

В.С. Григорків, М.В. Григорків

ЕКОНОМІЧНА КІБЕРНЕТИКА

ПРАКТИКУМ



Міністерство освіти і науки України
Чернівецький національний університет
імені Юрія Федьковича

В.С. Григорків, М.В. Григорків

ЕКОНОМІЧНА КІБЕРНЕТИКА

ПРАКТИКУМ

Навчальний посібник



Чернівці

Чернівецький національний університет
імені Юрія Федьковича

2024

УДК 330.46(075.8)

Г 831

Друкується за ухвалою Вченої ради
Чернівецького національного університету імені Юрія Федьковича
(протокол № 7 від 29 квітня 2024 р.)

Рецензенти:

- Ляшенко О.І.** – завідувач кафедри економічної кібернетики Київського національного університету імені Тараса Шевченка, доктор економічних наук, професор
- Маляр М.М.** – професор кафедри кібернетики і прикладної математики ДВНЗ «Ужгородський національний університет», доктор технічних наук, професор

Григорків В.С., Григорків М.В.

Г 831 Економічна кібернетика: практикум : навч. посіб. Чернівці : Чернівець. нац. ун-т ім. Ю. Федьковича. 2024. 200 с.

ISBN 978-966-423-844-8

Навчальний посібник містить типові завдання для практичних і лабораторних занять, самостійної та індивідуальної роботи, контролю і самоконтролю базових знань із економічної кібернетики та приклади їх виконання. Тематика запропонованих завдань охоплює елементи теорії систем, системного аналізу та систем управління, основи методу моделювання та проблематику адекватності моделей, елементи теорії інформації та її кількісного вимірювання, аналізу та синтезу економічних систем тощо.

Для студентів економічних і інших прикладних спеціальностей.

УДК 330.46(075.8)

ISBN 978-966-423-844-8

© Чернівецький національний університет,
імені Юрія Федьковича, 2024

© В.С. Григорків, 2024

© М.В. Григорків, 2024

ЗМІСТ

ПЕРЕДМОВА	5
Тема 1. Абстрактні структури систем і способи їх задання	8
Тема 2. Метричні характеристики та топологічний аналіз структур систем, заданих графами.....	19
Тема 3. Оптимальні маршрути та шляхи у мережних системах.....	26
Тема 4. Оптимальні потоки у мережних системах	34
Тема 5. Ізоморфні, гомеоморфні та гомоморфні системи. Зв'язок між системою-оригіналом і системою-моделлю	45
Тема 6. Кібернетичний підхід до моделювання систем типу «чорний ящик»	62
Тема 7. Відтворення (імітація) поведінки систем як окремий підхід до їх моделювання: диференціальні моделі неперервних динамічних систем.....	77
Тема 8. Відтворення (імітація) поведінки систем як окремий підхід до їх моделювання: різницеві моделі дискретних динамічних систем.....	88

Тема 9. Системи управління. Оптимальне управління динамічними системами	100
Тема 10. Ентропія систем зі скінченними кількостями можливих станів і кількісна міра інформації про них.....	118
Тема 11. Кількісне вимірювання невизначеності та інформації у випадку систем із неперервними множинами можливих станів	134
Тема 12. Вертикальна декомпозиція лінійних моделей оптимального планування виробництва методом Данцига-Вулфа	146
Тема 13. Вертикальна декомпозиція лінійних моделей оптимального планування діяльності виробничої системи методом Корнаї-Ліптака.....	157
Тема 14. Агрегування у моделях міжгалузевих балансів.....	166
Тема 15. Агрегування результатів функціонування багатопродуктового виробництва за допомогою методу апроксимації виробничих можливостей.....	187
Список використаних джерел	197

ПЕРЕДМОВА

Коротко та лаконічно під кібернетикою як окремою наукою треба розуміти науку про управління, оскільки об'єктом її вивчення і є системи управління різної природи, у тому числі соціально-економічні системи, якими займається економічна кібернетика. Сама назва «кібернетика» грецького походження і використовувалася вона у давньогрецькому суспільстві тоді, коли йшлося про управління у тому чи іншому контексті, а знаменитий філософ Платон ще й називав кібернетику мистецтвом управління, зокрема й мистецтвом управління людьми.

Незважаючи на те, що початковим періодом розвитку кібернетики вважають кінець першої половини ХХ століття і пов'язують його із відомими публікаціями Норберта Вінера, ще задовго до нього зустрічаються наукові праці з кібернетики як науки про управління системами, які мають певну складність і змінюються з плином часу, тобто є динамічними. Сьогодні кібернетика займає особливе місце у науці та освіті, а з огляду на її фундаментальну роль у процесах прийняття рішень та власне у процесах управління ще й у багатьох реальних сферах суспільної діяльності. У зв'язку з цим пов'язані з кібернетикою навчальні дисципліни входять до складу освітніх програм підготовки фахівців багатьох спеціальностей, у тому числі економістів-кібернетиків, для яких дисципліна з

однойменною назвою «Економічна кібернетика» має бути особливою у сенсі синергії знань про соціально-економічні системи й управління ними. При цьому здобуття таких знань і вміння застосовувати їх на практиці можливе тоді, коли здобувач освіти має у своєму розпорядженні повноцінне навчально-методичне забезпечення для відповідних лекційних, практичних чи лабораторних і самостійних занять. Саме з метою формування такого забезпечення розроблено і цей навчальний посібник.

Посібник містить практичні завдання та приклади їх виконання, спрямований найперше на поглиблення теоретичних знань із ключових понять, принципів і методів кібернетики, зокрема із теорії систем і системного аналізу, принципів і законів управління, методу моделювання та проблеми адекватності моделей, теорії інформації та її кількісного вимірювання, аналізу та синтезу економічних систем тощо. Якісне виконання цих завдань передбачає наявність деяких базових знань із інших дисциплін економіко-математичного циклу та власне теоретичного матеріалу з економічної кібернетики, викладеного у навчально-методичній літературі авторів цього посібника та інших літературних джерел, зокрема тих, що вказані у списку використаних на останніх сторінках посібника. Характерною особливістю пропонованих у посібнику завдань і відповідних прикладів є також акцент на самостійну та індивідуальну роботу, контроль і самоконтроль окремих знань із уведених до посібника розділів кібернетики.

Матеріал пропонованого навчального посібника охоплює лише частину тієї проблематики, яку можна віднести до сучасної кібернетики, але його цілком достатньо для оволодіння принаймні елементарними знаннями із кібернетики та розуміння її значення у сучасному освітньому та науковому просторі. Саме у цьому контексті даний посібник корисний для студентів прикладних економічних та інших спеціальностей.

ТЕМА 1. АБСТРАКТНІ СТРУКТУРИ СИСТЕМ І СПОСОБИ ЇХ ЗАДАННЯ

Завдання 1.1. Побудувати геометричні ілюстрації структур: лінійної (з послідовними зв'язками), паралельної (з паралельними зв'язками), повнозв'язаної, кільцевої, зіркової (типу зірки), матричної, ієрархічної (деревоподібної), з трьома центрами керування. Обґрунтувати кожну із них щодо характеру зв'язків між елементами та обміну інформацією між ними.

◀ Відповідні структури зображені на рис. 1.1.1 – 1.1.7.



Рис. 1.1.1. Лінійна структура

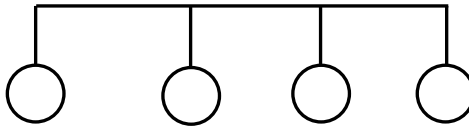


Рис. 1.1.2. Паралельна структура

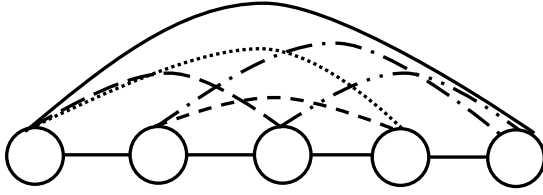


Рис. 1.1.3. Повнозв'язна структура

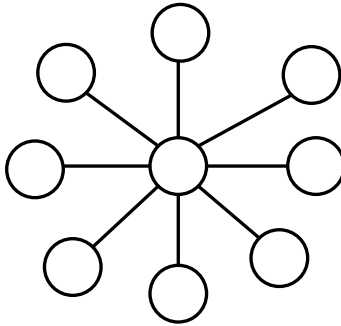


Рис. 1.1.4. Зіркова структура

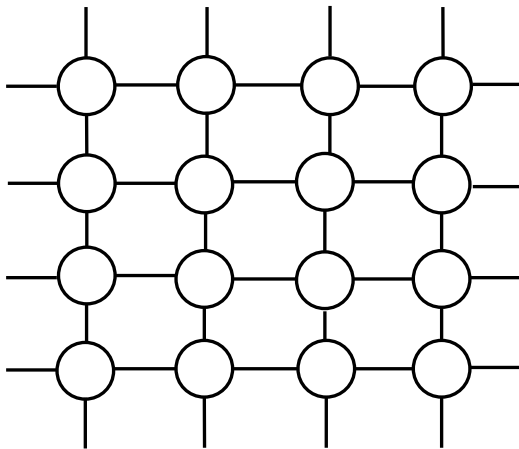


Рис. 1.1.5. Матрична структура

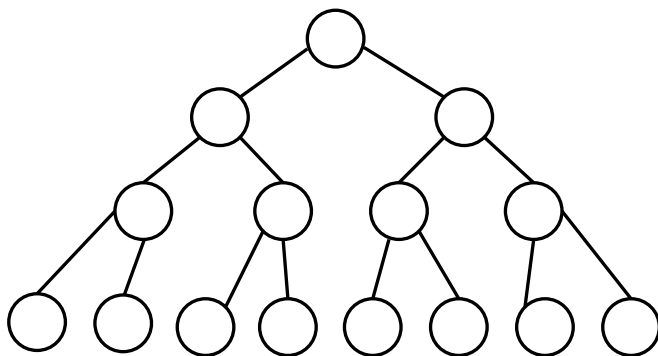


Рис. 1.1.6. Ієрархічна (деревоподібна) структура

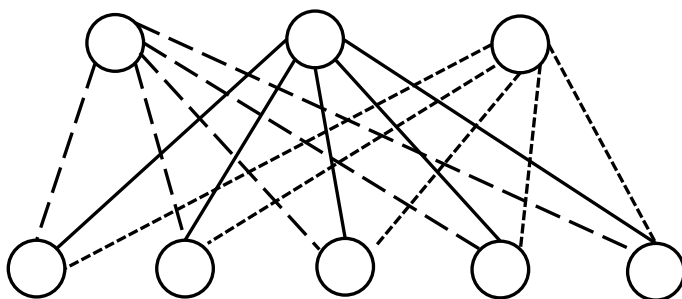


Рис. 1.1.7. Структура з трьома центрами керування

Уточнимо деякі особливості проілюстрованих вище структур. Для лінійної структури характерним є те, що всі її елементи, крім крайніх, мають зв'язок з двома сусідніми елементами. При цьому інформація, що передається з одного крайнього елемента до іншого крайнього елемента, є доступною для всіх елементів.

Система руйнується, якщо розривається будь-який зв'язок.

Порівняно з лінійною система з паралельною структурою має більшу швидкість обміну інформацією і є більш надійною.

Що стосується повнозв'язної структури, то тут кожний елемент зв'язаний з усіма іншими елементами, причому усі зв'язки вважаються рівноцінними, а швидкість поширення інформації та надійність – максимальними. Втім, такого типу системи неекономічні із-за надмірності зв'язків.

Для структур типу зірки характерним є наявність центрального елемента (елемента керування), за допомогою якого реалізуються усі зв'язки. Зіркова структура є частинним випадком ієрархії.

Висока надійність властива матричній структурі. Надійність тут забезпечується великою кількістю шляхів поширення інформації.

У ієрархічній структурі зв'язки здійснюються за принципом «керування – підлеглість», причому кількість зв'язків на вищих рівнях є меншою, хоча рівень агрегування збільшується.

Структури з багатьма (наприклад, з трьома) центрами керування за побудовою є певним узагальненням ромбовидної структури, у якої немає елемента керування вищого рівня. ►

Завдання 1.2. Для сукупності n ($n > 5$) довільно вибраних населених пунктів нашої країни потрібно побудувати допустиму мережу доріг (комунікацій), яка дозволяє проїхати із будь-якого одного населеного пункту у будь-який інший населений пункт, якщо:

а) для кожної пари цих пунктів можливий як мінімум один із типів зв'язків (прямий, опосередкований);

б) безпосередній (прямий) зв'язок заборонений між заданими двома пунктами;

в) три пункти мають з одним пунктом досліджуваної сукупності пунктів прямий зв'язок, а з усіма іншими – опосередкований зв'язок.

◀ Нехай $n = 6$, причому населені пункти не конкретизовані. Із умови завдання однозначно випливає, що проект чи модель допустимої мережі доріг є графом [16, 17, 20] (вершини графа – населені пункти, а ребра – дороги між ними), але залежно від тих чи інших обмежень його структура може бути різною. У випадку а) прикладом допустимої мережі може бути мережа, зображена на рис. 1.2.1.

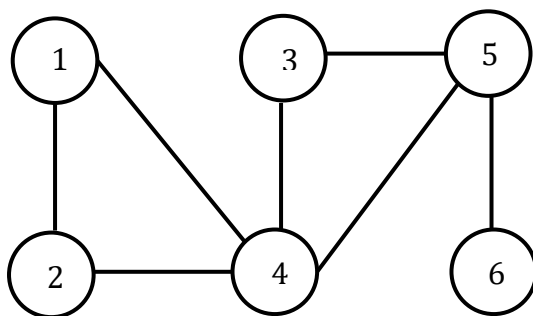


Рис. 1.2.1

Крім того, ця мережа (рис. 1.2.1) відображає також ситуацію б), коли існує заборона на будівництво прямої дороги між пунктами 1 та 3, а також 4 та 6.

Очевидно, що мережа (рис. 1.2.1) є неорієнтованим і зв'язним графом. Комунікація (зв'язок) між будь-якими двома пунктами тут може здійснюватися за допомогою простого ланцюга (ланцюга від пункту i до пункту j з різними ребрами та вершинами). У цьому графі (цій мережі) є не тільки дерева (зв'язні підграфи чи графи без циклів), але й покривне дерево (дерево, що містить усі вершини) з послідовними вершинами 1, 2, 4, 3, 5, 6. Наявність покривного дерева означає, що по існуючій мережі доріг можна проїхати між будь-якими пунктами. Є також замкнені маршрути з різними ребрами, тобто цикли, наприклад циклами є ланцюги, що зв'язують вершини 1, 2, 4 або 4, 3, 5.

У випадку обмеження в) прикладом допустимої мережі є граф, проілюстрований на рис. 1.2.2.

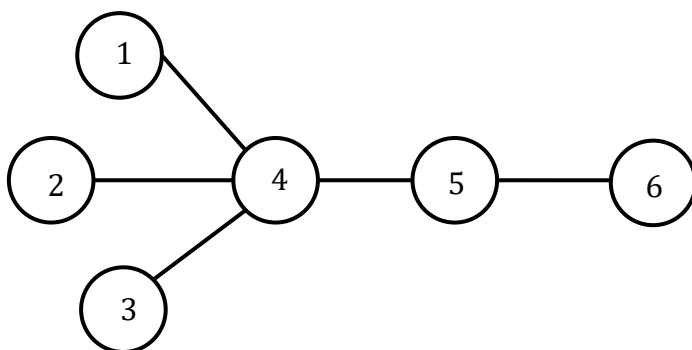


Рис. 1.2.2

Очевидно, що тут кожний із пунктів 1, 2, 3 має прямий зв'язок із пунктом 4, а з пунктами 5, 6 та між собою – опосередкований зв'язок. Граф (рис. 1.2.2) має як окремі дерева, так і сам є деревом, однак покривного дерева немає.

Зрозуміло, що мереж (графів) із зазначеними обмеженнями а), б), в) можна побудувати чимало. Зауважимо, що тут йдеться лише про допустимі, а не оптимальні мережі. ►

Завдання 1.3. Для зображеної на рис. 1.3.1 – 1.3.5 структури, яка представлена за допомогою графа з орієнтованими та неорієнтованими ребрами (орієнтовані ребра або дуги позначені відповідною стрілкою, а неорієнтовані – без стрілки, хоча вони можуть бути замінені парю орієнтованих дуг із протилежними напрямками), побудувати матрицю суміжностей вершин

$$U = (u_{ij})_{i,j=1}^n \quad (n - \text{кількість вершин}), \text{ кожний елемент}$$

якої визначається за формулою

$$u_{ij} = \begin{cases} 1, & \text{якщо } (i, j) - \text{ребро або дуга, що виходить} \\ & \text{з вершини } i \text{ та входить у вершину } j; \\ 0, & \text{якщо мають місце інші випадки} \end{cases} \quad (1.3.1)$$

та матрицю інцидентності $V = (v_{ij})_{i,j=1}^{n,m}$ (n – кількість вершин, m – кількість ребер чи дуг), кожний елемент якої визначається за формулою

$$v_{ij} = \begin{cases} -1, & \text{якщо дуга } j \text{ виходить з} \\ & \text{вершини } i \text{ та не є петлею;} \\ 1, & \text{якщо вершина } i \text{ є кінцем ребра} \\ & j \text{ або якщо дуга } j \text{ входить} \\ & \text{у вершину } i; \\ 0, & \text{якщо мають місце інші випадки,} \\ & \text{тобто вершина } i \text{ не є інцидентною} \\ & \text{ребру чи дузі } j. \end{cases} \quad (1.3.2)$$

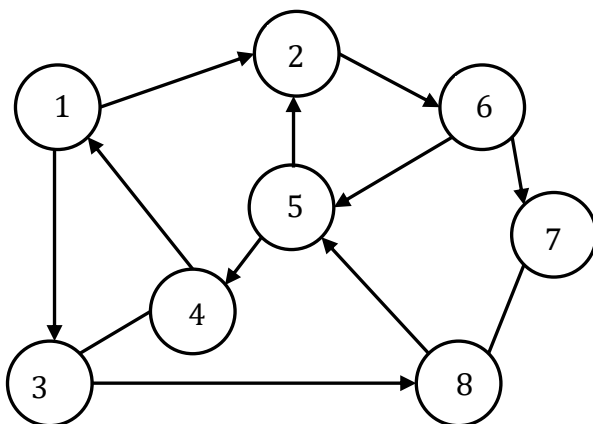


Рис. 1.3.1

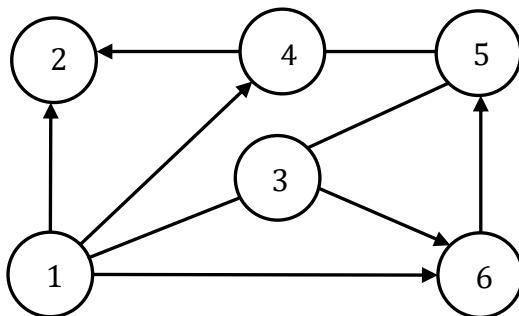
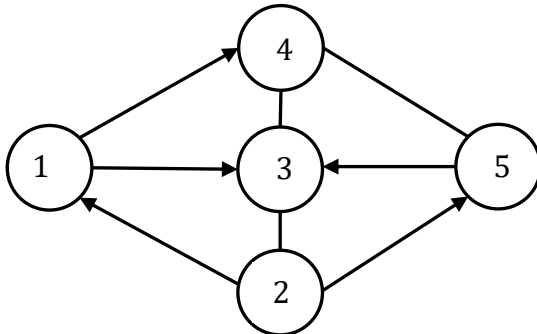
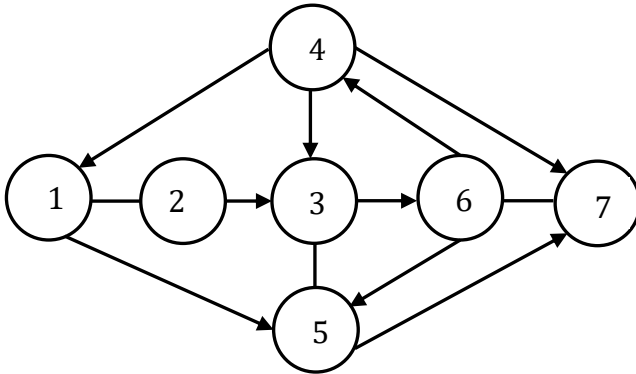


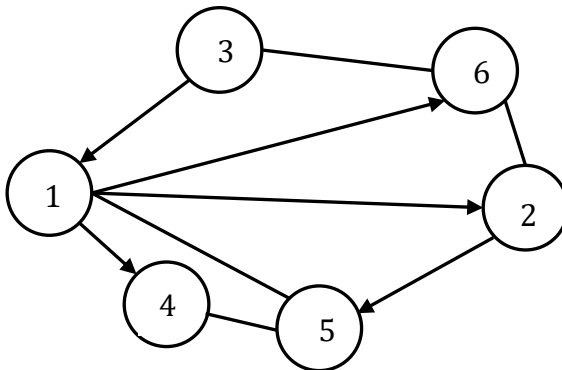
Рис. 1.3.2



Puc. 1.3.3



Puc. 1.3.4



Puc. 1.3.5

◀ Приклад побудови матриць U та V продемонструємо для зображеної на рис. 1.3.1 структури. Оскільки у структурі (графі) системи (рис. 1.3.1) є вісім елементів (вершин), то з урахуванням формули (1.3.1) матриця $U = (u_{ij})_{i,j=1}^8$, а саме:

$$U = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}.$$

Зауважимо, $u_{ii} = 0$ ($i = \overline{1,8}$) та $u_{ij} = u_{ji} = 1$, якщо ребро (i, j) є неорієнтованим, оскільки, як уже було сказано, це ребро замінюється парою орієнтованих дуг (i, j) та (j, i) . Побудова матриць суміжностей для інших структур (рис. 1.3.2 – 1.3.5) здійснюється аналогічно.

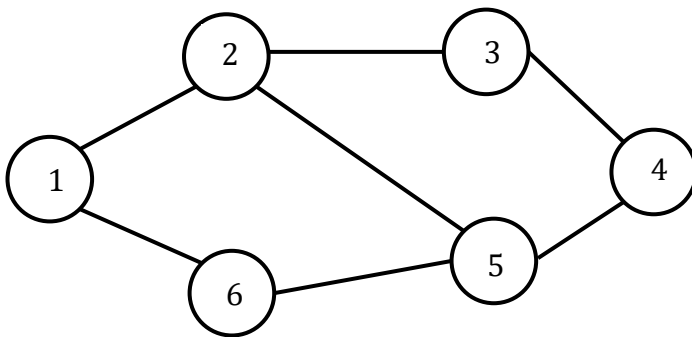
Зображений на рис. 1.3.1 граф має $n=8$ вершин і $m=12$ дуг і ребер (10 дуг і 2 ребра). Вершини пронумеровані, а дуги та ребра $(1,2)$, $(1,3)$, $(2,6)$, $(3,4)$, $(3,8)$, $(4,1)$, $(5,2)$, $(5,4)$, $(6,5)$, $(6,7)$, $(7,8)$, $(8,5)$ пронумеруємо відповідно так: 1, 2, 3, 4, 5, 6, 7, 8, 9,

10, 11, 12. Тоді згідно із формулою (1.3.2) матриця інцидентності $V = (v_{ij})_{i,j=1}^{8,12}$ для цього графа має вигляд:

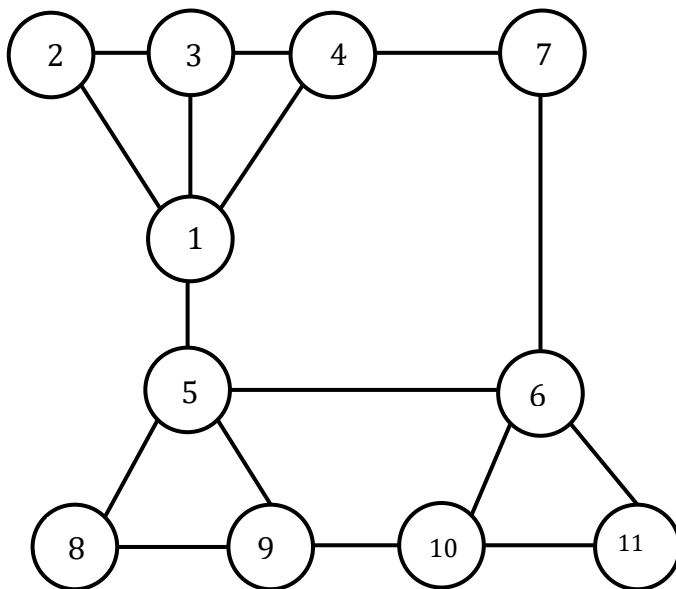
$$V = \begin{array}{c|cccccccccccc} & \begin{array}{c} (i,j) \\ i \end{array} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ \hline 1 & & -1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & & 1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 3 & & 0 & 1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 6 & & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ 7 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 8 & & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{array}$$


ТЕМА 2. МЕТРИЧНІ ХАРАКТЕРИСТИКИ ТА ТОПОЛОГІЧНИЙ АНАЛІЗ СТРУКТУР СИСТЕМ, ЗАДАНИХ ГРАФАМИ

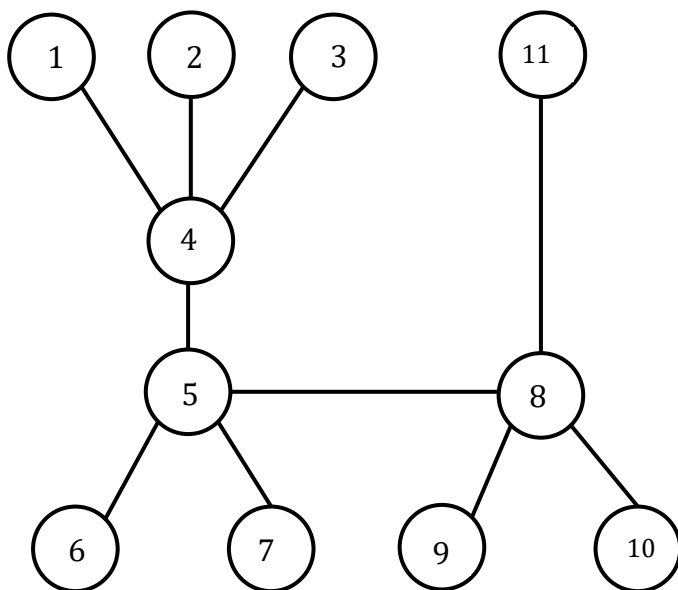
Завдання 2.1. Структури багатьох реальних систем, зокрема економічних систем є зв'язними неорієнтованими графами. Метричними характеристиками таких графів є: відстань між вершинами – довжина найкоротшого ланцюга, який з'єднує ці вершини; діаметр графа – максимальна відстань між вершинами; ексцентриситет вершини – відстань від цієї вершини до найбільш віддаленої від неї вершини; радіус графа – мінімальний ексцентриситет; центр графа – множина його вершин із мінімальним ексцентриситетом. Усі ці характеристики можна визначити на основі матриці відстаней між вершинами. Для проілюстрованих на рис. 2.1.1 – 2.1.5 графів (структур) необхідно побудувати матрицю відстаней $R = (\rho_{ij})_{i,j=1}^n$ (n – кількість вершин графа, ρ_{ij} – відстань від вершини i до вершини j) та уточнити зазначені вище метричні характеристики (довжини усіх ребер вважати такими, що дорівнюють одиниці).



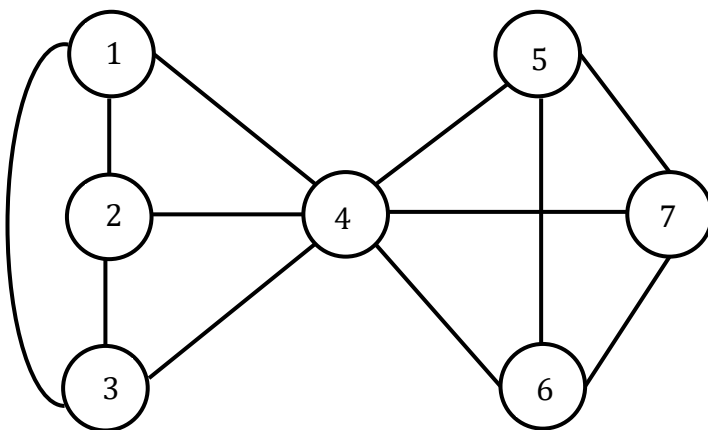
Puc. 2.1.1



Puc. 2.1.2



Puc. 2.1.3



Puc. 2.1.4

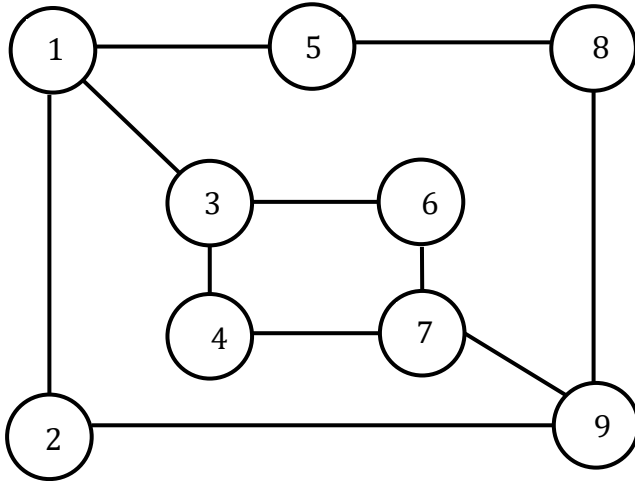


Рис. 2.1.5

◀ Продемонструємо виконання цього завдання на прикладі графа, зображеного на рис. 2.1.1.

$$R = \begin{array}{c|cccccc} & j & & & & & \\ \hline i & & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1 & & (0 & 1 & 2 & 3 & 2 & 1) \\ 2 & & 1 & 0 & 1 & 2 & 1 & 2 \\ 3 & & 2 & 1 & 0 & 1 & 2 & 3 \\ 4 & & 3 & 2 & 1 & 0 & 1 & 2 \\ 5 & & 2 & 1 & 2 & 1 & 0 & 1 \\ 6 & & 1 & 2 & 3 & 2 & 1 & 0 \end{array}$$

Отже, відстані між вершинами визначені (елементи матриці R). Оскільки максимальна відстань (максимальне значення елемента матриці R) дорівнює 3, то діаметр графа $d(H)=3$ (H - граф).

Ексцентриситети для кожної з вершин такі (найбільші елементи відповідних рядків матриці):

$$\varepsilon(1) = 3, \quad \varepsilon(2) = 2, \quad \varepsilon(3) = 3, \quad \varepsilon(4) = 3, \quad \varepsilon(5) = 2, \quad \varepsilon(6) = 3.$$

Мінімальний ексцентриситет або радіус графа $r(H) = 2$. Враховуючи, що $r(H) = \varepsilon(2) = \varepsilon(5) = 2$, центром графа є множина вершин $C(H) = \{2, 5\}$.

Аналогічно виконуються інші підзавдання. ►

Завдання 2.2. Важливе значення для дослідження структури системи має топологічний аналіз структур, який полягає у виявленні так званих ізольованих, висячих і тупикових елементів (або вершин на мові графів). Уточнимо, що ізольована вершина – це вершина, у яку не входить і з якої не виходить жодна дуга; висяча вершина – це вершина, у яку неможливо потрапити із жодної іншої вершини; тупикова вершина – це вершина, з якої неможливо потрапити у інші вершини. Необхідно здійснити топологічний аналіз структур, проілюстрованих на рис. 2.2.1 – 2.2.5.

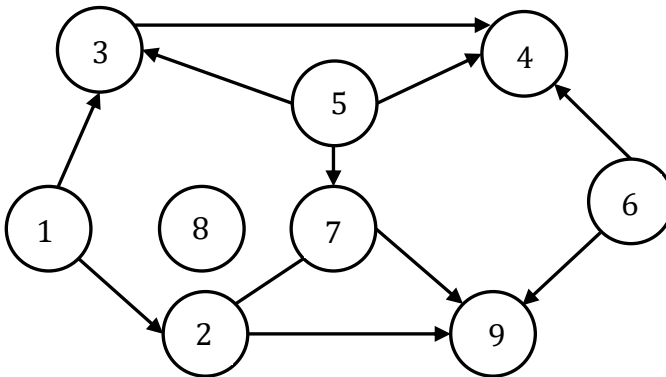
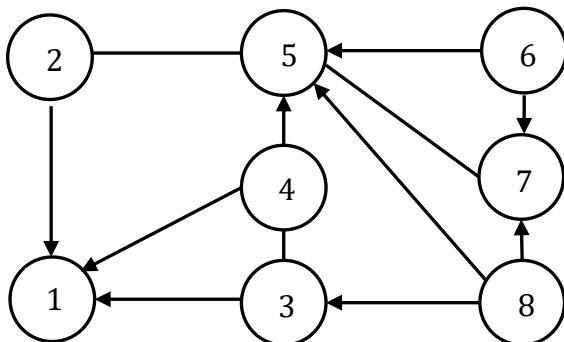
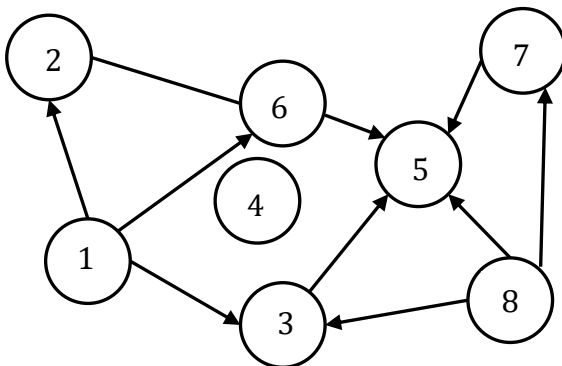


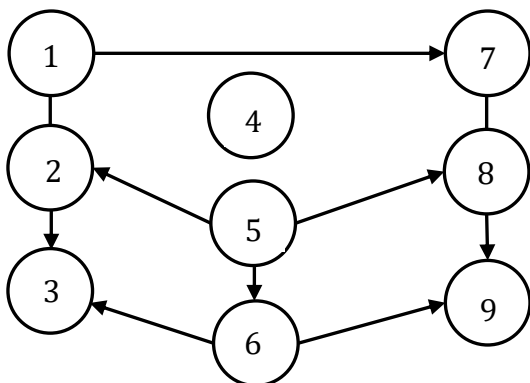
Рис. 2.2.1



Puc. 2.2.2



Puc. 2.2.3



Puc. 2.2.4

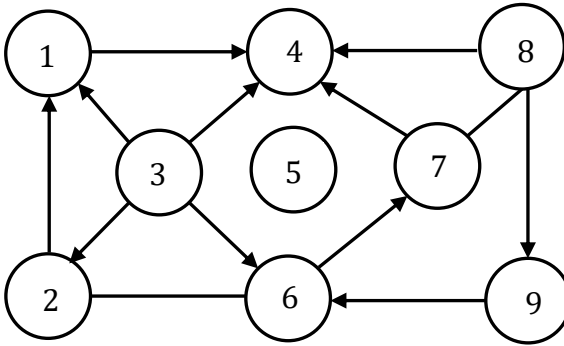


Рис. 2.2.5

◀ Наприклад, зображена на рис. 2.2.1 структура (граф) має одну ізольовану вершину 8, три висячі вершини 1, 5 та 6 і дві тупикові вершини 4 та 9. Якщо йдеться про моделі структур (реальних чи абстрактних), то наявність ізольованих вершин є причиною допущених під час побудови структур помилок, адже система є цілісною, а її елементи – взаємопов’язаними. Наявність висячих і тупикових вершин (елементів) пов’язана із наявністю у системі вхідних і вихідних елементів, за допомогою яких власне кажучи відбувається взаємодія із зовнішнім середовищем. У аналогічний спосіб здійснюється топологічний аналіз інших структур (рис. 2.2.2 – 2.2.5). ▶

ТЕМА 3. ОПТИМАЛЬНІ МАРШРУТИ ТА ШЛЯХИ У МЕРЕЖНИХ СИСТЕМАХ

Завдання 3.1. 1) Описати алгоритм Дейкстри для знаходження найкоротшого маршруту (шляху) у мережі.

2) За допомогою цього алгоритму для зображених на рис. 3.1.1 – 3.1.5 мереж знайти найкоротший маршрут (шлях) між вершинами:

- а) x_1 та x_4 ;
- б) x_1 та x_5 ;
- в) x_9 та x_3 .

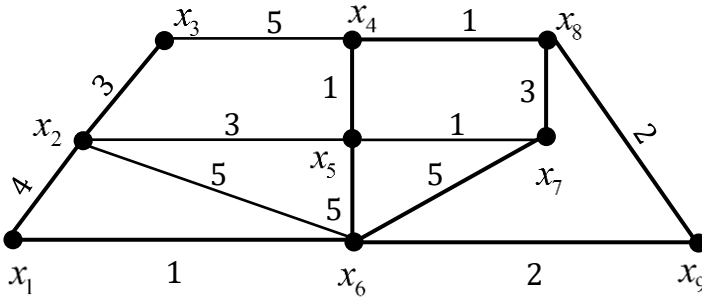
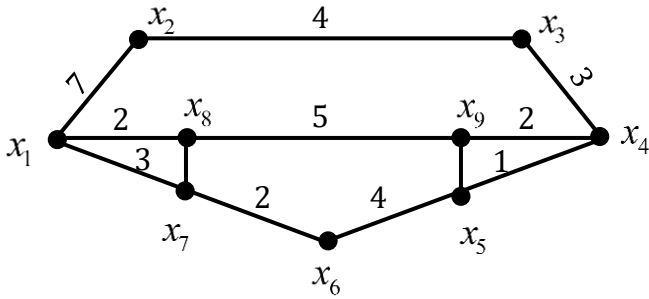
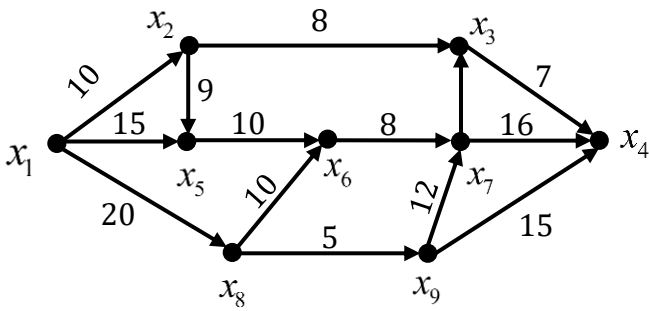


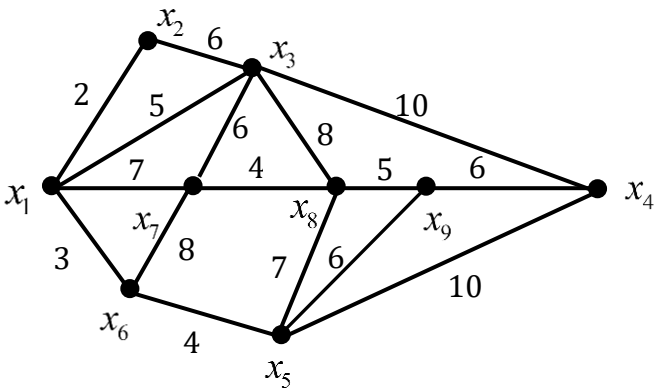
Рис. 3.1.1



Puc. 3.1.2



Puc. 3.1.3



Puc. 3.1.4

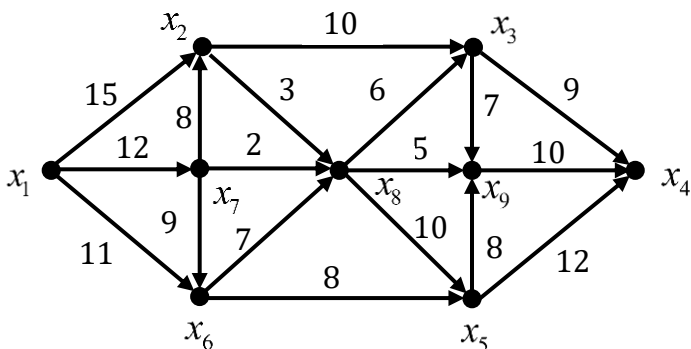


Рис. 3.1.5

◀ Перш за все підкреслимо, що алгоритм Дейкстри належить до класу найпростіших алгоритмів знаходження найкоротшого маршруту (шляху) у зваженому графі [16, 20]. Ідея цього алгоритму полягає у тому, щоб здійснювати рух через так зване найкоротше ребро або ребро з мінімальною вагою.

1) Нехай $\rho(x) \equiv \rho(\hat{x}, x)$ – довжина найкоротшого шляху від деякої початкової вершини \hat{x} до вершини x , що включає лише так звані зафарбовані вершини, а $l(x_i, x_j)$ – вага ребра, інцидентного зазначеним вершинам x_i та x_j . З урахуванням цих позначень конкретизуємо алгоритм пошуку найкоротшого маршруту (шляху) від вершини \hat{x} до вершини \tilde{x} .

1. Усі вершини та ребра не зафарбовані. Для початкової вершини \hat{x} фіксуємо $\rho(\hat{x})=0$, а для усіх інших вершин x ($x \neq \hat{x}$) $\rho(x)=\infty$. Зафарбовуємо

вершину \hat{x} і покладаємо $z = \hat{x}$ (z – остання зафарбована вершина).

2. Для кожної із незафарбованих вершин (насправді лише для тих, з якими вершина z має зв'язок) перерозраховуємо величину $\rho(x)$ за формулою

$$\rho(x) = \min\{\rho(x), \rho(z) + l(z, x)\},$$

де у фігурних дужках значення $\rho(x)$ є попереднім (старим).

У випадку, коли $\rho(x) = \infty$ для усіх незафарбованих вершин, робота алгоритму завершується, оскільки у графі відсутні шляхи з \hat{x} у незафарбовані вершини. У протилежному випадку, зафарбовуємо вершину x з найменшою величиною $\rho(x)$. Якщо таких вершин декілька, вибираємо будь-яку з них. Зафарбовуємо також ребро (чи дугу), що веде у вершину x на тому етапі, на якому вперше було одержано це значення $\rho(x)$. Далі покладаємо $z = x$.

3. Якщо $z = \tilde{x}$ (\tilde{x} – кінцева вершина), пошук найкоротшого маршруту (шляху) завершено. У протилежному випадку повертаємося до кроку 2.

Зауважимо, що після завершення роботи алгоритму одержимо дерево (або орієнтоване дерево, у випадку орієнтованої мережі). Крім того, у випадку, коли необхідно знайти найкоротші шляхи від вершини \hat{x} до усіх інших вершин x , то алгоритм треба продовжити до тих пір, доки не зафарбуються усі вершини графа.

У аналогічний спосіб алгоритм використовується для пошуку найкоротшого шляху у орієнтованих мережах.

2) Користуючись описаним вище алгоритмом, знайдемо, наприклад найкоротший маршрут (шлях) між вершинами x_1 та x_4 (рис. 3.1.1):

1. $\rho(x_1) = 0$, $\rho(x_i) = \infty$, $i = \overline{2,9}$. Зафарбовуємо x_1 та покладаємо $z = x_1$;

2. 2.1. Далі постійно уточнюємо величини $\rho(x)$ для кожної з незафарбованих вершин (для сусідніх користуємося формулою з п. 2 алгоритму) і рухаємося до кінцевої вершини.

$$\rho(x_2) = \min \{ \rho(x_2), \rho(z) + l(z, x_2) \} = \min \{ \infty, 0 + 4 \} = 4;$$

$$\rho(x_6) = \min \{ \rho(x_6), \rho(z) + l(z, x_6) \} = \min \{ \infty, 0 + 1 \} = 1;$$

$$\rho(x_3) = \rho(x_4) = \rho(x_5) = \rho(x_7) = \rho(x_8) = \rho(x_9) = \infty.$$

Зафарбовуємо вершину x_6 ($\rho(x_6)$ є найменшим) та ребро (x_1, x_6) , що веде з вершини x_1 до вершин x_6 , і покладаємо $z = x_6$.

$$2.2. \rho(x_9) = \min \{ \infty, 1 + 2 \} = 3, \quad \rho(x_7) = \min \{ \infty, 1 + 5 \} = 6,$$

$$\rho(x_5) = \min \{ \infty, 1 + 5 \} = 6, \quad \rho(x_2) = \min \{ 4, 1 + 5 \} = 4,$$

$$\rho(x_3) = \rho(x_4) = \rho(x_8) = \infty.$$

Вибираємо вершину x_9 , зафарбовуємо її та ребро (x_6, x_9) , що веде до x_9 з оцінкою $\rho(x_9) = 3$, вперше отриманою на цьому етапі (етапі 2.2), і покладаємо $z = x_9$.

$$2.3. \rho(x_8) = \min\{\infty, 3+2\} = 5,$$

$$\rho(x_7) = 6, \quad \rho(x_5) = 6, \quad \rho(x_2) = 4,$$

$$\rho(x_3) = \rho(x_4) = \infty.$$

Зафарбовуємо вершину x_2 та ребро (x_1, x_2) , оскільки мінімальна оцінка $\rho(x_2) = 4$ вперше була отримана на етапі 2.1, і покладаємо $z = x_2$.

$$2.4. \rho(x_3) = \min\{\infty, 4+3\} = 7,$$

$$\rho(x_5) = \min\{6, 4+3\} = 6,$$

$$\rho(x_7) = 6, \quad \rho(x_8) = 5, \quad \rho(x_4) = \infty.$$

Зафарбовуємо вершину x_8 і ребро (x_9, x_8) , оскільки оцінка $\rho(x_8) = 5$ вперше була отримана на етапі 2.3. Покладаємо $z = x_8$.

$$2.5. \rho(x_4) = \min\{\infty, 5+1\} = 6,$$

$$\rho(x_7) = \min\{6, 5+3\} = 6,$$

$$\rho(x_5) = 6, \quad \rho(x_3) = 7.$$

3. Для вершин x_4, x_5, x_7 оцінка $\rho(x)$ однакова, тому можна вибрати будь-яку з них, але $\tilde{x} = x_4$ є кінцевою вершиною, тому вибираємо її. Зафарбовуємо x_4 і ребро (x_8, x_4) , бо мінімальна оцінка $\rho(x_4) = 6$ отримана саме на цьому етапі (етапі 2.5).

Пошук найкоротшого маршруту (шляху) між вершинами x_1 та x_4 завершено. Цей шлях включає тільки зафарбовані вершини. ►

Завдання 3.2. Використовуючи алгоритм Дейкстри, побудувати дерева найкоротших маршрутів (шляхів) для мереж, зображених на рис. 3.1.1 – 3.1.5.

◀ Знову повернемося до мережі, зображеної на рис. 3.1.1. Щоб побудувати дерево найкоротших шляхів, алгоритм потрібно продовжити до тих пір, поки не будуть зафарбовані усі вершини. Повертаємося до другого пункту алгоритму і реалізуємо цей процес, поклавши $z = x_4$.

$$2.6. \rho(x_5) = \min \{6, 6+1\} = 6,$$

$$\rho(x_3) = \min \{7, 6+5\} = 7,$$

$$\rho(x_7) = 6.$$

Із вершин x_5 , x_7 з однаковою оцінкою $\rho(x)$ вибираємо вершину x_5 . Зафарбовуємо x_5 і ребро, що веде до x_5 з оцінкою $\rho(x_5) = 6$, яка вперше була отримана на етапі 2.2, тобто ребро (x_6, x_5) . Покладаємо $z = x_5$.

$$2.7. \rho(x_7) = \min \{6, 6+1\} = 6,$$

$$\rho(x_3) = 7.$$

Із двох вершин x_7 і x_3 вибираємо x_7 . Зафарбовуємо x_7 і ребро, що веде до x_7 з оцінкою $\rho(x_7) = 6$, яка вперше була отримана на етапі 2.2, тобто ребро (x_6, x_7) . Оскільки вершина x_7 тепер уже не зв'язана безпосередньо із жодною з незафарбованих вершин, то перерозрахунок

чисел $\rho(x)$ потрібно здійснювати для вершин, які залишилися, але з мінімальною оцінкою $\rho(x)$ (якщо їх є декілька, то вибираємо одну із них).

2.8. Залишилася одна (остання) вершина x_3 , яка не має зв'язку із незафарбованими вершинами. Тому вершина x_3 зафарбовується. Крім неї, зафарбовується ребро, що веде до x_3 з оцінкою $\rho(x_3)=7$, яка вперше отримана на етапі 2.4, тобто ребро (x_2, x_3) .

Дерево найкоротших шляхів побудоване. Воно проілюстроване на рис. 3.2.1.

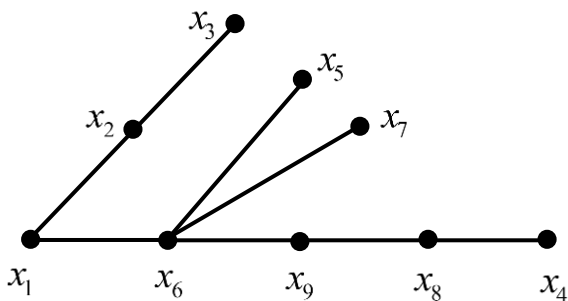


Рис. 3.2.1. Дерево найкоротших маршрутів для мережі, зображеної на рис. 3.1.1



ТЕМА 4. ОПТИМАЛЬНІ ПОТОКИ У МЕРЕЖНИХ СИСТЕМАХ

Завдання 4.1. *Описати та проаналізувати алгоритм Форда-Фалкерсона для пошуку максимального потоку у орієнтованій мережі з одним джерелом і з декількома джерелами.*

◀ Завдання має теоретичний характер, тому для його виконання потрібно опрацювати елементи теорії мереж і оптимальних потоків у мережах [16, 18, 20].

Уточнимо деякі необхідні нам поняття та факти з теорії оптимізації потоків у мережах. Потік – це переміщення об’єктів у просторі з однієї точки у іншу. Початкову точку (початковий пункт) називають *джерелом*, а кінцеву точку (кінцевий пункт) – *стоком*. На практиці *оптимізація потоків*, наприклад максимізація сумарної кількості переміщених від постачальника до споживача одиниць товару, є серйозною проблемою. Якщо (x, z) – деяка дуга мережі (йдеться про орієнтовані мережі), то через $f(x, z)$ позначимо потік на цій дузі (кількість одиниць потоку), $p(x, z)$ – пропускну здатність дуги, $v(x, z)$ – вартість переміщення одиниці потоку. Залежно від цих величин дуги можуть бути *порожніми* ($p(x, z) = 0$ або $v(x, z) \rightarrow \infty$), *збільшувальними* ($f(x, z) < p(x, z)$),

зменшувальними ($f(x, z) > 0$, тобто потік можна зменшити), проміжними ($0 < f(x, z) < p(x, z)$, тобто такі дуги одночасно є як збільшувальними, так і зменшувальними). Надалі множини збільшувальних, зменшувальних і проміжних дуг будемо позначати відповідно так:

$$M^+ = \{(x, z) / f(x, z) < p(x, z)\},$$

$$M^- = \{(x, z) / f(x, z) > 0\},$$

$$M^{+-} = \{(x, z) / 0 < f(x, z) < p(x, z)\}.$$

Крім того, потік по дузі (x, z) можна зменшити на максимальну величину $\mu(x, z) = f(x, z)$ (таке зменшення відбувається тоді, коли у поточному ланцюзі дуга (x, z) є зворотною) та збільшити – на максимальну величину $m(x, z) = p(x, z) - f(x, z)$.

Є різні способи переміщення додаткової кількості одиниць потоку від джерела у стік (використання ланцюга тільки із збільшувальних дуг, тільки із зменшувальних дуг, а третій спосіб – комбінація перших двох). У теорії оптимізації потоків однією з основних задач є *знаходження збільшувального ланцюга від джерела до стоку*. Алгоритм Форда-Фалкерсона призначений для пошуку максимального потоку у мережі. Реалізація цього алгоритму здійснюється поетапно у зазначеній нижче послідовності.

1) Конкретизація початкових значень потоків. У випадку відсутності інформації про ці значення усі

початкові потоки вважаються нульовими для усіх дуг (x, z) ($f(x, z) = 0$).

2) Уточнення усіх дуг щодо їх належності до множин M^+ , M^- чи M^{+-} (дуги із множини M_0 порожніх дуг виключаються з подальшого розгляду).

3) Знаходження у залишковій мережі збільшувального ланцюга від джерела до стоку із урахуванням правил фарбування вершин і дуг, а саме:

а) якщо вершина x зафарбована, то зафарбовується дуга $(x, z) \in M^+$ та вершина z або дуга $(x, z) \in M^-$ та вершина z (зафарбовуються прямі дуги із M^+ і зворотні дуги із M^- , іншого фарбування немає);

б) якщо процес фарбування вершин завершується на вершині стоку, то у мережі є збільшувальний ланцюг (він складається із зафарбованих дуг і вершин), а якщо вершина стоку у процесі фарбування залишається незафарбованою, то збільшувального ланцюга від джерела до стоку у мережі немає (у цій ситуації поточний потік у мережі є оптимальним або необхідно здійснити максимально можливе збільшення потоку і перейти до етапу 2).

Мережу називають *ефективною*, якщо по всіх проміжних дугах від джерела до стоку можна перемістити максимально можливу кількість одиниць потоку. Мережа також може мати не одне, а декілька джерел. Тоді вводиться додаткова вершина, яка додатковими дугами з'єднується з усіма цими джерелами (додатковим дугам призначаються

пропускні здатності, які є величинами, що перевищують сумарну пропускну здатність усіх тих дуг, що виходять із джерел мережі). Алгоритм Форда-Фалкерсона застосовується уже до новоствореної мережі, але після завершення роботи алгоритму додаткова вершина відкидається. При цьому одержаний максимальний потік для новоствореної мережі співпадає із максимальним сумарним потоком з усіх джерел для початкової мережі. ►

Завдання 4.2. Проілюстровані на рис. 4.2.1 – 4.2.5 мережі відображають комунікації між постачальником деякого товару (вершина x_1) та його споживачем (вершина x_5) через мережу посередників (вершини $x_2 - x_4$). Пропускні здатності дуг (одночасно і ваги дуг) зафіксовані поряд з ними. З'ясувати, який максимальний обсяг товару може дійти від постачальника x_1 до споживача x_5 , тобто знайти максимальний потік від x_1 до x_5 .

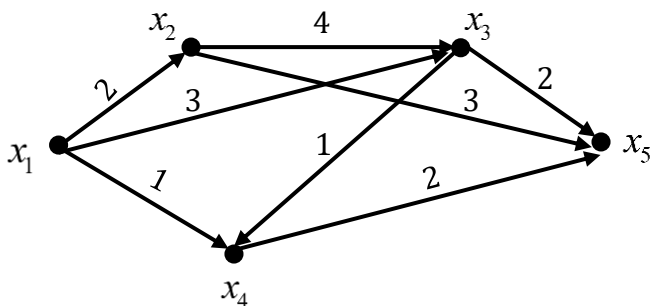
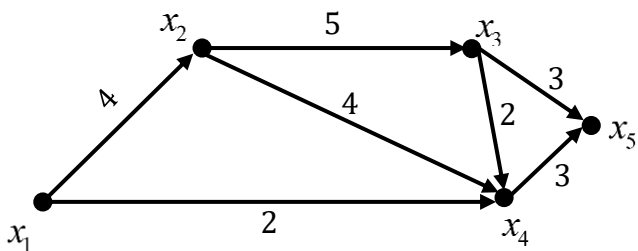
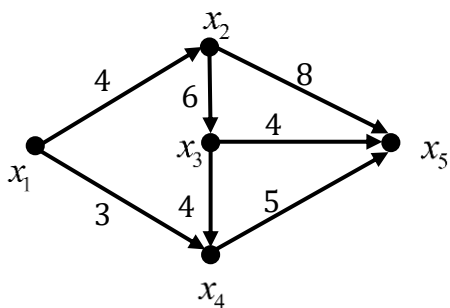


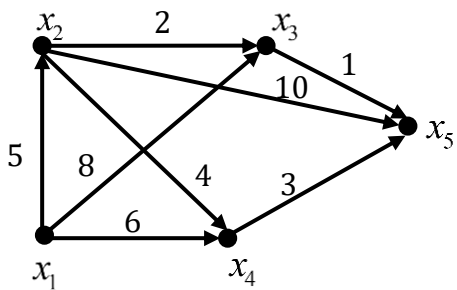
Рис. 4.2.1



Puc. 4.2.2



Puc. 4.2.3



Puc. 4.2.4

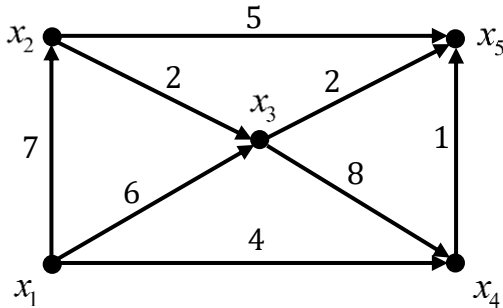


Рис. 4.2.5

◀ Виконання цього завдання (покроково) продемонструємо для мережі, зображеній на рис. 4.2.1. Усі кроки алгоритму конкретизуємо з урахуванням викладеного вище матеріалу з теорії оптимальних потоків у мережних системах (завдання 4.1).

1) На першому кроці усі початкові потоки вважаємо нульовими (вони не задані, тому $f(x, z) = 0$), а вихідну мережу такою, що збігається із залишковою.

2) Конкретизуємо множини M^- та M^+ (зменшувальну та збільшувальну множини дуг), де M^- – множина дуг (x, z) , для яких $f(x, z) \geq 0$ і $\mu(x, z) = f(x, z)$, а M^+ – множина дуг, для яких $f(x, z) < p(x, z)$ і $t(x, z) = p(x, z) - f(x, z)$. При цьому кожен дугу мітимо позначеннями « $f - p, *$ » (f – потік, p – пропускна здатність, $*$ – одна із множин M^- чи M^+ , які надалі

будемо позначати як «-» чи «+» відповідно), що власне проілюстровано на рис. 4.2.6.

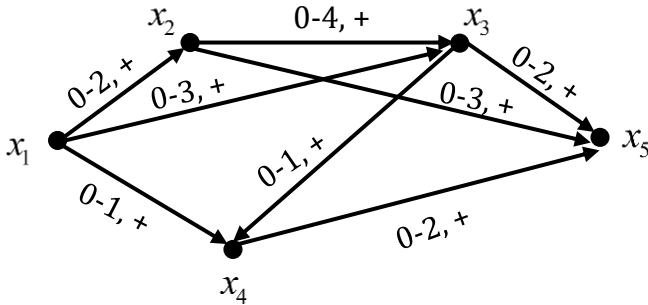


Рис. 4.2.6

3) 3.1. На третьому кроці вибираємо збільшувальний ланцюг. Оскільки він не єдиний, то можна вибрати будь-який, наприклад (x_1, x_4) , (x_4, x_5) . Із того, що

$$\min \{ m(x_1, x_4), m(x_4, x_5) \} = \min \{ 1, 2 \} = 1,$$

впливає, що по цьому ланцюзі можна перемістити одну одиницю потоку.

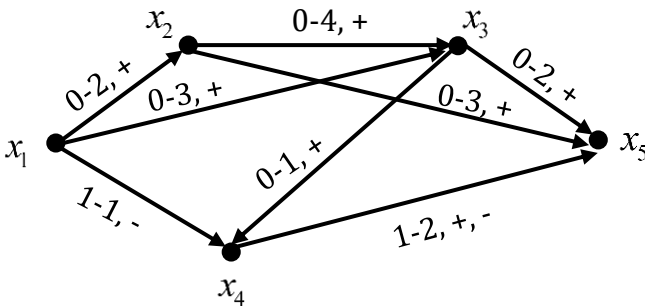


Рис. 4.2.7

Зафіксувавши поточний потік, отримаємо модифіковану мережу (рис. 4.2.7), у якій дуга (x_1, x_4) є зменшувальною, а дуга (x_4, x_5) – як зменшувальною, так і збільшувальною, тобто $(x_1, x_4) \in M^-$, а $(x_4, x_5) \in M^-$ і $(x_4, x_5) \in M^+$ ($\mu(x_4, x_5) = f(x_4, x_5) = 1$, $m(x_4, x_5) = p(x_4, x_5) - f(x_4, x_5) = 2 - 1 = 1$).

Усі інші дуги залишаються незмінними, тобто належать множині M^+ .

Зауважимо, що на цьому (третьому) кроці алгоритму якщо б не знайшлося збільшуваного ланцюга, то процедура його знаходження завершилася б.

3.2. Оскільки збільшувальні ланцюги у мережі є, то на модифікованій мережі (рис. 4.2.7) знову вибираємо збільшувальний ланцюг, наприклад (x_1, x_2) , (x_2, x_3) , (x_3, x_5) . По ньому можна пропустити 2 одиниці потоку, оскільки $\min \{2, 4, 2\} = 2$. Після цього отримаємо модифіковану мережу, зображену на рис. 4.2.8.

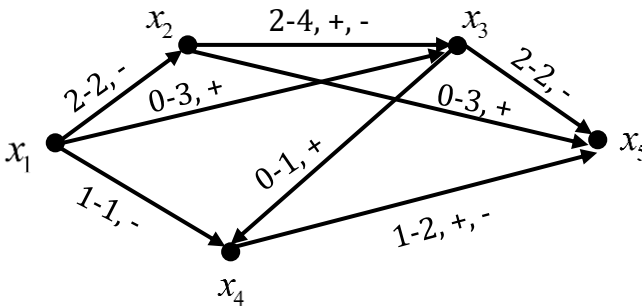


Рис. 4.2.8

3.3. Продовжуємо пошук збільшувальних ланцюгів. Вибираємо збільшувальний ланцюг (x_1, x_3) , (x_3, x_4) , (x_4, x_5) . Враховуючи, що $\min\{3, 1, 1\} = 1$, по цьому ланцюгу збільшимо потік на одну одиницю та отримаємо наступну мережу (рис. 4.2.9).

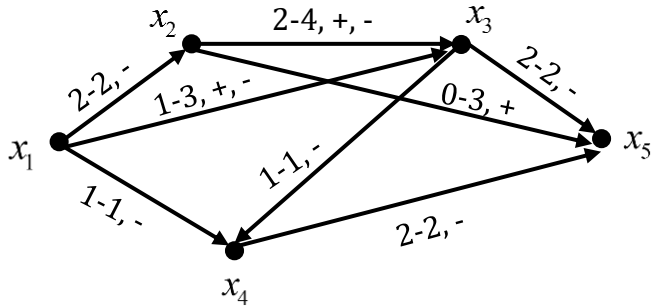


Рис. 4.2.9

3.4. Наступним збільшувальним ланцюгом буде ланцюг (x_1, x_3) , (x_2, x_3) , (x_2, x_5) (ланцюг – це довільна послідовність дуг чи ребер, зокрема послідовність прямих і зворотних дуг. Зважаючи на те, що $\min\{2, 2, 3\} = 2$, дійдемо висновку, що він дозволяє пропустити 2 одиниці потоку.

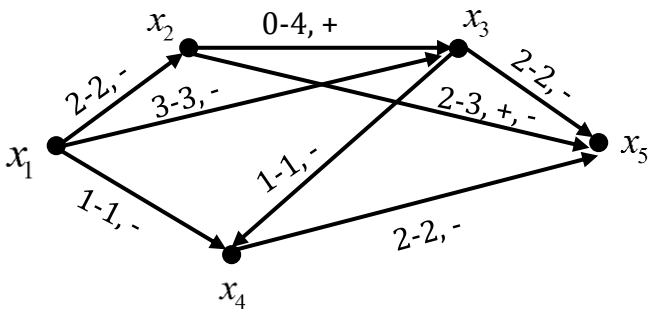


Рис. 4.2.10

Уточнимо, що для поточного ланцюга дуга (x_2, x_3) є зворотною, тому ми зменшили потік по ній на 2 одиниці.

3.5. Як впливає із рис. 4.2.10 усі дуги, що виходять із вершини x_1 (із джерела) насичені (потік дорівнює пропускній здатності), інших збільшувальних ланцюгів у мережі немає, тому *отриманий поточний потік є максимальним*.

Для перевірки правильності виконання завдання необхідно порівняти вихідний потік із джерела x_1 та вхідний потік у стік x_5 . Оскільки сума потоків по вихідних дугах із x_1 ($1+3+2=6$) дорівнює сумі потоків по вхідних дугах у x_5 ($2+2+2=6$), то це підтверджує, що алгоритм знаходження максимального потоку реалізований правильно. Крім того, враховуючи, що сума пропускних здатностей дуг, які виходять із x_1 , дорівнює 6 ($1+3+2=6$) і усі 6 одиниць потоку дійшли до x_5 , можна стверджувати, що мережа проміжних пунктів

(посередників) функціонує ефективно (мережа є ефективною). З іншого боку, сума усіх пропускних здатностей потоку, вхідного у x_5 , дорівнює 7 ($2+3+2=7$). Це означає, що максимально можливий потік у x_5 більший (можливості стоку x_5 більші), тому для його реалізації слід шукати додаткові джерела.

Попутно підкреслимо, що розглянутий приклад мережі показує, що всі одиниці потоку можна відправити із джерела x_1 у стік x_5 (рис. 4.2.10) певними шляхами (однонаправленими прямими дугами), а саме:

1 одиницю дугами (x_1, x_4) , (x_4, x_5) ;

2 одиниці – (x_1, x_3) , (x_3, x_5) ;

1 одиницю – (x_1, x_3) , (x_3, x_4) , (x_4, x_5) ;

2 одиниці – (x_1, x_2) , (x_2, x_5) .

Дуга (x_2, x_3) виявилася незадіяною, що власне кажучи і відображено на рис. 4.2.10 ($f(x_2, x_3) = 0$). ►

**ТЕМА 5. ІЗОМОРФНІ, ГОМЕОМОРФНІ ТА
ГОМОМОРФНІ СИСТЕМИ. ЗВ'ЯЗОК МІЖ СИСТЕМОЮ-
ОРИГІНАЛОМ І СИСТЕМОЮ-МОДЕЛЛЮ**

Завдання 5.1. Обґрунтувати поняття ізоморфних систем як графів і до наведених на рис. 5.1.1 – 5.1.5 систем (або їх структурних моделей, у яких точки x_i ($i = \overline{1,6}$) – це елементи систем, а лінійні ланки (x_i, x_j) – зв'язки між елементами x_i та x_j) побудувати не менше шести ізоморфних і не менше шести неізоморфних систем.

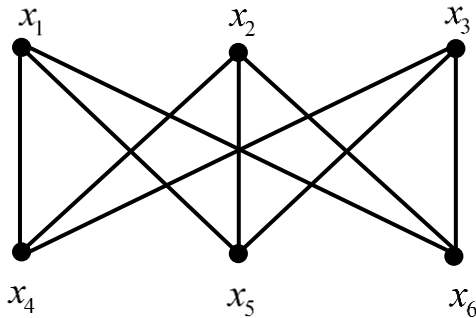
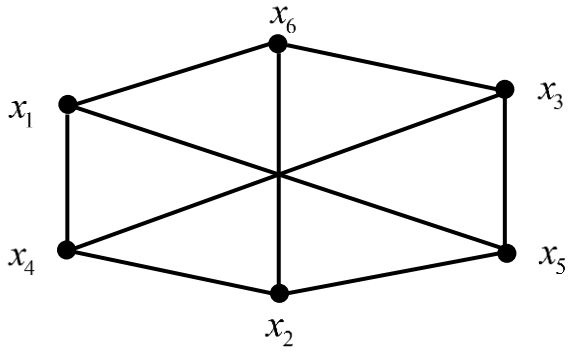
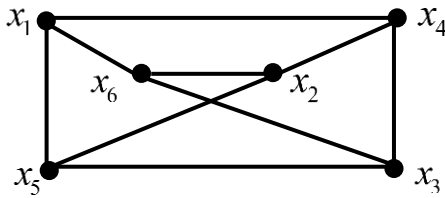


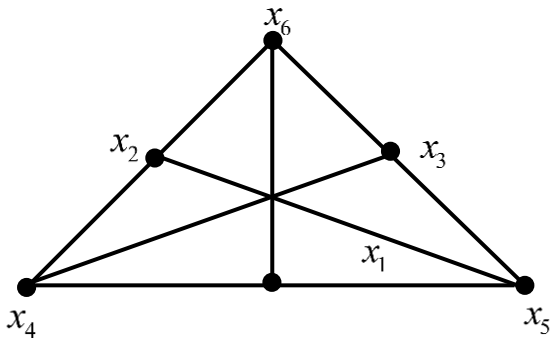
Рис. 5.1.1



Puc. 5.1.2



Puc. 5.1.3



Puc. 5.1.4

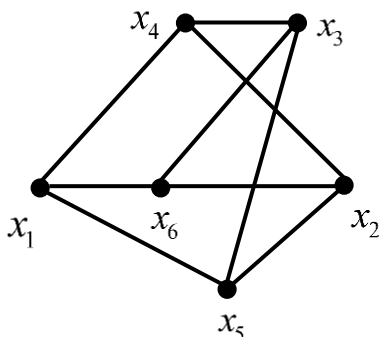


Рис. 5.1.5

◀ *Ізоморфізм графів* – це взаємнооднозначне відображення (бієкція) між множинами вершин графів, яке зберігає суміжність вершин. Інакше кажучи, два графи $H^{(1)} = (X^{(1)}, Y^{(1)})$ і $H^{(2)} = (X^{(2)}, Y^{(2)})$ ізоморфні, якщо відображення $f : X^{(1)} \rightarrow X^{(2)}$ є таким, що для будь-яких $x, z \in X^{(1)}$, $(x, z) \in Y^{(1)}$ відповідно $f(x), f(z) \in X^{(2)}$, $(f(x), f(z)) \in Y^{(2)}$ та навпаки.

По-іншому можна сказати ще так [2]: два графи $H^{(1)} = (X^{(1)}, Y^{(1)})$ і $H^{(2)} = (X^{(2)}, Y^{(2)})$ ізоморфні, якщо має місце така взаємно однозначна відповідність між множинами їх вершин $X^{(1)}$, $X^{(2)}$ та множинами ребер $Y^{(1)}$, $Y^{(2)}$, при якій зберігається відношення інцидентності, тобто для будь-яких вершин $x_1^{(1)}, x_2^{(1)} \in X^{(1)}$ і відповідних їм вершин $x_1^{(2)}, x_2^{(2)} \in X^{(2)}$ ребру

$(x_1^{(1)}, x_2^{(1)}) \in Y^{(1)}$ відповідає ребро $(x_1^{(2)}, x_2^{(2)}) \in Y^{(2)}$ і навпаки.

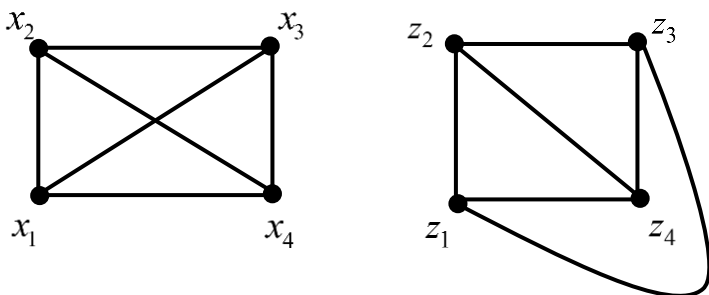


Рис. 5.1.6

Отже, граф називається ізоморфним, якщо він отриманий із деякого іншого графа за допомогою зміни його зображення на площині. Ізоморфні графи – це графи, які не розрізняють.

Приклад ізоморфних графів проілюстрований на рис. 5.1.6, де зліва – це *планарний граф* (тобто граф, який можна так зобразити на площині (тобто так змінити його зображення на площині), щоб його ребра перетиналися лише у вершинах), а справа – *плоский граф* (граф у якого ребра перетинаються лише у вершинах).

Побудова ізоморфних і неізоморфних систем на основі зображених на рис. 5.1.1 – 5.1.5 систем є елементарним завданням, тому зупинятися на цьому немає сенсу. ►

Завдання 5.2. Розкрити зміст гомеоморфізму і для зображених на рис. 5.2.1 – 5.2.5 систем побудувати як мінімум п'ять гомеоморфних систем (x_i – i -тий елемент системи, (x_i, x_j) – зв'язок між елементами x_i та x_j).

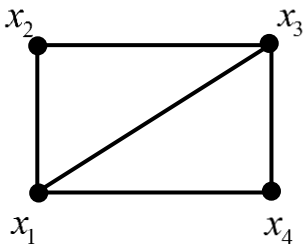


Рис. 5.2.1

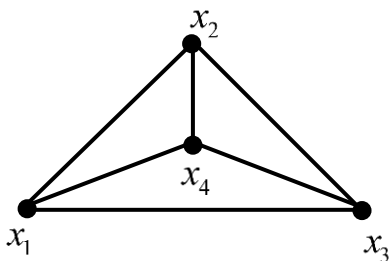
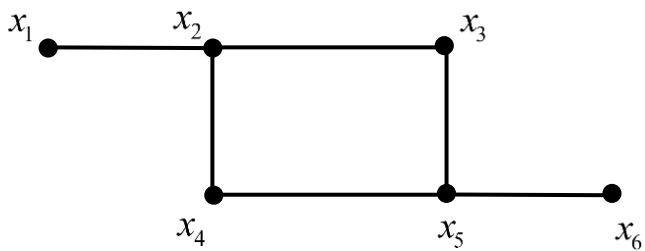
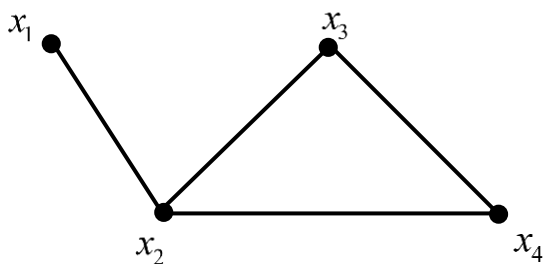


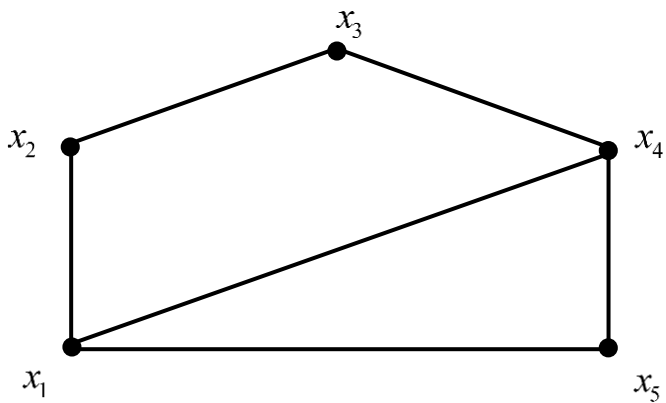
Рис. 5.2.2



Puc. 5.2.3



Puc. 5.2.4



Puc. 5.2.5

◀ 1) На відміну від ізоморфізму або відношення тотожності систем (графів) *гомеоморфізм* – це відношення еквівалентності на множині графів. Інакше кажучи, *гомеоморфні* графи – це еквівалентні графи у тому сенсі, що кожний з них може бути отриманий за допомогою *розбиття деякого графа або один з одного розбиттям його ребер*. Отже, якщо два графи *гомеоморфні*, то для кожного із них існують такі *розбиття*, які є ізоморфними.

Під *розбиттям графа* розуміють граф, отриманий за допомогою операції розбиття ребер, а розбиття ребра – це заміна одного ребра (x, z) на два (x, u) , (u, z) , де u – додана нова вершина. До прикладу, на основі графа (рис. 5.2.1) можна побудувати гомеоморфні графи, зображені на рис. 5.2.6 – 5.2.9 і інші.

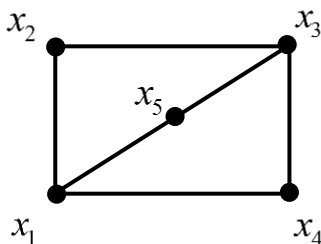


Рис. 5.2.6

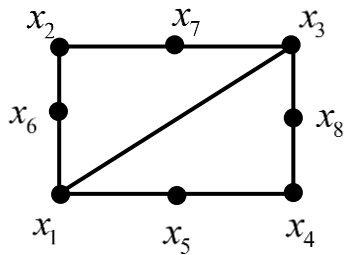


Рис. 5.2.7

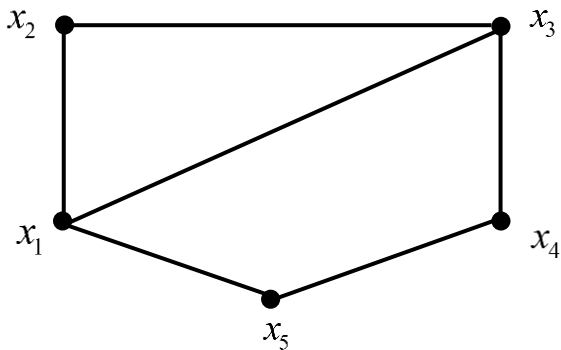


Рис. 5.2.8

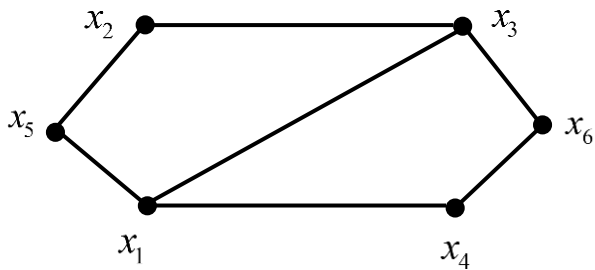


Рис. 5.2.9

Отже, побудова гомеоморфних графів (систем) є доволі простою процедурою, яку потрібно застосувати у цьому завданні до кожного із графів (рис. 5.2.1 – 5.2.5).

Зауважимо також, що для гомеоморфних графів завжди існують розбиття кожного із них, які є ізоморфні між собою, у чому легко переконатися. ►

Завдання 5.3. Обґрунтувати поняття гомоморфізму систем і для систем, зображених графами на рис. 5.3.1 – 5.3.5 за допомогою операції стягування підграфів в окремі вершини побудувати декілька графів (систем), які будуть гомоморфними образами вихідних систем.

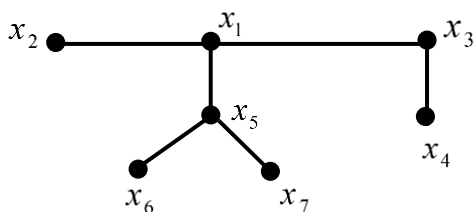


Рис. 5.3.1

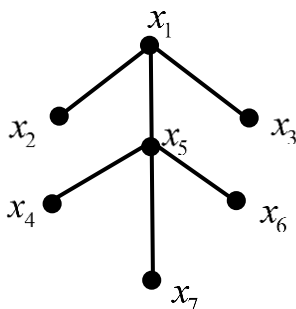


Рис. 5.3.2

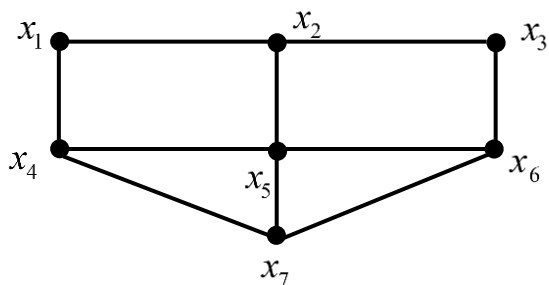


Рис. 5.3.3

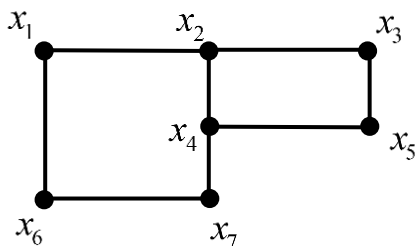


Рис. 5.3.4

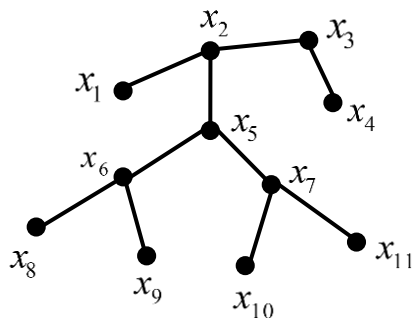


Рис. 5.3.5

◀ Гомоморфізм – це відношення подібності систем у певному структурному чи функціональному аспекті.

Гомоморфізм є узагальненням поняття ізоморфізму у тому розумінні, що однозначна відповідність між системами існує лише в один бік, тобто *дві системи гомоморфні, якщо кожному елементу і кожному зв'язку (відношенню) між елементами однієї системи відповідає один елемент і один зв'язок (одне відношення) іншої системи, але не навпаки.*

У термінах теорії графів *два графи будуть гомоморфними, якщо відповідне стягування підграфів графа з більшою кількістю вершин в окремі вершини графа дає граф, який буде ізоморфним більш простому графу (графу з меншою кількістю вершин).*

Операція *стягування (отождення, злиття)* суміжних вершин x і z полягає у тому, що із графа вилучаються ці вершини та інцидентні з ними ребра, але замість вилучених вершин додається нова вершина u (її можна також позначити через x чи z), яка з'єднується ребрами з тими вершинами, що були суміжними з вилученими. Граф $H^{(1)}$ стягується у граф $H^{(2)}$, якщо $H^{(2)}$ утворюється із $H^{(1)}$ шляхом скінченного числа застосування операції стягування вершин. До прикладу, якщо у графі (рис. 5.3.1) стягувати суміжні вершини x_3 і x_4 у вершину $u \equiv x_3$ (при цьому вилучаються ребра (x_3, x_4) та (x_3, x_1)) і з'єднати новоутворену вершину із вершиною x_1 (вона суміжна з x_3), то отримаємо граф, проілюстрований на рис. 5.3.6, який буде гомоморфним образом вихідного графа (рис. 5.3.1).

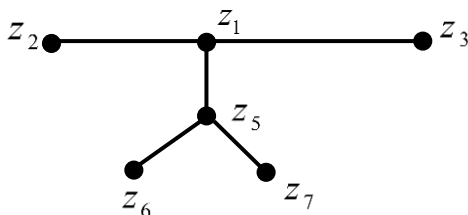


Рис. 5.3.6

Тому графи (рис. 5.3.1, 5.3.6) гомоморфні, оскільки стягування вершин x_3 і x_4 вихідного графа (рис. 5.3.1) дає граф, ізоморфний отриманому графу (рис. 5.3.6). Отже, між системами (графами) $S^{(1)}$ і $S^{(2)}$ (рис. 5.3.1 і 5.3.6) існує гомоморфізм, оскільки відображення $S^{(1)} \rightarrow S^{(2)}$ однозначне ($x_2 \rightarrow z_2, x_1 \rightarrow z_1, x_5 \rightarrow z_5, x_6 \rightarrow z_6, x_7 \rightarrow z_7, x_3, x_4 \rightarrow z_3, (x_2, x_1) \rightarrow (z_2, z_1), (x_1, x_5) \rightarrow (z_1, z_5), (x_5, x_6) \rightarrow (z_5, z_6), (x_5, x_7) \rightarrow (z_5, z_7), (x_1, x_3) \rightarrow (z_1, z_3), (x_3, x_4) \rightarrow z_3$ (ребро (x_3, x_4) вироджується у вершину z_3)), а відображення $S^{(2)} \rightarrow S^{(1)}$ багатозначне, оскільки $z_3 \mapsto x_3, x_4$.

Очевидно, що до кожної із систем (кожного із графів), зображених на рис. 5.3.1 – 5.3.5 можна побудувати різні гомоморфні образи, у чому власне кажучи і полягає зміст цього завдання.

Що стосується співвідношення між системою та її моделлю, то зайвий раз підкреслимо, що зв'язки між системою та її моделлю гомоморфні [2, 15, 17, 19], про

що йтиметься також у наступному завданні. *Однозначна відповідність тут існує лише в один бік – від моделі до системи.* ►

Завдання 5.4. *Розкрити зміст проблеми адекватності моделі досліджуваної системи-оригіналу (досліджуваного об'єкта) за допомогою відношень ізоморфізму та гомоморфізму. Навести приклади ізоморфізму та гомоморфізму.*

◀ Поняття ізоморфізму та гомоморфізму коротко обґрунтовані у попередньому завданні, тому використовуємо їх лише у контексті вивчення зв'язку між системо-оригіналом (об'єктом дослідження) та системою-моделлю (моделлю цього об'єкта).

Під моделлю досліджуваного об'єкта, тобто моделлю системи-оригіналу, у загальному випадку розуміють деяку штучну систему (штучний об'єкт), що за певних умов може бути зручним заміником оригінала, відтворюючи ті чи інші його властивості, які власне цікавлять дослідника. *Отже, модель – це допоміжний засіб чи об'єкт, а моделювання – це дослідження об'єкта-оригіналу за допомогою таких допоміжних засобів. Оскільки ізоморфізм є відношенням тотожності систем у певному структурному чи функціональному плані, то ізоморфна модель (система-модель) є у певному розумінні копією об'єкта дослідження (системи-оригіналу). Ізоморфізм між такими системами є показником високого ступеня адекватності моделі.*

Однак, адекватність моделі не означає, що модель є

повною копією об'єкта. *Адекватність моделі означає тільки те, що вимоги повноти, істинності та точності відтворені нею лише у тій мірі, яка цілком достатня для досягнення мети моделювання.* Відображення об'єкта (системи-оригіналу) у модель є багатозначним, тому модель є гомоморфним образом об'єкта пізнання (від об'єкта до моделі відображення багатозначне, а навпаки – однозначне), причому різні гомоморфні моделі цього об'єкта зазвичай не є ізоморфними (*гомоморфізм – відношення подібності систем у певному структурному чи функціональному плані*).

Отже, гомоморфізм – це відношення подібності між системою-оригіналом і системою-моделлю (деякої подібності структур чи функцій). Зазначимо також, що процес моделювання є ітераційним, оскільки у такий спосіб модель можна постійно коригувати та вдосконалювати.

Щодо прикладів ізоморфізму та гомоморфізму, то достатньо показовими можуть бути наведені нижче приклади, зокрема ізоморфними є:

1) два дійсних лінійних простори (простори, у яких введено операції додавання та множення на дійсне число) L та L' , якщо між ними є взаємно однозначна відповідність типу (зауважимо що, простори L і L' за своєю природою можуть бути абсолютно різними, але у сенсі властивостей введених у них операцій ці простори можуть бути тотожними, тобто їх неможливо розрізнити):

$$(a + b)' = a' + b' \quad (a, b \in L; a', b' \in L'),$$

$$(\alpha a)' = \alpha a' \quad (a \in L, a' \in L', \alpha \in \mathbb{R}),$$

де \mathbb{R} – числова пряма (простір дійсних чисел), «'» – символ, що вказує на те, що простір L' , елементи a' , b' , $(a+b)'$ є відповідно образами простору L , елементів a , b , $(a+b)$.

2) два евклідові простори E і E' (лінійні простори, у яких уведена операція скалярного добутку), якщо між ними є така взаємно однозначна відповідність:

- вони ізоморфні як лінійні простори;
- скалярний добуток зберігається:
 $\langle a, b \rangle = \langle a', b' \rangle \quad (a, b \in E; a', b' \in E');$

3) множина дійсних чисел із бінарним відношенням «менше» ($x_1 < x_2$, тобто перше число є меншим від другого) та множина точок числової осі \mathbb{R} із бінарним відношенням «лівіше» (точка T_1 знаходиться лівіше від точки T_2);

4) множина (простір) векторів-відрізків на площині, які виходять із початку координат, і множина із впорядкованих пар дійсних чисел (кожному вектору-відрізку, що виходить із початку координат ставиться у відповідність впорядкована пара його координат і навпаки).

Можна навести також чимало прикладів гомоморфізму, зокрема якщо для елементів множини \mathbb{R} (множини дійсних чисел) розглянути бінарне відношення «більше», а для елементів (точок) множини \mathbb{R}^2 (множина точок декартової площини) – бінарне

відношення «правіше» (точка B розміщена правіше від точки A , якщо $x_B > x_A$, тобто абсциса точки B більша від абсциси точки A), то відображення $\mathbb{R}^2 \rightarrow \mathbb{R}^1$ буде однозначним (відображення усіх точок площини з координатами (x, y) або точок паралельної до осі ординат прямої, що проходить через точку осі абсцис з координатою x , у точку осі абсцис із координатою x) буде однозначним, а обернене відображення $\mathbb{R} \rightarrow \mathbb{R}^2$ – неоднозначним або сюр'єктивним. Отже, тут між \mathbb{R} і \mathbb{R}^2 існує гомоморфізм.

Зазначимо також, що приклади ізоморфізму та гомоморфізму для систем, зображених графами наведено у завданнях 5.1 і 5.3. ►

Завдання 5.5. *Навести приклади ізоморфних і гомоморфних систем (системи-оригіналу та системи-моделі), які є системами типу «чорний ящик», що характеризуються лише входами та виходами (вхідними та вихідними характеристиками).*

◀ Два «чорних ящики», які характеризуються певними відображеннями входів у виходи, будуть ізоморфні, якщо існує взаємно однозначна відповідність між множинами їх входів і взаємно однозначна відповідність між множинами їх виходів, що будь-який вхід першої системи буде прообразом відповідного входу другої системи тоді і лише тоді, коли відповідний вихід першої системи є прообразом відповідного виходу другої системи і навпаки.

Одним із прикладів ізоморфних систем (оригіналу та моделі) типу «чорні ящики» можуть бути системи $S^{(1)}$ і $S^{(2)}$, які характеризуються відповідно лінійними відображеннями (функціональними залежностями) $y^{(1)} = f^{(1)}(x^{(1)}) = x^{(1)}$, $y^{(2)} = f^{(2)}(x^{(2)}) = x^{(2)}$ ($x^{(1)} \in X^{(1)}$, $x^{(2)} \in X^{(2)}$, $y^{(1)} \in Y^{(1)}$, $y^{(2)} \in Y^{(2)}$; $X^{(1)}, X^{(2)}$ – множини входів, а $Y^{(1)}, Y^{(2)}$ – множини виходів). Очевидно, що коли існує взаємно однозначна відповідність між множинами входів $X^{(1)}$ і $X^{(2)}$, то така ж взаємно однозначна відповідність існує між множинами їх виходів $Y^{(1)}$ і $Y^{(2)}$ та навпаки. Інакше кажучи, у наведеному прикладі обидві системи тотожні, оскільки відповідно множини виходів і самі відображення (перетворювачі сигналів) збігаються.

Зовсім інша ситуація буде спостерігатися, у випадку, коли відповідність між множинами входів взаємно однозначна, а відповідність між множинами виходів багатозначна, наприклад кожному $y^{(1)} \in Y^{(1)}$ відповідає два елементи $y^{(2)}, \tilde{y}^{(2)} \in Y^{(2)}$ (елементарний приклад: $X^{(1)} = X^{(2)} = \{x\}$, $Y^{(1)} = \{y^{(1)} / y^{(1)} = x\}$, $Y^{(2)} = \{y^{(2)} / y^{(2)} = \pm\sqrt{1-x^2}\}$). Тоді системи будуть гомоморфними, оскільки порушується тотожність відношення їх елементів. Аналогічних прикладів можна запропонувати чимало, що власне і вимагається зробити у цьому завданні. ►

ТЕМА 6. КІБЕРНЕТИЧНИЙ ПІДХІД ДО МОДЕЛЮВАННЯ СИСТЕМ ТИПУ «ЧОРНИЙ ЯЩИК»

Завдання 6.1. *Розкрити зміст кібернетичного підходу до моделювання систем, згідно з яким система розглядається як «чорний ящик». Обґрунтувати основні завдання та етапи ідентифікації моделі, побудова якої здійснюється методом «чорного ящика». Навести приклади найбільш відомих моделей цього класу.*

◀ Кібернетичний підхід у моделюванні базується на дослідженні системи за вхідними впливами на неї (входами) та спостережуваними результатами її функціонування (виходами). Інакше кажучи під час використання кібернетичного підходу ми маємо справу із так званою моделлю «чорного ящика», яка відображає лише зв'язок результатів функціонування системи із факторами зовнішнього впливу на неї і зовсім не відображає ті внутрішні процеси, що відбуваються у системі. Графічна ілюстрація цієї моделі зображена на рис. 6.1.1.

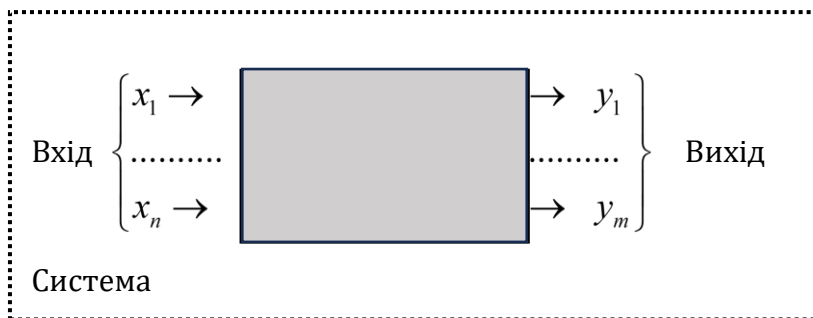


Рис. 6.1.1. Модель «чорного ящика»

Під час дослідження «чорного ящика» ми практично нічого однозначно про структуру «чорного ящика» та його поведінку, тобто поведінку досліджуваної системи, сказати не можемо, оскільки поведінка цієї системи не відрізняється від поведінки ізоморфних з нею систем. Системи можуть бути дуже різними за структурою і призначенням, але зв'язки між вхідними впливами (на рис. 6.1.1 вони позначені змінними x_1, \dots, x_n) і вихідними результатами (на рис. 6.1.1 вони позначені змінними y_1, \dots, y_m) можуть моделюватися однаковими функціональними залежностями. Саме у такому сенсі слід розуміти *принципову особливість «чорного ящика»*.

Інформаційною базою для побудови моделей зв'язку між входами та виходами є дані спостереження за поведінкою системи (детермінованої чи стохастичної). Для отримання цих даних часто проводяться активні експерименти із системою (зміна вхідних компонент приводить до зміни вихідних компонент), які дозволяють глибше вивчити

властивості системи. Очевидно, що процес побудови моделей зазначеного вище зв'язку передбачає наявність експериментальних даних для ідентифікації моделей [19]. *Мета ідентифікації полягає у тому, щоб на основі результатів спостережень за вхідними і вихідними компонентами побудувати модель, яка буде у певному сенсі оптимальною.*

Основними етапами ідентифікації є:

- вибір класу моделі та її описання;
- вибір типів вхідних компонент;
- уточнення критеріїв для обґрунтування подібності між системою та моделлю;
- конкретизація методу ідентифікації та розробка алгоритмів параметризації;
- перевірка побудованої моделі на адекватність.

Розрізняють завдання ідентифікації у широкому і вузькому розумінні, що власне кажучи залежить від обсягу інформації (апріорної) про досліджувану систему.

Якщо структура моделі системи невизначена («чорний ящик»), то йдеться про *ідентифікацію у широкому розумінні*. У цьому випадку клас і структура математичної моделі вибираються на основі аналізу наявних теоретичних знань про закономірності функціонування даної системи чи подібних до неї систем. Інколи таку ідентифікацію ще називають *непараметричною*, адже дослідник не має жодних знань про параметри системи.

Ідентифікація у вузькому розумінні – це оцінювання

параметрів системи, якщо структура моделі відома (так званий «сірий ящик»). Отже, тут використовується *параметрична ідентифікація* математичної моделі.

Загальна схема ідентифікації системи проілюстрована на рис. 6.1.2.

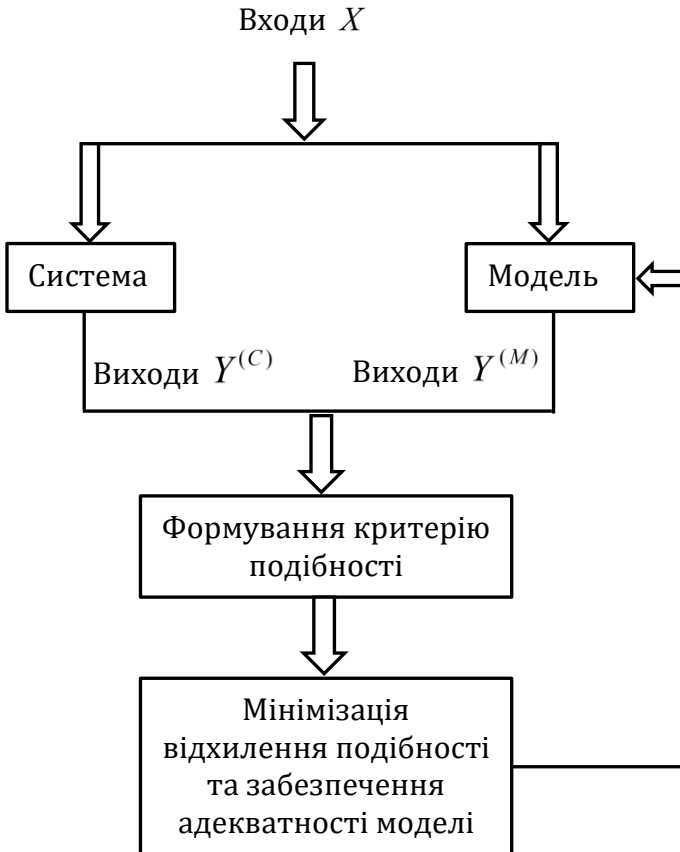


Рис. 6.1.2. Загальна схема ідентифікації системи

На схемі показано, що входи системи (X) та моделі (X) однакові, а виходи ($Y^{(C)}$ та $Y^{(M)}$) у загальному випадку різні, тому для їх порівняння обов'язково треба вибрати критерій подібності («близькості») і мінімізувати відхилення подібності у процесі налагодження моделі та досягнення її адекватності.

Кібернетичний підхід або метод «чорного ящика» (поняття модель і метод «чорного ящика» часто ототожнюються) нерідко використовується для побудови багатьох прикладних моделей. До найбільш відомих моделей, які будуються за допомогою цього підходу, можна віднести моделі типу «вхід–вихід» чи регресійні (у економіці економетричні) моделі для параметризації яких застосовують метод найменших квадратів, а також моделі планування експерименту та ін. ►

Завдання 6.2. *Обґрунтувати поняття регресійної функціональної моделі зв'язку між вхідними та вихідними компонентами системи та основних етапів її побудови. Крім того, побудувати:*

1) *лінійну функцію регресії вихідної компоненти у від вхідної компоненти x (або лінійне рівняння парної регресії y на x), що формалізує зв'язок між виробничими фондами (x) та випуском продукції (y), якщо вибіркові спостережувані значення x на y задано таблицею 6.2.1 ($k \geq 0$).*

Таблиця 6.2.1

Номер спостереження	Випуск продукції, млн. грн. (y)	Виробничі фонди, млн. грн. (x)
1	$16+k$	2
2	$18+k$	3
3	$19+k$	4
4	$20+k$	5
5	$22+k$	6
6	$25+k$	7
7	$26+k$	8
8	$27+k$	10
9	$29+k$	11
10	$30+k$	12

2) лінійну функцію регресії вихідної компоненти y від вхідних компонент x_1 та x_2 (або лінійне рівняння множинної (дві вхідних компоненти) регресії y на векторну величину (x_1, x_2)), що формалізує регресійну залежність випуску продукції (y) від виробничих фондів (x_1) і трудових ресурсів (x_2), вибіркові спостережувані значення яких задано таблицею 6.2.2 ($k \geq 0$).

Таблиця 6.2.2

Номер спостереження	Випуск продукції, млн. грн. (y)	Виробничі фонди, млн. грн. (x_1)	Трудові ресурси, млн. грн. (x_2)
1	$20+k$	10	12
2	$35+k$	15	10
3	$30+k$	20	9
4	$45+k$	25	11
5	$60+k$	40	12
6	$69+k$	37	11
7	$75+k$	43	13
8	$90+k$	35	15
9	$105+k$	38	16
10	$110+k$	55	18

◀ Як було зазначено у попередньому завданні функціональні моделі «вхід-вихід» детермінованих чи стохастичних систем будуються на основі даних спостереження за вхідними та вихідними компонентами (надалі змінними) x_1, \dots, x_n та y_1, \dots, y_m . Для економічних систем типу «чорний ящик» характерними є регресійні функціональні моделі зв'язку між вхідними та вихідними змінними. Ці моделі є важливим інструментом для аналізу та прогнозування економічних явищ. Особливість економічних (і не тільки) систем у тому, що вихідні показники (результати функціонування системи) зазвичай залежать як від суттєвих (надалі x_1, \dots, x_n), так і від несуттєвих (надалі v_1, \dots, v_k) факторів (змінних).

Останніх, тобто несуттєвих факторів, які можна вважати випадковими, зазвичай достатньо велика кількість, тому нехтувати їх сумарним впливом не можна. Це означає, що будь-яка результуюча змінна y (показник, вихідна компонента) є взагалі кажучи деякою функцією f від сукупності суттєвих і несуттєвих змінних, тобто

$$y = f(x_1, \dots, x_n, v_1, \dots, v_k). \quad (6.2.1)$$

Нехай функція (6.2.1) гладка і така, що відображає зв'язок вхідних і вихідної компонент системи в умовах відносно стабільного комплексу зовнішніх впливів на неї. Тоді в околі будь-якого набору значень $(x^*, v^*) = (x_1^*, \dots, x_n^*, v_1^*, \dots, v_k^*)$ вхідних факторів цю функцію можна розкласти у ряд Тейлора [10], окремо виділивши детерміновану та випадкову складові, а саме:

$$y = c_0 + \sum_{i=1}^n c_i \Delta x_i + \sum_{j=1}^k d_j \Delta v_j + o(\Delta x) + o(\Delta v). \quad (6.2.2)$$

У співвідношенні (6.2.2)

$$c_0 = f(x^*, v^*); \quad c_i = \frac{\partial f(x^*, v^*)}{\partial x_i}, \quad i = \overline{1, n};$$

$$d_j = \frac{\partial f(x^*, v^*)}{\partial v_j}, \quad j = \overline{1, k};$$

$$\Delta x_i = x_i - x_i^*, \quad \Delta v_j = v_j - v_j^*;$$

$$\Delta x = \max_{i=1, n} |\Delta x_i|, \quad \Delta v = \max_{j=1, k} |\Delta v_j|,$$

де $o(\Delta)$ – нескінченно мала вищого порядку малості,

ніж Δ , а математичне сподівання $M(v_j) = 0$, $j = \overline{1, k}$, тобто вважається, що усереднений вплив окремого несуттєвого фактора нульовий.

Несуттєві (випадкові) фактори незалежні, їх окремих вплив на результати незначний (практично нульовий), а з огляду на те, що їх достатньо багато, можна скористатися *центральною граничною теоремою (ЦГТ)* теорії ймовірностей, згідно з якою випадкова величина

$$\varepsilon = \sum_{j=1}^k d_j \Delta v_j + o(\Delta v). \quad (6.2.3)$$

має нормальний розподіл із математичним сподіванням $M(\varepsilon) = 0$ і дисперсією $D(\varepsilon) = \sigma^2$. Зрозуміло, що у випадку порушення умов ЦГТ розподіл величини ε може бути іншим, але такий випадок у класичному регресійному аналізі не розглядається. Отже, підсумовуючи вищесказане та співвідношення (6.2.1)–(6.2.3), приходимо до висновку, що лінійний кореляційний зв'язок між вихідною змінною y та вхідними змінними x_1, \dots, x_n можна формалізувати співвідношенням

$$y = \varphi(x_1, \dots, x_n) + \varepsilon, \quad (6.2.4)$$

де x_1, \dots, x_n – детерміновані величини, ε і y – випадкові величини. Якщо вхідна змінна одна, то зв'язок між x і y описують співвідношенням

$$y = g(x) + \varepsilon, \quad (6.2.5)$$

де x – детермінована величина, а ε – випадкова

величина. Що стосується y , то як у (6.2.4), так і у (6.2.5) ця величина випадкова.

Функції $\varphi(x_1, \dots, x_n)$ та $g(x)$ називаються відповідно *функціями множинної* ($n > 1$) *та парної регресії*. Оскільки

$$\varphi(x_1, \dots, x_n) = M(y/(x_1, \dots, x_n)), \quad g(x) = M(y/x),$$

де $M(y/(x_1, \dots, x_n))$ і $M(y/x)$ – умовні математичні сподівання величини y , то останні співвідношення у регресійному аналізі часто називають ще *рівняннями регресії* у на (x_1, \dots, x_n) і у на x .

Отже, поняття регресійної моделі обґрунтовано. У економічних науках цю модель називають економетричною, що пов'язано із специфікою економічних змінних (вхідних і вихідних) і економічного експерименту, повторити який багатократно із збереженням умов і наслідків функціонування економічної системи неможливо. До основних етапів побудови регресійних (економетричних) моделей належать етапи: *підготовчий, специфікації, параметризації, верифікації, застосування*. Детальніше розкриття етапів побудови цих моделей наведено, наприклад у [1].

Зауважимо також, що вище розкрито концептуальний підхід до обґрунтування функціональних моделей лінійного регресійного аналізу. Аналогічний підхід можна використати і до обґрунтування нелінійних функцій регресії, але для цього потрібно скористатись розкладом функції (6.2.1)

у ряд Тейлора більш високого порядку.

1) У цьому підзавданні йдеться про зв'язок (6.2.5), тобто про функцію парної лінійної регресії

$$g(x) = \alpha + \beta x, \quad (6.2.6)$$

де α та β – параметри регресії.

Очевидно, що побудувати функцію (6.2.6), тобто знайти значення її параметрів, можна лише у випадку наявності генеральної сукупності значень змінних x та y . Оскільки у таблиці 6.2.1 задана лише вибірка із $n=10$ значень кожної із змінних, то на основі цих значень можна побудувати не фактичну (теоретичну) функцію регресії (6.2.6) (теоретичне рівняння регресії (6.2.6)), а функцію

$$\bar{g}(x) = a + bx, \quad (6.2.7)$$

де параметри a та b є певними оцінками параметрів α та β .

Функція (6.2.7) (рівняння (6.2.7)) називається вибірковою функцією регресії (вибірковим рівнянням регресії). Тоді для кожного спостереження i ($i = \overline{1, n}$) співвідношення (6.2.5) конкретизується так:

$$y_i = \alpha + \beta x_i + \varepsilon_i = a + bx_i + \delta_i, \quad i = \overline{1, n}, \quad (6.2.8)$$

де δ_i – оцінка випадкової величини (випадкового відхилення) ε у i -тому спостереженні.

З огляду на базове припущення про нормальний розподіл величини ε для параметризації функції (6.2.7) можна використати метод найменших квадратів (МНК), який з урахуванням (6.2.8) зводить задачу

параметризації до безумовної задачі мінімізації (по a та b):

$$F(a, b) = \sum_{i=1}^n \delta_i^2 = \sum_{i=1}^n (y_i - a - bx_i)^2 \rightarrow \min. \quad (6.2.9)$$

Розв'язавши (6.2.9), отримаємо

$$a = \bar{y} - \bar{x}b, \quad (6.2.10)$$

$$b = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - (\bar{x})^2} = \bar{\rho} \frac{\bar{\sigma}_y}{\bar{\sigma}_x}, \quad (6.2.11)$$

де

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad \overline{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2, \quad \overline{y^2} = \frac{1}{n} \sum_{i=1}^n y_i^2,$$

$$\overline{xy} = \frac{1}{n} \sum_{i=1}^n x_i y_i, \quad \bar{\sigma}_x = \sqrt{\overline{x^2} - (\bar{x})^2}, \quad \bar{\sigma}_y = \sqrt{\overline{y^2} - (\bar{y})^2},$$

$$\bar{\rho} = \frac{\bar{\sigma}_{xy}}{\bar{\sigma}_x \cdot \bar{\sigma}_y}, \quad \bar{\sigma}_{xy} = \overline{xy} - \bar{x} \cdot \bar{y}.$$

Якщо використати дані таблиці 6.2.1, наприклад, при $k=0$, то із (6.2.10), (6.2.11) випливає, що $a=13,6439$; $b=1,4053$, тобто

$$\bar{g}(x) = 13,6439 + 1,4053x. \quad (6.2.12)$$

Підзавдання 1) виконане. Однак зауважимо, що повне дослідження функції (рівняння) парної лінійної регресії (6.2.12) потребує також обчислення та аналізу кореляційних характеристик зв'язку між x та y , стандартних помилок регресії та коефіцієнтів регресії, перевірки на деякому рівні значущості певних гіпотез щодо коефіцієнтів регресії та кореляції, побудову інтервальних оцінок для коефіцієнтів і лінії регресії,

перевірки значущості рівняння регресії у цілому тощо. Для ґрунтовного ознайомлення з усіма етапами цього дослідження можна використати відповідний матеріал посібника [1].

2) У сенсі принципу побудови регресійних моделей виконання цього підзавдання здійснюється аналогічно до попереднього. Нижче число незалежних вхідних змінних позначимо через m , а число спостережень, як і раніше, – через n . Тоді згідно із співвідношенням (6.2.4)

$$\varphi(x_1, \dots, x_m) = \alpha + \sum_{j=1}^m \beta_j x_j \quad (6.2.13)$$

та

$$\bar{\varphi}(x_1, \dots, x_m) = a + \sum_{j=1}^m b_j x_j \quad (6.2.14)$$

відповідно *теоретична та вибіркова функції множинної (m вхідних змінних) регресії (теоретичне та вибіркове рівняння множинної регресії)*. Параметри a , b_j ($j = \overline{1, m}$) функції (6.2.14) є оцінками параметрів α , β_j ($j = \overline{1, m}$) функції (6.2.13). У таблиці 6.2.2 задана вибірка сукупність значень змінних x_1 , x_2 , y , тому можна побудувати лише вибіркову функцію регресії. Для усіх спостережуваних значень змінних y і x_1, \dots, x_m виконуються аналогічно до (6.2.8) співвідношення

$$\tilde{y} = \tilde{X} \tilde{\beta} + \varepsilon = \tilde{X} \tilde{b} + \delta, \quad (6.2.15)$$

де

$$\tilde{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \tilde{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{m1} \\ 1 & x_{12} & \cdots & x_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_{1n} & \cdots & x_{mn} \end{pmatrix},$$

$$\tilde{b} = \begin{pmatrix} a \\ b_1 \\ \cdots \\ b_m \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdots \\ \varepsilon_n \end{pmatrix}, \quad \delta = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \cdots \\ \delta_n \end{pmatrix}, \quad \tilde{\beta} = \begin{pmatrix} \alpha \\ \beta_1 \\ \cdots \\ \beta_m \end{pmatrix},$$

причому \tilde{y} і \tilde{X} – вектор і матриця заданих спостережуваних значень (наприклад, у таблиці 6.2.2).

З огляду на (6.2.15) параметризація функції (6.2.14) зводиться до задачі безумовної мінімізації (по a та b_1, \dots, b_m)

$$\Phi(a, b_1, \dots, b_m) = \sum_{i=1}^n \delta_i^2 =$$

$$= \sum_{i=1}^n \left(y_i - a - \sum_{j=1}^m b_j x_j \right)^2 \rightarrow \min. \quad (6.2.16)$$

У векторно-матричній формі розв'язок задачі (6.2.16) має вигляд

$$\tilde{b} = \left(\tilde{X}^T \tilde{X} \right)^{-1} \tilde{X}^T \tilde{y}, \quad (6.2.17)$$

де T – транспонування.

До прикладу при $k=0$ для даних таблиці 6.2.2 шуканий вектор (6.2.17) конкретизується так:

$$\tilde{b} = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T \tilde{y} = \begin{pmatrix} a \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} -44,3653 \\ 1,1896 \\ 5,5462 \end{pmatrix}.$$

Знаходженням вектора \tilde{b} процес побудови вибіркової функції (чи вибіркового рівняння) завершується, але повне дослідження цієї функції (цього рівняння), пов'язане з перевіркою на адекватність, включає ще ряд етапів, обґрунтування яких можна знайти у [1]. ►

ТЕМА 7. ВІДТВОРЕННЯ (ІМІТАЦІЯ) ПОВЕДІНКИ СИСТЕМ ЯК ОКРЕМИЙ ПІДХІД ДО ЇХ МОДЕЛЮВАННЯ: ДИФЕРЕНЦІАЛЬНІ МОДЕЛІ НЕПЕРЕРВНИХ ДИНАМІЧНИХ СИСТЕМ

Завдання 7.1. Частіше всього кібернетичний підхід до моделювання систем використовується для побудови функціональних моделей зв'язку між входами і виходами (вхідними та вихідними змінними) на основі даних спостереження за ними. Якщо є деяка додаткова інформація про цей зв'язок (у більш загальному контексті – деяка додаткова інформація про процеси, що відбуваються у системі), то такі моделі можна побудувати без спостережуваних даних про входи та виходи, опираючись на цю інформацію (тобто на відповідні знання про вимоги до моделі зв'язку між входами та виходами) та широкі можливості математичного інструментарію, зокрема інструментарію диференціальних рівнянь, який використовується не тільки для моделювання динамічних систем. Прикладами таких ситуацій можуть бути ті, які наведені нижче у підзавданнях 1) та 2).

1) Обсяг випуску продукції y є функцією від обсягу витрат (ресурсу) x ($x, y > 0$). Відомо, що коефіцієнт еластичності випуску відносно витрат (k, m – додатні сталі) такий:

$$а) e_y^{(x)} = k; \quad б) e_y^{(x)} = mx;$$

$$в) e_y^{(x)} = mx^2; \quad з) e_y^{(x)} = m\sqrt{x};$$

$$д) e_y^{(x)} = m \ln x; \quad е) e_y^{(x)} = m \log_2 x.$$

Необхідно визначити функцію випуску $y = y(x)$, користуючись означенням коефіцієнта еластичності функції, та обсяг випуску при конкретно заданих значеннях x^* , $y^* = y(x^*)$, k , m і x (вибір цих значень здійснити самостійно).

2) Для виробничої системи на макрорівні входами є виробничі фонди (капітал) K і трудові ресурси L , а виходом – валовий внутрішній продукт або кінцевий продукт Y (усі величини додатні). Потрібно побудувати модель макровиробничої функції $Y = Y(K, L)$, якщо ($e_Y^{(K)}$ і $e_Y^{(L)}$ – коефіцієнти еластичності функції відносно K і L ; Y^* , α та β – деякі додатні величини):

$$а) e_Y^{(K)} = \alpha, \quad e_Y^{(L)} = \beta, \quad Y(1,1) = Y^*;$$

$$б) e_Y^{(K)} = \frac{1}{2}, \quad e_Y^{(L)} = \frac{1}{2}, \quad Y(4,4) = 8;$$

$$в) e_Y^{(K)} + e_Y^{(L)} = \frac{5}{6}, \quad e_Y^{(K)} e_Y^{(L)} = \frac{1}{6}, \quad Y(4,8) = 40;$$

$$г) e_Y^{(K)} + e_Y^{(L)} = \frac{7}{12}, \quad e_Y^{(L)} - e_Y^{(K)} = \frac{1}{12}, \quad Y(16,8) = 20;$$

$$д) e_Y^{(K)} = e_Y^{(L)}, \quad \left(e_Y^{(K)}\right)^2 - 3e_Y^{(K)} + \frac{8}{9} = 0, \quad Y(27,27) = 1;$$

$$е) e_Y^{(K)} + e_Y^{(L)} = 1, \quad \left(e_Y^{(K)} e_Y^{(L)}\right)^{\frac{1}{2}} = \frac{\sqrt{2}}{3}, \quad Y(8,27) = 36.$$

◀ 1) Виконаємо це підзавдання для випадку а).

Згідно з означенням коефіцієнта еластичності функції $y(x)$ у точці x [3, 6]

$$k = \frac{x}{y(x)} y'(x)$$

або

$$\frac{dy}{dx} = k \frac{y}{x} \left(\frac{dy}{y} = k \frac{dx}{x} \right). \quad (7.1.1)$$

Проінтегрувавши рівняння (7.1.1) (рівняння з відокремлюваними змінними), отримаємо

$$y = Cx^k, \quad C \neq 0, \quad (7.1.2)$$

де C – довільна стала.

Функція (7.1.2) – загальний розв'язок диференціального рівняння (7.1.1). Оскільки у точці x_0 відповідно $y_0 = y(x_0) = Cx_0^k$, тобто $C = \frac{y_0}{x_0^k}$, то із (7.1.2)

впливає, що

$$y = y_0 \left(\frac{x}{x_0} \right)^k. \quad (7.1.3)$$

Отже, функціональна залежність випуску продукції від витрат ресурсу при заданих x_0 , y_0 і k має вигляд (7.1.3). Якщо задано конкретне значення витрат x , то знайдемо також і відповідне конкретне значення y . Попутно підкреслимо, що коефіцієнт еластичності показує граничний процентний приріст випуску продукції на 1% нескінченно малого приросту витрат (ресурсу).

2) Для демонстрації виконання цього підзавдання

також розглянемо випадок а). Оскільки частинні коефіцієнти еластичності випуску продукції відносно капіталу та трудових ресурсів знаходяться за формулами

$$e_Y^{(K)} = \frac{K}{Y} \cdot \frac{\partial Y}{\partial K} = \alpha, \quad e_Y^{(L)} = \frac{L}{Y} \cdot \frac{\partial Y}{\partial L} = \beta,$$

то

$$\frac{\partial Y}{Y} = \alpha \frac{\partial K}{K}, \quad \frac{\partial Y}{Y} = \beta \frac{\partial L}{L}$$

або

$$\ln Y = \ln [A(L)K^\alpha], \quad \ln Y = \ln [B(K)L^\beta]$$

чи

$$Y = A(L)K^\alpha, \quad Y = B(K)L^\beta. \quad (7.1.4)$$

Із співвідношень (7.1.4) та умови $Y(1,1) = Y^*$ випливає, що

$$Y^* = A(1) = B(1) \text{ або } A(L) = Y^* L^\beta, \quad B(K) = Y^* K^\alpha \text{ тому}$$

$$Y = Y^* K^\alpha L^\beta. \quad (7.1.5)$$

Легко перевірити, що усі умови а) для функції (7.1.5) виконуються.

Функція (7.1.5) є степеневою макровиробничою функцією або так званою функцією Кобба-Дугласа [3, 6].

Виконання підзавдання 2) для випадків б)–е) з огляду на проаналізований випадок а) є достатньо тривіальною задачею. ►

Завдання 7.2. Диференціальні моделі є основним математичним інструментарієм для дослідження динаміки систем (об'єктів) [9], перехідні процеси у яких

неперервні у часі (час присутній явно чи опосередковано, змінна часу неперервна). Динаміка системи – це процес зміни її станів у часі. Кожний стан динамічної системи пов'язаний із часом і його можна описати сукупністю деяких показників. У цьому завданні йдеться саме про таку систему.

Отже, нехай макроекономічна система (наприклад, економіка країни) характеризується сукупністю залежних від часу показників $\{K, L, Y, I, C\}$ (K, L – входи, Y – вихід системи (I, C – його складники), причому K – капітал, L – трудові ресурси, Y – кінцевий продукт, I – валові інвестиції, C – невиробниче споживання). Крім того, нехай мають місце такі базові припущення:

- у початковий момент часу t_0 задані значення

$$K(t_0) = K^{(0)}, \quad L(t_0) = L^{(0)}; \quad (7.2.1)$$

- зміна у часі трудових ресурсів відбувається згідно із формулою

$$L = L^{(0)} e^{\eta(t-t_0)}, \quad (7.2.2)$$

де η – річний темп приросту робочої сили (стала величина);

- витрати на амортизацію капіталу

$$A = \mu K, \quad (7.2.3)$$

де $\mu > 0$ – заданий (сталий) темп амортизації;

- зв'язок між Y і K, L описується неокласичною функцією випуску [2, 3]

$$Y = F(K, L), \quad (7.2.4)$$

яка є лінійно однорідною;

- основними складниками кінцевого випуску є

невиробниче споживання та валові інвестиції, тобто

$$Y = C + I; \quad (7.2.5)$$

• валові інвестиції складаються із чистих інвестицій (J) та амортизаційних відрахувань:

$$I = J + A; \quad (7.2.6)$$

• чисті інвестиції збігаються із приростом капіталу ($J = \Delta K$), який у кожний момент час можна вважати граничним:

$$\frac{dK}{dt} = J. \quad (7.2.7)$$

• норма накопичення капіталу s задана ($0 \leq s \leq 1$), тобто задана величина (у загальному випадку залежна від часу)

$$s = \frac{I}{Y}. \quad (7.2.8)$$

Необхідно:

1) на основі цих припущень побудувати (формалізувати) модель економічного зростання у абсолютних і у відносних показниках (K, L, Y, C, I – абсолютні показники, $k = K/L$, $c = C/L$, $y = Y/L$, $i = I/L$ – відносні показники, тобто відповідні показники у розрахунку на одиницю праці);

2) конкретизувати модель економічного зростання у відносних показниках, якщо

$$F(K, L) = bK^\alpha L^{1-\alpha}, \quad (7.2.9)$$

де $0 < \alpha < 1$ (b, α – сталі величини), та знайти її розв'язок (траєкторію економічного зростання);

3) уточнити траєкторію економічного зростання

$k = k(t)$ у випадку, коли $t_0 = 0$, $k^{(0)} = n+1/n$, $\eta = 0,4+1/n$, $\mu = 0,2+1/n$, $s = 0,7+1/n$, $\alpha = 0,5+1/n$ ($n > 3$), та побудувати графічну ілюстрацію траєкторії.

◀ **1)** Умови (7.2.1)–(7.2.8) є основою для формалізації моделі економічного зростання. Із (7.2.8) та (7.2.5) отримуємо співвідношення

$$I = sY, \quad C = (1-s)Y, \quad (7.2.10)$$

а із (7.2.5) –(7.2.7) і (7.2.3), (7.2.4), (7.2.10) – співвідношення

$$\begin{aligned} \frac{dK}{dt} &= I - A = Y - C - A = F(K, L) - \\ &-(1-s)F(K, L) - \mu K = sF(K, L) - \mu K. \end{aligned} \quad (7.2.11)$$

Тепер, об'єднавши (7.2.4), (7.2.10), (7.2.11), (7.2.1), (7.2.2), отримаємо один із варіантів моделі економічного зростання [3, 4] у абсолютних показниках ($t \geq t_0$, динаміка системи досліджується на деякому проміжку часу):

$$\begin{cases} Y(t) = F(K(t), L(t)), \\ I(t) = sY(t), \\ C(t) = (1-s)Y(t), \quad L(t_0) = L^{(0)} e^{\eta(t-t_0)}, \\ \frac{dK(t)}{dt} = sF(K(t), L(t)) - \mu K(t), \\ K(t_0) = K^{(0)}, \\ 0 \leq s(t) \leq 1. \end{cases} \quad (7.2.12)$$

Ввівши у розгляд відносні показники (змінну t надалі опустимо, оскільки вона буде задана неявно) $k = K/L$ (фондоозброєність одиниці праці), $c = C/L$

(невиробниче споживання на одиницю праці), $y = Y/L$ (продуктивність праці), $i = I/L$ (питомі валові інвестиції), прийдемо до співвідношень

$$\begin{aligned} y &= \frac{Y}{L} = \frac{F(K, L)}{L} = \frac{F\left(L \cdot \frac{K}{L}, L \cdot 1\right)}{L} = \\ &= \frac{LF\left(\frac{K}{L}, 1\right)}{L} = f(k), \end{aligned} \quad (7.2.13)$$

$$y = \frac{Y}{L} = \frac{I}{L} + \frac{C}{L} = i + c, \quad (7.2.14)$$

$$\begin{aligned} \dot{k} &= \frac{dk}{dt} = \frac{d}{dt}\left(\frac{K}{L}\right) = \frac{\dot{K} \cdot L - K \cdot \dot{L}}{L^2} = \\ &= \frac{[sF(K, L) - \mu K] - K \cdot \eta L}{L \cdot L} = s \frac{F(K, L)}{L} - \\ &- \mu \frac{K}{L} - \eta \frac{K}{L} = sf(k) - (\mu + \eta)k, \end{aligned} \quad (7.2.15)$$

$$\frac{K^{(0)}}{L^{(0)}} = \frac{K(t_0)}{L(t_0)} = k(t_0) = k^{(0)}. \quad (7.2.16)$$

Із урахуванням (7.2.13)–(7.2.16) легко перейти від моделі (7.2.12) до моделі у відносних показниках, тобто моделі

$$\begin{cases} y = f(k), \\ i = sy, \quad c = (1-s)y, \\ \dot{k} = sf(k) - (\mu + \eta)k, \\ k(t_0) = k^{(0)}, \\ 0 \leq s \leq 1. \end{cases} \quad (7.2.17)$$

Очевидно, що співвідношення (7.2.15), (7.2.16) є основними у моделі (7.2.17). При цьому рівняння (7.2.15) називається основним диференціальним рівнянням неокласичної теорії економічного зростання [3, 4]. Розв'язавши задачу Коші (7.2.15), (7.2.16) при заданому s (у моделях оптимального економічного зростання ця величина є управлінням), отримуємо k , а потім усі інші відносні та абсолютні показники стану досліджуваної макроекономічної системи.

2) Конкретизація моделі (7.2.17) фактично зводиться до конкретизації рівняння (7.2.15), а знаходження траєкторії економічного зростання – до розв'язування задачі (7.2.15), (7.2.16). Виконаємо це підзавдання для випадку функції (7.2.9) при $b = 1$ та $s = const$. Тоді

$$f(k) = F(k, 1) = k^\alpha. \quad (7.2.18)$$

З огляду на (7.2.18) диференціальне рівняння (7.2.15) деталізується так:

$$\dot{k} = sk^\alpha - (\mu + \eta)k. \quad (7.2.19)$$

Рівняння (7.2.19) – це так зване рівняння Бернуллі [9]. Воно нелінійне (шукана функція $k = k(t)$ входить до рівняння нелінійно), але його можна звести до лінійного рівняння першого порядку. Справді, поділивши обидві частини (7.2.19) на k^α ($\alpha \neq 0$, $\alpha \neq 1$, $k \neq 0$), отримуємо рівняння

$$k^{-\alpha} \dot{k} = s - (\mu + \eta)k^{1-\alpha}. \quad (7.2.20)$$

Потім, зробивши заміну $z = k^{1-\alpha}$, від (7.2.20) прийдемо до рівняння

$$\dot{z} = -(\mu + \eta)(1 - \alpha)z + s(1 - \alpha). \quad (7.2.21)$$

Лінійне рівняння першого порядку (7.2.21) неоднорідне, тому його загальний розв'язок z_{3n} дорівнює сумі загального розв'язку z_{30} відповідного йому однорідного рівняння та частинного розв'язку $z_{чн}$ неоднорідного рівняння. Очевидно, що

$$z_{30} = C_1 e^{-(\mu + \eta)(1 - \alpha)t}, \quad (7.2.22)$$

де C_1 – довільна стала.

Враховуючи (7.2.22), для знаходження $z_{чн}$ використовуємо метод варіації довільної сталої, тобто

$$z_{чн} = C_1(t) e^{-(\mu + \eta)(1 - \alpha)t}. \quad (7.2.23)$$

Після підстановки (7.2.23) у (7.2.21) одержимо рівняння

$$dC_1 = s(1 - \alpha) e^{(\mu + \eta)(1 - \alpha)t} dt,$$

проінтегрувавши яке, знайдемо

$$C_1 = \frac{s}{\mu + \eta} e^{(\mu + \eta)(1 - \alpha)t} + C_2, \quad (7.2.24)$$

де C_2 – довільна стала.

Взявши до уваги (7.2.24), конкретизуємо розв'язок (7.2.23):

$$z_{чн} = C_2 e^{-(\mu + \eta)(1 - \alpha)t} + \frac{s}{\mu + \eta}. \quad (7.2.25)$$

Після цього знаходимо

$$\begin{aligned} z_{3n} &= z_{30} + z_{чн} = \\ &= \tilde{C} e^{-(\mu + \eta)(1 - \alpha)t} + \frac{s}{\mu + \eta}, \end{aligned} \quad (7.2.26)$$

де $\tilde{C} = C_1 + C_2$ – довільна стала.

Повернувшись до змінної k і врахувавши початкову умову (7.2.16), остаточно перепишемо (7.2.26), тобто розв'язок рівняння (7.2.19), у вигляді:

$$k = \left(\left[\left(k^{(0)} \right)^{1-\alpha} - \frac{s}{\mu + \eta} \right] e^{(\mu + \eta)(1-\alpha)(t_0 - t)} + \frac{s}{\mu + \eta} \right)^{\frac{1}{1-\alpha}}. \quad (7.2.27)$$

Отже, траєкторія економічного зростання, тобто траєкторія (7.2.27), побудована.

3) Нехай наприклад, $t_0 = 0$; $k^{(0)} = 4$; $\eta = 0,4$; $\mu = 0,2$; $s = 0,6$; $\alpha = 0,5$. Тоді формула (7.2.27) має вигляд:

$$k = \left(e^{-0,3t} + 1 \right)^2. \quad (7.2.28)$$

Очевидно, що функція (7.2.28) при $t > 0$ спадає ($\dot{k}(t) < 0$) та опукла вниз ($\ddot{k}(t) > 0$), а $k(0) = 4$. Побудова її графіка є елементарною задачею, зупинятися на якій немає сенсу. Зауважимо лише, що рівняння (7.2.19) має два стаціонарні (рівноважні) розв'язки, а саме: $k^{(1)} = 0$,

$k^{(2)} = \left(\frac{\mu + \eta}{s} \right)^{\frac{1}{\alpha-1}}$. Нескладно встановити [3], що $k^{(2)}$ є

асимптотично стійкою точкою рівноваги, причому $k(t) \rightarrow k^{(2)}$ при $t \rightarrow \infty$. Оскільки у нашому випадку (для вибраних даних) $k^{(2)} = 1$, то монотонне спадання траєкторії (7.2.28) є закономірним ($k^{(0)} = 4 > k^{(2)} = 1$). Якщо взяти $k^{(0)} < k^{(2)} = 1$, то траєкторія (7.2.28) буде зростати. ►

ТЕМА 8. ВІДТВОРЕННЯ (ІМІТАЦІЯ) ПОВЕДІНКИ СИСТЕМ ЯК ОКРЕМИЙ ПІДХІД ДО ЇХ МОДЕЛЮВАННЯ: РІЗНИЦЕВІ МОДЕЛІ ДИСКРЕТНИХ ДИНАМІЧНИХ СИСТЕМ

Завдання 8.1. *Динаміка або процес зміни станів дискретної системи характеризується сукупністю (ланцюжком) переходів з одного стану в інший у визначені дискретні моменти часу. Тому ефективним інструментарієм для їх дослідження є різницеві моделі, тобто моделі, представлені різницевиими рівняннями. Приклади формалізації та знаходження розв'язків таких моделей продемонстровані нижче у підзавданнях 1) та 2).*

У найпростішому агрегованому варіанті односекторну (однопродуктову) економіку можна описати також як систему, стан якої у довільний момент часу характеризується сукупністю показників $\{X, V, Y, I, C\}$, де X – валовий випуск, V – виробниче споживання (проміжна продукція), Y – кінцевий випуск (продукт), I – валові інвестиції, C – невиробниче споживання. Нехай час дискретний ($t \in Z_+ = \{0, 1, 2, \dots\}$) і мають місце припущення:

- *у початковий момент $t = 0$*

$$X|_{t=0} = X_0; \quad (8.1.1)$$

- *складниками валового випуску є виробниче споживання та кінцевий випуск, причому задана частка*

α ($\alpha = \text{const}$, $0 < \alpha < 1$) виробничого споживання у валовому випуску:

$$X = V + Y, \quad (8.1.2)$$

$$V = \alpha X; \quad (8.1.3)$$

• кінцевий випуск поділяється на валові інвестиції та невиробниче споживання, частка якого у валовому випуску дорівнює β ($\beta = \text{const}$, $0 < \beta < 1$):

$$Y = I + C, \quad (8.1.4)$$

$$C = \beta X; \quad (8.1.5)$$

• приріст продукції у періоді $[t-1, t]$ ($(t-1)$ і t – відповідно попередній та поточний моменти часу), тобто $\Delta X_t = X_t - X_{t-1}$, залежить від обсягу інвестицій у момент $(t-1)$, який є початковим для цього періоду, та прямо пропорційний цьому обсягу з коефіцієнтом γ ($\gamma = \text{const}$):

$$\Delta X_t = X_t - X_{t-1} = \gamma I_{t-1}. \quad (8.1.6)$$

Необхідно:

1) формалізувати модель (рівняння) динаміки валового випуску X для дискретних моментів часу $t \in Z_+$ і знайти відповідний розв'язок (послідовність X_t , $t \in Z_+$) у випадку, коли невиробниче споживання є:

а) динамічною змінною;

б) сталою величиною;

2) з'ясувати характер динаміки (тренду) валового випуску на основі розв'язків моделей а), б) (підзавдання 1)), здійснивши графічну ілюстрацію розв'язків для множини моментів часу

$$t \in \tilde{Z}_+ = \{0, 1, 2, \dots, n\} \quad (n \geq 10, \quad n < +\infty),$$

$$\text{якщо } X_0 = 3, \quad c^* = 1,$$

$$\alpha = \frac{1}{2} + \frac{t}{2(6t^2 + 1)}, \quad \beta = \frac{1}{3} + \frac{2t^2 - t}{2(6t^2 + 1)}, \quad \gamma = 2 + \frac{t}{t+2}.$$

◀ **1)** Найперше зауважимо, що економічна система у цьому завданні повністю аналогічна (і навіть еквівалентна у певному сенсі) системі, про яку йшлося у підзавданні 7.2, хоча деякі припущення та дискретний час змінюють спосіб формалізації та аналізу її моделі.

Підкреслимо також, що множину дискретних моментів часу $\{t_0, t_1, t_2, \dots\}$ зазвичай замінюють множиною цілих чисел, тобто множиною Z_+ , використання якої у багатьох випадках є повністю оправданим і не впливає на загальність результатів (для дискретних економічних систем доцільність такої заміни очевидна, оскільки під часовим періодом $[t_{i-1}, t_i]$ часто розуміють i -тий рік).

а) Оскільки час дискретний, то співвідношення (8.1.2)–(8.1.5) конкретизуються для довільного моменту часу t у вигляді:

$$\begin{cases} V_t = \alpha X_t, & X_t = V_t + Y_t. \\ Y_t = I_t + C_t, & C_t = \beta X_t. \end{cases} \quad (8.1.7)$$

Далі, враховуючи (8.1.7) для моменту часу $(t-1)$ та (8.1.6), отримаємо рівність

$$X_{t-1} = \alpha X_{t-1} + \beta X_{t-1} + \frac{1}{\gamma} (X_t - X_{t-1})$$

або

$$X_t = \gamma \left(1 - \alpha - \beta + \frac{1}{\gamma} \right) X_{t-1} = \left(\frac{1 - \alpha - \beta + \mu}{\mu} \right) X_{t-1}, \quad (8.1.8)$$

де $\mu = \frac{1}{\gamma}$. Рівняння (8.1.8) є різницеvim [11].

Приєднавши до (8.1.8) початкову умову (8.1.1), одержимо відповідну задачу Коші, розв'язком якої буде послідовність

$$X_t = X_0 \left(\frac{1 - \alpha - \beta + \mu}{\mu} \right)^t, \quad t \in Z_+. \quad (8.1.9)$$

Справедливість формули (8.1.9) легко перевірити, здійснивши її підстановку у рівняння (8.1.8). Зауважимо, що рівняння (8.1.8) є лінійним однорідним різницеvim рівнянням першого порядку [11].

б) Якщо $C_t = c^*$ ($c^* = const$), то із урахуванням цього факту із перших трьох співвідношень (8.1.7) для моменту часу $(t-1)$ впливає, що

$$X_{t-1} = \alpha X_{t-1} + c^* + \frac{1}{\gamma} (X_t - X_{t-1})$$

або

$$X_t = \gamma \left(1 - \alpha + \frac{1}{\gamma} \right) X_{t-1} - \gamma c^*. \quad (8.1.10)$$

На відміну від (8.1.8) рівняння (8.1.10) є неоднорідним різницеvim рівнянням першого порядку. Його загальний розв'язок (як і у випадку диференціальних рівнянь) є сумою загального розв'язку відповідного йому однорідного рівняння та будь-якого частинного розв'язку неоднорідного

рівняння. Загальний розв'язок однорідного рівняння

$$X_t = \gamma \left(1 - \alpha + \frac{1}{\gamma} \right) X_{t-1}$$

конкретизується формулою

$$X_t^{z.o.} = D \left(\frac{1 - \alpha + \mu}{\mu} \right)^t, \quad (8.1.11)$$

де D – довільна стала.

Із теорії різницевих рівнянь відомо [11], що частинний розв'язок рівнянь типу (8.1.10) можна шукати у вигляді $X_t^{u.n.} = D^*$, де $D^* = const$. Після підстановки цього розв'язку у (8.1.10), матимемо рівність

$$D^* = \gamma \left(1 - \alpha + \frac{1}{\gamma} \right) D^* - \gamma c^*$$

або

$$D^* = \frac{\gamma c^* \mu}{1 - \alpha} = \frac{c^* \mu}{\mu(1 - \alpha)} = \frac{c^*}{1 - \alpha},$$

тому з огляду на (8.1.11)

$$X_t^{z.n.} = X_t^{z.o.} + X_t^{u.n.} = D \left(\frac{1 - \alpha + \mu}{\mu} \right)^t + \frac{c^*}{1 - \alpha}. \quad (8.1.12)$$

Взявши до уваги початкову умову (8.1.1), отримаємо рівність $X_0 = D + \frac{c^*}{1 - \alpha}$, тобто $D = X_0 - \frac{c^*}{1 - \alpha}$ і розв'язок задачі (8.1.10), (8.1.1) матиме вигляд:

$$X_t = \left(X_0 - \frac{c^*}{1 - \alpha} \right) \left(\frac{1 - \alpha + \mu}{\mu} \right)^t + \frac{c^*}{1 - \alpha}, \quad t \in Z_+. \quad (8.1.13)$$

Підкреслимо, що розв'язки (8.1.9) та (8.1.13) (послідовності валових випусків на множині Z_+) дають можливість визначити аналогічні послідовності на множині Z_+ величин V, Y, I, C , а у кінцевому результаті – послідовність станів досліджуваної економічної системи, тобто траєкторію дискретної системи у фазовому просторі зазначених змінних.

2) Для виконання цього підзавдання потрібно уточнити усі члени послідовностей (8.1.9) та (8.1.13) для заданих значень n, X_0, c^* і відповідних значень параметрів α, β, γ ($t \in \tilde{Z}_+$), а потім множини точок $\{(t, X_t), t \in \tilde{Z}_+\}$ проілюструвати на площині координат t, X . Для більш чіткої візуалізації тренду бажано побудувати ламані кусково-лінійні криві з вершинами у цих точках. ►

Завдання 8.2. *Економічна система характеризується сукупністю показників $\{Y, C, I\}$, де Y – національний дохід, C – невиробниче споживання, I – інвестиції в економіку. Значення цих показників уточнюються у дискретні моменти часу ($t \in Z_+$).*

Вихідними для моделювання цієї системи є припущення:

- у початковий момент часу $t = 0$

$$Y|_{t=0} = Y_0, \quad I|_{t=0} = I_0 \quad (8.2.1)$$

- у будь-який момент часу t національний дохід розподіляється на споживання та інвестиції:

$$Y_t = C_t + I_t; \quad (8.2.2)$$

• обсяг споживання у момент t лінійно залежить від національного доходу у момент $(t-1)$ (враховується лаг запізнення, що дорівнює одному періоду), тобто

$$C_t = a + bY_{t-1}, \quad (8.2.3)$$

де a та b відповідно автономне споживання та гранична схильність населення до споживання ($a \geq 0$, $a < b < 1$).

Потрібно:

1) формалізувати модель динаміки національного доходу та її розв'язок у випадку, коли:

а) обсяг інвестицій не залежить від національного доходу, тобто є автономною і сталою величиною, що дорівнює I_0 , для усіх $t \in Z_+$:

$$I_t = I_0; \quad (8.2.4)$$

б) обсяг інвестицій пов'язаний з лагом запізнення в один період і у момент t лінійно залежить від національного доходу у момент $(t-1)$, тобто

$$I_t = I_0 + sY_{t-1}, \quad (8.2.5)$$

де s ($0 < s < 1$) – гранична схильність до інвестування;

2) для моделей динаміки національного доходу (випадки а), б) із підзавдання 1)) визначити рівноважні значення доходів та величини мультиплікаторів досліджуваної системи, а також здійснити графічну інтерпретацію рівноважних і загальних розв'язків на площині t, Y , якщо $a = 100 + 5k$, $b = 0,8 + 0,1/k$, $I_0 = 200 + 20k$, $Y_0 = 6000 - 100k$, $s = 0,04 + 0,01/k$

($1 \leq k \leq 50$).

◀ **1)** Перед виконанням цього завдання нагадаємо, що основним результатом функціонування будь-якої економіки є ВВП (кінцевий продукт), який у вартісній формі складається із національного доходу та амортизаційного фонду. Національний дохід включає споживання (невиробниче) та інвестиції (чисті інвестиції). Тому зміст величин Y та I у цьому завданні не слід ототожнювати із їх змістом у завданнях 7.2 та 8.1.

а) Припущення (8.2.1)–(8.2.3) дозволяють перейти до співвідношення

$$Y_t = a + bY_{t-1} + I_0$$

або

$$Y_t - bY_{t-1} = a + I_0, \quad (8.2.6)$$

яке насправді формалізує один із найпростіших варіантів моделі динаміки національного доходу. У сенсі теорії різницевих рівнянь [11] співвідношення (8.2.6) є різницевим неоднорідним рівнянням першого порядку зі сталими коефіцієнтами. Елементарно переконатися у тому, що функція

$$Y_t^{3.0.} = Mb^t,$$

де M – довільна стала, є загальним розв'язком відповідного до (8.2.6) однорідного рівняння

$$Y_t - bY_{t-1} = 0.$$

Оскільки характеристичним для цього рівняння є рівняння $\lambda - b = 0$ або $\lambda = b$, а отже $Y_t^{3.0.} = M\lambda^t = Mb^t$.

Далі потрібно знайти будь-який частинний розв'язок неоднорідного рівняння (8.2.6), який буде

деякою сталою величиною \tilde{M} . Після підстановки \tilde{M} у рівняння (8.2.6) отримаємо

$$Y_t^{ч.н.} = \tilde{M} = \frac{a + I_0}{1 - b}.$$

Отже, загальний розв'язок рівняння (8.2.6) такий:

$$Y_t^{з.н.} = Y_t^{з.о.} + Y_t^{ч.н.} = Mb^t + \frac{a + I_0}{1 - b}. \quad (8.2.7)$$

Оскільки $Mb^t \rightarrow 0$ при $t \rightarrow \infty$ ($0 < b < 1$), то $Y_t^{з.н.} \rightarrow \tilde{M}$ при $t \rightarrow \infty$, тобто у даному випадку національний дохід монотонно спадатиме до свого рівноважного стану або стаціонарного розв'язку рівняння (8.2.6)

$$\tilde{Y} = \tilde{M}.$$

Із урахуванням початкової умови (8.2.1) розв'язок (8.2.7) конкретизується у вигляді (8.2.8):

$$Y_t = \left(Y_0 - \frac{a + I_0}{1 - b} \right) b^t + \frac{a + I_0}{1 - b}. \quad (8.2.8)$$

Зауважимо, що мультиплікатор $\frac{1}{1 - b}$ є додатною величиною. Уточнимо поняття мультиплікатора. У широкому розумінні мультиплікатор – це коефіцієнт, що характеризує зміну доходу залежно від зміни інвестицій. У сенсі досліджуваної моделі мультиплікатор є коефіцієнтом, який характеризує величину збільшення рівноважного національного доходу у випадку збільшення так званих автономних (незалежних від величин доходу) інвестицій (витрат суб'єктів макроекономіки). Щоб детальніше розкрити зміст мультиплікатора, припустимо, що обсяг автономних інвестицій I_0 збільшився до величини

$(I_0 + \Delta I)$. Тоді модель динаміки національного доходу (8.2.6) трансформується у модель

$$Y_t - bY_{t-1} = a + I_0 + \Delta I, \quad (8.2.9)$$

а розв'язки (8.2.7), (8.2.8) – у розв'язки

$$Y_t^{з.н.} = Mb^t + \frac{a + I_0 + \Delta I}{1 - b}, \quad (8.2.10)$$

$$Y_t = \left(Y_0 - \frac{a + I_0 + \Delta I}{1 - b} \right) b^t + \frac{a + I_0 + \Delta I}{1 - b}. \quad (8.2.11)$$

Співвідношення (8.2.10)–(8.2.11), тобто загальний розв'язок рівняння (8.2.9) та розв'язок задачі Коші (8.2.9), (8.2.1), дуже чітко пояснюють зміст мультиплікатора (коефіцієнта) $1/(1-b)$, оскільки збільшення I_0 до $(I_0 + \Delta I)$ приводить до збільшення рівноважного національного доходу від $(a + I_0)/(1-b)$ до $(a + I_0 + \Delta I)/(1-b)$ або до $\Delta \tilde{Y} = \Delta I/(1-b)$. Інакше кажучи, зростання доходу на величину $\Delta \tilde{Y}$ відбувається з ефектом мультиплікатора $1/(1-b)$.

б) Якщо має місце залежність (8.2.5), то із (8.2.2), (8.2.3), (8.2.5) випливає, що

$$Y_t = a + bY_{t-1} + I_0 + sY_{t-1}$$

або

$$Y_t - (b + s)Y_{t-1} = a + I_0. \quad (8.2.12)$$

Рівняння (8.2.12), яке формалізує модель динаміки національного доходу у цьому випадку, також є різницеvim неоднорідним рівнянням першого порядку. Враховуючи, що характеристичне рівняння для однорідного рівняння

$$Y_t - (b + s)Y_{t-1} = 0$$

записується у вигляді $\lambda - (b + s) = 0$ (тобто $\lambda = b + s$), то загальний розв'язок однорідного рівняння

$$Y_t^{3.o.} = N(b + s)^t,$$

де N – довільна стала.

Далі, припустивши, що $Y^{ч.н.} = \tilde{N} = const$, і підставивши його у рівняння (8.2.12), остаточно матимемо

$$Y_t^{ч.н.} = \tilde{N} = \frac{a + I_0}{1 - b - s}.$$

Отже, загальний розв'язок рівняння (8.2.12) має такий вигляд:

$$Y_t^{3.н.} = Y_t^{3.o.} + Y_t^{ч.н.} = N(b + s)^t + \frac{a + I_0}{1 - b - s}. \quad (8.2.13)$$

Конкретизувавши сталу N за допомогою початкової умови (8.2.1), отримаємо розв'язок задачі Коші (8.2.12), (8.2.1):

$$Y_t = \left(Y_0 - \frac{a + I_0}{1 - b - s} \right) (b + s)^t + \frac{a + I_0}{1 - b - s}. \quad (8.2.14)$$

Очевидно, що повинна виконуватися умова $1 - b - s > 0$ або $s < 1 - b$, оскільки у реальній ситуації дохід завжди є невід'ємною величиною. Ця умова забезпечує монотонно спадаючу збіжність розв'язку (8.2.13) чи (8.2.14) при $t \rightarrow \infty$ до стаціонарного розв'язку

$$\bar{Y} = \tilde{N} = \frac{a + I_0}{1 - b - s}.$$

Інакше кажучи, якщо гранична схильність до заощаджень $(1-b)$ буде більшою від граничної схильності до інвестування, то ринкова рівновага буде стабільною. Умова $s < 1-b$ гарантує також додатність мультиплікатора $1/(1-b-s)$.

2) Виконання цього підзавдання зводиться до елементарних обчислювальних і графічних процедур, зупинятися на яких немає сенсу. ►

ТЕМА 9. СИСТЕМИ УПРАВЛІННЯ. ОПТИМАЛЬНЕ УПРАВЛІННЯ ДИНАМІЧНИМИ СИСТЕМАМИ

Завдання 9.1. Проаналізувати поняття управління, системи управління, типів зв'язків і управлінь у системах управління та розкрити зміст принципів регулювання, їх переваг і недоліків.

◀ Для виконання цього завдання необхідно ознайомитися з теоретичними основами систем управління та пов'язаними із поняттям управління (керування) іншими базовими поняттями. Відповідний матеріал можна знайти у літературі, присвяченій системам управління та проблемам кібернетики, зокрема у [2, 13-15, 17]. ▶

Завдання 9.2. Значна кількість прикладних задач, зокрема економічних, фінансових і ін., можуть бути сформульовані та розв'язані як задачі оптимального управління (керування) динамічними системами (процесами). До таких задач відносяться у тому числі і задачі про оптимальний розподіл інвестиційних вкладень. Завдання полягає у тому, щоб:

1) формалізувати задачу про оптимальний розподіл інвестиційних вкладень заданого обсягу s одиниць (од.) між m підприємствами, тобто задачу максимізації їх сумарного прибутку (від використання $u^{(i)}$ од. інвестицій i -те підприємство має прибуток, що

дорівнює $k^{(i)}(u^{(i)})$ од.), як задачу управління (керування) дискретним процесом зі скінченним числом кроків і описати алгоритм методу динамічного програмування, за допомогою якого цю задачу можна розв'язати;

2) знайти розв'язок цієї задачі у випадку $s = 5$, $m = 3$ і цілочислових значень $u^{(i)}$ та $k^{(i)} = k^{(i)}(u^{(i)})$, наведених у таблицях 9.2.1 – 9.2.5.

9.2.1

$u^{(i)}$	$k^{(1)}$	$k^{(2)}$	$k^{(3)}$
0	0	0	0
1	2	3	4
2	7	8	9
3	10	11	12
4	12	13	14
5	15	16	17

9.2.2

$u^{(i)}$	$k^{(1)}$	$k^{(2)}$	$k^{(3)}$
0	0	0	0
1	17	11	8
2	18	15	10
3	20	19	12
4	21	25	14
5	22	27	16

9.2.3

$u^{(i)}$	$k^{(1)}$	$k^{(2)}$	$k^{(3)}$
0	0	0	0
1	20	27	22
2	22	28	26
3	23	29	27
4	24	32	28
5	25	33	29

9.2.4

$u^{(i)}$	$k^{(1)}$	$k^{(2)}$	$k^{(3)}$
0	0	0	0
1	25	15	10
2	30	20	20
3	40	30	30
4	45	40	40
5	50	45	50

9.2.5

$u^{(i)}$	$k^{(1)}$	$k^{(2)}$	$k^{(3)}$
0	0	0	0
1	11	12	15
2	21	22	17
3	31	32	22
4	41	33	30
5	50	43	40

◀ **1)** Оскільки інвестор вкладає s од. своїх ресурсів у m підприємств, а його мета – максимізація сумарного прибутку від усіх вкладень, то задача про оптимальний розподіл інвестицій формалізується так:

$$\left\{ \begin{array}{l} k(u) = \sum_{i=1}^m k^{(i)}(u^{(i)}) \rightarrow \max, \\ \sum_{i=1}^m u^{(i)} = s, \\ u^{(i)} \geq 0, \quad u^{(i)} = \overline{1, m}. \end{array} \right. \quad (9.2.1)$$

Очевидно, що (9.2.1) є задачею математичного програмування. Якщо функції прибутків $k^{(i)}(u^{(i)})$, $i = \overline{1, m}$, задані аналітично та лінійні, то (9.2.1) – це задача лінійного програмування, а якщо хоча б одна із заданих аналітично функцій прибутків нелінійна, то (9.2.1) – задача нелінійного програмування. У кожному з цих випадків задачу (9.2.1) можна розв'язати відомими методами (наприклад симплекс-методом можна розв'язати задачу лінійного програмування). Однак є й інший підхід до розв'язування задачі (9.2.1), якщо її переформалізувати як задачу оптимального керування дискретним процесом зі скінченним числом кроків (тобто оптимального управління скінченною сукупністю станів системи, що відповідають певним дискретним моментам часу). Для цього потрібно поетапно аналізувати ефективність вкладень на одному, двох і т.д. – m підприємствах. Це приводить до m етапів на кожному з яких стан системи характеризується обсягом сумарних вкладень у перших i підприємств, тобто у підприємства з номерами $1, 2, \dots, i$, якщо розглядається i -тий етап. При цьому

керуванням на i -тому етапі буде величина $u^{(i)}$. Якщо через $x^{(i)}$ позначити сумарний обсяг вкладень на i -тому етапі, то задачу (9.2.1) можна звести до задачі

$$\begin{cases} \sum_{i=1}^m k^{(i)}(u^{(i)}) \rightarrow \max, \\ x^{(i)} = x^{(i-1)} + u^{(i)}, \quad i = \overline{1, m}, \\ 0 \leq u^{(i)} \leq s - x^{(i-1)}, \quad i = \overline{1, m}, \\ x^{(0)} = 0, \quad x^{(m)} = s. \end{cases} \quad (9.2.2)$$

Задача (9.2.2) є багатоетапною (багатокроковою) задачею оптимізації або задачею оптимального керування процесом інвестиційних вкладень і полягає у тому, щоб знайти такий розподіл вкладень $\{u^{(1)*}, \dots, u^{(m)*}\}$ та відповідну йому траєкторію станів системи $\{x^{(0)*}, x^{(1)*}, \dots, x^{(m)*}\}$, які дозволяють інвестору досягти максимального сумарного прибутку.

Метод динамічного програмування та власне алгоритм знаходження розв'язку задач типу (9.2.2) викладений у [10]. Цей метод ґрунтується на принципі оптимальності Беллмана, який можна сформулювати так [10]:

- *«оптимальне керування у будь-який момент часу (чи період для дискретного процесу) не залежить від передісторії системи і визначається тільки станом системи у цей момент часу (чи станом системи на початку періоду) та метою керування»;*

або

- *«яким би не був стан системи перед наступним*

кроком, керування на цьому кроці потрібно вибрати так, щоб сума ефективності процесу на цьому кроці й оптимальної ефективності процесу на всіх наступних кроках була максимальною».

Отже, якщо дискретний процес розбитий на m етапів чи кроків (на m часових періодів $[t_0, t_1], \dots, [t_{m-1}, t_m]$), то згідно із принципом Беллмана вводиться так звана функція Беллмана останніх $(m-i+1)$ кроків

$$B_i(x^{(i-1)}) = k^{(i)}(x^{(i-1)}, u^{(i)*}) + \sum_{j=i+1}^m k^{(j)}(x^{(j-1)*}, u^{(j)*}), \quad (9.2.3)$$

де $k^{(i)}(x^{(i-1)}, u^{(i)})$ – функція ефективності i -того кроку (етапу).

Із функції (9.2.3) випливає, що для останнього m -го кроку

$$B_m(x^{(m-1)}) = \max_{u^{(m)}} k^{(m)}(x^{(m-1)}, u^{(m)}), \quad (9.2.4)$$

де $u^{(m)} \in U^{(m)}(x^{(m-1)})$ ($U^{(m)}(x^{(m-1)})$ – задана множина допустимих керувань на m -му кроці),

а для усіх інших кроків

$$B_i(x^{(i-1)}) = \max_{u^{(i)}} \left\{ k^{(i)}(x^{(i-1)}, u^{(i)}) + B_{i+1}(f^{(i)}(x^{(i-1)}, u^{(i)})) \right\}, \quad i = \overline{1, m-1}, \quad (9.2.5)$$

де $u^{(i)} \in U^{(i)}(x^{(i-1)})$ ($U^{(i)}(x^{(i-1)})$ – задана множина допустимих керувань на i -му кроці), $f^{(i)}(x^{(i-1)}, u^{(i)})$ – задана функція i -того стану системи.

Співвідношення (9.2.4) і (9.2.5) прийнято називати

рівняннями Беллмана. Вони дозволяють знайти функції Беллмана $B_m(x^{(m-1)}), B_{m-1}(x^{(m-2)}), \dots, B_1(x^{(0)})$ та відповідні їм умовні оптимальні керування $u^{(m)*}(x^{(m-1)}), \dots, u^{(1)*}(x^{(0)})$. Зворотний процес, тобто безумовна оптимізація, починається із знаходження $x^{(0)*}$, тобто із розв'язування задачі

$$B_1(x^{(0)*}) = \max_{x^{(0)}} B_1(x^{(0)}), \quad (9.2.6)$$

де $x^{(0)} \in X^{(0)}$ ($X^{(0)}$ – задана множина допустимих початкових станів системи),

і продовжується послідовним уточненням оптимальних керувань і оптимальних станів за допомогою співвідношень

$$\begin{cases} u^{(1)*} = u^{(1)*}(x^{(0)*}), & x^{(1)*} = f^{(1)}(x^{(0)*}, u^{(1)*}), \\ u^{(2)*} = u^{(2)*}(x^{(1)*}), & x^{(2)*} = f^{(2)}(x^{(1)*}, u^{(2)*}), \\ \dots \\ u^{(m)*} = u^{(m)*}(x^{(m-1)*}), & x^{(m)*} = f^{(m)}(x^{(m-1)*}, u^{(m)*}). \end{cases} \quad (9.2.7)$$

Алгоритм методу динамічного програмування, який відображає окреслену вище умовну та безумовну оптимізацію, продемонструємо на конкретному прикладі у пункті 2).

2) Отже, нехай $s = 5$, $m = 3$, а можливі значення $u^{(i)}$ та $k^{(i)} = k^{(i)}(u^{(i)})$ наведені у таблиці 9.2.6 (функціональні залежності можуть задаватися як аналітично, так і таблично).

Таблиця 9.2.6

$u^{(i)} \backslash k^{(i)}$	0	1	2	3	4	5
$k^{(1)}$	0	1	3	4	5	6
$k^{(2)}$	0	2	4	6	7	8
$k^{(3)}$	0	3	5	7	8	10

а) Умовна оптимізація

Крок 1. Розв'язування задачі (9.2.4), яка у нашому випадку зводиться до побудови функції

$$B_3(x^{(2)}) = \max_{u^{(3)}} k^{(3)}(u^{(3)}), \quad (9.2.8)$$

де $0 \leq u^{(3)} \leq 5 - x^{(2)}$,

та знаходження $u^{(3)*}(x^{(2)})$, тобто розв'язку, що відповідає функції (9.2.8) для різних значень $x^{(2)}$ (табл. 9.2.7). Щодо побудови табл. 9.2.7 зауважимо, що із задачі (9.2.2) випливає, що $x^{(3)} = s = 5$ ($m = 3$), тому при $x^{(2)} = 0$ відповідно $u^{(3)} = x^{(3)} - x^{(2)} = 5$, при $x^{(2)} = 1$ відповідно $u^{(3)} = x^{(3)} - x^{(2)} = 4$ і т.д., що власне і відображено у цій таблиці.

Таблиця 9.2.7

$x^{(2)}$	0	1	2	3	4	5
$B_3(x^{(2)})$	10	8	7	5	3	0
$u^{(3)*}(x^{(2)})$	5	4	3	2	1	0

Крок 2. Реалізація співвідношення (9.2.5) при $m = 2$, тобто побудова функції

$$B_2(x^{(1)}) = \max_{u^{(2)}} \{k^{(2)}(u^{(2)}) + B_3(x^{(1)} + u^{(2)})\}, \quad (9.2.9)$$

де $0 \leq u^{(2)} \leq 5 - x^{(1)}$,

та знаходження $u^{(2)*}(x^{(1)})$ як розв'язку, що відповідає функції (9.2.9). З цією метою спочатку побудуємо на основі таблиць 9.2.6 і 9.2.7 таблицю 9.2.8 для значень суми $\{k^{(2)}(u^{(2)}) + B_3(x^{(1)} + u^{(2)})\}$.

Таблиця 9.2.8

$u^{(2)} \backslash x^{(1)}$	0	1	2	3	4	5
0	10	10	11	11	10	8
1	8	9	9	9	7	—
2	7	7	7	6	—	—
3	5	5	4	—	—	—
4	3	0	—	—	—	—
5	0	—	—	—	—	—

Виділені у таблиці 9.2.8 числа - це максимальні (по $u^{(2)}$) значення величини $\{k^{(2)}(u^{(2)}) + B_3(x^{(1)} + u^{(2)})\}$ при різних значеннях $x^{(1)}$. З урахуванням даних таблиці 9.2.8 будується таблиця 9.2.9.

Таблиця 9.2.9

$x^{(1)}$	0	1	2	3	4	5
$B_2(x^{(1)})$	11	9	7	5	3	0
$u^{(2)*}(x^{(1)})$	2,3	1,2,3	0,1,2	0,1	0	0

Як впливає із таблиці 9.2.9 $u^{(2)*}(x^{(1)})$ при $x^{(1)} = 0$, $x^{(1)} = 1$, $x^{(1)} = 2$, $x^{(1)} = 3$ знаходяться неоднозначно.

Крок 3. Реалізація співвідношень (9.2.5) при $m=1$, тобто побудова функції

$$\begin{aligned}
 B_1(x^{(0)}) &= \max_{u^{(1)}} \{k^{(1)}(u^{(1)}) + B_2(x^{(0)} + u^{(1)})\} = \\
 &= \max_{u^{(1)}} \{k^{(1)}(u^{(1)}) + B_2(u^{(1)})\}, \quad (9.2.10)
 \end{aligned}$$

де $0 \leq u^{(1)} \leq 5 - x^{(0)}$, $x^{(0)} = 0$,

та знаходження $u^{(1)*}(x^{(0)})$ як розв'язку, що відповідає функції (9.2.10). На основі таблиць 9.2.6 та 9.2.9 отримаємо таблицю 9.2.10.

Таблиця 9.2.10

$u^{(1)}$	0	1	2	3	4	5
$k^{(1)}(u^{(1)}) + B_2(u^{(1)})$	11	10	10	9	8	6

Із урахуванням таблиці 9.2.10 дійдемо висновку, що $B_1(x^{(0)}) = B_1(0) = 11$.

б) Безумовна оптимізація

Крок 1. Із того, що $x^{(0)*} = 0$ ($B_1(0) = 11$) є розв'язком задачі (9.2.6), і згідно із (9.2.7) $u^{(1)*}(x^{(0)*}) = u^{(1)*} = 0$ (табл. 9.2.10), маємо $x^{(1)*} = x^{(0)*} + u^{(1)*} = 0 + 0 = 0$.

Крок 2. Із урахуванням того, що $x^{(1)*} = 0$, згідно із співвідношеннями (9.2.7) із таблиці 9.2.9 випливає, що відповідно $u^{(2)*} = u^{(2)*}(x^{(1)*}) = u^{(2)*}(0) = 2$ і $u^{(2)*} = u^{(2)*}(x^{(1)*}) = u^{(2)*}(0) = 3$, тому також буде два варіанти $x^{(2)*}$, а саме $x^{(2)*} = x^{(1)*} + u^{(2)*} = 0 + 2 = 2$ і $x^{(2)*} = x^{(1)*} + u^{(2)*} = 0 + 3 = 3$.

Крок 3. Оскільки для $x^{(2)*} = 2$ і $x^{(2)*} = 3$ згідно із співвідношеннями (9.2.7) із таблиці 9.2.7 випливає, що відповідно $u^{(3)*} = u^{(3)*}(x^{(2)*}) = 3$ і $x^{(3)*} = x^{(2)*} + u^{(3)*} = 2 + 3 = 5$ та $u^{(3)*} = u^{(3)*}(x^{(2)*}) = 2$ і $x^{(3)*} = x^{(2)*} + u^{(3)*} = 3 + 2 = 5$.

Процес розв'язування завершено. Знайдено два варіанти оптимального розподілу інвестицій. У першому варіанті маємо оптимальне керування $\{u^{(1)*}, u^{(2)*}, u^{(3)*}\} = \{0, 2, 3\}$ та оптимальну траєкторію $\{x^{(0)*}, x^{(1)*}, x^{(2)*}, x^{(3)*}\} = \{0, 0, 2, 5\}$, а у другому – оптимальне керування $\{u^{(1)*}, u^{(2)*}, u^{(3)*}\} = \{0, 3, 2\}$ та оптимальну траєкторію $\{x^{(0)*}, x^{(1)*}, x^{(2)*}, x^{(3)*}\} = \{0, 0, 3, 5\}$. Це означає, що інвестору вигідно здійснювати вкладення лише у друге та

третє підприємства. При цьому максимальний сумарний прибуток ($k(u) = B_1(0) = 11$ од.) досягається тоді, коли другому і третьому підприємствам виділяються відповідно 2 од. і 3 од. інвестиційних вкладень або навпаки – відповідно 3 од. і 2 од. інвестиційних вкладень. ►

Завдання 9.3. *До задач оптимального управління (керування) дискретними системами (процесами) відноситься також задача про розподіл s од. інвестиційних вкладень між двома підприємствами Π_1 і Π_2 на t років планового періоду (s і t – задані величини). З огляду на те, що вкладених u од. інвестицій у підприємство Π_j дають йому у кінці року $k^{(j)}(u)$ од. прибутку та зменшуються до $f^{(j)}(u)$ од. ($j = \overline{1,2}$), що залишаються інвестиційним ресурсом для вкладень на наступний рік, потрібно:*

1) формалізувати цю задачу як задачу багатокрокового (багатоетапного) процесу оптимізації, яка полягає у тому, щоб визначити такий розподіл інвестицій між підприємствами, при якому за t років сумарний прибуток підприємств буде максимальним;

2) розв'язати її методом динамічного програмування для заданих у варіантах 9.3.1 – 9.3.5 значень t , s і функцій $k^{(1)} = k^{(1)}(u)$, $k^{(2)} = k^{(2)}(u)$, $f^{(1)} = f^{(1)}(u)$, $f^{(2)} = f^{(2)}(u)$.

дані варіант	m	s	$k^{(1)}(u)$	$k^{(2)}(u)$	$f^{(1)}(u)$	$f^{(2)}(u)$
9.3.1.	4	400	$1,7u$	$1,4u$	$0,22u$	$0,2u$
9.3.2.	4	300	$1,8u$	$1,6u$	$0,2u$	$0,7u$
9.3.3.	4	200	$0,9u$	$0,88u$	$0,3u$	$0,8u$
9.3.4.	4	100	$2u$	$4u$	$0,1u$	$0,8u$
9.3.5.	4	500	$0,5u$	$0,6u$	$0,8u$	$0,7u$

◀ **1)** Нехай $x^{(i-1)}$ ($i = \overline{1, m}$) – сумарний обсяг інвестиційних вкладень на початку i -того року ($[t_0, t_1]$ – перший рік, ..., $[t_{m-1}, t_m]$ – m -тий рік), $u^{(i)}$ – величина вкладень у підприємство Π_1 ($0 \leq u^{(i)} \leq x^{(i-1)}$), а $(x^{(i-1)} - u^{(i)})$ – величина вкладень у підприємство Π_2 у цьому ж i -тому році. Тоді загальний прибуток від вкладень у i -тому році дорівнює $\{k^{(1)}(u^{(i)}) + k^{(2)}(x^{(i-1)} - u^{(i)})\}$ од. Перехід від стану $x^{(i-1)}$ до стану $x^{(i)}$ здійснюється під впливом зазначених вище інвестиційних вкладень (управлінь), частина яких, а саме $(f^{(1)}(u^{(i)}) + f^{(2)}(x^{(i-1)} - u^{(i)}))$ од., у кінці i -того року залишається нереалізованою, тобто є ресурсом для інвестування на наступний $(i+1)$ -й рік, тому співвідношення $x^{(i)} = f^{(1)}(u^{(i)}) + f^{(2)}(x^{(i-1)} - u^{(i)})$ ($i = \overline{1, m}$) є рівняннями станів для досліджуваної системи

інвестування. З урахуванням уже введених формалізацій задачу про оптимальний розподіл інвестицій між двома підприємствами протягом m років запишемо у вигляді:

$$\begin{cases} \sum_{i=1}^m [k^{(1)}(u^{(i)}) + k^{(2)}(x^{(i-1)} - u^{(i)})] \rightarrow \max, \\ x^{(i)} = f^{(1)}(u^{(i)}) + f^{(2)}(x^{(i-1)} - u^{(i)}), \quad i = \overline{1, m}, \\ 0 \leq u^{(i)} \leq x^{(i-1)}, \quad i = \overline{1, m}, \\ x^{(0)} = s, \quad x^{(m)} \geq 0. \end{cases} \quad (9.3.1)$$

Наявність у задачі (9.3.1) умови $x^{(m)} \geq 0$ свідчить про те, що у кінці планового періоду (m -того року) деяка кількість інвестиційних ресурсів ще може бути не засвоєна, тобто може бути використана у майбутньому.

2) Наведемо приклад розв'язування задачі (9.3.1) методом динамічного програмування у випадку, коли $m = 4$; $s = 200$; $k^{(1)}(u) = 0,3u$; $k^{(2)}(u) = 0,6u$; $f^{(1)}(u) = u$; $f^{(2)}(u) = 0,8u$.

На відміну від прикладу, розглянутого у попередньому завданні 9.2, тут ми маємо справу з аналітично заданими функціями прибутків і станів, але розв'язування задачі методом динамічного програмування здійснюється за аналогічним алгоритмом, реалізацію якого наводимо нижче.

а) Умовна оптимізація

Крок 1. Згідно із (9.2.4) при $m = 4$

$$B_4(x^{(3)}) = \max_{u^{(4)}} [k^{(1)}(u^{(4)}) + k^{(2)}(x^{(3)} - u^{(4)})] =$$

$$\begin{aligned}
&= \max_{u^{(4)}} \left[0, 3u^{(4)} + 0,6(x^{(3)} - u^{(4)}) \right] = \\
&= \max_{u^{(4)}} \left[0, 6x^{(3)} - 0,3u^{(4)} \right] = 0,6x^{(3)},
\end{aligned}$$

оскільки $0 \leq u^{(4)} \leq x^{(3)}$. Отже, умовне оптимальне керування $u^{(4)*}(x^{(3)}) = 0$.

Крок 2. Конкретизація співвідношення (9.2.5) для $i = m - 1 = 3$ та відповідного умовного оптимального керування.

$$\begin{aligned}
B_3(x^{(2)}) &= \max_{u^{(3)}} \left[k^{(1)}(u^{(3)}) + k^{(2)}(x^{(2)} - u^{(3)}) + B_4(x^{(3)}) \right] = \\
&= \max_{u^{(3)}} \left[k^{(1)}(u^{(3)}) + k^{(2)}(x^{(2)} - u^{(3)}) + \right. \\
&\quad \left. + B_4(f^{(1)}(u^{(3)}) + f^{(2)}(x^{(2)} - u^{(3)})) \right] = \\
&= \max_{u^{(3)}} \left[0, 3u^{(3)} + 0,6(x^{(2)} - u^{(3)}) + \right. \\
&\quad \left. + 0,6(u^{(3)} + 0,8(x^{(2)} - u^{(3)})) \right] = \\
&= \max_{u^{(3)}} \left[0, 3u^{(3)} + 0,6x^{(2)} - 0,6u^{(3)} + \right. \\
&\quad \left. + 0,6u^{(3)} + 0,48x^{(2)} - 0,48u^{(3)} \right] = \\
&= \max_{u^{(3)}} \left[1,08x^{(2)} - 0,18u^{(3)} \right] = 1,08x^{(2)},
\end{aligned}$$

бо $0 \leq u^{(3)} \leq x^{(2)}$.

Умовне оптимальне керування $u^{(3)*}(x^{(2)}) = 0$.

Крок 3. Конкретизація (9.2.5) для $i = m - 2 = 2$ та відповідного умовного оптимального керування.

$$B_2(x^{(1)}) = \max_{u^{(2)}} \left[k^{(1)}(u^{(2)}) + k^{(2)}(x^{(1)} - u^{(2)}) + B_3(x^{(2)}) \right] =$$

$$\begin{aligned}
&= \max_{u^{(2)}} \left[k^{(1)}(u^{(2)}) + k^{(2)}(x^{(1)} - u^{(2)}) + \right. \\
&+ B_3 \left(f^{(1)}(u^{(2)}) + f^{(2)}(x^{(1)} - u^{(2)}) \right) \Big] = \\
&= \max_{u^{(2)}} \left[0, 3u^{(2)} + 0, 6(x^{(1)} - u^{(2)}) + \right. \\
&+ B_3 \left(u^{(2)} + 0, 8(x^{(1)} - u^{(2)}) \right) \Big] = \\
&= \max_{u^{(2)}} \left[0, 3u^{(2)} + 0, 6x^{(1)} - 0, 6u^{(2)} + \right. \\
&+ 1, 08(u^{(2)} + 0, 8(x^{(1)} - u^{(2)})) \Big] = \\
&= \max_{u^{(2)}} (0, 3u^{(2)} + 0, 6x^{(1)} - 0, 6u^{(2)} + \\
&+ 1, 08u^{(2)} + 0, 864x^{(1)} - 0, 864u^{(2)}) = \\
&= \max_{u^{(2)}} (1, 464x^{(1)} - 0, 084u^{(2)}) = 1, 464x^{(1)},
\end{aligned}$$

бо $0 \leq u^{(2)} \leq x^{(1)}$. Умовне оптимальне керування $u^{(2)*}(x^{(1)}) = 0$.

Крок 4. Конкретизація функції $B_1(x^{(0)})$ згідно із співвідношенням (9.2.5) та умовного оптимального керування $u^{(1)*}(x^{(0)})$.

$$\begin{aligned}
B_1(x^{(0)}) &= \max_{u^{(1)}} \left[k^{(1)}(u^{(1)}) + k^{(2)}(x^{(0)} - u^{(1)}) + B_2(x^{(1)}) \right] = \\
&= \max_{u^{(1)}} \left[k^{(1)}(u^{(1)}) + k^{(2)}(x^{(0)} - u^{(1)}) + \right. \\
&+ B_2 \left(f^{(1)}(u^{(1)}) + f^{(2)}(x^{(0)} - u^{(1)}) \right) \Big] = \\
&= \max_{u^{(1)}} \left[0, 3u^{(1)} + 0, 6(x^{(0)} - u^{(1)}) + \right.
\end{aligned}$$

$$\begin{aligned}
& +B_2\left(u^{(1)} + 0,8\left(x^{(0)} - u^{(1)}\right)\right) \Big] = \\
& = \max_{u^{(1)}} \left[0,3u^{(1)} + 0,6x^{(0)} - 0,6u^{(1)} + \right. \\
& \left. + 1,464\left(u^{(1)} + 0,8\left(x^{(0)} - u^{(1)}\right)\right) \right] = \\
& = \max_{u^{(1)}} \left[0,3u^{(1)} + 0,6x^{(0)} - 0,6u^{(1)} + \right. \\
& \left. + 1,464u^{(1)} + 1,1712x^{(0)} - 1,1712u^{(1)} \right] = \\
& = \max_{u^{(1)}} \left[1,7712x^{(0)} - 0,0072u^{(1)} \right] = 1,7712x^{(0)},
\end{aligned}$$

бо $0 \leq u^{(1)} \leq x^{(0)}$. Умовне оптимальне керування $u^{(1)*}(x^{(0)}) = 0$.

б) Безумовна оптимізація

Крок 1. При $x^{(0)} = s = 200$ розв'язком задачі (9.2.6) є $x^{(0)*} = 200$; $B_1(x^{(0)*}) = B_1(200) = 1,7712 \cdot 200 = 354,24$.

Із (9.2.7) випливає, що $u^{(1)*}(x^{(0)*}) = u^{(1)*} = 0$,

$$\begin{aligned}
x^{(1)*} &= f^{(1)}\left(u^{(1)*}\right) + f^{(2)}\left(x^{(0)*} - u^{(1)*}\right) = \\
&= u^{(1)*} + 0,8\left(x^{(0)*} - u^{(1)*}\right) = 0 + 0,8(200 - 0) = 160.
\end{aligned}$$

Крок 2. Згідно із (9.2.7)

$$\begin{aligned}
u^{(2)*} &= u^{(2)*}\left(x^{(1)*}\right) = 0, \\
x^{(2)*} &= f^{(1)}\left(u^{(2)*}\right) + f^{(2)}\left(x^{(1)*} - u^{(2)*}\right) = \\
&= u^{(2)*} + 0,8\left(x^{(1)*} - u^{(2)*}\right) = 0 + 0,8(160 - 0) = 128.
\end{aligned}$$

Крок 3. Аналогічно із (9.2.7) випливає, що

$$u^{(3)*} = u^{(3)*}\left(x^{(2)*}\right) = 0,$$

$$\begin{aligned}
 x^{(3)*} &= f^{(1)}(u^{(3)*}) + f^{(2)}(x^{(2)*} - u^{(3)*}) = \\
 &= u^{(3)*} + 0,8(x^{(2)*} - u^{(3)*}) = 0 + 0,8(128 - 0) = 102,4.
 \end{aligned}$$

Крок 4. Як і раніше, згідно із (9.2.7)

$$u^{(4)*}(x^{(3)*}) = 0,$$

$$\begin{aligned}
 x^{(4)*} &= f^{(1)}(u^{(4)*}) + f^{(2)}(x^{(3)*} - u^{(4)*}) = \\
 &= u^{(4)*} + 0,8(x^{(3)*} - u^{(4)*}) = 0 + 0,8(102,4 - 0) = 81,92.
 \end{aligned}$$

Підсумовуючи попередні обчислення, констатуємо, що оптимальними розподілами (оптимальними керуваннями) для першого та другого підприємств відповідно будуть $\{u^{(1)*}, u^{(2)*}, u^{(3)*}, u^{(4)*}\} = \{0, 0, 0, 0\}$ та $\{v^{(1)*}, v^{(2)*}, v^{(3)*}, v^{(4)*}\} =$
 $= \{x^{(0)*} - u^{(1)*}, x^{(1)*} - u^{(2)*}, x^{(2)*} - u^{(3)*}, x^{(3)*} - u^{(4)*}\} =$
 $= \{200; 160; 128; 102,4\}$. При цьому оптимальною траєкторією буде траєкторія $\{x^{(0)*}, x^{(1)*}, x^{(2)*}, x^{(3)*}, x^{(4)*}\} =$
 $= \{200; 160; 128; 102,4; 81,92\}$, а оптимальний сумарний прибуток $k^* = 0,6(200 + 160 + 128 + 102,4) = 354,24$. ►

ТЕМА 10. ЕНТРОПІЯ СИСТЕМ ЗІ СКІНЧЕННИМИ КІЛЬКОСТЯМИ МОЖЛИВИХ СТАНІВ І КІЛЬКІСНА МІРА ІНФОРМАЦІЇ ПРО НИХ

Завдання 10.1. *Розкрити поняття інформації, інформаційної системи, ентропії системи. Проаналізувати властивості ентропії системи із скінченною кількістю станів (системи дискретного типу) та формулу для визначення кількості інформації про систему. Знайти кількість інформації про систему, якщо до повідомлення про неї вона може перебувати у*

трьох станах із ймовірностями відповідно $p_1 = \frac{1}{k}$,

$p_2 = \frac{2}{k}$, $p_3 = \frac{k-3}{k}$ ($k \geq 4$), а після повідомлення про неї – у

двох станах із ймовірностями відповідно $\tilde{p}_1 = \frac{k-1}{2k}$,

$\tilde{p}_2 = \frac{k+1}{2k}$ ($k \geq 4$).

◀ Перша частина завдання пов'язана безпосередньо із обґрунтуванням деяких ключових понять теорії інформації, основи якої викладено у ряді літературних джерел, у тому числі й навчальному посібнику [2], тому зупинятися на цьому немає сенсу. Що стосується другої частини завдання, то для її виконання потрібно використати міру невизначеності системи, яку у свій час запропонував американський математик та інженер

Клод Шеннон. Якщо система може перебувати у N можливих станах (N – натуральне число), причому відомий ймовірнісний розподіл цих станів, де p_i ($i = \overline{1, N}$) – ймовірність появи i -го стану, то згідно із Шенноном міра H невизначеності системи знаходиться за формулою

$$H = - \sum_{i=1}^N p_i \log_2 p_i . \quad (10.1.1)$$

Ця формула строго обґрунтована. *Величину (10.1.1) ще називають ентропією системи.* Ентропія має ряд властивостей, що мають застосування при її визначеності та оцінюванні, зокрема ентропія:

- невід’ємна величина;
- обмежена величина;
- нульова, якщо стан системи відомий;
- максимальна, якщо всі стани системи мають однакову ймовірність;
- неперервна як функція від аргументів p_1, \dots, p_N ;
- не зміниться, якщо сукупність станів системи доповнити одним неможливим станом чи декількома неможливими станами.

Роль інформації полягає у зменшенні ступеня невизначеності системи. Кількість інформації про систему, що міститься у деякому повідомленні про неї, дорівнює різниці ентропій до повідомлення і після повідомлення про систему. Інакше кажучи, якщо H – ентропія системи до повідомлення про неї, а \tilde{H} – ентропія системи після повідомлення про неї, то

кількість інформації I про систему визначається за формулою

$$I = H - \tilde{H}, \quad (10.1.2)$$

де \tilde{N} – кількість можливих станів системи після повідомлення про неї, \tilde{p}_j – ймовірність i -того стану ($j=1, \tilde{N}$),

$$\tilde{H} = -\sum_{j=1}^{\tilde{N}} \tilde{p}_j \log_2 \tilde{p}_j. \quad (10.1.3)$$

Отже, знаючи величини (10.1.1) та (10.1.3), легко встановити величину (10.1.2). До прикладу, якщо згідно із умовою завдання $k=4$, то, скориставшись співвідношеннями (10.1.1) – (10.1.3), знайдемо:

$$\begin{aligned} H &= -\left(\frac{1}{4} \log_2 \frac{1}{4} + \frac{1}{2} \log_2 \frac{1}{2} + \frac{1}{4} \log_2 \frac{1}{4}\right) = \\ &= -(0,5 \log_2 0,25 + 0,5 \log_2 0,5) = \\ &= -0,5 \log_2 0,125 = -0,5 \cdot \ln 0,125 / \ln 2 = 1,5 \text{ (біт);} \end{aligned}$$

$$\begin{aligned} \tilde{H} &= -\left(\frac{3}{8} \log_2 \frac{3}{8} + \frac{5}{8} \log_2 \frac{5}{8}\right) = \\ &= -(0,375 \log_2 0,375 + 0,625 \log_2 0,625) = \\ &= -\frac{1}{\ln 2} (0,375 \ln 0,375 + 0,625 \ln 0,625) = 1,2 \text{ (біт);} \end{aligned}$$

$$I = H - \tilde{H} = 1,5 - 1,2 = 0,3 \text{ (біт).}$$

Отриманий результат у цьому конкретному випадку свідчить, що невизначеність системи після повідомлення про неї зменшилася на 0,3 біт. ►

Завдання 10.2. 1) уточнити поняття ймовірнісного розподілу одно- та двовимірних дискретних систем як відповідних розподілів одно- та двовимірних дискретних випадкових величин (д.в.в.);

2) обґрунтувати поняття та формули умовної ентропії однієї системи стосовно іншої та ентропії сукупної системи;

3) розкрити властивості умовної ентропії та ентропії сукупної системи;

4) проаналізувати формулу для визначення кількості інформації про одну систему, що міститься у повідомленні про іншу систему, та деталізувати її властивості.

◀ **1)** Метою цього завдання є уточнення тих фактів з теорії сукупності систем (тобто систем, складовими яких є окремі системи), які характеризують взаємний вплив однієї системи на іншу, зокрема у контексті кількісного вимірювання невизначеності систем та інформації про них. Оскільки йдеться про системи зі скінченною кількістю можливих станів, то підкреслимо, що ймовірнісні розподіли цих систем асоціюються із розподілами дискретних випадкових величин (д.в.в.) [2]. Інакше кажучи, надалі говорячи про систему X зі станами x_i та їх ймовірностями $P(X = x_i) = p_i$ ($i = \overline{1, N}$) чи про систему Y зі станами y_j та їх ймовірностями $P(Y = y_j) = \bar{p}_j$ ($j = \overline{1, M}$), ми будемо мати на увазі деякі дві одновимірні д.в.в. X та Y , закони розподілу яких задані відповідно таблицями 10.2.1 та 10.2.2.

Таблиця 10.2.1

X	x_1	x_2	\dots	x_N
P	P_1	P_2	\dots	P_N

Таблиця 10.2.2

Y	y_1	y_2	\dots	y_M
\bar{P}	\bar{P}_1	\bar{P}_2	\dots	\bar{P}_M

Якщо одновимірні д.в.в. X та Y формують двовимірну д.в.в. (X, Y) (або сукупну систему, складовими якої є X та Y), то закон її розподілу задається двовимірною таблицею 10.2.3 ($\tilde{p}_{ij} = P(X = x_i, Y = y_j)$, $i = \overline{1, N}$; $j = \overline{1, M}$).

Таблиця 10.2.3

$X \backslash Y$	y_1	y_2	\vdots	y_M
x_1	\tilde{p}_{11}	\tilde{p}_{12}	\vdots	\tilde{p}_{1M}
x_2	\tilde{p}_{21}	\tilde{p}_{22}	\vdots	\tilde{p}_{2M}
\dots	\dots	\dots	$\dots \vdots \dots$	\dots
x_N	\tilde{p}_{N1}	\tilde{p}_{N2}	\vdots	\tilde{p}_{NM}

Щодо д.в.в. X, Y та (X, Y) зауважимо, що їх реалізації (для систем стани) вважаються несумісними, тому виконуються співвідношення (10.2.1), (10.2.2):

$$\sum_{i=1}^N p_i = 1, \quad \sum_{j=1}^M \bar{p}_j = 1, \quad (10.2.1)$$

$$\sum_{i=1}^N \sum_{j=1}^M \tilde{p}_{ij} = \sum_{i=1}^N \left(\sum_{j=1}^M \tilde{p}_{ij} \right) = \sum_{j=1}^M \left(\sum_{i=1}^N \tilde{p}_{ij} \right) = 1. \quad (10.2.2)$$

Стосовно системи або вектора (X, Y) уточнимо також, що у випадку незалежності X та Y , тобто тоді, коли ймовірність будь-якого значення кожної з цих величин не залежить від того, яке значення набуває інша величина, мають місце рівності

$$\tilde{p}_{ij} = p_i \bar{p}_j, \quad i = \overline{1, N}; \quad j = \overline{1, M}. \quad (10.2.3)$$

Зовсім інша ситуація, коли величини X та Y залежні. Тоді, на відміну від (10.2.3), формула ймовірності добутку двох подій $(X = x_i)$ та $(Y = y_j)$ має вигляд

$$\begin{aligned} \tilde{p}_{ij} &= P(X = x_i) \cdot P(Y = y_j / X = x_i) = \\ &= p_i \cdot p^{**}(y_j / x_i) \end{aligned} \quad (10.2.4)$$

або

$$\begin{aligned} \tilde{p}_{ij} &= P(Y = y_j) \cdot P(X = x_i / Y = y_j) = \\ &= \bar{p}_j \cdot p^*(x_i / y_j), \end{aligned} \quad (10.2.5)$$

де $p^*(x_i / y_j) = P(X = x_i / Y = y_j)$ та $p^{**}(y_j / x_i) = P(Y = y_j / X = x_i)$ – так звані умовні ймовірності. Із

(10.2.4), (10.2.5) впливають співвідношення

$$\tilde{P}_{ij} = P_i P^{**}(y_j / x_i) = \bar{P}_j P^*(x_i / y_j), \quad (10.2.6)$$

а із (10.2.6) – співвідношення

$$P_i = \sum_{j=1}^M \tilde{P}_{ij} = \sum_{j=1}^M \bar{P}_j P^*(x_i / y_j), \quad (10.2.7)$$

$$\bar{P}_j = \sum_{i=1}^N \tilde{P}_{ij} = \sum_{i=1}^N P_i P^{**}(y_j / x_i). \quad (10.2.8)$$

Останні співвідношення (10.2.7), (10.2.8) важливі для з'ясування властивостей ентропії та кількості інформації у випадку, коли системи (в.в.) X , Y залежні.

2) За аналогією з формулою Шеннона (10.1.1) для одновимірної системи нескладно виписати формули для ентропії системи X за умови, що система Y перебуває у стані y_j ($Y = y_j$) та системи Y за умови, що система X перебуває у стані x_i ($X = x_i$). Ці формули мають вигляд:

$$H_{X/y_j} = - \sum_{i=1}^N P^*(x_i / y_j) \log_2 P^*(x_i / y_j), \quad (10.2.9)$$

$$H_{Y/x_i} = - \sum_{j=1}^M P^{**}(y_j / x_i) \log_2 P^{**}(y_j / x_i). \quad (10.2.10)$$

Крім *часткових умовних ентропій* (10.2.9), (10.2.10), знаходять також *повні умовні ентропії*, тобто *ентропію системи X за умови повної визначеності системи Y* або середню невизначеність системи X у випадку визначеності системи Y , та *ентропію системи Y за умови повної визначеності системи X* або середню невизначеність системи Y у випадку визначеності

системи X . Формули для цих ентропій відповідно такі:

$$\begin{aligned} H_{X/Y} &= \sum_{j=1}^M \bar{p}_j H(X/y_j) = \\ &= - \sum_{j=1}^M \sum_{i=1}^N \bar{p}_j p^*(x_i/y_j) \log_2 p^*(x_i/y_j), \end{aligned} \quad (10.2.11)$$

$$\begin{aligned} H_{Y/X} &= \sum_{i=1}^N p_i H(Y/x_i) = \\ &= - \sum_{i=1}^N \sum_{j=1}^M p_i p^{**}(y_j/x_i) \log_2 p^{**}(y_j/x_i). \end{aligned} \quad (10.2.12)$$

Ентропія сукупності системи (X, Y)

$$H_{(X,Y)} = - \sum_{i=1}^N \sum_{j=1}^M \tilde{p}_{ij} \log_2 \tilde{p}_{ij}. \quad (10.2.13)$$

3) Властивості умовних ентропій та ентропії сукупної системи (X, Y) встановлюються з урахуванням їх аналітичних виразів, тобто формул (10.2.9) – (10.2.13).

Ці властивості такі:

- для незалежних систем X і Y ентропія сукупності цих систем дорівнює сумі ентропій кожної з них:

$$H_{(X,Y)} = H_X + H_Y; \quad (10.2.14)$$

- ентропія сукупності двох залежних систем X і Y визначається так:

$$H_{(X,Y)} = H_X + H_{Y/X} = H_Y + H_{X/Y}; \quad (10.2.15)$$

- величина ентропії сукупної системи не змінюється під час перестановки систем-компонент:

$$H_{(X,Y)} = H_{(Y,X)}; \quad (10.2.16)$$

- умовна ентропія завжди не перевищує безумовну ентропію:

$$H_{X/Y} \leq H_X, \quad H_{Y/X} \leq H_Y. \quad (10.2.17)$$

Співвідношення (10.2.14) – (10.2.17), що відображають властивості відповідних ентропій, часто використовуються у практичних завданнях.

4) Оскільки згідно із означенням кількість інформації про систему X , що міститься у системі Y є величиною, на яку зменшиться невизначеність X після повідомлення про Y , то, позначивши цю величину через $I_{X/Y}$, отримаємо формулу для її визначення:

$$I_{X/Y} = H_X - H_{X/Y}. \quad (10.2.18)$$

Зауважимо, що H_X і $H_{X/Y}$ у формулі (10.2.18) ще називають відповідно *апріорною* та *апостеріорною* ентропією. Підставивши у (10.2.18) замість H_X і $H_{X/Y}$ вирази для її знаходження, отримаємо конкретизовану формулу (10.2.19) для $I_{X/Y}$:

$$I_{X/Y} = \sum_{i=1}^N \sum_{j=1}^M \tilde{p}_{ij} \log_2 \left(\frac{\tilde{p}_{ij}}{p_i \bar{p}_j} \right). \quad (10.2.19)$$

Величина $I_{X/Y}$ має такі властивості:

- вона невід'ємна:

$$I_{X/Y} \geq 0; \quad (10.2.20)$$

- інформація про систему X , що міститься у самій системі X , дорівнює ентропії цієї системи:

$$I_{X/X} = H_X; \quad (10.2.21)$$

- якщо $I_{Y/X}$ – кількість інформації про систему Y ,

що міститься у системі X , то:

$$I_{Y/X} = I_{X/Y}; \quad (10.2.22)$$

• кількість інформації $I_{X/Y}$ не перевищує ентропії H_X системи X :

$$I_{X/Y} \leq H_X; \quad (10.2.23)$$

• величина $I_{X/Y}$ дорівнює різниці між сумою безумовних ентропій H_X і H_Y та ентропією $H_{(X,Y)}$:

$$I_{X/Y} = H_X + H_Y - H_{(X,Y)}. \quad (10.2.24)$$

Співвідношення (10.2.20) – (10.2.24) також нерідко використовуються під час виконання практичних завдань.

Завершуючи викладення основного матеріалу щодо виконання даного завдання, підкреслимо, що для більш повнішого розкриття сформульованих підзавдань, зокрема для встановлення тих чи інших властивостей, виведення формул тощо потрібно скористатися відповідними літературними джерелами, зокрема посібником [2]. ►

Завдання 10.3. 1) У таблицях 10.3.1 та 10.3.2 задані безумовні розподіли двох систем X та Y , кожна з яких може перебувати у одному з двох станів ($k \geq 1$). Крім того, відомі умовні ймовірності:

$$p^*(x_1/y_1) = 0,3; \quad p^*(x_2/y_1) = 0,7;$$

$$p^*(x_1/y_2) = 0,6; \quad p^*(x_2/y_2) = 0,4.$$

Потрібно визначити умовну ентропію $H_{X/Y}$ та умовну ентропію $H_{Y/X}$, якщо $p^{**}(y_j/x_i) = p^*(x_j/y_i)$ ($i, j = \overline{1,2}$).

Таблиця 10.3.1

X	x_1	x_2
p	$\frac{k+1}{2k+1}$	$\frac{k}{2k+1}$

Таблиця 10.3.2

Y	y_1	y_2
\bar{p}	$\frac{2k+3}{3k+5}$	$\frac{k+2}{3k+5}$

2) Сумісний розподіл системи (X, Y) задано таблицями 10.3.3 – 10.3.7. Необхідно визначити ентропії $H_{(X,Y)}$, H_X , H_Y , $H_{X/Y}$, $H_{Y/X}$ та кількість інформації $I_{Y/X}$.

Таблиця 10.3.3

$X \backslash Y$	y_1	y_2	y_3	y_4	y_5
x_1	0,09	0,10	0,05	0,07	0,03
x_2	0,04	0,01	0,04	0,06	0,01
x_3	0,04	0,01	0,04	0,06	0,01
x_4	0,05	0,03	0,09	0,07	0,10

Таблиця 10.3.4

$X \backslash Y$	y_1	y_2	y_3	y_4	y_5
x_1	0,03	0,05	0,07	0,09	0,10
x_2	0,01	0,04	0,06	0,04	0,01
x_3	0,01	0,04	0,06	0,04	0,01
x_4	0,10	0,09	0,07	0,05	0,03

Таблиця 10.3.5

$X \backslash Y$	y_1	y_2	y_3	y_4	y_5
x_1	0,05	0,03	0,09	0,10	0,07
x_2	0,04	0,01	0,04	0,01	0,06
x_3	0,04	0,01	0,04	0,01	0,06
x_4	0,09	0,10	0,05	0,03	0,07

Таблиця 10.3.6

$X \backslash Y$	y_1	y_2	y_3	y_4	y_5
x_1	0,10	0,07	0,09	0,03	0,05
x_2	0,01	0,06	0,04	0,01	0,04
x_3	0,01	0,06	0,04	0,01	0,04
x_4	0,03	0,07	0,05	0,10	0,09

Таблиця 10.3.7

$X \backslash Y$	y_1	y_2	y_3	y_4	y_5
x_1	0,07	0,05	0,09	0,03	0,10
x_2	0,06	0,04	0,04	0,01	0,01
x_3	0,06	0,04	0,04	0,01	0,01
x_4	0,07	0,09	0,05	0,10	0,03

◀ **1)** Виконаємо підзавдання для $p_1 = \bar{p}_1 = 0,65$;
 $p_2 = \bar{p}_2 = 0,35$; $p^*(x_1/y_1) = 0,2$; $p^*(x_2/y_1) = 0,8$;
 $p^*(x_1/y_2) = 0,7$; $p^*(x_2/y_2) = 0,3$. Тоді, використовуючи
формули (10.2.11) та (10.2.12), отримаємо:

$$\begin{aligned}
H_{X/Y} &= -\sum_{j=1}^2 \sum_{i=1}^2 \bar{p}_j p^*(x_i/y_j) \log_2 p^*(x_i/y_j) = \\
&= -0,65(0,2 \log_2 0,2 + 0,8 \log_2 0,8) - \\
&\quad - 0,35(0,7 \log_2 0,7 + 0,3 \log_2 0,3) = \\
&= -\frac{1}{\ln 2} [0,65(0,2 \ln 0,2 + 0,8 \ln 0,8) + \\
&\quad + 0,35(0,7 \ln 0,7 + 0,3 \ln 0,3)] = 0,777 \text{ (біт)}; \\
H_{Y/X} &= -\sum_{i=1}^2 \sum_{j=1}^2 p_i p^{**}(y_j/x_i) \log_2 p^{**}(y_j/x_i) = \\
&= -0,65(0,2 \log_2 0,2 + 0,8 \log_2 0,8) - \\
&\quad - 0,35(0,7 \log_2 0,7 + 0,3 \log_2 0,3) =
\end{aligned}$$

$$= H_{X/Y} = 0,777 \text{ (біт)}.$$

2) Для демонстрації виконання цього підзавдання використаємо сумісний розподіл системи (X, Y) , заданий таблицею 10.3.8.

Таблиця 10.3.8

$X \backslash Y$	y_1	y_2	y_3	y_4	y_5
x_1	0,10	0,09	0,07	0,05	0,03
x_2	0,01	0,04	0,06	0,04	0,01
x_3	0,01	0,04	0,06	0,04	0,01
x_4	0,03	0,05	0,07	0,09	0,10

На основі даних таблиці 10.3.8 можна визначити ймовірності

$$p_i = \sum_{j=1}^5 \tilde{p}_{ij} \quad (i = \overline{1,4}), \quad \bar{p}_j = \sum_{i=1}^4 \tilde{p}_{ij} \quad (j = \overline{1,5}),$$

тобто отримати розподіли систем X та Y (таблиці 10.3.9, 10.3.10).

Таблиця 10.3.9

X	x_1	x_2	x_3	x_4
P	0,34	0,16	0,16	0,34

Таблиця 10.3.10

Y	y_1	y_2	y_3	y_4	y_5
\bar{p}	0,15	0,22	0,26	0,22	0,15

Після цього, застосувавши формули (10.2.13) та (10.1.1), отримаємо

$$\begin{aligned}
 H_{(X,Y)} &= -4(0,01 \log_2 0,01 + 0,04 \log_2 0,04) - \\
 &- 2(0,10 \log_2 0,10 + 0,09 \log_2 0,09 + 0,07 \log_2 0,07 + \\
 &+ 0,05 \log_2 0,05 + 0,03 \log_2 0,03 + 0,06 \log_2 0,06) = \\
 &= -\frac{4}{\ln 2}(0,01 \ln 0,01 + 0,04 \ln 0,04) - \\
 &- \frac{2}{\ln 2}(0,10 \ln 0,10 + 0,09 \ln 0,09 + 0,07 \ln 0,07 + \\
 &+ 0,05 \ln 0,05 + 0,03 \ln 0,03 + 0,06 \ln 0,06) = 4,06 \text{ (біт)}; \\
 H_X &= -2(0,34 \log_2 0,34 + 0,16 \log_2 0,16) = \\
 &= \frac{2}{\ln 2}(0,34 \ln 0,34 + 0,16 \ln 0,16) = 1,9 \text{ (біт)}; \\
 H_Y &= -2(0,15 \log_2 0,15 + 0,22 \log_2 0,22) - \\
 &- 0,26 \log_2 0,26 = -\frac{2}{\ln 2}(0,15 \ln 0,15 + 0,22 \ln 0,22) - \\
 &- \frac{1}{\ln 2} \cdot 0,26 \cdot \ln 0,26 = 2,29 \text{ (біт)}.
 \end{aligned}$$

Із формули (10.2.15) знаходимо умовні ентропії

$H_{X/Y}$ і $H_{Y/X}$:

$$H_{X/Y} = H_{(X,Y)} - H_Y = 4,06 - 2,29 = 1,77 \text{ (біт)};$$

$$H_{Y/X} = H_{(X,Y)} - H_X = 4,06 - 1,9 = 2,16 \text{ (біт)}.$$

Останні співвідношення підтверджують виконання нерівностей (10.2.17), тобто $H_{X/Y} \leq H_X$, $H_{Y/X} \leq H_Y$.

Кількість інформації $I_{X/Y}$ знаходимо за формулою (10.2.18):

$$I_{X/Y} = H_X - H_{X/Y} = 1,9 - 1,77 = 0,13 \text{ (біт)}.$$

Оскільки

$$I_{Y/X} = H_Y - H_{Y/X} = 2,29 - 2,16 = 0,13 \text{ (біт)},$$

то $I_{X/Y} = I_{Y/X}$, тобто властивість «симетричності» для міри інформації виконується. Легко перевірити також виконання рівності $I_{X/Y} = H_X + H_Y - H_{(X,Y)}$. ►

ТЕМА 11. КІЛЬКІСНЕ ВИМІРЮВАННЯ НЕВИЗНАЧЕНОСТІ ТА ІНФОРМАЦІЇ У ВИПАДКУ СИСТЕМ ІЗ НЕПЕРЕРВНИМИ МНОЖИНАМИ МОЖЛИВИХ СТАНІВ

Завдання 11.1. *Конкретизувати формули для визначення ентропії одно- та двовимірних систем з неперервними множинами можливих станів і кількості інформації про них. Обґрунтувати їх аналогію з відповідними формулами для систем дискретного типу та наявність тих самих властивостей ентропії і кількості інформації, що характерні для дискретних систем.*

◀ Найперше зазначимо, що у випадку систем з неперервними множинами можливих станів йдеться про ототожнення ймовірнісних розподілів станів цих систем з відповідними ймовірнісними розподілами неперервних випадкових величин (н.в.в.). Отож, при визначенні ентропії чи кількості інформації у цьому випадку під системами X , Y будемо розуміти н.в.в. [2], знання про які дозволяють виписати формули для ентропії та кількості інформації за аналогією з випадком дискретних випадкових величин (д.в.в.). Якщо $p(x)$, $\bar{p}(y)$, $\tilde{p}(x, y)$, $p^*(x/y)$, $p^{**}(y/x)$ – відповідно щільності розподілу н.в.в. X ; н.в.в. Y ; випадкового вектора (сукупної системи) (X, Y) ; н.в.в. X за умови, що

$Y = y$; н.в.в. Y за умови, що $X = x$, а \bar{H}_X , \bar{H}_Y , $\bar{H}_{X/Y}$, $\bar{H}_{Y/X}$, $\bar{I}_{X/Y}$, $\bar{I}_{Y/X}$, $\bar{H}_{(X,Y)}$ – відповідні аналоги величин H_X , H_Y , $H_{X/Y}$, $H_{Y/X}$, $I_{X/Y}$, $I_{Y/X}$, $H_{(X,Y)}$ для д.в.в., то

$$\bar{H}_X = - \int_{-\infty}^{\infty} p(x) \log_2 p(x) dx, \quad (11.1.1)$$

$$\bar{H}_Y = - \int_{-\infty}^{\infty} \bar{p}(y) \log_2 \bar{p}(y) dy, \quad (11.1.2)$$

$$\bar{H}_{X/Y} = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{p}(y) p^*(x/y) \log_2 p^*(x/y) dx dy, \quad (11.1.3)$$

$$\bar{H}_{Y/X} = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x) p^{**}(y/x) \log_2 p^{**}(y/x) dx dy, \quad (11.1.4)$$

$$\bar{I}_{X/Y} = \bar{I}_{Y/X} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{p}(x, y) \log_2 \left(\frac{\tilde{p}(x, y)}{p(x) \bar{p}(y)} \right) dx dy, \quad (11.1.5)$$

$$\bar{H}_{(X,Y)} = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{p}(x, y) \log_2 \tilde{p}(x, y) dx dy. \quad (11.1.6)$$

Усі інтеграли у формулах (11.1.1) – (11.1.6) невластні [2]. Припускається, що вони існують, тому зазначені співвідношення коректні у сенсі їх математичної формалізації та змісту.

Використовуючи формули (11.1.1) – (11.1.6), нескладно переконатися, що для систем неперервного типу також виконуються перелічені у прикладах завдань теми 10 властивості ентропії та кількості інформації. До прикладу, покажемо, що у випадку

залежних систем X та Y ентропія системи (X, Y) може бути обчислена за формулою

$$\bar{H}_{(X,Y)} = \bar{H}_X + \bar{H}_{Y/X} = \bar{H}_Y + \bar{H}_{X/Y}, \quad (11.1.7)$$

яка є аналогом формули (10.2.15) (завдання 10.2).

Доведемо лише першу частину рівності (11.1.7), але спочатку уточнимо, що у випадку н.в.в. для відповідних щільностей розподілу виконуються співвідношення (11.1.8) – (11.1.11), які потрібно знати для встановлення властивостей ентропії та кількості інформації:

$$\int_{-\infty}^{\infty} p(x)dx = 1, \quad \int_{-\infty}^{\infty} \bar{p}(y)dy = 1, \quad (11.1.8)$$

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{p}(x, y) dx dy &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \tilde{p}(x, y) dy \right] dx = \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \tilde{p}(x, y) dx \right] dy = 1, \end{aligned} \quad (11.1.9)$$

$$\tilde{p}(x, y) = p(x)p^{**}(y/x) = \bar{p}(y)p^*(x/y), \quad (11.1.10)$$

$$\int_{-\infty}^{\infty} \tilde{p}(x, y) dy = p(x), \quad \int_{-\infty}^{\infty} \tilde{p}(x, y) dx = \bar{p}(y). \quad (11.1.11)$$

Отже, оскільки

$$\begin{aligned} \int_{-\infty}^{\infty} p^{**}(y/x) dy &= \int_{-\infty}^{\infty} \frac{\tilde{p}(x, y)}{p(x)} dy = \frac{1}{p(x)} \int_{-\infty}^{\infty} \tilde{p}(x, y) dy = \\ &= \frac{1}{p(x)} \cdot p(x) = 1 \end{aligned}$$

і

$$-\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x)p^{**}(y/x) \log_2 p(x) dx dy =$$

$$\begin{aligned}
&= - \int_{-\infty}^{\infty} \left\{ p(x) \log_2 p(x) \left[\int_{-\infty}^{\infty} p^{**}(y/x) dy \right] dx \right\} = \\
&= - \int_{-\infty}^{\infty} p(x) \log_2 p(x) dx = \bar{H}_X,
\end{aligned}$$

то

$$\begin{aligned}
\bar{H}_{(X,Y)} &= - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{p}(x, y) \log_2 \tilde{p}(x, y) dx dy = \\
&= - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x) p^{**}(y/x) \log_2 [p(x) p^{**}(y/x)] dx dy = \\
&= - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x) p^{**}(y/x) [\log_2 p(x) + \log_2 p^{**}(y/x)] dx dy = \\
&= - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x) p^{**}(y/x) \log_2 p(x) dx dy - \\
&\quad - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x) p^{**}(y/x) \log_2 p^{**}(y/x) dx dy = \bar{H}_X + \bar{H}_{Y/X},
\end{aligned}$$

що й потрібно було встановити.

Аналогічно доводиться й друга частина співвідношення (11.1.7). Інші властивості ентропії та кількості інформації пропонується перевірити самостійно. ►

Завдання 11.2. Знайти:

1) ентропію системи X , щільність розподілу ймовірностей станів якої задається функціями (11.2.1) – (11.2.5):

$$p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}; \quad (11.2.1)$$

$$p(x) = \begin{cases} 0, & x \leq 0, \\ e^{-x}, & x > 0; \end{cases} \quad (11.2.2)$$

$$p(x) = \begin{cases} 0, & x \leq 0, \\ 4e^{-4x}, & x > 0; \end{cases} \quad (11.2.3)$$

$$p(x) = \begin{cases} 0, & x \leq 0, \\ 8e^{-8x}, & x > 0; \end{cases} \quad (11.2.4)$$

$$p(x) = \begin{cases} 0, & x \leq 0, \\ 16e^{-16x}, & x > 0. \end{cases} \quad (11.2.5)$$

2) 2.1) ентропію системи (X, Y) , компоненти якої мають нульові математичні сподівання та середні квадратичні відхилення відповідно σ_1 і σ_2 , а сумісний розподіл ймовірностей станів є нормальним і задається щільністю (11.2.6):

$$\tilde{p}(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2} - \frac{2\rho xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)\right], \quad (11.2.6)$$

де ρ – коефіцієнт кореляції X і Y ;

2.2) ентропії систем X та Y , попередньо встановивши щільності розподілу ймовірностей їх станів;

2.3) умовні ентропії кожної із систем-компонент і кількість інформації про кожну з них, що міститься в

іншій системі-компоненті, скориставшись отриманими у 2.1) і 2.2) результатами;

2.4) числові значення зазначених вище ентропій та кількості інформації, якщо $\rho = 0,7 + 0,05k$, $\sigma_1 = 0,5 + 0,25k$, $\sigma_2 = 0,6 + 0,3k$ ($k = \overline{0,4}$).

◀ **1)** Виконання цього завдання потрібно розпочати із перевірки умови нормування для щільності розподілу, тобто умови (11.1.8). Після цього треба застосувати формулу (11.1.1) для визначення ентропії системи з неперервною множиною можливих станів. Процес виконання завдання продемонструємо для випадку, коли

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \quad (11.2.7)$$

де $\sigma > 0$.

Функція (11.2.7) є щільністю нормального розподілу із нульовим математичним сподіванням ($M(X) = 0$) та середнім квадратичним відхиленням σ ($\sigma = \sqrt{D(X)}$, де $D(X)$ - дисперсія). Ввівши змінну $z = \frac{x}{\sigma}$ (тоді $dz = \frac{dx}{\sigma}$), для функції (11.2.7) отримаємо співвідношення

$$\begin{aligned} \int_{-\infty}^{\infty} p(x) dx &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} \cdot \sigma dz = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = 1, \end{aligned} \quad (11.2.8)$$

у якому враховано інтеграл Пуассона [2], тобто рівність

$$\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz = \sqrt{2\pi}.$$

Отже, умова нормування для щільності (11.2.7) виконується. Далі, враховуючи (11.2.8) і той факт, що для (11.2.7)

$$\begin{aligned} \sigma^2 &= D(X) = M(X^2) - (M(X))^2 = \\ &= \int_{-\infty}^{\infty} x^2 p(x) dx = \int_{-\infty}^{\infty} x^2 \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx, \end{aligned} \quad (11.2.9)$$

знаходимо шукану ентропію:

$$\begin{aligned} H_X &= - \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \log_2 \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \right] dx = \\ &= - \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} \left[\log_2 \frac{1}{\sigma\sqrt{2\pi}} + \log_2 e^{-\frac{x^2}{2\sigma^2}} \right] dx = \\ &= - \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} \left[-\log_2(\sigma\sqrt{2\pi}) - \frac{x^2}{2\sigma^2} \log_2 e \right] dx = \\ &= \log_2(\sigma\sqrt{2\pi}) \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx + \\ &+ \frac{\log_2 e}{2\sigma^2} \int_{-\infty}^{\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx = \\ &= \log_2(\sigma\sqrt{2\pi}) + \frac{\log_2 e}{2\sigma^2} \sigma^2 = \log_2(\sigma\sqrt{2\pi}) + \frac{\log_2 e}{2} = \end{aligned}$$

$$= \log_2(\sigma\sqrt{2\pi}) + \log_2\sqrt{e} = \log_2(\sigma\sqrt{2\pi e}) \text{ (біт)}. \quad (11.2.10)$$

Зауважимо, що співвідношення (11.2.9), яке використано у (11.2.10), фактично наперед задане для функції (11.2.7), але його можна перепроверити. Справді, зробивши у інтегралі (11.2.9) заміну $z = x/\sigma$ (тоді $dx = \sigma dz$, $x^2 = \sigma^2 z^2$) і застосувавши потім формулу інтегрування за частинами ($u = z$, $du = dz$, $dv = ze^{-z^2/2} dz$, $v = -e^{-z^2/2}$) матимемо:

$$\begin{aligned} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2\sigma^2}} dx &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma^2 z^2 e^{-\frac{z^2}{2}} \cdot \sigma dz = \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{z^2}{2}} dz = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z \cdot z e^{-\frac{z^2}{2}} dz = \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \left[-z e^{-\frac{z^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz \right] = \frac{\sigma^2}{\sqrt{2\pi}} [0 + \sqrt{2\pi}] = \sigma^2. \end{aligned}$$

Для інших функцій $p(x)$ завдання виконується аналогічно.

2) 2.1) Оскільки

$$\begin{aligned} \log_2 \tilde{p}(x, y) &= -\log_2 e \left[\frac{1}{2(1-\rho^2)} \left(\frac{x^2}{\sigma_1^2} - \frac{2\rho xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2} \right) \right] - \\ &-\log_2 \left(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2} \right), \end{aligned}$$

то, використавши формулу (11.1.6), отримаємо:

$$\begin{aligned}
\bar{H}_{(X,Y)} &= - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{p}(x,y) \log_2 \tilde{p}(x,y) dx dy = \\
&= \log_2 \left(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2} \right) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{p}(x,y) dx dy + \\
&+ \frac{\log_2 e}{2(1-\rho^2)\sigma_1^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 \tilde{p}(x,y) dx dy - \\
&- \frac{\rho \log_2 e}{(1-\rho^2)\sigma_1\sigma_2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \tilde{p}(x,y) dx dy + \\
&+ \frac{\log_2 e}{2(1-\rho^2)\sigma_2^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y^2 \tilde{p}(x,y) dx dy.
\end{aligned}$$

Далі, враховуючи співвідношення (11.1.9), (11.1.11) та співвідношення для коефіцієнта кореляції ρ між випадковими величинами X і Y ($\text{cov}(X,Y)$ – коваріація величин X і Y), тобто формулу

$$\begin{aligned}
\rho &= \frac{\text{cov}(X,Y)}{\sigma_1\sigma_2} = \\
&= \frac{1}{\sigma_1\sigma_2} \int_{-\infty}^{\infty} [x - M(X)] \left\{ \int_{-\infty}^{\infty} (y - M(Y)) \tilde{\rho}(x,y) dy \right\} dx = \\
&= \frac{1}{\sigma_1\sigma_2} \int_{-\infty}^{\infty} [x - 0] \left\{ \int_{-\infty}^{\infty} (y - 0) \tilde{\rho}(x,y) dy \right\} dx = \\
&= \frac{1}{\sigma_1\sigma_2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \tilde{\rho}(x,y) dx dy, \tag{11.2.11}
\end{aligned}$$

переконаємося, що

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{p}(x,y) dx dy = 1,$$

$$\begin{aligned}
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 \tilde{p}(x, y) dx dy &= \int_{-\infty}^{\infty} x^2 \left[\int_{-\infty}^{\infty} \tilde{p}(x, y) dy \right] dx = \\
&= \int_{-\infty}^{\infty} x^2 p(x) dx = \sigma_1^2, \\
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y^2 \tilde{p}(x, y) dx dy &= \int_{-\infty}^{\infty} y^2 \left[\int_{-\infty}^{\infty} \tilde{p}(x, y) dx \right] dy = \\
&= \int_{-\infty}^{\infty} y^2 \bar{p}(y) dy = \sigma_2^2, \\
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \tilde{p}(x, y) dx dy &= \rho \sigma_1 \sigma_2,
\end{aligned}$$

тому

$$\begin{aligned}
\bar{H}_{(X,Y)} &= \log_2 \left(2\pi\sigma_1\sigma_2\sqrt{1-\rho^2} \right) + \frac{\log_2 e}{2(1-\rho^2)} + \\
&+ \frac{\log_2 e}{2(1-\rho^2)} - \frac{\rho^2 \log_2 e}{1-\rho^2} = \log_2 \left(2\pi e\sigma_1\sigma_2\sqrt{1-\rho^2} \right).
\end{aligned}$$

Попутно зауважимо, що із (11.2.11) можна знайти коваріацію $\text{cov}(X, Y)$, що є елементом $\sigma_{12} = \sigma_{21}$ дисперсійної матриці двовимірної випадкової величини (X, Y) . Крім того, вище у формулах для дисперсій σ_1^2 і σ_2^2 враховано, що для розподілу (11.2.6) $M(X) = 0$ і $M(Y) = 0$.

2.2) Після обчислення інтегралів (11.1.11) отримаємо вже відомі нам функції

$$p(x) = \frac{1}{\sigma_1\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma_1^2}}, \quad \bar{p}(y) = \frac{1}{\sigma_2\sqrt{2\pi}} e^{\frac{-y^2}{2\sigma_2^2}}.$$

Переконаємось у цьому. Отже, з урахуванням заміни

$$z_1 = \frac{x}{\sigma_1}, \quad z_2 = \frac{y}{\sigma_2}, \quad dz_2 = \frac{dy}{\sigma_2} \text{ маємо:}$$

$$\begin{aligned} p(x) &= \int_{-\infty}^{\infty} \tilde{p}(x, y) dy = \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left\{-\frac{\frac{x^2}{\sigma_1^2} - \frac{2\rho xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}}{2(1-\rho^2)}\right\} dy = \\ &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left\{-\frac{z_1^2 - 2\rho z_1 z_2 + z_2^2}{2(1-\rho^2)}\right\} \sigma_2^2 dz_2 = \\ &= \frac{1}{\sigma_1\sqrt{2\pi}} \cdot \frac{\sigma_2}{\sigma_2\sqrt{2\pi(1-\rho^2)}} \times \\ &\times \int_{-\infty}^{\infty} \exp\left\{-\frac{z_2^2 - 2\rho z_1 z_2 + \rho^2 z_1^2 + z_1^2 - \rho^2 z_1^2}{2(1-\rho^2)}\right\} dz_2 = \\ &= \frac{1}{\sigma_1\sqrt{2\pi}} \cdot \frac{1}{\sqrt{2\pi(1-\rho^2)}} \times \\ &\times \int_{-\infty}^{\infty} \exp\left\{-\frac{(z_2 - \rho z_1)^2}{2(1-\rho^2)}\right\} \cdot \exp\left\{\frac{(1-\rho^2)z_1^2}{2(1-\rho^2)}\right\} dz_2 = \\ &= \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{z_1^2}{2}} \cdot \frac{1}{\sqrt{2\pi(1-\rho^2)}} \cdot \int_{-\infty}^{\infty} \exp\left\{\frac{(z_2 - \rho z_1)^2}{2(1-\rho^2)}\right\} dz_2 = \\ &= \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{z_1^2}{2}} \cdot 1 = \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_1^2}}. \end{aligned} \tag{11.2.12}$$

Аналогічно знаходиться $\bar{p}(y)$. Зауважимо лише, що у (11.2.12)

$$\frac{1}{\sqrt{2\pi(1-\rho^2)}} \int_{-\infty}^{\infty} \exp\left\{-\frac{(z_2 - \rho z_1)^2}{2(1-\rho^2)}\right\} dz_2 = 1$$

тому, що це інтеграл від нормально розподіленої одновимірної випадкової величини Z_2 із математичним сподіванням $M(Z_2) = \rho z_1$ і дисперсією $\sigma^2 = 1 - \rho^2$.

Що стосується ентропій, то із (11.2.10) випливає, що $H_X = \log_2(\sigma_1 \sqrt{2\pi e})$, $H_Y = \log_2(\sigma_2 \sqrt{2\pi e})$.

2.3) Тепер, скориставшись попередніми результатами й формулою (11.1.7), визначаємо умовні ентропії $\bar{H}_{Y/X}$, $\bar{H}_{X/Y}$:

$$\bar{H}_{Y/X} = \bar{H}_{(X,Y)} - \bar{H}_X = \log_2\left(2\pi e \sigma_1 \sigma_2 \sqrt{1-\rho^2}\right) -$$

$$-\log_2\left(\sigma_1 \sqrt{2\pi e}\right) = \log_2\left(\sigma_2 \sqrt{2\pi e(1-\rho^2)}\right);$$

$$\bar{H}_{X/Y} = \bar{H}_{(X,Y)} - \bar{H}_Y = \log_2\left(\sigma_1 \sqrt{2\pi e(1-\rho^2)}\right).$$

Згідно із формулою (10.2.18) для кількості інформації отримаємо

$$\bar{I}_{X/Y} = \bar{H}_X - \bar{H}_{X/Y} = \bar{I}_{Y/X} = \bar{H}_Y - \bar{H}_{Y/X} =$$

$$= \bar{H}_X + \bar{H}_Y - \bar{H}_{(X,Y)} = -\log_2 \sqrt{1-\rho^2}.$$

2.4) Виконання цього завдання є елементарним. До прикладу, якщо $k=0$ ($\rho=0,7$; $\sigma_1=0,5$; $\sigma_2=0,6$), то $H_X = 1,05$ (біт) і т.д. ►

ТЕМА 12. ВЕРТИКАЛЬНА ДЕКОМПОЗИЦІЯ ЛІНІЙНИХ МОДЕЛЕЙ ОПТИМАЛЬНОГО ПЛАНУВАННЯ ВИРОБНИЦТВА МЕТОДОМ ДАНЦИГА-ВУЛФА

Завдання 12.1.

1) *Формалізувати лінійну модель оптимального планування випуску продукції виробничою системою, яка складається з декілька локальних підсистем, кожна з яких має свої окремі цілі.*

2) *Для цієї моделі описати теоретичні основи методу декомпозиції Данцига-Вулфа, який дозволяє за скінченну кількість ітерацій знайти розв'язок моделі.*

◀ 1) Будемо вважати, що виробнича система складається з M підсистем (наприклад M підприємств), кожна з яких використовує як свої локальні ресурси, так і деякий глобальний ресурс, спільний для усіх підсистем. Для кожної k -тої підсистеми ($k = \overline{1, M}$) введемо позначення:

$x^{(k)}$ – вектор інтенсивності виробництва (вираз « $x_i^{(k)}$ – інтенсивність технологічного способу виробництва i », який використовується в економіко-математичному моделюванні, означає, що $x_i^{(k)}$ – це кількість продукції, яка випускається даним виробничим об'єктом, наприклад за один рік, при використанні технології i , тобто під інтенсивністю

технологічного способу слід розуміти величину, що характеризує міру його використання у плані);

$c^{(k)}$ – вектор локальних доходів;

$A^{(k)}$ – локальна матриця питомих витрат глобальних ресурсів;

b – вектор глобальних ресурсів;

$D^{(k)}$ – локальна матриця питомих витрат локальних ресурсів;

$d^{(k)}$ – вектор локальних ресурсів,

де $x^{(k)} \in \mathbb{R}^{n_k}$, $d^{(k)} \in \mathbb{R}^{m_k}$, $b \in \mathbb{R}^m$, $c^{(k)} \in \mathbb{R}^{n_k}$, $D^{(k)}$ – матриця розмірності $(m_k \times n_k)$, $A^{(k)}$ – матриця розмірності $(m \times n_k)$ (\mathbb{R}^l – l -вимірний точковий простір).

Тоді лінійна модель оптимального планування випуску продукції такою виробничою системою формалізується задачею лінійного програмування

$$\left\{ \begin{array}{l} \sum_{k=1}^M \langle c^{(k)}, x^{(k)} \rangle \rightarrow \max, \end{array} \right. \quad (12.1.1)$$

$$\left\{ \begin{array}{l} \sum_{k=1}^M A^{(k)} x^{(k)} = b, \end{array} \right. \quad (12.1.2)$$

$$\left\{ \begin{array}{l} D^{(k)} x^{(k)} = d^{(k)}, \quad k = \overline{1, M}, \end{array} \right. \quad (12.1.3)$$

$$\left\{ \begin{array}{l} x^{(k)} \geq 0, \quad k = \overline{1, M}, \end{array} \right. \quad (12.1.4)$$

де $0 \in \mathbb{R}^{n_k}$, $\langle \cdot, \cdot \rangle$ – операція скалярного добутку.

Легко побачити, що матриця задачі (12.1.1) – (12.1.4) (або матриця співвідношень (12.1.2) – (12.1.4) має блочно-діагональну структуру (рис. 12.1.1).

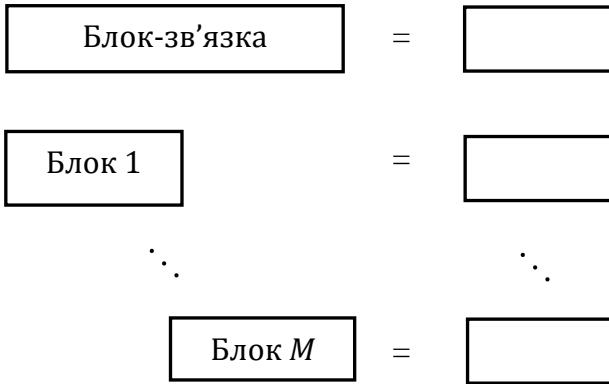


Рис. 12.1.1. Блочно-діагональна структура

Блок-зв'язка включає обмеження загального вигляду, тобто обмеження для усіх змінних задачі. Крім блоку-зв'язки, є також блоки локальних обмежень (блок 1, ..., блок M), що містять змінні, які характеризують локальні підсистеми. Якби не було блока-зв'язки (12.1.2), то задача (12.1.1), (12.1.3), (12.1.4) є сукупністю M самостійних, незалежних між собою підзадач, розв'язки яких у цілому формують розв'язок задачі (12.1.1), (12.1.3), (12.1.4). Наявність обмежень (12.1.2) не дозволяє це зробити. Тому можна виділити дві підзадачі: задачу центра на основі обмежень (12.1.2) і периферійну або локальну задачу на основі обмежень (12.1.3), (12.1.4).

2) Деталізуємо метод декомпозиції моделі (задачі) (12.1.1) – (12.1.4) на підзадачі, який прийнято називати методом Данцига-Вулфа. Нехай для кожного $k = \overline{1, M}$ відомі вектори $x^{(k,1)}, x^{(k,2)}, \dots, x^{(k,l_k)}$, які задовольняють

умови (12.1.3), (12.1.4). З теорії опуклого аналізу випливає, що довільна опукла комбінація

$$x^{(k)} = \sum_{l=1}^{l_k} \alpha^{(k,l)} x^{(k,l)}, \quad (12.1.5)$$

для якої $\sum_{l=1}^{l_k} \alpha^{(k,l)} = 1$ і $\alpha^{(k,l)} \geq 0$, $l = \overline{1, l_k}$, $k = \overline{1, M}$, також

задовольняє при кожному k обмеження (12.1.3), (12.1.4). Отже, це дозволяє сформулювати задачу центра як задачу знаходження таких опуклих комбінацій (12.1.5) (тобто знаходження таких коефіцієнтів $\alpha^{(k,l)}$, $l = \overline{1, l_k}$, $k = \overline{1, M}$), які задовольняють глобальні обмеження (12.1.2) й максимізують цільову функцію (12.1.1).

Тому задача центра має вигляд:

$$\left\{ \begin{array}{l} \sum_{k=1}^M \sum_{l=1}^{l_k} \alpha^{(k,l)} \langle c^{(k)}, x^{(k,l)} \rangle \rightarrow \max, \end{array} \right. \quad (12.1.6)$$

$$\left\{ \begin{array}{l} \sum_{k=1}^M \sum_{l=1}^{l_k} \alpha^{(k,l)} (A^{(k)} x^{(k,l)}) = b, \end{array} \right. \quad (12.1.7)$$

$$\left\{ \begin{array}{l} \sum_{l=1}^{l_k} \alpha^{(k,l)} = 1, \quad k = \overline{1, M}, \end{array} \right. \quad (12.1.8)$$

$$\left\{ \begin{array}{l} \alpha^{(k,l)} \geq 0, \quad l = \overline{1, l_k}, \quad k = \overline{1, M}. \end{array} \right. \quad (12.1.9)$$

Очевидно, що задача (12.1.6) – (12.1.9) є канонічною задачею лінійного програмування відносно змінних $\alpha^{(k,l)}$, $l = \overline{1, l_k}$, $k = \overline{1, M}$, але вона є задачею меншої розмірності порівняно із задачею (12.1.1) – (12.1.4).

Локальна задача k -ої підсистеми така:

$$\left\{ \begin{array}{l} \langle c^{(k)} - A^{(k)} \lambda, x^{(k)} \rangle \rightarrow \max, \\ D^{(k)} x^{(k)} = d^{(k)}, \\ x^{(k)} \geq 0. \end{array} \right. \quad (12.1.10)$$

$$\left\{ \begin{array}{l} D^{(k)} x^{(k)} = d^{(k)}, \\ x^{(k)} \geq 0. \end{array} \right. \quad (12.1.11)$$

$$\left\{ \begin{array}{l} x^{(k)} \geq 0. \end{array} \right. \quad (12.1.12)$$

У задачі (12.1.10) – (12.1.12) $\lambda \in \mathbb{R}^m$ – вектор оцінок глобальних ресурсів. За допомогою цього вектора центр керує плануванням діяльності локальних підсистем. Цільова функція (12.1.10) фактично формалізує прибуток, який одержить k -та підсистема (від доходу $\langle c^{(k)}, x^{(k)} \rangle$ віднімається вартість витрат на глобальні ресурси $\langle A^{(k)} \lambda, x^{(k)} \rangle$). Це означає, що центр стимулює вибір таких локальних планів, які опосередковано враховують дефіцитність глобальних ресурсів, незважаючи на те, що обмежень на них у локальних задачах немає.

Очевидно, що вектор λ оцінок глобальних ресурсів можна знайти на кожній ітерації, розв'язавши двоїсту до (12.1.6) – (12.1.9) задачу. Критерієм оптимальності поточного допустимого плану (або опорного чи базисного невід'ємного розв'язку) вихідної задачі (12.1.1) – (12.1.4) у випадку розв'язування її симплекс-методом, буде невід'ємність оцінок заміщення σ_j для всіх змінних задачі (у задачах на мінімум ці оцінки мають бути недодатними). Використання цього критерію у випадку розкладу задачі (12.1.1) – (12.1.4) на задачу центра (12.1.6) – (12.1.9) і локальні задачі (12.1.10) – (12.1.12) дозволяє сформулювати його так:

поточний опорний план задачі центра буде оптимальним, якщо при розв'язуванні локальних задач із вектором оцінок λ , що відповідає даному плану, для кожного $k = \overline{1, M}$ оптимальне значення цільової функції (12.1.10) збігається з оцінкою λ_{m+k} обмеження (12.1.8) у задачі центра.

Отже, метод Данцига – Вулфа дозволяє за скінченну кількість ітерацій одержати поточний розв'язок вихідної задачі. ►

Завдання 12.2. Продемонструвати процес вертикальної декомпозиції моделі та знаходження її розв'язку за допомогою методу Данцига – Вулфа на конкретному прикладі (k може набувати значень від 1 до 5):

$$\begin{cases} (1+k)x_1 + (2+k)x_2 + (3+k)x_3 + kx_4 \rightarrow \max, \\ 3x_1 + x_2 + 2x_3 + 2x_4 = 100, \\ 2x_1 + x_2 = 40, \\ x_3 + 2x_4 = 40, \\ x_j \geq 0, j = \overline{1, 4}. \end{cases}$$

◄ Виконаємо завдання у випадку, коли $k = 1$, тобто проілюструємо процес декомпозиції та знаходження розв'язку моделі (12.2.1) – (12.2.5):

$$\begin{cases} 2x_1 + 3x_2 + 4x_3 + x_4 \rightarrow \max, & (12.2.1) \\ 3x_1 + x_2 + 2x_3 + 2x_4 = 100, & (12.2.2) \\ 2x_1 + x_2 = 40, & (12.2.3) \\ x_3 + 2x_4 = 40, & (12.2.4) \\ x_j \geq 0, j = \overline{1,4}. & (12.2.5) \end{cases}$$

Якщо застосувати до матриці даної моделі (надалі задачі) так звані симплекс-перетворення, то ми від задачі (12.2.1) – (12.2.5) перейдемо до еквівалентної задачі

$$\begin{cases} 2x_1 + 3x_2 + 4x_3 + x_4 \rightarrow \max, \\ -\frac{1}{2}x_2 + x_3 = 0, \\ x_1 + \frac{1}{2}x_2 = 20, \\ \frac{1}{4}x_2 + x_4 = 20, \\ x_j \geq 0, j = \overline{1,4}. \end{cases} \quad (12.2.6)$$

Базисними змінними у задачі (12.2.6) є змінні x_1, x_3, x_4 . Розв'язавши (12.2.6) симплекс-методом, одержимо такий розв'язок: $x_1^* = 0, x_2^* = 40, x_3^* = 20, x_4^* = 10, f_{\max}^* = 210$.

Тепер розв'яжемо задачу (12.2.1) – (12.2.5) вище зазначеним методом декомпозиції Данцига–Вулфа. Порівнявши (12.2.1)–(12.2.5) з (12.1.1)–(12.1.4), отримаємо

$M = 2$, $x^{(1)} = (x_1, x_2)^T$, $x^{(2)} = (x_3, x_4)^T$,
 $c^{(1)} = (2, 3)^T$, $c^{(2)} = (4, 1)^T$, $A^{(1)} = (3, 1)$, $A^{(2)} = (2, 2)$,
 $D^{(1)} = (2, 1)$, $D^{(2)} = (1, 2)$, $b = 100$, $d^{(1)} = 40$, $d^{(2)} = 40$,
 де T – транспонування.

Перша ітерація. Враховуючи, що задача центра будеється за умови, що сукупність локальних планів, які задовольняють умови (12.2.3) – (12.2.5) уже відома, на першому кроці можна вибрати

$$x^{(1,1)} = (x_1, x_2)^T = (20, 0)^T, \quad x^{(2,1)} = (x_3, x_4)^T = (0, 20)^T.$$

Це означає, що для обох підсистем ($k = \overline{1, 2}$) опуклі комбінації (12.1.5) будуть складатися поки що лише з одного вектора ($\alpha^{(1,1)} = 1$, $\alpha^{(2,1)} = 1$):

$$x^{(1)} = \alpha^{(1,1)} x^{(1,1)} = (20, 0)^T,$$

$$x^{(2)} = \alpha^{(2,1)} x^{(2,1)} = (0, 20)^T.$$

Задача центра (задача (12.1.6)–(12.1.9)) формалізується так:

$$\begin{cases} 40\alpha^{(1,1)} + 20\alpha^{(2,1)} \rightarrow \max, & (12.2.7) \end{cases}$$

$$\begin{cases} 60\alpha^{(1,1)} + 40\alpha^{(2,1)} = 100, & (12.2.8) \end{cases}$$

$$\begin{cases} \alpha^{(1,1)} = 1, & (12.2.9) \end{cases}$$

$$\begin{cases} \alpha^{(2,1)} = 1. & (12.2.10) \end{cases}$$

Двоїстою задачею до (12.2.7) – (12.2.10) буде задача

$$\begin{cases} 100\lambda_1 + \lambda_2 + \lambda_3 \rightarrow \min, \\ 60\lambda_1 + \lambda_2 = 40, \\ 40\lambda_1 + \lambda_3 = 20. \end{cases} \quad (12.2.11)$$

Задача (12.2.11) має множину (безліч) розв'язків, зокрема $\lambda_1^* = 0, \lambda_2^* = 40, \lambda_3^* = 20$, де λ_1^* – оцінка обмеження (12.2.8), λ_2^* – оцінка обмеження (12.2.9), λ_3^* – оцінка обмеження (12.2.10). З урахуванням оцінки λ_1^* (вона є оцінкою глобального ресурсу) побудуємо локальні задачі (12.1.10) – (12.1.12):

для першої підсистеми –

$$\begin{cases} 2x_1 + 3x_2 \rightarrow \max, \\ 2x_1 + x_2 = 40, \\ x_1 \geq 0, x_2 \geq 0; \end{cases} \quad (12.2.12)$$

для другої підсистеми –

$$\begin{cases} 4x_3 + x_4 \rightarrow \max, \\ x_3 + 2x_4 = 40, \\ x_3 \geq 0, x_4 \geq 0. \end{cases} \quad (12.2.13)$$

Задачі (12.2.12) і (12.2.13) відповідно мають такі розв'язки: $x_1^* = 0, x_2^* = 40, \bar{f}^* = 120$ і $x_3^* = 40, x_4^* = 0, \bar{f}^* = 160$. Враховуючи, що $\bar{f}^* \neq \lambda_2^*$ і $\bar{f}^* \neq \lambda_3^*$, робимо висновок, що поточний план задачі центра (12.2.7) – (12.2.10) не є оптимальним. У зв'язку з цим переходимо до другої ітерації.

Друга ітерація. Тепер для побудови опуклих комбінацій для першої підсистеми візьмемо вектори $x^{(1,1)} = (20, 0)^T$ і $x^{(1,2)} = (0, 40)^T$, а для другої підсистеми – вектори $x^{(2,1)} = (0, 20)^T$ і $x^{(2,2)} = (40, 0)^T$, тобто

$$x^{(1)} = \alpha^{(1,1)} x^{(1,1)} + \alpha^{(1,2)} x^{(1,2)} = (20\alpha^{(1,1)}, 40\alpha^{(1,2)}),$$

$$x^{(2)} = \alpha^{(2,1)} x^{(2,1)} + \alpha^{(2,2)} x^{(2,2)} = (40\alpha^{(2,2)}, 20\alpha^{(2,1)}).$$

Тоді задача центра матиме вигляд:

$$\begin{cases} 40\alpha^{(1,1)} + 120\alpha^{(1,2)} + 20\alpha^{(2,1)} + 160\alpha^{(2,2)} \rightarrow \max, & (12.2.14) \end{cases}$$

$$\begin{cases} 60\alpha^{(1,1)} + 40\alpha^{(1,2)} + 40\alpha^{(2,1)} + 80\alpha^{(2,2)} = 100, & (12.2.15) \end{cases}$$

$$\begin{cases} \alpha^{(1,1)} + \alpha^{(1,2)} = 1, & (12.2.16) \end{cases}$$

$$\begin{cases} \alpha^{(2,1)} + \alpha^{(2,2)} = 1, & (12.2.17) \end{cases}$$

$$\begin{cases} \alpha^{(k,l)} \geq 0, \quad l = \overline{1,2}, \quad k = \overline{1,2}. & (12.2.18) \end{cases}$$

Очевидно, що розв'язком задачі (12.2.14) – (12.2.18)

будуть величини $\alpha^{(1,1)} = 0$, $\alpha^{(1,2)} = 1$, $\alpha^{(2,1)} = \frac{1}{2}$, $\alpha^{(2,2)} = \frac{1}{2}$.

Значення цільової функції при цьому дорівнює

$$f_{\max}^* = 210, \quad \text{а} \quad x^* = (20\alpha^{(1,1)}, 40\alpha^{(1,2)}, 40\alpha^{(2,1)}, 20\alpha^{(2,2)})^T \equiv$$

$$\equiv (0, 40, 20, 10)^T. \text{ Отже, вихідна задача (12.2.1) – (12.2.5)}$$

розв'язана. Для того, щоб з'ясувати зміст методу Данцига – Вулфа, знову побудуємо до задачі (12.2.14) – (12.2.18) двоїсту задачу, тобто задачу:

$$\begin{cases} 100\lambda_1 + \lambda_2 + \lambda_3 \mapsto \min, \\ 60\lambda_1 + \lambda_2 \geq 40, \\ 40\lambda_1 + \lambda_2 \geq 120, \\ 40\lambda_1 + \lambda_3 \geq 20, \\ 80\lambda_1 + \lambda_3 \geq 160. \end{cases} \quad (12.2.19)$$

Щоб розв'язати задачу (12.2.19) симплекс-методом, необхідно перейти до еквівалентної канонічної задачі з невід'ємними змінними (змінні λ_1 , λ_2 , λ_3 , які можуть

набувати довільних значень, замінити різницею невід'ємних змінних). Розв'язавши задачу (12.2.19) одержимо розв'язок: $\lambda_1^* = 3,5$; $\lambda_2^* = -20$; $\lambda_3^* = -120$; $\tilde{f}_{\min}^* = 210$.

Потім, врахувавши оцінку λ_1^* глобального ресурсу, формуємо нові локальні задачі:

для першої підсистеми –

$$\begin{cases} -8,5x_1 - 0,5x_2 \mapsto \max, \\ 2x_1 + x_2 = 40, \\ x_1 \geq 0, x_2 \geq 0; \end{cases} \quad (12.2.20)$$

для другої підсистеми –

$$\begin{cases} -3x_3 - 6x_4 \mapsto \max, \\ x_3 + 2x_4 = 40, \\ x_3 \geq 0, x_4 \geq 0. \end{cases} \quad (12.2.21)$$

Розв'язки задач (12.2.20), (12.2.21) такі:

$$x_1^* = 0, \quad x_2^* = 40, \quad x_3^* = 40(1 - \mu), \quad x_4^* = 20\mu, \quad \bar{f}^* = -20, \\ \overline{\overline{f}}^* = -120 \quad (\mu \in [0, 1]).$$

Враховуючи, що $\bar{f}^* = \lambda_2^*$ і $\overline{\overline{f}}^* = \lambda_3^*$, стверджуємо, що знайдений план центра оптимальний. Це означає, що процес розв'язування задачі (12.2.1) – (12.2.5) методом Данцига–Вулфа завершений. ►

ТЕМА 13. ВЕРТИКАЛЬНА ДЕКОМПОЗИЦІЯ ЛІНІЙНИХ МОДЕЛЕЙ ОПТИМАЛЬНОГО ПЛАНУВАННЯ ДІЯЛЬНОСТІ ВИРОБНИЧОЇ СИСТЕМИ МЕТОДОМ КОРНАЇ-ЛІПТАКА

Завдання 13.1. *На основі моделі оптимального планування випуску продукції виробничою системою, яка складається із сукупності локальних підсистем, розкрити теоретичні аспекти схеми вертикальної декомпозиції Корнаї-Ліптака.*

◀ Згідно із методом декомпозиції Корнаї-Ліптака, який реалізує схему безпосереднього прямого розподілу ресурсів у дворівневій системі (у методі Данцига-Вулфа розподіл ресурсу є опосередкованим і здійснюється через ціни), зв'язок центра та локальних підсистем полягає у тому, що центр розподіляє глобальні ресурси між локальними підсистемами у натуральних величинах, отримуючи після цього оцінки ефективності зазначених ресурсів. Що стосується декомпозиції Данцига-Вулфа, то у цьому випадку навпаки, центр повідомляє підсистемам значення оцінок глобальних ресурсів, після чого отримує відповідні плани підсистем, показники прибутків і потреби глобального ресурсу, які слугують основою для формування задачі центра.

Уточнимо зміст схеми Корнаї – Ліптака без її строгого математичного обґрунтування. Будемо

вважати, що центр виділив k -ій підсистемі вектор $b^{(k)} \in \mathbb{R}^m$ ($k = \overline{1, M}$) глобальних ресурсів, причому, як і раніше (завдання 12.1) будемо вважати, що:

M – кількість підсистем (зокрема M підприємств), кожна з яких використовує як свої локальні, так і деякі глобальні ресурси, спільні для усіх підсистем.

Для кожної k -тої підсистеми ($k = \overline{1, M}$) введемо позначення:

$x^{(k)}$ – вектор інтенсивності виробництва ($x_i^{(k)}$ – це кількість продукції, яка випускається даною виробничою підсистемою (данім виробничим об'єктом), наприклад за один рік, при використанні технології i);

$c^{(k)}$ – вектор локальних доходів;

$A^{(k)}$ – локальна матриця питомих витрат глобальних ресурсів;

b – вектор глобальних ресурсів;

$D^{(k)}$ – локальна матриця питомих витрат локальних ресурсів;

$d^{(k)}$ – вектор локальних ресурсів,

де $x^{(k)} \in \mathbb{R}^{n_k}$, $d^{(k)} \in \mathbb{R}^{m_k}$, $b \in \mathbb{R}^m$, $c^{(k)} \in \mathbb{R}^{n_k}$, $D^{(k)}$ – матриця розмірності $(m_k \times n_k)$, $A^{(k)}$ – матриця розмірності $(m \times n_k)$ (\mathbb{R}^l – l -вимірний точковий простір).

Тоді локальна задача для k -тої підсистеми має вигляд:

$$\begin{cases} \langle c^{(k)}, x^{(k)} \rangle \mapsto \max, & (13.1.1) \\ A^{(k)} x^{(k)} = b^{(k)}, & (13.1.2) \\ D^{(k)} x^{(k)} = d^{(k)}, & (13.1.3) \\ x^{(k)} \geq 0, & (13.1.4) \end{cases}$$

де $0 \in \mathbb{R}^{n_k}$, $\langle \cdot, \cdot \rangle$ – операція скалярного добутку.

Розв'язавши двоїсту до (13.1.1) – (13.1.4) задачу, отримаємо зокрема, вектор об'єктивно обумовлених оцінок для обмежень (13.1.2), які власне і є оцінками ефективності розподілу зазначених глобальних ресурсів. Достатніми умовами оптимальності розподілених ресурсів будуть умови [2]

$$\lambda_i^{(k_1)} = \lambda_i^{(k_2)}, \quad k_1, k_2 = \overline{1, M}, \quad i = \overline{1, m}. \quad (13.1.5)$$

Для перевірки умов (13.1.5) необхідно до кожної з M локальних задач побудувати двоїсту, а потім розв'язати її. Підкреслимо також, що необхідні та достатні умови є складнішими.

У тому випадку, коли умови оптимальності не виконуються, необхідно здійснити перерозподіл ресурсів, розв'язавши задачу центра

$$\begin{cases} \sum_{k=1}^M \langle b^{(k)}, \lambda^{(k)} \rangle \mapsto \max, & (13.1.6) \end{cases}$$

$$\begin{cases} \sum_{k=1}^M b^{(k)} = b, & (13.1.7) \end{cases}$$

$$\begin{cases} b^{(k)} \geq 0, \quad k = \overline{1, M}, & (13.1.8) \end{cases}$$

де $0 \in \mathbb{R}^m$, $\lambda^{(k)} \in \mathbb{R}^m$.

Враховуючи, що для кожної i -ої компоненти вектора глобальних ресурсів b (для кожного виду

ресурсу) рівняння (13.1.7) незалежні, розв'язок задачі центра (13.1.6) – (13.1.8) буде таким:

$$b_i^{(k)*} = \begin{cases} 0, & \lambda_i^{(k)} < \max_{k=1, M} \lambda_i^{(k)} = \lambda_i^{(k)*}, \\ b, & \lambda_i^{(k)} = \max_{k=1, M} \lambda_i^{(k)} = \lambda_i^{(k)*}. \end{cases} \quad (13.1.9)$$

У тому випадку, коли серед двоїстих оцінок є кілька максимальних, весь обсяг b_i виділяється одному із локальних об'єктів, для якого досягається максимум у (13.1.9), наприклад тому, хто має мінімальний індекс k .

Зазначимо, що метод Корнаї – Ліптака також ітераційний. Якщо t – номер ітерації, $\lambda_i^{(k)}(t)$ – оцінка ефективності використання i -го глобального ресурсу k -ою підсистемою на ітерації t , $b_i^{(k)}(t)$ – обсяг i -го ресурсу, що використовується k -ою підсистемою на ітерації t , то формули для розподілу глобальних ресурсів (усереднених за всіма реалізованими ітераціями) такі:

$$b_i^{(k)}(t) = \frac{t-1}{t} b_i^{(k)}(t-1) + \frac{1}{t} b_i^{(k)*}, \quad k = \overline{1, M}; \quad i = \overline{1, m}. \quad (13.1.10)$$

Встановлено, що безпосереднє використання розподілу (13.1.9) приводить до нестійкості процесу, тому величину $b_i^{(k)*}$ використовують тільки як величину корекції розподілу, реалізованого на попередній ітерації, що й відображено у (13.1.10). Крім того, розподіл (13.1.10) на деяких ітераціях може бути недопустимим. У цьому випадку користуються попередніми допустимими обсягами ресурсів. Потім формують нові локальні задачі (13.1.1) – (13.1.4) і

процес розв'язування вихідної задачі продовжується до тих пір, доки розв'язок не буде знайдений.

Зазначимо, що початковий розподіл глобальних ресурсів може бути довільним із допустимих (довільними із допустимих можуть бути величини $b_i^{(k)}(0)$, $k = \overline{1, M}$; $i = \overline{1, m}$). Підкреслимо також, що у загальному випадку цей метод є наближеним і дозволяє для заданої похибки наближення ε одержати розв'язок задачі за скінченну кількість ітерації $T(\varepsilon)$. Процес обчислень завершується, якщо різниця між верхньою та нижньою оцінками оптимального значення критерію (12.1.1) не перевищує заданої похибки ε . ►

Завдання 13.2. *Проілюструвати процес вертикальної декомпозиції Корнаї-Лінтака на конкретному прикладі у випадку, коли $M = 2$; $x^{(1)} = (x_1, x_2)^T$; $x^{(2)} = (x_3, x_4)^T$; $c^{(1)} = (1+k, 2+k)^T$, $c^{(2)} = (3+k, k)^T$ (k може набувати значень від 1 до 5); $A^{(1)} = (3, 1)$; $A^{(2)} = (2, 2)$; $D^{(1)} = (2, 1)$; $D^{(2)} = (1, 2)$; $b = 100$; $d^{(1)} = 40$; $d^{(2)} = 40$.*

◀ Нехай $k = 1$. На першій ітерації візьмемо, що

$$b^{(1)}(0) = 50, \quad b^{(2)}(0) = 50.$$

Використовуючи (13.1.1) – (13.1.4), отримаємо локальні задачі:

для першої підсистеми:

$$\begin{cases} 2x_1 + 3x_2 \mapsto \max, \\ 3x_1 + x_2 = 50, \\ 2x_1 + x_2 = 40, \\ x_1 \geq 0, \quad x_2 \geq 0; \end{cases} \quad (13.2.1)$$

для другої підсистеми:

$$\begin{cases} 4x_3 + x_4 \mapsto \max, \\ 2x_3 + 2x_4 = 50, \\ x_3 + 2x_4 = 40, \\ x_3 \geq 0, \quad x_4 \geq 0; \end{cases} \quad (13.2.2)$$

Задачі (13.2.1) і (13.2.2) мають відповідно такі розв'язки:

$$x_1^* = 10, \quad x_2^* = 20, \quad f^{(1)*} = 80 \quad \text{і} \quad x_3^* = 10, \quad x_4^* = 15, \quad f^{(2)*} = 55.$$

Двоїстою до (13.2.1) буде задача

$$\begin{cases} 50\lambda_1^{(1)} + 40\lambda_2^{(1)} \mapsto \min, \\ 3\lambda_1^{(1)} + 2\lambda_2^{(1)} \geq 2, \\ \lambda_1^{(1)} + \lambda_2^{(1)} \geq 3, \end{cases} \quad (13.2.3)$$

а до (13.2.2) – задача

$$\begin{cases} 50\lambda_1^{(2)} + 40\lambda_2^{(2)} \mapsto \min, \\ 2\lambda_1^{(2)} + \lambda_2^{(2)} \geq 4, \\ 2\lambda_1^{(2)} + 2\lambda_2^{(2)} \geq 1. \end{cases} \quad (13.2.4)$$

Розв'язками задач (13.2.3) і (13.2.4) є

$$\lambda_1^{(1)*} = -4, \quad \lambda_2^{(1)*} = 7, \quad \tilde{f}^{(1)*} = 80 \quad \text{і} \quad \lambda_1^{(2)*} = \frac{7}{2}, \quad \lambda_2^{(2)*} = -3,$$

$$\tilde{f}^{(2)*} = 55.$$

Значення оцінки так званого глобального ресурсу в обох задачах різні ($\lambda_1^{(1)*} \neq \lambda_1^{(2)*}$), тобто умова (13.1.5) не

виконується. Отже потрібно розв'язати задачу центра (13.1.6) – (13.1.8), яка у даному випадку має вигляд:

$$\begin{cases} -4b^{(1)} + 3,5b^{(2)} \mapsto \max, \\ b^{(1)} + b^{(2)} = 100, \\ b^{(1)} \geq 0, b^{(2)} \geq 0. \end{cases} \quad (13.2.5)$$

Розв'язок задачі (13.2.5), враховуючи (13.1.9), буде таким: $b^{(1)*} = 0$, $b^{(2)*} = 100$.

Далі, скориставшись співвідношенням (13.1.10), конкретизуємо розподіл ресурсів на наступній ітерації:

$$b^{(1)}(1) = \frac{1-1}{1} b^{(1)}(1-1) + \frac{1}{1} b^{(1)*} = 0 \cdot 50 + 1 \cdot 0 = 0,$$

$$b^{(2)}(1) = \frac{1-1}{1} b^{(2)}(1-1) + \frac{1}{1} b^{(2)*} = 0 \cdot 50 + 1 \cdot 100 = 100.$$

Переходимо до другої ітерації. Якщо для отриманих значень $b^{(1)}(1)$ і $b^{(2)}(1)$ побудувати задачі (13.1.1) – (13.1.4), то для першої підсистеми множина допустимих розв'язків буде порожньою, оскільки при $x_1 \geq 0$, $x_2 \geq 0$ рівності $3x_1 + x_2 = 0$ і $2x_1 + x_2 = 40$ несумісні. Відсутність глобального ресурсу тут оцінюється нескінченно великою величиною (можна вважати, що $\lambda_1^{(1)*}$ є нескінченно великою величиною), а отже, згідно з (13.1.9), $b^{(1)*} = 100$, $b^{(2)*} = 0$. Підставивши ці значення у (13.1.10), матимемо на другій ітерації

$$b^{(1)}(2) = \frac{2-1}{2} b^{(1)}(2-1) + \frac{1}{2} b^{(1)*} = 0 + \frac{1}{2} \cdot 100 = 50,$$

$$b^{(2)}(2) = \frac{2-1}{2}b^{(2)}(2-1) + \frac{1}{2}b^{(2)*} = \frac{1}{2} \cdot 100 + 0 = 50.$$

Це означає, що ми знову повернулися до початкового розподілу ресурсу, що у свою чергу приводить нас до задачі (13.2.5), розв'язком якої є $b^{(1)*} = 0$, $b^{(2)*} = 100$.

Після цього, як на третій, так і на четвертій ітераціях, розподіл глобального ресурсу, згідно з (13.1.10), приводить нас до локальних задач, які не мають розв'язку. У цьому випадку можна вважати, що $b^{(1)}(3) = b^{(2)}(3) = b^{(1)}(4) = b^{(2)}(4) = 50$, тобто розподіл залишається попереднім.

На п'ятій ітерації, почавши з $b^{(1)}(4) = b^{(2)}(4) = 50$, ми знову отримуємо локальні задачі (13.2.1), (13.2.2) та відповідну задачу центра (13.2.5), тому $b^{(1)*} = 0$, $b^{(2)*} = 100$. Тоді

$$b^{(1)}(5) = \frac{5-1}{5}b^{(1)}(4) + \frac{1}{5}b^{(1)*} = \frac{4}{5} \cdot 50 = 40,$$

$$b^{(2)}(5) = \frac{5-1}{5}b^{(2)}(4) + \frac{1}{5}b^{(2)*} = \frac{4}{5} \cdot 50 + \frac{1}{5} \cdot 100 = 60.$$

Тепер локальні задачі відповідно для першої і другої підсистем будуть такими:

$$\begin{cases} 2x_1 + 3x_2 \mapsto \max, \\ 3x_1 + x_2 = 40, \\ 2x_1 + x_2 = 40, \\ x_1 \geq 0, x_2 \geq 0, \end{cases} \quad (13.2.6)$$

$$\begin{cases} 4x_3 + x_4 \mapsto \max, \\ 2x_3 + 2x_4 = 60, \\ x_3 + 2x_4 = 40, \\ x_3 \geq 0, \quad x_4 \geq 0. \end{cases} \quad (13.2.7)$$

Ці задачі мають такі розв'язки:

$$x_1^* = 0, \quad x_2^* = 40, \quad f^{(1)*} = 120, \quad x_3^* = 20, \quad x_4^* = 10, \quad f^{(2)*} = 90.$$

Двоїстою до (13.2.6) є задача

$$\begin{cases} 40\lambda_1^{(1)} + 40\lambda_2^{(1)} \mapsto \min, \\ 3\lambda_1^{(1)} + 2\lambda_2^{(1)} \geq 2, \\ \lambda_1^{(1)} + \lambda_2^{(1)} \geq 3, \end{cases} \quad (13.2.8)$$

а до (13.2.7) – задача

$$\begin{cases} 60\lambda_1^{(2)} + 40\lambda_2^{(2)} \mapsto \min, \\ 2\lambda_1^{(2)} + \lambda_2^{(2)} \geq 4, \\ 2\lambda_1^{(2)} + 2\lambda_2^{(2)} \geq 1. \end{cases} \quad (13.2.9)$$

Розв'язком задачі (13.2.9) буде $\lambda_1^{(2)*} = \frac{7}{2}$, $\lambda_2^{(2)*} = -3$,

$\tilde{f}^{(2)*} = 90$, а задача (13.2.8) має множину (безліч)

розв'язків зі значенням цільової функції $\tilde{f}^{(1)*} = 120$.

Оскільки серед розв'язків задачі (13.2.8) є, зокрема,

$\lambda_1^{(1)*} = \frac{7}{2}$, $\lambda_2^{(1)*} = -\frac{1}{2}$, то це означає, що оцінка глобального

ресурсу в обох задачах однакова ($\lambda_1^{(1)*} = \lambda_1^{(2)*}$). Отже,

процес розв'язування завершений. Зауважимо, що поява неєдиності розв'язку двоїстої задачі в одній із підсистем – це характерне явище для схеми Корнаї – Ліптака. ►

ТЕМА 14. АГРЕГУВАННЯ У МОДЕЛЯХ МІЖГАЛУЗЕВИХ БАЛАНСІВ

Завдання 14.1. Нехай $X = (x_{ij})_{i,j=1}^4$ та

$x = (x_1, x_2, x_3, x_4)^T$, де T – транспонування (варіанти 14.1.1 – 14.1.5) – відповідно матриця міжгалузевих потоків продукції та вектор валового випуску балансової моделі чотиригалузевої економічної системи. Необхідно:

1) на основі матриці X і вектора x знайти матрицю прямих матеріальних витрат $A = (a_{ij})_{i,j=1}^4$;

2) визначити агреговані матриці міжгалузевих потоків $X^* = (x_{kl}^*)_{k,l=1}^3$ і прямих матеріальних витрат

$A^* = (a_{kl}^*)_{k,l=1}^3$ спрощеної нової балансової моделі трьохгалузевої економічної системи, утвореної у результаті об'єднання у одну галузь (або сектор) другої та третьої галузей зазначеної вище чотиригалузевої економіки.

14.1.1

$$X = \begin{pmatrix} 10 & 60 & 100 & 20 \\ 40 & 20 & 20 & 100 \\ 20 & 100 & 60 & 60 \\ 30 & 40 & 60 & 40 \end{pmatrix}, \quad x = \begin{pmatrix} 200 \\ 400 \\ 500 \\ 400 \end{pmatrix}$$

14.1.2

$$X = \begin{pmatrix} 10 & 40 & 25 & 60 \\ 20 & 20 & 100 & 20 \\ 30 & 40 & 25 & 40 \\ 20 & 20 & 50 & 60 \end{pmatrix}, \quad x = \begin{pmatrix} 100 \\ 200 \\ 250 \\ 200 \end{pmatrix}$$

14.1.3

$$X = \begin{pmatrix} 80 & 30 & 60 & 40 \\ 20 & 90 & 90 & 20 \\ 40 & 60 & 90 & 40 \\ 20 & 60 & 30 & 60 \end{pmatrix}, \quad x = \begin{pmatrix} 200 \\ 300 \\ 300 \\ 200 \end{pmatrix}$$

14.1.4

$$X = \begin{pmatrix} 20 & 30 & 40 & 10 \\ 20 & 20 & 20 & 30 \\ 40 & 10 & 60 & 20 \\ 40 & 20 & 20 & 30 \end{pmatrix}, \quad x = \begin{pmatrix} 200 \\ 100 \\ 200 \\ 100 \end{pmatrix}$$

14.1.5

$$X = \begin{pmatrix} 90 & 30 & 120 & 80 \\ 60 & 60 & 40 & 120 \\ 90 & 30 & 80 & 80 \\ 30 & 60 & 120 & 40 \end{pmatrix}, \quad x = \begin{pmatrix} 300 \\ 300 \\ 400 \\ 400 \end{pmatrix}$$

◀ **1)** Для виконання цього завдання найперше потрібно ознайомитися із елементами теорії балансових моделей міжгалузевих зв'язків або моделей міжгалузевого балансу [3]. Статична модель

міжгалузевого балансу у випадку n -галузевої економіки формалізується матрично-векторним співвідношенням

$$x = X + y = Ax + y, \quad (14.1.1)$$

де $x = (x_1, \dots, x_n)^T$ та $y = (y_1, \dots, y_n)^T$ - вектори валової та кінцевої продукції, а $X = (x_{ij})_{i,j=1}^n$ та $A = (a_{ij})_{i,j=1}^n$ - матриці міжгалузевих потоків продукції та прямих матеріальних (виробничих) витрат. При цьому

$$a_{ij} = \frac{x_{ij}}{x_j}, \quad i, j = \overline{1, n}. \quad (14.1.2)$$

Не заглиблюючись у теорію моделі (14.1.1), зазначимо лише, що у цій моделі (її ще називають моделлю Леонтьєва) заданими вважаються матриця X і вектор y (чи матриця A і вектор y), а шуканим - вектор валового випуску x . Однак, нас цікавить проблема агрегування, що полягає у об'єднанні (укрупненні) множини елементів у одне ціле, а у контексті моделювання - у переході від складнішої моделі до простішої моделі (наприклад, від моделі більшої розмірності до моделі меншої розмірності), тому перейдемо до проблеми агрегування моделі (14.1.1).

Враховуючи економічний зміст матриці X , легко зрозуміти, що об'єднання двох галузей в одну фактично перетворює матрицю X у матрицю X^* за допомогою елементарних лінійних операцій. До прикладу, якщо

$X = (x_{ij})_{i,j=1}^4$, то об'єднання другої та третьої галузей у галузь з номером q означає, що

$$x_{iq}^* = x_{i2} + x_{i3}, \quad i = \overline{1,4}, \quad i \neq 2, \quad i \neq 3; \quad (14.1.3)$$

$$x_{qj}^* = x_{2j} + x_{3j}, \quad j = \overline{1,4}, \quad j \neq 2, \quad j \neq 3; \quad (14.1.4)$$

$$x_{qq}^* = x_{22} + x_{23} + x_{32} + x_{33}. \quad (14.1.5)$$

Співвідношення (14.1.3) – (14.1.5) насправді означають «злиття» шляхом додавання елементів другого та третього стовпців матриці X у один стовпчик і другого та третього рядків матриці X у один рядок (тут логічно вважати $q = 2$).

Матриця X^* буде такою:

$$X^* = \begin{pmatrix} x_{11}^* & x_{12}^* & x_{13}^* \\ x_{21}^* & x_{22}^* & x_{23}^* \\ x_{31}^* & x_{32}^* & x_{33}^* \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} + x_{13} & x_{14} \\ x_{21} + x_{31} & x_{22} + x_{23} + x_{32} + x_{33} & x_{24} + x_{34} \\ x_{41} & x_{42} + x_{43} & x_{44} \end{pmatrix}.$$

Крім матриць $X = (x_{ij})_{i,j=1}^4$ та $X^* = (x_{kl}^*)_{k,l=1}^3$,

введемо у розгляд також матрицю

$$\tilde{X} = \begin{pmatrix} x_{11} & x_{12} + x_{13} & x_{14} \\ x_{21} & x_{22} + x_{23} & x_{24} \\ x_{31} & x_{32} + x_{33} & x_{34} \\ x_{41} & x_{42} + x_{43} & x_{44} \end{pmatrix}.$$

Тоді виконання операцій (14.1.3) – (14.1.5)

зводиться до двох кроків: а) додавання другого та третього стовпців матриці X і утворення матриці \tilde{X} розмірності (4×3) ; б) додавання другого і третього рядків матриці \tilde{X} і утворення матриці X^* розмірності (3×3) .

Зауважимо, що операції додавання другого і третього стовпців і рядків можна поміняти місцями.

Аналогічно вектор валового випуску $x = (x_1, x_2, x_3, x_4)^T$ агрегується у вектор валового випуску $x^* = (x_1^*, x_2^*, x_3^*)^T \equiv (x_1, x_2 + x_3, x_4)^T$.

Потім, використовуючи співвідношення (14.1.2), а більш точно співвідношення

$$a_{kl}^* = \frac{x_{kl}^*}{x_l^*}, \quad k, l = \overline{1, 3} \quad (14.1.6)$$

та матрицю X^* , уточнюємо матрицю $A^* = (a_{kl}^*)_{k,l=1}^3$.

Нескладно також перевірити, що матрицю A^* можна обчислити за формулою

$$A^* = I^* A V^*, \quad (14.1.7)$$

де

$$I^* = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

є матрицею, отриманою на основі одиничної матриці

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

шляхом додавання її другого і третього рядків (номери рядків відповідають номерам галузей, що об'єднуються у єдину агреговану галузь), а

$$V^* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & v_2 & 0 \\ 0 & v_3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

– матрицею, отриманою на основі матриці вагових коефіцієнтів (ваг) кожної із галузей

$$\left(v_1 = 1, v_4 = 1, v_2 = \frac{x_2}{x_2 + x_3}, v_3 = \frac{x_3}{x_2 + x_3} \right)$$

$$V = \begin{pmatrix} v_1 & 0 & 0 & 0 \\ 0 & v_2 & 0 & 0 \\ 0 & 0 & v_3 & 0 \\ 0 & 0 & 0 & v_4 \end{pmatrix}$$

шляхом додавання її другого та третього стовпців.

Порівняно з (14.1.6) формула (14.1.7) для визначення агрегованої матриці прямих матеріальних витрат у процедурах агрегування моделей міжгалузевих балансів використовується значно ширше, причому для різного числа галузей n .

Проілюструємо результати агрегування моделі міжгалузевих зв'язків у випадку, коли

$$X = \begin{pmatrix} 10 & 60 & 100 & 20 \\ 40 & 20 & 20 & 100 \\ 20 & 100 & 60 & 60 \\ 30 & 40 & 60 & 40 \end{pmatrix}, \quad x = \begin{pmatrix} 200 \\ 400 \\ 500 \\ 400 \end{pmatrix}.$$

1) Матрицю матеріальних витрат A знаходимо за допомогою співвідношень (14.1.2):

$$A = \begin{pmatrix} 0,05 & 0,15 & 0,2 & 0,05 \\ 0,2 & 0,05 & 0,04 & 0,25 \\ 0,1 & 0,25 & 0,12 & 0,15 \\ 0,15 & 0,1 & 0,12 & 0,1 \end{pmatrix}.$$

2) Об'єднання другої та третьої галузей у одну агреговану галузь приводить до матриці міжгалузевих потоків продукції

$$X^* = \begin{pmatrix} 10 & 160 & 20 \\ 60 & 200 & 160 \\ 30 & 100 & 40 \end{pmatrix}$$

та вектора валових випусків $x^* = (200, 900, 400)^T$.

Отже, з урахуванням (14.1.6) агрегована матриця прямих матеріальних витрат має вигляд:

$$A^* = \begin{pmatrix} \frac{1}{20} & \frac{8}{45} & \frac{1}{20} \\ \frac{3}{10} & \frac{2}{9} & \frac{2}{5} \\ \frac{3}{20} & \frac{1}{9} & \frac{1}{10} \end{pmatrix} \approx \begin{pmatrix} 0,05 & 0,18 & 0,05 \\ 0,3 & 0,22 & 0,4 \\ 0,15 & 0,11 & 0,1 \end{pmatrix}.$$

Застосування формули (14.1.7) також приводить до цього ж результату. ►

Завдання 14.2. *Моделі міжгалузевого балансу часто слугують для демонстрації типових прикладів агрегування в економіці. При цьому поширеним у прикладних дослідженнях є метод ітераційного агрегування в один продукт чи в t продуктів ($1 < t < n$, n – кількість галузей економіки). Потрібно:*

1) *конкретизувати алгоритм ітераційного агрегування у загальному випадку ($1 < t < n$) та в один продукт;*

2) *проілюструвати роботу цього алгоритму та його результати для балансової моделі (моделі Леонт'єва) трьохгалузевої економіки (у випадку агрегування в один продукт і у два продукти) із конкретно заданими матрицею прямих матеріальних витрат A та вектором кінцевої продукції y , попередньо знайшовши точний розв'язок системи (14.1.1) (варіанти 14.2.1 – 14.2.5):*

14.2.1

$$A = \begin{pmatrix} 0,2 & 0,15 & 0,1 \\ 0,25 & 0,3 & 0,1 \\ 0,15 & 0,15 & 0,3 \end{pmatrix}, \quad y = \begin{pmatrix} 130 \\ 70 \\ 60 \end{pmatrix}$$

14.2.2

$$A = \begin{pmatrix} 0,1 & 0,2 & 0,2 \\ 0,2 & 0,3 & 0,1 \\ 0,1 & 0,1 & 0,2 \end{pmatrix}, \quad y = \begin{pmatrix} 100 \\ 150 \\ 120 \end{pmatrix}$$

14.2.3

$$A = \begin{pmatrix} 0,2 & 0,2 & 0,3 \\ 0,3 & 0,1 & 0,2 \\ 0,1 & 0,3 & 0,1 \end{pmatrix}, \quad y = \begin{pmatrix} 200 \\ 100 \\ 200 \end{pmatrix}$$

14.2.4

$$A = \begin{pmatrix} 0,1 & 0,3 & 0,1 \\ 0,2 & 0,1 & 0,2 \\ 0,3 & 0,2 & 0,1 \end{pmatrix}, \quad y = \begin{pmatrix} 200 \\ 300 \\ 100 \end{pmatrix}$$

14.2.5

$$A = \begin{pmatrix} 0,4 & 0,2 & 0,2 \\ 0,1 & 0,3 & 0,4 \\ 0,2 & 0,1 & 0,2 \end{pmatrix}, \quad y = \begin{pmatrix} 120 \\ 180 \\ 200 \end{pmatrix}$$

◀ **1)** Алгоритм процесу ітераційного агрегування міжгалузевого балансу у $m > 1$ продуктів ($m < n$) для першого кроку ($k = 0$) такий:

1. вибір вектора $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})^T$ – початкового наближення шуканого вектора $x = (x_1, \dots, x_n)^T$ – розв'язку системи (14.1.1);

2. обчислення кожної компоненти $z_s^{(k)}$ s -того агрегованого продукту, тобто визначення агрегованого вектора $z^{(k)} = (z_1^{(k)}, \dots, z_m^{(k)})^T$:

$$z_s^{(k)} = \sum_{j \in J_s} x_j^{(k)}, \quad s = \overline{1, m}, \quad (14.2.1)$$

де J_s – множина індексів продукції деталізованої номенклатури, що входять в s -й агрегований продукт;

3. знаходження матриць $D^{(k)}$ і $Q^{(k)}$ – матриць напіваагрегованих і агрегованих нормативів (J_r – множина індексів продукції деталізованої номенклатури, що входять в r -й агрегований продукт):

$$D^{(k)} = \left(d_{is}^{(k)} \right)_{i,s=1}^{n,m}, \quad d_{is}^{(k)} = \frac{\sum_{j \in J_s} a_{ij} x_j^{(k)}}{z_s^{(k)}}, \quad (14.2.2)$$

$$Q^{(k)} = \left(q_{rs}^{(k)} \right)_{r,s=1}^m, \quad q_{rs}^{(k)} = \sum_{i \in J_r} d_{is}^{(k)}; \quad (14.2.3)$$

4. визначення наступних $(k+1)$ -их наближень векторів $z^{(k+1)} = \left(z_1^{(k+1)}, \dots, z_m^{(k+1)} \right)^T$ і $x^{(k+1)} = \left(x_1^{(k+1)}, \dots, x_n^{(k+1)} \right)^T$:

а) розв'язування відносно $z^{(k+1)}$ системи агрегованого міжгалузевого балансу

$$z^{(k+1)} = Q^{(k)} z^{(k+1)} + \tilde{y}, \quad (14.2.4)$$

де $\tilde{y} = \left(\tilde{y}_1, \dots, \tilde{y}_m \right)^T$ – вектор агрегованої кінцевої продукції, у якому

$$\tilde{y}_r = \sum_{i \in J_r} y_i; \quad (14.2.5)$$

б) обчислення $x^{(k+1)}$:

$$x^{(k+1)} = D^{(k)} z^{(k+1)} + y; \quad (14.2.6)$$

5. якщо бажана точність компонент вектора $x^{(k+1)}$ недосягнута, то перехід до наступного $(k+1)$ -го кроку.

Що стосується питання збіжності та швидкості збіжності ітераційного процесу (14.2.1)-(14.2.6), то його окремо проаналізуємо у випадку агрегування в один продукт ($m=1$) та агрегування в $m>1$ продуктів ($m < n$).

Агрегування в один продукт ($m=1, s=1, J_1 = \{1, 2, \dots, n\}$). У цій ситуації формули (14.2.1)-(14.2.6) стають простішими, зокрема матриця $D^{(k)}$ буде матрицею-стовпцем

$$\begin{cases} D^{(k)} = (d_i^{(k)})_{i=1}^n \equiv (d_{i1}^{(k)})_{i,s=1}^{n,1}, \\ d_i^{(k)} = \frac{\sum_{j=1}^n a_{ij} x_j^{(k)}}{z^{(k)}}, \end{cases} \quad (14.2.7)$$

де $z^{(k)} \equiv z_1^{(k)}$ – число, що знаходиться за формулою

$$z^{(k)} = \sum_{j=1}^n x_j^{(k)}. \quad (14.2.8)$$

Числом є також матриця $Q^{(k)}$, тобто

$$Q^{(k)} = q^{(k)} \equiv (q_{rs}^{(k)})_{r,s=1}^1 = \sum_{i=1}^n d_i^{(k)}. \quad (14.2.9)$$

Система (14.2.4) перетворюється в одне лінійне рівняння з одним невідомим

$$z^{(k+1)} = q^{(k)} z^{(k+1)} + \tilde{y}, \quad (14.2.10)$$

де

$$\tilde{y} = \sum_{i=1}^n y_i. \quad (14.2.11)$$

Співвідношення (14.2.6) для знаходження $(k+1)$ -го наближення залишається без змін.

Строго доведено, що ітераційний процес (14.2.6) – (14.2.11) збігається, якщо:

- вектор початкового наближення $x^{(0)} \geq 0$ (складається з невід'ємних елементів);

- матриця A є цілком продуктивною, тобто $a_{ij} \geq 0$, $i, j = \overline{1, n}$; $\sum_{i=1}^n a_{ij} = \alpha_j < 1$, $j = \overline{1, n}$;

- вектор $y \geq 0$ (складається з невід'ємних елементів) і $\tilde{y} = \sum_{i=1}^n y_i > 0$.

Крім того, швидкість збіжності ітераційного процесу збігається зі швидкістю геометричної прогресії, знаменником якої є число $\gamma = \max |\lambda(C)|$, де $C = A - \tilde{A}$,

$\tilde{A} = (\beta_i \alpha_j)_{i,j=1}^n$, $\beta_i = \frac{y_i}{\tilde{y}}$ ($i = \overline{1, n}$), $\lambda(C)$ – власне значення

матриці C , $\max |\lambda(C)|$ – максимальне за модулем власне значення матриці C .

2) Проілюструємо процес (14.2.6) – (14.2.11) на конкретному прикладі моделі (14.1.1), коли

$$A = \begin{pmatrix} 0,4 & 0,2 & 0,1 \\ 0,25 & 0,2 & 0,3 \\ 0,1 & 0,1 & 0,4 \end{pmatrix}, \quad y = \begin{pmatrix} 31,727 \\ 37,393 \\ 9,3483 \end{pmatrix}.$$

Отже, реалізуємо процес (14.2.6) – (14.2.11), тобто процес (14.2.1) – (14.2.6) для випадку агрегування в один продукт, й оцінимо швидкість його збіжності.

Зауважимо, що всі умови збіжності виконуються. Точний розв'язок системи (14.1.1) такий:

$$x_1^* = 91,4908; x_2^* = 92,6939; x_3^* = 46,2858.$$

За початкове наближення візьмемо, наприклад наближення:

$$x_1^{(0)} = 50; x_2^{(0)} = 150; x_3^{(0)} = 75.$$

$$\text{Оскільки } z^{(0)} = x_1^{(0)} + x_2^{(0)} + x_3^{(0)} = 275,$$

$$d_1^{(0)} = \left(\sum_{j=1}^3 a_{1j} x_j^{(0)} \right) / z^{(0)} = 57,5 / 275 = 0,2091;$$

$$d_2^{(0)} = \left(\sum_{j=1}^3 a_{2j} x_j^{(0)} \right) / z^{(0)} = 65 / 275 = 0,2364;$$

$$d_3^{(0)} = \left(\sum_{j=1}^3 a_{3j} x_j^{(0)} \right) / z^{(0)} = 50 / 275 = 0,1818;$$

$$D^{(0)} = (0,2091 \quad 0,2364 \quad 0,1818)^T;$$

$$Q^{(0)} = q^{(0)} = d_1^{(0)} + d_2^{(0)} + d_3^{(0)} = 0,6273;$$

$$\tilde{y} = y_1 + y_2 + y_3 = 78,4683,$$

$$\text{то } z^{(1)} = q^{(0)} z^{(0)} + \tilde{y}, \text{ тобто } z^{(1)} = \tilde{y} / (1 - q^{(0)}) = 210,5401$$

і, згідно з (14.2.6),

$$\begin{aligned}
 x^{(1)} &= D^{(0)}z^{(1)} + y = \begin{pmatrix} 0,2091 \\ 0,2364 \\ 0,1818 \end{pmatrix} \cdot 210,5401 + \begin{pmatrix} 31,727 \\ 37,393 \\ 9,3483 \end{pmatrix} = \\
 &= \begin{pmatrix} 75,7509 \\ 87,1647 \\ 47,6245 \end{pmatrix}.
 \end{aligned}$$

Отже, $x_1^{(1)} = 75,7509$; $x_2^{(1)} = 87,1647$; $x_3^{(1)} = 47,6245$.

Перейдемо до наступної (другої) ітерації. У цьому випадку $z^{(1)} = x_1^{(1)} + x_2^{(1)} + x_3^{(1)} = 210,5401$;

$$d_1^{(1)} = 52,4958/210,5401 = 0,2493;$$

$$d_2^{(1)} = 50,6580/210,5401 = 0,2406;$$

$$d_3^{(1)} = 35,3414/210,5401 = 0,1679;$$

$$D^{(1)} = (0,2493 \quad 0,2406 \quad 0,1679)^T;$$

$$Q^{(1)} = q^{(1)} = d_1^{(1)} + d_2^{(1)} + d_3^{(1)} = 0,6578;$$

$$z^{(2)} = \tilde{y}/(1 - q^{(1)}) = 229,3054.$$

Отже,

$$\begin{aligned}
 x^{(2)} &= D^{(1)}z^{(2)} + y = 229,3054 \cdot \begin{pmatrix} 0,2493 \\ 0,2406 \\ 0,1679 \end{pmatrix} + \begin{pmatrix} 31,727 \\ 37,393 \\ 9,3483 \end{pmatrix} = \\
 &= \begin{pmatrix} 88,8928 \\ 92,5639 \\ 47,8487 \end{pmatrix},
 \end{aligned}$$

тобто $x_1^{(2)} = 88,8928$; $x_2^{(2)} = 92,5639$; $x_3^{(2)} = 47,8487$.

Отже, вже на другій ітерації отримано наближений розв'язок, який є досить близьким до точного розв'язку.

Процес розв'язування продовжується доти, доки не буде досягнута бажана точність. Зокрема, на шостій ітерації результати будуть такими: $x_1^{(6)} = 91,46946$; $x_2^{(6)} = 92,6856$; $x_3^{(6)} = 46,28038$. Нескладно оцінити також швидкість збіжності методу ітераційного агрегування. Зважаючи на те, що $\alpha_1 = 0,75$; $\alpha_2 = 0,5$; $\alpha_3 = 0,8$ (тут α_j – сума елементів j -того стовпця матриці A); $\beta_1 = y_1 / \tilde{y} = 0,4043$; $\beta_2 = y_2 / \tilde{y} = 0,4765$; $\beta_3 = y_3 / \tilde{y} = 0,1192$, знаходимо

$$\tilde{A} = (\beta_i \alpha_j)_{i,j=1}^3 = \begin{pmatrix} \beta_1 \alpha_1 & \beta_1 \alpha_2 & \beta_1 \alpha_3 \\ \beta_2 \alpha_1 & \beta_2 \alpha_2 & \beta_2 \alpha_3 \\ \beta_3 \alpha_1 & \beta_3 \alpha_2 & \beta_3 \alpha_3 \end{pmatrix}$$

$$= \begin{pmatrix} 0,3032 & 0,2021 & 0,3234 \\ 0,3574 & 0,2382 & 0,3812 \\ 0,0894 & 0,0597 & 0,0954 \end{pmatrix},$$

$$C = A - \tilde{A} = \begin{pmatrix} 0,0968 & -0,0021 & -0,2234 \\ -0,1074 & -0,0382 & -0,0812 \\ 0,0106 & 0,0403 & 0,3046 \end{pmatrix}.$$

Власні значення матриці C є розв'язками кубічного рівняння $|C - \lambda I| = 0$, де $|G|$ – визначник матриці G . Наближені величини власних значень матриці C такі: $\lambda_1(C) = 0,2876$; $\lambda_2(C) = 0,0756$; $\lambda_3(C) = 0$. Тому $\gamma = \max |\lambda(C)| = 0,2876$. Легко переконатися у тому, що $|\lambda(C)| \leq \|C\|$, де $\|C\|$ – норма матриці C . За $\|C\|$ можна

вибрати будь-яку із трьох величин $\|C\|_I = \max_{i=1,n} \sum_{j=1}^n |c_{ij}|$,

$$\|C\|_{II} = \max_{j=1,n} \sum_{i=1}^n |c_{ij}|, \quad \|C\|_{III} = \left(\sum_{i=1}^n \sum_{j=1}^n c_{ij}^2 \right)^{\frac{1}{2}}.$$

У нашій ситуації

$\|C\|_I = 0,3555$; $\|C\|_{II} = 0,6092$; $\|C\|_{III} = 0,41$. Очевидно, γ і $\|C\|_I$ близькі величини. Оскільки $\gamma^2 = 0,2876^2 = 0,0827$, то можна стверджувати, що кожні дві ітерації даного методу ітераційного агрегування будуть зменшувати початкову похибку розв'язку практично в 10 разів.

Наведемо також приклад агрегування у два продукти ($m = 2$), причому першим в агрегованій номенклатурі будемо вважати продукт, що утворився в результаті агрегування 1-го і 3-го видів продукції (стовпців) деталізованої номенклатури. Другий вид продукції в агрегованій номенклатурі будемо вважати другим видом продукції у деталізованій номенклатурі (матриця A і вектор y ті ж самі).

За початкове наближення, як і раніше, візьмемо $x_1^{(0)} = 50$, $x_2^{(0)} = 150$, $x_3^{(0)} = 75$. Для реалізації першого кроку ітераційного процесу потрібно використати формули (14.2.1)-(14.2.6) при $k = 0$. Отже, $s = \overline{1,2}$;
 $J_1 = \{1,3\}$; $J_2 = \{2\}$; $z_1^{(0)} = \sum_{j \in J_1} x_j^{(0)} = x_1^{(0)} + x_3^{(0)} = 125$;

$$z_2^{(0)} = x_2^{(0)} = 150;$$

$$d_{11}^{(0)} = \frac{\sum_{j \in J_1} a_{1j} x_j^{(0)}}{z_1^0} = \frac{0,4 \cdot 50 + 0,1 \cdot 75}{125} = 0,22;$$

$$d_{21}^{(0)} = \frac{\sum_{j \in J_1} a_{2j} x_j^{(0)}}{z_1^0} = \frac{0,25 \cdot 50 + 0,3 \cdot 75}{125} = 0,28;$$

$$d_{31}^{(0)} = \frac{\sum_{j \in J_1} a_{3j} x_j^{(0)}}{z_1^0} = \frac{0,1 \cdot 50 + 0,4 \cdot 75}{125} = 0,28;$$

$$d_{12}^{(0)} = 0,2; \quad d_{22}^{(0)} = 0,2; \quad d_{32}^{(0)} = 0,1 \quad (\text{у цьому випадку}$$

$$d_{i2}^{(0)} = a_{i2} x_2^{(0)} / z_2^0 = a_{i2}, \quad i = \overline{1,3});$$

$$D^{(0)} = \begin{pmatrix} 0,22 & 0,2 \\ 0,28 & 0,2 \\ 0,28 & 0,1 \end{pmatrix};$$

$$q_{11}^{(0)} = \sum_{i \in J_1} d_{i1}^{(0)} = d_{11}^{(0)} + d_{31}^{(0)} = 0,5;$$

$$q_{12}^{(0)} = \sum_{i \in J_1} d_{i2}^{(0)} = d_{12}^{(0)} + d_{32}^{(0)} = 0,3;$$

$$q_{21}^{(0)} = \sum_{i \in J_2} d_{i1}^{(0)} = d_{21}^{(0)} = 0,28;$$

$$q_{22}^{(0)} = \sum_{i \in J_2} d_{i2}^{(0)} = d_{22}^{(0)} = 0,2;$$

$$Q^{(0)} = \begin{pmatrix} 0,5 & 0,3 \\ 0,28 & 0,2 \end{pmatrix};$$

$$\tilde{y}_1 = \sum_{i \in J_1} y_i = y_1 + y_3 = 41,0753;$$

$$\tilde{y}_2 = \sum_{i \in J_2} y_i = y_2 = 37,393.$$

Як впливає із уже зроблених проміжних обчислень, щоб отримати перший рядок матриці $Q^{(0)}$, потрібно додати перший та третій рядки матриці $D^{(0)}$, а другий рядок матриці $Q^{(0)}$ збігається із другим рядком матриці $D^{(0)}$.

Агрегована система (14.2.4) має вигляд $z^{(1)} = Q^{(0)}z^{(1)} + \tilde{y}$, тобто

$$\begin{aligned} z^{(1)} &= (I - Q^{(0)})^{-1} \tilde{y} = \begin{pmatrix} 4,3478261 & 1,6304348 \\ 3,9130435 & 2,7173913 \end{pmatrix} \cdot \begin{pmatrix} 41,0753 \\ 37,393 \end{pmatrix} = \\ &= \begin{pmatrix} 239,55511 \\ 262,34085 \end{pmatrix}. \end{aligned}$$

Тепер

$$\begin{aligned} x^{(1)} &= D^{(0)}z^{(1)} + y = \begin{pmatrix} 0,22 & 0,2 \\ 0,28 & 0,2 \\ 0,28 & 0,1 \end{pmatrix} \cdot \begin{pmatrix} 239,55511 \\ 262,34085 \end{pmatrix} + \begin{pmatrix} 31,727 \\ 37,393 \\ 9,3483 \end{pmatrix} = \\ &= \begin{pmatrix} 136,89729 \\ 156,9366 \\ 93,309516 \end{pmatrix}. \end{aligned}$$

Отже, $x_1^{(1)} = 136,8973$, $x_2^{(1)} = 156,9366$, $x_3^{(1)} = 93,3095$.

Перейдемо до наступного кроку ($k = 1$), здійснивши відповідні обчислення:

$$z_1^{(1)} = x_1^{(1)} + x_3^{(1)} = 230,2068; \quad z_2^{(1)} = x_2^{(1)} = 156,9366;$$

$$d_{11}^{(1)} = 0,28; \quad d_{21}^{(1)} = 0,27; \quad d_{31}^{(1)} = 0,22;$$

$$d_{12}^{(1)} = 0,2; \quad d_{22}^{(1)} = 0,2; \quad d_{32}^{(1)} = 0,1;$$

$$D^{(1)} = \begin{pmatrix} 0,28 & 0,2 \\ 0,27 & 0,2 \\ 0,22 & 0,1 \end{pmatrix};$$

$$q_{11}^{(1)} = 0,5; \quad q_{12}^{(1)} = 0,3; \quad q_{21}^{(1)} = 0,27; \quad q_{22}^{(1)} = 0,2;$$

$$Q^{(1)} = \begin{pmatrix} 0,5 & 0,3 \\ 0,27 & 0,2 \end{pmatrix},$$

$$\begin{aligned} z^{(2)} &= (I - Q^{(1)})^{-1} \tilde{y} = \begin{pmatrix} 2,507837 & 0,9404389 \\ 0,846395 & 1,5673981 \end{pmatrix} \cdot \begin{pmatrix} 41,0753 \\ 37,393 \end{pmatrix} = \\ &= \begin{pmatrix} 138,17599 \\ 93,375646 \end{pmatrix}. \end{aligned}$$

Тому

$$\begin{aligned} x^{(2)} &= D^{(1)} z^{(2)} + y = \begin{pmatrix} 0,28 & 0,2 \\ 0,27 & 0,2 \\ 0,22 & 0,1 \end{pmatrix} \begin{pmatrix} 138,17599 \\ 93,375646 \end{pmatrix} + \begin{pmatrix} 31,727 \\ 37,393 \\ 9,3483 \end{pmatrix} = \\ &= \begin{pmatrix} 89,091406 \\ 93,375646 \\ 49,084583 \end{pmatrix}, \end{aligned}$$

$$\text{тобто } x_1^{(2)} = 89,0914; \quad x_2^{(2)} = 93,3756; \quad x_3^{(2)} = 49,0846.$$

Далі процес обчислень продовжується до досягнення заданої точності.

Для оцінювання збіжності потрібно знайти $\|C\| = \|A - \tilde{A}\|$, де $\tilde{A} = (a_{ij})_{i,j=1}^n$, $\tilde{a}_{ij} = \beta_i \alpha_{rj}$ ($i \in J_r$, $r = \overline{1, m}$),

$$\beta_i = y_i / \tilde{y}_r \quad (i \in J_r, \quad r = \overline{1, m}), \quad \alpha_{rj} = \sum_{i \in J_r} a_{ij} \quad (j = \overline{1, n}; \quad r = \overline{1, m}).$$

Оскільки $J_1 = \{1, 3\}$, $J_2 = \{2\}$, то $\beta_1 = y_1 / \tilde{y}_1 = 0,7724$;

$$\beta_2 = y_2 / \tilde{y}_2 = 1; \beta_3 = y_3 / \tilde{y}_3 = 0,2276 \left(\tilde{y}_3 \equiv \tilde{y}_1 \right);$$

$$\alpha_{11} = a_{11} + a_{31} = 0,5; \alpha_{12} = a_{12} + a_{32} = 0,3;$$

$$\alpha_{13} = a_{13} + a_{33} = 0,5;$$

$$\alpha_{21} = a_{21} = 0,25; \alpha_{22} = a_{22} = 0,2; \alpha_{23} = a_{23} = 0,3;$$

$$\tilde{A} = \left(\tilde{\alpha}_{ij} \right)_{i,j=1}^3 \equiv \left(\beta_i \alpha_{ij} \right)_{i,j=1}^3 = \begin{pmatrix} \beta_1 \alpha_{11} & \beta_1 \alpha_{12} & \beta_1 \alpha_{13} \\ \beta_2 \alpha_{21} & \beta_2 \alpha_{22} & \beta_2 \alpha_{23} \\ \beta_3 \alpha_{11} & \beta_3 \alpha_{12} & \beta_3 \alpha_{13} \end{pmatrix} =$$

$$= \begin{pmatrix} \beta_1 (a_{11} + a_{31}) & \beta_1 (a_{12} + a_{32}) & \beta_1 (a_{13} + a_{33}) \\ \beta_2 a_{21} & \beta_2 a_{22} & \beta_2 a_{23} \\ \beta_3 (a_{11} + a_{31}) & \beta_3 (a_{12} + a_{32}) & \beta_3 (a_{13} + a_{33}) \end{pmatrix} =$$

$$= \begin{pmatrix} 0,7724 \cdot 0,5 & 0,7724 \cdot 0,3 & 0,7724 \cdot 0,5 \\ 1 \cdot 0,25 & 1 \cdot 0,2 & 1 \cdot 0,3 \\ 0,2276 \cdot 0,5 & 0,2276 \cdot 0,3 & 0,2276 \cdot 0,5 \end{pmatrix} =$$

$$= \begin{pmatrix} 0,3862 & 0,2317 & 0,3862 \\ 0,25 & 0,2 & 0,3 \\ 0,1138 & 0,0683 & 0,1138 \end{pmatrix},$$

$$C = A - \tilde{A} = \begin{pmatrix} 0,0138 & -0,0317 & -0,2862 \\ 0 & 0 & 0 \\ -0,0138 & 0,0317 & 0,2862 \end{pmatrix}.$$

Для цієї матриці

$$\lambda_1(C) = 0,3; \lambda_2(C) = \lambda_3(C) = 0, \|C\|_I = 0,3317;$$

$$\|C\|_{II} = 0,5724; \|C\|_{III} = 0,41.$$

Отже, всі власні значення можна оцінити зверху числом $\tilde{\gamma} = \|C\|_I = 0,3317$, звідки випливає, що швидкість збіжності тут рівносильна швидкості геометричної прогресії зі знаменником $\tilde{\gamma}$.

Метод ітераційного агрегування дозволяє будувати лінійні моделі балансів різної розмірності й знаходити розв'язок як у деталізованій, так і в укрупненій системах показників. Очевидно, що його застосування доцільне для систем лінійних алгебраїчних рівнянь великої розмірності. ►

ТЕМА 15. АГРЕГУВАННЯ РЕЗУЛЬТАТІВ ФУНКЦІОНУВАННЯ БАГАТОПРОДУКТОВОГО ВИРОБНИЦТВА ЗА ДОПОМОГОЮ МЕТОДУ АПРОКСИМАЦІЇ ВИРОБНИЧИХ МОЖЛИВОСТЕЙ

Завдання 15.1. *Розкрити зміст методу апроксимації виробничих можливостей багатопродуктової фірми-виробника, який дозволяє отримати єдиний узагальнений продукт (агрегований результат функціонування фірми) і спростити дослідження системних взаємозв'язків та характеристик фірми на виробничо-технологічному рівні.*

◀ Ідея методу апроксимації виробничих можливостей полягає у лінійній апроксимації виробничих можливостей фірми-виробника за допомогою одного обмеження. Це дає можливість зменшити розмірність вихідної задачі та сформувати відповідний агрегований варіант, який буде одним із найпростіших у випадку апроксимації області допустимих можливостей однією гіперплощиною.

Уточнимо зміст цього методу. Нехай $y = (y_1, \dots, y_n)^T$ – вектор випусків (T – транспонування), тобто вектор, компонентами якого є обсяги певних видів продукції чи послуг, що виробляються деякою фірмою-виробником. У просторі випусків (продуктів) область виробничих можливостей може бути описана співвідношенням-нерівністю $\varphi(y) \leq 0$, де φ – деяка зазвичай нелінійна

функція. Вважаючи область виробничих можливостей опуклою, можна у деяких заданих межах апроксимувати її за допомогою однієї гіперплощини, тобто замінити співвідношення $\varphi(y) \leq 0$ співвідношенням

$$\sum_{i=1}^n \alpha_i y_i \leq 1, \quad (15.1.1)$$

де $\alpha = (\alpha_1, \dots, \alpha_n)^T$ – нормований вектор коефіцієнтів

гіперплощини $\langle \alpha, y \rangle = \sum_{i=1}^n \alpha_i y_i$. Коефіцієнт α_i показує,

ту частину потужності виробничого об'єкта, яка зайнята виробництвом i -го виду продукції. Тому, компоненти вектора α – це коефіцієнти взаємного заміщення продукції. Уточнимо цей факт. Якщо $\bar{\varphi}(y) = \langle \alpha, y \rangle = \text{const}$ – гіперплощина рівня, а у випадку фіксації усіх видів продукції, крім i -того та j -того видів, – лінія рівня, то гранична норма заміщення j -того продукту i -тим продуктом

$$\bar{v}_{ij} = -\frac{dy_i}{dy_j} = \frac{\partial \bar{\varphi}(y) / \partial y_j}{\partial \bar{\varphi}(y) / \partial y_i} = \frac{\alpha_j}{\alpha_i},$$

тобто є постійною величиною яка зовсім не залежить від обсягів вироблених продуктів i та j . Зміст заміщення продукції тут такий: величина \bar{v}_{ij} конкретизує граничну кількість одиниць i -того продукту, яка необхідна для заміщення вибулої одиниці j -того продукту, щоб сумарний випуск $\bar{\varphi}(y)$ не змінився.

Незважаючи на те, що побудова апроксимації (15.1.1) практично може бути здійснена різними способами, усі вони пов'язані із визначенням деякого початкового плану та подальшим дослідженням області допустимих планів ($\varphi(y) \leq 0$) в околі цього початкового плану. Критерієм вибору можливого початкового плану може бути, наприклад, максимум випуску продукції у найбільш ймовірному асортименті. Такий вибір формалізується задачею

$$\begin{cases} \beta \rightarrow \max, \\ \varphi(y) \leq 0, \\ y = \beta \bar{y}, \end{cases} \quad (15.1.2)$$

де \bar{y} – вектор найбільш ймовірного асортименту.

Задача (15.1.2) належить до класу оптимізаційних задач [4, 5, 7, 8, 12], інструментарій розв'язування яких сьогодні є достатньо досконалим.

Розв'язок $y^* = (y_1^*, \dots, y_n^*)^T$ задачі (15.1.2) буде тим планом, який використовується для побудови апроксимації (15.1.1) в околі точки $y = y^*$. Якщо y^* належить поверхні $\varphi(y) = 0$, то найпростішою апроксимацією є гіперплощина, дотична до цієї поверхні у точці y^* . Однак доцільніше за апроксимуючу взяти таку гіперплощину, яка перетинає поверхню $\varphi(y) = 0$ та є паралельною до дотичної гіперплощини, оскільки вона буде наближенням поверхні в більш широкій області допустимих випусків продукції.

Нормальним вектором дотичної гіперплощини, що проходить через точку y^* , буде градієнт

$$h^* = (h_1^*, \dots, h_n^*)^T = \left(\frac{\partial \varphi(y^*)}{\partial y_1}, \dots, \frac{\partial \varphi(y^*)}{\partial y_n} \right)^T,$$

тому рівняння цієї гіперплощини формалізується у вигляді $\langle h^*, y - y^* \rangle = 0$ або $\langle h^*, y \rangle = \langle h^*, y^* \rangle$ або

$$\langle h^* / \langle h^*, y^* \rangle, y \rangle = 1, \quad (15.1.3)$$

де $\langle \cdot, \cdot \rangle$ – як і раніше операція скалярного добутку векторів.

Вектор α паралельний векторам h^* та $\alpha^* = h^* / \langle h^*, y^* \rangle$. Очевидно, що вектор α^* задовольняє умову $\langle \alpha^*, y^* \rangle = 1$. Тому з дотичної гіперплощини можна отримати будь-яку паралельну до неї січну площину, помноживши вектор α^* на число, яке є більшим за одиницю. Враховуючи (15.1.3), рівняння такої січної площини матиме вигляд

$$\langle \alpha, y \rangle = 1,$$

$$\text{де } \alpha = (1 + \varepsilon) \alpha^* = \frac{(1 + \varepsilon) h^*}{\langle h^*, y^* \rangle}, \quad \varepsilon - \text{ довільне додатне}$$

число. ►

Завдання 15.2. У просторі випусків продукції область виробничих можливостей двопродуктової фірми-виробника (y_1 та y_2 – обсяги випусків продукції

першого та другого видів) задається співвідношенням-нерівністю

$$\varphi(y) \equiv \varphi(y_1, y_2) \leq 0 \quad (y_1 \geq 0, y_2 \geq 0). \quad (15.2.1)$$

Крім того, відомий $\bar{y} = (\bar{y}_1, \bar{y}_2)^T = (1, 1)^T$ (T – транспонування) – вектор найбільш ймовірного асортименту випусків продукції кожного з видів (тобто $\bar{y}_1 = \bar{y}_2$ або $\bar{y}_1 : \bar{y}_2 = 1:1$). Можливі також коливання асортименту від $\bar{y}_1 : \bar{y}_2 = 1:2$ до $\bar{y}_1 : \bar{y}_2 = 2:1$. Користуючись критерієм максимізації випуску продукції у найбільш ймовірному асортименті (тобто при $\bar{y} = (1, 1)^T$) необхідно (варіанти обмежень 15.2.1 – 15.2.5):

а) знайти початковий план випуску продукції $y^* = (y_1^*, y_2^*)^T$;

б) побудувати у точці y^* рівняння дотичної та рівняння паралельних до неї січних, що проходять через точки максимального випуску при $\bar{y} = (1, 2)^T$ і $\bar{y} = (2, 1)^T$.

15.2.1

$$\varphi(y) = \frac{y_1^2}{9} + \frac{y_2^2}{4} - 1 \leq 0, \quad y_1 \geq 0, \quad y_2 \geq 0$$

15.2.2

$$\varphi(y) = \frac{y_1^2}{16} + \frac{y_2^2}{9} - 1 \leq 0, \quad y_1 \geq 0, \quad y_2 \geq 0$$

15.2.3

$$\varphi(y) = y_1^2 + y_2^2 - 1 \leq 0, \quad y_1 \geq 0, \quad y_2 \geq 0$$

15.2.4

$$\varphi(y) = \frac{y_1^2}{25} + \frac{y_2^2}{25} - 1 \leq 0, \quad y_1 \geq 0, \quad y_2 \geq 0$$

15.2.5

$$\varphi(y) = \frac{y_1^2}{64} + \frac{y_2^2}{49} - 1 \leq 0, \quad y_1 \geq 0, \quad y_2 \geq 0$$

◀ Виконаємо це завдання для варіанту 15.2.1. Область виробничих можливостей задається співвідношеннями-нерівностями

$$\varphi(y) = \frac{y_1^2}{9} + \frac{y_2^2}{4} - 1 \leq 0, \quad y_1 \geq 0, \quad y_2 \geq 0. \quad (15.2.2)$$

Потрібно на основі задачі (15.1.2) при $\bar{y} = (1,1)^T$ вибрати початковий план випуску продукції $y^* = (y_1^*, y_2^*)^T$ та побудувати рівняння дотичної гіперплощини (у даному випадку лінії) у точці y^* й рівняння паралельних до неї січних, які проходять через точки максимального випуску продукції при $\bar{y} = (1,2)^T$ і $\bar{y} = (2,1)^T$.

а) Задача (15.1.2) у цьому конкретному випадку деталізується так:

$$\begin{cases} \beta \rightarrow \max, \\ \varphi(y) = \frac{y_1^2}{9} + \frac{y_2^2}{4} - 1 \leq 0, \\ y_1 \geq 0, \quad y_2 \geq 0, \\ y_1 = \beta, \quad y_2 = \beta. \end{cases} \quad (15.2.3)$$

Підставивши у співвідношення (15.2.2) β замість y_1 і y_2 , зведемо задачу (15.2.3) до задачі вибору максимального додатного значення β на множині $13\beta^2 - 36 \leq 0$, тобто на відрізку $\left[-\frac{6}{\sqrt{13}}, \frac{6}{\sqrt{13}}\right]$.

Очевидно, розв'язком задачі (15.2.3) є $\beta^* = \frac{6}{\sqrt{13}}$ або $y_1^* = y_2^* = \beta^* = \frac{6}{\sqrt{13}}$.

б) Із отриманого результату випливає, що градієнт

$$h^* = \left(\frac{2}{9}y_1^*, \frac{1}{2}y_2^*\right)^T = \left(\frac{4}{3\sqrt{13}}, \frac{3}{\sqrt{13}}\right)^T,$$

а вектор α^* для гіперплощини (у даному випадку прямої), дотичної до $\varphi(y) = 0$ у точці $y = y^*$, відповідно буде таким:

$$\alpha^* = \frac{h^*}{\langle h^*, y^* \rangle} = \left(\frac{2}{3\sqrt{13}}, \frac{3}{2\sqrt{13}}\right)^T.$$

Отже, на площині $y_1 O y_2$ згідно із (15.1.3) рівняння дотичної до еліпса $\frac{y_1^2}{9} + \frac{y_2^2}{4} = 1$ у точці

$$(y_1^*, y_2^*)^T = \left(\frac{6}{\sqrt{13}}, \frac{6}{\sqrt{13}} \right)^T \text{ має вигляд}$$

$$\frac{2}{3\sqrt{13}} y_1 + \frac{3}{2\sqrt{13}} y_2 = 1 \text{ або } y_2 = -\frac{4}{9} y_1 + \frac{2\sqrt{13}}{3}.$$

Безпосередньою перевіркою також легко у цьому переконатися.

Тепер можна розширити область допустимої апроксимації, зсунувши пряму (гіперплощину) паралельно самій собі в області $\varphi(y) < 0$, тобто провести січну, яка буде паралельною до дотичної. Наприклад, можна провести січну через точку $\left(\frac{3}{\sqrt{10}}, \frac{6}{\sqrt{10}} \right)$, що відповідає максимальному випуску продукції у асортименті $1:2$ ($\bar{y} = (1, 2)^T$), або через точку $\left(\frac{12}{5}, \frac{6}{5} \right)$, що відповідає максимальному випуску продукції у асортименті $2:1$ ($\bar{y} = (2, 1)^T$) (для знаходження цих точок потрібно розв'язати задачу (15.2.3), у якій замість умов $y_1 = \beta$, $y_2 = \beta$ потрібно використати умови $y_1 = \beta$, $y_2 = 2\beta$ та умови $y_1 = 2\beta$, $y_2 = \beta$). Обидві прямі будуть проходити через зазначені точки й мати кутовий коефіцієнт, що дорівнює $\left(-\frac{4}{9} \right)$ (їх рівняння такі:

$y_2 = -\frac{4}{9}y_1 + C_1$ і $y_2 = -\frac{4}{9}y_1 + C_2$, де C_1 і C_2 – константи, які знаходяться із умови проходження прямої через конкретно задану точку). Наприклад, пряма, що проходить через точку $\left(\frac{3}{\sqrt{10}}, \frac{6}{\sqrt{10}}\right)$, має рівняння

$$\left(C_1 = \frac{22}{3\sqrt{10}}\right)$$

$$y_2 = -\frac{4}{9}y_1 + \frac{22}{3\sqrt{10}} \text{ або } \frac{2\sqrt{10}}{33}y_1 + \frac{3\sqrt{10}}{22}y_2 = 1.$$

При цьому

$$\alpha = (1+\varepsilon)\alpha^* = (1+\varepsilon)\left(\frac{2}{3\sqrt{13}}, \frac{3}{2\sqrt{13}}\right)^T = \left(\frac{2\sqrt{10}}{33}, \frac{3\sqrt{10}}{22}\right)^T,$$

$$\text{тобто } (1+\varepsilon) = \frac{\sqrt{10}\sqrt{13}}{11}, \text{ а } \varepsilon = \frac{\sqrt{10}\sqrt{13}}{11} - 1 \approx 0,0365.$$

Зауважимо, що ця пряма проходить нижче дотичної (оскільки $\frac{22}{3\sqrt{10}} < \frac{2\sqrt{13}}{3}$) і перетинає еліпс у двох точках,

що належать дузі еліпса, розміщеної у першій чверті площини y_1Oy_2 (тобто ця пряма є наближення еліпса у більш широкій області допустимих випусків продукції).

На завершення зазначимо таке. У межах допустимості апроксимації (15.1.1) для виробничого об'єкта (у контексті завантаження потужностей) конкретний набір випусків не важливий. Принципово лише, щоб цей набір задовольняв обмеження (15.1.1). З огляду на це, варто розглядати єдиний агрегований

(узагальнений) продукт $\langle \alpha, y \rangle$ і на його основі вивчати взаємозв'язки об'єктів. У цьому й полягає зміст агрегування як процедури укрупнення продуктів. ►

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Григорків В.С. Економетрика: Лінійні моделі парної та множинної регресії: навч. посібник. Чернівці: Чернівецьк. нац. ун-т, 2009. 224 с.
2. Григорків В.С. Економічна кібернетика: Навчальний посібник. Чернівці: Рута, 2006. 198 с.
3. Григорків В.С. Моделювання економіки: підручник. Чернівці : Чернівецьк. нац. ун-т, 2019. 360 с.
4. Григорків В.С. Оптимальне керування в економіці: навч. посібник. Чернівці: Чернівецьк. нац. ун-т, 2011. 200с.
5. Григорків В.С. Оптимізаційні методи та моделі: вибрані завдання для тематичного контролю: навч. посіб. / В.С. Григорків, М.В. Григорків, О.І. Ярошенко, О.Ю. Вінничук, Л.В. Скращук. Чернівці: Чернівецьк. нац. ун-т, 2022. 168 с.
6. Григорків В.С., Григорків М.В. Еластичність та її застосування в економічному аналізі: Навчальний посібник. Чернівці : ТОВ «ДрукАрт», 2008. 64 с.
7. Григорків В.С., Григорків М.В. Моделі прийняття рішень в економіці: навч. посібник. Чернівці: Чернівецьк. нац. ун-т, 2021. 256 с.
8. Григорків В.С., Григорків М.В. Моделі прийняття рішень в економіці: вибрані завдання для

тематичного контролю: навч. посіб. Чернівці : Чернівець. нац. ун-т, 2021. 48 с.

9. Григорків В.С., Григорків М.В., Скращук Л.В. Диференціальні моделі економічної динаміки: основи теорії та приклади: навч. посібник. Чернівці: Чернівець. ун-т, 2015. 224 с.

10. Григорків В.С., Григорків М.В., Ярошенко О.І. Оптимізаційні методи та моделі: підручник. Чернівці: Чернівець. нац. ун-т, 2022. 440 с.

11. Григорків В.С., Ярошенко О.І. Дискретні моделі економічної динаміки: навч. посібник. Чернівці: Чернівецький нац. ун-т, 2014. 96 с.

12. Григорків М.В. Оптимізаційні методи та моделі: тестові завдання: навч. посібник / М.В. Григорків, В.С. Григорків, О.І. Ярошенко, О.Ю. Вінничук, Л.В. Скращук. Чернівці: Чернівець. нац. ун-т, 2022. 136 с.

13. Зацеркляний М.М., Мельников О.Ф. Основи економічної кібернетики: Навч. посібник. Чернівці: ТОВ Видавництво «Наші книги», 2008. 392 с.

14. Інформаційні системи і технології в економіці: Посібник для студентів вищих навчальних закладів / За редакцією В.С. Пономаренка. К.: Видавничий центр «Академія», 2002. 544 с.

15. Катренко А.В. Системний аналіз об'єктів та процесів комп'ютеризації: Навчальний посібник. Львів: «Новий світ». 2000. 424 с.

16. Машина Н.І. Математичні методи в економіці: навчальний посібник. Київ: Центр навчальної літератури, 2003. 148 с.

17. Пономаренко О.І., Пономаренко В.О. Системні методи в економіці, менеджменті та бізнесі: Навч. посібник. К. Либідь, 1995. 240 с.

18. Снарський, А.О. Основи теорії складних мереж [Електронний ресурс] : навчальний посібник / Снарський А.О., Ланде Д.В., Субач І.Ю. ; ІСЗЗІ КПІ ім. Ігоря Сікорського. – Електронні текстові дані (1 файл: 6,31 Мбайт). Київ : ІСЗЗІ КПІ ім. Ігоря Сікорського, 2023. 225 с.

19. Томашевський В.М. Моделювання систем. К.: Видавнича група ВНУ, 2005. 352 с.

20. Ярошенко О.І. Дискретний аналіз: вибрані завдання для тематичного контролю: навч. посібник. Чернівці: Чернівець. ун-т, 2021. 104 с.

Навчальне видання

Григорків Василь Степанович
Григорків Марія Василівна

ЕКОНОМІЧНА КІБЕРНЕТИКА
ПРАКТИКУМ

Навчальний посібник

Відповідальний за випуск *В.С. Григорків*

Літературний редактор *О.В. Лукул*

Технічний редактор
та дизайн обкладинки *О.М. Кудрінська*

Підписано до друку 30.04.2024. Формат 60x84/16
Папір офсетний. Друк різнографічний. Ум.-друк. арк. 11,0.
Обл.-вид. арк. 11,8. Зам. Н-035.
Видавництво та друкарня Чернівецького національного університету
імені Юрія Федьковича
58002, Чернівці, вул. Коцюбинського, 2
e-mail: ruta@chnu.edu.ua

Свідоцтво суб'єкта видавничої справи ДК № 891 від 08.04.2002



ISBN 978-966-423-844-8



9 789664 238448