

КЛАСИЧНИЙ ПРИВАТНИЙ УНІВЕРСИТЕТ

Бахрушин В.Є.

МАТЕМАТИЧНІ ОСНОВИ МОДЕЛЮВАННЯ СИСТЕМ

Навчальний посібник

Запоріжжя
2009

УДК 338.27:311+519.2+681.3

ББК А.22.12

Б30

Рецензенти – д.т.н., професор О.І. Михальов;
д.т.н., професор А.П. Оксанич;
д.т.н., професор Д.М. Піза.

Рекомендовано Міністерством освіти і науки України як навчальний посібник для вищих навчальних закладів – Лист № 1.4.18-Г-2309 від 05 листопада 2008 р.

Бахрушин В.Є. Математичні основи моделювання систем: Навчальний посібник для студентів. – Запоріжжя: Класичний приватний університет, 2009. – 224 с.

ISBN 966-414-009-0

Навчальний посібник містить відомості про класифікацію систем та методів їх моделювання, теоретичні основи та математичні методи моделювання складних систем. У посібнику розглянуто приклади побудови й дослідження математичних моделей, що досліджуються за допомогою аналітичних і чисельних методів; критерії та методи визначення стійкості систем; постановки й алгоритми розв'язування оптимізаційних задач; застосування методу статистичних випробувань при моделюванні систем; особливості побудови й дослідження регресійних моделей.

УДК 338.27:311+519.2+681.3

ББК А.22.12

© Бахрушин В.Є., 2009

© Класичний приватний університет, 2009

ISBN 966-414-009-0

ВСТУП

Моделювання є одним з основних сучасних методів дослідження систем. Зазвичай воно передбачає створення концептуальної моделі об'єкта дослідження, її формалізацію та перетворення у математичну або комп'ютерну модель, перевірку адекватності й подальше дослідження отриманої моделі за допомогою аналітичних або чисельних методів і сучасних комп'ютерних технологій. Застосування методів моделювання часто дає змогу отримати більш точні відомості про поведінку й характеристики досліджуваних систем і процесів, ніж при їх безпосередньому вивченні, витрачаючи при цьому менше часу та коштів. Крім того у багатьох випадках використання інших методів дослідження взагалі виявляється неможливим через їх небезпеку, коштовність, тривалість або через відсутність відповідного обладнання і методик. Класичними прикладами таких досліджень є моделі кліматичних наслідків ядерного вибуху, розроблені М.М. Мойсеєвим та К. Саганом; модель світової динаміки Дж. Форрестера; модель протікання вірусного гепатиту Г.І. Марчука, кваркова модель будови елементарних частинок М. Гелл-Мана, моделі динаміки популяцій, моделі виникнення й еволюції Всесвіту тощо. Методи моделювання широко застосовують при проектуванні систем різного типу – атомних реакторів, двигунів літаків і ракет, металургійних печей, інтегральних схем, систем автоматичного й автоматизованого управління та інших.

Сьогодні моделювання систем найчастіше реалізують за допомогою сучасних комп'ютерних технологій. Такий підхід передбачає необхідність попередньої формалізації концептуальної моделі об'єкта дослідження та її подання у вигляді, придатному для реалізації тих чи інших алгоритмів чисельного аналізу або комп'ютерної імітації. Обидва підходи передбачають необхідність застосування сучасних математичних методів, що використовуються при створенні алгоритмів моделювання. Навіть при застосуванні спеціалізованих пакетів програмного забезпечення досліднику необхідно володіти основами відповідних математичних методів, оскільки користування такими пакетами зазвичай пе-

редбачає необхідність вибору оптимального алгоритму й певних параметрів його реалізації, іноді з декількох десятків можливих варіантів. Це зумовлює необхідність вивчення основних методів математичного моделювання систем майбутніми фахівцями.

Курс "Моделювання систем" є нормативною складовою стандартів підготовки бакалаврів, що навчаються за напрямками підготовки "Комп'ютерні науки", "Комп'ютеризовані системи, автоматика і управління", "Комп'ютерна інженерія" тощо. Зазвичай цей курс поділяють на дві основні складові – математичні методи моделювання систем й інструментальні засоби моделювання. У цьому посібнику в основному розглядаються питання першого з вказаних блоків.

Методи математичного моделювання широко застосовують при дослідженні самих різноманітних систем і процесів – природних, технічних, екологічних, економічних, соціальних тощо. В сучасних умовах їх найчастіше реалізують за допомогою комп'ютерних технологій. Тому формування знань та навичок, необхідних для побудови й дослідження математичних моделей складних систем, є необхідною складовою підготовки фахівців в галузі інформаційних і комп'ютерних технологій.

Передбачається, що студенти мають підготовку з математичного аналізу, лінійної алгебри, теорії диференціальних рівнянь, теорії інформації та математичної статистики, чисельних методів, програмування у межах, встановлених вимогами відповідних освітньо-професійних програм. Крім того, передбачається, що студенти знайомі з основами фізики в рамках програми нетехнічних спеціальностей вищих навчальних закладів.

Автор вдячний за обговорення матеріалу посібника й окремих питань моделювання систем чл.-кор. НАН України В.В. Сльозову та В.Г. Литовченку, професорам М.С. Блантеру, О.М. Горбаню, Г.В. Корнічу, Д.І. Левінзону, Л.М. Любчику, І.М. Лященко, О.І. Михальову, А.П. Оксаничу, Д.М. Пізі, Л.Н. Сергєєвій, доцентам В.М. Буйницькій, М.О. Ігнашіній, І.М. Нацкоку.

1. КЛАСИФІКАЦІЯ, ЗАГАЛЬНІ ВЛАСТИВОСТІ ТА ОСНОВНІ ТИПИ МОДЕЛЕЙ СИСТЕМ

1.1. Класифікація моделей

Моделюванням називають заміщення одного об'єкта іншим з метою отримання інформації про найважливіші властивості об'єкта-оригіналу шляхом здійснення експериментів з об'єктом-моделлю. **Модель** можна визначити як об'єкт, що у деяких відношеннях збігається з оригіналом і є засобом подання, пояснення та/або прогнозування його поведінки.

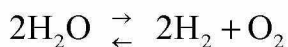
За формою подання вирізняють реальні та абстрактні моделі.

Реальні (матеріальні, речовинні, натурні) моделі є реальними об'єктами, які у тому чи іншому відношенні можуть замінювати об'єкти-оригінали. При цьому вони відтворюють ті характеристики та властивості останніх, що цікавлять дослідника, і є більш зручними для дослідження. Прикладами таких моделей можуть бути моделі літаків, призначені для аерогідродинамічних випробувань, макети архітектурних споруд, еквівалентні електричні схеми напівпровідникових приладів тощо.

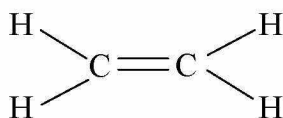
Абстрактні (ідеальні) моделі є ідеальними конструкціями, створеними засобами мислення, свідомості. Їх поділяють на мовні, графічні, знакові, математичні й алгоритмічні. Прикладами ідеальних моделей є рівняння, формули, схеми, плани тощо, що описують різноманітні реальні процеси та об'єкти. Наприклад:

$$\frac{d^2x}{dt^2} + \omega^2 x = 0$$

– модель гармонічних коливань;



– модель процесу дисоціації води;



– модель молекули етилену.

При побудові абстрактних моделей зазвичай використовують спеціальні мови, терміни, символи тощо. Ці мови створюють на основі тих чи інших природних мов, але значення наукових термінів можуть суттєво відрізнятися від значень відповідних слів у природній та професійних мовах. Прикладом є значення терміна "реакція" у хімії, біології та політології.

У сучасній науці і техніці велике значення мають **математичні** та **комп'ютерні** моделі. У природознавстві й технічних науках їх побудова та аналіз є одним з основних методів наукового дослідження і проектування. При вивченні екологічних та соціально-економічних макросистем математичні моделі використовують в основному для прогнозування тенденцій розвитку систем, визначення їх загальних властивостей, стійкості, критичних умов існування й таке інше.

При побудові моделей часто виникає необхідність встановлення **відношення подібності** між моделлю та оригіналом. Для цього використовують такі основні методи.

Пряма подібність цілеспрямовано встановлюється під час створення моделі. Прикладами моделей прямої подібності є макети споруд, моделі літаків, шаблони, викройки, фотографії тощо. Для їх побудови використовують теорію подібності, яка дає змогу перераховувати результати вимірювань, одержувані в експериментах з моделями, на результати, що можуть бути одержані у подібних умовах на оригіналі.

Непряма подібність не встановлюється дослідником, а існує об'єктивно. Вона проявляється як збіг або достатня близькість абстрактних моделей різних систем та процесів. Наприклад, можна говорити про математичну тотожність моделей гармонічних коливань у різноманітних механічних, електричних, хімічних та інших системах. Тому коливальні системи одного типу можна використовувати як моделі коливальних систем іншого типу. Зокрема, електричний коливальний контур можна використовувати як матеріальну модель довільного гармонічного осцилятора, завдяки зручності автоматичної реєстрації його вихідної характеристики. Прикладами моделей непрямої подібності є годинник

(аналог часу), піддослідні тварини у медичних дослідженнях (аналог організму людини) тощо. Існують спеціальні аналогові обчислювальні машини, за допомогою яких можна створювати електричні ланцюги, що є аналогами різноманітних динамічних систем.

Умовна подібність встановлюється як результат домовленості. Прикладами моделей умовної подібності є алгоритм (модель послідовності дій), гроші (модель вартості), паспорт особи (модель його власника), робочі креслення (модель готового виробу), географічні карти (модель місцевості), умовні сигнали тощо.

Залежно від мети моделювання, вирізняють пізнавальні та прагматичні моделі. **Пізнавальні (дескриптивні, описові)** моделі відповідають теоретичним цілям і є формою організації та подання знань. Вони відображають те, що вже існує, а їх основним завданням є наближення моделі до реальності. У структурному системному аналізі їх називають моделями "Як є" ("As is").

Метою розробки й застосування пізнавальних моделей зазвичай є:

- вивчення об'єкта (наукові дослідження), тобто найбільш повне і правильне відображення його состава, структури й властивостей;
- управління об'єктом, тобто найточніше відображення впливу певних факторів у визначеному діапазоні їх змінювання, на його істотні властивості;
- прогнозування, тобто найточніше визначення подальшої поведінки об'єкта і майбутніх значень його параметрів;
- навчання, тобто правильне відображення певних властивостей об'єкта.

Прикладами пізнавальних моделей є наукові поняття, рівняння залежностей тощо. Пізнавальні моделі зазвичай використовують як теоретичну основу для проектування алгоритмів, програмних систем, засобів автоматизації, систем управління, електронних систем тощо.

Прагматичні (нормативні, прескриптивні) моделі є засобом керування, організації практичної діяльності людей, уявленням про зразково правильні дії, порядок їх виконання або результати. Основним завданням таких моделей є наближення реальності до них, оскільки у цьому разі модель розуміється як пев-

ний стандарт. У структурному системному аналізі прагматичні моделі називають моделями "Як має бути" ("To be").

Типовими прикладами прагматичних моделей є: мета (модель бажаного майбутнього); алгоритм (модель майбутньої діяльності); проекти машин та споруд; складальні креслення; шаблони; закони; плани та програми дій; розклади; статuti й посадові інструкції організацій, бізнес-плани тощо.

В основі іншого підходу до класифікації моделей систем лежить їх поділ на статичні й динамічні. **Статичні** моделі призначені для визначення будови та характеристик, що відповідають певному відносно стабільному стану системи. Їх прикладами є структурні схеми, кристалічні решітки кристалів, географічні карти, розклади руху поїздів, кваліфікаційні характеристики фахівців тощо. **Динамічні** моделі відображають процеси змінювання стану та характеристик системи. Їх прикладами є рівняння фізичних процесів (рівняння дифузії, теплопровідності та інші), моделі хімічних реакцій, перспективні плани розвитку організацій, моделі циркуляції атмосфери й т. п.

1.2. Основні властивості моделей

Основними властивостями моделей вважають їх скінченність, наближеність, повноту, адекватність та істинність.

Скінченність моделі полягає у тому, що вона відображає лише деякі з характеристик та відношень, властивих оригіналу. Скінченність зумовлюється обмеженістю часу, пам'яті ЕОМ та інших ресурсів, потрібних для розробки та аналізу моделі. Іншою причиною є те, що із великої кількості різноманітних зв'язків досліджуваної системи з навколишнім середовищем, а також її компонентів та зв'язків між ними лише деякі істотно впливають на властивості, що вивчаються. Тому немає потреби у врахуванні інших зв'язків і компонентів. Більш того врахування в моделі несуттєвої інформації призведе до отримання надмірно складної моделі, результати дослідження якої можуть виявитися досить неточними через її математичну некоректність та похибки округлення при розрахунках. Зокрема, при створенні розкладу занять в університеті немає пот-

реби враховувати те, що студенти складаються з певних атомів, між якими діють ковалентні, водневі та інші хімічні зв'язки.

Наближеність моделі означає, що вона лише наближено відображає характеристики та відношення, що існують в об'єкті дослідження. Типовими прикладами наближень, які використовують при математичному моделюванні, є заміна дискретних систем неперервними та навпаки, задання обмежень на точність чисельних розрахунків, заміна нелінійних залежностей лінійними тощо.

Із скінченності та наближеності моделі випливає, що вона відображає оригінал **неповно**. Ступінь повноти моделі визначається метою та завданнями дослідження. Слід зазначити, що неповнота моделі не обов'язково є її недоліком, оскільки немає сенсу будувати модель, яка є точним повторенням оригіналу. Важливо, щоб модель достатньо повно відтворювала лише те риси об'єкта дослідження, що вивчаються, й ті його складові, зв'язки та зовнішні впливи, які можуть істотно впливати на них.

Адекватність моделі характеризує можливість реалізації цілей моделювання, а її **істинність** – відповідність моделі сукупності наявних знань про об'єкт дослідження. Істинність моделі не є гарантією її адекватності. Зокрема, неадекватність істинної моделі може бути зумовлена накопиченням помилок через необхідність виконання великого обсягу розрахунків. Тут іноді спрацює закон "краще є ворогом гарного". З іншого боку, адекватними можуть бути моделі, що не є істинними. Типовими прикладами таких моделей є регресійні моделі, які дають можливість прогнозувати поведінку досліджуваної системи у певному діапазоні зміни вхідних параметрів, але не відображають наявні знання про будову, зв'язки та внутрішні процеси в ній й тому свідомо не можуть бути істинними.

Спільною рисою всіх моделей є те, що **модель є цільовим відображенням оригіналу**. Це означає, що при побудові моделі у першу чергу слід враховувати цілі й завдання дослідження оригіналу. При зміні мети дослідження часто буває необхідно змінити й модель, що використовується. Наприклад, якщо вивчається орбітальний рух Землі навколо Сонця, моделлю Землі може бути

матеріальна точка. Якщо ж вивчаються геологічні процеси, використовують модель оболонкової структури Землі (ядро – мантія – літосфера – гідросфера – атмосфера). При дослідженні кліматичних процесів і прогнозуванні погоди суттєвими є конвекційні потоки у нижніх шарах атмосфери, течії в океані, а також коливання коефіцієнта відбиття сонячного світла у різних точках поверхні Землі. Таким чином, для одного й того самого об'єкта може існувати багато різних моделей. Тому часто говорять про **множинність моделей систем**.

Якість моделей залежить як від об'єктивних, так і від суб'єктивних факторів. До об'єктивних насамперед належать обмеження пов'язані з можливостями використовуваних методів і програмного забезпечення. Наприклад, до появи сучасних суперкомп'ютерів і методів паралельного програмування моделі аеро-й гідродинамічних процесів, що відбуваються під час руху у них з високою швидкістю твердих тіл складної форми (ракет, літаків, кораблів), не давали змоги дослідити деякі суттєві ефекти через неможливість застосування достатньо щільних сіток при виконанні розрахунків методами скінченних елементів.

Основними суб'єктивними факторами, що впливають на якість моделей систем, є:

- вибірковість відображення властивостей оригінала дослідником, яка залежить від рівня його кваліфікації, попереднього досвіду, самопочуття та інших чинників;

- властивість людського розуму заповнювати прогалини в наявній інформації на основі попереднього досвіду дослідника, який не завжди є адекватним ситуації, що вивчається;

- застосування для нового об'єкта так званих типових моделей, які можуть виявитися непридатними для його подання.

Існує декілька причин, що ведуть до необхідності вивчення моделей замість реальних об'єктів.

По-перше, реальні об'єкти зазвичай зазнають впливу великої кількості різноманітних факторів, що унеможлиблює отримання будь-яких корисних висновків за результатами їх дослідження. Тому виникає необхідність спростити

їх за допомогою моделей, що ігнорують вплив несуттєвих з погляду досягнення мети дослідження чинників.

По-друге, у багатьох випадках дослідження реальних систем виявляється або дуже коштовним, або занадто тривалим, або завдає неприпустимої шкоди навколишньому середовищу чи дослідникам. Типовими прикладами таких ситуацій є дослідження парникового ефекту, кліматичних наслідків ядерної війни, розповсюдження епідемій і т. п.

По-третє, у багатьох випадках метою дослідження є прогнозування поведінки об'єкта дослідження. У більшості випадків це неможливо без побудови їх моделей.

По-четверте, до необхідності розробки моделей часто веде неможливість безпосереднього вивчення реальних об'єктів через їх розмір (елементарні частинки, галактики), час протікання досліджуваних процесів або руйнування чи істотне спотворення під час проведення експериментальних досліджень.

1.3. Основні види та етапи моделювання

Відповідно до класифікації моделей, вирізняють такі види моделювання.

Концептуальне моделювання за допомогою природних або штучних мов визначає принципові аспекти побудови або функціонування об'єкта дослідження. Прикладами завдань концептуального моделювання є: побудова сценаріїв розвитку організації з метою прийняття стратегічних рішень; вибір інерціальної або неінерціальної системи відліку при записі рівнянь руху у механіці; врахування або неврахування динаміки системи при її вивченні тощо. Відомим з курсу фізики прикладом концептуальної моделі є модель ідеального газу. Згідно з нею ідеальним газом є сукупність молекул, що є матеріальними точками, які не взаємодіють одна з одною інакше, ніж через пружні зіткнення. Як бачимо, ця модель не містить жодних математичних елементів, але враховуючі відомі поняття й закони фізики її можна достатньо легко перетворити у математичну модель. Зазвичай, якщо це можливо, слід будувати концептуальну модель так, щоб її можна було формалізувати за допомогою тих чи інших матема-

тичних методів. Це дає змогу у подальшому використовувати комп'ютерні технології її дослідження.

Інтуїтивне моделювання передбачає прогнозування поведінки досліджуваної системи на основі уявних експериментів і врахування практичного досвіду дослідника. Класичним прикладом інтуїтивного моделювання через здійснення уявних експериментів є дискусія А. Ейнштейна і Н. Бора стосовно філософських засад квантової механіки. Завдання моделювання на основі практичного досвіду зазвичай вирішують фахівці-практики: інженери, лікарі, агрономи, кваліфіковані робітники тощо. Наприклад досвідчений сталевар часто може точніше спрогнозувати хід плавки та її результати, ніж це можна зробити на основі певних фізико-хімічних або математичних моделей. Але це стосується лише плавок добре відомих йому марок сталі у стандартних печах.

Фізичне моделювання здійснюють шляхом дослідження як моделі реального фізичного об'єкта, між яким і об'єктом-оригіналом встановлені відношення подібності, що ґрунтуються на подібності фізичних процесів, що відбуваються в них. Зокрема при проектуванні гідротехнічних споруд здійснюють попереднє дослідження моделей, що є їх зменшеними аналогами. Іншим прикладом фізичного моделювання є апробація нових лікарських засобів на піддослідних тваринах.

Структурно-функціональне моделювання передбачає побудову і використання як моделей графіків, таблиць, блок-схем тощо, для яких встановлено певні правила їх об'єднання й перетворення. Прикладами структурно-функціонального моделювання є: розробка блок-схем алгоритмів роботи комп'ютерних програм; побудова й перетворення електричних схем; створення класифікаційних схем; розробка розкладів занять у навчальних закладах і т. ін.

Математичне (логіко-математичне) моделювання передбачає побудову й дослідження моделей математичними та логічними засобами. Прикладами математичних моделей є: рівняння коливань як модель циклічного процесу; логістична крива як модель кривої зростання; рівняння рівномірного та рівноприскореного руху тощо.

Імітаційне (програмне) моделювання передбачає побудову логіко-математичної моделі об'єкта дослідження у вигляді певної послідовності дій (алгоритму), що реалізована як комп'ютерна програма. Зазвичай такий вид моделювання використовують, якщо певні дії відбуваються багаторазово і мають певні відомі статистичні характеристики. У таких випадках встановити зв'язок між входами й виходами досліджуваної системи у вигляді певної математичної моделі зазвичай буває доволі складно. Навпаки, відтворити за допомогою комп'ютерної імітації послідовність подій, що відбуваються в системі, і отримати у числовому вигляді значення її вихідних параметрів, буває набагато легше.

Початковим етапом дослідження складної системи є побудова **моделі чорного ящика** (рис. 1.1). При формальному описі вона задає множини вхідних та вихідних змінних системи, але не містить ніякої інформації про відношення між цими двома множинами.

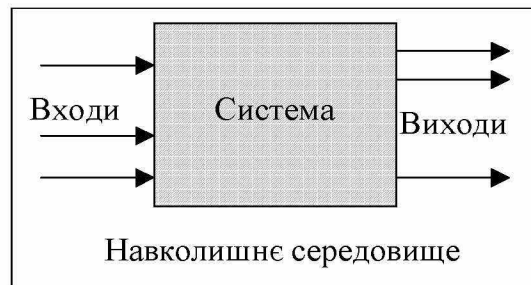


Рис. 1.1. Структура моделі чорного ящика

Під **вхідними змінними (входами)** розуміють впливи навколишнього середовища на досліджувану систему, а під **вихідними змінними (виходами)** – її впливи на навколишнє середовище. До вихідних змінних зокрема належать ті, що відображають контрольовані за допомогою вимірювальних приладів характеристики реальної системи, а до вхідних – управляючі впливи на неї та неконтрольовані впливи зовнішнього середовища. Побудова моделі чорного ящика є нетривіальним завданням, оскільки у реальних систем існує велика кількість входів та виходів, але залежно від мети моделювання деякі з них можуть виявитися суттєвими, а інші – несуттєвими. Завданням дослідника при побудові мо-

делі чорного ящика є визначення обмеженої кількості тих входів і виходів, суттєвих з погляду реалізації цілей моделювання. Крім того, завжди є невідомі дослідникові й тому невраховані ним впливи навколишнього середовища на досліджувану систему і навпаки – системи на навколишнє середовище.

У багатьох випадках модель чорного ящика є єдиною можливою моделлю системи. Зокрема, така ситуація виникає у випадках, коли внутрішня будова системи є невідомою (наприклад, при дослідженні елементарних частинок), а також якщо необхідно мінімізувати вплив дослідження на систему (наприклад, коли вивчається психіка людини). Подальше дослідження таких систем може відбуватися шляхом побудови регресійних моделей, які у математичній формі відображають зв'язок між їх вхідними та вихідними змінними. У деяких випадках такі моделі можуть потім стати математичним відображенням певних концептуальних моделей. Відомо, наприклад, що рівняння, які описують вільне падіння тіл у механіці, спочатку були отримані Галілео Галілеєм саме як регресійні моделі певного "чорного ящика", і лише значно пізніше теоретично виведені на основі законів класичної механіки.

Для вивчення динаміки систем, моделлю яких є "чорний ящик", часто застосовують методи аналізу часових рядів, рівні яких утворюють показники системи, що цікавлять дослідника.

Наступним етапом є побудова моделей складу та структури системи, а також її структурної схеми. При цьому необхідно мати на увазі, що для кожної системи можна побудувати багато таких моделей. Проте існує певна оптимальна модель, яка залежить від мети дослідження і дає змогу досягти цієї мети з найменшими витратами ресурсів дослідника.

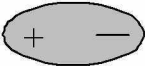
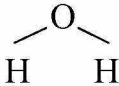
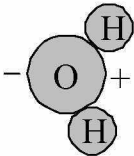
Моделлю складу є перелік основних компонентів (складових частин) досліджуваної системи. При цьому вирізняють два типи компонентів – елементи, які розглядають як неподільні складові, та підсистеми – складові, що з погляду досягнення цілей моделювання мають розглядатися як утворення, що складаються з інших компонентів. При побудові моделі складу необхідно виявити всі підсистеми та елементи, які суттєво впливають на досліджувані характеристики

і параметри системи. Модель складу відображає неоднорідність досліджуваної системи.

У **моделі структури** треба описати всі зв'язки між підсистемами та окремими елементами, та зв'язки всередині підсистем, що є важливими для досягнення мети аналізу. Якщо це можливо, зв'язки необхідно подати у математичному вигляді.

Після цього будують **структурну схему** досліджуваної системи, яка є сукупністю моделей чорного ящика, складу та структури. В ній вказують всі важливі компоненти системи, зв'язки між ними, зв'язки компонентів системи із зовнішнім середовищем (входи та виходи системи). Побудова структурної схеми дає можливість виявити тип структури (лінійна, деревоподібна, матрична, сітьова тощо) і наявність зворотних зв'язків, що має суттєве значення для обрання методу подальшого аналізу системи.

Як приклад розглянемо різні види моделей молекули води:

Модель чорного ящика	
Модель складу	H ₂ O (2 атоми водню й один атом кисню)
Модель структури	2 ковалентні полярні зв'язки O - H
Структурні схеми	H:O:H  

Модель чорного ящика не містить відомостей ані про склад, ані про структуру молекули, але відображає наявність в неї дипольного моменту, тобто те, що один бік молекули притягується до позитивно заряджених частинок, а інший – до негативно заряджених. Модель складу дає інформацію про те, з яких атомів складається молекула води, а модель структури – про те, як ці атоми зв'язані один з одним. Структурна схема першого типу відображає той факт, що утворення зв'язків пов'язано з утворенням спільних електронних пар між атомом кисню й кожним з двох атомів водню. Друга структурна схема відображає

те, що зв'язки існують тільки між атомами кисню та водню, при цьому є певні геометричні співвідношення у взаємному розташуванні атомів. Третя структурна схема додатково містить дані про наявність у молекули кисню дипольного моменту.

Всі структурні схеми мають загальні властивості, які дають можливість застосовувати до їх аналізу сучасні математичні методи, зокрема апарат теорії графів, теорії структур та інших розділів математики. На рис. 1.2 показані деякі типи структур, що найчастіше зустрічаються у реальних системах.

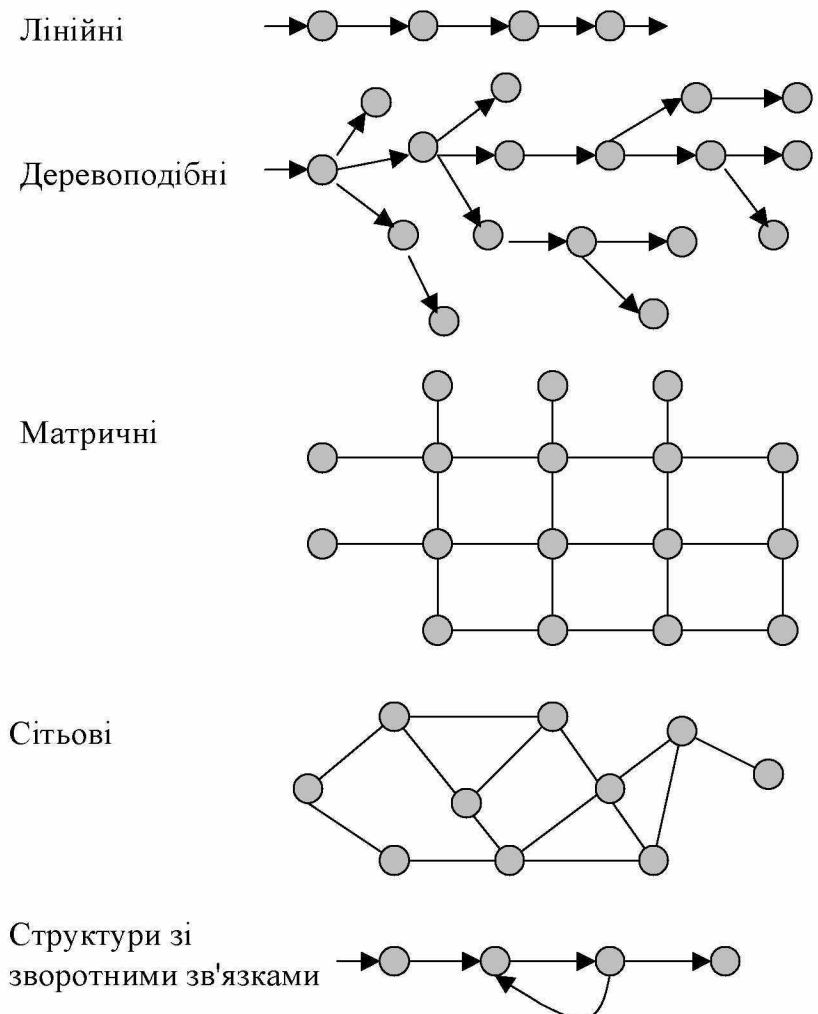


Рис. 1.2. Основні типи структур систем

1.4. Приклади структур систем

Розглянемо приклади структур деяких систем більш докладно. На рис. 1.3 показано структурну схему найпростішої системи з управлінням. **Об'єктом управління** є система, над якою здійснюють управління. Вектор вихідних змінних $\mathbf{X}_{out} = \{x_{out1}, \dots, x_{outn}\}$ характеризує стан об'єкта управління у певний момент часу. Вектор збурень $\mathbf{F} = \{f_1, \dots, f_k\}$ характеризує вплив зовнішнього середовища на об'єкт управління. Вектор вхідних змінних $\mathbf{X}_{in} = \{x_{in1}, \dots, x_{inm}\}$ характеризує управляючі впливи, що робляться для приведення системи до заданого стану.

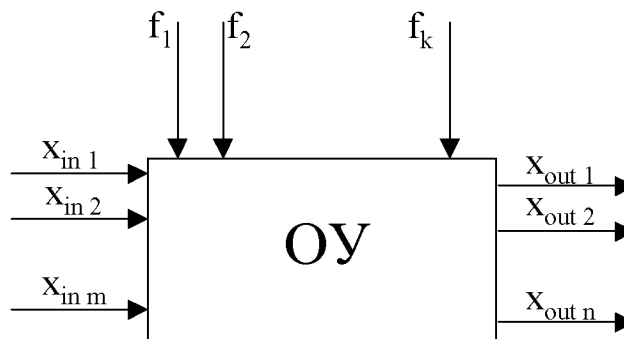


Рис. 1.3. Схема системи з управлінням

Наступні моделі відображають структури зі зворотними зв'язками. На рис. 1.4 показано схему управління за розімкненим й замкненим циклами. У першому випадку між вхідними та вихідними змінними немає безпосереднього зв'язку. Він здійснюється через регулятор (Р) і об'єкт управління (ОУ).

Як правило **управління за розімкненим циклом** реалізується за допомогою людини – оператора (Оп), яка, спостерігаючи за значеннями вихідної величини, встановлює потрібні значення вхідних параметрів (рис. 1.4.а).

При **управлінні за замкненим циклом** (рис. 1.4.б) зміни вихідних величин передають на входи системи за допомогою сукупності пристроїв, які називають зворотним зв'язком (ЗЗ). Вхідною величиною зворотного зв'язку є вихідна величина об'єкта управління \mathbf{X}_{out} .

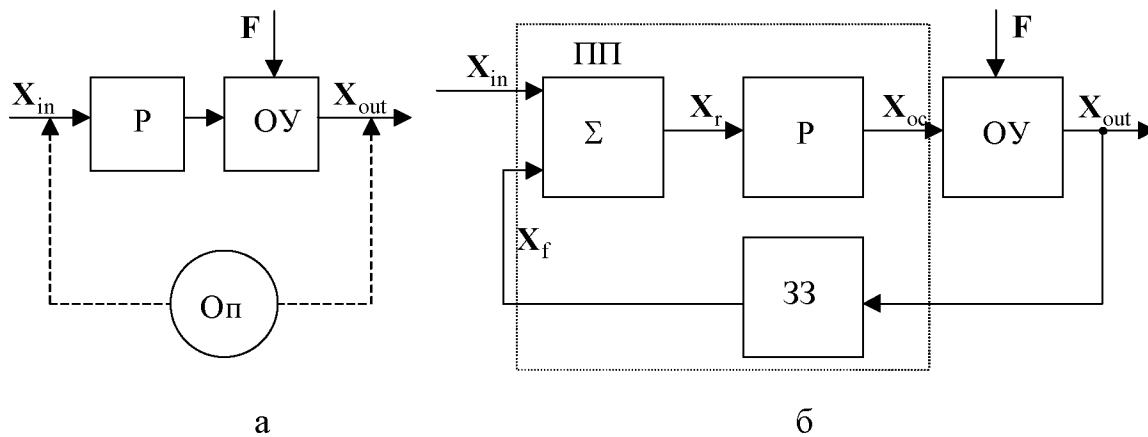


Рис. 1.4. Схеми управління за розімкненим (а) й замкненим (б) циклами

Вихідну величину зворотного зв'язку будемо позначати X_f . Вона сумісно з вхідною величиною формує регулюючий вплив X_r на вході регулятора, який у свою чергу формує управляючий вплив X_{oc} , що подається на вхід об'єкта управління.

Функціональну схему системи автоматичного управління у загальному випадку можна подати (рис. 1.5) як сукупність таких функціональних компонентів: управляючий пристрій (УП), підсумовуючий пристрій (ПП), пристрій послідовного корегування (ППК), підсилювач (П), виконавчий елемент (ВЕ), місцевий зворотний зв'язок (МЗЗ), об'єкт управління (ОУ) й головний зворотний зв'язок (ГЗЗ).

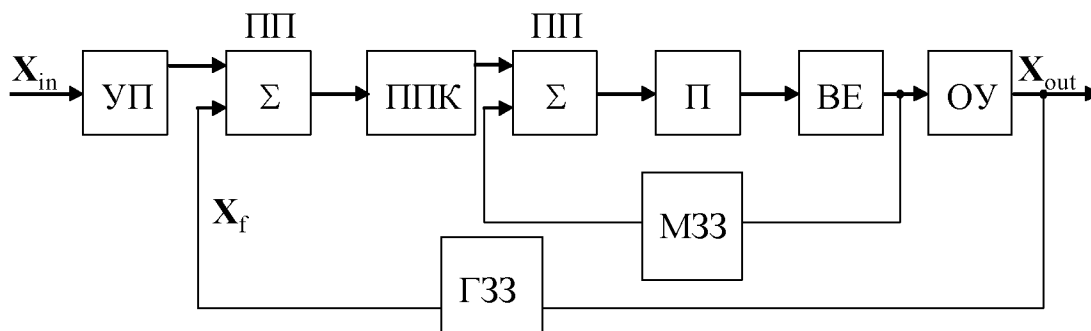


Рис. 1.5. Функціональна схема системи автоматичного управління

Загальна схема процесу автоматизованого управління показана на рис. 1.6. Згідно з нею, **автоматизоване управління** можна визначити як управління з розімкненим зворотним зв'язком, контур якого містить ЕОМ і особу, що приймає рішення.

Як структури зі зворотним зв'язком зазвичай можна розглядати типові алгоритми математичного моделювання. Прикладом є відома з курсу чисельних методів блок-схема алгоритму розв'язування нелінійного алгебраїчного рівняння методом ділення навпіл, наведена на рис. 1.7.

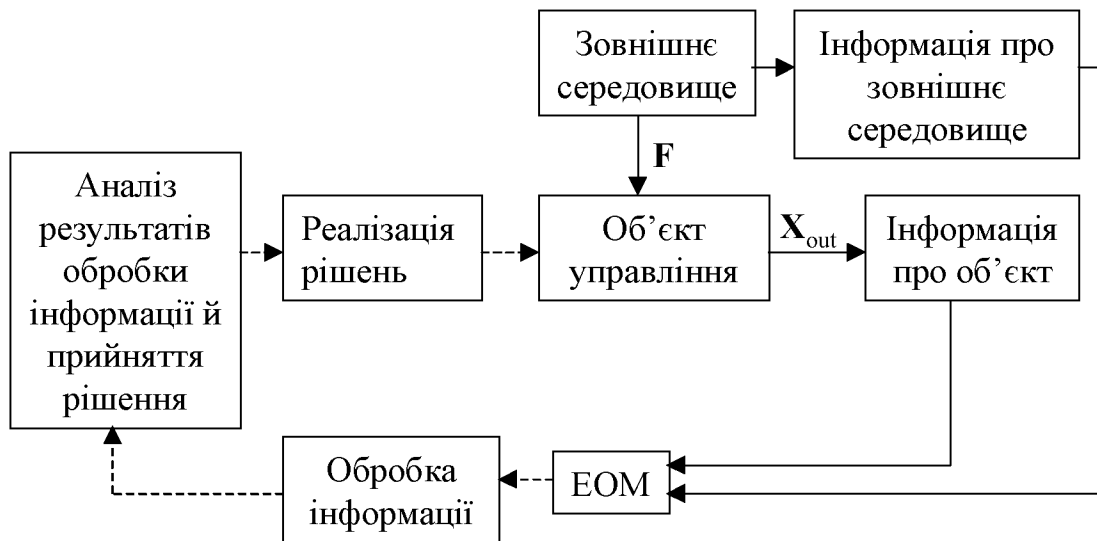


Рис. 1.6. Загальна схема автоматизованого управління

Наступні дві моделі відображають деревоподібні структурні схеми.

На рис. 1.8 зображено **дерево цілей** при розробці автоматизованої системи пошуку дефектів в рентгенівських знімках зварних швів¹. Побудова дерева цілей є одним з основних етапів дослідження або проектування систем. Вона передбачає визначення головної мети, а також декількох ієрархічних рівнів часткових цілей, досягнення кожної з яких необхідно для досягнення пов'язаної з нею цілі вищого ієрархічного рівня.

¹ Методичні рекомендації щодо побудови дерева цілей та дерева проблем. – НУ "Львівська політехніка" [Електронний ресурс]. – Режим доступу: <http://asu.we.lviv.ua/?part=library.metod>

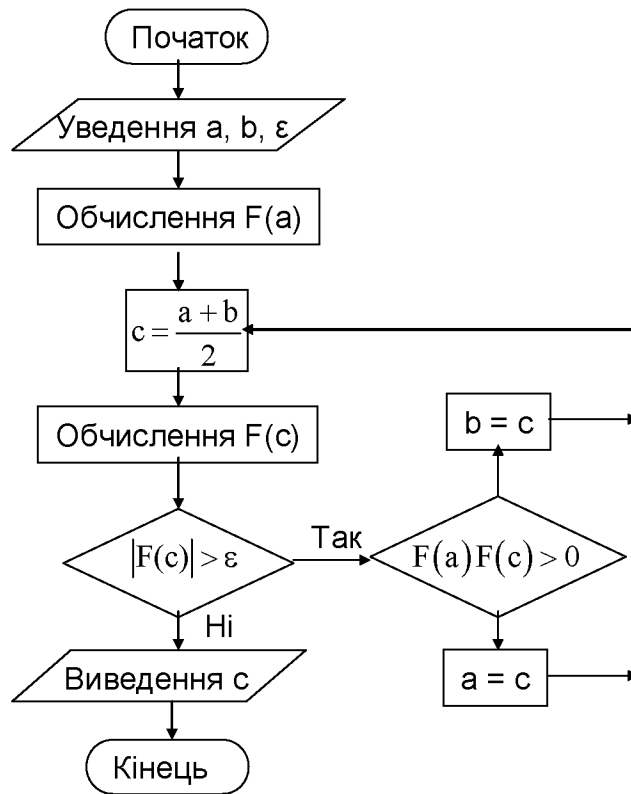


Рис. 1.7. Блок-схема алгоритму розв'язування нелінійного алгебраїчного рівняння методом ділення навпіл

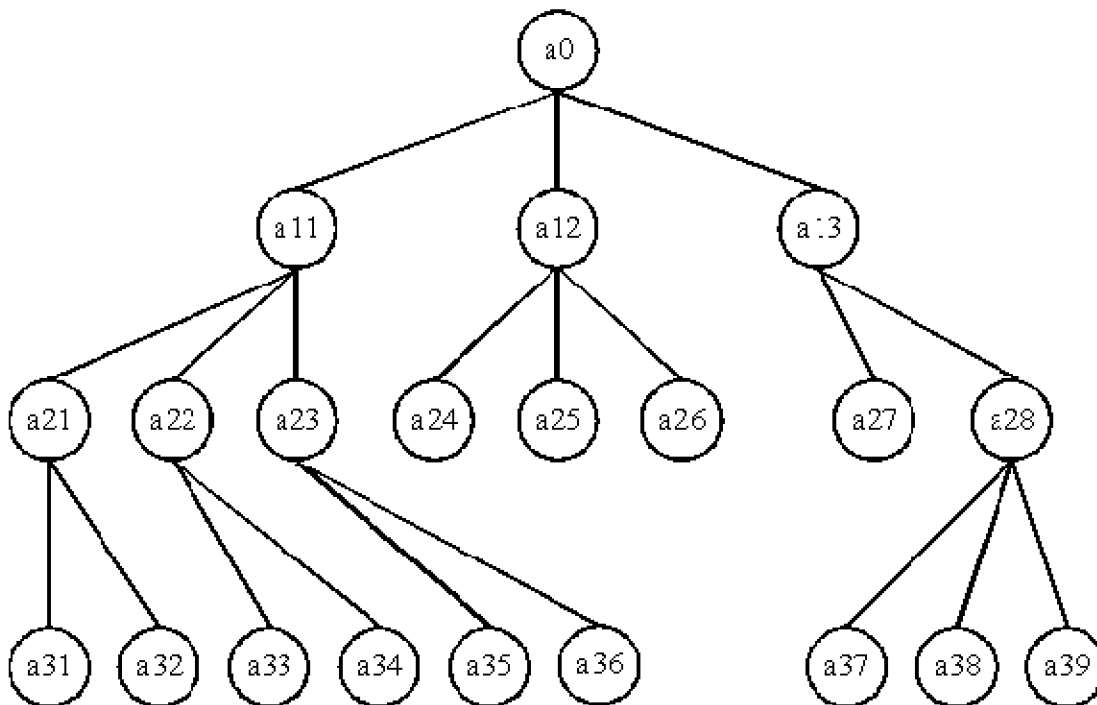


Рис. 1.8. Дерево цілей

Елементами такого дерева є:

- a_0 – головна мета – створення програмного пакета для пошуку дефектів у рентгенівських знімках зварних швів;
- a_{11} – математична модель системи;
- a_{21} – знаходження функції інформативності;
- a_{31} – знаходження значення контрасту;
- a_{32} – знаходження значення дисперсії;
- a_{22} – знаходження локальних максимумів;
- a_{33} – розбиття зображення на локальні регіони;
- a_{34} – оптимізація пошуку максимумів функції інформативності для локальних регіонів;
- a_{23} – створення бінарного зображення;
- a_{35} – стохастична гіпотеза наявності об'єкта в регіоні пошуку;
- a_{36} – визначення адаптивного порогу бінаризації;
- a_{12} – інформаційна база;
- a_{24} – опрацювання літературних джерел;
- a_{25} – забезпечення можливості доступу до нещодавно друківаних статей;
- a_{26} – забезпечення можливості доступу до Internet;
- a_{13} – програмна реалізація поставленої задачі;
- a_{27} – залучення спеціалістів, які мають необхідні навички роботи з сучасними технологіями;
- a_{28} – визначення апаратних засобів;
- a_{37} – забезпечення застосування сучасних ЕОМ;
- a_{38} – визначення засобів друку результатів роботи (різноманітні принтери);
- a_{39} – визначення засобів вводу графічних зображень (різноманітні сканери).

Іншим типовим прикладом деревоподібних структур є **класифікаційні моделі**. Необхідність їх побудови пов'язана з тим, що вони дають змогу визначати загальні властивості систем, що належать до певного класу, а також застосовувати типові моделі, методи дослідження й проектування таких систем. Зазвичай моделі класифікації на кожному рівні ієрархії вирізняють декілька підмножин досліджуваної системи (підсистеми), що не перетинаються один з одним, але у сукупності містять усі її компоненти. Це призводить до деревоподібної структури відповідної класифікаційної моделі. Добре відомими зі шкільного курсу біології прикладами таких моделей є класифікації живих і рослинних організмів. Іншим прикладом є показана на рис. 1.9 модель класифікації польових транзисторів.

Лінійні структури можуть використовуватися при побудові простих алгоритмів, наприклад алгоритмів, що пов'язані з необхідністю виконання певної визначеної послідовності дій. Їх також використовують при моделюванні деяких простих систем (RLC-контур в електродинаміці) або компонентів більш складних систем.

Типовими прикладами застосування матричних структур є організація структур даних, що використовуються у реляційних базах; розробка алгоритмів розв'язування задач аеро- та гідродинаміки методом скінчених елементів; проектування систем детектування випромінювань, розпізнавання й обробки зображень тощо.

Сітьові структури використовують при розробці алгоритмів складних розрахунків, що передбачають розгалуження й організацію циклів; проектуванні систем зв'язку, транспортних, електричних та інших мереж тощо.

1.5. Методи формалізованого подання систем

У більшості випадків моделювання передбачає формалізоване подання досліджуваної системи. Згідно з Ф.Є. Темніковим вирізняють такі основні методи формалізованого подання: аналітичні, статистичні, теоретико-множинні, логічні, лінгвістичні й семіотичні, графічні.



Рис. 1.9. Класифікація польових транзисторів

При застосуванні **аналітичних методів** властивості багатовимірної й багатозв'язної системи відображаються у n -вимірному просторі у вигляді однієї точки, що рухається з часом у фазовому просторі. Відображення задають певним оператором (функцією, функціоналом). Взаємодія систем та підсистем подається у вигляді рівнянь зв'язку. Поведінка точок та їх взаємодія описуються строгими співвідношеннями.

Основним математичним апаратом, що використовується у цій групі методів є: класичний математичний аналіз, теорія диференціальних рівнянь, математичне програмування, теорія ігор тощо.

Аналітичні методи застосовують при дослідженні систем, властивості яких можна відобразити за допомогою детермінованих величин і залежностей. Прикладами є задачі руху та стійкості систем, оптимального розміщення, розподілу робіт і ресурсів; оптимальної стратегії поведінки; автоматичного управління і т. ін. Для вивчення складних багатокомпонентних і багатокритеріальних систем аналітичні методи зазвичай не використовують через проблеми з отриманням необхідних аналітичних залежностей і обґрунтуванням адекватності моделей.

Статистичні методи передбачають відображення системи в n -вимірному просторі "розмітою" точкою, яка характеризується функцією розподілу (щільності розподілу) або математичним сподіванням та дисперсією. Рух цієї точки у фазовому просторі описують певною випадковою функцією.

Основним математичним апаратом статистичних методів є математична статистика, метод статистичних випробувань (Монте-Карло), методи перевірки статистичних гіпотез, теорія статистичних зв'язків.

Статистичні методи використовують для вирішення завдань аналізу даних, статистичного програмування, нових розділів теорії ігор, статистичної фізики і радіотехніки, теорії масового обслуговування, теорії надійності тощо. Основною проблемою, що обмежує застосування статистичних методів, є неможливість у багатьох випадках отримання репрезентативної вибірки для подання необхідних властивостей та зв'язків.

Теоретико-множинні методи ґрунтуються на поданні систем, як певних множин, а динаміки і зв'язків – як відношень, заданих на цих множинах. Їх основним математичним апаратом є дискретна математика, алгебра логіки та математична лінгвістика.

Методи цієї групи використовують для вивчення систем, які можна подати як сукупність деяких множин із заданими на них відношеннями, зокрема в

теорії автоматів, математичній теорії систем (за М. Месаровичем), теорії баз даних та інформаційних систем тощо. Основні обмеження теоретико-множинних методів пов'язані з можливою неповнотою або суперечливістю подання досліджуваної системи.

Логічні методи передбачають подання реальної системи та її зв'язків мовою однієї з алгебр логіки. Це дає змогу переходити від одного подання системи до іншого, яке є еквівалентним першому, але простішим. Основний математичний апарат логічних методів – висловлювання, предикати, логічні функції (операції), логічні закони, квантори тощо.

Ці методи використовують для дослідження нових систем і структур, інформація про які недостатня для застосування аналітичних або статистичних методів. Прикладами є розробка автоматів, автоматичних систем контролю, задачі розпізнання образів, теорія формальних мов та ін. Логічні методи придатні тільки для подання відношень, які передбачені законами алгебри логіки, що використовується, та задовольняють вимоги логічного базису.

Лінгвістичні й семіотичні методи застосовують для формування однозначних визначень систем, їх елементів і відношень, а також розробки формалізованих мов подання і моделювання систем. Основним математичним апаратом у цьому разі є теорія знакових систем й лінгвістичні змінні (тезаурус, граматики, семантика, прагматика).

Лінгвістичні методи використовують для аналізу мов і текстів (у т. ч. мов програмування й моделювання); відображення та аналізу складних систем, інформація про які є недостатньою для застосування аналітичних і статистичних методів, а також методів формальної логіки. Прикладами є прийняття рішень в задачах з високою початковою невизначеністю, мови моделювання й автоматичного проектування. Основними обмеженнями є проблеми алгоритмічної розв'язності й логічних парадоксів, а також обґрунтування достовірності одержуваних результатів.

Прикладом логічного парадоксу є завдання військовому перукарю поглотити всіх чоловіків у полку, які не голяться самі.

Графічні методи застосовують для наочного відображення складу, структури, динаміки й властивостей складних систем. Основним апаратом графічних методів є теорія графів, геометрія, теорія сітьових структур, графіки, діаграми, гістограми тощо.

Методи цієї групи використовують при дослідженні різноманітних структур і процесів, організації взаємодії у людино-машинних та організаційних системах. Прикладами є методика оцінювання й контролю програм (PERT), сітьове планування та управління, системи організаційного управління, метод діаграм Фейнмана у фізиці елементарних частинок, частотні критерії стійкості систем управління та ін. Основними обмеженнями при застосуванні графічних методів є висока частка ручного труда та неможливість вирішення завдань прогнозування.

Згідно з В.М. Волковою та А.А. Денисовим, можна встановити відповідність методів формалізованого подання і прикладних методів моделювання систем, показану у табл. 1.1.

Таблиця 1.1

Відповідність методів формалізованого подання і прикладних методів моделювання систем

Прикладні методи моделювання	Методи формалізованого подання систем					
	Аналітичні	Статистичні	Теоретико-множинні	Логічні	Лінгвістичні й семіотичні	Графічні
Економіко-математичні методи						
Виробничі функції	+	+				
Балансові моделі	+					
Моделі обсягового планування			+			
Моделі календарного	+		+			+

планування						
Потокові (транспортні) моделі	+					+
Моделі розподілення й призначення	+					
Моделі управління запасами	+	+				
Моделі зношування й заміни обладнання	+	+				
Моделі масового обслуговування		+				
Змагальні моделі	+	+				
Методи роботи з масивами даних						
Методи організації масивів			+	+		+
Методи обробки масивів (сортування, впорядкування, розміщення)			+	+		
Методи пошуку інформації			+	+	+	

Контрольні питання

1. Що називають моделюванням?
2. Що називають моделлю?
3. Які моделі називають реальними? Наведіть приклади таких моделей.
4. Які моделі називають абстрактними? Наведіть приклади абстрактних моделей.

5. Які моделі називають математичними? Наведіть приклади таких моделей.
6. Якими є основні методи встановлення подібності між моделлю та оригіналом? Наведіть приклади їх застосування.
7. Які моделі називають пізнавальними? Наведіть приклади пізнавальних моделей.
8. З якою метою розробляють пізнавальні моделі?
9. Які моделі називають прагматичними? Наведіть приклади таких моделей.
10. Для чого використовують прагматичні моделі?
11. Які моделі називають статичними? Наведіть приклади таких моделей.
12. Які моделі називають динамічними? Наведіть приклади динамічних моделей.
13. У чому полягає властивість скінченності моделей? Чим вона зумовлена?
14. У чому полягає властивість наближеності моделей? Наведіть приклади наближень, що найчастіше використовуються при створенні математичних моделей систем.
15. У чому полягає властивість неповноти моделей? Чим визначається необхідна ступінь повноти моделі?
16. Що називають адекватністю моделі? Наведіть приклади адекватних й неадекватних моделей.
17. Що називають істинністю моделі. Наведіть приклади істинних та помилкових моделей.
18. Яким є зв'язок між істинністю й адекватністю моделі? Чи можуть неадекватні моделі бути істинними, або помилкові моделі – адекватними? Чи обов'язково адекватні моделі є істинними, а істинні – адекватними?
19. Що розуміють, говорячи, що модель є цільовим відображенням оригіналу?

20. Які фактори визначають якість моделі? Наведіть приклади.
21. Чим зумовлена необхідність вивчення моделей систем замість оригіналів?
22. Який вид моделювання називають концептуальним? Наведіть приклади концептуальних моделей систем і постановки завдань концептуального моделювання.
23. Який вид моделювання називають інтуїтивним? Наведіть приклади інтуїтивного моделювання систем.
24. Який вид моделювання називають фізичним? Наведіть приклади фізичних моделей систем і постановки завдань фізичного моделювання.
25. Який вид моделювання називають структурно-функціональним? Наведіть приклади структурно-функціональних моделей систем і постановки завдань такого моделювання.
26. Який вид моделювання називають математичним? Наведіть приклади математичних моделей систем і постановки завдань математичного моделювання.
27. Який вид моделювання називають імітаційним? Наведіть приклади імітаційних моделей систем і постановки завдань імітаційного моделювання.
28. Який тип моделі називають чорним ящиком? Які завдання вирішують за допомогою таких моделей?
29. Що називають входами та виходами системи? Для чого вони потрібні?
30. Що називають моделлю складу системи? Наведіть приклади моделей складу.
31. Що називають елементами й підсистемами? Наведіть приклади підсистем і елементів певної системи.
32. Що називають моделлю структури системи? Наведіть приклади моделей структури.
33. Що називають зв'язками системи? Які основні типи зв'язків вирізняють при дослідженні систем? Наведіть приклади зв'язків різного типу.

34. Що називають структурною схемою системи? Наведіть приклади структурних схем.

35. У чому полягає різниця між управлінням за розімкненим циклом та управлінням за замкненим циклом? Наведіть структурні схеми, що реалізують відповідні способи управління.

36. Що називають деревом цілей?

37. У чому полягають особливості аналітичних методів формалізованого подання систем? Які завдання можна вирішувати за їх допомогою? Наведіть приклади.

38. У чому полягають особливості статистичних методів формалізованого подання систем? Які завдання можна вирішувати за їх допомогою? Наведіть приклади.

39. У чому полягають особливості теоретико-множинних методів формалізованого подання систем? Які завдання можна вирішувати за їх допомогою? Наведіть приклади.

40. У чому полягають особливості логічних методів формалізованого подання систем? Які завдання можна вирішувати за їх допомогою? Наведіть приклади.

41. У чому полягають особливості лінгвістичних і семіотичних методів формалізованого подання систем? Які завдання можна вирішувати за їх допомогою? Наведіть приклади.

42. У чому полягають особливості графічних методів формалізованого подання систем? Які завдання можна вирішувати за їх допомогою? Наведіть приклади.

2. МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ СИСТЕМ

2.1. Поняття, класифікація та основні властивості математичних моделей

Якщо об'єкт дослідження є не надто складним, достатньо вивченим, а його властивості й характеристики можна виявити на основі теоретичних уявлень і наявних у літературі даних, доцільно як метод дослідження обрати побудову математичної моделі. У цьому разі процеси функціонування елементів системи подають як певні функціональні співвідношення (алгебраїчні, диференціальні, інтегро-диференціальні, скінченно-різницеві тощо) або за допомогою логічних мов.

Під **математичною моделлю** реальної системи (процесу) розуміють сукупність співвідношень (формул, рівнянь, нерівностей, логічних умов, операторів тощо), які визначають характеристики станів системи залежно від її параметрів, зовнішніх умов (вхідних сигналів, впливів), початкових умов та часу. Загалом за визначенням В.М. Глушкова математична модель – це множина символічних математичних об'єктів і співвідношень між ними. За М.М. Амосовим, математична модель – це система, що відображає іншу систему.

Математичні моделі можна досліджувати такими методами:

а) аналітичними, що дають змогу отримати у загальному вигляді явні залежності для досліджуваних характеристик;

б) чисельними, що дають можливість одержати числові значення шуканих параметрів при конкретних початкових і межових умовах;

в) якісними, за допомогою яких можна визначити певні властивості розв'язку (стійкість, монотонність, характер змінювання, асимптотика тощо) без отримання його в явному вигляді;

г) аналоговими, що дають змогу вивчати властивості досліджуваної системи за допомогою певного реального об'єкта, наприклад електричної схеми, яку можна подавати за допомогою тієї самої математичної моделі.

Для складних об'єктів альтернативним є шлях **ідентифікації** моделі, через вибір найбільш придатної моделі, що належить до певного класу математичних моделей, і визначення її параметрів за наявними емпіричними даними. Наприклад, для вивчення багатьох практично важливих властивостей твердих тіл використовують метод механічної спектроскопії. При цьому отримують так званий релаксаційний спектр, що може бути поданий у вигляді суми декількох простих максимумів, кожний з яких відповідає певному процесу, що відбувається у досліджуваному об'єкті. Для інтерпретації одержуваних спектрів необхідно використовувати їх математичні моделі. Завдання ідентифікації математичної моделі у цьому разі полягає у визначенні рівняння, що описує один максимум (тут необхідно зробити вибір з декількох можливих варіантів таких рівнянь, які відповідають різним типам процесів), визначенні кількості максимумів та визначенні коефіцієнтів відповідних рівнянь.

Існує багато різноманітних класифікацій математичних моделей. Зокрема, вирізняють моделі: статичні й динамічні, диференціальні й інтегральні, детерміновані і статистичні, лінійні та нелінійні, геометричні, топологічні, імітаційні, оптимізаційні тощо.

Статичні моделі характеризують конкретний стан об'єкта у заданий момент часу. Як правило, їх застосовують для подання стійких рівноважних чи стаціонарних станів систем, або таких нерівноважних й нестаціонарних станів, змінюванням яких з часом при дослідженні або роботі системи можна знехтувати. Прикладами статичних математичних моделей систем є відомі з курсу фізики рівняння стану ідеального газу, закон Кулона, формула Планка тощо.

Якщо ж метою моделювання є опис не одного конкретного стану системи, а різниці між станами, динаміки змінювання стану з часом, певних процесів, що відбуваються у системі, й таке інше, то виникає потреба в моделях, які б відображали ці процеси. Відповідні математичні моделі називають **динамічними**. Їх прикладами є: рівняння дифузії $\vec{j} = \text{grad } c$ (\vec{j} – потік компонента, що дифундує, c – його концентрація), рівняння руху, закон Ома у диференціальній формі $\vec{j} = \frac{1}{\rho} \mathbf{E}$ (\vec{j}

– густина току, ρ – питомий електричний опір матеріалу, E – напруженість електричного поля), закон радіоактивного розпаду ($\frac{dN}{dt} = -\lambda N$, де N – кількість радіоактивних атомів, t – час, λ – коефіцієнт пропорційності) тощо.

Загальна математична модель динаміки системи задає множини входів, виходів та станів системи, а також зв'язків між ними. У найпростішому випадку вона встановлює зв'язок між функціями $X(t)$ та $Y(t)$, де X – вектор вхідних змінних, а Y – вектор вихідних змінних системи. У разі, коли система є безінерційною і миттєво перетворює вхід до виходу, цей зв'язок можна подати функцією $y = \Phi(x)$. На практиці ця функція зазвичай невідома. Вона може бути параметризованою або непараметризованою. У першому випадку про функцію Φ відомо, що вона належить до сім'ї функцій, визначених з точністю до значень окремих параметрів. Наприклад: $\Phi_1(x) = ax + b$, $\Phi_2(x) = A \exp(Bx)$. При цьому значення всіх або частини параметрів невідомі. У другому випадку невідомим є й вигляд функції, але можуть бути відомими деякі її загальні властивості – неперервність, гладкість, монотонність, симетричність й ін.

Детерміновані (жорсткі) моделі передбачають, що зв'язок між входами та виходами системи може бути описаний за допомогою відомих детермінованих функцій, а вхідні значення є точно визначеними. Поведінку детермінованої моделі у принципі можна передбачити однозначно, але у зв'язку із складністю багатьох детермінованих моделей при їх дослідженні застосовують чисельні й статистичні методи. Зазвичай детерміновані моделі призводять до отримання систем алгебраїчних, диференціальних, інтегро-диференціальних або скінченно-різницевих рівнянь. Типовими завданнями, для вирішення яких застосовують ці моделі є дослідження динаміки систем, балансу, оптимізація систем та ін.

Стохастичні (статистичні) моделі будують, якщо параметри об'єкта дослідження піддаються впливу неконтрольованих випадкових факторів, або він характеризується власною невизначеністю поведінки. Прикладами об'єктів першого типу є курси валют, фондові індекси, нестійкі технологічні процеси,

екосистеми тощо. Типовими прикладами об'єктів другого типу є квантові частинки, а також системи, що складаються з багатьох однотипних елементів, які мають власну поведінку. В обох випадках однозначно передбачити властивості таких об'єктів, їх зміну у часі та реакцію на певні зовнішні впливи неможливо, тому при їх вивченні застосовують методи теорії ймовірностей і математичної статистики. У цьому разі зв'язок між входами та виходами подають через випадкові чи кореляційні функції або через рівняння регресії, а не відомі точно параметри і вхідні змінні характеризують за допомогою відповідних функцій розподілу або математичних сподівань і стандартних відхилень. Окремими випадками статистичних моделей є моделі факторного аналізу, що встановлюють зв'язок між факторами, які впливають на досліджувану систему і дають змогу зменшити розмірність простору ознак, а також класифікаційні моделі, за допомогою яких у сукупності досліджуваних об'єктів або ознак виділяють групи подібних один до одного елементів. Прикладами статистичних моделей є відомі з курсу фізики розподіл Максвелла, що характеризує розподіл молекул ідеального газу за швидкостями їх руху, та рівняння Шредингера, яке дає змогу визначити щільність імовірності знаходження квантової частинки у певній точці простору.

Поділ моделей на дискретні й неперервні визначається характером області допустимих значень їх параметрів. Якщо остання є неперервною, то модель називають **неперервною**. Неперервні ознаки можуть набувати будь-яких значень у певному заданому діапазоні. Прикладами таких ознак є час, потрібний учню або студенту на виконання певного завдання, довжина, вага, температура тощо. Практично усі математичні моделі, що розглядаються у курсі фізики, є прикладами неперервних математичних моделей. Якщо ж область допустимих значень параметрів дискретна, то модель називають **дискретною**. Параметри дискретних моделей можуть набувати лише окремих визначених значень. Дискретними ознаками зокрема є кількість студентів у групі, кількість комп'ютерів в певній установі, обсяг пам'яті у комп'ютерних пристроях та ін. Прикладами дискретних моделей є різноманітні розклади; моделі даних для застосування у базах даних; моделі, що використовуються при вирішенні завдань оптимізації

маршрутів перевезень, призначень виконавців робіт тощо. У фізиці дискретні математичні моделі використовуються зокрема при вивченні структури й деяких структурно чутливих властивостей кристалів.

У деяких випадках важко визначити, чи є параметри системи неперервними або дискретними. Наприклад енергію протягом тривалого часу вважали неперервною величиною. Але на початку ХХ сторіччя було встановлено, що вона може змінюватися лише на величину кратну певній мінімальній кількості (кванту) енергії. Сьогодні у теоретичній фізиці продовжуються дискусії з приводу того, чи існують такі самі мінімальна довжина й мінімальний проміжок часу.

В інших випадках виявляється зручнішим при вивченні систем певного типу застосовувати відомі математичні моделі та методи дослідження, призначені для систем іншого типу. Тому при обробці інформації й дослідженні систем часто здійснюють перетворення неперервних ознак до дискретних і навпаки. Наприклад ознаки, що вимірюють у грошових одиницях, за своєю природою є дискретними. Але у багатьох випадках мінімальна грошова одиниця є настільки малою порівняно з вимірюваними значеннями, що відповідні параметри зручніше розглядати як неперервні.

Іноді виникає необхідність побудови так званих гібридних моделей. Вони містять як неперервні, так і дискретні компоненти. Найпростішим прикладом такої моделі є модель, що описує рух кинутого м'ячика, який час від часу відскакує від землі. У цьому разі його вільний політ між двома відскоками може бути поданий за допомогою неперервної моделі, а миттєва зміна руху під час відскоку – за допомогою дискретної.

Лінійні моделі характеризуються лінійним зв'язком між вхідними та вихідними змінними. На відміну від них у **нелінійних моделях** усі або деякі зв'язки є нелінійними. Найчастіше таку класифікацію застосовують для моделей, що подаються алгебраїчними та диференціальними рівняннями. Її необхідність пов'язана з тим, що для лінійних моделей зазвичай є стандартні процедури їх аналізу, а у багатьох випадках відомими є й основні властивості таких моделей.

Тому для отримання необхідних результатів за допомогою таких моделей часто буває достатньо підставити задані числові значення до готових формул або комп'ютерних програм. Для нелінійних моделей така можливість є лише в окремих найпростіших випадках. Крім того, для нелінійних моделей значно частіше виникають проблеми аналізу й забезпечення стійкості одержуваних розв'язків.

Прикладами лінійних математичних моделей є закон Гука ($F = -k\Delta x$, де F – сила пружності, k – коефіцієнт пропорційності, Δx – видовження пружини), другий закон Ньютона ($\frac{dp}{dt} = F$, де p – імпульс матеріальної точки, F – діюча на неї сила, t – час), моделі факторного аналізу, більшість моделей, що використовуються у кореляційному аналізі даних тощо. Прикладами нелінійних моделей є закон всесвітнього тяжіння ($F = \gamma \frac{m_1 m_2}{r_{12}^2}$, де F – сила взаємодії, γ – гравітаційна стала, m_1, m_2 – маси взаємодіючих тіл, r_{12} – відстань між ними), модель нормального розподілу ($f(x, m, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - m)^2}{2\sigma^2}\right)$, де f – щільність розподілу, m – математичне сподівання досліджуваної величини, σ^2 – її дисперсія) та інші.

Одним із поширених методів дослідження нелінійних моделей є їх зведення до лінійних. Але при його застосуванні слід пам'ятати, що таке перетворення, навіть коли воно допустимо, веде до звуження меж застосовності результатів дослідження.

Одновимірними називають моделі, що мають лише одну вихідну змінну (незалежно від кількості вхідних змінних). **Багатовимірні** моделі мають декілька вихідних змінних. У загальному випадку кожна із вихідних змінних багатовимірної системи залежить від декількох вхідних. Тому багатовимірні системи зазвичай є також і багатозв'язними. Наприклад, модель, що визначає собівартість виробництва певного конкретного виробу на підприємстві, буде одновимірною. Модель яка визначає собівартості виробництва усієї номенклатури ви-

робів цього підприємства буде багатовимірною. У другому випадку вона буде і багатозв'язною, через те, що одні й ті самі ресурси можуть використовуватися при виготовленні різних виробів. При цьому існують певні обмеження на їх загальний обсяг. Деякі з виробів можуть одночасно бути як кінцевою продукцією, так і сировиною для виготовлення інших виробів. Наприклад, на металургійному комбінаті частина чавуну відвантажується замовникам, а частина йде на виплавляння сталі. Так само і сталь: частина її є кінцевою продукцією, а частина використовується для виготовлення прокату.

Властивість **стаціонарності** або **нестаціонарності** відображає ступінь змінюваності об'єкта у часі. Стаціонарними є динамічні системи, параметри яких за певних зовнішніх умов змінюються у часі настільки повільно, що цими змінами за період спостереження можна зневажати. У математичних моделях це відображається через залежність коефіцієнтів нестаціонарних моделей від часу. Зокрема рівняння руху зарядженої частинки в електричному полі можна записати у вигляді $\frac{d^2\mathbf{r}}{dt^2} = \frac{q\mathbf{E}}{m}$ або як $\frac{d^2\mathbf{r}}{dt^2} = \frac{q\mathbf{E}(t)}{m}$, де \mathbf{r} , q , m – відповідно, радіус-вектор, заряд і маса частинки, t – час, \mathbf{E} – напруженість електричного поля. При цьому перше рівняння є стаціонарною, а друге – нестаціонарною математичними моделями руху. Типовими прикладами нестаціонарних математичних моделей є моделі, що використовуються при формуванні бюджетів різного рівня, оскільки вони мають враховувати динаміку цін, курсів валют, прожиткового мінімуму тощо.

Властивість зосередженості або розподіленості кількісних характеристик внутрішніх властивостей об'єкта дослідження (температура, механічні напруження, деформації, прожитковий мінімум, швидкості процесів тощо) характеризує врахування просторової протяжності системи й скінченної швидкості процесів, що відбуваються у неї, на адекватність моделі.

Якщо просторовою протяжністю можна зневажити й вважати, що всі процеси у системі відбуваються миттєво, то для подання системи можна використовувати математичні **моделі із зосередженими параметрами**. Такі моделі

часто подають у вигляді звичайних диференціальних рівнянь та їх систем. Прикладами математичних моделей систем із зосередженими параметрами є рівняння руху матеріальної точки та абсолютно твердого тіла, закони Архімеда й Паскаля в гідростатиці тощо.

У просторово протяжних об'єктах (гази, рідини, деформівні тверді тіла, складні технічні й соціально-економічні системи) необхідно враховувати залежність характеристик від координат. З погляду математики об'єкти з розподіленими параметрами є одновимірними або багатовимірними полями відповідних величин (температур, напружень тощо), які існують у часі й просторі певної розмірності (лінійні, плоскі, тривимірні). Вихідні змінні відповідних моделей є функціями часу й просторових координат, а також їх похідних. Як математичні моделі таких об'єктів часто використовують диференціальні рівняння у частинних похідних. Прикладами **моделей з розподіленими параметрами** є закони Фіка та Фур'є в теорії дифузії й теплопровідності.

З погляду практичного використання важливою є класифікація моделей за методами їх подальшого аналізу. При цьому вирізняють моделі, які досліджують **аналітично, чисельно** та за допомогою **апаратного моделювання** (аналогових обчислювальних машин).

У першому випадку вихідну математичну модель необхідно перетворити до такої системи співвідношень, яка дає можливість одержати необхідний результат аналітичними методами. Результатом аналітичного дослідження зазвичай є: побудова формул, що задають шукані величини в явній формі; перетворення рівнянь до вигляду, для якого аналітичний розв'язок є відомим тощо. Результатами аналітичного дослідження також можуть бути якісні висновки про наявність особливих точок, асимптотичну поведінку, монотонність та однозначність залежності, стійкість розв'язку та інші. Основні аналітичні методи дослідження математичних моделей вивчалися у курсі вищої математики, зокрема у таких його розділах, як математичний аналіз, лінійна алгебра, диференціальні рівняння й теорія імовірності. Серед них можна назвати методи диференціювання, інтегрування й дослідження функцій, методи Гауса та Крамера розв'язу-

вання систем лінійних алгебраїчних рівнянь, аналітичні методи розв'язування і якісні методи аналізу звичайних диференціальних рівнянь та їх систем тощо. Розвиток комп'ютерної техніки та програмного забезпечення дає сьогодні можливість використовувати для аналітичного дослідження математичних моделей прикладні пакети символічних обчислень, такі, як Maple, Mathcad, Reduce та інші, що значно розширює можливості аналітичних методів.

На відміну від інших методів, аналітичне дослідження моделі дає можливість зробити висновки про її загальні властивості, що не залежать від конкретних значень параметрів, вхідних змінних, початкових та межових умов. Утім, для одержання аналітичного розв'язку часто виникає необхідність спрощення вихідної математичної моделі. Типовими прикладами таких спрощень є: розкладання складної функції у ряд Тейлора або Фур'є та врахування лише кількох перших членів отриманого ряду; нехтування деякими величинами, які вважають малими тощо. У більшості реальних випадків математична модель не може бути перетворена до вигляду, що дає змогу одержати аналітичний розв'язок за умови збереження її адекватності. Тому для дослідження моделі використовують інші методи.

Чисельне дослідження має ширшу сферу застосування, ніж аналітичне, оскільки обчислювальні методи аналізу можуть застосовуватися до більш широкого класу математичних моделей. Основні методи чисельного дослідження математичних моделей розглядаються в курсах "Чисельні методи", "Обчислювальна математика" та ін. До них зокрема належать: методи інтерполювання функцій; різноманітні способи уточнення коренів алгебраїчних рівнянь; ітераційні алгоритми розв'язування систем алгебраїчних рівнянь; методи Рунге-Кутта розв'язування диференціальних рівнянь та інші. До недавнього часу використання чисельних методів стримувалося обмеженістю ресурсів пам'яті та швидкодії ЕОМ. Але сучасний стан розвитку обчислювальної техніки значною мірою знімає ці обмеження. Принциповим недоліком чисельних методів є те, що вони дають змогу виконувати лише аналіз окремих випадків. Тому одержувані висновки не будуть загальними та повними.

Ще однією проблемою при використанні обчислювальних методів може бути некоректність формулювання математичної задачі, тобто відсутність у неї розв'язку, його нестійкість або наявність декількох розв'язків. Прикладами коректних задач є знаходження коренів поліному, розв'язування системи лінійних алгебраїчних рівнянь (у випадку, коли кількість рівнянь дорівнює кількості невідомих, і детермінант матриці, складеної з коефіцієнтів рівнянь, є значимо відмінним від нуля), розв'язування задачі Коші для звичайних диференціальних рівнянь тощо.

Умову стійкості можна записати так:

$$\|\delta y\| \rightarrow 0, \text{ якщо } \|\delta x\| \rightarrow 0, \quad (2.1)$$

де δx – похибка вхідних даних, δy – похибка розв'язку. **Нестійкість розв'язку** означає, що малі похибки значень вихідних параметрів можуть призводити до великих похибок одержуваних результатів. У таких випадках застосування чисельних методів недоцільно.

Можливі ситуації **слабкої стійкості (поганої зумовленості)**, коли умова (2.1) формально виконується, але похибка результату є неприпустимо великою. Прикладом є задачі, що потребують обчислення визначників матриць у випадках, коли ці визначники є близькими до нуля.

При аналоговому моделюванні для дослідження математичних моделей використовують спеціальні установки. Як правило, це електричні ланцюги, що містять активні опори, конденсатори, котушки індуктивності та джерела струму. Залежності сили току та напруги від часу в таких ланцюгах описуються різноманітними диференціальними рівняннями. Це дає можливість одержувати розв'язки відповідних рівнянь у числовому вигляді, виконуючи вимірювання відповідних електричних величин.

2.2. Похибки математичних моделей

Важливою характеристикою результатів моделювання є **похибка** одержуваних результатів. Можна виділити такі її основні складові: похибка математичної моделі, похибка вихідних даних, похибка розрахункового алгоритму та похибка обчислень.

Похибка математичної моделі пов'язана з тим, що будь-яка модель є лише наближеним відображенням об'єкта моделювання. Джерелами цієї похибки є припущення, що роблять при розробці концептуальної моделі досліджуваної системи, а також наближення, використані при побудові її математичної моделі. Зокрема, наближеність моделі може визначатися припущеннями про однорідність твердого тіла, незалежність коефіцієнта дифузії від концентрації домішки та від часу, нехтуванням впливу внутрішніх та зовнішніх електромагнітних полів на поведінку електрично заряджених домішок тощо. Якщо необхідно, похибку моделі можна зменшити шляхом урахування додаткових факторів, що впливають на вимірювані характеристики.

Похибка емпіричних даних пов'язана з: погрішностями вимірювань; використанням наближених значень параметрів досліджуваної системи; заміною генеральних сукупностей вибірками обмеженого обсягу; використанням даних, що були піддані попередній статистичній обробці, тощо. Її можна зменшити тільки шляхом проведення повторних або додаткових експериментів та спостережень.

Похибка розрахункового алгоритму пов'язана зі спрощеннями та припущеннями, що робляться при заміні вихідної математичної моделі алгоритмом обчислювань. Зокрема, більшість методів чисельного знаходження інтегралів використовує заміну площі під кривою, що інтегрується, на суму площ відносно простих фігур. При цьому похибка інтегрування істотно залежить від того, на яку кількість ділянок розбивають інтервал інтегрування та які саме фігури будуть обрані для наближення.

Багато алгоритмів передбачають застосування ітераційних процедур, які теоретично можуть виконуватися нескінченно довго, поступово наближуючись до точного розв'язку. Але у практиці час виконання ітерацій має бути обмеже-

ним. Тому необхідно заздалегідь встановити точність результату, що задовольняє дослідника, й задати умови зупинення обчислень. При цьому виникає необхідність пошуку компромісу між точністю одержуваних результатів з одного боку й потрібними для цього ресурсами (часом обчислень, пам'яттю ЕОМ) з іншого.

Похибка обчислень пов'язана з необхідністю обмеження кількості значущих цифр у числах, з якими здійснюють розрахунки. Її існування у деяких випадках веде до необхідності відмови від використання "точних" формул й застосування наближених методів. Наприклад, якщо у відомій формулі для зна-

ходження коренів квадратного рівняння $x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ виявиться, що

$b^2 \gg 4ac$, то похибка визначення величини $x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$ може стати як

завгодно великою. У цьому разі для знаходження коренів доцільніше використовувати один із наближених алгоритмів.

Похибки алгоритму та обчислень є **усувними похибками**, оскільки їх можна зменшити до необхідного рівня шляхом внесення змін до алгоритму та підвищення точності розрахунків. Похибки вихідних даних та моделі є **неусувними похибками**. Останнє слід розуміти так: ці похибки не можуть бути усунені або зменшені математичними методами, але точність моделювання може бути підвищена шляхом обрання більш адекватної моделі (цей етап є неформальним, і для його реалізації необхідно використовувати не тільки математичні методи, а й дані відповідних конкретних наук), а також підвищення точності емпіричних даних.

При побудові моделі необхідно прагнути до балансу похибок різного типу. Це зумовлено тим, що похибка моделювання визначається найбільшою з них. Підвищувати точність розрахунків на деякому етапі моделювання недоцільно, якщо вихідні дані для нього мають суттєво меншу точність або якщо результати цього етапу потім використовують для розрахунків зі значно більшою похибкою. У цьому разі прагнення до надмірної точності призведе лише до не-

виправданого збільшення часу розрахунків та потрібної пам'яті ЕОМ, але точність результату при цьому не підвищиться.

Можна рекомендувати таку послідовність балансу похибок. На першому етапі необхідно визначити ступінь точності вихідних емпіричних даних, що використовуватимуться при розрахунках, та характер впливу їх похибки на результати моделювання. Загалом похибка результатів перевищує похибку вихідних даних і є тим вищою, ніж складніша модель. Але при оцінюванні можливих наслідків неточності вихідних даних слід розрізняти систематичну та випадкову похибку моделювання. У деяких випадках неточність визначення сталих коефіцієнтів, що входять до моделі, призводить лише до появи систематичної похибки кінцевого результату. Якщо такий коефіцієнт використовують у моделі як самостійний доданок, то його похибка зумовлює збільшення або зменшення результату на одну й ту саму величину. При цьому якісна поведінка моделі та всі інші її кількісні характеристики залишаються незмінними. Якщо коефіцієнт використовують як загальний множник, то відхилення вихідних результатів будуть пропорційні похибці його визначення. При цьому якісна поведінка моделі також залишається правильною. У більш складних випадках похибки вихідних даних можуть призводити до якісних змін у поведінці моделі.

Розглянемо, наприклад, модель, задану у вигляді диференціального рівняння:

$$\frac{dy}{dx} = ax^2 + bx + c, \quad (2.2)$$

для якої емпірично встановлено такі значення коефіцієнтів: $a = 2$; $b = 4$; $c = 2$.

Для цих значень величина $\frac{dy}{dx} = 0$ у точці $x = -1$, а в інших точках вона є додатною. Відповідно, функція $y(x)$ є зростаючою і має одну критичну точку – точку перегину $x = -1$.

Істинні значення коефіцієнтів моделі зазвичай відрізняються від емпірично встановлених, що містять певну похибку. Нехай, наприклад, вони становлять $a = 1,9$; $b = 4,1$; $c = 1,9$. Тоді маємо дві критичні точки $x_1 \approx (-1,484)$ та $x_2 \approx (-0,674)$, де похідна дорівнює нулю. При $x \in (-\infty, x_1) \cup (x_2, +\infty)$ функція зростає, а при $x \in (x_1, x_2)$ спадає. Відповідно, x_1 є точкою максимуму, а x_2 – точкою мінімуму функції $y(x)$. Якщо істинні значення коефіцієнтів моделі становлять $a = 2,14$; $b = 3,9$; $c = 1,9$, то похідна є додатною для всіх x , і функція не має критичних точок.

Така різниця поведінки функції $y(x)$ може бути істотною з точки зору змістової моделі системи. Зокрема, якщо $y(x)$ – потенціальна енергія фізичної системи, то ми одержуємо три суттєво різні фізичні моделі. Залежно від істинних значень коефіцієнтів система може: 1) мати тільки точку нестійкої рівноваги; 2) мати точки як стійкої, так і нестійкої рівноваги; 3) взагалі не мати точок рівноваги. Якщо похибка вихідних даних може призводити до якісно різних результатів моделювання, подальший аналіз такої моделі є доцільним лише у випадках, коли ставиться завдання визначити можливі сценарії розвитку системи, її можливі стани тощо. Інакше необхідно підвищити точність вихідних даних, побудувати іншу модель або змінити завдання дослідження.

Якщо похибка вихідних даних дає змогу досягти мети моделювання, то наступним етапом є оцінювання та порівняння похибок проміжних і підсумкових розрахунків. За результатами порівняння робиться висновок про збалансованість похибок, необхідність доробки алгоритму чи підвищення точності розрахунків на деякому етапі, якщо точність результатів не задовольняє дослідника, або про можливість спрощення алгоритму чи зменшення точності деяких розрахунків, коли одержувана на відповідному етапі точність є надмірною.

2.3. Методи формування й математичні схеми дослідження моделей систем

Зазвичай використовують такі два способи формування математичних моделей. У першому випадку досліджувану систему поділяють на декілька підсистем, моделі яких є відомими з попереднього досвіду та літературних джерел. Формальне математичне об'єднання відповідних моделей буде математичною моделлю цілої системи. Процедура дослідження такої моделі визначається специфікою задачі. Для багатьох випадків розроблено стандартні методи розв'язування подібних задач. Найчастіше використовують **блокову структуру** моделі. Процес відновлення системи за цими блоками, як правило, здійснюють за допомогою ЕОМ.

Інший спосіб передбачає ідентифікацію системи на основі реєстрації її вхідних та вихідних сигналів, тобто потребує проведення експериментального дослідження системи.

Вихідною інформацією при побудові математичних моделей процесів функціонування систем є дані про призначення й умови роботи досліджуваної (проектованої) системи S . Вона визначає головну мету моделювання і дає змогу сформулювати вимоги до розроблюваної моделі M .

Перехід від змістового до формального подання процесу функціонування системи з урахуванням впливу зовнішнього середовища здійснюють за допомогою певної **математичної схеми**. При її використанні дослідник має у першу чергу перевірити адекватність відображення нею реальних процесів. Наприклад, подання процесу функціонування інформаційно-обчислювальної системи колективного використання у вигляді мережі систем масового обслуговування дає змогу подавати процеси, що відбуваються у системі, але при складних законах вхідних потоків і потоків обслуговування результати такого моделювання неможливо отримати в явному вигляді.

При побудові математичної моделі системи необхідно вирішити питання про її повноту, яка регулюється, у першу чергу, вибором межі "система S – зовнішнє середовище E ". Також необхідно максимально спростити модель для її дослідження.

Модель системи S можна подати як множину величин, що описують процес її функціонування і утворюють у загальному випадку такі підмножини:

- сукупність вхідних впливів на систему: $x_i \in X, i = \overline{1, n_X}$;
- сукупність впливів зовнішнього середовища: $v_l \in V, l = \overline{1, n_V}$;
- сукупність внутрішніх (власних) параметрів системи: $h_k \in H, k = \overline{1, n_H}$;
- сукупність вихідних характеристик системи: $y_j \in Y, j = \overline{1, n_Y}$.

При цьому в перелічених підмножинах можна виокремити керовані й некеровані змінні. У загальному випадку x_i, v_l, h_k, y_j є елементами підмножин, що не перетинаються і містять як детерміновані, так і стохастичні складові.

При моделюванні системи S вхідні впливи, впливи зовнішнього середовища E та внутрішні параметри системи є незалежними (екзогенними) змінними, які у векторній формі мають, відповідно, такий вигляд:

$$\mathbf{x}(t) = (x_1(t), x_2(t), \dots, x_{n_X}(t)); \quad (2.3)$$

$$\mathbf{v}(t) = (v_1(t), v_2(t), \dots, v_{n_V}(t)); \quad (2.4)$$

$$\mathbf{h}(t) = (h_1(t), h_2(t), \dots, h_{n_H}(t)). \quad (2.5)$$

Вихідні характеристики системи є залежними (ендогенними) змінними. У векторній формі вони мають вигляд:

$$\mathbf{y}(t) = (y_1(t), y_2(t), \dots, y_{n_Y}(t)). \quad (2.6)$$

Процес функціонування системи S описується у часі оператором F_S , який у загальному випадку перетворює екзогенні змінні до ендогенних у відповідності із співвідношеннями вигляду:

$$\mathbf{y}(t) = F_S(\mathbf{x}, \mathbf{v}, \mathbf{h}, t). \quad (2.7)$$

Сукупність залежностей вихідних характеристик системи від часу $y_j(t)$ для всіх $j = \overline{1, n_Y}$ називають її **вихідною траєкторією** або **вихідною характеристикою** $y(t)$. Залежність (2.7) називають **законом функціонування** системи S і позначають F_S . У загальному випадку закон функціонування системи F_S можна задати як функцію, функціонал, логічні умови, в алгоритмічній або табличній формі, а також у вигляді словесного правила відповідності.

Для опису й дослідження системи S важливим є поняття алгоритму функціонування A_S , під яким розуміють метод отримання вихідних характеристик з урахуванням вхідних впливів $x(t)$, впливів зовнішнього середовища $v(t)$ і власних параметрів системи $h(t)$. Очевидно, що один і той самий закон функціонування F_S системи S можна реалізувати різними способами, тобто за допомогою певної множини алгоритмів функціонування A_S .

Співвідношення (2.7) є математичним поданням поведінки об'єкта моделювання у часі t , тобто відображають його динамічні властивості, тому вони є його динамічною моделлю.

Статичні моделі є відображенням між двома підмножинами властивостей модельованого об'єкта Y і $\{X, V, H\}$. У векторній формі їх можна записати у вигляді:

$$y = f(x, v, h). \quad (2.8)$$

Співвідношення (2.7) і (2.8) є математичними схемами загального вигляду й дають змогу описати широкий клас систем. На практиці доцільно на початковому етапі дослідження використовувати типові математичні схеми. Вони дають змогу використовувати вже розроблені методика й алгоритми дослідження і мають переваги простоти й наочності, але можливості їх застосування обмежені.

Як детерміновані моделі для подання систем, що функціонують у неперервному часі, для випадку, коли випадковими чинниками можна зневажати, використовують алгебраїчні, диференціальні, інтегральні, інтегро-диференціальні

й інші рівняння та системи рівнянь. При поданні систем, що функціонують у дискретному часі, застосовують скінченні автомати та скінченно-різницеві схеми. Як стохастичні моделі для систем з неперервним часом використовують системи масового обслуговування і стохастичні рівняння, а для систем з дискретним часом – імовірнісні автомати.

Для подання більш складних систем, зокрема для автоматизованих систем управління, у певних випадках доцільніше обирати **агрегативні** моделі. Вони дають змогу відображати широкий круг об'єктів дослідження як систем. Це здійснюють шляхом розбиття моделі на скінченну кількість підсистем із збереженням зв'язків, що забезпечують взаємодію компонентів реальної системи.

2.4. Загальна схема побудови математичної моделі

Загальна схема побудови математичної моделі є такою. Насамперед необхідно визначити, для чого потрібна модель, тобто визначити цілі й завдання моделювання. Це має принципове значення для обрання методів її побудови й дослідження.

На другому етапі необхідно визначити структуру моделі і способи математичного подання її окремих блоків. Для цього використовують два основні підходи. Перший передбачає побудову математичної моделі на основі відомих теоретичних даних про закономірності поведінки системи або протікання процесу. У цьому разі одержувана математична модель буде системою відомих з предметної області теоретичних моделей. Перевагами такого підходу є відповідність структури моделі реальній структурі об'єкта дослідження. Завдяки цьому всі параметри моделі мають реальний фізичний (економічний, технічний або інший) зміст. Такі моделі дають змогу аналізувати не тільки загальні властивості системи як цілого, але також і поведінку її окремих елементів, зміну структури, визначити відносні вклади різних факторів у властивості, що спостерігаються, тощо. Недоліками цього підходу зазвичай є складність одержува-

них моделей і, внаслідок цього, можливість їх некоректності та накопичення похибок при розрахунку вихідних характеристик досліджуваних систем.

У другому випадку будують статистичну модель, наприклад рівняння або систему рівнянь регресії. За допомогою таких моделей можна встановлювати зв'язки між окремими змінними; вирішувати завдання класифікації; прогнозувати, як будуть змінюватися характеристики системи або процесу при зміні вхідних змінних. Статистичні, зокрема регресійні й кореляційні, моделі зазвичай відповідають розгляду досліджуваної системи як чорного ящика. Тому їх принципово неможливо використовувати для оцінювання вкладу окремих підсистем у формування загальних властивостей системи, аналізу структури системи, зв'язків між її елементами тощо. Коефіцієнти рівнянь регресії часто не мають якогось реального змісту. Проте регресійні моделі вирізняються простою і у багатьох випадках дають можливість одержувати більш точні оцінки вихідних характеристик досліджуваних систем та процесів.

В багатьох випадках при побудові моделей складних систем і процесів використовують комбінації цих підходів. Зокрема, часто базову модель будують як систему відомих теоретичних моделей, окремі компоненти якої є рівняннями регресії.

Якщо модель будують як систему відомих теоретичних моделей, то на цьому етапі на основі відомих теоретичних та емпіричних даних про об'єкт дослідження розробляють його концептуальну модель. Це передбачає визначення суттєвих для вирішення поставлених завдань компонентів системи, взаємозв'язків між ними, взаємозв'язків системи і навколишнього середовища, можливих станів системи, закономірностей поведінки системи в цілому та її окремих елементів тощо. Потім переходять від змістового до формального опису, тобто відбирають теоретичні уявлення, на основі яких будуватиметься загальна математична модель об'єкта дослідження, визначають межі застосування зроблених у них припущень і спрощень. На цьому етапі широко застосовують метод аналогій та експериментальні дослідження системи.

Далі процедура розробки моделі залежить від обрання методики її подальшого аналізу. На сьогодні найбільш поширеним методом вивчення математичних моделей є їх чисельне дослідження за допомогою ЕОМ. Для цього можна використовувати різноманітні математичні, статистичні та інші прикладні пакети програмного забезпечення, зокрема Microsoft Excel, MathCad, MathLab, Mathematica, Statistica, SPSS тощо. Але при цьому слід мати на увазі, що не існує алгоритмів чисельних розрахунків, які б давали змогу отримати задовільні розв'язки для всіх задач певного класу. Якість роботи алгоритму залежить не тільки від типу задачі, але й від її конкретних умов та параметрів. Тому обрання алгоритму є нетривіальним завданням. Як правило, якість алгоритму (й моделі взагалі) визначають: порівнянням результатів, одержуваних за різними алгоритмами розрахунків; виконанням розрахунків для тестових моделей де кількісний результат є відомим заздалегідь; порівнянням результатів моделювання з відомими якісними даними про модельовану систему (закономірностями її поведінки, впливом певних вхідних параметрів на характеристики, що розраховуються, тощо). У типових прикладних пакетах, як правило, не вказано конкретний алгоритм, за яким виконуються розрахунки. Більше того, досить часто не вказано навіть математичний метод, на якому базується цей алгоритм. Це суттєво ускладнює попередній аналіз можливості застосування прикладних пакетів до дослідження тієї чи іншої моделі, а також пошук джерел похибок моделювання.

Наступним етапом є вивчення моделі. Залежно від мети дослідження, на цьому етапі можуть встановлюватися межі адекватності моделі, області оптимальних значень параметрів, наявність прихованих зв'язків між окремими змінними тощо.

На заключному етапі здійснюють експериментальну перевірку моделі й у випадку необхідності повертаються до одного з попередніх етапів.

Після побудови математичної моделі необхідно визначити її **адекватність**. Для цього використовують формальні й неформальні процедури. У першому випадку результати моделювання порівнюють з емпіричними даними, що

відповідають тим самим значенням вхідних параметрів, межових і початкових умов. У цьому разі залишки моделі, тобто різниці між емпіричними даними й результатами, що розраховані за моделлю, мають бути нормально розподіленими випадковими величинами з близьким до нуля середнім арифметичним. Не має бути часової, просторової або будь-якої іншої кореляції між цими залишками. Крім того має бути певна відповідність масштабів між стандартними відхиленнями залишків і похибок емпіричних даних. Неформальні процедури перевірки адекватності математичних моделей зазвичай передбачають якісне порівняння типових залежностей, передбачуваних моделлю, з наявними емпіричними даними та результатами, що одержані на інших вже апробованих моделях. Вони також мають встановлювати відповідність моделі більш загальним теоретичним законам і принципам, сформульованим у певній предметній області. Інколи для перевірки адекватності треба залучати незалежних експертів, які не брали участі в розробці моделі.

За результатами перевірки адекватності моделі приймають рішення щодо можливості її використання. Результатом перевірки може бути висновок про необхідність доробки (корегування) та оптимізації моделі. При корегуванні уточнюють перелік суттєвих параметрів моделі, обмеження, функціональні зв'язки між параметрами тощо. Під оптимізацією розуміють спрощення моделі при збереженні заданого рівня її адекватності. Основними критеріями оптимальності є витрати часу, пам'яті ЕОМ та інших ресурсів при використанні моделі. Як правило, основним критерієм оптимальності є співвідношення між можливими втратами, пов'язаними з неточністю моделі, й додатковими витратами, необхідними для їх усунення.

2.5. Приклади математичних моделей

З курсу фізики, відомо, що для опису руху матеріальної точки під дією постійної сили F можна використовувати таке рівняння:

$$\mathbf{r}(t) = \mathbf{r}_0 + \mathbf{v}_0 t + \frac{\mathbf{F}t^2}{2m}, \quad (2.9)$$

де $\mathbf{r}(t)$ – радіус-вектор точки у момент часу t , \mathbf{r}_0 – радіус-вектор точки у початковий момент часу, \mathbf{v}_0 – початкова швидкість, m – маса. Це рівняння можна розглядати як найпростішу математичну модель руху. Ця модель є: детермінованою (для кожного моменту часу при відомих значеннях параметрів можна встановити значення радіус-вектора $\mathbf{r}(t)$); динамічною (вона описує зміну радіус-вектора з часом); стаціонарною (коефіцієнти рівняння не залежать від часу); неперервною (значення часу і компонентів радіус-вектора можуть набувати будь-яких значень в певних діапазонах); нелінійною за часом (залежність радіус-вектора від часу є нелінійною). Очевидно, що така модель може використовуватися для подання руху тільки тих систем, що мають зосереджені параметри.

У більш складних випадках, коли сила не є сталою, модель (2.9) може бути перетворена різними способами. Зокрема, якщо відома величина сили для кожного моменту часу, математичну модель руху можна записати у вигляді системи, що складається з диференціального рівняння другого ступеня і двох початкових умов:

$$\frac{d^2 \mathbf{r}}{dt^2} = \frac{\mathbf{F}(\mathbf{r}, t)}{m}; \quad \mathbf{r}(0) = \mathbf{r}_0; \quad \mathbf{v}(0) = \mathbf{v}_0. \quad (2.10)$$

Моделі такого виду використовують для розрахунку траєкторії руху тіл під дією відомої сили. За більшістю класифікацій ця модель належить до тих самих типів, що й модель (2.9). Але на відміну від стаціонарної моделі (2.9) вона є нестаціонарною, оскільки передбачає можливість явної залежності сили \mathbf{F} від часу, і диференціальною. Лінійність або нелінійність моделі (2.10) буде залежати від характеру залежності сили від часу. Крім того, якщо залежність $\mathbf{F}(\mathbf{r}, t)$ буде подаватися не через певну відому функцію, а через випадкову фун-

кцію, наприклад, якщо для кожного набору (\mathbf{r}, t) нам будуть відомі лише математичне сподівання і стандартне відхилення сили, то модель (2.10) буде не детермінованою, а стохастичною.

Якщо кожного моменту часу відомий радіус-вектор точки, то математичну модель руху можна подати у вигляді ступеневого ряду:

$$\mathbf{r}(t) = \mathbf{r}_0 + t \left. \frac{d\mathbf{r}}{dt} \right|_{t=0} + \frac{t^2}{2!} \left. \frac{d^2\mathbf{r}}{dt^2} \right|_{t=0} + \frac{t^3}{3!} \left. \frac{d^3\mathbf{r}}{dt^3} \right|_{t=0} + \dots + \frac{t^n}{n!} \left. \frac{d^n\mathbf{r}}{dt^n} \right|_{t=0}. \quad (2.11)$$

Такі моделі застосовують, якщо функція $\mathbf{r}(t)$ є відомою але складною. Якщо цю залежність необхідно використовувати як складову іншої, більш загальної моделі, її розклад в ряд може виявитися необхідним для забезпечення можливості дослідження цієї загальної моделі.

У багатьох випадках математичні моделі складних процесів подають у вигляді систем диференціальних рівнянь. Зокрема Г.І. Марчук запропонував таку модель взаємодії імунної системи організму людини з антигеном:

$$\begin{cases} \frac{dV}{dt} = \beta V - \gamma FV; \\ \frac{dC}{dt} = \xi Q(t - \tau) - \mu_C (C - C_0); \\ \frac{dF}{dt} = \rho C - (\mu_f + \eta \gamma V) F; \\ \frac{dM}{dt} = \sigma V - \mu_M M, \end{cases} \quad (2.12)$$

де V – кількість вірусів у організмі; C – кількість плазмових клітин, що виробляють антитіла; F – кількість антитіл; M – маса тканини, пошкодженої вірусами; β – коефіцієнт розмноження вірусів, тобто кількість нових вірусів, що створює кожен з них у середньому протягом одиниці часу; γ – коефіцієнт, що відображає ймовірність нейтралізації вірусу при його зустрічі з антигеном

($\gamma = \text{const}$); ξ – коефіцієнт, що відображає ймовірність збудження реакції створення каскаду плазмових клітин одним антигеном; $Q(t - \tau)$ – швидкість утворення плазмових клітин; τ – інкубаційний період, тобто час, потрібний для організації каскаду клітин і початку масового синтезу антитіл; μ_C – коефіцієнт, обернено пропорційний часу життя плазмових клітин; C_0^* – імунологічний рівень плазмових клітин у організмі; ρ – коефіцієнт відтворення антитіл за одиницю часу з розрахунку на одну клітину; μ_f – коефіцієнт, обернено пропорційний середньому часу розпаду антитіл; η – коефіцієнт, що відображає ймовірність розпаду антитіла при нейтралізації ним вірусом; σ – коефіцієнт, що відображає ступінь пошкодження органу одним вірусом в одиницю часу; μ_M – коефіцієнт, обернено пропорційний часу відновлення пошкодженої частини органа в e разів. Модель (2.12) є детерміновано, нелінійною (деякі доданки у правих частинах містять добутки змінних), неперервною, динамічною, стаціонарною, із зосередженими параметрами.

У лівих частинах рівнянь системи (2.12) стоять похідні певних параметрів за часом. У змістових моделях такі похідні зазвичай мають значення швидкостей зростання відповідних величин. У системі (2.12) це швидкості зростання кількості вірусів, плазмових клітин, антитіл та маси пошкодженої тканини в організмі. У правих частинах рівнянь стоять величини, що характеризують швидкості утворення або розпаду цих компонентів при різних процесах. Доданки, що беруться зі знаком "+", відповідають утворенню компонентів, а доданки зі знаком "-" – їх розпаду. Наприклад, доданки правої частини першого рівняння мають значення швидкості зростання кількості вірусів внаслідок їх ділення та швидкості її згасання через зіткнення вірусів з антитілами.

Значне місце у квантовій фізиці посідає рівняння Шредингера, що дає змогу визначити хвильову функцію елементарної частинки $\psi(x, y, z, t)$ як функцію від координат частинки і часу. Відповідна математична модель є операторним рівнянням:

$$-\frac{\hbar}{i} \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \Psi + U(x, y, z, t) \Psi, \quad (2.13)$$

де \hbar – стала Планка, Ψ – хвильова функція, t – час, x, y, z – координати частинки, m – її маса, U – потенціальна енергія, $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ – оператор Лапласа. Ця модель є: неперервною; динамічною; лінійною; нестационарною; з розподіленими параметрами. Вона також є статистичною (значення хвильової функції є детермінованими величинами, але сама вона є щільністю імовірності, тобто статистичною характеристикою досліджуваного об'єкта).

Математична модель електростатичного поля, що відображає теорему Гауса, записується як інтегральне рівняння:

$$\oint (\mathbf{E} d\mathbf{S}) = 4\pi q, \quad (2.14)$$

де \mathbf{E} – напруженість електричного поля, $d\mathbf{S}$ – елемент поверхні. Ліва частина рівняння (2.14) виражає потік вектора напруженості електричного поля через замкнену поверхню, всередині якої знаходиться електричний заряд q . Ця теорема еквівалентна закону Кулона і широко використовується при розрахунку електростатичних систем. Модель (2.14) є детермінованою, статичною, неперервною, лінійною, з розподіленими параметрами (вона передбачає, що значення напруженості електричного поля є функцією координат).

Динамічна модель міжгалузевого балансу Леонт'єва має вигляд матричного рівняння:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\dot{\mathbf{x}}(t) + \mathbf{c}(t), \quad (2.15)$$

де $\mathbf{x}(t)$ – вектор-стовпчик обсягів виробництва, $\dot{\mathbf{x}}(t)$ – вектор-стовпчик абсолютних швидкостей приросту виробництва (похідних від обсягів виробництва за

часом), A – матриця коефіцієнтів прямих матеріальних витрат, B – матриця коефіцієнтів капіталомісткості приростів виробництва (витрати виробничого накопичення на одиницю приросту відповідних видів продукції), $c(t)$ – вектор-стовпчик споживання. Ця модель є лінійною, детермінованою, стаціонарною (якщо матриці A і B не залежать від часу), із зосередженими параметрами.

Математичні моделі складних систем і процесів можна подавати у вигляді графів. Наприклад, на рис. 2.1 наведено граф, що є математичною моделлю ділянки залізниці у Запорізькій області й прилеглих районах.

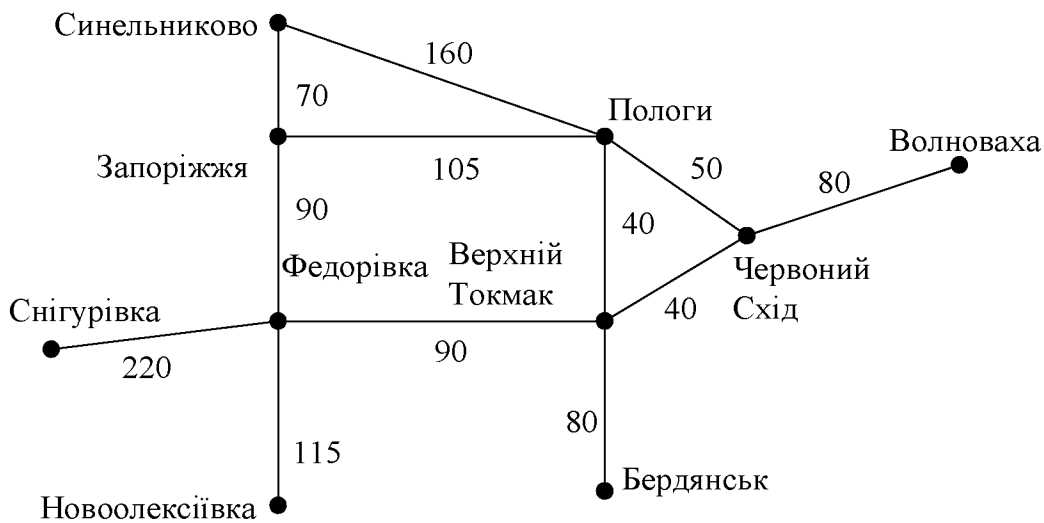


Рис. 2.1. Математична модель ділянки залізниці у Запорізькій області

Вершинами графу є вузлові станції. Його ребра характеризують їх з'єднання між собою, а також відстані між ними (наведені відстані є умовними). Ця модель є статичною й дискретною. Для наведеного графу існує еквівалентне подання у вигляді таблиці (2.1). За допомогою графів часто подають моделі певних послідовностей дій, зокрема моделі алгоритмів, розкладів та календарних планів. Математичний апарат, призначений для дослідження таких моделей, вивчався у курсах дискретної математики та теорії алгоритмів.

Таблиця 2.1

Математична модель ділянки залізниці у Запорізькій області

	Бер	ВТ	Вол	Зап	НО	Пол	Син	Сн	Фед	ЧС
Бер		80	—	—	—	—	—	—	—	—
ВТ	80		—	—	—	40	—	—	90	40
Вол	—	—		—	—	—	—	—	—	80
Зап	—	—	—		—	105	70	—	90	—
НО	—	—	—	—		—	—	—	115	—
Пол	—	40	—	105	—		160	—	—	50
Син	—	—	—	70	—	160		—	—	—
Сн	—	—	—	—	—	—	—		220	—
Фед	—	90	—	90	115	—	—	220		—
ЧС	—	40	80	—	—	50	—	—	—	

Як математичні моделі будови ідеальних кристалів використовують точкові групи симетрії. Кожна група містить певний властивий для неї набір елементів симетрії. Зокрема, точкова група O_h , що відповідає кристалам з кубічною решіткою, містить центр симетрії, 3 поворотні осі симетрії 4-го порядку, 4 поворотні осі 3-го порядку, 6 поворотних осей 2-го порядку, а також 9 площин дзеркального відображення. Під поворотними осями n -го порядку розуміють такі осі, при повороті стосовно яких на $\frac{2\pi}{n}$ градусів кристал суміщається з собою.

Така модель кристалу є статичною, дискретною (передбачає можливість знаходження атомів тільки у певних точках простору), із зосередженими параметрами (оскільки передбачає однаковість всіх характеристик кристалу у будь-якій його точці). Математичний опис неідеальних (реальних) кристалів у певних випадках роблять у вигляді моделі, що поєднує модель ідеального кристалу з певною моделлю відхилень від ідеальності. Наприклад, можна врахувати теп-

ловий рух атомів у кристалі через модель їх коливань навколо положень рівноваги, що визначаються моделлю ідеального кристалу.

2.6. Математичне подання систем із зосередженими параметрами

Розглянемо як приклад задачу про коливання пружинного маятника. З курсу фізики відома концептуальна (фізична) модель таких коливань. Згідно з нею маятник розглядають як систему, що складається з матеріальної точки маси m , підвішеної на абсолютно пружній невагомій пружині з коефіцієнтом жорсткості k . Існує певне положення рівноваги (величина розтягнення пружини), яка відповідає рівності абсолютних значень протилежно направлених сил – сили притягнення до Землі й сили пружності. При цьому значення першої сили є сталим, а значення другої залежить від величини розтягнення пружини. Якщо маятник змістити з положення рівноваги, рівнодіюча цих двох сил буде направлена так, щоб повернути його до положення рівноваги. Але коли маятник повернеться до цього положення, він матиме певну швидкість руху й не зможе зупинитися миттєво. Тому він проскакує рівноважне положення й продовжує рухатися, доки не зупиниться під впливом рівнодіючої сили, яка після проходження положення рівноваги змінює свій напрям на протилежний. Так виникають коливання. Якщо на маятник додатково діє сила тертя, направлена протилежно швидкості його руху, ці коливання будуть поступово згасати за амплітудою. Якщо ж тертя відсутнє, то виникають незгасаючі коливання. Математичною моделлю сили притягнення до Землі є формула $F_g = mg$, де m – маса маятника, g – прискорення вільного падіння у поверхні Землі. Зазвичай останнє розглядають як певну константу, значення якої ($9,81 \text{ м/с}^2$) наведено у підручниках і таблицях. Але насправді воно залежить від географічних координат тієї точки, де спостерігають коливання, а також від її висоти над поверхнею. Це треба враховувати за наявності підвищених вимог до точності вихідних даних і параметрів моделі. Математичною моделлю сили пружності є формула $F_e = -k\Delta x$ (закон Гука), де k – коефіцієнт жорсткості пружини, Δx – зміщення маятника від положення рівноваги. Коефіцієнт жорсткості, як і прискорення вільного руху,

зазвичай розглядають як табличну константу для певного матеріалу. Але за підвищених вимог до точності параметрів слід враховувати, що на неї впливають температура; флуктуації складу матеріалу, з якого виготовлено пружину; дефекти кристалічної структури цього матеріалу та інші фактори. Формула, що виражає закон Гука, є першим членом розкладу сили пружності в ряд Тейлора. Нехтування іншими доданками є обґрунтованим для багатьох металевих пружин, але може призводити до помилок при застосуванні до матеріалів, що зберігають пружні властивості при великих деформаціях. Враховуючи сказане, математичну модель коливань за відсутності тертя і за умови справедливості закону Гука можна записати у вигляді:

$$\begin{cases} m \frac{d^2(x(t) - x_e)}{dt^2} = m \frac{d^2x}{dt^2} = -kx(t) + mg = -k(x + x_e) + mg = -kx; \\ x(0) = x_0; \left. \frac{dx}{dt} \right|_{t=0} = v_0. \end{cases} \quad (2.16)$$

Перше рівняння цієї моделі є звичайним диференціальним рівнянням другого порядку зі сталими коефіцієнтами, а інші дві рівності – початковими умовами. Загальний розв’язок першого рівняння можна записати у вигляді:

$$x = A \sin(\omega t + \varphi_0), \quad (2.17)$$

де $\omega = \sqrt{k/m}$, а A , φ_0 – параметри, значення яких обирають відповідно до початкових умов. Зокрема, якщо $x_0 \neq 0$, $v_0 = 0$, то:

$$x_0 = A \sin(\varphi_0); v_0 = A \cos(\varphi_0) = 0. \quad (2.18)$$

Звідси маємо: $\varphi_0 = \pi/2$, $A = x_0$.

Як інший приклад розглянемо найпростішу модель динаміки популяцій. Нехай певна популяція у момент t_0 має біомасу x_0 . Будемо вважати, що швидкість розмноження біомаси пропорційна значенню біомаси у даний момент часу (це припущення відповідає ситуації, коли розмноження йде шляхом ділення – віруси, бактерії, найпростіші організми), а швидкість загибелі організмів пропорційна її квадрату. Відповідна математична модель має такий вигляд:

$$\frac{dx}{dt} = kx - \alpha x^2, \quad (2.19)$$

де α і k – певні сталі.

Розв’язок рівняння (2.19) можна знайти шляхом розділення змінних й інтегрування за умови $x(t_0) = x_0$. Він має такий вигляд:

$$x(t) = \frac{kx_0 e^{kt}}{k - \alpha x_0 (1 - e^{kt})}. \quad (2.20)$$

На рис. 2.2. показано залежність біомаси від часу для різних початкових умов.

Як ще один приклад розглянемо найпростішу модель епідемії у випадку, коли не здійснюються жодні карантинні заходи (природний хід епідемії).

Нехай N здорових осіб, і в момент часу $t = 0$ до цієї групи потрапляє одна хвора особа. Будемо також вважати, що інкубаційний період захворювання дорівнює нулю (хвора особа відразу стає джерелом інфекції), й не відбувається видалення захворілих (немає ані одужання, ані загибелі, ані ізоляції). Позначимо через $x(t)$ – кількість хворих осіб у момент часу t , а через $y(t)$ – кількість здорових (тобто тих, що можуть захворіти). У будь-який час виконується рівність:

$$x(t) + y(t) = N + 1, \quad (2.21)$$

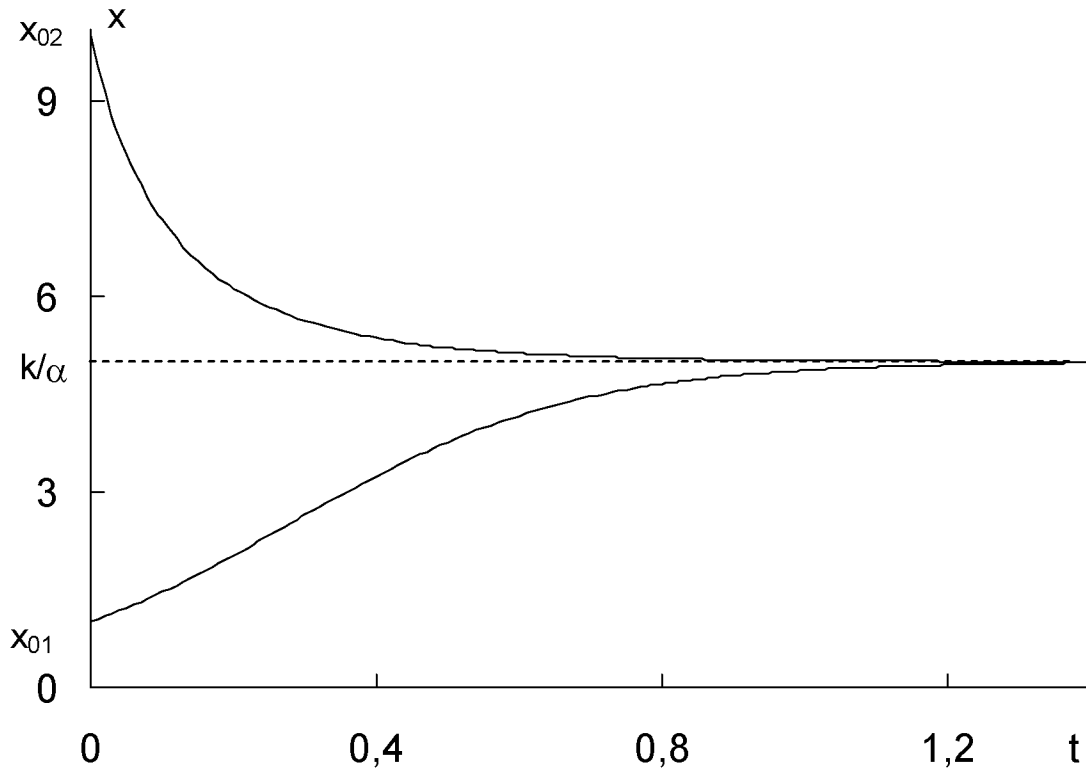


Рис. 2.2. Залежність біомаси від часу для різних початкових умов

яка відображає той факт, що загальна кількість осіб у групі, що розглядається, не змінюється з часом. При $t = 0$ виконується початкова умова $x(0) = 1$.

Розглянемо інтервал часу від t до $t + dt$. Кількість осіб, що захворіють у цей проміжок часу, можна взяти пропорційною величині dt , а також кількості зустрічей здорових і захворілих людей, тобто добутку $x(t)y(t)$, або $x(t)(N - x(t))$. Звідси швидкість зростання кількості хворих може бути описана математичною моделлю:

$$\frac{dx}{dt} = \alpha x [N + 1 - x], \quad (2.22)$$

де α – коефіцієнт пропорційності, що відображає імовірність захворювання здорової людини при її зустрічі з хворою.

Розв'язком цього рівняння при вказаній початковій умові є функція:

$$x(t) = \frac{N+1}{Ne^{-\alpha(N+1)t} + 1}, \quad (2.23)$$

графік якої показано на рис. 2.3.

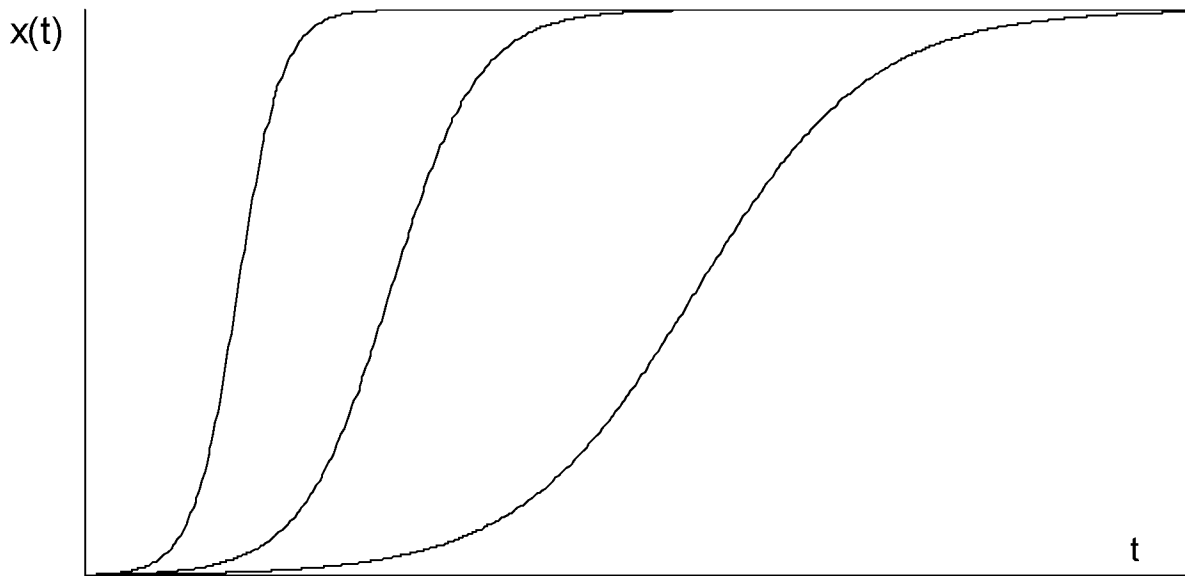


Рис. 2.3. Залежність кількості хворих від часу для різних значень α

У всіх розглянутих випадках ми мали справу з просторово однорідними системами, або просторово однорідними полями з погляду математичного подання параметрів цих систем. Зміну параметрів такого поля у часі можна подати як поведінку його окремої характерної точки.

2.7. Математичне подання систем з розподіленими параметрами

У багатьох випадках параметри поля можуть змінюватися не тільки у часі, а й у просторі. Відповідні системи називають системами з розподіленими параметрами. Математичним апаратом для подання таких систем є диференціальні рівняння у частинних похідних, які мають такий загальний вигляд:

$$A(x, y) \frac{\partial^2 u}{\partial x^2} + 2B(x, y) \frac{\partial^2 u}{\partial x \partial y} + C(x, y) \frac{\partial^2 u}{\partial y^2} = F(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}); \quad (2.24)$$

$$(A^2 + B^2 + C^2 \neq 0).$$

Залежно від знаку дискримінанта $D = B^2 - AC$ отримують один з таких видів рівнянь:

- еліптичні ($D < 0$);
- параболічні ($D = 0$);
- гіперболічні ($D > 0$).

Еліптичні рівняння, або рівняння Пуассона мають такий загальний вигляд:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = F(x, y). \quad (2.25)$$

При $F(x, y) = 0$ їх називають також **рівняннями Лапласа**.

Прикладами **параболічних рівнянь** є **рівняння теплопровідності та дифузії**. Вони мають такий загальний вигляд:

$$\frac{\partial u}{\partial t} - \alpha^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) = F(x, y, z, t). \quad (2.26)$$

Рівняння гіперболічного типу (хвильове рівняння) для одновимірного випадку записується у вигляді:

$$\frac{\partial^2 u}{\partial t^2} - \alpha^2 \frac{\partial^2 u}{\partial x^2} = 0. \quad (2.27)$$

Застосовність рівняння того або іншого типу для подання об'єкта дослідження визначається властивостями відповідного об'єкта, а також характером

зовнішніх впливів (межевими умовами). У деяких випадках відповідне рівняння для подання системи можна отримати, використовуючи відомі з предметної області закони, що описують поведінку об'єкта дослідження.

Розглянемо процес теплопередачі вздовж тонкого стрижня, що нагрівається з лівого боку. При цьому будемо нехтувати втратами теплоти. Візьмемо ділянку стрижня, обмежену точками x_i та x_{i+1} . Згідно із законом збереження енергії, зміна кількості теплоти $q(t_{j+1}) - q(t_j)$ у досліджуваній ділянці за проміжок часу від t_j до t_{j+1} визначається різницею теплових потоків через її бокові межі $w_{i+1} - w_i$. Таким чином рівняння балансу теплоти для ділянки, що розглядається, у j -й момент часу можна записати у вигляді:

$$\left[q(t_{j+1}) - q(t_j) \right] (x_{i+1} - x_i) = -(w_{i+1} - w_i) (t_{j+1} - t_j). \quad (2.28)$$

Звідси отримуємо різницеве рівняння:

$$\frac{q_i(t_{j+1}) - q_i(t_j)}{t_{j+1} - t_j} = - \frac{w_{i+1} - w_i}{x_{i+1} - x_i}, \quad (2.29)$$

Переходячи до границі при $(t_{j+1} - t_j) \rightarrow 0$, $(x_{i+1} - x_i) \rightarrow 0$, отримуємо відповідну модель у вигляді диференціального рівняння:

$$\frac{\partial q}{\partial t} = \frac{\partial w}{\partial x}. \quad (2.30)$$

Враховуючи, що $q = \rho c_p T$, $w = \lambda(dT/dx)$, де ρ , c_p , λ – густина, теплоємність й коефіцієнт теплопровідності, одержуємо рівняння теплопровідності у вигляді:

$$\frac{\partial}{\partial t} (\rho c_p T) = \frac{\partial}{\partial x} \left(\lambda \frac{\partial T}{\partial x} \right), \quad (2.31)$$

яке за умови сталості теплофізичних параметрів ρ , c_p та λ набуває вигляду, аналогічного (2.26) в одновимірному випадку:

$$\frac{\partial T}{\partial t} = \frac{\lambda}{\rho c_p} \cdot \frac{\partial^2 T}{\partial x^2}, \quad (2.32)$$

відрізняючись від нього лише відсутністю члена $F(x, t)$, який враховує межові умови.

2.8. Комп'ютерна реалізація математичних моделей систем

Сьогодні для дослідження математичних моделей складних систем найчастіше використовують різноманітні технології комп'ютерного моделювання. Поняття комп'ютерного моделювання не є тотожним поняттю математичного моделювання систем. З одного боку, важливою складовою математичного моделювання є дослідження загальних властивостей моделей, для чого необхідно використовувати аналітичні методи дослідження. З іншого боку комп'ютерне моделювання дає змогу досліджувати не тільки математичні моделі складних систем, а й моделі деяких інших типів, зокрема імітаційні.

Під комп'ютерною моделлю зазвичай розуміють:

а) умовний образ певної системи об'єктів або процесів, поданий за допомогою взаємозалежних комп'ютерних таблиць, блок-схем, діаграм, графіків, рисунків, анімаційних фрагментів, гіпертекстів тощо, який відображає структуру й взаємозв'язки між компонентами об'єкта дослідження; такі моделі називають **структурно-функціональними**;

б) окрему комп'ютерну програму, сукупність програм або програмний комплекс, що дає змогу за допомогою певної послідовності операцій і графічного відображення їх результатів відтворювати (імітувати) процеси функціонування досліджуваного об'єкта, як правило, за наявності впливу випадкових факторів; такі моделі називають **імітаційними**.

Основні підходи до структурно-функціонального моделювання спочатку було розроблено у теорії електричних ланцюгів, електроніки й радіотехніки, де вперше почали застосовувати блок-схеми. Подальшого розвитку вони дістали у теорії автоматичного управління. Вона розробила достатньо загальну методологію аналізу й синтезу структурних схем, яка ґрунтується на тому, що кожній математичній операції над сигналами приводиться у відповідність певний елементарний структурний блок.

Розроблений апарат має широке застосування в аналізі неперервних лінійних динамічних систем, що можуть бути подані за допомогою диференціальних рівнянь. Але він не придатний для опису процесів у соціально-економічних та організаційних системах, де зв'язки між окремими блоками є більш складними й зазвичай не можуть бути зведені до певної функції часу. Він також є незручним у деяких інших випадках, коли поняття елементарного блока є істотно іншим, ніж в теорії автоматичного управління. Зокрема у деяких випадках ці блоки мають відповідати основним операціям машинної обробки даних – нагромадженню, сортуванню, передачі тощо. На сьогодні при розробці блок-схем використовують стандарти **Єдиної системи програмної документації** (ЄСПД).

Подальшого розвитку методологія побудови й аналізу блок-схем дістала у теорії **автоматизованих систем управління виробництвом** (АСУВ). При цьому функції, виконувані блок-схемами, дещо змінилися порівняно із алгоритмами і програмами. Головним призначенням графічних символів при проектуванні АСУВ є саме моделювання об'єкта автоматизації й процесів функціонування самої системи. Основними символами є функціональні блоки, що відображають основні функції збору, нагромадження, передачі й обробки даних. Поряд з ними до умовних графічних позначень включено символи, що дають змогу подавати різноманітні структури об'єктів управління. Розроблено спеціальні стандарти, які регламентують склад, розміри й вигляд символів, а також правила їх застосування. У цілому сукупність символів для АСУВ і правил їх використання утворює найпростішу мову структурно-функціонального моде-

лювання, що застосовується при системному аналізі й проектуванні автоматизованих систем.

Сучасні методи структурно-функціонального аналізу й моделювання складних систем ґрунтуються на працях професора Массачусетського технологічного інституту Дугласа Росса, який вперше використав поняття структурного аналізу при створенні алгоритмічної мови модульного програмування АРТ. Подальший розвиток ідеї подання складних об'єктів як ієрархічних багаторівневих модульних систем призвів до створення технології структурного аналізу й проектування **SADT (Structured Analyses and Design Technique)**. Сьогодні вона успішно застосовується при вирішенні таких завдань, як вдосконалення управління фінансами й матеріально технічним забезпеченням крупних організацій; створення програмного забезпечення АСУ обчислювальними та телекомунікаційними мережами; довгострокове й стратегічне планування діяльності установ тощо.

Відмінною рисою методології SADT є принцип побудови моделі зверху вниз. На першому етапі побудову ієрархії моделей починають з розробки досить грубої (ескізної) моделі. Далі цю модель уточнюють, розробляючи більш докладні моделі окремих SADT-блоків вищого рівня ієрархії. При цьому інші блоки моделі можуть залишатися незмінними.

Ще однією відмінністю технології SADT є можливість одночасно із структуруванням проблеми розробляти структуру бази даних з використанням мови SQL. Виключно важливою особливістю технології SADT є можливість зведення моделі до **кольорової сітки Петрі**.

Таким чином, застосування технології SADT й відповідного програмного забезпечення дає змогу уніфікувати окремі блоки моделі складної системи, розпаралелити процес створення моделі та об'єднати побудовані модулі до єдиної ієрархічної динамічної моделі.

Одним з поширених засобів структурно-функціонального аналізу є пакет "**CASE-Аналітик**". Він призначений для автоматизації проектування і впровадження широкого класу систем обробки інформації й управління: інформацій-

но-обчислювальних мереж, організаційно-управлінських АСУ всіх рівнів, банківських та бухгалтерських систем, систем автоматизації експерименту й діловодства тощо.

В основі цього пакету лежать засоби побудови строгої та наочної структурно-функціональної моделі системи, яка є ієрархією **діаграм потоків інформації** та **функціональних зв'язків**, що автоматично відображаються у базі даних. Пакет дає змогу будувати й редагувати потокові діаграми, здійснювати пошук за діаграмами та даними, експортувати й імпортувати дані з інших пакетів, а також оформлювати проектну документацію відповідно до вимог стандартів.

Подібні програмні засоби зазвичай крім засобів структурно-логічного моделювання містять засоби моделювання й проектування баз даних, вартісного аналізу, оцінювання ризиків, контролю й управління реалізацією проектів.

Контрольні питання

1. Що називають математичною моделлю системи?
2. Назвіть та охарактеризуйте основні методи дослідження математичних моделей.
3. Які моделі називають статичними? Наведіть приклади статичних математичних моделей.
4. Які моделі називають динамічними? Наведіть приклади динамічних математичних моделей.
5. Які моделі називають детермінованими? Наведіть приклади детермінованих математичних моделей.
6. Який математичний апарат найчастіше застосовують для побудови детермінованих математичних моделей?
7. Які моделі називають стохастичними? Наведіть приклади стохастичних математичних моделей.
8. Який математичний апарат найчастіше застосовують для побудови стохастичних математичних моделей?

9. Які моделі називають неперервними? Наведіть приклади неперервних математичних моделей.

10. Які моделі називають дискретними? Наведіть приклади дискретних математичних моделей.

11. Які моделі називають лінійними? Наведіть приклади лінійних математичних моделей.

12. Які моделі називають нелінійними? Наведіть приклади нелінійних математичних моделей.

13. Які моделі називають одновимірними? Наведіть приклади одновимірних математичних моделей.

14. Які моделі називають стаціонарними? Наведіть приклади стаціонарних математичних моделей.

15. Які моделі називають нестаціонарними? Наведіть приклади нестаціонарних математичних моделей.

16. Якими є основні завдання аналітичного дослідження математичних моделей?

17. Якими є основні методи аналітичного дослідження математичних моделей?

18. Яке програмне забезпечення можна використовувати для аналітичного дослідження математичних моделей?

19. Якими є типові проблеми, що обмежують область застосовності аналітичних методів дослідження математичних моделей складних систем?

20. Які методи використовують для спрощення математичних моделей? Наведіть приклади.

21. Які завдання вирішують за допомогою чисельних методів аналізу математичних моделей?

22. Якими є основні проблеми, що виникають при застосуванні чисельних методів дослідження математичних моделей?

23. Якими є основні методи чисельного дослідження математичних моделей?

24. Які математичні задачі називають некоректними? Наведіть приклади некоректних математичних задач.

25. Що називають стійкістю розв'язку? Чому розв'язки задач, одержуваних при чисельному дослідженні математичних моделей мають бути стійкими?

26. Що називають аналоговим дослідженням математичних моделей? Як можна реалізувати таке дослідження на практиці?

27. Якими є основні складові похибки моделювання?

28. Що називають похибкою математичної моделі, якими є її основні джерела, як можна зменшити цю похибку?

29. Що називають похибкою емпіричних даних, якими є її основні джерела, як можна зменшити цю похибку?

30. Що називають похибкою розрахункового алгоритму, якими є її основні джерела, як можна зменшити цю похибку?

31. Що називають похибкою обчислювань, якими є її основні джерела, як можна зменшити цю похибку?

32. Які складові похибки моделювання є усувними, і які з них є неусувними?

33. Чому при побудові моделі прагнуть до балансу похибок різного типу?

34. Які способи найчастіше використовують для подання математичних моделей систем?

35. На які основні підмножини поділяють сукупність величин, що описує функціонування системи?

36. Наведіть приклади типових математичних схем, які використовують для подання систем різного типу.

37. Якими є основні етапи побудови математичної моделі системи?

38. Якими є основні підходи до визначення структури моделі і способів математичного подання її окремих блоків?

39. Як оцінюють якість алгоритму чисельного дослідження математичної моделі?

3. ПРИКЛАДИ ПОБУДОВИ Й ДОСЛІДЖЕННЯ МАТЕМАТИЧНИХ МОДЕЛЕЙ СИСТЕМ

3.1. Модель динаміки складу земної атмосфери

Сучасна атмосфера Землі за своїм складом істотно відрізняється від атмосфер найближчих до нас планет Сонячної системи – Венери та Марса, де основним компонентом є вуглекислий газ, концентрація якого дорівнює близько 95%. Кількість вільного кисню в атмосфері Марса всього 0,2%, а в атмосфері Венери вона ще менша. У земній атмосфері вміст вуглекислого газу дорівнює 0,032%, а кисню – 21%. Між тим вулканічна діяльність на Землі, також як на Марсі і Венері, призводить до утворення величезної кількості вуглекислоти. Різниця, що спостерігається, пов'язана з наявністю життя на нашій планеті. Її важливою причиною є життєдіяльність земних рослин і тварин. У зелених рослинах зосереджена основна маса живої речовини. За допомогою сонячного випромінювання вони синтезують глюкозу та інші вуглеводи із молекул CO_2 та H_2O . Молекули кисню, що вивільняються в такому процесі, неперервно потрапляють до атмосфери Землі. Кількість кисню, що виділяється земними рослинами дорівнює приблизно 2,5 млрд. т/рік. Водночас вони поглинають близько 3,5 млрд. т/рік вуглекислоти та 1,4 млрд. т/рік води.

Розвинене життя є потужним фактором, що перетворює атмосферу та кору Землі. Це дало В.І. Вернадському підґрунтя для введення поняття біосфери – оболонки земної поверхні, що містить живі організми у будь-якій концентрації. Вона охоплює верхній (приблизно 3 км) шар кори, усю гідросферу та нижні шари атмосфери. За сучасними оцінками повна маса живої речовини на Землі дорівнює приблизно 4220 млрд. т. З них близько 20 млрд. т припадає на тварин і мікроорганізми, а інше – на зелені рослини. Поверхневі шари земної кори перетворені життєдіяльністю різних організмів приблизно на 99%. Практично весь кисень земної атмосфери є результатом фотосинтезу рослин і за час існування біосфери він оновився, пройшовши біогенний цикл, не менш, ніж 1 млн. разів. Якщо б цей процес припинився, то за кілька тисяч років кисень повністю

б зник з атмосфери. Бактерії спроможні розкласти найміцніші хімічні сполуки, що входять до складу алюмосилікатів та гранітів, котрі є основною складовою земної кори. Одним із результатів їх життєдіяльності є утворення ґрунту.

Атмосфера Землі містить близько 10^{15} т вільного кисню. Ще порядку 10^{19} т кисню зв'язані у земній корі. Кількість вуглецю в атмосфері дорівнює $(2-3) \cdot 10^{12}$ т, в океані – близько 10^{14} т, у земній корі – $10^{16} - 10^{17}$ т. У літосфері вуглець розподілений нерівномірно: у верхніх шарах його вміст у 5–10 разів вищий, ніж у нижніх. Кількість азоту в атмосфері та літосфері дорівнює приблизно по $4 \cdot 10^{15}$ т.

Вивільнення та зв'язування кисню, вуглецю та азоту на поверхні Землі відбувається при утворенні та розкладанні органічних речовин, життєдіяльності живих організмів, ерозії, промислової діяльності людей, дисипації у міжпланетний простір.

Розглянемо модель кругообігу цих елементів, запропоновану В.О. Костициним в 1935 р. Вона не враховує промислову діяльність людства, для якої на час її побудови не існувало адекватних методів математичного опису та потрібної інформації. На думку В.О. Костицина, врахування цієї діяльності могло б призвести лише до зміни значень певних числових коефіцієнтів і початкових умов. Самі ж рівняння та їх розв'язки якісно б не змінилися. Крім того, згідно з сучасними даними, кількість вуглецю й азоту, що потрапляє до атмосфери внаслідок життєдіяльності мікроорганізмів, може істотно перевищувати обсяги промислових викидів цих елементів.

Загальна маса вуглекислоти в океані приблизно у 30–40 разів вища, ніж в атмосфері. У моделі приймається гіпотеза Т. Шлезинга, згідно з якою океан розглядається як регулятор вмісту вуглекислоти в атмосфері: він вивільняє вуглекислоту, коли її парціальний тиск в атмосфері зменшується, і поглинає її, якщо цей тиск зростає. Таке припущення відповідає термодинамічній рівновазі між океаном та атмосферою стосовно розчинення вуглекислоти.

Спочатку розглянемо кругообіг кисню та вуглецю. Уведемо такі позначення:

- x – маса вільного атмосферного кисню;
- y – загальна маса вуглекислоти в атмосфері та океані;
- v – загальна маса кисню та вуглецю у рослинах;
- u – загальна маса кисню та вуглецю у тваринах;
- s – загальна маса кисню та вуглецю в останках тварин та рослин, що розсіяні в земній корі.

Концептуальну модель кругообігу кисню та вуглецю можна зобразити схемою, яку подано на рис. 3.1.

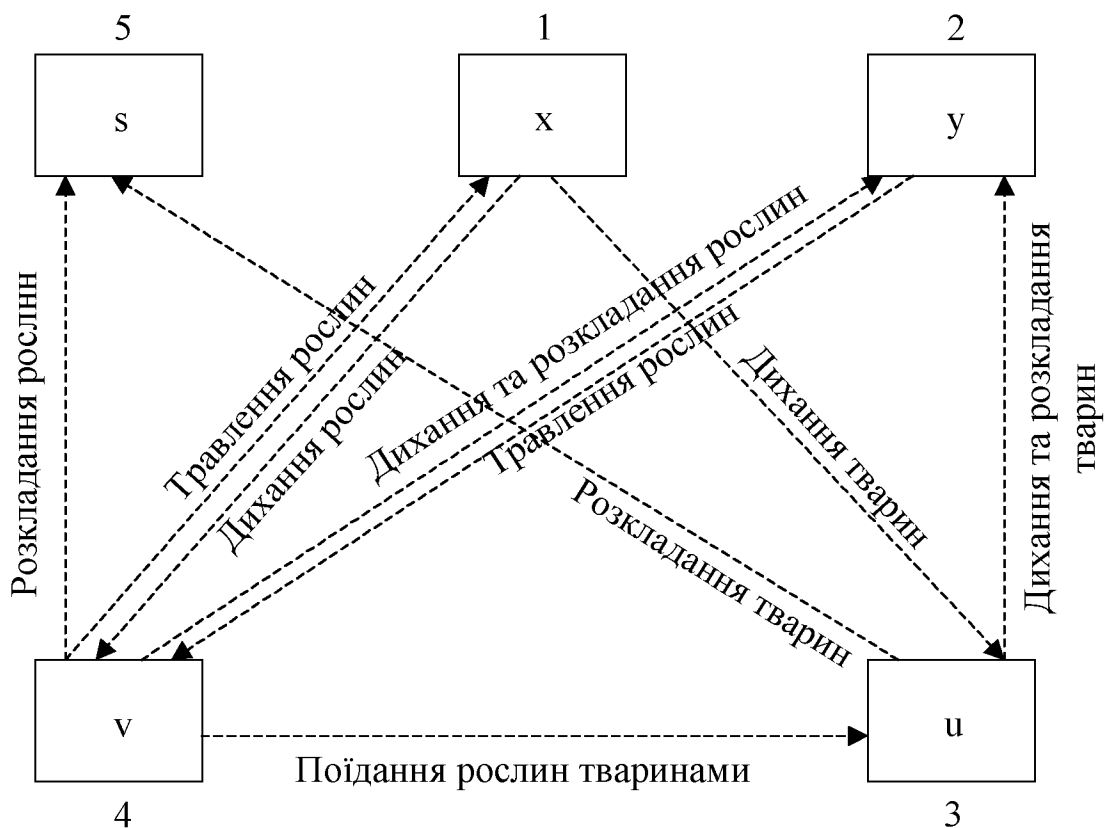


Рис. 3.1. Концептуальна модель кругообігу кисню та вуглецю

Їй відповідає така математична модель:

$$\begin{cases} \frac{dx}{dt} = -\alpha_{13}u - \alpha_{14}v + \alpha_{41}v; \\ \frac{dy}{dt} = \alpha_{32}u - \alpha_{24}v + \alpha_{42}v; \\ \frac{du}{dt} = \alpha_{13}u - \alpha_{32}u - \alpha_{35}u + \beta uv; \\ \frac{dv}{dt} = \alpha_{14}v - \alpha_{41}v + \alpha_{24}v - \alpha_{42}v - \alpha_{45}v - \beta uv; \\ \frac{ds}{dt} = \alpha_{35}u + \alpha_{45}v; \\ x + y + u + v + s = M. \end{cases} \quad (3.1)$$

Ця модель є динамічною, нелінійною, неперервною, детермінованою, із зосередженими параметрами. Вона стаціонарна, якщо коефіцієнти α_{ij} та β є сталими, і нестаціонарна, якщо хоча б один з них змінюється з часом.

Розглянемо систему (3.1) більш детально.

Згідно з першим рівнянням, атмосферний кисень витрачається на дихання тварин та рослин і вивільняється у процесі травлення рослин. Припускається, що швидкості цих процесів, тобто кількість кисню, який витрачається або вивільняється у них в одиницю часу, є пропорційними до загальних мас тварин і рослин, відповідно, які, у свою чергу, є пропорційними до загальних мас кисню та вуглецю, що містяться у них. Таке припущення відповідає розкладу невідомих залежностей швидкостей процесів у ряд Тейлора із врахуванням лише перших ненульових членів (із змістових уявлень можна зробити висновок, що за нульової біомаси рослин або тварин процеси їх дихання не відбуваються, тому вільні члени відповідних рядів Тейлора дорівнюють нулю). Також приймається, що швидкості вказаних процесів не залежать від вмісту кисню в атмосфері, що допустимо лише при розгляді відносно малих (у геологічному вимірі) інтервалів часу.

Друге рівняння відображає отримання атмосферою вуглецю, який вивільняється рослинами і тваринами у процесі дихання, а також при розкладанні живої матерії. Воно також враховує поглинання вуглекислоти рослинами при їх

травленні. Це рівняння ґрунтується на тих самих припущеннях, що й перше. Крім того, воно не враховує такі джерела вуглекислого газу, як вулканічна діяльність, промисловість, ерозія тощо.

Третє рівняння відображає засвоєння та віддачу газів тваринами внаслідок їх дихання та розкладання, збагачення ґрунту киснем та вуглецем при розкладанні трупів тварин, а також травлення тварин рослинами. Швидкість останнього процесу приймається пропорційною добутку біомас тварин та рослин.

Четверте та п'яте рівняння відображають швидкості потрапляння вуглецю та кисню до рослин і земної кори. Вони враховують ті самі процеси, що вже були розглянуті, оскільки баланс кисню та вуглецю у системі має зберігатися. Збільшення їх обсягу в однієї підсистемі внаслідок певного процесу має призводити до його зменшення в деякої іншій підсистемі завдяки тому самому процесу. Останнє рівняння фактично є формою запису закону збереження маси.

Якщо вважати океан певним стабілізатором кількості вуглекислоти в атмосфері, то можна вважати, що коефіцієнти $\alpha_{32}, \alpha_{24}, \alpha_{42}$ є менш мінливими порівняно із $\alpha_{13}, \alpha_{14}, \alpha_{41}$, тому у певному наближенні їх можна розглядати як сталі. Враховуючи велику масу атмосферного кисню, можна також припустити, що вона змінюється відносно повільно, і при розгляді достатньо малих проміжків часу коефіцієнти $\alpha_{13}, \alpha_{14}, \alpha_{41}$ також можна вважати сталими. Гіпотеза про сталість коефіцієнтів системи (3.1) суттєво спрощує її розв'язання і не впливає істотно на якісні результати. Але слід зазначити, що за зроблених припущень ми можемо отримати лише якісні оцінки.

Геологічні дані дають підстави вважати, що $s(t)$ є зростаючою функцією, а із закону збереження маси випливає:

$$u + v < M. \tag{3.2}$$

Праві частини диференціальних рівнянь моделі залежать тільки від змінних u та v . Тому спочатку можна розв'язати третє та четверте рівняння, а потім виразити з них інші змінні. Запишемо ці рівняння у простішому вигляді:

$$\begin{cases} \frac{du}{dt} = u(-\lambda + \beta v); \\ \frac{dv}{dt} = v(\mu - \beta u), \end{cases} \quad (3.3)$$

де $\lambda = \alpha_{32} + \alpha_{35} - \alpha_{31}$, $\mu = \alpha_{14} - \alpha_{41} + \alpha_{24} - \alpha_{42} - \alpha_{45}$.

Оскільки за відсутності рослин тварини існувати не можуть, маємо зробити висновок, що $\lambda > 0$. Відсутність тварин мала б призвести до збільшення біомаси рослин, звідси отримуємо, що $\mu > 0$. З (3.3) можна одержати:

$$\begin{aligned} u^\mu v^\lambda &= H \exp[\beta(u + v)]; \\ H &= u_0^\mu v_0^\lambda \exp[-\beta(u_0 + v_0)]. \end{aligned} \quad (3.4)$$

Система (3.3) є системою Ляпунова. Із загальної теорії диференціальних рівнянь відомо, що вона має періодичні розв'язки, яким відповідають замкнені криві на фазовій траєкторії, побудованої в координатах (v, u) .

Уведемо позначення: ω – період для $v(t)$ й $u(t)$, $u_c = \mu/\beta$, $v_c = \lambda/\beta$. Точка (u_c, v_c) є загальним центром S усіх кривих, які відповідають (3.4). Величина ω залежить як від коефіцієнтів λ, μ, β , так і від початкових умов, тобто від значень u_0, v_0 . Зокрема $\omega = 0$ при $u_0 = u_c, v_0 = v_c$, і $\omega = \infty$ при $H = 0$.

Систему (3.3) можна записати у такій спосіб:

$$\begin{cases} \frac{d}{dt}(\ln u) = -\lambda + \beta v; \\ \frac{d}{dt}(\ln v) = \mu - \beta u. \end{cases} \quad (3.5)$$

Розв'язуючи її отримаємо:

$$u(t + \omega) - u(t) = -\lambda \int_0^{\omega} u \, dt + \beta \int_0^{\omega} uv \, dt = 0;$$

$$v(t + \omega) - v(t) = \mu \int_0^{\omega} v \, dt - \beta \int_0^{\omega} uv \, dt = 0,$$

або

$$\int_0^{\omega} uv \, dt = v_c \int_0^{\omega} u \, dt = u_c \int_0^{\omega} v \, dt.$$

Далі:

$$\ln \frac{u(t + \omega)}{u(t)} = -\lambda \omega + \beta \int_0^{\omega} v \, dt = 0;$$

$$\ln \frac{v(t + \omega)}{v(t)} = \mu \omega - \beta \int_0^{\omega} u \, dt = 0,$$

звідки:

$$u_c = \frac{1}{\omega} \int_0^{\omega} u \, dt; \quad v_c = \frac{1}{\omega} \int_0^{\omega} v \, dt; \quad u_c v_c = \frac{1}{\omega} \int_0^{\omega} uv \, dt. \quad (3.6)$$

З інших рівнянь системи (3.1) отримуємо:

$$\begin{aligned} \frac{x(t + \omega) - x(t)}{\omega} &= -\alpha_{13} u_c - \alpha_{14} v_c + \alpha_{41} v_c; \\ \frac{y(t + \omega) - y(t)}{\omega} &= \alpha_{32} u_c - \alpha_{24} v_c + \alpha_{42} v_c; \\ \frac{s(t + \omega) - s(t)}{\omega} &= \alpha_{35} u_c + \alpha_{45} v_c, \end{aligned} \quad (3.7)$$

або

$$\frac{x(t+\omega) - x(t)}{\omega} = \frac{1}{\beta} [(\alpha_{32} + \alpha_{35})(\alpha_{41} - \alpha_{14}) + \alpha_{13}(\alpha_{45} - \alpha_{24} + \alpha_{42})];$$

$$\frac{y(t+\omega) - y(t)}{\omega} = \frac{1}{\beta} [-(\alpha_{24} - \alpha_{42})(\alpha_{35} - \alpha_{13}) - \alpha_{32}(\alpha_{45} + \alpha_{41} - \alpha_{14})].$$

(3.8)

Зробимо такі припущення:

– рослини виділяють більшу кількість кисню, ніж засвоюють, тобто $\alpha_{14} < \alpha_{41}$;

– рослини засвоюють більшу кількість вуглекислого газу, ніж виділяють, тобто $\alpha_{42} < \alpha_{24}$;

– рослини втрачають більшу кількість речовини внаслідок її розсіювання у земній корі, ніж вони отримують внаслідок засвоєння вуглекислоти, тобто $\alpha_{45} + \alpha_{42} > \alpha_{24}$;

– тварини розсіюють більшу кількість речовини, ніж вони отримують за рахунок засвоєння кисню, тобто $\alpha_{35} > \alpha_{13}$.

За таких припущень, якщо абстрагуватися від періодичних коливань, з (3.8) випливає, що вільний кисень неперервно накопичується в атмосфері, а вуглекислий газ – виводиться з неї. З іншого боку:

$$s(t + \omega) - s(t) = \text{const} > 0.$$

Це означає, що втрачені атмосферою гази накопичуються у земній корі.

Зробимо інші припущення:

– рослини виділяють більшу кількість кисню, ніж засвоюють, тобто $\alpha_{14} < \alpha_{41}$;

– рослини засвоюють меншу кількість вуглекислого газу, ніж виділяють, тобто $\alpha_{42} > \alpha_{24}$;

– тварини розсіюють меншу кількість речовини, ніж вони отримують за рахунок засвоєння кисню, тобто $\alpha_{35} < \alpha_{13}$.

У цьому разі, якщо зневажити періодичними коливаннями, з (3.8) випливає, що з часом запас вуглекислого газу в атмосфері буде зростати, а запас вільного кисню – зменшуватися.

За припущень, що:

- $\alpha_{14} > \alpha_{41}$;
- $\alpha_{24} - \alpha_{42} > \alpha_{45} > \alpha_{14} - \alpha_{41}$;
- $\alpha_{35} > \alpha_{13}$,

одержимо висновок, що запас обох газів у атмосфері з часом буде зменшуватися. Інші можливі варіанти співвідношень між коефіцієнтами ми розглядати не будемо, оскільки вони не відповідають відомим емпіричним даним.

Стосовно азоту органічний світ можна поділити на дві частини: організми, які спроможні зв'язувати атмосферний азот безпосередньо, та організми, що не можуть забезпечувати себе азотом без допомоги організмів попередньої групи. Припустимо існування певного симбіозу між організмами першої групи та рослинами й позначимо як z загальну кількість атмосферного азоту. Через x та y , як і u у попередньому випадку, будемо позначати загальні кількості вільного атмосферного кисню і вуглекислого газу в атмосфері та океані. Зміст змінних u , v , s дещо змінимо і будемо вважати їх загальними масами всіх трьох компонентів у тваринах, рослинах та розсіяних у земній корі, відповідно.

Кругообіг води враховувати не будемо, оскільки це призвело б до істотного ускладнення рівнянь без зміни основних властивостей їх розв'язків. Концептуальну модель динаміки складу атмосфери, що враховує кругообіг азоту, зображено на рис. 3.2.

Їй відповідає така математична модель:

$$\left\{ \begin{array}{l} \frac{dx}{dt} = -\alpha_{14}u - \alpha_{15}v + \alpha_{51}v; \\ \frac{dy}{dt} = \alpha_{42}u - \alpha_{25}v + \alpha_{52}v; \\ \frac{dz}{dt} = -\alpha_{35}v; \\ \frac{du}{dt} = \alpha_{14}u - \alpha_{42}u - \alpha_{46}u + \beta uv; \\ \frac{dv}{dt} = \alpha_{15}v - \alpha_{51}v + \alpha_{25}v - \alpha_{52}v + \alpha_{35}v - \alpha_{56}v - \beta uv; \\ \frac{ds}{dt} = \alpha_{46}u + \alpha_{56}v. \end{array} \right. \quad (3.9)$$

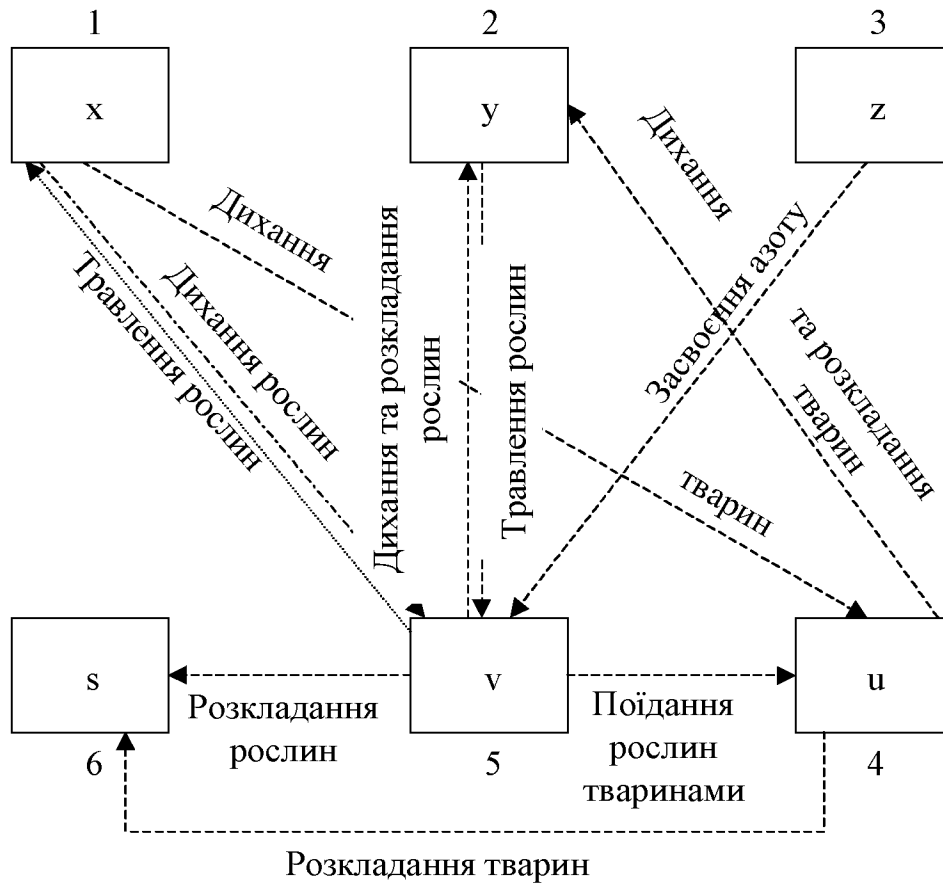


Рис. 3.2. Концептуальна модель динаміки складу атмосфери, що враховує кругообіг азоту

Додатково можна прийняти такі припущення:

- загальна біомаса тварин за відсутності рослин зменшується, тобто:

$$\lambda = \alpha_{42} + \alpha_{46} - \alpha_{14} > 0; \quad (3.10)$$

– загальна біомаса рослин за відсутності тварин зростає, тобто:

$$\mu = \alpha_{15} - \alpha_{51} + \alpha_{25} - \alpha_{52} + \alpha_{35} - \alpha_{56} > 0. \quad (3.11)$$

За цих умов розв'язками четвертого та п'ятого рівнянь системи (3.9) є періодичні функції u та v , які задовольняють співвідношення (3.4), s – зростає, а z – спадає. Рівняння (3.7) набувають вигляду:

$$\begin{cases} \frac{x(t+\omega) - x(t)}{\omega} = \frac{1}{\beta} [(\alpha_{46} + \alpha_{42})(\alpha_{51} - \alpha_{15}) - \alpha_{14}(\alpha_{25} - \alpha_{52} + \alpha_{35} - \alpha_{56})]; \\ \frac{y(t+\omega) - y(t)}{\omega} = \frac{1}{\beta} [\alpha_{42}(\alpha_{15} - \alpha_{51} + \alpha_{35} - \alpha_{56}) - (\alpha_{25} - \alpha_{52})(\alpha_{46} - \alpha_{14})]; \\ \frac{z(t+\omega) - z(t)}{\omega} = -\frac{\alpha_{35}(\alpha_{42} + \alpha_{46} - \alpha_{14})}{\beta}. \end{cases} \quad (3.12)$$

Звідси випливає, що:

$$\frac{x(t+\omega) + y(t+\omega) + z(t+\omega) - x(t) - y(t) - z(t)}{\omega} = -\frac{\lambda\alpha_{56} + \mu\alpha_{46}}{\beta} < 0, \quad (3.13)$$

тобто маса атмосфери зменшується з часом. Висновок про те, за рахунок яких газів відбувається це зменшення, залежить від припущень, аналогічних нерівностям, які розглядалися при аналізі кругообігу кисню та вуглецю.

Ці результати практично не змінюються, якщо замість сталих коефіцієнтів увести функції, які слабо залежать від часу або від змінних x , y , z , u , v , s . У

цьому разі періодичний процес стає квазіперіодичним, а точка (u, v) , що зображує органічний світ на фазовій площині, буде описувати певну нерегулярну криву. Її кутові точки відповідають різким змінам умов існування рослин та тварин, наприклад у періоди інтенсивної вулканічної діяльності або заледенінь. Збіднення атмосфери хоча б одним з потрібних для життя газів може призвести до швидкого вимирання значної частини органічного світу. Таке порушення рівноваги може також сприяти розвитку інших видів тварин і рослин, що відображається стрибкоподібною зміною коефіцієнтів моделі. Якщо після подібного збурення настає відносно тривалий спокійний період, то рівновага поступово відновлюється, коливання u та v згасають, і точка (u, v) прагне до стану стійкої рівноваги (u_c, v_c) . Режим наближення до рівноваги є близьким до періодичного, але супроводжується зменшенням маси атмосфери внаслідок розсіювання газів у земній корі.

У моделі (3.9) не враховано такі джерела та стоки газів, як вулканізм, промисловість, ерозія тощо. Позначимо через σ приріст загальної маси вуглекислого газу в атмосфері та океані за одиницю часу внаслідок дії цих джерел та стоків. Тоді друге рівняння системи (3.9) набуває вигляду:

$$\frac{dy}{dt} = \sigma + \alpha_{42}u - \alpha_{25}v + \alpha_{52}v, \quad (3.14)$$

а приріст кількості вуглекислоти за період дорівнюватиме:

$$\frac{y(t + \omega) - y(t)}{\omega} = \alpha_{42}u_c - \alpha_{25}v_c + \alpha_{52}v_c + \frac{1}{\omega} \int_t^{t+\omega} \sigma(s) ds. \quad (3.15)$$

Ця величина зазвичай є від'ємною константою, але у певні періоди часу вона може швидко змінюватися і навіть ставати великою додатною величиною, достатньою для забезпечення суттєвого збагачення атмосфери вуглекислим газом.

Відмітимо, що початок розвитку органічного світу не може бути описаний моделлю В.О. Костицина, оскільки вона дає ненульові значення u та v лише за умови, що u_0 та v_0 відмінні від нуля.

3.2. Аналіз періодичних розв'язків моделі В.О. Костицина

Система (3.3) є системою Лоттки – Вольтерра, що описує класичну систему "хижак – жертва". Наявність періодичних розв'язків впливає із загальної теорії систем Ляпунова. Праві частини її рівнянь є аналітичними функціями і допускають аналітичність першого інтегралу (3.4). Тоді, згідно із загальною теорією систем Ляпунова, ці рівняння мають періодичний розв'язок в околі точки спокою:

$$u^* = \mu/\beta, \quad v^* = \lambda/\beta, \quad (3.16)$$

який може бути поданим у вигляді рядів за степенями c , де c – початкове відхилення функції $u(t)$ або $v(t)$ від рівноваги.

Для побудови періодичних розв'язків можна застосовувати різноманітні аналітичні і графічні методи, зокрема методи Ван-дер-Поля, Крилова – Боголюбова й фазової площини.

Метод Ван-дер-Поля. Цей метод є найпростішим варіантом метода осереднення. Положимо: $u = u^* + \xi$, $v = v^* + \eta$. Тоді рівняння (3.3) можна записати у такій спосіб:

$$\dot{\xi} = \mu\eta + \beta\xi\eta, \quad \dot{\eta} = -\lambda\xi - \beta\xi\eta. \quad (3.17)$$

Для аналізу цієї системи уведемо змінні Ван-дер-Поля:

$$\xi = c\sqrt{\mu} \cos \varphi, \quad \eta = -c\sqrt{\lambda} \sin \varphi. \quad (3.18)$$

Підставимо (3.18) у (3.17) й розв'яжемо отримані рівняння стосовно \dot{c} та $\dot{\varphi}$. Отримаємо:

$$\begin{aligned}\dot{c} &= -\frac{\beta c^2}{\sqrt{\mu}} \left\{ \lambda \sqrt{\mu} \cos^2 \varphi \sin \varphi + \mu \sqrt{\lambda} \cos \varphi \sin^2 \varphi \right\}; \\ \dot{\varphi} &= \sqrt{\lambda \mu} + \frac{\beta c}{\sqrt{\lambda \mu}} \left\{ \lambda \sqrt{\mu} \cos \varphi \sin^2 \varphi - \mu \sqrt{\lambda} \cos^2 \varphi \sin \varphi \right\}.\end{aligned}\tag{3.19}$$

Праві частини рівнянь (3.19) є періодичними функціями "швидкої" змінної φ з періодом 2π . Тому для отримання асимптотичного розв'язку з точністю $O(c^2)$ достатньо розв'язати систему, одержану з (3.19) осередненням за φ . Оскільки:

$$\int_0^{2\pi} \cos^2 \varphi \sin \varphi \, d\varphi = \int_0^{2\pi} \cos \varphi \sin^2 \varphi \, d\varphi = 0,$$

після осереднення одержимо:

$$\dot{c} = 0, \quad \dot{\varphi} = \sqrt{\lambda \mu}.\tag{3.20}$$

Тоді у першому наближенні маємо:

$$T = 2\pi / \sqrt{\lambda \mu},\tag{3.21}$$

$$\xi = c \cos(\sqrt{\lambda \mu} t + t_0); \quad \eta = -c \sqrt{\frac{\lambda}{\mu}} \sin(\sqrt{\lambda \mu} t + t_0).\tag{3.22}$$

Таким чином, в межах методу Ван-дер-Поля розв'язки рівнянь (3.18) збігаються з розв'язками лінеаризованих рівнянь: змінні ξ та η гармонічно коли-

ваються з частотою $\sqrt{\lambda\mu}$, яка не залежить від амплітуди c . Цей результат впливає із загальної теорії систем Ляпунова. Точніше період можна визначити через ряд:

$$T = \frac{2\pi}{\sqrt{\lambda\mu}} (1 + hc^2 + \dots), \quad (3.23)$$

що не містить першого степеня амплітуди c . Це підтверджує, що за малих відхилень від рівноваги залежністю періоду від амплітуди можна зневажати без втрати точності.

Метод Крилова – Боголюбова є загальним методом осереднення. Розглянемо рівняння (3.19). Вважаючи відхилення від рівноваги малим, уведемо малий параметр ε і прийнемо $c = \varepsilon z$, де $\varepsilon = z(0)$. Тоді система (3.19) набуде вигляду:

$$\begin{aligned} \dot{z} &= -\frac{\varepsilon\beta z^2}{\sqrt{\mu}} \left\{ \lambda\sqrt{\mu} \cos^2 \varphi \sin \varphi + \mu\sqrt{\lambda} \cos \varphi \sin^2 \varphi \right\}; \\ \dot{\varphi} &= \sqrt{\lambda\mu} + \frac{\varepsilon\beta z}{\sqrt{\lambda\mu}} \left\{ \lambda\sqrt{\mu} \cos \varphi \sin^2 \varphi - \mu\sqrt{\lambda} \cos^2 \varphi \sin \varphi \right\}. \end{aligned} \quad (3.24)$$

При аналізі динаміки систем часто виникає необхідність відділення швидких процесів від повільних. Це дає змогу спростити математичну модель шляхом її декомпозиції й роздільного опису відповідних процесів. Для цього уведемо нові змінні:

$$z = \bar{z} + \varepsilon u(\bar{z}, \varepsilon), \quad \varphi = \bar{\varphi} + \varepsilon v(\bar{z}, \varepsilon), \quad (3.25)$$

де

$$\dot{\bar{z}} = \varepsilon A(t, \varepsilon), \quad \dot{\bar{\varphi}} = \sqrt{\lambda\mu} + \varepsilon B(t, \varepsilon). \quad (3.26)$$

На функції u та v додатково накладемо умову обмеженості при $\bar{\varphi} \rightarrow \infty$.

Підставляючи (3.25, 3.26) до (3.24), отримаємо:

$$\left\{ \begin{aligned} \frac{\partial u}{\partial \varphi} \sqrt{\lambda \mu} &= -\frac{\beta(\bar{z} + \varepsilon u)^2}{\sqrt{\mu}} \left\{ \lambda \sqrt{\mu} \cos^2(\bar{\varphi} + \varepsilon v) \sin(\bar{\varphi} + \varepsilon v) + \right. \\ &+ \mu \sqrt{\lambda} \cos(\bar{\varphi} + \varepsilon v) \sin^2(\bar{\varphi} + \varepsilon v) \left. \right\} - A - \varepsilon \frac{\partial u}{\partial \varphi} B - \varepsilon \frac{\partial u}{\partial z} A; \\ \frac{\partial v}{\partial \varphi} \sqrt{\lambda \mu} &= \frac{\beta(\bar{z} + \varepsilon u)}{\sqrt{\lambda \mu}} \left\{ \lambda \sqrt{\mu} \cos(\bar{\varphi} + \varepsilon v) \sin^2(\bar{\varphi} + \varepsilon v) - \right. \\ &- \mu \sqrt{\lambda} \cos^2(\bar{\varphi} + \varepsilon v) \sin(\bar{\varphi} + \varepsilon v) \left. \right\} - B - \varepsilon \frac{\partial v}{\partial \varphi} B - \varepsilon \frac{\partial v}{\partial z} A. \end{aligned} \right. \quad (3.27)$$

Послідовні наближення для розв'язання системи (3.27) можна розраховувати за такою схемою (k – номер ітерації):

$$\begin{aligned} \frac{\partial u^k}{\partial \varphi} \sqrt{\lambda \mu} &= \\ &= -\frac{\beta(\bar{z} + \varepsilon u^{k-1})^2}{\sqrt{\mu}} \left\{ \lambda \sqrt{\mu} \cos^2(\bar{\varphi} + \varepsilon v^{k-1}) \sin(\bar{\varphi} + \varepsilon v^{k-1}) + \right. \\ &+ \mu \sqrt{\lambda} \cos(\bar{\varphi} + \varepsilon v^{k-1}) \sin^2(\bar{\varphi} + \varepsilon v^{k-1}) \left. \right\} - \\ &- A^k - \varepsilon \frac{\partial u^{k-1}}{\partial \varphi} B^{k-1} - \varepsilon \frac{\partial u^{k-1}}{\partial z} A^{k-1}; \end{aligned} \quad (3.28)$$

$$\begin{aligned} \frac{\partial v^k}{\partial \varphi} \sqrt{\lambda \mu} &= \\ &= \frac{\beta(\bar{z} + \varepsilon u^{k-1})}{\sqrt{\lambda \mu}} \left\{ \lambda \sqrt{\mu} \cos(\bar{\varphi} + \varepsilon v^{k-1}) \sin^2(\bar{\varphi} + \varepsilon v^{k-1}) - \right. \\ &- \mu \sqrt{\lambda} \cos^2(\bar{\varphi} + \varepsilon v^{k-1}) \sin(\bar{\varphi} + \varepsilon v^{k-1}) \left. \right\} - \\ &- B^k - \varepsilon \frac{\partial v^{k-1}}{\partial \varphi} B^{k-1} - \varepsilon \frac{\partial v^{k-1}}{\partial z} A^{k-1}. \end{aligned} \quad (3.29)$$

Якщо покласти: $u^0 = v^0 = A^0 = B^0 \equiv 0$, то для першого наближення будемо мати:

$$\frac{\partial u^1}{\partial \varphi} \sqrt{\lambda \mu} = -\frac{\beta \bar{z}^{-2}}{\sqrt{\mu}} \left\{ \lambda \sqrt{\mu} \cos^2 \bar{\varphi} \sin \bar{\varphi} + \mu \sqrt{\lambda} \cos \bar{\varphi} \sin^2 \bar{\varphi} \right\} - A^1; \quad (3.30)$$

$$\frac{\partial v^1}{\partial \varphi} \sqrt{\lambda \mu} = \frac{\beta \bar{z}^{-2}}{\sqrt{\lambda \mu}} \left\{ \lambda \sqrt{\mu} \cos \bar{\varphi} \sin^2 \bar{\varphi} - \mu \sqrt{\lambda} \cos^2 \bar{\varphi} \sin \bar{\varphi} \right\} - B^1. \quad (3.31)$$

Праві частини (3.30, 3.31) є періодичними функціями $\bar{\varphi}$, тому для того, щоб ці рівняння допускали обмежені розв'язки за $\varphi \rightarrow \infty$, необхідно й достатньо, щоб середні значення правих частин дорівнювали нулю. Звідси маємо:

$$A^1 = 0, \quad B^1 = 0, \quad (3.32)$$

й, відповідно,

$$u^1(\bar{\varphi}) = -\frac{\beta \bar{z}^{-2}}{\sqrt{\mu}} \int_0^{\bar{\varphi}} \left\{ \lambda \sqrt{\mu} \cos^2 \bar{\varphi} \sin \bar{\varphi} + \mu \sqrt{\lambda} \cos \bar{\varphi} \sin^2 \bar{\varphi} \right\} d\bar{\varphi} + F_1(\bar{z}); \quad (3.33)$$

$$v^1(\bar{\varphi}) = \frac{\beta \bar{z}^{-2}}{\sqrt{\lambda \mu}} \int_0^{\bar{\varphi}} \left\{ \lambda \sqrt{\mu} \cos \bar{\varphi} \sin^2 \bar{\varphi} - \mu \sqrt{\lambda} \cos^2 \bar{\varphi} \sin \bar{\varphi} \right\} d\bar{\varphi} + F_2(\bar{z}). \quad (3.34)$$

Точність наближеного розв'язку не залежить від вибору $F_1(\bar{z})$ і $F_2(\bar{z})$.

Тому прийmemo їх рівними нулю і після інтегрування отримаємо:

$$u^1(\bar{\varphi}) = -\frac{\beta \bar{z}^{-2}}{3} \left\{ \sqrt{\frac{\lambda}{\mu}} (1 - \cos^2 \bar{\varphi}) - \sin^3 \bar{\varphi} \right\}, \quad (3.35)$$

$$v^1(\varphi) = \frac{\beta \bar{z}}{3} \left\{ \frac{1}{\sqrt{\mu}} \sin^3 \bar{\varphi} - \frac{1}{\sqrt{\lambda}} (1 - \cos^3 \varphi) \right\}. \quad (3.36)$$

З умов (3.26) та (3.32) випливає, що у цьому наближенні $\bar{z} = \text{const}$, $\bar{\varphi} = \sqrt{\lambda \mu t}$. Звідси:

$$z = \bar{z} - \frac{\varepsilon \beta \bar{z}^2}{3} \left(\sqrt{\frac{\lambda}{\mu}} (1 - \cos^3 \sqrt{\lambda \mu t}) - \sin^3 \sqrt{\lambda \mu t} \right), \quad (3.37)$$

$$\varphi = \sqrt{\lambda \mu t} + \frac{\varepsilon \beta \bar{z}}{3} \left(\frac{1}{\sqrt{\mu}} \sin^3 \sqrt{\lambda \mu t} - \frac{1}{\sqrt{\lambda}} (1 - \cos^3 \sqrt{\lambda \mu t}) \right). \quad (3.38)$$

Повертаючись до вихідних змінних ξ та η , отримаємо:

$$\begin{aligned} \xi = c_0 - \frac{\beta c_0^2}{3} \left\{ \sqrt{\frac{\lambda}{\mu}} (1 - \cos^3 \sqrt{\lambda \mu t}) - \sin^3 \sqrt{\frac{\lambda}{\mu}} (1 - \cos^3 \sqrt{\lambda \mu t}) \right\} \times \\ \times \cos \left\{ \sqrt{\lambda \mu t} + \frac{c_0 \beta}{3} \left(\frac{1}{\sqrt{\mu}} \sin^3 \sqrt{\lambda \mu t} - \frac{1}{\sqrt{\lambda}} (1 - \cos^3 \sqrt{\lambda \mu t}) \right) \right\}; \end{aligned} \quad (3.39)$$

$$\begin{aligned} \eta = -\sqrt{\frac{\lambda}{\mu}} \left\{ c_0 - \frac{\beta c_0^2}{3} \left(\sqrt{\frac{\lambda}{\mu}} (1 - \cos^3 \sqrt{\lambda \mu t}) - \right. \right. \\ \left. \left. - \sin^3 \sqrt{\lambda \mu t} \right) \right\} \sin \left\{ \sqrt{\lambda \mu t} + \frac{c_0 \beta}{3} \left(\frac{1}{\sqrt{\mu}} \sin^3 \sqrt{\lambda \mu t} - \frac{1}{\sqrt{\lambda}} (1 - \cos^3 \sqrt{\lambda \mu t}) \right) \right\}. \end{aligned} \quad (3.40)$$

Формули (3.39, 3.40) дають змогу точніше розрахувати характер зміни функцій ξ та η залежно від початкового відхилення c_0 від положення рівноваги, порівняно з (3.22).

3.3. Моделювання динаміки складу атмосфери у пакеті MathCad

Для ілюстрації основних закономірностей динаміки складу атмосфери, описаних вище, розглянемо деякі результати чисельного дослідження системи (3.5), отримані за допомогою пакету MathCad. Нижче наведено програму (написана С.В. Галенко) та графіки, що ілюструють динаміку складових і фазову траєкторію системи.

$$\lambda := 0.5$$

$$\mu := 0.05$$

$$\beta := 0.5$$

$$F(t, u, v) := u \cdot (-\lambda + \beta \cdot v)$$

$$G(t, u, v) := v \cdot (\mu - \beta \cdot u)$$

$$a := 0$$

$$\text{ORIGIN} := 0$$

$$b := 2000$$

$$N := 1000$$

$$h := \frac{(b - a)}{N}$$

$$h = 2$$

$$K1(t, u, v) := h \cdot F(t, u, v)$$

$$L1(t, u, v) := h \cdot G(t, u, v)$$

$$K2(t, u, v) := h \cdot F\left(t + \frac{h}{2}, u + \frac{K1(t, u, v)}{2}, v + \frac{L1(t, u, v)}{2}\right)$$

$$L2(t, u, v) := h \cdot G\left(t + \frac{h}{2}, u + \frac{K1(t, u, v)}{2}, v + \frac{L1(t, u, v)}{2}\right)$$

$$K3(t, u, v) := h \cdot F\left(t + \frac{h}{2}, u + \frac{K2(t, u, v)}{2}, v + \frac{L2(t, u, v)}{2}\right)$$

$$L3(t, u, v) := h \cdot G\left(t + \frac{h}{2}, u + \frac{K2(t, u, v)}{2}, v + \frac{L2(t, u, v)}{2}\right)$$

$$K4(t, u, v) := h \cdot F(t + h, u + K3(t, u, v), v + L3(t, u, v))$$

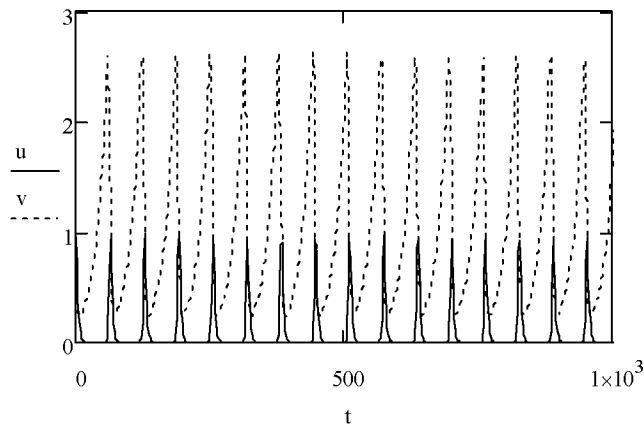
$$L4(t, u, v) := h \cdot G(t + h, u + K3(t, u, v), v + L3(t, u, v))$$

$$\Delta K(t, u, v) := \frac{(K1(t, u, v) + 2 \cdot K2(t, u, v) + 2 \cdot K3(t, u, v) + K4(t, u, v))}{6}$$

$$\Delta L(t, u, v) := \frac{(L1(t, u, v) + 2 \cdot L2(t, u, v) + 2 \cdot L3(t, u, v) + L4(t, u, v))}{6}$$

$$i := 0..N - 1$$

$$\begin{pmatrix} t_0 \\ u_0 \\ v_0 \end{pmatrix} := \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad \begin{pmatrix} t_{i+1} \\ u_{i+1} \\ v_{i+1} \end{pmatrix} := \begin{pmatrix} t_i + h \\ u_i + \Delta K(t_i, u_i, v_i) \\ v_i + \Delta L(t_i, u_i, v_i) \end{pmatrix}$$



$$F(t_1, u_1, v_1) = -0.195$$

$$K1(t_1, u_1, v_1) = -0.389$$

$$\Delta K(t_1, u_1, v_1) = -0.34$$

$$u_1 = 0.718$$

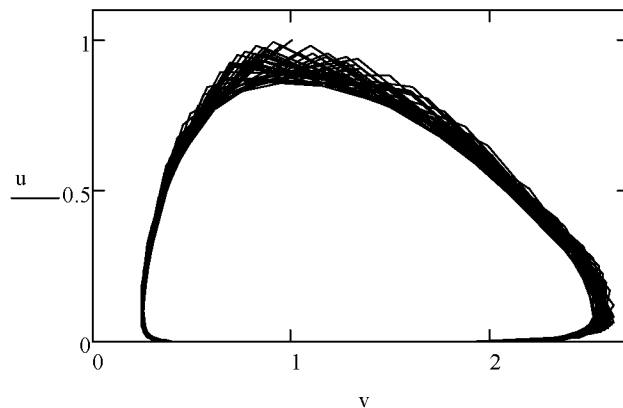
$$t_1 = 2$$

$$G(t_1, u_1, v_1) = -0.141$$

$$L1(t_1, u_1, v_1) = -0.283$$

$$\Delta L(t_1, u_1, v_1) = -0.163$$

$$v_1 = 0.458$$



З наведених результатів добре видно періодичний характер змінювання загальних обсягів вуглецю та кисню в тваринах і рослинах, що відповідає періодичним коливанням пропорційних цим величинам загальних біомас тварин і рослин. Другий графік ілюструє зазначену вище відповідність замкнених фазових траєкторій в координатах (v, u) періодичним розв'язкам системи (3.3).

3.4. Задача трьох тіл: чисельне моделювання траєкторії руху

Однією з класичних задач математичного моделювання є задача трьох тіл. Вона досліджує рухи у системі, яка складається із трьох тіл, що взаємодіють один з одним. Її прикладом може бути система "Сонце – планета – супутник планети".

Першим етапом виконання дослідження має бути побудова фізичної моделі досліджуваної системи. Розглянемо його більш детально на прикладі моделювання траєкторії руху Місяця навколо Сонця.

Розробка фізичної моделі починається з визначення складу системи. Як супутник Землі, Місяць входить до складу Сонячної системи, яка, у свою чергу, є підсистемою більш великих зіркових систем. Проте, не всі об'єкти цих систем будуть суттєво впливати на рух Місяця. При моделюванні його руху до складу системи, очевидно, слід включити сам Місяць (M), Землю (E) та Сонце (S). Питання щодо необхідності включення інших небесних тіл (B) до складу системи повинно вирішуватися залежно від мети моделювання. Критерієм впливу цих

тіл на рух Місяця може служити відношення $\frac{a_B}{a_E} = \frac{M_B r_E^2}{M_E r_B^2}$, де a_B і a_E – прискорення,

що набуває Місяць внаслідок його гравітаційної взаємодії з тілом, яке розглядається, та Землею, M_B та M_E – маси тіла і Землі, r_B та r_E – мінімальні відстані від Місяця до них. Маси планет Сонячної системи та Землі можна знайти в навчальній і довідковій астрономічній літературі. Дані про мінімальні відстані між Місяцем та іншими небесними тілами містяться у фахових виданнях, які є важкодоступними. Проте ці величини можна достатньо точно оцінити так. Відстані від Місяця до інших планет Сонячної системи великі порівняно з відстанню від нього до Землі. Тому вони є близькими до відстаней цих планет від Землі. Вони будуть мінімальні, якщо Земля та планета знаходяться з одного боку від Сонця на одній прямій з ним. Орбіти планет Сонячної системи у першому наближенні можна вважати коловими. Тому мінімальні відстані приблизно будуть дорівнювати різниці між радіусами орбіт планет і радіусом орбіти Землі. Значення цих параметрів є у навчальних та довідкових виданнях з астрономії.

Критерієм слабкості впливу планети на рух Місяця можна вважати виконання нерівності $a_B/a_E < \alpha$, де значення α обирається залежно від мети моделювання. Воно може, наприклад, дорівнювати 0,001–0,01. Значення параметра α має бути тим нижчим, чим більша потрібна точність результату, а також, чим більший період часу, протягом якого необхідно дослідити траєкторію руху.

Вплив небесних тіл, що не входять до складу сонячної системи, можна, у першому наближенні, розглядати як вплив зовнішнього середовища на систему в цілому. Тоді за необхідності більш точного аналізу руху Місяця в зоряній системі відліку ми можемо здійснити декомпозицію моделі і розглядати його як суму руху Місяця стосовно Сонця та руху сонячної системи як цілого стосовно зоряної системи координат.

Наступним кроком побудови моделі є визначення характеру зв'язків між елементами системи, тобто характеру взаємодії Місяця з іншими тілами, які ми зарахували до складу досліджуваної системи. Найбільш важливою буде гравітаційна взаємодія. Залежно від потрібної точності моделі її описують за допомогою закону всесвітнього тяжіння або за допомогою загальної теорії відносності. Можна вважати, що поправка, яку дає остання, в даному випадку не суттєва, тобто її можна не враховувати. Тоді сила, що діє на Місяць з боку іншої планети (P), буде дорівнювати:

$$\mathbf{F} = -\gamma \frac{M_P m}{r_{PM}^3} \mathbf{r}_{PM}, \quad (3.41)$$

де $\gamma = 6,673 \cdot 10^{-11} \text{ Н} \cdot \text{м}^2 \cdot \text{кг}^{-2}$ – гравітаційна стала, m – маса Місяця, \mathbf{r}_{PM} – радіус-вектор, проведений від планети до Місяця, r_{PM} – його довжина.

Закон всесвітнього тяжіння сформульовано для матеріальних точок, тому перед застосуванням треба визначити можливість його використання у вигляді (3.41). Є три основних джерела можливих помилок. По-перше, теорія дозволяє використовувати звичайну форму закону для опису взаємодії неточкових тіл, які мають форму кулі. Форми Місяця, Землі, Сонця та інших небесних тіл дещо

відрізняються від кульової. Цим можна зневажати при аналізі руху планет. Але, якщо розраховуються траєкторії штучних супутників Землі та інших об'єктів, що знаходяться на малих, порівняно з радіусами планет, відстанях від них, відмінність форми планети від кульової може стати суттєвим джерелом похибки. По-друге, внаслідок того, що досліджувані об'єкти не є точковими тілами, взаємодія може призводити до внутрішнього перерозподілу їх мас, а також до зміни їх обертового руху, що внаслідок дії законів збереження буде відображатися й на їх орбітальному русі. Прикладом є вплив земних припливів і відпливів, зумовлених гравітаційною взаємодією з Місяцем, на її орбітальний рух. Необхідність врахування відповідних ефектів залежить від мети моделювання і потрібної точності розрахунків. Зазвичай ці ефекти є суттєвими лише для великих (порівняно з періодом обертання) інтервалів часу. По-третє, слід враховувати вибір системи відліку, в якій розв'язується задача. Якщо обирають геліоцентричну систему, тобто Сонце поміщають до початку координат, то система відліку буде неінерційною, завдяки галактичному руху Сонця та його руху внаслідок взаємодії з планетами Сонячної системи. Тобто при розрахунку сил, що діють на Місяць та інші планети, необхідно до сили гравітаційної взаємодії додати силу інерції, зумовлену прискоренням рухом системи відліку. Якщо ж розглядати рух Місяця в інерційній системі відліку, то треба враховувати рух Сонця стосовно цієї системи. Ці поправки будуть відносно малими, тому у першому наближенні ними можна зневажати. Але при точних розрахунках їх потрібно враховувати.

Виходячи з наведеного вище, математичну модель руху Місяця (для випадку, коли враховується лише його взаємодія із Сонцем та Землею і нехтується неінерційністю геліоцентричної системи відліку) можна записати так:

$$\left\{ \begin{array}{l} \frac{d^2 \mathbf{r}_M}{dt^2} = -G \left(\frac{M_E \mathbf{r}_{EM}}{r_{EM}^3} + \frac{M_S \mathbf{r}_{SM}}{r_{SM}^3} \right), \\ \frac{d^2 \mathbf{r}_E}{dt^2} = -G \left(\frac{m \mathbf{r}_{ME}}{r_{ME}^3} + \frac{M_C \mathbf{r}_{SE}}{r_{SE}^3} \right), \\ \mathbf{r}_M(t = t_0) = \mathbf{r}_{M0}, \\ \mathbf{v}_M(t = t_0) = \mathbf{v}_{M0}, \\ \mathbf{r}_E(t = t_0) = \mathbf{r}_{E0}, \\ \mathbf{v}_E(t = t_0) = \mathbf{v}_{E0}. \end{array} \right. \quad (3.42)$$

Модель (3.42) є описовою, динамічною, стаціонарною, диференціальною, нелінійною, неперервною, із зосередженими параметрами. Незважаючи на її простоту ця модель має аналітичні розв'язки лише в деяких окремих випадках. Тому для реальних ситуацій її слід розв'язувати за допомогою чисельних методів. Враховуючи, що модель (3.42) є системою звичайних диференціальних рівнянь, як стандартну математичну схему її дослідження доцільно вибрати один із чисельних методів розв'язування задачі Коші.

Для чисельного аналізу цієї моделі необхідно перейти від векторної форми запису до скалярної. У загальному випадку кожне з рівнянь системи (3.42) розпадається ще на три. Наприклад, для першого рівняння одержимо:

$$\left\{ \begin{array}{l} \frac{d^2 x_M}{dt^2} = -G \left[\frac{M_E}{r_{EM}^3} (x_M - x_E) + \frac{M_S}{r_{SM}^3} x_M \right]; \\ \frac{d^2 y_M}{dt^2} = -G \left[\frac{M_E}{r_{EM}^3} (y_M - y_E) + \frac{M_S}{r_{SM}^3} y_M \right]; \\ \frac{d^2 z_M}{dt^2} = -G \left[\frac{M_E}{r_{EM}^3} (z_M - z_E) + \frac{M_S}{r_{SM}^3} z_M \right], \end{array} \right. \quad (3.43)$$

де x, y, z – декартові координати.

Для задачі, що розглядається, можна прийняти, що орбіти, за якими рухаються Місяць навколо Землі та Земля навколо Сонця, лежать в одній площині.

ні. Тоді ми одержимо плоску задачу і достатньо буде замінити кожне векторне рівняння на два скалярних.

Схема розв'язування системи (3.42) після її переведення до скалярної форми може бути такою. На першому етапі необхідно задати початкові умови, а саме початкові координати Місяця і Землі та початкові проекції їх швидкостей. Початкові координати зручно взяти такими, щоб у момент $t = 0$ Сонце, Місяць та Земля розміщувалися вздовж однієї прямої. Тоді цю пряму можна обрати як одну з осей координат, наприклад, як вісь x . Початковою координатою Землі x_E візьмемо середній радіус Земної орбіти. Для одержання початкової координати Місяця x_M до координати Землі додамо середній радіус орбіти Місяця. При такому виборі початкових координат ми нехтуємо відмінністю орбіт Землі та Місяця від кіл. Початкові швидкості Землі та Місяця будуть спрямовані перпендикулярно вісі x . Тому їх проекції на вісь Ox дорівнюють нулю. Зневажаючи несталість кутових швидкостей орбітальних рухів Землі та Сонця, можна прийняти:

$$v_{yE} = \frac{2\pi R_E}{T_E}; \quad v_{yM} = v_{yE} + \frac{2\pi R_M}{T_M}, \quad (3.44)$$

де R_E та R_M – відповідно, середні радіуси орбіт Землі та Місяця; T_E і T_M – періоди обертання Земля навколо Сонця та Місяця навколо Землі. Слід відзначити, що оскільки орбіти Місяця та Землі близькі до колових, їх початкові швидкості будуть близькими до значень перших космічних швидкостей для систем "супутник – Земля" і "супутник – Сонце" при відповідних відстанях до Землі і Сонця. Як відомо з курсу фізики, перші космічні швидкості визначають межу між замкненими (для $v \geq v_{1к}$) й незамкненими (для $v < v_{1к}$) траєкторіями руху супутника. Тому невелика помилка у визначенні початкових швидкостей може призвести до якісної зміни розрахованої траєкторії руху. Але така похибка легко може бути усунена при наступному корегуванні моделі.

Після визначення початкових координат та проекцій швидкостей можна розрахувати проекції прискорень Місяця та Землі у початковий момент часу за формулами систем (3.42) та (3.43). Далі потрібно задати крок Δt і організувати цикл за часом. Умовою закінчення циклу може бути досягнення заданого часу руху або проходження Місяцем одного чи декількох повних обертів навколо Сонця.

У найпростішому випадку застосування методу Ейлера на кожному кроці потрібно виконати такі дії.

1. Розрахувати нові значення координат Землі та Місяця:

$$\begin{cases} x_E^{i+1} = x_E^i + v_{xE}^i \Delta t + a_{xE}^i \Delta t / 2; \\ x_M^{i+1} = x_M^i + v_{xM}^i \Delta t + a_{xM}^i \Delta t / 2; \\ y_E^{i+1} = y_E^i + v_{yE}^i \Delta t + a_{yE}^i \Delta t / 2; \\ y_M^{i+1} = y_M^i + v_{yM}^i \Delta t + a_{yM}^i \Delta t / 2. \end{cases} \quad (3.45)$$

2. Розрахувати нові значення проекцій швидкостей Землі та Місяця:

$$\begin{cases} v_{xE}^{i+1} = v_{xE}^i + a_{xE}^i \Delta t; \\ v_{xM}^{i+1} = v_{xM}^i + a_{xM}^i \Delta t; \\ v_{yE}^{i+1} = v_{yE}^i + a_{yE}^i \Delta t; \\ v_{yM}^{i+1} = v_{yM}^i + a_{yM}^i \Delta t. \end{cases} \quad (3.46)$$

3. Розрахувати нові значення проекцій прискорень Землі та Місяця згідно з (3.42, 3.43).
4. Додати час Δt і перейти до наступного кроку циклу.

На рис. 3.3 наведено фрагменти траєкторій руху трьох взаємодіючих тіл. Точками і стрілками позначені початкові координати й напрямки їх руху.

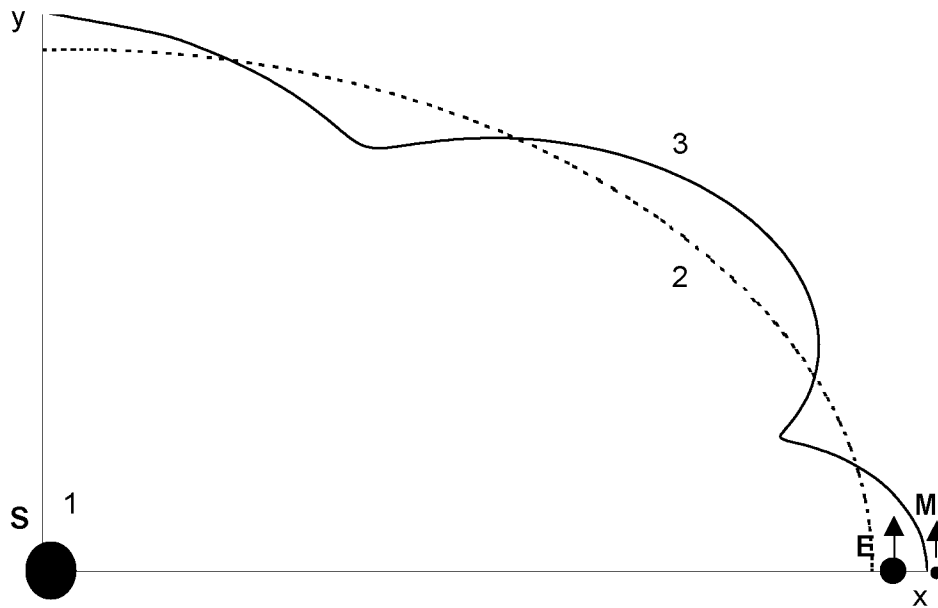


Рис. 3.3. Результати розрахунку фрагментів траєкторій руху трьох взаємодіючих тіл

Для першого тіла початкова швидкість дорівнювала нулю. Під час обчислень може виникнути така розрахункова помилка. Внаслідок скінченності кроку за часом i , відповідно, скінченності переміщень і приростів швидкостей, у деякий момент Місяць може виявитися занадто близьким до Землі, або Земля виявиться занадто близькою до Сонця. У цьому випадку на наступному кроці вони одержать великі прирости швидкостей, а далі змістяться на великі відстані від центрів орбіт. На цих відстанях сила їх притягнення до Землі або Сонця суттєво зменшиться і вони не зможуть повернутися на свої орбіти. Для запобігання виникненню такої помилки необхідно обирати невеликий крок за часом. Необхідно також ввести додаткові умови закінчення розрахунків у випадках, коли одне тіло наблизиться до іншого на відстань меншу, ніж певна заздалегідь задана для цієї пари величина. Виконання такої умови фактично буде означати зіткнення цих двох тіл та їх подальший рух як одного цілого. Якщо така ситуація виникає при моделюванні, це означає, що початкові проекції швидкостей Землі або Місяця були занадто малими і їх треба збільшити.

На практиці метод Ейлера застосовують лише у навчальних цілях при вивченні загальних принципів чисельного розв'язування диференціальних рівнянь. У реальних дослідженнях найчастіше використовують метод Рунге – Кутта 4-го порядку.

Для розв'язаного стосовно другої похідної диференціального рівняння загального вигляду:

$$\frac{d^2y}{dt^2} = f(t, y, y'); y(t_0) = y_0; y'(t_0) = z_0 \quad (3.47)$$

алгоритм розрахунку за цим методом є таким.

На першому етапі рівняння (3.47) перетворюють до системи двох рівнянь першого порядку, уводячи нову змінну:

$$\begin{aligned} \frac{dy}{dt} &= z(t, y); \\ \frac{dz}{dt} &= f(t, y, z). \end{aligned} \quad (3.48)$$

На другому етапі організують цикл за часом, основні розрахункові формули якого є такими:

$$\begin{aligned} y_{i+1} &= y_i + \frac{(k_1 + 2k_2 + 2k_3 + k_4)\Delta t}{6}; \\ z_{i+1} &= z_i + \frac{(l_1 + 2l_2 + 2l_3 + l_4)\Delta t}{6}; \\ k_1 &= z(t_i, y_i); k_2 = z\left(t_i + \frac{\Delta t}{2}, y_i + \frac{k_1}{2}\right); \\ k_3 &= z\left(t_i + \frac{\Delta t}{2}, y_i + \frac{k_2}{2}\right); k_4 = z(t_i + \Delta t, y_i + k_3); \end{aligned} \quad (3.49)$$

$$\begin{aligned} \ell_1 &= f(t_i, y_i, z_i); \ell_2 = f\left(t_i + \frac{\Delta t}{2}, y_i + \frac{k_1}{2}, z_i + \frac{\ell_1}{2}\right); \\ \ell_3 &= f\left(t_i + \frac{\Delta t}{2}, y_i + \frac{k_2}{2}, z_i + \frac{\ell_2}{2}\right); \\ \ell_4 &= f(t_i + \Delta t, y_i + k_3, z_i + \ell_3). \end{aligned}$$

Як рівняння (3.47) при моделюванні руху у системі "Сонце – Земля – Місяць" можна взяти рівняння системи (3.42), записані у скалярній формі. Наприклад, для координати x Місяця:

$$f(x_M, v_M) = f(x_M) = -G \left[\frac{M_E}{r_{EM}^3} (x_M - x_E) + \frac{M_S}{r_{SM}^3} x_M \right]. \quad (3.50)$$

Із математичним моделюванням руху планет пов'язана тривала дискусія між прихильниками геоцентричної та геліоцентричної моделей Всесвіту, яка виникла ще в античний період і закінчилася лише у XVIII–XIX сторіччях після побудови класичної механіки й експериментальної перевірки її астрономічних прогнозів. У своєму первинному вигляді ані модель Птолемея ані модель Коперніка були не в змозі адекватно описати спостережуваний рух планет. Тому вони змушені були використовувати так звані епіцикли.

У найпростішому випадку введення епіциклу передбачало, що по коловій траєкторії навколо Землі або Сонця рушиться не сама планета, а центр деякого іншого кола. Сама ж планета рушиться вздовж цього додаткового кола. Математичну модель такого руху можна записати у вигляді:

$$\begin{cases} x_1(t) = x_0 + R_1 \cos(\omega_1 t); & x_2(t) = x_1 + R_2 \cos(\omega_2 t); \\ y_1(t) = y_0 + R_1 \sin(\omega_1 t); & y_2(t) = y_1 + R_2 \sin(\omega_2 t), \end{cases} \quad (3.51)$$

де через x_0, y_0 позначено координати центру обертання, x_1, y_1 – координати поточного центру епіциклу, x_2, y_2 – координати тіла, що обертається, R_1, R_2 –

радіуси основної орбіти та епіциклу, ω_1, ω_2 – кутові швидкості обертання вздовж них. На рис. 3.4 показано вигляд руху планети за наявності одного епіциклу.

З рисунку видно, що вже при введенні одного епіциклу вдається відобразити важливу особливість реального руху планет з погляду земного спостерігача, яка полягає у наявності ділянок їх зворотного руху.

З погляду сучасної математики підхід Птолемея можна розглядати як своєрідний розклад функції, що описує траєкторію планети, у ряд за коловими траєкторіями, радіуси яких прямують до нуля.

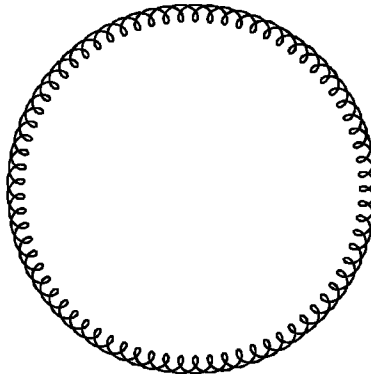


Рис. 3.4. Траєкторія руху планети за наявності одного епіциклу

Контрольні питання

1. Чому склад земної атмосфери відрізняється від складу атмосфер сусідніх планет Сонячної системи? У чому полягають основні відмінності?
2. У чому полягає роль океану в формуванні складу земної атмосфери?
3. Якими є основні компоненти системи, що визначає склад атмосфери Землі?
4. Якими є основні процеси, що визначають кругообіг кисню, вуглецю й азоту у біосфері?
5. Яким є змістове значення окремих рівнянь систем 3.1, 3.9?
6. Пояснить, чому $s(t)$ є зростаючою функцією?

7. Грунтуючись на концептуальній моделі доведіть правильність нерівності (3.2).
8. Якими є основні властивості систем Ляпунова?
9. Доведіть, що система рівнянь Лоттки – Вольтерра має періодичні розв'язки.
10. Що називають фазовою траєкторією системи? У якому випадку фазова траєкторія є замкненою?
11. Наведіть приклади співвідношень між коефіцієнтами системи (3.8), що не розглядалися і поясніть, чому вони не відповідають концептуальній моделі.
12. Яким є основне припущення методів осереднення?
13. У чому полягає метод Ван-дер-Поля? Для розв'язання яких задач його можна застосовувати?
14. У чому полягає сутність методу Крилова – Боголюбова? Для розв'язання яких задач його можна застосовувати?
15. Які засоби розв'язання диференціальних рівнянь та їх систем передбачені у пакеті MathCad?
16. Якою є загальна постановка задачі трьох тіл?
17. Чому у загальному випадку задача трьох тіл не має аналітичного розв'язку?
18. Які чисельні методи можна використовувати для розв'язування систем диференціальних рівнянь?
19. У чому полягає сутність методу Ейлера? Якими є його переваги й недоліки?
20. У чому полягає сутність методів Рунге – Кутта? Якими є їх переваги й недоліки?
21. Чому модель Птолемея можна розглядати як розклад траєкторії руху планети у ряд по коловим орбітам?

4. АНАЛІЗ СТІЙКОСТІ СИСТЕМ

4.1. Математична ознака стійкості

Систему називають **стійкою**, якщо після її виведення із стану рівноваги вона за відсутності зовнішніх впливів прагне повернутися до цього стану. Будь-яка динамічна система характеризується перехідним процесом, що виникає в ній при порушенні її рівноваги зовнішніми впливами (сигналами управління, завадами, збуреннями тощо). Перехідний процес залежить як від властивостей системи, так і від типу зовнішнього впливу. Його можна подати як суму двох складових:

$$y'(t) = y_{\text{int}}(t) + y_{\text{ext}}(t), \quad (4.1)$$

де $y_{\text{int}}(t)$ – вільні рухи системи, що визначаються її власними властивостями і початковими умовами; $y_{\text{ext}}(t)$ – змушені рухи, які визначаються збурюючим впливом та властивостями системи.

Для того, щоб система могла правильно реагувати на сигнали управління або зміну навантаження при перехідному процесі, складова, що описує вільні рухи, з часом має наближатися до нуля, тобто система є стійкою, якщо:

$$\lim_{t \rightarrow \infty} y_{\text{int}}(t) = 0. \quad (4.2)$$

Таким чином, характер вільного руху системи визначає її стійкість або нестійкість. Приклади вихідних характеристик, тобто залежностей вихідних параметрів від часу, для стійких та нестійких систем наведено на рис. 4.1.

Широкий клас динамічних систем може бути описано за допомогою диференціальних рівнянь або їх систем (нагадаємо, що будь-яке звичайне диференціальне рівняння n -го порядку можна перетворити до системи n звичайних диференціальних рівнянь першого порядку).

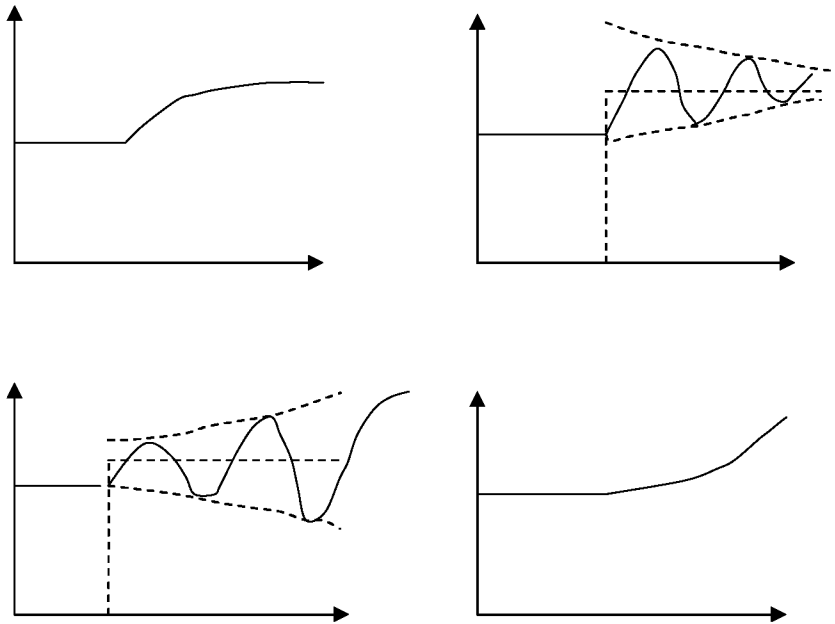


Рис. 4.1. Приклади вихідних характеристик стійких та нестійких САУ

У цьому випадку рівняння вільного руху лінійної системи, розв'язане стосовно досліджуваної величини, можна записати у вигляді:

$$a_0 \frac{d^n y_{\text{int}}}{dt^n} + a_1 \frac{d^{n-1} y_{\text{int}}}{dt^{n-1}} + \dots + a_{n-1} \frac{dy_{\text{int}}}{dt} + a_n y_{\text{int}}(t) = 0, \quad (4.3)$$

де a_0, a_1, \dots, a_n – сталі коефіцієнти, що визначаються параметрами системи. В операторній формі це рівняння набуває вигляду:

$$a_0 p^n + a_1 p^{n-1} + \dots + a_{n-1} p + a_n = 0, \quad (4.4)$$

де $p = \frac{d}{dt}$ – оператор диференціювання за часом.

З оператором можна виконувати ті самі дії, що і з іншими членами алгебраїчного рівняння (множити, ділити, виносити за дужки тощо). Можливість запису диференціального рівняння в операторній формі значно спрощує всі подальші розрахунки.

Такий перехід від диференціального рівняння до еквівалентного йому алгебраїчного називають **перетворенням Лапласа**. При цьому алгебраїчне рівняння є записом диференціального в операторній формі. Функцію $u(p)$ називають зображенням функції $u(t)$, а функцію $u(t)$ – оригіналом функції $u(p)$. Операцію переходу від вихідної функції $u(t)$ до її зображення називають прямим перетворенням Лапласа, а операцію переходу від зображення $u(p)$ до оригіналу – оберненим перетворенням Лапласа.

В математичній теорії стійкості об'єктом дослідження зазвичай є окремий розв'язок аналізованої системи диференціальних рівнянь. Основним завданням теорії стійкості є розробка методів визначення стійкості окремих розв'язків, які б не вимагали знання загального розв'язку відповідної системи. Як досліджувану величину найчастіше беруть відхилення регульованого параметра від заданого значення.

Розв'язок рівняння (4.3) при всіх дійсних коренях має вигляд:

$$y_{\text{int}}(t) = \sum_{i=1}^n A_i e^{\lambda_i t}, \quad (4.5)$$

де A_i – додатні сталі інтегрування, що визначаються параметрами системи й початковими умовами.

У більш загальному випадку нелінійних систем часто здійснюють їх лінеаризацію, яка полягає в заміні вихідного нелінійного рівняння ланки наближеною лінійною залежністю. Це можна здійснити, наприклад, розкладом нелінійної залежності в ряд Тейлора в околі точки усталеного режиму, нехтуючи при цьому членами вище першого порядку малості.

Розглянемо систему, структурна схема якої показана на рис. 4.2.

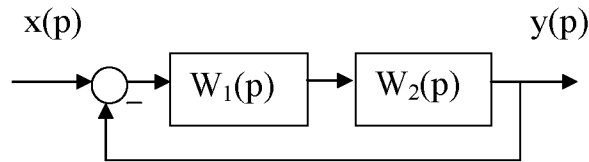


Рис. 4.2. Структурна схема системи

Передаточною функцією системи називають відношення її вихідного сигналу до вхідного. Для показаної системи, що містить дві послідовно з'єднані ланки з передавальними функціями $W_1(p)$ і $W_2(p)$, а також негативний зворотний зв'язок, згідно із загальними правилами отримання передаточних функцій з'єднань маємо:

$$W(p) = \frac{y(p)}{x(p)} = \frac{W_1 W_2}{1 + W_1 W_2}. \quad (4.6)$$

Припустимо, що обидві ланки є **інерційними**, тобто відповідають рівнянням:

$$W_1 = \frac{k_1}{1 + T_1 p}; \quad W_2 = \frac{k_2}{1 + T_2 p}, \quad (4.7)$$

де k_1, k_2, T_1, T_2 – певні константи. Підставляючи (4.7) до (4.6) і приводячи до спільного знаменника, отримаємо:

$$\left[T_1 T_2 p^2 + (T_1 + T_2) p + k_1 k_2 + 1 \right] y(p) = k_1 k_2 x(p). \quad (4.8)$$

Переходячи від зображення до оригіналу, одержимо:

$$T_1 T_2 \frac{d^2 y}{dt^2} + (T_1 + T_2) \frac{dy}{dt} + (1 + k_1 k_2) y(t) = k_1 k_2 x(t). \quad (4.9)$$

Якщо права частина цього рівняння дорівнює нулю, то ми отримаємо рівняння власних рухів, а якщо ні – рівняння змушених рухів. Для визначення стійкості системи достатньо розв'язати рівняння вільних рухів.

Позначимо: $a_0 = T_1 T_2$, $a_1 = T_1 + T_2$, $a_2 = 1 + k_1 k_2$. Тоді характеристичне рівняння матиме вигляд:

$$a_0 \lambda^2 + a_1 \lambda + a_2 = 0. \quad (4.10)$$

Його коренями є:

$$\lambda_{1,2} = -\frac{a_1}{2a_0} \pm \frac{\sqrt{a_1^2 - 4a_0 a_2}}{2a_0}. \quad (4.11)$$

Залежно від значень параметрів a_0, a_1, a_2 ці корені можуть бути додатними, від'ємними та уявними.

Розглянемо докладніше формулу (4.5). Оскільки всі коефіцієнти A_i є сталими величинами, то характер поведінки кожного з доданків $A_i e^{-\lambda_i t}$ залежить від значення λ_i . Для дійсних коренів: якщо $\lambda_i > 0$, то величина $A_i e^{-\lambda_i t}$ з часом збільшується до нескінченності; якщо $\lambda_i < 0$, вона наближається з часом до нуля (рис. 4.3).

У загальному випадку λ_i є комплексними величинами. Комплексні корені характеристичного рівняння завжди бувають комплексно спряженими: $\lambda_i = \alpha \pm i\beta$.

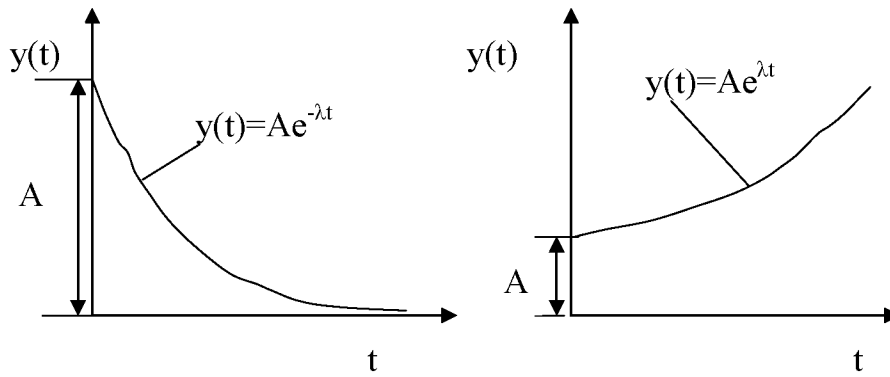


Рис. 4.3. Вигляд залежності вихідного параметра від часу для дійсних коренів характеристичного рівняння

У цьому випадку з (4.5) згідно з формулою Ейлера одержимо:

$$y_{\text{int}}(t) = A_1 e^{(\alpha+i\omega)t} + A_2 e^{(\alpha-i\omega)t} = A e^{\alpha t} \sin(\omega t + \varphi). \quad (4.12)$$

Якщо $\alpha > 0$, то амплітуда коливань з часом зростає, а у випадку $\alpha < 0$ вона наближається до нуля (рис. 4.4). Якщо $\alpha = 0$, то буде спостерігатися суто гармонічний процес з постійною амплітудою коливань.

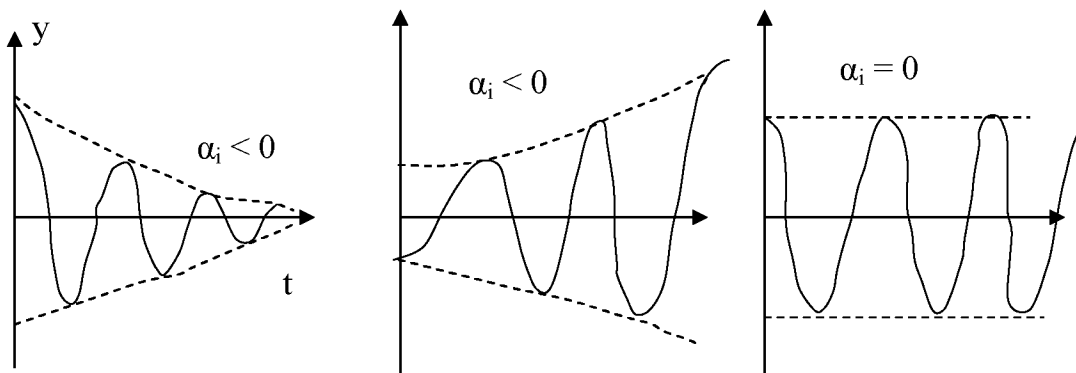


Рис. 4.4. Вигляд залежності вихідного параметра від часу для комплексних коренів характеристичного рівняння

Звідси впливає перша теорема А.М. Ляпунова:

1. Для забезпечення стійкості системи, що описується лінійними диференціальними рівняннями зі сталими коефіцієнтами, необхідно й достатньо, щоб дійсні корені характеристичного рівняння, яке відповідає вихідному диференціальному рівнянню, були від'ємними, а комплексні корені мали від'ємну дійсну частину. У цьому випадку $\lim_{t \rightarrow \infty} y_{\text{int}}(t) = 0$.

2. Якщо дійсна частина коренів є від'ємною, але є нульові або суто уявні корені, то система буде консервативною, тобто $\lim_{t \rightarrow \infty} y_{\text{int}}(t) = \text{const} \neq 0$.

3. Якщо є додатні дійсні корені або комплексні корені з додатною дійсною частиною, то система є нестійкою і $\lim_{t \rightarrow \infty} y_{\text{int}}(t) \rightarrow \infty$.

Таким чином, вигляд залежності $y_{\text{int}}(t)$ визначається знаком дійсної частини коренів характеристичного рівняння (рис. 4.5).

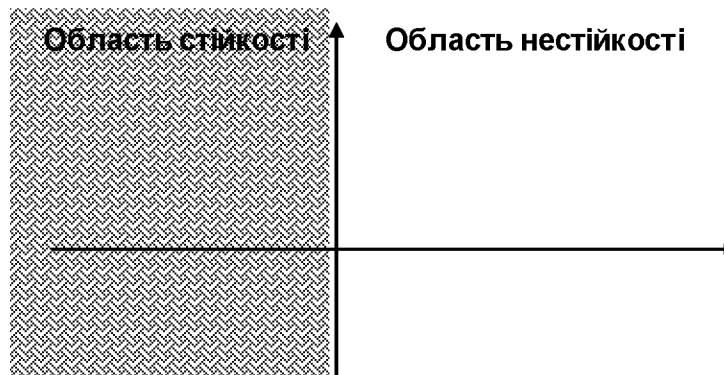


Рис. 4.5. Вплив коренів характеристичного рівняння на стійкість системи

Прямий метод аналізу стійкості, що ґрунтується на обчисленні коренів характеристичного рівняння, є досить складним, особливо у випадку рівнянь високих степенів. Тому розроблено інші методи, які дають змогу визначити стійкість системи без розв'язання характеристичного рівняння. На їх основі сформульовано певні критерії стійкості, що дають змогу одночасно визначити вплив окремих параметрів системи та структурних змін на її стійкість.

Вирізняють дві групи критеріїв стійкості – алгебраїчні (Рауса й Гурвиця), що ґрунтуються на аналізі коефіцієнтів характеристичного рівняння, і частотні (Михайлова й Найквіста), які базуються на дослідженні частотних характеристик. Частотні критерії дають змогу оцінювати стійкість системи не тільки на основі рівнянь динаміки, а й за експериментально визначеними частотними характеристиками у випадках, коли самі ці рівняння залишаються невідомими.

4.2. Критерій Гурвиця

Цей критерій дає змогу визначити місцезнаходження коренів характеристичного рівняння на комплексній площині.

Нехай характеристичне рівняння записується у вигляді:

$$a_0\lambda^n + a_1\lambda^{n-1} + a_{n-1}\lambda + a_n = 0. \quad (4.13)$$

Для перевірки стійкості системи спочатку будемо визначник Гурвиця за таким правилом: по головній діагоналі записуємо зліва направо всі коефіцієнти характеристичного рівняння від a_1 до a_n у порядку зростання індексу.

У кожному стовпчику вгору від головної діагоналі ставимо коефіцієнти характеристичного рівняння з послідовно зростаючими індексами, а вниз – коефіцієнти з послідовно згасаючими індексами. На місця, де значення індексів мають бути від'ємними або більшими, ніж n , ставимо нулі. Наприклад, для $n = 6$ отримуємо такий визначник Гурвиця:

$$\Delta_6 = \begin{vmatrix} a_1 & a_3 & a_5 & 0 & 0 & 0 \\ a_0 & a_2 & a_4 & a_6 & 0 & 0 \\ 0 & a_1 & a_3 & a_5 & 0 & 0 \\ 0 & a_0 & a_2 & a_4 & a_6 & 0 \\ 0 & 0 & a_1 & a_3 & a_5 & 0 \\ 0 & 0 & a_0 & a_2 & a_4 & a_6 \end{vmatrix}. \quad (4.14)$$

Далі виділяємо головні діагональні мінори й отримуємо визначники Гурвиця нижчих порядків. Для випадку (4.14) вони є такими:

$$\Delta_1 = a_1; \quad \Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix}; \quad \Delta_3 = \begin{vmatrix} a_1 & a_3 & a_5 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix}; \quad \Delta_4 = \begin{vmatrix} a_1 & a_3 & a_5 & 0 \\ a_0 & a_2 & a_4 & a_6 \\ 0 & a_1 & a_3 & a_5 \\ 0 & a_0 & a_2 & a_4 \end{vmatrix};$$

$$\Delta_5 = \begin{vmatrix} a_1 & a_3 & a_5 & 0 & 0 \\ a_0 & a_2 & a_4 & a_6 & 0 \\ 0 & a_1 & a_3 & a_5 & 0 \\ 0 & a_0 & a_2 & a_4 & a_6 \\ 0 & 0 & a_1 & a_3 & a_5 \end{vmatrix}.$$

Згідно з критерієм, що розглядається, для того, щоб система була стійкою, необхідно й достатньо, щоб визначник Гурвиця та його головні діагональні мінори мали той самий знак, що й коефіцієнт a_0 . Оскільки знак першого коефіцієнта характеристичного рівняння зазвичай вибирають додатним, то часто цю умову формулюють як додатність усіх головних діагональних мінорів визначника Гурвиця.

Для системи першого порядку характеристичне рівняння має вигляд:

$$a_0\lambda + a_1 = 0.$$

Згідно з критерієм Гурвиця, така система буде стійкою, якщо обидва коефіцієнти характеристичного рівняння є додатними.

Характеристичне рівняння системи другого порядку:

$$a_0\lambda^2 + a_1\lambda + a_2 = 0.$$

Визначник Гурвиця $\Delta_2 = \begin{vmatrix} a_1 & 0 \\ a_0 & a_2 \end{vmatrix} = a_1 a_2$. Він є додатним, якщо коефіцієнти

a_1 і a_2 мають однакові знаки. Оскільки знаки коефіцієнтів a_1 і a_0 також повинні збігатися, то умовою стійкості є додатність усіх трьох коефіцієнтів характеристичного рівняння.

Характеристичне рівняння системи третього порядку:

$$a_0 \lambda^3 + a_1 \lambda^2 + a_2 \lambda + a_3 = 0.$$

Для її стійкості перші два коефіцієнти мають бути додатними. Із умови:

$$\Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} = a_1 a_2 - a_0 a_3 > 0$$

впливає додаткова вимога: $a_1 a_2 > a_0 a_3$.

Визначник

$$\Delta_3 = \begin{vmatrix} a_1 & a_3 & 0 \\ a_0 & a_2 & 0 \\ 0 & a_1 & a_3 \end{vmatrix} = a_3 \Delta_2.$$

Оскільки $\Delta_2 > 0$, то з умови $\Delta_3 > 0$ впливає вимога $a_3 > 0$.

Таким чином, система третього порядку є стійкою, якщо коефіцієнти a_0 , a_1 , a_3 – додатні й $a_1 a_2 > a_0 a_3$.

Складність використання критерію Гурвиця істотно зростає із збільшенням порядку характеристичного рівняння.

4.3. Критерій Рауса

Критерій Рауса є системою нерівностей, що складені за певними правилами з коефіцієнтів характеристичного рівняння замкненої системи. Для перевірки стійкості САУ будують таблицю, вигляд якої показано у табл. 4.1.

До її першого рядка записують коефіцієнти з парними індексами (a_0, a_2, a_4, \dots), а до другого – з непарними (a_1, a_3, a_5, \dots). До наступних рядків заносять коефіцієнти, які розраховують за формулою:

$$a_{ki} = a_{k+1\ i-2} - r_i a_{k+1\ i-1}, \quad (4.15)$$

де $r_i = \frac{a_{1\ i-2}}{a_{1\ i-1}}$, $i \geq 3$ – індекс, що позначає номер рядка, k – індекс, який позначає

номер стовпчика.

Таблиця 4.1

Допоміжна таблиця для перевірки критерію Рауса

Коефіцієнт, r	Номер рядка	Номер стовпчика		
		1	2	3
Номер стовпчика				
–	1	$a_{11} = a_0$	$a_{21} = a_2$	$a_{31} = a_4$
–	2	$a_{12} = a_1$	$a_{22} = a_3$	$a_{32} = a_5$
$r_3 = \frac{a_{11}}{a_{12}}$	3	$a_{13} = a_{21} - r_3 a_{22}$	$a_{23} = a_{31} - r_3 a_{32}$	$a_{33} = a_{41} - r_3 a_{42}$
$r_4 = \frac{a_{12}}{a_{13}}$	4	$a_{14} = a_{22} - r_4 a_{23}$	$a_{24} = a_{32} - r_4 a_{33}$	$a_{34} = a_{42} - r_4 a_{43}$

Кількість рядків табл. 4.1 дорівнює $n + 1$, де n – степінь характеристичного рівняння.

Для того, щоб САУ була стійкою, необхідно й достатньо, щоб елементи a_{1i} мали один і той самий знак. Часто цей критерій формулюють як вимогу $a_{1i} > 0$, оскільки перший коефіцієнт характеристичного рівняння зазвичай беруть додатним.

Кількість від'ємних коефіцієнтів a_{1i} дорівнює кількості коренів з додатною дійсною частиною. Якщо $a_n = 0$, то є нульовий корінь. Якщо будь-який із проміжних коефіцієнтів дорівнює нулю, то це свідчить про наявність пари суто уявних коренів.

Критерій Рауса є зручним для алгоритмізації і часто використовується при комп'ютерних розрахунках. Але його важко застосовувати для перевірки впливу окремих параметрів на стійкість системи й визначення запасу стійкості. Іншим недоліком є те, що критерій не є наочним й тому він не завжди є зручним для користування.

4.4. Критерії Михайлова та Найквіста

Нехай характеристичне рівняння системи має вигляд:

$$a_0 p^n + a_1 p^{n-1} + \dots + a_{n-1} p + a_n = f(p). \quad (4.16)$$

Розглядаючи f як комплекснозначну функцію, перетворимо поліном $f(p)$ до вигляду:

$$f(p) = f(i\omega) = U(\omega) + iV(\omega). \quad (4.17)$$

Для кожного значення ω функція $f(i\omega)$ є певним вектором у комплексній площині. Криву, що є геометричним місцем кінців вектора $f(i\omega)$ при змінюванні ω від нуля до нескінченності, називають **годографом Михайлова**. За його розташуванням на комплексній площині можна визначити стійкість системи.

Умова стійкості за Михайловим (**критерій Михайлова**) є такою. Динамічна система, що описується лінійним диференціальним рівнянням порядку n , є стійкою, якщо годограф функції $f(i\omega)$ при змінюванні ω від нуля до нескінченності послідовно обходить у додатному напрямку (проти часової стрілки) n квадрантів комплексної площини й іде до нескінченності. Приклади годографів стійких та нестійких систем наведено, відповідно, на рис. 4.6 і 4.7.

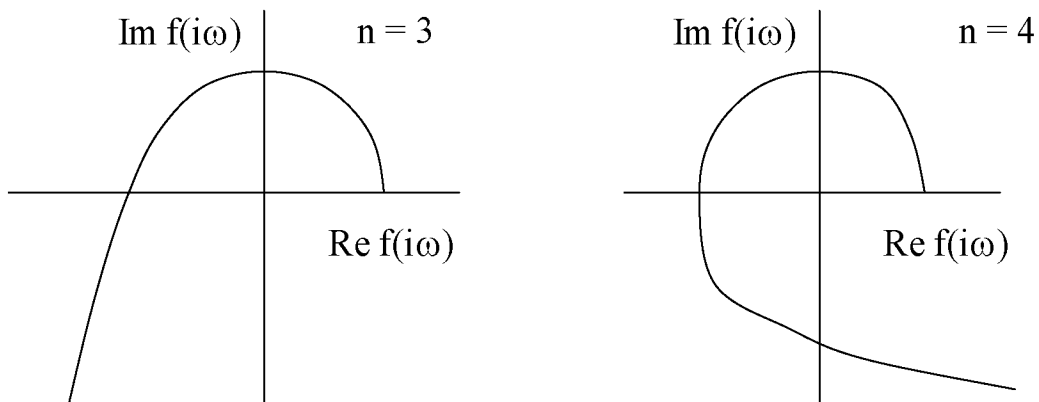


Рис. 4.6. Приклади характеристик стійких систем

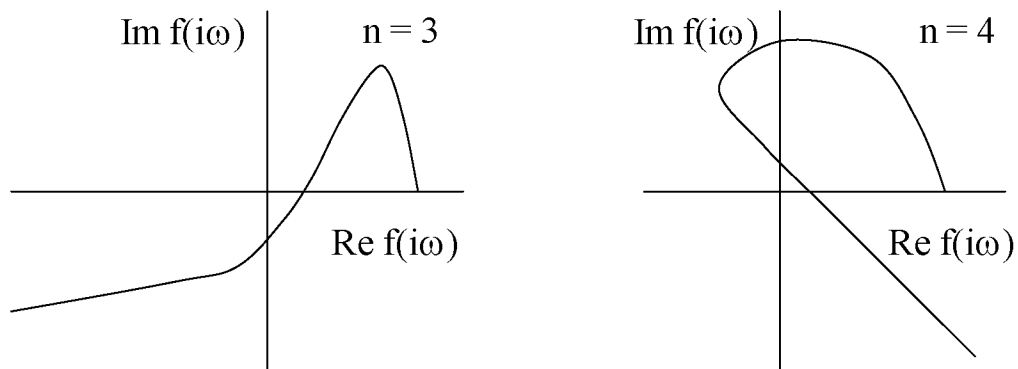


Рис. 4.7. Приклади характеристик нестійких систем

Запропонований Найквістом **амплітудно-фазовий критерій стійкості** дає змогу визначати стійкість системи в цілому за результатами дослідження лише її розімкненого стану, що істотно спрощує розрахунки. Згідно з ним, система, що є стійкою в розімкненому стані, буде стійкою і в замкненому стані,

якщо амплітудно-фазова характеристика розімкненої системи при змінюванні частоти від нуля до нескінченності не охоплює точку з координатами $(-1, 0)$ на комплексній площині.

Якщо ж система в розімкненому стані є нестійкою і має m коренів з додатною дійсною частиною, то для стійкості системи в замкненому стані амплітудно-фазова характеристика розімкненої системи при $\omega \in [0, +\infty)$ має $m/2$ разів охоплювати точку з координатами $(-1; 0)$, тобто приріст аргументу вектора $K(i\omega)$ повинен бути рівним $\varphi = m\pi$.

Приклад. Використовуючи розглянуті критерії, перевіримо стійкість системи, що має характеристичне рівняння:

$$p^3 + 50p^2 + 600p + 8000 = 0.$$

Запишемо це рівняння у вигляді:

$$(i\omega)^3 + 50(i\omega)^2 + 600(i\omega) + 8000 = 0.$$

Враховуючи, що $i^2 = -1$, отримаємо:

$$f(i\omega) = (8000 - 50\omega^2) + i(600\omega - \omega^3);$$

$$\operatorname{Re} f(i\omega) = 8000 - 50\omega^2; \operatorname{Im} f(i\omega) = 600\omega - 50\omega^3.$$

Результати обчислень наведено в табл. 4.2, а годограф Михайлова для досліджуваної системи показано на рис. 4.8.

Результати обчислень для перевірки стійкості системи
за критерієм Михайлова

ω	0	10	15	20	30	40	∞
Re $f(i\omega)$	8000	3000	-3250	-12000	-37000	-72000	$-\infty$
Im $f(i\omega)$	0	5000	5625	4000	-9000	-40000	$-\infty$

Годограф Михайлова послідовно проходить через перші три квадранти в додатному напрямку й іде до нескінченності. Тому система, що розглядається, є стійкою.

Перевіримо стійкість цієї системи за критерієм Гурвиця. Згідно з ним коефіцієнти a_0, a_1, a_3 повинні бути додатними і має виконуватися умова $a_1 a_2 > a_0 a_3$. Для прикладу, що розглядається: $a_0 = 1 > 0$; $a_1 = 50 > 0$; $a_3 = 8000 > 0$; $a_1 a_2 = 30000 > a_0 a_3 = 8000$. Таким чином, система є стійкою.

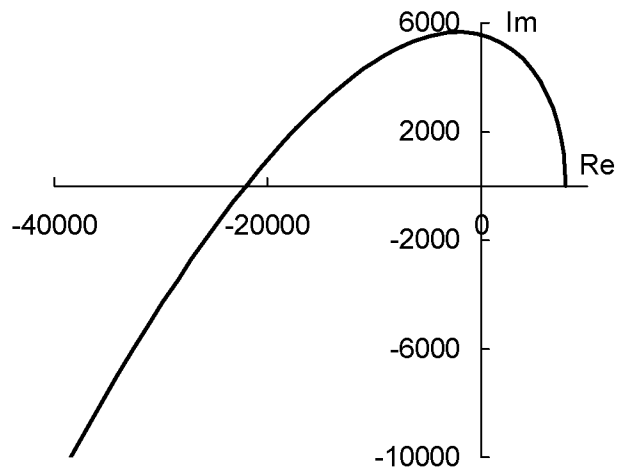


Рис. 4.8. Годограф Михайлова для системи, що розглядається у прикладі

Побудуємо табл. 4.3 для перевірки стійкості системи за критерієм Рауса.

Таблиця 4.3

Допоміжна таблиця для перевірки стійкості системи за критерієм Рауса

Коефіцієнт, r	Коефіцієнти Рауса		
–	1	600	0
–	50	8000	0
0,02	440	0	0
0,114	8000	0	0

Усі коефіцієнти Рауса в першому стовпчику є додатними, тому система є стійкою.

4.5. Дослідження впливу параметрів лінійних систем на їх стійкість

Усі критерії стійкості дають змогу при заданих параметрах системи робити висновки про її стійкість. За їх допомогою можна прослідкувати вплив деяких параметрів на стійкість системи, наприклад, визначити граничні значення коефіцієнтів підсилення системи й часу запізнення.

Для того, щоб дослідити вплив різних параметрів системи на її стійкість, розроблено спеціальні методи:

- аналіз переміщення коренів характеристичного рівняння у площині коренів й побудова кореневих годографів;
- побудова областей стійкості у просторі параметрів системи або коефіцієнтів характеристичного рівняння.

Метод кореневого годографа дає змогу оцінити вплив параметрів системи на розташування коренів характеристичного рівняння. Його було запропоновано К.Ф. Феодорчиком у 1948 р. Зазвичай розглядають вплив одного параметра (найчастіше, коефіцієнта підсилення) на розташування коренів.

Загальне рівняння, що описує динаміку лінійної системи можна подати у вигляді:

$$\begin{aligned} & (a_n p^n + a_{n-1} p^{n-1} + \dots + a_1 p + a_0) y(p) = \\ & + (b_m p^m + b_{m-1} p^{m-1} + \dots + b_1 p + b_0) x(p). \end{aligned} \quad (4.18)$$

Передаточною функцією відповідної системи є:

$$W(p) = \frac{y(p)}{x(p)} = \frac{b_m p^m + b_{m-1} p^{m-1} + \dots + b_1 p + b_0}{a_n p^n + a_{n-1} p^{n-1} + \dots + a_1 p + a_0} = \frac{B(p)}{A(p)}. \quad (4.19)$$

Будемо вважати, що $m \leq n$. Це справедливо для широкого класу фізично реалізованих систем, зокрема для систем автоматичного управління. Полюсами передаточної функції називають корені полінома знаменника функції (4.19), а її нулями – корені поліному її чисельника. Положення полюсів на комплексній площині визначає стійкість системи. Поліном $A(p)$ називають характеристичним поліномом передаточної функції.

Кореневим годографом називають геометричне місце полюсів передаточної функції на комплексній площині при змінюванні одного з параметрів системи від нуля до нескінченності. Метод кореневого годографа дає змогу визначати полюси й нулі передаточної функції замкненої системи за наявності інформації про полюси та нулі розімкненої системи. Під замкненою розуміється система на вхід якої подається її вихідний сигнал, узятий зі знаком "плюс" або "мінус". Тобто це вихідна система до якої додано зворотний зв'язок. Між передаточними функціями замкненої та розімкненої систем виконується співвідношення:

$$W(p) = \frac{W_p}{1 \pm W_p}. \quad (4.20)$$

Знак "+" у знаменнику відповідає негативному зворотному зв'язку, а знак "-" – позитивному.

За наявності негативного зворотного зв'язку для визначення полюсів замкненої системи необхідно розв'язати рівняння:

$$W_p(p) = -1. \quad (4.21)$$

Оскільки функція $W_p(p)$ є комплексозначною, то (4.21) розпадається на два рівняння – рівняння модулів та рівняння аргументів:

$$|W(p)| = 1; \quad (4.22)$$

$$\arg W(p) = \pm(2v+1)\pi, \quad v = 0, 1, 2, \dots \quad (4.23)$$

За наявності позитивного зворотного зв'язку замість (4.23) рівнянням аргументів буде:

$$\arg W(p) = \pm 2\pi v, \quad v = 0, 1, 2, \dots \quad (4.24)$$

Якщо певна точка p^* є полюсом замкненої системи, то проводячи до неї вектори з усіх нулів характеристичної функції розімкненої системи W_p (їх аргументи позначимо через θ_j^0), а також вектори з усіх її полюсів (їх аргументи позначимо як θ_i^*), із (4.23, 4.24) одержимо, відповідно:

$$\sum_{j=1}^m \theta_j^0 - \sum_{i=1}^n \theta_i^* = \pm(2v+1)\pi, \quad v = 0, 1, 2, \dots; \quad (4.25)$$

$$\sum_{j=1}^m \theta_j^0 - \sum_{i=1}^n \theta_i^* = \pm 2\pi v, \quad v = 0, 1, 2, \dots \quad (4.26)$$

Кути θ відраховуються від додатного напрямку дійсної осі проти годинникової стрілки.

Геометричне місце точок, що задовольняють умову (4.25) або (4.26) є кореневим годографом досліджуваної системи. Він має такі основні властивості (для випадку системи з негативним зворотним зв'язком).

1. Гілки кореневого годографа є неперервними й симетричними відносно дійсної осі на комплексній площині.

2. Кількість гілок дорівнює порядку системи n . Вони починаються в n полюсах розімкненої системи p_i^* . При зростанні коефіцієнта підсилення від нуля до нескінченності полюси замкненої системи рухаються вздовж гілок кореневого годографа.

3. Рівно m гілок кореневого годографа закінчуються в нулях розімкненої системи p_j^0 , а інші (якщо $m < n$) йдуть у нескінченність.

4. Асимптотою кореневого годографа є зірка, яка складається з $n - m$ променів, що виходять з однієї точки дійсної осі з координатою:

$$x_0 = \frac{\sum_{j=1}^m p_j^0 - \sum_{i=1}^n p_i^*}{n - m} \quad (4.27)$$

під кутами

$$\theta = \frac{2v + 1}{n - m} \pi, \quad v = 1, \dots, n - m - 1 \quad (4.28)$$

до дійсної осі.

Приклади таких зірок для $n - m = 1 \dots 4$ показано на рис. 4.9.

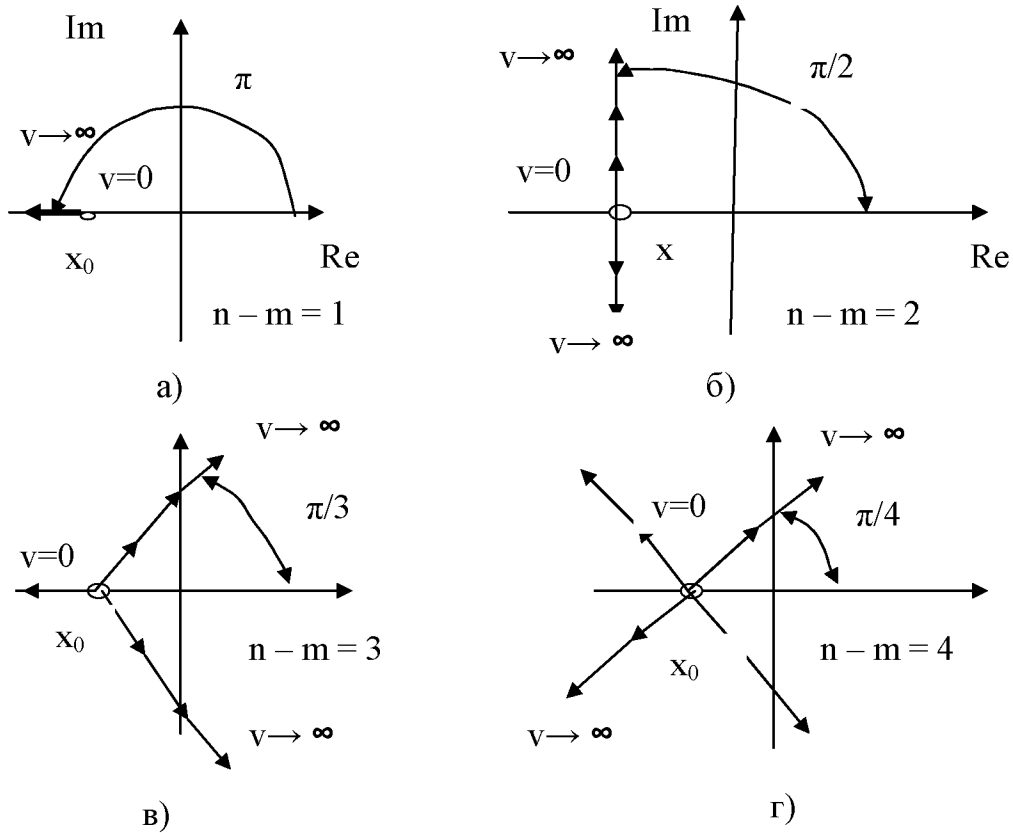


Рис. 4.9. Кореневі годографи різних типів

5. Система є стійкою, якщо гілки кореневого годографа розташовані у лівій півплощині. Замкнена система втрачає стійкість коли гілки кореневого годографа перетинають уявну вісь зліва направо. Відповідне значення коефіцієнта підсилення називають критичним. Цю властивість використовують при дослідженні впливу параметрів системи на її стійкість. Перетин годографа з уявною віссю може бути знайдено за допомогою одного з критеріїв стійкості.

Приклад. Розглянемо систему з передавальною функцією:

$$W(p) = \frac{k}{p(1+pT_1)(1+pT_2)} = \frac{P(p)}{Q(p)}$$

при змінюванні k від 0 до ∞ , якщо $T_1 = 0,1$ с і $T_2 = 0,02$ с.

У цьому випадку: $n = 3$, $m = 0$, $p_1^* = 0$, $p_2^* = -1/T_1 = -10$, $p_3^* = -1/T_2 = -50$.
Оскільки $n - m = 3$, то асимптоти кореневого годографа утворюють трипроменеву зірку з центром у точці:

$$x_0 = (-10 - 50)/3 = -20.$$

Для знаходження перетину з уявною віссю складемо таблицю Рауса. Спочатку запишемо рівняння системи в операторній формі:

$$p^3 + 60p^2 + 500p + 500k = 0$$

або

$$0,002p^3 + 0,12p^2 + p + k = 0.$$

$$\text{Звідси: } a_0 = 0,002; a_1 = 0,12; a_2 = 1; a_3 = k; c_{13} = 1 - \frac{0,002}{0,12}k; c_{14} = k.$$

$$\text{Умовою перетину з уявною віссю є: } c_{13} = 1 - \frac{0,002}{0,12}k, \text{ звідки маємо } k = 60.$$

Величину ω при перетині можна визначити за формулою:

$$\omega^2 = \frac{a_3}{c_{12}} = \frac{a_3}{a_1} = \frac{k}{0,12} = 500; \omega = 10\sqrt{5} \approx 22,5.$$

Кореневий годограф системи, що розглядається, показано на рис. 4.10.

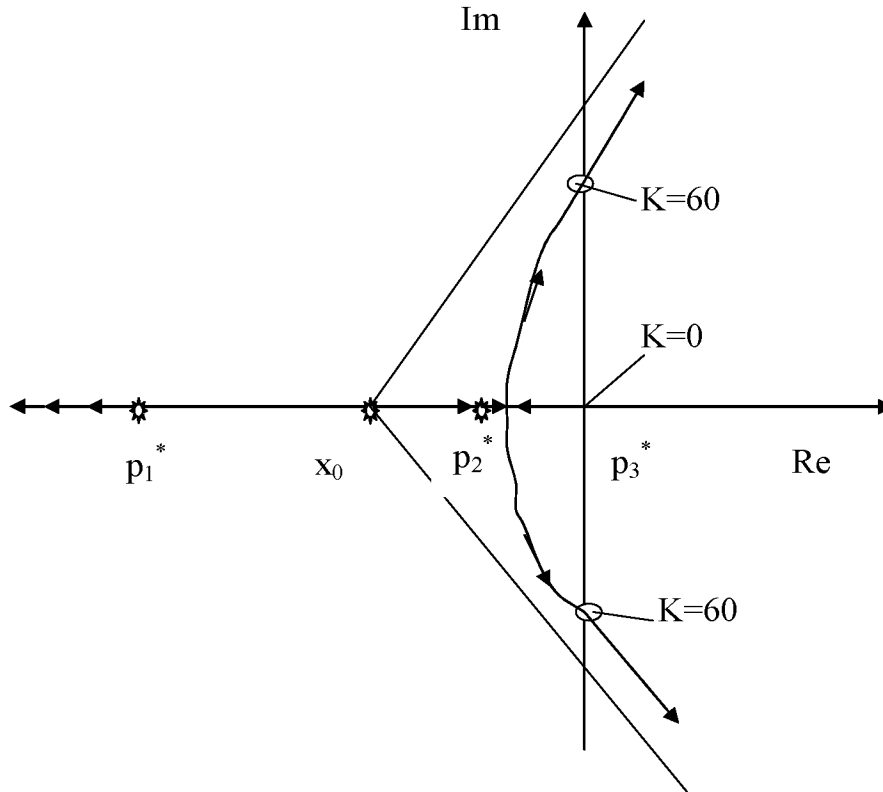


Рис. 4.10. Кореневий годограф для прикладу, що розглядається

Діаграма Вишнеградського. І.А. Вишнеградський у 1876 р. заснував класичну теорію автоматичного управління, дослідив системи третього порядку. Він описав умови стійкості для таких систем.

Нехай САУ описується рівнянням:

$$a_0 \frac{d^3 y}{dt^3} + a_1 \frac{d^2 y}{dt^2} + a_2 \frac{dy}{dt} + a_3 y = 0. \quad (4.29)$$

Характеристичне рівняння для (4.29) має вигляд:

$$a_0 p^3 + a_1 p^2 + a_2 p + a_3 = 0 \quad (4.30)$$

або

$$p^3 + c_1 p^2 + c_2 p + c_3 = 0. \quad (4.30a)$$

Коефіцієнти у рівняннях (4.29) і (4.30) для замкненої системи:

$$a_0 = T_1 T_2 T_3; a_1 = T_3 (T_1 + T_2); a_2 = T_3; a_3 = k. \quad (4.31)$$

На основі аналітичного дослідження І.А. Вишнеградський побудував діаграму, яка дає можливість, не розв'язуючи рівняння системи, а знаючи лише його числові параметри, визначити не тільки стійкість системи, а й форму перебігу процесу і прослідкувати за її зміною при змінюванні тих або інших параметрів процесу.

Зробимо підстановки: $Z = \frac{p}{\sqrt[3]{c_3}}$, $X = \frac{c_1}{\sqrt[3]{c_3}}$, $Y = \frac{c_2}{\sqrt[3]{c_3^2}}$. Тоді рівняння (4.30а)

набуває вигляду:

$$Z^3 + XZ^2 + YZ + 1 = 0. \quad (4.32)$$

Якщо система знаходиться на межі стійкості, то рівняння (4.30а) має один дійсний та два чисто уявних кореня, які позначимо: $p_1 = -\alpha$; $p_{2,3} = \pm i\beta$. Тоді з (4.30а) маємо:

$$\begin{aligned} p^3 + c_1 p^2 + c_2 p + c_3 &= (p + \alpha)(p - i\beta)(p + i\beta) = \\ &= p^3 + \alpha p^2 + \beta^2 p + \alpha \beta^2. \end{aligned}$$

Звідси, прирівнюючи коефіцієнти при однакових степенях p одержимо: $c_1 = \alpha$; $c_2 = \beta^2$; $c_3 = \alpha \beta^2$. Тоді $c_1 c_2 = c_3$. Виражаючи коефіцієнти через X та Y , отримаємо:

$$XY = 1, \quad (4.33)$$

що є рівнянням гіперболи, яка є межею стійкості досліджуваної системи. Її називають **гіперболою Вишнеградського**. Умовою стійкості системи є виконання нерівності:

$$XY > 1. \tag{4.34}$$

Якщо $XY < 1$, система є нестійкою. Якщо ж умова (4.34) виконується, то можливі різні режими встановлення рівноваги, деякі з котрих показано на рис. 4.11.

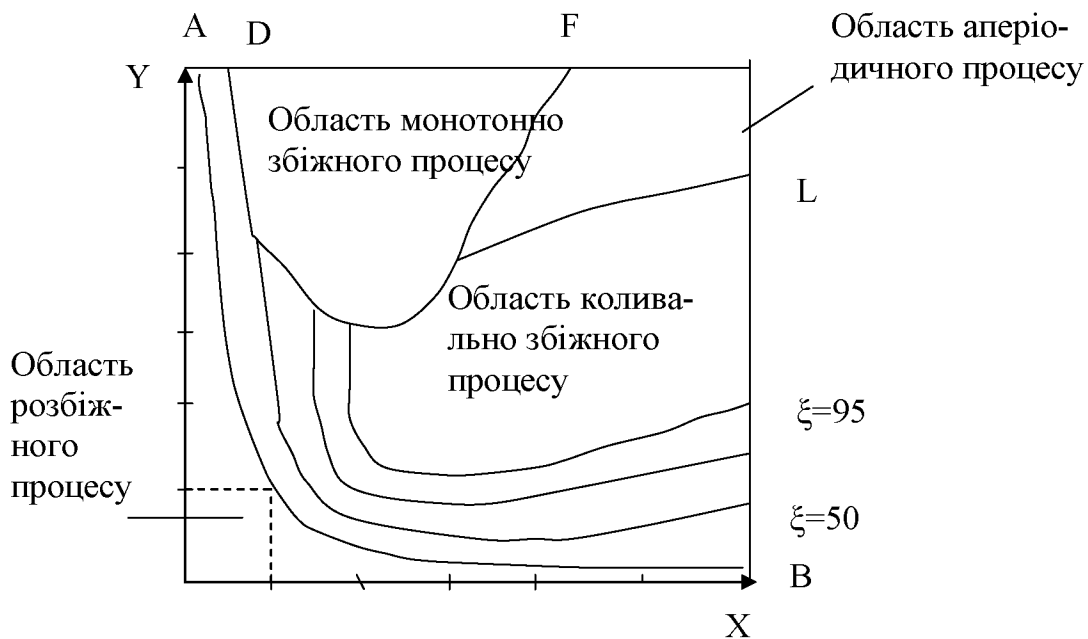


Рис. 4.11. Діаграма Вишнеградського, доповнена кривими рівного згасання

Метод Вишнеградського був розвинений М.А. Айзерманом, який доповнив діаграму кривими згасання. Він також увів поняття ступеня згасання:

$$\xi = \psi = 100 \frac{y_1 - y_2}{y_1}, \tag{4.35}$$

де y_1 та y_2 – відповідно, амплітуди першого й другого періодів коливань.

Контрольні питання

1. Які системи називають стійкими?
2. За яких припущень можна використовувати формулу (4.1).
3. Навести приклади вільних і змушених рухів реальних систем.
4. Чи залежить стійкість системи від характеру її змушених рухів? Чому?
5. Якими мають бути вільні рухи системи для того, щоб ця система була стійкою?
6. Яким є загальний вигляд рівняння вільного руху лінійної системи?
7. Якими мають бути корені характеристичного рівняння системи, щоб вона була стійкою?
8. Якими мають бути корені характеристичного рівняння, щоб система була нестійкою?
9. Перевірити стійкість заданої системи, використовуючи першу теорему Ляпунова.
10. Перевірити стійкість заданої системи, використовуючи критерій Гурвиця.
11. Перевірити стійкість заданої системи, використовуючи критерій Рауса.
12. У чому полягають основні переваги й недоліки критеріїв Рауса й Гурвиця?
13. Перевірити стійкість заданої системи, використовуючи критерій Михайлова.
14. У чому полягає критерій Найквіста перевірки стійкості лінійних систем?
15. Що називають кореневим годографом системи?

16. Як розташований кореневий годограф стосовно дійсної та уявної осей?

17. Скільки гілок містить кореневий годограф? Звідки виходять і де закінчуються ці гілки?

18. Який вигляд має асимптотичний кореневий годограф при $\nu \rightarrow \infty$?

19. Як визначають перетин кореневого годографа з уявною віссю?

20. Побудувати схематично кореневий годограф для системи, що задана передаточною функцією.

21. Як будують діаграму Вишнеградського?

22. Побудувати діаграму Вишнеградського для заданої системи.

5. СТАТИСТИЧНЕ МОДЕЛЮВАННЯ СИСТЕМ

5.1. Метод Монте-Карло

Важливим класом завдань математичного моделювання є побудова статистичних моделей. Для їх вирішення широко використовують **метод Монте-Карло (метод статистичних випробувань)**. Він є обчислювальним методом розв'язання математичних задач за допомогою випадкових величин. Його особливістю є проста структура алгоритму обчислень. Загальна схема реалізації цього методу передбачає розбиття досліджуваного процесу на елементарні події, для яких можна встановити необхідні ймовірнісні характеристики. Після цього здійснюють багаторазову прогонку елементарних подій, виконують частотний аналіз та обчислюють необхідні характеристики процесу, що вивчається. Іншою особливістю методу є його повільна збіжність. Як правило, вона пропорційна $\sqrt{1/N}$, де N – кількість випробувань. Тому ефективність методу тим вища, чим більшою є допустима похибка розрахунків. Зазвичай його використовують у випадках, коли допустима похибка не нижча за 1%. Типовою помилкою при застосуванні методу Монте-Карло є прийняття висновків (особливо при кількісних дослідженнях) за недостатньої кількості випробувань.

Надмірно велика кількість випробувань може призводити до помилок іншого типу. Змістові висновки при математичному моделюванні мають бути наслідками припущень вихідної змістової моделі. Одержувані ж при великій кількості випробувань результати можуть виявитися наслідками певних граничних теорем, що виконуються за будь-яких вихідних припущень.

Як правило, для однієї і тієї самої задачі можна розробити багато варіантів реалізації методу Монте-Карло, в яких моделюють різні випадкові величини. Це створює можливість обрання оптимального для дослідника варіанта з погляду очікуваних похибки розрахунків та потрібних ресурсів.

Найчастіше повторні випробування при незмінних умовах експерименту здійснюють для знаходження математичних сподівань випадкових величин. З імовірністю α математичне сподівання m_x знаходиться в інтервалі

$[\bar{x} - \varepsilon; \bar{x} + \varepsilon]$, де $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$ – середнє значення, x_i – реалізація випадкової ве-

личини X у i -му випробуванні, N – кількість випробувань, $\varepsilon = t_\alpha \sigma_x$, $\sigma_x^2 = \frac{\sigma_X^2}{N}$ –

дисперсія \bar{x} , $\sigma_x^2 \approx S^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2$ – дисперсія випадкової величини X , S –

емпірична оцінка дисперсії. Звідси маємо: $\varepsilon = \frac{t_\alpha \sigma_X}{\sqrt{N}}$. Величини t_α знаходять із

таблиць розподілу Стюдента для заданого рівня значущості α . Оскільки точ-

ність, з якою необхідно визначити математичне сподівання, задається, то для

оцінювання потрібної кількості опитів можна використовувати вираз $N = \frac{t_\alpha^2 \sigma_X^2}{\varepsilon^2}$,

або $N = \frac{t_\alpha^2 S^2}{\varepsilon^2}$ у разі, коли істинна дисперсія σ_x^2 є невідомою.

Метод Монте-Карло застосовують також для знаходження невідомих ймовірностей p настання певних випадкових подій. Для цього в N випробуваннях реєструють, скільки разів відбулася досліджувана подія, визначають її частоту d , яку й беруть за оцінку p . При заданих ε та α довірчим інтервалом є

$[d - t_\alpha \sigma_d; d + t_\alpha \sigma_d]$. Дисперсія $\sigma_d^2 = \frac{p(1-p)}{N} \approx \frac{d(1-d)}{N}$. Точність ε оцінки ймові-

рності p дорівнює $\varepsilon = t_\alpha \sqrt{\frac{d(1-d)}{N}}$. Відповідно, кількість випробувань, що треба

виконати для забезпечення точності ε при рівні значущості α ,

$N = \frac{t_\alpha^2 p(1-p)}{\varepsilon^2} \approx \frac{t_\alpha^2 d(1-d)}{\varepsilon^2}$. Звідси випливає, що для підвищення точності необ-

хідно значно збільшити кількість випробувань.

Іноколи можна досягти тієї самої точності при меншій кількості випробувань. Нехай, наприклад, досліджувана частота є функцією якогось параметра системи χ : $p = f(\chi)$. Унаслідок дії закону великих чисел можна очікувати, що із збільшенням N емпірична залежність $d = f(\chi)$ буде ставати все гладкішою. То-

ді, використовуючи відносно малу кількість випробувань, можна одержати емпіричні точки й потім побудувати апроксимуючу функцію, застосування якої дасть змогу зменшити кількість випробувань, потрібних для знаходження досліджуваної ймовірності p із заданою точністю.

5.2. Алгоритми генерування псевдовипадкових послідовностей, що відповідають рівномірному розподілу

Для реалізації методу Монте-Карло необхідно використовувати різноманітні алгоритми генерації послідовностей випадкових чисел та векторів. Як правило, вони передбачають два основних етапи. На першому етапі будують псевдовипадкову послідовність, що має рівномірний розподіл на відрізку $[0,1]$, а на другому – перетворюють її у послідовність, яка відповідає заданому закону розподілу. Тому розглянемо спочатку алгоритми генерування рівномірних випадкових послідовностей.

З курсу теорії ймовірностей і математичної статистики відомо, що щільність розподілу для випадкової величини, яка має рівномірний розподіл на відрізку $[a, b]$, задається виразом:

$$f(x) = \begin{cases} 0, & \text{якщо } x < a, \\ \frac{1}{b-a}, & \text{якщо } a \leq x \leq b, \\ 0, & \text{якщо } x > b. \end{cases} \quad (5.1)$$

Математичне сподівання цієї випадкової величини $m_x = \frac{a+b}{2}$, а її дисперсія $\sigma_x^2 = \frac{(b-a)^2}{12}$. Найчастіше у практиці використовують випадкові величини, що мають рівномірний розподіл на відрізках $[0, 1]$ або $[-1, 1]$. Для цих випадків щільність розподілу та його моменти дорівнюють відповідно:

$$f(x) = \begin{cases} 0, & \text{якщо } x < 0, \\ \frac{1}{b-a}, & \text{якщо } 0 \leq x \leq 1, \quad m_x = 0,5; \quad \sigma_x^2 = 1/12; \\ 0, & \text{якщо } x > 1, \end{cases} \quad (5.2)$$

та

$$f(x) = \begin{cases} 0, & \text{якщо } x < -1; \\ \frac{1}{b-a}, & \text{якщо } -1 \leq x \leq 1; \quad m_x = 0; \quad \sigma_x^2 = 1/3; \\ 0, & \text{якщо } x > 1. \end{cases} \quad (5.3)$$

Випадкову величину, що має рівномірний розподіл на відрізку $[0, 1]$ далі будемо позначати літерою ξ . Ймовірність її потрапляння до певного інтервалу, що лежить усередині відрізка $[0, 1]$,

$$P(\alpha \leq \xi \leq \beta) = \int_{\alpha}^{\beta} f(\xi) d\xi = \int_{\alpha}^{\beta} d\xi = \beta - \alpha, \quad (5.4)$$

тобто дорівнює довжині цього інтервалу. Цю властивість можна використовувати як необхідну й достатню умову того, що деяка випадкова величина має рівномірний розподіл на відрізку $[0, 1]$.

Генерування чергових реалізацій рівномірно розподіленої випадкової величини ξ можна здійснювати за допомогою такого перетворення:

$$\xi = z_1 \cdot 2^{-1} + z_2 \cdot 2^{-2} + \dots + z_i \cdot 2^{-i} + \dots, \quad (5.5)$$

де z_i – реалізація випадкової величини z , яка з рівною ймовірністю $p_z = 0,5$ може набувати значення 0 або 1.

Випадкова величина ξ може мати нескінченну кількість реалізацій. Але при запису чисел у ЕОМ використовується скінченна кількість двійкових розрядів. Тому кількість реалізацій випадкової величини, які можна використовувати при моделюванні, також буде скінченною. Максимальна кількість випадкових чисел, що записуються за допомогою k двійкових розрядів і не збігаються одне з одним, дорівнює 2^k . Із сукупності чисел $0, 1, 2, \dots, 2^k - 1$ можна одержати такі значення дискретної випадкової величини ξ :

$$\xi_i = \frac{i}{2^k - 1} \quad (i = 0, 1, 2, \dots, 2^k - 1). \quad (5.6)$$

Їх імовірності дорівнюють $p_i = 1/2^k$. Розподіл такої випадкової величини є квазірівномірним. Її моменти дорівнюють:

$$m_1 = \sum_{i=0}^{2^k-1} \frac{i}{2^k - 1} \cdot \frac{1}{2^k} = \frac{1}{2}; \quad \sigma_{\xi}^2 = \sum_{i=0}^{2^k-1} \frac{1}{2^k} \left(\frac{i}{2^k - 1} - \frac{1}{2} \right)^2 = \frac{1}{12} \cdot \frac{2^k + 1}{2^k - 1} \approx \frac{1}{12}. \quad (5.7)$$

Для генерування рівномірно розподілених послідовностей випадкових чисел розроблено три способи: табличний, фізичний та програмний. У першому випадку використовують таблиці випадкових величин, одержаних за допомогою фізичного або програмного генератора випадкових чисел. Такі таблиці можуть містити велику кількість випадкових чисел і зберігатися на паперових або електронних носіях. Їх використовують при ручних розрахунках, а також при налагодженні програм та перевірці особливо важливих результатів.

Для фізичної генерації випадкових чисел можна використовувати довільний випадковий фізичний процес. Найчастіше користуються джерелами радіоактивного випромінювання або власними шумами електронних компонентів ЕОМ. У першому випадку задають певний проміжок часу Δt . Потім підрахову-

ють кількість k частинок, випромінених за проміжок часу Δt_i . Значення ξ_i розраховують по формулі 5.5. При цьому беруть:

$$z_i = \begin{cases} 0, & \text{якщо } k - \text{ непарне,} \\ 1, & \text{якщо } k - \text{ парне.} \end{cases}$$

Проміжок Δt має бути достатньо великим, для того щоб імовірності одержання парних і непарних значень k були рівними з достатнім ступенем точності.

Якщо використовують шуми електронних компонентів, то необхідно спочатку побудувати залежність вихідної напруги від часу за достатньо великий проміжок часу (рис. 5.1).

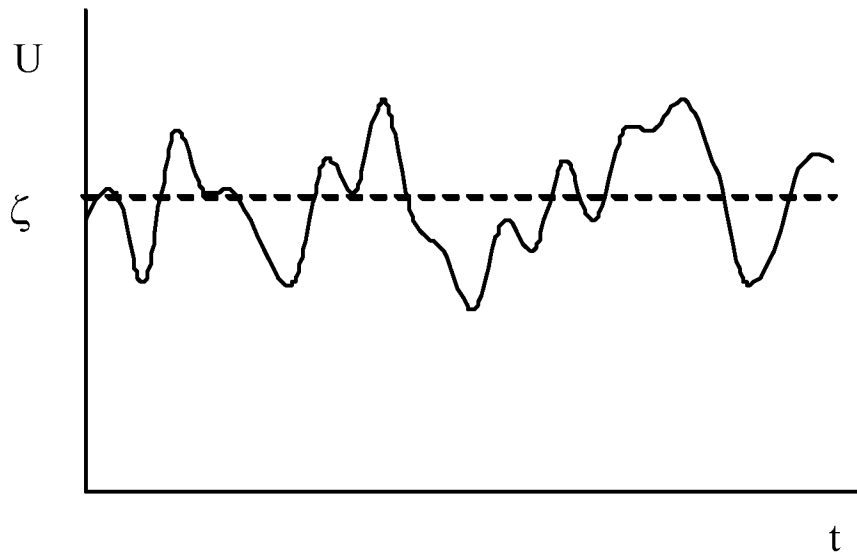


Рис. 5.1. Схематична залежність вихідної напруги від часу, що використовується для генерування випадкових чисел

Після цього знаходять рівень відтинання ζ , для якого виконується умова $P(U > \zeta) = P(U < \zeta) = 0,5$. Значення випадкової величини z_i визначають через рівні проміжки часу Δt за умовою:

$$z_i = \begin{cases} 0, & \text{якщо } U(t_i) < \zeta, \\ 1, & \text{якщо } U(t_i) > \zeta. \end{cases}$$

Зазвичай використовують m таких генераторів, що працюють паралельно (m – кількість двійкових розрядів). Тоді на кожному такті здійснюється генерація одного m -розрядного числа. Одержувані таким методом послідовності можуть відрізнятися від рівномірної випадкової послідовності, якщо на вихідну характеристику впливають деякі неконтрольовані зовнішні фактори. Для усунення цієї проблеми необхідно періодично корегувати кожен з генераторів.

При використанні програмного методу генерування випадкових чисел значення $i + 1$ -го випадкового числа визначають, використовуючи значення i -го числа за рекурентною формулою $\xi_{i+1} = f(\xi_i)$, тому одержувані числа не є дійсно випадковими. Їх називають псевдовипадковими числами. Функція f має бути достатньо складною для того, щоб взаємозв'язок сусідніх елементів послідовності ξ_i та ξ_{i+1} не впливав на результати. Існує багато алгоритмів, які використовуються для генерації псевдовипадкових послідовностей рівномірно розподілених чисел. Розглянемо два з них.

Алгоритм фон Неймана. Нехай ξ_k є m -розрядним двійковим числом, що задовольняє умову $0 < \xi_k < 1$ і має вигляд:

$$\xi_k = \varepsilon_1 \cdot 2^{-1} + \varepsilon_2 \cdot 2^{-2} + \dots + \varepsilon_m \cdot 2^{-m}. \quad (5.8)$$

Його квадрат може бути записаний у вигляді:

$$\xi_k^2 = \delta_1 \cdot 2^{-1} + \delta_2 \cdot 2^{-2} + \dots + \delta_{2m} \cdot 2^{-2m}. \quad (5.9)$$

Будемо вважати m парним числом. Така умова завжди виконується для сучасних ЕОМ. Наступне псевдовипадкове число ξ_{k+1} одержимо, використовуючи коефіцієнти середніх членів ряду 5.9, по формулі:

$$\xi_{k+1} = f(\xi_k) = \delta_{\frac{m}{2}+1} \cdot 2^{-1} + \delta_{\frac{m}{2}+2} \cdot 2^{-2} + \dots + \delta_{\frac{3m}{2}} \cdot 2^{-m}. \quad (5.10)$$

Одержувана за таким алгоритмом послідовність псевдовипадкових чисел за своїми властивостями близька до рівномірної випадкової послідовності, але кількість малих чисел в ній є дещо вищою, ніж це має бути для останньої.

Метод остач. У цьому випадку для одержання рівномірної послідовності псевдовипадкових чисел використовують рекурентне співвідношення $\xi_{k+1} = \{M\xi_k\}$, де хвилястими дужками позначено дробову частину числа. Як початкове значення ξ_0 можна взяти число 2^{-m} , де m є числом двійкових розрядів у мантиї комірки ЕОМ. M має бути достатньо великим цілим числом. Рекомендується використовувати $M = 5^{2p+1}$, де p є максимальним із цілих чисел, для яких виконується умова $5^{2p+1} < 2^m$.

У таблиці 5.1 вказано деякі переваги та недоліки основних методів генерування випадкових чисел.

Після одержання рівномірної випадкової послідовності необхідно перевірити її якість. Використовують такі методи перевірки.

Перевірка за моментами розподілу. Для її здійснення розраховують моменти одержаного розподілу:

$$m_\xi = \frac{1}{N} \sum_{i=1}^N \xi_i ; \quad \sigma_i^2 = \frac{1}{N} \sum_{i=1}^N \xi_i^2 - m_\xi^2. \quad (5.11)$$

Переваги й недоліки основних методів генерування випадкових чисел

Метод	Переваги	Недоліки
Табличний	<ol style="list-style-type: none"> 1. Висока швидкість одержання чисел, якщо таблиця записана в оперативному запам'ятовуючому пристрої (ОЗП). 2. Можливість здійснення повторних розрахунків. 3. Однократна перевірка якості рівномірної випадкової послідовності. 	<ol style="list-style-type: none"> 1. Потребує великого обсягу пам'яті в ОЗП. 2. Обмежений запас чисел. 3. Необхідність зовнішньої пам'яті.
Фізичний	<ol style="list-style-type: none"> 1. Висока швидкість генерування чисел. 2. Не потребує місця в ОЗП. 3. Необмежений запас випадкових чисел. 	<ol style="list-style-type: none"> 1. Неможливість повторних розрахунків. 2. Необхідно періодично корегувати датчик. 3. Потрібен спеціальний пристрій.
Програмний	<ol style="list-style-type: none"> 1. Потребує мало місця в ОЗП. 2. Можливість здійснення повторних розрахунків. 3. Однократна перевірка якості рівномірної випадкової послідовності. 4. Не потребує зовнішніх пристроїв. 	<ol style="list-style-type: none"> 1. Порівняно низька швидкість генерації рівномірної випадкової послідовності. 2. Обмежений запас випадкових чисел.

Для послідовності псевдовипадкових чисел, що розподілені на відрізок $[0,1]$ їх значення мають бути близькими відповідно до $1/2$ та $1/12$.

Перевірку гіпотези про рівномірний розподіл можна здійснити за допомогою критерію χ^2 , порівнюючи частоти потрапляння елементів отриманої послідовності до певних інтервалів з їх очікуваними значеннями, або за критерієм згоди Колмогорова-Смірнова, що ґрунтується на визначенні найбільшого за модулем відхилення емпіричної функції розподілу від теоретичної.

Вказані критерії розглядалися в курсі теорії ймовірностей та математичної статистики. Зауважимо лише, що при обранні рівня значущості слід враховувати, яка помилка для конкретної задачі є більш суттєвою – прийняття помилкової гіпотези чи неприйняття правильної гіпотези.

Перевірка за гістограмою. У цьому випадку відрізок $[0, 1]$ розбивають на n рівних частин. Потім підраховують емпіричні частоти потрапляння чисел, що генеруються, в одержані інтервали: $p_i = \frac{m_i}{N}$ (m_i – кількість чисел, що потрапили до i -го інтервалу, N – загальна кількість чисел у послідовності). Після цього будують гістограму розподілу. За великих N гістограма має наближатися до горизонтальної прямої $p = n/N$. Кількість відрізків розбиття та розмір послідовності визначаються конкретною задачею. У більшості випадків рекомендується обирати $n = 20 \dots 50$ та $N = (10^2 \dots 10^3)n$.

Перевірка на періодичність. Випадкові числа, що генеруються за програмним методом, рано чи пізно починають повторюватися. Якщо серед елементів послідовності $\xi_0, \xi_1, \xi_2, \dots, \xi_{q-1}$ немає однакових чисел, а елемент ξ_q збігається з одним із попередніх, то q називають відрізком аперіодичності. При використанні рівномірних випадкових послідовностей він має бути не меншим, ніж кількість чисел, потрібна для здійснення розрахунків. Наявність періодичності й величину періоду можна встановити при дослідженні автокореляційної функції отриманої послідовності. Її значення

$$\rho(\tau) = \frac{\frac{1}{n-\tau} \sum_{k=1}^{n-\tau} (\xi_k - \bar{\xi})(\xi_{k+\tau} - \bar{\xi})}{\frac{1}{n} \sum_{k=1}^n (\xi_k - \bar{\xi})^2} \quad (5.12)$$

характеризують ступінь статистичного зв'язку між значеннями досліджуваної послідовності, відстань між якими дорівнює τ . Свідченням періодичності у цьому разі буде наявність максимумів цієї функції, розташованих на однакої відстані один від одного. Для побудови графіків автокореляційних функцій зручно використовувати спеціальні пакети статистичного аналізу, зокрема пакети SPSS та Statistica. На рис. 5.2 показано приклад графіка автокореляційної функції для послідовності, період якої дорівнює 10.

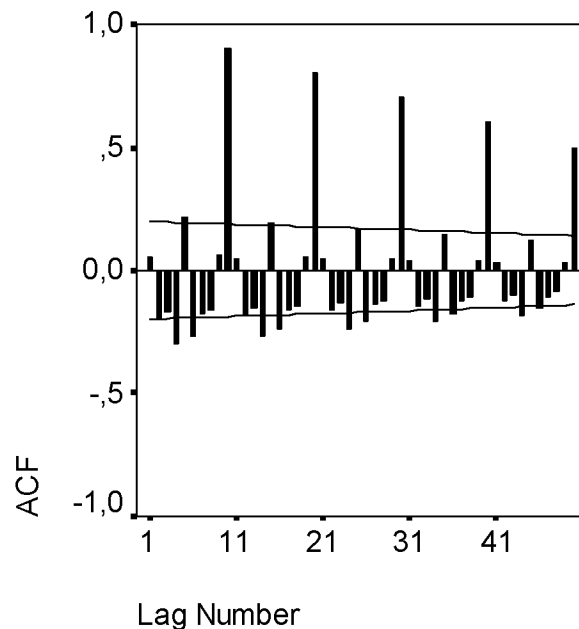


Рис. 5.2. Схематичний вигляд автокореляційної функції для періодичної послідовності

Перевірка на випадковість. Серед елементів послідовності, що генерується, може знайтись серія елементів $\xi_{k+1}, \xi_{k+2}, \dots, \xi_{k+\tau}$, які знаходяться в одній половині відрізка $[0, 1]$, і при цьому елементи ξ_k та $\xi_{k+\tau+1}$ розміщуються в іншій

половині цього відрізка. При використанні такої послідовності необхідно, щоб для кількості N чисел, що використовуються в розрахунках, виконувалася умова $r_{\max} \ll N$.

Існує багато типів відхилень від випадковості й критеріїв, що перевіряють їх наявність. Один з найпростіших тестів передбачає підрахунок кількості серій із двох, трьох, чотирьох і т. д. послідовних елементів, що розташовані на відрізках $[0; 0,5]$ і $[0,5; 1]$. Потім ці кількості порівнюють з очікуваними за умови випадковості значеннями за допомогою χ^2 -критерію.

Інший тест ґрунтується на застосуванні критерію Дарбіна – Уотсона, що перевіряє наявність автокореляції першого порядку. Його значення розраховують за формулою:

$$d = \frac{\sum_{k=2}^n (\xi_k - \xi_{k-1})^2}{\sum_{k=1}^n \xi_k^2}. \quad (5.13)$$

Можна показати, що

$$d \approx 2(1 - \rho_1), \quad (5.14)$$

де ρ_1 – коефіцієнт автокореляції першого порядку (коефіцієнт кореляції між послідовностями елементів $\xi_1, \xi_2, \dots, \xi_{n-1}$ та $\xi_2, \xi_3, \dots, \xi_n$). Звідси випливає, що за наявності додатної автокореляції значення d наближається до нуля, від'ємної – до чотирьох, а за її відсутності – до двох.

Більш складні тести передбачають дослідження властивостей автокореляційної та частинної автокореляційної функції досліджуваної послідовності. В окремих випадках вони дають змогу не тільки встановити факт випадковості, а й визначити тип випадкового процесу.

Перевірка у роботі. Як підсумковий тест для перевірки якості генератора випадкових чисел будують імітаційну модель, результат роботи якої є відомим з теорії. Порівняння результатів моделювання з теоретичними висновками дає змогу зробити висновок про придатність генератора для розв'язування певного класу задач.

5.3. Загальні методи генерування випадкових послідовностей із заданими законами розподілу

На практиці часто виникає необхідність генерування послідовностей випадкових чисел з довільними законами розподілу. Для цього застосовують різні методи перетворення рівномірних випадкових послідовностей у послідовності з іншими законами розподілу. У найпростіших випадках це можна зробити, використовуючи таку властивість: якщо ξ – випадкова величина, рівномірно розподілена на відрізку $[0, 1]$, то випадкова величина X , що є розв'язком рівняння

$$\int_{-\infty}^x f(x)dx = \xi, \quad (5.15)$$

має щільність розподілу $f(x)$.

Нехай треба одержати послідовність чисел, рівномірно розподілених на відрізку $[a, b]$. Тоді необхідно для кожного елемента ξ_i рівномірної випадкової послідовності знайти розв'язок рівняння:

$$\xi_i = \int_{-\infty}^{x_i} f(x)dx = \int_a^{x_i} \frac{dx}{b-a} = \frac{x_i - a}{b-a}. \quad (5.16)$$

Звідси маємо:

$$x_i = \xi_i (b - a) + a. \quad (5.17)$$

Для послідовності чисел, що відповідає експоненціальному розподілу:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{якщо } 0 \leq x < \infty; \\ 0, & \text{якщо } x < 0. \end{cases} \quad (5.18)$$

Звідси:

$$\xi_i = \int_{-\infty}^x f(x) dx = \int_0^{x_i} \lambda e^{-\lambda x} dx = 1 - e^{-\lambda x_i} \quad (5.19)$$

та

$$x_i = -\frac{1}{\lambda} \ln(1 - \xi_i). \quad (5.20)$$

Якщо величина ξ_i має рівномірний розподіл на $[0, 1]$, величина $1 - \xi_i$ також буде мати рівномірний розподіл на $[0, 1]$. Тому замість виразу (5.20) можна використовувати простішу формулу:

$$x_i = -\frac{1}{\lambda} \ln \xi_i. \quad (5.21)$$

Описаний спосіб не завжди є прийнятним. Як правило, із рівняння 5.15 неможливо одержати формулу для знаходження шуканого значення x в явному вигляді. Звичайно, це рівняння можна розв'язати за допомогою чисельних методів, але це веде до істотного збільшення обсягу розрахунків. Тому більш доцільним у таких випадках є використання інших, більш ефективних методів генерування послідовностей випадкових чисел, що мають задану щільність розподілу $f(x)$.

Один із них (**метод відбору**) ґрунтується на застосуванні такої властивості. Нехай ξ_1 та ξ_2 – це чергові числа, які належать до рівномірної випадкової послідовності, заданої на відрізку $[0, 1]$. Зробимо перетворення:

$$x' = \xi_1 (b - a) + a; \quad y' = \xi_2 c. \quad (5.22)$$

Випадкові величини x' та y' мають рівномірний розподіл на відрізках $[a, b]$ і $[0, c]$ відповідно. Можна довести, що випадкова величина, яка визначається умовою: $x = x'$, якщо $y' < f(x')$, має щільність розподілу $f(x)$. Тому для генерування випадкової послідовності зі щільністю розподілу $f(x)$, можна використовувати такий алгоритм.

Генерують чергові два числа рівномірної випадкової послідовності ξ_{2i-1} та ξ_{2i} . Потім розраховують відповідні значення x_i' та y_i' . Після цього перевіряють виконання умови $y_i' < f(x_i')$. Якщо ця умова виконується, то $x_k = x_i'$ є черговим випадковим числом послідовності, яку треба побудувати. Якщо умова не виконується, то переходять до перевірки наступної пари значень рівномірної випадкової послідовності.

Ще один спосіб генерування послідовності випадкових чисел із заданим розподілом ґрунтується на можливості наближення істинного розподілу рівномірним на достатньо малих відрізках. Нехай ми маємо функцію щільності розподілу $f(x)$, задану графічно на відрізку $[a, b]$. Розіб'ємо цей відрізок на n рівних частин так, щоб виконувалася умова:

$$\int_{a_0}^{a_1} f(x) dx = \int_{a_1}^{a_2} f(x) dx = \dots = \int_{a_{n-1}}^{a_n} f(x) dx = \frac{1}{n}, \quad (5.23)$$

де $a_0 = a$, $a_n = b$. При такому розбитті потрапляння випадкової точки на будь-який відрізок є рівноймовірним. У загальному випадку відрізки $[a_i, a_{i+1}]$ будуть нерівними. Апроксимуємо $f(x)$ такою східчастою функцією $\phi(x)$, що для будь-якого відрізка розбиття виконується рівність:

$$\int_{a_i}^{a_{i+1}} f(x)dx = \int_{a_i}^{a_{i+1}} \varphi(x)dx = \varphi_i(a_{i+1} - a_i) = \frac{1}{n}. \quad (5.24)$$

Тобто:

$$\varphi(x) = \begin{cases} 0, & \text{якщо } x < a; \\ \frac{1}{n(a_{i+1} - a_i)}, & \text{якщо } x \in [a_i, a_{i+1}); \\ 0, & \text{якщо } x \geq b. \end{cases} \quad (5.25)$$

Координату випадкової точки M можна записати у вигляді:

$$x_M = a_i + \rho_M, \quad (5.26)$$

де a_i – координата лівого кінця інтервалу, в якому знаходиться точка, а ρ_M – його відстань до точки M . Значення ρ утворюють рівномірну випадкову послідовність. Тому їх можна одержати за формулою:

$$\rho = \xi(a_{i+1} - a_i). \quad (5.27)$$

На відміну від 5.17, ця формула не містить доданка a_i , оскільки значення ρ відраховуються саме від точки a_i .

Загальна схема побудови випадкової послідовності із заданою щільністю розподілу в даному випадку є такою. По-перше, визначають номер інтервалу, в якому знаходиться чергова точка M . Для цього знаходять черговий елемент ξ_{2j-1} рівномірної випадкової послідовності. Номер інтервалу i розраховують за формулою:

$$i = \lceil n\xi_{2j-1} \rceil, \quad (5.28)$$

де квадратними дужками позначено цілу частину числа. Після цього генерують наступний елемент рівномірної випадкової послідовності ξ_{2j} і розраховують випадкове число:

$$x = a_i + \xi_{2j} (a_{i+1} - a_i). \quad (5.29)$$

Недоліком цього способу є те, що таблиця значень меж інтервалів a_i має бути сформована заздалегідь.

5.4. Спеціальні методи генерування випадкових послідовностей з деякими законами розподілу

Загальна формула 5.15 є незручною для генерування випадкових чисел з нормальним законом розподілу, оскільки призводить до необхідності розв'язування відносно x_i рівняння:

$$\xi_i = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{x_i} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right) dx. \quad (5.30)$$

Інші загальні методи також не завжди бувають достатньо ефективними. Тому для побудови псевдовипадкових послідовностей з деякими законами розподілу доцільніше використовувати спеціальні методи, розроблені саме для таких послідовностей.

Один із найбільш поширених і простих методів побудови псевдовипадкових послідовностей, що задовольняє нормальний закон розподілу, ґрунтується на використанні центральної граничної теореми. Нехай ξ_i – елементи рівномірної випадкової послідовності, заданої на відрізку $[0, 1]$. Побудуємо з них послідовність:

$$\alpha_i = 2\xi_i - 1. \quad (5.31)$$

Вона також буде рівномірною і випадковою, але заданою на відріжку $[-1, 1]$. Розрахуємо величину:

$$\beta = \sum_{i=1}^n \alpha_i. \quad (5.32)$$

Згідно з центральною граничною теоремою, за достатньо великих n величина β буде мати розподіл, близький до стандартного нормального з параметрами:

$$m_\beta = M\left[\sum_{i=1}^n \alpha_i\right] = \sum_{i=1}^n M[\alpha_i] = 0; \quad (5.33)$$

$$\sigma_\beta^2 = D\left[\sum_{i=1}^n \alpha_i\right] = \sum_{i=1}^n D[\alpha_i] = \frac{n}{3}. \quad (5.34)$$

Величину $\gamma = \frac{\beta}{\sigma_\beta}$, що має приблизно стандартний нормальний розподіл з параметрами $m_\gamma = 0$ та $\sigma_\gamma = 1$, можна розрахувати за формулою:

$$\gamma = \sqrt{\frac{3}{n}} \sum_{i=1}^n (2\xi_i - 1), \quad (5.35)$$

яка дає задовільні для практичного користування результати при $n \geq 8$.

У практиці математичного моделювання складних систем велике значення має можливість скорочення кількості обчислень. При генеруванні нормально

розподілених величин для цього можна використовувати такі емпіричні формули:

$$\mu_k = \gamma_k - \frac{1}{20n} (3\gamma_k - \gamma_k^3); \quad (5.36)$$

та

$$\varsigma_k = \gamma_k - \frac{41}{13440n^2} (\gamma_k^5 - 10\gamma_k^3 + 15\gamma_k). \quad (5.37)$$

Використання 5.36 дає змогу скоротити величину n для обчислення γ_k до п'яти, а використання 5.37 – до двох.

Для одержання елементів нормальної послідовності використовують також перетворення:

$$\gamma_j = \frac{\sum_{i=1}^k \xi_i - k/2}{\sqrt{k/12}}. \quad (5.38)$$

На практиці часто буває достатнім взяти $k = 12$.

Для одержання послідовності випадкових чисел v , що мають нормальний розподіл з параметрами $m_v = a$ та $\sigma_v = b$, необхідно зробити перетворення:

$$v_k = b\gamma_k + a. \quad (5.39)$$

Дві незалежні послідовності з нормальним законом розподілу й параметрами $(0; 1)$ можна одержати з двох незалежних рівномірних випадкових послідовностей за допомогою перетворень:

$$\begin{cases} N_{1i} = (-2 \ln \xi_{2i})^{1/2} \sin(2\pi \xi_{1i}); \\ N_{2i} = (-2 \ln \xi_{2i})^{1/2} \cos(2\pi \xi_{1i}). \end{cases} \quad (5.40)$$

Такий спосіб дає можливість суттєво зменшити обсяг потрібних розрахунків завдяки тому, що один і той самий елемент кожної рівномірної послідовності використовується двічі.

Стисло зупинимося на методах побудови псевдовипадкових послідовностей з деякими іншими законами розподілу.

Чергові реалізації B_i випадкової величини, що має **розподіл Бернуллі** з параметром p (імовірність успіху), одержують з відповідних елементів рівномірної¹ послідовності, за допомогою такого перетворення:

$$B_i = \begin{cases} 1, & \text{якщо } \xi_i < p; \\ 0, & \text{якщо } \xi_i > p. \end{cases} \quad (5.41)$$

Реалізації G_i випадкової величини, що має **геометричний розподіл**, можна одержати з елементів рівномірної послідовності $\xi_i \in [0, 1]$ за таким алгоритмом. Черговий елемент G_i дорівнює значенню $\frac{\ln \xi_i}{\ln(1-p)}$, округленому у більший бік до найближчого цілого числа (p – імовірність успіху при одному випробуванні).

Реалізації B_j випадкової величини, що має **біноміальний розподіл** з параметрами n, p (n – кількість незалежних випробувань), можна одержати з елементів рівномірної послідовності так. Візьмемо перші n елементів рівномірної послідовності й перевіримо виконання умови $\xi_i < p$. Кількість елементів, для яких ця умова виконується, є першим елементом послідовності з біноміальним розподілом B_1 . Потім здійснюємо таку саму перевірку для наступних n елементів рівномірної послідовності й одержуємо другий елемент B_2 . Продовжуючи цю процедуру, отримаємо необхідну кількість елементів псевдовипадкової послідовності з біноміальним законом розподілу.

¹ Якщо не оговорено інше, під рівномірною послідовністю далі у цьому розділі ми будемо розуміти послідовність, елементи якої рівномірно розподілені на відрізок $[0; 1]$.

За малих p попередній алгоритм є не ефективним. У цьому випадку можна застосовувати такий метод. Будемо послідовно додавати елементи псевдовипадкової послідовності, починаючи з першого, доки їх сума не перевищить p . Перший елемент послідовності з біноміальним розподілом дорівнюватиме кількості доданків мінус одиниця. Аналогічно одержуємо інші елементи біноміальної послідовності.

Реалізації Y_j випадкової величини, що має **від’ємний біноміальний розподіл** з параметрами x, p (x – кількість успіхів), можна одержати з елементів рівномірної послідовності $\xi_i \in [0, 1]$ за методом бракування. Підраховуємо кількість перших елементів рівномірної послідовності, які менше p . Нехай вона вперше стане рівною x після того, як ми врахуємо m -й елемент. Тоді Y_1 дорівнюватиме кількості чисел ξ_i , які більші за p , серед перших m елементів рівномірної послідовності. Чергові елементи Y_j одержуємо за аналогічною процедурою, використовуючи наступні елементи рівномірної послідовності.

При малих p більш ефективним є алгоритм, що використовує перетворення:

$$Y_j = \sum_{i=1}^x G_i - x, \quad (5.42)$$

де G_i – елементи послідовності з геометричним законом розподілу.

Реалізації C_j випадкової величини, що має **розподіл Паскаля** з параметрами x, p , можна одержати, використовуючи перетворення:

$$C_j = \sum_{i=1}^x G_i, \quad (5.43)$$

де G_i – чергові елементи послідовності з геометричним законом розподілу.

Реалізації P_i **розподілу Пуассона** одержують з відповідних елементів рівномірної послідовності за таким алгоритмом. Розраховують функцію розподілу $F(x)$ для $x = 0, 1, 2, \dots, N$ за формулою:

$$F(x) = \sum_{k=0}^x \lambda^k \exp(-\lambda) / k!, \quad (5.44)$$

де λ – параметр розподілу (середнє), N має бути достатньо великим. Значення P_i вважають рівним x , для якого виконується умова: $F(x) < \xi_i < F(x+1)$.

За малих λ для прискорення процедури використовують спрощений алгоритм. Значення P_i беруть рівним x , для якого виконується умова: $\xi_1 \xi_2 \dots \xi_{x+1} < e^{-\lambda} < \xi_1 \xi_2 \dots \xi_x$.

Реалізації H_j випадкової величини, що має **гіпергеометричний розподіл** з параметрами N, X, n (X – кількість елементів, які мають задану властивість, N – загальна кількість елементів у вибірці, n – кількість витягнутих без повернення елементів), можна одержати з елементів рівномірної послідовності $\xi_i \in [0, 1]$ так. Будемо вважати успіхом подію $\xi_i < p_i$ ($i = 1, 2, \dots, n$), де:

$$\begin{aligned} p_1 &= X/N; \\ N_1 &= N; \\ p_{i+1} &= (N_i p_i - d_i) / (N - i); \\ d_i &= \begin{cases} 0, & \text{якщо } \xi_i \geq p_i; \\ 1, & \text{якщо } \xi_i < p_i; \end{cases} \\ N_i &= N - i + 1. \end{aligned} \quad (5.45)$$

Шуканим випадковим числом H_j буде кількість успіхів.

Чергові реалізації L_j **логнормального розподілу** з параметрами m, σ ($m = \exp \mu$ – медіана, σ – стандартне відхилення для $\ln L$, μ – математичне спо-

дівання для $\ln L$) можна одержати з елементів нормальної з параметрами $(0, 1)$ або рівномірної послідовностей за допомогою перетворень:

$$\begin{cases} L_j = m \exp(\sigma N_j); \\ L_j = m \exp \left[\sigma \left(\sum_{i=1}^{12} \xi_i - 6 \right) \right]. \end{cases} \quad (5.46)$$

Чергові реалізації X_j **розподілу χ^2** з ν степенями вільності можна одержати із елементів незалежних рівномірної та нормальної з параметрами $(0, 1)$ послідовностей за допомогою таких перетворень.

Для парних ν :

$$X_j = -\frac{1}{2} \ln \left(\prod_{i=1}^{\nu/2} \xi_i \right). \quad (5.47)$$

Для непарних ν :

$$X_j = -\frac{1}{2} \ln \left(\prod_{i=1}^{(\nu-1)/2} \xi_i \right) + (N_j)^2. \quad (5.48)$$

Для $\nu = 1$:

$$X_j = (N_j)^2. \quad (5.49)$$

Чергові реалізації Π_i **розподілу Парето** з параметром форми c можна одержати із відповідних елементів рівномірної послідовності за допомогою перетворення:

$$\Pi_i = (1/\xi_i)^{1/c}. \quad (5.50)$$

Чергові реалізації W_i **розподілу Вейбула** з параметрами масштабу b та форми c можна одержати із відповідних елементів рівномірної послідовності за формулою:

$$W_i = b(-\ln \xi_i)^{1/c}. \quad (5.51)$$

Чергові реалізації L_i **логістичного розподілу** можна одержати із відповідних елементів рівномірної послідовності за допомогою перетворення:

$$L_i = a + k \ln \left(\frac{\xi_i}{1 - \xi_i} \right), \quad (5.52)$$

де a – параметр розміщення, $k = \sqrt{3}b / \pi$ – коефіцієнт масштабу, b – стандартне відхилення.

5.5. Моделювання багатовимірних випадкових векторів

N -вимірний випадковий вектор можна задати як сукупність його компонент $\mathbf{X} = (\eta_1, \eta_2, \dots, \eta_n)$, де компоненти вектора η_i є випадковими величинами із заданим законом розподілу. Завданням моделювання є одержання послідовності чергових реалізацій випадкового вектора $X_1, X_2, \dots, X_i, \dots$ у вигляді послідовності чисел $(\eta_1^1, \eta_2^1, \dots, \eta_n^1), (\eta_1^2, \eta_2^2, \dots, \eta_n^2), \dots, (\eta_1^i, \eta_2^i, \dots, \eta_n^i), \dots$. Якщо компоненти вектора незалежні одна від одної, то доцільно здійснювати генерування кожної компоненти як незалежної випадкової послідовності із заданим законом розподілу. Перед використанням отриманих реалізацій випадкового вектора необхідно перевірити відсутність кореляції між сукупностями випадкових чисел, що генеруються.

Складнішою є задача моделювання випадкового вектора за наявності кореляції (залежності) між його компонентами. Кореляційну залежність між ком-

понентами нормального випадкового вектора можна задати через матрицю других моментів:

$$\mathbf{K} = \{\rho_{ij}\} \quad (i, j = 1, 2, \dots, n), \quad (5.53)$$

де

$$\rho_{ij} = \begin{cases} \sigma_i^2, & \text{якщо } i = j \\ M[(\eta_i - a_i)(\eta_j - a_j)], & \text{якщо } i \neq j. \end{cases} \quad (5.54)$$

Матриця \mathbf{K} є симетричною стосовно головної діагоналі оскільки $\rho_{ij} = \rho_{ji}$.

Нехай \mathbf{X} – це двовимірний випадковий вектор, компоненти якого η_1 і η_2 мають математичні сподівання a_1 і a_2 . Матриця других моментів:

$$\mathbf{K} = \begin{pmatrix} \sigma_1^2 & \rho \\ \rho & \sigma_2^2 \end{pmatrix}, \quad (5.55)$$

де $\rho = \rho_{12} = \rho_{21}$.

Згенеруємо дві незалежні нормальні послідовності випадкових чисел μ_1 і μ_2 , математичні сподівання яких дорівнюють нулю, а дисперсії – одиниці. З них можна одержати дві нові нормальні послідовності за допомогою перетворень:

$$\begin{cases} \overline{\eta_1} = a\mu_1, \\ \overline{\eta_2} = b\mu_1 + c\mu_2. \end{cases} \quad (5.56)$$

Оберемо коефіцієнти a, b, c такими, щоб виконувалися умови:

$$M[\overline{\eta_1}^2] = \sigma_1^2; \quad M[\overline{\eta_2}^2] = \sigma_2^2; \quad M[\overline{\eta_1}\overline{\eta_2}] = \rho. \quad (5.57)$$

Враховуючи (5.56), одержимо:

$$\begin{cases} M[\overline{\eta_1^2}] = a^2 M[\mu_1^2] = a^2; \\ M[\overline{\eta_2^2}] = b^2 M[\mu_1^2] + 2bcM[\mu_1\mu_2] + c^2 M[\mu_2^2] = b^2 + c^2; \\ M[\overline{\eta_1\eta_2}] = abM[\mu_1^2] + acM[\mu_1\mu_2] = ab. \end{cases} \quad (5.58)$$

Звідси маємо систему:

$$\begin{cases} a^2 = \sigma_1^2; \\ b^2 + c^2 = \sigma_2^2; \\ ab = \rho, \end{cases} \quad (5.59)$$

розв'язком якої є:

$$a = \sigma_1, \quad b = \frac{\rho}{\sigma_1}, \quad c = \frac{\sqrt{\sigma_1^2\sigma_2^2 - \rho^2}}{\sigma_1}. \quad (5.60)$$

Тому:

$$\begin{cases} \overline{\eta_1} = \sigma_1\mu_1; \\ \overline{\eta_2} = \frac{\rho}{\sigma_1}\mu_1 + \frac{\sqrt{\sigma_1^2\sigma_2^2 - \rho^2}}{\sigma_1}\mu_2. \end{cases} \quad (5.61)$$

Реалізації вихідних випадкових величин η_1 і η_2 можна одержати за допомогою перетворень:

$$\begin{cases} \eta_1 = \bar{\eta}_1 + a_1; \\ \eta_2 = \bar{\eta}_2 + a_2. \end{cases} \quad (5.62)$$

Для тривимірного випадкового вектора (η_1, η_2, η_3) матриця других моментів має вигляд:

$$\mathbf{K} = \begin{pmatrix} \sigma_1^2 & \rho_{12} & \rho_{13} \\ \rho_{12} & \sigma_2^2 & \rho_{23} \\ \rho_{13} & \rho_{23} & \sigma_3^2 \end{pmatrix}. \quad (5.63)$$

Реалізації тривимірного випадкового вектора можна одержати за допомогою перетворень:

$$\begin{cases} \eta_1 = \bar{\eta}_1 + a_1 = a\mu_1 + a_1; \\ \eta_2 = \bar{\eta}_2 + a_2 = b\mu_1 + c\mu_2 + a_2; \\ \eta_3 = \bar{\eta}_3 + a_3 = d\mu_1 + e\mu_2 + f\mu_3 + a_3. \end{cases} \quad (5.64)$$

де μ_1, μ_2, μ_3 – незалежні нормально розподілені випадкові послідовності, для яких математичні сподівання дорівнюють нулю, а дисперсії – одиниці.

Другі моменти розподілу дорівнюють:

$$\begin{aligned} M[\bar{\eta}_1^2] &= M[a^2\mu_1^2] = a^2; \\ M[\bar{\eta}_2^2] &= M[b^2\mu_1^2 + 2bc\mu_1\mu_2 + c^2\mu_2^2] = b^2 + c^2; \\ M[\bar{\eta}_3^2] &= M[d^2\mu_1^2 + e^2\mu_2^2 + f^2\mu_3^2 + 2de\mu_1\mu_2 + 2df\mu_1\mu_3 + 2ef\mu_2\mu_3] = \\ &= d^2 + e^2 + f^2; \\ M[\bar{\eta}_1\bar{\eta}_2] &= M[ab\mu_1^2 + ac\mu_1\mu_2] = ab; \end{aligned}$$

$$\begin{aligned}
M[\overline{\eta_1 \eta_3}] &= M[ad\mu_1^2 + ae\mu_1\mu_2 + af\mu_1\mu_3] = ad; \\
M[\overline{\eta_2 \eta_3}] &= M[bd\mu_1^2 + be\mu_1\mu_2 + bf\mu_3\mu_3 + cd\mu_1\mu_2 + ce\mu_2^2 + cf\mu_2\mu_3] = bd + ce.
\end{aligned}
\tag{5.65}$$

Звідси одержуємо систему шести рівнянь з шістьма змінними:

$$\begin{cases}
a^2 = \sigma_1^2; \\
b^2 + c^2 = \sigma_2^2; \\
d^2 + e^2 + f^2 = \sigma_3^2; \\
ab = \rho_{12}; \\
ad = \rho_{13}; \\
bd + ce = \rho_{23},
\end{cases}
\tag{5.66}$$

розв'язуючи яку можна знайти коефіцієнти a, b, c , необхідні для побудови тривимірного випадкового вектора (η_1, η_2, η_3) .

5.6. Порівняння деяких алгоритмів генерування випадкових послідовностей

Як було зазначено вище, переважна більшість алгоритмів генерування випадкових послідовностей із заданими законами розподілу передбачає попередню побудову послідовностей, елементи яких підпорядковуються рівномірному розподілу на відрізку $[0; 1]$. Існує велика кількість методів їх генерування, що зумовлено суперечливими вимогами до одержуваних послідовностей. З одного боку, відповідний алгоритм повинен забезпечувати необхідну якість послідовності. При цьому поняття якості істотно залежить від того, де вона буде використовуватися надалі. З іншого боку, враховуючи, що побудова випадкової послідовності є лише попереднім етапом дослідження, застосовуваний алгоритм має бути таким, що не виявляє істотного впливу на потрібні для моделювання ресурси пам'яті ЕОМ, часу тощо. Далі будуть наведені результати порівняння якості рівномірних випадкових послідовностей, заданих на відрізку $[0, 1]$, що отримані методом остач; за допомогою вбудованого генератора елект-

ронних таблиць MS Excel; а також за допомогою алгоритму, що ґрунтується на використанні рекурентної формули, яка утворює наступний елемент із попереднього за допомогою трикутного відображення $x_{n+1} = C(1 - 2|0,5 - x_n|)$, де C є близьким до одиниці.

Як один із методів перевірки якості випадкових послідовностей було використано χ^2 -тест електронних таблиць MS Excel, який дає імовірність p того, що досліджувана вибірка задовольняє заданий закон розподілу. Для 100 послідовностей обсягом від 200 до 10000 елементів, отриманих за допомогою вбудованого генератора випадкових послідовностей MS Excel, значення імовірності p варіювали в інтервалі від 0,013 до 0,980 і не залежали від кількості елементів послідовності (рис. 5.3). Середнє значення імовірності p для кожної з отриманих серій випадкових послідовностей є близьким до 0,5, а її стандартне відхилення – до 0,3. Це вказує на те, що важко заздалегідь передбачити якість рівномірної послідовності, згенерованої за допомогою вбудованого засобу електронних таблиць MS Excel, тому одержувані послідовності використовувати у реальних дослідженнях тільки після їх тестування.

Для послідовностей обсягом від 200 до 10000 елементів, отриманих методом остач, значення імовірності p варіювали в інтервалі від 0 до 1 і не залежали від кількості елементів послідовності (рис. 5.4).

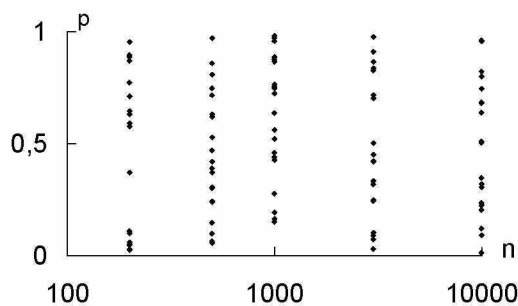


Рис. 5.3. Залежність значень імовірності χ^2 -тесту від обсягу n для послідовностей, отриманих за допомогою вбудованого генератора MS Excel

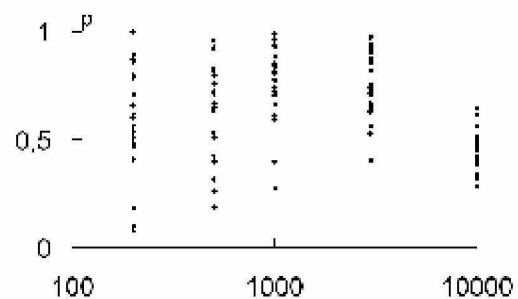


Рис. 5.4. Залежність значень імовірності χ^2 -тесту від обсягу n для послідовностей, отриманих методом остач

Але ці значення істотно й немонотонно залежать від параметрів ξ_0, M (рис. 5.5, 5.6). Змінюючи їх, можна дещо підвищити середнє значення і зменшити стандартне відхилення для імовірності p .

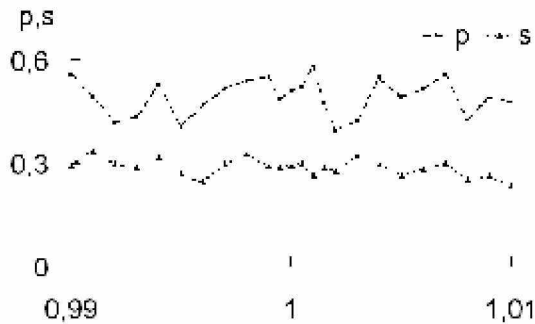


Рис. 5.5. Залежність середнього значення p та стандартного відхилення s імовірності χ^2 -тесту від параметра **A** для послідовностей, отриманих методом остач

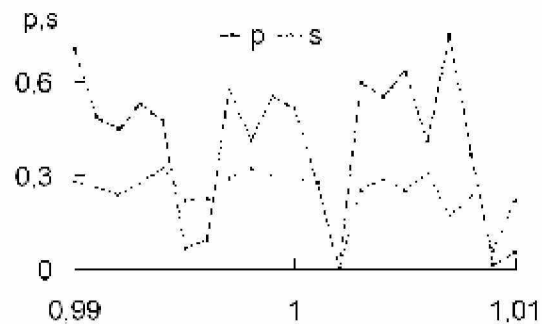


Рис. 5.6. Залежність середнього значення p та стандартного відхилення s імовірності χ^2 -тесту від параметра **B** для послідовностей, отриманих методом остач

У нашому випадку максимальні значення імовірності p в серіях із 20 послідовностей дорівнювали 0,7–0,75, а мінімальні значення її стандартного відхилення – 0,1–0,15. Втім і у цьому разі, одержувані послідовності можна використовувати лише після їх попередньої перевірки.

Для послідовностей, отриманих за допомогою трикутного відображення, результат χ^2 -тесту істотно залежить від обрання початкового елемента x_1 (рис. 5.7). Ця залежність є хаотичною, тому придатні для досягнення мети моделювання значення x_1 мають бути визначені попередньо.

Як було вказано вище, останньою перевіркою якості випадкової послідовності має бути її випробування при моделюванні реальної системи або процесу, результат дослідження для якого є відомим з попередніх незалежних даних.

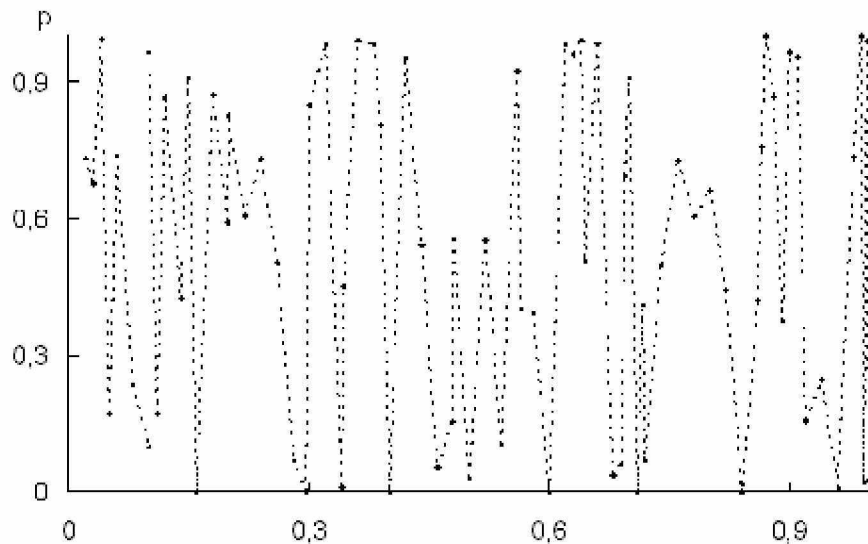


Рис. 5.7. Залежність значення χ^2 -тесту від величини x_1 при використанні трикутного відображення

Як приклад розглянемо проблему генерування випадкової послідовності, що задовольняє розподіл Максвелла. Він характеризує рівноважний розподіл частинок ідеального газу за швидкостями при певній температурі і є добре вивченим за допомогою різноманітних експериментальних і теоретичних методів.

Для генерування випадкових послідовностей, що відповідають розподілу Максвелла при кімнатній температурі використовували метод відбору з параметрами: $a = 10$; $b = 10000$. Параметр c брали удвічі більшим, ніж максимальне значення теоретичної функції щільності розподілу. На рис. 5.8 показано графік функції розподілу отриманої випадкової послідовності. З нього видно, що у цілому вона відповідає заданому закону розподілу. Але є систематичні відхилення функції розподілу у бік малих значень всередині дослідженого інтервалу швидкостей і в бік великих швидкостей по його бокам. Для більш ґрунтовного висновку про якість отриманої послідовності перевіримо її за допомогою інших методів.

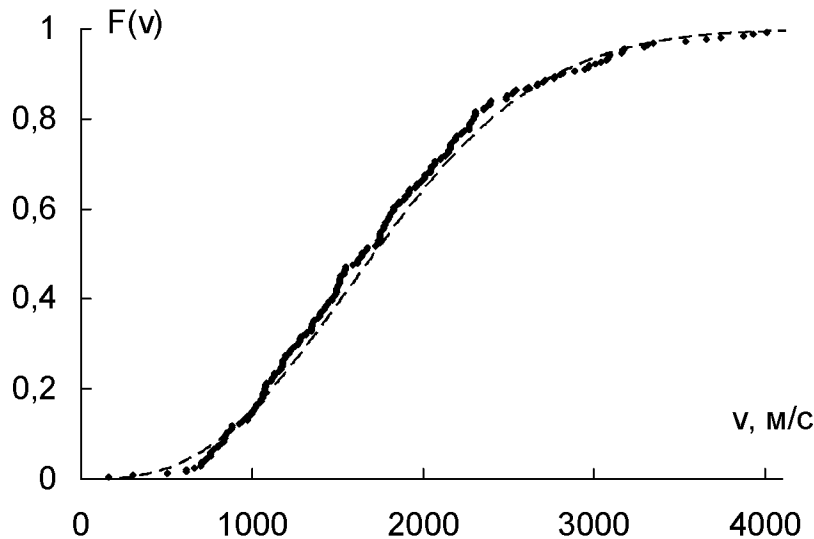


Рис 5.8. Графік функції розподілу побудованої псевдовипадкової послідовності (пунктиром показано теоретичну функцію розподілу)

У цьому випадку ми можемо розрахувати математичне сподівання, для якого відомо його теоретичне значення $v_m = \sqrt{\frac{8kT}{\pi m}}$, де k – стала Больцмана, T – температура, m маса частинки газу. У нашому випадку відмінність середнього арифметичного вибірки відрізняється від теоретичного математичного сподівання лише на 1,1%, а різниця вибіркової середньоквадратичної швидкості від відповідного теоретичного значення є ще меншою – близько 0,7%. Але застосувати тести, що ґрунтуються на порівнянні середнього арифметичного вибірки з теоретичним значенням математичного сподівання у цьому разі неможливо, оскільки ці тести передбачають, що вибірка підпорядковується нормальному закону розподілу. Найпростішим з непараметричних тестів є χ^2 -критерій, який у даному випадку дає імовірність відповідності генерованої послідовності розподілу Максвелла для температури 300 К, рівну 0,19. Тобто більш імовірним є те, що вона не відповідає заданому закону розподілу. Це збігається із зробленим раніше висновком про наявність систематичних відхилень емпіричної функції розподілу від теоретичної.

Іншим тестом, який можна використати для перевірки відповідності згенерованої послідовності заданому закону розподілу є критерій згоди Колмогорова – Смирнова, що ґрунтується на визначенні максимальної за модулем різниці між емпіричною та теоретичною функціями розподілу. У нашому випадку він дає імовірність відповідності не більше, ніж 0,43. Це також свідчить про недостатню якість отриманої послідовності. Причинами цього можуть бути недостатня кількість елементів рівномірної вихідної послідовності, а також надмірна простота використаних алгоритмів. Слід зазначити, що придатність отриманої послідовності для подальшого використання залежить від конкретного завдання дослідження. У багатьох випадках вона може виявитися цілком задовільною. Зокрема, вона є достатньою для отримання основних якісних закономірностей процесів, що вивчаються.

Контрольні питання

1. У чому полягає сутність методу Монте-Карло?
2. Чим можуть бути зумовлені похибки результатів при використанні методу Монте-Карло?
3. Як можна підвищити точність результатів моделювання при використанні методу Монте-Карло?
4. Якою є функція щільності довільного рівномірного розподілу?
5. Якими є математичне сподівання й дисперсія для довільного рівномірного розподілу?
6. Як визначити імовірність потрапляння випадкової величини до певного інтервалу при відомій функції щільності розподілу?
7. Як визначити імовірність потрапляння випадкової величини до певного інтервалу при відомій функції розподілу?
8. У чому полягає сутність табличного способу генерування послідовностей випадкових чисел, якими є основні переваги й недоліки цього способу?
9. У чому полягає сутність фізичних способів генерування послідовностей випадкових чисел, якими є їх основні переваги й недоліки?

10. У чому полягає сутність програмних способів генерування послідовностей випадкових чисел, якими є їх основні переваги й недоліки?

11. У чому полягає сутність алгоритму фон Неймана генерування рівномірних випадкових послідовностей?

12. Яким є алгоритм методу остач генерування рівномірних випадкових послідовностей?

13. Які властивості використовують для перевірки якості згенерованих псевдовипадкових послідовностей?

14. Як перевірити випадковість елементів побудованої випадкової послідовності?

15. Як перевірити відповідність елементів згенерованої випадкової послідовності заданому закону розподілу?

16. Як можна отримати аналітичну формулу, що перетворює елементи рівномірної випадкової послідовності в елементи випадкової послідовності із заданим законом розподілу?

17. Яким є алгоритм методу відбору генерування випадкових послідовностей із заданим законом розподілу?

18. Яким є алгоритм генерування випадкових послідовностей із заданим законом розподілу, що ґрунтується на наближенні істинного розподілу рівномірним на малих відрізках?

19. Які спеціальні методи використовують для генерування випадкових послідовностей із стандартним нормальним розподілом?

20. Як перетворити послідовність із стандартним нормальним розподілом до послідовності, що відповідає нормальному розподілу з довільними математичним сподіванням і дисперсією?

21. Поясніть алгоритми генерування випадкових послідовностей, що підпорядковуються основним типам законів розподілу?

22. Які вбудовані засоби генерування випадкових послідовностей передбачені в електронних таблицях MS Excel, мовах програмування Pascal та C++?

6. ОПТИМІЗАЦІЙНЕ МОДЕЛЮВАННЯ СИСТЕМ

Оптимізаційне моделювання передбачає побудову математичної моделі досліджуваної системи з наступним визначенням значень параметрів моделі, що відповідають мінімуму або максимуму певної цільової функції. Для природничих систем такою функцією часто бувають потенціальна або вільна енергія системи (принцип мінімуму енергії) та її ентропія (принцип максимуму ентропії). В економічних системах як цільові функції найчастіше застосовують собівартість виробництва певної продукції, прибуток підприємства, витрати часу або інших ресурсів тощо.

6.1. Постановка задачі оптимізації. Основні поняття

Функцію $f(\mathbf{X})$, визначену на опуклій множині X , називають опуклою, якщо для будь-яких $\mathbf{X}_1, \mathbf{X}_2 \in X$ та усіх $\alpha \in [0, 1]$ виконується нерівність:

$$f(\alpha\mathbf{X}_1 + (1-\alpha)\mathbf{X}_2) \leq \alpha f(\mathbf{X}_1) + (1-\alpha)f(\mathbf{X}_2). \quad (6.1)$$

Функцію $f(\mathbf{X})$, визначену на опуклій множині X , називають угнутою, якщо функція $-f(\mathbf{X})$ є опуклою, тобто, якщо для будь-яких $\mathbf{X}_1, \mathbf{X}_2 \in X$ та усіх $\alpha \in [0, 1]$ виконується нерівність:

$$f(\alpha\mathbf{X}_1 + (1-\alpha)\mathbf{X}_2) \geq \alpha f(\mathbf{X}_1) + (1-\alpha)f(\mathbf{X}_2). \quad (6.2)$$

Якщо нерівності (6.1, 6.2) виконуються строго, то функцію називають, відповідно, **строго опуклою** та **строго угнутою**.

Опукла функція $f(\mathbf{X})$, визначена на опуклій множині X , є неперервною у кожній внутрішній точці цієї множини і має у кожній внутрішній точці похідну за будь-яким напрямом s .

Нехай функція $f(\mathbf{X})$ визначена в області O N -вимірного простору R^N , де $\mathbf{X} = (x_1, x_2, \dots, x_N) \in O$, і набуває значення на певному інтервалі O_f дійсної вісі.

Точку $\mathbf{Y} = (y_1, y_2, \dots, y_N)$ називають **точкою локального мінімуму (максимуму)** $f(\mathbf{X})$, якщо існує таке $\varepsilon > 0$, що нерівність

$$f(\mathbf{Y}) \leq f(\mathbf{Y} + \delta\mathbf{X}) \quad (f(\mathbf{Y}) \geq f(\mathbf{Y} + \delta\mathbf{X})) \quad (6.3)$$

виконується для усіх $\delta\mathbf{X} = (\delta x_1, \delta x_2, \dots, \delta x_N)$, що задовольняють нерівності $0 < |\delta x_i| < \varepsilon$. Значення функції у цій точці називають **локальним мінімумом (максимумом)** функції.

Якщо нерівність (6.3) виконується строго, то мінімум (максимум) називають **строгим**, у протилежному випадку – **нестрогим**.

Точку $\mathbf{Y} \in O$ називають **точкою глобального мінімуму (максимуму)** функції $f(\mathbf{X})$, якщо нерівність (6.3) виконується для усіх $\mathbf{X} \in O$. Точка глобального екстремуму функції завжди є й точкою її локального екстремуму, а число $f(\mathbf{Y})$ – точною нижньою (верхньою) межею множини значень $f(\mathbf{X})$. Зворотні твердження можуть виявитися помилковими. Функції, що мають єдиний мінімум (максимум), називають **унімодальними**.

Очевидно, що строгий глобальний максимум є єдиним.

У загальному вигляді задача оптимізації ставиться так. Нехай задані:

- функція $f(\mathbf{X})$, визначена на множині $O \subseteq R^N$;
- множина $D \subseteq R^N$ (ця множина відображає наявність змістових обмежень на значення незалежних змінних, наприклад те, що термін виконання певної операції не може бути від'ємною величиною).

Необхідно знайти точку $\mathbf{Y} = (y_1, y_2, \dots, y_N) \in D$, в якій функція $f(\mathbf{X})$ сягає екстремального значення, тобто:

$$f(\mathbf{Y}) = \text{extr} f(\mathbf{X}), \quad \mathbf{Y} \in D. \quad (6.4)$$

Функцію $f(\mathbf{X})$ називають **цільовою функцією**, змінні \mathbf{X} – **керованими змінними**, D – **допустимою множиною**, будь-який набір значень $\mathbf{Y} \in D$ – **допустимим розв'язком** задачі оптимізації.

Теорема Вейєрштрасса. Нехай допустима множина D є компактною (тобто обмеженою та замкненою) й не пустою. Тоді визначена на неї неперервна цільова функція $f(\mathbf{X})$ набуває глобального максимуму у внутрішній або межовій точці множини D .

Доведення цього твердження випливає з того, що образ неперервної функції, визначеної на компактній множині, також є компактим. Будь-яка компактна множина дійсних чисел має точні верхню та нижню грані. Вони досягаються у певних точках множини D , які і є, відповідно, точками глобальних максимуму та мінімуму $f(\mathbf{X})$ на заданій множині.

Умови теореми Вейєрштрасса є достатніми але не необхідними для існування глобального екстремуму. Зокрема не існує глобальних мінімуму та максимуму функції $f(x) = x$ на інтервалі $(0, 1)$, який є незамкненою множиною. Це пов'язано з тим, що точні нижня та верхня грані множини значень функції досягаються у точках 0 та 1 , відповідно, які не належать інтервалу $(0, 1)$. З іншого боку, функція $f(x) = \sin x$ має глобальні максимум та мінімум на інтервалі $(-2, 1)$, який також є незамкненим.

Точка \mathbf{Y} , у якій $f(\mathbf{X})$ сягає екстремуму, має належати $O \cap D$. Якщо $O = D = \mathbb{R}^N$, тобто множини O та D збігаються з простором \mathbb{R}^N і немає будь-яких обмежень на значення компонентів вектора \mathbf{X} , то задачу оптимізації називають задачею на безумовний екстремум. Якщо хоча б одна із множин O або D є власною підмножиною \mathbb{R}^N ($O \subset \mathbb{R}^N$ або $D \subset \mathbb{R}^N$, тобто значення компонентів вектора \mathbf{X} обмежені або областю визначення функції $f(\mathbf{X})$, або допустимою множиною), і $O \cap D \neq \emptyset$, то її називають задачею на умовний екстремум. Якщо $O \cap D = \emptyset$, точка екстремуму не існує. Якщо перетином множин O і D є єдина точка \mathbf{Y} , то вона і є точкою екстремуму.

У задачах на умовний екстремум часто задають не допустиму множину D , а систему співвідношень:

$$\psi_j(x_1, x_2, \dots, x_N) \leq (\geq, =) 0, \quad j=1, 2, \dots, M, \quad (6.5)$$

яка її визначає.

Функції ψ_j називають **обмеженнями**. Якщо деякі обмеження є наслідками інших, їх називають **надлишковими**. Якщо у точці умовного екстремуму певне обмеження (нерівність) виконується як рівність, то ресурс, що його породжує, називають **дефіцитним**, а обмеження – **активним**. Якщо ж воно виконується як строга нерівність, відповідний ресурс називають **бездефіцитним**, а обмеження – **пасивним**. Обмеження-рівності завжди є активними.

Якщо точка Y безумовного локального екстремуму функції $f(X)$ знаходиться всередині множини допустимих значень D , то обмеження виконуються автоматично. Якщо ж Y знаходиться поза межами D , то точка умовного локального екстремуму потрапляє на межу множини D , а обмеження-нерівності перетворюються у рівності.

У задачах тільки з обмеженнями-рівностями кількість обмежень M має бути меншою за кількість змінних N . При $M > N$ задача не має розв'язків, або частина обмежень є надлишковою. Якщо $M = N$, то допустимий розв'язок, як правило, є єдиним.

Необхідна умова екстремуму. Нехай функція $f(X)$ задана на R^N і диференційована у певній точці $Y = (y_1, y_2, \dots, y_N) \in R^N$. Якщо у цій точці функція має локальний екстремум, то:

$$\text{grad} f(Y) = e_1 \frac{\partial f(Y)}{\partial x_1} + e_2 \frac{\partial f(Y)}{\partial x_2} + \dots + e_N \frac{\partial f(Y)}{\partial x_N} = 0, \quad (6.6)$$

де e_i – одиничні вектори, що задають напрями координатних осей, та

$$\frac{\partial f(\mathbf{Y})}{\partial x_1} = \frac{\partial f(\mathbf{Y})}{\partial x_2} = \dots = \frac{\partial f(\mathbf{Y})}{\partial x_N} = 0. \quad (6.7)$$

Для функції однієї змінної умова (6.7) переходить у $\frac{df(y)}{dx} = 0$, що відповідає відомій з курсу диференціального числення теоремі Ферма.

Точку $\mathbf{Y} \in \mathbb{R}^N$ називають **стаціонарною**, або **критичною** точкою функції $f(\mathbf{X})$, якщо її часткові похідні у цій точці задовольняють умови (6.6, 6.7).

Точки локального екстремуму функції слід шукати серед її критичних точок.

Розглянемо функцію $f(\mathbf{X}) = 2x_1^2 + 4x_2^2$. Її графіки наведено на рис. 6.1.

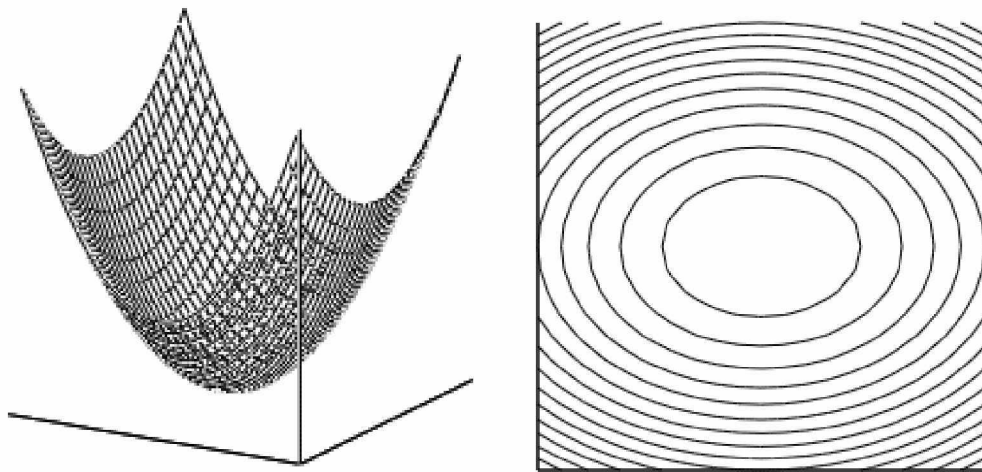


Рис. 6.1. Поверхневий та контурний графіки функції $f(\mathbf{X}) = 2x_1^2 + 4x_2^2$

Знайдемо стаціонарні точки:

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= 4x_1, & \frac{\partial f}{\partial x_2} &= 8x_2; \\ 4x_1 &= 0, & 8x_2 &= 0. \end{aligned} \quad (6.8)$$

Звідси маємо єдину стаціонарну точку $(0; 0)$, яка є точкою мінімуму.

Нехай $f(\mathbf{X}) = 5x_1^2 - 3x_2^2$. Графіки цієї функції наведено на рис. 6.2. Її стаціонарна точка $\mathbf{Y} = (0; 0)$ не є ані точкою мінімуму, ані точкою максимуму: у точці \mathbf{Y} $f(\mathbf{X})$ зростає на множині точок $\{x_1; 0\}$ і згасає на множині $\{0; x_2\}$. Таку критичну точку називають **сідловою**, а графік $f(\mathbf{X})$ – **сідловою поверхнею**.

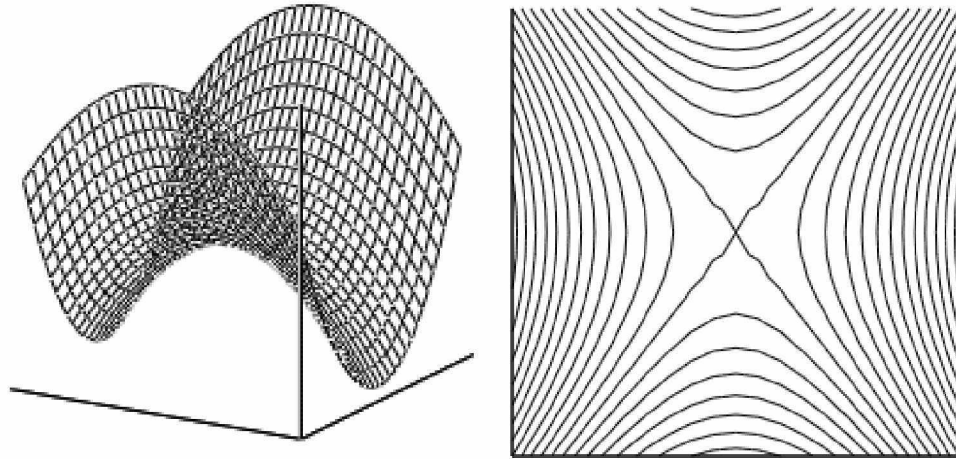


Рис. 6.2. Поверхневий та контурний графіки функції $f(\mathbf{X}) = 5x_1^2 - 3x_2^2$

Розглянемо функцію $f(\mathbf{X}) = (x_1 - 1)^3 + (x_2 + 1)^5$. Її графіки наведено на рис.

6.3. Її критичні точки визначаються рівняннями:

$$\frac{\partial f}{\partial x_1} = 3(x_1 - 1)^2 = 0; \quad \frac{\partial f}{\partial x_2} = 5(x_2 + 1)^4 = 0. \quad (6.9)$$

Єдиною критичною точкою є $\mathbf{Y} = (1; -1)$. У цій точці $f(\mathbf{X})$ зростає як за змінною x_1 на множині $\{x_1; -1\}$, так і за змінною x_2 на множині $\{1; x_2\}$. У цьому разі критична точка є точкою перегину.

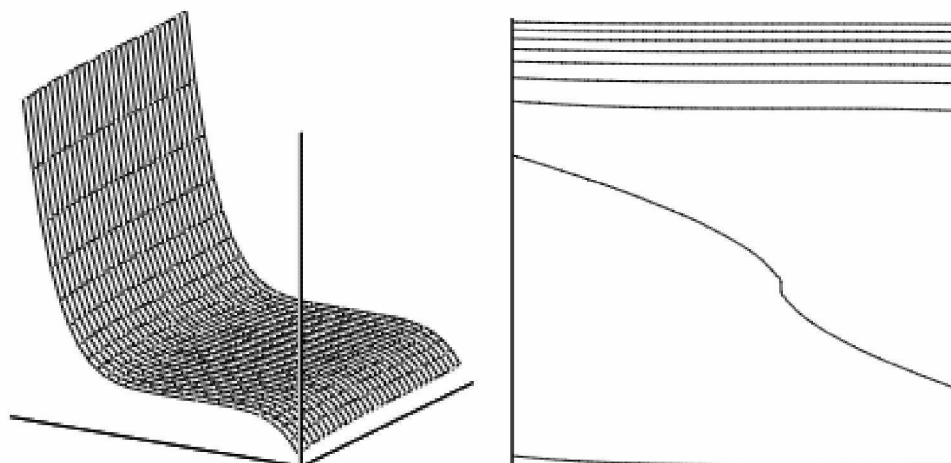


Рис. 6.3. Поверхневий та контурний графіки функції $f(\mathbf{X}) = (x_1 - 1)^3 + (x_2 + 1)^5$

Для двічі диференційованої у точці \mathbf{X} функції $f(\mathbf{X})$ визначена **матриця Гессе**:

$$H(\mathbf{X}) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{X})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{X})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{X})}{\partial x_1 \partial x_N} \\ \frac{\partial^2 f(\mathbf{X})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{X})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\mathbf{X})}{\partial x_2 \partial x_N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f(\mathbf{X})}{\partial x_N \partial x_1} & \frac{\partial^2 f(\mathbf{X})}{\partial x_N \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{X})}{\partial x_N^2} \end{pmatrix}, \quad (6.10)$$

яка є симетричною квадратною матрицею порядку $N \times N$.

Квадратичною формою матриці Гессе у точці \mathbf{X}^0 називають однорідний багаточлен другого степеня:

$$Q_0(\mathbf{X}) = \sum_{i=1}^N \sum_{j=1}^N x_i \frac{\partial^2 f(\mathbf{X}^0)}{\partial x_i \partial x_j} x_j = \sum_{i=1}^N \sum_{j=1}^N h_{ij} x_i x_j, \quad (6.11)$$

де h_{ij} є компонентами матриці Гессе. Матриця Гессе називається **додатно (від'ємно) визначеною** в точці \mathbf{X}^0 , якщо для будь-якої точки \mathbf{X} , крім $\mathbf{X} = 0$, її квадратична форма $Q_0(\mathbf{X}) > 0$ ($Q_0(\mathbf{X}) < 0$). Матриця Гессе називається **додатно (від'ємно) напіввизначеною**, або, відповідно, **невід'ємно (недодатно) визначеною** в точці \mathbf{X}^0 , якщо для будь-якої точки \mathbf{X} , крім $\mathbf{X} = 0$, її квадратична форма $Q_0(\mathbf{X}) \geq 0$ ($Q_0(\mathbf{X}) \leq 0$). Матриця Гессе називається **невизначеною** в точці \mathbf{X}^0 , якщо знак її квадратичної форми є різним у різних точках \mathbf{X} .

Нагадаємо, що **головними мінорами** довільної квадратної матриці A порядку $N \times N$ називають її мінори порядку $1, 2, \dots, N$, розташовані у її лівому верхньому куті:

$$M_1 = a_{11}; \quad M_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}; \quad M_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}; \quad \dots \quad (6.12)$$

Знаковизначеність будь-якої матриці A може бути перевірено за допомогою умов Сильвестра, згідно з якими вона є:

- додатно визначеною, якщо усі її головні мінори $M_i(A)$ є додатними;
- від'ємно визначеною, якщо знак її k -го головного мінору M_k збігається із знаком $(-1)^k$;
- додатно напіввизначеною, якщо A є виродженою матрицею ($|A| = 0$), та усі її головні мінори $M_i(A)$ є невід'ємними;
- від'ємно напіввизначеною, якщо знак її k -го головного мінору M_k збігається із знаком $(-1)^k$ або його значення дорівнює нулю.

Розглянемо функцію $f(\mathbf{X}) = x_1^3 + x_2^2 x_1 - 27x_1$. Її графіки наведено на рис.

6.4. Визначимо критичні точки:

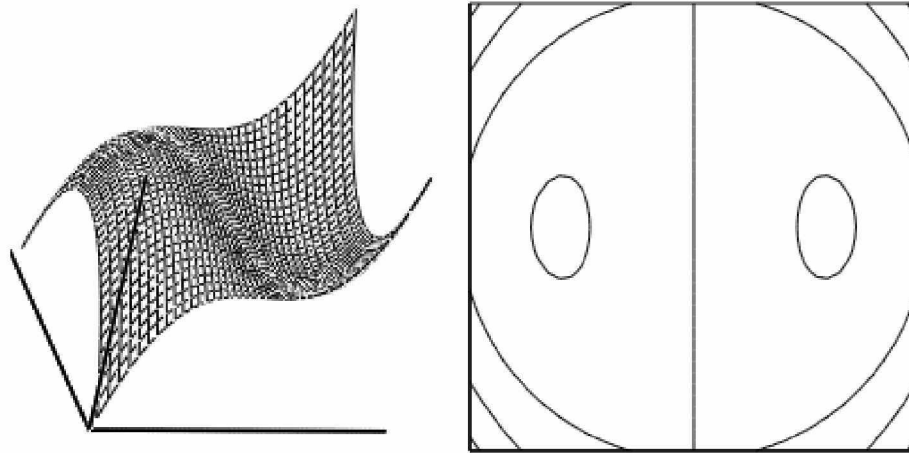


Рис. 6.4. Поверхневий та контурний графіки функції $f(\mathbf{X}) = x_1^3 + x_2^2 x_1 - 27x_1$

$$\begin{cases} \frac{\partial f}{\partial x_1} = 3x_1^2 + x_2^2 - 27 = 0; \\ \frac{\partial f}{\partial x_2} = 2x_1 x_2 = 0. \end{cases} \Rightarrow \begin{matrix} 1. \begin{cases} x_1 = 0; \\ x_2 = \pm 3\sqrt{3}. \end{cases} \\ 2. \begin{cases} x_1 = \pm 3; \\ x_2 = 0. \end{cases} \end{matrix} \quad (6.13)$$

Звідси: $\mathbf{X}^1 = (0; 3\sqrt{3})$, $\mathbf{X}^2 = (0; -3\sqrt{3})$, $\mathbf{X}^3 = (3; 0)$, $\mathbf{X}^4 = (-3; 0)$.

Матриця Гессе: $\mathbf{H}(\mathbf{X}) = \begin{pmatrix} 6x_1 & 2x_2 \\ 2x_2 & 2x_1 \end{pmatrix}$. Її квадратичні форми у критичних

точках мають вигляд:

$$\begin{aligned} 1. Q_1(\mathbf{X}) &= (x_1, x_2) \begin{pmatrix} 0 & 6\sqrt{3} \\ 6\sqrt{3} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \\ &= (6\sqrt{3}x_2 \quad 6\sqrt{3}x_1) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 12\sqrt{3}x_1 x_2 \stackrel{\leq}{\geq} 0; \end{aligned}$$

$$\begin{aligned}
2. Q_2(\mathbf{X}) &= (x_1, x_2) \begin{pmatrix} 0 & -6\sqrt{3} \\ -6\sqrt{3} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \\
&= (-6\sqrt{3}x_2 \quad -6\sqrt{3}x_1) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -12\sqrt{3}x_1x_2 \begin{matrix} \geq 0; \\ \leq 0; \end{matrix}
\end{aligned}$$

$$\begin{aligned}
3. Q_1(\mathbf{X}) &= (x_1, x_2) \begin{pmatrix} 18 & 0 \\ 0 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \\
&= (18x_1 \quad 6x_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 18x_1^2 + 6x_2^2 > 0;
\end{aligned}$$

$$\begin{aligned}
4. Q_1(\mathbf{X}) &= (x_1, x_2) \begin{pmatrix} -18 & 0 \\ 0 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \\
&= (-18x_1 \quad -6x_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -18x_1^2 - 6x_2^2 < 0;
\end{aligned}$$

(згідно з визначенням ми не враховуємо точку $(0; 0)$ при знаходженні знаку квадратичної форми у випадках 3 та 4). Як впливає з наведених вище даних, матриця Гессе є невизначеною у точках \mathbf{X}^1 та \mathbf{X}^2 , додатно визначеною у точці \mathbf{X}^3 і від'ємно визначеною у точці \mathbf{X}^4 .

Необхідна умова локального екстремуму. Нехай функція $f(\mathbf{X})$ задана на \mathbb{R}^N і двічі диференційована у певному околі точки $\mathbf{Y} \in \mathbb{R}^N$. Якщо у цій точці $f(\mathbf{X})$ має локальний мінімум (максимум), то матриця Гессе $H(\mathbf{Y})$ є недодатно (невід'ємно) визначеною.

Достатня умова екстремуму. Нехай функція $f(\mathbf{X})$ задана на \mathbb{R}^N і має критичну точку \mathbf{Y} , в якій усі другі частинні похідні існують і є неперервними. Якщо матриця Гессе $H(\mathbf{Y})$ у цій точці є додатно (від'ємно) визначеною, то \mathbf{Y} є точкою локального мінімуму (максимуму) $f(\mathbf{X})$. Якщо $H(\mathbf{Y})$ є невизначеною, то екстремуму у точці \mathbf{Y} немає. Зокрема, для функції $f(\mathbf{X}) = x_1^3 + x_2^2x_1 - 27x_1$, матриця

Гессе якої розглядалася вище, точки X^1 та X^2 не є точками екстремумів, точка X^3 є точкою локального мінімуму, а X^4 – локального максимуму.

6.2. Чисельні методи оптимізації

Одновимірна оптимізація. Задача одновимірної оптимізації передбачає знаходження найменшого або найбільшого значення функції $f(x)$ на заданому відрізку $[a; b]$. Їх існування для будь-якої неперервної на $[a; b]$ функції випливає з відомої із курсу математичного аналізу теореми Вейерштрасса.

Функція $f(x)$ може досягати своїх найменшого та найбільшого значень на відрізку $[a; b]$ у крайніх точках відрізка, а також у точках локальних мінімуму та максимуму.

Найпростішим методом пошуку екстремуму унімодальної функції $f(x)$ є **метод сканування**. У цьому разі весь відрізок поділяють на n ($n = 10 \dots 20$) рівних частин, отримуючи точки x_1, x_2, \dots, x_{n-1} такі, що:

$$x_1 - x_0 = x_2 - x_1 = \dots = x_i - x_{i-1} = x_n - x_{n-1} = h_1 = \frac{b - a}{n}, \quad (6.14)$$

де $x_0 = a, x_n = b$.

Після цього знаходимо точку x_k , якій відповідає найменша із величин $f(a), f(x_1), \dots, f(x_{n-1}), f(b)$. Якщо $x_k \neq a, b$, то внаслідок унімодальності $f(x)$ ми маємо зробити висновок, що точка екстремуму належить відрізку $[x_{k-1}, x_{k+1}]$. У випадках, коли $x_k = a$ або $x_k = b$, вона знаходиться, відповідно, на відрізках $[a, x_1]$ або $[x_{n-1}, b]$.

Якщо довжина отриманого відрізка перевищує задану допустиму похибку ε , то необхідно уточнити значення точки екстремуму. Для цього розбиваємо знайдений відрізок на n нових частин. Позначаємо їх крайні точки як x_0, x_1, \dots, x_n і знов знаходимо точку x_k , якій відповідає найменша із величин $f(a), f(x_1), \dots, f(x_{n-1}), f(b)$. Після цього визначаємо новий інтервал локаліза-

ції точки екстремуму. Цю процедуру продовжуємо, доки не буде виконана умова $x_{k+1} - x_{k-1} < \varepsilon$. Після цього беремо точку x_k за точку екстремуму. У разі, коли на всіх кроках ітерацій найменшим (найбільшим) постійно виявляється значення функції у точці a або точці b , точкою екстремуму після закінчення ітераційної процедури доцільно обрати саме цю точку.

Метод золотого перерізу дає можливість зменшити кількість обчислень без істотного ускладнення алгоритму. У цьому випадку на відрізку $[a, b]$ обирають дві внутрішні точки x_1 та x_2 , які задовольняють умови:

$$\frac{b - x_1}{b - a} = \frac{x_1 - a}{b - x_1}, \quad \frac{x_2 - a}{b - a} = \frac{b - x_2}{x_2 - a}. \quad (6.15)$$

Звідси маємо:

$$\frac{b - x_1}{b - a} = \frac{b - x_2}{x_2 - a} = \frac{-1 + \sqrt{5}}{2} \approx 0,618. \quad (6.16)$$

Цю величину називають **золотим перерізом**, що й дало назву методу. Розв'язуючи рівняння (6.16), одержуємо:

$$x_1 \approx a + \frac{3 - \sqrt{5}}{2}(b - a); \quad x_2 \approx a - \frac{1 - \sqrt{5}}{2}(b - a). \quad (6.17)$$

Далі ітераційну процедуру будують так. Спочатку знаходять значення функції у точках a , x_1 , x_2 , b . Потім вихідний відрізок звужують, відкидаючи точку a у разі, коли найменшим значенням є $f(x_2)$ чи $f(b)$, або точку b в інших випадках. Потім за вказаною процедурою на новому відрізку знов визначають дві внутрішні точки. При цьому, завдяки властивостям золотого перерізу, одна з них збігатиметься з точкою x_1 або x_2 , отриманою на попередньому кроці ітерацій. Тому на кожному кроці, крім першого, необхідно розраховувати лише

одну нову точку та відповідне значення функції. Ітерації продовжують, доки не буде виконано умову $\Delta^{(i)} < \varepsilon$, де $\Delta^{(i)}$ – довжина відрізка локалізації точки екстремуму після i -го кроку. Як значення точки екстремуму можна взяти середину цього відрізка.

Багатовимірна оптимізація. Чисельні методи багатовимірної оптимізації використовують різноманітні ітераційні процедури. При цьому наступне наближення отримують з попереднього за рекурентною формулою:

$$\mathbf{X}^{k+1} = \mathbf{X}^k + t^{k+1}\mathbf{S}^k, \quad (6.18)$$

де вектор \mathbf{S} визначає напрям пошуку, а скалярний параметр t – відстань між точками \mathbf{X}^k та \mathbf{X}^{k+1} у цьому напрямі.

У всіх методах необхідно задати умову зупинення ітерацій. Зазвичай використовують такі умови:

$$\left| \mathbf{X}^{k+1} - \mathbf{X}^k \right| < \varepsilon_1; \quad \left| f(\mathbf{X}^{k+1}) - f(\mathbf{X}^k) \right| < \varepsilon_2; \quad \left| \frac{\partial f(\mathbf{X}^{k+1})}{\partial x_i} \right| < \varepsilon_3, \quad (6.19)$$

де $\varepsilon_1, \varepsilon_2, \varepsilon_3$ є достатньо малими додатними числами.

Методи, що розглядаються нижче, є методами мінімізації функції. Якщо потрібно знайти максимум, то необхідно внести відповідні зміни до алгоритмів, або увести допоміжну функцію, мінімум якої збігається з максимумом вихідної функції (наприклад, помножити вихідну функцію на -1), а потім використати наведені алгоритми.

Метод найшвидшого спуску (градієнтного спуску). Цей метод ґрунтується на тому, що градієнт функції у точці має напрям, у якому функція найшвидше зростає. Тому обирають $\mathbf{S}^k = -\text{grad} f(\mathbf{X}^k)$, тобто формула (6.18) набуває вигляду:

$$\mathbf{X}^{k+1} = \mathbf{X}^k - t^{k+1} \text{grad} f(\mathbf{X}^k). \quad (6.20)$$

Параметр t^{k+1} визначають із умови, що він відповідає мінімуму $f(\mathbf{X})$ на промені $\mathbf{X}^k - t^{k+1} \text{grad} f(\mathbf{X}^k)$. Тоді його можна знайти, розглядаючи $f(\mathbf{X})$ як функцію однієї змінної:

$$\varphi(t^{k+1}) = f(\mathbf{X}^k - t^{k+1} \text{grad} f(\mathbf{X}^k)) \quad (6.21)$$

і записуючи умову мінімуму у вигляді:

$$\frac{d\varphi}{dt^{k+1}} = 0. \quad (6.22)$$

Достатньо простим та універсальним є такий спосіб визначення t^{k+1} , що ґрунтується на апроксимації похідної в (6.22) її різницеvim поданням з наступним застосуванням розкладу у ряд Тейлора по t^{k+1} :

$$\begin{aligned} \frac{d\varphi(t^{k+1})}{dt^{k+1}} &= \frac{f(\mathbf{X}^k - t^{k+1} \text{grad} f(\mathbf{X}^k)) - f(\mathbf{X}^k)}{t^{k+1}} = \\ &= -\text{grad} f(\mathbf{X}^k) \cdot \text{grad}^T f(\mathbf{X}^k) + \\ &+ \frac{1}{2} t^{k+1} \text{grad} f(\mathbf{X}^k) \cdot \mathbf{H}(\mathbf{X}^k) \cdot \text{grad}^T f(\mathbf{X}^k) = 0. \end{aligned} \quad (6.23)$$

Звідси маємо:

$$t^{k+1} = 2 \frac{\text{grad} f(\mathbf{X}^k) \cdot \text{grad}^T f(\mathbf{X}^k)}{\text{grad} f(\mathbf{X}^k) \cdot \mathbf{H}(\mathbf{X}^k) \cdot \text{grad}^T f(\mathbf{X}^k)}. \quad (6.24)$$

Якщо на певному кроці ітерацій не вдається знайти корінь рівняння (6.22), то як чергове значення параметра t можна взяти будь-яке число, для якого $f(\mathbf{X}^{k+1}) < f(\mathbf{X}^k)$.

Метод покоординатного спуску (релаксації). У цьому разі як черговий напрям пошуку мінімуму обираємо напрям координатної вісі, вздовж якої функція у точці поточного наближення змінюється найшвидше. Для цього обчислюємо компоненти вектора градієнта $\frac{\partial f(\mathbf{X}^k)}{\partial x_i}$ та обираємо номер i тієї з них, що є максимальною за модулем.

Нове наближення шукаємо на промені:

$$\mathbf{X}^k + t^{k+1} \mathbf{X}^k = \mathbf{X}^k - t^{k+1} \left(0, 0, \dots, 0, \frac{\partial f(\mathbf{X}^k)}{\partial x_i}, 0, \dots, 0 \right). \quad (6.25)$$

Для знаходження t^{k+1} можна використати метод, аналогічний розглянутому у методі найшвидшого спуску. Тоді отримаємо:

$$t^{k+1} = \frac{2(s_i^k)^2}{(s_i^k)^2 \frac{\partial^2 f(\mathbf{X}^k)}{\partial x_i^2}}, \quad s_i^k = -\frac{\partial f(\mathbf{X}^k)}{\partial x_i} \quad (6.26)$$

Метод Ньютона – Рафсона. При знаходженні критичних точок отримуємо систему рівнянь $\varphi(\mathbf{X}) = \frac{\partial f(\mathbf{X})}{\partial x_i} = 0$. Нехай $f(\mathbf{X})$ є двічі диференційованою функцією з невідродженою матрицею Гессе. Тоді $\varphi(\mathbf{X})$ можна розкласти у ряд Тейлора в околі точки \mathbf{X}^k :

$$\varphi(\mathbf{X}) = \varphi(\mathbf{X}^k) + \text{grad} \varphi(\mathbf{X}^k) \delta \mathbf{X} + o(\delta \mathbf{X}). \quad (6.27)$$

Приймаючи, що в одержуваній точці \mathbf{X} виконується умова $\varphi(\mathbf{X})=0$, отримаємо:

$$\varphi(\mathbf{X}^k) + \text{grad} \varphi(\mathbf{X}^k) \delta \mathbf{X} = 0 \quad (6.28)$$

або

$$\frac{\partial f(\mathbf{X}^k)}{\partial x_j} + \sum_{i=1}^N \frac{\partial^2 f(\mathbf{X}^k)}{\partial x_j \partial x_i} (x_i - x_i^k) = 0. \quad (6.29)$$

Цю систему можна записати в матричній формі у вигляді:

$$\mathbf{H}(\mathbf{X}^k) \delta \mathbf{X} = -\text{grad} f(\mathbf{X}^k). \quad (6.30)$$

При невиродженій матриці Гессе її розв'язком, відомим з курсу алгебри, є:

$$\delta \mathbf{X} = -\mathbf{H}^{-1}(\mathbf{X}^k) \cdot \text{grad} f(\mathbf{X}^k). \quad (6.31)$$

Звідси для наступного наближення маємо:

$$\mathbf{X}^{k+1} = \mathbf{X}^k - \mathbf{H}^{-1}(\mathbf{X}^k) \cdot \text{grad} f(\mathbf{X}^k). \quad (6.32)$$

Зазначимо, що на відміну від методу найшвидшого спуску градієнт у формулі (6.32) помножується на матрицю, а не на скаляр, тому напрям $\mathbf{X}^{k+1} - \mathbf{X}^k$ не збігається з напрямом антиградієнта.

Метод спряжених напрямів. У цьому методі черговий напрям пошуку \mathbf{S}^k обирається таким, щоб він був спряженим з усіма попередніми.

Два ненульових вектори \mathbf{Z}_1 та \mathbf{Z}_2 називають **спряженими** стосовно додатно визначеної симетричної матриці A , якщо скалярний добуток $(A\mathbf{Z}_1^T)\mathbf{Z}_2 = 0$. Систему відмінних від нуля векторів \mathbf{Z}_k ($k = 1, 2, \dots, K$) називають взаємно спряженою стосовно матриці A , якщо $(A\mathbf{Z}_i^T)\mathbf{Z}_j = 0$ при $0 \leq i, j \leq K, i \neq j$.

У методі спряжених напрямів обирають:

$$\begin{aligned} \mathbf{S}^0 &= -\text{grad} f(\mathbf{X}^0); \\ \mathbf{S}^k &= -\text{grad} f(\mathbf{X}^k) + \beta^{k-1} \mathbf{S}^{k-1}. \end{aligned} \tag{6.33}$$

Параметр

$$\beta^{k-1} = \frac{\mathbf{H} \cdot (\mathbf{S}^{k-1})^T \cdot \text{grad} f(\mathbf{X}^k)}{\mathbf{H} \cdot (\mathbf{S}^{k-1})^T \cdot \mathbf{S}^{k-1}} \tag{6.34}$$

забезпечує спряженість напрямів \mathbf{S}^k та \mathbf{S}^{k-1} стосовно додатно визначеної симетричної матриці \mathbf{H} порядку $N \times N$. Крок у напрямі \mathbf{S}^k для одержання наступного наближення розраховують за формулою:

$$t^{k+1} = -\frac{\text{grad} f(\mathbf{X}^k) \cdot (\mathbf{S}^k)^T}{\mathbf{S}^k \cdot \mathbf{H} \cdot (\mathbf{S}^k)^T}. \tag{6.35}$$

Як правило, за напрями, що спрягаються, обирають градієнти функції у точках чергових наближень. Тоді замість формул (6.34, 6.35) застосовують вирази:

$$f(\mathbf{X}) = c_1x_1 + c_2x_2 + \dots + c_Nx_N \rightarrow \max. \quad (6.44)$$

Система обмежень матиме вигляд:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1N}x_N \leq b_1; \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2N}x_N \leq b_2; \\ \dots\dots\dots; \\ a_{M1}x_1 + a_{M2}x_2 + \dots + a_{MN}x_N \leq b_M; \\ x_i \geq 0 \quad (i=1,2,\dots,N); \quad b_j \geq 0 \quad (j=1,2,\dots,M). \end{cases} \quad (6.45)$$

Задача про раціон. У раціоні харчування необхідно передбачити споживання M видів харчувальних речовин у кількості не менше b_j для j -ої речовини. При цьому можна використовувати N видів вихідних продуктів. Є відомими кількості харчувальної речовини j -го типу у продукті i -го типу (a_{ij}) та вартості одиниці i -го вихідного продукту (c_i).

Необхідно визначити раціон мінімальної вартості.

Математична модель має такий вигляд:

$$\begin{cases} f(\mathbf{X}) = c_1x_1 + c_2x_2 + \dots + c_Nx_N \rightarrow \min; \\ a_{11}x_1 + a_{12}x_2 + \dots + a_{1N}x_N \geq b_1; \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2N}x_N \geq b_2; \\ \dots\dots\dots; \\ a_{M1}x_1 + a_{M2}x_2 + \dots + a_{MN}x_N \geq b_M; \\ x_i \geq 0 \quad (i=1,2,\dots,N); \quad b_j \geq 0 \quad (j=1,2,\dots,M). \end{cases} \quad (6.46)$$

Множину невід'ємних розв'язків системи обмежень називають допустимими планами ЗЛП. Можна довести, що множина L усіх допустимих планів є або опуклою, або порожньою. З погляду геометрії вона є відкритим або замкненим опуклим багатогранником.

Будь-яку ЗЛП можна звести до канонічного вигляду. Для цього вводять додаткові (так звані "слабкі") змінні. Нехай вихідне j -те обмеження має вигляд:

$$\sum_{i=1}^N a_{ij}x_i \leq b_j \quad (b_j \geq 0). \quad (6.47)$$

Уводячи нову слабку змінну $y_j \geq 0$, ми можемо записати його як рівність:

$$\sum_{i=1}^N a_{ij}x_i + y_j = b_j. \quad (6.48)$$

Якщо доданки у (6.48) є певним ресурсом, то значення y_j матиме значення залишку цього ресурсу.

У разі, коли вихідне j -те обмеження задається як:

$$\sum_{i=1}^N a_{ij}x_i \geq b_j \quad (b_j \geq 0), \quad (6.49)$$

то, уводячи нову слабку змінну $y_j \geq 0$, можна перетворити його до рівності:

$$\sum_{i=1}^N a_{ij}x_i - y_j = b_j. \quad (6.50)$$

Якщо права частина якогось із обмежень є від'ємною, то її завжди можна зробити додатною, помножуючи обидві частини на -1 .

Якщо на певну керовану змінну за змістом задачі не накладають обмеження $x_i \geq 0$, то її завжди можна подати як різницю двох невід'ємних змінних x_a та x_b :

$$x_i = x_a - x_b \quad (x_a \geq 0, x_b \geq 0). \quad (6.51)$$

Цю підстановку необхідно застосовувати у всіх обмеженнях, які містять змінну x_i , а також у цільовій функції. При цьому для від'ємних значень x_i беруть: $x_a = 0$; $x_b > 0$, а для додатних – $x_a > 0$; $x_b = 0$.

Канонічну систему обмежень називають:

- ◆ **сумісною**, якщо вона має хоча б один розв'язок;
- ◆ **несумісною**, якщо вона не має жодного розв'язку;
- ◆ **надлишковою**, якщо хоча б одне із обмежень можна виразити через інші.

У подальшому будемо вважати, що система обмежень є сумісною й не надлишковою.

Звідси випливає, що у задачах лінійного програмування кількість змінних має бути більшою за кількість рівнянь у системі обмежень ($N > M$). У зв'язку з цим для розв'язання ЗЛП серед усіх змінних виділяють M **базисних** та $P = N - M$ **вільних**, так, щоб матриця коефіцієнтів A_B при базисних змінних була невиродженою, і записують систему обмежень у вигляді:

$$A_B \mathbf{Y}^T + A_F \mathbf{Z}^T = \mathbf{B}^T, \quad (6.52)$$

де $\mathbf{Y} = (y_1, y_2, \dots, y_M)$ – базисні змінні, $\mathbf{Z} = (z_1, z_2, \dots, z_M)$ – вільні змінні, A_B, A_F – матриці коефіцієнтів розміром $M \times M$ та $M \times P$, відповідно.

У подальшому вільні змінні розглядають як незалежні і переносять їх до правої частини лінійних обмежень:

$$A_B \mathbf{Y}^T = \mathbf{B}^T - A_F \mathbf{Z}^T. \quad (6.53)$$

Базисні змінні розглядають як залежні і визначають їх значення за допомогою перетворення:

$$\mathbf{Y}^T = (\mathbf{A}_B)^{-1} (\mathbf{B}^T - \mathbf{A}_F \mathbf{Z}^T). \quad (6.54)$$

Систему (6.54) часто називають системою обмежень, записаною у **базисному вигляді**, а змінні \mathbf{Y} – **базисом**.

Система (6.53) може мати нескінченну кількість розв'язків (допустимих планів). Оскільки множина (непуста) допустимих планів є опуклою, будь-який знайдений екстремум буде глобальним. Якщо область допустимих значень є необмеженою за певними змінними, то вводять додаткові (замикальні) обмеження вигляду $\sum_{i=1}^N x_i \leq K$. При обранні достатньо великих значень K такі обмеження практично не спотворюють змісту задачі і не впливають на її розв'язок.

Допустимий розв'язок, для якого усі вільні змінні дорівнюють нулю, називають **базисним розв'язком** або **опорним планом**. Кількість базисних розв'язків дорівнює C_N^M . Якщо у базисному розв'язку кількість додатних змінних є меншою за M , то його називають виродженим. Можна довести, що базисні розв'язки є вершинами опуклої багатогранної області допустимих планів.

Допустимий розв'язок, за якого цільова функція досягає свого мінімуму або максимуму, називають **оптимальним планом**. Оптимальний план завжди є базисним розв'язком.

Допустимий розв'язок, за якого цільова функція досягає свого мінімуму або максимуму, називають **оптимальним планом**. Оптимальний план завжди є базисним розв'язком.

Базисний розв'язок канонічної ЗЛП називають **невиродженим**, якщо усі вільні члени системи обмежень (6.53) є додатними, і **виродженим**, якщо серед них є хоча б одне нульове значення.

Канонічну ЗЛП називають **невиродженою**, якщо усі її базисні розв'язки є невивродженими, і **виродженою**, якщо є виродженим хоча б один з базисних розв'язків.

Якщо цільова функція має екстремум у декількох вершинах багатогранника допустимих розв'язків, то вона досягає того самого значення у будь-якій точці, яка є їх лінійною комбінацією.

де α_{ij}, β_j – параметри, одержувані при розв’язанні системи (6.55).

Система (6.56) визначає допустимі розв’язки, якщо базисні змінні x_1, \dots, x_M є невід’ємними за невід’ємних значень вільних змінних x_{M+1}, \dots, x_{M+P} . Якщо ж невід’ємними є параметри β_j , то за умови $x_{M+1} = \dots = x_{M+P} = 0$ ця система визначає базисний розв’язок. Від’ємні значення одного або декількох параметрів β_j свідчать про те, що поділ змінних на базисні та вільні був невдалим, і необхідно обрати інший варіант поділу.

Цільову функцію необхідно виразити через вільні змінні і записати у вигляді:

$$f(\mathbf{X}) = d_0 + d_1 x_{M+1} + \dots + d_p x_{M+P}, \quad (6.57)$$

де d_k – коефіцієнти, одержувані при підстановці (6.56) до (6.44).

Оскільки екстремум цільової функції досягається на одному з базисних розв’язків, то мінімуму (максимуму) має відповідати невід’ємність (неододатність) усіх коефіцієнтів d_k . Якщо ця умова не виконується, то наявність від’ємних (додатних) коефіцієнтів призведе до зменшення (збільшення) значення цільової функції порівняно з базисним розв’язком при невід’ємних значеннях відповідних вільних змінних. У разі, коли отриманий набір коефіцієнтів d_k містить хоча б одне від’ємне (додатне) значення, то необхідно обрати інший набір базисних змінних і повторити процес розв’язання. Це здійснюється, доки не буде отримано базисний розв’язок, для якого всі коефіцієнти d_k цільової функції будуть невід’ємними (недодатними). Повна кількість можливих варіантів збігається з кількістю можливих базисних планів C_N^M , тому метод перебору можна використовувати лише за достатньо малих значень M та N , що досить рідко виконується для реальних задач.

Графічний метод базується на геометричній інтерпретації ЗЛП. У випадку двох змінних будь-яке рівняння вигляду

$$a_{i1}x_1 + a_{i2}x_2 = b_i \quad (6.58)$$

задає на площині пряму, а нерівність вигляду

$$a_{i1}x_1 + a_{i2}x_2 \geq (\leq) b_i \quad (6.59)$$

– півплощину.

Побудувавши усі півплощини, які задає система обмежень ЗЛП, одержимо, як їх перетин, багатокутник розв'язків. З урахуванням неосновних обмежень він буде розміщуватися у першій чверті координатної площини.

Лінії рівня цільової функції задає рівняння:

$$c_1x_1 + c_2x_2 = \text{const} . \quad (6.60)$$

Напрямок найшвидшого зростання цільової функції задає її градієнт, а напрям її найшвидшого згасання – антиградієнт. Обидва ці вектори перпендикулярні лініям рівня. З визначення градієнта випливає, що $\mathbf{g} = (c_1; c_2)$.

Для знаходження точки мінімуму цільової функції необхідно спочатку побудувати одну з її ліній рівня. Зазвичай беруть лінію $c_1x_1 + c_2x_2 = 0$, яка проходить через початок координат. Після цього її переміщують у напрямі градієнта або антиградієнта доки вона не перетне першу або останню (залежно від вигляду цільової функції та шуканого екстремуму) у відповідному напрямі точку багатокутника розв'язків. Є можливою ситуація, коли через цю точку проходить сторона багатокутника розв'язків, паралельна лініям рівня. Тоді усі точки, що лежать на цій стороні, будуть точками відповідного екстремуму цільової функції. Також можливо, що багатокутник розв'язків є необмеженим, й тоді цільова функція не має екстремуму, тобто $\min f(\mathbf{X}) = -\infty$ або $\max f(\mathbf{X}) = +\infty$.

ЗЛП не має розв'язків, якщо:

◆ не існує багатокутника розв'язків, тобто система обмежень не має області невід'ємних розв'язків;

◆ лінія рівня перетинає необмежений багатокутник розв'язків, не торкаючись його у жодній кутовій точці.

Розглянемо такий приклад. Для виготовлення двох видів продукції P_1 та P_2 використовують три види сировини: S_1 , S_2 та S_3 . Цільовою функцією є $f(\mathbf{X}) = 50x_1 + 30x_2$. Система обмежень:

$$\begin{cases} 5x_1 + 6x_2 \leq 30; \\ 7x_1 + 5x_2 \leq 35; \\ 4x_1 + 9x_2 \leq 36. \end{cases} \quad (6.61)$$

Необхідно знайти, якими мають бути обсяги вироблення продуктів, щоб цільова функція була максимальною.

На рис. 6.5 показано графічний розв'язок ЗЛП. Спочатку будемо відокремлювальну пряму $L_1 : 5x_1 + 6x_2 = 30$. Для обрання півплощини розв'язку нерівності обираємо будь-яку точку, наприклад $(0, 0)$, і перевіряємо, чи виконується для неї нерівність. Якщо так, обираємо півплощину, у якій лежить ця точка, якщо ні – обираємо іншу півплощину.

Потім аналогічно будемо відокремлювальні прямі $L_2 : 7x_1 + 5x_2 = 35$ та $L_3 : 4x_1 + 9x_2 = 36$ і обираємо півплощини розв'язків. Враховуючи неосновні обмеження, отримуємо п'ятикутник, який на рисунку виділений жирними лініями.

Будемо лінію рівня $f(\mathbf{X}) = 50x_1 + 30x_2 = 0$. Вона проходить через точку $(0; 0)$ перпендикулярно до градієнта $\mathbf{g} = (50; 30)$. На рис. 6.5 замість градієнта побудовано соннаправлений йому вектор $\mathbf{g}_1 = (5; 3)$.

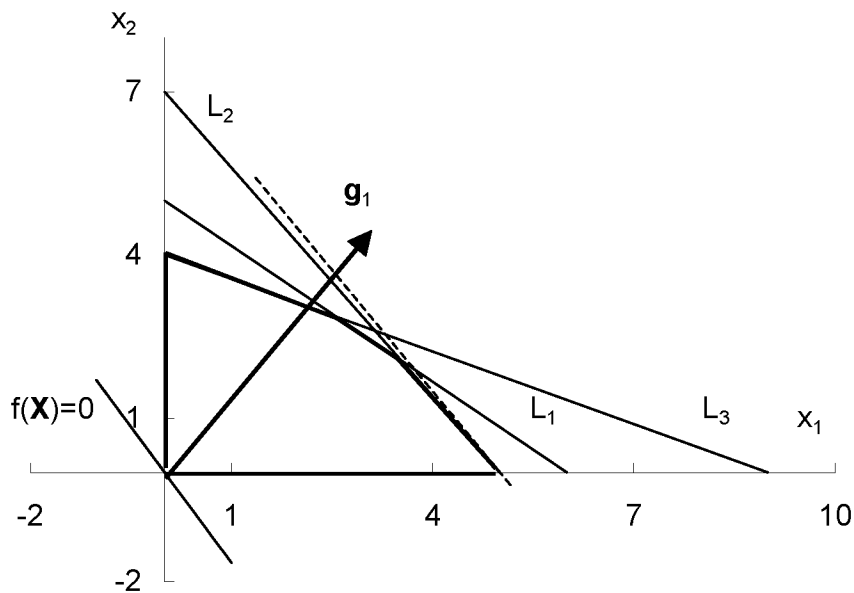


Рис. 6.5. Графічний розв'язок ЗЛП

Переміщуючи лінію рівня вздовж градієнта, одержимо, що максимум цільової функції $f(\mathbf{X}) = 250$ досягається у точці перетину відокремлювальної прямої L_2 з віссю абсцис $(0; 5)$. Водночас можна помітити, що мінімум цільової функції $f(\mathbf{X}) = 0$ був би досягнутий у точці $(0; 0)$.

У випадку, коли точкою шуканого екстремуму є точка перетину двох відокремлювальних прямих, її координати можна знайти як розв'язок відповідної системи рівнянь. Наприклад, координати точки перетину прямих L_1 і L_2 можна отримати, розв'язуючи систему:

$$\begin{cases} 5x_1 + 6x_2 = 30; \\ 7x_1 + 5x_2 = 35. \end{cases} \quad (6.62)$$

При розв'язанні цієї задачі методом перебору потрібно було б перевірити $C_5^3 = \frac{5!}{3!2!} = 10$ варіантів, зокрема значення цільової функції ще у трьох точках, які відповідають вершинам отриманого багатокутника розв'язків, що залиши-

лися. Останні дорівнюють: 120 для точки (0; 4) перетину відокремлювальної прямої L_3 з віссю ординат, приблизно 238 для точки перетину відокремлювальних прямих L_1 і L_2 , координати якої є близькими до (3,53; 2,06), та приблизно 214 для точки перетину відокремлювальних прямих L_1 та L_3 з координатами, близькими до (2,57; 2,86). Усі вони є меншими за 250, що свідчить про правильність отриманого розв'язку ЗЛП. Більша кількість базисних розв'язків при застосуванні методу перебору пов'язана з тим, що у цьому випадку ми розглядаємо більшу кількість базисних змінних (три а не дві) і, відповідно, маємо не плоский багатокутник розв'язків, а тривимірний багатогранник.

У загальному випадку для розв'язання даної задачі методом перебору слід перевести її до канонічного вигляду. Для цього треба увести три додаткові "слабкі" змінні $y_1, y_2, y_3 \geq 0$. Тоді розв'язувана задача може бути записана у такій спосіб:

$$\begin{cases} f(\mathbf{X}) = 50x_1 + 30x_2 \rightarrow \max; \\ 5x_1 + 6x_2 + y_1 = 30; \\ 7x_1 + 5x_2 + y_2 = 35; \\ 4x_1 + 9x_2 + y_3 = 36; \\ x_1, x_2, y_1, y_2, y_3 \geq 0. \end{cases} \quad (6.63)$$

З наявних п'яти змінних треба виділити 3 базові та 2 вільні. Оберемо як базові змінні x_1, x_2, y_1 . Перепишемо систему обмежень у вигляді:

$$\begin{cases} 5x_1 + 6x_2 + y_1 = 30; \\ 7x_1 + 5x_2 = 35 - y_2; \\ 4x_1 + 9x_2 = 36 - y_3. \end{cases} \quad (6.64)$$

Розв'язуючи її, отримуємо:

$$\begin{cases} 5x_1 + 6x_2 + y_1 = 30; \\ 28x_1 + 20x_2 = 140 - 4y_2; \\ 28x_1 + 63x_2 = 252 - 7y_3. \end{cases} \quad (6.65)$$

$$\begin{cases} 28x_1 = \frac{3780}{43} - \frac{252}{43}y_2 + \frac{140}{43}y_3; \\ 43x_2 = 112 + 4y_2 - 7y_3; \\ y_1 = -\frac{57}{43} + \frac{69}{43}y_2 - \frac{67}{43}y_3. \end{cases} \quad (6.66)$$

Вільний член у виразі для змінної y_1 є від'ємним. Це означає, що поділ змінних на вільні та базисні був невдалим і необхідно обрати інший варіант поділу змінних.

Оберемо як базові змінні x_1, x_2, y_2 . Тоді:

$$\begin{cases} 5x_1 + 6x_2 = 30 - y_1; \\ 7x_1 + 5x_2 + y_2 = 35; \\ 4x_1 + 9x_2 = 36 - y_3. \end{cases} \quad (6.67)$$

Розв'язуючи її, отримуємо:

$$\begin{cases} 20x_1 + 24x_2 = 120 - 4y_1; \\ 7x_1 + 5x_2 + y_2 = 35; \\ 20x_1 + 45x_2 = 180 - 5y_3. \end{cases} \quad (6.68)$$

$$\begin{cases} x_1 = \frac{18}{7} - \frac{3}{7}y_1 + \frac{2}{7}y_3; \\ x_2 = \frac{60}{21} + \frac{4}{21}y_1 - \frac{5}{21}y_3 \\ y_2 = \frac{19}{7} + \frac{43}{21}y_1 - \frac{17}{21}y_3. \end{cases} \quad (6.69)$$

У цьому випадку вільні члени у виразах для усіх змінних є додатними.
Цільова функція:

$$f(\mathbf{X}) = \frac{1500}{7} - \frac{110}{7}y_1 + \frac{50}{7}y_3. \quad (6.70)$$

Коефіцієнт при змінній y_3 є додатним, тому отриманий базисний розв'язок $X = \left(\frac{18}{7}, \frac{60}{21}, \frac{19}{7}\right)$ не є оптимальним. Він відповідає точці перетину відокремлювальних прямих L_1 та L_3 , одержаній раніше при графічному розв'язанні ЗЛП. Тому необхідно розглядати інший варіант поділу змінних на базисні та вільні. Ми не будемо розглядати усі інші варіанти, підкреслимо лише, що процес перебору можна зупинити після одержання розв'язку, для якого коефіцієнти при змінних у виразі для цільової функції виявляться від'ємними при пошуку максимуму або додатними при пошуку мінімуму.

Симплекс-метод. Методи перебору та графічний можна застосовувати лише за невеликих значень кількості змінних та обмежень ЗЛП. У реальних задачах вони можуть сягати кількох десятків або сотень. Основним методом розв'язання таких задач є симплекс-метод. У його основі лежить ідея скорочення кількості потрібних обчислень за рахунок цілеспрямованого руху до екстремуму при проведенні ітерацій. При цьому на кожному кроці ітерацій здійснюється перехід до тієї з сусідніх вершин багатогранника розв'язків, якій відповідає найменше (при пошуку мінімуму) або найбільше (при пошуку максимуму) значення цільової функції.

На початковому етапі реалізації симплекс-методу ЗЛП записують у канонічному вигляді, виокремлюють M базисних та $P = N - M$ вільних змінних і виражають базисні змінні і цільову функцію через вільні змінні згідно з (6.56, 6.57). У подальшому будемо застосовувати позначення y_1, y_2, \dots, y_M для базисних змінних і x_1, \dots, x_P – для вільних. Тоді (6.56, 6.57) набувають вигляду:

$$\begin{cases} y_1 = \beta_1 + \alpha_{11}x_1 + \dots + \alpha_{1P}x_P; \\ \dots; \\ y_M = \beta_M + \alpha_{M1}x_1 + \dots + \alpha_{MP}x_P, \end{cases} \quad (6.71)$$

$$f(\mathbf{X}) = d_0 + d_1x_1 + \dots + d_Px_P. \quad (6.72)$$

У разі, коли усі β_i є невід'ємними, отримуємо допустимий розв'язок. При цьому відповідний базисний розв'язок матиме вигляд:

$$y_1 = \beta_1; y_2 = \beta_2; \dots; y_M = \beta_M; f = d_0. \quad (6.73)$$

Якщо ж хоча б одне із значень β_i є від'ємним, необхідно змінити поділ змінних на базисні та вільні і заново виконати процедуру одержання базисного розв'язку.

Базисний розв'язок, для якого усі коефіцієнти d_k ($k = 1, \dots, P$) є додатними (від'ємними), відповідає мінімуму (максимуму) цільової функції. Якщо ж один або декілька коефіцієнтів є від'ємними (додатними), то функція цілі буде згасати (зростати) при збільшенні відповідних вільних змінних. Такі вільні змінні доцільно перевести у базисні. Цей процес у симплекс-методі здійснюють поетапно, починаючи з тієї змінної, яка має найбільший за модулем коефіцієнт. Позначимо її як x_B .

Одночасно треба перевести одну з базисних змінних у вільні. Для визначення, яку саме з базисних змінних обрати для цього, врахуємо, що із збільшенням змінної x_B водночас з потрібним зменшенням (збільшенням) цільової функції деякі з базисних змінних можуть зменшуватися, що призведе до порушення умови їх невід'ємності. Для визначення змінної, яка має бути виключеною з базису, перепишемо систему обмежень (6.71) у припущенні, що всі вільні змінні, крім x_B , дорівнюють нулю:

$$\begin{cases} y_1 = \beta_1 + \alpha_{1B}x_B; \\ \dots \dots \dots; \\ y_M = \beta_M + \alpha_{MB}x_B, \end{cases} \quad (6.74)$$

і розв'яжемо відносно x_B рівняння вигляду:

$$y_i = 0. \quad (6.75)$$

При цьому одержимо: $x_B = -\beta_i / \alpha_{iB}$. Змінну y_F , для якої величина $-\beta_F / \alpha_{FB}$ має мінімальне додатне значення, переведемо із базисних до вільних. Відповідне рівняння у (6.71) називають головним, або розв'язувальним рівнянням. Якщо певна з величин $\beta_i = 0$, то відповідну змінну y_i виключають із базису за $\alpha_{iB} < 0$ і не розглядають при визначенні мінімального додатного значення $-\beta_i / \alpha_{iB}$ за $\alpha_{iB} > 0$.

У разі, коли усі розв'язки x_B рівнянь (6.75) є від'ємними, то цільова функція необмежена знизу (зверху), і оптимальний план ЗЛП не існує.

Після визначення x_B та y_F необхідно виразити нові базисні змінні і функцію цілі через новий набір вільних змінних. Невід'ємність нових значень β_i забезпечує алгоритм симплекс-методу. Далі цикл повторюють, доки всі коефіцієнти d_k не будуть додатними (від'ємними), що відповідає мінімуму (максимуму) цільової функції.

Розглянемо такий приклад. Знайти значення змінних x_1, x_2, x_3 , що мінімізують цільову функцію $f(\mathbf{X}) = -8x_1 - 10x_2 + 3x_3$ за обмежень:

$$\begin{cases} 3x_1 + 4x_2 - 2x_3 \leq 300; \\ 3x_1 + x_2 \leq 120; \\ x_1 + 2x_2 \leq 500; \\ x_3 \leq 200 \end{cases} \quad (6.76)$$

з урахуванням природних обмежень $x_1, x_2, x_3 \geq 0$.

Запишемо цю задачу у канонічній формі. При цьому система обмежень набуває вигляду:

$$\begin{cases} 3x_1 + 4x_2 - 2x_3 + y_1 = 300; \\ 3x_1 + x_2 + y_2 = 120; \\ x_1 + 2x_2 + y_3 = 500; \\ x_3 + y_4 = 200. \end{cases} \quad (6.77)$$

Кількість змінних $N = 7$, кількість рівнянь $M = 4$. Таким чином, треба виділити 4 базисних та 3 вільних змінних. Як базисні змінні зручно взяти y_1, y_2, y_3, y_4 , котрі легко виразити через змінні x_1, x_2, x_3 , що у такому разі залишаються вільними:

$$\begin{cases} y_1 = 300 - 3x_1 - 4x_2 + 2x_3; \\ y_2 = 120 - 3x_1 - x_2; \\ y_3 = 500 - x_1 - 2x_2; \\ y_4 = 200 - x_3. \end{cases} \quad (6.78)$$

Далі необхідно перевірити невід'ємність коефіцієнтів β_i . У даному випадку $\beta_1 = 300, \beta_2 = 120, \beta_3 = 500, \beta_4 = 200$, тобто усі $\beta_i \geq 0$. Тому отриманий розв'язок є допустимим і йому відповідає базисний розв'язок $Y = (300, 120, 500, 200)$.

Вираз для цільової функції $f(X) = -8x_1 - 10x_2 + 3x_3$ залишається незмінним, оскільки він містить тільки вільні змінні. Значення цільової функції, яке відповідає знайденому базисному розв'язку: $f(X) = 0$.

Від'ємні коефіцієнти у виразі для цільової функції відповідають змінним x_1 та x_2 , при цьому більшим за модулем є коефіцієнт при x_2 , тому цю змінну переводимо до базису.

Для визначення змінної, яка виключається з базису на основі заданої системи обмежень, складаємо рівності:

$$\begin{cases} y_1 = 300 - 4x_2 = 0; \\ y_2 = 120 - x_2 = 0; \\ y_3 = 500 - 2x_2 = 0. \end{cases} \Rightarrow \begin{cases} x_2 = 75; \\ x_2 = 120; \\ x_2 = 250. \end{cases} \quad (6.79)$$

Мінімальне додатне значення відповідає змінній y_1 , тому саме її переводимо із базисних до вільних. Таким чином, новими базисними змінними є x_2, y_2, y_3, y_4 , а вільними – x_1, x_3, y_1 . Далі переходимо до наступного кроку ітерацій.

Виражаємо нову базисну змінну через вільні:

$$x_2 = 75 - 0,75x_1 + 0,5x_3 - 0,25y_1. \quad (6.80)$$

Підставляючи цей вираз до формул для інших базисних змінних, отримаємо:

$$\begin{cases} x_2 = 75 - 0,75x_1 + 0,5x_3 - 0,25y_1; \\ y_2 = 45 - 2,25x_1 - 0,5x_3 + 0,25y_1; \\ y_3 = 350 + 0,5x_1 - x_3 + 0,5y_1; \\ y_4 = 200 - x_3. \end{cases} \quad (6.81)$$

Цільова функція набуває вигляду:

$$f(\mathbf{X}) = -750 - 0,5x_1 - 2x_3 + 2,5y_1. \quad (6.82)$$

Таким чином, новим базисним розв'язком є $x_2 = 75$, $y_2 = 45$, $y_3 = 350$, $y_4 = 200$, а відповідне значення цільової функції $f = -750$.

У виразі для цільової функції від'ємними є значення коефіцієнтів при x_1 та x_3 , при цьому більшим за модулем є коефіцієнт при x_3 . Тому саме цю змінну переводимо до базису.

Для визначення змінної, яку необхідно виключити з базису, складаємо систему:

$$\begin{cases} x_2 = 75 + 0,5x_3 = 0; \\ y_2 = 45 - 0,5x_3 = 0; \\ y_3 = 350 - x_3 = 0; \\ y_4 = 200 - x_3. \end{cases} \quad (6.83)$$

Звідси маємо:

$$\begin{cases} x_3 = -150; \\ x_3 = 90; \\ x_3 = 350; \\ x_3 = 200. \end{cases} \quad (6.84)$$

Мінімальне додатне значення відповідає змінній y_2 , яку ми і виключаємо із базису. Таким чином, новим набором базисних змінних є x_2 , x_3 , y_3 , y_4 , а нові вільні змінні – x_1 , y_1 , y_2 .

Переходимо до третього кроку ітерацій. Виражаємо нову базисну змінну x_3 через вільні:

$$x_3 = 90 - 4,5x_1 + 0,5y_1 - 2y_2. \quad (6.85)$$

Підставляючи цю формулу у вирази для інших базових змінних, одержуємо:

$$\begin{cases} x_2 = 120 - 3x_1 - y_2; \\ y_3 = 260 + 5x_1 + 2y_2; \\ y_4 = 110 + 4,5x_1 - 0,5y_1 + 2y_2. \end{cases} \quad (6.86)$$

Цільова функція з новими вільними змінними набуває вигляду:

$$f(\mathbf{X}) = -930 + 8,5x_1 + 1,5y_1 + 4y_2 \quad (6.87)$$

Таким чином новим базисним розв'язком є: $x_2 = 120$, $x_3 = 90$, $y_3 = 260$, $y_4 = 110$. Йому відповідає значення цільової функції $f = -930$. Усі коефіцієнти β_i у формулі для цільової функції є додатними. Таким чином ми можемо зробити висновок, що отриманий базисний розв'язок є оптимальним планом. У даному випадку ми одержали його на третьому кроці ітерацій. Загальна ж кількість базисних планів цієї задачі $C_7^4 = \frac{7!}{3! \cdot 4!} = 35$.

Якщо задача оптимізації є невивроженою (усі $\beta_i > 0$), то послідовність значень цільової функції є строго монотонною і збігається до мінімуму або максимуму. В іншому випадку вона є нестрого монотонною, і можливе зациклювання алгоритму. Найпростіший спосіб виходу з такої ситуації полягає в обранні, якщо це можливо, нової змінної, яка уводиться до базису або виключається з нього. Але у загальному випадку ця проблема залишається невирішеною.

6.4. Транспортна задача

Окремі задачі лінійного програмування допускають простіші способи розв'язання порівняно із розглянутим вище симплекс-методом. Однією з них є транспортна задача.

У загальному випадку її формулюють так. Є M пунктів відправлення (постачальників) A_1, A_2, \dots, A_M та N пунктів призначення (споживачів) B_1, B_2, \dots, B_N . Постачальники володіють запасами певного вантажу у кількості $a_1, a_2, \dots,$

a_M , відповідно. Споживачам цей вантаж потрібний у кількості b_1, b_2, \dots, b_N одиниць. Вартість перевезення вантажу від постачальника a_i до споживача b_j дорівнює c_{ij} . Необхідно розробити план перевезень, що забезпечує задоволення замовлень усіх споживачів або постачання споживачам усього наявного запасу вантажу за мінімальної сумарної вартості перевезень.

Позначимо кількість вантажу, що перевозиться від i -го постачальника до j -го споживача як x_{ij} . Величини x_{ij} у подальшому будемо називати перевезеннями. Тепер ми можемо записати транспортну задачу у математичній формі:

$$f(\mathbf{X}) = \sum_{i=1}^M \sum_{j=1}^N c_{ij} x_{ij} \rightarrow \min ; \quad (6.88)$$

$$\sum_{j=1}^N x_{ij} \leq a_i \quad (i = 1, \dots, M); \quad (6.89)$$

$$\sum_{i=1}^M x_{ij} \leq b_j \quad (j = 1, \dots, N); \quad (6.90)$$

$$x_{ij} \geq 0. \quad (6.91)$$

Вираз (6.88) є цільовою функцією, нерівності (6.89) – обмеженнями, що пов'язані з обмеженістю запасу у постачальників, а нерівності (6.90) пов'язані із задоволенням замовлень. Розгляд ситуації $\sum_{i=1}^M x_{ij} \geq b_j$ є недоцільним, оскільки в цьому випадку, очевидно, мінімум цільової функції буде досягнуто за умови $\sum_{i=1}^M x_{ij} = b_j$, яку враховують нерівності (6.90). Нерівності (6.91) відображають природні обмеження.

Можливість спрощення методів розв'язання транспортної задачі зумовлена тим, що усі коефіцієнти при змінних в обмеженнях (6.89, 6.90) є сталими і дорівнюють одиниці.

Для розв'язання транспортної задачі складають **транспортну таблицю** (табл. 6.1).

Транспортна таблиця

П/С	B_1	B_2	...	B_j	...	B_N	Запас
A_1	c_{11}	c_{12}	...	c_{1j}	...	c_{1N}	a_1
...
A_M	c_{M1}	c_{M2}	...	c_{Mj}	...	c_{MN}	a_M
Замовлення	b_1	b_2	...	b_j	...	b_N	$\sum b_j \setminus \sum a_i$

У правому верхньому куті внутрішніх комірок записують величини c_{ij} , а у нижніх частинах цих комірок – значення перевезень, що відповідають накладеним обмеженням. У крайніх правих комірках вказують запаси постачальників, а у нижніх – потреби споживачів. Внутрішні комірки, які відповідають ненульовим перевезенням, називають базисними, а ті, що залишилися, – вільними.

Якщо $\sum_{i=1}^M a_i = \sum_{j=1}^N b_j$, транспортну задачу називають **збалансованою**, а у протилежному випадку – **незбалансованою**, або **відкритою**. У випадку заборони перевезень між деякими пунктами, це можна відобразити призначенням для них достатньо високої вартості.

Для збалансованої транспортної задачі система обмежень має вигляд:

$$\sum_{j=1}^N x_{ij} = a_i; \quad \sum_{i=1}^M x_{ij} = b_j. \quad (6.92)$$

Вона містить $M \times N$ змінних та $M + N$ основних обмежень, одне з яких не є незалежним, оскільки впливає з інших з урахуванням рівняння балансу. Таким чином, ранг системи обмежень $r = M + N - 1$, і треба обрати r базисних змінних. Кількість вільних змінних

$$p = MN - M - N + 1 = (M - 1)(N - 1). \quad (6.93)$$

Допустимий розв'язок збалансованої транспортної задачі має задовольняти такі вимоги:

- ◆ сума перевезень за кожним рядком таблиці повинна дорівнювати запасам відповідного постачальника;
- ◆ сума перевезень за кожним стовпцем повинна дорівнювати замовленню відповідного споживача.

Як і для інших ЗЛП, загальна схема розв'язання транспортної задачі передбачає знаходження допустимого розв'язку та перехід від нього до оптимального плану за певною ітераційною процедурою. На відміну від довільної ЗЛП, розв'язок збалансованої транспортної задачі завжди існує.

Для знаходження допустимого розв'язку, що буде використовуватися як початкове наближення, найчастіше застосовують методи **північно-західного кута**, **мінімальної вартості перевезень** та **Фогеля**. Розглянемо перший з них.

Алгоритм побудови допустимого плану у цьому випадку є таким. Спочатку перевозимо вантаж із пункту A_1 до пункту B_1 . Цьому перевезенню надають максимально можливого значення. Якщо замовлення споживача B_1 є меншим за запас a_1 , то залишок перевозиться до споживача B_2 . У випадку, коли й після цього у постачальника A_1 залишається певний запас вантажу, то його перевозять у такому самому порядку до споживачів $B_3, B_4 \dots$. Якщо запас a_1 виявляється недостатнім для задоволення замовлення b_1 , то необхідну кількість вантажу перевозять від постачальника A_2 . У випадку, коли і він виявиться неспроможним забезпечити виконання замовлення, то до пункту B_1 послідовно перевозять вантаж від постачальників $A_3, A_4 \dots$. Після виконання замовлення B_1 , за аналогічною схемою послідовно задовольняють замовлення споживачів B_2, B_3 й т. д. Метод північно-західного кута завжди дає можливість отримати допустимий розв'язок, але цей розв'язок може виявитися досить далеким від оптимального.

Розглянемо таку задачу. Запаси певного вантажу у постачальників A_1 – A_4 дорівнюють, відповідно, 35, 40, 25 та 25 одиниць. Споживачам B_1 – B_5 необхідно отримати, відповідно, 30, 40, 13, 23 й 19 одиниць цього вантажу. Матриця вартості перевезень дорівнює:

$$C = \begin{pmatrix} 9 & 6 & 5 & 7 & 7 \\ 4 & 5 & 4 & 3 & 8 \\ 7 & 6 & 5 & 6 & 7 \\ 8 & 7 & 6 & 7 & 6 \end{pmatrix}.$$

Треба побудувати допустимий план перевезень.

Спочатку заносимо вихідні дані до транспортної таблиці 6.2

Таблиця 6.2

Вихідна транспортна таблиця для задачі, що розглядається

ПС	B_1	B_2	B_3	B_4	B_5	Запас
A_1	9	6	5	7	7	35
A_2	4	5	4	3	8	40
A_3	7	6	5	6	7	25
A_4	8	7	6	7	6	25
Замовлення	30	40	13	23	19	125

Споживачу B_1 потрібна менша кількість вантажу, ніж та, що є у постачальника A_1 . Тому повністю задовольняємо його замовлення за рахунок перевезення 30 одиниць вантажу від A_1 до B_1 . Залишок запасу постачальника A_1 у кількості 5 одиниць перевозимо до споживача B_2 . Йому необхідно додатково пос-

тачати 35 одиниць. Така кількість вантажу є у постачальника A_2 , звідки й перевозимо його до B_2 . Залишок запасу постачальника A_2 у кількості 5 одиниць перевозимо до споживача B_3 . Додатково для повного задоволення замовлення B_3 перевозимо до нього 8 одиниць вантажу від постачальника A_3 . Залишок запасу постачальника A_3 у кількості 17 одиниць перевозимо до споживача B_4 . Для того, щоб задовольнити його потреби, необхідно додати ще 6 одиниць вантажу, які перевозимо від постачальника A_4 . Залишок запасу A_4 , який повністю задовольняє потребу споживача B_5 , перевозимо до останнього. У підсумку отримуємо наведену нижче транспортну таблицю 6.3, яка дає допустимий розв'язок задачі.

Таблиця 6.3

Транспортна таблиця, що дає допустимий розв'язок задачі

ПС	B_1	B_2	B_3	B_4	B_5	Запас
A_1	9 30	6 5	5	7	7	35
A_2	4	5 35	- 4 5	+ 3	8	40
A_3	7	6	+ 5 8	- 6 17	7	25
A_4	8	7	6	7 6	6 19	25
Замовлення	30	40	13	23	19	125

Знаходимо значення цільової функції:

$$f = 9 \cdot 30 + 6 \cdot 5 + 5 \cdot 35 + 4 \cdot 5 + 5 \cdot 8 + 6 \cdot 17 + 7 \cdot 6 + 6 \cdot 19 = 793 \text{ одиниці.}$$

Наступним етапом розв'язування задачі є оптимізація отриманого допустимого плану. В основу методу оптимізації, який ми будемо застосовувати, по-

кладено ідею, що в оптимальному плані мають бути максимально використані перевезення з мінімальними тарифами.

Для реалізації цієї ідеї знаходимо вільну комірку з мінімальним тарифом і переводимо її до базису. У нашій таблиці мінімальним є тариф для перевезення від постачальника A_2 до споживача B_4 , який дорівнює 3 одиницям. Черговим етапом оптимізації є побудова на основі визначеної комірки циклу, який не порушує умов допустимості плану, але надає можливість зменшити значення цільової функції. Основні правила побудови циклу є такими:

- ◆ цикл має містити парну кількість комірок, одна з яких є вільною, а інші – базисними;

- ◆ комірки у циклі можна з'єднувати за вертикаллю або за горизонталлю.

На комірці A_2B_4 ми можемо побудувати лише один цикл $A_2B_4 \rightarrow A_3B_4 \rightarrow A_3B_3 \rightarrow A_2B_3$. Позначаємо його у транспортній таблиці пунктирною лінією. Легко перевірити, що при збільшенні перевезень A_2B_4 та A_3B_3 і одночасному зменшенні перевезень A_3B_4 та A_2B_3 на одну й ту саму величину, обмеження (6.92) не порушуються. Для того, щоб визначити, зменшиться, чи збільшиться значення цільової функції при переміщенні одиниці вантажу за циклом, достатньо обчислити вартість такого переміщення. Якщо вона буде додатною, значення цільової функції збільшиться, якщо від'ємною – зменшиться. У випадку, коли вартість перевезення за циклом дорівнює нулю, значення цільової функції не змінюється. Для обчислення вартості перевезення за циклом ставимо знак "+" у вільну комірку транспортної таблиці, на якій будується цикл (у даному випадку це комірка A_2B_4), а потім за чергою ставимо знаки "-" та "+" в інші комірки циклу. Знак "+" означає, що відповідне перевезення буде збільшено, а знак "-", що воно буде зменшено. Вартість побудованого нами циклу, тобто вартість перевезення за ним одиниці вантажу, дорівнює:

$$q = +3 - 6 + 5 - 4 = -2.$$

Це свідчить про те, що його застосування дає можливість зменшити значення цільової функції. Максимальна кількість вантажу, яку ми можемо перемістити за цим циклом, очевидно, дорівнює мінімальному значенню перевезення у тих комірках, де ми зменшуємо обсяг перевезень. На даному етапі вона дорівнює $\min\{x_{23}, x_{34}\} = 5$ одиницям вантажу. З погляду загального підходу до розв'язання ЗЛП, зменшення перевезення x_{23} на 5 одиниць означає переведення відповідної змінної з базисних до вільних. Здійснивши відповідні зміни у плані перевезень, отримуємо новий допустимий план (табл. 6.4):

Таблиця 6.4

Покращений допустимий план для задачі, що розглядається

П/С	B ₁	B ₂	B ₃	B ₄	B ₅	Запас
A ₁	- 9	+ 6	5	7	7	35
	30	5				
A ₂	+ 4	- 5	4	3	8	40
		35		5		
A ₃	7	6	5	6	7	25
			13	12		
A ₄	8	7	6	7	6	25
				6	19	
Замовлення	30	40	13	23	19	125

Значення цільової функції:

$$f = 9 \cdot 30 + 6 \cdot 5 + 5 \cdot 35 + 3 \cdot 5 + 5 \cdot 13 + 6 \cdot 12 + 7 \cdot 6 + 6 \cdot 19 = 783 \text{ одиниці.}$$

У новому допустимому плані мінімальний тариф, що дорівнює чотирьом одиницям, мають вільні комірочки A_2B_1 і A_2B_3 . До базису будемо переводити комірочку A_2B_1 , оскільки A_2B_3 стала базисною тільки на попередньому етапі. Єдиним циклом, який можна побудувати на цій комірці, є $A_2B_1 \rightarrow A_1B_1 \rightarrow A_1B_2 \rightarrow$

A_2B_2 . Помічаємо значком "+" комірки A_2B_1 та A_1B_2 , значком "-" – комірки A_1B_1 та A_2B_2 і розраховуємо вартість циклу $q = +4 - 9 + 6 - 5 = -4$ од. Вона є від'ємною, таким чином, використання циклу дає можливість зменшити сумарну вартість перевезень. Максимальний обсяг вантажу, що можна перерозподілити за цим циклом, дорівнює $\min\{x_{11}, x_{22}\} = 30$. Вносимо відповідні зміни і одержуємо новий допустимий план (табл. 6.5):

Таблиця 6.5

Ще кращий допустимий план для задачі, що розглядається

П/С	B_1	B_2	B_3	B_4	B_5	Запас
A_1	9	6 35	5	7	7	35
A_2	4 30	- 5 5	4	+ 3 5	8	40
A_3	7	+ 6	5 13	6 12	7	25
A_4	8	7	6	7 6	6 19	25
Замовлення	30	40	13	23	19	125

Значення цільової функції для нього:

$$f = 4 \cdot 30 + 6 \cdot 35 + 5 \cdot 5 + 5 \cdot 13 + 3 \cdot 5 + 6 \cdot 12 + 7 \cdot 6 + 6 \cdot 19 = 663 \text{ одиниці.}$$

У новій таблиці мінімальний тариф, що дорівнює чотирьом одиницям, має вільна комірка A_2B_3 . Але на її основі ми можемо побудувати лише цикл $A_2B_3 \rightarrow A_2B_4 \rightarrow A_3B_4 \rightarrow A_3B_3$, який ми вже використовували, і вартість якого у такій послідовності буде додатною. Наступними за величиною тарифу є вільна комірка A_1B_3 , де він дорівнює п'яти одиницям. На її основі ми можемо побуду-

вати цикл $A_1B_3 \rightarrow A_3B_3 \rightarrow A_3B_4 \rightarrow A_2B_4 \rightarrow A_2B_2 \rightarrow A_1B_2$ з вартістю $+2$. Його використання не призведе до покращення плану перевезень. Тому розглянемо наступні за величиною тарифу комірки A_3B_2 та A_4B_3 , де він дорівнює шести одиницям.

Комірка A_4B_3 породжує цикл $A_4B_3 \rightarrow A_3B_3 \rightarrow A_3B_4 \rightarrow A_4B_4$ з вартістю $q = +6 - 5 + 6 - 7 = 0$, застосування якого не зменшить значення цільової функції. Комірка A_3B_2 породжує цикл $A_3B_2 \rightarrow A_2B_2 \rightarrow A_2B_4 \rightarrow A_3B_4$ з вартістю $q = +6 - 5 + 3 - 6 = -2$ одиниці. Максимальний обсяг вантажу, що можна перерозподілити за цим циклом, дорівнює $\min\{x_{22}, x_{34}\} = 5$. Вносимо відповідні зміни і одержуємо новий допустимий план (табл. 6.6).

Таблиця 6.6

Оптимальний план для задачі, що розглядається

П/С	B ₁	B ₂	B ₃	B ₄	B ₅	Запас
A ₁	9	6 35	5	7	7	35
A ₂	4 30	5	4	3 10	8	40
A ₃	7	6 5	5 13	6 7	7	25
A ₄	8	7	6	7 6	6 19	25
Замовлення	30	40	13	23	19	125

Значення цільової функції:

$$f = 4 \cdot 30 + 6 \cdot 35 + 6 \cdot 5 + 5 \cdot 13 + 3 \cdot 10 + 6 \cdot 7 + 7 \cdot 6 + 6 \cdot 19 = 653 \text{ одиниці.}$$

Легко перевірити, що серед циклів, які породжують вільні комірки нової транспортної таблиці, немає таких, вартість яких є від'ємною. Тому отриманий план є оптимальним. Зазначимо, що оптимальний план ми отримали на четвертому кроці ітерацій. Загальна ж кількість допустимих планів даної задачі, що містить 20 невідомих та 8 незалежних основних обмежень, дорівнює

$$C_{20}^8 = \frac{20!}{12!8!} = 125970.$$

6.5. Динамічне програмування

Динамічне програмування є методом оптимізації, розробленим для багатокрокових (багатоетапних) операцій. Його було запропоновано Р. Беллманом у 50-х рр. XX в.

Розглянемо певну операцію O , що розпадається на m послідовних кроків. Нехай її ефективність характеризується певним показником W – виграшем, який можна подати як суму виграшів на окремих кроках w_i :

$$W = \sum_{i=1}^m w_i \quad (6.93)$$

Критерій, що виявляє таку властивість, називають **адитивним**.

Досліджувана операція є керованим процесом, тобто існують певні регульовані параметри, які впливають на його результати. При цьому на кожному кроці приймається рішення, що впливає як на величину виграшу на даному кроці, так і на величину виграшу за операцію у цілому. Це рішення далі будемо називати **кроковим управлінням**. Сукупність усіх крокових управлінь є управлінням операцією у цілому. Позначимо його як x , а крокові управління – як x_i ($i = 1, \dots, m$):

$$x = (x_1, x_2, \dots, x_m). \quad (6.94)$$

Крокові управління у загальному випадку можуть бути числами, векторами, функціями й т. д.

Необхідно визначити управління, що максимізує виграш:

$$W = \sum_{i=1}^m w_i \Rightarrow \max. \quad (6.95)$$

Відповідне управління x^* називають **оптимальним**. Воно складається з сукупності оптимальних крокових управлінь:

$$x^* = (x_1^*, x_2^*, \dots, x_m^*). \quad (6.96)$$

Виграш, що забезпечує оптимальне управління x^* :

$$W^* = \max_x \{W(x)\}. \quad (6.97)$$

В основі методу динамічного програмування лежить ідея покрокової оптимізації. При цьому управління на кожному кроці x_i вибирають так, щоб максимізувати суму виграшів на цьому кроці й усіх кроках, що залишилися до кінця операції (принцип оптимальності).

Але останній крок можна планувати так, щоб він самий давав найбільший виграш. Тому процедура динамічного програмування зазвичай передбачає попереднє планування від кінця до початку для визначення умовно оптимального управління і подальше повторне планування у зворотному напрямі, що дає змогу отримати безумовно оптимальне управління x^* .

Метод динамічного програмування є придатним для цілочислових розв'язків, нелінійних цільових функцій й різноманітних обмежень, що накладаються на розв'язки. Але, це веде до того, що на відміну від лінійного програмування у цьому випадку немає стандартної обчислювальної процедури.

Розв'язання задач динамічного програмування зручно виконувати у такій послідовності.

1. Обирають параметри (фазові координати), що характеризують стан досліджуваної системи перед кожним кроком.

2. Визначають кроки, з яких складається досліджувана операція.

3. Виявляють можливий набір крокових керувань x_i , для кожного кроку й обмеження, що накладаються на них.

4. Визначають виграш, який дає управління x_i на i -му кроці, якщо перед цим система перебувала у стані S , тобто записати функцію виграшу:

$$w_i = f_i(S, x_i). \quad (6.98)$$

5. Визначають, як змінюється стан системи під впливом управління x_i на i -му кроці. Новий стан:

$$S' = \varphi_i(S, x_i). \quad (6.99)$$

У загальному випадку аргументами функцій (6.98, 6.99) є вектори.

6. Записують основне рекурентне рівняння динамічного програмування, що відображає умовний оптимальний виграш $W_i(S)$ (починаючи з i -го кроку й до кінця) через вже відому функцію $W_{i+1}(S)$:

$$W_i(S) = \max_{x_i} \{f_i(S, x_i) + W_{i+1}(\varphi_i(S, x_i))\}. \quad (6.100)$$

Цьому виграшу відповідає умовне оптимальне управління на i -му кроці $x_i(S)$. У вже відому функцію $W_{i+1}(S)$ необхідно замість S підставити змінений стан $S' = \varphi_i(S, x_i)$.

7. Здійснюють умовну оптимізацію останнього (m -го) кроку, задаючись набором станів S , із яких можна за один крок дійти до кінцевого стану, обчислюючи для кожного з них умовний оптимальний виграш за формулою:

$$W_m(S) = \max_{x_m} \{f_m(S, x_m)\} \quad (6.101)$$

і визначаючи умовне оптимальне управління $x_m(S)$, для якого цей максимум досягається.

8. Здійснюють умовну оптимізацію $(m - 1)$ -го, $(m - 2)$ -го й т. д. кроків за формулою (6.100), беручи в ній $i = (m - 1), (m - 2), \dots, i$ для кожного із кроків вказують умовне оптимальне управління $x_i(S)$, за якого досягається максимум.

Якщо стан системи у початковий момент S_0 є відомим, то на першому кроці не треба варіювати стан системи. Оптимальний виграш знаходять для цього відомого стану. Він і буде оптимальним виграшем за всю операцію: $W^* = W_1(S_0)$.

9. Здійснюють безумовну оптимізацію управління. Для цього беруть знайдене оптимальне управління на першому кроці $x_1^* = x_1(S_0)$; змінюють стан по формулі (6.99); для отриманого стану визначають оптимальне управління на другому кроці x_2^* й т. д. до кінця.

Досі йшлося про адитивні задачі динамічного програмування, але метод можна застосовувати і до задач з мультиплікативним критерієм:

$$W = \prod_{i=1}^m w_i, \quad (6.102)$$

якщо виграші w_i є додатними. У цьому випадку в формулі (6.100) замість додавання необхідно виконувати множення.

Контрольні питання

1. Які завдання вирішують шляхом оптимізаційного моделювання систем?
2. Яку функцію називають опуклою, угнутою?

3. Які точки називають точками локального та глобального мінімуму і максимуму функції?
4. Які функції називають унімодальними? Наведіть приклади унімодальних функцій та функцій, що не є унімодальними.
5. Якою є загальна постановка задачі оптимізації?
6. Що називають цільовою функцією задачі оптимізації, допустимою множиною, допустимим розв'язком, керованими змінними?
7. У чому полягають достатні умови існування екстремуму функції?
8. Яку задачу називають задачею безумовної оптимізації?
9. Яку задачу називають задачею умовної оптимізації?
10. Що називають обмеженнями задачі умовної оптимізації? Які обмеження називають активними, пасивними?
11. У чому полягає необхідна умова локального екстремуму функції?
12. Які точки називають стаціонарними (критичними)?
13. Які точки називають сідловими?
14. Яку матрицю називають матрицею Гессе? Що є квадратичною формою матриці Гессе?
15. Як формулюються необхідна й достатня умови локального екстремуму функції через властивості її матриці Гессе?
16. Якими є алгоритми методів сканування й золотого перерізу пошуку локального екстремуму функції однієї змінної? У чому полягають переваги й недоліки цих методів? Чи існують певні обмеження на їх застосування?
17. Якою є загальна схема ітераційних методів розв'язування задач безумовної багатовимірної оптимізації?
18. Охарактеризуйте основні алгоритми розв'язування задач безумовної багатовимірної оптимізації? У чому полягають їх переваги й недоліки? Якими є обмеження на застосування окремих методів?
19. Якою є загальна постановка ЗЛП?
20. Записати задану ЗЛП у канонічному та у стандартному вигляді.

21. Яку задачу називають двоїстою до заданої ЗЛП? Назвіть основні правила побудови двоїстих задач.
22. Записати двоїсту задачу до ЗЛП.
23. Наведіть приклади ЗЛП та двоїстих до них задач.
24. Якими є можливі варіанти існування розв'язку ЗЛП?
25. Яким є алгоритм методу перебору розв'язування ЗЛП? У чому полягають його переваги й недоліки?
26. Як розв'язати ЗЛП графічним методом? Якими є переваги й недоліки графічного методу?
27. Яким є алгоритм симплекс-методу розв'язування ЗЛП? У чому полягають його переваги й недоліки?
28. Якою є загальна постановка транспортної ЗЛП?
29. Які властивості транспортної задачі дають змогу розробити для її розв'язання більш ефективні алгоритми, порівняно із загальними алгоритмами розв'язування ЗЛП.
30. Охарактеризуйте основні алгоритми розв'язування транспортної ЗЛП, їх переваги, недоліки й обмеження.
31. У чому полягає сутність методу динамічного програмування, для вирішення яких завдань його застосовують?

7. КОРЕЛЯЦІЙНИЙ АНАЛІЗ

Кореляцією (кореляційним зв'язком) між випадковими величинами (ознаками) називають наявність статистичного або імовірнісного зв'язку між ними. При цьому закономірна зміна певних ознак призводить до закономірної зміни середніх значень інших, пов'язаних з ними ознак. **Кореляційним аналізом** називають сукупність методів виявлення кореляційного зв'язку. Тому його можна застосовувати для формалізованого подання моделей зв'язків між окремими компонентами системи або між окремими процесами, що відбуваються в ній. Наявність кореляційного зв'язку не означає існування причинно-наслідкового зв'язку між досліджуваними ознаками. Зокрема, вона може бути зумовлена наявністю невідомих факторів, що впливають на них.

Замість гіпотези про наявність кореляційного зв'язку часто розглядають протилежну гіпотезу про відсутність зв'язку між досліджуваними величинами. Нехай ознака A має r рівнів A_1, A_2, \dots, A_r , а ознака B – s рівнів B_1, B_2, \dots, B_s . Їх вважають **незалежними**, якщо події "ознака A набуває значення A_i " та "ознака B набуває значення B_j " є незалежними для усіх можливих пар i, j , тобто:

$$P(A_i, B_j) = P(A_i)P(B_j). \quad (7.1)$$

Це можна сформулювати в інший спосіб: ознаки є незалежними, якщо значення ознаки A не впливає на ймовірності реалізації можливих значень ознаки B :

$$P(B_j / A_i) = P(B_j), \quad \forall(A_i, B_j). \quad (7.2)$$

Кореляційний аналіз здійснюють на початковому етапі вирішення всіх трьох основних проблем статистичного аналізу даних. У проблемі статистичного аналізу залежностей і побудови регресійних моделей він дає змогу встановити сам факт існування зв'язку між змінними та оцінити ступінь його прояву. У проблемі класифікації даних за допомогою кореляційного аналізу отримують вихідну інформацію у вигляді коваріаційних й кореляційних матриць та інших характеристик парних порівнянь. Це дає змогу визначити подібні один до одного або до певних еталонів об'єкти, сформувати класи подібних об'єктів й здійснити класифікацію. У проблемі зменшення розмірності досліджуваного простору ознак також за допомогою коваріаційних й кореляційних матриць визначають ознаки, що можуть бути без втрати істотної інформації подані через інші наявні дані.

Загальна методика перевірки гіпотези про існування зв'язку між ознаками передбачає три основних етапи: визначення типу даних; перевірку гіпотези про відсутність зв'язку і, при її відхиленні, оцінювання сили зв'язку. Тип вихідних даних істотно впливає на вибір методів і критеріїв, які можна застосовувати на наступних етапах аналізу.

7.1. Кореляційний аналіз кількісних ознак

Універсальною характеристикою ступеня тісноти зв'язку між кількісними ознаками є **коефіцієнт детермінації**. Вибірковий коефіцієнт детермінації певної ознаки y за вектором незалежних ознак

$\mathbf{X} = (x^{(1)}, x^{(2)}, \dots, x^{(p)})$ можна розрахувати як:

$$K_d(y; \mathbf{X}) = 1 - \frac{S_\varepsilon^2}{S_y^2}, \quad (7.3)$$

де

$$s_y^2 = \sum_{i=1}^n \frac{(y_i - \bar{y})^2}{n}, \quad (7.4)$$

n – кількість спостережень, а вибіркоче значення дисперсії нев'язок ε обчислюють за однією з таких формул:

$$s_\varepsilon^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{f}(\mathbf{X}_i))^2, \quad (7.5)$$

де $\hat{f}(\mathbf{X}_i)$ є статистичною оцінкою невідомого значення функції регресії $f(\mathbf{X})$ у точці \mathbf{X}_i , або:

$$s_\varepsilon^2 = \frac{1}{m} \sum_{j=1}^m \sum_{i=1}^{v_j} (y_{ij} - \bar{y}_{j*})^2, \quad (7.6)$$

де v_j – кількість даних, що потрапили до j -го інтервалу групування, y_{ji} – значення i -го спостереження досліджуваної ознаки, що потрапило до j -го

інтервалу, $\bar{y}_{j*} = \frac{\sum_{i=1}^{v_j} y_{ji}}{v_j}$ – її середнє значення за спостереженнями, які потрапили до j -го інтервалу, m – кількість інтервалів.

Формулу (7.5) застосовують у випадку, коли за результатами попереднього аналізу встановлено, що умовна дисперсія $D(\varepsilon | \mathbf{X}) = \sigma_\varepsilon^2 = \text{const}$, тобто не залежить від \mathbf{X} . Формулу (7.6) використовують, якщо ця умова не

виконується, а також у всіх випадках, коли обчислення здійснюють за згрупованими даними.

Величина коефіцієнта детермінації може змінюватися в межах від нуля до одиниці й відображає частку загальної дисперсії досліджуваної ознаки, яка зумовлена зміною функції регресії $f(X)$. При цьому нульове значення коефіцієнта детермінації відповідає відсутності будь-якого зв'язку, а його рівність одиниці – наявності строго функціонального зв'язку. Слід зазначити, що для обмеженого набору даних часто можна побудувати декілька різних адекватних регресійних моделей. Групування даних також можна здійснювати різними способами. Тому існує певна невизначеність коефіцієнтів детермінації: при застосуванні різних регресійних моделей або різних способів групування ми будемо отримувати дещо різні значення коефіцієнта детермінації.

Інші поширені характеристики ступеня тісноти зв'язку між ознаками можна розглядати як окремі випадки коефіцієнта детермінації, отримані для конкретних математичних моделей зв'язку.

Розрізняють парні та частинні кореляційні характеристики. Парні характеристики розраховують за результатами вимірювань тільки досліджуваної пари ознак. Тому вони не враховують опосередкованого або сумісного впливу інших ознак. Частинні характеристики є очищеними від впливу інших факторів, але для їх розрахунку необхідно мати вихідну інформацію не тільки про досліджувані ознаки, а й про всі інші, вплив яких необхідно усунути.

Для кількісних ознак найширше застосовуються коефіцієнти Пірсона і Фехнера. **Коефіцієнт кореляційного відношення Пірсона (парний коефіцієнт кореляції, вибіркового коефіцієнта кореляції, коефіцієнт Бравайса – Пірсона)** вимірює ступінь лінійного кореляційного зв'язку між кількісними скалярними ознаками. Його розраховують за формулою:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}. \quad (7.7)$$

Коефіцієнт Пірсона можна виразити також через дисперсії σ_y і $\sigma_{\Delta y}$, друга з яких характеризує розкид емпіричних точок стосовно рівняння лінійної регресії $y = ax + b$, де a та b – коефіцієнти, визначені методом найменших квадратів:

$$r = \frac{1}{\sqrt{1 + (\sigma_{\Delta y} / \sigma_y)^2}}. \quad (7.8)$$

Застосування коефіцієнта Пірсона як міри зв'язку є обґрунтованим лише за умови, що сумісний розподіл пари ознак є нормальним. Тому перед його розрахунком слід перевірити виконання цієї гіпотези. Якщо вона виконується, то квадрат коефіцієнта кореляції Пірсона дорівнює коефіцієнту детермінації.

Значення коефіцієнта кореляції може змінюватися від -1 до $+1$. Значення -1 та $+1$ відповідають строгій лінійній функціональній залежності, яка в першому випадку є згасаючою, а у другому – зростаючою. Для функціональної залежності $y = \text{const}$ коефіцієнт кореляції, як видно з наведеної формули, є невизначеним, оскільки в цьому випадку знаменник дорівнює нулю. Чим ближчим є значення коефіцієнта кореляції до -1 або $+1$, тим більш обґрунтованим є припущення про наявність лінійного зв'язку. Наближення його значення до нуля свідчить про відсутність лінійного зв'язку, але не є доказом відсутності статистичного зв'язку взагалі.

На рис. 7.1 показано дві кореляційні залежності. Для обох наборів тестових даних є очевидним існування статистичного зв'язку між параметрами x та y . Але коефіцієнти кореляції для них дорівнюють, відповідно, $r_1 = 0,995$ і $r_2 = 0,006$. Близькість коефіцієнта кореляції для другого набору до нуля пов'язана не з відсутністю зв'язку, а з його нелінійністю. Для порівняння, коефіцієнти детермінації для тих самих наборів дорівнюють приблизно 0,98 та 1,00.

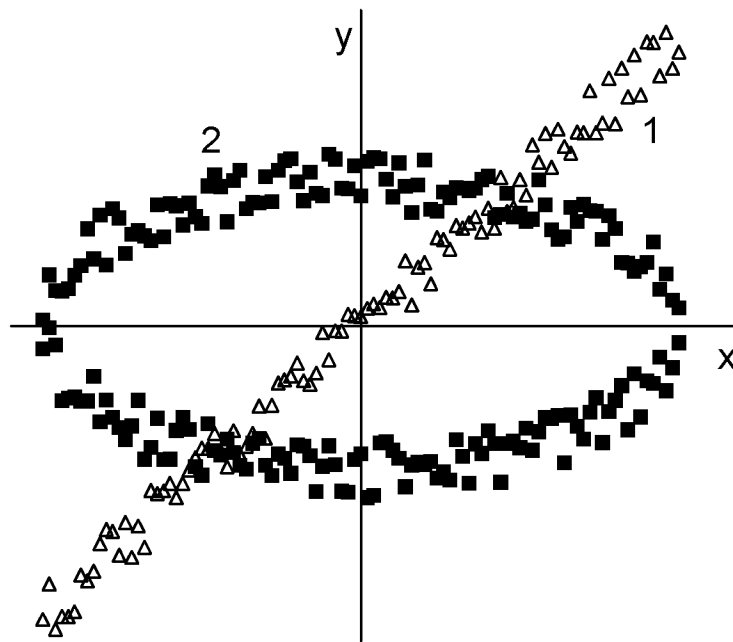


Рис. 7.1. Графічне зображення двох наборів тестових даних

Показаний приклад свідчить, що у багатьох випадках для попереднього аналізу припущення про наявність та тип зв'язку між певними ознаками доцільно нанести наявні дані на графік.

Як видно, близькість коефіцієнта кореляції до нуля в загальному випадку не є доказом незалежності ознак. Але можна довести, що у випадку, коли сумісний розподіл випадкових величин (x, y) є нормальним, рівність $r = 0$ свідчить про статистичну незалежність x і y .

Коефіцієнт кореляції Пірсона часто розглядають як універсальну міру кореляційного зв'язку. У багатьох пакетах загального призначення, зокрема в електронних таблицях MS Excel не передбачено інших засобів його вимірювання. Але як впливає з наведених вище даних насправді сфера його обгрунтованого застосування є доволі вузькою, оскільки лінійність залежності й нормальний розподіл даних навколо неї є скоріше виключенням, ніж правилом.

При дослідженні багатовимірних сукупностей випадкових величин із коефіцієнтів кореляції, обчислених попарно між ними, можна побудувати квадратну симетричну кореляційну матрицю з одиницями на головній діагоналі. Вона є основним елементом при побудові багатьох алгоритмів багатовимірної статистики, наприклад у факторному аналізі. Довірчий інтервал вибіркової оцінки коефіцієнта кореляції для двовимірної нормальної генеральної сукупності:

$$r \in \left[\tanh \left(z(r) - \frac{N_{\frac{1+p}{2}}}{\sqrt{n-3}} \right); \tanh \left(z(r) + \frac{N_{\frac{1+p}{2}}}{\sqrt{n-3}} \right) \right], \quad (7.9)$$

де n – обсяг вибірки, $N_{\frac{1+p}{2}}$ – квантіль нормального розподілу, p – стандартне значення довірчого рівня, $z(r)$ – z -перетворення (перетворення Фішера) вибіркового коефіцієнта кореляції r .

Коефіцієнт кореляції можна застосовувати для перевірки гіпотези про значущість зв'язку. Для нормально розподілених вихідних даних величину вибіркового коефіцієнта кореляції вважають значимо відмінною від нуля, якщо виконується нерівність:

$$r^2 > [1 + (n-2)/t_\alpha^2]^{-1}, \quad (7.10)$$

де t_α – критичне значення t-розподілу з $(n-2)$ ступенями вільності.

Статистика $\sqrt{n-1}r$ має r-розподіл зі щільністю:

$$\varphi_{r(n)}(r) = \frac{1}{\sqrt{\pi}} \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-2}{2}\right)} (1-r^2)^{\frac{n-4}{2}} \quad (-1 < r < 1). \quad (7.11)$$

Для великих за обсягом вибірок статистика $\sqrt{n-1}r$ наближається до стандартного нормального розподілу.

У випадку, коли між двома наборами ознак існує нелінійний зв'язок, для оцінювання ступеня його тісноти часто використовують **кореляційне відношення**. Це можливо, якщо щільність розміщення емпіричних точок на координатній площині дає можливість їх групування за однією із змінних і підрахунку групових середніх значень другої змінної для кожного інтервалу. Тоді кореляційне відношення залежної змінної y за незалежною змінною x можна розрахувати за формулою:

$$\rho_{yx}^2 = s_{y(x)}^2 / s_y^2, \quad (7.12)$$

де

$$s_{y(x)}^2 = \frac{1}{n} \sum_{j=1}^s v_j (\bar{y}_{j*} - \bar{y})^2;$$

$$s_y^2 = \frac{1}{n} \sum_{j=1}^s \sum_{i=1}^{v_j} (y_{ji} - \bar{y})^2;$$

$$\bar{y} = \frac{1}{n} \sum_{j=1}^s v_j \bar{y}_{j*};$$

$$\bar{y}_{j*} = \left(\sum_{i=1}^{v_j} y_{ij} \right) / v_j,$$

n – обсяг вибірки, s – кількість інтервалів групування за віссю абсцис, v_j – кількість точок, що потрапили до j -го інтервалу.

Кореляційне відношення може змінюватися в інтервалі від нуля до одиниці. Із $\rho_{yx} = 1$ випливає наявність строго функціонального зв'язку між досліджуваними ознаками, і навпаки, однозначний функціональний зв'язок між ними свідчить про те, що $\rho_{yx} = 1$. За відсутності зв'язку $\rho_{yx} = 0$, і навпаки, коли $\rho_{yx} = 0$, це означає, що для всіх інтервалів групування $\bar{y}_{j*} = \bar{y}$, тобто групові середні \bar{y}_{j*} не залежать від x .

На відміну від коефіцієнта кореляції, кореляційне відношення не є симетричним: у загальному випадку $\rho_{yx} \neq \rho_{xy}$. Більше того, можливі ситуації, коли один із цих коефіцієнтів дорівнює нулю, другий – одиниці. Зокрема, це може спостерігатися для парних функцій за умови, що функція розподілу значень незалежної змінної є симетричною стосовно нуля. Для даних, що наведені на рис. 7.1, кореляційне відношення першої серії дорівнює приблизно 0,98 і в межах похибки обчислень збігається з коефіцієнтами детермінації і кореляції. Для другої серії $\rho_{yx} \approx 0,63$ і $\rho_{xy} \approx 0,72$.

Можна довести, що кореляційне відношення збігається із модулем коефіцієнта кореляції між тими самими змінними за наявності лінійного зв'язку, а також за відсутності зв'язку. В інших випадках воно перевищує модуль коефіцієнта кореляції. Це дає можливість використовувати їх різ-

ницю як характеристику ступеня відхилення зв'язку від лінійності. Для цього розраховують величину:

$$v^2 = \frac{(n-k)(\rho_{yx}^2 - r^2)}{(k-2)(1-\rho_{yx}^2)}, \quad (7.13)$$

де n – кількість емпіричних точок, k – кількість невідомих параметрів моделі. Ця величина приблизно підпорядковується F -розподілу з параметрами $s - 2$ та $n - s$. Якщо розраховане за формулою (7.13) значення перевищує точку v_α^2 розподілу $F(s-2, n-s)$, то гіпотезу про лінійний зв'язок вважають необґрунтованою з імовірністю похибки α . Слід зазначити, що у зв'язку з можливістю різних способів групування даних значення кореляційного відношення, як і значення коефіцієнта детермінації, у загальному випадку є дещо невизначеним.

Коефіцієнт кореляції Фехнера розраховують за формулою:

$$r_F = \frac{C - H}{C + H} = \frac{2C - n}{n} = \frac{2C}{n} - 1, \quad (7.14)$$

де C – кількість збігів знаків відхилень варіант від відповідних середніх, H – кількість знаків, що не збігаються. Значення коефіцієнта Фехнера можуть змінюватися в межах від -1 до $+1$. Як і коефіцієнт Пірсона, він показує наявність лінійного зв'язку: чим ближчим до одиниці за модулем є значення коефіцієнта, тим сильнішим є зв'язок. Малі значення абсолютної величини коефіцієнта свідчать про відсутність лінійного зв'язку, але цього недостатньо для твердження про відсутність будь-якого зв'язку взагалі. Зокрема, для наведених на рис. 7.1 наборів даних значення коефіцієнта

Фехнера дорівнюють, відповідно, $r_{F1} = 0,941$ і $r_{F2} = -0,010$. Застосування для обчислення коефіцієнта лише кількості збігів або незбігів знаків відхилень від середніх значень можна розглядати як зведення первинної кількісної шкали до номінальної, що має призвести до втрати частини корисної інформації. Тому цей критерій застосовують досить рідко, але у певних випадках, коли інформація про збіги й незбіги знаків відхилень потрібна й для інших цілей, він може виявитися зручнішим за критерій Пірсона.

Коваріацією називають змішаний момент другого порядку. На відміну від інших показників, що характеризують наявність статистичного зв'язку, вона не є безрозмірною величиною. Її розраховують за формулою:

$$\text{Cov}(X, Y) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}). \quad (7.15)$$

Коваріація вибірки із самою собою є дисперсією. З наведеної формули можна отримати, що коефіцієнт кореляції Пірсона $r = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$, де

σ_X^2, σ_Y^2 – дисперсії вибірок.

При аналізі багатовимірних вибірок часто застосовують **коваріаційні матриці**:

$$C = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{pmatrix}, \quad (7.16)$$

де $c_{ij} = \text{cov}(x_i, x_j)$. Діагональні елементи матриці (7.16) є дисперсіями $c_{ii} = \sigma^2(x_i)$ відповідних рядів спостережень. Коваріаційна матриця є симетричною, тобто $c_{ij} = c_{ji}$.

7.2. Кореляційний аналіз порядкових ознак

Під **ранговою кореляцією** розуміють статистичний зв'язок між порядковими ознаками. Вихідні дані зазвичай подають у вигляді табл. 7.1.

Таблиця 7.1

Таблиця вихідних даних для рангового кореляційного аналізу

Порядковий номер об'єкта	Порядковий номер досліджуваної ознаки						
	0	1	2	...	k	...	p
1	x_{10}	x_{11}	x_{12}	...	x_{1k}	...	x_{1p}
2	x_{20}	x_{21}	x_{22}	...	x_{2k}	...	x_{2p}
...
i	x_{i0}	x_{i1}	x_{i2}	...	x_{ik}	...	x_{ip}
...
n	x_{n0}	x_{n1}	x_{n2}	...	x_{nk}	...	x_{np}

У цій таблиці елемент x_{ik} дає ранг i-го об'єкта за k-ю властивістю.

Завданнями аналізу в цьому випадку можуть бути: вивчення структури досліджуваних об'єктів; перевірка сукупної узгодженості ознак та умовне ранжирування об'єктів за ступенем тісноти зв'язку кожної з них з іншими ознаками; побудова єдиного групового впорядкування об'єктів (задача регресії на порядкових змінних).

У першому випадку кожену послідовність впорядкованих за k -ю ознакою n об'єктів подають як точку $\mathbf{X}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$, $k = 0, 1, \dots, p$ у n -вимірному просторі ознак. Найхарактернішими типами структури є такі.

1. Аналізовані точки рівномірно розкидані по всій області їх можливих значень. Це означає відсутність будь-якого зв'язку між досліджуваними ознаками.

2. Частина точок утворює ядро (кластер) із точок, що розташовані близько одна до одної, а інші випадково розкидані навколо цього ядра. Це відповідає існуванню підмножини узгоджених ознак.

3. Аналізовані точки утворюють декілька кластерів, розташованих відносно далеко один від одного. Це відповідає наявності декількох таких підмножин ознак, що існує істотний статистичний зв'язок між ознаками, які належать до однієї і тієї самої підмножини, і не існує значущого зв'язку між ознаками, які належать до різних підмножин.

Прикладом завдання другого типу є визначення узгодженості думок групи експертів з наступним впорядкуванням їх за рівнем компетентності. Для цього розраховують коефіцієнти конкордації для різних сукупностей досліджуваних змінних.

Вирішення завдань третього типу зводиться до побудови такого впорядкування, яке б у певному значенні було б найближчим до кожного з наданих впорядкувань досліджуваних ознак. Для цього часто застосовують середнє арифметичне або медіану наявних базових рангів. Це можна розглядати як задачу найкращого у певному розумінні відновлення невідомого ранжирування за наявними емпіричними даними, що зумовлює можливість її розгляду як задачі регресії.

Коефіцієнт рангової кореляції Спірмена (показник кореляції рангів Спірмена, коефіцієнт кореляції рангів) використовують, якщо досліджується зв'язок між рядами даних, вимірними за порядковою шкалою. Його можна застосовувати також і для кількісних даних, але, як правило, це буває недоцільним. У найпростішому випадку досліджувані об'єкти класифікують за двома ознаками. Наприклад, ми можемо спочатку впорядкувати групу учнів за їх здібностями до математики, а потім – до іноземних мов. Місця, які i -й учень займе в обох списках, будуть його рангами r_i та s_i . Якщо досліджувані ознаки взаємозв'язані, то послідовність рангів r_1, r_2, \dots, r_n певною мірою впливає на послідовність рангів s_1, s_2, \dots, s_n .

Ступінь близькості двох послідовностей відображує величина:

$$S_p = \sum_{i=1}^n (r_i - s_i)^2. \quad (7.17)$$

Якщо для нумерації об'єктів попередньо впорядкувати їх за однією із ознак, наприклад за зростанням рангів r_i , то формула (7.17) може бути записана так:

$$S_p = \sum_{k=1}^n (k - s_k)^2. \quad (7.18)$$

Величина S_p набуде найменшого можливого значення $S_p = 0$ тоді й тільки тоді, коли послідовності повністю збігатимуться. Найбільше можливе значення $S_p = \frac{1}{3}(n^3 - n)$ відповідає випадку, коли послідовності є повністю протилежними, тобто для будь-яких i, j з нерівності $r_i > r_j$ впливає $s_i < s_j$, і послідовності рангів першої ознаки $r_i = \{1, 2, \dots, n\}$ відповідає

послідовність рангів другої $s_i = \{n, n-1, \dots, 1\}$. Величину S_p незручно застосовувати як міру зв'язку, оскільки на її значення впливає кількість пар варіант досліджуваних рядів n .

Тому як міру зв'язку використовують **коефіцієнт рангової кореляції Спірмена**, значення якого розраховують за формулою:

$$\rho_s = 1 - \frac{6(S_p + B_x + B_y)}{n^3 - n}, \quad (7.19)$$

де B_x, B_y – поправки на об'єднання рангів у відповідних рядах, які обчислюють за формулою:

$$B_i = \frac{1}{12} \sum_{i=1}^m n_i (n_i^2 - 1), \quad (7.20)$$

де m – кількість груп об'єднаних рангів у вибірці, n_i – кількість рангів в i -й групі.

Значення коефіцієнта можуть змінюватися в межах від -1 до $+1$, при цьому -1 відповідає повній протилежності послідовностей рангів, а $+1$ – їх повному збігу.

Коефіцієнт рангової кореляції Спірмена можна застосовувати як показник некорельованості вибірок. У цьому випадку розраховують величину:

$$t_p = \sqrt{n-2} \frac{\rho_s}{\sqrt{1-\rho_s^2}}. \quad (7.21)$$

Для великих за обсягом вибірок ($n > 50$) критичні значення мають розподіл Стьюдента з $(n - 2)$ ступенями вільності. Статистика $\sqrt{n - 1}r_s$ для великих вибірок наближається до стандартного нормального розподілу.

Інший підхід використовує як міру подібності двох вибірок мінімальну кількість перестановок сусідніх об'єктів, потрібну для переведення однієї вибірки до іншої. Можна показати, що вона дорівнює кількості інверсій в однієї з рангових послідовностей у випадку, коли інша рангова послідовність впорядкована за зростанням. Нехай, наприклад, $n = 4$, послідовність r_i впорядкована за зростанням, а $s_i = \{4, 3, 1, 2\}$. Інверсіями є: $4 > 3$; $4 > 1$; $4 > 2$; $3 > 1$; $3 > 2$. Їх кількість $K = 5$. Найменше можливе значення кількості інверсій $K = 0$ відповідає повному збігу рангових послідовностей, а найбільше $K = \frac{n(n-1)}{2}$ – їх повній протилежності.

Як і в попередньому випадку кількість інверсій залежить від обсягу вибірки і є незручною для застосування як показника кореляції. Для цього використовують **коефіцієнт рангової кореляції Кендалла (коефіцієнт кореляції рангів, ранговий коефіцієнт кореляції)**, який розраховують за формулою:

$$\tau = 1 - \frac{2K}{\sqrt{\left(\frac{n(n-1)}{2} - B_x\right)\left(\frac{n(n-1)}{2} - B_y\right)}}, \quad (7.22)$$

де r_j, s_i – масиви рангів аналізованих рядів, n – кількість пар варіант у них. B_x, B_y – поправки на об'єднання рангів у відповідних рядах, які обчислюють за формулою:

$$B_i = \frac{1}{2} \sum_{i=1}^m n_i (n_i - 1), \quad (7.23)$$

де m – кількість груп об'єднаних рангів у вибірці, n_i – кількість рангів в i -й групі.

Для коефіцієнта рангової кореляції Кендалла у випадку великих вибірок статистика:

$$\tau \sqrt{\frac{9n(n+1)}{2(2n+5)}} \quad (7.24)$$

має розподіл, близький до стандартного нормального закону.

Коефіцієнт рангової кореляції Кендалла призначений для визначення сили кореляційного зв'язку між двома рядами даних за тих самих умов, що і коефіцієнт рангової кореляції Спірмена. Як і для коефіцієнта Спірмена його значення можуть змінюватися в межах від -1 до $+1$, при цьому -1 відповідає повній протилежності послідовностей рангів, а $+1$ – їх повному збігу. Слід зазначити, що обчислення коефіцієнта Кендалла є більш трудомістким, але з іншого боку, він має ряд переваг порівняно із коефіцієнтом Спірмена. Основними з них є:

- кращий рівень вивченості його статистичних властивостей, зокрема його вибіркового розподілу;
- можливість його застосування для визначення частинної кореляції;
- більша зручність перерахунку при додаванні нових даних.

7.3. Кореляційний аналіз номінальних ознак

Типовою ситуацією, де зустрічається необхідність перевірки зв'язку між номінальними ознаками, є обробка результатів соціологічних досліджень, що можуть містити такі комбінації ознак, як освіта, стать, професія, підтримка певної політичної партії, регіон мешкання тощо.

При дослідженні зв'язків між **категоризованими ознаками** вихідні дані подають у вигляді таблиці спряженості (табл. 7.2). До категоризованих зараховують номінальні ознаки, а також порядкові ознаки, для яких є відомим скінченний набір можливих градацій.

Таблиця 7.2

Таблиця спряженості категоризованих ознак

Рівні ознаки 1	Рівні ознаки 2				Разом
	1	2	...	r	
1	f_{11}	f_{12}	...	f_{1r}	n_1
2	f_{21}	f_{22}	...	f_{2r}	n_2
...
c	f_{c1}	f_{c2}	...	f_{cr}	n_c
Разом	m_1	m_2	...	m_r	S

Величини f_{ij} показують скільки разів зустрічалася комбінація ознак, за якої рівень першої має значення i , а рівень другої – значення j , m_j є сумами стовпців, а n_i – сумами рядків. За даними таблиці можна оцінити значення імовірностей, що входять до формули (7.1):

$$p_{ij} = P(A_i B_j) = \frac{f_{ij}}{S}; \quad p_i = P(A_i) = \sum_{j=1}^r p_{ij} = \frac{n_i}{S}; \quad p_j = P(B_j) = \sum_{i=1}^c p_{ij} = \frac{m_j}{S}. \quad (7.25)$$

Звідси для незалежних ознак маємо:

$$f_{ij} \approx n_i m_j / S. \quad (7.26)$$

Величини $\varphi_{ij} = n_i m_j / S$ є очікуваними частотами. Нульову гіпотезу про відсутність зв'язку відхиляють, якщо різницю між ними й частотами, що спостерігаються, не можна пояснити випадковими чинниками. Як критерій можна використовувати величину:

$$\chi^2 = \sum_{i=1}^c \sum_{j=1}^r \frac{(f_{ij} - \varphi_{ij})^2}{\varphi_{ij}} = n \left[\sum_{i=1}^c \sum_{j=1}^r \frac{f_{ij}^2}{n_i m_j} - 1 \right], \quad (7.27)$$

яка при достатньо великому обсязі вибірки наближається до розподілу χ^2 з кількістю ступенів вільності $(r-1)(c-1)$. На практиці для можливості застосування критерію часто вважають достатнім, щоб усі значення f_{ij} були не меншими ніж п'ять. При збільшенні кількості ступенів вільності мінімальні значення f_{ij} можуть бути дещо меншими.

Існує велика кількість показників ступеня тісноти статистичного зв'язку, призначених для категоризованих змінних, які не є універсальними, а відображають окремі властивості такого зв'язку.

Коефіцієнт Крамера розраховують за формулою:

$$C = \left[\frac{\sum_{i=1}^c \sum_{j=1}^r \frac{f_{ij}^2}{n_i m_j} - 1}{\min(c-1, r-1)} \right]^{1/2}. \quad (7.28)$$

Він змінюється в межах від нуля до одиниці. При цьому значення $C = 0$ свідчить про статистичну незалежність аналізованих ознак, а значення $C = 1$ – про можливість однозначного відтворення значень однієї ознаки за відомими значеннями другої. Дисперсію оцінки коефіцієнта Крамера можна отримати з виразу:

$$\sigma_C^2 \approx \frac{1}{n \min(c-1, r-1)}. \quad (7.29)$$

Її довірчий інтервал:

$$[C - u_{1-\alpha} \sigma_C; C + u_{1-\alpha} \sigma_C], \quad (7.30)$$

де u_q – q -квантиль стандартного нормального розподілу.

Поліхоричний коефіцієнт спряженості Чупрова призначений для дослідження кореляції номінальних ознак у таблиці спряженості $r \times c$. Його значення розраховують за формулою:

$$T = \frac{J-1}{\sqrt{(r-1)(c-1)}}; \quad J = \sum_{i=1}^c \sum_{j=1}^r \frac{f_{ij}^2}{n_i m_j}. \quad (7.31)$$

Існує велика кількість коефіцієнтів, що характеризують кореляцію між ознаками у випадку, коли кожна з двох ознак може мати лише два рівні, які найчастіше відповідають наявності та відсутності ознаки. У цьому випадку таблиця спряженості має розмір 2×2 і її елементи позначають так: $a = f_{11}$, $b = f_{12}$, $c = f_{21}$, $d = f_{22}$.

Коефіцієнт (показник подібності) Жаккара обчислюють за формулою:

$$J = \frac{a}{a + b + c}. \quad (7.32)$$

Значення цього коефіцієнта можуть змінюватися в межах від нуля до одиниці.

Простий коефіцієнт зустрічальності (показник подібності Сокала й Міченера) розраховують за формулою:

$$J = \frac{a + d}{n} = \frac{a + d}{a + b + c + d}. \quad (7.33)$$

Як і в попередньому випадку, значення коефіцієнта можуть змінюватися в межах від нуля до одиниці.

Показник подібності Рассела і Рао обчислюють як:

$$J = \frac{a}{n} = \frac{a}{a + b + c + d}. \quad (7.34)$$

Його значення також можуть змінюватися в межах від нуля до одиниці.

Коефіцієнт спряженості Бравайса (φ-коефіцієнт Пірсона, показник подібності Чупрова) розраховують за формулою:

$$C = \frac{ad - bc}{\sqrt{(a + b)(a + c)(d + b)(d + c)}}. \quad (7.35)$$

Значення цього коефіцієнта може змінюватися в межах від -1 до $+1$.

Коефіцієнт асоціації Юла визначають із співвідношення:

$$Q = \frac{ad - bc}{ad + bc}. \quad (7.36)$$

Коефіцієнт коліації Юла обчислюють як:

$$Y = \frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}}. \quad (7.37)$$

Він не має переваг порівняно з коефіцієнтом асоціації. Значення обох коефіцієнтів змінюються в межах від -1 до $+1$.

Хеммінгова відстань (метрика Хеммінга) $H = a + d$ також може застосовуватися для визначення кореляції, але, як і коваріація, вона не є безрозмірною величиною і може набувати будь-яких невід'ємних значень (верхньою межею є загальна кількість спостережень n).

7.4. Кореляційний аналіз змішаних ознак

Коефіцієнт Гауера застосовують у тому випадку, коли досліджувані ознаки виміряні в різних шкалах. Обчислення елементів матриці подібності здійснюють за формулою:

$$s_{ij} = \frac{\sum_{k=1}^p S_{ijk}}{\sum_{k=1}^p W_{ijk}}, \quad (i = 1, \dots, n; j = 1, \dots, n), \quad (7.38)$$

де S_{ijk} ($i, j = 1, \dots, n; k = 1, \dots, p$) – внесок ознаки у подібність об'єктів, W_{ijk} – вагова змінна ознаки, p – кількість ознак, що характеризують об'єкт, n – кількість об'єктів.

Для дихотомічних ознак алгоритм підрахунку внеску ознаки і визначення вагових коефіцієнтів збігається з коефіцієнтом Жаккара. Для порядкових ознак алгоритм підрахунку внеску ознаки збігається з Хеммінговою відстанню, узагальненою на порядкові змінні, а вагові коефіцієнти беруть рівними одиниці для всіх ознак. Для кількісних ознак:

$$S_{ijk} = 1 - \frac{|x_{ik} - x_{jk}|}{R_k}, \quad (7.39)$$

де x_{ik}, x_{jk} – значення k -ї змінної для об'єктів i та j , R_k – розкид k -ї ознаки, обчислений за всіма об'єктами.

Бісеріальний коефіцієнт кореляції призначений для дослідження кореляції в таблицях розміром $2 \times n$, які є дихотоміями за певною номінальною ознакою і класифікаціями за номінальною або порядковою ознакою, яка класифікується за q класами і може бути впорядкованою або не-впорядкованою. Вихідний розподіл має бути двовимірним нормальним.

У разі класифікації за порядковою ознакою бісеріальний коефіцієнт:

$$r_b = \frac{(\bar{x}_1 - \bar{x})n_1}{ns_x z_k}, \quad (7.40)$$

де \bar{x}_1 – середнє за першим рядком, \bar{x} – загальне середнє за всією таблицею, s_x – вибіркоче середнє квадратичне відхилення, n_1 – чисельність

першого рядка, n – загальна чисельність усіх вибірок, z_k – ордината щільності нормального розподілу в точці k , де k – розв'язок рівняння:

$$1 - F(k) = n_1 / n. \quad (7.41)$$

Похибку бісеріального коефіцієнта можна визначити за формулою:

$$m_{r_b} = \frac{1 - r_b}{\sqrt{n}}. \quad (7.42)$$

Вона має t -розподіл з кількістю ступенів вільності $(n - 2)$.

Бісеріальний коефіцієнт кореляції за таблицею Келлі – Вуда розраховують за формулою:

$$r_b = \frac{|\bar{x}_1 - \bar{x}_2| pq}{s_x \zeta}, \quad (7.43)$$

де $p = n_1 / n$ – частка частот у рядку, що визначається умовою $p > q$, q – частка частот в іншому рядку, ζ – ордината в точці межі класів частот першого та другого рядків, яка визначається за таблицею Келлі – Вуда. Похибку коефіцієнта визначають за формулою:

$$m_r = \frac{\frac{\sqrt{pq}}{\zeta} - r_b^2}{\sqrt{n}}. \quad (7.44)$$

У випадку класифікації за номінальною ознакою обчислення бісеріального коефіцієнта кореляції можна здійснити за формулою:

$$r_{\eta} = \sqrt{\frac{\frac{1}{n} \sum_{i=1}^q n_i \left(\frac{\mu_i}{\sigma_i} \right)^2 - \left(\frac{\mu_y}{\sigma_y} \right)^2}{1 + \frac{1}{n} \sum_{i=1}^q n_i \left(\frac{\mu_i}{\sigma_i} \right)^2}}, \quad (7.45)$$

де n загальна чисельність таблиці; n_i – чисельність таблиці в i -му перетині; μ_i/σ_i – оцінка в перетині i , одержувана за таблицею нормального інтеграла від відносної частоти першої з двох якісних ознак; μ_y/σ_y – оцінка, одержувана за таблицею нормального інтеграла від відносної частоти першої якісної ознаки за всією таблицею.

У випадку, коли одна із змінних дихотомізована, а інша – вимірjana в кількісній шкалі, обчислюють **точково-бісеріальний коефіцієнт кореляції**, який визначається за формулою:

$$r_{pb} = \frac{|\bar{x}_p - \bar{x}|}{s_x} \sqrt{\frac{n_p}{n_q}}, \quad (7.46)$$

де \bar{x}_p – середнє варіант кількісної вибірки, які відповідають подіям верхнього (першого) рівня дихотомічної вибірки, \bar{x} – середнє кількісної вибірки, s_x – середнє квадратичне кількісної вибірки, n_p – кількість подій у верхній (з рівнем 1) групі, n_q – кількість подій у нижній (з рівнем 2) групі. При цьому передбачається, що дихотомічна змінна може набувати лише два значення: 1 (верхній рівень) та 0 (нижній рівень). З погляду теорії точ-

ково-бісеріальну кореляцію можна розглядати як окремий випадок кореляційного відношення Пірсона.

Величину точково-бісеріального коефіцієнта кореляції вважають відмінною від нуля на рівні значущості α , якщо виконується нерівність:

$$r_{pb} \sqrt{\frac{n-2}{1-r_{pb}^2}} \geq t_{\alpha}, \quad (7.47)$$

де t_{α} – критичне значення t-розподілу з $(n-2)$ ступенями вільності.

7.5. Множинна кореляція

Про множинну кореляцію говорять, коли певна ознака пов'язана не з однією, а із сукупністю декількох інших ознак.

У випадку, коли досліджувані ознаки задовольняють багатовимірний нормальний розподіл, **частинний коефіцієнт кореляції** між двома ознаками при фіксованих значеннях інших ознак розраховують за формулою:

$$r_{ijX^{(i,j)}} = -\frac{R_{ij}}{\sqrt{R_{ii}R_{jj}}}, \quad (7.48)$$

де R_{kl} – алгебраїчне доповнення для елемента r_{kl} у кореляційній матриці.

Для тривимірної ознаки звідси можна отримати:

$$r_{01(2)} = \frac{r_{01} - r_{02}r_{12}}{\sqrt{(1-r_{02}^2)(1-r_{12}^2)}}. \quad (7.49)$$

Частинні коефіцієнти кореляції порядку k , тобто такі, що не враховують опосередкований вплив k інших змінних, можна розрахувати за коефіцієнтами порядку $k - 1$, використовуючи рекурентну формулу:

$$r_{01(2,3,\dots,k+1)} = \frac{r_{01(2,\dots,k)} - r_{0k+1(2,\dots,k)}r_{1k+1(2,\dots,k)}}{\sqrt{(1 - r_{0k+1(2,\dots,k)}^2)(1 - r_{1k+1(2,\dots,k)}^2)}}. \quad (7.50)$$

Частинні коефіцієнти кореляції мають всі властивості парних коефіцієнтів кореляції. Вони є показниками наявності лінійного зв'язку між двома незалежними ознаками, який не залежить від впливу інших ознак.

Тісноту зв'язку між декількома змінними у випадку множинної регресії можна оцінити за допомогою **коефіцієнта множинної кореляції**:

$$R = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (7.51)$$

де Y_i – значення змінної, взяті з кореляційної таблиці, y_i – відповідні значення, розраховані за рівнянням регресії.

Крім того застосовують множинний коефіцієнт кореляції, який є мірою лінійної кореляції між певною змінною y та сукупністю величин X_1, X_2, \dots, X_n і визначається як звичайний парний коефіцієнт кореляції між y та множинною лінійною регресією за X_1, \dots, X_n . При цьому припускають, що досліджувана сукупність підпорядковується багатовимірному нормальному закону.

Множинний коефіцієнт кореляції можна розрахувати за формулою:

$$R_{yX}^2 = 1 - \frac{|R|}{R_{00}}, \quad (7.52)$$

де $|R|$ – визначник кореляційної матриці.

Його також можна визначити за частинними коефіцієнтами кореляції:

$$R_{yX}^2 = 1 - (1 - r_{01}^2)(1 - r_{02(1)}^2)(1 - r_{03(12)}^2) \dots (1 - r_{0n(1,2,\dots,n-1)}^2). \quad (7.53)$$

Наприклад множинний коефіцієнт кореляції між певною ознакою z та двома іншими ознаками (x, y) дорівнює:

$$R_z = R_{z/xy} = \sqrt{\frac{r_{zx}^2 + r_{zy}^2 - 2r_{xy}r_{zx}r_{zy}}{1 - r_{xy}^2}}.$$

Множинний коефіцієнт кореляції мажорує будь-який парний або частинний коефіцієнт кореляції, що характеризує статистичні зв'язки досліджуваної ознаки. Як видно з формули (7.52), додавання нових ознак не може зменшувати коефіцієнт множинної кореляції.

Для багатовимірних нормальних сукупностей виконується рівність:

$$K_d(y, X) = R_{yX}^2. \quad (7.54)$$

Подання математичних об'єктів називають **канонічним**, якщо кожному об'єкту однієї множини відповідає один і тільки один об'єкт іншої множини й ця відповідність є взаємно однозначною. **Канонічний коре-**

ляційний аналіз здійснюють між двома сукупностями (групами) вибірок. Він призначений для визначення лінійної функції від перших p компонент і лінійної функції від q компонент, що залишилися, таких, щоб коефіцієнт кореляції між цими лінійними функціями набув найбільшого можливого значення. Чисельності груп (кількість вибірок у першій та другій групах, p та q) можуть розрізнятися, але необхідною умовою є рівна кількість варіант у всіх вибірках, що становлять обидві групи. Матриця взаємної кореляції двох груп вибірок має вигляд:

$$R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix}, \quad (7.55)$$

де R_{11} – матриця взаємної кореляції p змінних першої групи розміром $p \times p$, R_{22} – матриця взаємної кореляції q змінних другої групи розміром $q \times q$, R_{12} – матриця взаємної кореляції змінних першої та другої груп розміром $p \times q$.

Розв'язання зводиться до узагальненої проблеми власних значень:

$$R_{12}^T R_{11}^{-1} R_{12} v = \lambda R_{22} v, \quad (7.56)$$

де λ – вектор власних значень розміром q .

Квадратні корені з власних значень називають **канонічними кореляціями**.

Для випадкової вибірки обсягу n з $(p+1)$ -вимірного нормального розподілу коефіцієнт множинної кореляції вважають відмінним від нуля на рівні значущості α , якщо виконується нерівність:

$$R \geq \sqrt{\frac{pF}{v + pF}}, \quad (7.57)$$

де F – значення оберненої функції F -розподілу для довірчого рівня $(1 - \alpha)$ та кількості ступенів вільності p та $(n - p - 1)$.

Коефіцієнт конкордації призначений для дослідження зв'язків між порядковими ознаками, кількість яких є більшою, ніж два. Як міру узгодженості беруть суму квадратів відхилень сум рангів спостережень (об'єктів) від їх спільного середнього рангу:

$$S_w = \sum_{i=1}^n (S_i - \bar{S})^2 = \sum_{i=1}^n S_i^2 - \frac{\left(\sum_{i=1}^n S_i\right)^2}{n}; \quad (7.58)$$

$$\bar{S} = \frac{\sum_{i=1}^n S_i}{n}; \quad S_i = \sum_{j=1}^k R_{ij},$$

де R_{ij} – ранг i -го спостереження за j -ою ознакою.

Коефіцієнт конкордації Кендалла обчислюють за формулою:

$$W = \frac{12S_w}{k^2(n^3 - n)}. \quad (7.59)$$

Його значення може змінюватися в межах від нуля до одиниці, при цьому він дорівнює одиниці лише за умови, що всі досліджувані ранжирування збігаються. Коефіцієнт конкордації дорівнює нулю, якщо $k \geq 3$ і всі ранжирування є випадковими впорядкуваннями вихідної вибірки.

Середнє за всіма можливими парами ранжирувань значення коефіцієнта Спірмена за відсутності об'єднаних рангів пов'язано з коефіцієнтом конкордації співвідношенням:

$$\bar{\rho}_s = \frac{kW - 1}{k - 1}. \quad (7.60)$$

Величина $(k - 1) \frac{W}{1 - W}$ має F-розподіл з кількостями ступенів вільності $(n - 1)$ та $((n - 1)(k - 1) - 2)$. Високі значення функції F-розподілу свідчать про високий рівень узгодженості між ранжируваннями.

При $n > 7$ величина $k(n - 1)W$ за відсутності зв'язку між ознаками має розподіл, близький до χ^2 з $(n - 1)$ ступенем вільності. Якщо:

$$k(n - 1)W > \chi_{\alpha}^2(n - 1), \quad (7.61)$$

то гіпотезу про відсутність рангової кореляції можна відкинути при рівні значущості α .

Контрольні питання

1. Що називають кореляцією двох випадкових величин?
2. Які ознаки вважають статистично незалежними?
3. Які проблеми аналізу даних потребують попередньої перевірки наявності статистичного зв'язку між досліджуваними ознаками?
4. Якою є загальна методика перевірки гіпотези про наявність статистичного зв'язку?
5. З якою метою на початковому етапі кореляційного аналізу перевіряють тип даних?

6. Що є універсальною характеристикою статистичного зв'язку між кількісними ознаками?
7. Якими є переваги й недоліки застосування коефіцієнта детермінації?
8. Для заданого набору даних визначити коефіцієнт детермінації і зробити висновок про наявність кореляційного зв'язку.
9. У чому полягає різниця між парними та частинними кореляційними характеристиками?
10. Що вимірює парний коефіцієнт кореляції Пірсона?
11. Для заданого набору даних розрахувати значення парного коефіцієнта кореляції Пірсона і зробити висновок про наявність кореляційного зв'язку.
12. Що називають кореляційним відношенням двох випадкових величин? Які властивості зв'язку характеризує цей показник?
13. Що вимірює коефіцієнт кореляції Фехнера?
14. Для заданого набору даних розрахувати значення коефіцієнта кореляції Фехнера і зробити висновок про наявність кореляційного зв'язку.
15. Що називають коваріацією випадкових величин? Як цей показник пов'язаний з парним коефіцієнтом кореляції Пірсона?
16. Яких значень можуть набувати показники корельованості кількісних ознак? Які висновки можна зробити на основі значень цих показників?
17. Що називають ранговою кореляцією?
18. На якій властивості корельованих ознак ґрунтується коефіцієнт рангової кореляції Спірмена?
19. На якій властивості корельованих ознак ґрунтується коефіцієнт рангової кореляції Кендалла?
20. Яких значень можуть набувати показники рангової кореляції і про що свідчать їх значення?
21. Для заданого набору даних розрахувати значення коефіцієнтів рангової кореляції Спірмена й Кендалла і зробити висновок про наявність кореляційного зв'язку.

22. Які показники використовують для перевірки корельованості номінальних ознак? У чому полягають особливості застосування окремих показників?

23. Для заданого набору даних перевірити корельованість даних за допомогою критерію χ^2 , а також коефіцієнта Крамера й поліхоричного коефіцієнта спряженості Чупрова.

24. Які показники використовують для перевірки корельованості ознак, що виміряні у шкалах різного типу? Якими є особливості їх застосування?

25. Які показники використовують для дослідження корельованості декількох ознак? Якими є особливості їх застосування?

8. МЕТОДИ ПОБУДОВИ Й ДОСЛІДЖЕННЯ РЕГРЕСІЙНИХ МОДЕЛЕЙ

Завданням дослідження складних систем і процесів часто є перевірка наявності й встановлення типу зв'язку між незалежними змінними x_i (предикторами, факторами), значення яких можуть змінюватися дослідником і мають певну заздалегідь задану похибку, та залежною змінною (відгуком) z . Розв'язання таких завдань є предметом регресійного аналізу. На практиці завдання регресійного аналізу зазвичай ставиться так: необхідно підібрати достатньо просту функцію, що у певному розумінні найкращим чином описує наявну сукупність емпіричних даних.

8.1. Загальна характеристика методів та задач регресійного аналізу

Класичний регресійний аналіз включає методи побудови математичних моделей досліджуваних систем, методи визначення параметрів цих моделей і перевірки їх адекватності. Він припускає, що регресія є лінійною комбінацією лінійно незалежних базисних функцій від факторів з невідомими коефіцієнтами (параметрами). Фактори й параметри є детермінованими, а відгуки – рівноточними (тобто мають однакові дисперсії) некорельованими випадковими величинами. Припускають також, що всі змінні вимірюють у неперервних числових шкалах.

Звичайна процедура класичного регресійного аналізу є такою. Спочатку обирають гіпотетичну модель, тобто формулюють гіпотези про фактори, які суттєво впливають на досліджувану характеристику системи, і тип залежності відгуку від факторів. Потім за наявними емпіричними даними про залежність відгуку від факторів оцінюють параметри обраної моделі. Далі за статистичними критеріями перевіряють її адекватність.

При побудові регресійних моделей реальних систем і процесів вказані вище припущення виконуються не завжди. У більшості випадків їх невиконання веде до некоректності застосування процедур класичного регресійного аналізу і потребує застосування більш складних методів аналізу емпіричних даних.

Постулат про рівноточність і некорельованість відгуків не є обов'язковим. У випадку його невиконання процедура побудови регресійної моделі певною мірою змінюється, але суттєво не ускладнюється.

Більш складною проблемою є обрання моделі та її незалежних змінних. У класичному регресійному аналізі припускають, що набір факторів задається однозначно, всі суттєві змінні наявні в моделі й немає ніяких альтернативних способів обрання факторів. На практиці це припущення не виконується. Тому виникає необхідність розробки формальних та неформальних процедур перетворення й порівняння моделей. Для пошуку оптимальних формальних перетворень використовують методи факторного та дискримінантного аналізу. На сьогодні розроблено комп'ютеризовані технології послідовної побудови регресійних моделей.

Фактори в класичному регресійному аналізі вважають детермінованими, тобто припускають, що дослідник має про них всю необхідну інформацію з абсолютною точністю. На практиці це припущення часто не виконується. Відмова від детермінованості незалежних змінних приводить до необхідності застосування моделей кореляційного аналізу. В окремих випадках можна використовувати компромісні методи конфлюентного аналізу, які припускають можливість нормально розподіленого та усіченого розкиду значень факторів. Якщо ця умова виконується, побудову моделі можна звести до багаторазового розв'язування регресійної задачі.

Відмова від припущення про детермінованість параметрів моделей у регресійному аналізі призводить до суттєвих ускладнень, оскільки порушує його статистичні основи. Але на практиці це припущення виконується не завжди. У деяких випадках можна вважати параметри випадковими величинами із заданими законами розподілу. Тоді як оцінки параметрів можна брати їх умовні математичні сподівання для відгуків, що спостерігалися. Умовні розподіли та математичні сподівання розраховують за узагальненою формулою Байєса. Тому відповідні методи називають **байєсівським регресійним аналізом**.

Регресійні моделі часто використовують для опису процесів, що розвиваються у часі. У певних випадках це призводить до необхідності переходу від випадкових значень відгуків до випадкових послідовностей, випадкових процесів або випадкових полів. Однією з поширених і найпростіших моделей такого типу є **модель авторегресії**. Вона припускає, що відгук залежить не тільки від факторів, але також і від часу. Якщо останню залежність можна виявити, то проблема зводиться до стандартної задачі побудови регресії для модифікованого відгуку. В інших випадках необхідно використовувати більш складні прийоми.

Процедури класичного регресійного аналізу припускають, що закон розподілу відгуків є нормальним. Проте на практиці найчастішими є випадки, коли цей закон невідомий чи відомо, що він не є нормальним. Їх дослідження зумовило виникнення непараметричного регресійного аналізу, який не передбачає необхідності попереднього задання функції розподілу.

Важливою проблемою, яка виникає при оцінюванні параметрів регресійних моделей, є наявність грубих помилок серед набору аналізованих даних. Ці помилки можуть виникати внаслідок неправильних дій дослідника, збоїв у роботі апаратури, неконтрольованих короткотривалих сильних зовнішніх впливів на досліджувану систему тощо. У таких випадках використовують два підходи, що дають змогу зменшити вплив грубих помилок на результати аналізу. У першому з них розробляють критерії та алгоритми пошуку помилкових даних. Потім ці дані відкидають. У другому підході розробляють алгоритми аналізу, які є нечутливими до наявних помилкових даних (алгоритми робастного оцінювання параметрів).

Одним з основних постулатів класичного регресійного аналізу є припущення, що найкращі оцінки параметрів можна одержати, використовуючи метод найменших квадратів. На практиці оцінки, одержані за допомогою цього методу, часто бувають недостатньо точними і містять великі похибки. Причиною цього може бути структура регресійної моделі. Якщо вона є лінійною комбінацією експонент або поліномом високого степеня, то це призводить до пога-

ної зумовленості матриці системи нормальних рівнянь і нестійкості оцінок параметрів. Підвищення стійкості оцінок можна досягти шляхом відмови від вимоги їх незміщеності. Розвиток цього напрямку досліджень призвів до виникнення ребеневого, або рідж-регресійного аналізу.

Найчастіше **задачу побудови регресійної моделі** формують так. Необхідно знайти функцію заданого класу, для якої функціонал:

$$F(\alpha) = \sum_{i=1}^n (z_i(\boldsymbol{\beta}, \mathbf{X}) - y_i)^2 \rightarrow \min. \quad (8.1a)$$

У цьому виразі $z_i(\boldsymbol{\beta}, \mathbf{X})$ – значення функції, що апроксимує залежність, в i -ої точці, y_i – відповідне значення емпіричної залежності, $\boldsymbol{\alpha}$ – вектор параметрів, які треба знайти, \mathbf{X} – вектор незалежних змінних. Одержану функцію $z(\boldsymbol{\beta}, \mathbf{X})$ називають **(середньоквадратичною) регресійною моделлю**. Метод її пошуку, оснований на застосуванні критерію (8/1a), називають методом найменших квадратів.

Іноді замість функціонала (8.1a) для визначення параметрів регресійних моделей розв'язують задачі мінімізації інших функціоналів, зокрема:

$$F(\alpha) = \sum_{i=1}^n |z_i(\boldsymbol{\beta}, \mathbf{X}) - y_i| \rightarrow \min. \quad (8.1б)$$

$$F(\alpha) = \max |z_i(\boldsymbol{\beta}, \mathbf{X}) - y_i| \rightarrow \min. \quad (8.1в)$$

Одержувані при цьому регресійні моделі називають, відповідно, **середньоабсолютними (медіанними)** та **мінімаксними**. Ці моделі найчастіше використовують при побудові робастних алгоритмів регресійного аналізу, але їх практичне застосування обмежується поганою збіжністю таких алгоритмів.

Функцію, що апроксимує, у випадку однієї незалежної змінної (моделі простої регресії) часто шукають у вигляді полінома $z(x) = \sum_{j=0}^M \alpha_j x^j$, оберненого

полінома $z(x) = \frac{1}{\sum_{j=0}^M \alpha_j x^j}$, експоненціальних або показникових функцій $z = \alpha e^x$

чи $z = \alpha b^x$, степеневій функції $z = \alpha x^b$, лінійно-логіарифмічній функції $z = \alpha_1 + \alpha_2 x + \alpha_3 \ln x$, тригонометричного ряду Фур'є тощо. За наявності декількох незалежних змінних (моделі множинної регресії) найчастіше використовують функції, лінійні як за параметрами, так і за незалежними змінними

$z = \alpha_0 + \sum_{i=1}^p \alpha_i x_i$, а також поліноміальні моделі, що є лінійними за параметрами,

але нелінійними за незалежними змінними:

$$z = \alpha_0 + \sum_{i=1}^p \alpha_i x_i + \sum_{\substack{i,j=1 \\ i \geq j}}^p \alpha_{ij} x_i x_j + \sum_{\substack{i,j,k=1 \\ i \geq j \\ j \geq k}}^p \alpha_{ijk} x_i x_j x_k + \dots$$

Останні відповідають розкладу функції відгуку в ряд Тейлора. Проте можливе й використання для апроксимації інших видів залежностей.

Регресійні моделі називають **лінійними** або **нелінійними**, якщо вони є, відповідно, лінійними або нелінійними за параметрами. При цьому визначення "лінійна" часто опускають. Значення найвищого степеня предиктора в поліноміальних моделях називають **порядком моделі**. Наприклад,

$$z = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \varepsilon, \quad (8.2)$$

де ε – похибка моделі, є лінійною моделлю третього порядку.

Обрання типу регресійної моделі є нетривіальним завданням. Для моделей, що містять одну незалежну змінну, рекомендують спочатку нанести наявні

емпіричні дані на графік. Це дає можливість визначити наявність чи відсутність залежності між досліджуваними величинами, а також зробити певні припущення про тип залежності.

На рис. 8.1 як приклади подано певні набори емпіричних точок, для яких потрібно побудувати регресійні моделі. З наведених графіків видно, що ці моделі доцільно будувати у вигляді лінійної, квадратичної та експоненціальної функцій, відповідно. Але, як правило, визначення типу моделі за графіком емпіричних даних є не настільки очевидним. Тому зазвичай доводиться перевіряти декілька варіантів моделі і вибрати кращий з них за певними критеріями.

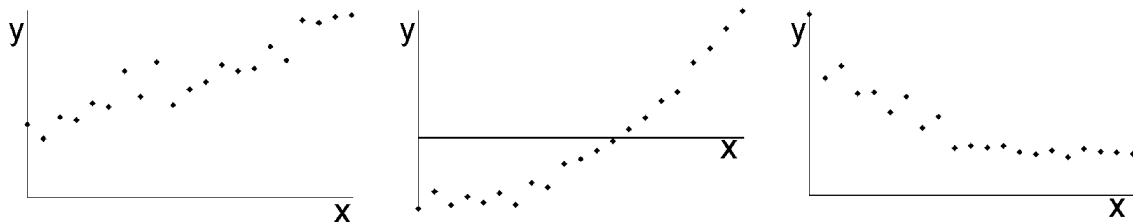


Рис. 8.1. Приклади наборів даних, для яких треба побудувати регресійні моделі

Часто як попередній етап регресійного аналізу рекомендують методами кореляційного аналізу перевіряти наявність значимого зв'язку між досліджуваними змінними. Але при цьому слід враховувати, що звичайні методи кореляційного аналізу дають змогу перевіряти лише гіпотезу про наявність лінійного зв'язку. Якщо зв'язок є, але він нелінійний, висновки, отримані за допомогою кореляційного аналізу, можуть бути помилковими.

Важливою особливістю регресійних моделей є те, що їх не можна застосовувати поза межами тієї області значень вихідних параметрів, для якої вони були побудовані. При використанні регресійних моделей типу полінома, оберненого полінома, тригонометричного ряду та деяких інших слід мати на увазі, що, збільшуючи кількість членів ряду, можна одержати скільки завгодно близькі до нуля значення функціоналів (8.1). Проте це не завжди свідчить про якість апроксимації, оскільки ці функціонали не дають інформації про ступінь

наближення моделі до емпіричної залежності у проміжках між наявними точками.

Іншою проблемою може бути наявність декількох екстремумів функціоналів (8.1). У цьому випадку необхідно мати на увазі, що більшість стандартних алгоритмів дає можливість знаходити локальні, а не глобальні екстремуми функціоналів, і результат мінімізації залежатиме від вибору початкових умов пошуку. Це часто призводить до необхідності встановлення додаткових критеріїв вибору моделі, серед яких можуть бути як формальні критерії їх адекватності, так і неформальні критерії, що ґрунтуються на сукупності відомих даних про об'єкт дослідження.

Поліноміальні регресійні моделі, як правило, є формальними. Їх використовують для опису систем і процесів, теорію яких розроблено недостатньо. При цьому спираються на відомі властивості ряду Тейлора для аналітичних функцій. Більш цікавими для дослідників зазвичай є математичні моделі, які відображають структуру та зв'язки у системах, сутність та механізми процесів, що відбуваються у них. Якщо теоретичні основи досліджуваних систем і процесів достатньо розроблені, часто постає проблема визначення окремих параметрів моделі за наявними емпіричними даними. Для її вирішення у багатьох випадках можна використовувати формальні процедури регресійного аналізу.

На практиці часто доводиться користуватися нелінійними за параметрами та багатовимірними моделями. Під багатовимірними тут розуміють моделі, що розглядають декілька відгуків. Задачам, що розв'язуються у межах відповідних напрямів регресійного аналізу, властиві й інші ускладнення. Зокрема у багатовимірних моделях окремі відгуки можуть бути пов'язані один з одним. Сама регресійна модель часто задається у неявному вигляді та є неаналітичним розв'язком певної системи алгебраїчних або диференціальних рівнянь. Нестійкість оцінок параметрів для нелінійних моделей різко зростає. Як правило, такі задачі мають декілька розв'язків або не мають розв'язків взагалі.

8.2. Лінійні однофакторні моделі

Найпростішим для аналізу і найбільш дослідженим є випадок лінійної кореляційної залежності між двома змінними X та Y . Наявність лінійного зв'язку можна перевірити, розрахувавши коефіцієнт парної кореляції Пірсона (7.7).

Розглянемо детальніше задачу підбору параметрів лінійної моделі:

$$z(x) = \alpha_0 + \alpha_1 x + \varepsilon, \quad (8.3)$$

за набором наявних емпіричних точок (x_i, y_i) .

У методі найменших квадратів виходять з припущення, що найкращими значеннями параметрів α_0 і α_1 будуть ті, для яких сума квадратів відхилень емпіричних значень y_i від розрахункових значень $z(x_i)$ набуває мінімального можливого значення. Для знаходження таких значень параметрів необхідно розв'язати систему:

$$\begin{cases} \frac{\partial}{\partial \alpha_0} \sum_{i=1}^n [z(x_i) - y_i]^2 = \frac{\partial}{\partial \alpha_0} \sum_{i=1}^n [\alpha_1 x_i + \alpha_0 - y_i]^2 = 0, \\ \frac{\partial}{\partial \alpha_1} \sum_{i=1}^n [z(x_i) - y_i]^2 = \frac{\partial}{\partial \alpha_1} \sum_{i=1}^n [\alpha_1 x_i + \alpha_0 - y_i]^2 = 0. \end{cases} \quad (8.4)$$

З (8.4) можна одержати такі вирази для оцінок α_0^* і α_1^* коефіцієнтів лінійної залежності:

$$\alpha_1^* = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2};$$

$$\alpha_0^* = \bar{Y} - \alpha_1^* \bar{X} = \frac{\sum_{i=1}^n y_i - \alpha_1^* \sum_{i=1}^n x_i}{n} = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}. \quad (8.5)$$

Перше рівняння в (8.5) є відношенням коваріації

$\text{cov}(x, y) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})$ до дисперсії $\sigma_x^2 = \frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{X})^2$, тобто:

$$\alpha_1^* = \frac{\text{cov}(x, y)}{\sigma_x^2} = \frac{\sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{X})^2}. \quad (8.6)$$

У багатьох випадках завдяки особливостям округлення чисел у ЕОМ останній вираз дає змогу отримати точніші оцінки параметрів, ніж (8.5).

Як приклад розглянемо таку задачу. У табл. 8.1 подано емпіричні дані, для яких треба побудувати регресійну модель, а також дані, необхідні для розрахунку її параметрів. На рис. 8.2 наведено емпіричні точки та графік залежності, побудований за одержаною методом найменших квадратів моделлю. Як видно з рисунка, одержана модель задовільно описує наявні емпіричні дані.

Таблиця 8.1

Приклад розрахунку регресійної моделі

№ випробування	x_i	y_i	x_i^2	$x_i \cdot y_i$	z	Залишок
1	0	0,310	0	0	0,516	0,206
2	0,3	1,037	0,09	0,311	1,506	0,470
3	0,6	2,513	0,36	1,508	2,497	-0,017
4	0,9	3,843	0,81	3,459	3,487	-0,356
5	1,2	4,840	1,44	5,807	4,477	-0,363
6	1,5	6,020	2,25	9,030	5,467	-0,553
7	1,8	5,865	3,24	10,557	6,457	0,592
8	2,1	7,470	4,41	15,686	7,447	-0,022
9	2,4	8,889	5,76	21,332	8,438	-0,451

10	2,7	9,25399	7,29	24,98577	9,427681	0,174
11	3	10,39294	9	31,17882	10,41785	0,025
12	3,3	11,11287	10,89	36,67247	11,40801	0,295
Σ	19,8	71,54526	45,54	160,5277		0
α_1^*	3,300548	α_0^*	0,516201			

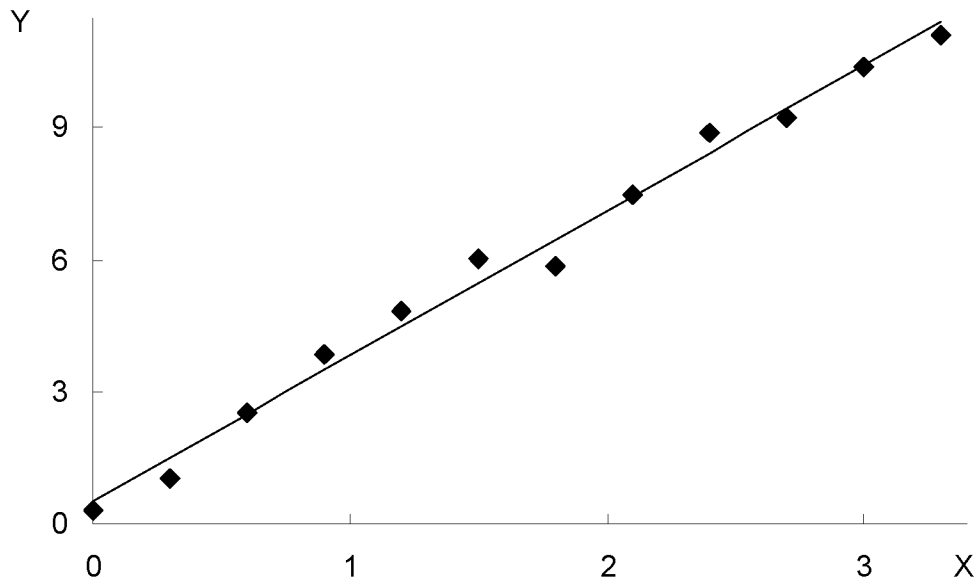


Рис. 8.2. Графіки досліджуваної залежності й лінійної моделі

Для розглянутої регресійної моделі сума залишків $\sum_{i=1}^n (y_i - z(x_i))$ дорівнює нулю, якщо модель містить вільний член α_0 . Виключення вільного члена з моделі зазвичай є невиправданим. Задання моделі з $\alpha_0 = 0$ доцільно лише у випадках, коли з теорії відомо, що для нульових значень предикторів відгук має бути рівним нулю. Якщо це невідомо, але бажано одержати модель, що не містить вільного члена, більш доцільним є застосування центрування даних.

Підставляючи до моделі $z(x) = \alpha_0 + \alpha_1 x + \varepsilon$ оцінку коефіцієнта α_0^* з (8.5) можна одержати:

$$Y^* = \alpha_0^* + \alpha_1^* x + \varepsilon = \bar{Y} + \alpha_1^* (x - \bar{X}) + \varepsilon. \quad (8.7)$$

Звідси отримуємо центровану модель:

$$Y - \bar{Y} = \alpha_1^* (x - \bar{X}) + \varepsilon, \quad (8.8)$$

яка не містить вільного члена.

Незважаючи на те, що, як правило, реальні залежності відгуків від факторів є нелінійними, розглянутий випадок широко використовують у практиці побудови регресійних моделей. Це пов'язано з трьома основними причинами. По-перше, він є найбільш простим і дослідженим. Зокрема, для нього достатньо повно розроблені процедури визначення статистичних характеристик одержуваних оцінок параметрів (дисперсії, довірчих інтервалів тощо) та перевірки адекватності моделей. По-друге, у багатьох випадках складні залежності можна подати як набір лінійних (на малих відрізках змінювання факторів) залежностей. По-третє, нелінійні залежності у деяких випадках можна перетворити до лінійного вигляду шляхом заміни змінних. Деякі приклади такого перетворення наведено у табл. 8.2.

Таблиця 8.2

Приклади лінеаризації нелінійних залежностей

Вихідна залежність	Лінеаризована залежність	Нові змінні
$z = \alpha_0 \exp(-\alpha_1 x)$	$\ln z = \ln \alpha_0 - \alpha_1 x$	$x, \ln z$
$z = \alpha_0 [1 - \exp(-\alpha_1 x)]$	$\ln \frac{\alpha_0}{\alpha_0 - z} = \alpha_1 x$	$x, \ln \frac{\alpha_0}{\alpha_0 - z}$
$z = \alpha_0 \exp(-\alpha_1/x)$	$\ln z = \ln \alpha_0 - \alpha_1/x$	$1/x, \ln z$
$z = \alpha_0 x^{\alpha_1}$	$\ln z = \ln \alpha_0 + \alpha_1 \ln x$	$\ln x, \ln z$
$z = \alpha_0 x + \alpha_1 x^2$	$z/x = \alpha_0 + \alpha_1 x$	$x, z/x$

$z = \alpha_0 \sin(\alpha_1 x)$	$\arcsin (z/\alpha_0) = \alpha_1 x$	$x, \arcsin (z/\alpha_0)$
---------------------------------	-------------------------------------	---------------------------

Перетворення нелінійних залежностей до лінійних є строго обґрунтованим, якщо вихідні дані є точними. На практиці вони завжди вимірюються з деякою похибкою. Розглянемо модель:

$$z = \alpha_0 x^{\alpha_1} + \varepsilon, \quad (8.9)$$

де ε – похибка вимірювань.

Її лінеаризована форма матиме вигляд:

$$\ln z = \ln \alpha_0 + \alpha_1 \ln x + \varepsilon', \quad (8.10)$$

де ε' є невідомою випадковою величиною. Використання як лінеаризованої форми виразу:

$$\ln z = \ln \alpha_0 + \alpha_1 \ln x \quad (8.11)$$

є коректним лише у тому випадку, коли величина ε' є малою порівняно з іншими доданками правої частини (8.11).

Розглянемо питання про точність оцінок. Для цього запишемо таку тотожність:

$$(y_i - \bar{Y}) = (y_i^* - \bar{Y}) + (y_i - y_i^*). \quad (8.12)$$

Тут y_i^* є оцінкою значення величини y при $x = x_i$. Якщо піднести обидві частини цієї тотожності до квадрата та взяти суму від $i = 1$ до n , то одержимо:

$$\sum_{i=1}^n (y_i - \bar{Y})^2 = \sum_{i=1}^n (y_i^* - \bar{Y})^2 + \sum_{i=1}^n (y_i - y_i^*)^2. \quad (8.13)$$

У цієї рівності немає члена $2 \sum_{i=1}^n (y_i^* - \bar{Y})(y_i - y_i^*)$, оскільки:

$$y_i - \bar{Y} = \alpha_1 (x_i - \bar{X});$$

$$y_i - y_i^* = y_i - \bar{Y} - \alpha_1 (x_i - \bar{X});$$

$$\sum_{i=1}^n (y_i - \bar{Y})(y_i - y_i^*) = \alpha_1 \sum_{i=1}^n (x_i - \bar{X})(y_i - \bar{Y}) - \alpha_1^2 \sum_{i=1}^n (x_i - \bar{X})^2 = 0,$$

враховуючи (8.6).

Розглянемо склад виразу (8.13). Його ліва частина є сумою квадратів відхилень значень, що спостерігалися, стосовно загального середнього. Перший доданок правої частини є сумою квадратів відхилень оцінок цих значень, зроблених на основі обраної моделі регресії, від загального середнього. Її часто називають сумою квадратів стосовно регресії. Другий доданок правої частини є сумою квадратів відхилень значень, що спостерігалися, від їх оцінок, одержаних з використанням обраної моделі. Цей доданок називають сумою квадратів, що зумовлена регресією. Для того, щоб модель була придатною для завбачення значень досліджуваної величини, необхідно, щоб він був малим порівняно із сумою квадратів стосовно регресії. У граничному випадку він має дорівнювати нулю.

Будемо вважати дисперсію залишків $\sigma_{\varepsilon_i}^2$ і, відповідно, дисперсію відгуків $\sigma_{Y_i}^2$ сталими. Дисперсія емпіричних точок стосовно середнього $\sigma_{Y_i}^2$ буде дорівнювати їх дисперсії σ_{YX}^2 стосовно лінії регресії у випадку, коли постульована модель є істинною. У протилежному випадку $\sigma_{Y_i}^2 > \sigma_{YX}^2$. Оцінкою величини σ_{YX}^2 є відношення суми квадратів відхилень спостережень від середнього до кілько-

сті степенів вільності. Останнє беруть рівним різниці між кількістю випробувань і кількістю констант, які визначаються незалежно одна від одної за їх результатами. У випадку, що розглядається, воно дорівнює $n - 2$, оскільки з емпіричних даних визначають два параметри прямої регресії. Тобто:

$$\sigma_{XY}^2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n - 2}. \quad (8.14)$$

Для прикладу, що наведений у таблиці 8.1, можна одержати $\sigma_{YX}^2 = 14,02$, $\sigma_{Y_i}^2 = 14,17$, $\sigma_{\varepsilon_i}^2 = 0,15$. Таким чином рівність $\sigma_{Y_i}^2 = \sigma_{YX}^2 + \sigma_{\varepsilon_i}^2$, яка впливає з (8.13), є виконаною.

Нехай p_i є кількістю повторних вимірювань величини \bar{Y}_i при заданому значенні x_i . Тоді квадратична форма, що мінімізується в методі найменших квадратів:

$$Q = \sum_{i=1}^n (\bar{Y}_i - z(x_i))^2 p_i = \sum_{i=1}^n p_i \left[\bar{Y}_i - \alpha_0 - \alpha_1 (x_i - \bar{x}) \right]^2. \quad (8.15)$$

Розглядаючи її як функцію параметрів α_0 , α_1 , як і у попередньому випадку, одержуємо оцінки параметрів:

$$\begin{cases} \alpha_0^* = \sum_{i=1}^n p_i \bar{Y}_i / \sum_{i=1}^n p_i = \bar{Y}; \\ \alpha_1^* = \sum_{i=1}^n p_i \bar{Y}_i (x_i - \bar{x}) / \sum_{i=1}^n p_i (x_i - \bar{x})^2. \end{cases} \quad (8.16)$$

У припущенні, що умовний розподіл величини \bar{Y}_i при заданому x_i є нормальним, оцінкою дисперсії буде величина:

$$\sigma_{Y_i}^{2*} = \frac{1}{n} \sum_{i=1}^n p_i (\bar{Y}_i - Y_i^*)^2, \quad (8.17)$$

де $Y_i^* = \alpha_0^* + \alpha_1^*(x_i - \bar{x}_i)$.

Слід зазначити, що висновки, одержувані на основі мінімізації дисперсії похибки, є правильними тільки тоді, коли постульована модель коректна. В інших випадках вони можуть виявитися помилковими. Перед прийняттям рішення стосовно моделі треба перевірити гіпотезу, що лінійна модель $z = \alpha_0 + \alpha_1(x - \bar{x})$ задовільно описує емпіричні дані із заданою точністю при заданому рівні значущості η . Для цього визначають міру похибки емпіричних даних:

$$S_a^2 = \frac{1}{n-2} \sum_{i=1}^n p_i (\bar{Y}_i - Y_i^*)^2. \quad (8.18)$$

Ця величина є зміщеною оцінкою дисперсії $\sigma_{Y_i}^2$. Незміщеною оцінкою цієї дисперсії є величина:

$$S_e^2 = \frac{\sum_{i=1}^n \sum_{j=1}^{p_i} (Y_{ij} - \bar{Y}_i)^2}{\sum_{i=1}^n p_i - n}, \quad (8.19)$$

де Y_{ij} – j -те одиничне вимірювання при $x = x_i$. Критерієм адекватності моделі при заданій надійності $1 - \eta$ є виконання нерівності:

$$S_a^2 / S_e^2 \leq F_{1-\eta}, \quad (8.20)$$

де $F_{1-\eta}$ – відповідне значення функції розподілу Фішера, для кількостей степенів вільності $n_1 = n_2 = n - 1$.

Довірчі інтервали для параметрів α_0, α_1 можна знайти за допомогою коефіцієнтів t-розподілу Стюдента з кількістю степенів вільності $\sum_{i=1}^n p_i - 2$:

$$\begin{cases} \alpha^* - t_{1-\eta/2} S_{\alpha^*} \leq \alpha \leq \alpha^* + t_{1-\eta/2} S_{\alpha^*} ; \\ S_{\alpha^*} = \sqrt{\frac{1}{n-2} \sum_{i=1}^n p_i (\bar{\alpha} - \alpha^*)^2} . \end{cases} \quad (8.21)$$

8.3. Поліноміальні моделі

У багатьох випадках емпіричні залежності можна описати поліноміальними моделями вигляду:

$$z = \sum_{i=1}^q \alpha_i x_i^i . \quad (8.22)$$

Оцінки параметрів таких моделей отримують шляхом розв'язування нормальних рівнянь виду:

$$\begin{pmatrix} n & \sum x_i & \sum x_i^2 & \dots & \sum x_i^q \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \dots & \sum x_i^{q+1} \\ \dots & \dots & \dots & \dots & \dots \\ \sum x_i^q & \sum x_i^{q+1} & \sum x_i^{q+2} & \dots & \sum x_i^{2q} \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \dots \\ \alpha_q \end{pmatrix} = \begin{pmatrix} \sum Y_i \\ \sum Y_i x_i \\ \dots \\ \sum Y_i x_i^q \end{pmatrix} . \quad (8.23)$$

Зазвичай стовпці, що утворюють матрицю X , не є ортогональними. Це призводить до того, що за необхідності збільшення степеня полінома буває необхідним перераховувати оцінки всіх його коефіцієнтів. Тому для поліномів

високих степенів більш раціональним методом побудови регресійної моделі є заміна вихідного рівняння (8.22) іншим:

$$z = \sum_{i=1}^q \alpha'_i \zeta_i, \quad (8.24)$$

де $\zeta_i = \zeta_i(x)$ є поліномами i -го степеня за x , які задовольняють умови ортогональності:

$$\begin{cases} \sum_{j=1}^n \zeta_{ij} = 0, \quad i = 1, 2, \dots, q; \\ \sum_{j=1}^n \zeta_{ij} \zeta_{i'j} = 0, \quad i \neq i', \end{cases} \quad (8.25)$$

ζ_{ij} є i -м поліномом для точки x_j .

Квадратична форма, що мінімізується у методі найменших квадратів має вигляд:

$$Q = \sum_{j=1}^n \left(Y_j - \alpha'_0 - \alpha'_1 \zeta_{1j} - \dots - \alpha'_q \zeta_{qj} \right)^2. \quad (8.26)$$

Значення, що відповідають мінімуму (8.26), можна знайти, розв'язавши систему:

$$\begin{cases} \frac{\partial Q}{\partial \alpha'_0} = -2 \sum_{j=1}^n \left(Y_j - \alpha_0^* - \alpha_1^* \zeta_{1j} - \dots - \alpha_q^* \zeta_{qj} \right) = 0; \\ \frac{\partial Q}{\partial \alpha'_i} = -2 \left(\sum_{j=1}^n Y_j \zeta_{ij} - \alpha_0^* \sum_{j=1}^n \zeta_{ij} - \alpha_1^* \sum_{j=1}^n \zeta_{1j} \zeta_{ij} - \dots - \alpha_q^* \sum_{j=1}^n \zeta_{qj} \zeta_{ij} - \dots - \alpha_q^* \sum_{j=1}^n \zeta_{qj} \zeta_{ij} \right) = 0, \quad i = 1, 2, \dots, q. \end{cases} \quad (8.27)$$

Звідси, використовуючи умови ортогональності (8.25), одержуємо:

$$\alpha'_{0*} = \bar{Y}; \quad \alpha'_{i*} = \frac{\sum_{j=1}^n Y_j \zeta_{ij}}{\sum_{j=1}^n \zeta_{ij}^2} . \quad (8.28)$$

Використовуючи умови ортогональності, можна одержати явний вигляд поліномів для випадку, коли значення x змінюються з рівним кроком ω :

$$\begin{cases} \zeta_{0j} = 1; \\ \zeta_{1j} = v_j - \bar{v}_j; \\ \zeta_{2j} = \zeta_{1j}^2 - (n^2 - 1)/12, \end{cases} \quad (8.29)$$

де $v_j = (x_{j+1} - x_1) / \omega$.

Поліноми вищих степенів одержують за рекурентною формулою:

$$\zeta_{r+1,j} = \zeta_{1j} \zeta_{rj} - \frac{r^2 (n^2 - r^2)}{4(4r^2 - 1)} \zeta_{r-1,j} . \quad (8.30)$$

Розглянемо такий приклад. У табл. 8.3 наведено результати вимірювання деякої величини.

Таблиця 8.3

Емпіричні дані для побудови поліноміальної регресійної моделі

j	1	2	3	4	5	6	7	8	9	10
x_j	0	10	20	30	40	50	60	70	80	90
Y_j	23	29	41	60	79	88	83	61	33	27

Побудуємо модель досліджуваної залежності у вигляді полінома 5-го степеня:

$$z = \alpha'_0 + \alpha'_1 \zeta_1 + \alpha'_2 \zeta_2 + \alpha'_3 \zeta_3 + \alpha'_4 \zeta_4 + \alpha'_5 \zeta_5,$$

де $\zeta_i = \sum_{t=1}^i \beta_{it} x^t$ – поліноми і-го степеня, які задовольняють умови ортогональності. Дані, необхідні для розрахунку значень ζ_{ij} і коефіцієнтів α_i , наведено в табл. 8.4.

Таблиця 8.4

Результати розрахунку допоміжних параметрів для поліноміальної регресійної моделі та оцінок значень досліджуваної величини

v_j	1	2	3	4	5	6	7	8	9	10
ζ_{0j}	1	1	1	1	1	1	1	1	1	1
ζ_{1j}	-4,5	-3,5	-2,5	-1,5	-0,5	0,5	1,5	2,5	3,5	4,5
ζ_{2j}	12	4	-2	-6	-8	-8	-6	-2	4	12
ζ_{3j}	-25,2	8,4	21	18,6	7,2	-7,2	-18,6	-21	-8,4	25,2
ζ_{4j}	43,2	-52,8	-40,8	7,2	43,2	43,2	7,2	-40,8	-52,8	43,2
ζ_{5j}	-60	140	-10	-110	-60	60	110	10	-140	60
$y_j \zeta_{0j}$	23	29	41	60	79	88	83	61	33	27
$y_j \zeta_{1j}$	-103,5	-101,5	-102,5	-90	-39,5	44	124,5	152,5	115,5	121,5
$y_j \zeta_{2j}$	276	116	-82	-360	-632	-704	-498	-122	132	324
$y_j \zeta_{3j}$	-579,6	243,6	861	1116	568,8	-633,6	-1544	-1281	-277,2	680,4
$y_j \zeta_{4j}$	993,6	-1531	-1673	432	3413	3802	597,6	-2489	-1742	1166
$y_j \zeta_{5j}$	-1380	4060	-410	-6600	-4740	5280	9130	610	-4620	1620
Y_i^*	24,09	30,1	42,23	60,96	79,8	89,74	83,86	61,81	34,38	28,03

Легко перевірити, що для одержаних даних виконуються умови ортогональності, тобто суми значень ζ_{ij} ($i \neq 0$) у кожному рядку й суми за i добутоків вигляду $\zeta_{ij}\zeta_{kj}$ ($i \neq k$) дорівнюють нулю. За даними таблиці розраховуємо коефіцієнти α'_i (табл. 8.5).

Таблиця 8.5

Результати розрахунку параметрів побудованої регресійної моделі

i	0	1	2	3	4	5
$\sum_{j=1}^n \zeta_{ij}^2$	10	82,5	528	3089	16474	78000
$\sum_{j=1}^n y_j \zeta_{ij}$	524	121	-1550	-845,4	2969	2950
α'_i	52,4	1,467	-2,936	-0,2737	0,1802	0,03782
$\sigma_{\alpha'_j}^2$	0,038	0,0036	0,00072	0,00012	0,000023	0,000005
$\sigma_{\alpha'_j}$	0,20	0,068	0,027	0,011	0,0048	0,0022

Оцінками дисперсії цих коефіцієнтів є величини $\sigma_{\alpha'_i}^2 = \frac{\sum_{j=1}^n (y_j - y_j^*)^2}{(n - q - 1) \sum_{j=1}^n \zeta_{ij}^2}$.

Звідси за формулами (8.24, 8.28) одержуємо оцінки Y_i^* значень досліджуваної величини Y у точках $x = x_j$, які наведені у табл. 8.3 і є досить близькими до її емпіричних значень. На рис. 8.3 наведено графіки вихідних даних та побудованої регресійної моделі.

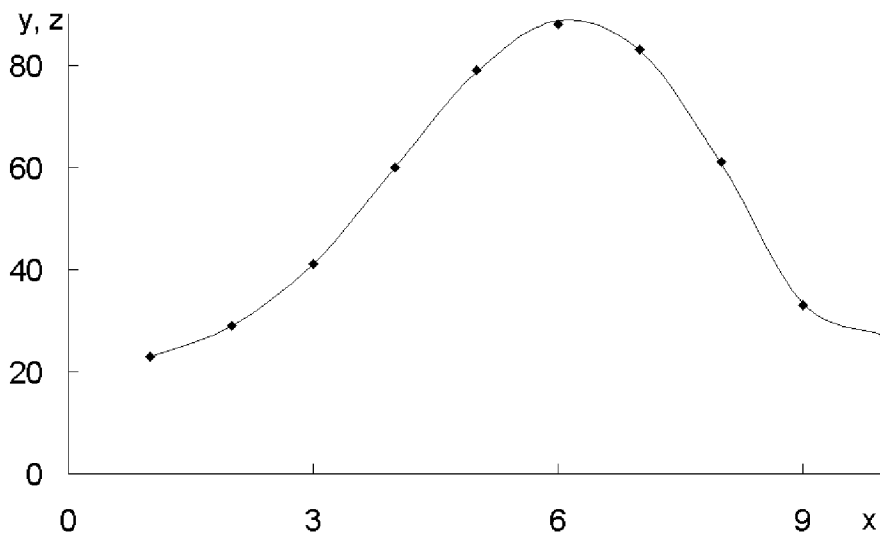


Рис. 8.3. Графіки вихідних даних (крапки) й побудованої поліноміальної регресійної моделі (лінія)

8.4. Нелінійні однофакторні моделі інших типів

Апроксимацію емпіричних залежностей тригонометричними багаточленами називають **гармонічним аналізом**. У цьому випадку модель має вигляд:

$$z(x) = \alpha_0 + \sum_{k=1}^r \alpha_k \cos \frac{2\pi kx}{T} + \sum_{k=1}^r \beta_k \sin \frac{2\pi kx}{T}, \quad (8.32)$$

де T – період спостереження апроксимованої залежності, r – кількість гармонік ($r < n/2$), n – кількість частин, на які розділений період T .

Її параметри визначають за формулами:

$$\begin{aligned} \alpha_0 &= \frac{1}{n} \sum_{k=0}^n y_k; \\ \alpha_m &= \frac{2}{n} \sum_{k=0}^n y_k \cos \frac{2\pi km}{n}, \quad m = 1, 2, \dots, r; \\ \beta_m &= \frac{2}{n} \sum_{k=0}^n y_k \sin \frac{2\pi km}{n}, \quad m = 1, 2, \dots, r; \end{aligned} \quad (8.33)$$

де y_k – значення апроксимованої функції у точках $x_k = \frac{kT}{n}$.

На рис. 8.4 показано графіки емпіричних даних і відповідних регресійних моделей, побудованих у вигляді тригонометричних рядів для m рівних 2, 3, 4 й 5. Видно, що зі збільшенням кількості членів тригонометричного ряду різниця між моделлю та емпіричними точками зменшується. Добре видно також, що найбільшу похибку модель дає поблизу меж відрізка, на якому визначені емпіричні дані.

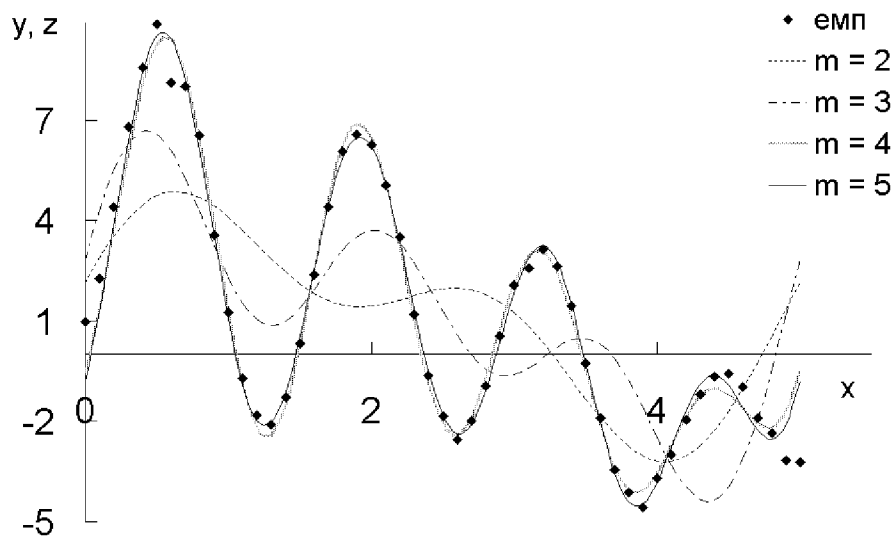


Рис. 8.4. Графіки вихідних даних і побудованої тригонометричної регресійної моделі

Моделі, що мають вигляд **модифікованої показникової функції**, записують як $z(x) = a + bc^x$ або $z(x) = a + bc^{-x}$. Якщо емпіричні точки є рівновіддаленими одна від одної, тобто крок емпіричної залежності за x є сталим, то можна перейти від незалежної змінної x до порядкового номера відповідної точки i й розрахувати параметри моделі першого типу за формулами:

$$c = \frac{(N-1) \sum_{i=1}^{N-1} y_i y_{i+1} - \sum_{i=1}^{N-1} y_i \sum_{i=1}^{N-1} y_{i+1}}{(N-1) \sum_{i=1}^{N-1} y_i^2 - \left(\sum_{i=1}^{N-1} y_i \right)^2};$$

$$b = \frac{N \sum_{i=1}^N c^i y_i - \sum_{i=1}^N c^i \sum_{i=1}^N y_i}{N \sum_{i=1}^N c^{2i} - \left(\sum_{i=1}^N c^i \right)^2}; \quad a = \frac{\sum_{i=1}^N y_i - b \sum_{i=1}^N c^i}{N}. \quad (8.34)$$

У моделі другого типу параметри необхідно визначати за однією з ітераційних процедур мінімізації функціонала, що характеризує суму квадратів відхилень рівнів емпіричних точок від моделі. Зокрема у цьому випадку можна використовувати метод деформівного багатогранника.

Крива Гомперца описується рівняннями $\hat{y}_t = ab^{c^t}$ або $\hat{y}_t = ab^{c^{-t}}$, які логарифмуванням зводяться до узагальненої показникової функції першого або другого типу, відповідно.

Логістична функція $\hat{y}_t = \frac{1}{a + bc^t}$ або $\hat{y}_t = \frac{1}{a + bc^{-t}}$ зводиться до модифікованої показникової перетворенням $y^* = 1/\hat{y}_t = a + bc^{\pm t}$.

Модифіковану показникову функцію використовують як модель у випадках, коли досліджуваній залежності властиве насичення, тобто при збільшенні значень незалежної змінної відгук поступово наближається до певного граничного значення, а його прирости прямують до нуля. У таких випадках існує певний обмежувальний фактор, вплив якого збільшується із зростанням досягнутого рівня. Значення рівня насичення, як правило, можна задати, виходячи з наявних даних про об'єкт дослідження. У такому разі інші параметри моделі можна визначити методом найменших квадратів після її лінеаризації.

Якщо вплив обмежувального фактора виявляється лише після досягнення певного рівня розвитку процесу, слід використовувати моделі **S-подібного зро-**

стання, до яких належать крива Гомперца і логістична функція. Вони описують процеси в яких темп зростання поступово збільшується на початкових стадіях і поступово зменшується наприкінці. При цьому слід враховувати, що крива Гомперца є асиметричною, а логістична крива симетрична стосовно точки перегику. Процес побудови й дослідження логістичної моделі називають логістичним аналізом. Логістичну криву часто називають законом зростання, оскільки вона описує залежність кількості популяції або її біомаси від часу.

При побудові нелінійних регресійних моделей загального вигляду часто використовують методи, основані на мінімізації функціоналів виду (8.1). Для функціонала (8.1а) це призводить до необхідності розв'язання системи:

$$\left\{ \frac{\partial}{\partial \alpha_k} \sum_{i=1}^n [z(x_i) - y_i]^2 = 0. \right. \quad (8.35)$$

Якщо вона є системою лінійних рівнянь, застосовують звичайні алгоритми розв'язання таких систем – Гауса, простих ітерацій, Зейделя тощо. В окремих випадках, зокрема при логістичному аналізі, доцільніше розробляють спеціальні алгоритми. Якщо ж система (8.35) є нелінійною, використовують алгоритми нелінійної оптимізації: Ньютона–Рафсона, квазіньютонівські, спряжених градієнтів та інші.

8.5. Лінійні багатофакторні моделі

Як було зазначено вище лінійна як за параметрами, так і за незалежними змінними регресійна модель може бути записана у вигляді:

$$Y = \alpha_0 + \sum_{j=1}^p \alpha_j x_j + \varepsilon = \mathbf{Xb} + \mathbf{e}, \quad (8.36)$$

де Y – вектор-стовпчик відгуків, який має розмірність n ($n > p$);

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix} - \text{матриця значень } p \text{ незалежних змінних при } n \text{ ви-}$$

мірюваннях; α – вектор-стовпчик невідомих параметрів моделі, що має розмірність $p + 1$; ε – вектор-стовпчик похибок моделі, який має розмірність n .

Оцінки параметрів моделі у методі найменших квадратів отримують мінімізацією скалярного добутку:

$$Q = (\mathbf{Y} - \mathbf{X}\beta)^T (\mathbf{Y} - \mathbf{X}\beta), \quad (8.37)$$

де символ "Т" позначає транспонування.

Система (8.35) у цьому випадку набуває вигляду:

$$-2\mathbf{X}^T (\mathbf{Y} - \mathbf{X}\beta) = 0.$$

Звідси маємо:

$$\mathbf{X}^T \mathbf{Y} = \mathbf{X}^T \mathbf{X}\beta.$$

Помножуючи обидві частини цієї рівності ліворуч на матрицю $(\mathbf{X}^T \mathbf{X})^{-1}$

одержимо:

$$(\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{Y}) = (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{X})\beta.$$

Добуток $(\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{X}) = \mathbf{E}$, де \mathbf{E} – одинична матриця розмірності $p + 1$.

Тому остаточно маємо:

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}. \quad (8.38)$$

Одним з ускладнень, що можуть виникати при побудові регресійних моделей за наявності декількох предикторів, є можлива нерівноточність спостережень. У найпростішому випадку це можна врахувати за допомогою коваріаційної матриці такого вигляду:

$$\Omega = \begin{pmatrix} \sigma_1^2 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \sigma_n^2 \end{pmatrix}, \quad (8.39)$$

де σ_i^2 – дисперсія i -го спостереження. У цьому випадку формула (8.38) набуває вигляду:

$$\mathbf{b} = (\mathbf{X}^T \Omega^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Omega^{-1} \mathbf{Y}. \quad (8.40)$$

Такий метод побудови регресійних моделей називають **зваженим методом найменших квадратів**.

Іншим можливим ускладненням є наявність взаємозв'язку між ними, яку називають **мультиколінеарністю**. Припустимо, що існує лінійна залежність між компонентами вектора \mathbf{X} :

$$v_1 x_1 + v_2 x_2 + \dots + v_m x_m = 0. \quad (8.41)$$

Чим ближчою є ліва частина (8.41) до нульового вектора, тим сильнішою є мультиколінеарність. Граничний випадок точного виконання рівності (8.41) називають **строгою мультиколінеарністю**. Мультиколінеарність призводить до нестійкості обчислювальної процедури у зв'язку з високою похибкою обчис-

лення оберненої матриці. Коефіцієнти регресійної моделі у цьому випадку виявляються сильно корельованими один з одним, а довірчий рівень і дисперсія їх оцінок – підвищеними. Внаслідок цього інтерпретація результатів стає неможливою, а значення коефіцієнтів – статистично незначимими.

Існує декілька способів корегування мультиколінеарності. Одним з них є збільшення обсягу вибірки, за якою оцінюють значення параметрів моделі. Але, як правило, це неможливо. Інший підхід передбачає виявлення найсильніше корельованих змінних і об'єднання їх в один фактор. Він дає позитивні результати лише за умови, що таке об'єднання є теоретично обґрунтованим. В окремих випадках для усунення мультиколінеарності відкидають одну чи декілька сильно зв'язаних змінних, що призводить до появи нових похибок. У такому разі необхідно визначити, яка з похибок є більш істотною. Для цього можна по чергово виключати зв'язані змінні й порівнювати одержувані результати. Жоден з перелічених методів не дає змоги суттєво підвищити точність моделі.

Альтернативою їм є **методи зміщеного оцінювання**, зокрема гребеневі оцінки, редуковані оцінки, оцінки Марквардта, оцінки Хоккінса тощо. Якщо коваріаційні матриці є погано зумовленими, при обчисленні оцінок коефіцієнтів регресії застосовують також метод регуляризації О.М. Тихонова. Завданням оцінювання при їх застосуванні є одержання значень параметрів регресії, які були б стійкими навіть за умови сильної спряженості незалежних змінних.

У граничному випадку, коли матриця $X'X$ є одиничною, оцінки, одержувані за формулою (8.41) є незміщеними і мають мінімальну дисперсію. Із збільшенням власних чисел матриці $X'X$ відстань між оцінками $\hat{\beta}$ та істинними β збільшується. Крім того стає можливою зміна напрямку впливу вхідних змінних на вихідні.

На сьогодні розроблено понад вісімдесят алгоритмів зміщеного оцінювання параметрів лінійних регресійних моделей. Їх поділяють на такі групи: методи звичайних гребневих оцінок, методи узагальнених гребневих оцінок, методи оцінок дробового рангу, методи стиснутих оцінок. Всі вони є лінійними перетвореннями оцінок, одержуваних МНК.

У методах гребеневого аналізу ставиться завдання отримання оцінок з мінімальною дисперсією. Оцінка гребеневої регресії \mathbf{b}_k є лінійним перетворенням оцінки \mathbf{a} , отриманої МНК, і залежить від параметра k і матриці вихідних даних \mathbf{X} . Її можна записати у вигляді:

$$\mathbf{b}_k = (\mathbf{X}^T \mathbf{X} - k\mathbf{E})^{-1} \mathbf{X}^T \mathbf{Y}. \quad (8.42)$$

Ефективність методів гребневих оцінок залежить від статистичних характеристик вихідної інформації і оптимального вибору параметра k . Метод гребеневого аналізу, як і інші методи зміщеного оцінювання параметрів регресійних моделей, за певних умов можуть бути обґрунтовані теоретично, але у більшості практичних ситуацій перевірити виконання цих умов неможливо. Тому застосування цих методів потребує певної обережності.

Параметри нелінійних регресійних моделей за наявності декількох незалежних змінних зазвичай оцінюють з використанням чисельних методів нелінійної мінімізації функціоналу (8.1а). У багатьох випадках це завдання може бути істотно спрощено з використанням методів планування експерименту, які дають змогу визначити оптимальні для подальшого аналізу плани, тобто точки простору незалежних ознак, у яких потрібно здійснити вимірювання значень відгуків.

8.6. Перевірка адекватності регресійних моделей

Основні методи перевірки адекватності регресійних моделей ґрунтуються на таких трьох властивостях їх залишків.

По-перше, для адекватної моделі стандартне дисперсія залишків має бути близькою до дисперсії емпіричних точок. При цьому припускають, що дисперсії усіх емпіричних точок є однаковими. Причиною неадекватності при невиконанні цієї властивості є використання надмірно спрощених або ускладнених регресійних моделей. Відомо, наприклад, що за наявності n емпіричних точок

можна побудувати поліноміальну модель $n - 1$ порядку, яка пройде строго через всі ці точки. Але використовувати такий поліном як регресійну модель за наявності похибок емпіричних даних, очевидно, недоцільно. З іншого боку, якщо ступінь полінома буде надто малою, то він не відтворюватиме істотних рис досліджуваної залежності. Тому існує певне оптимальне значення степені такого полінома. Як критерій виконання цієї властивості часто застосовують такий критерій:

$$\frac{S}{\Delta^2} \leq F, \quad \frac{\Delta^2}{S} \leq F, \quad (8.43)$$

де S – значення цільового функціонала (8.1а), Δ^2 – сума квадратів похибок емпіричних даних по всім точкам, F – критичне значення критерію Фішера для вибраного рівня значущості й кількостей ступенів вільності, що дорівнюють $n - 1$.

Невиконання першої умови свідчить про надмірну спрощеність моделі, зокрема про необхідність збільшення порядку поліноміальної моделі. Невиконання другої умови є свідченням того, що модель треба спростити, наприклад зменшити порядок полінома.

Інша властивість залишків, яку перевіряють при визначенні адекватності моделі, полягає у тому, що вони мають підпорядковуватися нормальному закону розподілу з нульовим математичним сподіванням.

Найпростіші способи перевірки нормальності ряду залишків ґрунтуються на розрахунку значень коефіцієнтів його асиметрії та ексцесу. Вважають, що розподіл є близьким до нормального, якщо $|A|, |E| < 0,1$, і сильно відрізняється від такого, коли $|A|, |E| > 0,5$.

Для перевірки нульової гіпотези про відповідність розподілу залишків моделі нормальному закону розподілу можна використовувати критерій χ^2 . Він оперує не первинними даними, а їх розподілом за класами рів-

ної ширини. Тому необхідно враховувати вимогу щодо мінімальних обсягів ряду спостережень та кількості класів. За різними оцінками мінімальна допустима кількість класів знаходиться у межах 4–7, а кількість елементів у ряді спостережень – в межах 20–40.

Для формування класових інтервалів спочатку визначають найменше й найбільше значення у ряді залишків $e_i = z(x_i) - y_i$. Потім розраховують ширину класу $h = \frac{\max(e_i) - \min(e_i)}{N}$, де N – кількість класів. Далі призначають межі класів:

$$h = \frac{\max(e_i) - \min(e_i)}{N}$$

начають межі класів:

$$1\text{-й} - [\min(e_i); \min(e_i) + h);$$

$$2\text{-й} - [\min(e_i + h); \min(e_i) + 2h);$$

...

$$n\text{-й} - [\min(e_i) + (n - 1)h; \min(e_i) + nh = \max(e_i)]$$

і підраховують частоти потрапляння результатів до цих інтервалів.

Значення критерію розраховують за формулою:

$$\chi^2 = \sum_{k=1}^N \frac{(v_k - np'_k)^2}{np'_k}, \quad (8.44)$$

де v_k – абсолютні частоти для N класів; p'_k – теоретичні ймовірності потрапляння до k -го класу для нормального закону розподілу з математичним сподіванням рівним нулю й стандартним відхиленням, що відповідає стандартному відхиленню ряду залишків; n – кількість спостережень. Розраховану величину порівнюють з критичним значенням розподілу χ^2_c для відповідного рівня значущості й кількості степенів вільності $N - 3$. Гіпотезу про відповідність залишків нормальному закону розподілу приймають за умови $\chi^2 \leq \chi^2_c$.

Перевірку нормальності розподілу ряду залишків моделі можна здійснити також за результатами аналізу квантилей. Квантилем порядку q функції розподілу $F(x)$ називають корінь X_q рівняння:

$$P(X \leq X_q) = F(X_q) = q. \quad (8.45)$$

Розподіл ряду залишків можна вважати близьким до нормального, якщо значення окремих квантилей є близькими до наведених нижче величин:

- ◆ 2,5% процентіль $\approx -2\sigma$;
- ◆ 16% процентіль $\approx -\sigma$;
- ◆ 50% процентіль ≈ 0 ;
- ◆ 84% процентіль $\approx \sigma$;
- ◆ 97,5% процентіль $\approx 2\sigma$.

Якщо ряд залишків підпорядковується нормальному закону розподілу з відомою дисперсією, то як критерій рівності його середнього значення нулю можна використовувати величину:

$$Z = \frac{\bar{x}\sqrt{n}}{\sigma}, \quad (8.46)$$

де \bar{x} – середнє арифметичне ряду залишків, σ – його дисперсія.

Критерій Z є випадковою величиною, що підпорядковується стандартному нормальному розподілу. Якщо конкуруючою гіпотезою є $\bar{X}_1 \neq \bar{X}_2$, то праву критичну точку можна визначити з умови:

$$P(0 < Z < z_{\text{прав.кр.}}) = \Phi(z) = (1 - \alpha) / 2, \quad (8.47)$$

де $\Phi(z)$ є функцією Лапласа (інтегралом ймовірностей), що пов'язана з функцією стандартного нормального розподілу $F(z)$ співвідношенням $\Phi(z) = F(z) - 1/2$. Ліва й права критичні точки для одного й того самого рівня значущості пов'язані умовою $Z_{\text{лів.кр.}} = -Z_{\text{прав.кр.}}$.

Нехай, наприклад, середнє арифметичне залишків $\bar{x} = 3$, $\sigma^2 = 121$, $n = 40$. Як нульову гіпотезу візьмемо твердження, що математичне сподівання залишків дорівнює нулю. Тоді:

$$Z = \frac{3\sqrt{40}}{121\sqrt{2}} \approx 1,22.$$

Беручи рівень значущості рівним 0,05 (5%), отримаємо для визначення критичної точки умову:

$$\Phi(z_{\text{прав.кр.}}) = (1 - \alpha) / 2 = 0,475.$$

Звідси $z_{\text{прав.кр.}} = 1,96$. Оскільки $Z < z_{\text{прав.кр.}}$, ми можемо прийняти нульову гіпотезу на рівні значущості 0,05.

Для перевірки гіпотези про рівність середнього значення нулю при невідомій дисперсії залишків застосовують t-критерій Стьюдента. Його значення розраховують за формулою:

$$t = \frac{\bar{x}\sqrt{n}}{s}, \tag{8.48}$$

де s – стандартне відхилення залишків. Відповідна статистика має розподіл Стьюдента з $k = n - 1$ ступенями вільності. Критичні точки симетричні стосовно нуля. Нульову гіпотезу відхиляють: якщо $|t| > t_{1-\alpha/2, k}$.

Третьою властивістю є те, що залишки адекватної регресійної моделі мають бути некорельованими випадковими величинами. Наявність автокореляції першого порядку перевіряють за допомогою критерію Дарбіна-Уотсона:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}, \quad (8.49)$$

де n – кількість емпіричних точок. Для адекватної моделі має виконуватися умова $d \approx 2$. Близькі до нуля значення d свідчать про наявність додатної автокореляції, а значення, що наближаються до 4, – про наявність від'ємної автокореляції.

Наявність автокореляції вищих порядків перевіряють шляхом дослідження автокореляційної функції. Про наявність автокореляції у цьому випадку свідчить збільшення абсолютних значень коефіцієнта автокореляції при певних значеннях параметра зсуву. На рис 8.5 показано приклади автокореляційних функцій для ряду, що є білим шумом, (ліворуч) та рядом, який змінюється за синусоїдальним законом, (праворуч).

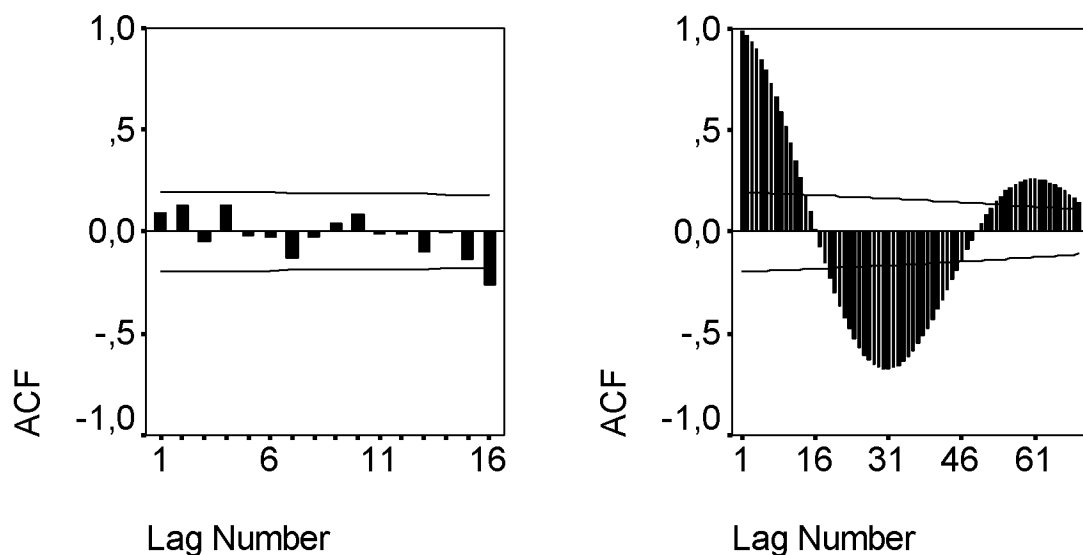


Рис. 8.5. Приклади автокореляційних функцій деяких рядів

Горизонтальними лініями на цих графіках показано довірчі інтервали для нульових значень коефіцієнта автокореляції. З наведених графіків добре видно, що у першому випадку автокореляція є практично відсутньою, а у другому – для певних значень параметра зсуву спостерігається істотна додатна або від'ємна автокореляція.

Контрольні питання

1. Яким є основне завдання регресійного аналізу?
2. У чому полягають основні припущення класичного регресійного аналізу?
3. Якою є звичайна процедура класичного регресійного аналізу?
4. Як формулюється задача побудови регресійної моделі?
5. Які функціонали використовують для визначення параметрів регресійних моделей? У чому полягають переваги й недоліки різних типів таких функціоналів?
6. Якими є основні типи функцій, що використовуються для побудови однофакторних регресійних моделей?
7. Які моделі називають лінійними? Що називають порядком регресійної моделі?
8. Чому регресійні моделі не рекомендують використовувати поза межами тієї області значень вихідних параметрів, для якої вони побудовані?
9. Для заданого набору даних побудувати однофакторну лінійну регресійну модель і перевірити її адекватність.
10. У яких випадках нелінійні однофакторні моделі можна звести до лінійних? Навести приклади відповідних перетворень.
11. Для заданого набору даних побудувати однофакторну нелінійну регресійну модель і перевірити її адекватність.
12. Як використовують критерій Фішера для перевірки адекватності регресійних моделей?

13. Як визначають довірчі інтервали для коефіцієнтів однофакторних регресійних моделей?
14. Яким є загальний вигляд поліноміальної регресійної моделі?
15. Яким є загальний алгоритм визначення порядку і параметрів поліноміальних регресійних моделей?
16. Для заданого набору даних побудувати поліноміальну регресійну модель і перевірити її адекватність.
17. У яких випадках використовують регресійні моделі у вигляді тригонометричних поліномів? Яким є загальний алгоритм побудови таких моделей?
18. Для заданого набору даних побудувати регресійну модель у вигляді тригонометричного поліному і перевірити її адекватність.
19. Якими є загальні алгоритми побудови однофакторних регресійних моделей у вигляді модифікованої показникової функції, кривої Гомперця та логістичної кривої?
20. Яким є загальний алгоритм побудови багатофакторної лінійної регресійної моделі?
21. Для заданого набору даних побудувати багатофакторну лінійну регресійну модель і перевірити її адекватність.
22. Що називають мультиколінеарністю даних? Навести приклади.
23. Для чого застосовують алгоритми зміщеного оцінювання параметрів багатофакторних лінійних регресійних моделей? Навести приклади.
24. За якими властивостями перевіряють адекватність регресійних моделей? Якими є основні критерії адекватності?

ЛІТЕРАТУРА

Монографії та навчальні посібники

1. Ажогин В.В., Згуровский М.З. Моделирование на цифровых, аналоговых и гибридных ЭВМ. – Киев: Вища шк, 1983.
2. Бахрушин В.Є. Математичне моделювання. – Запоріжжя: ГУ "ЗІДМУ", 2004.
3. Бахрушин В.Є. Аналіз даних. – Запоріжжя: ГУ "ЗІДМУ", 2005.
4. Бахрушин В.Є. Часові ряди. – Запоріжжя: ГУ "ЗІДМУ", 2006.
5. Березин Ю.А. Моделирование нелинейных волновых процессов. – Новосибирск: Наука, 1982.
6. Бешелев С.Д., Гурвич Ф.Г. Математико-статистические методы экспертных оценок. – М.: Статистика, 1980.
7. Бусленко Н.П. Моделирование сложных систем. – М.: Наука, 1968.
8. Глушков В.М., Иванов В.В., Яненко В.М. Моделирование развивающихся систем. – М.: Наука, 1983.
9. Згуровский М.З., Мельник В.С. Нелинейный анализ и управление бесконечномерными системами. – Киев: Наукова думка, 1999.
10. Згуровский М.З., Новиков А.Н. Системный анализ стохастических распределенных процессов (моделирование, оценивание состояний, идентификация): Учеб. пособ. – К: Выща школа, 1988.
11. Ивахненко А.Г., Юрачковский Ю.П. Моделирование сложных систем по экспериментальным данным. – М.: Радио и связь, 1986.
12. Кафаров В.В., Глебов М.Б. Математическое моделирование основных процессов химического производства. – М.: Высшая школа, 1991.
13. Костицын В.А. Эволюция атмосферы, биосферы и климата. – М.: Наука, 1984.
14. Марчук Г.И. Математическое моделирование в проблемах окружающей среды. – М.: Наука, 1982.

15. Месарович М., Такахара Я. Общая теория систем: математические основы. – М.: Мир, 1978.
16. Моисеев Н.Н. Математика ставит эксперимент. – М.: Наука, 1979.
17. Моисеев Н.Н. Математические задачи системного анализа. – М.: Наука, 1981.
18. Моисеев Н.Н., Александров В.В., Тарко А.М. Человек и биосфера. Опыт системного анализа и эксперименты с моделями. – М.: Наука, 1985.
19. Мухин О.И. Моделирование систем: Конспект лекций. – Ч. 1. – Пермь: Пермский ГТУ, 1999.
20. Николис Дж. Динамика иерархических систем. – М.: Мир, 1989.
21. Советов Б.Я., Яковлев С.А. Моделирование систем. – М.: Высшая школа, 2006.
22. Советов Б.Я. Моделирование систем. Практикум. – М.: Высшая школа, 1999.
23. Тихонов А.Н., Кальнер В.Д., Гласко В.Б. Математическое моделирование технологических процессов и метод обратных задач в машиностроении. – М.: Машиностроение, 1990.
24. Томашевский В.М. Моделирование систем. – К.: ВНУ, 2005.
25. Хазанов Л.Э. Математическое моделирование в экономике. – М.: БЕК, 1998.
26. Экономико-математические методы и прикладные модели / Под ред. В.В. Федосеева. – М.: Юнити, 1999.

Періодичні наукові та науково-популярні видання

1. Автоматика и телемеханика.
2. Динамические системы.
3. Дискретный анализ и исследование операций.
4. Дифференциальные уравнения и процессы управления.
5. Журнал вычислительной математики и математической физики.
6. Известия РАН. Сер. Техническая кибернетика.

7. Кібернетика та системний аналіз.
8. Математическая физика. Анализ. Геометрия.
9. Математичне моделювання.
10. Наука и жизнь.
11. Природа.
12. Проблемы кибернетики.
13. Проблемы управления и информатики.
14. Системні технології.
15. Системные исследования.
16. Складні системи і процеси.
17. Теоретическая и математическая физика.
18. Управляющие системы и машины.
19. Экономика и математические методы.
20. Экономическая кибернетика.
21. Econometrica.
22. Economic Modeling.
23. Electronic Transactions on Numerical Analysis.
24. Journal of Association of Computing Machinery.
25. Journal of Economic Behavior and Organization.
26. Journal of Economic Dynamic and Control.
27. Journal of Mathematical Economics.
28. Journal of Molecular Modeling.
29. Management Science.
30. Management Technology.
31. Operation Research.
32. The Journal of Non-linear Oscillations.