

**для МІНІСТЕРСТВО ТРАНСПОРТУ ТА ЗВ'ЯЗКУ УКРАЇНИ
ДЕРЖАВНА АДМІНІСТРАЦІЯ ЗВ'ЯЗКУ**

Одеська національна академія зв'язку ім. О.С. Попова

Кафедра управління проектами та системного аналізу

Бобровнича Н.С., Борисевич Є. Г.

Е К О Н О М Е Т Р І Я

Навчальний посібник

7.050107 – Економіка підприємства

7.050201 – Менеджмент організації

Одеса 2010

УДК 330.43

ББК 65.051.73

Б72

План НМВ 2010 р.

Бобровнича Н. С., Борисевич Є. Г.

Б72 Економетрія: навч. посіб./ Н. С. Бобровнича, Є. Г. Борисевич. – Одеса: ОНАЗ ім. О.С. Попова, 2010. – 180 с.

Викладаються методологічні основи моделювання складання складних економічних систем і вивчення економічних процесів за допомогою причинового та регресійного аналізу. Розглядаються проблеми прогнозування та перевірки їх на адекватність реальним економічним процесам.

У посібнику також розглядаються основні задачі та алгоритми сіткового планування та управління.

Навчальний посібник призначений для студентів, що навчаються на економічному факультеті і може бути використаний при вивченні питань економетрії на технічних факультетах.

Схвалено на засіданні кафедри Управління проектами та системного аналізу
Протокол №11 від 03.07.09
Затверджено методичною радою академії зв'язку
Протокол № 1 від 08.09.2009 р.

© Бобровнича Н. С.,

© Борисевич Є. Г.,

© ОНАЗ ім. О. С. Попова, 2010

ЗМІСТ

ВСТУП	5
Розділ 1. Історичні відомості про становлення предмета економетрії та зв'язок із математико-статистичними методами	7
1.1. Історичні відомості про становлення предмета економетрії.....	7
1.2. Постановка завдань та цілей дослідження економетрії. Будова економетрії.....	11
Розділ 2. Методологічні основи моделювання складних економічних систем, процесів	19
2.1. Поняття системи. Математичне моделювання.....	19
2.2. Класифікація систем.....	30
Розділ 3. Причиново-наслідкові відношення. Причиновий аналіз	35
3.1. Про причиновість у соціально-економічних явищах і процесах.....	35
3.2. Необхідність формалізації причиново-наслідкових відношень у вивченні економічних процесів.....	36
3.3. Про алгебру каузаций. Причиновий аналіз як один із засобів формалізації причиново-наслідкових відносин на різних класах економічних моделей.....	37
Розділ 4. Регресійний аналіз. Його особливості та різновиди	56
4.1. Про економічні величини.....	56
4.2. Про регресію взагалі.....	57
4.3. Основи регресійного аналізу. Його задачі.....	58
4.4. Класифікація форм регресії.....	61
4.5. Діаграми розсіювання та побудова регресійних функцій.....	64
Розділ 5. Метод найменших квадратів. Його сильні та слабкі сторони Узагальнений МНК	67
5.1. Метод найменших квадратів (МНК).....	67
5.2. Узагальнення МНК.....	76
5.3. Альтернативні методи для оцінок параметрів моделі.....	80
5.4. Абсурдні рішення, чуттєвість МНК, усереднення „абсурдностей”.....	85
5.5. Методи відбору найбільш значимих факторів і найбільш значимих членів адитивної моделі. Їх машинна інтерпретація.....	87
5.6. Економічна інтерпретація результатів, які отримані за допомогою МНК.....	89
Розділ 6. Перевірка моделі на адекватність. Проблема прогнозування	91
6.1. Математичні проблеми перевірки на адекватність регресійних Моделей реальним економічним процесам.....	91

6.2. Статистичні аспекти, які пов'язані з обчисленням надійних інтервалів моделі, її поведінки у майбутньому.....	101
6.3. Питання, що пов'язані з обчисленням надійних інтервалів моделі, її поведінки у майбутньому.....	108
Розділ 7. Мультиколінеарність. Гомо- і гетероскедастичність.....	111
7.1. Мультиколінеарність. Її наслідки.....	111
7.2. Гомо- та гетероскедастичність.....	116
Розділ 8. Автокореляція.....	120
8.1. Автокореляція. Її наслідки.....	120
8.2. Критерії Дарбіна-Уотсона (d-статистика) і фон Неймана.....	123
8.3. Додаткові моменти щодо автокореляції.....	124
8.4. Оцінка параметрів моделі з автокорельованими залишками.....	125
Розділ 9. Сіткове планування та управління.....	128
9.1. Основні поняття та правила побудови сіткових графіків.....	128
9.2. Параметри сіткового графіка та способи їх розрахунку.....	131
9.3. Аналіз та оптимізація сіткового графіка.....	135
Додаток А – Структура мережі зв'язку.....	139
Додаток Б – Алгоритми пошуку шляхів у мереж.....	147
Список літератури.....	176

ВСТУП

На початку ХХ століття тенденції розвитку економічних теорій за умов безперервно зростаючої складності соціально-економічних процесів, бурхливого розвитку науки і промисловості, форм господарювання і, як наслідок, ускладнення управлінських функцій, практично на всіх рівнях ієрархії економічних систем, зумовили створення нового напрямку в науці про соціально-економічні явища, що дістав назву "економетрія" (від німецької "Okonometrie"). Вперше цей термін запропонував у 1910 році львівський вчений П. Чомпа в німецькій книзі з бухгалтерського обліку – „Нариси економетрії і природної теорії бухгалтерії”, яка ґрунтується на політичній економії. В економічну науку, термін "економетрія" був уведений норвезьким статистиком Рагнарм Фрішем (у подальшому він став, разом з нідерландським вченим Яном Тінбергеном першим лауреатом Нобелівської премії в галузі економіки за їхні праці, які пов'язані з економетричними моделями прийняття рішень, 1969 р.), де у доповіді, опублікованій у 1926 році, ввів його для позначення самостійної галузі наукових досліджень в економіці. Метою цих досліджень було установлення конкретних кількісних закономірностей та отримання їх оцінок в економічних процесах за допомогою математичних методів (в основному статистичних). Зазначимо, що іноді використовується термін "економетрика", хоча це, на наш погляд, неправомірно, оскільки ні про яку метрику в цій дисципліні не йдеться. Поширення терміна "економетрика" в сучасній вітчизняній літературі, на наш погляд, пов'язане з неточним перекладом цього терміна з англійської мови "Econometric" російською та українською мовами. Ситуація може скластися так, як це було (див. розділ 2) з перекладом англійського слова "system analysis", яке переклали не як "аналіз систем", а "системний аналіз". До чого це призвело? До появи нового наукового напрямку – "системного аналізу" - дослідження складних систем будь-якої організації та орієнтації у найширшому розумінні цього слова. Якщо подивитися у „Словнику іншомовних слів” (К.:АН УРСР, 1974. – с.430), то – "метрія" (від грец. μετρεω – вимірюю, у складних словах відповідає поняттю „вимірювання”), а „метрика” (грец. μετρίχη від грец. μετρον – міра).

Для вирішення конкретних економічних задач Р.Фріш почав залучати вчених різних країн, які об'єдналися в економетричне міжнародне товариство – Econometric Society (Міжнародне об'єднання для розвитку економічної теорії і її зв'язку зі статистикою та математикою). Для обміну досвідом воно почало видавати з 1933 року щомісячний журнал "*Econometrica*".

Історія виникнення нової галузі знань і формування кола вчених, які займалися проблематикою потреб, які виникли в економіці, описується у [13]: "Тридцятичотиритлітній норвезький викладач економіки університету в Осло Р.Фріш в 1928 р. зустрів Чарльза Ф.Руса – молодого співробітника математичного факультету Принстонського університету (секретаря секції з економіки, соціології і статистики Американського товариства розвитку наук). Ф. Рус і Р.Фріш вирішили діяти разом. На першому етапі вони звернулись за допомогою до Ірвінга Фішера (одного з найвідоміших американських вчених, який за-

ймався математичною статистикою), і в квітні 1928 року всі троє зустрілись у Нью-Хейвені. І.Фішер не поділяв оптимізму колег, але обіцяв співробітничати, якщо Ф. Рус і Р. Фріш знайдуть 100 осіб, які зацікавлені у створенні такого товариства і стануть його членами. Складений на початку список включав близько як 80 осіб. Однак було вирішено продовжувати роботу і розпочати листування з ними. Листи з пропозиціями адресатів було зустрінuto доброзичливо і залучено ще майже 80 нових осіб". Таким чином, 29 грудня 1930 року в Клівленді було засновано економетричне товариство, яке дало змогу вченим різних країн сконцентрувати свої зусилля на нових напрямках вивчення складних економічних систем та процесів і явищ, що в них відбуваються.

Економісти, переконавшись у достовірності та науковому обґрунтуванні отримуваних результатів і ефективності використовуваних методів, визнали економетрію як самостійну наукову дисципліну, основним завданням якої є розвиток економічної теорії в її зв'язку зі статистикою і математикою.

РОЗДІЛ 1

ІСТОРИЧНІ ВІДОМОСТІ ПРО СТАНОВЛЕННЯ ПРЕДМЕТА ЕКОНОМЕТРІЇ ТА ЗВ'ЯЗОК ІЗ МАТЕМАТИКО-СТАТИСТИЧНИМИ МЕТОДАМИ

1.1 Історичні відомості про становлення предмета економетрії

Визначаючи предмет економетрії, наведемо її визначення різними вченими. Так, польський економіст і статистик О. Ланге писав, що економетрія займається визначенням спостережуваних у економічному житті конкретних кількісних закономірностей, застосовуючи для цієї мети статистичні методи. Французький вчений Е. Маленво розглядав економетрію як потужний інструмент для виявлення найбільш стійких характеристик у поведінці реальних економічних одиниць серед різних підходів: "Метою економетрії є емпіричне виведення законів. Економетрія доповнює економічну теорію, використовуючи реальні дані для перевірки та уточнення постульованих відношень". Таким чином, Е. Маленво розглядав економетрію, як частину позитивної економіки, яка займається "аналізом фактів (даних), на основі яких формуються принципи економічної поведінки".

За визначенням Т. Шаттеліса "...економетрія – це термін, який надається поняттю, сфера якого належить до дуже складної наукової дисципліни, яка народилась у галузі інтерференції математики, статистики й економіки. Предметом економетрії є економіка в кількісному аспекті. Об'єкт вивчення – зв'язки між народногосподарським ансамблем, його найпростішими компонентами і поведінкою цих компонент. Таким чином, зі сфери економетрії виключаються, наприклад, економіка підприємств, кількісні аспекти якої стосуються дослідження операцій. Галузь економетрії визначається моделями математичної економіки. Економетрія шукає числові значення параметрів моделей математичної економіки. Вона займається вимірюванням економічних об'єктів як аналогічних еквівалентів математичної моделі. Методом економетрії є метод, запропонований математичною статистикою".

Лоуренс Р. Клейн (лауреат Нобелівської премії в галузі економіки (1980 р.), за видатний внесок у розвиток створення великих економетричних моделей і їх широкого застосування для аналізу кон'юнктурних коливань і економічної політики) визначав економетрію як науку, яка вивчає вимірювання зв'язків у відповідному економічному аналізі: "Економетрія, як наука, вивчає вимірювання зв'язків у відповідному економічному аналізі. Основне завдання економетрії – наповнення емпіричним змістом апріорних економічних суджень".

Г. Тінтер – ототожнював економетрію з економічною статистикою. Г. Хансен під економетрією розумів застосування математико-статистичних методів в економіці. Видатний економіст Самуельсон (лауреат Нобелівської премії (1970 р.) економетрію визначив як науку, яка дає можливість здійснювати кі-

лькисний аналіз реальних економічних явищ. Дж. Джонстон пише, що "найбільш суттєвим завданням економетричного дослідження є оцінка і перевірка економічної моделі" [8, с.5].

В енциклопедії кібернетики (1974 р., с.544) економетрія визначається як: "...напрям в економіці, заснований на використанні математичних моделей для аналізу і прогнозування економічних явищ і пов'язаний з визначенням та оцінкою адекватності реальних явищ математичним уявленням про них. ...Важко розділити математичну економіку й економетрію, з одного боку, економетрію і економічну статистику, – з іншого. Можна лише підкреслити зв'язок математичної економіки й економетрії. Побудова математичної моделі економіки завжди підтверджується оцінками адекватності такої моделі реальній дійсності. Економічна статистика має справу з ustalеними і відносно нескладними економічними обчисленнями. Поява економетрії пов'язана з твердженнями про недостатність таких економіко-статистичних обчислень для економічного аналізу і прогнозування. Найбільшого розвитку набули в економетрії методи множинної кореляції. Висновки, отримані за допомогою економетричних побудов, мають обмежене значення, у будь-якому разі, надійними можна вважати оцінки при незначних змінах параметрів. Досвід свідчить, що досить точне прогнозування характеристик і показників вимагає внесення у моделі факторів соціального значення".

Російські вчені С. А. Айвазян і В. С. Мхитарян [1] визначають економетрію як: "економетрика – економіко-математична наукова дисципліна, яка розробляє і використовує методи, моделі, прийоми, що дозволяють додавати конкретне кількісне вираження загальним (якісним) закономірностям економічної теорії на базі економічної статистики і з використанням математико-статистичного інструментарію".

Зазначимо, що до недавнього часу в літературі з економетрії спостерігалася справжній "голод", тому що ця дисципліна не викладалася у вузах колишнього СРСР і подавалася лише як критичний аналіз буржуазних економічних теорій. Наприклад, як відзначалося у [69]: "Критичне дослідження сучасних економетричних концепцій міжнародних економічних відношень необхідне для нас також і тому, що ці концепції і сконструйовані на їхній основі моделі, у більшості випадків суперечать марксистсько-ленінській теорії світової торгівлі і міжнародного поділу праці". У [1] з цього приводу також зазначалося, що: "...із трьох складових економетрики – економічної теорії, економічної статистики і математико-статистичного інструментарію дві перші були представлені в нашій країні явно незадовільно. Не було якісної економічної теорії, не було систем національних рахунків і необхідного інформаційного забезпечення економетричного моделювання".

У літературі, яка з'явилася в Україні, щодо визначення економетрії теж немає єдиного погляду. Наприклад, в [22], економетрія визначається як наука, що вивчає методи оцінювання параметрів економетричних моделей, які характеризують її кількісні взаємозв'язки між економічними показниками, а також розглядає основні напрямки застосування цих моделей в економічних дослідженнях". У [18] економетрика визначається як наука, що вивчає кількісні

закономірності та взаємозв'язки економічних об'єктів і процесів за допомогою математико-статистичних методів та моделей. У [25] під економетрикою розуміється "фундаментальна економіко-математична наука, яка на основі статистичних даних про соціально-економічні процеси вивчає методику побудови економічних моделей для відображення закономірностей, кількісних зв'язків, динаміки цих процесів в економічному просторі з метою прогнозування, аналізу взаємного впливу явищ та прийняття оптимальних рішень щодо планування, розподілу матеріальних, трудових, фінансових ресурсів".

Існують й інші визначення економетрії, але в більшості випадків вони (так чи інакше) повторюють вже існуючі визначення.

Із вищевикладеного випливає, що досі, не зважаючи на нібито міцно сталий термін "економетрія", однозначного визначення його немає. Оскільки економетрія – це самостійна наукова дисципліна, то з метою виключення в подальшому неоднозначності в розумінні її суті наведемо найбільш повне, на наш погляд, її визначення, що відображає мету, об'єкт і завдання досліджень, методологічні принципи, які застосовуються, і методи та засоби, що використовуються.

Узагальнюючи всі визначення економетрії і додаючи до них в явному вигляді вплив параметрів зовнішнього середовища (у тому числі соціальних) на економічні системи, що досліджуються, модельні експерименти над економетричними моделями, які моделюють ці системи з їх кінцевою метою, остаточно визначимо предмет нашого дослідження.

Економетрія – прикладна економіко-математична дисципліна, яка вивчає динаміку реальних мікро- та макроекономічних явищ і процесів в економічних просторах та економічних системах за умов впливу стохастичних соціально-економічних факторів із застосуванням спеціальних математичних моделей і методів, які дають можливість здійснити відтворювані модельні досліди для квантифікованого (кількісного та якісного) аналізу і прогнозування результатів розвитку економічних систем, процесів і явищ, що в них відбуваються, з метою подальшого управління ними.

Економетрія є базовою основою для визначення класів альтернативних рішень у системах підтримки прийняття управлінських рішень на рівнях мікро- та макроекономіки.

За умов трансформованої, перехідної або ринкової економіки (до цього часу використовують для позначення сучасного стану української економіки всі ці терміни), економетрія повинна відіграти одну з найважливіших ролей у стратегічному і перспективному плануванні та проведенні економічної політики країни на різних рівнях ієрархії управління. Це пов'язане з підвищенням значення та новим осмисленням таких економічних категорій, як податок, процент, кредит, прибуток, ціна та ін., з посиленням ролі розподільних відношень у народному господарстві країни, яке обумовлене виникненням нових форм власності (створенням малих, кооперативних, спільних підприємств, фондових бірж, акціонерних товариств, орендних колективів тощо), поглибленням економічної кризи, інфляцією, порушенням традиційних економічних зв'язків, входженням країни у світове економічне співтовариство і т. ін. Процеси, що

породжуються в умовах перехідної економіки, потребують науково обґрунтованого соціально-економічного осмислення їх суті та взаємозв'язку в умовах нових економічних відносин, демократизації державної системи.

Економетрія займається квантифікаційною (кількісною та якісною) оцінкою характеристик процесів і явищ, які відбуваються в різних економічних сферах (розподіл і кінцеве використання валового національного продукту, створення первинних і кінцевих доходів домашніх господарств, державних закладів і громадських організацій, накопичення основних фондів і матеріальних оборотних засобів, капіталовкладення у житлове будівництво, попит державних структур і різних фінансових інститутів, виробничі капіталовкладення, експорт, імпорт, зайнятість, безробіття, заробітна плата та ціни, доходи та витрати державних структур, виробництво, інвестування тощо), виробленням альтернативних науково обґрунтованих рекомендацій щодо їх управління.

Розвиток економетрії відбувається в тісному зв'язку із загальними соціально-економічними завданнями, які вирішуються державою на сучасному етапі трансформації економіки України. В економетрії мабуть більш, ніж у будь-якому іншому науковому напрямку, накопичилось багато невирішених проблем науково-методологічного характеру, які особливо проявилися в умовах трансформації економіки (яка функціонує в умовах невизначеності, конфлікту та жорстких ресурсних обмежень), корінної перебудови "живого організму" всієї соціально-економічної сфери держави.

Місце, яке посідає економетрія у системі управління економікою на будь-якому рівні її організації можна уявити з таких міркувань. Процес обґрунтування оперативних вимог до економічної системи будь-якої складності включає два взаємопов'язані боки: методологічний і цільовий.

З методологічного боку – в основу планування розвитку економічної системи покладено програмно-цільовий підхід з використанням організаційно-функціонального інструменту його реалізації у вигляді певної системи планування її розвитку. Під програмою розуміється сукупність заходів, які забезпечують досягнення кінцевих цілей даного етапу розвитку (етапу довгострокового планування) [59].

З цільового боку сутність планування розвитку економічної системи відбиває його кінцевий конкретний підсумок: створення збалансованої системи, яка оптимально поєднує різні за своїм призначенням формування, що забезпечують розв'язання економічних задач. У кінцевому рахунку цільовий бік сутності планування економічною системою має знаходити своє відображення у відповідній частині програми її розвитку. В основу цільового боку покладено головні цілі та інтереси економічної системи, а також вимоги до стратегії щодо їх забезпечення.

Процес планування розвитку економічної системи повинен починатися з визначення її цілей, які пов'язуються зі спектром загроз щодо її економічної безпеки (рис. 1.1).

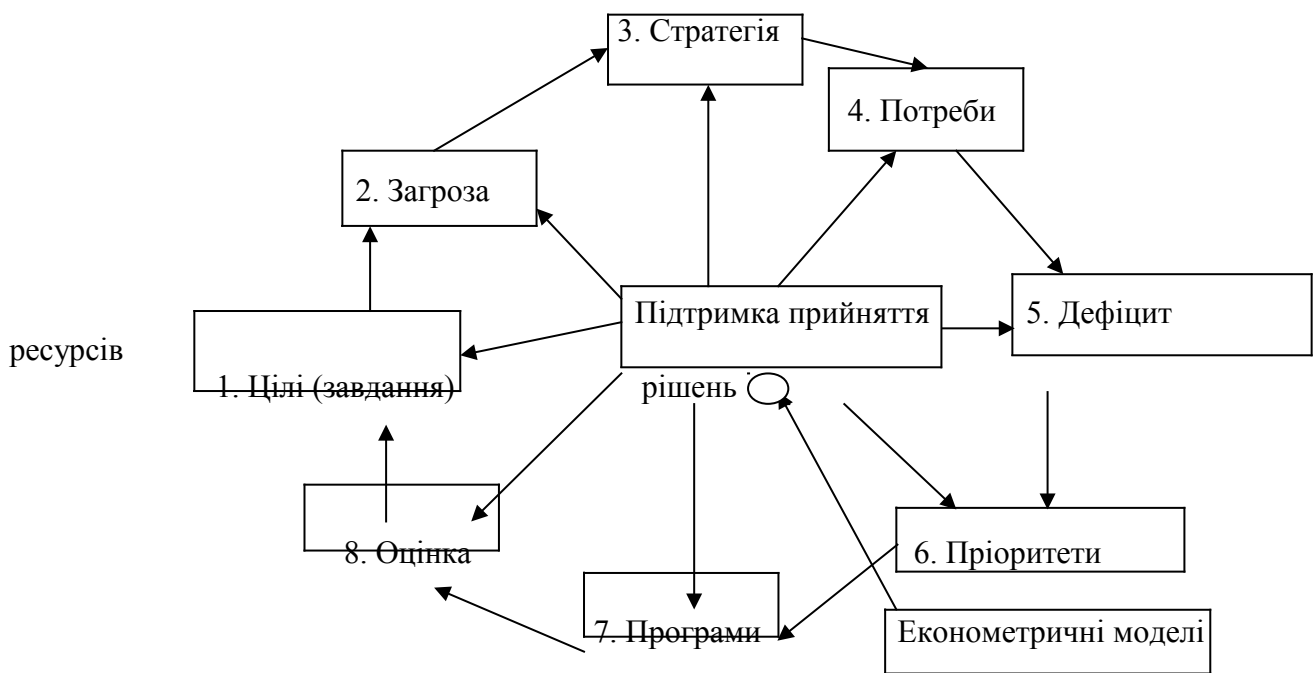


Рисунок 1.1 – Процес планування розвитку економічної системи

Це забезпечує можливість розробки стратегії протистояння і усунення зазначених загроз. На певному спектрі загроз відповідно до розробленої стратегії визначаються потреби в ресурсах, необхідних для вирішення поставлених завдань. Необхідні ресурси порівнюються з наявними, в результаті чого визначається дефіцит ресурсів. Останні (якщо вони мають місце) ранжуються за ступенем важливості й необхідності (визначення пріоритетності) в інтересах забезпечення програм розвитку досліджуваної економічної системи. Нарешті, ці програми оцінюються відносно первісних цілей, що дає змогу визначити реалістичність і необхідність продовження здійснення, коригування або їх закриття.

Єдиним засобом вирішення всього комплексу перелічених завдань планування розвитку економічної системи й управління нею у процесі її функціонування може бути тільки створення відповідної системи підтримки прийняття управлінських рішень (СППУР). "Серцем" такої системи є економіко-математичні (економетричні) моделі, які включаються в контури управління як безпосередньо економічної системи, так і, передусім, надсистеми (наприклад, державних органів управління). Оскільки саме визначення і вчасне коригування цілей функціонування системи, безсумнівно, є найбільш пріоритетними завданнями, використання СППУР передбачає необхідність перекладення всього процесу управління економічною системою на суттєво вищий рівень – системної організації процесів прийняття управлінських рішень.

1.2 Постановка завдань та цілей дослідження економетрії.

Будова економетрії

Як впливає з вищенаведеного визначення предмету економетрії, її про-

блем і завдань, при побудові економетричних моделей можна виділити такі етапи:

1. Постановка конкретного економічного завдання у термінах, якими описуються процеси, що відбуваються у економічній системі, з відображенням існуючих тенденцій їх змін, їх якісного економічного аналізу. Визначення мети дослідження, досягнення якої потребує створення моделей (з урахуванням обмежень та припущень). Визначення складності моделі. Вибір суттєвих параметрів. Визначення середовища, в якому функціонує досліджувана економічна система і модель, яка повинна її імітувати.
2. Формалізація задачі – побудова математичної моделі економічної системи і процесів, що в ній відбуваються, з урахуванням усіх обмежень (якщо дозволяє обчислювальне середовище, оскільки розмірність задачі може виявитися дуже великою при урахуванні великої кількості обмежень на параметри).
3. Перевірка і коректування моделі та з'ясування ступеня адекватності її реальним процесам та явищам, що досягається: отриманням множини спостережень над змінними, які входять до моделі, у відповідні моменти часу; вибором методу оцінювання невідомих параметрів з урахуванням специфіки даних спостережень.
4. Реалізація процедури оцінювання, яка забезпечує найкраще наближення модельних значень змінних значенням, які спостерігалися в реальній економічній системі (реальному процесі) – адекватність моделі реальному економічному процесу. Інтерпретація моделі з точки зору закладеної в неї апріорі динаміки старіння та загибелі. Виявлення додаткових напрямів використання моделі в інших сферах економіки.

Використання економетричної моделі, як невід'ємної частини системи підтримки рішень при управлінні досліджуваною економічною системою, у тому числі при: проведенні аналізу впливу зовнішнього середовища на модельовану систему; удосконаленні її структури; прогнозуванні її функціонування на короткострокові та довгострокові проміжки часу; здійсненні організації оперативного контролю поведінки та діагностування стану основних економічних показників, результати яких зменшують рівень невизначеності інформації про досліджуваний об'єкт; створенні дослідного "полігону" для перевірки та випробування розроблених моделей економічних об'єктів і процесів, що в них відбуваються; виборі методів і засобів прийняття управлінських рішень, аналізу можливих наслідків їх виконання; розробці стратегій розвитку перспективних досліджень у різних сферах економіки на основі прогнозованих тенденцій її розвитку і відповідних економічних експериментів.

На рис. 1.2 подано будову економетрії, методи її дослідження та, в кінцевому рахунку, використання отриманих результатів для дослідження реальних економічних систем різної складності та організації, побудови "прогнозного телескопа" та отримання множини науково обґрунтованих рішень щодо управління відповідними системами. Використовуючи рис. 1.2, відповідно до загальних принципів дослідження будь яких складних об'єктів (рис. 1.3), визначимо: мету, завдання, методологію, методи та засоби, які притаманні

економетрії як окремій дисципліні.

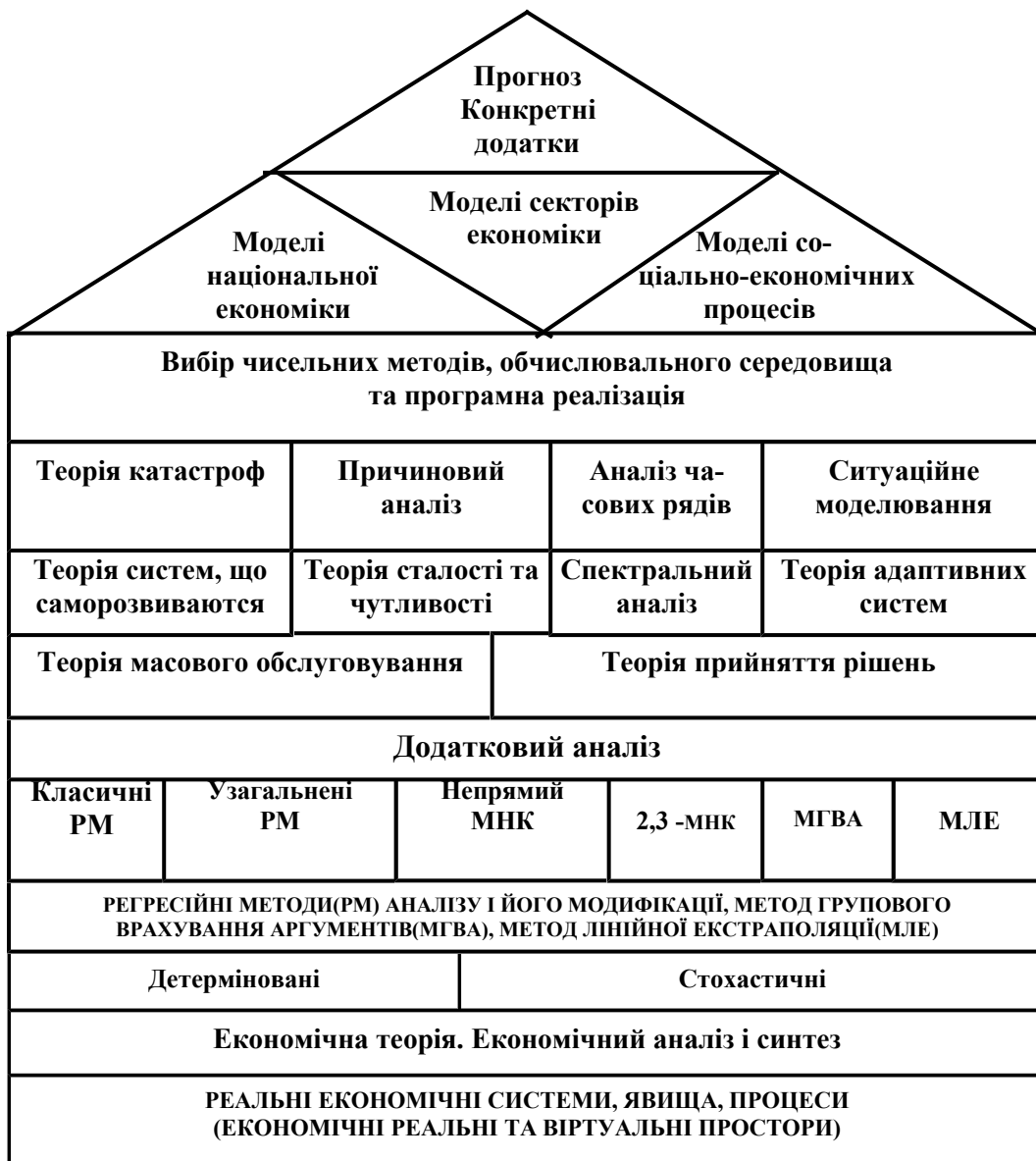


Рисунок – 1.2 Будова економетрії або мету досягнуто, або її потрібно коригувати

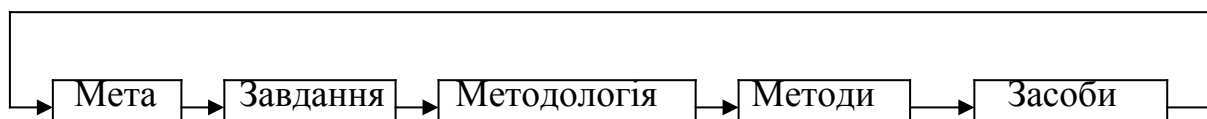


Рисунок 1.3 – Загальні принципи дослідження складних систем

Об'єктом дослідження є економічні системи та простори різного рівня складності та орієнтації: економіка галузей, регіонів, держави і світу в цілому, процеси, які в них відбуваються.

Предметом економетрії – є спектр проблем методів побудови та використання економетричних моделей, аналізу взаємозв'язків між економічними процесами на мікро- і макроекономічному рівні та управління ними, що виникають при розробці науково обґрунтованої стратегії прогнозування, перспективного програмно-цільового планування і оперативного управління економікою на рі-

зних рівнях її ієрархії та організації.

Метою дослідження є аналіз (у квантифікованому аспекті) реальних економічних систем та процесів, що в ній відбуваються, побудова і дослідження відповідних економетричних моделей, їх використання в конкретних економічних системах і просторах при системній організації управління ними.

До категорій та економічних величин, з якими працює економетрія, належать: валовий внутрішній продукт, труд, капітал, вартість, ціна, заробітна плата, прибутки населення, ресурс, витрати на виробництво продукту, чистий прибуток, норма прибутку, процент, кредит, грошовий обіг, товарооборот та ін.

Завдання економетрії можна поділити на базові, які пов'язані з головними функціями економічної системи, і нові завдання, які виникають чи можуть виникнути в неї у зв'язку з необхідністю прийняття ефективних, науково обґрунтованих управлінських рішень з урахуванням економічних, соціальних і політичних аспектів в умовах її трансформації.

Базовим завданням економетрії на рівні макроекономіки є дослідження реальних процесів та явищ (з оцінкою параметрів і перевіркою на значущість відповідної їм економетричної моделі): у виробництві, розподілі, перерозподілі та кінцевому використанні валового внутрішнього продукту, у якому беруть участь усі галузі фінансово-кредитної сфери – державний бюджет, податкова політика, страхування, кредит, ощадна справа; виявлення на цій основі пропорцій і закономірностей у розподільчих фінансово-кредитних відношеннях у народному господарстві; формування податкової політики тощо. Від діяльності всіх перелічених галузей, їх системної узгодженості залежить ефективність розподільчих відношень, збалансованість доходів і витрат у народному господарстві, забезпечення процесів відтворення грошових ресурсів, фінансової захищеності державного, колективного і особистого майна від можливих інфляційних та інших лих. На мікрорівні завдання, які постають перед економетрією, включають питання, що пов'язані з науково обґрунтованим прийняттям управлінських рішень на підприємствах різного рівня складності та власності.

Одним з важливих завдань економетрії є також виявлення і прогнозування статистичних закономірностей, що притаманні: грошовому обігу; розподілу і використанню ВВП; складу і динаміці доходів і витрат державного бюджету; фінансовим ресурсам; інвестиціям та їх цільовому використанню тощо. У цей час у статистиці виявляється недосконалість методології обчислення багатьох економічних показників, наприклад, розподілу доходів комерційних структур, податкових доходів, характеристик ефективності податкової і процентної політики, оцінки рівня інфляції, виміру фінансової стійкості тощо. Недосконалість методології обчислення економічних показників, відмінності її від прийнятої у світовій практиці, призвели до перекручування фактів, породили сумніви в їх дієвості і стали однією з важливих причин виникнення помилок при вирішенні завдань у різних сферах економіки, що негативно вплинуло на достовірність прогнозованих соціально-економічних оцінок. Тому однією з головних проблем сучасного розвитку економетрії є також подолання цих недоліків. Як? За допомогою, наприклад, виконання діагностичних процедур.

Крім того, економетрія повинна (за допомогою відповідних моделей, адаптації до реального середовища їх функціонування) виконувати: контрольні

функції, реалізація результатів яких дасть змогу забезпечити відновлюваність економічних процесів за умов, наприклад, залучення необхідних грошових засобів і фондів (такі завдання включають контроль за виконанням держбюджету, касових і кредитних планів банківськими установами, за акумуляцією тимчасово вільних грошових коштів підприємств, установ, населення та ін.); дати характеристику виконання планів кредитними і фінансовими установами будь-яких форм власності; оцінити невикористані фінансові, кредитні та інші ресурси; встановити причини і наслідки існуючого економічного становища тощо.

Доповнюючи традиційні завдання, що вирішуються за допомогою економетрії, вона може допомогти у визначенні повноти й об'єктивності кількісних економічних даних, кількісній оцінці цілого спектра допоміжних характеристик економічної сфери (наприклад, виявленні джерел "радіоактивного випромінювання" грошей, спрямованих на руйнування економічного простору, в якому вони знаходяться в обігу).

Можливість вирішення традиційних і нових завдань економетрії базується на методологічних принципах, які включають: системно-процесуальне вивчення стану і розвитку економіки держави в цілому, окремих її регіонів, галузей, підприємств; визначення характеру змін в їх діяльності; виявлення і оцінку ступеня впливу на функціонування різних сфер економіки факторів, обумовлених нестабільністю динаміки протікання процесів в економіці та різним рівнем часу їх життя (включаючи і наслідки прийнятих управлінських рішень і час їх прийняття).

Слід особливо відзначити вихідну принципову орієнтацію пропонованих методологічних принципів економетрії на їх прикладне використання у складі системного комплексу інструментальних обчислювальних засобів. Саме за таких умов забезпечується: найбільш ефективно функціонування економічних структур з адаптацією під рівень вимог, можливостей, компетенції, прерогатив і відповідальності осіб, які приймають управлінські рішення; можливість отримання на кожному рівні ієрархії управлінських структур альтернативних варіантів науково обґрунтованих рішень щодо управління економічними системами; можливість прогнозувати кількісний і якісний ефект від їх реалізації з урахуванням конкретних вимог і об'єктивних умов управління; вибір найбільш ефективних управлінських рішень тощо.

Методи економетрії орієнтовані на визначення і обчислення значень основних мікро- та макроекономічних показників, їх класифікацію за ступенем важливості тощо. Вони виступають як додаткова складова частина до загального комплексу традиційних методів і засобів моделювання структури і поведінки економічних систем, який включає власне статистичні методи: однокроковий метод найменших квадратів (1-МНК), дво- і трискроковий метод найменших квадратів (2-МНК, 3-МНК); метод максимальної правдоподібності; методи, побудовані на визначенні різних статистичних оцінок - математичного сподівання, дисперсії, автокореляції, мультиколінеарності, довірчих інтервалів тощо для оцінювання параметрів моделі, їх верифікації, перевірки відповідних гіпотез тощо.

Крім того, методи економетрії включають в себе: спеціальні методи аналізу і моделювання власних часів протікання процесів і явищ в економічних

просторах і системах; методи мікро- та макро-економіки; лінійної екстраполяції (МЛЕ); причинного аналізу; групового врахування параметрів (МГВА); моделі, спектрального аналізу, адаптивного та ітераційного моделювання і управління; отримання класів альтернативних рішень (зокрема, Парето-аналіз); організаційні методи прийняття рішень; методи каузальної алгебри (причинного аналізу), загальної теорії систем; спеціальні методи прийняття рішень; теорії адаптивних систем, актуарної математики тощо. Для опису економічних систем використовують мову теорії ресурсних класів систем, і каузальної алгебри, а для дослідження поведінки таких систем – методи, які покладені в основу системного, процесного і економетричного підходу.

Моделі економетрії повинні бути адаптивними, такими, що самоорганізуються залежно від умов і обмежень, які обумовлені внутрішньою структурою економічної системи та її навколишнім середовищем, надавати змогу виявляти закономірності протікання динамічних процесів, що в них відбуваються.

Інтердепедентні (одночасні, взаємопов'язані, симультаивні) моделі економетрії в комплексі з підсистемою підготовки і підтримки процесів прийняття рішень забезпечують формування класів альтернативних рішень, отриманих з використанням підсистем моделювання власних часів "життя" і динаміки протікання процесів у різних сферах економіки з урахуванням різноманітних комбінацій параметрів навколишнього середовища.

Підсумовуючи викладене можна зазначити, що практичне значення економетрії визначається такими аспектами:

- 1) вирішенням задач системної організації та управління економічними об'єктами і системами різної складності;
- 2) проведенням досліджень, що пов'язані з моделюванням штатних і екстремальних ситуацій в економіці на всіх рівнях ієрархії її управління, оцінкою стратегічних і оперативних рішень, які приймаються;
- 3) подальшим розвитком теорії позитивної та нормативної економіки;
- 4) створенням системи навчання і тренінгу при підготовці осіб, які приймають управлінські рішення на різних рівнях ієрархії економіки;
- 5) подальшим розвитком системних досліджень у галузі економічного аналізу; управління економікою тощо.

Вирішення перелічених задач забезпечується реалізацією економетричних методів і засобів на відповідних програмно-математичних продуктах, які дають можливість проводити відновлювальні модельні досліди різного ступеня складності для розв'язання задач підготовки прийняття управлінських рішень в економічних системах різної складності. Для економіки в цілому ці продукти мають бути включені як підсистема в Державну економіко-математичну модель країни, створення якої є найбільш актуальним і нагальним заходом для удосконалення процесів управління економікою держави.

На закінчення цієї теми розглянемо питання застосування математичного апарату як засобу дослідження економічних систем та процесів, що в них відбуваються.

Як відомо будь-яка економічна модель включає опис математичними

методами реально діючої економічної системи, процесів та явищ, що в ній відбуваються. Моделі економіки будуються для: теоретичних цілей економічного аналізу; визначення прогностичних оцінок функціонування економічної системи; оперативного і стратегічного управління процесами, що в ній відбуваються, плануванням її стійкого розвитку тощо. На змістовному рівні моделі економіки об'єднують такі основні процеси, як виробництво, споживання, планування, управління, фінанси та ін.

Однак часто в економічних моделях, майже завжди, наголос робиться на будь-який один процес (наприклад, фінанси), тоді як усі інші подаються у спрощеному вигляді. Залежно від того, якому економічному процесу приділяється увага при побудові математичної моделі, використовують той чи інший математичний апарат.

Так, наприклад, в моделях планування в основному використовуються системи алгебраїчних (лінійних) рівнянь та нерівностей, оскільки основним завданням планування є балансова ув'язка виробництва та споживання, різних складових частин, що математично можна виразити у вигляді системи рівнянь і (або) нерівностей.

Моделі оптимального планування математично являють собою екстремальні задачі з обмеженнями. В основному, це задачі лінійного програмування, їх розширення або узагальнення. Наприклад, при визначенні інтенсивності виробничих можливостей таким чином, щоб були виконані планові завдання, не перевитрачені ресурси, а деяка виділена складова частина була вироблена у максимальній кількості, приходимо до задачі лінійного програмування.

Економічні моделі управління базуються на різного роду екстремальних задачах, зокрема, задачах оптимального управління в розумінні Понтрягіна.

Моделі зростання породжують особливого роду екстремальні задачі. Вони будуються з метою вивчення максимально можливих темпів зростання економічної системи за тих, або інших умов, зокрема, при якомога великому інтервалі часу. Найбільш відомою моделлю зростання є модель розширюваної економіки, що запропонована та вивчена видатним американським математиком Дж. фон Нейманом. Вона задається двома невід'ємними матрицями витрат і випуску та призначена для знаходження максимального технологічного темпу зростання, який система може витримати скільки завгодно довго (мінімаксі задачі).

Модель рівноваги, головним об'єктом моделювання якої є взаємодія протиборствующих економічних сил або факторів, базується на теорії ігор.

У моделях прогнозу використовують апарат кореляційного та регресійного аналізу (наприклад, в економічних моделях, які описують складні макроекономічні процеси), теорію випадкових процесів, теорію масового обслуговування, теорію статистичних рішень тощо.

Дослідженням економічних систем різного рівня складності й орієнтації (крім економетрії), займаються такі самостійні математичні дисципліни, як економічна кібернетика та математична економіка. Багато з перелічених вище моделей є об'єктом дослідження цих дисциплін.

Економічна кібернетика – напрям у кібернетиці, який використовує її

методи і засоби з метою дослідження і організації процесів в економічних системах.

Математична економіка – напрям у теоретичній економіці, який складено на основі використання математичних моделей і методів для вияву різного роду закономірностей і ефектів в економічних системах.

Навіть з цього обмеженого переліку економічних завдань, необхідність застосування різноманітного математичного інструментарію для їх розв'язання не залишає й тіні сумніву. Зрозуміло, що всі поставлені завдання у цьому посібнику розглянути неможливо не тільки через обмежений обсяг, а й тому що багато з них ще потребують свого переосмислення та нового вирішення.

РОЗДІЛ 2

МЕТОДОЛОГІЧНІ ОСНОВИ МОДЕЛЮВАННЯ СКЛАДНИХ ЕКОНОМІЧНИХ СИСТЕМ (ПРОЦЕСІВ)

2.1 Поняття системи. Математичне моделювання

У межах загальної теорії систем як наукового напрямку, який пов'язаний з розробкою сукупності філософських, методологічних, конкретно-наукових і прикладних проблем аналізу та синтезу складних систем будь-якої природи, поняття системи вперше було визначене у 1937 році. (і з 1950 р. посіло своє особливе місце в науці) одним з її засновників, біологом-теоретиком Л. Берталанфі як комплекс елементів, які перебувають у взаємозв'язку, а також А. Холлом і Р. Ф. Фейджін, які визначали її як множину об'єктів разом з відповідними співвідношеннями між ними і між їх атрибутами.

Сам Л. Берталанфі був представником організмичної течії у філософії. Він так описує запропоновану ним загальну теорію систем: „Оскільки фундаментальна ознака живого є організація, традиційні засоби дослідження окремих частин і процесів не можуть дати повного опису живих явищ. Такі дослідження не включають інформацію про координацію частин і процесів. Тому головним завданням біології повинно бути відкриття законів, які діють у біологічних системах (на всіх рівнях організації). Спроби виявити основи теоретичної біології вказують на фундаментальні зміни в уявленні світу”. Подібний підхід, коли він служить методологічною базою дослідження, може бути названий „органічною біологією”, а коли він застосовується при концептуальному поясненні життєвих явищ – „системною теорією організмів”.

Відомий американський вчений М. Месарович, один із засновників теорії, сформулював основні вимоги, яким повинна задовольняти ця теорія: 1) бути настільки загальною, щоб охоплювати вже існуючі теорії, що стосуються теорії систем; 2) мати науковий характер, її терміни і визначення повинні бути математично однозначні і відповідати її призначенню – вивчати абстрактні моделі відповідних систем; 3) мати наукове підґрунтя настільки фундаментальне, щоб її висновки мали безсумнівну практичну цінність при вивченні конкретних систем, що зустрічаються в житті .

В усіх визначеннях (як засновників загальної теорії систем, так і авторів подальших розробок) завжди зазначалося, що система являє собою цілісний комплекс взаємопов'язаних елементів, має визначену структуру та взаємодіє з деяким середовищем.

Проте невизначеність розвитку системного підходу полягає в тому, що з приводу визначення терміна "система" існує багато розбіжностей. Ці розбіжності породили різноманітні варіанти загальної теорії систем, які часом якісно різняться між собою (ситуацію можна порівняти з тією, яка склалася через нечітке визначення терміна "множина").

Наприклад, терміном **система** пропонують визначати [70]:

1) множину елементів, які перебувають у відношеннях і зв'язках між

собою (Философский словарь; под ред. М. Розенталя. – 3-е изд., 1975);

2) порядок, теорія, устрій, систематизація-класифікація (Александров З. Е. Словарь синонимов русского языка, 1975);

3) сукупність елементів, взаємопов'язаних між собою (Горский Д. П и др. Краткий словарь по логике, 1991);

4) об'єктивно існуючий комплекс процесів і явищ, а також інструмент, засіб розгляду таких процесів і явищ (Лопатников Л. И. Популярный экономико-математический словарь, 1990);

5) ціле, складене з частин, є з'єднання – об'єктивна єдність закономірно пов'язаних один з одним предметів, явищ. У науці і техніці – множина елементів (вузлів, агрегатів, приладів) (Политехнический словарь; под ред. И. И. Артоболевского, 1982);

6) спільність, що складається із взаємозалежних частин, кожна з яких привносить щось конкретне в унікальні характеристики цілого (Мескон М. Х., Альберт М., Хедоури Ф. Основы менеджмента, 1992) тощо.

Таким чином, аналізуючи викладене, можна зробити висновок, що поняття "система" містить такі притаманні їй основні фактори:

– система як річ або матерія, як сукупність речей (логічна точка зору);

– система як множина елементів (теоретико-множинова точка зору), спочатку без зв'язків, а потім зі зв'язками між елементами;

– система як процес (тавтологічна або тотожна точка зору – вона є найбільш визначеною) та ін.

У подальшому, щоб запобігти термінологічній невизначеності, під **системою** у **широкому розумінні** будемо розуміти *процес* (поняття якого буде розкрито далі), а під **системою** у **вузькому розумінні** набір множин:

$$\mathfrak{S} = \langle \{S_i\}_{i \in I}, L, F \rangle,$$

де $\{S_i\}_{i \in I}$, $i \in I$ - множина елементів різної фізичної природи (неподільних стосовно самої системи); L – множина зв'язків між елементами; F – множина функцій (операторів), які виконуються цією множиною елементів.

Як впливає з визначення системи у вузькому розумінні, *виділення елементів системи* є одним з перших кроків при побудові її формалізованого (математичного) опису.

Функцією системи будемо називати алгоритм (правило) отримання результатів, приписаних метою (призначенням) системи.

При цьому природним уявляється визначення **структури системи** \mathfrak{X} як сукупності елементів $\{S_i\}_{i \in I}$, $i \in I$ і зв'язків між ними L , тобто

$\mathfrak{X} = \langle \{S_i\}_{i \in I}, L \rangle$ **а моделі системи** – як математичні співвідношення, які опи-

сують досліджувану систему (процеси або явища). Тобто, **модель системи** – це опис траєкторій її поведінки у деякому фазовому просторі параметрів. Складання математичної моделі системи називається **математичним моделюванням**. Поведінка системи, яка моделюється, може мати детермінований (суворо визначений), імовірний або індетермінований (взагалі невизначений) характер.

Таким чином, математичний опис системи характеризує власне систему, незалежно від будь-яких дій, що впливають на неї, а модель функціонування

дає опис її поведінки в умовах впливу цих дій на систему з боку деяких параметрів внутрішнього та зовнішнього середовища.

Основними методами дослідження систем є методи математичного моделювання, які дають можливість створити опис структури і поведінки реальних (соціальних, економічних, біологічних, технічних та ін.) систем. До цих методів належать такі, що розроблені у теоріях: ймовірностей і математичної статистики, графів, математичного аналізу, статистичних рішень, причинного і багатофакторного аналізу, часових рядів тощо.

Відмінною особливістю математичного моделювання є те, що воно дає можливість вивчення, прогнозування і оптимізації управління реальними складними системами, для яких фізичний (натурний) експеримент є утрудненим або економічно не вигідним, а іноді навіть небезпечним, оскільки може призвести до великих матеріальних, соціальних або інших втрат, а інколи й до загибелі системи (наприклад, коли діюча економічна політика не виправдала себе).

Крім того, слід зазначити, що для проведення натурального експерименту часто необхідно мати великий проміжок часу (наприклад, у сільсько-господарському виробництві), а це суттєво подовжує строки виконання науково-дослідних робіт і не дає, таким чином, можливості прийняття коректного вирішення проблеми у встановлені строки, що часто є вирішальним з точки зору старіння і загибелі самої досліджуваної системи.

Вирішення будь-яких науково-дослідних задач (проблем) звичайно обумовлює таку послідовність кроків: 1) визначення цілі дослідження; 2) спостереження та експеримент; 3) теоретичні дослідження; 4) власне саму організацію системи (процесу).

Як зазначалося у [60], специфіка теоретичної дослідницької роботи, своєю чергою, зумовлює такі **етапи моделювання складних систем**:

1) **постановка конкретної задачі (проблеми)** з урахуванням цілей дослідження у термінах опису системи (процесів), які їй (їм) притаманні, вибір методології дослідження;

2) **формалізація задачі (проблеми)** – побудова математичної моделі задачі (системи, процесу);

3) **перевірка та коригування моделі**, визначення ступеня адекватності моделі реальному об'єкту (процесу);

4) **знаходження оптимального вирішення** задачі (проблеми) на основі уточненої моделі за допомогою того чи іншого чисельного методу оптимізації, побудова алгоритму вирішення цієї задачі (проблеми).

Після цього складається програма реалізації зазначеного алгоритму на ЕОМ і виконується обчислювальний експеримент, в результаті якого дослідник отримує сукупність даних для опису поведінки об'єкта;

5) **аналіз отриманих даних**, надання їм необхідної змістовної форми та **практичне їх використання**.

Як впливає з наведених етапів, робота, яка пов'язана з дослідженням моделей складних систем, зводиться до різних за змістом (через різну специфіку) постановкам задач, до вибору (для кожної конкретної системи) методології,

визначених математичних методів і засобів їх реалізації. Тому володіння *високою математичною культурою* є необхідним атрибутом для кожного дослідника складних систем.

Ми досить часто користуємося терміном "складна система". Але він здебільшого має чисто суб'єктивний характер, бо досі не існує достатньо загально-го визначення такої системи. Тому, залежно від типу об'єкта дослідження, використовують те або інше поняття складної системи, яке звичайно справедливе відносно конкретного об'єкта, але не завжди справедливе для іншого. У зв'язку з цим наведемо тільки **характерні ознаки складних систем** [60], які притаманні всім системам цього класу:

- 1) *велика кількість* взаємопов'язаних між собою елементів і підсистем;
- 2) *складність функцій, що виконуються системою* у процесі реалізації мети функціонування;
- 3) *багатовимірність системи*, яка обумовлена наявністю великої кількості зв'язків між підсистемами;
- 4) *взаємодія із внутрішнім та зовнішнім середовищем* і функціонування в умовах випадкових факторів;
- 5) *наявність множини критеріїв оцінки* якості функціонування системи та її підсистем;
- 6) *різноманітність структури*, що обумовлена як різноманітністю структур її підсистем, так і різноманітністю структур їх об'єднання;
- 7) *наявністю управління*, яке часто має ієрархічну структуру, а також розгалуженої інформаційної мережі та інтенсивних інформаційних потоків;
- 8) *різноманітність фізичної природи* підсистем, яка обумовлена їх різною фізичною суттю;
- 9) *існування інтегративних ознак*, які притаманні системі в цілому, але не характерні для кожного її елемента зокрема;
- 10) *відсутністю можливості* отримання достовірної інформації про властивості системи у цілому при вивченні її окремих елементів;
- 11) *велика розмірність* і складність моделі системи, що вимагає необхідності використання для її дослідження сучасних математичних методів декомпозиції, макромодельовання, імітаційного модельовання тощо.

Таким чином, **складна система** являє собою множину взаємопов'язаних і взаємодіючих елементів та підсистем різної фізичної природи (гетерогенна структура), які складають неподільне ціле, що забезпечує виконання системою деякої складної функції.

Розглянемо такі поняття, як "системний аналіз" і "системний підхід".

"Системний аналіз" – це дисципліна, яка розвиває методи побудови складних господарських, екологічних, технічних, організаційних структур тощо [59]. Відзначимо, що російськомовний та український термін "системний аналіз" не має точного аналогу в іноземних мовах. На початку 60-х років у США з'явився термін "system analysis" для позначення техніки аналізу складних систем, яка саме тоді виникла. Необхідність її виникнення була зумовлена вимогами розвитку, перш за все, методів дослідження операцій і вивчення засобів подання інформації, що полегшують формулювання дослідником цілей

цих операцій.

Дослідник операцій у зарубіжній літературі звичайно називався "analyst". Для того, щоб відзначити особливість кваліфікації спеціаліста, який займається аналізом і проектуванням складних систем, почали використовувати термін "system analyst". Таким чином, термін "system analysis" належало б перекласти як "аналіз систем", але його було перекладено як "системний аналіз", оскільки з англійської мови на російську обидва ці терміни перекладаються однаково. І вже з російської наукової термінології цей термін потрапив в українську. Російською та українською термін "системний аналіз" має значно більше змістовне навантаження, а саме – ця дисципліна, яка містить у собі не тільки конкретні засоби подання інформації, але й фундаментальні розділи теорії.

"Системний підхід" – це ще більш розпливчате й неточне поняття, яке вперше було запропоноване радянськими вченими.

У розвитку науки завжди чітко простежувалися дві тенденції – аналіз і синтез. Ми завжди бачимо прагнення до аналізування – вивчення конкретних фактів, проникнення у глибину факту, що вивчається, розкриття тонкої структури явищ і т. ін. Але завжди існує прагнення створити синтезуючі теорії, які дають можливість об'єднувати різноманітні факти, побачити перспективи розвитку того або іншого процесу, його зв'язки з іншими явищами, врахувати їх взаємну обумовленість тощо. Впродовж останніх десятиліть роль синтезуючих побудов стала особливо великою. Потреба не тільки вивчати явища, факти, але й встановлювати їх зв'язок з іншими фактами призвела до появи спеціального терміна "системний підхід". *Системний підхід – це є методологія дослідження складних систем.*

Введення "**системних законів**" для опису та дослідження складних економічних систем повинно стати цементуючим матеріалом між теоретичною економічною наукою та реальним життям і сприяти, таким чином, піднесенню якості вивчення та пізнання економічних явищ і процесів на більш високому формалізованому рівні. Іншими словами, **системний підхід** в економіці повинен **сприяти "законотворчому" процесу** в розвитку економічної науки на якісно новому формалізованому рівні. Крім того, "...загальна теорія систем, виводить наукове передбачення на більш високий рівень, ніж це дозволяє зробити будь-яка з окремих наукових дисциплін, оскільки вона базується у такому разі на системі законів, та пізнається лише завдяки проведенню міждисциплінарних досліджень".

І на завершення обговорення питань, пов'язаних зі складними системами та застосуванням для їх дослідження системного аналізу з позицій системного підходу, вкажемо на ті **основні види невизначеностей**, з якими стикається кожен дослідник, коли використовує такі теорії [70]:

- *теорія великих систем*: невизначеність меж таких систем, яка не знімається навіть положеннями теорії "розмитих" (нечітких) множин;
- *теорія організації*: невизначеність переходу від однієї системи до іншої, проблема декомпозиції;
- *наука про поведінку*: невизначеність біхевіористичних часових послідовностей;

- *групова динаміка*: невизначеність малих груп та їх меж;
- *методи сіткового планування*: невизначеність в умовах змін середовища або послідовності дій (невизначеність мереж);
- *теорія масового обслуговування*: невизначеність у врахуванні нерегулярних масивів;
- *теорія автоматичного керування*: невизначеність в урахуванні виробничого процесу, проблема багатовимірності;
- *теорія інформації*: невизначеність оцінки цінності інформації, меж застосування формул Шеннона;
- *математична статистика*: проблеми взаємодії факторів;
- *теорія ігор*: наявність сингулярних формул (вироджених);
- *теорія черг*: наявність сингулярних формул;
- *теорія статистичних рішень*: складність застосування в умовах, що змінюються;
- *теорія алгоритмів*: невизначеність мети алгоритмів;
- *теорія автоматів*: невизначеність малоймовірних ситуацій при роботі автоматів, абстракція без меж;
- *системотехніка*: невизначеність інструментальних висновків і моментів переходу від проектування до експлуатації [43] тощо.

Всі ці невизначеності призводять до складності застосування теорії на практиці. Тому у подальшому при розгляді моделей реальних економічних систем (процесів) ми завжди будемо накладати обмеження на них, зводячи до мінімуму ті невизначеності, які характерні для реально існуючих систем.

Що стосується дослідження **складних економічних систем**, то слід враховувати ту обставину, що економічна система повинна виступати як єдність продуктивних сил та виробничих відносин. Розвиток і функціонування економіки відбувається відповідно до її законів та принципів, що з них випливають.

Реальним економічним системам притаманні всі перелічені вище ознаки складних систем, які (у межах загальної класифікації теорії складних систем) зумовлюють такі **загальносистемні властивості складних економічних систем**:

1. **Складність структури економічних систем**. Найбільш характерною їх рисою є ієрархічність і антропогенність (наявність людей), але не у чистому вигляді, а у вигляді переплетення деяких ієрархічних систем. Тут характерне подвійне підпорядкування, перетин контурів управління, сукупність вертикальних, горизонтальних і перехресних зв'язків, що виражають відношення управління або взаємодії. Ці складні структури породжують адміністративно-виробничі, територіально-регіональні, виробничо-галузеві, міжгалузеві, соціальні та інші відношення. Поділ системи на підсистеми, тобто виділення її елементів, має, як правило, евристичний характер і неформалізований.

2. **Цілісність системи**. Порушення цілісності системи, тобто виділення з неї окремих фрагментів, порушує емерджентні властивості системи.

3. **Складність інформаційних процесів**. Передбачає необхідність

комбінаторних побудов при описі системи для відображення різноманітних економічних, соціальних, організаційних, технічних, наукових та інших аспектів функціонування систем.

4. **Множинність цілей та багатовимірність критеріїв ефективності системи.** Цілі та критерії підсистем можуть не збігатися, конфліктувати між собою або навіть суперечити загальним критеріям і цілям. Це обумовлено багатофункціональністю економічних систем (облік, контроль, регулювання, аналіз, коротко- та довгострокове планування – особливо у перехідній економіці, яка функціонує за умов невизначеності, конфлікту і жорстких ресурсних обмежень).

5. **Динамічність процесів, що відбуваються у системах.** Їх динаміка має відмінний характер; швидкодія деяких процесів системи змінюється у широких межах – від надвеликого до умовно-постійних змін (наприклад, зміни інфляційних процесів у 1993-1994 рр.).

6. **Багатообразність структури та природи економічних систем.** Обумовлена багатообразністю форм економічних проявів.

Крім того, можна виділити такі особливі процесні властивості, які характерні саме для складних економічних систем:

1. **Велику питому вагу суб'єктивних факторів,** що впливають на функціонування економічних систем. В економічній системі люди виконують роль регуляторів; людина одночасно є елементом продуктивних сил і виробничих відносин. Людині притаманні різноманітні інтереси. Тому в економічних системах існують сильні зв'язки між виробничими та соціальними процесами, і ці зв'язки та їх характер визначають виробничі відносини. У реальних економічних системах з різним функціональним призначенням і з різним місцем в ієрархії керування працює велика кількість людей (антропогенність економічної системи).

2. **Складність інформаційних процесів,** яка значною мірою збільшується завдяки участі людей у функціонуванні системи. Людина реалізує, як правило, неформалізовані алгоритми, її діяльність не завжди прогнозується. Крім того, інформація, яка іде від людини, також не вкладається у суворі схеми структури даних. Зворотні зв'язки, що діють у системах керування, також проявляються незвично, оскільки діяльність людей не завжди можна описати математичними залежностями.

3. **Тісні зв'язки між суспільною та економічною системами.** У діяльності та розвитку економічних систем велике значення має суспільний досвід людей, урахування якого потребує формулювання відповідних специфічних проблем, які практично не виникають при дослідженнях технічних та деяких інших систем. При формулюванні цих проблем необхідно, зокрема, враховувати неформальні елементи свідомої погодженості окремих осіб і груп людей щодо напрямків розвитку системи.

Все це робить економічні системи вкрай складними для опису та дослідження. За великих масштабів економічних систем, їх слабкоструктурованості, наявності в них неформалізованих процесів і локальних непередбачених управлінських рішень неможливо обмежитися при їх описі умовами дете-

рмінованості. Тому для опису цих систем необхідно використовувати стохастичні моделі, особливо при розв'язанні задач коротко- та довгострокового планування та прогнозування, оперативного та стратегічного управління тощо.

Досвід досліджень складних систем (як систем, що підлягають управлінню) незалежно від їх складності, характеру внутрішніх взаємозв'язків, галузевої орієнтації тощо дає змогу сформулювати **три основних принципи дослідження, створення і використання складних систем** [43]:

- 1) принцип фізичності;
- 2) принцип модельованості;
- 3) принцип цілеспрямованості.

Принцип фізичності: будь-якій системі властиві фізичні закономірності (можливо унікальні), що визначають специфіку її внутрішніх причиново-наслідкових зв'язків, які обумовлюють характер існування та функціонування системи під впливом зовнішнього середовища. Ніяких інших закономірностей для пояснення дії систем будь-якої природи (у тому числі соціально-економічної) не потрібно.

Принцип фізичності включає два постулати.

Постулат цілісності: складна система повинна розглядатися як єдине ціле.

Постулат автономності: різні класи фізичних явищ можуть бути поставлені у відповідність до різних груп перетворень, кожна група породжує свою геометрію.

Принцип модельованості: складна система відображається *кінцевою множиною* моделей, кожна з яких відображає певну грань суттєвості системи. Цей важливий принцип дає можливість досліджувати певну властивість або групу властивостей складної системи за допомогою однієї або декількох спрощених (вужкоорієнтованих) моделей.

Доведення існування та стабільності орієнтованих (як завгодно вузько) моделей спирається на постулат додатковості, а оцінка меж цієї стабільності – на постулат невизначеності.

Постулат додатковості: складні системи, що перебувають у різних середовищах (ситуаціях), можуть проявляти різні системні властивості, у тому числі альтернативні (тобто такі, що взаємно виключають одна одну).

Постулат невизначеності: підвищення точності визначення (вимірювання) якої-небудь кількісної властивості складної системи понад деяку межу тягне за собою зниження можливої точності визначення (виміру) іншої властивості. Іншими словами, одночасно виміряти значення двох (або більше) параметрів із точністю, що перевищує певний рівень, неможливо, тобто існує область невизначеності, у межах якої властивості можуть бути описані тільки ймовірнісними характеристиками.

Принцип цілеспрямованості: сукупність функцій складної системи не послаблює процеси, які стимулюють певні стани системи. Цей принцип визначає особливе місце та роль складних систем. Цілеспрямованість – функціональна тенденція, яка спрямована на досягнення системою деякого стану або на посилення, зменшення, або збереження деякого процесу. При цьому система виявляється

здатною протистояти зовнішньому впливу, а також використовувати середовище та випадкові події.

Наслідком принципу цілеспрямованості є *постулат вибору*: складні системи мають здатність до вибору поведінки і, отже, однозначно передбачити засіб дії та екстраполювати їх стан неможливо за будь-якого апріорного знання властивостей системи та ситуації. Складна система будує свою поведінку у суттєвому (хоча і неоднозначному) зв'язку із ситуацією. Отже, на цю поведінку можна впливати. Можна чекати, що ступінь неоднозначності залежить від ситуації, тобто зовнішніх зв'язків. Більш того, не виключено, що за певних умов неоднозначність буде близькою до нуля.

Всі складні системи, у тому числі й економічні, мають такі *загальні властивості* [43]:

1) унікальність: кожна складна система не має повних аналогів поведінки (у будь-якому випадку аналоги настільки виняткові, що їх наявністю практично можна знехтувати);

2) слабкопередбачливість: ніяке, скільки завгодно докладне знання морфології і функцій елементів (підсистем) не дає можливості визначити функції об'єкта, що досліджується; ніяке, скільки завгодно докладне і точне знання поведінки об'єкта на інтервалі $[-T, 0]$ не дає можливості точно передбачити його поведінку на інтервалі $[0, T]$;

3) негентропійність або цілеспрямованість: система спроможна (у визначених межах) керувати своєю енергією (зменшувати її, зберігати, збільшувати) при випадковому і несприятливому впливові середовища або (і) здійснювати поведінку, яка спрямована на досягнення визначеної мети.

Негентропія (міра ймовірності перебування в даному стані) визначає "прагнення" системи до виконання основного процесу, спроможність усувати при цьому наслідки зовнішніх і внутрішніх випадкових впливів.

Цілеспрямованість ("прагнення" досягнення мети) висловлює саме цю тенденцію до зменшення, збереження або підсилення процесу, який веде до мети. Тому поняття "негентропійність" і "цілеспрямованість" – споріднені.

Отже, складні системи мають властивості унікальності, слабкопередбачуваності та негентропійності (цілеспрямованості). Властивість унікальності є зовнішньою по відношенню до системи і впливає на відношення до неї дослідника (особи, яка приймає рішення). Властивість негентропійності (цілеспрямованості) є внутрішньою, такою, що важко розпізнається, і не завжди є доступною до зрозуміння дослідника, особливо на відносно короткому (порівняно з часом існування системи) інтервалі часу.

Для дослідника складної системи на перший план виступає властивість слабкої передбачуваності поведінки, яка, за суттю, і є практичною ознакою складної системи, інші ж властивості можуть бути виявлені тільки у процесі дослідження, яке їх виявить.

Розглянемо питання, які пов'язані з виділенням **рівнів абстракції опису моделей складних систем (у тому числі й економічних)**.

1. *Лінгвістичний або символічний рівень.* Це вищий рівень абстракції, який дає змогу відобразити найбільш загальні властивості системи, а також ви-

ди її зв'язків із зовнішнім середовищем. Система розглядається як єдине ціле або у вигляді обмеженої кількості підсистем. Структура може бути задана у загальному вигляді шляхом визначення наявності або відсутності взаємозв'язків між підсистемами та зовнішнім середовищем. При описі моделей можна використовувати мову алгебри, логіки та теорії предикатів.

2. *Теоретико-множинний рівень*. Дає змогу провести деяку деталізацію моделі при більш глибокій її декомпозиції. Система складається із множини елементів, між якими існують відношення різних класів. Моделі цього рівня дають змогу досліджувати функціональні та структурні властивості великих економічних систем.

Однією із специфічних і зручних мов теоретико-множинного рівня для опису економічних моделей, зокрема, організаційних систем керування великими економічними системами, є алгебра каузаций (алгебра, яка вивчає причиново-наслідкові співвідношення).

3. *Абстрактно-алгебраїчний рівень*. Базується на застосуванні алгебраїчного апарату (теорії груп і підгруп), сімей, відношень (унарних, бінарних, тернарних і т.ін.), алгебраїчних моделей та методів (теорія поліномів, позиномів, векторна алгебра тощо).

4. *Топологічний рівень*. Дозволяє відобразити розміщення елементів системи у фізичному просторі. Крім звичайного геометричного простору, можна розглядати функціональні простори поведінки і т. ін.

5. *Логіко-математичний рівень*. Дозволяє залучити спеціальні моделі для дослідження конкретних систем керування. На цьому рівні застосовуються, наприклад, економетричні й імітаційні моделі, а також описуються як неперервні, так і дискретні системи.

6. *Теоретико-інформаційний рівень*. На цьому рівні описуються не тільки комунікативні відношення та задаються кількісні характеристики потоків інформації, але й вирішуються питання оброблення інформації, її кодування, раціонального використання, здійснюється формування логічних інформаційних описів (баз даних) та ін.

7. *Динамічний рівень*. Будь-яка реальна система не залишається нерухомою, в ній постійно циркулюють та обмінюються із зовнішнім середовищем матеріальні, інформаційні та енергетичні потоки. Вони просуваються у часі та просторі. Цей рух відображається різними моделями, для чого використовуються системи диференціальних та інтегральних рівнянь тощо. Для дискретних систем вводиться поняття стану системи й описується траєкторія її поведінки у вигляді переходів або послідовностей станів. Застосовується теорія автоматів, алгоритмічні моделі для опису динаміки системи та ін.

8. *Евристичний рівень*. Зумовлює використання умоглядних та вербальних (словесних) моделей людських здібностей та досвіду, які дають можливість різко скоротити кількість варіантів, що розглядаються з метою прийняття рішень. Ці методи і моделі важко формалізувати повною мірою, вони поки що доступні тільки людському інтелекту. Моделі евристичного рівня практично завжди використовуються при управлінні великими економічними об'єктами та невід'ємні від них.

Найважливішими **задачами дослідження складних систем є задачі синтезу**, які полягають у знаходженні структури і визначальних параметрів системи за її властивостями, що задаються, і задачі аналізу, при вирішенні яких (за умов знання структури і параметрів системи) вивчається її поведінка, тобто

досліджуються властивості системи та її характеристики. Задачі синтезу й аналізу взаємозворотні та звичайно вирішуються сумісно, зокрема, задачі синтезу, як більш складні, частіше за все вирішуються з використанням результатів розв'язання задач аналізу.

Вирішення питань організації та взаємодії різних економічних систем в одному і тому ж економічному просторі, прогнозування динаміки їх розвитку у майбутньому, організація взаємодії відмінних одне від одного економічних просторів створює реальні передумови для вирішення безконфліктної (або, принаймні, квазібезконфліктної) взаємодії економічних систем і просторів найрізноманітнішого характеру й орієнтації. Особливо це актуально тепер, коли спостерігаються процеси інтеграції різних економічних просторів.

Концепція системного підходу для аналізу та синтезу економічних систем і просторів, їх взаємодії являє собою той базовий рівень, на якому концептуально і категоріально повинна будуватися вся теорія економічної науки. При цьому треба пам'ятати, що при вивченні та моделюванні економічних систем одним із основних принципів є **принцип причиновості**, як один із основних принципів існування економічних систем, відповідно з яким будь-який процес, що призводить до зміни стану економічної системи, пов'язаний із визначеною сукупністю умов (причин), що породжують цей процес. Таким чином, **системний підхід** в економіці являє собою саме той, принципово відмінний від традиційного підходу, рівень, коли термінам "економічна система", "економічний простір" та ін. надається не тільки інтуїтивно-умоглядний характер, але й по можливості суворо формалізоване поняття.

Безсумнівно, усвідомлення можливості та необхідності застосування системного підходу до вивчення взаємодія економічних систем і просторів та трансформації їх з одного виду в інший є, якщо й не визначальним, то одним із основних. У межах системного підходу вже можна здійснювати постановку та вирішування таких проблем економічних систем, які раніше неможливо було навіть формалізувати. Наприклад, питання системної надійності, стійкості тощо ставляться вже цілком на іншій, суворо формалізованій мові, що, своєю чергою, дає змогу математично обчислювати відповідні показники. Крім того, такий підхід дає можливість здійснювати декомпозицію систем на формалізованому рівні, і розв'язувати проблеми "узгодженості" та оптимізації структури та функцій економічних систем, забезпечення їх адаптивності, "живучості" за умов впливу різних зовнішніх факторів (екстерналій), самоорганізації, саморозвитку тощо.

І нарешті, найбільш плідним наслідком застосування системного підходу може бути можливість постановки і вирішування проблем анамнезу і діагностування структури та поведінки економічного простору, тобто, проблем "**економічної генетики**", що вивчає у часі перехідні процеси в економічних системах і просторах, у тому числі і процеси переходу від одного економічного простору до іншого, від однієї економічної системи до іншої. Так, вирішування проблеми "розкриття" генетично закладених "батьківських" основ у нову економічну систему, яка породжена старою, може не тільки допомогти у вирішенні такої актуальної, хоча і традиційної задачі, як прогнозування поведінки

системи, але й у розв'язанні проблем їх "старіння" та "загибелі", проблем, які пов'язані із розкриттям невизначеностей і позамежного ризику, що властиві будь-якій економічній системі, та, особливо, трансформованій економіці, нарешті, у вирішенні проблеми соціальної орієнтації суспільства.

З урахуванням можливостей, які мають системний підхід до досліджень в економіці та застосування математичного моделювання, відкриваються широкі перспективи забезпечення вирішення всього комплексу економічних задач і апробації його наслідків на основі проведення економічних експериментів над моделлю, а не над реально існуючою економічною системою.

Під **економічним експериментом** розуміється сукупність вимірювань параметрів досліджуваної економічної системи, результати яких піддаються обліку (або можуть бути математично оцінені) з використанням будь-яких доступних засобів вимірювання [54].

Основною властивістю економічного експерименту є його відтворюваність в умовах різноманітних варіацій параметрів середовища функціонування економічної системи, яка досліджується.

Згідно із вищевикладеним, дослідження моделей складних економічних систем зводиться, зважаючи на їх специфічність, до різних за змістом постановок задач, вибору (для кожної конкретної системи) визначених математичних методів та засобів.

Обґрунтування вибору методів та засобів, які застосовуються для дослідження економічних систем, повинно враховувати слабку їх структурованість (наприклад, на відміну від технічних систем), що обумовлена специфікою підвищеного впливу людського фактора на поведінку цих систем.

2.2 Класифікація систем

Про класифікацію взагалі. **Класифікація** – впорядковуючий розподіл систем (або їх елементів) за загальними (об'єднувальними) і розрізнявальними ознаками, особливим випадком якого виступає їх стандартне групування за класами і під класами, яке дає змогу встановлювати родовидові співвідношення (градації) систем. *Мета класифікації* – згрупувати схожі системи (або їх елементи) для обґрунтування загальних методів дослідження. Класифікація створює наочність, оглядність і перспективу, відображаючи високий рівень знання досліджуваної системи (або її елементів), процесу, явища. Зауважимо, що для багатьох описових наук класифікація є не тільки основним засобом дослідження, але й майже самоціллю.

У пізнанні класифікація здійснюється через побудову систем співвідпорядкованих понять, кожне з яких отримує суворо фіксоване місце відносно вищих і нижчих класифікаційних рубрик (класів, родів, видів і т.д.). При цьому повинні виконуватись певні логічні правила розподілу обсягу понять, обліку співрозмірності елементів класифікаційної системи та інші. Суттєвим є вибір основ класифікації, для чого відбираються найбільш важливі в практичному й теоретичному відношенні ознаки. Результати класифікації відбиваються у вигляді таблиць та схем, у вигляді переліку і т.п.

Прийоми класифікації використовуються при статистичній обробці економічної інформації. **Класифікація – важлива пізнавальна процедура, що дає можливість правильно орієнтуватися в складних ситуаціях, багато-**

елементних взаємодіях, вивчати ієрархічні зв'язки складних систем, їх структурні та системні характеристики, виявляти закономірності в безлічі явищ.

Розглянемо тепер питання, які пов'язані з **класифікацією систем** з точки зору системного підходу. Загальна теорія систем використовує класифікацію систем за різними ознаками. Найбільш типовими є класифікації за [70]:

1) *видами (класами) систем*: матеріальні; енергетичні; інформаційні тощо;

2) *природою систем*: фізичні; біологічні; технічні; математичні; економічні; соціальні тощо;

3) *особливостями структури систем*: унітарні; мультиплікативні; центральні; ієрархічні; змінні; оптимальні; надлишкові; нерівномірні тощо;

4) *за властивостями систем*: ефективність; надійність; спостережуваність; керованість; стійкість; вартість; організованість; емерджентність; цілісність; технологічність; системність; функціональність; інформаційність; комунікабельність; економічність; агрегативність; автономність; відкритість; зовнішність; швидкодія; окупність; дефіцитність; динамічність; циклічність тощо;

5) *поведінкою систем*: примітивні; програмні; рефлексно-адаптивні; евристичні; що розвиваються тощо.

Найбільш загальною класифікаційною ознакою систем по відношенню до зовнішнього середовища можна вважати **розподіл систем на системи відкритого і закритого видів**. Для систем закритого виду дією об'єкта на середовище і навпаки нехтують, а для систем відкритого виду – береться до уваги те, що характеристики середовища (у загальному випадку) залежать від реакцій об'єкта керування і навпаки. Зовнішнє середовище може впливати на об'єкт керування, змінюючи множини його вхідних і вихідних змінних та внутрішніх станів (рис. 2.1), де U – множина можливих вхідних впливів на систему відповідно до її цілей; V – множина вихідних реакцій системи; A – множина внутрішніх станів системи; Z – множина зовнішніх впливів на систему (середовище, у якому функціонує система); F_A – оператор (множина операторів), які переводять систему з одного стану в інший; F_V – оператор (множина операторів), який відображає множину вхідних впливів, станів системи і впливів зовнішнього середовища на множину вихідних реакцій системи. Зазначимо, що такий автоматний підхід часто використовується при імітаційному моделюванні у економіці [36].



Рисунок. 2.1 – Типи систем

Дії об'єкта на середовище і середовища на об'єкт можуть мати як *детермінований*, так і *ймовірнісний* характер.

ВИЗНАЧЕННЯ 2.1. Детермінованими назвемо такі системи, в яких процеси взаємопов'язані таким чином, що простежується ланцюг причин і наслідків, яким притаманна "жорстка" причиновість.

ВИЗНАЧЕННЯ 2.2. Імовірнісними (стохастичними) назвемо такі системи, в яких немає визначеного конкретного взаємозв'язку між вхідними та вихідними змінними, проте можна встановити деякі ймовірнісні співвідношення між ними.

За допомогою стохастичних моделей будуються прогностні оцінки систем, що вивчаються. При цьому аналітичні вирази статистичних закономірностей визначаються, наприклад, за допомогою методів математичної статистики, статистичних рішень та економетрії.

ВИЗНАЧЕННЯ 2.3. Індетермінованими назвемо такі системи, для яких встановлення причиново-наслідкових зв'язків на початковому етапі досліджень не є можливим за умов застосування будь-якого відомого апарату математичних засобів.

Прикладами таких систем можуть бути системи з великою розмірністю параметрів, неповнотою, невизначеністю або взагалі відсутністю опису параметрів тощо. Інакше кажучи, ці системи являють собою так звані "чорні ящики", для дослідження яких можуть бути застосовані, наприклад, методи імітаційного моделювання, деякою мірою методи експертних оцінок, різні евристичні методи і

т. ін. Результатом використання цих методів є розширення знань про систему з метою залучення у подальшому формалізованих методів дослідження, що дає змогу у підсумку віднести систему, яка розглядається, до класу детермінованих або стохастичних. Класифікація систем з точки зору відкритості та закритості систем наведена на рис. 2.2.

Проблема ідентифікації (у широкому розумінні) зводиться до побудови опису досліджуваної системи або її елементів у вигляді структурно-функціональних відношень. Ця проблема тісно пов'язана із задачами параметричної надійності та теорії статистичних рішень.

Професор Лінчепнінського університету (Швеція) Л. Льюнг, предмет теорії ідентифікації визначив таким чином: *"Формування моделей на основі результатів спостережень і дослідження їх властивостей – от, по суті, основний зміст науки. Моделі ("гіпотези", "закони природи", "парадигми" і тощо.) можуть бути більш-менш формалізованими, але всі мають ту головну особливість, що зв'язують спостереження в деяку загальну картину. Вирішення завдання побудови математичних моделей динамічних систем за даними спостережень за їх поведінкою становить предмет теорії ідентифікації, що тим самим стає елементом загальної наукової методології"*.

Проблема аналізу, тобто дослідження поведінки моделі за її описом. У межах цієї проблеми можна виділити два основних кола задач:

- вивчення поведінки моделі за описом, що задається;
- дослідження впливу на поведінку моделі характеристик її опису.

До першого типу належать такі задачі: знаходження подій, які подані в моделі, оцінка якості її функціонування, контроль цієї якості і т. ін. Другий тип

включає задачі оцінки чуттєвості показників якості до варіацій параметрів зовнішнього середовища, аналіз інваріантності поведінки моделі до її параметрів та структури тощо.

Проблема синтезу, тобто побудови опису моделі за її поведінкою. Ця проблема також має два аспекти: параметричний і структурний. У першому випадку використовується можливість впливу на поведінку моделі шляхом зміни її параметрів або ймовірнісних характеристик. Таке "параметричне" управління дає можливість ввести визначені критерії оптимальності, які формалізують процедуру параметричного синтезу. У теоретичному і практичному аспекті наслідки реалізації параметричного синтезу, без сумніву, становили б інтерес при дослідженні взаємодії економічних систем (наприклад, банків) із зовнішнім середовищем і їх колективної поведінки.

У другому випадку вирішуються, по суті, задачі введення такої надмірності в опис моделі, яка забезпечує потрібну поведінку модельованої системи. Тут безсумнівний тісний зв'язок прогностичних моделей і моделей, які реально існують.

Проблема складності моделі і її зв'язок з економічними проблемами. Вирішення цієї проблеми є однією з першочергових базових задач при визначенні пріоритетності системних досліджень. Зазначимо, якщо внутрішня складність системи є абсолютною і визначається множиною елементів, зв'язків, перехідних процесів і т. ін. саме цієї системи, то зовнішня (по відношенню до спостерігача або дослідника) складність системи є відносною оцінкою. Внутрішня та зовнішня складність системи, своєю чергою, визначаються її структурною і функціональною складністю.

Проблема розкладання збурювальної змінної (яка є складовою моделі системи) на складові – це визначення "*вагових коефіцієнтів*" параметрів, тобто часток, які вносять ці складові у загальну величину збурювальної змінної, відповідно впливаючи також і на "прогностичну" модель.

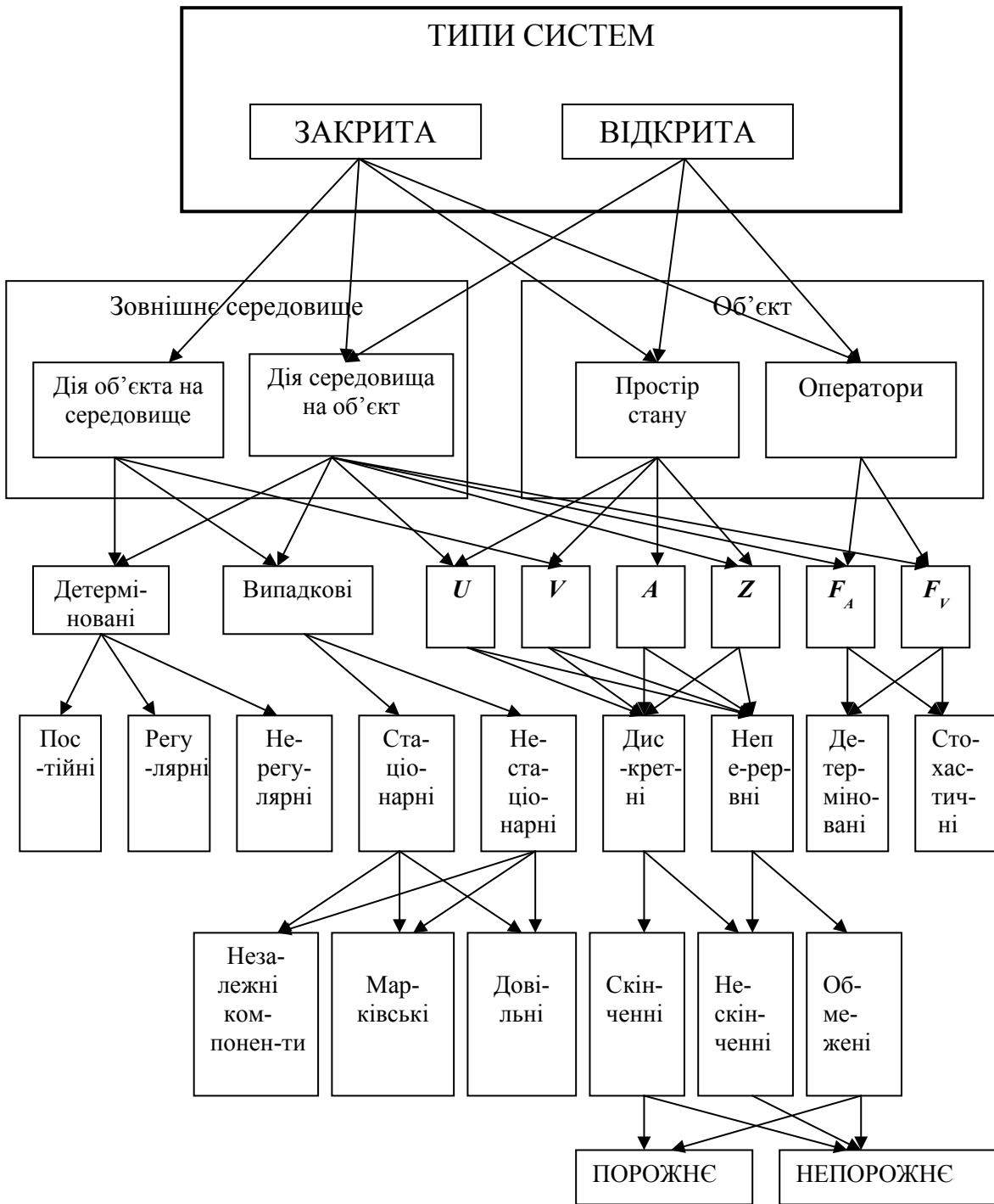


Рисунок 2.2 – Класифікація типів систем

РОЗДІЛ 3

Причиново-наслідкові відношення. Причиновий аналіз

3.1. Про причиновість у соціально-економічних явищах і процесах

Поняття процесу – складна категорія. У [70] наведені визначення категорії процесу в енциклопедіях і словниках. Наведемо деякі з них:

1) процес – послідовна зміна чогось, рух (Енциклопедичний словник Ф.А. Брокгауза і І.А. Єфрона, 1907);

2) процес – закономірна, послідовна зміна явища, його перехід в інші явища (Філософський словник; за ред. М. Розенталя, 1975);

3) процес – рух уперед: послідовна зміна чогось; сукупність дій, які спрямовані на досягнення якогось результату; розгляд судової справи: встановлений законом порядок розгляду і розв'язання судових справ (МРЕ, ВРЕ, 1949 – 1959 рр.);

4) процес соціальний – проходження, просування – послідовна зміна стану суспільства або його окремих систем (Соціологічний довідник; за ред. В.І. Воловича, 1990 р.);

5) визначення або поняття процесу немає (Політехнічний словник; за ред. І.І. Артоблевського, 1982; Д.П.Горський та ін., Короткий словник з логіки, 1991; Л.І. Лопатников, Популярний економіко-математичний словник, 1992; Філософська енциклопедія: В 5-ти т., 1960-1970 рр.).

Таким чином, поняття процесу трактується та визначається неоднозначно, а у деяких словниках воно взагалі не згадується. Аналізуючи п. 1...4, можна зробити висновок (хоч він і буде деякою мірою спрощеним) основними ознаками будь-якого процесу є:

- послідовна зміна;
- перехід з одного явища в інше;
- одnobічний рух (просування);
- зміна стану об'єктів або заміна однієї системи на іншу.

На наш погляд, наведене у [70] визначення процесу є найбільш повним, таким, що визначає всі його характерні ознаки:

– процес і рух – це не одне й те саме, оскільки поняття "процесу" є історично більш пізнім і другорядним, охоплює "рух того, що рухається", а не просто рух;

– усі процеси є двобічні, тобто являють собою сукупність не менш, як двох полярно або вузько протилежних сторін (дія та протидія, притягання та відштовхування, зростання та зменшення, збурення та гальмування, монтаж та демонтаж і т. ін.);

– будь-який процес це завжди кількісно-якісний перехід, прихований, як правило, сферою невизначеності;

– невизначеність процесу є момент переходу від однієї сторони процесу до іншої;

– процес взагалі (в цілому) являє собою зміну взагалі, тобто невизначену

зміну. Аналізуючи викладене, можна дати таке визначення процесу.

ВИЗНАЧЕННЯ 3.1. *Процесом* назвемо таку двосторонню зміну форми руху матерії, під час якої спостерігається перехід від однієї сторони явища до іншої, зовнішньо прихований сферою невизначеності. З урахуванням цього можна зробити висновок, що невизначеність притаманна будь-якому процесу.

ВИЗНАЧЕННЯ 3.2. Першу сторону явища в процесі назвемо *причиною*, а другу – *наслідком*.

Основною метою процесного аналізу в галузі економічного управління є допомога у вирішенні не простих задач, а проблем, пов'язаних з граничним ризиком і максимальною невизначеністю.

3.2 Необхідність формалізації причиново-наслідкових відношень у вивченні економічних процесів

Для вивчення процесів взагалі та економічних процесів зокрема існують різноманітні формальні методи. Один з них – метод *причинового аналізу* – має на меті створити такий опис системи пов'язаних між собою змінних, який дає змогу вказати змінні, що є "причинами" і "наслідками" та прогнозувати значення останніх по перших. Введемо наступне визначення.

ВИЗНАЧЕННЯ 3.3. *Подія* – це вектор, компонентами якого можуть бути різноманітні параметричні зміни, які виникають внаслідок процесів, що відбуваються у системі. Надалі будемо позначати події великими латинськими літерами *A, B, C* та ін.

ВИЗНАЧЕННЯ 3.4. *Причиновий зв'язок* – це таке поєднання процесів реальної дійсності, за якого зміна одного з них є наслідком зміни іншого. З визначення 3.4 випливає, що наслідок має визначене розміщення у просторі та часі.

ВИЗНАЧЕННЯ 3.5. *Оператор* – це речовинний пристрій або наділений структурний процес, який забезпечує виконання причинового зв'язку. Оператор складається з *організованих компонентів*. Вони також оператори, бо являють собою пристрої для перетворення однієї події в іншу. Для того, щоб оператор знаходився в області попередньої події, він повинен бути відповідним чином погоджений у просторі та часі.

З метою спрощення події можуть розумітися як поля, що беруть початок в операторі, який їх утворює, але розповсюджуються у просторі та часі від свого початкового просторово-часового місцезнаходження. Таким чином, між двома подіями може існувати причиновий зв'язок тільки у тому випадку, коли деякий оператор здійснює перетворення. Тому перша подія є причиною другої тільки тоді, коли перша погоджується з деяким оператором, який на неї реагує і має здатність спричиняти другу [28].

Тепер ми можемо дати більш формальне трактування причиновості, застосовуючи наступне визначення.

ВИЗНАЧЕННЯ 3.6. Подію *A* назвемо *причиною* події *B* тоді і тільки тоді, коли виконуються такі умови:

- 1) існує оператор, який породжує *B*, реагує на *A* і створений таким чи-

ном, що зв'язок між A і B можна розкласти у послідовність сумісних компонентів з полями подій, які перетинаються;

2) здійснення події A погоджене з існуванням оператора умови 1), - він завжди існує у полі події A ;

3) коли умови 1) і 2) виконуються, оператор є ізольованим від полів подій, які відрізняються від A , причому ні A , ні B не присутні заздалегідь, тоді здійснення A завжди починається до початку здійснення B ;

4) коли умови 1) і 2) виконуються, A тягне за собою B ; тобто протягом деякого проміжку часу здійснення A завжди супроводжується здійснюванням B , хоча B може бути присутнім без A або ж обидві події можуть бути відсутні.

Виходячи з вищевикладеного, сформулюємо стисло основні особливості, притаманні **причинному висновку**:

а) одна подія не є безпосередньою причиною іншої, якщо немає жодного діючого оператора, який міг би бути основою цього зв'язку;

б) подія не може бути причиною іншої події, якщо перша подія не узгоджується з існуючими операторами;

в) причинами події не можуть бути інші події, які здійснюються після цієї події;

г) якщо подія A здійснюється без наступного здійснення події B , то A не є причиною B за даних обставин.

Таким чином, процедури причинного аналізу акцентують увагу на конфігураціях подій у часі або в конкретні моменти часу, а не на змінах взагалі. Причинова обумовленість породжує моделі подій, а вивчення останніх може забезпечити розуміння причинових відношень, які їх породжували. Це має дуже важливе значення з точки зору питань вирішення задач управління [28].

3.3 Про алгебру каузаций. Причинний аналіз як один із засобів формалізації причинно-наслідкових відносин на різних класах економічних моделей

Розглянемо зараз питання, які безпосередньо пов'язані із побудовою *алгебри каузаций (каузація* – встановлення зв'язку двох явищ або фактів). У суворій формі каузація встановлює причинно-наслідкові зв'язки між явищами або фактами. У більш широкому розумінні каузація встановлює вплив одних явищ або фактів на інші і відображується у моделях знань у вигляді каузальних сіток і сценаріїв. У вузькому розумінні каузація в тих же моделях призводить до причинно-наслідкових сіток.

У [28] подано одну з можливих формальних методик побудови каузальних теоретичних моделей, яка базується на застосуванні спеціального виду математичних формулювань – сіткових графів, які являють собою рівняння, що описують ці моделі у наочній формі. При застосуванні невеликої кількості правил, що піддаються інтерпретації, можна отримати формалізовані математичні висновки, використовуючи лише процедури накреслювання і вивчення цих графів.

Сіткові графи – це міст між якісними теоріями і більш абстрактними

представленнями інших теорій, які використовують мову рівнянь. Побудова сіткового графа починається із введення визначень усіх змінних, які підлягають розгляду. Кожна змінна позначена на діаграмі символом або аббревіатурою. Фактично процес конструювання сіткового графа припускає розгляд однієї змінної у деякий момент часу і дослідження її причинових відношень з іншими змінними системи. Процедура побудови діаграми, яку розглянуто нижче, застосовується тільки для **безпосередніх причинових відношень**.

ВИЗНАЧЕННЯ 3.7. Причинове відношення назовемо **безпосереднім**, якщо зміна значень однієї із змінних зумовлює зміну значень другої змінної без необхідної зміни будь-якої третьої змінної між ними, яка спеціально вказується.

Суцільна (пряма) лінія між символами позначає безпосередній причиновий зв'язок для даної пари змінних. Стрілка вказує напрямок причиновості - у бік залежної змінної (наслідку). Записувати це будемо у вигляді: $X \rightarrow Y$, тобто X є причиною для наслідку Y .

Причину бажано розміщувати на графі ліворуч або зверху, а наслідок - праворуч або знизу.

Кожний причиновий шлях позначається єдиним визначаючим символом.

Шляховий символ припускає дві інтерпретації. Перша – він ясно виражає, що оператор повинен встановлювати визначене причиново-наслідкове відношення і дає способи специфікації оператора при обговоренні. Друга – символ виражає кількісний опис оператора.

Як кількісна характеристика символ може означати характеристичну таблицю, складну формулу або число.

У цьому розділі розглядаються лише лінійні перетворення, і шляховий символ є структурним коефіцієнтом часткового причинового відношення. Звичайно, для позначення шляхових символів, використовуються стандартні літери a, c, e, p, f, g, q тощо (символ "b" резервується для інших цілей). Якщо літер недостатньо, додаються індекси. При індексуванні структурному коефіцієнту завжди приписуються два індекси: перший індекс характеризує залежну (тобто вхідну) змінну (наслідок) даного відношення; другий – незалежну (тобто вихідну) змінну (причину).

Наприклад, $X \xrightarrow{a} Y$ означає, що X є причиною Y , а символ "a" характеризує тип операції, яка перетворює змінну X у змінну Y . У подальшому припускається, що це перетворення - лінійне. У цьому випадку "a" виявляється також структурним коефіцієнтом, який описує це саме лінійне перетворення. Запис $X \xrightarrow{a} Y$ означає те саме, що було наведено вище, але тут структурний коефіцієнт визначається індексом більш конкретно.

Подання причинового зв'язку у вигляді $Z_1 \xrightarrow{a_{21}} Z_2$ означає, що Z_1 є причиною Z_2 і $\xrightarrow{a_{21}}$ знову позначає структурний коефіцієнт, що описує як величина Z_1 перетворюється у величину Z_2 .

Якщо числова оцінка значення структурного коефіцієнта невідома, але відомий його знак, то корисна інформація на пояснюючій діаграмі також буде наведена. Наприклад, якщо зростання X призводить до зростання Y , то це будемо записувати у вигляді:

$$X \xrightarrow{a} Y$$

Розглянемо деякі можливі ситуації, які виникають при побудові теоретичної моделі економічного процесу, у вигляді сіткових графів і процедури укладання по них структурних рівнянь:

1. Витки

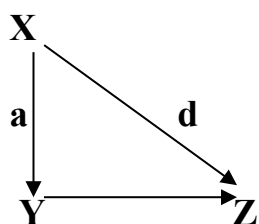
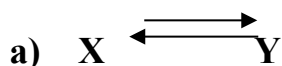


Рисунок 3.1 – Витки

Наведена на рис. 3.1 діаграма, означає, що X – безпосередня причина як для Z , так і для Y . Крім того, Y – безпосередня причина Z .

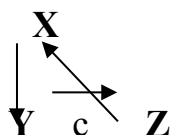
2. Петлі

Діаграми, які наведені на рис. 3.2 а), б), в) ілюструють:



а) X – причина Y , а Y – причина X ;

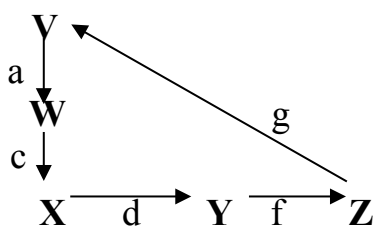
б)



б) опосередкована петля зворотного зв'язку (Y

залежить від X , Z залежить від Y , а X залежить від Z);

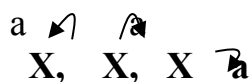
в)



в) опосередкована петля зворотного зв'язку, яка може залишитись невідзначеною у якісній теорії, стає явною, як тільки діаграма буде накреслена.

Рисунок 3.2 – Петлі

3. Прості петлі



Наведені на рис. 3.3 всі три зображення еквівалентні.

Рисунок 3.3 – Прості петлі

Символом "а" позначена операція або множина операцій, які перетворюють змінну X у наступну, але ту саму, змінну X .

4. Відсутність стрілок

Відсутність стрілки свідчить про відсутність (за наявності серйозних аргументів) або невизначеність причинового зв'язку у даній системі. У випадку невизначеності зв'язку, тобто кожного разу, коли виникають сумніви, краще додати стрілку. При подальшому вивченні системи накопичуються серйозні аргументи, що дозволяють або ліквідувати стрілку, або ж залишити її, або ж змінити її напрямком. Процедура накопичування вказаних аргументів визначається принципами причинового висновку.

5. Збурення

Збурення позначається окремим символом поряд із збурювальною змінною: напрямком стрілки – від збурення до змінної.

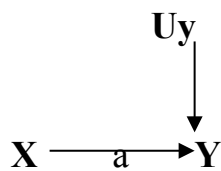


Рисунок 3.4 – Збурення

Зображення на діаграмі (рис. 3.4) зовнішніх збурень означає, що X впливає на Y , але на Y впливають ще інші невизначені фактори, які наведені сукупно через U_y .

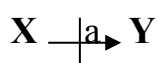
Збурювальні члени звичайно позначаються літерою U із нижнім індексом. Індекс визначається тією змінною, на яку впливає це збурення.

Для ідентифікації стрілки, яка іде від збурень до залежних змінних користуються двома угодами. З одного боку, збурювальний фактор можна розглядати у вигляді змінної зі своїми внутрішніми властивостями, яка вимірюється у деякій специфічній шкалі. Стрілки позначаються літерами – a, c, p тощо із двома нижніми індексами: перший збігається із символом залежної змінної, другий – із літерою U , показуючи, який саме фактор розглядається. Наприклад, стрілка від U до Y позначається, як a_{yu} . З іншого боку, збурювальний фактор може бути виражений у тій самій шкалі, що і залежна змінна, на яку він впливає. У цьому випадку немає необхідності позначати стрілку, бо вона завжди означає одне і те саме перетворення: залежна змінна змінюється на ту саму величину, що і збурювальна. Це – тотожне перетворення зі структурним коефіцієнтом, який дорівнює одиниці, що є умовою відсутності мітки.

6. Нелінійні співвідношення

Для зазначення на діаграмі ситуації, коли перетворення має такий вигляд лише при перевищенні вхідної змінної деякого порогу, вводиться спеціальне правило. Стрілка перекреслюється короткою поперечною рисою, позначається, з якого значення вхідної змінної діє причиновий оператор. На рис. 3.5 наведено приклад, на якому обмежувач на діаграмі вказує на нелінійність.

Нелінійні відношення досить складні та нами не розглядаються.



Однак, зауважимо, що єдиний обмежувач може

Рисунок 3.5 – Нелінійні співвідношення

бути легко вилучений шляхом перетворення діаграми у дві інші – одна застосовується нижче значення граничного, а друга – вище (рис. 3.6).

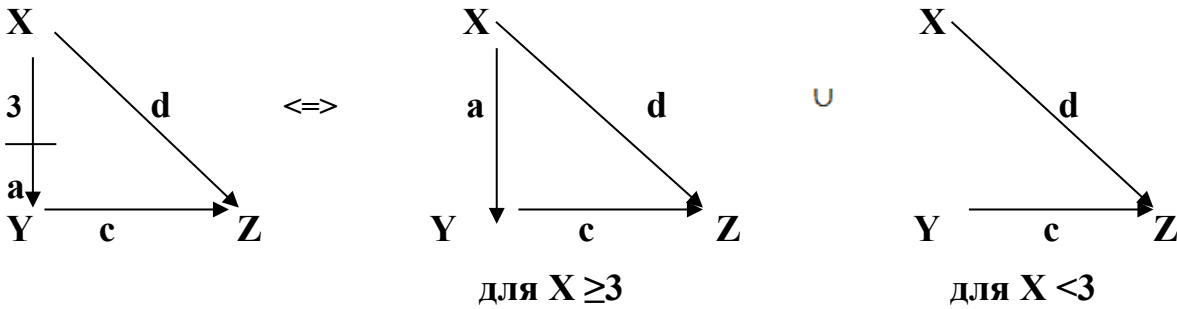


Рисунок 3.6 – Нелінійні співвідношення

Для зображення взаємодії змінних іноді використовують інший вид побудови діаграм. У кінці стрілок від взаємодіючих змінних вміщується коло, яке відзначає залежну змінну зовні і містить всередині символ (звичайно f або g), що виражає спеціальне перетворення значення вхідних змінних у значення залежної змінної. Наведена на рис. 3.7 діаграма означає вираз виду: $Z = X \cdot Y$. Альтернативою зображення мультиплікативної взаємодії схеми, яка представлена на рис. 3.7 є рис. 3.8:

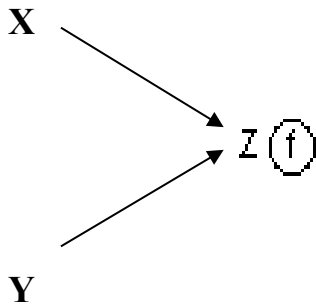


Рисунок 3.7 – Нелінійне співвідношення

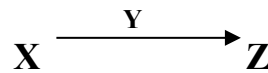


Рисунок 3.8 – Нелінійне співвідношення (альтернативне зображення)

7. Аналіз сіткових графів

У подальшому розглядаються тільки стійкі системи, всі змінні яких під час спостереження були стабілізовані в незмінюваних або статичних значеннях. Обмеження нелінійності відношень потребує виміру змінних у інтервальній шкалі, так, що виконуються властивості адитивності та пропорційності. Взагалі не потрібно, щоб шкала була інтервальною, достатньо її деякого наближення. Техніка вимірювання в соціальних та економічних науках досить розвинута, щоб забезпечити прийнятну апроксимацію інтервального вимірювання для аналізу багатьох каузальних систем.

8. Складання рівнянь

Значення змінної, що визначається єдиним входом, дорівнює значенню входу, помноженому на структурний коефіцієнт:

$$X \quad a \quad Y$$

а) $a \longrightarrow \Leftrightarrow Y = aX;$

б)
$$\begin{array}{ccc} & 3 & \\ & \downarrow & \\ X & \xrightarrow{a} & Y \end{array} \Leftrightarrow Y = aX + 3$$

Значення змінної, що визначається однією або декількома вхідними величинами, дорівнює сумі вхідних значень, помножених відповідно на їх структурні коефіцієнти. Порядок підсумовування не має ніякого значення, тобто володіє властивістю комутативності:

в)
$$\begin{array}{ccc} X & \searrow a & \\ & & Z \\ Y & \nearrow c & \end{array} \Leftrightarrow Z = aX + cY \text{ або } Z = cY + aX;$$

д)
$$\begin{array}{ccc} & Uy & \\ & \downarrow & \\ X & \xrightarrow{a} & Y \end{array} \Leftrightarrow Y = aX + 1 \cdot Uy = aX + Uy \text{ або } Y = Uy + aX$$

9. Зображення структурних рівнянь

Для кожної залежної змінної шляхи, які відходять від змінної, при написанні рівняння для цієї змінної, не враховуються, але кожна вхідна стрілка вказує член, який повинен бути врахований. Змінна із петлею зворотного зв'язку розглядається так само, як й інша змінна: рівняння для неї містить члени для кожного з її безпосередніх входів навіть у тому випадку, коли ця вхідна змінна міститься в деякій іншій петлі зворотного зв'язку. В економічній літературі (у тому числі і економетричній) залежні змінні називаються *ендогенними змінними*, а змінні, значення яких визначаються зовнішніми для системи числами – екзогенними.

10. Правило редуції

Якщо одна змінна визначає іншу, а та, в свою чергу, визначає третю, то значення третьої змінної може бути виражене як значення першої змінної помножене на добуток структурних коефіцієнтів вздовж ланцюга. Цей самий принцип застосовується, коли ланцюг має більше ніж дві ланки:

$$X \xrightarrow{a} Y \xrightarrow{c} Z \Leftrightarrow Z = acX \Leftrightarrow X \xrightarrow{ac} Z$$

Щоб виразити значення залежної змінної через численні безпосередні та посередні входи, спочатку отримують окремі впливи вздовж кожного ланцюга за правилами п. 8 – складання рівнянь і п.10 – правилу редуції. Потім підсумовують усі впливи відповідно до правила 8 п. с і д. Приклад наведений на рис. 3.9.

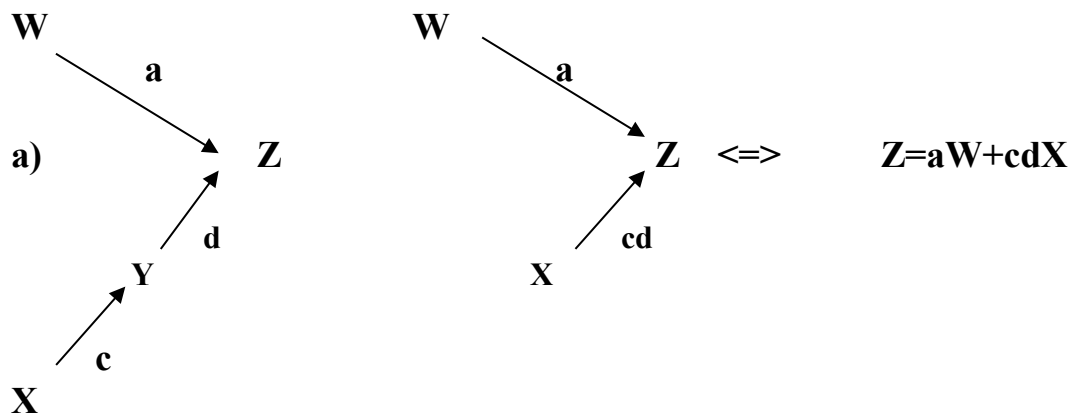


Рисунок 3.9 – Правило редуції(а)

Коли на вхідну змінну діють декілька непрямих вхідних джерел через одну й ту саму проміжну змінну, ланцюг від кожного джерела визначається автономно, так якби інших джерел не було. Наприклад, рис. 3.10.

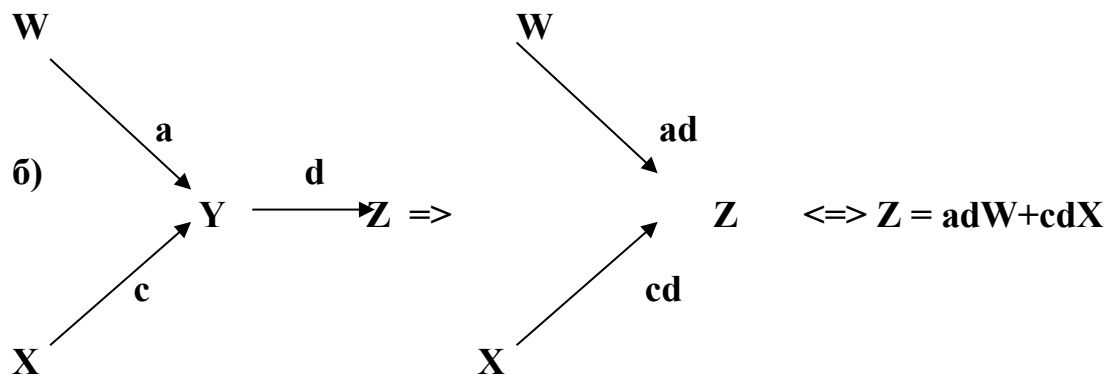


Рисунок 3.10 – Правило редукції (б)

Це правило краще застосовувати, коли існує декілька вхідних змінних, а не одна. Якщо проміжна змінна схильна до збурення, збурювальний член інтерпретується як ще одне джерело.

Якщо джерело і залежна змінна пов'язані багатьма шляхами, то зв'язок між ними дорівнює сумі впливів вздовж кожного окремого шляху. Вплив вздовж шляху знаходиться за правилами складання рівнянь – п. 8 і правилом редукції. Порядок підсумовування не має значення. Ситуацію, що відповідає цьому випадку, наведено на рис. 3.11.

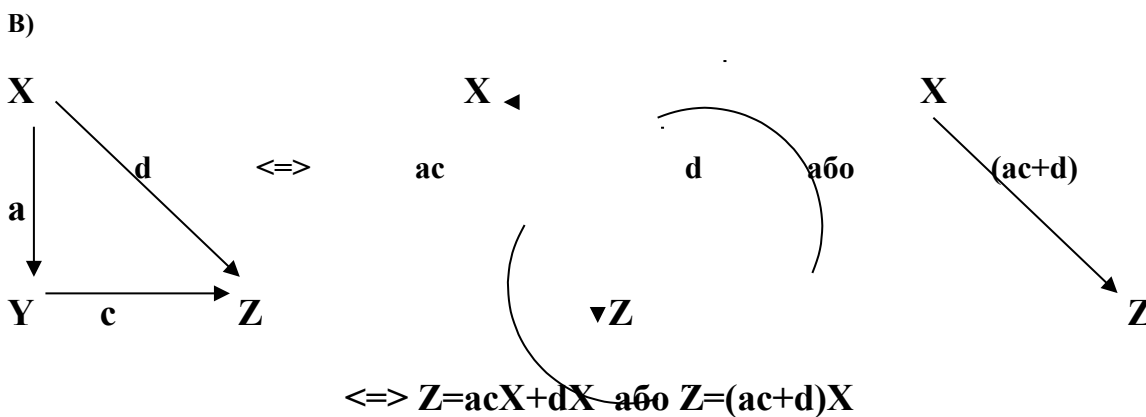


Рисунок 3.11 – Правило редукції (в)

Правило, наведене на рис. 3.11 свідчить, що вплив однієї змінної на іншу відбувається по усіх ланцюгах, які зв'язують обидві ці змінні.

11. Додаткова інтерпретація правил

Правило редукції для ланцюга означає що, якщо всі коефіцієнти вздовж ланцюга додатні, то зростання джерела зумовлює зростання кінцевої змінної, а спадання джерела зумовлює спадання кінцевої змінної. З іншого боку, єдиний від'ємний знак де-небудь вздовж ланцюга призводить до протилежного правила напрямку зміни від входу до виходу. У загальному випадку, якщо в ланцюгу відсутні від'ємні коефіцієнти або число їх парне, то і зміна у джерелі веде до зміни залежної змінної того самого знаку. Якщо у ланцюгу з'являється непарне число від'ємних коефіцієнтів, зміни у значенні вхідної змінної зумовлюють зміни кінцевої змінної у протилежних напрямках.

12. Сіткові графи з петлями:

ВИЗНАЧЕННЯ 3.8. Ланцюг стрілок, що відходить від деякої змінної і повертається, врешті-решт, назвемо петлею, за умови, що вздовж усього шляху жодного разу не змінюються напрями стрілок і, що увесь шлях не проходить ні через одну змінну більше, ніж один раз, за винятком змінної, яка прийнята за початок, через яку, вважається, шлях проходить двічі.

На рис. 3.12 – 3.15 наведені приклади сіткових графів з петлями.

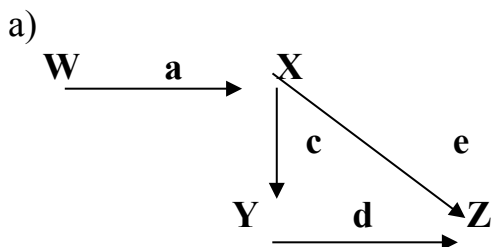
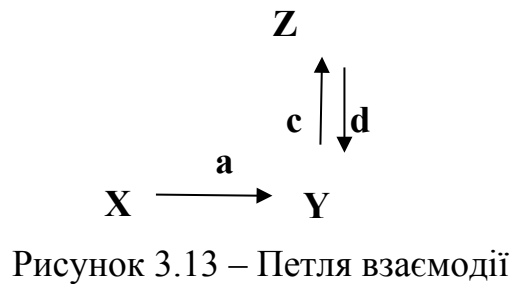
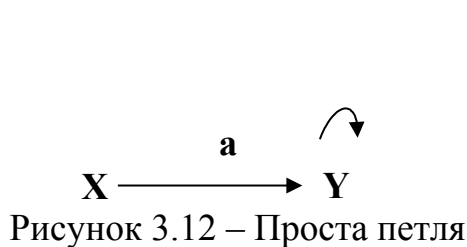


Рисунок 3.14 - Опосередковані петлі зворотного зв'язку

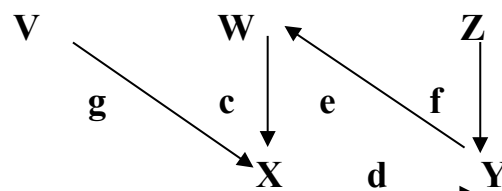
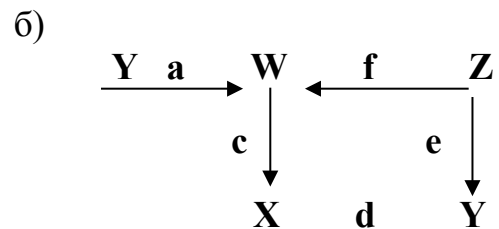


Рисунок 3.15 – Граф з однією петлею (WXY)

На рис 3.16 наведені більш складані графи систем.

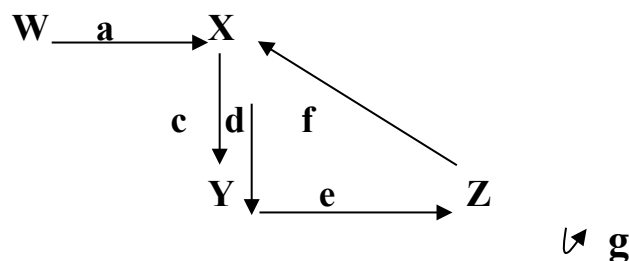


Рисунок 3.16 – Система із трьома петлями (XY), (XYZ), (Z)

На системі, яку наведено на рис. 3.16 показано *гніздо петель*, тобто множину петель, ланцюжки яких частково перехрещуються. Петлю гнізда необхідно ідентифікувати як окрему, якщо тільки ця петля відмінна від інших, єдиною стрілкою. Петлі, які є суміжними, не породжують гнізда.

13. Ефект зворотного зв'язку

"Ефект зворотного зв'язку" петлі, що позначається літерою **L**, дорівнює добутку коефіцієнтів вздовж петлі.

Приклад обчислення **L** наведено на рис. 3.17.

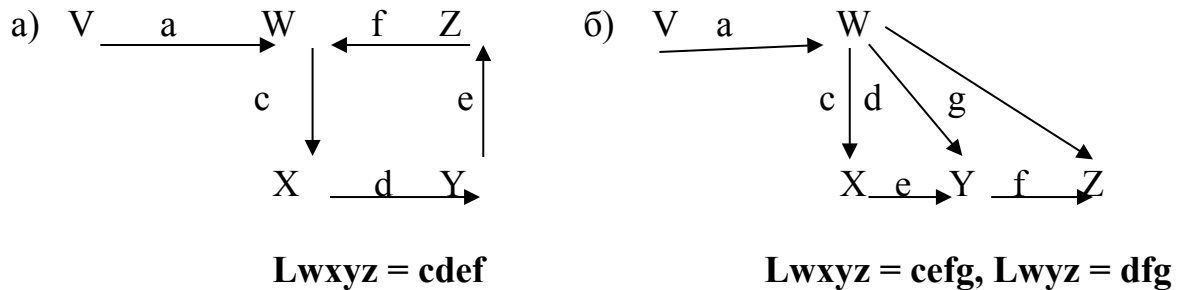


Рисунок 3.17 – Ефект зворотного зв'язку

Ефект зворотного зв'язку, "зворотний ефект", допускає змістовну інтерпретацію. Він вказує, на скільки одиниць зміниться змінна петлі протягом одного циклу змін вздовж всієї петлі, що пройшов після початкової зміни значення на одну одиницю.

Ланцюжок стрілок являє собою "відкритий" шлях, якщо вздовж неї немає жодного обертання напрямку і, якщо стрілки цього шляху не направляються ні на одну змінну більше одного разу. Приклад відкритого шляху для системи, яка має один відкритий шлях між **W** та **Z** (**WXYZ**) і два відкриті шляхи між **Y** та **Z** (**YZ**)(**ZXY**) наведено на рис. 3.18:

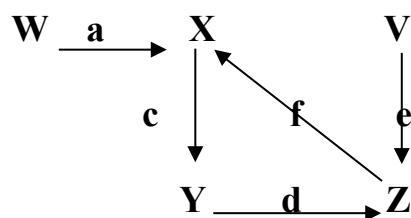


Рисунок 3.18 – Приклад відкритого шляху

14. Ефект відкритого шляху:

"Ефект відкритого шляху", що позначається через **E**, дорівнює добутку коефіцієнтів вздовж усього ланцюга відкритого шляху. Приклади обчислення ефекту відкритого шляху, наведено на рис. 3.19 і 3.20.

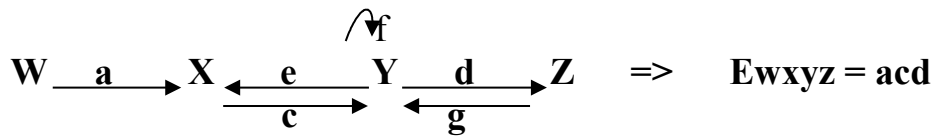


Рисунок 3.19 – Приклад обчислення ефекту відкритого шляху

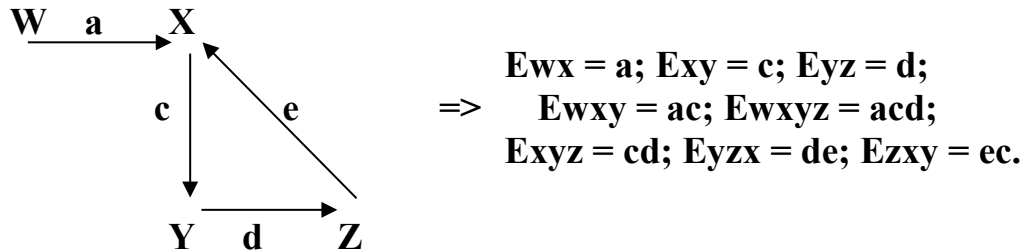


Рисунок 3.20 – Приклад обчислення ефекту відкритого шляху

15. Стикові шляхи:

Дві петлі назвемо "*стиковими*", якщо їх ідентифікатори містять один або декілька загальних символів.

Петля дотикається відкритого шляху, якщо будь-який символ з ідентифікатора відкритого шляху (за винятком найпершого) присутній також в ідентифікаторі петлі.

Ідентифікатори петель і відкритих шляхів являють собою список змінних, які ідуть вздовж відповідних шляхів. На рис. 3.21 наведено приклад, у якому петля (WX) дотикається відкритого шляху (VWZ), але не (WZ), а петлі (WX) і (XY) стикаються.

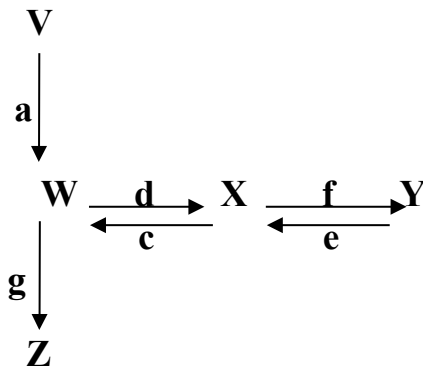


Рисунок 3.21 – Приклад стикових шляхів

Концепція "дотику" дає змогу визначити *релевантний зворотний зв'язок* таким чином. Нехай змінна X буде входною змінною для Y, тобто X впливає на Y безпосередньо або опосередковано, а Y не впливає на X ні прямо, ні опосередковано. Очевидно, від X до Y повинен бути один або декілька відкритих шляхів. Якщо в цій системі присутні такі петлі, то релевантний зворотний зв'язок для X-Y – відношення є множина всіх петель, що дотикаються хоча б одного із відкритих шляхів від X до Y, плюс додаткові петлі, які дотикаються петель першої множини, плюс нові петлі, що дотикаються попередніх і тому інше.

16. Редукція петель

Основні принципи аналізу графів систем із петлями були запропоновані для технічних додатків Самуелем Месоном у 1951 році. Основне правило формулюється таким чином:

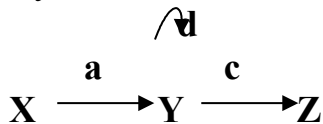
Повний ефект впливу T деякої вхідної змінної на довільну залежну змінну може бути розрахований таким чином: нехай **E**, **E₁**, **E₂**,...суть окремі ефекти відкритих шляхів входу на залежну змінну; **L**, **L₁**, **L₂**,... – зворотні ефекти для всіх окремих петель, які забезпечують релевантні зворотні зв'язки.

Тоді

$$T = \left[\frac{(E + E_1 + E_2 + \dots)(1 - L)(1 - L_1)(1 - L_2)\dots}{(1 - L)(1 - L_1)(1 - L_2)\dots} \right]^*$$

де $[\bullet]^*$ є спеціальний оператор, який позначає, що спочатку і в чисельнику, і в знаменнику відбувається множення, потім викреслюються члени, які дорівнюють добуткам ефектів дотичних шляхів, і тільки потім відбувається ділення.

Розглянемо приклад обчислення повного ефекту впливу. Нехай задано діаграму системи



Оскільки між **X** і **Z**, існує тільки один відкритий шлях з ефектом **E_{xyz} = ac** і тільки одна петля зі своїм зворотним ефектом **L_y = d**, отримаємо

$$T = \left[\frac{E_{xyz}(1 - L_{y1})}{1 - L_y} \right]^* = \left[\frac{E_{xyz} - E_{xyz}L_y}{1 - L_y} \right]^*$$

Відкритий шлях від **X** до **Z** і петля (**Y**) дотикаються один одного, тому оператор $[\bullet]^*$ вимагає знищення члена **E_{xyz}L_y**. У результаті отримаємо:

$$T_{zx} = \frac{E_{xyz}}{1 - L_y} = \frac{ac}{1 - d}$$

Про властивості стабільності з єдиною петлею свідчить значення ефекту зворотного зв'язку. Якщо абсолютна величина зворотного ефекту більша за одиницю або дорівнює їй (тобто, величина зворотного ефекту ≤ -1 або $\geq +1$), система буде нестійкою. Якщо величина зворотного ефекту знаходиться між "-1" і "+1", то система з єдиною петлею буде стійкою.

Система з багатьма петлями може бути нестійкою, навіть якщо кожна індивідуальна петля окремо породжувала б стійкість. І, навпаки, якщо на діаграмі присутня тільки одна петля, то її вплив позначається просто у поправці всіх коефіцієнтів відкритих шляхів її дотичних на множник $1/[1-L]$.

На рис. 3.22 наведено приклад, який включає найбільш важливі моменти стикових шляхів і редукції петель.

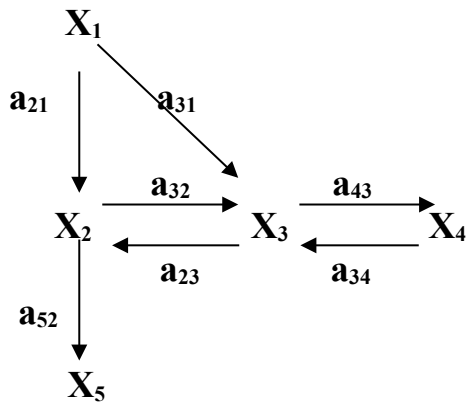


Рисунок 3.22 – Приклад стикових петель і редуції петель

Від X_1 до X_5 існують два відкритих шляхи:

$$E_{125} = a_{21}a_{52};$$

$$E_{1325} = a_{31}a_{23}a_{52}.$$

У системі дві петлі:

$$L_{23} = a_{32}a_{23} \text{ і } L_{34} = a_{43}a_{34}.$$

Тому повний ефект від X_1 до X_5 дорівнює:

$$T = \left[\frac{(E_{125} + E_{1325})(1 - L_{23})(1 - L_{34})}{(1 - L_{23})(1 - L_{34})} \right]^*.$$

Спочатку розкриємо дужки знаменника, тому що він є і в чисельнику і в знаменнику.

У результаті перемноження дужок отримаємо:

$$1 - L_{23} - L_{34} + L_{23} \cdot L_{34}.$$

Два ефекти в останньому члені належать до дотичних петель, тому даний член відкидається. Отже:

$$[1 - L_{23} - L_{34} + L_{23} \cdot L_{34}]^* = 1 - L_{23} - L_{34}.$$

З урахуванням цього співвідношення перепишемо формулу для T_{51} у вигляді

$$T_{51} = \frac{[(E_{125} + E_{1325})(1 - L_{23})(1 - L_{34})]^*}{1 - L_{23} - L_{34}}.$$

Добуток у верхній частині відношення має вигляд:

$$E_{125} + E_{1325} - \frac{E_{125}L_{23}}{1 - L_{23}} - \frac{E_{1325}L_{23}}{1 - L_{23}} - E_{125}L_{34} - \frac{E_{1325}L_{34}}{1 - L_{34}}.$$

Підкреслені члени виразу містять співмножник - ефекти, які належать до дотичних шляхів, тому ми їх опускаємо. У результаті такої процедури отримаємо:

$$T_{51} = \frac{E_{125} + E_{1325} - E_{125}L_{34}}{1 - L_{23} - L_{34}} = \frac{E_{125}(1 - L_{34}) + E_{1325}}{1 - L_{23} - L_{34}}.$$

Підставляючи тепер замість кожного ефекту їх більш конкретне значення, отримаємо:

$$T_{51} = \frac{a_{21}a_{53}(1 - a_{43}a_{34}) + a_{31}a_{23}a_{32}}{1 - a_{32}a_{23} - a_{43}a_{34}}$$

17. Зведена форма системи

Як тільки будуть отримані всі повні ефекти, які пов'язують кожну залежну змінну зі всіма вхідними змінними, можна побудувати зведену форму діаграми, щоб показати зв'язок кожної залежної змінної безпосередньо із вхідними і відкинути проміжні змінні. На рис. 3.23 наведено приклад, в якому в діаграмі поданої форми петлі відсутні.

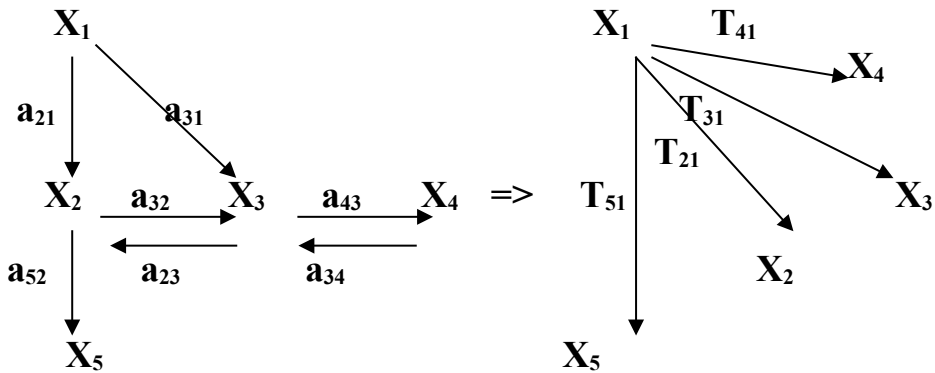


Рисунок 3.23 – Приклад діаграми за відсутності петель

18. Частковий добуток

Петлі можуть бути еліміновані із графа перетворенням їх у гіпотетичні "вхідні" змінні, що лежать між петлею і рештою частини графа. На рис. 3.24 подана діаграма, яка введенням нових – гіпотетичних змінних Y_0 та Z_0 елімінує її петлі.

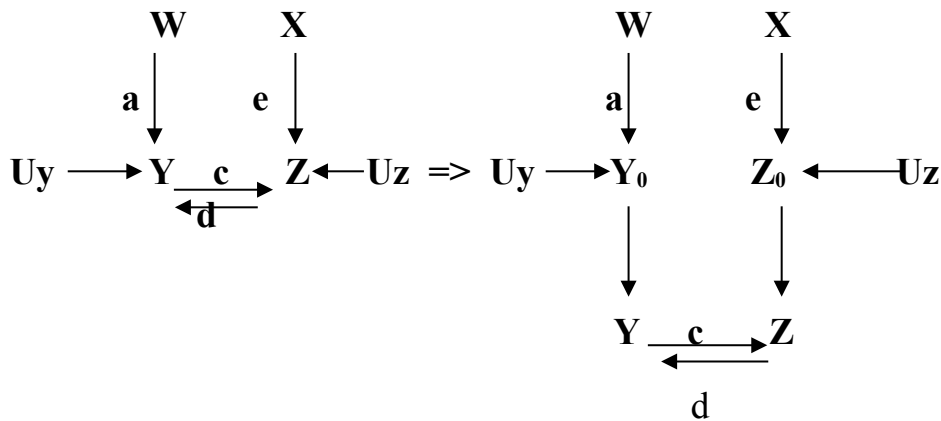
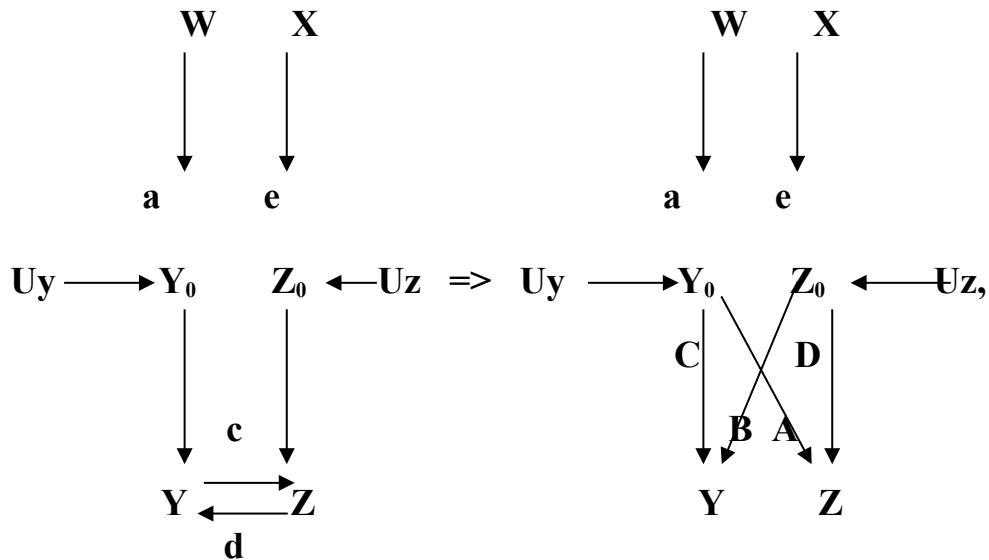


Рисунок 3.24 – Діаграма з введенням нових гіпотетичних змінних Y_0 та Z_0

Стрілки від гіпотетичних вхідних змінних символами не позначаються. Це означає, що коефіцієнти при цих стрілках дорівнюють одиниці. Тепер, використовуючи звичайні процедури зведення стосовно отриманої діаграми, яка включає нові – гіпотетичні змінні і петлю, можна виразити у зведеній формі значення

ня первісних змінних із петлі (рис. 3.25).

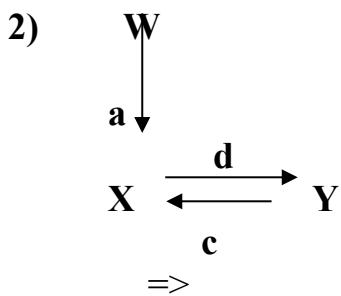
Якщо у змінних петлі тільки єдине зовнішнє джерело, гіпотетичні змінні стають зайвими і можуть бути виключені (рис. 3.26). Таким чином, виключення петлі зберігає правильні відношення між змінними петлі і всіма первісними входами. Більш того, наведені петлі зберігають належні відношення між змінними "вище" і "нижче", як це показано на рис. 3.27.



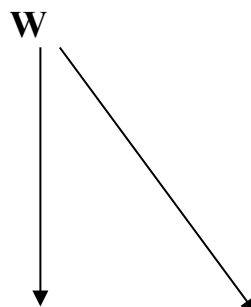
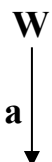
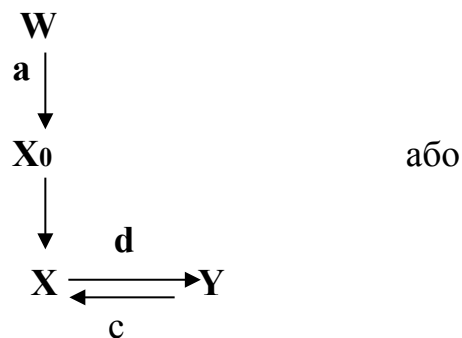
Це: $A = c/(1-L)$, $B = d/(1-L)$, $C = D = 1/(1-L)$.

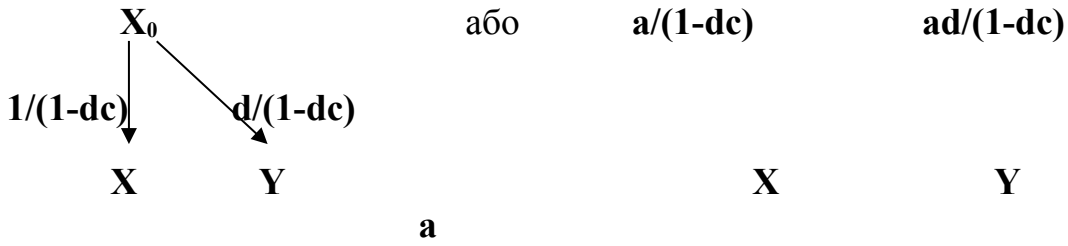
Рисунок 3.25 – Зведена форма значення первісних змінних із петлі

1) $X \xrightarrow{a} Y \Rightarrow X \xrightarrow{a} Y_0 \xrightarrow{\quad} Y \Rightarrow X \xrightarrow{a} Y_0 \xrightarrow{1/(1-c)} Y \Rightarrow X \xrightarrow{a/(1-c)} Y;$



Y не має зовнішнього джерела і для нього не потрібно гіпотетичної змінної. Тому зведена діаграма породжує наступну:





3) $U_y \longrightarrow Y \rightleftharpoons Z \longleftarrow U_z$ перетворюється на діаграму

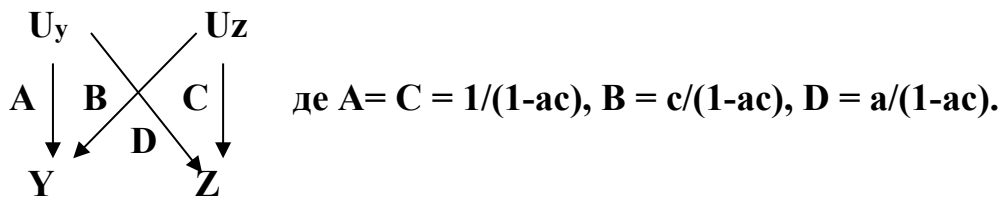
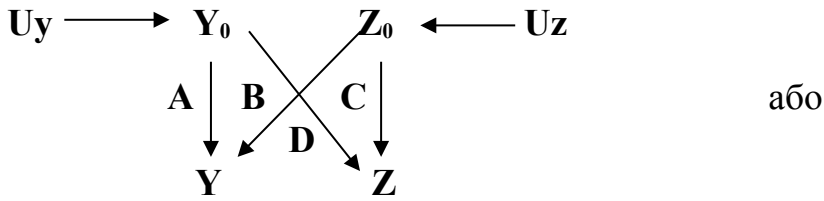
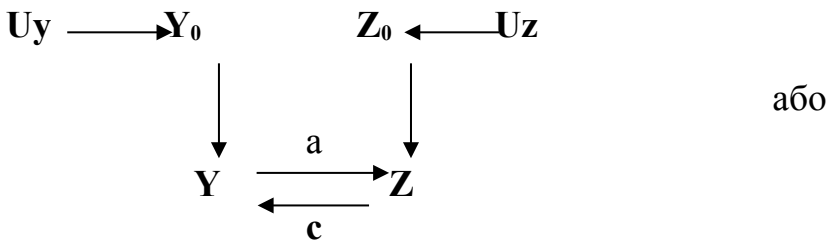


Рисунок 3.26 – Діаграми з виключенням гіпотетичних змінних

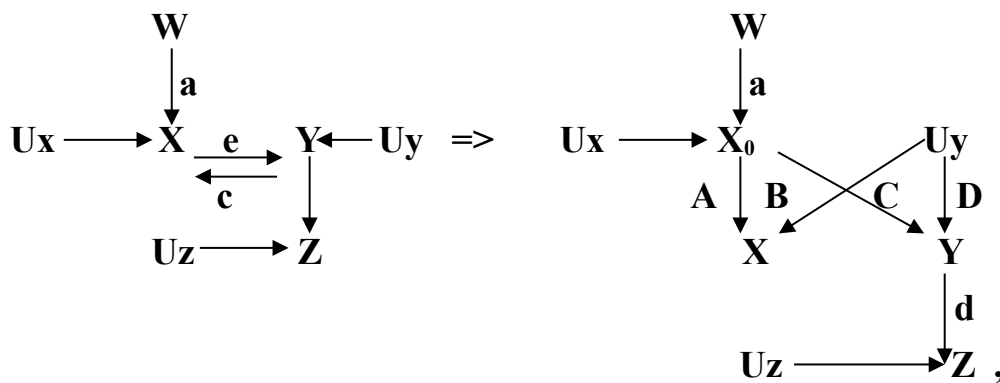
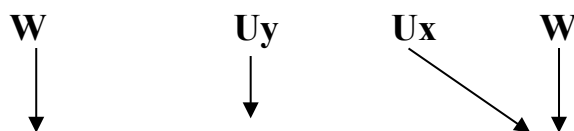
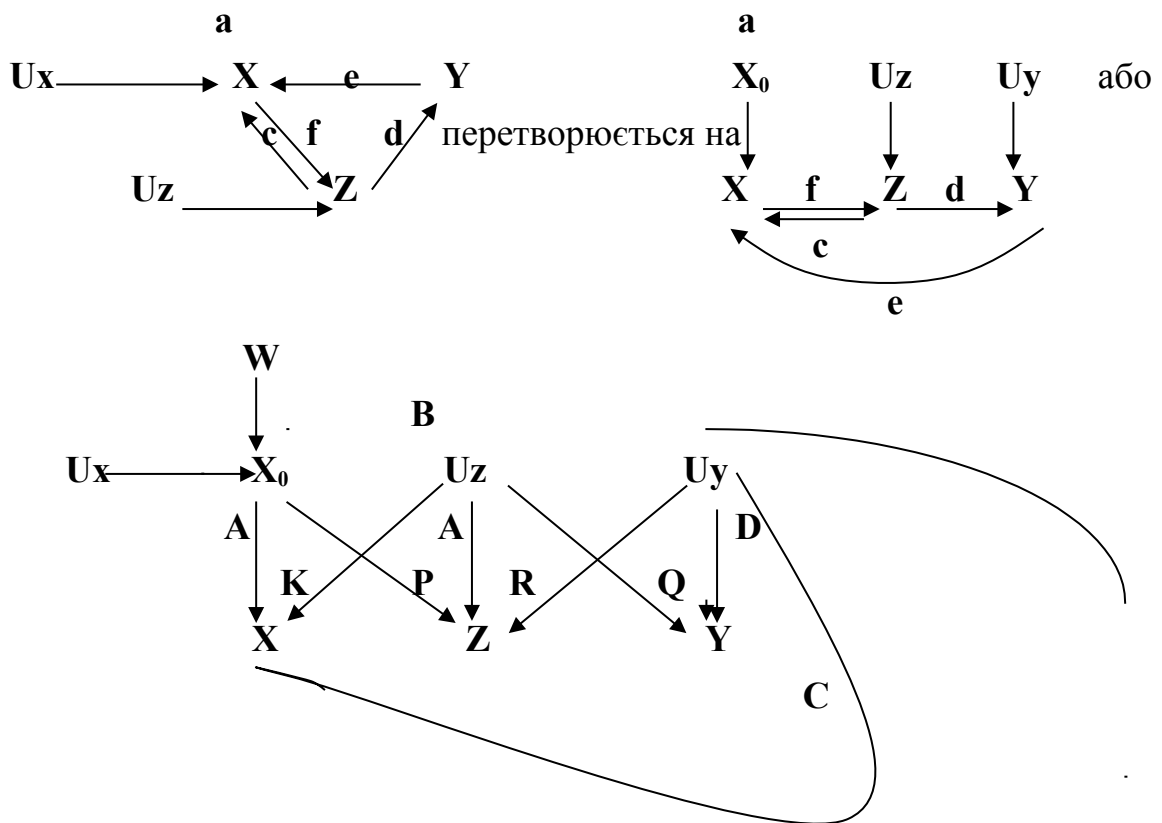


Рисунок 3.27 – Діаграма з виключенням петлі
Також можуть бути частково редуковані гнізда петель (рис. 3.28).





де $A = 1/(1-L_1 -L_2)$, $B = fdA$, $C = eA$, $D = (1-L_1)A$, $K = (c+de)A$, $P = fA$, $R = efA$, $q = dA$, а $L_1 = L_{xz} = fc$, $L_2 = L_{zy} = fde$.

Рисунок 3.28 – Діаграма частково редукованих гнізд петель

19. "Виведення" – інша форма часткового добутку

Вибрані змінні можуть бути "виведені" з петлі виключенням будь-яких стрілок зворотного зв'язку, що йдуть від них, і представленням усіх інших залежних змінних системи у зведеній формі.

Приклад, відповідний "виведенню", подано на рис. 3.29.

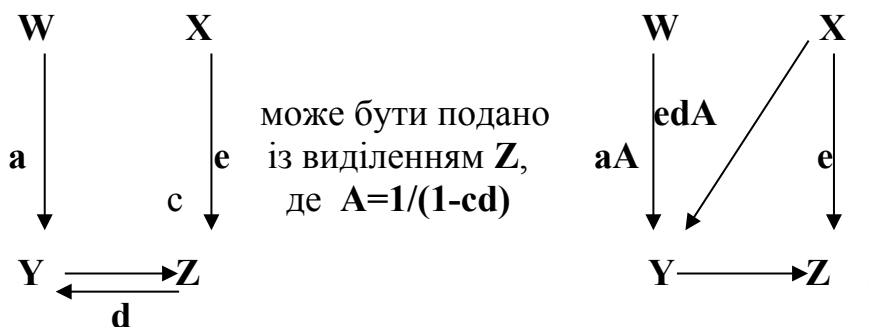


Рисунок 3.29 – Приклад "виведення"

Залежність Y від Z прихована у змінній діаграмі, але всі повні ефекти від джерел наведені правильно:

$$A = 1/(1-cd), T_{yw} = aA, T_{yx} = edA, T_{zw} = a cA, T_{zx} = e+edcA = eA.$$

Система, що наведена на рис. 3.30, може бути зображена із виведеною змінною

Y (див. рис. 3.31).

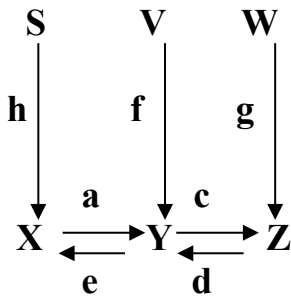


Рисунок 3.30 – Діаграма без виведеної змінної

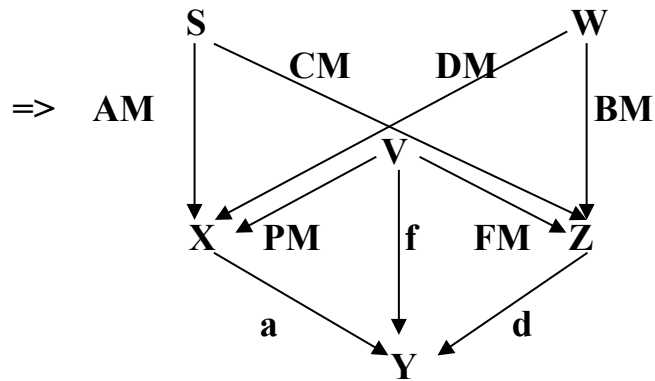


Рисунок 3.31 – Діаграма з виведеною змінною

На рис. 3.31 введено такі позначення: множник $M = 1/(1-ae-cd)$, $A = h(1-cd)$, $B = g(1-ae)$, $C = hac$, $D = gde$, $P = fe$, $F = fc$.

Повні ефекти у цій змінній діаграмі наведено правильно:

$$Tys = [h(1-cd)M]a + [hacM]d = [haM](1-cd+cd) = haM.$$

Виведення змінних дає ще один шлях редукції петель; воно зберігає без змін більше відношень первісних змінних петель, ніж часткове зведення, яке розглянуте раніше (див. п.18).

Однак, редукція за допомогою виведення не симетрична, і завжди можливий інший спосіб редукції, який розглянутий на прикладі, що наведено на рис. 3.32. У ньому перший приклад (див. рис. 3.29) може бути наведено із виведенням Y-змінної.

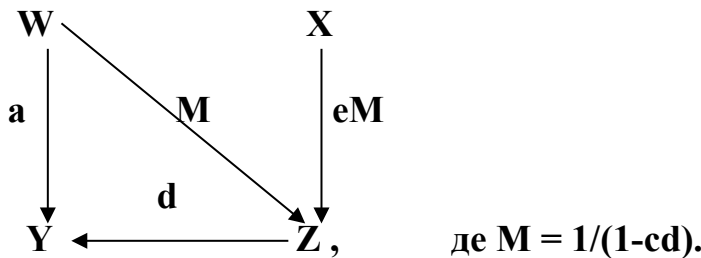


Рисунок 3.32 – Приклад з виведенням Y-змінної

Аналогічно, для другого прикладу може бути виведена змінна X (рис. 3.33).

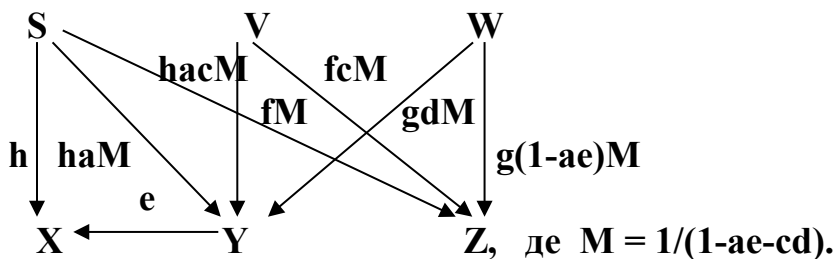


Рисунок 3.33 – Приклад з виведенням X-змінної

Виведення дає можливість побачити відношення змінних петлі до зовнішніх джерел (але не з петлі).

На закінчення побудови правил, які наведено вище, зазначимо, що вони дають змогу зображувати графічно та аналізувати системи різноманітних

структур. Однак слід зазначити, що складні системи потребують часто таких заплутаних графів, що практично зручніше перетворити їх на множину структурних рівнянь і для аналізу користуватися відповідними стандартними математичними засобами.

Вирази зведеної форми, які отримані з аналізу графа, визначають для системи рівняння "вхід-вихід". Якщо на цей момент часу будуть задані індивідуальні значення вхідних змінних, то остаточні наслідки для всіх залежних змінних можуть бути легко розраховані. Таким чином, теорія, яка доповнена кількісними оцінками структурних коефіцієнтів, може бути використана для передбачення майбутніх подій і ситуацій, які будуть наслідком від поточних змін. Отже можна зробити висновок, що сіткові графи і відповідні їм вирази зведеної форми можуть орієнтувати прийняття практичних рішень, а саме:

- 1) вказуючи ті вихідні змінні, зміна яких може призвести до бажаного результату;

- 2) підказуючи величину дій, що необхідні для досягнення бажаних змін;

- 3) вказуючи ті додаткові ефекти, які можуть виникнути всупереч бажаним.

Регресійний аналіз і отримані в результаті його рівняння регресії, які більш докладно розглядатимуться далі, є лише засіб для переведення інформації одного виду в інший. Причинновий аналіз процесів може обумовлювати можливість такого переведення: але це не означає, що рівняння регресії виявляє ці процеси для будь-якого простого випадку. І, нарешті, зазначимо, що у причинновому аналізі регресії корисні, але на рівняння регресії неможливо дивитися шаблонно, як на структурні рівняння, які безпосередньо представляють причинові зв'язки. Причинновий аналіз і регресійний аналіз є попередніми кроками для економічного аналізу складних економічних явищ і процесів (систем).

РОЗДІЛ 4 РЕГРЕСІЙНИЙ АНАЛІЗ. ЙОГО ОСОБЛИВОСТІ ТА РІЗНОВИДИ

4.1 Про економічні величини

Введемо декілька визначень для уникнення невизначеності у використованій термінології [54, ч. 2].

ВИЗНАЧЕННЯ 4.1 Кількісну характеристику властивостей економічної системи (процесу) або явища назвемо **економічною величиною**.

Економічна величина проявляється у вигляді її конкретних реалізацій. Наприклад, величина, яка визначає випуск продукції, маса цієї продукції, її вартість і т. ін.

ВИЗНАЧЕННЯ 4.2. Часткові реалізації однієї і тієї самої економічної величини назвемо **однорідними економічними величинами**.

Однорідні економічні величини відрізняються одна від одної кількісно. Порівняння розмірів двох однорідних економічних величин виконується у процесі вимірювання.

ВИЗНАЧЕННЯ 4.3. Отримане експериментально за допомогою міри порівняння однієї економічної величини з іншою, однорідною з нею економічною величиною, яка прийнята за одиницю вимірювання, назвемо **вимірюванням економічної величини**.

ВИЗНАЧЕННЯ 4.4. **Одиницею вимірювання** економічної величини назвемо конкретне її значення, що прийняте за основне порівняння для кількісної оцінки величин того самого роду.

ВИЗНАЧЕННЯ 4.5. **Мірою** назвемо пристрій, тіло, яке призначене для матеріального відтворення одиниці вимірів.

Наприклад, лінійка – міра довжини, гиря – міра ваги та ін.

Залежно від методів отримання результатів вимірювання розділимо на прямі та непрямі.

ВИЗНАЧЕННЯ 4.6. **Вимірювання** назвемо **прямими**, якщо величина, що вимірюється, порівнюється з мірою безпосередньо або за допомогою вимірювальних пристроїв (приборів), градуйованих у тих одиницях, в яких вимірюється дана величина.

ВИЗНАЧЕННЯ 4.7. **Вимірювання** назвемо **непрямими**, якщо безпосередньо вимірюється не сама величина, а інша, яка пов'язана з нею функціонально.

Числове значення величини, яка підлягає вимірюванню, при непрямому вимірюванні отримується шляхом відповідних розрахунків на основі залежностей, які існують між величинами і виражених у математичній формі. Непрямі вимірювання застосовуються в тому випадку, коли прямі вимірювання утруднені або неможливі.

Результат вимірювання деякої частинної реалізації економічної величини X може бути наведений у вигляді добутку двох множників – $X = \{X\} \cdot [X]$,

де $[X]$ - величина вимірювань економічної величини X , $\{X\}$ - числове значення величини, що вимірюється, якщо вона виражена в одиницях $[X]$.

Числове значення – це абстрактне число, яке дорівнює відношенню величини, що вимірюється, до одиниці її вимірювання.

Одиниця вимірювання економічної величини $[X]$, як частинна реалізація цієї величини X , також може бути виражена у вигляді добутку множників $\{X\}$ і $[X]$. При цьому числове значення $\{X\}$ одиниці вимірювання дорівнює одиниці.

Одиниці вимірювання $[x_1], [x_2], \dots, [x_n]$ однієї і тієї самої економічної величини X , тобто, однорідні величини, відрізняються одна від одної розміром. Наприклад, \$1 в 10 раз менше ніж \$10, 1 кг бананів в 1000 раз більше ніж їх 1 г. Таки одиниці будемо називати **-частковими одиницями економічної величини**.

ВИЗНАЧЕННЯ 4.8. *Розміром одиниці вимірювань* назвемо кількість економічної величини, яка міститься в одиниці вимірювань.

Взагалі X одиницею $[x_1]$ отримане числове значення $\{x_1\}$, а при вимірюванні одиницею $[x_2]$ отримане числове значення $\{x_2\}$, то

$$\frac{\{x_1\}}{\{x_2\}} = \frac{[x_2]}{[x_1]},$$

тобто $\{x_1\} \cdot [x_1] = \text{const}$.

ВИЗНАЧЕННЯ 4.9. *Економічним експериментом* назвемо сукупність вимірювань, в яких всі впливи на економічну систему, що досліджується, піддаються обліку (або можуть бути математично оцінені) з використанням будь-яких доступних засобів вимірювання.

Основна властивість економічного експерименту – відтворюваність в умовах різних варіацій параметрів середовища, в якому розглядається функціонування економічної системи, що досліджується, заради якої проводиться експеримент.

З урахуванням можливостей, які відкриваються, що надає системний підхід в економіці, застосування математичного моделювання забезпечує вирішення цілого комплексу економічних завдань і апробацію їх наслідків на основі проведення економічних експериментів над моделлю, а не над економічною системою, яка реально існує, оскільки це не завжди можна здійснити, часто економічно не вигідно, а інколи й небезпечно, бо може приховувати в собі великі соціально-економічні потрясіння.

4.2 Про регресію взагалі

Проблема невизначеності є однією з головних перепон до застосування системного підходу в усіх сферах економічного аналізу – світової економіки в цілому та її регіонів (Північна Америка, Латинська Америка, Європа, Азія, Японія, Азія, Африка - нафтопереробні країни, Пустельна Африка, Тропічна Африка тощо), національної економіки, окремих її регіонів, секторів, фірм і підприємств з різною формою власності. Тому її усунення є першочерговим завданням кожного дослідника будь-якої реальної економічної системи, хоча це

дуже складне завдання, яке не завжди вирішується. Системні спеціалісти корпорації RAND навіть назвали проблему невизначеності "звіром", з яким ми не можемо впоратися протягом багатьох років.

Виявлення різних випадкових залежностей (звичайно, йдеться не про всі, а тільки про якісь їх складові частини, оскільки в реальних складних економічних системах урахувати всю множину таких факторів неможливо) та їх вивчення на предмет взаємозалежності і взаємозв'язку знімає деякою мірою проблему невизначеності досліджуваної економічної моделі.

У цій темі розглядається один із можливих підходів до вирішення проблеми усунення невизначеності, який має назву – регресійний аналіз у сферах економічного аналізу.

Введемо наступне визначення.

ВИЗНАЧЕННЯ 4.10. *Регресією* назвемо однобічну стохастичну залежність однієї випадкової змінної від другої або декількох інших випадкових змінних.

Однобічна стохастична залежність виражається за допомогою функції, яка на відміну від суворої математичної залежності, називається *функцією регресії*. Принциповою відмінною між суворою функціональною залежністю і функцією регресії є те, що у першому випадку аргумент (незалежна змінна) повністю визначає значення функції, і ця функція зворотна (наприклад, функція $y = 8x$ і $x = y/8$). У другому випадку (для функції регресії) цього сказати неможливо. Виходячи з вищевикладеного, встановлюють регресію y на x або x на y залежно від того, досліджують стохастичну залежність y на x або x на y , відповідно.

Слід зазначити, що функція регресії тільки формально встановлює відповідність між змінними, хоча вони можуть і не перебувати в причинно-наслідкових відношеннях. У цьому випадку можуть виникнути так *звані нонсенс-регресії (хибні, абсурдні)*, які не мають не тільки будь-якої практичної цінності, але і сенсу взагалі. Прикладом такої нонсенс-регресії може слугувати зв'язок між розоренням банку "АПЕКС" у 1998 році і кількістю блондинок, які народилися у Києві у 1997 році. Тому, при складанні рівнянь регресії треба завжди виходити з реальних задач, які мають конкретний практичний додаток і несуть змістовне навантаження.

4.3 Основи регресійного аналізу. Його задачі.

Залежно від числа змінних, що розглядаються, розрізняють просту (парну) і множинну регресії.

ВИЗНАЧЕННЯ 4.11. *Регресією* між двома змінними, назвемо простою (парною) регресією і будемо записувати у вигляді

$$\hat{Y} = f(x). \quad (4.1)$$

Змінну \hat{Y} називають по-різному - залежна, результативна, регресанд, тобто, змінна, що підлягає поясненню, а також, завбачувана змінна, або цільова функція. Змінну x називають незалежною, пояснюючою або передбачуваною змінною, регресором або факторною ознакою.

ВИЗНАЧЕННЯ 4.12. Регресією між залежною змінною та декількома причиново зумовленими пояснюючими змінними x_1, x_2, \dots, x_n назовемо **множинною регресією** та будемо записувати у вигляді

$$y = f(x_1, x_2, \dots, x_n) \quad (4.2)$$

Ковпачок над y означає оцінку залежної змінної, отриманої з виразу (4.1) чи (4.2) при деяких усереднених значеннях. У подальшому будемо називати результативною змінною або регресією.

Знаходження функціональної залежності (4.1) та (4.2) є одним з основних завдань регресійного аналізу.

Знаходження залежностей для (4.1) та (4.2) виконується за емпіричними даними, отриманими в результаті експерименту чи зібраними при вивченні реальної економічної системи (незалежно від її виду). Але ці дані включають в себе складові і випадкові другорядні причини та ті, що їх спричинили. Тому отримані емпіричним шляхом дані носять випадковий характер, що не завжди збігаються один з одним при повторенні економічного експерименту. У зв'язку з цим не враховувати ці другорядні фактори ("шум") при виявленні істинної залежності між причиною та наслідком просто було б рівнозначно "прикидці на око" чи, що те саме, отриманню необ'єктивної картини реального економічного процесу. Таким чином, функція регресії чисельно оцінює усереднену залежність між досліджуваними змінними – пояснюючими та результативними, тобто таку залежність між наслідком та причиною, за якої інші фактори не змінюються і тим самим не накладають "шум" на основну залежність.

ВИЗНАЧЕННЯ 4.13. Випадкову змінну u назовемо **збурюваною (збуренням)**, якщо

$$u = y - \hat{y} \quad (4.3)$$

де \hat{y} – функція регресії, а y – реальні спостережувані значення процесу, якій досліджується.

Таким чином, якщо казати на змістовному рівні, u вміщує вплив неврахованих факторів - змінних, помилок спостережень та різноманітних перешкод, спричинених неточністю їх отримання через різні об'єктивні причини, наприклад, погрішності вимірювального пристрою, її значення змінюються для кожного спостереження y . Із (4.3) з врахуванням (4.2) отримаємо, що

$$y = f(x_1, x_2, \dots, x_n) + u \quad (4.4)$$

Із (4.4) випливає, що, змінна y – випадкова величина при заданих пояснюючих змінних x_1, x_2, \dots, x_n , тому що їй неможна поставити у відповідність тільки одне значення через випадковість змінної u .

Статистичні залежності будуть тим точніше відбивати причиново-наслідковий зв'язок між змінними, чим більше даних отримано при багаторазовому повторенні експерименту (спостережень) і чим краще ("чисто") ми поставили експеримент або спостерігали, збирали та здійснювали попередню обробку даних про досліджуваний процес.

Але, не зважаючи навіть на ту обставину, що нами вже отримано досить представницьку вибірку залежної та пояснюючих змінних (за умови, що ми "чисто" поставили експеримент), значення u безпосередньо отримати неможли-

во, тому що воно являє собою інтегровану випадкову змінну різноманітних, не завжди враховуваних випадкових сторонніх факторів.

Це відбувається з різні причини, основними з яких є такі [22].

1. Будь-яка математична модель є спрощеним описом реальної дійсності. Тому співвідношення (4.4) також є спрощеним (існують інші фактори, які в ньому не враховані). Існує багато факторів, які можна було б включити в модель, але їх не можна виміряти, наприклад, це можуть бути психологічні фактори, менталітет населення і т. ін. Більше того, є багато факторів, які є суттєвими, але через відсутність досвіду вони в моделі не можуть бути враховані. Немає впевненості, що входить у дану сукупність, а що ні. Це призводить до того, що y і відрізняються в більшості випадків один від одного. Для їх вирівнювання треба додавати випадкову величину (від'ємну або додатну).

2. Часто співвідношення (4.4) об'єднує між собою агреговані змінні. Наприклад, сумарний прибуток торговельного підприємства за місяць складається з його прибутків за кожний день цього місяця. Однак, усереднене значення сумарного прибутку лише до деякої міри збігається з конкретним значення прибутку за кожен день. Тому функція, яка відображає взаємозв'язок поміж загальним прибутком та витратами підприємства, лише апроксимує цей прибуток та витрати за кожен день цього місяця.

3. Залежність між залежною змінною y та фактором x структурно визначена неправильно. Наприклад, якщо розглядаються часові ряди, то може бути, що y залежить від x : не сьогоднішнього часу, а попереднього. Прикладом такої залежності, може бути залежність ВВП від капіталовкладень держави та зовнішніх інвестицій у промисловість по регіонах.

4. Залежність між y та x : функціонально визначена неправильно. Наприклад, ми вибрали цю залежність квадратичною, а вона кубічна.

5. Виміри, щодо змінних, які використовує модель (4.4) здійснені неправильно, або не зовсім точно. Наприклад, ми за обсягом продукції, яка вироблена на підприємстві, робимо висновок, як воно працює. Але ми при цьому не враховуємо, що більшість продукції була використана для бартеру з іншими підприємствами або розкрадена.

Сумарна сукупність усіх цих факторів і є тією випадковою змінною, яку ми не можемо врахувати у моделі. У зв'язку з вищевикладеним, змінній u відводиться особлива роль і на відміну від інших випадкових змінних, вона називається **латентною** змінною.

Надалі числові оцінки латентної змінної будемо позначати через u і називати **залишками**.

Зараз ми в змозі сформулювати **основну задачу регресійного аналізу та етапи її вирішення**.

ЗАДАЧА 4.1. Встановити вид функції, яка б найкраще описувала усереднену масову течію економічного процесу.

1-й етап (попередній): проведення обґрунтованого якісного економічного аналізу досліджуваного процесу; попередня обробка статистичної інформації про цей процес;

II-й етап: на основі проведеного аналізу сформулювати гіпотезу про вид функції, яка описує цей процес;

III-й етап: прийняття гіпотези або відмова від неї шляхом статистичної перевірки за емпіричними даними;

IV-й етап: визначення кінцевої функції регресії, оцінка невідомих значень залежної змінної, отриманих теоретично обґрунтованих і статистично надійних різноманітних прогнозів, щодо поведінки досліджуваного процесу.

Перейдемо тепер безпосередньо до класифікації форм регресії.

4.4 Класифікація форм регресії

ВИЗНАЧЕННЯ 4.14. Лінійною регресією назвемо регресію виду:

$$= b_0 + b_1x_1 + b_2x_2 + \dots + b_mx_m \quad (4.5)$$

де x_1, x_2, \dots, x_m – пояснювальні змінні; b_0, b_1, \dots, b_m – оцінки відповідних параметрів регресії $\beta_0, \beta_1, \dots, \beta_m$ – в лінійній моделі

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_mx_m + u$$

тобто $b_0 = \widehat{\beta}_0, b_1 = \widehat{\beta}_1, b_2 = \widehat{\beta}_2, \dots, b_m = \widehat{\beta}_m.$

У випадку парної (простої) регресії вираз (4.5) набуває вигляду:

$$\widehat{y} = b_0 + b_1x,$$

де $b_0 = \widehat{\beta}_0, b_1 = \widehat{\beta}_1, \dots$

Одним з можливих методів отримання оцінок (знаходження числових значень коефіцієнтів b_0, b_1, \dots, b_m у регресійній моделі) є *метод найменших квадратів* (МНК), його розглянемо в наступному розділі.

ВИЗНАЧЕННЯ 4.15. *Нелінійною регресією* назвемо регресію, в якій вид функції (4.1) або (4.2) має нелінійний характер.

Наприклад, нелінійною регресією можуть виступати поліноміальні регресії виду

$$\widehat{y} = b_0 + b_1x + b_2x^2 + \dots + b_mx^m \quad (4.6)$$

або показникові регресії виду $\widehat{y} = ab^x$ і т. ін.

Гіпотеза про вибір виду функції регресії є основним завданням другого етапу розв'язання задачі 4.1.

Оскільки для більшості економічних процесів вид функції регресії носить складний нелінійний характер, розглянемо такі регресії більш докладно.

ВИЗНАЧЕННЯ 4.16. Нелінійну регресією назвемо *квазілінійною*, якщо вона може бути представлена у вигляді:

$$\widehat{y} = b_0 + b_1F_1(x) + b_2F_2(x) + \dots + b_mF_m(x), \dots \dots \dots (4.7)$$

де $F_1(x), F_2(x), \dots, F_m(x)$ – нелінійні функції від пояснювальних змінних x .

До квазілінійних регресій безпосередньо також можна застосовувати МНК, представивши її у виді лінійної множинної регресії.

Наприклад, вираз (4.6) зводиться до виразу (4.5) за допомогою заміни:

$$X = x_1, x^2 = x_2, x^3 = x_3, \dots, x^m = x_m;$$

$$b_0 = b_0^*, b_1 = b_1^*, b_2 = b_2^*, \dots, b_m = b_m^*.$$

У результаті такої заміни, отримуємо:

$$\hat{Y} = b_0^* + b_1^* x_1 + \dots + b_m^* x_m$$

що повністю відповідає виразу (4.5).

ВИЗНАЧЕННЯ 4.17. Нелінійну регресію *назвемо суттєво нелінійною*, якщо вона нелінійна за оцінювальними параметрами.

Прикладами суттєво нелінійних регресій можуть бути:

1) степенева функція $\hat{Y} = ax^b$;

2) показникова функція $\hat{Y} = ab^x$;

3) логістична функція $\hat{Y} = \frac{a}{1+be^{-cx}}$ або $\hat{Y} = \frac{a}{1+e^{b-cx}}$ і т.ін.

Незважаючи на те, що *суттєво нелінійні функції зустрічаються досить часто при вивченні економічних процесів, застосування до них безпосередньо МНК неприпустимо.*

Для оцінки параметрів суттєво нелінійних регресій застосовують ітераційні методи або ж вдаються до їх апроксимації.

Але найбільш розповсюджений метод полягає в лінійному перетворенні функції суттєво нелінійної регресії, що дає можливість застосовувати до них МНК.

Наприклад, для степеневої функції це можливо зробити таким чином. Прологарифмувавши обидві частини по основі e степеневу функцію $=ax^b$, отримаємо:

$$\ln = \ln a + b \ln x$$

Вводячи нові змінні $\hat{Z} = \ln y, b_0 = \ln a, x^* = \ln x$, ,отримаємо: $\hat{Z} = b_0 + bx^*$

З суттєво нелінійних регресій значний економічний інтерес становлять виробничі функції.

ВИЗНАЧЕННЯ 4.18. *Виробничою функцією* назвемо функцію, яка описує кількісну залежність причинно-наслідкових відносин між результатом економічного процесу і умовами його , хоча б частина з яких керована.

Вихідну інформацію для виробничої функції отримують у результаті збору статистичних даних або експериментальним шляхом, коли дослідник контролює хід дослідження і визначає, які величини повинні бути змінними. Припустимий функціональний зв'язок повинен відображати сутність вивченого економічного процесу.

Першою *виробничою функцією* була *степенева функція Кобба-Дугласа*, яка в загальному випадку має вигляд

$$\hat{Y} = b_0 x_1^{b_1} x_2^{b_2} \dots x_m^{b_m} \quad (4.8)$$

де \hat{Y} – випуск продукції, або національний дохід і т. ін.;– x_1, x_2, \dots, x_m впливові

фактори; b_0 – нормований множник; b_1, b_2, \dots, b_m – коефіцієнти еластичності.

Для приведення виразу (4.8) до лінеаризованого вигляду прологарифмуємо частину цієї рівності за основою e . Отримаємо

$$\ln \hat{y} = \ln b_0 + b_1 \ln x_1 + b_2 \ln x_2 + \dots + b_m \ln x_m.$$

Вводячи нові змінні:

$$\hat{z} = \ln \hat{y}; \quad a_0 = \ln b_0; \quad \xi_1 = \ln x_1; \quad \xi_2 = \ln x_2; \quad \dots; \quad \xi_m = \ln x_m$$

приходимо до лінійного виду регресії (4.5).

Аналогічно можна зробити і з функцією Кобба-Дугласа-Тінбергена, яка має вигляд:

$$\hat{y} = b_0 x_1^{b_1} x_2^{b_2} \dots x_k^{b_k} e^{\gamma t}$$

Ця виробнича функція відображає вплив науково-технічного прогресу на економічні процеси, де t – час, а γ – показник його темпу розвитку.

У табл. 4.1 зведені до лінійних $\hat{y} = b_0' + b_1' x^*$ багато нелінійних залежностей з двома параметрами b_0 і b_1 , за допомогою перетворення $x \rightarrow x^1$ і $y \rightarrow y^1$.

Після приведення до лінійної регресії, отримаємо за допомогою МНК значення b_0^1 і b_1^1 , які після перетворення b_0^1 у b_0 і b_1^1 у b_1 дають потрібні оцінки параметрів b_0 і b_1 нелінійної залежності (в деяких випадках легко допускається така процедура на узагальнення з m -параметрами).

І, наприкінці, зазначимо ще один важливий клас лінійно-гомотетичних функцій: CES (Constant Elasticity of Substitution) – функцій з постійною еластичністю заміщення та функцій зі змінюваною з часом еластичністю заміни VES (Variable Elasticity of Substitution). Цей клас функцій дає змогу моделювати зміни еластичності заміни факторів виробництва із зміною рівня випуску:

$$\hat{y} = b_0 [b_2 x_2^{-p} + (1 - b_2) x_1^{-p}]^{-1/p}$$

$$\hat{y} = b_0 [b_2 x_2^{-p} + (1 - b_2) q \left(\frac{x_2}{x_1}\right)^{-s(1+p)} x_1^{-p}]^{-1/p}$$

де b_1, b_2 – коефіцієнти еластичності випуску за першим та другим фактором ($b_1 = 1 - b_2$ – умова лінійної гомотетичності); p, q і s часто встановлюються виходячи з економічних вимог (чи визначаються виходячи із диференціальних рівнянь цих функцій).

Таблиця 4.1

№	Функція	x^1	y^1	b_1	b_0
1	$b_1 x + b_0$	x	y	b_1^1	b_0^1
2	$1/(b_1 x + b_0)$	x	$1/y$	b_1^1	b_0^1
3	$b_0 + b_1/x$	$1/x$	y	b_1^1	b_0^1
4	$x/(b_1 x + b_0)$	x	x/y	b_1^1	b_0^1
5	$b_0 b_1^x$	x	$\lg y$	$10^{b_1^1}$	$10^{b_0^1}$
6	$b_0 e^{b_1 x}$	x	$\ln y$	b_1^1	$e^{b_0^1}$

7	$b_0 10^{b_1 x}$	x	$\lg y$	b_1^1	$10^{b_0^1}$
8	$1/(b_1 e^x + b_0)$	e^{-x}	$1/y$	b_1^1	b_0^1
9	$b_0 x^{b_1}$	$\lg x$	$\lg y$	b_1^1	$10^{b_0^1}$
10	$b_0 + b_1$	$\lg x$	y	b_1^1	b_0^1
11	$b_0 + b_1$	$\lg y$	y	b_1^1	b_0^1
12	$b_0/(b_1 + x)$	x	$1/y$	b_0^1 / b_1^1	$1 / b_1^1$

Продовження таблиці 4.1

13	$b_0 x/(b_1 + x)$	$1/x$	$1/y$	b_0^1 / b_1^1	$1 / b_1^1$
14	$b_0 e^{b_1/x}$	$1/x$	$\ln x$	b_1^1	$e^{b_0^1}$
15	$b_0 10^{b_1/x}$	$1/x$	$\lg x$	b_1^1	$10^{b_0^1}$
16	$b_0 + b_1 x^n$	x^n	y	b_1^1	b_0^1

4.5 Діаграми розсіювання та побудова регресійних функцій

Для виявлення причинно-наслідкових зв'язків між змінними будують діаграми розсіювання, які являють собою наочну форму такого взаємозв'язку. Для цього, використовуючи декартову систему координат, результат кожного спостереження відображають точкою на ній. Сукупність усіх точок називають "хмарою", яка визначає залежність двох змінних (по осі абсцис x – пояснюючою змінною, по осі ординат y – результативною). Іноді для діаграми розсіювання в літературі зі статистики використовуються терміни: "поле кореляції" чи "поле розсіювання". По "хмарі" можна судити про ступінь тісноти зв'язку між змінними та про вид регресії – додатної чи від'ємної (рис. 4.1-4.4).

За побудованою "хмарою" можна судити про вид функції регресії і висунути гіпотезу про її математичний запис (II-й і III-й етапи регресійного аналізу). Процес знаходження функції регресії назвемо **вирівнюванням** окремих значень залежної змінної. Побудовою самої функції регресії та встановленням впливу пояснюючих змінних на залежну змінну закінчується четвертий етап регресійного аналізу.

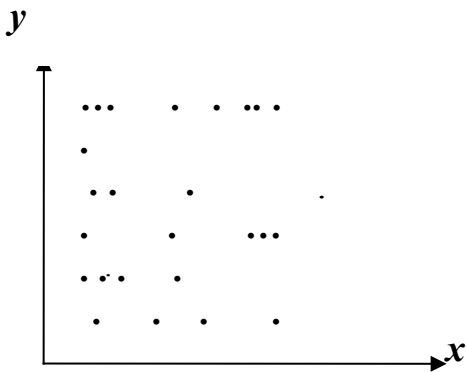


Рисунок 4.1 – Діаграма розсіювання у випадку відсутності зв'язків

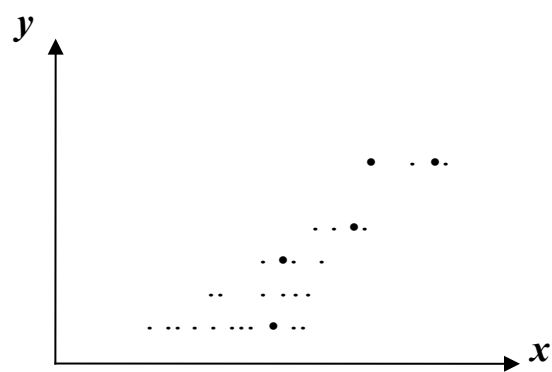
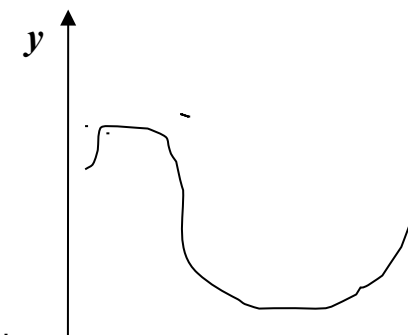
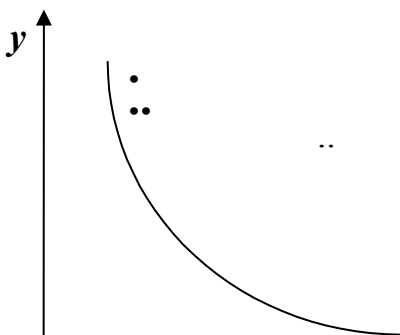


Рисунок 4.2 – Додатна регресія



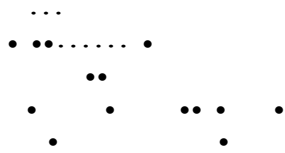


Рисунок 4.3 – Від’ємна регресія

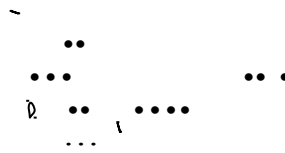


Рисунок 4.4 – Комбінована форма регресії

Незважаючи на всі переваги, діаграми розсіювання мають такі суттєві недоліки:

- 1) побудовані за ними регресії носять суб'єктивний характер (є приблизними) і тому є наближеними і не зовсім точно відображають характер змін емпіричних даних;
- 2) за великої кількості змінних (більше трьох) побудувати геометрично діаграму розсіювання неможливо.

Але, незважаючи на зазначені недоліки, висування гіпотези "на око" про вид функції регресії (для однієї та двох пояснюючих змінних) та її погодження з теоретично побудованою лінією регресії часто дає можливість досліднику судити про адекватність вибраної математичної моделі реальному економічному процесу, тобто погодження цієї моделі з реальними економічними умовами протікання процесу.

Наприкінці розділу зупинимося більш докладно на парній лінійній регресії

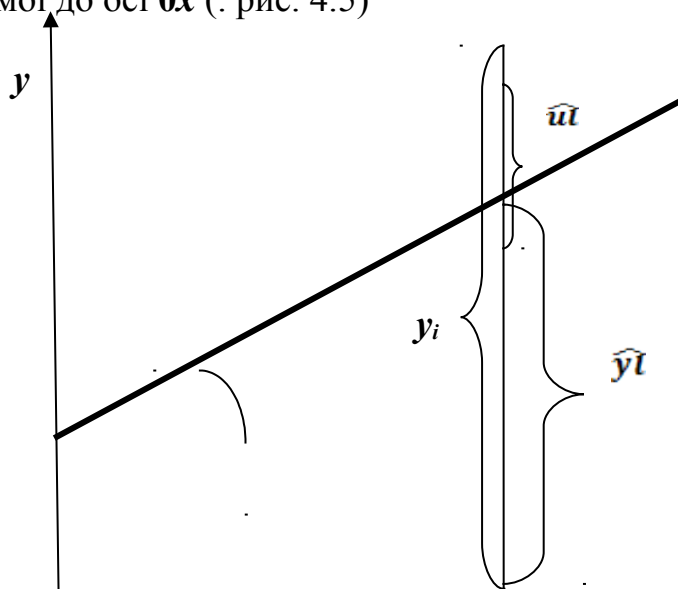
$$\hat{y} = b_0 + b_1x,$$

яку іноді записують у вигляді

$$\hat{y} = b_0x_0 + b_1x,$$

де x_0 – фіктивна змінна; b_0 – постійна регресії, що визначає точку перетину прямої з віссю ординат; b_1 – коефіцієнт регресії.

У рівнянні для b_0 економічна інтерпретація часто не тільки утруднена, але й взагалі неможлива, оскільки b_0 є середнім значенням y в точці $x = 0$. Вона лише виконує в рівнянні функцію вирівнювання. Рівняння регресії інтерпретуємо тільки в області накопичення точок. Постійна регресії b_0 має розмірність y , а b_1 – розмірність, яка дорівнює співвідношенню розмірності залежної змінної до розмірності пояснюючої змінної, тобто коефіцієнт регресії характеризує нахил прямої до осі $0x$ (. рис. 4.5)



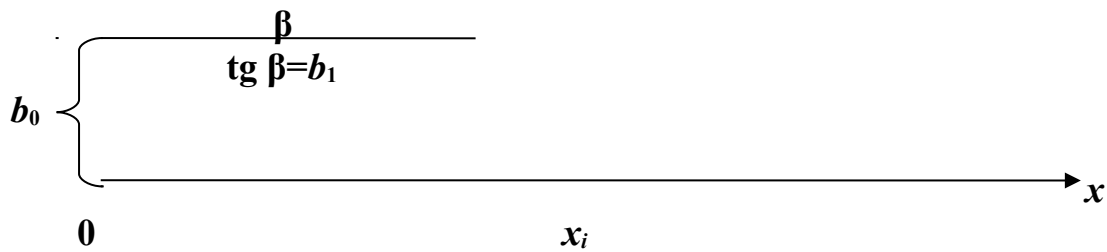


Рисунок – 4.5

З рис. 4.5 випливає економічна інтерпретація \hat{y} – значення якого вказують середнє значення залежної змінної y_i при заданому x_i пояснюючою змінною x з припущенням, що єдиною причиною зміни змінної y є змінна x , а випадкова збурювана змінна $u = 0$. Більш формалізовано, \hat{y} — це математичне сподівання залежної змінної,); при заданих значеннях пояснюючих змінних x .

Відхилення y_i від \hat{y}_i зумовлено різноманітними не підлягаючими суворому обліку і контролю причинами. Тому числове значення u_i визначається як

$$\hat{u}_i = y_i - \hat{y}, i=1,2,\dots,m.$$

Зрозуміло, що чим менше \hat{u}_i для всіх $i=1,2,\dots, m$, тим краще вибрана пряма. Для цього необхідно розрахувати правильно коефіцієнти b_0 і b_1 , а в загальному випадку b_0, b_1, b_2,\dots, b_m . Їх вибір припускає знання особливих процедур, одна із яких носить назву методу найменших квадратів (МНК).

РОЗДІЛ 5

МЕТОД НАЙМЕНШИХ КВАДАТІВ. ЙОГО СИЛЬНІ ТА СЛАБКІ СТОРОНИ. УЗАГАЛЬНЕНИЙ МНК

5.1 Метод найменших квадратів

Усереднення несумісних розв'язання надлишкової системи рівнянь може бути здійснено різними способами, наприклад, "на око", методом медіанних центрів, методом максимальної правдоподібності і т. ін. Однак одним з найбільш потужних методів є розроблений в 1795-1805 рр. французьким математиком Андрієн М. Лежандром (1752-1833) і німецьким математиком Карлом Фрідрихом Гауссом (1777-1855) метод регресійного аналізу, чи, як часто його називають, метод найменших квадратів (МНК).

Зазначимо, що застосування МНК, якому близько 200 років, гальмувалося через труднощі, пов'язані зі значними обсягами обчислень, і тільки в середині ХХ століття цей метод набув широкого розповсюдження у зв'язку з появою ЕОМ.

На сьогодні існує досить багато різноманітних програмних продуктів, які дають можливість реалізувати на ЕОМ цей метод. Тому їх застосування настільки різноманітне: статистика, економетрія, оцінка похибок вимірювань тощо.

Для розуміння цього методу розглянемо простий приклад з використанням МНК.

Припустимо, після попереднього аналізу досліджуваного економічного процесу на основі статистичних даних про його поведінку ми вирішили, що математична модель цього процесу має вигляд

$$\hat{y} = \beta_0 + \beta_1 x + \beta_2 x^2.$$

Знайдемо "найкращі" значення параметрів β_0 , β_1 , β_2 вибраної нами моделі за допомогою МНК. Це нелінійна (квазілінійна) модель – є x^2 . Зведемо її до лінійної. Для цього покладемо $x_1 = x$ і $x_2 = x^2$. У результаті отримаємо лінійну модель (для лінійних моделей існують потужні алгебраїчні засоби для їх дослідження), яка має вигляд $\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$

Між розрахованими за моделлю значеннями \hat{y}_i та експериментальними відрахуваннями y_i будуть, як це підкреслювалося у розділі 4, спостерігатись відхилення. Позначимо їх як

$$\hat{u}_i = y_i - \hat{y}_i, i = 1, 2, \dots, n. \quad (5.1)$$

Далі будемо називати їх *залишками*. Вони включають вплив неврахованих факторів - змінних, випадкових перешкод та помилок спостереження тощо (більш докладно це питання висвітлено у розділі 4). Їхні значення можуть змінюватись від одного спостереження до іншого.

МНК дає змогу знайти такі значення (оцінки) b_0, b_1, b_2 вихідних параметрів $\beta_0, \beta_1, \beta_2$ моделі, для якої (це пов'язано з тим, що необхідний критерій для підбору коефіцієнтів моделі повинен враховувати ту обставину, за якої одержана функція регресії (якщо її представити на графіку) буде якомога ближче проходити між експериментально отриманими змінними)

$$U = \sum_{i=1}^n U_i^2 \rightarrow \min. \quad (5.2)$$

У подальшому таку модель будемо записувати у вигляді

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 \quad (5.3)$$

Беручи частинні похідні по b_0, b_1, b_2 , і прирівнюючи їх до нуля, одержимо систему з трьох рівнянь з трьома невідомими b_0, b_1, b_2 , вирішенням якої і є шукані значення оцінок. У нашому випадку, оскільки

$$U = \sum_{i=1}^n U_i^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_{i1} - b_2 x_{i2})^2 \Rightarrow$$

$$\left\{ \begin{array}{l} \frac{\partial U}{\partial b_0} = -2 \sum_{i=1}^n (y_i - b_1 x_{i1} - b_2 x_{i2}) = 0, \\ \frac{\partial U}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_1 x_{i1} - b_2 x_{i2}) x_{i1} = 0, \\ \frac{\partial U}{\partial b_2} = -2 \sum_{i=1}^n (y_i - b_1 x_{i1} - b_2 x_{i2}) x_{i2} = 0. \end{array} \right. \quad (5.4)$$

$$\left\{ \begin{array}{l} \frac{\partial U}{\partial b_1} = -2 \sum_{i=1}^n (y_i - b_1 x_{i1} - b_2 x_{i2}) x_{i1} = 0, \\ \frac{\partial U}{\partial b_2} = -2 \sum_{i=1}^n (y_i - b_1 x_{i1} - b_2 x_{i2}) x_{i2} = 0. \end{array} \right. \quad (5.5)$$

З виразів завдань (5.4) – (5.6) отримаємо систему нормальних (лінійних) рівнянь:

$$\left\{ \begin{array}{l} b_0 n + b_1 \sum_{i=1}^n x_{i1} + b_2 \sum_{i=1}^n x_{i2} = \sum_{i=1}^n y_i \\ b_0 \sum_{i=1}^n x_{i1} + b_1 \sum_{i=1}^n x_{i1}^2 + b_2 \sum_{i=1}^n x_{i1} x_{i2} = \sum_{i=1}^n x_{i1} y_i \\ b_0 \sum_{i=1}^n x_{i2} + b_1 \sum_{i=1}^n x_{i1} x_{i2} + b_2 \sum_{i=1}^n x_{i2}^2 = \sum_{i=1}^n x_{i2} y_i. \end{array} \right. \quad (5.7)$$

Застосовуючи для розв'язання системи (5.7) один з відомих методів щодо розв'язання системи лінійних рівнянь (наприклад, метод Гаусса, або правило Крамера - метод визначників, або матричний метод) знаходимо невідомі коефіцієнти b_0, b_1, b_2 :

$$b_2 = \frac{\sum_{i=1}^n (x_{i2} - \bar{x}_2)(y_i - \bar{y}) \cdot \sum_{i=1}^n (x_{i1} - \bar{x}_1)^2}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 \sum_{i=1}^n (x_{i2} - \bar{x}_2)^2 - \left(\sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \right)^2}$$

$$b_1 = \frac{\frac{\sum_{i=1}^n (x_{i1} - \bar{x}_1)(y_i - \bar{y}) \cdot \sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2)}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 \sum_{i=1}^n (x_{i2} - \bar{x}_2)^2 - \left(\sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \right)^2}, \frac{\sum_{i=1}^n (x_{i1} - \bar{x}_1)(y_i - \bar{y}) \cdot \sum_{i=1}^n (x_{i2} - \bar{x}_2)^2}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 \sum_{i=1}^n (x_{i2} - \bar{x}_2)^2 - \left(\sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \right)^2}}{\frac{\sum_{i=1}^n (x_{i2} - \bar{x}_2)(y_i - \bar{y}) \cdot \sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2)}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 \sum_{i=1}^n (x_{i2} - \bar{x}_2)^2 - \left(\sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \right)^2}, \frac{\sum_{i=1}^n (x_{i2} - \bar{x}_2)(y_i - \bar{y}) \cdot \sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2)}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 \sum_{i=1}^n (x_{i2} - \bar{x}_2)^2 - \left(\sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \right)^2}},$$

$$b_0 = \bar{y} - b_1 \bar{x}_1 - b_2 \bar{x}_2.$$

Підставляючи їх значення в (5.3) отримуємо лінію, яку називають *лінією регресії*. Коефіцієнти b_1 і b_2 називаються коефіцієнтами регресії \hat{y} по x_1 , і \hat{y} по x_2 відповідно.

Тут ми не наводимо доведення (щоб не перевантажувати основний зміст тексту) того, що знайдена точка (b_0, b_1, b_2) є точкою, яка задовольняє умові (5.2). Але, у математичному аналізі є теорема, за допомогою якої можна визначити достатні умови екстремуму функції, у нашому випадку – мінімуму функції $U(b_0, b_1, b_2)$.

Така сама методика і міркування застосовуються у випадку знаходження m коефіцієнтів регресії $b_0, b_1, b_2, \dots, b_m$ (в подальшому в цьому розділі ми розглянемо, як ця задача вирішується за допомогою використання теорії матриць, довідковий матеріал до якої наведено у Д. 1).

Розглянемо тепер властивості МНК. Для наочності міркувань розглянемо більш просту лінійну регресійну модель виду

$$\hat{y} = b_0 + b_1 x \quad . \quad (5.8)$$

Застосовуючи МНК отримаємо (пропонуємо Вам переконатися в цьому самостійно):

$$b_0 = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad (5.9)$$

$$b_1 = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad (5.10)$$

які є розв'язання системи

$$\begin{cases} b_0 n + b_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \end{cases} \quad (5.11)$$

$$\begin{cases} b_0 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i \end{cases} \quad (5.12)$$

Поділивши (5.11) на n отримаємо:

$$b_0 + b_1 \frac{\sum_{i=1}^n x_i}{n} = \frac{\sum_{i=1}^n y_i}{n}$$

або

$$b_0 + b_1 \bar{x} = \bar{y}$$

де \bar{x} – середні арифметичні спостережуваних значень x_1, x_2, \dots, x_n та \bar{y} – середні арифметичні спостережуваних значень y_1, y_2, \dots, y_n відповідно.

Таким чином, перше рівняння системи потребує, щоб лінійна регресія проходила через точку (\bar{x}, \bar{y}) , тобто через центр ваги поля експериментальних точок, або інакше (5.11) визначає середній рівень шуканої лінії регресії. Це наочно проявляється, якщо перенести початок координат в точку \bar{x} , оскільки при $\bar{x} = 0$ маємо, що $\bar{y} = b_0$.

При $\bar{x} = 0$ з рівняння (5.12) отримаємо

$$(5.13)$$

Друге рівняння (5.12) визначає коефіцієнт b_1 , як нахил шуканої прямої відносно осі ОХ (див. розд. 4), а тому і (5.13) має той самий сенс.

Інтерпретуючи коефіцієнти моделі треба мати на увазі наступне.

По-перше, коефіцієнт b_0 з рівняння регресії, формально показує рівень значення y при $x = 0$. Інколи це має сенс, особливо, якщо вибіркові значення x_i достатньо близькі від цієї точки. Якщо це не так, то інтерпретація такої формальної залежності може призвести до неправильних висновків.

По-друге. Тому, що обчислені параметри моделі – є оцінками, то інтерпретація всієї моделі теж має сенс оцінки.

По-третє. Рівняння регресії відображає тільки загальну тенденцію для вибірки. При цьому кожне окреме спостереження підпадає під вплив випадковості.

По-четверте. Треба бути дуже уважним при побудові (специфікації) аналітичної форми рівняння регресії, оскільки при неправильній специфікації її інтерпретація досліджуваного економічного процесу може призвести до вкрай заплутаних висновків щодо його поведінки і розвитку у майбутньому.

Вираз (5.10) можна записати і в іншому вигляді, а саме через коефіцієнт кореляційного моменту [вибіркова коваріація] $K(X, Y)$ [Cov(X, Y)]. Він характеризує не лише зв'язок між випадковими величинами X та Y , але й їх розсіюван-

ня) та дисперсію

$$b_1 = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} = \frac{\frac{\sum_{i=1}^n x_i y_i}{n} - \bar{x} \bar{y}}{\frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2} = \frac{K(X, Y)}{\sigma^2(X)},$$

де

$$K(X, Y) = \text{Cov}(X, Y) = \frac{M(X, Y)}{n} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n} = \left[\frac{1}{n} \sum_{i=1}^n x_i y_i \right] - \bar{x} \bar{y}.$$

Другий запис для обчислення b_1 можна записати в такому вигляді:

$$b_1 = \frac{K(X, Y)}{\sigma^2(X)} = \frac{K(X, Y)}{\sigma(X)\sigma(Y)} \cdot \frac{\sigma(Y)}{\sigma(X)} = r(X, Y) \cdot \frac{\sigma(Y)}{\sigma(X)},$$

де

$$r(X, Y) = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \sqrt{n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2}}$$

це коефіцієнт кореляції між змінними X та Y .

Наведемо властивості для виразу $K(X, Y)$ (доказ їх впливає з визначення $K(X, Y)$). Радимо Вам це зробити самостійно)-

1. $K(X, Y+W) = K(X, Y) + K(X, W)$,
2. $K(X, aY) = aK(X, Y)$,
3. $K(X, a) = 0$, якщо $a = \text{const}$;
4. $K(X, X) = \sigma^2(X) = \sigma_x^2$

За допомогою цих властивостей, наприклад, можна довести важливе співвідношення, а саме $K(\hat{Y}, u) = 0$, яке засноване на тому, що

$$b_1 \sigma^2(X) = b_1 \left[\frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 \right] = K(X, Y) = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y},$$

$$u = Y - b_0 - b_1 X$$

$$\begin{aligned} [K(\hat{Y}, u) &= K(b_0 + b_1 X, Y - b_0 - b_1 X) = K(b_0, Y - b_0 - b_1 X) + K(b_1 X, Y - b_0 - b_1 X) = \\ &= 0 + b_1 K(X, Y - b_0 - b_1 X) = b_1 K(X, Y) - b_1 K(X, b_0) - b_1 K(X, b_1 X) = b_1 K(X, Y) - \\ &- 0 - b_1^2 K(X, X) = b_1 b_1 \sigma^2(X) - b_1^2 \sigma^2(X) = 0]. \end{aligned}$$

За великої кількості незалежних змінних x_1, x_2, \dots, x_n розв'язання нормальної системи є площина в багатовимірному просторі. Як і в розглянутому прикладі з парної регресії, перше рівняння в цьому випадку визначає поєднання цієї площини з центром ваги поля експериментальних точок – «хмари», а окремі рівняння системи спільно визначають коефіцієнти регресії, тобто коефіцієнти нахилу цієї площини до осей x_1, x_2, \dots, x_n .

Якщо до обчислення за МНК всі початкові дані центрувати, тобто знайти \bar{x}, \bar{y} і перенести початок системи координат до центру ваги хмари, то в нових координатах і $\bar{x} = 0$ маємо, що $\bar{y} = b_0 = 0$.

У цьому випадку з повної системи нормальних рівнянь МНК опускаються як перше рівняння (воно перетворюється на тотожність), так і перший стовпець (оскільки $b_0 = 0$) і система набуде вигляду:

$$\begin{cases} b_1 \sum_{i=1}^n x_{i1}^2 + b_2 \sum_{i=1}^n x_{i1}x_{i2} + \dots + b_n \sum_{i=1}^n x_{i1}x_{in} = \sum_{i=1}^n x_{i1}y_i \\ b_1 \sum_{i=1}^n x_{i2}x_{i1} + b_2 \sum_{i=1}^n x_{i2}^2 + \dots + b_n \sum_{i=1}^n x_{i2}x_{in} = \sum_{i=1}^n x_{i2}y_i \\ \dots \\ b_1 \sum_{i=1}^n x_{in}x_{i1} + b_2 \sum_{i=1}^n x_{in}x_{i2} + \dots + b_n \sum_{i=1}^n x_{in}^2 = \sum_{i=1}^n x_{in}y_i \end{cases} \quad (5.14)$$

Система (5.14) є більш загальною, а з іншого боку, якщо $\bar{y} = b_0$ і \bar{x} знаходяться раніше, то число визначених коефіцієнтів зменшиться на одиницю. Це має суттєве практичне значення, оскільки розв'язання за МНК для трьох шуканих коефіцієнтів можна запрограмувати на калькуляторі і розв'язати нескладні задачі. Наведено обчислювані алгоритми за МНК на мікрокалькуляторах. Для чотирьох і більше коефіцієнтів їх розв'язують тільки на ЕОМ.

І, нарешті, зазначимо, що система (5.14) зручна для розрахунків коефіцієнтів регресійної моделі в тих випадках, коли попередньо відомо, що поверхня відгуку повинна проходити через якусь певну точку, наприклад, початок координат.

Особливістю МНК є те, що отримані цим методом розв'язання незворотні. Це впливає з того, що (наприклад, для лінійної моделі (5.8) коефіцієнт b_i регресії \hat{y} по x визначається за співвідношенням (5.13). Якщо ж обчислювати зворотну регресію \hat{x} по y , то коефіцієнт a_i в моделі $\hat{x} = a_0 + a_1 y$ обчислюється як

$$(5.15)$$

Якщо порівняти (5.13) і (5.15), побачимо, що знаменники у них різні. Оскільки $b_1 \neq a_1$, то і $b_1 a_1 \neq 1$. Позначимо через ρ^2 вираз

$$b_1 \cdot a_1 = \frac{\left[\sum_{i=1}^n x_i y_i \right]^2}{\sum_{i=1}^n x_i^2 \cdot \sum_{i=1}^n y_i^2} = \rho^2.$$

Коефіцієнт ρ^2 називається коефіцієнтом взаємної кореляції значень x_i та y_i .

При $\rho \rightarrow 1$, тобто при незначному розсіюванні експериментальних точок і великої протяжності поля точок (хмари) обидві прямі регресії "близькі" одна до одної і їх різницею можна зневажати.

При малому $\rho < 0,95$ ця різниця стає досить суттєвою, а сам МНК не-ефективним. У цьому випадку можна скористатися методом ортогональної регресії, який розглянуто, наприклад, в [61]. В економіці існують задачі, для яких одна і та сама змінна може виступати і як факторна величина і як залежна. Тому питанням дослідження таких спряжених регресій приділяється у математичній статистиці велика увага. Більш докладно це питання буде розглянуто далі в цьому розділу.

Однак, для застосування МНК потрібне виконання таких умов (умови Гауса-Маркова) [6, 8, 22]:

1. Математичне сподівання залишків повинне дорівнювати нулю, тобто

$$M(u) = 0 \quad (5.16)$$

Якщо ця умова не виконується, то існує систематичний вплив на пояснювальну змінну y , і тому до моделі, яка описує досліджуваний процес або економічну систему не введено всіх основних пояснювальних змінних x_i . Тому необхідно розглянути іншу специфікацію моделі. Але, якщо регресійна модель має вільний член, то майже завжди за допомогою його значення можна скоригувати відповідне рівняння так, щоб виконувалася умова: $M(u) = 0$. Тому, при розв'язанні практичних задач, можна вважати, що (5.23) виконується майже завжди.

2. Залишки u_1 – компоненти вектора u некорельовані між собою і мають постійну дисперсію, тобто

$$M(u, u^r) = \sigma^2 E, \quad (5.17)$$

де E – одинична матриця, u^r – транспонований вектор u . Ця умова визначає, що залишки u є помилками вимірювання. Якщо залишки являють собою загальний вплив пояснюючих змінних, які не враховані в моделі, то умова сталості дисперсії залишків (гомоскедастичність) не виконується (гетероскедастичність), оскільки дисперсія змінюється для окремих груп спостережень і тому впливає на методи оцінювання параметрів моделі.

Якщо умова (5.17) не виконується, то регресія, яка оцінена за звичайним МНК дає неефективні оцінки (ефективність пов'язана з величиною дисперсії оцінок; функція оцінювання для вектора параметрів B моделі буде найкращою за мінімальної дисперсії, тобто коли для всіх параметрів b_1 регресії y , де – дисперсія оцінок B , визначених за звичайним МНК, а – дисперсія оцінок вектора B визначених будь-якими іншими методами).

3. Пояснюючі змінні (регресори) x_i не корельовані із залишками u , тобто

$$M(x^{\prime r}, u) = 0. \quad (5.18)$$

Це означає, що значення будь-якого фактора в кожному експерименті повинно бути екзогенною змінною, тобто визначатись повністю зовнішніми причинами, які не враховуються в рівнянні регресії.

У більшості випадків використовують більш загальне припущення – пояснювальні змінні не є стохастичними, тобто не мають випадкової складової.

Умова (5.18) насамперед визначає, що коли вона порушується, то в моделі присутні лагові (затримані у часі) змінні, або модель була сконструйована на базі одночасних рівнянь. У цьому випадку для оцінювання параметрів моделі використовують дво- або трикроковий метод найменших квадратів. Математично це можна обґрунтувати, наприклад, для парної регресії, таким чином. Тому що

$$\begin{aligned} K(X, Y) &= K(X, \beta_0 + \beta_1 X + u) = K(X, \beta_0) + K(X, \beta_1 X) + K(X, u) = \\ &= 0 + \beta_1 K(X, X) + K(X, u) = \beta_1 \sigma^2(X) + K(X, u) \end{aligned}$$

і $b_1 = \frac{K(X, Y)}{\sigma^2(X)}$, маємо

$$b_1 = \frac{K(X, Y)}{\sigma^2(X)} = \frac{\beta_1 \sigma^2(X) + K(X, u)}{\sigma^2(X)} = \beta_1 + \frac{K(X, u)}{\sigma^2(X)}.$$

Якщо $K(X, Y) = 0$, то оцінка b_1 співпадає зі значенням коефіцієнта β_1 .

4. Пояснюючі змінні x_i не корельовані між собою (умова мультиколінеарності), тобто

$$\begin{aligned} \sigma^2(x_j^{\prime r}, x_i) &= 0, \text{ якщо } j \neq i; \\ \sigma^2(x_j^{\prime r}, x_i) &= 1, \text{ якщо } j = i. \end{aligned} \quad (5.19)$$

Умова (5.19) свідчить про незалежність усіх пояснюючих змінних між собою. Однак в економіці дуже важко знайти такі фактори, які б, так чи інакше не були б залежні між собою. Тому на перших кроках конструювання моделі треба з'ясувати, чи не впливає залежність факторів на оцінку параметрів моделі. Тому ця умова має назву мультиколінеарності. Якщо існує мультиколінеарність, то існує *тісний* лінійний зв'язок між двома чи більше факторами x_1 . Перевірка на мультиколінеарність здійснюється, наприклад, за допомогою алгоритма Феррара-Глобера. Крім розглянутих чотирьох умов Гаусса-Маркова також повинна виконуватись умова про нормальний розподіл збурення u . У цьому випадку коефіцієнти регресії b_1 , також мають нормальний розподіл. Все це дає змогу використовувати загально відомі перевірки різних гіпотез, а також визначати довірчі інтервали.

При побудові регресійних функцій будемо вважати, що збурення мають нормальний розподіл. Це твердження ґрунтується на висновку з *центральної граничної теореми*: якщо ми маємо нескінченну суму досить малих випадкових величин, дія кожної з яких на загальну суму незначна, то незважаючи на те, мають вони нормальний розподіл, чи ні, загальна їх сума буде мати нормальний розподіл. Оскільки збурення – це в загальному випадку множина

неврахованих впливів на залежну змінну y (вважаємо, що модель правильно специфікована), то ми можемо практично вважати, що вона нормально розподілена.

Уведемо декілька визначень, які нам знадобляться у подальшому для одержання властивостей оцінок параметрів регресійних моделей.

ВИЗНАЧЕННЯ 5.1. Вибіркова оцінка B параметрів β називається незміщеною, якщо виконується умова $M(B) = \beta$.

ВИЗНАЧЕННЯ 5.2. Вибіркова оцінка B параметрів β називається ефективною, якщо дисперсія цієї оцінки (при заданому обсязі вибірки) є найменшою.

ВИЗНАЧЕННЯ 5.3. Вибіркова оцінка B параметрів β називається обґрунтованою, якщо вона задовольняє закону великих чисел, тобто чим більші будуються вибірки, тим більша імовірність того, що помилка оцінки не буде перевищувати достатньо малу наперед задану величину ξ , тобто

$$\lim_{n \rightarrow \infty} \{ |B - \beta| < \xi \} = 1.$$

ВИЗНАЧЕННЯ 5.4. Вибіркова оцінка B параметрів β називається інваріантною, якщо для будь-якої наперед заданої функції ψ оцінка параметрів функції $\psi(\beta)$ буде дорівнювати $\psi(B)$.

За допомогою МНК отримані оцінки є незміщеними і мають найменшу дисперсію, тобто є ефективними. Останнє підтверджує теорема Гаусса-Маркова, яка доводить, що функція оцінювання за методом МНК покомпонентно мінімізує дисперсію на класі всіх лінійних незміщених оцінок.

5.2 Узагальнення МНК

Розглянемо зараз узагальнення МНК. Для цього нам знадобиться математичний апарат матричної алгебри.

При дослідженні залежності (за аналогією як це мало місце у випадку лінійної парної регресії) однієї залежної змінної від декількох змінних де x_1, x_2, \dots, x_n у випадку лінійної регресії, маємо:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m. \quad (5.20)$$

Для уніфікованого вигляду виразу (5.20) помножимо постійну β_0 на фіктивну змінну x_0 , яка завжди тотожно дорівнює одиниці ($x_0 = 1$). Тоді (5.20) набуде вигляду

$$\hat{y} = \beta_0 x_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m. \quad (5.21)$$

У подальшому (щоб не ставити позначку " $\hat{}$ " над невідомими параметрами, для позначення його оцінки) рівняння (5.21) будемо записувати у вигляді

$$\hat{y} = b_0 x_0 + b_1 x_1 + b_2 x_2 + \dots + b_m x_m, \quad (5.22)$$

де b_i – це відповідні оцінки невідомих параметрів β_i , $i = 0, 1, 2, \dots, m$.

Результати спостережень y_1, y_2, \dots, y_n запишемо у вигляді вектор-стовпця Y , відповідні обчислені значення регресії через \hat{Y} , значення змінних x_1, x_2, \dots, x_m у вигляді матриці X розмірності $n(m+1)$, а решту функції регресії - у вигляді вектор - стовпця U (\hat{U} – оцінки вектора U), тобто

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}; \quad X = \begin{pmatrix} x_{10} & x_{11} & \dots & x_{1m} \\ x_{20} & x_{21} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ x_{n0} & x_{n1} & \dots & x_{nm} \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1m} \\ 1 & x_{21} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n1} & \dots & x_{nm} \end{pmatrix}$$

$$B = \begin{pmatrix} b_0 \\ b_1 \\ \dots \\ b_m \end{pmatrix}; \quad \hat{Y} = \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \dots \\ \hat{y}_n \end{pmatrix}; \quad U = \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix}; \quad \hat{U} = \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \dots \\ \hat{u}_n \end{pmatrix};$$

Тому лінійну модель регресії (5.21) в матричній формі запишемо у вигляді:

$$Y = X\beta + U,$$

де $\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \dots \\ \beta_m \end{pmatrix}$ – вектор невідомих параметрів регресії, який маємо оцінити.

Вектор-стовпець B є оцінкою вектора β тобто $b_i = \hat{\beta}_i, i=0, 1, \dots, m$.

Оскільки $y - \hat{y} = \hat{u}$, то можна записати, що

$$y = b_0x_0 + b_1x_1 + \dots + b_mx_m + u$$

або в матричній формі

$$Y = XB + U, \text{ де } XB = X\beta = \hat{Y}.$$

Для знаходження оцінок вектора B знову застосуємо МНК, який припускає в матричній формі (за "методом чайника"), що

$$U(B) = (Y - \hat{Y})'(Y - \hat{Y}) = U'U = \sum_{i=1}^n u_i^2 \rightarrow \min. \quad (5.23)$$

Розкриваючи дужки в лівій частині (5.23), запишемо останній вираз у вигляді

$$U(B) = (Y - XB)'(Y - XB) = (Y' - B'X')(Y - XB) = Y'Y - 2B'X'Y + B'X'XB.$$

Беручи частинні похідні за компонентами вектора B та прирівнюючи їх до нуля, отримаємо: $-2X'Y + 2X'XB = 0$, що рівнозначно рівності $X'XB = X'Y$.

Якщо для матриці $X'X$ існує зворотна їй матриця $(X'X)^{-1}$, то отримаємо після рішення системи (помноживши зліва на $(X'X)^{-1}$ обидві частини останнього рівняння) нормальних рівнянь шуканий вектор-стовпець оцінок параметрів регресії у вигляді:

$$B(X'X)^{-1}X'Y. \quad (5.24)$$

Значимо, що для застосування МНК у цьому випадку також повинні виконуватись умови 1–4, які були розглянуті раніше – (5.16) – (5.19).

Тому, наприклад, умова (5.17) буде мати такий вигляд

$$\begin{aligned}
 M(U, U') &= M \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix} \cdot (u_1, u_2, \dots, u_n) = M \begin{pmatrix} u_1^2 & u_1 u_2 & \dots & u_1 u_n \\ u_2 u_1 & u_2^2 & \dots & u_2 u_n \\ \dots & \dots & \dots & \dots \\ u_n u_1 & u_n u_2 & u_1 u_2 & u_n^2 \end{pmatrix} = \\
 &= \begin{pmatrix} M(u_1^2) & M(u_1 u_2) & \dots & M(u_1 u_n) \\ M(u_2 u_1) & M(u_2^2) & \dots & M(u_2 u_n) \\ \dots & \dots & \dots & \dots \\ M(u_n u_1) & M(u_n u_2) & \dots & M(u_n^2) \end{pmatrix} = \begin{pmatrix} \sigma_n^2 & 0 & \dots & 0 \\ 0 & \sigma_n^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_n^2 \end{pmatrix} = \\
 &= \sigma_n^2 \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix} = \sigma_n^2 E,
 \end{aligned}$$

де E – одинична матриця. Скористаємося тим, що $M(u_i, u_j) = 0$, $V_i \neq j$ (умова відсутності зв'язку між випадковими факторами), а $M(u_i^2) = \sigma_i^2$ (умова гомоскедастичності).

Знайдені оцінки B є лінійними незміщеними для вектора стовпця β тобто для них виконується умова $M(B) = \beta$. Насправді:

$$\begin{aligned}
 M(B) &= M[(X'X)^{-1}X'Y] = M[(X'X)^{-1}X'(X\beta + U)] = M[(X'X)^{-1}X'X\beta + (X'X)^{-1}X'U] = \\
 &= M[(X'X)^{-1}X'XB] + M[(X'X)^{-1}X'U] = M(E\beta + (X'X)^{-1}X'M(U)) = EM(\beta + M(X'X)^{-1}X' \cdot 0) = \beta,
 \end{aligned}$$

оскільки відповідно до умови (5.16), $M(U) = 0$, а E – одинична матриця.

Ця умова означає, що при багаторазовому проведенні експерименту, за умови, що для окремих вибірок якщо і були помилки оцінки, середнє значення цих помилок буде дорівнювати нулю. Значимо, що у випадку, коли оцінка зміщена, то $\theta = M(B) - \beta$. Ця величина є сталою порівняно з випадковою

помилкою ξ , яка дорівнює різниці оцінки і параметрів моделі, які треба оцінити, тобто

$$\xi = B - \beta.$$

Узагальнимо отриманий результат (5.24) для знаходження оцінок параметрів регресії, припускаючи, що умова (5.25) не виконується і $M(U, U') = \sigma^2_n S$, де S – відома симетрична додатно визначена матриця порядку n , тобто дисперсія та коваріація елементів утворюючих збурювання U , відомі з точністю до деякого множника. Порухення передумови рівності дисперсій збурень, як ми вже зазначали називають *гетероскедастичністю*.

Нехай $M(u-u_j) = 0, \forall(i \neq j)$. Ситуація, за якої виникає гетероскедастичність, характерна, наприклад, при помилковій специфікації форми залежності між змінними (ми знаходимо лінійну залежність між y та x , хоча дійсна залежність, наприклад, є квадратичною або кубічною тощо).

Щоб знову запрацював МНК, необхідно перетворити спостереження Y та X у нові змінні Y^* та X^* , які задовольняли б умовам:

- 1) $Y^* = X^* \beta + U^*$
- 2) $M(U^*) = 0, M(U^* U^{*'}) = \sigma^2_n E.$

Крім того, для критеріїв перевірки значущості параметрів регресії і в подальшому побудові довірчих інтервалів повинна виконуватись умова про нормальність розподілу обурень U^* .

Застосовуючи вже відомий нам МНК за цих припущень до Y^* та X^* і отримані оцінки, знову виражаючи через вихідні змінні, знаходимо шукане рівняння регресії. Для такого перетворення існує декілька еквівалентних методів оцінювання. Ми розглянемо метод, який ґрунтується на розкладанні матриці S .

Як відомо з теорії матриць, додатно визначена матриця припускає представлення її у вигляді добутку двох матриць. У матричній алгебрі показано, що може бути знайдена єдина невиврождена симетрична матриця P , така що

$$S = PP'. \quad (5.26)$$

Перемножуючи обидві частини (5.26) справа на $P^{1'}$ і зліва на P^{-1} , отримаємо

$$P^{-1} S P^{1'} = E \quad (5.27)$$

Враховуючи, що

$$P^{1'} P^{-1} = S^{-1} \quad (5.28)$$

та перемножуючи зліва на матрицю P^{-1} рівняння $Y = X \beta + U$, отримаємо

$$Y^* = X^* \beta + U^* , \quad (5.29)$$

де

$$Y^* = P^{-1} Y; X^* = P^{-1} X \text{ і } U^* = P^{-1} U. \quad (5.30)$$

Використовуючи (5.27) можна показати (рекомендуємо Вам це зробити самостійно), що

$$M(U^*, U^{*\prime}) = \sigma^2_n E,$$

а тому (5.29) задовольняє передумовам застосування вже відомого нам МНК. Тому можна скористатися вже отриманими раніше рішеннями системи нормальних рівнянь для отримання шуканих оцінок параметрів регресії

$$B = (X^{*\prime} X^*)^{-1} X^{*\prime} Y^*. \quad (5.31)$$

Підставляючи у (5.31) відповідні вирази із (5.30) з урахуванням (5.28), в результаті отримаємо, що

$$\begin{aligned} B &= (X^{*\prime} X^*)^{-1} X^{*\prime} Y^* = [(P^{-1} X) P^{-1} X]^{-1} (P^{-1} X) (P^{-1} Y) = \\ &= [X' (P^{-1}) P^{-1} X]^{-1} X' (P^{-1}) P^{-1} Y = [X' S^{-1} X]^{-1} X S^{-1} Y \end{aligned} \quad (5.32)$$

Отримані оцінки B за формулою (5.32) являють собою оцінки узагальненого методу найменших квадратів і називаються оцінками Ейткена.

Знайдені оцінки є лінійними незміщеними для вектора-стовпця β тобто для них виконується як і для (5.24) умова

$$M(B) = \beta$$

(рекомендуємо це ствердження довести самостійно).

Наприкінці зазначимо, що для застосування узагальненого МНК, потрібно визначити матрицю S .

5.3. Альтернативні методи для оцінок параметрів моделі

5.3.1. Ми розглянули питання, пов'язані з оцінкою параметрів регресійної моделі МНК. Але крім цього методу існує ще багато інших. Один з них пов'язаний з використанням дисперсійного аналізу і має назву покрокової регресії. В основу ідеї цього методу покладено той факт, що існує математичний зв'язок між параметрами оцінюваної моделі та коефіцієнтами парної кореляції, а саме

$$b = r_{yx} \frac{\sigma_y}{\sigma_x}.$$

5.3.2. Другим методом для оцінки параметрів моделі є метод максимальної правдоподібності (ММП). Р.Фішер показав, що у визначеному змісті оцінки, отримані на основі ММП, найкраще використовують інформацію про параметри, що містяться в спостереженнях*, що зробило цей метод дуже популярним. Було відкрито, що для багатьох задач різної статистичної природи, ММП дає гарні результати. Ці задачі настільки різноманітні, що не можна їх вирішувати за допомогою єдиної теорії, яка описувала б властивості методу і вказувала межі його застосування. Однак не в усіх практичних завданнях ММП дає задовільні результати. Це пов'язане, по-перше, з припущенням про точність належності невідомої щільності розподілу визначеному параметричному сімейству (нормальному, показниковому або ін.), що на практиці виконується лише приблизно. Тому і результати, отримані за ММП *можуть* призвести до ви-

сновків, які не будуть відповідати навіть приблизно реальній дійсності. Так може відбуватися за визначених, хоч і не великих, відхилень від початкових припущень. Особливо чутливі до таких порушень повинні бути оптимальні методи. Тому у цьому випадку знаходять оцінку, яка хоч і не є найкращою, але має досить сталі властивості у більш широкому класі розподілів, які включають у себе досліджуваний закон розподілу, як частковий. Отримані таким чином оцінки називаються сталими або робастними (від англ. robust estimation – грубе або стале оцінювання) По-друге, їх якісні властивості проявляються часто лише за достатньо великих обсягів вибірок, тобто є асимптотичними. По-третє, оцінки, які отримані за ММП можуть бути навіть не обґрунтованими, якщо число оцінюваних за вибіркою параметрів має той самий порядок що і обсяг вибірки. Щоб було більш-менш повне уявлення про цей метод, розглянемо його більш докладніше [8, с.33].

При використанні МНК параметри регресійної моделі і дисперсія залишків невідома. Тому спочатку оцінюються оцінки параметрів моделі, а потім обчислюються залишки, за якими знаходять їх дисперсію. Якщо задати функцію розподілу залишків, то використовуючи ММП, можна одночасно знайти оцінки всіх невідомих параметрів, включаючи і дисперсію залишків. Розглянемо це на прикладі для парної регресії.

Нехай ми знаємо, що залишки розподілені за нормальним законом. Запишемо у загальному вигляді це твердження через відповідну функцію

$$P(U) = \frac{1}{\sqrt{(\sigma_u^2 2\pi)^n}} e^{-\frac{1}{2\sigma_u^2} \sum_{i=1}^n u_i^2} du_1 du_2 \dots du_n$$

З урахуванням вигляду залишків $u_i = y_i - \beta_0 - \beta_1 x_i$ запишемо функцію правдоподібності у вигляді

$$L(\beta_0, \beta_1, \sigma_u) = \frac{1}{\sqrt{(\sigma_u^2 2\pi)^n}} e^{-\frac{1}{2\sigma_u^2} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2}$$

Прологарифмуємо цей вираз за основою e , отримаємо

$$\ln L = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma_u^2 - \frac{1}{2\sigma_u^2} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$$

Взявши часткові похідні останнього виразу за β_0 , β_1 , σ_u^2 прирівнюючи їх до нуля отримаємо наступну систему рівнянь

$$\begin{cases} \frac{\partial (\ln L)}{\partial \beta_0} = \frac{1}{\sigma_u^2} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) = 0; \\ \frac{\partial (\ln L)}{\partial \beta_1} = \frac{1}{\sigma_u^2} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) x_i = 0; \\ \frac{\partial (\ln L)}{\partial \sigma_u^2} = -\frac{n}{2\sigma_u^2} + \frac{1}{2\sigma_u^4} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 = 0. \end{cases}$$

Остання система після нескладних перетворень набуває такого вигляду

$$\begin{cases} n\beta_0 + \beta_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i; \\ \beta_0 \sum_{i=1}^n x_i + \beta_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i; \\ \frac{1}{n} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 = \sigma_u^2. \end{cases}$$

Перші два рівняння цієї системи дають нам значення для оцінок параметрів моделі β_0 і β_1 , які точно збігаються з оцінками моделі, отриманими за допомогою МНК. З третього рівняння системи отримуємо оцінку для дисперсії, яка буде мати вигляд

$$\sigma_u^2 = \frac{\sum_{i=1}^n u_i^2}{n}$$

Нескладно показати, що одержані оцінки максимізують функцію правдоподібності. Те саме можна сказати для випадку, коли ми розглядаємо багатомірну регресійну модель. Дійсно, з урахуванням вигляду залишків $u_i = y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \dots - \beta_m x_{im}$ запишемо функцію правдоподібності у вигляді

$$L(\beta_0, \beta_1, \dots, \beta_m, \sigma_u) = \frac{1}{\sqrt{(\sigma_u^2 2\pi)^2}} e^{-\frac{1}{2\sigma_u^2} (Y - XB)'(Y - XB)}.$$

Прологарифмуємо останній вираз за основою e . Отримаємо

$$\ln L = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma_u^2 - \frac{1}{2\sigma_u^2} (Y - XB)'(Y - XB).$$

Взявши часткові похідні останнього виразу по \mathbf{B} і σ_u^2 і прирівнюючи їх до нуля отримуємо таку систему рівнянь

$$\begin{cases} \frac{\partial (\ln L)}{\partial \mathbf{B}} = -\frac{1}{2\sigma_u^2} (-2X'Y + 2X'XB) = 0 \\ \frac{\partial (\ln L)}{\partial \sigma_u^2} = -\frac{n}{2\sigma_u^2} + \frac{1}{2\sigma_u^4} (Y - XB)'(Y - XB) = 0 \end{cases}$$

З першого рівняння отримуємо, що $\mathbf{B} = (\mathbf{X}'\mathbf{X}^{-1})\mathbf{X}'\mathbf{Y}$ (порівняйте ці оцінки з оцінками, які ми отримали за допомогою МНК). З другого рівняння отримуємо

$$\sigma_u^2 = \frac{(Y - XB)'(Y - XB)}{n}$$

5.3.3. Крім розглянутих методів, які використовуються для оцінки параметрів моделі окреме місце посідає метод, який базується на експериментах за методом Монте-Карло (ММК). Він називається ММК мабуть тому,

що базується на генерації випадкових чисел (якщо Ви бачили як грають у рулетку в казино, то уявляєте собі як відбувається ця генерація).

Розглянемо ММК більш докладно. Експеримент за ММК дає контрольований експеримент, який надає можливість перевірити отримані оцінки параметрів моделі і таким чином відповісти на питання: чи хороші (задовольняють нас) отримані оцінки, чи ні. ММК у найпростішому випадку складає таку послідовність кроків:

Крок 1 (підготовка до використання регресійного аналізу):

- вибираються істинні значення параметрів моделі;
- у кожному окремому спостереженні вибирається значення x ;
- використовується деякий процес генерації випадкових чисел для отримання випадкових значень u у кожному зі спостережень (для цього потрібно мати програму-генератор випадкових чисел на ЕОМ, яка існує в деяких сучасних програмних продуктах як невід'ємна їх частина);
- у кожному спостереженні генеруються значення y з використанням співвідношення $y = \beta_0 + \beta_1 x + u$ та значень β_0, β_1, x, u ;

Крок 2 (висновок про можливість використання методів регресійного аналізу, тобто чи хороші оцінки параметрів ми отримуємо, якщо використовуємо тільки інформацію про x, y):

- використовується регресійний аналіз для оцінки параметрів β_0, β_1 моделі з використанням тільки отриманих зазначеним чином значень y та даних для x ;
- порівнюються отримані оцінки з істинними значеннями параметрів моделі і робиться висновок, наскільки вони хороші та чи можливо використовувати регресійні методи для їх отримання.

У подальшому нами цей метод використовуватися не буде.

5.3.4. Для аналітичних досліджень в економіці дуже плідним є застосування і методу групового врахування аргументів (МГВА), який був розроблений видатним українським вченим А.Г. Івахненко ще у 60-х роках (Інститут кібернетики НАН України). МГВА дає змогу розв'язувати широке коло задач математичного моделювання, у тому числі, задачу ідентифікації досліджуваних економічних процесів за даними спостережень та експериментів, що відтворюються, і результати яких являють собою зафіксовані значення вихідної характеристики об'єкта як функції його параметрів.

Під математичною моделлю, синтезованою (сконструйованою) за допомогою МГВА, розуміється поліноміальна (або позиноміальна) модель, вид якої визначається опорною функцією, яка використовується. Як опорні функції можуть вибиратися поліноми (або позиноми) різних класів, гармонічні та експоненційні функції, диференційні рівняння тощо. Необхідно зазначити, що при синтезі моделі за допомогою МГВА немає необхідності визначати її початкову структуру (як, наприклад, у регресійному аналізі), що підвищує об'єктивність синтезованої моделі і знижує суб'єктивізм при конструюванні моделі процесу.

Математичні моделі, синтезовані за допомогою МГВА, називають такими моделями, що самоорганізуються. Під самоорганізацією у даному випадку розуміється процес машинного пошуку глобального мінімуму деяких критеріїв синтезу (відносного і середньоквадратичного відхилення та ін.) при поступовому ускладненні моделі. При цьому критерій синтезу має властивість "зовнішнього доповнення", що полягає у наступному. В алгоритмах МГВА вихідна вибірка поділяється на дві частини, які названі відповідно послідовностями, що навчають і перевіряють. Синтез моделі відбувається за даними послідовності, що навчає, і виконується у декілька етапів, які називаються рядами селекції. Шукана математична модель досліджуваного процесу, яка називається "повним описом" і відображає залежність вихідного показника ефективності функціонування об'єкта від усіх досліджуваних факторів

$$Y = f(X_1, X_2, \dots, X_n),$$

замінюється деяким набором "приватних описів", які є функціями тільки двох аргументів:

$$B_{ij} = f(A_i, A_j),$$

де $i = 1, 2, \dots, n - 1$; $j = i + 1, i + 2, \dots, n$; f – функція, яка обумовлена видом опорної функції.

Система рівнянь, що отримана за МГВА, її порівняння з регресійними рівняннями дає змогу отримати нові знання про об'єкт, що досліджується, і тим самим зменшити ступінь невизначеності. При цьому можуть бути виявлені такі важливі відомості про об'єкт, як його структура, про яку на першому етапі досліджень (при аналізі вхідної інформації) практично нічого не було відомо; можуть бути виділені ті дані, вірогідність яких викликає сумніви, і визначені можливі джерела дезінформації тощо.

На основі інформації, яка отримана в процесі конструювання моделі, а також при введенні нових вхідних даних процес синтезу моделі повинен бути повторений. Таким чином, постійно відбувається адаптація моделі до нових даних, умов і цілей функціонування і, отже, – до нових задач.

Нижче, тільки для уявлення про вид функції, яка може бути синтезована (сконструйована) за допомогою МГВА наводиться один з її можливих виглядів

$$U = +7.86849 - 0.072734 * (13.125 - X_2) * (1.1 - X_3) * (30.8 - X_{11}) + 9.36082e-08 * (11.4 - X_1) * (13.125 - X_2)^3 * (1.1 - X_3)^3 * (2.28048 - X_4)^2 * (30.8 - X_{11})^3 - 26.5144 * (2.28048 - X_4) * (0.87 - X_6) * (0.070274 - X_9) + 1.84586e-11 * (11.4 - X_1) * (13.125 - X_2)^4 * (1.1 - X_3)^4 * (2.28048 - X_4)^3 * (0.87 - X_6) * (30.8 - X_{11})^5 + 0.00109495 * (13.125 - X_2) * (1.1 - X_3) * (3.26806 - X_5) * (30.8 - X_{11}) + 0.000196732 * (13.125 - X_2) * (1.1 - X_3) * (3.26806 - X_5)^2 * (30.8 - X_{11}) + 0.00214097 * (13.125 - X_2) * (1.1 - X_3)^2 * (2.28048 - X_4)^3 * (3.26806 - X_5)^2 * (0.87 - X_6) * (0.070274 - X_9) * (30.8 - X_{11}) + 8.42179e-06 * (13.125 - X_2)^2 * (1.1 - X_3)^3 * (2.28048 - X_4)^6 * (3.26806 - X_5)^2 * (0.87 - X_6) * (0.070274 - X_9) * (30.8 - X_{11})^2 - 1.28608e-08 * (11.4 - X_1) * (13.125 - X_2)^4 * (1.1 - X_3)^4 * (2.28048 - X_4)^5 * (0.87 - X_6)^6 * (0.070274 - X_9)^2 * (30.8 - X_{11})^5 + 0.00028946 * (13.125 - X_2)^2 * (1.1 - X_3)^4 * (2.28048 - X_4)^9 * (3.26806 - X_5)^2 * (0.87 - X_6)^2 * (0.070274 - X_9)^2 * (30.8 - X_{11})^2$$

5.4. Абсурдні рішення, чуттєвість МНК, усереднення „абсурдностей”

Треба особливо зауважити, що МНК – формальний метод. Тому інколи при розв’язанні задач за МНК отримуємо просто абсурдні результати, які ставлять дослідника у безвихідність. У зв’язку з цим розглянемо питання, які пов’язані з отриманням абсурдних рішень більш докладніше.

5.4.1 Найчастіша причина, що трапляється на практиці – виникнення абсурдних рішень при використанні МНК – *неоднорідність статистики експериментальних даних*. Це, наприклад, пов’язано з тим, що неправильно помічені і введені в ЕОМ протоколи вимірювань або існує систематична помилка при проведенні частини вимірювань і т. ін. (тобто отримана неоднорідна статистика). На рис. 5.1 – 5.3 наведено декілька можливих таких варіантів для однієї і тієї самої задачі.

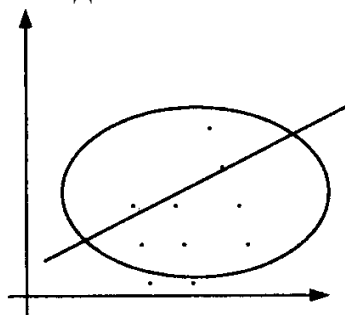


Рисунок 5.1

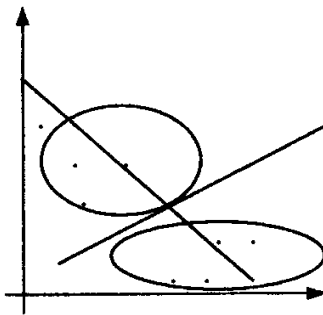


Рисунок 5.2

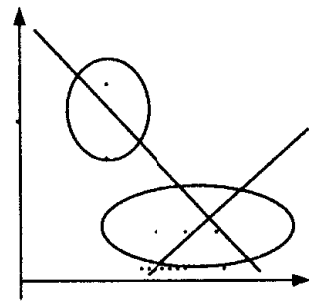


Рисунок 5.3

Для того, щоб судити про абсурдність або припустимість отриманого розв’язання за МНК, доцільно ще до розв’язання МНК дізнатися будь-яким наближенням, або графічним методом (або із апріорних припущень), який має бути порядок шуканих величин.

5.4.2 *Розв’язання задачі, отримане за МНК, в якій є випадкові помилки, є також випадковим.* Завдяки усередненню результатів багаторазових підрахунків розв’язання стає більш визначеним, більш стійким. *Перевірку на стійкість можна здійснювати таким чином.* Якщо, наприклад, у нас лінія регресії побудована за 25 експериментальними даними, і ми одну точку відкинули і побудували за 24 точками нову лінію регресії, і вона мало "відрізняється" від вже побудованої, то розв’язання стійке для 24 точок. Цю процедуру повторюємо знову дотих пір, поки "різниця" стане суттєвою. Але може статися, що за коефіцієнтами, наприклад, b_0 , b_1 розв’язання стійке, а по b_2 не стійке. Подібна погана обумовленість розв’язання свідчить, як правило, *межу переускладнення моделі.*

5.4.3 *Абсурдні розв’язання можуть виникнути і з причини, наприклад, коли в МНК на певному етапі виконуються системи рівнянь, у процесі якого завжди обчислюються числа, які можуть бути великими та близькими між собою, а через найменшу помилку в первинних даних їх різниця може змінюватись у декілька разів, і рішення будуть погано обумовленими.*

Для виключення абсурдних рішень, які отримують МНК, існує багато різних методів. Розглянемо деякі з них .

1. *Оригінальна робастна програма обробки даних* – програма стійка до присутності "промахів", запропонована І.Л. Місюченко (ЛШ, Ленінград) і заснована на використанні медіани п'яти оцінок центра.

Алгоритм цього методу полягає в наступному. Вся «хмара» експериментальних даних, яка складається з n точок, поділяється на k груп. У кожній групі точок знаходиться її центр, але не просто у вигляді медіани, а як медіана п'яти оцінок центра x_5 за абсцисами і ординатами точок цієї групи. Отже, по центрах, які знайдені таким способом, МНК визначають апроксимуючу функцію.

Ця програма працює в десятки разів швидше, ніж звичайний МНК по всьому масиву точок, оскільки *час обробки за МНК зменшується пропорційно квадрату числа відрахувань, що обробляються, і квадрату числа визначених коефіцієнтів моделі.* У результаті така програма не тільки усуває вплив промахів, але й за той самий час дає можливість прорахувати декілька варіантів і вибрати найкращий. *Цей метод достатньо ефективний при будь-якому законі розподілу похибок первинних даних, що є також дуже важливим моментом переваги в застосуванні цього методу.*

Труднощі застосування цього методу полягають у тому, що необхідно завчасно вибрати груп – m . Це особливо незручно, якщо вид моделі завчасно не відомий. Тому удаються до перебору варіантів, через що збільшується час виконання програми.

2. *Самоадаптований робастний метод обробки даних, який є суттєвим розвитком МНК, розроблений І.С.Кириловою та В.Я.Крейновичем (ВНШП, Ленінград, 1985 р.). Він ґрунтується на використанні запропонованого в 1965 році І. А. Назаровим (ЛСТІ) єдиного математичного опису всього різновиду класу експоненційних розподілів формулою виду*

$$P(x) = A(\alpha)e^{-|x|^\alpha},$$

яка при зміні α від $+\infty$ до 0 послідовно описує всі розподіли від рівномірного з ексцесом $\varepsilon = 1,8$ еліпсу розсіювання і $\alpha = +\infty$, через нормальне ($\varepsilon = 3, \alpha = 2$), до розподілів з нескінченно великим ексцесом ($\varepsilon \rightarrow \infty, \alpha \rightarrow 0$).

Ідея самоадаптованого алгоритму полягає в тому, що для $\alpha = 2$ (нормальний розподіл) найбільш ефективний МНК. Для розподілу Лапласа ($\alpha = 1$) найбільш ефективний метод найменших модулів і т.д.

Алгоритм аналогічний алгоритму МНК, але для будь-яких α як цілих, так і дробових. Наприклад, при $\alpha = 2$ він дає відповідь, відповідний до МНК, і для будь-якого α від $+\infty$ до 0 забезпечує отримання оцінок, ефективних саме для даного розподілу похибок первинних даних. Розрахунок відбувається у вигляді ітераційного процесу. *Ця процедура формальна, але не потребує втручання оператора і може просто інформувати його про хід розв'язання на кожній ітерації.*

3. *Метод ортогональної регресії.* Основним недоліком МНК є його несиметричність відносно аргументу і функції, тобто змінні при МНК нерівноправні між собою, і перенос їх з однієї частини в іншу призводить до різних

розв'язань. Але при підготовці апроксимуючих рівнянь до розв'язання за МНК такі переноси іноді необхідні.

У зв'язку з цим *метод ортогональної регресії*, що описаний Лінником Ю.В, дозволяє усунути цей недолік.

Пряма ортогональної регресії у цьому випадку збігається з віссю симетрії хмари (еліпса) розсіювання експериментальних точок (рис. 5.4).

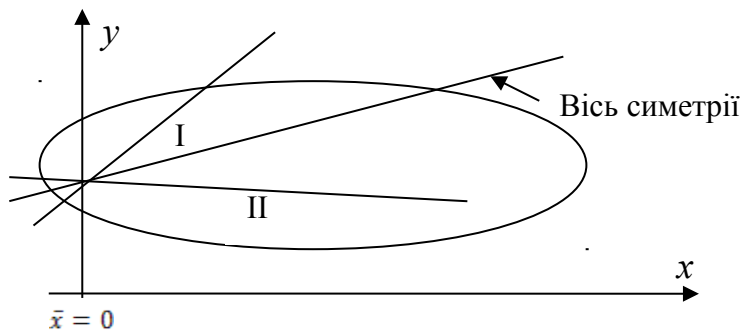


Рисунок 5.4 (лінія I – $\hat{x} = \bar{x} + a_1 y$, а лінія II – $\hat{y} = \bar{y} + b_1 x$)

У подальшому цей метод було розвинуто стосовно багатofакторного експерименту (для «хмари» розсіювання в багатовимірному просторі), при якому знаходяться рівняння всіх площин симетрії. У цьому випадку всі змінні виявляються рівноправними.

5.5. Методи відбору найбільш значущих факторів і найбільш значущих членів адитивних моделей, їх машинна інтерпретація

Розглянемо в загальних рисах методи відбору найбільш значущих факторів і найбільш значущих членів адитивної моделі.

Забезпечення компактності адитивної моделі припускає не введення в неї додаткових членів і факторів, а облік лише тих, які дійсно підвищують точність опису досліджуваного явища або процесу. Для цього існують методи, деякі з яких ми і розглянемо у загальному вигляді.

5.5.1. Метод визначення значущості членів за зміненням коефіцієнта множинної кореляції полягає в тому, що після того як вид моделі вже вибраний, з неї по чергово вилучаються кожний з членів, вибирається нове розв'язання за МНК і визначаються відповідні оцінки коефіцієнта множинної кореляції $\rho_{y\hat{y}}$.

Порівнюючи ці оцінки між собою, можна встановити порядкові місця за значущістю кожного з членів моделі.

5.5.2. Метод покрокової регресії є деяким спрощенням вищенаведеного методу з одночасною повною або частковою його автоматизацією. ЕОМ по чергово шукає розв'язання за МНК для моделей $y_0 = b_0, y_1 = b_0 + b_1 x, y_2 = b_0 + b_1 x_1 + b_2 x_2$ і т. д, визначаючи для кожної моделі залишкові дисперсії D_0, D_1, D_2, \dots та їх різниці $\Delta_1 = D_0 - D_1, \Delta_2 = D_1 - D_2, \dots$. В автоматичному режимі оператор заздалегідь задає бажане значення залишкової дисперсії $D_{з\text{ап}}$. Програма, як тільки $D_i \leq D_{з\text{ап}}$ зупиняється і видає результат.

У напівавтоматичному, діалоговому режимі виводяться всі $\Delta_1, \Delta_2, \dots$. Порівнюючи їх між собою, робимо висновок про відносну значущість факторів.

Недоліком цього методу є те, що значення Δ_i залежать від того, в якому порядку вводились в модель різні дані, хоча є програми, які передбачають порядок їх введення в модель.

5.5.3 Метод наближеного розрахунку коефіцієнтів значущості (метод β – коефіцієнтів) полягає у визначенні відносної ваги членів рівняння за їх відносними вкладками у результуюче значення відгуку y . Розглянемо у загальних рисах цей метод.

При зміні системи розрахунку в точку $\bar{x}_1, \bar{x}_2, \dots$,

$\hat{y}_i - \bar{y} = b_1(x_{i1} - \bar{x}_1) + b_2(x_{i2} - \bar{x}_2) + \dots$ впливає, що

$$\frac{\hat{y}_i - \bar{y}}{\sigma_y} = b_1 \frac{\sigma_{x1}}{\sigma_y} \cdot \frac{x_{i1} - \bar{x}_1}{\sigma_{x1}} + b_2 \frac{\sigma_{x2}}{\sigma_y} \cdot \frac{x_{i2} - \bar{x}_2}{\sigma_{x2}} + \dots$$

Таким чином, отримуємо, що

$\frac{\hat{y}_i - \bar{y}}{\sigma_y}, \frac{x_{i1} - \bar{x}_1}{\sigma_{x1}}, \dots, i = 1, 2, \dots, n$ – центровані і нормовані величини представлені ви-

бірками обсягу n , а $\beta_i = b_i \frac{\sigma_{xi}}{\sigma_y}$ ваги їх відносного вкладу в значення відгуку.

Зручність використання методу β – коефіцієнтів обумовлено простотою їх розрахунку. Але потрібно мати на увазі наступне:

- оскільки $\sigma_{xi} > 0, \sigma_y > 0$ то β_i має знак b_i . Тому знаки при β_i не враховуються, а значущість оцінюється за відношенням їх абсолютних величин;

- якщо β_i – коефіцієнти двох членів моделі різняться між собою менш ніж в 1,3 – 1,5 разу, то такі члени визнаються рівнозначними; якщо ця умова виконується, то такі члени вважаються нерівнозначними.

5.5.4. Зазначимо, якщо нами побудована послідовність моделей за однією і тією самою вибіркою і з них обирається найкраща, то потрібно вибрати ту, яка задовольняє фізичній суті. Можна спробувати це зробити, якщо порівняти всі отримані розв'язання за значеннями залишкової дисперсії, тобто за окремими значеннями і вибрати ту, яка має мінімальну залишкову дисперсію. Ця гіперплощина називається *площиною ортогональної регресії*. Але на практиці досить часто вона відповідає абсурдному з фізичної точки зору, рішення задачі через те, що $R_{y\hat{y}} < 0,96$. Тому в цьому випадку потрібно удаватися не до мінімуму дисперсії, а модель, яку шукаємо, вибрати за економічною суттю розв'язаної задачі.

5.6. Економічна інтерпретація результатів, які отримані за допомогою МНК

5.6.1. Коефіцієнти регресії b_i є мультиплікаторами. Вони показують, на скільки одиниць зміниться y , якщо x_i зміниться на одиницю *у припущенні*, що немає впливу інших факторів. Насправді, оскільки

$$\widehat{y} = b_0 + b_1x_1 + \dots + b_ix_i + b_mx_m; \quad (*)$$

$$\widehat{y}_1 = b_0 + b_1x_1 + \dots + b_i(x_i + 1) + b_mx_m. \quad (**)$$

Різницею виразів (***) і (*) буде

$$\Delta y = \widehat{y}_1 - y = b_i.$$

Це свідчить про те, що при зміні x_i на одиницю (*у припущенні*, що немає впливу інших факторів) y зміниться на Δy одиниць.

5.6.2 Середня, наприклад, продуктивність для фактора X_i :

$$\mu_i = \frac{\widehat{Y}}{X_i},$$

5.6.3 Гранична, наприклад, продуктивність фактора X_i .

$$E_i = \frac{\partial \widehat{Y}}{\partial X_i}$$

Гранична продуктивність свідчить, що при збільшенні X_i на одиницю на величину E_i , зміниться \widehat{Y} за всіх незмінних інших факторів.

5.6.4 Еластичність регресора \widehat{Y} за фактором X_i :

$$\alpha_i = \frac{\Delta \widehat{y}_i}{\widehat{y}_i} / \frac{\Delta \widehat{x}_i}{\widehat{x}_i} = \frac{E_i}{\mu_i}.$$

Ці коефіцієнти свідчать, що регресанд y зміниться на α %, якщо фактор x_i зміниться на один процент *у припущенні*, що немає впливу інших факторів. Таким чином, *коефіцієнт еластичності є показником впливу зміни питомої ваги x_i на y у припущенні*, що немає вплив інших факторів.

Коефіцієнти еластичності для лінійної регресійної моделі обчислюються за формулою

$$\alpha_i = b_i \frac{\bar{x}_i}{\bar{y}}, \bar{x}_i = \frac{\sum_{j=1}^n x_{ji}}{n}, \bar{y} = \frac{\sum_{j=1}^n y_j}{n}.$$

5.6.5. Загальна еластичність Y від усіх факторів X_i дорівнює:

$$\alpha = \sum_{i=1}^m \alpha_i.$$

Цей показник свідчить, що на α процентів зміниться Y , якщо одночасно збільшити на один процент всі фактори X_i .

5.6.6 Міра ефективності використання X_i :

$$\gamma_i = \frac{E_i}{X_i}.$$

Вона показує, який буде граничний приріст Y на одну одиницю фактора X_i .

5.6.7 Гранична норма заміщення фактора X_i фактором X_j :

$$h_{ij} = \frac{\partial Y / \partial X_i}{\partial Y / \partial X_j}.$$

Цей показник характеризує можливість взаємного заміщення одного фактора на інший за постійного (незмінного) значення Y .

І, на закінчення цього розділу зазначимо, що визначення оцінок параметрів регресійної моделі, це є один з перших кроків побудови математичної моделі досліджуваного економічного процесу. Після того, як модель побудована, постає питання: наскільки вона адекватна реальному вивчаному процесу; наскільки оцінки, які ми знайшли "хороші"; можливо чи ні визначити, як буде поводити себе процес у майбутньому (прогнозувати його поведінку), і наскільки цей прогноз буде відповідати реаліям життя тощо.

РОЗДІЛ 6

ПЕРЕВІРКА МОДЕЛІ НА АДЕКВАТНІСТЬ. ПРОБЛЕМА ПРОГНОЗВАННЯ

6.1 Математичні проблеми перевірки на адекватність регресійних моделей реальним економічним процесам

6.1.1 У другому розділі ми розглянули моделі системи. Тепер мова йтиме не взагалі про моделі, а про їх деяку підмножину – математичні моделі, які являють собою систему математичних співвідношень, що описують досліджуваний економічний процес або явище. І тут відразу постає багато нових питань. Наскільки повно і точно відображає наша модель реальний процес? Наскільки ті спрощення й обмеження, що ми прийняли при укладанні математичної моделі процесу, не вплинули на повноту і вірогідність відображуваної моделлю реальної дійсності? Чи не надто ми все спростили? Звичайно, як говорив К. Прутков – не можна осягнути неосяжне. Але тоді постає питання: чи маємо ми взагалі право ("моральне", "правове", "етичне" і т. ін.) відображати моделлю ті процеси, що нас цікавлять, і після цього результати, отримані на основі моделі, підносити у вигляді "меню" для "клієнта"? Чи не є створена "нашими ж руками" модель, тим самим "кухарем-отруйником", який радий бачити, як корчаться в муках страждань його "клієнти"? Таких питань може бути безліч і, найгірше те, що усі вони не тільки мають право на існування, але і "живучі" настільки, що "усунути" їх не можливо. Тому з найперших кроків конструювання моделі ми завжди повинні пам'ятати про "кухаря-отруйника" і по можливості дуже акуратно "куштувати блюда", які "приготовлені" моделлю. При цьому, звичайно, гарантії, що Ви уникнете отруєння ніхто не дасть. Але щоб у випадку чого, хоча б знайти "винуватця" Ваших бід необхідно, щоб "рецепт приготування" моделі був у Вас більш-менше повним, а саме - наскільки "приготовлена кухарем" модель "адекватно" відображає замовлене Вами блюдо з запропонованого меню. Звичайно, Ви не в аптеці і тому деякі (а може бути й усі) складові частини моделі можуть коливатися в тих або інших межах, але це не завжди страшно для Вашого здоров'я (звичайно з деякою заданою імовірністю).

Адекватність (від лат. *adaequatus* – прирівняний, однаковий), що відповідає, правильне, точне: у теорії пізнання – правильне відтворення в мисленні зв'язків і відношень об'єктивного світу. Проблема адекватності – проблема, яка має давні корені, їй приділяли велику увагу великі філософи і вчені протягом усієї історії розвитку науки. Зараз ця проблема також не втратила своєї актуальності. Так А. І. Ракітов писав: "Язык считается адекватным, если получаемые в нем предложения могут описать все существующие или возможные ситуации в области объектов, информацию о которых выражает, хранит и передает данный язык". Але тоді, як правильно відзначив А. Ю. Цофнас: "... когда язык выражает не все ситуации, то он адекватен тем из них, которые он выражает". Далі "Тогда и язык может быть адекватным не множеству "ситуаций" объектов, а только осмыслению ситуации в этом же или ином языке". Напевно в цьому ракурсі і потрібно роз-

глядати мову математичної статистики, як "адекватне" відображення поведінки реальних економічних систем (процесів, явищ) і на цією мовою говорити про "дійсність", "адекватність" отриманих результатів реальним процесам. У цьому зв'язку не можна утриматися від спокуси, щоб не згадати творця науки "тектології" А. А. Богданова (власне кажучи "батька" загальної теорії систем) і не навести його слова з приводу "істини": "...Человеческое мышление постоянно работает по этому способу, постоянно ведет подбор понятий и их комбинаций, испытывает их одну за другой, отбрасывает те, удерживает эти и сохраняющийся остаток обозначает как "истину".

6.1.2. Розглянемо питання, яке пов'язане з якістю оцінки регресійної моделі взагалі, тобто "адекватність" моделі реальному досліджуваному процесу – щільність зв'язку між показниками моделі. Для цього розглянемо наступне співвідношення (загальне відхилення реальних даних від їх середнього)

$$y_i - \bar{y} = (y_i - \bar{y}) + (y_i - \hat{y}_i).$$

Перший доданок у правій частині характеризує відхилення, яке впливає із самої регресійної моделі, а другий - відхилення, яке не можна пояснити з регресійної моделі. Піднесемо до квадрату обидві частини цього співвідношення, підсумуємо, і одержимо такий вираз

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2 + 2 \sum_{i=1}^n (\hat{y}_i - \bar{y})(y_i - \hat{y}_i). \quad (6.1)$$

Оскільки

$$\begin{aligned} \sum_{i=1}^n (\hat{y}_i - \bar{y})(y_i - \hat{y}_i) &= \sum_{i=1}^n (b_0 + b_1 x_{i1} + \dots + b_m x_{im} - (b_0 + b_1 \bar{x}_1 + \dots + b_m \bar{x}_m))(y_i - \hat{y}_i) = \\ &= b_1 \sum_{i=1}^n (x_{i1} - \bar{x}_1)(y_i - \hat{y}_i) + \dots + b_m \sum_{i=1}^n (x_{im} - \bar{x}_m)(y_i - \hat{y}_i) = \\ &= b_1 \sum_{i=1}^n x_{i1} (y_i - \hat{y}_i) - b_1 \bar{x}_1 \sum_{i=1}^n (y_i - \hat{y}_i) + \dots + b_m \sum_{i=1}^n x_{im} (y_i - \hat{y}_i) - b_m \bar{x}_m \sum_{i=1}^n (y_i - \hat{y}_i) = \\ &= b_1 \sum_{i=1}^n x_{i1} u_i - b_1 \bar{x}_1 \sum_{i=1}^n u_i + \dots + b_m \sum_{i=1}^n x_{im} u_i - b_m \bar{x}_m \sum_{i=1}^n u_i = 0, \end{aligned}$$

тому вираз (6.1) набуде вигляду

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

Перша сума правої частини – сума квадратів, що пояснює регресію, а друга – сума квадратів помилок. Поділивши обидві частини на n отримаємо відповідні дисперсії:

$$\sigma_{\text{заг}}^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n} \quad \text{– загальна варіація;} \quad \sigma_{\text{рег}}^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{n} \quad \text{– варіація, що}$$

$$\text{пояснює регресію;} \quad \sigma_{\text{пом}}^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad \text{– варіація помилок.}$$

З урахуванням викладеного маємо, що

$$\sigma_{\text{заг}}^2 = \sigma_{\text{рег}}^2 + \sigma_{\text{пом}}^2. \quad (6.2)$$

З виразу (6.2) отримаємо $1 = \frac{\sigma_{рег}^2}{\sigma_{заг}^2} + \frac{\sigma_{ном}^2}{\sigma_{заг}^2}$. Позначимо перший доданок через R^2 , тобто

$$R^2 = \frac{\text{пояснююча варіація}}{\text{загальна варіація}} = \frac{\sigma_{рег}^2}{\sigma_{заг}^2}.$$

Вираз для R^2 можна також записати у такому вигляді

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\sum_{i=1}^n (y_i - \bar{y})^2 - \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n u_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

або

$$R^2 = \frac{\sum_{i=1}^n \frac{\sum_{j=1}^n (y_j - \bar{y})(\hat{y}_j - \bar{y})}{n}}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\sum_{i=1}^n \frac{\sum_{j=1}^n (y_j - \bar{y})(y_j - \bar{y})}{n}}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\sum_{i=1}^n \frac{\sum_{j=1}^n (y_j - \bar{y})^2}{n}}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1$$

Вираз для R^2 називають вибіркоvim коефіцієнтом множинної детермінації (у подальшому – просто коефіцієнтом детермінації). Він характеризує тісноту зв'язку загального впливу всіх незалежних факторів x_i на залежну змінну. Його числове значення змінюється від нуля до одиниці, тобто $R^2 \in [0,1]$. Наведений коефіцієнт детермінації характеризує, якою мірою варіація залежної змінної визначається варіацією незалежних факторів. Чим ближчий він до одиниці, тим більша варіація залежної змінної визначається варіацією незалежних змінних. Тому R^2 показує, наскільки якісно побудована регресійна модель до значень Y , які спостерігаються.

У випадку коли $R^2 = 1$ то рівняння регресії точно відповідає всім наглядам і тому $y = \hat{y}$ (всі $u_i = 0$, тобто всі догляди знаходяться на регресійній площині). Якщо ж у вибірці відсутній взаємозв'язок між пояснюючою і пояснювальними змінними, то $R^2 \rightarrow 0$.

Тому при побудові регресії необхідно, щоб R^2 якомога ближче наближався до одиниці.

Наведемо ще один вираз для обчислення коефіцієнта детермінації.

Для цього обчислимо коефіцієнт кореляції $r_{y,\hat{y}}$ між y і \hat{y} :

$$\begin{aligned}
r_{y,\hat{y}} &= \frac{M(y, \hat{y})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \cdot \sum_{i=1}^n (\hat{y}_i - \bar{y})^2}} = \frac{M(\hat{y} + u, \hat{y})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \cdot \sum_{i=1}^n (\hat{y}_i - \bar{y})^2}} = \\
&= \frac{M(\hat{y}, \hat{y}) + M(u, \hat{y})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \cdot \sum_{i=1}^n (\hat{y}_i - \bar{y})^2}} = \frac{M(\hat{y}, \hat{y})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \cdot \sum_{i=1}^n (\hat{y}_i - \bar{y})^2}} = \frac{\sigma^2(\hat{y})}{\sqrt{\sigma^2(y)\sigma^2(\hat{y})}} = \\
&= \frac{\sqrt{\sigma^2(\hat{y})}}{\sqrt{\sigma^2(y)}} = \sqrt{\frac{\sigma^2(\hat{y})}{\sigma^2(y)}} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}} = \sqrt{R^2}..
\end{aligned}$$

Значення $R = \sqrt{R^2}$ називають множинним коефіцієнтом кореляції. Він характеризує тісноту лінійного зв'язку усіх незалежних факторів x_i із залежною змінною y . Для нього з урахуванням та без урахування числа ступенів свободи характерна така сама зміна числового значення, як і для коефіцієнта детермінації.

Запишемо R^2 у матричній формі. Для цього запишемо його складові у такому вигляді:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - 2\bar{y} \sum_{i=1}^n y_i + n\bar{y}^2 = Y'Y - 2\bar{y}n\bar{y} + n\bar{y}^2 = Y'Y - n\bar{y}^2,$$

а

$$\begin{aligned}
\sum_{i=1}^n u_i^2 &= U'U = Y'Y - \hat{Y}'\hat{Y} = Y'Y - (XB)'(XB) = Y'Y - B'X'XB = Y'Y - B'X'X(X'X)^{-1}X'Y = \\
&= Y'Y - B'EX'Y = Y'Y - B'X'Y.
\end{aligned}$$

Тому вираз для R^2 набуде такого вигляду:

$$R^2 = 1 - \frac{Y'Y - B'X'Y}{Y'Y - n\bar{y}^2} = \frac{B'X'Y - n\bar{y}^2}{Y'Y - n\bar{y}^2}. \quad (6.3)$$

Оскільки R випадкове, то постає питання про надійність меж, в яких вона може змінюватись. Для вибраного рівня значущості α і відповідного ступеня свободи $k = n - m - 1$ межі надійності обчислюються за формулою (без доведення)

$$\Delta R = \pm t_{\alpha/2, k} \cdot \frac{1 - R}{\sqrt{n}}.$$

Однак, при використанні коефіцієнта детермінації як "адекватної" регресійної моделі реальному економічному процесу можуть виникнути деякі проблеми, а саме:

1) при введенні додаткового регресора у регресійну модель, перерахований коефіцієнт детермінації буде принаймні не менший ніж попередній. Тому може скластися враження, що остання модель краща за попередню;

2) при введенні додаткового регресора у регресійну модель втрачається

один ступінь свободи, але вираз (6.3) цей факт не враховує.

Тому при обчисленні R^2 необхідно його обчислювати таким чином, щоб для моделі, яку ми конструємо, враховувались перелічені недоліки і він забезпечував би компенсацію такого автоматичного збільшення незалежних факторів за рахунок деякого "штрафу". З цією метою ще раз перепишемо вираз для R^2

$$R^2 = 1 - \frac{\sum_{i=1}^n u_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Поділимо чисельник та знаменник у R^2 на число відповідних ступенів свободи (а не просто на n , бо у цьому випадку ми отримаємо зміщені оцінки). У результаті у чисельнику і знаменнику будуть відповідні незміщені дисперсії σ_u^2 (оцінена дисперсія залишків) та σ_y^2 (вибіркова дисперсія залежної змінної). Тому вираз для R^2 набуде вигляду

$$R^2 = 1 - \frac{\sigma_u^2}{\sigma_y^2}, \quad (6.4)$$

де

$$\sigma_u^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - m - 1}; \quad \sigma_y^2 = \frac{\sum_{i=1}^n u_i^2}{n - m - 1}; \quad \sigma_{\hat{y}}^2 = \frac{\sum_{i=1}^n \hat{y}_i^2}{n - m - 1}; \quad \sigma_{\bar{y}}^2 = \frac{\sum_{i=1}^n \bar{y}^2}{n - m - 1}; \quad \sigma_{\bar{y}}^2 = \frac{\sum_{i=1}^n \bar{y}^2}{n - m - 1}$$

а

$$\sigma_y^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - 1} = \frac{Y'Y - n\bar{y}^2}{n - 1},$$

де m – число незалежних факторів.

З урахуванням цього отримуємо, що

$$R^2 = 1 - \frac{Y'Y - B'X'Y}{n - m - 1} \cdot \frac{Y'Y - n\bar{y}^2}{n - 1} = 1 - \frac{Y'Y - B'X'Y}{Y'Y - n\bar{y}^2} * \frac{n - 1}{n - m - 1}. \quad (6.5)$$

Для того, щоб відрізнити отриманий коефіцієнт R^2 – (6.5) від (6.3), його позначають через \bar{R}_T^2 і називають скоригованим коефіцієнтом детермінації за Тейлором. Цей коефіцієнт пов'язаний з коефіцієнтом, обчисленим за формулою (6.3), таким співвідношенням

$$\bar{R}_T^2 = 1 - (1 - R^2) \frac{n-1}{n-m-1}. \quad (6.6)$$

Існує ще один вираз для скоригованого коефіцієнта детермінації, який називається скоригованим коефіцієнтом за Аменією і позначається як \bar{R}_A^2 . Він має вигляд.

$$\bar{R}_A^2 = 1 - (1 - R^2) \frac{n+m+1}{n-m-1}. \quad (6.7)$$

При введенні додаткового регресора коефіцієнт \bar{R}_A^2 змінюється на більшу величину, ніж \bar{R}_T^2 і тому відображає ступінь свободи в цьому випадку більш чітко. Тому при меншому числі регресорів перевага надається \bar{R}_A^2 , а не \bar{R}_T^2 .

Але зауважимо, що аналізуючи все вищевикладене можна стверджувати, що зростання скоригованих коефіцієнтів і додавання нової незалежної змінної в рівняння регресії не обов'язково свідчить, що її коефіцієнт значно відрізняється від нуля. Тому його зростання не обов'язково свідчить, що рівняння найбільше відповідає досліджуваному процесу, ніж за відсутності цієї змінної. Це і є одна з причин того, чому скориговані коефіцієнти не стали широко використовуватись як перевірна процедура. Друга причина цього є зменшення уваги до самого коефіцієнта детермінації R^2 . Це пов'язане з тим, виникають ситуації, коли сама модель погано визначена, але має високий рівень значення R^2 , тому цей показник на сьогодні розглядається як один з цілого сімейства діагностичних показників для перевірки при побудові моделі, а не як один з головних. Тому і коригування його мало що дає.

І ще одне, що стосується коефіцієнта детермінації. Іноді використовують співвідношення (частковий коефіцієнт детермінації)

$$\Delta R_k^2 = \frac{(1 - R^2)}{n - k} t_k^2 \quad (6.8)$$

для обчислення граничного вкладу k -го регресора у коефіцієнт детермінації, де R^2 – коефіцієнт детермінації, який обчислений для усіх k регресорів, включених у регресійну модель;

$$t_k = \frac{b_k - b_k^*}{\sigma_{b_k}} - t - \text{статистика для } k\text{-го коефіцієнта регресії (} b_k^* \text{ – довільне}$$

(задане та обґрунтоване дослідником) число; у нашому випадку $b_k^* = 0$);

σ_{b_k} – стандартна похибка оцінки b_k (k -го регресійного коефіцієнта), її обчислення розглянемо далі.

Обчислене значення ΔR_k^2 показує, наскільки k -й регресор впливає на якість моделі, а саме: наскільки зменшиться R^2 , якщо k -й регресор буде виключений з моделі.

6.1.3 Крім розглянутого коефіцієнта детермінації, для перевірки ре-

гресійної моделі на "адекватність" використовують F -критерій Фішера. При цьому виходять з того, що залишки u розподілені нормально, тобто користуються фундаментальною теоремою про те, що для нормально розподіленої випадкової величини з нульовою середньою і одиничною дисперсією сума квадратів її n

випадково вибраних значень має розподіл χ^2 з n ступенями свободи.

За допомогою F -критерію Фішера перевіряються гіпотези про значущість кількох коефіцієнтів або кількох лінійних комбінацій, або їх композиції. Наприклад, часто перевіряють гіпотезу про те, що всі коефіцієнти при незалежних змінних дорівнюють нулю, тобто $b_i = 0$ ($i = 1, 2, \dots, m$). Мовою математичної статистики це формулюється таким чином: перевірити нульову гіпотезу H_0 :

$$H_0 : b_1 = b_2 = \dots = b_m = 0.$$

Для її перевірки використовують F -критерій Фішера, з m та $(n-m-1)$ ступенями свободи. Цей критерій розрахований на перевірку тільки двосторонніх гіпотез. Зазначимо, що розглянута вище t -статистика (тести) розрахована на перевірку як односторонніх, так і двосторонніх гіпотез.

У зв'язку з тим, що ми часто будемо використовувати такі гіпотези, розглянемо більш докладно питання, які пов'язані з прийняттям або неприйняттям таких гіпотез.

На першому кроці вибирається критерій, який розбиває всю множину можливих значень на дві неперетинні множини. На наступному кроці висувається гіпотеза: до якої з цих множин належить цей критерій. Висунуту гіпотезу називають нульовою і позначають як H_0 . Тоді альтернативною (конкуруючою) до неї є гіпотеза, яка протилежна нульовій. Оскільки гіпотеза, яку ми перевіряємо може бути як правильною, так і неправильною, потрібна її перевірка. Оскільки це робиться за допомогою статистичних методів, така перевірка називається статистичною. Якщо за допомогою перевірки ми відкидаємо перевірену гіпотезу H_0 , а вона є правильна, то ми робимо помилку, яку називають помилкою першого роду. Якщо ми приймаємо гіпотезу, але вона є помилкова, то говорять про помилку другого роду. Від цих помилок та кола вирішуваних задач часто залежать в економіці дуже важливі рішення.

Для статистичної перевірки нульової гіпотези використовують спеціально підібрані випадкові величини, закон розподілу яких точно або наближено ми знаємо. Такі випадкові величини називаються статистичними критеріями.

Критичною областю називають сукупність значень критерію, за яких H_0 відхиляють, інакше таку область називають областю припустимих значень, або областю прийняття гіпотези. Якщо ми приймаємо гіпотезу, то говорять, що дані, з якими ми працюємо, узгоджуються з нульовою гіпотезою, тобто ми не маємо підстав нею знехтувати.

Зараз ми вже в змозі сформулювати основний принцип щодо статистичної перевірки гіпотез. Він полягає в наступному. Якщо значення критерію належить критичній області, то гіпотезу треба відхилити, якщо значення критерію належить області припустимих значень, то її приймають. Оскільки критерій є випадкова величина, то і її можливі значення належать деякому інтервалу. Тому як критична область, так і область прийняття гіпотези є також деякими інтервалами. Точ-

ку, яка поділяє розглядуваний інтервал на два інтервали, називають критичною точкою.

Залежно від виду розглядуваних інтервалів можна говорити і відповідно про види гіпотез, а саме, якщо через K позначити критерій, а через $k_{кр}$ – критичну точку, то:

– одностороння гіпотеза перевіряє лівосторонню ($K < k_{кр(1)}$), або правосторонню ($K > k_{кр(2)}$) критичну області ($k_{кр(1)} < 0, k_{кр(2)} > 0$);

– двостороння гіпотеза перевіряє двосторонню критичну область ($K < k_{кр(1)}, K > k_{кр(2)}$).

Якщо $|k_{кр(1)}| = k_{кр(2)} = k_{кр}$, то двосторонню критичну область іноді записують у вигляді $|K| > k_{кр}$ і називають двосторонніми симетричними критичними областями. Найбільш часто використовують саме двосторонні симетричні критичні області.

Таким чином, для знаходження критичної області необхідно знайти критичну точку $k_{кр}$. Для цього задають достатньо малу ймовірність α , яку називають у зв'язку з її особливою значущістю - рівнем значущості. Після цього з рівнянь $P(|K| > k_{кр}) = \alpha_1$ або $P(K > k_{кр}) = \alpha$, або $P(K < -k_{кр}) = \alpha$, або знаходять критичну точку. Для відповідних критеріїв існують таблиці, які наведено у більшості книг із статистики і по яких знаходять необхідні критичні значення, задовольняють цьому критерію.

Зазначимо, що якщо $P(K < -k_{кр}) = P(K > k_{кр})$, то з того, що для двосторонньої критичної області $P(K < -k_{кр}) + P(K > k_{кр}) = \alpha$, то

$$P(K < -k_{кр}) = P(K > k_{кр}) = \frac{\alpha}{2}.$$

Останній вираз дає змогу обчислювати критичні точки для двосторонніх критичних областей за двосторонніми.

Оскільки події ($K < -k_{кр}$), ($K > k_{кр}$) і ($-k_{кр} \leq K \leq k_{кр}$) являють собою повну групу несумісних подій, то

$$P(K < -k_{кр}) + P(K > k_{кр}) + P(-k_{кр} \leq K \leq k_{кр}) = 1.$$

Звідси маємо, що ймовірність прийняття нульової гіпотези дорівнює:

$$P(-k_{кр} \leq K \leq k_{кр}) = 1 - P(K < -k_{кр}) - P(K > k_{кр}) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha.$$

Якщо ми обчислюємо ймовірність того, що критерій попадає у критичну область, а істина – конкуруюча гіпотеза, то така ймовірність має назву потужності критерію. Таким чином, потужність критерію – це ймовірність того, що H_0 не буде прийнята, якщо істина – альтернативна гіпотеза.

Зараз ми вже можемо розглянути питання, які пов'язані з різними статистичними перевірками критеріїв та їх модифікацій, які характеризують модель на "адекватність" реальним досліджуваним процесам.

Розглянемо ситуацію, коли пояснювальна варіація є великою порівняно з непояснюючою. Тому можна вважати, що всі або деякі коефіцієнти моделі суттєво відрізняються від нуля. А це, своєю чергою, дає нам право говорити про гіпотезу: усі коефіцієнти моделі дорівнюють нулю проти альтернативної до неї – не всі коефіцієнти моделі дорівнюють нулю. Цю гіпотезу можна перевірити за допомогою F -критерію. Для цього розглянемо відношення, яке ми позначимо

через F

$$F = \frac{\text{пояснювальна_варіація/число_ступенів_свободи}}{\text{непояснювальна_варіація/число_ступенів_свободи}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / m}{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n - m - 1)} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \cdot \frac{n - m - 1}{m} \quad (6.9)$$

Обчислене значення F -критерію (Фішера) порівнюється з табличним при ступенях свободи m і $(n - m - 1)$ та вибраному рівні значущості α (рівень помилки), або $p = 1 - \alpha$ (рівень довіри). Якщо $F(m, n - m - 1) > F_{\text{табл}}(m, n - m - 1, \alpha)$, то гіпотезу про сутності зв'язку між залежною і незалежними змінними моделі приймаємо, інакше – відкидаємо.

Нагадаємо, що у статистиці під кількістю ступенів свободи розуміють різницю між кількістю наглядів і кількістю параметрів, які встановлені у результаті цих наглядів, незалежно один від одного.

Наприклад, для парної регресії вигляду $y = b_0 + b_1x + u$ маємо:

1) вираз $\sum_{i=1}^n (y_i - \bar{y})^2$, має один ступінь свободи, оскільки для утворення цієї суми квадратів потрібно $(n-1)$ незалежних змінних, а саме $(y_1 - \bar{y}), (y_2 - \bar{y}), \dots, (y_n - \bar{y})$, де одна з цих змінних є лінійною комбінацією інших завдяки виконанню рівності $\sum_{i=1}^n (y_i - \bar{y}) = \sum_{i=1}^n y_i - n\bar{y} = n\bar{y} - n\bar{y} = 0$; 2) вираз $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ має $(n-1)$ ступінь свободи, оскільки для утворення цієї суми квадратів потрібна тільки одна незалежна змінна, а саме b_1 , що випливає з наступної рівності $\hat{y}_i - \bar{y} = b_1(x_i - \bar{x})$; 3) вираз $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ має $(n - 2)$ ступінь свободи, оскільки для утворення цієї суми квадратів потрібно тільки дві незалежні змінні, а саме b_0, b_1 .

Тому, наприклад, для парної регресії F -критерій буде мати такий вигляд

$$F = \frac{\frac{\sum_{i=1}^n (\hat{y}_i - y)^2}{1}}{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n-2}} = \frac{\sum_{i=1}^n (\hat{y}_i - y)^2}{\sum_{i=1}^n (\hat{y}_i - y_i)^2} \cdot \frac{n-2}{1} \quad (6.10)$$

Між F -критерієм Фішера і коефіцієнтом детермінації R^2 існує безпосередній зв'язок. Знайдемо його. Для цього вираз (6.9) для обчислення F -критерію Фішера перетворимо таким чином

$$F = \frac{\frac{\sum_{i=1}^n (\hat{y}_i - y)^2}{n-m-1}}{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{m}} = \frac{\sum_{i=1}^n (\hat{y}_i - y)^2}{\sum_{i=1}^n (\hat{y}_i - y)^2 \cdot \sum_{i=1}^n (\hat{y}_i - y_i)^2} \cdot \frac{n-m-1}{m} \quad (6.11)$$

$$= \frac{\frac{\sum_{i=1}^n (\hat{y}_i - y)^2}{\sum_{i=1}^n (\hat{y}_i - y)^2}}{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\hat{y}_i - y)^2}} \cdot \frac{n-m-1}{m} = \frac{R^2}{1-R^2} \cdot \frac{n-m-1}{m}$$

Для того, щоб визначити вплив на знайдену регресію з m регресорами, k регресорів ($m > k$) за допомогою F -критерію Фішера використовують співвідношення

$$F = \frac{\sum_{i=1}^n \varepsilon_i^2 - \sum_{i=1}^n u_i^2}{\sum_{i=1}^n u_i^2} \cdot \frac{n - m - 1}{k} \quad (6.12)$$

де u_i – відхилення для регресії, у яку включені всі $(m + 1)$ регресорів; ε_i – відхилення для регресії, у яку не включені k регресорів, тобто для регресії з $(m - k - 1)$ регресорами.

Запишемо вираз (6.12) у матричному вигляді

$$F = \frac{\varepsilon' \varepsilon - u' u}{u' u} \cdot \frac{n - m - 1}{k} \quad (6.13)$$

Для заданого рівня значущості α і ступенів свободи $k_1 = k$ і $k_2 = n - m - 1$ знайдемо табличне значення F – критерію $F_{\text{табл.}}(k_1, k_2, \alpha)$. Порівнюючи табличне та обчислене значення F – критерію робимо висновок: якщо $F_{\text{табл.}}(k_1, k_2, \alpha) < F$, то при вибраному рівні значущості α , або з надійністю $p = 1 - \alpha$ можна вважати, що виключені з моделі k регресорів суттєво впливають на модель і їх не можна виключати з розгляду; інакше, тобто якщо $F_{\text{табл.}}(k_1, k_2, \alpha) > F$, то виключені з моделі фактори з заданою надійністю не суттєво впливають на модель.

6.1.4 І, нарешті, розглянемо питання, яке пов'язане зі значущістю коефіцієнта кореляції. Оскільки коефіцієнт кореляції є також вибірковою характеристикою, яка може відхилитись від свого "істинного" значення, значущість коефіцієнта кореляції також потребує перевірки. Вона базується на t -критерії

$$t = \frac{R \sqrt{n - m - 1}}{\sqrt{1 - R^2}}, \quad (6.14)$$

де R^2 – коефіцієнт детермінації моделі, якій характеризує, якою мірою варіація залежної змінної визначається варіацією незалежних змінних; $R = \sqrt{R^2}$ – множинний коефіцієнт кореляції, який характеризує тісноту зв'язку всіх незалежних змінних із залежною; $(n - m - 1)$ – число ступенів вільності.

Якщо $|t| > t_{\text{табл.}}(\alpha / 2, n - m - 1)$, де $t_{\text{табл.}}(\alpha / 2, n - m - 1)$ – відповідне табличне значення t -розподілу з $(n - m - 1)$ ступенями свободи, то можна зробити висновок про значущість коефіцієнта кореляції між залежною і незалежними змінними моделі. Для парної регресії маємо

$$t = \frac{R \sqrt{n - 1 - 1}}{\sqrt{1 - R^2}} = \frac{R \sqrt{n - 2}}{\sqrt{1 - R^2}} \quad (6.15)$$

Тому при $|t| > t_{\text{табл.}}(\alpha / 2, n - 2)$ можемо зробити висновок про значущість коефіцієнта кореляції між залежною Y і незалежною змінною X у парній регресійній моделі.

6.2. Статистичні аспекти, які пов'язані з обчисленням надійних

інтервалів моделі, її поведінки у майбутньому

Розглянемо питання, яке пов'язане з визначенням дисперсії та стандартні похибки оцінок параметрів b_i . (Це пов'язано з тим, що оскільки випадкова змінна u безпосередньо впливає на y , то y є також випадковою величиною. Але для знаходження оцінок B ми використовуємо співвідношення (5.24), а саме $B = (X'X)^{-1} X'Y$, у яке також входить Y і тому оцінки B також мають випадковий характер). Для цього знайдемо коваріаційну матрицю, яка визначає дисперсію оцінок параметрів. Ця матриця має вигляд:

$$\Sigma_B = \text{Cov}B = M[(B - \beta) \cdot (B - \beta)'] \quad (6.16)$$

Оскільки

$$B - \beta = (X'X)^{-1} X'Y - \beta = (X'X)^{-1} X'(X\beta + u) - \beta = (X'X)^{-1} X'X\beta + (X'X)^{-1} X'u - \beta = E\beta + (X'X)^{-1} X'u - \beta = (X'X)^{-1} X'u.$$

То вираз (6.16) набуває вигляду

$$\begin{aligned} \Sigma_B = \text{Cov}B &= M[(B - \beta) \cdot (B - \beta)'] = M\left\{[(X'X)^{-1} X'u][(X'X)^{-1} X'u]'\right\} = \\ &= M[(X'X)^{-1} X'u u' X (X'X)^{-1}] = (X'X)^{-1} X' M[u u'] X (X'X)^{-1} = \\ &= (X'X)^{-1} X' \sigma_u^2 X (X'X)^{-1} = \sigma_u^2 (X'X)^{-1} X' X (X'X)^{-1} = \sigma_u^2 (X'X)^{-1} E = \sigma_u^2 (X'X)^{-1} \end{aligned}$$

(у випадку, коли ми застосовуємо оцінки Ейткена, відповідна кореляційна матриця має вигляд

$$\text{Cov}B = \sigma_u^2 (X^* X^*)^{-1} = \sigma_u^2 (X' S^{-1} X)^{-1},$$

де

$$\sigma_u^2 = \frac{u' S^{-1} u}{n - m - 1}. \quad (6.18)$$

Оскільки дисперсія залишків σ_u^2 невідома, то як статистичну її оцінку використовуємо вибірккову дисперсію залишків $\hat{\sigma}_u^2$ при $n - (m + 1)$ ступенів свободи (у випадку парної регресії маємо тільки один коефіцієнт b_1 при x , тому $m = 1$; нагадаємо, що одиницю до m додаємо за рахунок вільного коефіцієнта b_0 при $x_0 \equiv 1$)

$$\hat{\sigma}_u^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - m - 1} = \frac{\sum_{i=1}^n u_i^2}{n - m - 1} = \frac{u'u}{n - m - 1} = \frac{Y'Y - \hat{Y}'\hat{Y}}{n - m - 1} = \frac{Y'Y - B'X'Y}{n - m - 1}. \quad (6.19)$$

Нагадаємо, що у статистиці під кількістю ступенів свободи розуміють різницю між кількістю наглядів і кількістю параметрів, які встановлені у результаті цих наглядів, незалежно один від одного.

Обчислюючи окремо добутки відповідних матриць $Y'Y$, $B'X'Y$, та підставляючи їх значення в (6.19), обчислюємо $\hat{\sigma}_u^2$. Обчислюючи обернену матри-

цю $C = (X'X)^{-1}$:

$$C = (X'X)^{-1} = \begin{pmatrix} c_{00} & c_{01} & \dots & c_{0m} \\ c_{10} & c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots & \dots \\ c_{m0} & c_{m1} & \dots & c_{mm} \end{pmatrix}, \quad (6.20)$$

(цю матрицю також називають матрицею помилок) можна обчислити коваріаційну матрицю (6.17), яка матиме у загальному випадку такий вигляд:

$$\Sigma_B = \sigma_u^2 C = \begin{pmatrix} \hat{\sigma}_{b_0}^2 & \hat{\sigma}_{b_{01}}^2 & \dots & \hat{\sigma}_{b_{0m}}^2 \\ \hat{\sigma}_{b_{10}}^2 & \hat{\sigma}_{b_1}^2 & \dots & \hat{\sigma}_{b_{1m}}^2 \\ \dots & \dots & \dots & \dots \\ \hat{\sigma}_{b_{m0}}^2 & \hat{\sigma}_{b_{m1}}^2 & \dots & \hat{\sigma}_{b_m}^2 \end{pmatrix}, \quad (6.21)$$

де $\hat{\sigma}_{b_{ij}} = \hat{\sigma}_{b_i b_j}$ - оцінка коваріації між параметрами моделі b_i та b_j , $\hat{\sigma}_{b_i}^2$ - оцінка дисперсії параметра моделі b_i .

Зараз ми вже в змозі визначати дисперсії і стандартні помилки оцінок параметрів відповідної функції регресії:

$$\hat{\sigma}_{b_i}^2 = \sigma_u^2 c_{ij} \quad ; \quad (6.22)$$

$$\hat{\sigma}_{b_{ij}} = \hat{\sigma}_{b_i b_j} = \sigma_u^2 c_{ij}, \quad (6.23)$$

де c_{ij} - діагональний елемент матриці $C=(X'X)^{-1}$, а c_{ij} - елемент матриці C , який знаходиться на перетині i -го рядка та j -го стовпця.

Позначимо через S_{b_j} - стандартизовану помилку оцінки параметра b_j регресійної моделі. Вона обчислюється як $S_{b_j} = \sqrt{\hat{\sigma}_{b_j}^2}$.

Відношення

$$\sigma_{b_j} = \frac{S_{b_j}}{|b_j|} \cdot 100\% \quad (6.24)$$

характеризує, наскільки добре обчислена відповідна оцінка параметра моделі. Якщо вона мала, то це свідчить про її незміщеність, якщо велика, то вона зміщена. Наслідком такої зміщеності є невиконання умови (5.16), тобто $M(u) \neq 0$. Це, своєю чергою, виникає через те, що залишки мають систематичну складову, яка зумовлена некоректною специфікацією моделі досліджуваного процесу.

Перед тим, як розглянути задачу, яка пов'язана з обчисленням прогнозу за

моделлю, проаналізуємо питання щодо значущості оцінок параметрів моделі та їх довірчих інтервалів.

Припускаючи, що залишки і моделі розподілені за нормальним законом (це була одна з умов, яка дозволяє використання звичайного МНК), тобто $u \in N(0, \sigma^2 E)$ перевіримо на значущість оцінки параметрів B та знайдемо для них довірчі інтервали. Як впливає з умови, параметри B моделі задовольняють багатовимірному нормальному розподілу, тобто

$$B \in N(a, \sigma_u^2 (X'X)^{-1})$$

Якщо була б відома величина дисперсії σ_u^2 , то перевірку на значущість коефіцієнтів регресії та оцінювання довірчих інтервалів можна було б зробити за відомими співвідношеннями з математичної статистики. Однак, у нашому випадку дисперсія σ_u^2 невідома. Тому для розв'язання задачі знайдемо оцінку цієї невідомої дисперсії. Для цього визначимо залишки y :

$$\begin{aligned} u &= Y - XB = XB + e - X[(X'X)^{-1} X'(XB + e)] = \\ &= XB + e - X(X'X)^{-1} X'XB + X(X'X)^{-1} X'e = \\ &= XB + e - XB + X(X'X)^{-1} X'e = e + X(X'X)^{-1} X'e = \\ &= [E_n - X(X'X)^{-1} X']e = Ne, \end{aligned} \quad (6.25)$$

де e – неспостережувана випадкова величина для всієї генеральної сукупності даних, тоді як u – випадкова величина для вибірки.

Таким чином, залишки, які дістали на підставі експериментальних даних, записано у вигляді лінійних функцій від невідомих залишків e . Помножимо з права (6.25) на транспонований вектор залишків u . Тоді сума квадратів відхилень набуде такого вигляду:

$$u'u = e'N'Ne = e'Ne = e'[E_n - X(X'X)^{-1}X']e,$$

де N – симетрична ідемпотентна матриця, тобто помножена сама на себе дає знову матрицю N (оскільки E_n – одинична матриця, а $X(X'X)^{-1}X'$ симетрична матриця)*.

$$\begin{aligned} * N' &= [E_n - X(X'X)^{-1} X']' = E_n - X(X'X)^{-1} X' = N; \\ N'N &= [E_n - X(X'X)^{-1} X'] \cdot [E_n - X(X'X)^{-1} X'] = \\ &= E_n - X(X'X)^{-1} X' - X(X'X)^{-1} X' + [X(X'X)^{-1} X'] [X(X'X)^{-1} X'] = \\ &= E_n - X(X'X)^{-1} X' = N. \end{aligned}$$

Тому математичне сподівання $M(uu')$ з урахуванням того, що $M(e'Ne) = \sigma^2 \text{tr}(N)$, де $\text{tr}(N)$ – слід матриці (сума елементів, які знаходяться на її головній діагоналі) N і властивостей обчислення сліду матриць буде мати наступний вигляд:

$$\begin{aligned} M(uu') &= \sigma^2 \text{tr} [E_n - X(X'X)^{-1} X'] = \sigma_e^2 (\text{tr} E_n - \text{tr} [X(X'X)^{-1} X']) = \\ &= \sigma_e^2 (n - \text{tr} [(X'X)^{-1} X'X]) = (n - m - 1) \sigma_e^2 \end{aligned}$$

У цьому співвідношенні матриця $X'X$ має порядок $(m+1)$, добуток

$(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X}) = \mathbf{E}_n$, а її слід дорівнює $m+1$. Тому

$$\sigma_e^2 = \frac{u'u}{n-m-1}.$$

Останнє співвідношення дає нам змогу отримати незміщену оцінку дисперсії залишків.

Покажемо, що сума квадратів розподілена незалежно від \mathbf{B} . Для цього знайдемо коваріацію залишків:

$$\begin{aligned} \mathbf{M}[u(\mathbf{B}-\beta)'] &= \mathbf{M}[(\mathbf{E}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{u}\mathbf{u}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}] = \\ &= \sigma_u^2 \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} - \sigma_u^2 \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \mathbf{0}. \end{aligned}$$

Оскільки \mathbf{u} і \mathbf{B} є лінійні функції від нормально розподілених змінних, то вони також нормально розподілені і, як було показано, їх коваріації дорівнюють нулю.

Тому можна скористатися t - розподілом для перевірки гіпотез відносно дійсності кожного з параметрів регресійної моделі

$$b_j \in N(a_j, \sigma_u^2 c_{jj}).$$

Перевірку гіпотези виконуємо згідно з t -критерієм:

$$t_j = \frac{b_j}{\sqrt{\sigma_u^2 c_{jj}}} = \frac{b_j}{S_{b_j}}$$

де c_{jj} – діагональний елемент матриці $(\mathbf{X}'\mathbf{X})^{-1}$, S_{b_j} – стандартизована помилка оцінки параметра моделі.

Значення t_j – критерію порівнюється з табличними при $(m-n-1)$ ступенях свободи і рівні значущості α ; якщо $t_j > t_{\text{табл}}$, то відповідно оцінка параметра регресійної моделі є достовірною.

Довірчі інтервали для параметра b_j обчислюються як

$$\left(b_j - t_{\alpha/2,k} \sqrt{\sigma_u^2 c_{jj}}; b_j + t_{\alpha/2,k} \sqrt{\sigma_u^2 c_{jj}}; b_j \right)$$

Довірчий інтервал при рівні надійності $(1-\alpha)$ є інтервал з випадково залежними межами і накриває дійсне значення коефіцієнта регресії b_j з рівнем довіри $(1-\alpha)$.

Це впливає з того, що якщо задано рівень значущості α , то згідно з t – статистикою Стьюдента для відповідної кількості ступенів свободи випадкова змінна

$$t_j = \frac{b_j - \beta_j}{\sigma_{b_j}} = \frac{b_j - \beta_j}{\sqrt{\sigma_u^2 c_{jj}}}$$

є центрованою t -розподіленою змінною з $k=n-m-1$ ступенями свободи. З таблиці t - розподілу Стьюдента знаходимо критичне значення $t_{\alpha/2,k}$ для $k = n - m - 1$ ступенів вільності. Використовуючи мову теорії ймовірностей запишемо умову, яка задає необхідні довірчі інтервали:

$$P(-t_{\alpha/2,k} < t < t_{\alpha/2,k}) = 1 - \alpha.$$

Підставляючи в останній вираз значення для t -статистики, отримаємо

$$P\left(-t_{\alpha/2,k} < \frac{b_j - \beta_j}{\sigma_{b_j}} < t_{\alpha/2,k}\right) = 1 - \alpha.$$

Звідки остаточно маємо

$$P(b_j - \sigma_{b_j} t_{\alpha/2,k} < \beta_j < b_j + \sigma_{b_j} t_{\alpha/2,k}) = 1 - \alpha,$$

тобто довірчі інтервали для параметра \mathbf{b}_j обчислюються як

$$\left(b_j - t_{\alpha/2,k} \sqrt{\sigma_u^2 c_{jj}}; b_j + t_{\alpha/2,k} \sqrt{\sigma_u^2 c_{jj}} \right).$$

Якщо стандартні помилки параметрів більші за абсолютні значення оцінки цих параметрів, то це може означати, що оцінка параметра є зміщеною.

Зауважимо, що для парної регресії отримуємо вирази для параметрів регресії $\mathbf{b}_0, \mathbf{b}_1$ будуть мати вигляд

$$\left(b_0 - t_{\alpha/2,k} \sigma_u \sqrt{c_{00}}; b_0 + t_{\alpha/2,k} \sigma_u \sqrt{c_{00}} \right) \\ \left(b_1 - t_{\alpha/2,k} \sigma_u \sqrt{c_{11}}; b_1 + t_{\alpha/2,k} \sigma_u \sqrt{c_{11}} \right),$$

де

$$\sigma_u = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-2}}, \quad c_{00} = \frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}, \quad c_{11} = \frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

Це можна було б отримати ще з таких міркувань. Оскільки парної регресії маємо $\mathbf{b} = \mathbf{y} - \mathbf{ax}$, то виконуються наступні відношення:

$$\sigma^2(\bar{y}) = \sigma^2 \left(\frac{\sum_{i=1}^n y_i}{n} \right) = \frac{1}{n^2} \sigma^2 \left(\sum_{i=1}^n y_i \right) = \frac{\sum_{i=1}^n \sigma^2(y_i)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$$\sigma^2(b_0) = \sigma^2(\bar{y} - b_1 \bar{x}) = \sigma^2(\bar{y}) + \bar{x}^2 \sigma^2(b_1) = \frac{\sigma^2}{n} + \frac{\sigma^2 \bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} = \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) =$$

$$= \sigma^2 \left(\frac{\sum_{i=1}^n x_i^2 - 2\bar{x} \sum_{i=1}^n x_i + n\bar{x}^2 + n\bar{x}^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} \right) = \frac{\sigma^2 \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\sigma^2(b_1) = \sigma^2 \left(\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) = \sigma^2 \left(\frac{\sum_{i=1}^n (x_i - \bar{x})y_i - \bar{y} \sum_{i=1}^n (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) =$$

$$= \sigma^2 \left(\frac{\sum_{i=1}^n (x_i - \bar{x})y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \sigma^2(y_i)}{\left[\sum_{i=1}^n (x_i - \bar{x})^2 \right]^2} = \frac{\sigma^2 \sum_{i=1}^n (x_i - \bar{x})^2}{\left[\sum_{i=1}^n (x_i - \bar{x})^2 \right]^2} = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

Тому довірчі інтервали для параметра \mathbf{b}_j обчислюються за формулою (з урахуванням того, що він розподілений за нормальним законом з дисперсією σ^2 і нульовим математичним сподіванням)

$$\left(b_1 - t_{\alpha/2,k} \sqrt{\frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}; b_1 + t_{\alpha/2,k} \sqrt{\frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \right)$$

або з урахуванням того, що незсуненою і обґрунтованою оцінкою для (σ^2) буде з урахуванням числа ступенів свободи $k = n - 2$

$$S = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - 2}}$$

остаточно отримаємо

$$\left(b_1 - t_{\alpha/2,k} \cdot \frac{S}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}; b_1 + t_{\alpha/2,k} \cdot \frac{S}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} \right)$$

Для параметра b_0 відповідно отримаємо

$$\left(b_0 - t_{\alpha/2,k} S \sqrt{\frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}}; b_0 + t_{\alpha/2,k} S \sqrt{\frac{\sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2}} \right)$$

Розглянемо зараз питання з визначення надійної зони регресії.

У зв'язку з цим, по-перше, вирішимо це питання для парної регресії, а потім на основі "методу чайника" зробимо це для багатовимірної регресійної моделі. Для цього обчислимо її дисперсію

$$\begin{aligned} \sigma^2(\hat{y}_i) &= \sigma^2(\bar{y}) + (x_i - \bar{x})^2 \sigma^2(b_1) = \frac{\sigma^2}{n} + \frac{\sigma^2 \cdot (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} = \\ &= \sigma^2 \left(\frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) \end{aligned}$$

Тому з урахуванням, що незсуненою і обґрунтованою оцінкою для σ^2 буде з урахуванням числа ступенів свободи S^2 , отримаємо надійну зону регресії (яка отримується в результаті з'єднання отриманих точок)

$$\hat{y}_i - t_{\alpha/2,k} \cdot \frac{S}{\sqrt{n}} \cdot \sqrt{1 + \frac{(x_i - \bar{x})^2}{\sigma^2(x)}}; \hat{y}_i + t_{\alpha/2,k} \cdot \frac{S}{\sqrt{n}} \cdot \sqrt{1 + \frac{(x_i - \bar{x})^2}{\sigma^2(x)}}.$$

Тоді для загального випадку будемо мати:

$$\left(\hat{y}_i - t_{\alpha/2,k} \cdot S \cdot \sqrt{X_i (X'X)^{-1} X_i'}; \hat{y}_i + t_{\alpha/2,k} \cdot S \cdot \sqrt{X_i (X'X)^{-1} X_i'} \right),$$

де $X_i = (1, x_{1i}, x_{2i}, \dots, x_{mi})$ – вектор, для якого знаходиться надійна зона.

6.3. Питання, що пов'язані з обчисленням надійних інтервалів моделі, її поведінкою у майбутньому

Розглянемо питання, які пов'язані з обчисленням прогнозу, що сприяє передбаченню імовірнісних шляхів поведінки досліджуваної економічної системи (процесу) у майбутньому. Це є дуже важливою і складною задачею, вирішення якої повною мірою використовується, наприклад, при прийнятті відповідальних стратегічних управлінських рішень щодо управління досліджуваним економічним процесом з урахуванням деякого періоду упередження.

Отже, нехай регресійне рівняння в матричній формі має вигляд

$$Y = XB + u,$$

де Y – вектор значень залежної змінної; X – матриця незалежних змінних розміром $n \times (m + 1)$; B – вектор оцінок параметрів моделі; u – вектор залишків.

Знайдемо прогнозні значення вектора Y_p , якій відповідатиме очікуваним значенням матриці незалежних змінних X_p . Цей прогноз може бути точковим або інтервальним.

Оскільки $Y = XB + u$, то незміщена оцінка прогнозу

$$M[Y_p(X_p)] = X_p \beta + u.$$

Дисперсія прогнозу є (без доведення)

$$M\left\{\widehat{Y}_p - M\left[\widehat{Y}_p(X_p)\right]\right\}^2 = \sigma_u^2 \left[\frac{1}{n} + \frac{(x_{pj} - \bar{x}_j)^2}{\sum_{i=1}^n x_{ij}^2} \right].$$

При віддаленні значення x_{ij} від відповідного середнього значення вибірки вона зростає.

Запишемо цю дисперсію прогнозу у матричному вигляді

$$\sigma_u^2 = \sigma_u^2 X_p (X'X)^{-1} X_p'$$

Запишемо середньоквадратичну помилку прогнозу

$$S = \sqrt{\widehat{\sigma}_u^2 \left[\frac{1}{n} + \frac{(x_{pj} - \bar{x}_j)^2}{\sum_{i=1}^n x_{ij}^2} \right]} = \widehat{\sigma}_u \sqrt{\left[\frac{1}{n} + \frac{(x_{pj} - \bar{x}_j)^2}{\sum_{i=1}^n x_{ij}^2} \right]} = \widehat{\sigma}_u \sqrt{X_p (X'X)^{-1} X_p'}$$

і запишемо критерій t – розподілу

$$t_{j(\alpha)} = \frac{\widehat{Y}_p - M(Y_p(X_p))}{\widehat{\sigma}_u \sqrt{\frac{1}{n} + \frac{(x_{pj} - \bar{x}_j)^2}{\sum_{i=1}^n x_{ij}^2}}} = \frac{\widehat{Y}_p - M(Y_p(X_p))}{\widehat{\sigma}_u \sqrt{X_p (X'X)^{-1} X_p'}}$$

при $(n - m - 1)$ ступенях свободи рівні значущості α .

Довірчий інтервал для прогнозних значень має вигляд

$$\widehat{Y}_p - t_{j/(\alpha/2)} \sigma_u \sqrt{\frac{1}{n} + \frac{(x_{pj} - \bar{x}_j)^2}{\sum_{i=1}^n x_{ij}^2}} \leq \bar{Y}_p \leq \widehat{Y}_p + t_{j/(\alpha/2)} \sigma_u \sqrt{\frac{1}{n} + \frac{(x_{pj} - \bar{x}_j)^2}{\sum_{i=1}^n x_{ij}^2}},$$

або в матричному вигляді

$$\widehat{Y}_p - t_{\alpha/2} \sigma_u \sqrt{X_p (X'X)^{-1} X_p'} \leq M(Y_p(X_p)) \leq \widehat{Y}_p + t_{\alpha/2} \sigma_u \sqrt{X_p (X'X)^{-1} X_p'}$$

Зауважимо, що $Y_p = X_p B$ можна розглядати як точкову оцінку математичного сподівання прогнозного значення Y_p , а також як індивідуальне значення Y_p для вектора незалежних змінних $X_p = (1, x_{1p}, x_{2p}, \dots, x_{mp})$, що лежать за межами базового періоду.

Для визначення інтервального прогнозу індивідуального значення Y_p необхідно знайти відповідну стандартну похибку, яка має вигляд

$$\sigma_{u(i)}^2 = \sigma_u^2 + \sigma_n^2 = \sigma_u^2 + \sigma_n^2 X_p (X'X)^{-1} X_p' = \sigma_u^2 \left(1 + X_p (X'X)^{-1} X_p'\right)$$

Тому інтервальний прогноз індивідуального значення визначається як

$$\widehat{Y}_p - t_{\alpha} \sigma_{n(i)} \leq Y_p \leq \widehat{Y}_p + t_{\alpha} \sigma_{n(i)},$$

або

$$\widehat{Y}_p - t_{\alpha/2} \sigma_u \sqrt{1 + X_p (X'X)^{-1} X_p'} \leq Y_p \leq \widehat{Y}_p + t_{\alpha/2} \sigma_u \sqrt{1 + X_p (X'X)^{-1} X_p'}$$

У наступному розділі ми досліджуватимемо регресійні моделі, розглянемо питання, які пов'язані з порушенням деяких умов, притаманних класичному аналізу, наприклад, умови гомоскедастичності, наявності у масиві даних мультиколінеарності тощо.

Для лінійної парної регресії: $y_p = b_0 + b_1 x$ — для знаходження прогнозних значень необхідно підставити значення x_{pi} у рівняння регресії.

Для знаходження межі надійних інтервалів індивідуальних прогнозованих значень для парної регресії обчислимо значення

$$\Delta y_{pi} = t_{\alpha, k} S \sqrt{1 + \frac{1}{n} + \frac{(x_{pi} - \bar{x})^2}{\sum_{i=1}^n (x_{pi} - \bar{x})^2}},$$

а межі надійних інтервалів індивідуальних прогнозованих значень тоді набудуть наступного вигляду

$$(y_{pi} - \Delta y_{pi}; y_{pi} + \Delta y_{pi}).$$

Це впливає з того, що

$$\begin{aligned} \sigma^2(y_{pi}) &= \sigma^2(b_0 + b_1 x_{pi} + u_i) = \sigma^2(b_0 + b_1 x_{pi}) + \sigma^2(u_i) = \\ &= \frac{\sigma^2}{n} + \frac{\sigma^2 (x_{pi} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + \sigma^2 = \sigma^2 \left[\frac{1}{n} + \frac{(x_{pi} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + 1 \right]. \end{aligned}$$

РОЗДІЛ 7

МУЛЬТИКОЛІНЕАРНІСТЬ. ГОМО- І ГЕТЕРОСКЕДАСТИЧНІСТЬ

7.1 Мультиколінеарність. Її наслідки

У розд. 5 згадувалося про мультиколінеарність (термін впроваджено Р. Фрішем), коли розглядали умови застосування МНК. Це стосується четвертої умови Гаусса-Маркова (5.19) (нагадаємо, що мультиколінеарність означає існування тісної лінійної залежності, або кореляції, між двома чи більше регресорами):

$$\sigma^2(x^T_j, x_i) = 0, \text{ якщо } j \neq i,$$

$$\sigma^2(x^T_j, x_i) = 1, \text{ якщо } j = i.$$

Однак в економічній системі дуже важко знайти параметри, які її характеризують та процеси, що в ній відбуваються, так чи інакше не пов'язані між собою. Крім того основна властивість системи – це її цілісність. Тому за визначенням вони повинні бути пов'язаними між собою. Але, таке становище може призвести до непередбачених наслідків. Отже спочатку перед конструюванням моделі треба з'ясувати, чи існує залежність між пояснюючими змінними, а також якщо вона й існує, то чи впливає така залежність на оцінку параметрів моделі. Тому ця умова має назву "мультиколінеарності", а не "колінеарності", як за наявності залежності між випадковими величинами (згадайте розділ з теорії імовірностей та математичної статистики). Зазначимо, що мультиколінеарність, крім того, що вона може негативно впливати на оцінки параметрів моделі (наприклад, мультиколінеарність призводить до зміщення отриманих оцінок моделі зі всіма наслідками для подальших досліджень), які отримані за допомогою МНК, може призвести до того, що модель взагалі побудувати буде неможливо. Останнє пов'язане з тим, що оцінки, отримані за МНК, обчислюються за формулою $\mathbf{B}=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$, а за "дуже сильної" мультиколінеарності, або функціонального зв'язку між регресорами моделі, матриця $\mathbf{X}'\mathbf{X}$ може бути виродженою, а тому неможливо буде побудувати обернену до неї матрицю (для знаходження оберненої матриці, одним з кроків її побудови є обчислення визначника, а після цього поділ відповідних алгебраїчних доповнень на нього).

Крім того, оскільки мультиколінеарність – це існування тісного лінійного зв'язку між двома чи більше факторами, які визначають пояснюючу (залежну) змінну, то модель, яка сконструйована взагалі буде не тільки практично недієздатною, а і беззмістовною взагалі через те, що отримані оцінки регресії не є надійними.

Підсумовуючи викладене, можна зробити висновок, що мультиколінеарність, це є не що інше як термін, який використовують для того, щоб окреслити коло проблем, які пов'язані з тим, що лінійна залежність між поясню-

вальними та пояснюючою змінними є не суворою і як наслідок це, своєю чергою, призводить до небажаних наслідків, що можуть позначитися на оцінках моделі.

Таким чином, мультиколінеарність може призвести до того, що [22]:

- знижується точність оцінювання моделі (помилки деяких оцінок стають занадто значними; досить корельованими одна з одною;
- дисперсії оцінок параметрів моделі різко збільшуються). При цьому інтервали довіри збільшуються, а t -статистики стають незначними;
- оцінки параметрів моделі для деяких змінних можуть бути незначними через наявність їх взаємозв'язку з іншими змінними, хоча вони і впливають суттєво на регресанд. Все це призводить до того, що вихідна множина вибірових даних не в змозі цей вплив виявити;
- оцінки параметрів стають досить чутливими до обсягів сукупності спостережень, оскільки збільшення останніх може призвести до суттєвих змін самих оцінок моделі. Більш того, іноді в умовах мультиколінеарності можна отримати незміщені та стійкі оцінки [19]. Але ефект мультиколінеарності виявляється у тому, що складно отримати значення параметрів з малою стандартною помилкою. Такий саме ефект спостерігається тоді, коли ми маємо малу кількість спостережень, або при невеликій зміні значень.

Зазначимо, що коли фактори розглядаються у часі, то вони у більшості випадків тим більш можуть бути тісно корельовані, що, природно, призведе до мультиколінеарності. Тому при розгляданні часових рядів, вирішення цієї проблеми набуває більшої ваги і без сумніву стає ще більш важливою та актуальною.

Покажемо на простому прикладі, як може вплинути мультиколінеарність на оцінку параметрів моделі. Нехай між деякими факторами x_i і x_k у регресійній моделі існує залежність, тобто $x_j = f(x_k)$. Якщо функція f – лінійна, то можна записати, що $x_k = \gamma_i x_i + \varepsilon_k$.

Припустимо, що ми розглядаємо двофакторну модель, тобто

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$ ($i = 1, k = 2$). З урахуванням залежності між факторами вона матиме вигляд:

$$\begin{aligned} y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u = \beta_0 + \beta_1 x_1 + \beta_2 (\gamma_i x_i + \varepsilon_k) + u \\ &= \beta_0 + (\beta_1 + \beta_2 \gamma_1) x_1 + \beta_2 \varepsilon_2 + u \end{aligned}$$

Таким чином, кількість змінних зменшилося на одиницю, але при факторі X_1 ми маємо другий коефіцієнт, який враховує спільний вплив факторів x_1 і x_2 . Крім того, збурення мають зовсім інший вигляд ніж u , тобто які у нас були у попередній моделі, яка була сконструйована спочатку. Тому, якщо ми знайшли оцінку для двофакторної моделі, то при x_1 вона не буде збігатися з оцінкою отриманої моделі при x_1 після врахування залежності між x_1 і x_2 .

Більш того, якщо припустити, що $\varepsilon_k = 0$, то оцінки отримані за МНК можна, як відомо обчислити за формулами:

$$\begin{aligned}
b_2 &= \frac{\sum_{i=1}^n (x_{i2} - \bar{x}_2)(y_i - \bar{y}) \sum_{i=1}^n (x_{i1} - \bar{x}_1)^2}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 \sum_{i=1}^n (x_{21} - \bar{x}_2)^2 - \left[\sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \right]^2} - \\
&\quad \frac{\sum_{i=1}^n (x_{i1} - \bar{x}_1)(y_i - \bar{y}) \sum_{i=1}^n (x_{i2} - \bar{x}_2)}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 \sum_{i=1}^n (x_{21} - \bar{x}_2)^2 - \left[\sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \right]^2} = \\
&= \frac{\gamma \sum_{i=1}^n (x_{i1} - \bar{x}_1)(y_i - \bar{y}) \sum_{i=1}^n (x_{i1} - \bar{x}_1)^2}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 \sum_{i=1}^n (x_{21} - \bar{x}_2)^2 - \left[\sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \right]^2} - \\
&\quad \frac{\gamma \sum_{i=1}^n (x_{i1} - \bar{x}_1)(y_i - \bar{y}) \sum_{i=1}^n (x_{i1} - \bar{x}_1)^2}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 \sum_{i=1}^n (x_{21} - \bar{x}_2)^2 - \left[\sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \right]^2} = \frac{0}{0}; \\
b_1 &= \frac{\sum_{i=1}^n (x_{i1} - \bar{x}_1)(y_i - \bar{y}) \sum_{i=1}^n (x_{i2} - \bar{x}_2)^2}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 \sum_{i=1}^n (x_{21} - \bar{x}_2)^2 - \left[\sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \right]^2} - \\
&\quad \frac{\sum_{i=1}^n (x_{i2} - \bar{x}_2)(y_i - \bar{y}) \sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2)}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 \sum_{i=1}^n (x_{21} - \bar{x}_2)^2 - \left[\sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \right]^2} = \\
&= \frac{\gamma^2 \sum_{i=1}^n (x_{i1} - \bar{x}_1)(y_i - \bar{y}) \sum_{i=1}^n (x_{i2} - \bar{x}_2)^2}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 \sum_{i=1}^n (x_{21} - \bar{x}_2)^2 - \left[\sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \right]^2} - \\
&\quad \frac{\gamma^2 \sum_{i=1}^n (x_{i1} - \bar{x}_2)(y_i - \bar{y}) \sum_{i=1}^n (x_{i1} - \bar{x}_1)^2}{\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 \sum_{i=1}^n (x_{21} - \bar{x}_2)^2 - \left[\sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \right]^2} = \frac{0}{0}.
\end{aligned}$$

З останніх виразів маємо, що визначити параметри моделі неможливо, якщо між її факторами існує функціональна залежність.

З цього простого прикладу можна уявити, до чого може призвести не врахування мультиколінеарності між вихідними факторами, які ми з економіко-

логіко-теоретичних міркувань вирішили включити до регресійної моделі.

Але, якщо перед дослідником постає лише задача отримання прогнозу, то наявність мультиколінеарності у масиві пояснювальних даних не завжди може призвести до проблеми щодо її усунення. Річ у тому, що чим більше пояснюючих змінних у регресійному рівнянні, тим вищий коефіцієнт детермінації, а тому і точніший прогноз. Правда, у цьому випадку має здійснюватись умова однакової лінійної залежності з пояснюючими змінними. Однак, ця умова не завжди виконується, особливо у трансформованих економіках. Крім того, якщо коефіцієнт детермінації наближається до одиниці і параметри регресії є значущими, то мультиколінеарність завжди призводить до проблеми, яка обов'язково повинна бути вирішена, оскільки у цьому випадку t -статистика стає великою [8,18].

Однак зазначимо, що в разі мультиколінеарності може існувати проблема через порушення не тільки умов про припущення, які дають нам змогу використовувати МНК, а й деякі інші, а саме:

– незміщеність – це властивість часто повторюваної вибірки. Але навіть за наявності мультиколінеарності ця властивість може бути стійкою;

– колінеарність не порушує властивостей мінімуму дисперсії на класі лінійних незміщених оцінок, тобто вони є ефективними. Однак це не означає, що дисперсія МНК-оцінки буде незмінно малою у будь-який з наведених вибірок;

– мультиколінеарність призводить до того, що неможна виявити індивідуальний вплив кожного регресора на регресанд у моделі, хоча в ній, на наш погляд, повинні бути всі регресори, з точки зору економіко-логіко-теоретичного обґрунтування.

Що ж робити у випадку, коли при дослідженні статистичної бази даних виявлена мультиколінеарність? Якщо число наглядів і вибіркові дисперсії великі, а дисперсія збурень мала, то ми можемо отримати достатньо непогані оцінки. Але проблема все ж таки залишається. Тому у разі мультиколінеарності треба:

а) збільшити ступінь виконання умов Гаусса-Маркова для забезпечення надійності отриманих оцінок;

б) використовувати зовнішню інформацію (теоретичні обмеження та емпіричні оцінки). Якщо спиратися на статистичну інформацію, треба (якщо це можливо) збільшити загальний обсяг вибірки (при цьому отримаємо менші дисперсії оцінок МНК), яку використовували для побудови регресійної моделі. При цьому, зрозуміло, постає друга проблема – знаходження додаткових даних (але тут постає ще інша проблема – вартості інформації). Для часових рядів це можна зробити завдяки, наприклад, скорочення часових інтервалів досліджуваних факторів, або за рахунок об'єднання крос-секцій та часових рядів. Це, однак, може призвести до автокорельованості (це питання більш докладно розглядається у наступному розділі). Крім того, скорочення часових інтервалів може призвести до зміщення отриманих оцінок, або її посилення, якщо зміщення вже існувало. Це пов'язане з тим, що отримані статистичні дані мають похибки вимірювання. Ця проблема більш ускладнена, ніж попередня, але вона може стати несуттєвою при дослідженні;

в) усунути мультиколінеарність за рахунок зміни специфікації моделі, а саме змінити одну або декілька змінних на інші (якщо це не призведе до зміни суті досліджуваного процесу). Крім того, можна виключити ті змінні, які "сильно" корелюють з іншими. Однак при цьому виникає ще одна проблема: що робити, якщо змінні, які ми хочемо виключити з розгляду, були включені у модель на основі теоретичних міркувань? Не можна ж дозволяти "накинути на нас вуздечку" лише для того, щоб зробити статистичні результати кращими з будь-яких "політичних" міркувань?

З вищенаведеного випливає, що перед тим, як використовувати МНК для оцінки параметрів, треба всі фактори, які ми вибрали для побудови моделі (крім залежної змінної), перевірити на мультиколінеарність і, якщо вона існує, позбавитись її. Тому у багатофакторному регресійному аналізі перевірка на мультиколінеарність є першим кроком на етапі первинної обробки вхідної статистичної інформації. Але, нажаль, на цей час не існує універсальних діагностичних тестів та алгоритмів, які б давали змогу повно провести дослідження вихідних даних на мультиколінеарність. Однак, якщо б не було проблеми, то не було б і досліджень щодо її вирішення. Тому на сьогодні вже існує декілька алгоритмів для перевірки на мультиколінеарність. Серед них, який найбільш часто використовується і дає більш повне вирішення цієї задачі є алгоритм Феррара-Глобера [22] і включає три види статистичних критеріїв, на основі яких перевіряється мультиколінеарність: а) всього масиву незалежних змінних (критерій χ^2);

б) кожної незалежної змінної з усіма іншими (F-критерій); в) кожної пари незалежних змінних (t-критерій).

Якщо мультиколінеарність виявлено, то існує декілька методів її усунення. До таких методів належать [18]:

1. Використання первинної інформації.
2. Об'єднання міжгалузевої та динамічної інформації.
3. Вилучення змінної або змінних та помилки специфікації.
4. Перетворення змінних.
5. Збільшення спостережень.
6. Факторний аналіз.
7. Гребенева регресія.
8. Метод головних компонентів.

На закінчення розгляду мультиколінеарності зазначимо, що жоден з розроблених методів для вирішення цієї проблеми не є досконалим. Більш того, застосування кожного з них може призвести до появи додаткових проблем. Так, наприклад, при перетворенні змінних може порушуватись умова про некорельованість похибок, хоча на першому етапі, до застосування методу, вони і були некорельовані. Крім того, при цьому ми втрачаємо один ступінь свободи, що у випадку малої потужності вибірки може призвести до суттєвих помилок оцінок моделі.

Але, якщо, наприклад, за допомогою методу та алгоритму Феррара-Глобера позбутися мультиколінеарності неможливо, то параметри моделі можна спробувати оцінити за одним з перелічених вище методів. Наприклад,

методу головних компонентів, використання якого дає можливість оцінювати параметри моделі, яка має велику вимірність, або фактори, які входять до цієї моделі є мультико-лінійними та іншими відомими засобами позбавитись її неможливо.

Суть методу полягає в тому, щоб перетворити множину змінних X на нову множину, яка б включала множину некорельованих змінних, причому, перша змінна з цієї множини мала максимально можливу дисперсію, друга змінна з цієї множині також мала максимально можливу дисперсію, але в підпросторі, якій є ортогональним до першого і т. ін. Алгоритм цього методу докладного наведений у [22, с.167-168]. Для більш змістовного ознайомлення з поясненнями щодо цього методу треба добре володіти мовою теорії матриць.

7.2 Гомо- та гетероскедастичність

Ми вже зазначали, що для того щоб використовувати МНК потрібно виконання умов Гаусса-Маркова (5.16)...(5.19). Одна з них (5.17) – це умова виконання рівності дисперсій збурень, тобто $M(u, u^{tr}) = \sigma^2 E$ (умова гомоскедастичності). Якщо ця умова порушується (умова гетероскедастичності), то оцінки параметрів регресії, які отримані за звичайним МНК будуть неефективними, хоча і незміщеними та обґрунтованими. Нагадаємо, що обґрунтованість оцінки означає, що вона задовольняє закону великих чисел, тобто чим більші будуються вибірки, тим більша ймовірність того, що помилка оцінки не буде перевищувати достатньо малу наперед задану величину ε , тобто

$$\lim_{n \rightarrow \infty} P \{ |B - \beta| < \varepsilon \} = 1$$

Якщо регресійна модель має лише дві змінні, то явище гетероскедастичності може не впливати на оцінки параметрів за рахунок перетворення певним чином змінних моделі. Однак, якщо економетрична модель має більше ніж дві змінні, то явище гетероскедастичності буде суттєво впливати на оцінки параметрів моделі. Тому необхідно у цьому випадку з'ясувати наявність гетероскедастичності і, якщо це можливо, позбавитись її.

Розглянемо цю проблему більш докладніше. Оскільки гомоскедастичність означає, що варіація кожної похибки навколо її математичного сподівання не залежить від пояснювальної змінної, то у разі порушення цієї властивості можуть бути такі випадки: 1) при збільшенні значень x значень дисперсія залишків зростає; 2) при збільшенні значень x дисперсія залишків спадає; 3) при збільшенні значень x дисперсія залишків починає спадати, але з деякого значення x дисперсія залишків починає збільшуватись.

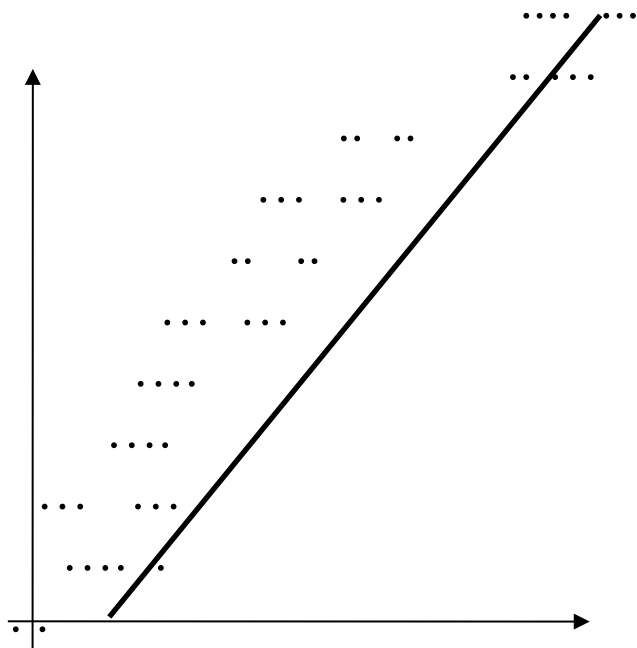


Рис. 7.1.

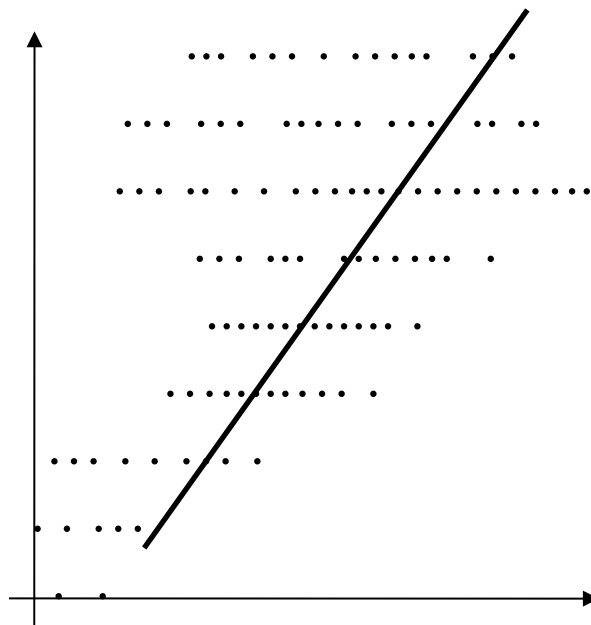


Рис. 7.2.

Випадок гомоскедастичності ілюструється на рис. 7.1. Випадку 1 відповідає рис.7.2

Таким чином, рис. 7.1 відповідає випадку гомоскедастичності, тобто коли $M(u_i^2)=\sigma^2$, а рис. 7.2 відповідає випадку гетероскедастичності, тобто коли $M(u_i^2)=\sigma_i^2$ ($\forall i=1,2,\dots,n$).

Оскільки залишки характеризують неврахований вплив пояснювальних змінних, тому у випадку гетероскедастичності неможна сказати, яку саму ситуацію ми маємо. Більш того, якщо припущення про гомоскедастичність порушується, то співвідношення для обчислення дисперсії параметрів регресії для оцінки їх значущості та побудови інтервалів довіри застосовувати неможливо, оскільки у цьому випадку вони не збігаються з такими самими оцінками при обчисленні їх за умови, що гетероскедастичність відсутня. Це пов'язане з тим, що дисперсія залишків у першому випадку буде змінним числом, а як відомо вона є складовою при обчисленні дисперсій параметрів моделі.

На цей час існує декілька тестів, які дають можливість перевірити модель на наявність гетероскедастичності, а саме:

- графічний метод;
- рангової кореляції Спірмена
- Голдфелда-Квондта
- Глейсера;
- на основі μ -критерію.

На закінчення, ще декілька слів, що стосується проблеми гетероскедастичності.

При застосуванні узагальненого методу найменших квадратів, застосування тих співвідношень, які були одержані за умови, що не існує проблеми гетероскедастичності потребують деякої модифікації. Особливо це стосується обчислення прогнозу. Розглянемо це питання більш докладніше [22, с93-195].

Отже, нехай ми вже сконструювали регресійне рівняння $Y = XA + u$, але для якого умову гомоскедастичності порушено, тобто $M(uu') = \sigma_u^2 S$. Позначимо останню рівність через V . Нам потрібно для заданого вектора X_p знайти значення Y_p .

З урахуванням знайденого рівняння регресії ми можемо записати, що $Y_p = X_p A + u_p$, де u_p - це невідоме значення відхилень у прогностовому періоді. Припустимо, що для них виконуються умови: 1) $M(u_p) = 0$; 2) $M(u_p u_p') = \sigma_u^2 E$. Позначимо вектор коваріацій поточних і прогностичних значень залишків через W , тобто

$$M(u_p u) = \begin{pmatrix} M(u_1 u_p) \\ M(u_2 u_p) \\ \vdots \\ M(u_n u_p) \end{pmatrix} = W.$$

Задачу лінійного прогнозу формально запишемо у вигляді: $p = c'Y$, де c - n -мірний вектор, який має мінімальну дисперсію прогнозу, а саме

$$\sigma_p^2 = M(p - Y_p)^2.$$

З останнього виразу випливає, що найменшого значення воно набуває при $M(p - Y_p)^2 = 0$. З урахуванням викладеного можна записати, що

$$p - Y_p = c'Y - X_p A - u_p = c'(XA + u) - X_p A - u_p = c'XA + c'u - X_p A - u_p = (c'X - X_p)A + (c'u - u_p).$$

Оскільки ми вважаємо, що умова незміщеності для прогнозу виконується, то вектор c повинен задовольняти умові: $c'X = X_p$. Тому помилка прогнозу матиме вигляд: $p - Y_p = c'u - u_p$. З урахуванням того, що $(p - Y_p)$ є скалярна величина, дисперсія прогнозу матиме такий вигляд:

$$\begin{aligned} \sigma_p^2 &= M(p - Y_p)^2 = M[(p - Y_p)(p - Y_p)'] = M[(c'u - u_p)(c'u - u_p)'] = M(c'uu'c - c'uu_p - c'uu_p + u_p^2) = \\ &= M(c'uu'c - 2c'uu_p + u_p^2) = c'Vc + \sigma_0^2 - 2c'W. \end{aligned}$$

Оскільки вірогідність прогнозу буде досягнута, коли дисперсія σ_p^2 буде мінімальною, то задачу сформуємо так: Знайти

$$\sigma_p^2 = c'Vc + \sigma_0^2 - 2c'W \rightarrow \min$$

за умови, що виконується рівність

$$c'X - X_p = 0.$$

Для розв'язання цієї задачі, побудуємо функцію Лагранжа

$$\Psi = c'Vc - 2c'W - 2\lambda(c'X - X_p),$$

де λ – це $(m-1)$ -вимірний вектор, компонентами якого є множники Лагранжа.

Диференціюючи цю функцію по c та λ і прирівнюючи одержані похідні до нуля, дістанемо рівняння:

$$\begin{pmatrix} c \\ \lambda \end{pmatrix} = \begin{pmatrix} V & X \\ X' & 0 \end{pmatrix}^{-1} \begin{pmatrix} W \\ X'P \end{pmatrix}.$$

Розв'язавши його, знаходимо

$$c = V^{-1}[E - X(X'V^{-1}X)^{-1}X'V^{-1}]W + V^{-1}X(X'V^{-1}X)^{-1}X'P.$$

Тому найкращий лінійний незміщений прогноз буде мати такий вигляд

$$p = X_p B + W'V^{-1}(Y - XA)$$

або з урахуванням того, що $B = (X'V^{-1}X)^{-1}X'V^{-1}Y$ одержимо

$$p = X_p B + W'V^{-1}u$$

де $u = Y - XB$ – вектор залишків, який відповідає оцінці параметрів моделі за звичайним МНК.

Зазначимо, що цей прогноз має такі особливості:

– вектор прогнозних значень множиться на вектор оцінок, обчислений згідно з узагальненим МНК;

– для оцінювання невідомих прогнозних залишків u_p застосовується матриця V , яка містить інформацію про взаємозалежність залишків базисного періоду.

РОЗДІЛ 8

АВТОКОРЕЛЯЦІЯ

8.1 Автокореляція. Її наслідки

Ми вже згадували про автокореляцію (автокореляція – це взаємозв'язок послідовних елементів часового чи просторового ряду даних). Як зазначалося у розд. 5, для того щоб можна було застосовувати МНК, потрібно виконання умов Гаусса-Маркова. Однією з цих умов є умова (5.17), яка свідчить, що залишки

u_i – компоненти вектора u повинні бути некорельованими між собою мати постійну дисперсію, тобто $M(uu') = \sigma^2 E$, де E – одинична матриця, u' – транспонований вектор u , σ^2 – стала дисперсія залишків. Однак може бути і така ситуація, коли дисперсія залишків σ^2 стала (гомоскедастичність), але спостерігається їх коваріація. У цьому випадку кажуть, що спостерігається автокореляція залишків. Зазначимо, що автокореляція звичайно розглядається у регресійному аналізі тільки при дослідженні часових рядів.

Деякі випадки існування автокореляції наведені на рис. 8.1 та 8.2 [9, с.218].

Рис. 8.1 відповідає випадку, коли автокореляція залишків позитивна, тобто коли $M(u_i u_{i+1}) = r > 0$

Рис. 8.2 ілюструє випадок, коли автокореляція залишків від'ємна, тобто коли $M(u_i u_{i+1}) = r < 0$ (тут після кожного позитивного значення одного спостереження іде від'ємне значення наступного).

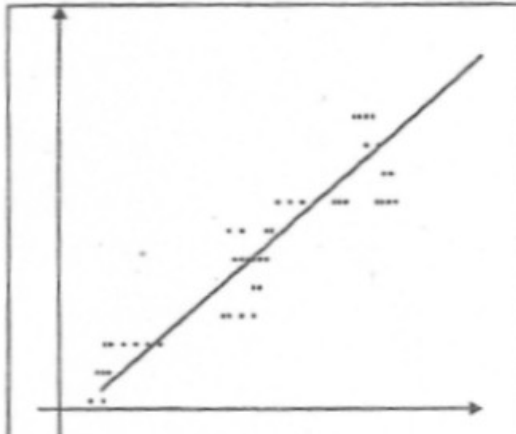


Рисунок 8.1

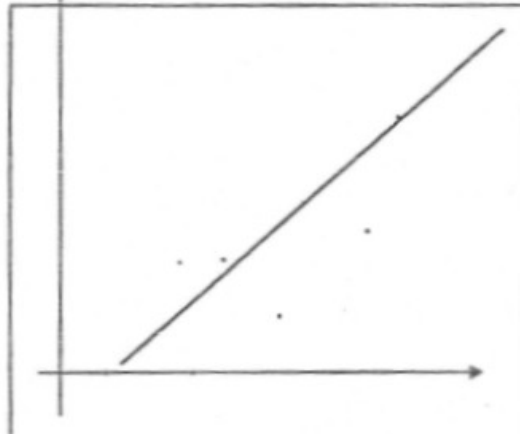


Рисунок 8.2

Розглянемо зараз основні причини, які породжують автокореляцію.

Як зазначається в [22], автокореляція виникає з таких причин:

1) якщо існує кореляція між послідовними значеннями деякої незалежної змінної у часі, то спостерігається і кореляція між послідовними значеннями залишків;

2) як наслідок помилкової специфікації моделі.

У випадку 2 може бути така ситуація, коли необхідно ввести до моделі

нову змінну. Але, зрозуміло, що це неможливо завжди, іноді не бажано, а інколи й *небезпечно* (поміркуйте чому).

Якщо не врахувати автокореляцію залишків при оцінці параметрів моделі, можна зіштовхнутися з таким:

1. Оцінки параметрів моделі можуть бути незміщеними, але неефективними, тобто вибіркові дисперсії векторів оцінок \mathbf{B} можуть бути невиправдано великими.

2. t і F -статистики, які знайдено для моделі, практично не можуть бути використані в дисперсійному аналізі.

3. Можна отримати некоректні прогнози (з великою вибірковою дисперсією) тому, що оцінки параметрів моделі неефективні.

Розглянемо більш докладно, що ми можемо отримати у разі існування автокореляції залишків, наприклад, у разі парної лінійної регресійної моделі. Отже, нехай модель задана у вигляді $y_t = \beta_0 + \beta_1 x_t + u_t$ і для неї виконується умова: $u_t = \rho u_{t-1} + w_t$, $|\rho| < 1$. Крім того, виконуються наступні співвідношення:

$$M(w_t) = 0, \quad (8.1)$$

$$M(w_t w_{t+s}) = \begin{cases} \sigma_w^2, & s = 0 \\ 0, & s \neq 0. \end{cases} \quad (8.2)$$

Залишки u_t запишемо у вигляді наступного співвідношення (у загальному випадку при $t \rightarrow \infty$):

$$\begin{aligned} u_t &= \rho u_{t-1} + w_t = \rho(\rho u_{t-2} + w_{t-1}) + w_t = \dots = \\ &= w_t + \rho w_{t-1} + \rho^2 w_{t-2} + \dots = \sum_{r=1}^{\infty} \rho^k w_{t-k}, \end{aligned}$$

де

$$1) \quad M(u_t) = 0 \text{ (оскільки } M(w_t) = 0);$$

$$2) \quad M(u_t^2) = M(w_t^2) + \rho^2 M(w_{t-1}^2) + \rho^4 M(w_{t-2}^2) + \dots$$

$$= (1 + \rho^2 + \rho^4 + \dots) \sigma_w^2 = \frac{1}{1 - \rho^2} \sigma_w^2;$$

$$3) \quad M(u_t u_{t-1}) = M[(w_t + \rho w_{t-1} + \rho^2 w_{t-2} + \dots)(w_{t-1} + \rho w_{t-2} + \rho^2 w_{t-3} + \dots)] =$$

$$= M\{[w_t + \rho(w_{t-1} + \rho w_{t-2} + \dots)](w_{t-1} + \rho w_{t-2} + \rho^2 w_{t-3} + \dots)\} =$$

$$= \rho \rho M[(w_{t-1} + \rho w_{t-2} + \rho^2 w_{t-3} + \dots)^2] = \rho \sigma_u^2.$$

$$M(u_t u_{t-2}) = \rho^2 \sigma_u^2.$$

$$M(u_t u_{t-s}) = \rho^s \sigma_u^2.$$

Як бачимо, умова про незалежність послідовних збурень не виконується.

Все вищевикладене у загальному вигляді можна записати таким чином:

$$M(uu') = V = \sigma_u^2 \begin{pmatrix} 1 & p & p^2 & p^3 & p^4 & \dots & \dots & p^{n-1} \\ p & 1 & p & p^2 & p^3 & \dots & \dots & p^{n-2} \\ p^2 & p & 1 & p & p^2 & \dots & \dots & p^{n-3} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ p^{n-1} & p^{n-2} & p^{n-3} & p^{n-4} & p^{n-5} & \dots & \dots & 1 \end{pmatrix}$$

З останнього виразу випливає, що умова $M(uu') = \sigma^2 E$ не виконується і тому застосування звичайного МНК неправомірне.

Можна показати, що у цьому випадку дисперсія оцінок параметрів за звичайним МНК відрізняється на деякий множник, якщо її обчислюють з урахуванням автокореляції. Тому за наявності автокореляції у разі застосування звичайного МНК ми втрачаємо цей множник.

Розглянемо випадок існування автокореляції між пояснювальними даними та залишками, тобто

$$x_t = \lambda x_{t-1} + v_t, \quad |\lambda| < 1, \quad (8.3)$$

$$u_t = \rho u_{t-1} + w_t, \quad |\rho| < 1. \quad (8.4)$$

Якщо коефіцієнти λ і ρ при пояснювальних змінних і залишках відповідно – додатні, то кажуть про додатну автокореляцію, якщо навпаки – від'ємну. Зауважимо, що від'ємна автокореляція спостерігається при дослідженні економічних процесів рідко.

Нехай помилки v_t і w_t взаємно незалежні і їх автокореляційні матриці діагональні. Тоді можна показати [8, 22], що звичайній МНК приблизно дає при достатньо великому n таку оцінку дисперсії параметрів B :

$$\sigma^2(B) = \sigma_u^2 (X'X)^{-1} \left(\frac{1+\rho\lambda}{1-\rho\lambda} \right). \quad (8.5)$$

Вираз (8.5) свідчить, що чим більші значення λ і ρ , тим зміщення дисперсії параметрів більше.

Якщо додатна автокореляція спостерігається і в залишках, і в незалежній змінній, то звичайний МНК дає зміщення і для залишкової дисперсії. Припустимо, як і раніше, для x_t і u_t виконуються умови (8.3) і (8.4). Тоді, як показано у [22]:

$$M(uu') = \sigma_u^2 \left(n - \frac{1+\rho\lambda}{1-\rho\lambda} \right)$$

і у цьому разі дисперсію залишків необхідно коригувати на величину зміщення. Але, у цьому випадку, не завжди можуть бути коректні рівні значущості для t і F -статистик.

Як бачимо, перевірка на автокореляцію потрібна завжди. Тому знаходження відповідних критеріїв для перевірки на автокореляцію і у разі її наявності

позбавлення (зрозуміло, якщо це можливо) є актуальним завданням для будь-якого дослідника, який займається моделюванням економічних процесів.

8.2 Критерії Дарбіна-Уотсона (d-статистика) і фон Наймана

Для перевірки на автокореляцію найчастіше використовують d -статистику, або як її часто називають критерій Дарбіна-Уотсона (DW -критерій). Для цього обчислюють d -статистику за формулою

$$d = \frac{\sum_{i=2}^T (u_i - u_{i-1})^2}{\sum_{i=1}^T (u_i)^2}$$

Після чого знаходять верхню (d_u) та нижню (d_n) межі для числа спостережень T , числа незалежних у часі змінних $(m+1)$ – включаючи і фіктивну змінну при вільному члені у регресійному рівнянні і рівня значущості α (з дод. 6, або в будь-якому підручнику з економетрії – d – статистика (DW -статистика) для критерію Дарбіна- Уотсона). Після цього робиться висновок про автокореляцію:

- 1) якщо $0 < d < d_n$ – ряд додатно автокорельований;
 - 2) якщо $d_n < d < d_u$ – необхідні подальші дослідження (необхідно збільшити число спостережень);
 - 3) якщо $d_u < d < 4 - d_u$ – ряд не містить автокореляції;
 - 4) якщо $4 - d_u < d < 4 - d_n$ – необхідні подальші дослідження (необхідно збільшити число спостережень);
 - 5) якщо $4 - d_n < d < 4$ – ряд відхилень від'ємно автокорельований.
- Все викладене можна зобразити графічно (рис. 8.3)

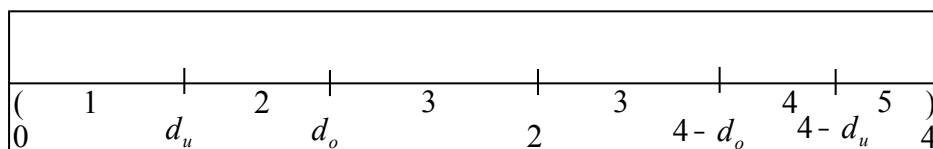


Рисунок 8.3

Таким чином, за відсутності автокореляції в ряді динаміки d звичайно наближене до двох. Останнє твердження використовують тоді, якщо потрібно тільки впевнитися у тому, є автокореляція чи ні.

Крім розглянутого DW -критерію для перевірки на автокореляцію залишків використовується також критерій фон Наймана:

$$Q = \frac{\sum_{t=2}^n (u_t - u_{t-1})^2}{\sum_{t=1}^n u_t^2} \quad \text{або} \quad Q = \frac{\sum_{t=2}^n (u_t - u_{t-1})^2}{\sum_{t=1}^n u_t^2} \cdot \frac{n}{n-1} = DW \cdot \frac{n}{n-1}$$

Зрозуміло, що $\lim_{n \rightarrow \infty} Q = DW$ і тому коли n є досить великим, критерії Дарбіна-Уотсона і фон Неймана збігаються.

Фактичне значення критерію фон Неймана порівнюється з табличним для вибраного рівня значущості α і заданого числа спостережень. Якщо $Q_{\text{факт}} < Q_{\text{табл}}$, то існує додатна автокореляція.

Розглянемо зараз деякі додаткові моменти щодо автокореляції і що стосується оцінки параметрів моделі у цьому випадку. Ці питання досить докладно викладені у [8, 22]. Наведемо їх майже без скорочення.

8.3 Додаткові моменти щодо автокореляції

1. Нециклічний коефіцієнт автокореляції. Він виражає ступінь взаємозв'язку залишків кожного наступного значення з попереднім, а саме ряд: u_1, u_2, \dots, u_{n-1} з рядом u_2, u_3, \dots, u_n і обчислюється за формулою:

$$r^* = \frac{\sum_{t=2}^n (u_t u_{t-1}) - \frac{1}{n-1} \left(\sum_{t=1}^n u_t \right) \left(\sum_{t=2}^n u_{t-1} \right)}{\sqrt{\left[\sum_{t=1}^n u_t^2 - \frac{1}{n-1} \left(\sum_{t=1}^n u_t \right)^2 \right] \left[\sum_{t=2}^n u_{t-1}^2 - \frac{1}{n-1} \left(\sum_{t=2}^n u_{t-1} \right)^2 \right]}}, \quad r^* \in (-1; 1).$$

Якщо $r^* < 0$, то маємо від'ємну автокореляцію, якщо $r^* > 0$ – додатну. Якщо $r^* \approx 0$, то автокореляція відсутня. Однак, імовірнісний розподіл r^* не завжди можна встановити. Тому, часто використовують циклічний коефіцієнт автокореляції.

2. Циклічний коефіцієнт автокореляції. Він виражає ступінь взаємозв'язку рядів: u_1, u_2, \dots, u_n і $u_2, u_3, \dots, u_n, u_n$ та обчислюється за формулою:

$$r^0 = \frac{\sum_{t=2}^n u_t u_{t-1} + u_n u_n - \frac{1}{n} \left(\sum_{t=1}^n u_t \right)^2}{\sum_{t=1}^n u_t^2 - \frac{1}{n} \left(\sum_{t=1}^n u_t \right)^2}$$

Якщо маємо довгі ряди, то вплив циклічних членів на величину коефіцієнта r^0 невизначений і можна вважати, що імовірнісний розподіл r^0 наближається до розподілу r^* . Якщо $u_1 = u_n$, то $r^0 = r^*$.

Знайдене значення r^0 порівнюється з табличним для вибраного рівня значущості α і довжини ряду n . Якщо, $r_{\text{факт.}}^0 \geq r_{\text{тфбл.}}^0$, то існує автокореляція.

Припускаючи, що, $\sum_{t=1}^n u_t \cong \sum_{t=2}^n u_{t-1} \approx 0$, то

$$r^0 = \frac{\sum_{t=2}^n u_t u_{t-1}}{\sum_{t=1}^n u_t^2}.$$

На практиці іноді користуються співвідношенням

$$r^0 = \frac{\frac{1}{n-1} \sum_{t=2}^n u_t u_{t-1}}{\frac{1}{n} \sum_{t=1}^n u_t^2}.$$

Розглянемо зараз обчислення оцінок параметрів регресійної моделі у разі існування автокореляції [22].

8.4 Оцінка параметрів моделі з автокорельованими залишками

Існує декілька методів для оцінки параметрів моделі з автокорельованими залишками. Розглянемо деякі з них.

1. Метод Ейткена

Розглянемо регресійну модель виду:

$$y_t = a_0 + a_1 x_t + u_t, \quad (8.6)$$

$$u_t = \hat{\rho} u_{t-1} + \varepsilon_t, \quad (8.7)$$

де ε_t – нормально розподілені випадкові залишки.

Для усунення автокореляції залишків u_t , перетворимо вираз (8.6) таким чином. Помножимо обидві частини моделі для попереднього періоду $y_{t-1} = a_0 + a_1 x_{t-1} + u_{t-1}$, на $\hat{\rho}$. Отримаємо:

$$\hat{\rho} y_{t-1} = \hat{\rho} a_0 + \hat{\rho} a_1 x_{t-1} + \hat{\rho} u_{t-1}. \quad (8.8)$$

Віднімемо від рівняння (8.6) рівняння (8.8). Одержимо:

$$y_t - \hat{\rho} y_{t-1} = a_0 (1 - \hat{\rho}) + a_1 (x_t - \hat{\rho} x_{t-1}) + (u_t - \hat{\rho} u_{t-1}). \quad (8.9)$$

До виразу (8.9) вже можна застосувати звичайний МНК. Значення $\hat{\rho}$ можна знайти на основі залишків за допомогою циклічного коефіцієнта кореляції r^0 . На практиці, приймається $\hat{\rho} = r^0$, однак r^0 коригується на величину зміщення.

При оцінці параметрів моделі (8.6) (у загальному випадку для m параметрів B моделі) для якої виконується умова (8.7) застосовують, наприклад, узагальнений метод найменших квадратів (метод Ейткена):

$$B = (X' V^{-1} X)^{-1} X' V^{-1} Y \quad (B = (X' S^{-1} X)^{-1} X' S^{-1} Y),$$

де: $V = \sigma_{\varepsilon}^2 S$; S – матриця кореляції залишків.

Запишемо загальний вигляд матриці S :

$$S = \begin{pmatrix} 1 & p & p^2 & p^3 & p^4 & \dots & \dots & p^{n-1} \\ p & 1 & p & p^2 & p^3 & \dots & \dots & p^{n-2} \\ p^2 & p & 1 & p & p^2 & \dots & \dots & p^{n-3} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ p^{n-1} & p^{n-2} & p^{n-3} & p^{n-4} & p^{n-5} & \dots & \dots & 1 \end{pmatrix},$$

де ρ^s – коефіцієнт автокореляції s -го порядку для залишків u_t . Можна вважати, що $\rho^s = 0$ при $s > 2$. Тому матриця S набуде такого вигляду:

$$S = \begin{pmatrix} 1 & \rho & \rho^2 & 0 & 0 & \dots & \dots & 0 \\ \rho & 1 & \rho & \rho^2 & 0 & \dots & \dots & 0 \\ \rho^2 & \rho & 1 & \rho & \rho^2 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & \dots & 1 \end{pmatrix},$$

а обернена до неї – набуде вигляду:

$$S^{-1} = \frac{1}{1-\rho^2} \begin{pmatrix} 1 & -\rho & 0 & 0 & 0 & \dots & \dots & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & 0 & \dots & \dots & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & \dots & 1 \end{pmatrix}.$$

Для обчислення ρ можна скористатися циклічним коефіцієнтом кореляції. Скоригуємо r^0 на величину його зміщення, отримаємо:

$$r_{\text{скор.}}^0 = \frac{n}{n-1} * \frac{\sum_{t=2}^n u_t \cdot u_{t-1}}{\sum_{t=1}^n u_t^2} + \frac{m+1}{n} (r_{\text{скор.}}^0 = r^0 + \frac{r + \lambda}{(n-1) - \frac{1+r\lambda}{1-r\lambda}}),$$

де $\frac{m+1}{n}$ – величина зміщення (m – кількість незалежних змінних).

Тому, дисперсія залишків $\sigma_u^2 = \frac{1}{n-m-1} u' u$ для матриці $V = \sigma_u^2 S$ набуде вигляду: $\sigma_u^2 = \frac{1}{n-m-1} u' u \left[n - \frac{1+\lambda\rho}{1-\lambda\rho} \right]$, де λ можна обчислити з рівняння $x_t = \lambda x_{t-1} + \varepsilon_t$ використовуючи звичайний МНК.

2. Метод перетворення вихідної інформації.

Метод передбачає: перетворення вихідної інформації і використання для оцінки параметрів на основі перетворених даних звичайного МНК.

Для цього знаходять таку матрицю перетворення T , щоб для регресійної моделі $TU = TXA + Tu$, $M(Tu, u^{tr} T') = \sigma^2 E$.

Як матрицю (розміром $n \times m$) вибирають

$$T_1 = \begin{pmatrix} \sqrt{1-\rho^2} & 0 & 0 & \dots & 0 \\ -\rho & 1 & 0 & \dots & 0 \\ 0 & -\rho & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix},$$

для якої виконується умова:

$M(T_1 u, u^{tr} T_1') = \sigma^2 E$. Тому (як вже відомо) можна застосовувати звичайний МНК (у нашому випадку до даних, які перетворені – $T_1 Y$):

$$T_1 Y = \begin{pmatrix} \sqrt{1-\rho^2} y^2 \\ y_2 - \rho y_1 \\ y_3 - \rho y_2 \\ \dots \\ y_n - \rho y_{n-1} \end{pmatrix},$$

$$T_1 X = \begin{pmatrix} \sqrt{1-\rho^2} & \sqrt{1-\rho^2} x_1^1 & \sqrt{1-\rho^2} x_1^2 & \dots & \sqrt{1-\rho^2} x_1^m \\ 1-\rho & x_2^1 - \rho x_1^1 & x_2^2 - \rho x_1^2 & \dots & x_2^m - \rho x_1^m \\ 1-\rho & x_3^1 - \rho x_2^1 & x_3^2 - \rho x_2^2 & \dots & x_3^m - \rho x_2^m \\ \dots & \dots & \dots & \dots & \dots \\ 1-\rho & x_n^1 - \rho x_{n-1}^1 & x_n^2 - \rho x_{n-1}^2 & \dots & x_n^m - \rho x_{n-1}^m \end{pmatrix}.$$

Якщо застосування вищезгаданих методів утруднено через відсутність інформації про авторегресійну модель, користуються наближеними методами Кочрена-Оркатта і Дарбіна.

3. Метод Кочрена-Оркатта

Метод призначений для наближеного пошуку параметрів моделі $y_t - \rho y_{t-1} = a_0(1-\rho) + a_1(x_t - \rho x_{t-1}) + \varepsilon_t$, за якого мінімізується величина $\sum_{t=2}^n \varepsilon_t^2$, яка є глобальним оптимумом. Альтернативний до нього – Дарбіна-Уотсона. Останній ґрунтується на ідеї припинення ітерації наближеного пошуку, коли на основі критерію Дарбіна-Уотсона робиться висновок про відсутність автокореляції залишків. І, на закінчення, наведемо ще один алгоритм, який дає можливість отримати оцінки параметрів авторегресійної моделі [22].

4. Метод Дарбіна. Метод включає два кроки.

Крок 1. Залишки $u_t = \rho u_{t-1} + \varepsilon_t$ ($|\rho| < 1$) підставимо в рівняння: $y_t = a_0 + a_1 x_t + u_t$, $t = 1, 2, \dots, n$. Одержимо: $y_t = a_0 + a_1 x_t + \rho u_{t-1} + \varepsilon_t$, де $u_{t-1} = y_{t-1} - a_0 - a_1 x_{t-1}$. Тому $y_t = a_0(1-\rho) + \rho y_{t-1} + a_1 x_t - a_1 \rho x_{t-1} + \varepsilon_t$.

За звичайним МНК визначимо параметри цієї моделі. У результаті обчислень маємо ρ .

Крок 2. Значення $\rho = r$ використовуємо для перетворення змінних ($y_t - r y_{t-1}$) і ($x_t - r x_{t-1}$), а звичайний МНК застосовуємо до перетворених даних.

Коефіцієнт при $(x_t - rx_{t-1})$ є оцінкою a_1 , а вільний член, поділений на $-r$ – оцінкою a_0 .

Розглянутий метод може бути поширений і для більш загальних випадків.

Наведемо деякі результати порівняльного аналізу (Грилїхес, Рао) розглянутих методів зі звичайним МНК:

- МНК дає менш ефективні оцінки, якщо сукупність спостережень $n = 20$, а $|\rho| > 0,3$;
- якщо $|\rho| < 0,3$, то ефективність оцінок приблизно однакова.

РОЗДІЛ 9

СІТКОВЕ ПЛАНУВАННЯ ТА УПРАВЛІННЯ

9.1 Основні поняття та правила побудови сіткових графіків

В основі сіткового планування та управління (СПУ) лежить сіткова модель – це графічне зображення плану, яке називають сітковим графіком.

Сітковий графік є схемою, на якій у певному порядку наочно показано всі операції зі створення будь-якого об'єкта або здійснення будь-якої програми діяльності.

Основними поняттями, які використовують при побудові сіткової моделі, є робота, подія та шлях.

Роботами називають будь-які процеси, дії, то сприяють досягненню певних результатів (виготовлення креслень, монтаж обладнання тощо). Роботою слід вважати також «очікування», яке потребує витрат часу; існують так звані фіктивні роботи.

Фіктивною роботою (залежністю) називають зв'язок між будь-якими результатами робіт, що не потребують витрат часу.

На сітковому графіку роботи зображують стрілками. Роботи, що потребують витрат часу, зображують суцільними лініями, фіктивні роботи – пунктирними,

Подія є підсумком якоїсь діяльності, проміжним або закінченим результатом виконання однієї або кількох попередніх робіт, що дозволяє розпочати виконання подальших робіт.

Подія на відміну від роботи не є процесом. Вона не має довжини і не супроводжується затратами часу та коштів. На сітковому графіку події зображують кружечками з порядковим номером.

Шлях – це послідовність робіт на сітковому графіку, в якій завершення кожної попередньої роботи збігається з початком подальшої.

Взаємозв'язок кружечків і стрілок, що є графічними символами сіткової моделі, потрібно здійснювати за певними правилами:

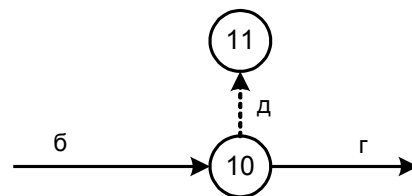
- 1) сітку викреслюють зліва направо;
- 2) нумерують події зліва направо та зверху вниз, тобто кожна подія з більшим порядковим номером позначається правіше або нижче від попередньої. Рекомендується нумерувати події після побудови сітки;
- 3) стрілка передбачає лише логічний взаємозв'язок; як правило, ні довжина стрілки, ні її напрям не мають значення. Проте їм слід все-таки надавати спільного напрямку також зліва направо;
- 4) при виконанні робіт, що йдуть одна за одною (а, б, в), кожен наступну роботу можна розпочинати лише після виявлення результатів попередніх робіт, тобто після здійснення визначеної події. На сітці це слід зобразити так:
- 5) якщо здійснення будь-якої події дає змогу розпочати кілька робіт (а, б, в), то на графіку це зображують так:

6) якщо будь-яка подія може статися лише внаслідок виконання ряду робіт (а, б, в), то на сітці це зображують так:

7) якщо в попередньому випадку для початку роботи а не треба очікувати здійснення події б, а можна обмежитися проміжним результатом, то його потрібно подати у вигляді самостійної події 5, і робота має починатися від неї, отже,

8) якщо дві роботи а і б мають початися після здійснення однієї й тієї самої події і завершитися одночасно наступною подією, то необхідно ввести додаткову подію і фіктивну роботу в:

9) якщо робота в йде після двох робіт а і б, а за роботою б – робота г, яка не залежить від а, то для зображення такого положення також звертаються до фіктивної роботи д:



Будувати сіткові графіки потрібно дуже уважно. Оскільки на графіку логічно пов'язані в єдину сітку всі роботи та події, невірне зображення навіть окремих робіт може призвести до великих помилок у наступних розрахунках.

Розглянемо деякі найхарактерніші з помилок:

1) у сітці не повинно бути «тупикових» подій (подія 7), тобто тих, від яких не починається жодної роботи (за винятком завершальної події всього комплексу робіт, наприклад події 10). Поява «тупикової» події свідчить або про припущення помилки в побудові графіка, або про те, що робота, яка передувала цій події, виявилася непотрібною;

2) у сітці не повинно бути так званих хвостових подій (подія 5), яким не передують жодна інша робота (за винятком вихідної події сітки). Ця подія виникла або в результаті помилки, або внаслідок того, що якісь роботи виявилися пропущеними;

3) у сітці не повинно бути і «замкнених контурів», тобто шляхів, які повертаються до вихідної події. Такий контур проходить через події 2, 3, 4, 2. Логічно він свідчить про те, що робота 2 може розпочатися лише після завершення роботи 2. «Замкнений контур» може виявитися лише внаслідок логічної помилки, яку після перевірки слід виправити.

Отже, з огляду на викладене розпочинаючи будувати сітку, необхідно спочатку встановити:

- які роботи мають бути завершені раніше, ніж розпочнеться дана робота;
- які роботи можуть бути розпочаті після завершення даної роботи;
- які роботи можуть бути виконані водночас із даною роботою.

З'ясувавши цю логічну інформацію, за допомогою правила побудови можна розпочинати будову сіткового графіка.

Нехай унаслідок аналізу взаємозв'язку робіт, за допомогою яких створюють будь-який об'єкт, здобуто таку залежність між ними, як наведено в табл. 9.1.

Таблиця 9.1 – Аналіз взаємозв'язку робіт

Робота	Попередні роботи	Наступні роботи
а	–	Г
б	–	в, д
в	б	Г
Г	а, в	е, ж
д	б	ж
е	Г	–
ж	Г, д	–

Тоді згідно з правилами побудови сітки маємо

Оскільки робота г залежить також від роботи в, то слід перебудувати графік:

У цьому разі знову слід накреслити графік, оскільки двічі викреслено роботу ж, яку можна розпочинати після завершення робіт г і д:

Така будова невірна, оскільки на графіку робота е виходить залежною від робіт г і д, а насправді вона залежить лише від роботи г.

У цьому разі за правилами побудови рекомендується скористатися фіктивною роботою. Тоді графік набуде такого вигляду:

Оскільки після робіт е і ж немає жодної роботи, то вони завершуються завершальною подією, отже, сітковий графік має вигляд:

Пронумеруємо події отриманого графіка. Отже, комплекс, заданий у табл. 9.1, показано у вигляді сіткового графіка.

9.2 Параметри сіткового графіка та способи їх розрахунку

Як зазначалося, усі роботи, крім фіктивних, потребують певних витрат часу. Існують два способи застосування часових оцінок – детерміністичний і ймовірнісний (стохастичний)

Детерміновану оцінку застосовують тоді, коли час виконання робіт можна визначити досить точно (за наявності довідково-нормативних матеріалів). У цьому разі для кожної роботи вказують одну часову оцінку.

Ймовірнісні оцінки застосовують тоді, коли є досить велика невизначеність часу виконання робіт.

Довжиною шляху є сума термінів виконання його робіт. Найдовший шлях називають критичним. Він визначає час, необхідний для виконання програми всіх робіт, що входять до сіткового графіка. Критичний шлях є одним із найважливіших розрахункових параметрів сіткового графіка.

Роботи, що перебувають на критичному шляху, звичайно виділяють жирними або подвійними стрілками.

Властивості, характерні для робіт критичного шляху, а також можливості, які відкриваються при використанні цих властивостей, значною мірою і визначають ефективність сіткових графіків.

1. Зміна протяжності будь-якої роботи, що лежить на критичному шляху, відповідним чином змінює (скорочує або подовжує) термін настання завершальної події, тобто дату досягнення кінцевої мети, що ставиться при плануванні розробки.

2. При плануванні комплексу операцій за критичним шляхом можна визначити термін настання завершальної події.

3. У процесі управління ходом розробки керівництво зосереджує увагу на головному напрямку – на роботах критичного шляху. Це дає змогу ефективно та оперативно контролювати обмежену кількість робіт, що впливають на термін розробки.

Нехай у наведеному раніше прикладі тривалість робіт характеризується такими даними:

Таблиця 9.2 – Аналіз взаємозв'язку робіт

Робота	Тривалість роботи, днів (Т)
а	7,16
б	4,16
в	2,0
г	5,0
д	6,83
е	3,16
ж	5,0

Тоді на сітковому графіку, який побудовано раніше, запишемо T над відповідними роботами:

Тут можна відокремити такі шляхи:

1-3-4-6= 15,32 дня;

1-3-4-5-6= 17,16 дня;

1-2-3-4-5-6= 16,16 дня; 1-2-3-4-6- 14,32 дня;

1-2-5-6= 15,99 дня.

Різницю між довжиною критичного шляху та будь-якого іншого шляху називають резервом часу шляху; ця різниця показує, наскільки в сумі може збільшитися тривалість усіх робіт, що належать даному шляху, без зміни терміну виконання програми.

Крім критичного шляху та резерву часу шляху найважливішими параметрами сіткового графіка є також резерви часу подій і робіт.

Ці параметри є вихідними для отримання ряду додаткових характеристик, а також для аналізу сіткового графіка.

Існує ряд методів, за допомогою яких можна розрахувати параметри сіткової моделі. Якщо подій небагато, найзручнішим є метод обчислень на сітковому графіку. При цьому кожний кружечок, що зображує подію, поділяють на чотири сектори.

Верхній сектор відводиться для номера події, лівий – для розрахованих ранніх термінів завершення події T_p , нижній – для номера події, через яку до даної події проходить шлях, максимальний за довжиною, правий – для розрахунку пізніх термінів завершення події T_n .

Ранній термін здійснення події – це термін, потрібний для виконання всіх робіт, що передують даній події. Ранній термін здійснення події визначається для кожної події послідовним переходом від найбільш ранніх подій до пізніших, тобто зліва направо. При цьому згідно з визначенням слід вибирати найдовший із попередніх даній події шлях.

Пізній термін здійснення події – це такий термін, перевищення якого спричинить відповідну затримку настання завершальної події. Пізній термін здійснення події визначається протилежним ходом – від завершальної події до вихідної, тобто справа наліво. Пізній термін завершення події є різницею між протяжністю критичного шляху та найдовшого з наступних за даною подією шляхів.

Для графіка, розглянутого раніше, це матиме такий вигляд:

Різницю між пізнім T_n і раннім T_p термінами здійснення події називають резервом часу події. Він показує, наскільки може бути відстроювано здійснення цієї події без зміни загального терміну виконання програми. Оскільки резерв

часу подій, що лежать на критичному шляху, дорівнює нулю, найпростішим способом виявлення критичного шляху є визначення всіх послідовно розміщених подій, резерв часу яких дорівнює нулю.

Знаючи ранні та пізні терміни здійснення подій, можна для будь-якої роботи визначити ранні та пізні терміни початку та завершення роботи, а на основі цього розрахувати резерви часу робіт.

Якщо подію, яка передуює роботі, позначити I , а наступну подію J , то кожну роботу можна позначити $i - j$.

Найбільш ранній із можливих термінів початку роботи

$$T_{pi-j} = T_{pi}$$

Найпізніший із допустимих термінів початку цієї роботи

$$T_{pi-j} = T_{nj} - t_{i-j}$$

Найпізніший із допустимих термінів завершення роботи

$$T_{pzi-j} = T_{nj}$$

Найбільш ранній із можливих термінів завершення роботи

$$T_{pzi-j} = T_{pi} + t_{i-j}$$

Наприклад, як було з'ясовано, усі шляхи, крім критичного, мають визначений резерв часу. Резерв часу шляху показує, наскільки в сумі може бути збільшена тривалість усіх робіт, що належать даному шляху, без зміни критичного шляху.

Отже, роботи, що не лежать на критичному шляху, також мають резерв часу.

Існують поняття загального (повного) та особистого (вільного) резервів часу роботи.

Повний резерв часу роботи – це максимальний час, на який можна подовжити дану роботу, не змінюючи при цьому протяжність критичного шляху:

$$R_{ni-j} = T_{nj} - T_{pi} - t_{i-j}$$

Важлива властивість повного резерву часу роботи полягає в тому, що коли його використовувати частково або повністю для збільшення тривалості будь-якої роботи, то відповідно зменшиться резерв часу всіх інших робіт, що

лежать на цьому шляху.

Вільний (особистий) резерв часу це максимальна кількість часу, на який можна збільшити її тривалість або перенести початок роботи без зміни раннього початку наступних робіт:

$$R_{bi-j} = T_{pj} - T_{pi} - t_{i-j}$$

Особистий резерв є незалежним і його використання не впливає на терміни виконання інших робіт.

Розглянуті параметри сіткових графіків (критичний шлях, резерви часу шляху, події, робіт) використовують розроблювачі та керівництво для аналізу і оптимізації сіткових графіків.

9.3 Аналіз та оптимізація сіткового графіка

Розрахований критичний шлях, що характеризує час виконання всіх робіт, які входять у сітковий графік, порівнюється із заданим директивним терміном виконання всіх робіт.

Для детермінованого сіткового графіка достатньо, щоб критичний шлях за тривалістю був меншим від директивного терміну.

Для стохастичного (ймовірнісного) сіткового графіка внаслідок наявності дисперсії оцінок виконання робіт, з яких складається критичний шлях, навіть якщо критичний шлях менший від директивного, необхідно розрахувати ймовірність успіху, тобто ймовірність того, що відбудеться саме так, як це запроектовано.

У загальному вигляді задача зводиться до визначення ймовірності P потрапляння неперервної випадкової величини x в інтервал (a, b) .

Випадковою величиною є термін здійснення завершальної події T_k .

Розподіл величини T_k виявляється близьким до нормального за такого припущення. Тривалість робіт, що лежать на критичному шляху, є незалежними випадковими величинами. У теорії ймовірностей встановлюється, що нормальний закон є граничним для суми таких складових.

Для нормального закону розподілу диференціальна функція розподілу має вигляд

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}}$$

де a – математичне сподівання неперервної випадкової величини (для даного випадку $a = T_k$ розрахункове); σ – середньоквадратичне відхилення, оскільки дисперсія суми випадкових величин дорівнює сумі дисперсій складових, для критичного шляху $\sigma_k = \sqrt{\sum \sigma_i^2}$, а загальна дисперсія дорівнює сумі дисперсій робіт критичного шляху.

Тоді

$$P(\alpha < x < \beta) = \frac{1}{\sigma \sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{(x-a)^2}{2\sigma^2}} dx$$

або

$$P(\alpha < x < \beta) = \frac{1}{2} \cdot \left[\Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right) \right]$$

Згідно з припущенням про нормальний закон розподілу величини T_k можна розрахувати межі α і β (3 зміни очікуваного терміну здійснення завершальної події за «правилом трьох сигм». Оскільки центром розсіювання в даному випадку є T_k то межі зміни очікуваного терміну завершення робіт дорівнюють $(T_k - 3\sigma)$ і $(T_k + 3\sigma)$; при $T_q < T_k - 3\sigma$ імовірність завершення події в заданий термін дорівнює нулю, а при $T_q > T_k + 3\sigma$ ця імовірність дорівнює одиниці. Отже, має сенс розглядати задачу лише при

$$T_k - 3\sigma < T_q < T_k + 3\sigma.$$

Отже, задачу можна сформулювати в такому вигляді: яка імовірність того, що час здійснення завершальної події, то є випадковою величиною x ; лежить у межах $\alpha = T_k - 3\sigma; \beta = T_q$, тоді

$$P = \frac{1}{2} \cdot \left[\Phi\left(\frac{T_q - T_k}{\sigma}\right) - \Phi\left(\frac{T_k - 3\sigma - T_k}{\sigma}\right) \right]$$

Згідно з даними таблиці значень функції Лапласа отримаємо:

$$P = \frac{1}{2} \cdot \left[\Phi\left(\frac{T_q - T_k}{\sigma}\right) + 0,9973 \right] \quad (9.1)$$

Якщо критичний шлях перевищує заданий директивний термін (для детерміністичних графіків), або мала ймовірність успіху (для стохастичних графіків), необхідно коригувати сітковий графік з метою зведення його параметрів згідно із заданими термінами.

Процес коригування сіткового графіка іноді називають оптимізацією, вважаючи під цим. послідовне поліпшення сітки з метою досягнення заданого терміну виконання робіт.

Існує кілька способів приведення сіткового графіка у відповідність із заданими термінами:

- зміна часу виконання робіт за рахунок перерозподілу ресурсів, як часових, так і матеріальних;
- зміна часу виконання робіт за рахунок інтенсифікації виконання робіт критичного шляху (додаткова кількість виконавців, обладнання, матеріальне стимулювання тощо);
- розподіл робіт та їх сумісність у часі;
- конструктивна зміна комплексу робіт, зміна графічного зображення внаслідок перегляду технології виконання робіт.

Загальний термін виконання робіт слід скорочувати насамперед завдяки зміні тривалості робіт критичного шляху. Це один із найпоширеніших прийомів, оскільки не пов'язаний зі зміною графічного зображення (топології) сітки.

Зменшення часових оцінок за критичними роботами забезпечується насамперед за рахунок перекидання відповідних ресурсів (людей, механізмів тощо) з робіт, що мають великі резерви часу.

Унаслідок перерозподілу ресурсів тривалість одних робіт зменшується, а

інших – збільшується; отже, необхідно знову розрахувати часові параметри сіткового графіка.

Якщо внутрішніх ресурсів недостатньо, слід ставити питання про залучення їх зовні. Наприклад, увести другу та третю зміну тощо.

Третій спосіб скорочення тривалості виконання робіт – паралельне їх виконання, що досягається розподілом робіт більшої тривалості. У ряді випадків для цього також потрібні додаткові ресурси, наприклад додаткова бригада для паралельного виконання однорідних робіт.

Якщо не вдається повною мірою зменшити термін виконання розробки за рахунок формування робіт, то звертаються до зміни графічного зображення (топології) сітки. Це можливо завдяки тому, що окремі роботи можна виконувати різними методами. Багатоваріантна технологія дає змогу відшукати нову послідовність виконання робіт і нові взаємозв'язки. При цьому необхідно повністю виконати всі роботи, пов'язані з побудовою та розрахунком сіткового графіка.

Якщо після всіх вжитих заходів щодо скорочення тривалості виконання робіт директивного терміну не досягнуто, це свідчить про його нереальність і перед вищестоящою організацією ставиться питання про зміну директивного терміну.

Визначення можливого терміну виконання всіх робіт, якого можна досягти з імовірністю P , у цьому разі розглядають як задачу, протилежну тій, яку розглянуто раніше. Задаючись бажаною імовірністю P із рівняння (9.1) можна визначити значення функції $\Phi\left(\frac{T_q - T_k}{\sigma}\right)$, знаючи T_q і σ , підібрати T_k .

Раніше вже було коротко розглянуто оптимізацію графіка за часом. Можливі також інші критерії оптимізації: «за вартістю», «час – вартість», «надійність», «завантаженість персоналу» та ін.

Таким в загальному вигляді виявляється СПУ. Підбиваючи підсумки, слід зупинитися на перевагах і недоліках СПУ.

Переваги СПУ:

Наочність зображення. Багато авторів стверджують: якщо з усіх переваг СПУ використати лише принцип зображення процесу у вигляді сітки, то й тоді матимемо великий ефект.

Можливе використання сітки з будь-яким бажаним ступенем деталізації (у керівника укрупнена сітка, чим нижчий рівень, тим більший ступінь деталізації).

Спочатку увага керівників зосереджується на найважливіших ділянках (роботи, що перебувають на критичному шляху), другорядні та мало важливі роботи вилучаються з фокуса уваги керівника. Високий ступінь концентрації управлінських зусиль забезпечує ефективність застосовуваних заходів.

Можливість визначення імовірності успіху.

Можливість оптимізації робіт, висока динамічність методу. Порівнюючи звичайні методи управління із сітковим, спеціалісти проводять аналогію з артилерійським снарядом і керованою ракетою.

Після вильоту снаряда з дула гармати керувати траєкторією його польоту

неможливо. Керована ж ракета може змінювати напрям залежно від змінюваних умов польоту, виявлених помилок прицілювання або переміщення цілі. Такою самою є відмінність між звичайними методами управління та сітковим.

Діапазон використання цього методу дуже широкий – будівництво, реконструкція, капітальні ремонти, створення об'єктів нової техніки, організація дослідницьких робіт та ін.

Найповніше переваги методу виявляються тоді, коли створюється щось нове, невипробуване, що погребує розв'язання багатьох задач, пов'язаних як одна з одною, так і із загальною метою.

Недоліки методу:

1. Складність графічного зображення. Велика кількість кружечків і стрілок може призвести до втрати наочності, тобто однієї з найголовніших переваг методу.

2. Застосування сіткового графіка порушує встановлені організаційні зв'язки (традиції), що склалися на підприємствах чи в організаціях.

3. Графік координує і кооперує зусилля всупереч налагодженим відносинам і часто бажанням окремих керівників.

4. Застосування сіткового графіка неможливе без встановлення строгої дисципліни на всіх ділянках запланованого процесу (зменшення ступенів свободи).

Найменше відхилення від встановлених правил може призвести до дискредитації цього дуже прогресивного способу планування та управління складними процесами. Саме такими є загальні положення СПУ.

ДОДАТОК А

А.1 ОСНОВНІ ПОНЯТТЯ З ТЕОРІЇ ГРАФІВ

А.1.1 Основні поняття теорії графів

Теорія графів є розділом математики, що має широкі практичні додатки. Багато проблем, що виникають в таких дуже різних областях знань, як психологія, управління, планування, хімія, електроніка, торгівля, освіта і ін., можуть бути

сформульовані як задачі теорії графів. Зважаючи на це теорія графів цікава не тільки сама по собі, але також і тим, що представляє загальну основу, на якій результати, отримані в різних областях знань, можуть бути зібрані, узагальнені і поширені.

На відміну від інших наукових дисциплін, теорія графів має цілком певну дату народження. Перша робота по теорії графів, написана математиком Леонардом Ейлером, була опублікована в 1736 році в працях Академії наук в Санкт-Петербурзі. Ним була сформульована і розв'язана задача про Кенігсберзькі мости [3].

До 50-х років нашого століття в теорії графів склалися два цілком різних напрями: алгебраїчний і оптимізаційний. Останній набув широкого розвитку завдяки появі ЕОМ і у зв'язку з розробкою методів лінійного програмування. У даному навчальному посібнику розглядатимемо оптимізаційний напрям теорії графів.

Граф G задається безліччю точок або вершин x_1, x_2, \dots, x_n (які позначимо X) і безліччю ліній або ребер (гілок) a_1, a_2, \dots, a_m (які позначимо A), об'єднуючих між собою всі або частину цих точок. Таким чином, граф G повністю задається парою (X, A) , тобто $G = (X, A)$ [4]. Замість позначень X і A можна використовувати й інші.

Якщо ребра з множини A орієнтовані, що звичайно показується стрілкою, то граф називається *орієнтованим* графом, або *орграфом*. Ребра орграфа звичайно називають дугами. Орграф наведений на рис. А.1.1. Якщо ребра не мають орієнтації, то граф називається *неорієнтованим* (рис.А.1.2). Граф, що має як орієнтовані, так і неорієнтовані ребра, називається *змішаним* (рис.А.1.3). Якщо в графі між двома будь-якими вершинами є декілька ребер, то такий граф називається *мультиграфом* (рис.А.1.4).

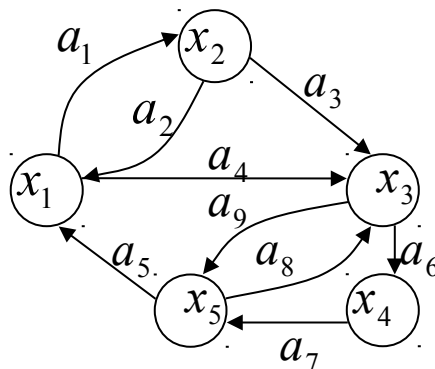


Рисунок А.1.1 – Зображення орграфа

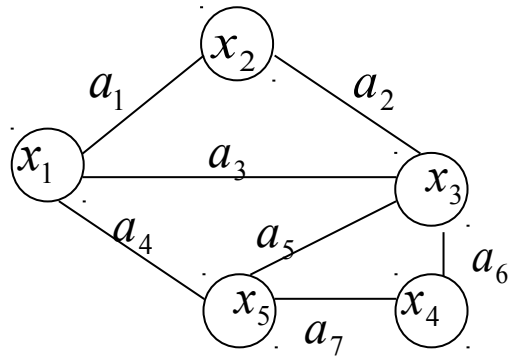


Рисунок А.1.2 – Зображення неорієнтованого графа

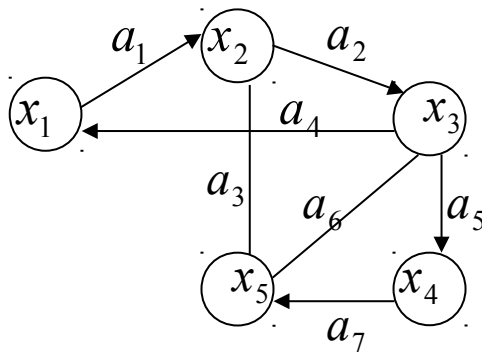


Рисунок А.1.3 – Зображення змішаного графа

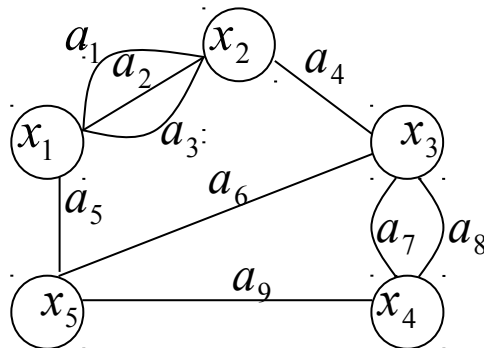


Рисунок А.1.4 – Зображення мультиграфа

Зазвичай вершинам графа приписуються номери (або позначення x_i, x_j), щоб розрізнити їх і мати нагоду ними оперувати в процесі аналізу. Ребра графа позначаються звичайно $a = (x_i, x_j)$, де x_i, x_j – початкова і кінцева вершини, зв'язані даним ребром. Окрім цього позначення, застосовують також $a(x_i, x_j)$ або $a(i, j)$, або a_{ij} . Іноді ребро (дугу) позначають (i, j) .

У використовуваних позначеннях i, j – номери вершин. Крім того, ребра можна позначати a_i , де i – номер ребра.

Шляхом μ_{ij} графа називається послідовність ребер, що починається у вершині i , закінчується у вершині j , у якій кінцева вершина кожного ребра, відмінного від останнього, є початковою вершиною наступного.

Так, на рис. А.1.5 послідовності дуг $a_6, a_5, a_9, a_8, a_4; a_1, a_6, a_5, a_9$ є шляхами.

Ці послідовності можна також записати у формі:

$$a(x_2, x_5), a(x_5, x_4), a(x_4, x_3), a(x_3, x_5), a(x_5, x_6);$$

$$a(x_1, x_2), a(x_2, x_5), a(x_5, x_4), a(x_4, x_3).$$

Більш простою є така форма запису:

$$a_{25}, a_{54}, a_{43}, a_{35}, a_{56};$$

$$a_{12}, a_{25}, a_{54}, a_{43}.$$

Шлях називається *простим*, або *само непересічним*, якщо в ньому не повторюються ні вершини, ні ребра.

У розглянутому прикладі другий шлях є простим.

Цикл – це шлях, початкова і кінцева вершини якого співпадають.

Граф називається зв'язковим, якщо для будь-яких різних вершин існує, принаймні, один, який їх з'єднує.

В орграфі шлях називають також ланцюгом.

Нагадаємо, що граф – це безліч вершин і ребер, позначається $G = (X, A)$. Будь-яка підмножина множини G називається *підграфом*, позначається $\bar{G} = (\bar{X}, \bar{A})$.

Дерево графа визначається як зв'язна підмножина \bar{G} , що не містить циклів. Отже, для будь-яких двох вершин графа в дереві існує єдиний шлях, що сполучає їх.

У графі, що містить n вершин, підграф із k вершин ($k \leq n$) є деревом, якщо виконані будь-які дві з наступних умов:

- 1) підграф є зв'язковим;
- 2) підграф не має циклів;
- 3) число ребер в підграфі дорівнює $k - 1$.

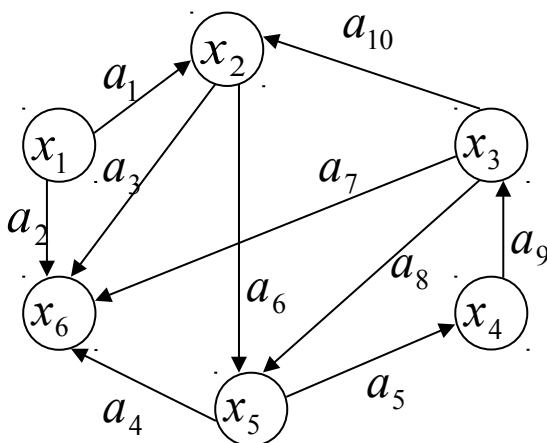


Рисунок А.1.5 – Зображення послідовності дуг $a_6, a_5, a_9, a_8, a_4; a_1, a_6, a_5, a_9$, які є шляхами

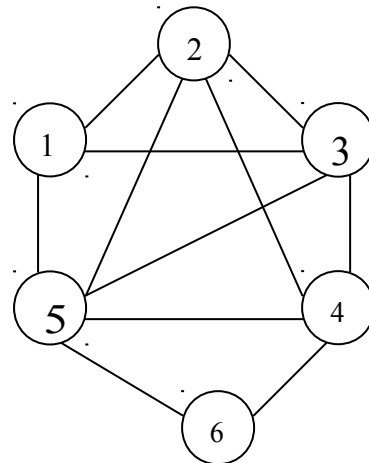


Рисунок А.1.6 – Зображення графа

Остовним „деревом”, або *остовом*, називається „дерево”, що містить всі вершини графа. Отже, якщо граф містить n вершин, то „дерево” з n вершинами і $n - 1$ ребром є остовом. „Дерево” для графа (рис. А.1.6) представлено на рис. А.1.7, остовне „дерево” – на рис. А.1.8. Тут для спрощення вершини графа позначені номерами, гілки ж не позначені. Їх позначатимемо за необхідності.

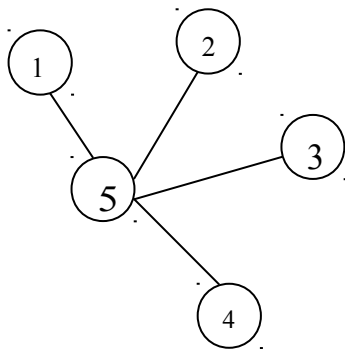


Рисунок А.1.7 – Зображення „дерева” графа для графа (рис. А.1.6)

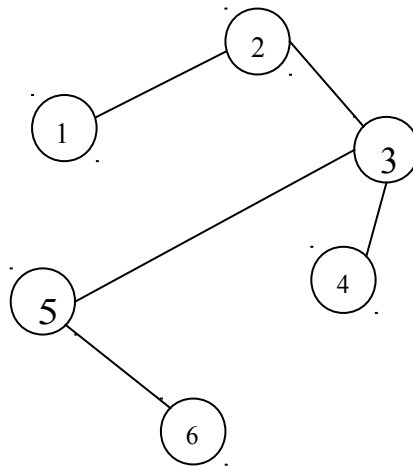


Рисунок А.1.8 – Зображення остовного „дерева” для графа (рис. А.1.6)

Помітно, що для графа можна побудувати декілька остовних „дерев”.

Надалі для спрощення, якщо це не обумовлено особібно, остовне „дерево” називатимемо просто „деревом”.

Якщо для вказівки деяких звичайних обмежень, що накладаються на ребра (дуги) графа, зіставити останнім числа, які називаються вагами ребер, або їх кількісними характеристиками, відповідними, наприклад, відстані, вартості, часу або іншим адитивним величинам, то можна ввести поняття мінімального (або максимального) „дерева”. Вага „дерева” визначається як сума ваг його ребер (звідси і вимога до адитивності характеристик ребер графа).

Найкоротшим (максимальним) „деревом” називається „дерево” з мінімальною (максимальною) вагою. Іноді таке „дерево” називають „деревом” мінімальної (максимальної) вартості.

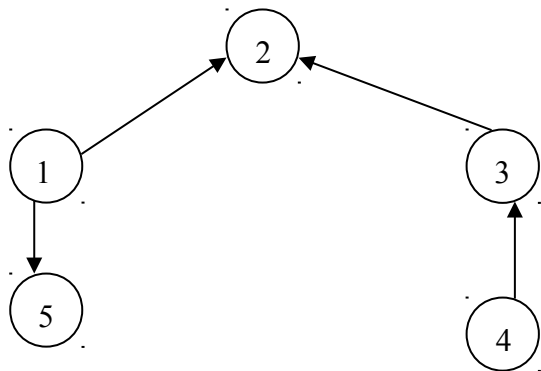


Рисунок А.1.9 – Зображення дерева графа рис. А.1.5

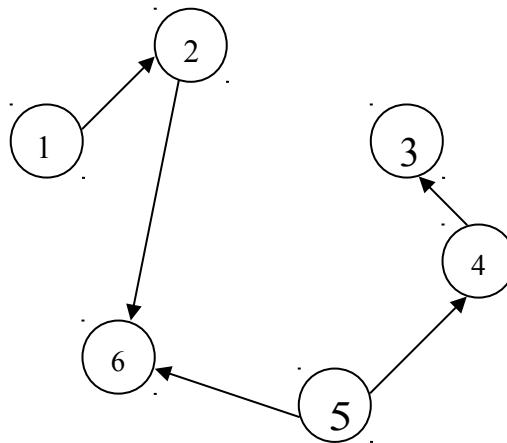


Рисунок А.1.10 – Зображення остова графа рис. А.1.5

„Дерева” і остови визначаються як для орієнтованих, так і для неорієнтованих графів.

На рис. А.1.9 зображено „дерево” графа рис. А.1.5, на рисунку А.1.10 – остов графа рис. А.1.5.

Січенням (розрізом) графа називається сукупність ребер, при видаленні яких граф розпадається на два незв'язні підграфи.

Перетин σ_{st} , що розділяє вершини s і t , є безліч ребер (A, \bar{A}) , де $s \in A, t \in \bar{A}$. Число $C(A, \bar{A})$ називається пропускнуою здатністю (величиною) січення. Пропускна здатність січення визначається як сума пропускнух здатностей ребер, утворюючих січення. На графах січення звичайно позначаються лініями, що пересікають ребра, утворюючих січення.

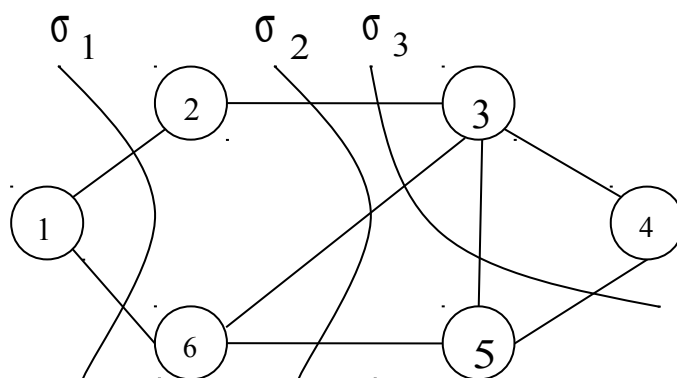


Рисунок А.1.11 – Розподіл графа на підграфи січеннями $\sigma_1, \sigma_2, \sigma_3$

На рис. А.1.11 січення σ_1 включає ребра $(1, 2), (1, 6)$ і розділяє граф на два підграфи, в одному з яких вершина 1, в іншому – вся решта вершин і ребер. Січення σ_2 (ребра $(2, 3), (6, 3), (6, 5)$) розділяє граф на два підграфи, в одному з яких вершини 1, 2, 6 і ребра $(1, 2), (1, 6)$, в іншому – вершини 3, 4, 5 і ребра $(3, 4), (3, 5)$ і $(4, 5)$. Січення σ_3 розділяє граф на два підграфи, в одному з яких вершини 1, 2, 5, 6 і ребра $(1, 2), (1, 6), (6, 5)$, в іншому – вершини 3 і 4 і ребро $(3, 4)$.

Модель мережі зв'язку

Структура мережі зв'язку, її топологія – це сукупність пунктів (вузлів, станцій і т. д.) і з'єднуючих їх ліній (каналів).

Для математичного опису структури мережі зв'язку зручно користуватися мережною моделлю. В якості моделі використовується граф $G = (X, A)$, де $X = \{x_i\}$ – сукупність вершин графа, які ставляться у відповідність пунктам мережі, а $A = \{a_{ij}\}$ – сукупність ребер графа, які виставляються відповідно лініям зв'язку. Оскільки канали мережі можуть бути односторонніми і двосторонніми, ребра графа можуть бути орієнтованими і неорієнтованими. Таким чином, як модель мережі зв'язку можуть бути використані орієнтовані, неорієнтовані, змішані графи, а також мультиграфи.

Мережні моделі широко використовуються на практиці при проектуванні систем електрозв'язку, систем космічного і радіозв'язку, телетрансляційних мереж, транспортних мереж і т.д.

Мережний аналіз відіграє все зростаючу роль, оскільки за допомогою мереж (графів) можна досить просто побудувати модель системи. Розширення області використання мереж пов'язане з тим, що методи мережного аналізу дозволяють [1]:

- 1) побудувати модель складної системи як сукупність простих;
- 2) скласти формальні процедури для визначення якісних і кількісних характеристик системи;
- 3) вказати механізм взаємодії компонентів системи з метою опису останньої в термінах її основних характеристик;
- 4) визначити, які дані необхідні для дослідження системи.

Основна перевага мережного підходу полягає у тому, що він може бути успішно застосований до розв'язання практично будь-якої задачі, коли дослідник володіє необхідними знаннями і здатністю точно побудувати мережну модель.

А.1.3 Матричні представлення графів

Для алгебраїчного задання графів зручно користуватися топологічними матрицями і матрицями характеристик ребер графа (гілок мережі).

А.1.3.1 Топологічні матриці

Матриця суміжності. Матриця суміжності (зв'язності) графа G – квадратна матриця $A = \{a_{ij}\}$ розміру n (n – число вершин графа). Визначається таким чином:

$$a_{ij} = \begin{cases} 1, & \text{якщо в } G \text{ існує ребро } (i, j), \\ 0, & \text{якщо в } G \text{ немає ребра } (i, j). \end{cases}$$

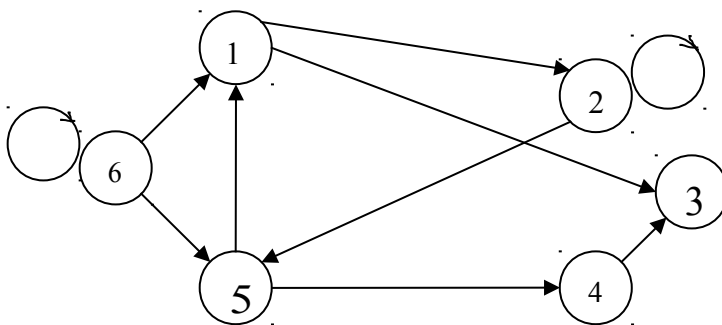


Рисунок А.1.12 – Зображення графа матриці суміжності табл. А.1.1

Елементи головної діагоналі матриці A звичайно приймаються рівними нулю ($a_{ij} = 0$), за винятком випадків, коли в деяких вершинах є петлі.

Матрицю суміжності графа, зображеного на рис. А.1.12, має вигляд табл. А.1.1.

У вершинах 2 і 6 є „петлі”, тому відповідні елементи головної діагоналі $a_{22} = 1$ і $a_{66} = 1$. Надалі розглядатимемо графи, що не містять „петель” у вершинах.

Таблиця А.1.1 – Матриця суміжності графа, зображеного на рис. А.1.12

$$A = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{matrix}$$

Відзначимо, що матриця A для орієнтованого графа несиметрична щодо головної діагоналі, симетричною вона буде для неорієнтованого графа.

Структурна матриця. Для спрощення запису структури і мережі ребрам графа привласнюються позначення, наприклад, a, b, c, \dots . Ці позначення використовуються як елементи структурної матриці.

Структурна матриця графа G – квадратна матриця $B = [b_{ij}]$ розміру n . Визначається таким чином:

$$b_{ij} = \begin{cases} 1, \text{ якщо } i = j; \\ 0, \text{ якщо в } G \text{ нема ребра } (i, j); \\ x, \text{ якщо ребро } (i, j) \text{ в } G \text{ існує; } i < j; \\ x, \text{ якщо ребро } (i, j) \text{ в } G \text{ існує; } i > j. \end{cases}$$

Тут x – позначення ребра на графі.

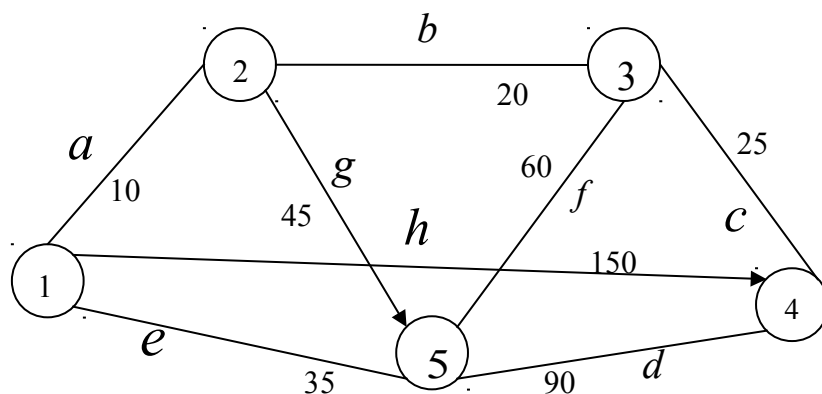


Рисунок А.1.13 – Зображення графа структурної матриці B табл. А.1.2

Для графа, зображеного на рис А.1.13, структурна матриця B представлена

в табл. А.1.2

Таблиця А.1.2 – Структурна матриця графа, зображеного на рис. А.1.13

$$B = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 1 & a & 0 & h & e \\ \bar{a} & 1 & b & 0 & g \\ 0 & \bar{b} & 1 & 0 & f \\ 0 & 0 & \bar{c} & 1 & d \\ \bar{e} & 0 & \bar{f} & \bar{d} & 1 \end{bmatrix} \end{matrix}$$

Крім розглянутих топологічних матриць, можуть бути використані матриці інцидентій «вершини-дуги», «дуги-дуги» [1,4, 5].

А.1.3.2 Матриці кількісних характеристик ребер графа

Для різних кількісних оцінок кожному ребру графа приписується деяка вага - число, що характеризує будь-яку властивість даного ребра, наприклад, довжину, вартість, пропускну здатність, каналну ємність, час передачі інформації, надійність і т.п. Вказані характеристики ребер графа представляються у формі відповідних квадратних матриць розміру n - довжин, вартостей і т.д.

Для побудови, наприклад, матриці довжин $L = [l_{ij}]$ користуються правилом:

$$l_{ij} = \begin{cases} 1, \text{ якщо } i = j; \\ \infty, \text{ якщо в } G \text{ нема ребра } (i, j); \\ \text{довжині ребра } (i, j), \text{ якщо ребро } (i, j) \text{ в } G \text{ існує.} \end{cases}$$

Вважаючи, що числа, зображені біля ребер графа рис А.1.13, є довжини ребер, матрицю довжин L представимо в табл. А.1.3

Таблиця А.1.3 – Матриця довжин L

$$L = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 0 & 10 & \infty & 150 & 35 \\ 10 & 0 & 20 & \infty & 45 \\ \infty & 20 & 0 & 25 & 60 \\ \infty & \infty & 25 & 0 & 90 \\ 35 & \infty & 60 & 90 & 0 \end{bmatrix} \end{matrix}$$

Таблиця А. 1.4 – Матриця каналних ємностей C

$$C = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} \infty & 10 & 0 & 150 & 35 \\ 10 & \infty & 20 & 0 & 45 \\ 0 & 20 & \infty & 25 & 60 \\ 0 & 0 & 25 & \infty & 90 \\ 35 & 0 & 60 & 90 & \infty \end{bmatrix} \end{matrix}$$

Матрицю каналних ємностей ребер $C=[c_{ij}]$ отримують за правилом:

$$C_{ij} = \begin{cases} \infty, & \text{якщо } i = j; \\ 0, & \text{якщо в } G \text{ нема ребра } (i, j); \\ \text{числу каналів ребра } (i, j), & \text{якщо ребро } (i, j) \text{ в } G \text{ існує.} \end{cases}$$

Вважаючи, що числа, вказані біля ребер графа рис. А.1.13 є число каналів ребер, відповідну матрицю каналних ємностей C представимо в табл. А.1.4.

Аналогічно виходять і інші матриці характеристик ребер графа. Якщо G – неорієнтований граф, то всі матриці симетричні щодо головної діагоналі, тобто їх елементи $x_{ij} = x_{ji}$.

Б.2 АЛГОРИТМИ ПОШУКУ ШЛЯХІВ У МЕРЕЖАХ

Основне призначення мережі зв'язку полягає в тому, щоб надавати абонентам сполучні шляхи для передачі повідомлень відповідно до адрес і необхідних якісних показників (якістю обслуговування, швидкістю передачі, надійністю і т. д.); при цьому вибір сполучних шляхів повинен здійснюватися так, щоб забезпечити найефективніше використання устаткування мережі, мінімально можливу довжину шляхів, число транзитних ділянок в шляхах, необхідне число каналів у шляхах і т.д. [6].

При розв'язанні різних задач аналізу і синтезу мереж зв'язку виникає необхідність в пошуку безлічі шляхів, існуючих між заданою парою вузлів мережі (вершин графа). Всі методи пошуку безлічі шляхів в мережі можна розділити на два класи – матричні і мережні. Матричні методи засновані на перетворенні різних матриць – топологічних або матриць характеристик ребер графа, мережні – на привласненні вершинам графа деяких позначень, які називають позначками (або індексами).

Б.2.1 Алгоритми пошуку безлічі шляхів

Б.2.1.1 Матричні алгоритми пошуку безлічі шляхів

Розкладання булевого визначника структурної матриці V . Якщо в мережі є можливість передачі інформації з вузла i у вузол j (по прямих каналах або через проміжні вузли), то говорять, що існує зв'язок від i до j . Для здійснення зв'язку повинен існувати відповідний шлях або шляхи.

Введемо такі позначення: μ_{ij} – шлях з вузла i у вузол j (з вершини i у вершину j графа); m_{ij} – безліч шляхів з i в j .

Шлях μ_{ij} – це впорядкований набір ребер, його складових.

Довжина шляху μ_{ij} – L_{ij} – це сума довжин l_{xy} ребер, його складових:

$$L_{ij} = .$$

Ранг шляху R_{ij} – число ребер, утворюючих шлях.

Для графа, зображеного на рис. Б.2.1, для шляху $\mu_{14} = a_{12}, a_{25}, a_{53}, a_{34}$ ранг $R_{14} = 4$, для шляху $\mu_{14} = a_{14}$ ранг $R_{14} = 1$, для шляху $\mu_{14} = a_{12}, a_{23}, a_{34}$ ранг $R_{14} = 3$.

Структурна матриця графа рис.Б.2.1, представлена в табл. Б.2.1, може розглядатися як булева матриця, оскільки її елементами є булеві змінні. Дійсно, будь-який елемент матриці B – це змінна, її інверсія, константи 0 або 1. Таким чином, шлях як послідовність ребер може бути представлений кон'юнкцією (логічним множенням) ребер, його складових.

Таблиця Б.2.1 – Структурна матриця графа, зображеного на рис. Б.2.1

$$B = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 1 & a & 0 & h & e \\ \bar{a} & 1 & b & 0 & g \\ 0 & \bar{b} & 1 & 0 & f \\ 0 & 0 & \bar{c} & 1 & d \\ \bar{e} & 0 & \bar{f} & \bar{d} & 1 \end{bmatrix} \end{matrix}$$

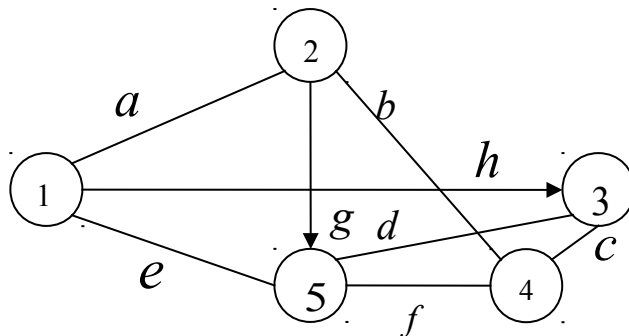


Рисунок Б.2.1 – Зображення графа структурної матриці табл. Б.2.1

Розглянуті шляхи μ_{14} запишемо у формі

$$\mu_{14} = agfc, \mu_{14} = h, \mu_{14} = abc .$$

Безліч шляхів можна записати в диз'юнктивній нормальній формі, застосувавши операцію диз'юнкції (логічного складання) до шляхів, що становлять множину.

Так, в даному прикладі

$$m_{14} = agfc + h + abc .$$

Тут для позначення логічного множення і складання застосовуються символи алгебраїчного множення і складання. Окрім вказаних символів, використовуються також символи: \wedge – для позначення кон'юнкції, \vee – для позначення диз'юнкції.

Безліч шляхів m_{ij} може бути отримане розкладанням визначника матриці B , з якої заздалегідь викреслюється стовбець з номером i і рядок з номером j , тобто [7]:

$$m_{ij} = \det B_{ij} = b_{k1}|B_{k1}| + b_{k2}|B_{k2}| + \dots + b_{kl}|B_{kl}| + \dots + b_{kn}|B_{kn}|. \quad (\text{Б.2.1})$$

Розкладання визначника матриці B_{ij} проведено по рядку з номером k . Розкладання (Б.2.1) можна проводити по будь-якому рядку (стовбцю). Тут B_{ij} – матриця, отримана з початкової структурної матриці після викреслювання i -го стовпця і j -го рядка.

У стовпцях кожної вершини графа записані ребра, що в неї входять; у рядках кожної вершини графа записані ребра, які виходять з вершини.

Викресливши стовпець i і рядок j , ми тим самим вилучимо ті ребра, які не можуть утворювати шляхів множини m_{ij} . $|B_{kl}|$ – визначник, отриманий з $\det B_{ij}$ після викреслювання k -го рядка і l -го стовпця.

Розкладання визначника (Б.2.1) слід проводити за відомими правилами лінійної алгебри, проте операції алгебраїчного множення і складання замінюються на операції логічного множення і складання, відповідно.

При перетворенні матриці і визначників слід використовувати основні правила і закони булевої алгебри :

- 1) $1 + a = 1; \quad 1 \cdot a = a;$
- 2) $0 + a = a; \quad 0 \cdot a = 0;$
- 3) $a + a = 1; \quad a \cdot a = 0;$
- 4) $a + a = a; \quad a \cdot a = a$ – закон повторення;
- 5) $a + ab = a$ – закон поглинання;
- 6) $(a + x) \cdot (a + y) = a + xy$ – розподільний закон.

При розкладанні булевих визначників необхідно також користуватися такими правилами:

- 1) визначник, що має одиницю в кожному рядку і стовпці, тотожно дорівнює одиниці;
- 2) якщо будь-який рядок або стовпець складається з нулів, то визначник тотожно дорівнює нулю;
- 3) якщо перед визначником записаний множник « a », то всі елементи « a » у визначнику можна замінити на «1», а всі елементи « a » у визначнику можна замінити на «0»;
- 4) якщо у визначнику поміняти місцями два рядки або два стовпці, або виробити його транспонування, то значення визначника не зміниться.

Розглянемо отримання безлічі шляхів m_{14} з матриці B (див. табл Б.2.1):

$$m_{14} = \det B_{ij} = \begin{matrix} & 2 & 3 & 4 & 5 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 5 \end{matrix} & \begin{bmatrix} a & 0 & h & e \\ 1 & b & 0 & g \\ \bar{b} & 1 & c & f \\ 0 & \bar{f} & \bar{d} & 1 \end{bmatrix} \end{matrix}$$

Розкладання визначника проведемо по стовпцю з номером 4:

$$m_{14} = h \begin{vmatrix} 1 & b & d \\ \bar{b} & 1 & f \\ 0 & \bar{f} & 1 \end{vmatrix} + c \begin{vmatrix} a & 0 & e \\ 1 & b & g \\ 0 & \bar{f} & 1 \end{vmatrix} + \bar{d} \begin{vmatrix} a & 0 & e \\ 1 & b & g \\ \bar{b} & 1 & f \end{vmatrix} = h + c(ab + e\bar{f} + ag\bar{f}) + \bar{d}(abf + e + b\bar{b}e + ag) =$$

$$= h + cab + ce\bar{f} + cag\bar{f} + \bar{d}abf + \bar{d}e + \bar{d}ag.$$

Упорядкуємо отримані шляхи за рангом, врахувавши послідовність ребер у шляхах. Отримаємо:

$$m_{14} = h + e\bar{d} + abc + e\bar{f}c + ag\bar{d} + ag\bar{f}c + abf\bar{d}.$$

Той самий результат отримаємо, розклавши визначника, наприклад, по першому рядку:

$$m_{14} = a \begin{vmatrix} b & 0 & d \\ 1 & c & f \\ \bar{f} & \bar{d} & 1 \end{vmatrix} + h \begin{vmatrix} 1 & b & g \\ \bar{b} & 1 & f \\ 0 & \bar{f} & 1 \end{vmatrix} + e \begin{vmatrix} 1 & b & 0 \\ \bar{b} & 1 & c \\ 0 & \bar{f} & \bar{d} \end{vmatrix} = a(bc + g\bar{d} + \bar{f}cg + \bar{d}bf) + h + e(\bar{d} + \bar{f}c) =$$

$$= h + e\bar{d} + abc + e\bar{f}c + ag\bar{d} + ag\bar{f}c + abf\bar{d}.$$

Зведення структурної матриці B до степеня. Початкова структурна матриця B є матрицею прямих зв'язків, тобто в ній записані всі шляхи першого рангу, існуючі в графі. При зведенні матриці до квадрату кожен її елемент $b_{ij}^{(2)}$ представлятиме безліч шляхів до другого рангу включно, існуючих між вершинами i і j . Зведення матриці B до степеня $R_{\max} = n - 1$ приводить до отримання всіх шляхів, існуючих між кожною парою (i, j) вершин графа, тобто

$$m_{ij} = b_{ij}^{(n-1)}.$$

Тут n – число вершин графа, R_{\max} – максимально можливий ранг шляхів в графі. $R_{\max} = n - 1$ у разі, коли який-небудь шлях проходить через всі вершини графа.

При зведенні матриці до степеня користуються відомими правилами лінійної алгебри:

$$B^{(2)} = B^{(1)} \cdot B^{(1)},$$

$$B^{(3)} = B^{(2)} \cdot B^{(1)} \text{ і т.д.}$$

Елемент матриці-співмножників $C = A \cdot D$ в лінійній алгебрі отримують відповідно до виразу

$$C_{ij} = \sum_{k=1}^n a_{ik} d_{kj}, \quad (\text{B.2.2})$$

n – загальний розмір матриць A і D (число стовпців матриці A і рядків матриці D , у разі зведення квадратної матриці B до степеня n – число вершин).

Оскільки матриця B – булева, зведення її до степеня здійснюється з ви-

користанням законів і тотожностей булевої алгебри. Крім того, у виразі (Б.2.2) операції алгебраїчного множення і складання замінюються на операції логічного множення і складання, відповідно. При зведенні матриці B до степеня q кожен елемент $b_{ij}^{(q)}$ матриці $B^{(q)}$ обчислюється за такими правилами:

1) елементи i -го рядка матриці $B^{(q-1)}$ логічно множаться на відповідні елементи j -го стовпця матриці $B^{(1)}$;

2) отримані логічні співмножники логічно складаються, утворюючи шукане значення елемента $b_{ij}^{(q)}$. Якщо при зведенні до деякого степеня q виявиться, $B^{(q)} = B^{(q-1)}$ обчислення слід припинити. При цьому $B^{(q)} = B^{(q-1)} = B^{(R_{\max})}$.

Покажемо процес отримання матриці $B^{(2)}$ для графа рис. Б.2.1, матриця B якого представлена в табл. Б.2.1.

Для отримання елемента $b_{12}^{(2)}$ необхідно використовувати елементи першого рядка і другого стовпця матриці $B^{(1)}$:

$$b_{12}^{(2)} = \frac{1 \quad a \quad 0 \quad h \quad e}{a \quad 1 \quad \bar{b} \quad 0 \quad 0} \quad - \text{I рядок } B^{(1)} - \text{II рядок } B^{(1)}.$$

$$a + a + 0 + 0 + 0 = a$$

Аналогічно отримуємо решту елементів матриці $B^{(2)}$:

$$b_{13}^{(2)} = \frac{1 \quad a \quad 0 \quad h \quad e}{0 \quad b \quad 1 \quad \bar{c} \quad \bar{f}}$$

$$ab + h\bar{c} + e\bar{f}$$

$$b_{15}^{(2)} = \frac{1 \quad a \quad 0 \quad h \quad e}{e \quad b \quad f \quad d \quad 1}$$

$$e + ag + hd$$

$$b_{14}^{(2)} = \frac{1 \quad a \quad 0 \quad h \quad e}{h \quad 0 \quad c \quad 1 \quad \bar{d}}$$

$$h + e\bar{d}$$

$$b_{21}^{(2)} = \frac{\bar{a} \quad 1 \quad b \quad 0 \quad g}{1 \quad \bar{a} \quad 0 \quad 0 \quad \bar{e}}$$

$$\bar{a} + g\bar{e}$$

$$b_{23}^{(2)} = \frac{\bar{a} \quad 1 \quad b \quad 0 \quad g}{0 \quad b \quad 1 \quad \bar{c} \quad \bar{f}}$$

$$b + g\bar{f}$$

$$b_{24}^{(2)} = \frac{\bar{a} \quad 1 \quad b \quad 0 \quad g}{h \quad 0 \quad c \quad 1 \quad \bar{d}}$$

$$\bar{a}h + bc + g\bar{d}$$

і т.д.

Отримані результати зведемо в матрицю $B^{(2)}$ (табл Б.2.2)

Таблиця Б.2.2 – Матриця $B^{(2)}$

$$B^{(2)} = \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \left[\begin{array}{ccccc} 1 & a & ab + h\bar{c} + e\bar{f} & h + e\bar{d} & e + ag + hd \\ \bar{a} + g\bar{e} & 1 & b + g\bar{f} & \bar{a}h + bc + g\bar{d} & g + \bar{a}e + bf \\ \bar{b}a + f\bar{e} & \bar{b} & 1 & c + f\bar{d} & f + \bar{b}g + cd \\ \bar{d}e & \bar{c}\bar{b} & \bar{c} + d\bar{f} & 1 & d + \bar{c}f \\ \bar{e} & \bar{e}d + \bar{f}\bar{b} & \bar{f} + \bar{d}\bar{c} & d + \bar{e}h + \bar{f}c & 1 \end{array} \right]$$

З порівняння $B^{(2)}$ і $B^{(1)}$ (табл.Б.2.2 і Б.2.1) випливає, що отримана матриця $B^{(2)}$ не є результуючою, оскільки $B^{(2)} \neq B^{(1)}$ і процес обчислень слід продовжувати, а саме:

1. Обчислити $B^{(3)} = B^{(2)} \cdot B^{(1)}$.

2. Якщо $B^{(3)} = B^{(2)}$, то $B^{(3)} = B^{(R_{\max})}$, результат отриманий.

Якщо $B^{(3)} \neq B^{(2)}$, продовжити обчислення, тобто отримувати $B^{(4)}$

Оскільки $R_{\max} = 4$ для даного графа мережі, матриця буде результуючою у будь-якому випадку, навіть якщо $B^{(4)} \neq B^{(3)}$.

В отриманий результуючій матриці $B^{(R_{\max})}$, як указувалося раніше, кожен елемент $b_{ij}^{(R_{\max})} = m_{ij}$, тобто визначає множену всіх шляхів з вершини i у вершину j .

Так,

$$b_{12}^{(R_{\max})} = m_{12} = a + \overline{hcb} + \overline{efb} + \overline{edcb} + \overline{hffb};$$

$$b_{13}^{(R_{\max})} = m_{13} = ab + \overline{hc} + \overline{ef} + \overline{agf} + \overline{hdf} + \overline{edc};$$

$$b_{14}^{(R_{\max})} = m_{14} = h + \overline{ed} + \overline{abc} + \overline{efc} + \overline{agd} + \overline{agfc} + \overline{abfd} \quad \text{і т.д.}$$

Б.2.1.2 Мережевий алгоритм пошуку безлічі шляхів

Мережні методи визначення множини шляхів між заданими вузлами мережі являються графічним еквівалентом матричних методів

Визначення безлічі шляхів засноване на побудові „дерева” шляхів з фіксованої вершини-витоку (коріння дерева) в усю решту вершин-стоків графа. Мережеві методи визначення множини шляхів між заданими вузлами мережі являються графічним еквівалентом матричних методів.

Для побудови „дерева” шляхів необхідно перш за все відзначити «яруси» дерева. На ярусі $R = 0$ поміщається вершина-виток – «корінь» дерева. На ярусі $R = 1$ розташовуються вершини, суміжні вершині-витоку, тобто вершини, шляхи в які з вершини-витоку мають ранг, рівний одиниці. На ярусі $R = 2$ розташовуються вершини, суміжні вершинам, розташованим на ярусі $R = 1$, і т.д.

При записі вершин на ярус $R = 2$ і подальші яруси необхідно стежити за тим, щоб шляхи рангів, що утворюються, 2, 3 і т.д. були простими (самонепересічними), тобто щоб жодна вершина в шляху не повторювалася більше одного разу.

Максимальне значення ярусу (рангу шляху) $R_{\max} = n-1$ (n – число вершин графа).

Дерево шляхів містить всі шляхи з фіксованої вершини-витоку в усю решту вершин. При цьому на ярусі 1 містяться всі шляхи першого рангу, на ярусі 2 – всі шляхи другого рангу і т.д.

Таким чином, на деякому k -му ярусі міститься інформація про всі шляхи k -го рангу з фіксованої вершини-витоку графа у всю решту вершин.

Для графа рис. Б.2.1 побудуємо „дерево” шляхів з вершини 1. Дана вершина є „корінням дерева” і поміщається на нульовий ярус. Значення рангу шляху тут $R = 0$. На перший ярус $R = 1$ розміщуються вершини 2, 4, 5, що мають безпосередній зв'язок з вершиною 1. Далі на другий ярус від вершини 2 розміщуються вершини, пов'язані з вершиною 2, а саме 3 і 5. Вершина 1 виключається з розгляду, оскільки шлях у вершину 2 пройшов через вершину 1.

Від вершини 4 на другий ярус записуються вершини 3 і 5, від вершини 5-вершини 4 і 3. Аналогічно записуються вершини на решту ярусів дерева.

Побудоване дерево для вершини-витоку 1 представлено на рис. Б.2.2. Як видно, „дерево” має три шляхи першого рангу ($R=1$): a , h , e ; шість шляхів другого рангу ($R = 2$):

ab , ag , hc , hd , $e\overline{f}$, $e\overline{d}$, на третьому ярусі ($R = 3$) записані шляхи третього

рангу: $ab\bar{f}$, abc , $ag\bar{f}$, $ag\bar{a}$, $hc\bar{b}$, hcf , $hd\bar{f}$, $e\bar{f}\bar{b}$, $e\bar{f}c$, $e\bar{a}\bar{c}$. Нарешті, шляхи четвертого рангу ($R = 4$): $ab\bar{f}\bar{a}$, $abcd$, $ag\bar{f}c$, $ag\bar{a}\bar{c}$, $hc\bar{b}g$, $hd\bar{f}b$, $e\bar{a}\bar{c}\bar{b}$.

Звичайно, з дерева можна отримати безліч шляхів з фіксованої вершини в будь-яку вершину графа послідовним переглядом ярусів „дерева”. Так,

$$m_{14} = h + e\bar{a} + abc + ag\bar{a} + e\bar{f}c + ab\bar{f}\bar{a} + ag\bar{f}c;$$

$$m_{15} = e + ag + hd + ab\bar{f} + hcf + abcd + hc\bar{b}g.$$

Б.2.2 Алгоритми пошуку найкоротших шляхів

Розподіл каналів і потоків інформації на мережі зв'язку проводиться з урахуванням характеристики шляху. Для оцінки вагової характеристики шляху використовуються критерії, наприклад, довжина шляху, число транзитних ділянок (ранг) шляху, якість тракту передачі, надійність, передачі інформації і т.д. Найпоширенішими методами розподілу є методи, що використовують «найкоротші» шляхи. Термін «найкоротший шлях» є достатньо умовним, оскільки під найкоротшим розуміють шляхи, найкоротші за довжиною, з мінімальним числом транзитних ділянок, з максимальною пропускною здатністю, шляхи мінімальної вартості, максимальної надійності.

У графі, що представляє мережу, виділена деяка вершина i , називана витокком.

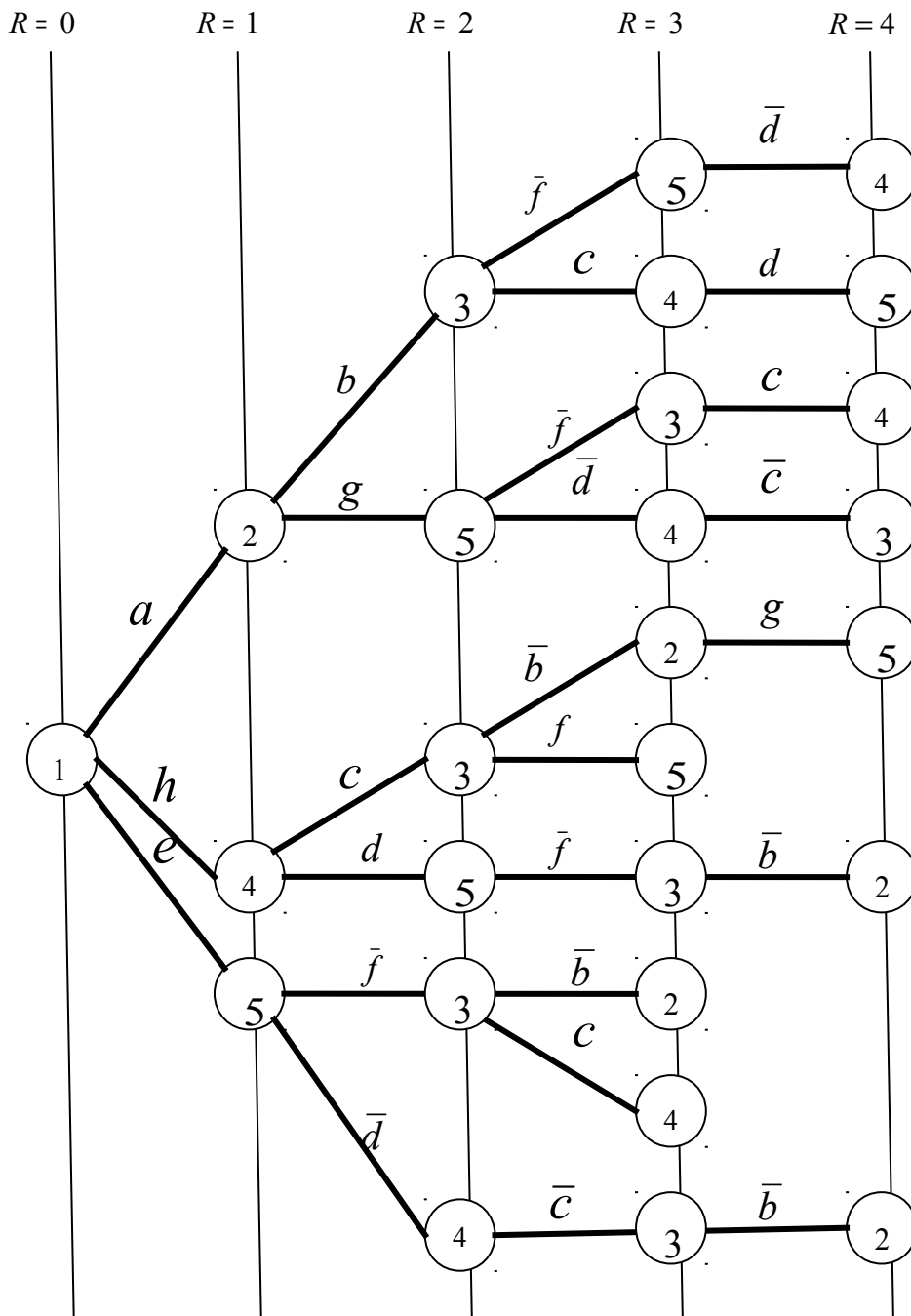


Рисунок Б.2.2 – Зображення „дерева” для вершини-витоку 1

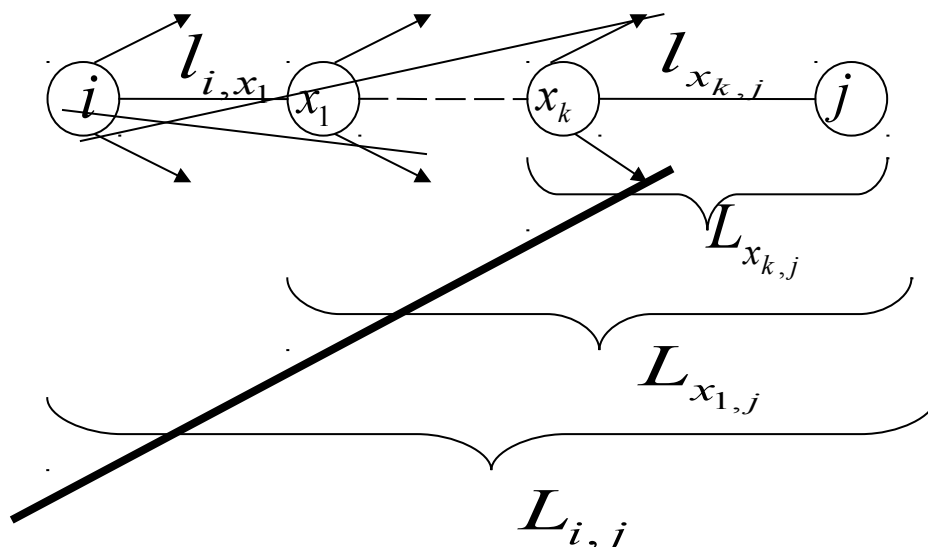


Рисунок Б.2.3 – Зображення шляху від вершини i до вершини j через проміжні вершини x_1, \dots, x_k

Кожному шляху (i, \dots, j) , що з'єднує вершину-витік i з вершиною-стоком j , ставиться у відповідність величина $w(i, \dots, j)$, звана вагою шляху. Таким чином, на безлічі всіх шляхів графа $G = (X, A)$, з'єднуючих витік i зі стоком j , визначена вагова функція $W(i, j)$ вершини j набуваюча значення $w(i, \dots, j)$ на шляху $\mu_{ij} = (i, \dots, j)$. Шлях, на якому $W(i, j)$ досягає екстремального значення, називається екстремальним (найкоротшим, або оптимальним) шляхом з i в j .

Всі методи вибору найкоротших шляхів, розвинені в теорії потоків, оснований на достатньо очевидному твердженні про те, що якщо найкоротший шлях μ_{ij} від довільної вершини i до вершини j проходить через проміжні вершини x_1, \dots, x_k (рис. Б.2.3), то найкоротші шляхи $\mu_{x_1 j}, \dots, \mu_{x_k j}$ від вершин x_1, \dots, x_k до вершини j , відповідно, є частинами найкоротшого шляху від i до j [6].

Якщо довжина шляху $\mu_{x_l j}$ дорівнює $L_{x_l j}$, то довжина шляху з i до j

$$L_{ij} = l_{i,x_l} + L_{x_l j}.$$

Тут l_{i,x_l} – довжина гілки зв'язку вершин i та x_l . Оскільки шлях μ_{ij} є найкоротшим, то

$$L_{ij} = \min_{k = x_1, \dots, x_n} (l_{i,x_k} + L_{x_k j}),$$

де n – число вершин графа.

Легко зрозуміти, що для знаходження найкоротшого шляху від вершини x_0 до вершини x_m необхідно проглянути всі можливі шляхи і вибрати той, який має екстремальну вагу.

Найчастіше використовуються шляхи мінімальної довжини і рангу. Нині існує ряд методів, що дозволяють упорядкувати процедуру визначення довжини, рангу, або інших характеристик шляхів. Умовно ці методи можна розділити на дві групи – матричні і мережні, або індексні.

Б.2.2.1 Матричні алгоритми пошуку найкоротших шляхів

Зведення матриці L до стееія. Матричний метод визначення довжин найкоротших шляхів, що дозволяє знайти довжини найкоротших шляхів між усіма вершинами графа одночасно, ґрунтується на застосуванні операцій над матрицями довжин L [6, 7].

Нехай задана мережа зв'язку у вигляді графа, зображеного на рис.Б.2.4, цифри на ребрах якого відповідають довжинам гілок. Тоді матриця довжин L матиме вид табл. Б.2.3.

По суті матриця L – це матриця відстаней безпосередніх зв'язків, тобто шляхів першого рангу. Позначимо матрицю $L^{(1)}$. Зведемо матрицю $L^{(1)}$ до квадрата:

$$L^{(2)} = L^{(1)} \cdot L^{(1)}$$

$i - j$ -й елемент матриці $L^{(2)}$ визначається за правилом:

$$l_{ij}^{(2)} = \sum_{k=1}^n l_{ik}^{(1)} \cdot l_{kj}^{(1)} = l_{i_1}^{(1)} \cdot l_{1j}^{(1)} + l_{i_2}^{(1)} \cdot l_{2j}^{(1)} + \dots + l_{i_m}^{(1)} \cdot l_{mj}^{(1)}. \quad (\text{Б.2.3})$$

Таблиця Б.2.3 – Матриця довжин L

$$L = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 0 & 50 & \infty & 80 & \infty \\ 50 & 0 & 40 & 90 & 130 \\ \infty & 40 & 0 & \infty & 60 \\ 80 & 90 & \infty & 0 & 25 \\ \infty & 130 & 60 & 25 & 0 \end{bmatrix} \end{matrix}$$

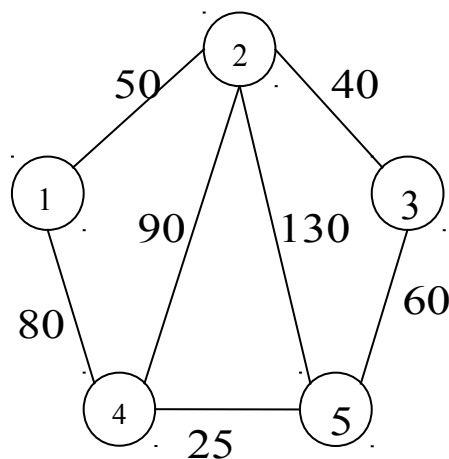


Рисунок Б.2.4 – Зображення мережі зв'язку у вигляді графа

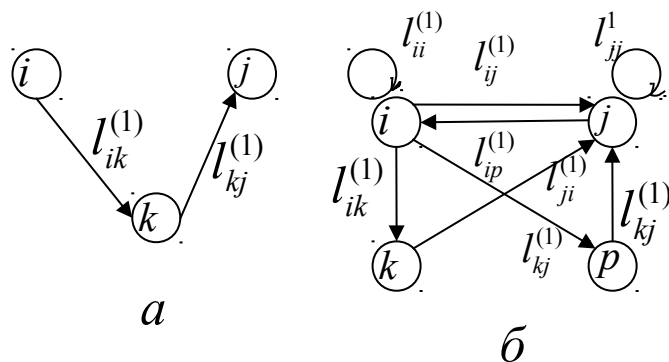


Рисунок Б.2.5 – Зображення а) двотранзитного шляху з вершини i до вершини j через проміжну вершину k ; б) чотирьох двотранзитних шляхів

Інтерпретуючи тепер множення як послідовне, а складання – як паралельне з'єднання гілок (за аналогією із зведенням структурної матриці B до степеня, розділ Б.2.1.1), легко зрозуміти, що співмножник $l_{ik}^{(1)} \cdot l_{kj}^{(1)}$ відповідає двотранзитному шляху (шляхи другого рангу), який проходить з вершини i у вершину j через проміжну вершину k (рис. Б.2.5,а), а сума, наприклад, чотирьох співмножників

$$l_{ii}^{(1)} \cdot l_{ij}^{(1)} + l_{ij}^{(1)} \cdot l_{jj}^{(1)} + l_{ik}^{(1)} \cdot l_{kj}^{(1)} + l_{ip}^{(1)} \cdot l_{pj}^{(1)}$$

– чотирьом двотранзитним шляхам (рис. Б.2.5,б). Відзначимо, що співмножники $l_{ii}^{(1)} \cdot l_{ij}^{(1)}$ і $l_{ij}^{(1)} \cdot l_{jj}^{(1)}$ фактично відповідають одотранзитним шляхам (тобто шляхам першого рангу, включаючим тільки одну гілку).

Оскільки довжина шляху, що складається з декількох гілок, визначається сумою довжин гілок, то необхідно при множенні матриці $L^{(1)}$ операцію множення у виразі (2.3) замінити на операцію складання, тобто замість $l_{ik}^{(1)} \cdot l_{kj}^{(1)}$ матимемо $l_{ik}^{(1)} + l_{kj}^{(1)}$.

За наявності декількох паралельних одно- і двотранзитних шляхів (див. рис. Б.2.5,б) для визначення довжини найкоротшого шляху необхідно операцію складання у виразі (Б.2.3) замінити на операцію вибору мінімуму з довжин всіх одно- і двотранзитних шляхів, тобто замість (Б.2.3) матимемо:

$$l_{ij}^{(2)} = \min_{k=1, \dots, n} (l_{ik}^{(1)} + l_{kj}^{(1)}) = \min (l_{i1}^{(1)} + l_{1j}^{(1)}, (l_{i2}^{(1)} + l_{2j}^{(1)}), \dots, (l_{in}^{(1)} + l_{nj}^{(1)})). \quad (\text{Б.2.4})$$

Таким чином, елемент $l_{ij}^{(2)}$ матриці $L^{(2)}$ дорівнює довжині найкоротшого шляху з i в j серед всіх одно- і двотранзитних шляхів.

При зведенні матриці $L^{(1)}$ до степеня q при використанні операції (Б.2.4)

отримаємо $L^{(q)} = L^{(q-1)} \cdot L^{(1)}$, елемент якої $l_{ij}^{(q)} = \min(l_{ik}^{(q-1)} + l_{kj}^{(1)})$ $k = 1, \dots, n$ визначає довжину найкоротшого шляху між вершинами i та j серед всіх шляхів першого, другого, q -го рангу.

Зведення матриці $L^{(1)}$ до степеня максимального рангу R_{\max} приводить до отримання матриці найкоротших за довжиною шляхів між усіма парами вершин графа, або матриці оптимальних шляхів $L_{\text{opt}} = L^{(R_{\max})}$.

Якщо при зведенні матриці $L^{(1)}$ до деяку степеня q виявиться, що

$$L^{(q)} = L^{(q-1)} \quad (\text{Б.2.5})$$

процес обчислень слід припинити, оскільки при рівності (Б.2.5) завжди виконується рівність $L^{(q)} = L^{(q+1)}$.

Таким чином, $L_{\text{opt}} = L^{(q)} = L^{(q-1)}$

Для даного прикладу (див. табл. Б.2.3) обчислимо $L^{(2)}$.

$$l_{12}^{(2)} = \min(l_{11}^{(1)} + l_{12}^{(1)}, l_{12}^{(1)} + l_{22}^{(1)}, l_{13}^{(1)} + l_{32}^{(1)}, l_{14}^{(1)} + l_{42}^{(1)}, l_{15}^{(1)} + l_{52}^{(1)}) = \min(50, 50, \infty, 170, \infty) = 50;$$

$$l_{13}^{(2)} = \min(l_{11}^{(1)} + l_{13}^{(1)}, l_{12}^{(1)} + l_{23}^{(1)}, l_{13}^{(1)} + l_{33}^{(1)}, l_{14}^{(1)} + l_{43}^{(1)}, l_{15}^{(1)} + l_{53}^{(1)}) = \min(\infty, 90, \infty, \infty, \infty) = 90;$$

$$l_{14}^{(2)} = \min(l_{11}^{(1)} + l_{14}^{(1)}, l_{12}^{(1)} + l_{24}^{(1)}, l_{13}^{(1)} + l_{34}^{(1)}, l_{14}^{(1)} + l_{44}^{(1)}, l_{15}^{(1)} + l_{54}^{(1)}) = \min(80, 140, \infty, 80, \infty) = 80$$

і т.д.

Обчисливши решту елементів матриці $L^{(2)}$, отримаємо

$$L^{(2)} = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 0 & 50 & 90 & 80 & 105 \\ 50 & 0 & 40 & 90 & 100 \\ 90 & 40 & 0 & 85 & 60 \\ 80 & 90 & 85 & 0 & 25 \\ 105 & 100 & 60 & 25 & 0 \end{bmatrix} \end{matrix}$$

Оскільки $L^{(2)} \neq L^{(1)}$ і показник степеня «2» $< R_{\max} = 4$, процес обчислення L_{opt} необхідно продовжити, обчисливши $L^{(3)} = L^{(2)} \cdot L^{(1)}$.

Елемент матриці $L^{(3)}$

$$\begin{aligned} l_{12}^{(3)} &= \min(l_{1k}^{(2)} + l_{k2}^{(1)}) = \min(l_{11}^{(2)} + l_{12}^{(1)}, l_{12}^{(2)} + l_{22}^{(1)}, l_{13}^{(2)} + l_{32}^{(1)}, l_{14}^{(2)} + l_{42}^{(1)}, l_{15}^{(2)} + l_{52}^{(1)}) = \\ &= \min(0 + 50, 50 + 0, 90 + 40, 80 + 90, 105 + 130) = 50. \end{aligned}$$

Обчисливши решту елементів матриці $L^{(3)}$, отримаємо:

$$\begin{matrix} & 1 & 2 & 3 & 4 & 5 \end{matrix}$$

$$L^{(3)} = \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \begin{bmatrix} 0 & 50 & 90 & 80 & 105 \\ 50 & 0 & 40 & 90 & 100 \\ 90 & 40 & 0 & 85 & 60 \\ 80 & 90 & 85 & 0 & 25 \\ 105 & 100 & 60 & 25 & 0 \end{bmatrix}$$

З порівняння $L^{(3)}$ і $L^{(2)}$ випливає, що $L^{(3)} = L^{(2)}$. Таким чином, $L_{\text{опт}} = L^{(3)} = L^{(2)}$. Матрицю $L_{\text{опт}}$ називають також дистанційною і позначають $D = [d_{ij}]$ [6].

Процедуру (Б.2.4) можна застосувати для отримання шляхів мінімальної вартості, часу передачі інформації і т.д. адитивних критеріїв оптимальності шляхів, додавши елементам l_{ij} матриці L відповідне значення вартості, часу передачі інформації по гілці (i,j) і інших характеристик.

Для отримання матриці шляхів мінімального рангу $R_{\text{опт}}$ між всіма вершинами графа для початкового графа необхідно побудувати матрицю безпосередніх зв'язків, або матрицю шляхів першого рангу $R = [r_{ij}]$. Матриця R будується аналогічно матриці L , проте замість довжини гілки зв'язку у відповідній позиції записується «1»:

$$r_{ij} = \begin{cases} 0, & \text{якщо } i = j; \\ 1, & \text{якщо між } i \text{ і } j \text{ існує безпосередній зв'язок}; \\ \infty, & \text{якщо між } i \text{ і } j \text{ зв'язку немає}. \end{cases}$$

Для графа рис. Б.2.4 матриця R має вид табл. Б.2.4.

Таблиця Б.2.4 – Матриця R

$$R = \begin{matrix} & 1 & 2 & 3 & 4 & 5 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 0 & 50 & 90 & 80 & 105 \\ 50 & 0 & 40 & 90 & 100 \\ 90 & 40 & 0 & 85 & 60 \\ 80 & 90 & 85 & 0 & 25 \\ 105 & 100 & 60 & 25 & 0 \end{bmatrix} \end{matrix}$$

Всі правила, розглянуті для отримання матриці $L_{\text{опт}}$ справедливі і для матриці $R_{\text{опт}}$, а саме:

$$R_{\text{опт}} = R^{(R_{\text{max}})}$$

і, крім того, якщо для деякого степеня q справедливе $R^{(q)} = R^{(q-1)}$, то з цього виходить, що $R_{\text{опт}} = R^{(q)} = R^{(q-1)}$.

Значимо також, що при зведенні матриці $R^{(1)}$ до будь-якого степеня q $R^{(q)} = R^{(q-1)}$ операцію (Б.2.4) слід застосовувати лише до тих елементів матриці $R^{(q-1)}$, які ще не визначені, тобто $r_{ij}^{(q-1)} = \infty$. Як тільки при зведенні до

чергового степеня q матриці $R^{(l)}$ всі елементи $r_{ij}^q \neq \infty$ ($i, j = \overline{1, n}$, граф G – зв'язний), процес обчислення $R_{\text{опт}}$ слід припинити.

Отримана матриця $R^{(q)} = R_{\text{опт}}$.

Слід зазначити, що розглянутий метод дозволяє визначити довжини (ранги, вартості і т. д.) найкоротших шляхів між усіма вершинами графа, але не вказує ті гілки, які входять в найкоротші шляхи.

Алгоритм Флойда. Алгоритм Флойда [1, 6, 8] дає можливість отримати як матрицю довжин найкоротших відстаней $L_{\text{опт}} = D$, так і маршрутну матрицю, що визначає вершини (гілки) графа, утворюючі шляхи.

Алгоритм Флойда працює таким чином. Спочатку за довжину найкоротшого шляху між довільними вершинами i і j приймається довжина гілки l_{ij} , що з'єднує ці вершини. Потім послідовно перевіряються будь-які проміжні вершини, розміщені між i і j . Якщо довжина шляху, що проходить через деяку проміжну вершину k , менше попереднього (поточного) значення, яке було позначене, наприклад, d_{ij} , то змінній d_{ij} привласнюється нове значення. Значимо, що початкове значення $d_{ij} = l_{ij}$. Дана процедура повторюється для всіх пар вершин, поки не будуть набуті всі значення матриці $L_{\text{опт}} = D$.

Для будь-яких трьох різних вершин i, k і j сформульовані умови можуть бути записані у вигляді нерівності

$$d_{ik} + d_{kj} \geq d_{ij} \quad (i \neq k \neq j) \quad (\text{Б.2.6})$$

оскільки інакше найкоротший шлях з вершини i у вершину j повинен містити вершину k і тоді величина d_{ij} не дорівнювала би довжині найкоротшого шляху. Рис. Б.2.6 ілюструє виконання процедури порівняння (Б.2.6) для пари вершин i і j .

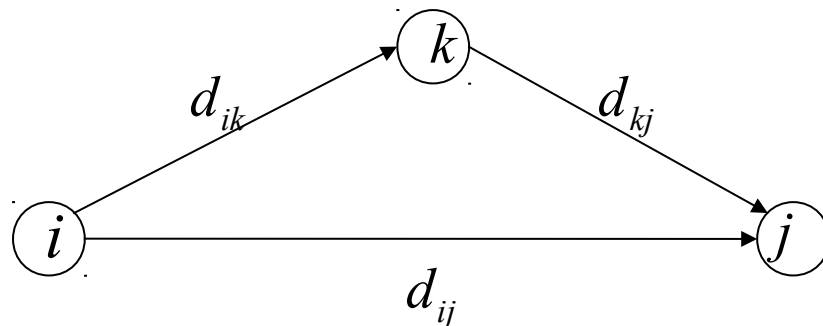


Рисунок Б.2.6 – Зображення виконання процедури порівняння (Б.2.6) для пари вершин i і j

Позначимо символом d_{ij}^q оцінку довжини найкоротшого шляху з i у j , отриману на q -й ітерації.

Алгоритм Флойда дозволяє розв'язувати задачу отримання найкоротших шляхів між усіма вершинами графа для графа з n вершин за n ітерацій.

Спосіб отримання на q -й ітерації величини d_{ij}^q може бути описаний за допомогою наступної тримісної операції (операції трикутника):

$$d_{ij}^q = \min(d_{ij}^{q-1}; d_{ik}^{q-1} + d_{kj}^{q-1}), i \neq k \neq j. \quad (\text{Б.2.7})$$

Якщо операцію (Б.2.7) виконувати з даною парою вершин i і j та усіма вершинами k ($i \neq k \neq j$) у порядку зростання номерів k , то значення d_{ij}^q , отримане на останній ітерації, буде дорівнювати довжині найкоротшого шляху з вершини i у вершину j .

На кожній ітерації алгоритму будується дві матриці.

Перша з них, так звана матриця довжин найкоротших шляхів, містить поточні оцінки довжин найкоротших шляхів, тобто на q -й ітерації дана матриця визначається як $D^q = [d_{ij}^q]$. Алгоритм починає роботу при $D^0 = [d_{ij}^0] = [l_{ij}]$.

Потім, виконуючи тримісну операцію (Б.2.7) з усіма елементами матриці D^0 , отримуємо D^1 і т. д., до тих пір, поки для кожної пари вершин не буде виконаний критерій оптимальності (Б.2.6).

Друга матриця, звана матрицею маршрутів (позначимо її Γ), служить для знаходження проміжних вершин (якщо такі є) найкоротших шляхів. На q -й ітерації вона визначається як $\Gamma^q = [\gamma_{ij}^q]$, де γ_{ij}^q – перша проміжна вершина найкоротшого шляху з i в j , вибрана з вершин множини $\{1, 2, \dots, k\}$ ($i \neq k \neq j$). Алгоритм починає роботу при $\Gamma^0 = [\gamma_{ij}^0]$, де $\gamma_{ij}^0 = j$. На q -й ітерації вершина γ_{ij}^q може бути отримана із співвідношення

$$\gamma_{ij}^q = \begin{cases} k, \text{ якщо } d_{ij}^{q-1} > d_{ik}^{q-1} + d_{kj}^{q-1}, \\ \gamma_{ij}^{q-1} \cdot \text{ в протилежному випадку.} \end{cases}$$

Після побудови матриць D^0 і Γ^0 потрібно для кожного $k = 1, 2, \dots, n$, використовуючи для обчислень елементи матриці D^{q-1} , отриманої на $(q-1)$ -й ітерації, виконати наступну процедуру.

Крок 1. Викреслити елементи k -го рядка і k -го стовпця. Назвемо цю безліч елементів базовим рядком і базовим стовпцем.

Крок 2. Для кожного елемента d_{ij}^{q-1} ($i \neq k$) матриці D^{q-1} , починаючи з першого, розташованого в лівому верхньому кутку матриці, перевірити виконання нерівності

$$d_{ik}^{q-1} + d_{kj}^{q-1} \geq d_{ij}^{q-1} \quad (\text{Б.2.8})$$

Якщо нерівність (Б.2.8) виконується, то вибрати нові значення i і j і знову виконати крок 2.

Якщо нерівність (Б.2.8) не виконується, то елементу d_{ij}^q матриці D^q привласнити значення $d_{ij}^q = d_{ik}^{q-1} + d_{kj}^{q-1}$, а елементу γ_{ij}^q матриці Γ^q привласнити значення $\gamma_{ij}^q = k$.

Після перевірки нерівності (Б.2.8) для всіх елементів (i, j) знов перейти до виконання кроку 1 при $q = q + 1$ до $q = n$.

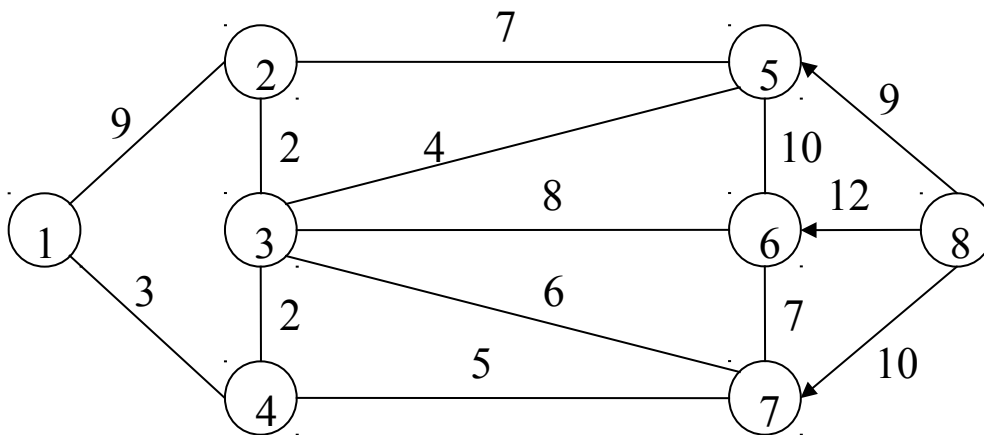


Рисунок Б.2.7 – Зображення графа для ілюстрації роботи алгоритму Флойда

Відповідно до (Б.2.7), при отриманні D^q з D^{q-1} тримісну операцію, що дозволяє встановити факт приналежності базового вузла k найкоротшого шляху, слід виконувати тільки з елементами матриці D^{q-1} тими, що не належать ні базовому рядку, ні базовому стовпцю.

Для ілюстрації роботи алгоритму Флойда визначимо найкоротші шляхи між усіма парами вершин графа, зображеного на рис. Б.2.7.

Початкові значення елементів матриці D^0 представлені в табл. Б.2.5, матриці Γ^0 – в табл. Б.2.6.

Таблиця Б.2.5 – Початкові значення елементів матриці D^0

	1	2	3	4	5	6	7	8
1	0	9	∞	3	∞	∞	∞	∞
2	9	0	2	∞	7	∞	∞	∞
3	∞	2	0	2	4	8	6	∞
4	3	∞	2	0	∞	∞	5	∞
5	∞	7	4	∞	0	10	∞	∞
6	∞	∞	8	∞	10	0	7	∞
7	∞	∞	6	5	∞	7	0	∞
8	∞	∞	∞	∞	9	12	10	0

Таблиця Б.2.6 – Початкові значення матриці Γ^0

$$\Gamma^0 = \begin{matrix} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix} & \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} \end{matrix}$$

Ітерація 1. Вершина $k = 1$ визначається як базова, отже, в матриці D^0 викреслюємо перший рядок і перший стовпець. Оскільки в базовому рядку в стовпцях 3, 5, 6, 7, 8 містяться елементи, які дорівнюють ∞ , то елементи цих стовпців можна не досліджувати відповідно до виразу (Б.2.7).

	1	2	3	4	5	6	7	8
1		9		3				
2	9	0		∞				
3								
4	3	∞		0				
5								
6								
7								
8								

← Базовий

рядок

↑ Базовий стовпець

Рисунок Б.2.8 – Зображення елементів матриці D^0

У базовому першому стовпці в рядках 3, 5, 6, 7, 8 містяться елементи, рівні ∞ , тому, відповідно до (Б.2.7), елементи цих рядків можна не досліджувати. Для того, щоб визначити, чи приведе використання вершини 1 до найбільш коротких шляхів, ніж ті, які записані в D^0 , слід досліджувати за допомогою тримісної операції (Б.2.7) лише елементи матриці D^0 , зображені на рис. Б.2.8, а саме елементи $d_{22}^0, d_{24}^0, d_{42}^0, d_{44}^0$.

Але оскільки елементи головної діагоналі ($i = j$) можна не розглядати, застосування операції (Б.2.7) дає наступні результати:

$$d_{24}^1 = \min(d_{24}^0; d_{21}^0 + d_{14}^0) = \min(\infty, 9 + 3) = 12;$$

$$d_{42}^1 = \min(d_{42}^0; d_{41}^0 + d_{12}^0) = \min(\infty, 3 + 9) = 12.$$

Оскільки

$$d_{24}^1 = 12 < d_{24}^0 = \infty \quad \text{і} \quad d_{42}^1 = 12 < d_{42}^0 = \infty,$$

то вказані елементи розміщуються в матриці D^1 , вся решта елементів

матриці D^0 переноситься в D^1 без зміни. Елементи матриці Γ^1 при цьому: $\gamma_{24}^1 = 1$ і $\gamma_{42}^1 = 1$ ($q = 1$), решта елементів Γ^1 дорівнює відповідним елементам матриці Γ^0 .
Матриці D^1 і Γ^1 представлені в таблицях Б.2.7 і Б.2.8, відповідно.

Таблиця Б.2.7 – Матриця D^1

$$D^1 = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix} & \begin{bmatrix} 0 & 9 & \infty & 3 & \infty & \infty & \infty & \infty \\ 9 & 0 & 2 & 12 & 7 & \infty & \infty & \infty \\ \infty & 2 & 0 & 2 & 4 & 8 & 6 & \infty \\ 3 & 12 & 2 & 0 & \infty & \infty & 5 & \infty \\ \infty & 7 & 4 & \infty & 0 & 10 & \infty & \infty \\ \infty & \infty & 8 & \infty & 10 & 0 & 7 & \infty \\ \infty & \infty & 6 & 5 & \infty & 7 & 0 & \infty \\ \infty & \infty & \infty & \infty & 9 & 12 & 10 & 0 \end{bmatrix} \end{matrix}$$

Таблиця Б.2.8 – Матриця Γ^1

$$\Gamma^1 = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix} & \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} \end{matrix}$$

Ітерація 2. Визначаємо вершину 2 як базову і викреслюємо в матриці D^1 другий рядок і другий стовпець. Оскільки в другому рядку в стовпцях 6, 7 і 8 містяться ∞ , відповідні стовпці не розглядаються. Оскільки в другому стовпці в рядках 6, 7 і 8 містяться ∞ , відповідні рядки не розглядаються. Крім того, не розглядаються елементи головної діагоналі. В результаті необхідно досліджувати елементи $d_{13}^1, d_{14}^1, d_{15}^1, d_{31}^1, d_{34}^1, d_{35}^1, d_{41}^1, d_{43}^1, d_{45}^1, d_{51}^1, d_{53}^1, d_{54}^1$.

Нові оцінки отримують лише такі елементи:

$$\begin{aligned} d_{13}^2 &= \min(d_{13}^1; d_{12}^1 + d_{23}^1) = \min(\infty; 9 + 2) = 11; \\ d_{15}^2 &= \min(d_{15}^1; d_{12}^1 + d_{25}^1) = \min(\infty; 9 + 7) = 16; \\ d_{31}^2 &= \min(d_{31}^1; d_{32}^1 + d_{21}^1) = \min(\infty; 2 + 9) = 11; \\ d_{45}^2 &= \min(d_{45}^1; d_{42}^1 + d_{25}^1) = \min(\infty; 12 + 7) = 19; \\ d_{51}^2 &= \min(d_{51}^1; d_{52}^1 + d_{21}^1) = \min(\infty; 7 + 9) = 16; \\ d_{54}^2 &= \min(d_{54}^1; d_{52}^1 + d_{24}^1) = \min(\infty; 7 + 12) = 19. \end{aligned}$$

Отже $\gamma_{13}^2 = \gamma_{15}^2 = \gamma_{31}^2 = \gamma_{45}^2 = \gamma_{51}^2 = \gamma_{54}^2 = 2$

Матриці D^2 і Γ^2 наведені в табл. Б.2.9 і Б.2.10.

Таблиця Б.2.9 – Матриця D^2

$$D^2 = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix} & \begin{bmatrix} 0 & 9 & 11 & 3 & 16 & \infty & \infty & \infty \\ 9 & 0 & 2 & 12 & 7 & \infty & \infty & \infty \\ 11 & 2 & 0 & 2 & 4 & 8 & 6 & \infty \\ 3 & 12 & 2 & 0 & 19 & \infty & 5 & \infty \\ 16 & 7 & 4 & 19 & 0 & 10 & \infty & \infty \\ \infty & \infty & 8 & \infty & 10 & 0 & 7 & \infty \\ \infty & \infty & 6 & 5 & \infty & 7 & 0 & \infty \\ \infty & \infty & \infty & \infty & 9 & 12 & 10 & 0 \end{bmatrix} \end{matrix}$$

Таблиця Б.2.10 – Матриця Γ^2

$$\Gamma^2 = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix} & \begin{bmatrix} 1 & 2 & 2 & 4 & 2 & 6 & 7 & 8 \\ 1 & 2 & 3 & 1 & 5 & 6 & 7 & 8 \\ 2 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 1 & 3 & 4 & 2 & 6 & 7 & 8 \\ 2 & 2 & 3 & 2 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{bmatrix} \end{matrix}$$

Виконуючи аналогічні обчислення на ітераціях 3, 4, 5, 6, 7, 8, отримаємо результат – матриці D і Γ . Розв'язок, отриманий на ітерації 5, не поліпшується на наступних ітераціях. Отже, оптимальним розв'язком є: $D = D^5$, $\Gamma = \Gamma^5$.

Таблиця Б.2.11 – Матриця D

$$D = D^5 = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix} & \begin{bmatrix} 0 & 7 & 5 & 9 & 16 & 13 & 8 & \infty \\ 7 & 0 & 2 & 6 & 7 & 10 & 8 & \infty \\ 5 & 2 & 0 & 4 & 4 & 8 & 6 & \infty \\ 3 & 4 & 2 & 6 & 19 & 10 & 5 & \infty \\ 9 & 6 & 4 & 0 & 0 & 10 & 10 & \infty \\ 13 & 10 & 8 & 10 & 10 & 0 & 7 & \infty \\ 8 & 8 & 6 & 10 & \infty & 7 & 0 & \infty \\ 18 & 15 & 13 & 9 & 9 & 12 & 10 & 0 \end{bmatrix} \end{matrix}$$

Таблиця Б. 2.12 – Матриця Γ

$$\begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{matrix}$$

$$\Gamma = \Gamma^5 = \begin{matrix} 1 & \left[\begin{array}{cccccccc} 1 & 4 & 4 & 4 & 4 & 4 & 4 & 8 \\ 2 & 4 & 2 & 3 & 3 & 3 & 3 & 8 \\ 3 & 4 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 1 & 3 & 3 & 4 & 3 & 3 & 7 & 8 \\ 5 & 4 & 3 & 3 & 3 & 5 & 6 & 3 & 8 \\ 6 & 4 & 3 & 3 & 3 & 5 & 6 & 7 & 8 \\ 7 & 4 & 3 & 3 & 4 & 3 & 6 & 7 & 8 \\ 8 & 5 & 5 & 5 & 5 & 5 & 6 & 7 & 8 \end{array} \right] \end{matrix}$$

Матриці D і Γ наведені в таблицях Б.2.11 і Б.2.12

Нехай необхідно визначити найкоротший шлях d_{15} . Довжина шляху $d_{15} = d_{15}^5 = 9$. Для знаходження відповідної послідовності вершин, що становлять шлях, звернемося до матриці Γ . Оскільки $\gamma_{15} = \gamma_{15}^5 = 4$ то вершина 4 є першою проміжною в найкоротшому шляху з вершини 1 до вершини 5. Потім знаходимо наступну вершину на шляху до вершини 5, звертаючись до елементу $\gamma_{45} = 3$, далі $\gamma_{35} = 5$, отже, послідовність вершин шляху μ_{15} визначена: $\mu_{15} = \{1, 4, 3, 5\}$.

Аналогічно визначається решта найкоротших шляхів. Наприклад, довжина шляху $\mu_{82} = d_{82} = 15$, шлях складається з вершин: $\mu_{82} = \{8, 5, 3, 2\}$.

Алгоритм, заснований на застосуванні операцій булевої алгебри.

Алгоритм пошуку найкоротших за рангом (транзитності) шляхів заснований на застосуванні операцій булевої алгебри (логічних операцій) до матриці зв'язності графа [9]. Пошук найкоротших шляхів в алгоритмі здійснюється за матрицею $A' = [a_{ij}]$, отримуваної шляхом перетворення початкової матриці зв'язності $A = [a_{ij}]$. У матриці A' ненульовими являються тільки елементи, відповідні ребрам графа, що входить хоча б в один найкоротший за транзитністю (рангу) шлях від фіксованої вершини-витоку в усю решту вершин-стоків графа. Перетворення матриці A в матрицю A' здійснюється виконанням операцій логічного складання, логічного множення, заперечення рівнозначності, інверсії рядків матриці A .

Перетворення матриці A в матрицю A' завжди можливо, оскільки початкова мережа представляється кінцевим графом. Число етапів, необхідних для перетворення матриці, визначається максимальним значенням рангу в усіх найкоротших шляхах від вершини-витоку до решти вершин мережі. Максимально можливе число етапів, необхідних для перетворення матриці A , матиме місце у разі представлення мережі графом-деревом з двома кінцевими вершинами і складе $(n - 1)$ для графа з n вершин.

«Етапом» назвемо процес вибору з вершин графа вершин, суміжних вже вибраним, тобто що мають з вибраними безпосередній зв'язок. Так, перший етап полягає у виборі вершин, суміжних вершині-витоку, другий – у виборі всіх вершин, суміжних вибраним на першому етапі, і так далі, до тих пір, поки всі вершини графа не будуть вибрані.

Початковими для роботи алгоритму є: матриця A і три допоміжні двійкові слова довжиною n – накопичення H , гасіння Γ і ознаки закінчення формування матриці A' – ПФ.

Алгоритм полягає у виконанні кроків:

Крок 1. Запис початкового стану Γ – нуль в розряді з номером вершини-витоку, одиниці – у всій решті розрядів.

Крок 2. Запис рядка вершини-витоку i з матриці A в матрицю A' , в H і ПФ.

Крок 3. Логічне множення інверсії H і Γ , результат залишається в Γ : $\Gamma = \overline{H} \wedge \Gamma$. Запис нулів в усі розряди H .

Крок 4. Визначення номера чергової вершини k відповідно до вмісту одиниці в k -му розряді ПФ.

Крок 5. Логічне множення k -го рядка матриці A і Γ , результат розміщується в k -й рядок матриці A' : $A'_k = A_k \wedge \Gamma$.

Крок 6. Логічне складання k -го рядка матриці A' і H , результат розміщується в H :

$$H = A'_k \vee H.$$

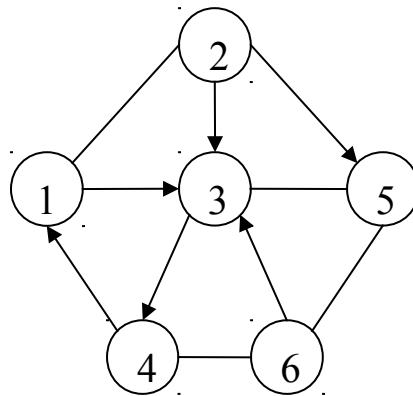


Рисунок б.2.9 – Зображення мережі матриці табл. Б.2.13

Крок 7. Перевірка проглядання всіх рядків матриці A відповідно до змісту ПФ. Перехід до кроку 8 при закінченні перегляду, інакше – повернення до кроку 4.

Крок 8. Запис вмісту H в ПФ: $H \rightarrow \text{ПФ}$.

Крок 9. Перевірка закінчення формування матриці A' . Якщо ПФ = 0, кінець алгоритму, інакше – повернення до кроку 3.

Для мережі, зображеної на малюнку Б.2.9, матриця $A = [a_{ij}]$ представлена в табл. Б.2.13.

Таблиця Б.2.13 – Матриця $A = [a_{ij}]$

$$A = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \end{matrix}$$

У матриці A елементи головної діагоналі $a_{ii} = 0$.

Після першого етапу для вершини-витоку 1 в результаті перетворення рядків 2 і 3, відповідним вершинам, суміжним вершині-витоку 1, матриця A' матиме вид табл. 2.14. Результуюча матриця A' представлена в табл. Б.2.15. Для пошуку за матрицею A' найкоротших шляхів необхідно виконати такі дії.

Нехай визначається шлях в деяку вершину j - μ_{ij} .

Таблиця Б.2.14 – Матриця A'

$$A' = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \end{matrix}$$

Таблиця Б.2.15 – Результуюча матриця A'

$$A' = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{matrix}$$

У стовпці j матриці A' знаходиться чергова одиниця, фіксується номер рядка k , який визначить вершину k , що входить в шлях, потім проглядається стовпець з номером k , знаходиться в ньому чергова одиниця, фіксується номер рядка p , тобто визначається номер чергової вершини p , що входить в шлях μ_{ij} , і так далі, до тих пір, поки черговий номер вибраної вершини не виявиться рівним початковому i . Отже, утворювана послідовність вершин представляється в зворотному порядку. Представити послідовність в прямому порядку не дуже важко.

Так, за матрицею A' отримаємо такі шляхи мінімального рангу з вершини 1 в усю решту вершин графа: 1 – 2; 1 – 3; 1 – 3 – 4; 1 – 2 – 5, 1 – 3 – 5; 1 – 3 –

4 – 6, 1 – 2 – 5 – 6, 1 – 3 – 5 – 6.

Б.2.2.2 Мережні алгоритми пошуку найкоротших шляхів

Задача пошуку найкоротшого шляху (шляхи мінімальної вартості) з деякого витoku s в деякий стік t може бути сформульована як задача лінійного програмування (ЛП).

Вважатимемо, як і раніше, що кожному ребру (i, j) графа поставлено у відповідність деяке число c_{ij} зване узагальненою вартістю ребра (в якості цієї вартості може використовуватися і значення l_{ij} – довжини ребра, t_{ij} – часу передачі інформації по ребру і т. д.). Фіктивним, або «некоштовним» ребрам приписується вартість $c_{ij} = 0$, а кожній парі вершин (i, j) , що не мають прямого зв'язку, – вартість $c_{ij} = \infty$.

Задача знаходження найкоротшого шляху з витoku s в стік t є задача знаходження в заданому графі такого шляху, для якого вартість проходження одиниці потоку мінімальна.

Математично ця задача може бути записана як наступна задача ЛП:
мінімізувати

$$\sum_i \sum_j c_{ij} f_{ij} \quad (\text{Б.2.9})$$

За умови, що

$$\sum_j f_{sj} - \sum_j f_{js} = 1; \quad (\text{Б.2.10})$$

$$\sum_j f_{ij} - \sum_j f_{ji} = 0; i \neq s; i \neq t; \quad (\text{Б.2.11})$$

$$\sum_j f_{tj} - \sum_j f_{jt} = -1; \quad (\text{Б.2.12})$$
$$f_{ij} \geq 0.$$

Де f_{ij} – величина потоку, що протікає по ребру (i, j) .

Згідно з рівністю (2.10), одиниця потоку витікає з витoku (тут j – суміжні з витocom s вершини).

Рівність (Б.2.11) гарантує збереження даної одиниці потоку при протіканні по графу $((i, j)$ – будь-яке ребро графа).

Згідно з рівністю (Б.2.12), одиниця потоку втікає в стік t (тут j – суміжні стоку t вершини).

Як найкоротший шлях може бути взята послідовність суміжних ребер (i, j) , для яких $f_{ij} = 1$.

Для розв'язання поставленої задачі був розроблений спеціальний метод, відомий під назвою алгоритму Дейкстри [1, 8, 11].

Алгоритм заснований на приписуванні вершинам або тимчасових, або постійних позначок. Первинно кожній вершині графа, виключаючи витік, приписується тимчасова позначка, відповідна довжині найкоротшого шляху, що веде з витoku в дану вершину. Витoku приписується постійна позначка, значення якої дорівнює нулю.

Кожній вершині, в яку не можна потрапити безпосередньо з витoku, при-

писується тимчасова позначка ∞ , а всій решті вершин – тимчасові позначки c_{sj} , $j \neq s$. Якщо визначено, що вершина належить найкоротшому шляху, її позначка стає постійною. Алгоритм заснований на простому твердженні: частина найкоротшого шляху є найкоротший шлях.

Алгоритм починає працювати при $j = s$. Потім величина j послідовно збільшується на одиницю і при $j = t$ алгоритм завершує роботу.

Для заданої вершини j через L_j позначатимемо оцінку довжини найкоротшого шляху з витoku s у вершину j . Якщо ця оцінка не може бути поліпшена, то відповідне значення назвемо постійною позначкою і позначатимемо L_j'' , інакше вершина має тимчасову позначку.

Спочатку постійна позначка приписується тільки витoku. Кожна інша позначка є тимчасовою і її величина рівна довжині ребра, що веде з витoku у відповідну вершину. Для визначення найближчої до витoku вершини вибирається тимчасова позначка з мінімальним значенням і оголошується постійною. Потім до тих пір, поки стоку не буде приписана постійна позначка, необхідно виконувати дві процедури.

1. Розглянути вершини, що залишилися, з тимчасовою позначкою. Порівняти величину кожної тимчасової позначки з сумою величини останньої з постійних позначок і довжини ребра, що веде з відповідної постійно позначеної вершини в дану вершину. Мінімальна з двох порівнюваних величин визначається як нова тимчасова позначка даної вершини. Відзначимо, що якщо величина старої тимчасової позначки менше за другу з порівнюваних величин, то позначка залишається колишньою.

2. Серед тимчасових позначок вибрати ту, значення якої мінімальне, і оголосити її постійною позначкою. Якщо при цьому постійна позначка приписується вершині t , то алгоритм завершує роботу. Інакше перейти на крок 1.

Алгоритм може бути виконаний за допомогою таблиці розв'язання, в якій стовпці відповідають вершинам графа, рядки - крокам обчислювального процесу, а її елементи – постійним і тимчасовим позначкам.

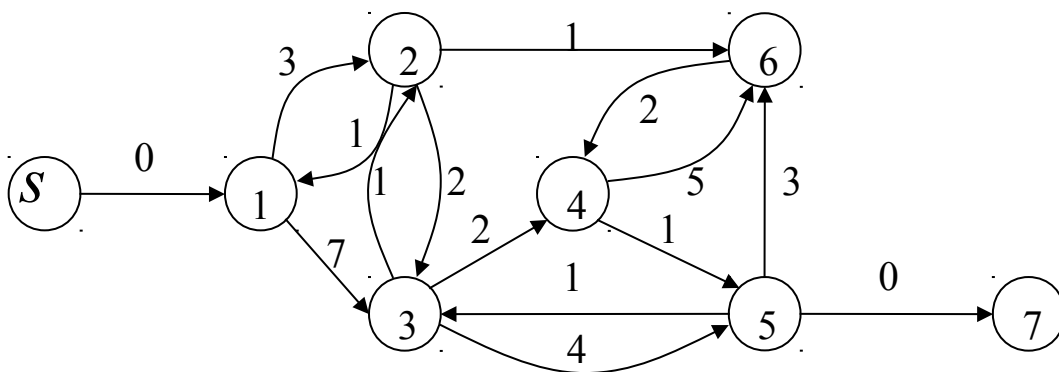


Рисунок Б.2.10 – Зображення графа

Розглянемо граф, зображений на рис. Б.2.10. Вершина s тут є виток, t – стоком, а числа c_{ij} , приписані ребрам, відповідають їх довжині або вартості,

або часу і т.д.

Робота алгоритму починається з того, що виток s приписується постійна позначка 0^n , а вершинам $1, 2, \dots, t$ – тимчасові позначки $L_j = c_{sj}$. Дані позначки записуються в табл. Б.2.16 як нульовий крок.

Таблиця Б.2.16 – Значення позначок

Крок	Вершина							
	s	1	2	3	4	5	6	t
0	0^n	0^n	∞	∞	∞	∞	∞	∞
1	0^n	0^n	3	7	∞	∞	∞	∞
2	0^n	0^n	3^n	5	∞	∞	4	∞
3	0^n	0^n	3^n	5	6	∞	4^n	∞
4	0^n	0^n	3^n	5^n	6	9	4^n	∞
5	0^n	0^n	3^n	5^n	6^n	7	4^n	∞
6	0^n	0^n	3^n	5^n	6^n	7^n	4^n	7
7	0^n	0^n	3^n	5^n	6^n	7^n	4^n	7^n

У табл. Б.2.16 на кроці 0 $L_1 = 0$, оскільки довжина ребра $c_{s1} = 0$, решта вершин має позначки ∞ , оскільки вони не мають безпосереднього зв'язку з вершиною s .

На кроці 1 вершина 1 отримує постійну позначку, оскільки $L_1 = 0$ є мінімальною серед всіх тимчасових позначок.

Вершини 2 і 3 безпосередньо пов'язані з вершиною 1, останньої з постійно позначених вершин. Для них:

$$L_2 = L_1 + c_{12} = 0 + 3 < \infty \quad \text{і} \quad L_3 = L_1 + c_{13} = 0 + 7 < \infty.$$

Тому вершинам 2 і 3 приписуються нові тимчасові позначки $L_2 = 3$ і $L_3 = 7$.

Далі на кроці 2 вершина 2 отримує постійну позначку 3^n , від неї змінюються тимчасові позначки вершин 3 і 6, а саме:

$$L_3 = L_2 + c_{23} = 3 + 2 = 5 < 7,$$

$$L_6 = L_2 + c_{26} = 3 + 1 = 4 < \infty.$$

У результаті $L_3 = 7$, $L_6 = 4$.

На кроці 3 постійну позначку отримує вершина 6 – 4^n і т.д.

У результаті виконання сьомого кроку вершині t приписується постійна позначка $L_8 = 7^n$. Отже, довжина найкоротшого шляху з s в t дорівнює 7. Цей шлях складається з ребер (i, j) , для кожного з яких різниця між значеннями постійних позначок його кінцевих вершин i і j рівна довжині цього ребра. Тобто умову, за якої вершини i і j належать найкоротшому шляху, може бути записано:

$$L_j = L_i + c_{ij} \quad (Б.2.13)$$

Співвідношення (2.13) можна використовувати рекурсивно, рухаючись від вершини t до вершини s . Визначивши вершину, яка безпосередньо передає t у найкоротшому шляху, повторюватимемо дану процедуру до тих пір, поки не досягнемо вершини s .

Для даного прикладу найкоротший шлях μ_{st} утворюється послідовністю вершин:

$$\mu_{st} = \{ s, 1, 2, 6, 4, 5, t \}.$$

Для визначення передостанньої вершини 5 в шляху μ_{st} необхідно серед постійних позначок табл. Б.2.16 в її останньому рядку вибрати ту, для якої справедливе співвідношення (Б.2.13):

$$L_t = L_5 + c_{5t}, \quad 7 = 7 + 0.$$

Далі серед вершин, суміжних вершині 5, шукаємо ту, для якої справедливо (Б.2.13):

$$\begin{aligned} L_5 &= L_4 + c_{45}, & 7 &= 6+1; \\ L_5 &\neq L_3 + c_{34}, & 7 &\neq 5 + 4. \end{aligned}$$

Цією вершиною є вершина 4. Далі робимо так само, до тих пір, поки не визначимо весь шлях.

Ефективний обчислювальний алгоритм, що дозволяє визначити найкоротші шляхи з фіксованої вершини-витоку s у всю решту вершин-стоків мережі, представлений в [12]. Алгоритм одночасно з пошуком найкоротших шляхів визначає і послідовність вершин, утворюючих шляхи.

Алгоритм полягає в привласненні кожній вершині графа j ($j = \overline{1, n}$, n – число вершин) трьох позначок виду $L_j, N_j, *_{j}$.

Де L_j – довжина найкоротшого по заданому критерію (довжина, вартість, час і т. д.) шляху з вершини-витоку s в дану вершину j ; N_j – номер вершини, яка є передостанньою в найкоротшому шляху з s в j ; $*_{j}$ – службова позначка, формована в процесі пошуку найкоротших шляхів для збільшення швидкості збіжності алгоритму. Приймається $*_{j} = 0$, якщо найкоротший шлях μ_{sj} з s в j ще не визначений; $*_{j} = 1$, якщо найкоротший шлях μ_{sj} визначений.

Суть алгоритму зводиться до виконання наступних кроків.

Крок 1. Привласнення початкових позначок всім вершинам графа:

$$\begin{aligned} L_s &= 0, & N_s &= 0; & *_{s} &= 0 & (s - \text{вершина-виток}); \\ L_j &= \infty, & N_j &= 0, & *_{j} &= 0 & (j = \overline{1, n}; j \neq s). \end{aligned}$$

Номери передостанніх вершин позначені нулями. Мається на увазі, що в нумерацію вершин графа число «0» не включене.

Крок 2. Серед всіх вершин графа j ($j = \overline{1, n}$), в які найкоротший шлях ще не визначений, тобто у яких $*_j = 0$, вибирається вершина i , що має мінімальне значення L_i (при першому виборі завжди $i = s$).

Крок 3. Для всіх вершин q , суміжних вибраній i , у яких $*_q = 0$, формується нова позначка

$$L_q^u = L_i + l_{iq}.$$

Тут L_i – позначка вибраної вершини i , l_{iq} – довжина (вартість, час і т. д.) ребра зв'язку вершини i та перевіреної вершини q .

Нова позначка L_q^u порівнюється з позначкою L_q , яка вже привласнена вершині q .

Якщо $L_q^u < L_q$ відбувається заміна позначок L_q і N_q на вершині q : $L_q = L_q^u$; $N_q = i$; $*_q$ – не змінюється.

Крок 4. Після перевірки всіх вершин, суміжних вершині i , вершині i привласнюється позначка $*_i = 1$.

Крок 5. Перевірка закінчення роботи алгоритму за умовою: $n - 1$ вершина отримала позначку $*$. Якщо умова виконується, то алгоритм роботу закінчує, якщо ні, то здійснюється перехід до кроку 2 алгоритму.

У результаті роботи алгоритму на кожній вершині одержана інформація про довжину найкоротшого шляху з вершини s у відповідну вершину. Послідовність вершин, що становлять шлях, визначається позначками N_j . Ця послідовність буде зворотною. Замінити її на пряму не дуже важко.

Так, нехай необхідно отримати шлях μ_{si} – з s у j . На вершині j визначається передостання вершина p у шляху $p = N_j$, далі на вершині p визначається чергова вершина $q = N_p$ і т. д., до тих пір, поки в якості чергової певної вершини не виявиться виток s .

Розглянемо граф, зображений на рис. Б.2.11, числа якого біля ребер відповідають значенням l_{ij} . Нехай в якості виток вибрана вершина $s = 1$.

Початкові позначки: $L_1 = 0$, $N_1 = 0$, $*_1 = 0$.

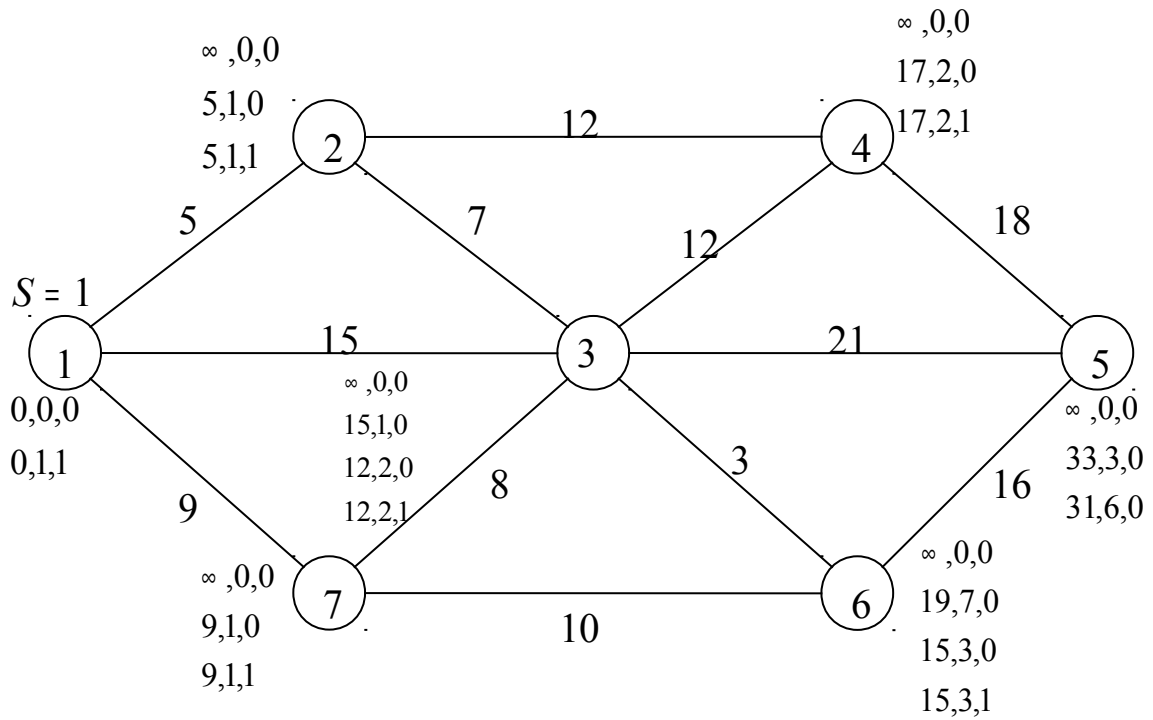


Рисунок Б.2.11 – Зображення графа

Вся решта вершин отримує позначки $L_j = \infty$,

$$N_j = 0, *_{j} = 0, (j = \overline{1, n}; j \neq s).$$

Біля кожної вершини на рис. Б.2.11 представлені позначки, отримувані в процесі роботи алгоритму. Самі верхні позначки – початкові.

Спочатку всі вершини графа мають третю позначку $*=0$, серед вершин вибирається $i = s = 1$, оскільки $L_1 = 0$, а для всієї решти вершин $L_j = \infty$.

З вершиною 1 мають безпосередні зв'язки вершини 2, 3, 7. Для них формуються нові позначки:

$$L_2'' = L_1 + l_{12} = 0 + 5 = 5 < L_2 = \infty,$$

$$L_3'' = L_1 + l_{13} = 0 + 15 = 15 < L_3 = \infty,$$

$$L_7'' = L_1 + l_{17} = 0 + 9 = 9 < L_7 = \infty.$$

Як бачимо, для всіх вершин нова позначка повинна замінити стару, при цьому $N_2 = 1, N_3 = 1, N_7 = 1$. Нові позначки на вершинах 2, 3, 7 записані в другому рядку.

Вершині 1 приписується позначка $*_1 = 1$, і вершина 1 зі всіх подальших дій вибуває.

Далі вибирається вершина 2: $i = 2, L_2 = 5$. Від неї проглядаються вершини 3 і 4:

$$L_3''' = L_2 + l_{23} = 5 + 7 = 12 < L_3 = 15,$$

$$L_4^n = L_2 + l_{24} = 5 + 10 = 15 < L_4 = \infty.$$

В результаті на вершинах 3 і 4 записуються нові позначки:

$$\begin{array}{ll} L_3=12; & N_3=2; \\ L_4=15; & N_4=2. \end{array}$$

Вершина 2 отримує позначку $*_2 = 1$ і з подальшого розгляду вибуває. Процес продовжується до тих пір, поки шість $(n - 1)$ вершин не отримають позначку $*_j = 1$.

Результуючі позначки записані останніми в стовпцях біля кожної вершини. Одержані позначки містять всю необхідну інформацію про найкоротші шляхи з вершини 1. Так, на вершині 2 маємо результуючі позначки: $L_2 = 5$; $N_2=1$. Це означає, що довжина шляху $\mu_{12} = L_2 = 5$, шлях проходить через вершину 1, але $s = 1$, отже, шлях прямої, тобто $L_{12} = 5$; $\mu_{12} = \{1, 2\}$ (L_{12} – довжина шляху μ_{12}).

Для вершини 5: $L_5 = 31$, $N_5 = 6$. Довжина шляху $L_{15} = 31$, шлях проходить через вершину б, але $s \neq 6$, отже, далі звертаємося до вершини 6 і визначаємо чергову вершину в шляху: $N_6 = 3$. Оскільки $3 \neq s$, звертаємося до вершини 3 і визначаємо $N_3 = 2$, далі $N_2 = 1$. Але $s = 1$, отже, шлях визначений. Отже, $L_{15} = 31$; $\mu_{15} = \{1, 2, 3, 6, 5\}$.

Аналогічно отримуємо решту шляхів:

$$\begin{array}{l} L_{13} = 12; \mu_{13} = \{1, 2, 3\}; \\ L_{14} = 17; \mu_{14} = \{1, 2, 4\}; \\ L_{16} = 15; \mu_{16} = \{1, 2, 3, 6\}; \\ L_{17} = 9; \mu_{17} = \{1, 7\}. \end{array}$$

Розглянутий алгоритм може бути використаний для пошуку найкоротшого шляху між будь-якими двома заданими вершинами s і t . При цьому вершина s фіксується як джерело, алгоритм працює відповідно до представленої послідовності кроків і закінчує роботу, як тільки в якості чергової вибраної на кроці 2 вершини виявляється вершина t , тобто $i = t$.

Даний алгоритм може бути також використаний для пошуку шляхів мінімального рангу. Для цього достатньо всім ребрам графа надати вагу, яка дорівнює одиниці: $l_{ij} = 1$. Тоді в результаті роботи алгоритму позначка L_j – на вершинах графа ($j = \overline{1, n}$) визначить мінімальний ранг шляху з джерела у вершину j .

СПИСОК ОСНОВНОЇ ЛІТЕРАТУРИ

1. Айвазян С.А. Прикладная статистика и основы эконометрики: [учеб. для вузов / С.А. Айвазян, В.С. Мхитарян – М.: ЮНИТИ, 1998. – 1022 с.
2. Болч. Многомерные статистические методы для экономики. Болч, К.Дж. Хуань – М.: Статистика, 1979. – 318 с.
3. Венецкий И.Г. Основные математико-статистические понятия и формулы в статистическом анализе: [Справочник]. / И.Г. Венецкий, В.И. Венецкая – 2-е изд., перераб. и доп. – М.: Статистика, 1979. – 447 с.
4. Вини Р. Введение в прикладной экономической анализ / Р. Вини, К. Холден. – М.: Финансы и статистика, 1981. – 294 с.
5. Вучков И. Прикладной линейный регрессионный анализ / И. Вучков, Л. Бояджева, Е. Солаков. – М.: Финансы и статистика, 1987. – 239 с.
6. Грубер Й. Эконометрія: Вступ до множинної регресії та економітриї / Грубер Й. В 2-х т. – К.: Нічлава, 1998. – Т.1. (384 с.), 1999. – Т.2 (308 с.).
7. Демиденко Е.З. Линейная и нелинейная регрессии / Демиденко Е.З. – М.: Финансы и статистика, 1981. – 302 с.
8. Джонстон Дж. Эконометрические методы / Джонстон Дж. – М.: Статистика, 1980. – 444с.
9. Введение в эконометрику / Доугерти К.; Пер. с англ. – М.: ИНФРА-М, 1997. XIV. – 402 с.
10. Дрейпер Н. Прикладной регрессионный анализ / Н. Дрейпер, Г. Смит – М.: Финансы и статистика, 1986. – Т.1 – 365 с.; Т.2 – 379 с.
11. Емельянов А.С. Эконометрия и прогнозирование / Емельянов А.С. – М.: Экономика, 1985. – С. 82-89.
12. Єлейко В. Основи економітриї / Єлейко В. – Львів: “Марка Лтд”, 1995. – 191 с.
13. Иванова В.М. Экономическая теория / Иванова В.М. Основы бизнеса: Ч.ІУ: Эконометрика; Ред. совет: А.Д. Смирнов, В.Ф. Максимова и др. – М.: СОМИНТЭК, 1991. – 158 с.
14. Кейн Э. Экономическая статистика и эконометрия. Введение в количественный экономический анализ / Кейн Э. – М.: Статистика, 1997. – 254 с.
15. Кокрен У. Методы выборочного исследования / Кокрен У. – М.: Статистика, 1976. – 410 с.
16. Ланге О. Введение в эконометрию / Ланге О. – М.: Прогресс, 1964. – 60 с.
17. Теоретико-ймовірнісні та статистичні методи в економітриці та фінансовій математиці / [Леоненко М.М., Мішура Ю.С., Пархоменко В.М., Ядренко М.Й.] – К.: ІНФОРМТЕХНІКА, 1995. – 380 с.
18. Лук'яненко І.Г. Эконометрика: [Підручник] / Лук'яненко І.Г., Краснікова Л.І. – К.: Т-во "Знання", КОО, 1998. – 494 с.
19. Магнус Я.Р. Эконометрика: начальный курс / Магнус Я.Р. Катышев П.К., Пересецкий А.А. – М.: Дело, 1997. – 248 с.
20. Маленво Э. Статистические методы эконометрии / Маленво Э. – М.: Статистика, 1975. – 423 с.

21. Мартышюс С. Методологические проблемы построения и применения эконометрических моделей / Мартышюс С. – Вильнюс: Максчас, 1979. – 170 с.
22. Наконечний С.І. Економетрія: навч. посіб. / Наконечний С.І., Терещенко Т.О., Романюк Т.П. – К.: КНЕУ, 1997. – 352 с.
23. Пономаренко О.І. Основи математичної економіки / Пономаренко О.І., Перестюк М.О., Бурим В.М. – К.: ШФОРМТЕХНІКА, 1995. – 319 с.
24. Тинтнер Г. Введение в эконометрию / Тинтнер Г. – М.: Статистика, 1965. – 361 с.
25. Толбатов Ю.А. Економетрика: [підруч. для студ. екон. спец, вищ.навч. закл.] / Толбатов Ю.А. – К.: Четверта хвиля, 1997. – 320 с.
26. Устойчивые статистические методы оценки данных; под ред. Р.Л. Лонера, Г.Н. Уилкинсона; пер. с англ. – М.: Машиностроение, 1984. – 232 с.
27. Фишер Ф. Проблема идентификации в эконометрии / Фишер Ф. – М.: Статистика, 1978. – 223 с.
28. Хейс Д. Причинный анализ в статистических исследованиях / Хейс Д. – М.: Финансы и статистика, 1981. – 225 с.
29. Шаттелис Т. Современные эконометрические методы / Шаттелис Т. – М.: Статистика, 1975. – 161 с.
30. Hendry D.F. Dynamic econometrics / Hendry D.F. New York: Oxford University Press Inc., 1995. – P. 3-619.
31. Johnston J. Econometrics Methods 2nd edition / Johnston J. New York: McGraw-Hill, 1972, – P. 16-80.
32. Judge, G. G. and ets. The Theory and Practice of Econometrics edition / Judge, G. G. New York: John Wiley, 1985, – P. 3-120.

СПИСОК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

33. Аганбегян А. Г. Система моделей народнохозяйственного планирования / Аганбегян А. Г., Багриновский К.А., Гранберг А.Г. – М.: Мысль 1972. – 351 с.
34. Адирим И.Г. Система моделей прогнозирования роста народного хозяйства республики / Адирим И.Г., Янов Я.А., Почс Р.Я. – Рига: Зинатне, 1975. – 191 с.
35. Аткинсон Э.Б. Лекции по экономической теории государственного сектора / Э.Б. Аткинсон, Дж.Э. Стиглиц – М.: Аспект Пресс, 1995. – 832 с.
36. Бакаев А.А. Имитационные модели в экономике / Бакаев А.А., Костина Н.И., Яровицкий Н.В. – К.: Наук, думка, 1978. – 304 с.
37. Геец В.М. Отраслевое прогнозирование: методологический и организационный аспекты / Геец В.М. – К.: Наук, думка, 1990. – 120 с.
38. Гранберг А.Г. Динамические модели народного хозяйства / Гранберг А.Г. – М.: Экономика, 1985. – 204 с.
39. Гранберг А.Г. Статистическое моделирование и прогнозирование / Гранберг А.Г. – М.: Финансы и статистика, 1990. – 378 с.
40. Дадаян В.С. Моделирование глобальных экономических процес сов /

Дадаян В.С. – М.: Экономика, 1984. – 278 с.

41. Данилов-Данильян В.И Система оптимального перспективного планирования народного хозяйства. Проблемы теории и методологии / В.И. Данилов-Данильян, М.Г. Завельский – М.: Наука, 1975. – 320 с.

42. Дружинин В.В. Проблемы системологии. Проблемы теории сложных систем / В.В. Дружинин, Д.С. Конторов – М.: Радио и связь, 1986. – 296 с.

43. Дружинин В.В. Системотехника / В.В. Дружинин, Д.С. Конторов – М.: Радио и связь, 1985. – 200 с.

44. Дюран Б. Кластерный анализ / Б. Дюран, П. Одел – М.: Статистика 1977. – 128 с.

45. Емельянов А.С. Общественное производство: Динамика, тенденции, модели / Емельянов А.С. – К.: Наук, думка, 1980. – С.347-409.

46. Емельянов А.С. Многорегиональная эконометрическая модель УКР-3: Плановое управление экономикой развитого социализма / А.С. Емельянов, В.П. Кузьменко В 5 т. – К.: Наук. думка, 1985. – Т.1: Народнохозяйственные процессы, их планирование и прогнозирование. – С.285-289.

47. Региональное управление: методология и моделирование / [В.А. Забродский, Ю.Н. Донченко, Н.А. Кизим и др.]. – Харьков: "Основа", 1991. –96 с.

48. Ивахненко А.Г. Долгосрочное прогнозирование и управление сложными системами / Ивахненко А.Г. – К.: Техника, 1975. – 311 с.

49. Ивахненко А.Г. Алгоритмы метода группового учета аргументов (МГУА) при непрерывных и бинарных признаках / Ивахненко А.Г. – К.: ИК-НАНУ. –1992. –49 с.

50. Ивахненко А.Г. Моделирование сложных систем по экспериментальным данным / А.Г. Ивахненко, Ю.П. Юрачковский – М.: Радио и связь, 1987. – 115 с.

51. Ивахненко А.Г. Исследование многокритериального подхода при селекции пар аргументов и восстановления решающего правила в вероятностных алгоритмах МГУА / Ивахненко А.Г., Ивлев М.М., Сегал В.В. //Автоматика. – 1981. –№1. – С. 88-91.

52. Имитационное моделирование производственных систем; /под. общ. ред. чл.-кор. АН СССР А.А.Вавилова. – М.: Машиностроение; Берлин: Техника, 1983. –416 с.

53. Колек Ю. Эконометрические модели в социалистических странах / Ю. Колек, И. Шуян; пер. со словац. – М.: Экономика, 1978. – 151с.

54. Королев О.А. Проблемы конструирования и использования макроэкономических эконометрических моделей переходной экономики: на примере Украины / Королев О.А. – К.: ТОВ "Міжнар. фін. агенція", 1997. – 219 с.

55. Корольов О.А. Економетрія в задачах, ситуаціях та проблемах для студентів спеціальності «Фінанси і кредит»: конспект лекцій і практикум / Корольов О.А. В 3-х ч. – К.: Київ. держ. торг.-екон. ун-т. – Ч.1 (1996. – 66 с.); Ч.2 (1996. – 72 а); Ч.3 (1997. – 81 с).

56. Кулинич О.І. Економетрія / Кулинич О.І. навч. посіб. – Хмельницький.: Поділля, 1997. – 115 с.

57. Михалевич М.В. Динамические макромоделли нестабильных процес-

сов при переходе к рыночной экономике / М.В. Михалевич, А.Ю. Чижевская // Кибернетика и системный анализ. – 1993. – №4. – С.81-88.

58. Михалевич В.С. Динамические макромоделли процессов ценообразования в переходной экономике / В.С. Михалевич, М.В. Михалевич // Кибернетика и системный анализ. – 1995. – №3. – С.116-129.

59. Моисеев Н.Н. Математические задачи системного анализа / Моисеев Н.Н. – М: Наука, 1981. – С.3-225.

60. Молчанов А.А. Моделирование и проектирование сложных систем / Молчанов А.А. – К.: Выща шк., 1988. – С. 6-44.

61. Новицкий П.В. Оценка погрешностей результатов измерений / П.В. Новицкий, И.А. Зограф – Л.: Энергоатомиздат, 1991. – 303 с.

62. Панасюк Б. Концептуальные основы экономического прогнозирования и планирования / Панасюк Б. // Экономика Украины. – 1996. – №5. – С. 7-17.

63. Панасюк Б. Прогнозирование развития экономики Украины / Б. Панасюк, И. Сергиенко, Гуляницкий // Экономика Украины. – 1996. – №1. – С.20-31.

64. Петров А.А. Системный анализ развивающейся экономики: Многосекторная модель и учет природных ресурсов / А.А. Петров, И.Г. Поспелов III / Изв. АН СССР. Техническая кибернетика. – 1979. – №4. – С.11-23.

65. Сегал В.В. Анализ и синтез сложных систем: Опыт системного исследования / Сегал В.В. – К.: ЦЭМИ "Тридента", 1994. – С. 59-153.

66. Методы организации адаптивного планирования и управления в экономикопроизводственных системах / [В.И. Скурихин, В.А. Забродский, П.А. Ивашенко и др.] – К.: Наук, думка, 1980. – 272 с.

67. Трансформація моделі економіки України (ідеологія, протиріччя, перспективи) / ІЕП: за ред. акад. НАНУ В.М. Гейця. – К.: Логос, 1999. – 500 с.

68. Фишберн П. Теория полезности для принятия решений / Фишберн П. – М.: Наука, 1978. – 352 с.

69. Фомин Б.С. Эконометрические теории и модели международных экономических отношений / Фомин Б.С. – М.: Мысль, 1970. – 268 с.

70. Хохлов Е.Н. Приоритетные идеи в области управления / Е.Н. Хохлов, Н.А. Бурьгин – К.: Либра-НМЦПА, 1993. – 102 с.

Навчальне видання

БОБРОВНИЧА Наталія Серафимівна
БОРИСЕВИЧ Євгенія Георгіївна

ЕКОНОМЕТРІЯ

Навчальний посібник

Редактор

Кодрул Л. А.

Комп'ютерне верстання
та макетування

Корнійчук Є. С.

Сдано в набір 06.05.2010 Підписано к печати 06.07.2010.
Формат 60/88/16 Зак. № 4237
Тираж 200 экз. Об'єм: 11,25 усл. печ. л.; 13,9 усл. авт. печ. л.
Отпечатано на издательском оборудовании фирмы RISO
в типографии редакционно-издательского центра ОНАС им. А.С. Попова
ОНАС, 2010