1}{n} \sum_{k=0}^{n-1} f''(\eta_k) = f''(\eta)$ і залишковий член узагальненої формули трапецій набирає остаточно вигляду

$$R_2(f) = -\frac{h^2}{12} (b - a) f''(\eta), \quad a \leq \eta \leq b. \quad (6.27)$$

З (6.27) випливає, що узагальнена формула трапецій точна для лінійної функції, бо друга похідна від лінійної функції дорівнює нулю. Обчислити похибку чисельного інтегрування за узагальненою формулою трапецій не можна через те, що точка $\eta \in [a; b]$ формули (6.27) невідома. Але легко оцінити абсолютну похибку наближеного інтегрування за формулою (6.26)

$$|R_2(f)| \leq \frac{h^2}{12} (b - a) M_2 = \frac{(b - a)^3}{12n^2} M_2, \quad (6.28)$$

де $h = \frac{b - a}{n}$, $M_2 = \max_{[a; b]} |f''(x)|$.

Якщо наближене значення інтеграла треба обчислити з наперед заданою точністю $\epsilon > 0$, то досить відрізок $[a; b]$ поділити на n рівних частин так, щоб виконувалася нерівність

$$\frac{(b - a)^3}{12n^2} M_2 < \epsilon \iff n > \sqrt{\frac{(b - a)^3 M_2}{12\epsilon}}, \quad (6.29)$$

або, що те саме,

$$h < \sqrt{\frac{12\epsilon}{(b - a)M_2}}.$$

Обчислимо наближене значення інтеграла (6.19) за формулою (6.26), якщо $n = 10$, і оцінимо повну абсолютну похибку чисельного інтегрування за формулою (6.3).

Використавши табл. 6.1 за формулою (6.26) знаходимо $I_T = 0,38068869 \approx 0,380689$. За формулою (6.28) ($a = 0$; $b = 1$; $h = 0,1$; $M_2 = 2,3$) маємо: $|R(f)| \leq \frac{0,01}{12} \cdot 2,3 \approx 0,00191667$, похибка остаточного округлення $\Delta_0 = 0,31 \cdot 10^{-6}$, а неусувна похибка, як і у випадку формул прямокутників,

$$\bar{R} = \Delta_f \cdot \sum_{k=1}^n |A_k| = \Delta_f \cdot \frac{h}{2} \cdot 2n = nh \Delta_f = (b - a) \Delta_f = 0,5 \cdot 10^{-7}.$$

Тому за формулою (6.3) дістанемо таку оцінку для повної абсолютної величини похибки чисельного інтегрування за формулою (6.26):

$$\Delta_I = |R(f)| + \bar{R} + \Delta_0 = 0,00191703 < 0,002.$$

А це означає, що наближення I_T має дві правильні значущі десяткові цифри і тому, зберігаючи одну сумнівну цифру, маємо:

$$\int_0^1 x \cos x dx = 0,381 \pm 0,002.$$

З оцінки Δ_I випливає, що найбільший внесок у Δ_I залишкового члена (похибки методу) $|R(f)|$. Тому, для визначення кількості відрізків n , на яку треба поділити відрізок інтегрування $[a; b]$, щоб забезпечити обчислення наближеного значення інтеграла з точністю $\epsilon > 0$, досить скористатися нерівністю (6.29), а всі проміжні обчислення виконувати з точністю, більшою за ϵ . Так, для інтеграла (6.19), якщо $\epsilon = 0,5 \cdot 10^{-4}$, відрізок $[a; b]$ треба поділити не менш як на 62 рівні частини. Справді, за формулою (6.29) маємо

$$n > \sqrt{\frac{1 \cdot 2,3 \cdot 10^4}{12 \cdot 0,5}} = 10^2 \sqrt{\frac{2,3}{6}} \approx 0,619 \cdot 10^2 \approx 62.$$

Обчисливши значення інтеграла (6.19) за формулою трапецій (6.26), коли $n = 25, 50, 100, 200$, що відповідає значенням $h = 0,04; 0,02; 0,01; 0,005$, матимемо: $I_{25} = 0,38159981$; $I_{50} = 0,38172994$; $I_{100} = 0,38176250$; $I_{200} = 0,38177064$.

Отже, I_{25} має три точних значущих десяткових знаки, I_{50} і I_{100} по чотири. Звідси видно, що розбиття відрізка $[0; 1]$ як на 50, так і на 100 частин дає чотири правильних значущих десяткових цифри. А це означає, що обчислене за формулою (6.29) значення $n = 62$ завищено приблизно на 20 %.

У § 6.6 показано, що правило Рунге і екстраполяція за Річардсоном дають зручний спосіб практичної оцінки залишкового члена (похибки методу) $R(f)$ і уточнення наближеного значення інтеграла I .

П р и м і т к а 1. Квадратурна сума узагальненої формули трапецій дорівнює півсумі квадратурних сум узагальнених формул лівих і правих прямокутників. Справді,

$$I_n = h (f(x_0) + f(x_1) + \dots + f(x_{n-1})),$$

$$I_n = h (f(x_1) + \dots + f(x_{n-1}) + f(x_n)).$$

Звідки

$$I_{\text{тр}} = \frac{1}{2} (I_n + I_n) = \frac{h}{2} (f(x_0) + 2(f(x_1) + \dots + f(x_{n-1})) + f(x_n)).$$

Саме тому для наближеного значення інтеграла (6.19) одержано однакові значення $I \approx 0,380689$ за формулою (6.26), як і в примітці § 6.3.

П р и м і т к а 2. Залишкові члени узагальнених формул середніх прямокутників і трапецій (див. (6.16) і (6.27)) мають протилежні знаки.

Отже, ці формули наближують інтеграл $I = \int_a^b f(x) dx$ з недостачею і з надлишком, якщо $f'(x)$ зберігає знак на $[a; b]$. Тому за наближене значення інтеграла I можна взяти півсуму цих двосторонніх наближень, тобто покласти $I \approx 0,5(I_{\text{сп}} + I_{\text{тр}}) = I^*$. Тоді абсолютна похибка наближення I^*

$$|I - I^*| < \frac{1}{2} |I_{\text{сп}} - I_{\text{тр}}|.$$

Так, для інтеграла (6.19), коли $n = 25$, $I_{\text{сп}} = 0,38186006$; $I_{\text{тр}} = 0,38159981$, тобто мають по три правильних десяткових знаки. А їх півсума $I^* = 0,38172994$ має вже чотири правильних знаки.

§ 6.5. ФОРМУЛА СІМПСОНА

Щоб побудувати триточкову квадратурну формулу з рівновіддаленими вузлами для обчислення наближеного значення $\int_{x_0-h}^{x_0+h} f(x) dx$, де $f(x)$ — неперервна на $[x_0 - h; x_0 + h]$ разом зі своїми похідними до четвертого порядку включно, можна використати інтерполяційний многочлен Лагранжа 2-го порядку, графік якого проходить через точки $(x_0 - h; f(x_0 - h))$, $(x_0; f(x_0))$ і $(x_0 + h; f(x_0 + h))$ і проінтегрувати його по x у межах від $x_0 - h$ до $x_0 + h$.

Проте таку квадратурну формулу будуватимемо тут, користуючись методом невизначених коефіцієнтів. Цей метод, крім того, дає змогу досить просто обчислити її залишковий член. Отже, побудуємо квадратурну формулу вигляду

$$\int_{x_0-h}^{x_0+h} f(x) dx = 2h (A(f(x_0 - h) + f(x_0 + h)) + B f(x_0)) + R(f), \quad (6.30)$$

де A і B — невідомі коефіцієнти, а $R(f)$ — залишковий член.

Щоб дістати рівняння, з яких можна визначити коефіцієнти A і B , подамо функції $f(x)$, $f(x_0 - h)$ і $f(x_0 + h)$ в околі точки x_0 за допомогою формули Тейлора. Маємо

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2!} f''(x_0) + \frac{(x - x_0)^3}{3!} f'''(x_0) + \frac{(x - x_0)^4}{4!} f^{(4)}(x_0 + \theta h),$$

$$f(x_0 - h) = f(x_0) - h f'(x_0) + \frac{h^2}{2!} f''(x_0) - \frac{h^3}{3!} f'''(x_0) + \frac{h^4}{4!} f^{(4)}(x_0 - \theta_2 h),$$

$$f(x_0 + h) = f(x_0) + h f'(x_0) + \frac{h^2}{2!} f''(x_0) + \frac{h^3}{3!} f'''(x_0) + \frac{h^4}{4!} f^{(4)}(x_0 + \theta_3 h),$$

$$0 < \theta, \theta_2, \theta_3 < 1.$$

Підставляючи ці значення функцій $f(x)$, $f(x_0 - h)$, $f(x_0 + h)$ у (6.30) і беручи до уваги, що

$$\int_{x_0-h}^{x_0+h} f(x) dx = 2hf(x_0) + \frac{2h^3}{3!} f''(x_0) + \frac{2h^5}{5!} f^{(4)}(x_0 + \theta_1 h), \quad 0 < \theta_1 < 1,$$

(тут за загальною теоремою про середнє

$$\int_{x_0-h}^{x_0+h} \frac{(x - x_0)^4}{4!} f^{(4)}(x_0 + \theta h) dx = \frac{2h^5}{5!} f^{(4)}(x_0 + \theta_1 h), \quad 0 < \theta_1 < 1),$$

для залишкового члена $R(f)$ дістанемо:

$$R(f) = 2h \left((1 - 2A - B) f(x_0) + h^2 \left(\frac{1}{3!} - A \right) f''(x_0) + \frac{h^4}{4!} \left(\frac{1}{5} f^{(4)}(x_0 + \theta_1 h) - A (f^{(4)}(x_0 - \theta_2 h) + f^{(4)}(x_0 + \theta_3 h)) \right) \right).$$

Невідомі коефіцієнти A і B доберемо так, щоб

$$\begin{cases} 1 - 2A - B = 0, \\ \frac{1}{3!} - A = 0. \end{cases}$$

Звідси знаходимо $A = \frac{1}{6}$, $B = \frac{2}{3}$.

За цих значень A і B залишковий член квадратурної формули (6.30)

$$R(f) = \frac{2h^5}{4!} \left(\frac{1}{5} f^{(5)}(x_0 + \theta_1 h) - \frac{1}{6} f^{(5)}(x_0 - \theta_2 h) + f^{(5)}(x_0 + \theta_3 h) \right).$$

Але $f^{(5)}$ неперервна на $[x_0 - h, x_0 + h]$, тому існує точка $\xi \in [x_0 - h, x_0 + h]$ така, що

$$\frac{\frac{1}{5} f^{(5)}(x_0 + \theta_1 h) - \frac{1}{6} f^{(5)}(x_0 - \theta_2 h) + f^{(5)}(x_0 + \theta_3 h)}{\frac{1}{5} + 2\left(-\frac{1}{6}\right)} = f^{(5)}(\xi).$$

Отже,

$$R(f) = -\frac{h^5}{90} f^{(5)}(\xi), \quad x_0 - h \leq \xi \leq x_0 + h. \quad (6.31)$$

Таким чином, триточкову квадратурну формулу (6.30) можна записати так:

$$\int_{x_0-h}^{x_0+h} f(x) dx = \frac{h}{3} (f(x_0 - h) + 4f(x_0) + f(x_0 + h)) - \frac{h^5}{90} f^{(5)}(\xi). \quad (6.32)$$

Це і є квадратурна формула Сімпсона, або формула парабол із залишковим членом. Вона точна для многочлена третього степеня, бо похідна четвертого порядку від такого многочлена дорівнює нулю. З формули (6.31) легко знайти таку оцінку для абсолютної похибки чисельного інтегрування за формулою Сімпсона:

$$|R(f)| \leq \frac{h^5}{90} M_4, \quad M_4 = \max_{[x_0-h; x_0+h]} |f^{(5)}(x)|.$$

Якщо треба обчислити $\int_a^b f(x) dx$ з достатньою точністю, то відрізок

$[a, b]$ ділять на $2n$ рівних відрізків завдовжки $h = \frac{b-a}{2n}$ і до кожного з відрізків $[x_{2k}; x_{2k+2}]$ ($k = 0, 1, \dots, n-1$) застосовують формулу Сімпсона (6.32). Тоді

$$\int_a^b f(x) dx = \sum_{k=0}^{n-1} \int_{x_{2k}}^{x_{2k+2}} f(x) dx =$$

$$\begin{aligned} &= \frac{h}{3} \sum_{k=0}^{n-1} (f(x_{2k}) + 4f(x_{2k+1}) + f(x_{2k+2})) - \frac{h^5}{90} \sum_{k=0}^{n-1} f^{(5)}(\xi_k) = \\ &= \frac{h}{3} (f(a) + f(b) + 2 \sum_{k=1}^{n-1} f(x_{2k}) + 4 \sum_{k=0}^{n-1} f(x_{2k+1})) - \frac{h^5}{90} \sum_{k=0}^{n-1} f^{(5)}(\xi_k), \end{aligned}$$

де $\xi_k \in [x_{2k}; x_{2k+1}]$ ($k = 0, 1, \dots, n-1$).

Оскільки $f^{(5)}(x)$ неперервна на відрізку $[a; b]$, то існує точка $\xi \in [a; b]$ така, що

$$\frac{1}{n} \sum_{k=0}^{n-1} f^{(5)}(\xi_k) = f^{(5)}(\xi), \quad \xi \in [a; b].$$

Таким чином дістаємо узагальнену формулу Сімпсона (парабол) із залишковим членом вигляду:

$$\begin{aligned} \int_a^b f(x) dx &= \frac{h}{3} (f(a) + f(b) + 2 \sum_{k=1}^{n-1} f(x_{2k}) + \\ &+ 4 \sum_{k=0}^{n-1} f(x_{2k+1})) - \frac{h^4}{180} (b-a) f^{(5)}(\xi). \end{aligned} \quad (6.33)$$

Залишковий член узагальненої формули Сімпсона

$$R(f) = -\frac{h^4}{180} (b-a) f^{(5)}(\xi) = -\frac{(b-a)^5}{2880n^4} f^{(5)}(\xi), \quad h = \frac{b-a}{2n}. \quad (6.34)$$

Звідси дістаємо таку оцінку абсолютної похибки чисельного інтегрування за узагальненою формулою Сімпсона:

$$\begin{aligned} |R(f)| &\leq \frac{h^4}{180} (b-a) M_4 = \frac{(b-a)^5}{2880n^4} M_4, \\ M_4 &= \max_{[a; b]} |f^{(5)}(x)|. \end{aligned} \quad (6.35)$$

Якщо наближене значення інтеграла треба обчислити з точністю $\epsilon > 0$, то відповідний крок інтегрування h визначається нерівністю

$$h \leq \sqrt[5]{\frac{180\epsilon}{(b-a)M_4}},$$

або, що те саме, відрізок $[a; b]$ треба поділити на n рівних частин, де

$$n \geq \sqrt[5]{\frac{(b-a)^5 M_4}{2880\epsilon}}. \quad (6.36)$$

За узагальненою формулою Сімпсона обчислимо наближене значен-

ня інтеграла (6.19) з кроком $h = 0,1$ і оцінимо повну абсолютну похибку Δ_I . Користуючись табл. 6.1, за формулою (6.33) знайдемо:

$$I_{\text{см}} = 0,38177448 \approx 0,3817745.$$

Щоб оцінити залишковий член $R(f)$ формули Сімпсона за формулою (6.35), треба знайти похідну четвертого порядку від функції $f(x) = x \cos x$. Маємо

$$f^{(4)}(x) = 4 \sin x + x \cos x,$$

звідси

$$|f^{(4)}(x)| = 4 |\sin x| + |x| |\cos x| < 4 + 1 = 5,$$

$$M_4 = \max_{(0;1)} |f^{(4)}(x)| = 5.$$

Тому для залишкового члена $R(f)$ за формулою (6.35) ($a = 0$; $b = 1$; $h = 0,1$; $M_4 = 5$) дістанемо $|R(f)| \leq 0,278 \cdot 10^{-5}$.

Похибка остаточного округлення $\Delta_0 = 0,2 \cdot 10^{-7}$, а неусувна похибка

$$\tilde{R} = \Delta_f \sum_{k=1}^n |A_k| = \Delta_f \cdot \frac{h}{3} \cdot 6n = 2nh \Delta_f = (b-a)\Delta_f = 0,5 \cdot 10^{-7},$$

бо $\frac{b-a}{2n} = h$, а значення підінтегральної функції f у вузлах x_k ($k = 0, 1, \dots, 10$) обчислювали з точністю $0,5 \cdot 10^{-7}$, тобто $\Delta_f = 0,5 \cdot 10^{-7}$.

За формулою (6.3) для повної абсолютної похибки чисельного інтегрування функції $f(x) = x \cos x$ знаходимо таку оцінку:

$$\Delta_I = 0,278 \cdot 10^{-5} + 0,5 \cdot 10^{-7} + 0,2 \cdot 10^{-7} = 0,285 \cdot 10^{-5} < 0,3 \cdot 10^{-5}.$$

Отже, обчислене за формулою Сімпсона для $n = 10$, $h = 0,1$ наближене значення інтеграла (6.19) має п'ять правильних значущих цифр, тобто

$$\int_0^1 x \cos x dx = 0,381774 \pm 0,000003$$

(у відповіді збережено одну сумнівну цифру).

Найбільший внесок у повну абсолютну похибку узагальненої формули Сімпсона вносить залишковий член $R(f)$. Тому для визначення кількості відрізків n розбиття $[a; b]$, яке гарантує обчислення наближеного значення інтеграла з точністю $\epsilon > 0$, досить скористатися формулою (6.36). Звичайно, всі проміжні обчислення при цьому слід проводити з точністю, більшою за ϵ .

Наприклад, щоб обчислити наближене значення інтеграла (6.19) з точністю $\epsilon = 0,5 \cdot 10^{-4}$ (з чотирма правильними десятковими знаками), треба відрізок $[0; 1]$ поділити не менш як на три рівні частини, бо за формулою (6.36) ($a = 0$, $b = 1$, $M_4 = 5$) маємо

$$n > \sqrt[4]{\frac{5 \cdot 10^4}{0,5 \cdot 2880}} = 5 \sqrt[4]{\frac{1}{18}} = 5 \cdot 0,48549176 = 2,4274588 \approx 3.$$

Обчислимо інтеграл (6.19) за формулою (6.33), поклавши $n = 2, 4, 8, 16$ (це відповідає $h = 0,25; 0,125; 0,0625; 0,03125$). Знайдемо $I_2 = 0,38182200$; $I_4 = 0,38177633$; $I_8 = 0,38177346$; $I_{16} = 0,38177333$. А це означає, що I_2 має три, I_4 — п'ять, I_8 — шість правильних значущих десяткових цифр. Що ж до I_{16} , то тут усі вісім цифр правильні.

У § 6.6 описано ще один спосіб практичної оцінки похибки чисельного інтегрування.

§ 6.6. ПОРІВНЯННЯ І ПРАКТИЧНА ОЦІНКА ПОХИБКИ КВАДРАТУРНИХ ФОРМУЛ

У §§ 6.3–6.5 розглянуто деякі квадратурні формули. Узагальнені формули трапецій і Сімпсона — це формули замкнутого типу, середніх прямокутників — відкритого типу, а лівих і правих прямокутників — напівзамкнутого і напіввідкритого типу. Як було уже зазначено, точне значення інтеграла визначають за формулами: лівих і правих прямокутників, якщо підінтегральна функція стала; середніх прямокутників і трапецій, якщо підінтегральна функція лінійна, і за формулою Сімпсона, якщо підінтегральною функцією є многочлен степеня, не вищого від третього.

Точність квадратурної формули характеризується порядком залишкового члена $R(f)$ стосовно степеня кроку інтегрування h . З формул (6.14) — (6.16), (6.27) і (6.34) видно, що залишковий член $R(f)$ квадратурних формул залежить від кроку інтегрування h і $R(f) \rightarrow 0$ ($h \rightarrow 0$). Кажуть, що залишковий член $R(f)$ має порядок p (p — натуральне число) відносно h , якщо існує скінченна границя

$$\lim_{h \rightarrow 0} \frac{R(f)}{h^p} = M \neq 0,$$

і записують це так: $R(f) = O(h^p)$.

Отже, залишкові члени формул лівих і правих прямокутників відносно h мають перший порядок: $R(f) = O(h)$, середніх прямокутників і трапецій — другий: $R(f) = O(h^2)$, а Сімпсона — четвертий: $R(f) = O(h^4)$. Квадратурну формулу вважають тим точнішою, чим більший порядок її залишкового члена $R(f)$. З розглянутих раніше квадратурних формул найточнішою є формула Сімпсона, а найменш точними — формули лівих і правих прямокутників. Точність формул середніх прямокутників і трапецій однакова.

Як було вище зазначено, іноді оцінити залишковий член квадратурної формули дуже важко або й неможливо, наприклад тоді, коли функцію задано графічно або таблично і аналітичний вираз її невідомий, або коли функцію задано складним аналітичним виразом і її похідні важко

оцінити. Але якщо похідну певного порядку знайдено, то оцінити її за модулем на відрізку інтегрування завжди можна, побудувавши за допомогою ЕОМ таблицю значень похідної. Проте оцінити залишковий член $R(f)$ квадратурної формули можна й тоді, коли не вдається оцінити зверху модуль похідної підінтегральної функції. Важливо лише знати порядок залишкового члена $R(f)$ відносно кроку інтегрування h . Для цього використовують *метод подвійного перерахунку*.

Нехай залишковий член деякої квадратурної формули має порядок p відносно кроку інтегрування h , тобто $R(f) = O(h^p)$, $p \in \mathbb{N}$. Припустимо також, що похідна, яка входить до залишкового члена $R(f)$, на відрізку інтегрування $[a; b]$ змінюється мало, а тому наближено її можна вважати сталою. Тоді залишковий член $R(f)$ набере вигляду

$$R(f) = Mh^p,$$

де M — деяка невідома стала.

Якщо відрізок $[a; b]$ поділити на n і $2n$ рівні частини ($h = \frac{b-a}{n}$, $\frac{h}{2} = \frac{b-a}{2n}$) і обчислити за квадратурною формулою наближені значення I_n та I_{2n} інтеграла $I = \int_a^b f(x) dx$, а відповідні їм залишкові члени позначити через $R_n(f)$ і $R_{2n}(f)$, то дістанемо дві рівності

$$\begin{aligned} I &= I_n + R_n(f) = I_n + Mh^p, \\ I &= I_{2n} + R_{2n}(f) = I_{2n} + M\left(\frac{h}{2}\right)^p. \end{aligned} \quad (6.37)$$

Ці рівності можна розглядати як лінійну систему рівнянь відносно I і M . Виключивши з цієї системи I , знайдемо для M значення

$$M = \frac{2^p}{2^p - 1} \cdot \frac{I_{2n} - I_n}{h^p}.$$

Підставивши це значення M у вираз для $R_{2n}(f)$, дістанемо

$$R_{2n}(f) = \frac{I_{2n} - I_n}{2^p - 1}. \quad (6.38)$$

Отже, залишковий член квадратурної формули пропорційний різниці двох наближених значень інтеграла, обчислених за цією ж квадратурною формулою з кроками h і $h/2$. Таку оцінку похибки квадратурної формули називають *правилом Рунге*. Якщо тепер (6.38) підставити у друге рівняння системи (6.37), то знайдемо уточнене значення інтеграла

$$I_{n,2n} = I_{2n} + \frac{I_{2n} - I_n}{2^p - 1}. \quad (6.39)$$

Обчислення наближеного значення інтеграла за формулою (6.39) називають *екстраполяцією за Річардсоном*. Якщо $I_n \neq I_{2n}$, то уточнене значення $I_{n,2n}$ ніколи не лежить між I_n і I_{2n} . Якщо $I_{2n} > I_n$, то з формули (6.39) випливає, що $I_{n,2n} > I_{2n} = \max\{I_n, I_{2n}\}$. А якщо $I_{2n} < I_n$, то $I_{n,2n} < I_{2n} = \min\{I_n, I_{2n}\}$. Отже, наближення $I_{n,2n}$ визначають з наближень I_n та I_{2n} в результаті операції екстраполяції, тому й сам спосіб обчислення $I_{n,2n}$ назвали екстраполяцією.

У табл. 6.4 для квадратурних формул, розглянутих у § 3–5, подано значення порядку p залишкового члена відносно кроку h , формули для обчислення значень залишкового члена $R_{2n}(f)$ і уточненого значення інтеграла $I_{n,2n}$.

Таблиця 6.4

Узагальнена квадратурна формула	p	$R_{2n}(f)$	$I_{n,2n}$
лівих і правих прямокутників	1	$I_{2n} - I_n$	$I_{2n} + (I_{2n} - I_n)$
середніх прямокутників і трапецій	2	$\frac{1}{3}(I_{2n} - I_n)$	$I_{2n} + \frac{1}{3}(I_{2n} - I_n)$
Сімпсона	4	$\frac{1}{15}(I_{2n} - I_n)$	$I_{2n} + \frac{1}{15}(I_{2n} - I_n)$

З табл. 6.4 видно, що для обчислення наближеного значення інтеграла з точністю $\varepsilon > 0$ методом подвійного перерахунку треба:

1. Обчислити наближені значення інтеграла I_n та I_{2n} з кроком

$$h = \frac{b-a}{n} \quad \text{і} \quad \frac{h}{2} = \frac{b-a}{2n}.$$

2. За формулою (6.38) обчислити наближене значення похибки $R_{2n}(f)$ чисельного інтегрування.

3. Порівняти $R_{2n}(f)$ з ε . Якщо $|R_{2n}(f)| < \varepsilon$, то за формулою (6.39) обчислити уточнене значення інтеграла $I_{n,2n}$ і процес обчислень припинити. Якщо $|R_{2n}(f)| \geq \varepsilon$, то, зберігши значення I_{2n} , відрізок $[a; b]$ поділити на $4n$ рівних частин і обчислити $R_{4n}(f)$, яке знову порівнюється з ε . Цей процес послідовного збільшення вдвічі числа вузлів квадратурної формули (зменшення вдвічі кроку інтегрування) продовжують доти, поки на певному кроці k не виконуватиметься нерівність $|R_{2^k n}(f)| < \varepsilon$.

Очевидно, що цей алгоритм має циклічний характер, Бейсік-програма 0.2 якого для формули Сімпсона має такий вигляд:

Програма 6.2

```

10 REM ЧИСЕЛЬНЕ ІНТЕГРУВАННЯ МЕТОДОМ СІМПСОНА
20 REM З НАПЕРЕД ЗАДАНОЮ ТОЧНІСТЮ
30 REM
40 INPUT "Ввести нижню межу інтегрування A=";A
50 INPUT "Ввести верхню межу інтегрування B=";B
60 INPUT "Ввести число інтервалів інтегрування M=";M
70 INPUT "Ввести точність обчислення E=";E
80 K=M
90 H = (B-A) / (2*M) : N=0
100 T=A: GOSUB 230 : I=F
110 T=T+H: GOSUB 230 : I=I+4*F
120 N=N+1: IF M=N THEN 140
130 T=T+H: GOSUB 230: I=I+2*F: GOTO 110
140 T=B: GOSUB 230 : I=I*(I+F)/3
150 IF M=K THEN 160 ELSE 170
160 U=I : M=M+M: GOTO 90
170 R=(I-U)/15
180 IF ABS (R)>E THEN 160
190 I=I+R
200 PRINT "Для A="A; "B="B; "M="M; "E="E
210 PRINT "значення інтеграла I="I; "похибка R="R
220 END
230 F= F(T) — вираз підінтегральної функції
240 RETURN
    
```

Приклад 2. Використавши правило Рунге і екстраполяцію за Річардсоном, обчислити довжину дуги еліпса $\frac{x^2}{9} + \frac{y^2}{4} = 1$ за квадратурними формулами § 6.3–6.5.

Розв'язання. Відомо, що довжина l дуги еліпса

$$l = a \int_0^{2\pi} \sqrt{1 - \varepsilon^2 \cos^2 t} dt = 4a \int_0^{\pi/2} \sqrt{1 - \varepsilon^2 \cos^2 t} dt,$$

де $\varepsilon = \frac{\sqrt{a^2 - b^2}}{a}$ — ексцентриситет еліпса. Якщо в останньому інтегралі замінити незалежну змінну за формулою $t = \frac{\pi}{2} - \tau$, то дістанемо:

$$\int_0^{\pi/2} \sqrt{1 - \varepsilon^2 \cos^2 t} dt = \int_0^{\pi/2} \sqrt{1 - \varepsilon^2 \sin^2 \tau} d\tau = E(\varepsilon^2),$$

де $E(\varepsilon^2)$ — так званий повний еліптичний інтеграл другого роду. Отже, довжина дуги еліпса $l = 4a E(\varepsilon^2)$.

За умовою $a = 3, b = 2, \varepsilon^2 = \frac{5}{9}$, тому $l = 4a \int_0^{\pi/2} \sqrt{1 - \frac{5}{9} \sin^2 t} dt = 4a E(\frac{5}{9})$. Щоб обчислити наближене значення інтеграла $E(\frac{5}{9}) = \int_0^{\pi/2} \sqrt{1 - \frac{5}{9} \sin^2 t} dt$, знайдемо спочатку значення функції $f(t) = \sqrt{1 - \frac{5}{9} \sin^2 t}$ на відрізку $[0; \frac{\pi}{2}]$ з кроком $h = \frac{\pi}{24}$ (табл. 6.5).

Таблиця 6.5

№ пор.	t_i	$\sin t_i$	$f_i = f(t_i)$		
0	0	0	1		
1	$\frac{\pi}{24} = 0,13089969$	0,13052619		0,99525624	
2	$\frac{\pi}{12} = 0,26179938$	0,25881903			0,98121598
3	$\frac{\pi}{8} = 0,39269907$	0,38268343		0,95845751	
4	$\frac{\pi}{6} = 0,52359876$	0,50000000			0,92796072
5	$\frac{5\pi}{24} = 0,65449845$	0,60876144		0,89113208	
6	$\frac{\pi}{4} = 0,78539815$	0,70710681			0,84983657
7	$\frac{7\pi}{24} = 0,91629783$	0,79335334		0,80642922	
8	$\frac{\pi}{3} = 1,0471975$	0,86602543			0,76376259
9	$\frac{3\pi}{8} = 1,1780972$	0,92387955		0,72512323	
10	$\frac{5\pi}{12} = 1,3089969$	0,96592582			0,69401700
11	$\frac{11\pi}{24} = 1,4398965$	0,99144484		0,67372806	
12	$\frac{\pi}{2} = 1,5707963$	1	0,66666667		
		$2n = 12$	$\Sigma_0 = 1,6666667$	$\Sigma_1 = 5,0501264$	$\Sigma_2 = 4,21679286$
		$2n' = 6$	$\Sigma'_0 = 1,6666667$	$\Sigma'_1 = 2,5250696$	$\Sigma'_2 = 1,6917233$

Користуючись цією таблицею, знаходимо:

1. За формулою лівих прямокутників:

а) $n = 3, h = \frac{\pi}{6} = 0,52359876,$

$$E_3 = h(f_0 + f_4 + f_8) = \frac{\pi}{6} \cdot 2,6917233 = 1,4093830;$$

б) $n = 6, h = \frac{\pi}{12} = 0,26179938,$

$$E_6 = h(f_0 + f_2 + f_4 + f_6 + f_8 + f_{10}) = h(f_0 + \Sigma_2) = \frac{\pi}{12} \cdot 5,2167929 =$$

$$= 1,3657531, R_6(f) = (E_6 - E_3) = -0,0436299,$$

$$E_{3,6} = E_6 + R_6(f) = 1,3221232;$$

в) $n = 12, h = \pi/24 = 0,13089969,$

$$E_{12} = h \sum_{k=0}^{11} f_k = h(f_0 + \Sigma_1 + \Sigma_2) = \frac{\pi}{24} \cdot 10,266919 = 1,3439365,$$

$$R_{12}(f) = E_{12} - E_6 = -0,0218166,$$

$$E_{6,12} = E_{12} + R_{12}(f) = 1,3221199.$$

Отже, $E_{3,6}$ і $E_{6,12}$ мають по шість правильних значущих цифр, тому $E(5/9) = 1,32212$.

2. За формулою правих прямокутників:

а) $n = 3, h = \pi/6 = 0,52359876,$

$$E_3 = h(f_4 + f_8 + f_{12}) = \frac{\pi}{6} \cdot 2,3583900 = 1,2348501;$$

б) $n = 6, h = \pi/12 = 0,26179938,$

$$E_6 = h(f_2 + f_4 + f_6 + f_8 + f_{10} + f_{12}) = h(\Sigma_2 + f_{12}) =$$

$$= \frac{\pi}{12} \cdot 4,8834596 = 1,2784867, R_6(f) = E_6 - E_3 = 0,0436366,$$

$$E_{3,6} = E_6 + R_6(f) = 1,3221233;$$

в) $n = 12, h = \pi/24 = 0,13089969,$

$$E_{12} = h \sum_{k=1}^{12} f_k = h(\Sigma_1 + \Sigma_2 + f_{12}) = \frac{\pi}{24} \cdot 9,9335859 = 1,3003033,$$

$$R_{12}(f) = E_{12} - E_6 = 0,0218163,$$

$$E_{6,12} = E_{12} + R_{12}(f) = 1,3221196.$$

Як бачимо, і за формулою правих прямокутників наближене значення $E\left(\frac{5}{9}\right) = 1,32212$, де всі цифри правильні.

3. За формулою трапецій:

а) $n = 3, h = \pi/6 = 0,52359876, h/2 = \pi/12 = 0,26179938;$

$$E_3 = \frac{h}{2} (f_0 + 2(f_4 + f_8) + f_{12}) = (\pi/12) \cdot 5,0501133 = 1,3221165;$$

б) $n = 6, h = \pi/12, h/2 = \pi/24 = 0,13089969,$

$$E_6 = \frac{h}{2} (f_0 + 2(f_2 + f_4 + f_6 + f_8 + f_{10}) + f_{12}) = \frac{h}{2} (\Sigma_0 + 2\Sigma_2) =$$

$$= (\pi/24) \cdot 10,10025239 = 1,3221195,$$

$$R_6(f) = \frac{1}{3}(E_6 - E_3) = 0,0000011,$$

$$E_{3,6} = E_6 + R_6(f) = 1,3221209;$$

в) $n = 12, h = \pi/24, h/2 = \pi/48 = 0,065449845 = 6,5449845 \cdot 10^{-2},$

$$E_{12} = \frac{h}{2} (f_0 + 2 \sum_{k=1}^{11} f_k + f_{12}) = \frac{h}{2} (\Sigma_0 + 2(\Sigma_1 + \Sigma_2)) =$$

$$= (\pi/48) \cdot 20,2005051 = 1,3221204,$$

$$R_{12}(f) = \frac{1}{3}(E_{12} - E_6) = \frac{1}{3} \cdot 0,00000006 = 0,00000002 \approx 0,$$

$$E_{6,12} = E_{12} = 1,3221204.$$

Отже, $E(5/9) = 1,32212$, де всі цифри правильні.

4. За формулою Сімпсона:

а) $n = 3, h = \frac{\pi}{12}, \frac{h}{3} = \frac{\pi}{36} = 0,087266461 = 8,7266461 \cdot 10^{-2},$

$$E_3 = \frac{h}{3} (f_0 + 4(f_2 + f_6 + f_{10}) + 2(f_4 + f_8) + f_{12}) = \frac{h}{3} (\Sigma'_0 + 4\Sigma'_1 + 2\Sigma'_2) =$$

$$= \left(\frac{\pi}{36}\right) \cdot (1,66666667 + 4 \cdot 10,100278 + 2 \cdot 1,6917233) =$$

$$= \frac{\pi}{36} \cdot 15,150391 = 1,3221210;$$

б) $n = 6, h = \pi/24, \frac{h}{3} = \frac{\pi}{72} = 4,3633230 \cdot 10^{-2},$

$$E_6 = \frac{h}{3} (\Sigma_0 + 4\Sigma_1 + 2\Sigma_2) = \left(\frac{\pi}{72}\right) \cdot 30,300758 = 1,3221199,$$

$$R_6(f) = \frac{1}{15} (E_6 - E_3) = \frac{1}{15} \cdot 0,1 \cdot 10^{-6} \approx 0,$$

$$E_{3,6} = E_6 = 1,3221199 \approx 1,32212.$$

Для порівняння наведемо значення інтеграла $E\left(\frac{5}{9}\right)$, обчислене за формулою лінійного інтерполювання (5.32) на основі табличних даних для функції $E(\epsilon^2)$, взятих із "П'ятизначних математичних таблиць" Б.І.Сегала і К.А.Семендяєва:

k	ϵ^2	E	ΔE
0	0,55226	1,32384	-0,00303
1	0,55805	1,32081	-0,00303
2	0,56328	1,31778	

Наше значення $\epsilon^2 = \frac{5}{9} = 0,(5)$ лежить між $\epsilon_0^2 = 0,55226$ і

$\varepsilon_1^2 = 0,55805$. Значення $E_0 = 1,32384$, $\Delta E_0 = -0,00303$, $h = \varepsilon_1^2 - \varepsilon_0^2 = -0,00579$, $\varepsilon^2 - \varepsilon_0^2 = \frac{5}{9} - 0,55226 = 0,00330$. Тоді за формулою (5.32) маємо:

$$E\left(\frac{5}{9}\right) = E_0 + \frac{\Delta E_0}{h} (\varepsilon^2 - \varepsilon_0^2) = 1,32384 + \frac{-303}{579} \cdot 0,00330 = 1,322113.$$

Після округлення до п'ятого десяткового розряду маємо: $E\left(\frac{5}{9}\right) = 1,32211$, що добре узгоджується з виконаними вище обчисленнями.

Отже, екстраполяція за Річардсоном — могутній і універсальний алгоритм підвищення точності чисельного інтегрування функцій. Навіть у таких методах низького порядку точності, як методи лівих і правих прямокутників, її застосування дає змогу дістати результат досить високого порядку точності при мінімальних обсягах обчислювальної роботи. Саме тому метод екстраполяції за Річардсоном з розвитком ЕОМ дістав значний розвиток у багатьох розділах обчислювальної математики.

ЛАБОРАТОРНА РОБОТА № 5

Тема. Чисельне інтегрування функцій.

Завдання 1. За узагальненою формулою середніх прямокутників (6.13): а) обчислити значення інтеграла за вказаним розвитком відрізка інтегрування $[a; b]$ на N_0 рівних частин і оцінити залишковий член; б) за формулою (6.21) із точністю $\varepsilon > 0$ знайти значення n і обчислити наближене значення інтеграла для $N \geq n$ (N вибрати так, якщо це дозволяють умови задачі, щоб $(b-a)$ ділилося на нього точно).

2. За узагальненою формулою Сімпсона (6.33) обчислити значення інтеграла із заданою точністю $\varepsilon > 0$, виходячи з вказаного значення N_0 . Крок інтегрування h визначається рівністю $h = \frac{b-a}{2N_0}$. Значення кроку інтегрування h , що забезпечує вказану точність, визначити за допомогою подвійного перерахунку з поділом відрізка $[a; b]$ на N і $2N$ частин. Надрукувати межі інтегрування, послідовність пар (N, I_N) , $(2N, I_{2N})$ і т.д., а також остаточне значення $R = \frac{1}{15} |I_{2N} - I_N|$.

3. Побудувати алгоритм обчислення наближеного значення інтеграла за узагальненою формулою трапецій (6.26) і написати відповідну програму мовою Бейсік. За цієї програмою обчислити наближене значення інтеграла при заданому N_0 (кількість відрізків інтегрування) і оцінити залишковий член за формулою (6.28); обчислити наближене значення інтеграла з точністю $\varepsilon > 0$ подвоєнням кількості вузлів інтегрування $N_0, 2N_0, 4N_0$ і т.д.

до виконання нерівності $\left| R(f) \right| \approx \frac{|I_{2n} - I_n|}{3} < \varepsilon$. Надрукувати межі інтегрування,

послідовні пари чисел (N_0, I_{N_0}) , $(2N_0, I_{2N_0})$ і т.д., а також $\left| R(f) \right| \approx \frac{|I_{2N} - I_N|}{3}$. Назвати правильні цифри обчисленого значення інтеграла. Порівняти обчислене значення інтеграла з тими значеннями, які дістають за формулою Ньютона—Лейбніца.

Варіант	$f(x)$	a	b	N_0	ε
1	$e^x \cos x$	0	1	5	$0,5 \cdot 10^{-6}$
2	$x^2 e^{-x}$	0	1	5	$0,5 \cdot 10^{-8}$
3	$x^2 \arctg x$	0	1	5	$0,5 \cdot 10^{-8}$
4	$e^{\sin x} \sin 2x$	0	1	5	$0,5 \cdot 10^{-8}$
5	$e^{\cos^2 x} \cos x$	0	1	5	$0,5 \cdot 10^{-8}$
6	$e^x / (1+x)$	0	1	5	$0,5 \cdot 10^{-8}$
7	$(1+x)^{3/2} e^{-x}$	0	1	5	$0,5 \cdot 10^{-8}$
8	$e^x / (3+2\cos x)$	0	1	5	$0,5 \cdot 10^{-8}$
9	$e^{\cos x} (1+x)$	0	1	5	$0,5 \cdot 10^{-8}$
10	$(1+x) \ln^2(1+x)$	0	1	5	$0,5 \cdot 10^{-6}$
11	$(1+x^2) e^{\arctg x}$	0	1	5	$0,5 \cdot 10^{-7}$
12	$e^{\sin x} \sin(\cos x)$	0	1	6	$0,5 \cdot 10^{-7}$
13	$\sin x / \sqrt{1+x^2}$	0	1	6	$0,5 \cdot 10^{-8}$
14	$x^2 + 2 \ln x$	1	2	8	$0,5 \cdot 10^{-8}$
15	$x^2 \cos(x/2)$	1	2	8	$0,5 \cdot 10^{-8}$
16	$x^2 \sin(1+x^2)$	0	1	6	$0,5 \cdot 10^{-6}$
17	$x \cos x^2$	0	1	6	$0,5 \cdot 10^{-6}$
18	$\cos x / (1 + \sin^3 x)$	0	1	6	$0,5 \cdot 10^{-8}$
19	$(1-x^2) \ln(1+x)$	0	1	5	$0,5 \cdot 10^{-6}$
20	$\sqrt{1+x} \cos x$	0	1	5	$0,5 \cdot 10^{-8}$
21	$\sqrt{1+x} \sin x$	0	1	5	$0,5 \cdot 10^{-8}$
22	$e^{-x} / (1+x)$	1	2	10	$0,5 \cdot 10^{-6}$
23	$\arctg x / (1+x^2)$	0	1	10	$0,5 \cdot 10^{-4}$
24	$x \ln^2 x$	2	3	8	$0,5 \cdot 10^{-8}$
25	$\sqrt{1 - (7/16) \sin^2 x}$	0	$\pi/2$	6	$0,5 \cdot 10^{-8}$
26	$\arcsin e^{-(x/8)}$	1	7	6	$0,5 \cdot 10^{-8}$
27	$\arccos e^{-(x/6)}$	2	7	5	$0,5 \cdot 10^{-8}$
28	$\ln(2 + \lg(x/10))$	0	14	7	$0,5 \cdot 10^{-8}$
29	$5x \ln x^2$	0	2	10	$0,5 \cdot 10^{-8}$
30	$3x^3 \cos x^2$	0	1	5	$0,5 \cdot 10^{-8}$

Розділ 7. РОЗВ'ЯЗУВАННЯ ЗВИЧАЙНИХ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ

§ 7.1. ПОСТАНОВКА ЗАДАЧІ

Часто задачі техніки і природознавства математично зводяться до відшукування розв'язку певного диференціального рівняння (або системи таких рівнянь), який задовольняє певні початкові умови (задача Коші). Проінтегрувати таке рівняння в скінченному вигляді вдається досить рідко. При цьому дістають здебільшого такий вираз, до якого шукана функція входить неявно, а тому користуватися ним незручно.

На практиці застосовують здебільшого наближене інтегрування диференціальних рівнянь. Воно дає змогу знайти наближений розв'язок задачі Коші або у вигляді певного аналітичного виразу (наприклад, ряду Тейлора), або у вигляді деякої таблиці значень.

Розглянемо окремі методи чисельного розв'язування задачі Коші для звичайного диференціального рівняння першого порядку, розв'язаного відносно похідної. Наближений розв'язок задачі Коші записують у вигляді певної таблиці значень.

Задача Коші полягає в тому, щоб знайти розв'язок $y(x)$ диференціального рівняння

$$y' = f(x, y), \quad (7.1)$$

який задовольняє початкову умову

$$y(x) \Big|_{x=x_0} = y_0. \quad (7.2)$$

Геометрично це означає, що треба знайти ту інтегральну криву $y(x)$ рівняння (7.1), яка проходить через точку (x_0, y_0) .

Задача Коші (7.1)—(7.2) має єдиний розв'язок, наприклад при виконанні умов такої теореми.

Теорема (Пікара). Якщо функція $f(x, y)$ двох змінних x і y неперервна в замкнутому прямокутнику

$$\bar{D} = \{(x, y) : |x - x_0| \leq l, |y - y_0| \leq b\}$$

з центром у точці (x_0, y_0) і задовольняє в ньому умову Ліпшиця по змінній y , тобто існує число $K > 0$, яке не залежить від x і y , таке, що

$$|f(x_1, y_1) - f(x_2, y_2)| \leq K|y_1 - y_2| \quad (7.3)$$

для будь-яких точок $(x_1, y_1) \in \bar{D}$ і $(x_2, y_2) \in \bar{D}$, то існує єдина диференційовна функція $y = \varphi(x)$, яка є розв'язком диференціального рівняння (7.1), що задовольняє початкову умову (7.2). Цей розв'язок визначений і неперервно диференційовний принаймні на відрізку $[x_0 - h; x_0 + h]$, де

$$h = \min \left\{ l, \frac{b}{M} \right\}, \quad M = \max_{(x,y) \in \bar{D}} |f(x, y)|. \quad (7.4)$$

Розглянемо так звані однокрокові чисельні методи розв'язування задачі Коші (7.1)—(7.2), в яких, щоб знайти наближений розв'язок у точці $x_{k+1} = x_k + h$, досить знати її розв'язок у точці x_k . І оскільки розв'язок задачі в точці x_0 відомий з початкових умов, то ці методи дають змогу обчислити послідовно значення розв'язку в наступних точках $x_1 = x_0 + h$, $x_2 = x_1 + h$, ... З однокрокових чисельних методів розглянемо лише явні методи Рунге—Кутта. Окремими представниками цих методів є методи типу Ейлера, які хронологічно передували методам Рунге—Кутта.

Надалі припускатимемо, що функція $f(x, y)$ рівняння (7.1) задовольняє умови теореми Пікара.

§ 7.2. МЕТОД ЕЙЛЕРА

Нехай на відрізку $[x_0; x_0 + l]$ треба знайти чисельний розв'язок задачі Коші (7.1)—(7.2). Для цього відрізок $[x_0; x_0 + l]$ поділимо на n (для простоти) рівних частин точками $x_0, x_1, x_2, \dots, x_n = x_0 + l$, де $x_k = x_0 + kh$ ($k = 0, 1, 2, \dots, n$), $h = \frac{l}{n}$.

Величину h називають кроком чисельного інтегрування диференціального рівняння (7.1).

Розв'язати задачу (7.1)—(7.2) чисельно — це означає для заданої послідовності $x_0, x_1, \dots, x_n = x_0 + l$ незалежної змінної x і числа y_0 знайти числову послідовність y_1, y_2, \dots, y_n , тобто для заданої послідовності значень незалежної змінної $x_k = x_0 + kh$ ($k = 0, 1, \dots, n$) побудувати таблицю наближених значень шуканого розв'язку задачі Коші.

Якщо наближений розв'язок задачі (7.1)—(7.2) в точці x_k відомий, то, проінтегрувавши рівняння (7.1) в межах від x_k до x_{k+1} , знайдемо його розв'язок в точці x_{k+1} за формулою

$$y(x_{k+1}) = y(x_k) + \int_{x_k}^{x_{k+1}} f(x, y(x)) dx. \quad (7.5)$$

Саме ця формула є вихідною для побудови багатьох чисельних методів розв'язування задачі (7.1)—(7.2).

Метод Ейлера. Якщо інтеграл у правій частині формули (7.5) обчислити за формулою лівих прямокутників, то знайдемо

$$y(x_{k+1}) = y(x_k) + hf(x_k, y(x_k)) + O(h^2). \quad (7.6)$$

Відкинувши в цій рівності доданок порядку $O(h^2)$, дістанемо розрахункову формулу

$$y_{k+1} = y_k + hf(x_k, y_k) \quad (k = 0, 1, 2, \dots, n-1), \quad h = x_{k+1} - x_k, \quad (7.7)$$

яку називають *формулою Ейлера*. Тут і далі скрізь y_k і $y(x_k)$ — відповідно наближене і точне значення шуканого розв'язку задачі (7.1) — (7.2) у точці x_k . Різницю $y_k - y(x_k)$ називають *похибкою* наближеного значення y_k в точці x_k .

Оскільки дотична до графіка функції $y(x)$ в точці (x_k, y_k) має кутовий коефіцієнт k , який дорівнює значенню похідної $y'_k = f(x_k, y_k)$, то рівняння дотичної до інтегральної кривої $y(x)$ задачі (7.1) — (7.2) в точці (x_k, y_k) матиме вигляд

$$y - y_k = y'_k(x - x_k) \quad \text{або} \quad y - y_k = f(x_k, y_k)(x - x_k).$$

Звідси для ординати точки y_{k+1} перетину цієї дотичної з прямою $x = x_{k+1}$ дістанемо формулу (7.7). А це означає, що на кожному з відрізків $[x_k, x_{k+1}]$, $(k = 0, 1, \dots, n-1)$ інтегральна крива наближено замінюється відрізком дотичної до неї в точці (x_k, y_k) . Якщо в площині Oxy позначити точки $M_k(x_k, y_k)$, $k = 0, 1, 2, \dots, n$ і сполучити їх по порядку відрізками, то дістанемо ламану (її називають *ламанною Ейлера*), яка наближено зображує графік шуканого розв'язку задачі (7.1) — (7.2). У цьому й полягає геометричний зміст методу Ейлера (рис. 7.1). Значимо, що похибка методу Ейлера на кожному кроці є величина порядку $O(h^2)$. Точність методу досить мала і з переходом від точки x_k до точки x_{k+1} її похибка систематично зростає.

Графічну схему алгоритму чисельного розв'язування задачі (7.1) — (7.2)

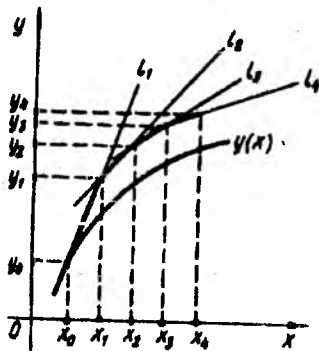


Рис. 7.1

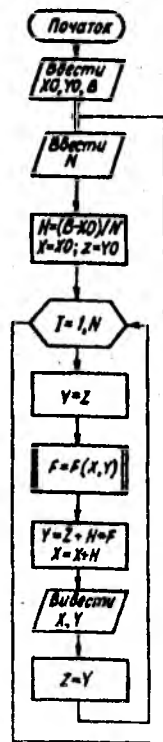


Рис. 7.2

за методом Ейлера подано на рис. 7.2. Відповідна програма мовою Бейсік має вигляд:

Програма 7.1

```

10 PRINT "РОЗВ'ЯЗУВАННЯ ДИФЕРЕНЦІАЛЬНОГО РІВНЯННЯ "
20 PRINT "          МЕТОДОМ ЕЙЛЕРА "
30 REM
40 INPUT "Ввести початкове значення X0="; X0
50 INPUT "Ввести початкове значення Y0="; Y0
60 INPUT "Ввести кінець відрізка інтегрування B="; B
70 INPUT "Ввести число поділу відрізка інтегрування N="; N
80 H=(B-X0)/N
90 X=X0 : Z=Y0
100 FOR I=1 TO N
110 Y=Z : GOSUB 170
120 Y=Z+H*F : X=X+H
130 PRINT "Y("X")="; Y
140 Z=Y
150 NEXT I
160 GOTO 70
170 F=f(x,y): 'F = вираз правої частини рівняння
180 RETURN
    
```

Програму 7.1, як і всі інші в цьому розділі, побудовано так, щоб забезпечувалось виконання обчислень з подвійним перерахунком (докладніше див. § 7.4). Рядок програми за номером 170 забезпечує обчислення значення правої частини рівняння (7.1) в точці (x_k, y_k) і присвоєння його значення змінній F .

Удосконалений метод Ейлера. Якщо інтеграл у правій частині формули (7.5) обчислити за формулою середніх прямокутників, тобто значення підінтегральної функції $f(x, y(x))$ обчислити в точці $x_{k+\frac{1}{2}} = x_k + \frac{1}{2}h$, то знайдемо

$$y(x_{k+1}) = y(x_k) + hf(x_{k+\frac{1}{2}}, y(x_{k+\frac{1}{2}})) + O(h^2). \quad (7.8)$$

Величину невідомого значення функції $y(x_{k+\frac{1}{2}})$ обчислимо за формулою (7.6) з кроком $\frac{1}{2}h$. Матимемо

$$y(x_{k+\frac{1}{2}}) = y(x_k) + \frac{1}{2}hf(x_k, y(x_k)) + O(h^2).$$

Підставивши це значення $y(x_{k+\frac{1}{2}})$ в (7.8), дістанемо

$$y(x_{k+1}) = y(x_k) + hf(x_{k+\frac{1}{2}}, y(x_k)) + \frac{1}{2}hf(x_k, y(x_k)) + O(h^2) +$$

$$+ O(h^3) = y(x_k) + hf(x_{k+\frac{1}{2}}, y(x_k)) + \frac{h}{2} f(x_{k+\frac{1}{2}}, y(x_k)) + O(h^3).$$

Відкинувши тут доданок, пропорційний h^3 , матимемо

$$y_{k+1} = y_k + hf(x_{k+\frac{1}{2}}, y_k) + \frac{1}{2} hf(x_k, y_k).$$

Розрахункові формули вдосконаленого методу Ейлера можна записати у вигляді

$$y_{k+\frac{1}{2}} = y_k + \frac{1}{2} hf(x_k, y_k), \quad (7.9)$$

$$y_{k+1} = y_k + hf(x_{k+\frac{1}{2}}, y_{k+\frac{1}{2}}). \quad (7.10)$$

Отже, в удосконаленому методі Ейлера спочатку за методом Ейлера (формула (7.9)) обчислюють наближений розв'язок $y_{k+\frac{1}{2}}$ задачі (7.1)–

(7.2) в точці $x_{k+\frac{1}{2}} = x_k + \frac{1}{2}h$, а потім за формулою (7.10) — наближений розв'язок y_{k+1} у точці x_{k+1} ; на кожному кроці інтегрування праву частину рівняння (7.1) обчислюють двічі (у точках (x_k, y_k) і $(x_{k+\frac{1}{2}}, y_{k+\frac{1}{2}})$).

Геометрично це означає, що на відрізку $[x_k, x_{k+1}]$ графік інтегральної кривої задачі (7.1)–(7.2) замінюється відрізком прямої, яка проходить через точку (x_k, y_k) і має кутовий коефіцієнт $k = f(x_{k+\frac{1}{2}}, y_{k+\frac{1}{2}})$. Іншими словами, ця пряма (рис. 7.3) утворює з додатним напрямом осі Ox кут $\varphi = \arctg f(x_{k+\frac{1}{2}}, y_{k+\frac{1}{2}})$.

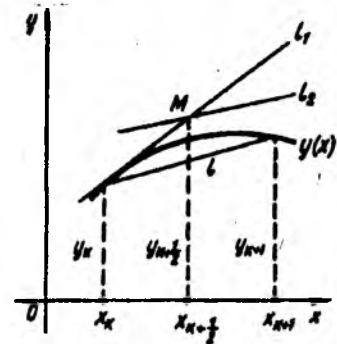


Рис. 7.3

Що ж до точки $(x_{k+\frac{1}{2}}, y_{k+\frac{1}{2}})$, то це точка перетину дотичної до інтегральної кривої задачі (7.1)–(7.2) в точці (x_k, y_k) з прямою $x = x_k + \frac{1}{2}h$. Похибка вдосконаленого методу Ейлера на кожному кроці має порядок $O(h^3)$.

Удосконалений метод Ейлера—Копші.

Якщо інтеграл у правій частині формули

(7.5) обчислити за формулою трапецій, то матимемо

$$y(x_{k+1}) = y(x_k) + \frac{h}{2} (f(x_k, y(x_k)) + f(x_{k+1}, y(x_{k+1}))) + O(h^3). \quad (7.11)$$

Невідоме значення $y(x_{k+1})$, що входить до правої частини цієї рівності, можна обчислити за формулою (7.7). Підставивши його в праву частину рівності (7.11), дістанемо рівність

$$y(x_{k+1}) = y(x_k) + \frac{h}{2} (f(x_k, y(x_k)) +$$

$$+ f(x_{k+1}, y(x_k) + hf(x_k, y(x_k)) + O(h^2))) + O(h^3) =$$

$$= y(x_k) + \frac{h}{2} (f(x_k, y(x_k)) + (f(x_{k+1}, y(x_k) + hf(x_k, y(x_k))) + O(h^3))).$$

Звідси для удосконаленого методу Ейлера—Копші матимемо такі розрахункові формули:

$$\bar{y}_{k+1} = y_k + hf(x_k, y_k), \quad (7.12)$$

$$y_{k+1} = y_k + \frac{h}{2} (f(x_k, y_k) + f(x_{k+1}, \bar{y}_{k+1})). \quad (7.13)$$

Отже, і в цьому методі на кожному кроці інтегрування праву частину рівняння (7.1) обчислюють двічі: спочатку за методом Ейлера (формула (7.12)) обчислюють наближене значення шуканого розв'язку \bar{y}_{k+1} у точці x_{k+1} , яке потім уточнюють за формулою (7.13). Похибка методу на кожному кроці має порядок $O(h^3)$.

Така побудова наближеного розв'язку задачі (7.1)–(7.2) з геометричної точки зору означає, що на відрізку $[x_k, x_{k+1}]$ графік інтегральної кривої наближають відрізком прямої, яка проходить через точку (x_k, y_k) і має кутовий коефіцієнт $k = \frac{1}{2}(f(x_k, y_k) + f(x_{k+1}, \bar{y}_{k+1}))$. Тобто ця пряма утворює з додатним напрямом осі Ox кут $\varphi = \arctg \frac{f(x_k, y_k) + f(x_{k+1}, \bar{y}_{k+1})}{2}$.

Координати точки (x_{k+1}, \bar{y}_{k+1}) визначають як точку перетину дотичної $y = y_k + f(x_k, y_k)(x - x_k)$ до графіка інтегральної кривої задачі (7.1)–(7.2) в точці (x_k, y_k) з прямою $x = x_k + \frac{1}{2}h$ (рис. 7.4). Графічну схему алгоритму наближеного розв'язання задачі Копші (7.1)–(7.2) за удосконаленим методом Ейлера—Копші зображено на рис. 7.5. Відповідна програма 7.2, за якою можна виконати відповідний розрахунок, набирає вигляду:

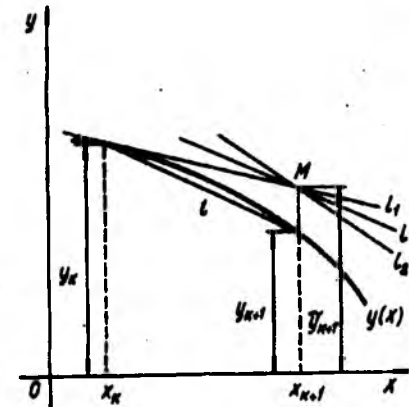


Рис. 7.4

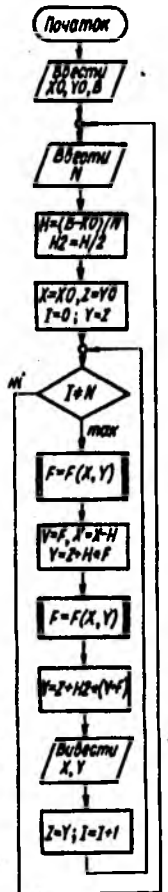


Рис. 7.5

Програма 7.2

```

10 PRINT "Розв'язування диференціального рівняння"
20 PRINT "удосконаленим методом Ейлера—Копі"
25 REM
30 INPUT "Ввести початкове значення X0="; X0
40 INPUT "Ввести початкове значення Y0="; Y0
50 INPUT "Ввести кінець відрізка інтегрування B="; B
60 INPUT "Ввести число поділу відрізка інтегрування N="; N
70 H = (B-X0)/N : H2=H/2
80 X=X0 : Z=Y0 : I=0 : Y=Z
90 IF I=N THEN 60
100 GOSUB 160 : V=F
110 Y=Z+H*F : X=X+H : GOSUB 160
120 Y=Z+H2*(V+F)
130 PRINT "Y("X") ="; Y
140 Z=Y : I=I+1
150 GOTO 90
160 F = f(x,y) : F = вираз правої частини рівняння
170 RETURN

```

Удосконалений метод Ейлера з ітераційною обробкою. Якщо в рівності (7.11) відкинути доданок, пропорційний h^3 , то для знаходження значення невідомого розв'язку u_{k+1} в точці x_{k+1} дістанемо формулу:

$$u_{k+1} = u_k + \frac{h}{2}(f(x_k, u_k) + f(x_{k+1}, u_{k+1})). \quad (7.14)$$

Невідома u_{k+1} входить до обох частин рівності (7.14). Тому метод, що визначається формулою (7.14), належить до *неявних методів* чисельного інтегрування задачі (7.1)-(7.2). Формули (7.12)-(7.13) саме й пропонують один із способів наближеного розв'язування рівняння (7.14). Проте його розв'язок завжди можна обчислити із наперед заданою точністю $\epsilon > 0$, якщо скористатися методом ітерацій. Послідовні наближення можна обчислювати за формулою

$$u_{k+1}^{(i+1)} = u_k + \frac{h}{2}(f(x_k, u_k) + f(x_{k+1}, u_{k+1}^{(i)})) \quad (i = 0, 1, 2, \dots). \quad (7.15)$$

За нульове наближення $u_{k+1}^{(0)}$ можна взяти значення u_k або значення u_{k+1} , обчислене за формулою Ейлера (7.7). Процес ітерацій за формулою (7.15) припиняють з виконанням умови

$$|u_{k+1}^{(i+1)} - u_{k+1}^{(i)}| < \epsilon,$$

тобто коли модуль різниці двох послідовних наближень до шуканого значення величини u_{k+1} буде меншим за наперед задану точність $\epsilon > 0$. За наближене значення величини u_{k+1} в точці x_{k+1} беруть значення $u_{k+1}^{(i+1)}$.

Легко встановити умови, в разі виконання яких ітераційний процес, що задається формулою (7.15), збігається. Для цього від рівності (7.14) віднімемо почленно рівність (7.15). Дістанемо

$$u_{k+1} - u_{k+1}^{(i+1)} = \frac{h}{2}(f(x_{k+1}, u_{k+1}) - f(x_{k+1}, u_{k+1}^{(i)})).$$

Звідси, користуючись умовою Лібшиця (7.3), знаходимо

$$|u_{k+1} - u_{k+1}^{(i+1)}| \leq \frac{h}{2} |f(x_{k+1}, u_{k+1}) - f(x_{k+1}, u_{k+1}^{(i)})| \leq \frac{h}{2} K |u_{k+1} - u_{k+1}^{(i)}|,$$

або

$$|u_{k+1} - u_{k+1}^{(i+1)}| \leq \left(\frac{h}{2}K\right)^{i+1} |u_{k+1} - u_{k+1}^{(0)}|.$$

Отже, ітераційний процес (7.15) збігається, тобто $u_{k+1}^{(i+1)} \rightarrow u_{k+1}$, коли $i \rightarrow \infty$, якщо крок інтегрування h вибрано так, щоб виконувалась нерівність $\frac{1}{2}hK < 1$. При цьому збіжність буде тим швидшою, чим меншою буде величина $\frac{1}{2}hK$.

Якщо алгоритм уточнення числового значення шуканого розв'язку після виконання кількох ітерацій не веде до рівності відповідної кількості десяткових знаків у двох послідовних наближеннях або розбігається, то крок інтегрування зменшують.

Графічну схему алгоритму чисельного інтегрування задачі (7.1)-(7.2) за удосконаленим методом Ейлера з ітераційною обробкою пропонуємо читачеві побудувати самостійно, а відповідна програма має вигляд:

Програма 7.3

```

10 PRINT "Розв'язування задачі Копі"
20 PRINT "удосконаленим методом Ейлера"
30 PRINT "з ітераційною обробкою"
35 REM
40 INPUT "Ввести початкове X0="; X0
50 INPUT "Ввести початкове Y0="; Y0
60 INPUT "Ввести кінець відрізка інтегрування B="; B
70 INPUT "Ввести задану точність EPS="; EPS
80 INPUT "Ввести кількість поділу відрізка N="; N
90 H=(B-X0)/N : H2=H/2
100 X=X0 : Z=Y0 : W=Z
110 FOR I=1 TO N : Y=Z
120 GOSUB 200 : V=F : X=X+H

```

```

130 V=W : GOSUB 200
140 P=Z+H2*(V+F) : R=P-W
150 IF ABS(R) < EPS THEN 170
160 W=P : GOTO 130
170 Y=P : Z=P : PRINT "Y("X")=";Y
180 NEXT I
190 END
200 F=f(x,y) : F = вираз правої частини рівняння
210 RETURN

```

У цій програмі змінні W і P визначають відповідно попередню і наступну ітерації, що обчислюються за формулою (7.15), а R — їх різницю. На кожному кроці інтегрування за нульове наближення беремо значення розв'язку в попередній точці.

Нарешті зазначимо, що неявний метод чисельного інтегрування задачі (7.1)–(7.2) використовують також тоді, коли інтеграл у правій частині рівності (7.5) обчислюють за формулою правих прямокутників

$$y(x_{k+1}) = y(x_k) + hf(x_{k+1}, y_{k+1}) + O(h^2).$$

Відкинувши член, пропорційний h^2 , дістають

$$y_{k+1} = y_k + hf(x_{k+1}, y_{k+1}). \quad (7.16)$$

Для обчислення невідомого значення величини y_{k+1} використовують метод ітерацій. Якщо, наприклад, функція $f(x, y)$ при фіксованому значенні x має частинну похідну по y і якщо ця похідна в деякому околі точки (x_{k+1}, y_k) задовольняє умову $|f'_y(x_{k+1}, y)| \leq q < 1$, то процес ітерацій, що визначається формулою $y_{k+1}^{(i+1)} = y_k + hf(x_{k+1}, y_{k+1}^{(i)})$ ($i = 0, 1, 2, \dots$), $y_{k+1}^{(0)} = y_k$, збігається, причому тим швидше, чим менше q .

Оскільки похибка методу (7.16) на кожному кроці інтегрування має порядок $O(h^2)$, а для розв'язування рівняння (7.16) на кожному кроці треба будувати певну ітераційну процедуру, то в практиці обчислень цей метод використовується дуже рідко.

Збіжність методу Ейлера. Нехай функція $f(x, y)$ з рівняння (7.1) задовольняє умови теореми Пікара і, крім того, нерівність

$$\left| \frac{df(x, y)}{dx} \right| = \left| \frac{\partial f}{\partial x} + f \frac{\partial f}{\partial y} \right| \leq N_1, \quad (x, y) \in \bar{D}. \quad (7.17)$$

Якщо $y(x_k)$ і y_k відповідно точний і наближений розв'язки задачі (7.1)–(7.2) в точці x_k , то різницю

$$\varepsilon_k = y_k - y(x_k) \quad (k=1, 2, \dots, n, \varepsilon_0=0) \quad (7.18)$$

називають *похибкою наближеного значення* y_k в точці x_k , або *похибкою чисельного методу*.

Оскільки значення точного розв'язку $y(x_k)$ задачі (7.1)–(7.2) в точці x_k невідоме, то й похибку ε_k наближеного розв'язку y_k в цій точці

обчислити не можна. Проте оцінити зверху модуль цієї похибки можна. З цієї оцінки можна зробити висновок про збіжність методу Ейлера в кожній точці x_k : $y_k \rightarrow y(x_k)$, якщо $h \rightarrow 0$.

Для цього від рівності (7.7) віднімають почленно рівність (7.5). Врахувавши (7.18), знаходимо

$$\varepsilon_{k+1} = \varepsilon_k + hf(x_k, y_k) - \int_{x_k}^{x_{k+1}} f(x, y(x)) dx,$$

або, після того жного перетворення,

$$\varepsilon_{k+1} - \varepsilon_k = h(f(x_k, y_k) - f(x_k, y(x_k))) + hf(x_k, y(x_k)) - \int_{x_k}^{x_{k+1}} f(x, y(x)) dx. \quad (7.19)$$

Інтеграл цієї рівності проінтегруємо частинами. Матимемо

$$\begin{aligned} \int_{x_k}^{x_{k+1}} f(x, y(x)) dx &= (x - x_{k+1})f(x, y(x)) \Big|_{x_k}^{x_{k+1}} - \\ &- \int_{x_k}^{x_{k+1}} (x - x_{k+1}) \frac{df(x, y(x))}{dx} dx = \\ &= hf(x_k, y(x_k)) - \int_{x_k}^{x_{k+1}} (x - x_{k+1}) \frac{df(x, y(x))}{dx} dx. \end{aligned}$$

Підставивши значення інтеграла в (7.19), дістанемо

$$\varepsilon_{k+1} - \varepsilon_k = h(f(x_k, y_k) - f(x_k, y(x_k))) + \int_{x_k}^{x_{k+1}} (x - x_{k+1}) \frac{df(x, y(x))}{dx} dx.$$

Звідси, використавши умову Ліпшиця (7.3) і нерівність (7.17), послідовно знаходимо

$$\begin{aligned} |\varepsilon_{k+1} - \varepsilon_k| &\leq h |f(x_k, y_k) - f(x_k, y(x_k))| + \\ &+ \int_{x_k}^{x_{k+1}} (x - x_{k+1}) \left| \frac{df(x, y(x))}{dx} \right| dx \leq hK |y_k - y(x_k)| + \\ &+ N_1 \int_{x_k}^{x_{k+1}} (x_{k+1} - x) dx = hK |\varepsilon_k| + \frac{h^2}{2} N_1. \end{aligned}$$

Але

$$|\varepsilon_{k+1}| - |\varepsilon_k| \leq |\varepsilon_{k+1} - \varepsilon_k|,$$

тому з попередньої нерівності дістаємо

або

$$|\varepsilon_{k+1}| - |\varepsilon_k| \leq hK|\varepsilon_k| + \frac{h^2}{2} N_1,$$

$$|\varepsilon_{k+1}| \leq (1+hK)|\varepsilon_k| + \frac{h^2}{2} N_1 \quad (k=0,1,2,\dots,n-1). \quad (7.20)$$

Отже, якщо значення величини модуля похибки наближеного розв'язку u_k в точці x_k відоме, тобто відоме $|\varepsilon_k|$ (а $\varepsilon_0=0$), то нерівність (7.20) дає змогу оцінити зверху величину модуля похибки $|\varepsilon_{k+1}|$ наближеного розв'язку u_{k+1} в точці x_{k+1} .

Проте з нерівності (7.20), яка дає рекурентну оцінку похибки, можна дістати й незалежну оцінку величини модуля похибки наближеного розв'язку задачі (7.1)-(7.2) методом Ейлера в будь-якій точці x_k . Щоб упевнитись у цьому, введемо позначення

$$a = 1+hK, \quad b = \frac{1}{2}h^2 N_1, \quad a>0, \quad b>0, \quad \varepsilon_0 = y_0 - y(x_0) = 0 \quad (7.21)$$

і подамо нерівність (7.20) в такому вигляді

$$|\varepsilon_{k+1}| \leq a|\varepsilon_k| + b \quad (k=0,1,2,\dots,n-1). \quad (7.22)$$

Звідси дістанемо

$$|\varepsilon_1| \leq b,$$

$$|\varepsilon_2| \leq a|\varepsilon_1| + b \leq ab + b = b(1+a),$$

$$|\varepsilon_3| \leq a|\varepsilon_2| + b \leq ab(1+a) + b = b(1+a+a^2),$$

Припустимо, що справедлива нерівність:

$$|\varepsilon_{k-1}| \leq b(1+a+a^2+\dots+a^{k-2})$$

і доведемо, що

$$|\varepsilon_k| \leq b(1+a+a^2+\dots+a^{k-1}).$$

Справді, за нерівністю (7.22) маємо:

$$\begin{aligned} |\varepsilon_k| &\leq a|\varepsilon_{k-1}| + b \leq ab(1+a+a^2+\dots+a^{k-2}) + b = \\ &= b(1+a+a^2+\dots+a^{k-1}). \end{aligned}$$

Оскільки $a = 1+hK \neq 1$, то

$$1+a+a^2+\dots+a^{k-1} = \frac{a^k-1}{a-1}$$

$$|\varepsilon_k| \leq \frac{b(a^k-1)}{a-1} \quad (k=0,1,2,\dots,n).$$

Підставивши в цю формулу значення a і b з (7.21), знайдемо незалежну оцінку похибки в точці x_k :

$$|\varepsilon_k| \leq \frac{hN_1}{2K} \left((1+hK)^k - 1 \right) \quad (k=0,1,2,\dots,n).$$

Скориставшись нерівностями $1+u \leq e^u$ ($u > 0$) і $kh \leq nh = l$, незалежну оцінку для величини модуля похибки наближеного розв'язку в точці x_k ($k=0,1,2,\dots,n$) можна записати остаточно в такому вигляді

$$|\varepsilon_k| \leq \frac{hN_1}{2K} (e^{lK} - 1) \quad (k=0,1,2,\dots,n). \quad (7.23)$$

Права частина цієї нерівності пропорційна кроку інтегрування h і тому на будь-якому скінченному відрізку похибка наближеного розв'язку прямуватиме до нуля, коли $h \rightarrow 0$, тобто метод Ейлера збігається.

Слід зазначити, що в практичних обчисленнях похибки наближеного розв'язку задачі (7.1)-(7.2) за допомогою формули (7.23) оцінювати часто незручно й недоцільно. Незручно тому, що треба спочатку обчислити сталі K і N_1 , а недоцільно, бо ця оцінка значно завищена. З розвитком і використанням сучасних ЕОМ для оцінки похибки наближеного розв'язку задачі (7.1)-(7.2) здебільшого вдаються до подвійного перерахунку і правила Рунге. Докладніше про це йдеться в § 7.4. Аналогічними міркуваннями, хоча і дещо складнішими, можна оцінити величини похибки наближеного розв'язку задачі (7.1)-(7.2) удосконаленими методами Ейлера і Ейлера-Коші. З цих оцінок впливатиме збіжність згаданих методів.

Справедливі такі твердження.

Теорема 1. Якщо права частина рівняння (7.1) задовольняє умови теореми Пікара і нерівності

$$\left| \frac{df(x,y(x))}{dx} \right| \leq N_1, \quad \left| \frac{d^2f(x,y(x))}{dx^2} \right| \leq N_2, \quad (x,y) \in \bar{D},$$

то для похибки ε_k наближеного значення u_k удосконаленого методу Ейлера справедлива оцінка

$$|\varepsilon_k| \leq \frac{1}{8} h^2 \left(N_1 + \frac{N_2}{3K} \right) \frac{\left(1+hK + \frac{1}{2} h^2 K^2 \right)^n - 1}{1 + \frac{1}{2} hK} \quad (k=1,2,\dots,n).$$

З цієї нерівності випливає, що $|\varepsilon_k| = |u_k - y(x_k)| \rightarrow 0$, коли $h \rightarrow 0$, тобто удосконалений метод Ейлера збігається, якщо $h \rightarrow 0$, причому порядок збіжності дорівнює двом.

Теорема 2. Якщо права частина рівняння (7.1) задовольняє умови теореми Пікара і, крім того, нерівності

$$\left| f'_x(x,y) \right| \leq M_1, \quad \left| f'_y(x,y) \right| \leq M_2, \quad \left| \frac{d^2f(x,y(x))}{dx^2} \right| \leq N_2, \quad (x,y) \in \bar{D},$$

а крок інтегрування h задовольняє нерівність $\frac{1}{2}hK < 1$, то для похибки ϵ_k наближеного розв'язку y_k в точці x_k удосконаленого методу Ейлера—Коші справедлива оцінка

$$|\epsilon_k| \leq \frac{h^2}{12} \left(\frac{N_2}{K} + 3(M_1 + MM_2) \right) \left(\left| \frac{1 + \frac{1}{2}hK}{1 - \frac{1}{2}hK} \right|^n - 1 \right),$$

з якої випливає збіжність методу, причому порядок збіжності дорівнює двом.

§ 7.3. МЕТОДИ РУНГЕ—КУТТА

З появою і розвитком ЕОМ у чисельному інтегруванні звичайних диференціальних рівнянь бурхливого розвитку набули методи типу Рунге—Кутта. Такі методи побудовано до 10-го порядку точності включно. В обчислювальній практиці їх широко застосовують завдяки тому, що вони:

однокрокові, тобто для обчислення розв'язку задачі (7.1)—(7.2) в точці x_{k+1} треба знати її розв'язок лише в точці x_k ;

дають змогу здійснювати чисельне інтегрування із змінним кроком; особливо зручні для програмування на ЕОМ, оскільки обчислення за ними має циклічний характер.

Поряд з перевагами ці методи мають і недоліки. Так, дуже важко оцінити похибки наближеного розв'язку задачі (7.1)—(7.2), бо відомі гарантовані оцінки локальної похибки здебільшого значно завищені, а тому практична їх цінність незначна. До недоліків слід віднести й те, що треба обчислювати в кількох точках праву частину рівняння (7.1) (функцію $f(x, y)$) на кожному кроці інтегрування. В удосконалених методах Ейлера і Ейлера—Коші на кожному кроці інтегрування значення функції $f(x, y)$ обчислювали в двох точках, а в методах Рунге—Кутта третього і четвертого порядку точності його обчислюють відповідно в трьох і чотирьох точках. Що ж до методів Рунге—Кутта, порядок точності яких більший від чотирьох, то праву частину рівняння (7.1) на кожному кроці інтегрування завжди обчислюють у точках, кількість яких більша від порядку точності методу. Уявлення про це дає табл. 7.1.

Таблиця 7.1

s	1	2	3	4	5	6	7	8	9	10
n	1	2	3	4,5	6	7,8	9	11	14	18

Тут s — порядок точності методу Рунге—Кутта, а n — кількість точок, в яких обчислюється функція $f(x, y)$ на кожному кроці інтегрування.

Далі задачу Коші (7.1)—(7.2) розв'язуватимемо чисельно методами Рунге—Кутта, порядок точності яких не перевищує чотирьох. Припустимо, що функція $f(x, y)$ в прямокутнику $\bar{\Delta}$ має неперервні частинні похідні до деякого порядку n . Тоді розв'язок $y(x)$ матиме неперервні похідні до $(n+1)$ -го порядку і для досить малих значень h у точці $x_{k+1} = x_k + h$, $h > 0$ його можна подати у вигляді розкладу

$$y(x_{k+1}) = y(x_k) + hy'(x_k) + \frac{h^2}{2!} y''(x_k) + \dots + \frac{1}{n!} h^n y^{(n)}(x_k) + O(h^{n+1}). \quad (7.24)$$

Похідні $y^{(i)}(x_k)$ ($i = 1, 2, \dots, n$) у правій частині формули (7.24) можна виразити через значення функції f та її частинних похідних у точці (x_k, y_k) .

Маємо

$$y'(x_k) = f(x_k, y_k) = f_k, \quad (7.25)$$

$$y''(x_k) = f_x'(x_k, y_k) + y f_y'(x_k, y_k) = (f_x' + f f_y')_k, \quad (7.26)$$

$$y'''(x_k) = (f_x'^2 + 2f f_{xy}' + f_y'^2 + f f_{xx}')_k, \quad (7.27)$$

де індекс k тут і далі означає, що відповідні функції обчислюють у точці (x_k, y_k) .

Аналогічно обчислюють і похідні вищих порядків, хоч їх вирази із зростанням порядку значно ускладнюються. Тому безпосереднє використання їх для обчислення наближеного значення розв'язку $y(x_{k+1}) = y(x_k + h)$ у точці x_{k+1} за формулою (7.24) навряд чи може бути доцільним.

Слід зазначити, що для ЕОМ розроблено комплекси програм, які дають змогу машині самостійно виконувати формальне диференціювання функцій, якщо лише задано програму обчислення функції $f(x, y)$. Таким чином, відпадає потреба створювати програми обчислення похідних функції $f(x, y)$. Проте затрати машинного часу на виконання цих алгоритмів часто можуть бути такі великі, що доцільніше користуватися чисельними методами інтегрування звичайних диференціальних рівнянь, зокрема методами Рунге—Кутта.

Замість похідних $y_k^{(i)}$ ($i = 1, 2, \dots, n+1$), що входять у праву частину формули (7.24), Рунге запропонував значення наближеного розв'язку y_{k+1} в точці x_{k+1} обчислювати за формулами вигляду:

$$y_{k+1} = y_k + \sum_{i=1}^r w_i k_i, \quad (7.28)$$

де функції $k_i = h^i f(\xi_i, \eta_i)$ ($i = 1, 2, \dots, r$), $\xi_i = x_k + \alpha_i h$, $\eta_i = y_k + \sum_{j=1}^{i-1} \beta_{ij} k_j$, а w_i, α_i ($\alpha_1 = 0$) і β_{ij} — деякі сталі, які називатимемо параметрами формул Рунге—Кутта.

Невідомі параметри w_1, α_1, β_1 визначають з умови, щоб розклад за степенями h правої частини формули (7.28) збігався з розкладом (7.24) до якомога вищих степенів h для довільної функції $f(x, y)$ і довільного кроку інтегрування h .

Перед побудовою розрахункових формул методів Рунге—Кутта введемо поняття порядку точності цих методів. Нехай $y(x_{k+1}) = y(x_k + h)$ і y_{k+1} — відповідно точний і наближений розв'язки задачі (7.1)—(7.2) в точці x_{k+1} , причому точний розв'язок обчислено за формулою (7.24), а наближений — за формулою (7.28), яку після розкладу за степенями h можна записати так:

$$y_{k+1} = y_k + ha_1 + \frac{h^2}{2} a_2 + \frac{h^3}{3!} a_3 + \dots + \frac{h^{n+1}}{(n+1)!} a_{n+1}, \quad (7.29)$$

де a_l ($l = 1, 2, \dots$) — деякі числові коефіцієнти.

Найбільший з показників степеня h , при якому коефіцієнти розкладів (7.24) і (7.29) рівні між собою, називають *порядком точності наближеного методу*. Отже, якщо порядок точності наближеного методу дорівнює s , то це означає, що в розкладах (7.24) і (7.29) коефіцієнти $a_1 = y'_k, a_2 = y''_k, a_3 = y'''_k, \dots, a_s = y_k^{(s)}$, але $a_{s+1} \neq y_k^{(s+1)}$.

Зазначимо, що похибка методу на одному кроці інтегрування завжди має порядок на одиницю більший, ніж порядок точності методу. Тому порядок точності методу Ейлера дорівнює одиниці, бо похибка методу на одному кроці має порядок h^2 , а порядок точності удосконалених методів Ейлера і Ейлера—Коші дорівнює двом, бо порядок похибки цих методів на одному кроці h^3 .

Формули Рунге—Кутта першого порядку точності. Нехай функція $f(x, y)$ в околі точки (x_k, y_k) має неперервні частинні похідні першого порядку. Тоді розв'язок $y(x)$ задачі Коші в околі точки x_k матиме неперервні похідні до другого порядку включно. Тому для досить малих значень h у точці $x_{k+1} = x_k + h$ його можна подати за степенями h , врахувавши (7.25), так:

$$y(x_{k+1}) = y(x_k) + hf(x_k, y_k) + O(h^2). \quad (7.24)$$

Формули Рунге—Кутта першого порядку точності будуватимемо в такому вигляді:

$$y_{k+1} = y_k + w_1 k_1, \quad (7.30)$$

де $k_1 = hf(x_k, y_k)$, а w_1 — невідомий параметр. Розклавши за степенями h праву частину виразу (7.30), матимемо

$$y_{k+1} = y_k + hw_1 f(x_k, y_k). \quad (7.31)$$

Прирівнявши між собою коефіцієнти при h у формулах (7.24) і (7.31), знаходимо $w_1 = 1$. Підставивши це значення w_1 в рівність (7.30), впевнюємося в тому, що формулою Рунге—Кутта першого порядку точності ($s=1$) є формула Ейлера. Похибка цього методу $y_{k+1} - y(x_{k+1})$ на кожному кроці є величиною порядку h^2 .

Формули Рунге—Кутта другого порядку точності. Нехай функція $f(x, y)$ в околі точки (x_k, y_k) має неперервні частинні похідні до другого порядку включно. Тоді розв'язок $y(x)$ задачі Коші (7.1)—(7.2) в околі точки x_k матиме неперервні похідні до третього порядку включно. Тому для досить малих значень h у точці $x_{k+1} = x_k + h$ його можна подати у вигляді такого розкладу за степенями h

$$y(x_{k+1}) = y(x_k) + hf(x_k, y_k) + \frac{h^2}{2!} y''(x_k) + O(h^3),$$

або, взявши до уваги формули (7.25) і (7.26)

$$y(x_{k+1}) = y(x_k) + hf(x_k, y_k) + \frac{h^2}{2} (f_x' + ff_y')_k + O(h^3). \quad (7.32)$$

Формули Рунге—Кутта другого порядку точності ($s=2$) шукатимемо в такому вигляді:

$$y_{k+1} = y_k + w_1 k_1 + w_2 k_2, \quad (7.33)$$

де

$$k_1 = hf(x_k, y_k), \quad (7.34)$$

$$k_2 = hf(x_k + \alpha_2 h, y_k + \beta_{21} k_1), \quad (7.35)$$

а w_1, w_2, α_2 і β_{21} — невідомі параметри, які визначають так, щоб (7.33) була формулою другого порядку точності ($s=2$), тобто, щоб коефіцієнти в розкладах (7.32) і (7.33) за степенями h були рівними між собою при h і h^2 .

Функції k_1 і k_2 розкладають за степенями h так:

$$\begin{aligned} k_1 &= hf_k \\ k_2 &= h(f_k + \alpha_2 h(f_x')_k + \beta_{21} k_1 (f_y')_k + \dots) = \\ &= hf_k + \alpha_2 h^2 (f_x')_k + h^2 \beta_{21} (ff_y')_k + O(h^3). \end{aligned}$$

Підставивши ці розклади k_1 і k_2 в (7.33), дістанемо розклад за степенями h (залишимо лише члени, пропорційні h^2) виду:

$$y_{k+1} = y_k + h(w_1 + w_2) f_k + h^2 (w_2 \alpha_2 (f_x')_k + w_2 \beta_{21} (ff_y')_k) + O(h^3). \quad (7.36)$$

Прирівнявши коефіцієнти при $hf_k, h^2 (f_x')_k$ і $h^2 (ff_y')_k$ у формулах (7.32) і (7.36), для визначення чотирьох невідомих параметрів w_1, w_2, α_2 і β_{21} дістанемо алгебраїчну систему трьох рівнянь вигляду

$$\begin{cases} w_1 + w_2 = 1, \\ w_2 \alpha_2 = \frac{1}{2}, \\ w_2 \beta_{21} = \frac{1}{2}. \end{cases} \quad (7.37)$$

З системи (7.37) видно, що $w_2 \neq 0, \alpha_2 \neq 0, \beta_{21} \neq 0$ і $\alpha_2 = \beta_{21}$. Вона має нескінченну множину розв'язків, що залежить від одного параметра,

причому кожний її розв'язок дає формули Рунге—Кутта другого порядку точності.

Якщо покласти $w_2=1$, тоді $w_1=0$, $\alpha_2=\beta_{21}=\frac{1}{2}$, а формули (7.33)–(7.35) набирають вигляду

$$y_{k+1} = y_k + k_2, \\ k_1 = hf(x_k, y_k), \quad k_2 = hf\left(x_k + \frac{h}{2}, y_k + \frac{k_1}{2}\right).$$

А це формули удосконаленого методу Ейлера.

Якщо покласти $w_2=\frac{1}{2}$, тоді $w_1=\frac{1}{2}$, $\alpha_2=\beta_{21}=1$ і формули (7.33)–(7.35) запишемо:

$$y_{k+1} = y_k + \frac{1}{2}(k_1 + k_2), \\ k_1 = hf(x_k, y_k), \quad k_2 = hf(x_k + h, y_k + k_1),$$

тобто дістали формули удосконаленого методу Ейлера—Коші.

Якщо покласти $w_2 = \frac{3}{4}$, то тоді $w_1 = \frac{1}{4}$, $\alpha_2 = \beta_{21} = \frac{2}{3}$ і з формул (7.33)–(7.35) дістанемо розрахункові формули

$$y_{k+1} = y_k + \frac{1}{4}(k_1 + 3k_2), \\ k_1 = hf(x_k, y_k), \quad k_2 = hf\left(x_k + \frac{2}{3}h, y_k + \frac{2}{3}k_1\right).$$

Формули Рунге—Кутта третього порядку точності. Нехай функція $f(x, y)$ в околі точки (x_k, y_k) має неперервні частинні похідні до третього порядку включно. Тоді розв'язок $y(x)$ задачі (7.1)–(7.2) в околі точки x_k матиме неперервні похідні до четвертого порядку включно. І тому для досить малих значень h у точці $x_{k+1} = x_k + h$ його можна подати у вигляді такого розкладу за степенями h :

$$y(x_{k+1}) = y(x_k) + hy'(x_k) + \frac{h^2}{2!}y''(x_k) + \frac{h^3}{3!}y'''(x_k) + O(h^4).$$

Підставивши сюди значення $y'(x_k)$, $y''(x_k)$ і $y'''(x_k)$ з (7.25)–(7.27), дістанемо

$$y(x_{k+1}) = y(x_k) + hf_k + \frac{h^2}{2}(f'_x + ff'_y)_k + \\ + \frac{h^3}{6}(f''_{xx} + 2f'_x f'_y + f^2 f''_{yy} + f'_y(f'_x + ff'_y))_k + O(h^4). \quad (7.38)$$

Формули Рунге—Кутта третього порядку точності ($s=3$) будують у такому вигляді:

$$y_{k+1} = y_k + w_1 k_1 + w_2 k_2 + w_3 k_3, \quad (7.39)$$

де

$$k_1 = hf(x_k, y_k), \quad k_2 = hf(x_k + \alpha_2 h, y_k + \beta_{21} k_1),$$

$$k_3 = hf(x_k + \alpha_3 h, y_k + \beta_{31} k_1 + \beta_{32} k_2), \quad (7.40)$$

а $w_1, w_2, w_3, \alpha_2, \alpha_3, \beta_{21}, \beta_{31}, \beta_{32}$ — деякі дійсні параметри, які добирають так, щоб (7.39) була формулою третього порядку точності, тобто, щоб у розкладах (7.38) і (7.39) за степенями h коефіцієнти при h, h^2 і h^3 були рівні між собою. Щоб розкласти за степенями h праву частину формули (7.39), треба спочатку розкласти за степенями h функції k_i ($i = 1, 2, 3$) (формули (7.40)). Маємо

$$k_1 = hf_k, \\ k_2 = hf_k + h(\alpha_2 h f'_x)_k + \beta_{21} k_1 (f'_y)_k + \frac{h}{2}(\alpha_2^2 h^2 f''_{xx})_k + 2\alpha_2 h \beta_{21} k_1 (f''_{xy})_k + \\ + \beta_{21}^2 k_1^2 (f''_{yy})_k + O(h^4), \\ k_3 = hf_k + h(\alpha_3 h f'_x)_k + (\beta_{31} k_1 + \beta_{32} k_2) (f'_y)_k + \frac{h}{2}(\alpha_3^2 h^2 f''_{xx})_k + \\ + 2\alpha_3 h (\beta_{31} k_1 + \beta_{32} k_2) (f''_{xy})_k + (\beta_{31} k_1 + \beta_{32} k_2)^2 (f''_{yy})_k + O(h^4).$$

До правої частини розкладу за степенями h функції k_2 входить функція k_1 , а функції k_3 — функції k_1 і k_2 . Тому підставимо $k_1 = hf_k$ у вираз k_2 , а потім k_1 і k_2 у вираз k_3 . Дістанемо

$$k_1 = hf_k, \\ k_2 = hf_k + h^2(\alpha_2 f'_x)_k + \beta_{21}(ff'_y)_k + \frac{h^3}{2}(\alpha_2^2 f''_{xx})_k + 2\alpha_2 \beta_{21}(ff''_{xy})_k + \\ + \beta_{21}^2 (f^2 f''_{yy})_k + O(h^4), \\ k_3 = hf_k + h^2(\alpha_3 f'_x)_k + (\beta_{31} + \beta_{32})(ff'_y)_k + h^3(\alpha_2 \beta_{32}(f'_x f'_y)_k + \\ + \beta_{21} \beta_{32}(ff''_{yy})_k) + O(h^4).$$

Підставивши ці розклади за степенями h функцій k_1, k_2 і k_3 в (7.39), знайдемо розклад за степенями h наближеного розв'язку y_{k+1} . Маємо

$$y_{k+1} = y_k + h(w_1 + w_2 + w_3)f_k + h^2(\alpha_2 w_2 + \alpha_3 w_3)(f'_x)_k + \\ + h^2(\beta_{21} w_2 + (\beta_{31} + \beta_{32})w_3)(ff'_y)_k + \frac{1}{2}h^3(\alpha_2^2 w_2 + \\ + \alpha_3^2 w_3)(f''_{xx})_k + h^3(\alpha_2 \beta_{21} w_2 + \alpha_3(\beta_{31} + \beta_{32})w_3)(ff''_{xy})_k + \\ + \frac{h^3}{2}(\beta_{21}^2 w_2 + (\beta_{31} + \beta_{32})^2 w_3)(f^2 f''_{yy})_k + h^3 \alpha_2 \beta_{32} w_3 (f'_x f'_y)_k + \\ + h^3 \beta_{21} \beta_{32} w_3 (ff''_{yy})_k + O(h^4). \quad (7.41)$$

Прирівнявши коефіцієнти при однакових степенях h у розкладах (7.38) і (7.41), для знаходження невідомих параметрів $w_1, w_2, w_3, \alpha_2, \alpha_3, \beta_{21}, \beta_{31}, \beta_{32}$ дістанемо нелінійну систему таких алгебраїчних рівнянь:

$$hf_k: \quad w_1 + w_2 + w_3 = 1, \quad (7.42)$$

$$h^2(f'_x)_k: \quad \alpha_2 w_2 + \alpha_3 w_3 = \frac{1}{2}, \quad (7.43)$$

$$h^2(ff_y')_k: \beta_{21}w_2 + (\beta_{31} + \beta_{32})w_3 = \frac{1}{2}, \quad (7.44)$$

$$\frac{1}{2}h^3(f_x^2)_k: \alpha_2^2w_2 + \alpha_3^2w_3 = \frac{1}{3}, \quad (7.45)$$

$$\frac{1}{2}h^3(f_x^2 f_y^2)_k: \beta_{21}^2w_2 + (\beta_{31} + \beta_{32})^2w_3 = \frac{1}{3}, \quad (7.46)$$

$$h^3(ff_{xy}')_k: \alpha_2\beta_{21}w_2 + \alpha_3(\beta_{31} + \beta_{32})w_3 = \frac{1}{3}, \quad (7.47)$$

$$h^3(f_x' f_y')_k: \alpha_2\beta_{32}w_3 = \frac{1}{6}, \quad (7.48)$$

$$h^3(ff_y'^2)_k: \beta_{21}\beta_{32}w_3 = \frac{1}{6}. \quad (7.49)$$

З рівнянь (7.48) і (7.49) випливає, що w_3 , α_2 , β_{21} і β_{32} не дорівнюють нулю і $\alpha_2 = \beta_{21}$, а з рівнянь (7.43) і (7.44) маємо, що $\alpha_3 = \beta_{31} + \beta_{32}$.

Отже, система (7.42)–(7.49) набуває вигляду

$$\beta_{21} = \alpha_2, \quad \alpha_2w_2 + \alpha_3w_3 = \frac{1}{2},$$

$$\beta_{31} + \beta_{32} = \alpha_3, \quad \alpha_2^2w_2 + \alpha_3^2w_3 = \frac{1}{3},$$

$$w_1 + w_2 + w_3 = 1, \quad \alpha_2\beta_{32}w_3 = \frac{1}{6}.$$

Необхідною умовою сумісності лінійних відносно w_2 і w_3 трьох останніх рівнянь цієї системи є рівність нулю визначника

$$\begin{vmatrix} \alpha_2 & \alpha_3 & \frac{1}{2} \\ \alpha_2^2 & \alpha_3^2 & \frac{1}{3} \\ 0 & \alpha_2\beta_{32} & \frac{1}{6} \end{vmatrix} = 0.$$

Звідси дістанемо

$$\alpha_2(2 - 3\alpha_2)\beta_{32} = \alpha_3(\alpha_3 - \alpha_2).$$

Отже, для визначення п'яти невідомих параметрів α_2 , α_3 , β_{21} , β_{31} і β_{32} маємо систему трьох рівнянь

$$\begin{cases} \beta_{21} = \alpha_2, \\ \beta_{31} + \beta_{32} = \alpha_3, \\ \alpha_2(2 - 3\alpha_2)\beta_{32} = \alpha_3(\alpha_3 - \alpha_2). \end{cases} \quad (7.50)$$

Що ж до невідомих w_1 , w_2 і w_3 , то їх знаходимо з системи лінійних рівнянь

$$\begin{cases} \alpha_2\beta_{32}w_3 = \frac{1}{6}, \\ \alpha_2w_2 + \alpha_3w_3 = \frac{1}{2}, \\ w_1 + w_2 + w_3 = 1. \end{cases} \quad (7.51)$$

Отже, існує нескінченна множина формул Рунге—Кутта третього порядку точності виду (7.39)–(7.40), яка залежить від двох параметрів. Невідомі параметри цих формул визначають із систем (7.50) і (7.51). Наведемо приклади окремих формул цієї множини.

1. Якщо покласти $\alpha_2 = \frac{1}{2}$, $\alpha_3 = 1$, тоді $\beta_{21} = \frac{1}{2}$, $\beta_{32} = 2$, $\beta_{31} = -1$, $w_3 = \frac{1}{6}$, $w_2 = \frac{2}{3}$, $w_1 = \frac{1}{6}$.

А формули (7.39) і (7.40) запишуться так:

$$y_{k+1} = y_k + \frac{1}{6}(k_1 + 4k_2 + k_3),$$

$$k_1 = hf(x_k, y_k), \quad k_2 = hf(x_k + \frac{h}{2}, y_k + \frac{k_1}{2}), \quad (7.52)$$

$$k_3 = hf(x_k + h, y_k - k_1 + 2k_2).$$

2. Якщо покласти $\alpha_2 = \frac{1}{2}$, $\alpha_3 = \frac{3}{4}$, то $\beta_{21} = \frac{1}{2}$, $\beta_{32} = \frac{3}{4}$, $\beta_{31} = 0$, $w_3 = \frac{4}{9}$, $w_2 = \frac{1}{3}$, $w_1 = \frac{2}{9}$. А формули (7.39) і (7.40) набувають вигляду

$$y_{k+1} = y_k + \frac{1}{9}(2k_1 + 3k_2 + 4k_3),$$

$$k_1 = hf(x_k, y_k), \quad k_2 = hf(x_k + \frac{h}{2}, y_k + \frac{k_1}{2}), \quad (7.53)$$

$$k_3 = hf(x_k + \frac{3}{4}h, y_k + \frac{3}{4}k_2).$$

3. Нехай $\alpha_2 = \frac{1}{3}$, $\alpha_3 = \frac{2}{3}$. Тоді $\beta_{21} = \frac{1}{3}$, $\beta_{32} = \frac{2}{3}$, $\beta_{31} = 0$, $w_3 = \frac{3}{4}$, $w_2 = 0$, $w_1 = \frac{1}{4}$, і для формул (7.39) і (7.40) дістанемо

$$y_{k+1} = y_k + \frac{1}{4}(k_1 + 3k_2),$$

$$k_1 = hf(x_k, y_k), \quad k_2 = hf(x_k + \frac{h}{3}, y_k + \frac{k_1}{3}), \quad (7.54)$$

$$k_3 = hf(x_k + \frac{2}{3}h, y_k + \frac{2}{3}k_2).$$

Усі формули Рунге—Кутта третього порядку точності на кожному кроці інтегрування мають похибку порядку h^4 .

Методи Рунге—Кутта четвертого порядку точності ($s=4$) можна побудувати аналогічно до попередніх, хоча це вимагатиме громіздкіших викладок і розв'язування нелінійної системи алгебраїчних рівнянь складнішої структури. Похибка цих методів на кожному кроці інтегрування є величиною порядку h^5 . Самі формули мають вигляд

$$y_{k+1} = y_k + \sum_{i=1}^4 w_i k_i, \quad (7.55)$$

де

$$\begin{aligned} k_1 &= hf(x_k, y_k), \\ k_2 &= hf(x_k + \alpha_2 h, y_k + \beta_{21} k_1), \\ k_3 &= hf(x_k + \alpha_3 h, y_k + \beta_{31} k_1 + \beta_{32} k_2), \\ k_4 &= hf(x_k + \alpha_4 h, y_k + \beta_{41} k_1 + \beta_{42} k_2 + \beta_{43} k_3). \end{aligned} \quad (7.56)$$

До цих формул входять 13 невідомих параметрів $w_i, \alpha_i, \beta_{ij}$.

Доведено, що існує нескінченна множина формул виду (7.55), (7.56), яка залежить від двох довільних дійсних параметрів і має четвертий порядок точності.

Обмежимося такими окремими прикладами формул, що належать до цієї множини:

1.

$$\begin{aligned} y_{k+1} &= y_k + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \\ k_1 &= hf(x_k, y_k), \quad k_2 = hf(x_k + \frac{h}{2}, y_k + \frac{k_1}{2}), \\ k_3 &= hf(x_k + \frac{1}{2}h, y_k + \frac{1}{2}k_2), \\ k_4 &= hf(x_k + h, y_k + k_3). \end{aligned} \quad (7.57)$$

Цю формулу найчастіше використовують в обчислювальній практиці. З геометричної точки зору формули (7.57) інтегральну криву задачі (7.1)–(7.2) на відрізку $[x_k, x_{k+1}]$ замінюють відрізком прямої, яка проходить через точку (x_k, y_k) і утворює з додатним напрямом осі Ox кут

$$\varphi = \arctg \frac{k_1 + 2k_2 + 2k_3 + k_4}{6h}.$$

2.

$$\begin{aligned} y_{k+1} &= y_k + \frac{1}{6}(k_1 + 4k_2 + k_4), \\ k_1 &= hf(x_k, y_k), \quad k_2 = hf(x_k + \frac{h}{4}, y_k + \frac{k_1}{4}), \end{aligned} \quad (7.58)$$

$$\begin{aligned} k_3 &= hf(x_k + \frac{h}{2}, y_k + \frac{k_2}{2}), \\ k_4 &= hf(x_k + h, y_k + k_1 - 2k_2 + 2k_3). \end{aligned}$$

3.

$$\begin{aligned} y_{k+1} &= y_k + \frac{1}{8}(k_1 + 3k_2 + 3k_3 + k_4), \\ k_1 &= hf(x_k, y_k), \quad k_2 = hf(x_k + \frac{h}{3}, y_k + \frac{k_1}{3}), \\ k_3 &= hf(x_k + \frac{2}{3}h, y_k - \frac{1}{3}k_1 + k_2), \\ k_4 &= hf(x_k + h, y_k + k_1 - 2k_2 + k_3). \end{aligned} \quad (7.59)$$

Програма 7.4 методу Рунге—Кутта, що визначається формулами (7.57), має вигляд:

Програма 7.4

```

10 PRINT "      Розв'язування задачі Коші"
20 PRINT " методом Рунге—Кутта 4-го порядку точності"
30 PRINT " "
40 INPUT "Ввести X0=";X0
50 INPUT "Ввести Y0=";Y0
60 INPUT "Ввести кінець відрізка інтегрування B=";B
70 INPUT "Ввести число поділу відрізка N=";N
80 H=(B-X0)/N : H2=H/2
90 X=X0: Z=Y0: Y=Z: I=0
100 IF I>N THEN 200
110 GOSUB 210: K1=H*F
120 X=X+H2: Y=Z+K1/2
130 GOSUB 210: K2=H*F
140 Y=Z+K2/2: GOSUB 210: K3=H*F
150 X=X+H2: Y=Z+K3: GOSUB 210: K4=H*F
160 Y=Z+(K1+2*(K2+K3)+K4)/6
170 PRINT "Y("X")="Y
180 Z=Y: I=I+1
190 GOTO 100
200 GOTO 70
210 F=F(X,Y): RETURN
220 END

```

Аналогічно можна побудувати формули Рунге—Кутта вищого порядку точності, проте обчислення їх параметрів значно ускладнюється. Щоб мати уявлення про труднощі обчислювального характеру, які при цьому виникають, зазначимо, що для знаходження значень 76 пара-

метрів $\alpha_i, w_i, \beta_{ij}$ формул Рунге—Кутта 8-го порядку точності треба розв'язати нелінійну систему алгебраїчних рівнянь, до якої входить 200 рівнянь. Та й саму систему ще треба побудувати. Значення правої частини рівняння (7.1) (функції $f(x, y)$) у цьому методі на кожному кроці інтегрування обчислюється в 11 точках.

Отже, в методах високого порядку точності на кожному кроці інтегрування значення функції $f(x, y)$ обчислюють в n точках, де $n > s$ ($s \geq 5$) — порядок точності формули. І оскільки основна частина розрахунків припадає на обчислення значень правої частини рівняння (7.1), то формулами Рунге—Кутта, порядок яких більше чотирьох, користуються значно рідше. Проте не треба забувати, що збільшення порядку точності формул Рунге—Кутта в свою чергу дає змогу виконувати обчислення з більшим кроком інтегрування h . І тому сумарні затрати машинного часу на розв'язування задачі (7.1)—(7.2) можуть бути менші для методів вищого порядку точності.

Виходячи з початкових умов (x_0, y_0) , формули Рунге—Кутта дають змогу обчислити значення шуканого розв'язку y_1 в точці $x_1 = x_0 + h$, тобто знайти нову пару чисел (x_1, y_1) . Аналогічно, використавши нову точку (x_1, y_1) , можна просунутися ще на один крок такої самої чи іншої довжини. Продовжуючи цей процес обчислень, можна скласти таблицю значень шуканого розв'язку задачі Коші в певних точках.

На закінчення зазначимо, що розглянуті вище методи Рунге—Кутта застосовують також до нормальних систем диференціальних рівнянь першого порядку. Німецький математик Р.Цурмюль узагальнив методи Рунге—Кутта на випадок диференціальних рівнянь другого, третього і вищих порядків.

§ 7.4. ПРО ОЦІНКУ ПОХИБКИ НАБЛИЖЕНОГО РОЗВ'ЯЗКУ ЗАДАЧІ КОШІ

Для методів Рунге—Кутта застосовують ряд *априорних* оцінок похибки наближеного розв'язку задачі Коші (7.1)—(7.2). У § 7.2 такі оцінки наведено для методу Ейлера та удосконалених методів Ейлера й Ейлера—Коші. Проте ці оцінки, як було вже зазначено, здебільшого значно завищені. Тому навряд чи доцільно рекомендувати їх для контролю в практичних розрахунках. Їх значення не стільки практичне, скільки теоретичне, бо з них безпосередньо випливає висновок про збіжність цих методів. Крім того, априорні оцінки містять у собі ряд сталих, для знаходження яких часто треба виконувати досить складні обчислення.

Тому, щоб оцінити похибку наближеного розв'язку задачі (7.1)—(7.2), намагаються використати інформацію, яку дістають у процесі чисельного розрахунку (такі оцінки називають *апостеріорними*). Найефективніші оцінки з подвійним перерахунком.

Спинимось докладніше на методі подвійного перерахунку. Розглянемо такі три випадки:

1. Задано крок інтегрування h і треба визначити точні цифри наближеного розв'язку в кожній вузловій точці x_k .

2. Задано точність $\varepsilon > 0$, з якою треба обчислити наближений розв'язок задачі, добираючи належним чином як сам метод, так і крок інтегрування h .

3. Оцінити похибку $\varepsilon_k = y_k - y(x_k)$, де y_k і $y(x_k)$ — відповідно наближений і точний розв'язки задачі, в кожній вузловій точці x_k .

Для цього розв'язок задачі (7.1)—(7.2) у кожній вузловій точці x_k обчислюють двічі: з кроком h і $h/2$. Позначатимемо їх відповідно y_k і y_k^* . Десяткові розряди наближень y_k і y_k^* , які збігаються між собою, вважають точними цифрами наближеного розв'язку в точці x_k .

Якщо наближений розв'язок задачі (7.1)—(7.2) треба обчислити з наперед заданою точністю $\varepsilon > 0$, то, використовуючи метод певного порядку точності, інтегрування з кроками h і $h/2$ доцільно вести паралельно, щоб вчасно встановити неузгодженість між значеннями y_k й y_k^* і, можливо, перейти до нового кроку. Якщо в точці x_k значення y_k і y_k^* задовольняють нерівність $|y_k^* - y_k| < \varepsilon$, то крок інтегрування для наступної точки x_{k+1} можна збільшити, наприклад подвоїти. Якщо $|y_k^* - y_k| \geq \varepsilon$, то крок інтегрування ділять навпіл. Цим забезпечується автоматичний вибір кроку інтегрування.

Нарешті, наявність наближених значень y_k і y_k^* , обчислених відповідно з кроком h і $h/2$, дає змогу наближено оцінити похибку методу $\varepsilon_k = y_k^* - y(x_k)$ у точці x_k . Щоб вивести таку оцінку похибки, припустимо, що виконуються такі умови:

1) на кожному кроці інтегрування h похибка методу приблизно пропорційна h^{s+1} ($s \geq 1$), де s — порядок точності методу;

2) похибка методу на кожному кроці інтегрування однакова;

3) на кожному наступному кроці інтегрування сумарна похибка методу включає також усі похибки, зроблені на попередніх кроках. Тому, якщо $y_1 - y(x_1) = Mh^{s+1}$, де M — невідомий коефіцієнт пропорційності, то

$$y_2 - y(x_2) = 2Mh^{s+1},$$

$$y_3 - y(x_3) = 3Mh^{s+1},$$

$$\dots\dots\dots$$

$$y_n - y(x_n) = nMh^{s+1}.$$

Отже, для похибки в точці x_k при інтегруванні з кроком h маємо рівність

$$y_k - y(x_k) = kMh^{s+1}, \tag{7.60}$$

а при інтегруванні з кроком $h/2$ — рівність:

$$y_k^* - y(x_k) = 2kM \left(\frac{h}{2}\right)^{s+1} \quad (7.61)$$

Віднявши почленно (7.61) від рівності (7.60) і розв'язавши здобуту рівність відносно невідомого коефіцієнта M , знайдемо

$$M = \frac{2^s(y_k - y_k^*)}{kh^{s+1}(2^s - 1)}$$

Підставивши це значення M в (7.61), маємо

$$y_k^* - y(x_k) = \frac{y_k - y_k^*}{2^s - 1}$$

Звідси для абсолютної похибки в точці x_k остаточно дістаємо таку рівність:

$$|\epsilon_k^*| = |y_k^* - y(x_k)| = \frac{|y_k - y_k^*|}{2^s - 1}$$

Оцінку абсолютної похибки методу за допомогою величини $|y_k - y_k^*| / (2^s - 1)$ називають *правилом Рунге*.

В табл. 7.2 подано оцінки похибок методів Рунге—Кутта порядку точності від 1 до 4 за допомогою правила Рунге.

Саме такими оцінками найчастіше користуються в обчислювальній практиці. Проілюструємо це на

Таблиця 7.2

Порядок точності s	Похибка методу $ \epsilon_k^* $
1	$ y_k - y_k^* $
2	$\frac{1}{3} y_k - y_k^* $
3	$\frac{1}{7} y_k - y_k^* $
4	$\frac{1}{15} y_k - y_k^* $

прикладі конкретної задачі Коші.

П р и к л а д. Методами Ейлера (7.7), удосконаленими методами Ейлера (7.9)—(7.10) і Ейлера—Коші (7.12)—(7.13), Рунге—Кутта 4-го порядку точності з кроком $h = 0,05$ на відрізку $[1,6; 2,6]$ знайти чисельний розв'язок такої задачі Коші: $y' = x + \cos \frac{y}{3}$,

$y(1,6) = 4,6$.

Р о з в ' я з а н н я. Результати обчислень задачі з кроками $h =$

$= 0,05$ і $h/2 = 0,025$ подано в табл. 7.3 лише для п'яти точок цього відрізка.

Всі значення розв'язку наведено з вісьмома знаками, хоча не всі вони є точними. Подвійний перерахунок показує, що за методом Ейлера знайдено три точні знаки, за удосконаленими методами Ейлера і Ейлера—Коші — п'ять, а за методом Рунге—Кутта 4-го порядку точності — всі вісім знаків (не враховуємо неминучих похибок округлення, які

здійснює машина). У табл. 7.4 подано наближений розв'язок цієї задачі з точними знаками.

Таблиця 7.3

Метод Ейлера				Удосконалений метод Ейлера		
x_k	y_k	y_k^*	$ y_k - y_k^* $	y_k	y_k^*	$ y_k - y_k^* $
1,6	4,6	4,6	0	4,6	4,6	0
1,8	4,9342303	4,9353109	0,0010806	4,9363916	4,9363827	0,0000089
2,0	5,2853057	5,2878386	0,0020329	5,2898725	5,2898564	0,0000161
2,2	5,6539279	5,6568218	0,0028939	5,6597183	5,6596969	0,0000215
2,4	6,0382439	6,0419480	0,0037041	6,0456572	6,0456317	0,0000255
2,6	6,4389013	6,4434090	0,0045077	6,4479251	6,4478969	0,0000282
Удосконалений метод Ейлера—Коші				Метод Рунге—Кутта 4-го порядку точності		
x_k	y_k	y_k^*	$ y_k - y_k^* $	y_k	y_k^*	$ y_k - y_k^* $
1,6	4,6	4,6	0	4,6	4,6	0
1,8	4,9363920	4,9363828	0,0000092	4,9363798	4,9363798	0
2,0	5,2898756	5,2898573	0,0000183	5,2898511	5,2898512	$1 \cdot 10^{-7}$
2,2	5,6597269	5,6596991	0,0000278	5,6596896	5,6596897	$1 \cdot 10^{-7}$
2,4	6,0456746	6,0456361	0,0000385	6,0456233	6,0456234	$1 \cdot 10^{-7}$
2,6	6,4479548	6,4479044	0,0000504	6,4478878	6,4478877	$1 \cdot 10^{-7}$

Таблиця 7.4

	Метод Ейлера	Удосконалений метод Ейлера	Удосконалений метод Ейлера—Коші	Метод Рунге—Кутта 4-го порядку точності
x_k	y_k	y_k	y_k	y_k
1,6	4,6	4,6	4,6	4,6
1,8	4,94	4,9364	4,9364	4,9353798
2,0	5,29	5,2899	5,2899	5,2898511
2,2	5,66	5,6597	5,6597	5,6596896
2,4	6,04	6,0456	6,0456	6,0456233
2,6	6,44	6,4479	6,4479	6,4478878

За формулами табл. 7.2 дістаємо такі оцінки для абсолютної похибки розв'язку в будь-якій точці $x_k \in [1,6; 2,6]$. Для методу Ейлера

$$|y_k - y(x_k)| = |y_k - y_k^*| \leq \max_{1 \leq k \leq 20} |y_k - y_k^*| = 0,0045077 < 0,005;$$

для удосконаленого методу Ейлера

$$|y_k - y(x_k)| = \frac{1}{3} |y_k - y_k^*| \leq \frac{1}{3} \max_{1 \leq k \leq 20} |y_k - y_k^*| = \\ = \frac{0,0000282}{3} = 0,0000094 < 0,00001;$$

для удосконаленого методу Ейлера—Коші

$$|y_k - y(x_k)| = \frac{1}{3} |y_k - y_k^*| \leq \frac{1}{3} \max_{1 \leq k \leq 20} |y_k - y_k^*| = \\ = \frac{0,0000504}{3} = 0,0000168 \approx 0,00002;$$

для методу Рунге—Кутта 4-го порядку точності

$$|y_k - y(x_k)| = \frac{1}{15} |y_k - y_k^*| \leq \frac{1}{15} \max_{1 \leq k \leq 20} |y_k - y_k^*| = \frac{1}{15} \cdot 10^{-7} < 0,00000001.$$

Нарешті, розглянемо ще один апостеріорний спосіб оцінки похибки методу, в якому використовується машинний розв'язок задачі Коші. Йдеться про двосторонні методи Рунге—Кутта, в яких на кожному кроці інтегрування для шуканого розв'язку задачі дістають дві оцінки: зверху і знизу, тобто розв'язок задачі Коші на кожному кроці затискується у "вилку".

На прикладі методу Рунге—Кутта 2-го порядку точності покажемо, як це зробити в точці $x_1 = x_0 + h$. З формул (7.32) і (7.36) при $k = 0$ для різниці $y_1 - y(x_1)$ дістаємо рівність

$$y_1 - y(x_1) = h(w_1 + w_2 - 1)f_0 + h^2(w_2\alpha_2 - \frac{1}{2})(f_x')_0 + \\ + h^2(w_2\beta_{21} - \frac{1}{2})(ff_y')_0 + O(h^3).$$

Невідомі параметри w_1, w_2, α_2 і β_{21} визначають з умови, щоб похибка $y_1 - y(x_1)$ починалася з максимально можливого степеня h , причому головна частина (відносно h) цієї похибки була пропорційна деякому дійсному параметру ω . Іншими словами, параметри w_1, w_2, α_2 і β_{21} вибирають так, щоб похибка методу

$$y_1 - y(x_1) = \omega h^2 \Psi(f)_0 + O(h^3), \quad (7.62)$$

де $\Psi(f)_0$ — певний диференціальний оператор функції f , обчислений у точці (x_0, y_0) .

З рівності (7.62) випливає, що для досить малих значень кроку h , якщо лише $\Psi(f)_0 \neq 0$, завжди можна побудувати пару формул Рунге—Кутта першого порядку точності, яка відповідає двом різним значенням параметра ω , що відрізняються лише знаком. Одна з цих формул наближатиме розв'язок задачі Коші (7.1)–(7.2) в точці x_1 зверху (верхнє наближення y_1^+), а друга — знизу (нижнє наближення y_1^-), причому відповідно до формули (7.62) маємо

$$y_1^+ - y(x_1) = \omega h^2 \Psi(f)_0 + O(h^3), \\ y_1^- - y(x_1) = -\omega h^2 \Psi(f)_0 + O(h^3).$$

За наближений розв'язок задачі (7.1)–(7.2) в точці x_1 береться півсума цих розв'язків, тобто

$$y_1 = \frac{1}{2}(y_1^+ + y_1^-).$$

Тоді відповідно до формули (7.62) маємо

$$y_1 - y(x_1) = \frac{1}{2}(O^+(h^3) + O^-(h^3)),$$

тобто двосторонній метод Рунге—Кутта має другий порядок точності, хоч верхнє і нижнє наближення — перший. Зрозуміло, що абсолютна величина похибки наближеного розв'язку y_1 в точці x_1 задовольняє нерівність

$$|y_1 - y(x_1)| \leq |y_1^+ - y_1^-|,$$

бо обидва розв'язки y_1^+ і y_1^- лежать між y_1^+ і y_1^- .

Для двосторонніх методів Рунге—Кутта другого порядку точності диференціальний оператор $\Psi(f)_0$ може набувати лише трьох значень

$$\Psi(f)_0 = (f_x')_0, \text{ або } \Psi(f)_0 = (ff_y')_0, \\ \text{або } \Psi(f) = (f_x' + ff_y')_0, \text{ якщо } \alpha_2 = \beta_{21}.$$

Нехай $\Psi(f)_0 = (f_x')_0$. Тоді для визначення параметрів $w_1, w_2, \alpha_2, \beta_{21}$ і ω двостороннього методу Рунге—Кутта 2-го порядку точності дістанемо таку систему рівнянь:

$$\begin{cases} w_1 + w_2 = 1, \\ w_2\alpha_2 = \frac{1}{2} + \omega, \\ w_2\beta_{21} = \frac{1}{2}. \end{cases}$$

з якої видно, що $w_2 \neq 0, \alpha_2 \neq 0, \beta_{21} \neq 0$. Це система трьох рівнянь з п'ятьма невідомими (ω також невідоме). Отже, два параметри можна вибирати довільно. Якщо довільно вибрати параметри w_2 і ω , то з останньої системи рівнянь дістанемо $w_1 = 1 - w_2, \alpha_2 = \frac{1 + 2\omega}{2w_2},$

$\beta_{21} = \frac{1}{2w_2}$. Звідси, зокрема, якщо $w_2 = 1$ (параметр ω поки не фіксуємо!),

то $w_1 = 0, \alpha_2 = \frac{1 + 2\omega}{2}, \beta_{21} = \frac{1}{2}$ і дістанемо формули двостороннього удосконаленого методу Ейлера вигляду

$$y_1 = y_0 + k_2, \quad k_1 = hf(x_0, y_0), \quad k_2 = hf(x_0 + \frac{1 + 2\omega}{2}h, y_0 + \frac{1}{2}k_1).$$

А якщо покласти $w_2 = \frac{1}{2}$, то $w_1 = \frac{1}{2}$, $\alpha_2 = 1+2\omega$, $\beta_{21} = 1$ і дістанемо формули двостороннього удосконаленого методу Ейлера—Коші

$$y_1 = y_0 + \frac{1}{2}(k_1 + k_2), \quad k_1 = hf(x_0, y_0), \\ k_2 = hf(x_0 + (1+2\omega)h, y_0 + k_1).$$

Нехай $\Psi(f)_0 = (ff')_0$. Тоді параметри w_1 , w_2 , α_2 , β_{21} і ω двостороннього методу Рунге—Кутта визначають із системи рівнянь

$$\begin{cases} w_1 + w_2 = 1, \\ w_2 \alpha_2 = \frac{1}{2}, \\ w_2 \beta_{21} = \frac{1}{2} + \omega, \end{cases}$$

в якій, як і раніше, $\alpha_2 \neq 0$, $\beta_{21} \neq 0$, $w_2 \neq 0$. Звідси, вибираючи довільно параметри w_2 і ω , маємо $w_1 = 1 - w_2$, $\alpha_2 = \frac{1}{2w_2}$, $\beta_{21} = \frac{1+2\omega}{2w_2}$. Зокрема,

якщо $w_2 = 1$, то $w_1 = 0$, $\alpha_2 = \frac{1}{2}$, $\beta_{21} = \frac{1+2\omega}{2}$ і дістанемо формули двостороннього методу Ейлера

$$y_1 = y_0 + k_2, \quad k_1 = hf(x_0, y_0), \\ k_2 = hf(x_0 + \frac{1}{2}h, y_0 + \frac{1+2\omega}{2}k_1),$$

(параметр ω поки не фіксується).

Якщо $w_2 = \frac{1}{2}$, то $w_1 = \frac{1}{2}$, $\alpha_2 = 1$, $\beta_{21} = 1+2\omega$ і дістанемо формули двостороннього удосконаленого методу Ейлера—Коші

$$y_1 = y_0 + \frac{1}{2}(k_1 + k_2), \quad k_1 = hf(x_0, y_0), \\ k_2 = hf(x_0 + h, y_0 + (1+2\omega)k_1).$$

Нехай $\alpha_2 = \beta_{21}$ і $\Psi(f)_0 = (f_x' + ff_x')$. Тоді для визначення параметрів w_1 , w_2 , α_2 , і ω дістанемо систему рівнянь

$$\begin{cases} w_1 + w_2 = 1, \\ w_2 \alpha_2 = \frac{1}{2} + \omega. \end{cases}$$

Вибравши, як і раніше, довільно параметри w_2 і ω , знаходимо $w_1 = 1 - w_2$, $\alpha_2 = \frac{1+2\omega}{2w_2}$.

Звідси, якщо $w_2 = 1$, то $w_1 = 0$, $\alpha_2 = \beta_{21} = \frac{1+2\omega}{2}$ і дістанемо формули двостороннього удосконаленого методу Ейлера

$$y_1 = y_0 + k_2, \quad k_1 = hf(x_0, y_0), \\ k_2 = hf(x_0 + \frac{1+2\omega}{2}h, y_0 + \frac{1+2\omega}{2}k_1).$$

Якщо $w_2 = \frac{1}{2}$, то $w_1 = \frac{1}{2}$, $\alpha_2 = \beta_{21} = 1+2\omega$ і дістанемо формули двостороннього удосконаленого методу Ейлера—Коші вигляду

$$y_1 = y_0 + \frac{1}{2}(k_1 + k_2), \quad k_1 = hf(x_0, y_0), \\ k_2 = hf(x_0 + (1+2\omega)h, y_0 + (1+2\omega)k_1).$$

Щоб надати певних числових значень параметрам α_2 і β_{21} в щойно побудованих формулах двостороннього методу Рунге—Кутта другого порядку точності, треба спочатку зафіксувати значення параметра ω . Якщо в цих формулах покласти $\omega = 0$, то дістанемо відомі класичні формули удосконалених методів Ейлера та Ейлера—Коші.

Зазначимо ще, що на першому кроці інтегрування за формулами побудованих двосторонніх методів значення функції f (правої частини рівняння (7.1)) обчислюється в трьох точках, тоді як у методі подвійного перерахунку — в шести.

Для двосторонніх методів Рунге—Кутта запропоновано різні методи продовження розрахунку, які забезпечують як двосторонність наближень, так і гарантовану точність. Досягається це за рахунок варіювання значень кроку h і параметра ω . Спочатку крок h добирають по можливості більшим, а параметр ω — по можливості меншим. Щоб забезпечити двосторонність наближень, параметр ω треба збільшувати, а крок h зменшувати, а щоб досягти наперед заданої точності, крок інтегрування h зменшують. Аналогічно побудовано двосторонні методи Рунге—Кутта вищих порядків точності (до шостого включно) як для диференціальних рівнянь першого порядку, так і для рівнянь вищих порядків.

ЛАБОРАТОРНА РОБОТА № 6

Тема. Чисельне інтегрування звичайних диференціальних рівнянь першого порядку, розв'язаних відносно похідної однокроковими методами.

Завдання 1. Користуючись програмами методів Ейлера, удосконаленого методу Ейлера—Коші і Рунге—Кутта 4-го порядку точності, поданих у §7.2 і §7.3, на відрізку $[x_0; b]$ знайти чисельний розв'язок задачі Коші: $y' = f(x, y)$, $y(x_0) = y_0$ з кроками $h = (b - x_0)/N$ і $h/2$ та визначити правильні цифри наближених розв'язків.

2. Для удосконаленого методу Ейлера (7.9)-(7.10) і методу Рунге—Кутта 3-го порядку точності (7.52) побудувати алгоритм, скласти програми мовою Бейсік і за цими програмами розв'язати на ЕОМ задачу Коші: $y' = f(x, y)$, $y(x_0) = y_0$. Користуючись подвійним перерахунком з кроками $h = (b - x_0)/N$ і $h/2$, визначити правильні цифри наближених розв'язків. Порівняти розв'язки між собою.

3. Для методу Рунге—Кутта 4-го порядку точності (7.58) побудувати графічну схему алгоритму, скласти програму мовою Бейсік і за цією програмою розв'язати задачу Коші: $y' = f(x, y)$, $y(x_0) = y_0$. Використавши подвійний перерахунок з кроками $h = (b - x_0)/N$ і $h/2$, визначити правильні цифри наближених розв'язку і порівняти його з розв'язками завдання 2.

У табл. 7.5 подано 30 варіантів робіт і використано такі позначення: $f(x, y)$ — права частина рівняння (7.1), $(x_0; y_0)$ — початкова умова (7.2), $[x_0; b]$ — відрізок чисельного інтегрування, N — кількість інтервалів, на яку треба поділити відрізок $[x_0; b]$.

Таблица 7.5

Вариант	$f(x,y)$	x_0	b	y_0	N
1	2	3	4	5	6
1	$x + 1 + 2y^2$	1	2	0,5	10
2	$1,4x - \sin(x+2y^2)$	1	2	1,2	10
3	$2x + \cos(x^2+y)$	2	3	1,4	10
4	$1,5y + \sin(y^2+0,7x)$	1	2	1,6	10
5	$\exp(-x-y) + 0,5y^2$	0,5	1	2	10
6	$x + \sqrt{y^2+1,5x^2}$	0	1	0,5	10
7	$0,5x + \sqrt[3]{2x+y^5}$	0	1	0,4	10
8	$xy + \sqrt[3]{x^2+0,125y}$	0	1	1	10
9	$2x + \sqrt{1+y^2+x^2}$	0	1	2	10
10	$(1-y^2)\cos x + 0,5xy$	0	1	0	10
11	$1 + (1-x)\sin y - (2+x)y$	0	1	0	10
12	$y\cos^2(y-0,1x) + 0,5(x^2+1)$	0	1	0	10
13	$\cos(0,5x+y)+x-y$	0	1	0	10
14	$\frac{\cos y}{1+x} + xy^2$	0	1	0	10
15	$\exp(-1-xy) + x^2 + y$	0	1	0	10
16	$xy + \frac{1}{\sqrt{x+y}}$	1	2	1	10
17	$y(y \ln(x-0,5))/x$	1	1,5	0,5	10
18	$2+0,1y \sin x - 0,5y^2$	0	1	0	10
19	$xy + y^2 + \sin(2-x)$	0	1	0,1	10
20	$x^2 + 0,1y^2 + \cos xy$	0	1	0,2	10
21	$xy + x^2 + \cos y$	0	1	0,3	10
22	$2\exp(-x^2-4x)-2y(x-2)$	1	3	10	10

Продолжения табл. 7.5

23	$3\exp(-x^2+2x)-3y(x-3)$	1	4	10	10
24	$1+x \ln y - y \ln x$	1	2	1	10
25	$(1+y^2)\exp(-0,8x)+x$	0	1	0	10
26	$0,25\exp(-\cos(x+1))-2y(x-1)$	1	3	4	10
27	$1 + \frac{0,6y}{2+x} - \sin(y+2x+0,5)$	0	1	0	10
28	$\sin^2(0,6x)\exp(-x^2+2x)-2y(x-2)$	1	5	8	10
29	$1 + y \sin x - 0,8y^2$	0	1	0	10
30	$\cos(0,8x)\exp(-x^2)-4y(x-3)$	1	5	8	10

Якщо серед значень x_i і y_i є від'ємні, то завжди можна знайти такі додатні числа p і q , що $\bar{x}_i = x_i + p > 0$ і $\bar{y}_i = y_i + q > 0$ ($i = 1, 2, \dots, n$).

Тому розв'язування поставленої задачі завжди можна звести до побудови емпіричної формули для додатних значень (\bar{x}_i, \bar{y}_i) .

Побудова лінійної емпіричної формули. Нехай між даними (x_i, y_i) ($i = 1, 2, \dots, n$) існує лінійна залежність. Шукатимемо емпіричну формулу у вигляді

$$y = ax + b, \quad (8.4)$$

де коефіцієнти a і b невідомі.

Знайдемо значення a і b , за яких функція $S(a, b) = \sum_{i=1}^n (y_i - ax_i - b)^2$ матиме мінімальне значення. Щоб знайти ці значення, прирівняємо до нуля частинні похідні функції $S(a, b)$

$$\begin{cases} \frac{\partial S}{\partial a} = 2 \sum_{i=1}^n (y_i - ax_i - b)(-x_i) = 0, \\ \frac{\partial S}{\partial b} = 2 \sum_{i=1}^n (y_i - ax_i - b)(-1) = 0. \end{cases}$$

Звідси, врахувавши, що $\sum_{i=1}^n b = nb$, маємо

$$\begin{cases} a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i, \\ a \sum_{i=1}^n x_i + nb = \sum_{i=1}^n y_i. \end{cases} \quad (8.5)$$

Розв'язавши відносно a і b останню систему, знайдемо

$$a = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \quad (8.6)$$

$$b = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}. \quad (8.7)$$

Зазначимо, що, крім графічного, є ще й аналітичний критерій виявлення лінійної залежності між значеннями x і y .

Покладемо $\Delta x_i = x_{i+1} - x_i$, $\Delta y_i = y_{i+1} - y_i$, $k_i = \Delta y_i / \Delta x_i$ ($i = 1, 2, \dots, n-1$).

Якщо $k_i = \text{const}$, то залежність між x і y лінійна, бо точки (x_i, y_i) лежатимуть на одній прямій. Якщо $k_1 \approx k_2 \approx \dots \approx k_{n-1}$, то між x і y існує майже лінійна залежність, оскільки точки (x_i, y_i) лежатимуть близько до деякої прямої.

П р и к л а д 1. Побудувати лінійну функцію $y = ax + b$ для залежності, поданої таблично.

Таблиця 8.2

x	0,6	0,8	1,1	1,4	1,8	2,0
y	0,194	0,603	1,213	1,788	2,621	2,981

Р о з в ' я з а н н я . Перевіримо залежність на лінійність: $k_1 = 2,045$; $k_2 = 2,033$; $k_3 = 1,917$; $k_4 = 2,083$; $k_5 = 1,8$. Усі обчислені k_i ($i=1, 2, \dots, 5$) відрізняються одне від одного на величину, яка не перевищує 0,3. Отже, для даної залежності можна будувати лінійну емпіричну формулу.

Для обчислення коефіцієнтів a і b за формулами (8.6) і (8.7) складемо табл. 8.3.

Таблиця 8.3

i	x_i	y_i	$x_i y_i$	x_i^2	δ
1	0,6	0,194	0,1164	0,36	-0,007
2	0,8	0,603	0,4824	0,64	-0,002
3	1,1	1,213	1,3343	1,21	0,013
4	1,4	1,788	2,5032	1,96	-0,011
5	1,8	2,621	4,7178	3,24	0,023
6	2,0	2,981	5,962	4	-0,016
	$\Sigma=7,7$	$\Sigma=9,4$	$\Sigma=15,1161$	$\Sigma=11,41$	$\Sigma \delta^2=0,00116$

Значення коефіцієнтів дорівнюють: $a = 1,9974 \approx 1,997$, $b = -0,9967 \approx -0,997$. Отже, шуканою прямою є $y = 1,997x - 0,997$.

За цією формулою обчислюємо значення функції y і знаходимо нев'язки, які також записуємо в таблиці 8.3. Сума квадратів нев'язок $\sum_{i=1}^6 \delta_i^2 \approx 0,0012$. Щоб перевірити правильність проведених обчислень, перевіримо контрольну суму

$$\sum_{i=1}^n y_i^2 = \sum_{i=1}^n \delta_i^2 + b \sum_{i=1}^n y_i + a \sum_{i=1}^n x_i y_i. \quad (8.8)$$

Щоб вивести формулу (8.8), знайдемо вираз для суми квадратів нев'язок.

$$\sum_{i=1}^n \delta_i^2 = \sum_{i=1}^n (y_i - ax_i - b)^2 = \sum_{i=1}^n y_i^2 + a^2 \sum_{i=1}^n x_i^2 + \sum_{i=1}^n b^2 + 2ab \sum_{i=1}^n x_i - 2a \sum_{i=1}^n x_i y_i - 2b \sum_{i=1}^n y_i.$$

Перегрупувавши члени і винісши a і b за дужки, маємо

$$\sum_{i=1}^n \delta_i^2 = \sum_{i=1}^n y_i^2 + a(a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i - 2 \sum_{i=1}^n x_i y_i) + b(a \sum_{i=1}^n x_i + nb - 2 \sum_{i=1}^n y_i).$$

Врахувавши рівняння нормальної системи (8.5), знайдемо

$$\sum_{i=1}^n \delta_i^2 = \sum_{i=1}^n y_i^2 - a \sum_{i=1}^n x_i y_i - b \sum_{i=1}^n y_i.$$

звідки дістанемо формулу (8.8). Перевіримо тепер за допомогою формули (8.8) результати, обчислені в прикладі 1.

Справді,

$$\sum_{i=1}^6 y_i^2 = 20,82556, \quad \sum_{i=1}^6 \delta_i^2 + b \sum_{i=1}^6 y_i + a \sum_{i=1}^6 x_i y_i = 20,82507.$$

Отже, на підставі формули (8.8) виконані вище обчислення правильні, оскільки різниця в результатах цілком допустима за такої точності обчислень.

Побудова квадратичної емпіричної залежності. Нехай функціональна залежність між x та y — квадратична. Шукатимемо емпіричну формулу у вигляді

$$y = ax^2 + bx + c. \quad (8.9)$$

Тоді формулу (8.2) запишемо так:

$$S(a, b, c) = \sum_{i=1}^n (y_i - ax_i^2 - bx_i - c)^2.$$

Для знаходження коефіцієнтів a , b , c , за яких функція $S(a, b, c)$ мінімальна, обчислимо частинні похідні $\frac{\partial S}{\partial a}$, $\frac{\partial S}{\partial b}$, $\frac{\partial S}{\partial c}$ і прирівняємо їх до нуля. В результаті дістанемо систему рівнянь

$$\begin{cases} \sum_{i=1}^n (y_i - ax_i^2 - bx_i - c) \cdot x_i^2 = 0, \\ \sum_{i=1}^n (y_i - ax_i^2 - bx_i - c) \cdot x_i = 0, \\ \sum_{i=1}^n (y_i - ax_i^2 - bx_i - c) = 0. \end{cases}$$

Після рівносильних перетворень маємо систему

$$\begin{cases} a \sum_{i=1}^n x_i^4 + b \sum_{i=1}^n x_i^3 + c \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i^2 y_i, \\ a \sum_{i=1}^n x_i^3 + b \sum_{i=1}^n x_i^2 + c \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i, \\ a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i + nc = \sum_{i=1}^n y_i. \end{cases} \quad (8.10)$$

Розв'язок цієї системи і визначає єдину параболу, яка краще від усіх інших парабол (8.9) подає на розглядуваному проміжку задану таблично функціональну залежність.

Формула для контролю обчислень, яку виводимо аналогічно до формули (8.8), має вигляд

$$\sum_{i=1}^n y_i^2 = a \sum_{i=1}^n x_i^2 y_i + b \sum_{i=1}^n x_i y_i + c \sum_{i=1}^n y_i + \sum_{i=1}^n \delta_i^2 \quad (8.11)$$

Сформулюємо аналітичний критерій для квадратичної залежності. Для цього введемо поділені різниці першого й другого порядку

$$[x_i, x_{i+1}] = \frac{\Delta y_i}{\Delta x_i}$$

і

$$[x_i, x_{i+1}, x_{i+2}] = \frac{[x_{i+1}, x_{i+2}] - [x_i, x_{i+1}]}{x_{i+2} - x_i} = \frac{\Delta(\Delta y_i / \Delta x_i)}{\Delta_1 x_i},$$

де

$$\Delta_1 x_i = x_{i+2} - x_i = \Delta x_i + \Delta x_{i+1}.$$

Точки (x_i, y_i) розміщені на параболі (8.9) тоді і тільки тоді, коли всі поділені різниці другого порядку зберігають сталі значення.

Якщо точки x_i ($i=1, 2, \dots, n$) рівновіддалені, тобто $\Delta x_i = h = \text{const}$, то для існування квадратичної залежності (8.9) необхідно й достатньо, щоб була сталою скінченна різниця другого порядку

$$\Delta^2 y_i = y_{i+2} - 2y_{i+1} + y_i \quad (i=1, 2, \dots, n-2),$$

причому $\Delta^2 y_i = 2h^2 a$.

Приклад 2. Методом найменших квадратів побудувати квадратичну функцію $y = ax^2 + bx + c$ для залежності, поданої в табл. 8.4.

Таблиця 8.4

i	x_i	y_i	\bar{y}_i	x_i^2	x_i^3	x_i^4	$x_i \cdot \bar{y}_i$	$x_i^2 \cdot \bar{y}_i$
1	0	-4.281	0.019	0	0	0	0	0
2	0.2	-4.117	0.183	0.04	0.008	0.0016	0.0366	0.0073
3	0.4	-3.755	0.545	0.16	0.064	0.0256	0.218	0.0872
4	0.6	-3.195	1.105	0.36	0.216	0.1296	0.663	0.3978
5	0.8	-2.437	1.863	0.64	0.512	0.4096	1.4904	1.1923
6	1.0	-1.481	2.819	1	1	1	2.819	2.819
7	1.2	-0.325	3.975	1.44	1.728	2.0736	4.77	5.724
8	1.4	1.028	5.328	1.96	2.744	3.8416	7.4592	10.4429
9	1.6	2.581	6.881	2.56	4.096	6.5536	11.0096	17.6154
10	1.8	4.331	8.631	3.24	5.832	10.4976	15.5358	27.9644
11	2.0	6.278	10.578	4	8	16	21.156	42.312
Σ	11	-5.373	41.927	15.4	24.2	40.5328	65.1576	108.56232

Таблиця 8.5

i	x_i	y_i	Δy	$\Delta^2 y$
1	0	-4.281	0.164	
2	0.2	-4.117	0.362	0.198
3	0.4	-3.755	0.560	0.198
4	0.6	-3.195	0.758	0.198
5	0.8	-2.437	0.956	0.198
6	1.0	-1.481	1.156	0.2
7	1.2	-0.325	1.353	0.197
8	1.4	1.028	1.553	0.2
9	1.6	2.581	1.750	0.197
10	1.8	4.331	1.947	0.147
11	2.0	6.278		

Розв'язання. Оскільки точки x_i ($i = 1, 2, \dots, 11$) рівновіддалені, причому $h = 0,2$, то складемо табл. 8.5 скінченних різниць до другого порядку.

З табл. 8.5 видно, що скінченні різниці другого порядку майже сталі, тому залежність між x і y можна вважати квадратичною.

Серед значень y_i є від'ємні. Тому до всіх y_i ($i = 1, 2, \dots, 11$) додамо, наприклад, число 4,3. Це означає, що початок координат перенесено в точку $(0; 4,3)$. Нові значення координат $\bar{y}_i = y_i + 4,3$ заносимо в табл. 8.4. Щоб побудувати систему (8.10), обчислимо суми при невідомих a, b, c . Проміжні обчислення запишемо в табл. 8.4.

Шукана нормальна система має вигляд

$$\begin{cases} 40,5328a + 24,2b + 15,4c = 108,5623, \\ 24,2a + 15,4b + 11c = 65,1576, \\ 15,4a + 11b + 11c = 41,927. \end{cases}$$

Розв'язування цієї системи мето-

дом Гаусса дає: $a = 2.47909 \approx 2.479$, $b = 0.32149 \approx 0.321$, $c = 0.0193 \approx 0.019$.

Тоді шуканий тричлен у системі координат xOy має вигляд:

$$y = 2,479x + 0,321x - 4,281. \quad (8.12)$$

Коефіцієнти a, b, c обчислюємо так, щоб знайдена формула забезпечувала точність, яку мають вихідні дані.

Тепер за формулою (8.12) визначаємо значення $y = y_{обч}$ і знаходимо нев'язки $\delta = y - y_{обч}$. В результаті обчислень маємо

$$\sum_{i=1}^{11} \delta_i^2 = 0,000015; \quad \sum_{i=1}^{11} y_i = -5,373; \quad \sum_{i=1}^{11} y_i^2 = 133,7118;$$

$$\sum_{i=1}^{11} x_i y_i = 17,8576; \quad \sum_{i=1}^{11} x_i y_i^2 = 42,3423,$$

$$a \sum_{i=1}^{11} x_i^2 y_i + b \sum_{i=1}^{11} x_i y_i + c \sum_{i=1}^{11} y_i + \sum_{i=1}^{11} \delta_i^2 = 133,7116.$$

Звідси видно, що рівність (8.12) у межах точності вихідних даних задовольняється.

Побудова емпіричних формул найпростіших нелінійних залежностей. Нехай у системі координат xOy маємо нелінійну залежність $y = F(x, a, b)$, неперервну і монотонну на відрізку $[x_1; x_n]$.

Введемо змінні $X = \varphi(x)$, $Y = \Psi(y)$ так, щоб у новій системі координат XOY задана емпірична нелінійна залежність стала лінійною

$$Y = AX + B, \quad (8.13)$$

Тоді точки з координатами $(\varphi(x_i), \Psi(y_i))$ в площині XOY лежатимуть на прямій лінії.

Покажемо, як від нелінійних залежностей

$$\begin{aligned} 1) y = ax^b, \quad 2) y = ab^x, \quad 3) y = a \ln x + b, \\ 4) y = \frac{a}{x} + b, \quad 5) y = \frac{1}{ax + b}, \quad 6) y = \frac{x}{ax + b} \end{aligned}$$

перейти до лінійних.

1) Розглянемо степеневу залежність $y = ax^b$, де $x > 0$, $a > 0$, $b > 0$.

Логарифмуючи її, знаходимо $\ln y = b \ln x + \ln a$. Звідси, поклавши $X = \ln x$, $Y = \ln y$, $B = \ln a$, $A = b$, маємо $Y = AX + B$.

2) Логарифмуючи показникову залежність $y = ab^x$, маємо $\ln y = \ln a + x \ln b$. Поклавши $Y = \ln y$, $X = x$, $A = \ln b$, $B = \ln a$ в системі координат XOY дістанемо залежність (8.13).

Зазначимо, що замість показникової залежності $y = ab^x$ часто шукують залежність $y = ae^{bx}$. Остання перетвориться в лінійну, якщо позначити $X = x$, $Y = \ln y$, $B = \ln a$, $A = b$.

3) Щоб перейти від логарифмічної залежності $y = a \ln x + b$ до лінійної $Y = aX + b$, досить зробити підстановку $Y = y$, $X = \ln x$.

4) У гіперболічній залежності замінимо змінні $1/x = X$, $y = Y$. Тоді

гіперболічна залежність перетвориться в лінійну (8.13), в якій $A = a$, $B = b$.

5) Розглянемо дробово-лінійну функцію $y = \frac{1}{ax + b}$. Знайдемо обернену функцію $1/y = ax + b$. Тоді, ввівши нові координати $Y = 1/y$, $X = x$, дістанемо лінійну залежність (8.13), де $A = a$, $B = b$.

6) Нехай маємо дробово-раціональну залежність $y = \frac{x}{ax + b}$. Оберненою до неї буде залежність $1/y = a + b/x$. Ввівши нові змінні $Y = 1/y$, $X = 1/x$, дістанемо лінійну залежність (8.13) з коефіцієнтами $A=b$, $B=a$.

Отже, для побудови будь-якої з емпіричних формул 1)-6) треба:

а) за вихідною таблицею даних (x_i, y_i) побудувати нову таблицю (X_i, Y_i) , використавши відповідні формули переходу до нових координат;

б) за новою таблицею даних знайти методом найменших квадратів коефіцієнти A і B лінійної функції (8.13);

в) за відповідними формулами знайти коефіцієнти a і b даної нелінійної залежності.

Вибрати емпіричну формулу для нелінійних залежностей графічним методом часто буває важко. Тоді вдаються до перевірки аналітичних критеріїв існування певної залежності. Для цього зводять її до лінійної і перевіряють виконання критерію лінійної залежності між перетвореними вихідними даними (X_i, Y_i) . Але є й власні аналітичні критерії наявності кожної з розглянутих вище нелінійних залежностей. Найпростіші необхідні умови їх наявності подано в табл. 8.6.

Таблиця 8.6

№ пор.	Емпірична формула	\bar{x}_s	\bar{y}_s	Спосіб порівнявання
1	$y = ax + b$	$\frac{x_1 + x_n}{2}$	$\frac{y_1 + y_n}{2}$	
2	$y = ax^p$	$\sqrt{x_1 \cdot x_n}$	$\sqrt{y_1 \cdot y_n}$	$Y = AX + B$, де $X = \ln x$, $Y = \ln y$, $A = b$, $B = \ln a$
3	$y = ab^x$	$\frac{x_1 + x_n}{2}$	$\sqrt{y_1 \cdot y_n}$	$Y = AX + B$, де $Y = \ln y$, $A = \ln b$, $B = \ln a$
4	$y = a \ln x + b$	$\sqrt{x_1 \cdot x_n}$	$\frac{y_1 + y_n}{2}$	$y = aX + b$, де $X = \ln x$
5	$y = \frac{a}{x + b}$	$\frac{2x_1 x_n}{x_1 + x_n}$	$\frac{y_1 + y_n}{2}$	$y = aX + b$, де $X = 1/x$
6	$y = \frac{1}{ax + b}$	$\frac{x_1 + x_n}{2}$	$\frac{2y_1 y_n}{y_1 + y_n}$	$Y = ax + b$, де $Y = 1/y$
7	$y = \frac{x}{ax + b}$	$\frac{2x_1 x_n}{x_1 + x_n}$	$\frac{2y_1 y_n}{y_1 + y_n}$	$Y = bX + a$, де $X = 1/x$, $Y = 1/y$

Умови перевіряють у такий спосіб. На заданому відрізку зміни незалежної змінної x вибирають дві точки, досить надійні і розміщені якомога далі одна від одної. Нехай, наприклад, це будуть точки x_1, x_n . Потім, залежно від типу емпіричної формули, що перевіряється, обчислюють значення \bar{x}_s , яке є або середнім арифметичним, або середнім геометричним, або середнім гармонічним значень x_1, x_n . Маючи значення y_1 і y_n , аналогічно обчислюють і відповідне значення \bar{y}_s . Далі, користуючись даною таблицею значень (x_i, y_i) ($i = 1, 2, \dots, n$), для значення \bar{x}_s знаходять відповідне йому значення y_s^* . Якщо \bar{x}_s немає в таблиці, то y_s^* знаходять наближено з побудованого графіка даної залежності або за допомогою лінійної інтерполяції

$$y_s^* = y_1 + \frac{y_{i+1} - y_i}{x_{i+1} - x_i} (\bar{x}_s - x_i),$$

де x_i і x_{i+1} — проміжні значення, між якими лежить \bar{x}_s ($x_i < \bar{x}_s < x_{i+1}$). Обчисливши y_s^* , знаходять величину $|\bar{y}_s - y_s^*|$. Якщо ця величина велика, то відповідна емпірична формула не придатна для апроксимації заданих табличних даних. З кількох придатних емпіричних формул перевагу надають тій, для якої відхилення $|\bar{y}_s - y_s^*|$ якомога менше.

П р и к л а д 3. Знайти і побудувати емпіричну формулу для залежності, яку подано в табл. 8.7.

Таблиця 8.7

x	8	10	15	20	30	40	60	80
y	13,8	14,0	15,4	16,3	17,2	17,8	18,5	18,8

Р о з в ' я з а н н я. Нехай крайні табличні значення досить надійні. Користуючись табл. 8.6, проведемо відповідні обчислення і занесемо їх у табл. 8.8.

Таблиця 8.8

№ пор.	Емпірична формула	\bar{x}_s	\bar{y}_s	y_s^*	$ \bar{y}_s - y_s^* $	Висновок
1	$y = ax + b$	44	15,9	18,08	2,18	не придатна
2	$y = ax^p$	25,298	15,633	16,777	1,144	не придатна
3	$y = ab^x$	44	15,633	18,08	2,447	не придатна
4	$y = a \ln x + b$	25,298	15,9	16,777	0,877	мало придатна
5	$y = a/x + b$	14,545	15,9	15,273	0,627	мало придатна
6	$y = 1/(ax + b)$	44	15,371	18,08	2,709	не придатна
7	$y = x/(ax + b)$	14,545	15,371	15,273	0,098	найбільш придатна

Значення y_i обчислювались за формулою лінійної інтерполяції. Таблиця 8.8 показує, що для апроксимації вихідних даних найбільш придатна дробово-раціональна функція $y = x / (ax + b)$. Ввівши нові змінні $X = 1/x$, $Y = 1/y$, зведемо її до лінійної $Y = AX + B$, де $A = b$, $B = a$.

Значення $X = 1/x$ і $Y = 1/y$ помістимо в таблиці 8.9. Для знаходження коефіцієнтів A і B виконаємо проміжні обчислення і занесямо їх в табл. 8.9.

Таблиця 8.9

i	x_i	y_i	$x_i \cdot y_i$	x_i^2	δ_i
1	0,125	0,0769	0,0096	0,0156	0,00001
2	0,1	0,0714	0,0071	0,01	-0,00027
3	0,667	0,0649	0,0043	0,0044	0,00019
4	0,05	0,0613	0,0031	0,0025	0,00008
5	0,0333	0,0581	0,0019	0,0011	0,00035
6	0,025	0,0562	0,0014	0,0006	0,00013
7	0,0177	0,0541	0,0009	0,0003	-0,00026
8	0,0125	0,0532	0,0006	0,0002	-0,00025
Σ	0,4292	0,4962	0,0291	0,0347	$3,85 \cdot 10^{-7}$

Коефіцієнти A і B обчислимо за формулами (8.6) і (8.7). Маємо $A = 0,2086$, $B = 0,0508$.

Обчисливши нев'язки $\delta_i = |Y_i - AX_i - B|$, знайдемо $S = \sum_{i=1}^8 \delta_i^2 = 3,85 \cdot 10^{-7}$.

Врахувавши, що $a = B$, $b = A$, шукана емпірична формула має вигляд $y = x / (0,0508x + 0,2086)$.

Сума квадратів відхилень між вихідними значеннями y і значеннями, обчисленими за останньою формулою, дорівнює 0,0335. Алгоритм побудови емпіричної формули за методом найменших квадратів подано в програмі 8.1.

Програма 8.1

```

10 REM ----- Побудова емпіричної формули -----
20 REM ----- методом найменших квадратів -----
30 INPUT "Ввести кількість пар заданих точок"; N
40 DIM X(N), Y(N), XX(N), YY(N), YE(N)
50 PRINT "задати парами X і Y через кому"
60 FOR I=1 TO N
70 INPUT X(I), Y(I)

```

```

80 NEXT I
90 PRINT "Вибрати номер формули"
100 PRINT " 1) y=a*x+b          5) y=a*log(x)+b"
110 PRINT " 2) y=a*x^b          6) y=a/x+b"
120 PRINT " 3) y=a*b^x          7) y=1/(a*x+b)"
130 PRINT " 4) y=a*exp(b*x)       8) y=x/(a*x+b)"
140 PRINT: INPUT "Ввести номер формули"; K$
150 P=INSTR("12345678", K$)
160 IF P=0 THEN 140
170 K=VAL(K$)
180 SX=0: SY=0: S1=0: S2=0
185 FOR I=1 TO N
190 ON K GOTO 200, 210, 220, 230, 240, 250, 260
200 XX(I)=X(I): YY(I)=Y(I):GOTO 270
210 XX(I)=LOG(X(I)): YY(I)=LOG(Y(I)):GOTO 270
220 XX(I)=X(I): YY(I)=LOG(Y(I)):GOTO 270
230 XX(I)=LOG(X(I)): YY(I)=Y(I):GOTO 270
240 XX(I)=1/X(I): YY(I)=Y(I):GOTO 270
250 XX(I)=X(I): YY(I)=1/Y(I): GOTO 270
260 XX(I)=1/X(I): YY(I)=1/Y(I)
270 SX=SX+XX(I): SY=SY+YY(I)
280 S1=S1+XX(I)*YY(I): S2=S2+XX(I)^2
290 NEXT I
300 Z=N*S2-SX^2
310 AA=(N*S1-SX*SY)/Z: BB=(SY*S2-S1*SX)/Z
320 ON K GOTO 330, 340, 350, 340, 340, 340, 330, 360
330 A=AA: B=BB: GOTO 370
340 A=EXP(BB): B=AA: GOTO 370
350 A=EXP(BB): B=EXP(AA): GOTO 370
360 A=BB: B=AA
370 FOR I=1 TO N
380 ON K GOTO 390, 400, 410, 420, 430, 440, 450, 460
390 YE(I) = A*X(I)+B: GOTO 470
400 YE(I)=A*X(I)^B:GOTO 470
410 YE(I)=A*B^X(I): GOTO 470
420 YE(I)=A*EXP(B*X(I)): GOTO 470
430 YE(I)=A*LOG(X(I))+B: GOTO 470
440 YE(I)=A/X(I)+B: GOTO 470
450 YE(I)=1/(A*X(I)+B): GOTO 470
460 YE(I)=X(I)/(A*X(I)+B)
470 S=S+(Y(I)-YE(I))^2
480 NEXT I
485 PRINT "k=" K; "aa=" AA;"bb=" BB
490 PRINT "a=" A; "b=" B
500 PRINT "Сума квадратів нев'язок =" S
510 END

```

§ 8.2. РОЗПОДІЛИ ВИПАДКОВИХ ВЕЛИЧИН

На практиці часто доводиться мати справу з величинами, які набувають певних значень, але заздалегідь не можна достовірно передбачити, яких саме. Такі величини називають *випадковими*. Вони характеризують кількісно можливі результати досліду чи спостереження. Випадкові величини діляться на *дискретні* і *неперервні*.

Випадкову величину називають *дискретною*, якщо множина її значень скінченна або зчисленна. Випадкову величину називають *неперервною*, якщо вона може набувати будь-яких значень з деякого скінченного або нескінченного інтервалу.

Щоб повністю характеризувати випадкову величину з імовірнісної точки зору, треба знати не лише значення випадкової величини, а й імовірність її появи. Якщо для дискретної випадкової величини X відомі всі можливі її значення x_1, x_2, \dots, x_n і відповідні ймовірності p_1, p_2, \dots, p_n , то кажуть, що задано теоретичний закон розподілу випадкової величини X .

Якщо певні значення випадкової величини X утворюють повну групу n несумісних подій, то сума ймовірностей їх появи дорівнює одиниці

$$\sum_{i=1}^n p_i = 1.$$

Розподіл дискретної випадкової величини задають таблицею, яку називають *рядом розподілу*. Розподіл неперервної випадкової величини можна задавати або функцією, або щільністю розподілу. Функція розподілу визначає ймовірність того, що випадкова величина X набула значення, меншого від фіксованого дійсного числа x , тобто

$$F(x) = P(X < x).$$

Зазначимо, що для дискретної випадкової величини X , заданої рядом розподілу, завжди можна знайти функцію розподілу $F(x)$, підсумувавши ймовірності значень, менші від x . Функцію розподілу називають ще *інтегральною функцією*, або *інтегральним законом розподілу*.

Границя відношення ймовірності попадання неперервної величини X з функцією розподілу $F(x)$ в інтервал $(x, x + \Delta x)$ до довжини цього інтервалу, коли остання прямує до нуля, є похідною від функції розподілу

$$\lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = F'(x). \quad (8.14)$$

Функцію (8.14) називають *щільністю розподілу*, її позначають $f(x)$. Вона характеризує щільність, з якою розподіляються ймовірності значень випадкової величини на деякому досить малому інтервалі. Щільність розподілу називають також *диференціальною функцією*, або *диференціальним законом розподілу*.

Розподіл імовірностей повністю характеризує випадкову величину. Але здебільшого закон її розподілу невідомий, тому зручніше користуватися деякими кількісними характеристиками, які відображають найістотніші властивості даної випадкової величини. Основними числовими характеристиками випадкової величини є *математичне сподівання*, *дисперсія*, *середнє квадратичне відхилення* тощо.

Математичне сподівання характеризує положення випадкової величини на числовій осі, визначаючи собою центр розподілу, або центр розсіяння. Математичне сподівання $M(X)$ дискретної випадкової величини X дорівнює сумі добутків окремих значень, яких набуває X , на їх відповідні ймовірності:

$$M(X) = \sum_{i=1}^n x_i p_i.$$

Математичне сподівання неперервної випадкової величини X , заданої диференціальною функцією $f(x)$, визначають за формулою

$$M(X) = \int_{-\infty}^{\infty} x f(x) dx,$$

а математичне сподівання неперервної випадкової величини X , яка набуває значень з інтервалу $(a; b)$, дорівнює

$$M(X) = \int_a^b x f(x) dx.$$

Другою важливою характеристикою випадкової величини є *дисперсія*, яка характеризує ступінь розсіяння значень випадкової величини навколо її центра. *Дисперсією* називають математичне сподівання квадрата відхилення випадкової величини від її математичного сподівання.

Для дискретної випадкової величини дисперсію обчислюють за формулою

$$D(X) = \sum_{i=1}^n (x_i - M(X))^2 \cdot p_i,$$

а для неперервної випадкової величини, розподіл якої задано щільністю ймовірності $f(x)$, — за формулою

$$D(X) = \int_{-\infty}^{\infty} (x - M(X))^2 f(x) dx.$$

Недолік дисперсії в тому, що вона має розмірність квадрата випадкової величини і нею незручно користуватися для характеристики розсіяння. Цього недоліку не має *середнє квадратичне відхилення* випадкової величини, яке обчислюється за формулою

$$\sigma(X) = \sqrt{D(X)}.$$

Дискретні випадкові величини характеризуються біномним розподілом та розподілом Пуассона.

Неперервні випадкові величини характеризуються нормальним законом розподілу, розподілами χ^2 (хі-квадрат), Стюдента, Фішера та деякими іншими.

Нормальний закон розподілу. Неперервні випадкові величини найчастіше розподілені за нормальним законом розподілу. Цьому закону підпорядковані похибки вимірювань, різноманітні властивості живих організмів (наприклад зріст людини), дальність польоту снаряда під час стрільби та багатьох інших кількісних ознак.

Неперервна випадкова величина X має нормальний закон розподілу, якщо щільність розподілу задано виразом

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (8.15)$$

З (8.15) випливає, що $f(x) > 0$ і $\int_{-\infty}^{\infty} f(x) dx = 1$. Графік щільності розподілу має вигляд, зображений на рис. 8.1; його називають *нормальною кривою*. Графік кривої симетричний відносно прямої $x = \mu$, максимальна ордината графіка (для $x = \mu$) дорівнює $1/(\sigma\sqrt{2\pi})$ і вісь абсцис є асимптотой цієї кривої. Оскільки $M(X) = \int_{-\infty}^{\infty} xf(x) dx = \mu$, то параметр μ є ма-

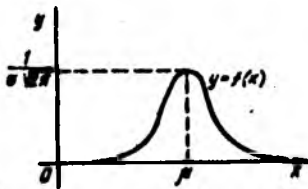


Рис. 8.1

тематичним сподіванням неперервної випадкової величини X . З іншого боку

$D(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx = \sigma^2$, тобто параметр σ є середнім квадратичним відхиленням величини X .

Форма кривої щільності розподілу не змінюється від зміни величини параметра μ . Якщо математичне сподівання зменшується чи збільшується, то графік кривої зсувається вліво чи вправо.

Якщо параметр σ змінюється, то змінюється й форма кривої. Із зростанням σ максимальна ордината кривої розподілу зменшується, а із зменшенням — зростає. Вся площа, обмежена кривою розподілу і віссю абсцис, дорівнює одиниці. Тому із зростанням σ крива розподілу стискається до осі абсцис і розтягується вздовж неї; із зменшенням σ крива розподілу розтягується вздовж осі ординат.

Щільність розподілу з параметрами $\mu = 0$ і $\sigma = 1$ називають *нормованою щільністю*, а її графік — *нормованою нормальною кривою*. Нормована нормальна крива — це крива розподілу нормованої випадкової величини

$$T = \frac{X - \mu}{\sigma}.$$

З нормальним розподілом тісно пов'язані розподіли χ^2 (хі-квадрат), Стюдента і Фішера.

Розподіл χ^2 (хі-квадрат). Нехай t_1, t_2, \dots, t_ν — незалежні випадкові величини, кожна з яких має нормований нормальний закон розподілу з параметрами $\mu = 0$ і $\sigma = 1$. Тоді випадкова величина

$$\chi^2 = t_1^2 + t_2^2 + \dots + t_\nu^2,$$

тобто сума квадратів незалежних випадкових нормованих величин t_i ($i = 1, 2, \dots, \nu$), кожна з яких підпорядкована нормальному закону розподілу, називається *випадковою величиною з розподілом χ^2* і з ν степенями вільності.

Щільність розподілу цієї випадкової величини задається функцією

$$f(x) = \begin{cases} 0, & \text{якщо } x \leq 0, \\ \frac{1}{\Gamma(\frac{\nu}{2})2^{\nu/2}} e^{-\frac{x}{2}} x^{\frac{\nu}{2}-1}, & \text{якщо } x > 0. \end{cases} \quad (8.16)$$

де ν — параметр розподілу, а $\Gamma(\nu/2)$ — значення гамма-функції, яка для будь-якого натурального числа n визначається так: $\Gamma(n) = \int_0^{\infty} x^{n-1} e^{-x} dx$. Розподіл χ^2 залежить лише від одного параметра — числа степенів вільності ν .

Розподіл Стюдента. Нехай U — нормована випадкова величина з нормальним законом розподілу, а V — випадкова величина з розподілом χ^2 з ν степенями вільності. Тоді випадкова величина

$$T = \frac{U}{\sqrt{V/\nu}}$$

має розподіл, який називають *t-розподілом*, або *розподілом Стюдента*.

Щільність розподілу Стюдента задається функцією

$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi} \Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}},$$

яка залежить лише від числа степенів вільності ν . Якщо $\nu \geq 30$,

то розподіл переходить у нормальний розподіл з параметрами $\mu = 0$ і

$$\sigma = \sqrt{\frac{\nu}{\nu - 2}}$$

Розподіл Фішера. Якщо випадкові величини U і V незалежні й кожна з них розподілена за законом χ^2 із степенями вільності відповідно ν_1 і ν_2 , тоді величина

$$F = \frac{U / \nu_1}{V / \nu_2}$$

має так званий розподіл Фішера (F -розподіл), який залежить від двох параметрів — ν_1 і ν_2 . F -розподіл використовують в основному в задачах, пов'язаних з дисперсіями.

§ 8.3. ДЕЯКІ ЧИСЛОВІ ХАРАКТЕРИСТИКИ ВИБІРКИ

Сукупність усіх можливих значень досліджуваної величини (ознаки) називають *генеральною сукупністю*. Генеральна сукупність може бути скінченною й нескінченною. Результати обмеженого ряду спостережень x_1, x_2, \dots, x_n випадкової величини називають *вибіркою* з генеральної сукупності. Кількість елементів вибіркової сукупності називають її *обсягом*. Окреме значення ознаки називають її *варіантою*. Число, яке показує, скільки разів зустрічається та чи інша варіанта, називають *частотою*. Сума всіх частот дорівнює обсягу сукупності. Щоб вивчити закономірності частоти появи варіант, їх розміщують у зростаючому або спадному порядку і вказують частоту появи кожної з варіант даної сукупності. При цьому дістають таблицю, яку називають *варіаційним рядом*, або *емпіричним розподілом*.

Нехай, наприклад, маємо 50 чисел, які є значеннями зросту (в см) студентів-юнаків третього курсу: 183, 170, 176, 178, 176, 180, 176, 185, 184, 174, 168, 174, 189, 172, 175, 176, 179, 176, 172, 178, 169, 171, 170, 177, 176, 179, 174, 176, 188, 178, 172, 176, 167, 166, 180, 183, 176, 182, 178, 172, 185, 183, 175, 174, 180, 166, 170, 171, 178, 169.

Оскільки під час комплектування лекційних потоків менше всього враховують зріст студентів, тому вибірку можна вважати випадковою. Тоді таблиця емпіричного розподілу матиме вигляд (табл. 8.10).

Таблиця 8.10

Варіанта	166	167	168	169	170	171	172	174	175	176
Частота	2	1	1	2	3	2	4	4	2	9
Варіанта	177	178	179	180	182	183	184	185	188	189
Частота	1	5	2	3	1	3	1	2	1	1

Дискретною варіацією ознаки називають таку, коли окремі значення варіанти відрізняються одне від одного на деяку скінченну величину. *Неперервна* — це така варіація, коли значення варіанти можуть відрізнитися одне від одного на будь-яку нескінченно малу величину. Розподіл неперервної варіації називають *інтервальним*. У цьому разі для побудови варіаційного ряду інтервал розподілу ділять на k рівних частин. Число k обчислюють за формулою

$$k = 1 + 3,32 \lg n,$$

де n — обсяг вибірки.

Довжина інтервалу

$$h = \frac{x_{\max} - x_{\min}}{k},$$

де x_{\max} і x_{\min} — відповідно максимальне і мінімальне значення варіанти. Якщо h — дробове число, то за величину інтервалу беруть найближче ціле число або найближчий нескладний дріб. До кожного інтервалу включають варіанти, більші від нижньої границі і не більші від верхньої. Нижню границю першого інтервалу визначають так, щоб мінімальна варіанта вибірки потрапляла приблизно в середину цього інтервалу. Потім підраховують частоти m_i для кожного інтервалу поділу. Сума всіх частот дорівнює кількості всіх значень ознаки, тобто обсягу вибірки. Часто інтервальний варіаційний ряд умовно замінюють дискретним. Тоді за значення варіанти x_i беруть середину інтервалу (x_i, x_{i+1}).

Для побудованого вище варіаційного ряду $k = 1 + 3,32 \lg 50 \approx 6,64$. Врахувавши, що $x_{\max} - x_{\min} = 189 - 166 = 23$, візьмемо $h = 23/6 \approx 4$. Тоді згрупований інтервальний ряд можна подати таблично (табл. 8.11).

Таблиця 8.11

№ пор.	1	2	3	4	5	6
Інтервал	165–169	169–173	173–177	177–181	181–185	185–189
Середина інтервалу	167	171	175	179	183	187
Частота	6	9	16	10	7	2

Варіаційні ряди дають уявлення про варіацію ознаки у вибірковій сукупності. Для повнішої характеристики вибірки використовують узагальнюючі числові характеристики положення, розсіяння та асиметрії емпіричного розподілу. Характеристики теоретичних розподілів можна розглядати як такі, що існують у генеральній сукупності, а характеристики емпіричних розподілів — як вибіркові характеристики. Характеристики розподілів імовірностей у генеральній сукупності називають ще *параметрами*, а вибіркові (емпіричні) — *оцінками*, або *статистиками*.

Нехай маємо експериментальні значення x_1, x_2, \dots, x_n неперервної випадкової величини (ознаки) X . Середнє значення ознаки визначають за формулою

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad (8.17)$$

тобто \bar{x} є *емпіричним*, або *вибірковим*, *середнім*. Воно є наближенням значенням (оцінкою) математичного сподівання $M(X)$ нормально розподіленої ознаки X генеральної сукупності.

Якщо за даними спостережень побудовано варіаційний ряд, то вводять *середню зважену*

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k m_i x_i, \quad (8.18)$$

де x_i — варіанта, якщо ряд дискретний, і центр інтервалу, якщо ряд інтервальний; m_i — частота варіанти x_i або статистична вага, k — кількість інтервалів.

Характеристикою розсіювання ознаки навколо середньої є *емпірична дисперсія*, яку обчислюють за формулою

$$S^2 = \sum_{i=1}^n (x_i - \bar{x})^2 / (n-1). \quad (8.19)$$

Зважена емпірична дисперсія дорівнює

$$S^2 = \frac{1}{n-1} \sum_{i=1}^k m_i (x_i - \bar{x})^2. \quad (8.20)$$

Якщо середнє значення і емпіричну дисперсію обчислюють одночасно, то формули (8.19) і (8.20) набирають відповідно вигляду

$$S^2 = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \right). \quad (8.21)$$

$$S^2 = \frac{1}{n-1} \left(\sum_{i=1}^k m_i x_i^2 - \frac{1}{n} \left(\sum_{i=1}^k m_i x_i \right)^2 \right). \quad (8.22)$$

Середнім квадратичним відхиленням називають корінь квадратний з дисперсії:

$$S = \sqrt{S^2}. \quad (8.23)$$

Середнє квадратичне відхилення має таку саму розмірність, що й значення ознаки і є абсолютною характеристикою коливання ознаки навколо середнього значення. Відносним показником розсіювання ознаки є *коефіцієнт варіації* V , який дорівнює відношенню середнього квадратичного відхилення до середнього

$$V = \frac{S}{\bar{x}}. \quad (8.24)$$

Коефіцієнт варіації можна обчислити і у відсотках:

$$V = \frac{S}{\bar{x}} \cdot 100\%.$$

Так, вибіркові характеристики наведеного вище емпіричного незгрупованого розподілу мають такі значення: $\bar{x} = 176,04$; $S^2 = 30,45$; $S = 5,52$; $V = 3\%$, а для згрупованого інтервального ряду — $\bar{x} = 175,72$; $S^2 = 28,53$; $S = 5,34$; $V = 3\%$.

§ 8.4. СТАТИСТИЧНА ПЕРЕВІРКА ГІПОТЕЗ

У наукових дослідженнях часто для встановлення справедливості певного факту висловлюють гіпотези, які можна перевірити статистично, тобто виходячи з результатів спостережень випадкової вибірки з генеральної сукупності.

Під *статистичною гіпотезою* розуміють будь-яке твердження щодо генеральної сукупності, яке перевіряється на основі вибірки. Статистичні гіпотези висловлюють як відносно законів розподілу, так і відносно параметрів розподілу. Наприклад, гіпотеза про те, що зріст двадцятирічних юнаків-студентів третього курсу підпорядкований нормальному закону розподілу, є гіпотезою про закон розподілу. Гіпотеза про те, що середні розміри деталей, які виготовляються на однотипних, паралельно працюючих станках, приблизно однакові, є гіпотезою про параметри розподілу.

Зроблений на основі статистичних даних висновок про те, що між кількома генеральними сукупностями або між емпіричним і теоретичним розподілом істотних відмінностей немає, називають *нульовою* (основною) гіпотезою. Гіпотезу, яка заперечує нульову, називають *альтернативною* гіпотезою. Нульову гіпотезу позначають буквою H_0 , альтернативну — H_a .

В результаті перевірки статистичної гіпотези, що ґрунтується на вибіркових спостереженнях, можна прийняти або відхилити нульову гіпотезу. Помилкове рішення можна допустити в двох випадках. Тому розрізняють помилки двох видів. Помилка першого виду полягає в тому, що нульова гіпотеза заперечується, тоді як насправді вона правильна. Помилка другого виду полягає в тому, що нульова гіпотеза приймається, тоді як правильною є альтернативна гіпотеза.

Ймовірність допустити помилку першого виду називають *рівнем значущості* і позначають α . Рівень значущості встановлює дослідник залежно від важливості досліджуваної задачі. Рівень значущості — це та мінімальна ймовірність, починаючи з якої можна вважати подію практично неможливою. Найчастіше її беруть такою, що дорівнює 0,1; 0,05; 0,01.

Є два типи задач перевірки гіпотез. Задачі першого типу пов'язані з перевіркою гіпотез про достовірність істотної відмінності між параметрами статистичних сукупностей. Відмінність між параметрами статистичних сукупностей вважають істотною, якщо вона перевищує ту, яку можна було б пояснити випадковими коливаннями. Прикладом задачі першого типу є, наприклад, оцінка достовірності істотної відмінності між дисперсіями двох вибірок або між їх середніми значеннями.

Задачі другого типу пов'язані з оцінкою ступеня розбіжності емпіричного й теоретичного розподілів. Прикладом задачі цього типу може бути перевірка гіпотези про те, що емпіричний розподіл узгоджується з нормальним законом розподілу.

Перевірка гіпотези полягає в тому, що за вибірковими даними обчислюють значення деякої величини, яка має відомий стандартний розподіл (нормальний, Стюдента, Пірсона й ін.). Цю величину називають *статистикою критерію* або просто *значенням критерію*.

Якщо обчислене за вибіркою значення критерію не перевищує граничного (критичного) значення, взятого з відповідних таблиць, то нульову гіпотезу визнають на заданому рівні значущості α . У цьому разі отриману за вибірковими даними відмінність можна пояснити тільки випадковістю вибірки. Але прийняття гіпотези зовсім не означає, що рівність параметрів генеральних сукупностей доведена, або те, що теоретичний закон відповідає емпіричному розподілу. Наявний статистичний матеріал не дає підстав для відхилення нульової гіпотези. Якщо обчислене значення критерію буде більше від критичного значення на заданому рівні значущості α , то відмінність генеральних сукупностей не можна пояснити тільки випадковістю вибірки. У цьому разі нульову гіпотезу відхиляють і кажуть, що на заданому рівні значущості відмінність істотна.

Розглянемо деякі задачі на перевірку гіпотез.

Порівняння двох вибірових дисперсій. Щоб перевірити гіпотезу про те, що дві незалежні вибірки, взяті з генеральних сукупностей X і Y , мають однакові дисперсії σ_x^2 і σ_y^2 , використовують F -критерій Фішера.

Умовою застосування F -критерію є припущення про те, що обидві вибірки незалежні і їх взято з нормально розподілених генеральних сукупностей з параметрами μ_x, σ_x і μ_y, σ_y .

F -критерій застосовують у такий спосіб.

1. Приймають припущення про нормальність розподілу генеральних сукупностей.

2. Формулюють нульову гіпотезу $H_0: \sigma_x^2 = \sigma_y^2$ і альтернативну гіпотезу $H_a: \sigma_x^2 \neq \sigma_y^2$.

3. Задають рівень значущості α .

4. Розглядають дві незалежні вибірки із сукупностей X і Y обсягом відповідно n_x і n_y .

5. Обчислюють значення вибірових дисперсій S_x^2 і S_y^2 . Більшу з них позначають S_1^2 , меншу — S_2^2 .

6. Знаходять значення F -критерію за формулою

$$F = S_1^2 / S_2^2. \quad (8.25)$$

7. З таблиці розподілу Фішера беруть критичне значення F_α для заданого рівня значущості α і числа степенів вільності $\nu_1 = n_1 - 1$ і $\nu_2 = n_2 - 1$.

8. Якщо обчислене значення F -критерію більше або дорівнює критичному, тобто $F \geq F_\alpha$, то це означає, що для заданого рівня значущості α дисперсії істотно відрізняються і нульову гіпотезу відхиляють; якщо $F < F_\alpha$, то вибіркові спостереження не суперечать нульовій гіпотезі.

Зазначимо, що F -критерій дуже чутливий до відхилень розподілу від нормального. Якщо припущення про нормальний розподіл не буде прийняте, то F -критерій застосовувати не варто. До нормально розподілених сукупностей F -критерій застосовують у випадку малого ($n < 30$) і середнього ($n < 100$) обсягів вибірок.

Порівняння двох вибірових середніх значень для незалежних вибірок.

Нехай маємо дві генеральні сукупності X і Y , кожна з яких має нормальний закон розподілу з математичним сподіванням відповідно μ_x і μ_y . З цих сукупностей отримано дві незалежні вибірки обсягом відповідно n_x і n_y . Середні значення вибірок позначимо \bar{x} і \bar{y} . Треба оцінити істотність відмінності між двома вибірковими середніми \bar{x} і \bar{y} . Для розв'язання цієї задачі перевіряють гіпотезу про рівність математичних сподівань генеральних сукупностей X і Y . Для цього використовують t -критерій Стюдента. Його застосування різне за різних припущень щодо дисперсій генеральних сукупностей X і Y . У математичній статистиці розглядають випадки відомих і невідомих генеральних дисперсій. Але на практиці генеральні дисперсії, як правило, не відомі, тому розглянемо тільки випадок невідомих дисперсій. Якщо дисперсії не відомі, то можна припустити, що вони рівні між собою, а можна такого припущення й не робити. Щоб вибрати потрібний варіант, на практиці спочатку за вибірковими даними перевіряють гіпотезу про рівність генеральних дисперсій, використовуючи F -критерій, а потім вибирають відповідний варіант t -критерію.

Порядок застосування t -критерію такий.

1. Припускають, що обидві вибірки незалежні і їх взято з генеральних сукупностей X і Y , які мають нормальний закон розподілу з параметрами μ_x, σ_x і μ_y, σ_y .

2. Формулюють нульову гіпотезу $H_0: \mu_x = \mu_y$ і альтернативну їй $H_a: \mu_x \neq \mu_y$.

3. Задають рівень значущості α .

4. Розглядають дві незалежні вибірки із генеральних сукупностей X і Y обсягом n_x і n_y .

5. Обчислюють вибіркові характеристики $\bar{x}, S_x, \bar{y}, S_y$.

6. Використовують F -критерій для перевірки гіпотези про рівність генеральних дисперсій.

7. За результатом застосування F -критерію приймають або не приймають припущення про рівність генеральних дисперсій.

8. Обчислюють значення t -критерію і число степенів вільності ν . У табл. 8.12 подано формули для обчислення t і ν , залежно від зробленого припущення щодо дисперсій і співвідношення між обсягами вибірок n_x і n_y .

9. З таблиці розподілу Стьюдента беруть критичне значення t_α t -критерію при заданому рівні значущості α і числі степенів вільності ν .

10. Якщо $t \geq t_\alpha$, то на заданому рівні значущості α відмінність між вибірковими середніми не випадкова, а істотна. Якщо $t < t_\alpha$, то немає підстав вважати, що відмінність між середніми істотна.

Оцінка ступеня розбіжності між емпіричним і теоретичним розподілами. Розглянуті вище критерії значущості забезпечують найвищу достовірність статистичних висновків для випадків, коли вибірки взято з нормально розподілених генеральних сукупностей. При відхиленнях від нормального розподілу точність критеріїв значущості значно зменшується. Тому перед тим, як ними користуватися, треба перевірити припущення про нормальний розподіл генеральної сукупності. Для цього використовують критерії згоди. Розглянемо один з таких критеріїв — критерій χ^2 (хі-квадрат) Пірсона.

Нехай теоретичний розподіл задано формулою $f(x)$, а m_1', m_2', \dots, m_k' теоретичні частоти відповідних значень x_1, x_2, \dots, x_k ознаки x . Треба встановити, якою мірою розподіл $f(x)$ відображає емпіричний ряд. Щоб зробити висновок про міру узгодження між емпіричним і теоретичним розподілом, розглядають сумарну розбіжність між емпіричними m_1, m_2, \dots, m_k і теоретичними m_1', m_2', \dots, m_k' частотами. Сумарна розбіжність між частотами залежить від розподілу $f(x)$, який дає нам теоретичні частоти, і від випадкових причин, внаслідок яких маємо емпіричний розподіл. Якщо сумарна розбіжність мала, то можна припустити, що вона зумовлена випадковими причинами, а тому теоретичний розподіл $f(x)$ добре відображає емпіричний ряд. Якщо сумарна розбіжність велика, то вона зумовлена істотними причинами, а саме тим, що теоретичний розподіл $f(x)$ погано відображає емпіричний ряд.

Критерій згоди Пірсона полягає в тому, що за міру розбіжності між емпіричними і теоретичними частотами беруть величину

$$\chi^2 = \sum_{i=1}^k \frac{(m_i - m_i')^2}{m_i'} \quad (8.26)$$

Величина χ^2 — середня зважена квадратів відхилень емпіричних і теоретичних частот, при цьому вагою є величини, обернені теоретичним частотам. Статистика χ^2 є випадковою величиною, яка має свій закон розподілу. Пірсон показав, що χ^2 не залежить ні від функції $f(x)$, ні від обсягу вибірки, а залежить лише від параметра ν — числа степенів вільності й дорівнює різниці між числом частот k , які порівнюються, і

Таблиця 8.12

№ пор.	Примітка щодо дисперсій σ_x, σ_y	Обсяг вибірок n_x, n_y	Формула t -критерію	Стандартна похибка різниці $S_{\bar{x}-\bar{y}}$	Число степенів вільності
1	$\sigma_x^2 = \sigma_y^2$	$n_x = n_y = n$	$t = \frac{\bar{x} - \bar{y}}{S_{\bar{x}-\bar{y}}}$	$S_{\bar{x}-\bar{y}} = \sqrt{\frac{S_x^2 + S_y^2}{n}}$	$\nu = 2n - 2$
2	$\sigma_x^2 = \sigma_y^2$	$n_x \neq n_y$		$S_{\bar{x}-\bar{y}} = \sqrt{\frac{n_x + n_y}{n_x n_y} \times \sqrt{\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2}}}$	$\nu = n_x + n_y - 2$
3	$\sigma_x^2 \neq \sigma_y^2$	$n_x = n_y = n$		$S_{\bar{x}-\bar{y}} = \sqrt{\frac{S_x^2 + S_y^2}{n}}$	$\nu = (n - 1) \frac{(S_x^2 + S_y^2)^2}{S_x^2 + S_y^2}$
4	$\sigma_x^2 \neq \sigma_y^2$	$n_x \neq n_y$		$S_{\bar{x}-\bar{y}} = \sqrt{\frac{S_x^2 + S_y^2}{n_x} + \frac{S_y^2}{n_y}}$	$\nu = \frac{\left(\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}\right)^2}{\frac{S_x^2}{n_x^2(n_x - 1)} + \frac{S_y^2}{n_y^2(n_y - 1)}}$

числом зв'язків, які на ці частоти накладено. При застосуванні критерію Пірсона вважають, що сума теоретичних частот дорівнює сумі емпіричних, а теоретичні середня і дисперсія дорівнюють вибірковій середній і дисперсії. Тому число степенів вільності $\nu = k - 3$.

Критерій згоди Пірсона застосовують у такий спосіб.

1. Формулюють гіпотезу H_0 — емпіричний розподіл відповідає нормальному і альтернативну гіпотезу — емпіричний розподіл не відповідає нормальному.

2. Задють рівень значущості α .

3. Розглядають вибірку обсягом $n \geq 40$ незалежних спостережень і емпіричний розподіл подають у вигляді інтервального варіаційного ряду.

4. Обчислюють вибіркові характеристики \bar{x} і S . Їх використовують замість генеральних параметрів μ і σ нормального розподілу, з яким порівнюватимуть емпіричний розподіл.

5. Обчислюють значення теоретичних частот m_i' для кожного з інтервалів групування. Для цього можна скористатися формулою

$$m_i' = n'f(t) = n' \cdot \frac{1}{\sqrt{2\pi}} e^{-t^2/2},$$

де $t = (x_i - \bar{x})/S$, $n' = nh/S$, n — обсяг вибірки, x_i — середня i -го інтервалу групування, h — довжина інтервалу, \bar{x} , S — вибіркові середнє та середнє квадратичне відхилення.

Якщо буде встановлено, що обчислені частоти m_i' деяких інтервалів групування менші п'яти, то сусідні інтервали об'єднують так, щоб сума обчислених теоретичних частот була не менша п'яти. Додають також частоти відповідних об'єднаних інтервалів.

6. Обчислюють значення χ^2 -критерію за формулою (8.26).

7. Знаходять табличне критичне значення χ_{α}^2 для заданого рівня значущості α і числа степенів вільності ν .

8. Якщо $\chi^2 \geq \chi_{\alpha}^2$, то емпіричний розподіл не відповідає нормальному розподілу на заданому рівні значущості α . Якщо $\chi^2 < \chi_{\alpha}^2$, то це дає право стверджувати, що гіпотеза H_0 допустима, тобто припущення про те, що в генеральній сукупності теоретичний розподіл не суперечить дослідним даним.

Наведемо Бейсік-програму 8.2 порівняння емпіричного розподілу з нормальним, за допомогою якої перевіримо гіпотезу про те, що емпіричний розподіл, поданий у табл. 8.10, відповідає нормальному закону розподілу.

Програма 8.2

```

10 PRINT "ПОРІВНЯННЯ ЕМПІРИЧНОГО"
20 PRINT "РОЗПОДІЛУ З НОРМАЛЬНИМ"
30 REM-----
40 READ N

```

```

50 DIM X(N), Y(N), M(N), ME(N)
60 XS=0: S=0
70 REM Введення заданого масиву
80 FOR I=1 TO N: READ X(I): NEXT I
90 REM Впорядкування X(I) за зростанням (100-170)
100 FOR I=1 TO N-1
110 R=X(I): K=I
120 FOR J=I+1 TO N
130 IF R<X(J) THEN 150
140 R=X(J): K=J
150 NEXT J
160 X(K)=X(I): X(I)=R
170 NEXT I
180 REM Групування в інтервали та
190 REM підрахунок емпіричних частот
200 REM (рядки 210-340)
210 A=X(I): B=X(N)
220 K=1+INT(3.32*LOG(N)/LOG(10))
240 H=(B-A)/K: H=INT(H+.5)
250 Y(0)=X(N)-K*H
260 FOR I=1 TO K-1
270 Y(I)=Y(I-1)+H
280 NEXT I
290 Y(K)=B: I=1
300 FOR J=0 TO K-1
310 IF X(I)>Y(J+1) THEN 340
320 ME(J)=ME(J)+1: I=I+1
330 IF I<=N THEN 310
340 NEXT J
350 REM Обчислення вибіркових середнього
360 REM значення та середнього квадратичного
370 REM відхилення (380-460)
380 H1=H/2: PI=3.141593: XS=0: X2=0
390 FOR I=1 TO K
400 C=Y(I-1)+H1: XS=XS+C*ME(I-1)
410 X2=X2+ME(I-1)*C*C
420 NEXT I
430 S=(X2-XS*XS/N)/(N-1)
440 XS=XS/N
450 S=SQR(S)
460 PRINT:PRINT "XS=" XS; "x2=" X2; "S=" S: PRINT
470 REM Обчислення теоретичних частот
480 REM та значення критерію XI
490 N1=H*N/S: XI=0: L=N1/SQR(2*PI)
495 PRINT " i C ME(i)*C ME(i)*C^2 C-XS (C-XS/S M(i)"
500 FOR I=0 TO K-1

```

```

510 C=Y(I)+H1 : T=(C-XS)/S
520 M(I)=L*EXP(-T*T/2)
530 XI=XI+(ME(I)-M(I))^2/M(I)
531 PRINT I+1; C; ME(I); TAB(15); ME(I)*C; TAB(24);
ME(I)*C*C; TAB(35);
532 PRINT USING"###.## ";C-XS;
533 PRINT USING"###.## ";(C-XS)/S;
534 PRINT USING"###.## ";M(I)
540 NEXT I
550 PRINT : PRINT "Значення критерію =";XI
560 END
565 DATA 50
570 DATA 183, 170, 176, 178, 176, 180, 176, 185, 184, 174
580 DATA 168, 174, 189, 172, 175, 176, 179, 176, 172, 178
590 DATA 169, 171, 170, 177, 176, 179, 174, 176, 188, 178
600 DATA 172, 176, 167, 166, 180, 183, 176, 182, 178, 172
610 DATA 185, 183, 175, 174, 180, 166, 170, 171, 178, 169

```

Виконавши програму, EOM надрукувала таку інформацію:
 $XS = 175.72 \times 2 = 1545274$ $S = 5.341558283681$

I	C	ME(I)	ME(I)*C	ME(I)*C^2	C-XS	(C-XS)/S	M(I)
1	167	6	1002	167334	-8.72	-1.63	3.94
2	171	9	1539	263169	-4.72	-0.88	10.11
3	175	16	2800	490000	-0.72	-0.13	14.80
4	179	10	1790	320410	3.28	0.61	12.37
5	183	7	1281	234423	7.28	1.36	5.90
6	187	2	374	69938	11.28	2.11	1.61

Значення критерію = 2.0500875442969

Обчислене значення критерію $\chi^2 = 2.05$. Число степенів вільності $\nu = n_{in} - 1 - 2 = 4 - 3 = 1$, оскільки оцінюється два параметри \bar{x} і S (n_{in} — кількість інтервалів після об'єднання першого з другим і шостого з п'ятим). Критичне значення χ_{α}^2 для $\alpha = 0,1$ і $\nu = 1$ має значення $\chi_{\alpha}^2 = 2,706 \geq 2,05 = \chi^2$.

Таким чином, гіпотеза про те, що емпіричні частоти розподілені за нормальним законом, приймається на 10%-му рівні значущості.

§ 8.5. ОСНОВИ РЕГРЕСІЙНОГО І КОРЕЛЯЦІЙНОГО АНАЛІЗІВ

Якщо значення однієї величини однозначно визначає значення іншої, то в цьому разі наявна функціональна залежність між величинами. Наприклад, площа круга і довжина кола однозначно залежать від радіуса. Але частіше величини зв'язані залежностями, коли одному значенню незалежної величини може відповідати не одне, а кілька значень іншої величини. Якщо в такій ситуації розглядати одну величину як незалежну, а іншу — як залежну від першої, то залежна величина веде себе як випадкова і її можна описати деяким імовірнісним законом розподілу. Але цей розподіл (або його параметри) закономірно змінюється від зміни значень незалежної величини. У цьому разі кажуть про *статистичну* (або *ймовірнісну*) залежність між величинами. При вивченні статистичних залежностей розрізняють регресію і кореляцію.

Регресія — це залежність математичного сподівання (середнього значення) випадкової величини (ознаки) Y від величини (ознаки) X . При цьому кажуть: "регресія Y на X ". Незалежна величина X не обов'язково має бути випадковою.

Кореляція — це залежність між двома випадковими величинами Y і X . Вона характеризується *коефіцієнтами кореляції*.

Регресійний аналіз встановлює форму залежності між випадковою величиною Y і значенням однієї або кількох змінних величин, значення яких вважають заданими точно. Така залежність, як правило, визначається деякою математичною моделлю — рівнянням регресії, яке містить кілька невідомих параметрів. На основі вибірових даних знаходять оцінки цих параметрів, потім визначають статистичні похибки оцінок і перевіряють відповідність прийнятої моделі експериментальним даним.

Кореляційний аналіз полягає у визначенні тісноти зв'язку між двома випадковими величинами. За міру зв'язку береться коефіцієнт кореляції, який обчислюють на основі вибірки пар значень (x_i, y_i) обсягу n з генеральної сукупності X і Y . Потім перевіряють гіпотези або встановлюють межі достовірного інтервалу для генерального коефіцієнта кореляції.

Регресійний аналіз. Одним з важливих етапів регресійного аналізу є встановлення рівняння регресії, яке описує зв'язок між значеннями y_i залежної змінної Y і значеннями x_i незалежної змінної X . Рівняння регресії відображає залежність, яка існує в генеральній сукупності, з якої взято експериментальні дані.

Регресію, яка описується рівнянням

$$y = \beta_0 + \beta_1 x, \quad (8.27)$$

називають *простою лінійною регресією*, оскільки вона враховує залежність тільки від однієї незалежної змінної x .

Часто поведінку результативної ознаки Y не вдається пояснити тіль-

ки однією незалежною змінною x . Тоді використовують модель *множинної лінійної регресії*:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k,$$

в якій значення випадкової величини Y залежить від k незалежних змінних x_1, x_2, \dots, x_k .

Термін "лінійна регресія" означає лінійність по відношенню до коефіцієнтів $\beta_1, \beta_2, \dots, \beta_k$, а не до змінних x_1, x_2, \dots, x_k які можуть бути навіть функціями від інших величин, в тому числі від x_1, x_2, \dots, x_k .

Окремим випадком множинної лінійної регресії є *поліноміальна регресія*

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_k x^k.$$

Розглянуті рівняння регресії описують залежність середнього значення випадкової величини Y від незалежної величини x в генеральній сукупності Y . Оскільки вся генеральна сукупність Y недоступна для спостережень, тому істинне рівняння регресії не відоме, і будь-яка регресійна модель є лише наближенням до справжньої. Для вибору моделі вдаються до предметного аналізу явища, графічного зображення експериментальних даних та досвіду дослідника.

Розглянемо найпростіший випадок регресійного аналізу — просту лінійну модель. Лінійний регресійний аналіз ґрунтується на таких припущеннях:

- 1) в генеральній сукупності Y справді існує лінійна регресія (8.27);
- 2) усі значення y_i випадкової величини Y статистично незалежні для всіх значень x_i незалежної змінної x ;
- 3) генеральна сукупність Y , з якої взято вибіркові дані, підпорядкована нормальному закону розподілу з сталою дисперсією σ^2 .

Оскільки справжнє рівняння регресії невідоме, треба знайти оцінки (наближені значення) генеральних параметрів β_0 і β_1 . Позначимо оцінки генеральних параметрів відповідно через b і a . Щоб визначити оцінки a і b , проводять вибірку обсягу n значень випадкової величини Y , які відповідають значенням x_i незалежної змінної x . Підставивши у формулу (8.27) замість параметрів β_0 і β_1 їх оцінки b і a , дістають рівняння регресії

$$\bar{y} = ax + b, \quad (8.28)$$

коефіцієнти якого знаходять з умови мінімуму суми квадратів відхилень дослідних значень y_i від обчислених за формулою (8.28):

$$\sum_{i=1}^n (y_i - ax_i - b)^2 \rightarrow \min.$$

За методом найменших квадратів

$$a = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \quad b = \bar{y} - a\bar{x},$$

де \bar{x} і \bar{y} — вибіркові середні.

У рівнянні (8.28) коефіцієнт a називають *коефіцієнтом регресії*. Він показує середню зміну залежної величини при зміні на одиницю незалежної величини. Коефіцієнт регресії завжди число іменоване. Якщо $a > 0$, то зв'язок між Y і x прямий, якщо $a < 0$, то зв'язок обернений, якщо $a = 0$, то зв'язку немає, тобто залежність між Y і x функціональна. Коефіцієнт b — *початок відліку*, тобто є значенням \bar{y} , коли $x = 0$.

Оцінкою точності опису реальної залежності між Y і x за допомогою рівняння лінійної регресії (8.28) є середнє квадратичне відхилення

$$S_{yx} = \frac{1}{\sqrt{n-2}} \sqrt{\sum_{i=1}^n y_i^2 - b \sum_{i=1}^n y_i - a \sum_{i=1}^n x_i y_i}.$$

Величина S_{yx} є мірою точності передбачення значень випадкової величини Y , тому S_{yx} називають також *стандартною похибкою передбачення*. Чим менша величина S_{yx} , тим вірогідніший лінійний зв'язок між Y і x .

Щоб перевірити відповідність математичної моделі експериментальним даним, використовують критерій Фішера. Нульовою гіпотезою в цьому разі є твердження про те, що регресія в генеральній сукупності лінійна.

Для перевірки нульової гіпотези обчислюють величину

$$F = \frac{(n-2) \left(\sum_{i=1}^n y_i^2 - \frac{1}{n} \left(\sum_{i=1}^n y_i \right)^2 \right)}{(n-1) \sum_{i=1}^n (y_i - ax_i - b)^2},$$

яка має F -розподіл Фішера з степенями вільності $\nu_1 = 1$ і $\nu_2 = n - 2$. Для заданого рівня значущості α і числа степенів вільності $\nu_1 = 1$ і $\nu_2 = n - 2$ знаходять з таблиць F -розподілу критичне значення F_α . Якщо $F \geq F_\alpha$, то нульову гіпотезу відхиляють; якщо $F < F_\alpha$, то немає достатніх підстав для її відхилення.

Якщо лінійна модель адекватна емпіричним даним, то встановлюють достовірність істотної відмінності від нуля коефіцієнта регресії a . Нульову гіпотезу про те, що $a = 0$, перевіряють за допомогою t -критерію Стьюдента. Статистика

$$t = \frac{|a|}{S_{ay}} \sqrt{\sum_{i=1}^n x_i^2 - n\bar{x}^2} \quad (8.29)$$

має розподіл Стьюдента з $\nu = n - 2$ степенями вільності. Обчислене за формулою (8.29) значення t -критерію порівнюють з табличним t_α при заданому рівні значущості α і числі степенів вільності $\nu = n - 2$.

Якщо значення критерію $t \geq t_\alpha$, то нульова гіпотеза про рівність нулю

коефіцієнта регресії відхиляється і вважають, що коефіцієнт регресії істотно відмінний від нуля. Якщо $t < t_{\alpha}$, то для відхилення нульової гіпотези немає достатніх підстав.

Кореляційний аналіз. Якщо обидві змінні X і Y випадкові і мають спільний нормальний розподіл, то зв'язок між ними вивчають за допомогою кореляційного аналізу. Особливістю кореляційного аналізу є те, що за його допомогою вивчають лінійну залежність між X і Y . Тіснота лінійного зв'язку в кореляційному аналізі характеризується спеціальним відносним показником, який називають *коефіцієнтом кореляції*. Значення коефіцієнта кореляції ρ належить відрізку $[-1; 1]$. Якщо $\rho = \pm 1$, то між випадковими величинами існує лінійний функціональний зв'язок. Якщо $\rho = 0$, то між величинами X і Y кореляції немає і їх називають *некорельованими*. Для нормально розподіленої сукупності X і Y некорельованість означає, що величини X і Y незалежні. Додатний знак ρ вказує на прямий зв'язок між X і Y , а від'ємний — на обернений зв'язок. Чим ближче коефіцієнт кореляції до одиниці, тим зв'язок між X і Y тісніший.

Для генеральної сукупності коефіцієнт кореляції, як правило, невідомий. Тому його обчислюють на основі експериментальних даних, які є вибіркою обсягу n пар значень (x_i, y_i) при спільному вимірюванні двох ознак X і Y . Коефіцієнт кореляції, обчислений за вибірковими даними, називають *вибірковим коефіцієнтом кореляції* (часто просто коефіцієнтом кореляції); його позначають r .

Вибірковий коефіцієнт кореляції обчислюють за формулою

$$r = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{(n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2)(n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2)}} \quad (8.30)$$

або за формулою

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{S_x \cdot S_y} \quad (8.31)$$

де \overline{xy} — середнє значення добутку x на y ; \bar{x} , \bar{y} — середні значення відповідних ознак; S_x , S_y — вибіркові середні квадратичні відхилення, знайдені за ознаками X і Y .

Знаючи коефіцієнт кореляції r , можна легко визначити коефіцієнт регресії

$$a = \frac{S_y}{S_x} \cdot r \quad (8.32)$$

Маючи вибірковий коефіцієнт кореляції, ще не можна зробити висновок про достовірність кореляції між ознаками X і Y . Щоб з'ясувати, чи справді випадкові величини X і Y перебувають у кореляційному

зв'язку, треба перевірити достовірність істотної відмінності коефіцієнта r від нуля. Для цього перевіряють нульову гіпотезу $H_0: \rho = 0$. При цьому припускають, що випадкові величини X і Y підпорядковані двовимірному нормальному розподілу.

Для перевірки нульової гіпотези обчислюють статистику

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad (8.33)$$

де n — обсяг вибірки. Статистика t має розподіл Стьюдента з $\nu = n - 2$ степенями вільності. Обчислене за формулою (8.33) значення t -критерію порівнюють з критичним значенням t_{α} , знайденим за таблицею розподілу Стьюдента при заданому рівні значущості α і числі степенів вільності $\nu = n - 2$.

Якщо $|t| \geq t_{\alpha}$, то нульову гіпотезу про те, що між змінними X і Y немає кореляційного зв'язку, відхиляють. Змінні X і Y вважають залежними. Якщо $|t| < t_{\alpha}$, то випадкові величини X і Y не корельовані.

Програма 8.3 містить регресійний та кореляційний аналізи залежності між величинами X і Y .

Програма 8.3

```

10 PRINT "РЕГРЕСІЙНИЙ ТА"
20 PRINT "КОРЕЛЯЦІЙНИЙ ЗВ'ЯЗОК МІЖ X І Y"
30 REM -----
40 REM -----Введення даних (рядки 50—90)-----
50 READ N
60 DIM X(N), Y(N)
70 FOR I=1 TO N
80 READ X(I), Y(I)
90 NEXT I
100 REM -----Обчислення регресійних та кореляційних --
110 REM характеристик (рядки 120-310)
120 X1=0: Y1=0: XY=0
130 X2=0: Y2=0
140 FOR I=1 TO N
150 X1=X1+X(I): X2=X2+X(I)^2
160 Y1=Y1+Y(I): Y2=Y2+Y(I)^2
170 XY=XY+X(I)*Y(I)
180 NEXT I
190 XC=X1/N: YC=Y1/N
200 SX=SQR((X2-X1*X1/N)/(N-1))
210 SY=SQR((Y2-Y1*Y1/N)/(N-1))
220 R=(XY/N-XC*YC)/(SX*SY)
230 A=R*SY/SX: B=(Y1-A*X1)/N

```

```

240 S=SQR(Y2-B*Y1-A*XY)/SQR(N-2)
250 TR=ABS(A)*SQR(X2-N*XC*XC)/S
260 TK=R*SQR(N-2)/SQR(1-R*R)
270 C=0
280 FOR I=1 TO N
290 C=C+(Y(I)-A*X(I)-B)^2
300 NEXT I
310 F=(N-2)*SY*SY/C
320 PRINT "a="A; "b="B; "r="R
330 PRINT "F="F; "tr="TR; "tk="TK
340 END
350 DATA 50
360 DATA 183, 72, 170, 70, 176, 83, 178, 68, 176, 69, 180, 83
370 DATA 176, 74, 185, 79, 184, 71, 174, 68, 168, 70, 174, 70
380 DATA 189, 85, 172, 83, 175, 85, 176, 73, 179, 74, 176, 69
390 DATA 172, 70, 178, 74, 169, 65, 171, 69, 170, 65, 177, 79
400 DATA 176, 84, 179, 80, 174, 71, 176, 76, 188, 73, 178, 72
410 DATA 172, 68, 176, 67, 167, 65, 166, 74, 180, 70, 183, 82
420 DATA 176, 71, 182, 72, 178, 68, 172, 78, 185, 74, 183, 72
430 DATA 175, 72, 174, 71, 180, 75, 166, 62, 170, 68, 171, 69
440 DATA 174, 73, 169, 67

```

За програмою 8.3 виконано лінійний регресійний та кореляційний аналізи залежності між зростом і масою 50 юнаків-студентів. Значення зросту та маси студентів подано в програмі в рядках 360–440. В результаті обчислень за програмою 8.3 маємо $a = 0,464$; $b = -8,840$. Тобто, рівняння регресії $y = 0,464x - 8,840$.

Для перевірки адекватності виведеного рівняння лінійної моделі в програмі обчислено значення критерію Фішера $F = 1,24$, яке менше від табличного тільки при 25 %-му рівні значущості. Для перевірки нульової гіпотези про рівність нулю коефіцієнта регресії a в програмі обчислено статистику t_r за формулою (8.29). Її значення $t_r = 3,49$ більше від табличного $t_{\alpha} = 1,67$, коли $\alpha = 0,05$, тобто при 5 %-му рівні значущості. Це означає, що коефіцієнт регресії істотно відмінний від нуля.

Коефіцієнт кореляції $r = 0,449$, а значення статистики, обчисленої за формулою (8.33), дорівнює $tk = 3,484$; воно більше від табличного значення 1,67 на 5 %-му рівні значущості. Це означає, що на цьому рівні значущості між зростом і масою студентів існує кореляційний зв'язок.

ЛАБОРАТОРНА РОБОТА № 7

Тема. Статистична обробка експериментальних даних.

Завдання. 1. Знайти і побудувати емпіричну формулу для залежності, поданої таблицею.

x	x_1	x_2	...	x_n
y	y_1	y_2	...	y_n

Оцінити ступінь розбіжності між даними значеннями y_i і значеннями, обчисленими за побудованою формулою.

2. Звести дану залежність до лінійної і провести лінійний регресійний та кореляційний аналізи між вирівняними даними (X_i, Y_i) .

Для цього:

а) побудувати рівняння регресії $Y = aX + b$;

б) за допомогою критерію Фішера перевірити адекватність виведеного рівняння лінійної моделі;

в) визначити тісноту лінійного зв'язку між вирівняними даними (X_i, Y_i) та оцінити достовірність кореляції між ними.

Значення (x_i, y_i) подано в табл. 8.13.

Таблиця 8.13

Варіант	1	2	3	4	5	6
1	x_i	y_i	y_i	y_i	y_i	y_i
2	.7	2.4	3	.3	-8	7
3	.9	2.7	3.5	.4	.6	4.1
4	1.1	2.9	4.1	.5	1.6	2.6
5	1.3	3.1	4.8	.6	2.4	1.5
6	1.5	3.3	5.6	.7	3	.8
7	1.7	3.4	6.5	.9	3.6	.3
8	1.9	3.6	7.6	1.1	4.1	-1
9	2.1	3.8	8.9	1.3	4.6	-4
10	2.3	3.9	10.5	1.6	5	-6
11	2.5	4	12.3	2	5.3	-8
12	2.7	4.2	14.4	2.4	5.7	-1
13	2.9	4.3	16.8	3	6	-1.1
14	3.1	4.4	19.7	3.6	6.3	-1.3
15	3.3	4.5	23	4.4	6.5	-1.4
16	3.5	4.6	27	5.4	6.8	-1.5
17	3.7	4.7	31.6	6.6	7	-1.6
18	3.9	4.8	37	8.1	7.2	-1.6
19	4.1	4.9	43.3	9.9	7.4	-1.7
20	4.3	5	50.7	12.1	7.6	-1.8
21	4.5	5.1	59.4	14.7	7.8	-1.8
22	4.7	5.2	69.5	18	8	-1.9
23	4.9	5.3	81.4	22	8.2	-1.9
24	5.1	5.4	95.3	26.9	8.4	-2
25	5.3	5.5	111.5	32.8	8.5	-2
26	5.5	5.6	130.6	40.1	8.7	-2.1
27	5.7	5.7	152.9	48.9	8.8	-2.1
28	5.9	5.8	179	59.8	9	-2.1
29	6.1	5.9	209.6	73	9.1	-2.2
30	6.3	5.9	245.4	89.2	9.2	-2.2
31	6.5	6	287.3	108.9	9.4	-2.2
32	6.7	6.1	336.3	133	9.5	-2.2
33	6.9	6.2	393.8	162.5	9.6	-2.3
34	7.1	6.3	461	198.5	9.7	-2.3
35	7.3	6.3	539.8	242.4	9.8	-2.3
35	7.5	6.4	632	296.1	10	-2.3

Таблица 8.13

Версия		7	8	9	10	11	12
<i>i</i>	x_i	y_i	y_i	y_i	y_i	y_i	y_i
1	1	1	6	.5	.7	2	7
2	1.2	1.05	6.8	.5	.7	2.5	6.5
3	1.4	1.09	7.5	.6	.8	3	6.14
4	1.6	1.12	8.1	.6	.9	3.4	5.87
5	1.8	1.15	8.7	.7	1	3.8	5.66
6	2	1.17	9.3	.8	1.1	4.1	5.5
7	2.2	1.19	9.9	.8	1.2	4.4	5.36
8	2.4	1.21	10.4	.9	1.3	4.6	5.25
9	2.6	1.22	10.9	1	1.5	4.9	5.15
10	2.8	1.23	11.4	1.1	1.6	5.1	5.07
11	3	1.25	11.9	1.2	1.8	5.3	5
12	3.2	1.26	12.3	1.3	2	5.5	4.93
13	3.4	1.26	12.8	1.5	2.2	5.7	4.88
14	3.6	1.27	13.2	1.6	2.4	5.8	4.83
15	3.8	1.28	13.6	1.8	2.7	6	4.78
16	4	1.29	14	2	3	6.2	4.75
17	4.2	1.29	14.4	2.2	3.3	6.3	4.71
18	4.4	1.3	14.8	2.4	3.6	6.4	4.68
19	4.6	1.3	15.2	2.6	4	6.6	4.65
20	4.8	1.31	15.5	2.9	4.4	6.7	4.62
21	5	1.31	15.9	3.1	4.9	6.8	4.6
22	5.2	1.32	16.2	3.5	5.4	6.9	4.57
23	5.4	1.32	16.6	3.8	6	7.1	4.55
24	5.6	1.32	16.9	4.2	6.6	7.2	4.53
25	5.8	1.33	17.3	4.6	7.3	7.3	4.51
26	6	1.33	17.6	5	8	7.4	4.5
27	6.2	1.33	17.9	5.5	8.9	7.5	4.48
28	6.4	1.33	18.2	6.1	9.8	7.6	4.46
29	6.6	1.34	18.6	6.7	10.8	7.7	4.45
30	6.8	1.34	18.9	7.3	12	7.8	4.44
31	7	1.34	19.2	8.1	13.2	7.8	4.42
32	7.2	1.34	19.5	8.8	14.6	7.9	4.41
33	7.4	1.35	19.8	9.7	16.2	8	4.4
34	7.6	1.35	20.1	10.7	17.9	8.1	4.39
35	7.8	1.35	20.3	11.7	19.8	8.2	4.38

Таблица 8.13

Версия		13	14	15	16	17	18
<i>i</i>	x_i	y_i	y_i	y_i	y_i	y_i	y_i
1	1	3	4.4	.3	-.8	8	1
2	1.3	3.3	5.6	.4	.5	6.8	.52
3	1.6	3.5	7.1	.4	1.1	6.1	.35
4	1.9	3.8	8.9	.5	1.8	5.6	.27
5	2.2	4	11.3	.6	2.4	5.3	.21
6	2.5	4.2	14.4	.7	2.9	5	.18
7	2.8	4.3	18.2	.8	3.3	4.8	.15
8	3.1	4.5	23	.9	3.8	4.6	.13
9	3.4	4.7	29.2	1.1	4.1	4.5	.12
10	3.7	4.8	37	1.3	4.5	4.4	.11

Версия		13	14	15	16	17	18
11	4	5	46.9	1.5	4.8	4.3	.1
12	4.3	5.1	59.4	1.7	5.1	4.2	.09
13	4.6	5.3	75.2	2	5.3	4.1	.08
14	4.9	5.4	95.3	2.3	5.6	4	.07
15	5.2	5.6	120.7	2.7	5.8	4	.07
16	5.5	5.7	152.9	3.1	6	3.9	.06
17	5.8	5.8	193.7	3.6	6.3	3.9	.06
18	6.1	5.9	245.4	4.2	6.5	3.8	.06
19	6.4	6.1	310.8	4.9	6.7	3.8	.05
20	6.7	6.2	393.8	5.7	6.8	3.7	.05
21	7	6.3	498.9	6.6	7	3.7	.05
22	7.3	6.4	632	7.7	7.2	3.7	.05
23	7.6	6.5	800.6	8.9	7.3	3.7	.04
24	7.9	6.6	1014.3	10.4	7.5	3.6	.04
25	8.2	6.7	1285	12.1	7.6	3.6	.04
26	8.5	6.8	1627.9	14	7.8	3.6	.04
27	8.8	6.9	2062.3	16.3	7.9	3.6	.04
28	9.1	7	2612.6	18.9	8.1	3.5	.04
29	9.4	7.1	3309.8	22	8.2	3.5	.03
30	9.7	7.2	4193	25.5	8.3	3.5	.03
31	10	7.3	5312	29.7	8.4	3.5	.03
32	10.3	7.4	6729.5	34.5	8.6	3.5	.03
33	10.6	7.5	8525.3	40.1	8.7	3.5	.03
34	10.9	7.6	10800.3	46.6	8.8	3.5	.03
35	11.2	7.7	13682.5	54.1	8.9	3.4	.03

Таблица 8.13

Версия		19	20	21	22	23	24
1	2	1.17	9.3	.6	.9	3.7	5.5
2	2.5	1.22	10.6	.7	1.1	4.8	5.2
3	3	1.25	11.9	.8	1.2	5.7	5
4	3.5	1.27	13	1	1.4	6.5	4.85
5	4	1.29	14	1.2	1.7	7.1	4.75
6	4.5	1.3	15	1.4	1.9	7.7	4.66
7	5	1.31	15.9	1.6	2.2	8.3	4.6
8	5.5	1.32	16.8	1.9	2.6	8.7	4.54
9	6	1.33	17.6	2.3	3	9.2	4.5
10	6.5	1.34	18.4	2.7	3.5	9.6	4.46
11	7	1.34	19.2	3.2	4.1	9.9	4.42
12	7.5	1.35	19.9	3.7	4.7	10.3	4.4
13	8	1.35	20.6	4.4	5.5	10.6	4.37
14	8.5	1.36	21.3	5.2	6.4	10.9	4.35
15	9	1.36	22	6.2	7.4	11.2	4.33
16	9.5	1.36	22.7	7.3	8.6	11.5	4.31
17	10	1.37	23.3	8.7	10	11.7	4.3
18	10.5	1.37	23.9	10.3	11.7	12	4.28
19	11	1.37	24.5	12.1	13.6	12.2	4.27
20	11.5	1.37	25.1	14.4	15.8	12.4	4.26
21	12	1.37	25.7	17	18.3	12.6	4.25
22	12.5	1.38	26.3	20.1	21.3	12.8	4.24
23	13	1.38	26.8	23.8	24.7	13	4.23
24	13.5	1.38	27.4	28.2	28.7	13.2	4.22

Вариант		19	20	21	22	23	24
25	14	1.38	27.9	33.3	33.3	13.4	4.21
26	14.5	1.38	28.5	39.4	38.7	13.6	4.2
27	15	1.38	29	46.7	45	13.8	4.2
28	15.5	1.39	29.5	55.2	52.3	13.9	4.19
29	16	1.39	30	65.3	60.8	14.1	4.18
30	16.5	1.39	30.5	77.3	70.6	14.2	4.18
31	17	1.39	31	91.5	82	14.4	4.17
32	17.5	1.39	31.5	108.2	95.3	14.5	4.17
33	18	1.39	31.9	128.1	110.7	14.7	4.16
34	18.5	1.39	32.4	151.5	128.6	14.8	4.16
35	19	1.39	32.9	179.3	149.4	14.9	4.15

Таблица 8.13

Вариант		25	26	27	28	29	30
<i>i</i>	x_i	y_i	y_i	y_i	y_i	y_i	y_i
1	5	6.57	28.5	7	.7	5.1	9
2	5.5	6.62	30.2	7.4	.8	5.4	8.72
3	6	6.66	31.7	7.9	.9	5.6	8.5
4	6.5	6.7	33.2	8.4	1	5.9	8.3
5	7	6.73	34.7	9	1.2	6.1	8.14
6	7.5	6.75	36.1	9.7	1.4	6.3	8
7	8	6.78	37.4	10.5	1.6	6.5	7.87
8	8.5	6.8	38.7	11.4	1.9	6.7	7.76
9	9	6.81	40	12.5	2.2	6.8	7.66
10	9.5	6.83	41.2	13.7	2.5	7	7.57
11	10	6.84	42.4	15	3	7.2	7.5
12	10.5	6.86	43.6	16.6	3.4	7.3	7.42
13	11	6.87	44.7	18.3	4	7.4	7.36
14	11.5	6.88	45.9	20.3	4.6	7.6	7.3
15	12	6.89	47	22.6	5.4	7.7	7.25
16	12.5	6.9	48	25.3	6.3	7.8	7.2
17	13	6.91	49.1	28.2	7.3	7.9	7.15
18	13.5	6.92	50.1	31.6	8.4	8.1	7.11
19	14	6.93	51.1	35.5	9.8	8.2	7.07
20	14.5	6.93	52.1	39.9	11.4	8.3	7.03
21	15	6.94	53.1	44.9	13.2	8.4	7
22	15.5	6.95	54.1	50.7	15.4	8.5	6.96
23	16	6.95	55	57.2	17.9	8.6	6.93
24	16.5	6.96	55.9	64.7	20.8	8.7	6.9
25	17	6.96	56.8	73.2	24.1	8.8	6.88
26	17.5	6.97	57.7	82.9	28	8.8	6.85
27	18	6.97	58.6	94	32.6	8.9	6.83
28	18.5	6.98	59.5	106.6	37.9	9	6.81
29	19	6.98	60.4	121	44	9.1	6.78
30	19.5	6.98	61.2	137.3	51.1	9.2	6.76
31	20	6.99	62.1	156	59.4	9.2	6.75
32	20.5	6.99	62.9	177.4	69	9.3	6.73
33	21	7	63.7	201.7	80.1	9.4	6.71
34	21.5	7	64.6	229.4	93.1	9.5	6.69
35	22	7	65.4	260.9	108.2	9.5	6.68

СПИСОК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

1. Бабушка И., Витасек Э., Прагер М. Численные процессы решения дифференциальных уравнений. М., 1969. 368 с.
2. Бахвалов Н.С. Численные методы: Анализ, алгебра, обыкновенные дифференциальные уравнения. М., 1975. 631 с.
3. Березин И.С., Жидков Н.П. Методы вычислений: в 2 т. М., 1962. Т.1. 464 с. 1962. Т.2. 620 с.
4. Вивальнюк Л.М. Элементы лінійного програмування: Навч. посібник. К., 1975. 191 с.
5. Гутер Р.С., Овчинников Б.В. Элементы численного анализа и математической обработки результатов опыта. М., 1970. 432 с.
6. Давидов М.О. Курс математичного аналізу: В 3 ч. К., 1990. Ч.1. 368 с. 1991. Ч. 2. 392 с. 1992. Ч. 3. 383 с.
7. Демидович Б.П., Марон И.А. Основы вычислительной математики. М., 1970. 664 с.
8. Демидович Б.П., Марон И.А., Шувалова Э.З. Численные методы анализа. М., 1967. 367 с.
9. Информатика / А.Р.Есаян, В.И.Ефимов, Л.П.Лапицкая и др. М., 1991. 288 с.
10. Жалдак М.І., Рамський Ю.С. Чисельні методи математики: Посіб. для самоосвіти вчителів. К., 1984. 206 с.
11. Коллатц Л. Численные методы решения дифференциальных уравнений. М., 1953. 460 с.
12. Компьютеры, модели, вычислительный эксперимент: Введение в информатику с позиций математического моделирования. М., 1988. 176 с.
13. Кузнецов Ю.Н., Кузубов В.И., Волощенко А.Б. Математическое программирование. М., 1980. 302 с.
14. Мысовских И.И. Лекции по методам вычислений. М., 1962. 342 с.

Передмова	3	§4.5. Знаходження початкового опорного плану	102
Розділ 1. МАТЕМАТИЧНІ МОДЕЛІ І ЧИСЕЛЬНІ МЕТОДИ		§4.6. Використання симплекс-таблиць	108
§1.1. Роль математичного моделювання в розв'язуванні задач навколишнього світу	4	<i>Лабораторна робота № 3</i>	124
§1.2. Обчислювальний експеримент та його технологічні етапи	6	Розділ 5. ІНТЕРПОЛЮВАННЯ ФУНКЦІЙ. ЧИСЕЛЬНЕ ДИФЕРЕНЦІЮВАННЯ	
§1.3. Деякі застосування обчислювального експерименту	8	§5.1. Постановка задачі	129
§1.4. Структура похибки розв'язку задачі	10	§5.2. Інтерполяційний многочлен Лагранжа	130
§1.5. Поняття стійкості та коректності	12	§5.3. Параболічне інтерполювання за схемою Ейткіна	140
Розділ 2. ЧИСЕЛЬНІ МЕТОДИ РОЗВ'ЯЗУВАННЯ РІВНЯНЬ З ОДНІЄЮ ЗМІННОЮ		§5.4. Скінченні різниці та їх властивості	145
§2.1. Постановка задачі	14	§5.5. Перший інтерполяційний многочлен Ньютона	157
§2.2. Відокремлення коренів. Теорема про оцінку похибки наближеного значення кореня.	15	§5.6. Другий інтерполяційний многочлен Ньютона	163
§2.3. Уточнення кореня методом поділу відрізка пополам	19	§5.7. Екстраполювання й обернене інтерполювання	167
§2.4. Метричні простори і принцип стискуючих відображень	21	§5.8. Інтерполювання функцій за допомогою сплайнів	170
§2.5. Метод ітерації	26	§5.9. Чисельне диференціювання функцій	176
§2.6. Метод Ньютона і його модифікації	34	<i>Лабораторна робота № 4</i>	185
§2.7. Метод хорд	39	Розділ 6. ЧИСЕЛЬНЕ ІНТЕГРУВАННЯ ФУНКЦІЙ	
§2.8. Комбінований метод дотичних і хорд	44	§6.1. Постановка задачі	187
<i>Лабораторна робота № 1</i>	45	§6.2. Квадратурні формули Ньютона—Котеса	190
Розділ 3. РОЗВ'ЯЗУВАННЯ СИСТЕМ ЛІНІЙНИХ АЛГЕБРАІЧНИХ РІВНЯНЬ		§6.3. Формула прямокутників	191
§3.1. Постановка задачі	47	§6.4. Формула трапецій	198
§3.2. Метод послідовного виключення змінних (метод Гаусса)	49	§6.5. Формула Сімпсона	202
§3.3. Обчислення оберненої матриці методом Гаусса	58	§6.6. Порівняння і практична оцінка похибки квадратурних формул	207
§3.4. Метод квадратних коренів	60	<i>Лабораторна робота № 5</i>	214
§3.5. Метод простої ітерації	65	Розділ 7. РОЗВ'ЯЗУВАННЯ ЗВИЧАЙНИХ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ	
§3.6. Метод Зейделя	73	§7.1. Постановка задачі	216
<i>Лабораторна робота № 2</i>	75	§7.2. Метод Ейлера	217
РОЗДІЛ 4. РОЗВ'ЯЗУВАННЯ ЗАДАЧ ЛІНІЙНОГО ПРОГРАМУВАННЯ		§7.3. Методи Рунге—Кутта	228
§4.1. Приклади задач лінійного програмування	77	§7.4. Про оцінку похибки наближеного розв'язку задачі Коші	238
§4.2. Загальна задача лінійного програмування	82	<i>Лабораторна робота № 6</i>	245
§4.3. Геометричний зміст задач лінійного програмування	87	Розділ 8. МЕТОДИ ОБРОБКИ ЕКСПЕРИМЕНТАЛЬНИХ ДАНИХ	
§4.4. Загальні відомості про симплекс-метод	96	§8.1. Метод найменших квадратів	248
		§8.2. Розподіли випадкових величин	260
		§8.3. Деякі числові характеристики вибірки	264
		§8.4. Статистична перевірка гіпотез	267
		§8.5. Основи регресійного і кореляційного аналізу	275
		<i>Лабораторна робота № 7</i>	280
		<i>Список рекомендованої літератури</i>	285

Підручник

Лященко Микола Якович
Головань Микола Степанович

ЧИСЕЛЬНІ МЕТОДИ

Зав. редакцією *А.С.Кривошея*
Художник обкладинки *Г.Т.Задніпряний*
Художній редактор *Т.О.Щур*
Технічний редактор *І.М.Лукашенко*
Коректор *Н.В.Лісова*

НБ ПНУС



601116

Здано до набору 14.02.95. Підп. до друку 25.12.95. Формат 60×84/16.
Папір офс. Гарн. Тип Таймс. Офсет. друк. Ум.-друк. арк.16,74. Ум.фарбовідб.17,09.
Обл.-вид.арк. 21,98. Вид. №3650. Зам. № *6-10*.

Оригінал-макет виготовлений Київським ПКБ АСУ
252034 Київ, Рейтарська, 86

Видавництво "Либідь" при Київському університеті
252001 Київ, Хрещатик, 10

Свідоцтво про державну реєстрацію № 05591690 від 23.04.94

Київська книжкова друкарня наукової книги
252004 Київ, вул. Б.Хмельницького, 19