

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»

# МАТЕМАТИЧНІ МЕТОДИ ОПТИМІЗАЦІЇ ЧАСТИНА 1. СТАТИЧНІ СИСТЕМИ

## Навчальний посібник

Рекомендовано Методичною радою КПІ ім. Ігоря Сікорського  
як навчальний посібник для здобувачів ступеня бакалавра  
за освітньою програмою «Інженерія інноваційних інформаційно-телекомунікаційних  
технологій та систем»  
спеціальності G5 Електроніка, електронні комунікації, приладобудування та радіотехніка

Електронне мережеве навчальне видання

Київ  
КПІ ім. ІГОРЯ СІКОРСЬКОГО  
2025

УДК 621  
М11

Укладачі: *Лисенко Олександр Іванович*, д-р техн. наук, проф.  
*Тачиніна Олена Миколаївна*, д-р техн. наук, проф.  
*Алексєєва Ірина Віталіївна*, канд. фіз.-мат н., доц.  
*Явіся Валерій Сергійович*, канд. техн. наук, доц.  
*Новіков Валерій Іванович*, канд. техн. наук.  
*Михайло Михайлович Потьомкін*, д-р техн. наук, проф.  
*Станіслав Вікторович Валуйський*, канд. техн. наук.

Рецензент *Носков В.І.*, канд. техн. наук, доц., доцент кафедри електронних комунікацій та інтернету речей НН ІТС

Відповідальний редактор *Руренко О.Г.*, канд. фіз.-мат. наук

*Гриф надано Методичною радою КПІ ім. Ігоря Сікорського  
(протокол № 2 від 06.11.2025 р.)  
за поданням вченої ради Навчально-наукового інституту телекомунікаційних систем  
(протокол № 8 від 29.09.2025 р.)*

**М11 Математичні методи оптимізації.** Частина 1. Статичні системи [Електронний ресурс]: навч. посіб. для здобувачів ступеня бакалавра за освіт. програмою «Інженерія інноваційних інформаційно-телекомунікаційних технологій та систем» спец. G5 Електроніка, електронні комунікації, приладобудування та радіотехніка / КПІ ім. Ігоря Сікорського; уклад.: О.І. Лисенко та ін. – Електрон. текст. дані (1 файл). – Київ: КПІ ім. Ігоря Сікорського, 2025. – 360 с.

Метою посібника «Математичні методи оптимізації. Частина 1. Статичні системи» є навчити здобувачів ступеня магістра зі спеціальності G5 Електроніка, електронні комунікації, приладобудування та радіотехніка методам синтезу цифрових алгоритмів автоматичного керування процесом функціонування мобільних компонент телекомунікаційних систем та мереж. Навчальний посібник призначено для вивчення дисципліни «Математичні методи оптимізації» для здобувачів ступеня магістра зі спеціальності G5 Електроніка, електронні комунікації, приладобудування та радіотехніка.

УДК 621

Реєстр. № НП 25/26-073. Обсяг 23,6 авт. арк.

Національний технічний університет України  
«Київський політехнічний інститут імені Ігоря Сікорського»  
проспект Берестейський, 37, м. Київ, 03056  
<https://kpi.ua>

Свідоцтво про внесення до Державного реєстру видавців, виготовлювачів і розповсюджувачів видавничої продукції ДК № 5354 від 25.05.2017 р.

© КПІ ім. Ігоря Сікорського, 2025

## ЗМІСТ

Вступ.....	8
Розділ 1. Класичні математичні методи оптимізації .....	11
Глава 1. Загальна методологія математичного програмування та дослідження операцій.....	11
1.1. Основні поняття та означення. Ефективність операцій.....	11
1.2. Математичні моделі операцій.....	13
Глава 2. Структура методики дослідження операцій .....	21
2.1. Складові методики дослідження операцій .....	21
2.2. Типові класи задач дослідження операцій .....	24
Глава 3. Предмет математичного програмування .....	29
3.1. Загальна задача математичного програмування.....	29
3.2. Класифікація задач математичного програмування.....	29
3.3. Теорема про достатні умови глобального максимуму.....	33
Глава 4. Типові задачі математичного програмування .....	37
4.1. Приклади типових задач математичного програмування.....	37
Глава 5. Задачі лінійного програмування .....	50
5.1. Поняття про лінійне програмування .....	50
5.2. Форми запису задачі лінійного програмування.....	50
5.3. Геометрична інтерпретація задач лінійного програмування.....	52
5.4. Графічний метод розв'язання задач лінійного програмування.....	53
Глава 6. Симплекс-метод розв'язання задачі лінійного програмування.....	63
6.1. Поняття про симплекс-метод.....	63
6.2. Зведення стандартної форми обмежень нерівностей до обмежень рівностей (рівнянь обмежень) основної задачі лінійного програмування.....	64
6.3. Основні прийоми та способи симплекс-методу розв'язання задач лінійного програмування.....	66
Глава 7. Алгоритм пошуку опорних розв'язків основної задачі лінійного програмування .....	74
7.1. Табличний алгоритм заміни базисних змінних (стандартні таблиці) .....	74
7.2. Алгоритм пошуку опорного розв'язку за допомогою стандартних таблиць .....	78
Глава 8. Алгоритм пошуку оптимального розв'язку основної задачі лінійного програмування .....	84

8.1. Алгоритм пошуку оптимального розв'язку за допомогою стандартних таблиць .....	84
8.2. Розв'язання задачі лінійного програмування з використанням системи комп'ютерної математики Matlab.....	90
Глава 9. Транспортна задача лінійного програмування.....	92
9.1. Постановка транспортної задачі .....	92
9.2. Табличний метод пошуку опорного плану (метод «північно-західного кута»).....	96
9.3. Метод пошуку оптимального плану за допомогою циклічного перерахунку (метод циклічного перерахунку або розподільний метод) .....	100
Глава 10. Метод потенціалів розв'язання транспортної задачі.....	106
10.1. Обґрунтування методу потенціалів.....	106
10.2. Побудова потенціального плану.....	109
10.3. Методика розв'язання транспортної задачі методом потенціалів .....	111
Глава 11. Деякі окремі випадки транспортних задач .....	119
11.1. Транспортні задачі з неправильним балансом .....	119
11.2. Розв'язок транспортної задачі за критерієм часу .....	124
Глава 12. Післяоптимізаційний аналіз задачі лінійного програмування .....	130
12.1. Ідея аналізу параметричної чутливості.....	130
12.2. Ідея фізичного змісту побудови математичної моделі двоїстої задачі лінійного програмування.....	132
12.3. Загальна постановка і правила побудови двоїстої задачі .....	132
12.4. Основні теореми двоїстості.....	137
Глава 13. Аналіз розв'язку задачі лінійного програмування.....	141
13.1. Аналіз розв'язку задачі лінійного програмування на основі двоїстих оцінок.....	141
13.2. Аналіз розв'язку задачі лінійного програмування на основі використання параметричного програмування.....	148
Глава 14. Цілочислові задачі лінійної оптимізації .....	155
14.1. Постановка цілочислових задач лінійної оптимізації .....	155
14.2. Метод Гоморі розв'язання цілочислових задач лінійного програмування .....	155
14.3. Метод гілок і границь розв'язання цілочислової задачі лінійного програмування .....	161
Глава 15. Нелінійне програмування .....	166
15.1. Класичні умови екстремуму задачі нелінійного програмування.....	166

Глава 16. Найпростіша задача нелінійного програмування в умовах невід’ємності змінних .....	182
Глава 17. Задачі опуклого та квадратичного програмування .....	191
17.1. Умови та теорема Куна-Таккера .....	191
17.2. Квадратичне програмування .....	194
Глава 18. Огляд основних підходів до побудови числових методів розв’язання задач нелінійного програмування .....	200
18.1. Модельно-тестовий приклад .....	200
18.2. Метод сканування (метод повного перебирання) .....	201
18.3. Метод Гаусса – Зейделя (метод покоординатного спуску) .....	201
18.3.1. Класичний метод мінімізації функції однієї змінної у разі його застосування в методі Гаусса–Зейделя .....	203
18.3.2. Метод екстраполяційного пошуку .....	204
18.4. Метод градієнта .....	207
18.5. Метод найшвидшого спуску .....	211
18.6. Метод випадкового пошуку .....	213
18.7. Застосування системи комп’ютерної математики MATLAB для розв’язка задач квадратичного та нелінійного програмування .....	214
Розділ 2. Еволюційні методи оптимізації та підходи до їх вдосконалення	226
Глава 19. Еволюційні методи оптимізації .....	226
19.1. Генетичні алгоритми .....	228
19.2. Метод LARES .....	233
19.3. Метод моделювання відпалу .....	235
19.4. Метод оптимізації колонією бджіл .....	238
19.5. Метод оптимізації роєм частинок .....	242
19.6. Метод оптимізації мурашиною колонією .....	245
19.7. Метод оптимізації колонією бактерій .....	250
Глава 20. Методичні підходи до вдосконалення еволюційних методів оптимізації та приклади їх використання .....	256
20.1. Аналіз напрямків удосконалення еволюційних методів оптимізації ...	256
20.2. Удосконалений генетичний алгоритм .....	259
20.3. Удосконалений метод оптимізації роєм частинок .....	262
20.4. Удосконалений метод оптимізації колонією бджіл .....	265
Розділ 3. Приклади практичного застосування математичних методів оптимізації в телекомунікаціях та риадіотехніці .....	271
Глава 21. Моделі побудови та функціонування епізодичних радіомереж із БПЛА в зоні надзвичайної ситуації .....	271

21.1. Особливості побудови та функціонування епізодичних радіомереж на основі БПЛА .....	271
21.2. Аналіз задач управління мережею БПЛА в зоні надзвичайної ситуації .....	281
21.3. Аналіз методів підвищення пропускної здатності епізодичних радіомереж із застосуванням БПЛА.....	284
21.4. Загальна постановка задачі підвищення пропускної здатності епізодичних радіомереж із застосуванням БПЛА .....	286
21.5. Висновки до глави 21.....	291
Глава 22. Математичні моделі оцінки зв'язності вузлів епізодичних радіомереж із БПЛА в зоні надзвичайної ситуації.....	293
22.1. Вихідні дані та обмеження .....	293
22.2. Методика оцінки зв'язності вузлів епізодичних радіомереж із БПЛА	294
22.3. Математична модель оцінки максимальної дальності радіозв'язку.....	295
22.4. Формування зони обслуговування БПЛА .....	301
22.5. Розрахунок тривалості зв'язності рятувальників в зоні надзвичайної ситуації .....	308
22.5.1 Розрахунок на випадок прямої радіо видимості між вузлами.....	309
22.5.2 Розрахунок на випадок ретрансляції через проміжний вузол.....	310
22.5.3 Моделювання переміщення вузлів мережі.....	313
22.6. Висновки до глави 22.....	315
Глава 23. Математичні моделі оцінки показників функціонування епізодичних радіомереж із БПЛА в зоні надзвичайної ситуації.....	319
23.1. Вихідні дані, обмеження та допущення.....	319
23.2. Методика оцінки показників функціонування епізодичних радіомереж із БПЛА .....	320
23.3. Розрахунок інтенсивності трафіку в каналах епізодичних радіомереж із БПЛА .....	320
23.4. Дослідження пропускної здатності каналів епізодичних радіомереж із БПЛА .....	322
23.5. Дослідження затримки передачі в каналах епізодичних радіомереж із БПЛА .....	331
23.6. Визначення кількості абонентів в стільниках МА та ТА.....	336
23.7. Висновки до глави 23.....	338
Глава 24. Метод підвищення пропускної здатності епізодичних радіомереж з управлінням положенням БПЛА в зоні надзвичайної ситуації.....	340
24.1. Постановка задачі у математичному вигляді .....	340

24.2. Опис методу.....	341
24.3. Удосконалений алгоритм пошуку квазіоптимального положення БПЛА .....	346
24.4. Оцінка ефективності запропонованого методу.....	353
24.5. Висновки до глави 24.....	357
Навчально-методичні матеріали.....	359

## Вступ

Розумне (раціональне або оптимальне) розв'язання складних наукових та науково-прикладних(технічних) завдань значно спрощується при застосуванні математичних методів, що базуються на математичному моделюванні, тобто заміщенні одних об'єктів іншими, що забезпечують збереження найбільш суттєвих властивостей і особливостей тих об'єктів, які заміщують. Математичні методи використовуються як для аналізу, так і для синтезу складних об'єктів та процесів. В сучасних умовах математичні методи знаходить найширше застосування у багатьох областях науки і техніки у тому числі і при наукових дослідженнях у телекомунікація та радіотехніці, тобто досліджені явищ, які мають місце при проектуванні та експлуатації як засобів телекомунікацій та радіотехніки, так і телекомунікаційних систем та мереж. За допомогою математичних методів вдається знаходити як найкращі (екстремальні), так і прийнятні рішення. Математичні методи оптимізації в телекомунікаціях та радіотехніці базуються на методах дослідження операцій, які включають в себе методи математичного моделювання і пошуку найкращих форм структури телекомунікаційних систем і мереж та значень їх параметрів. Математичні методи оптимізації складають фундамент прикладної математичної підготовки наукових магістрів та докторів філософії зі спеціальності G5 Електроніка, електронні комунікації, приладобудування та радіотехніка.

Навчальний посібник дозволяє здобувачам із більшою якістю засвоювати матеріал навчальної дисципліни «Математичні методи оптимізації».

Для успішного засвоєння дисципліни здобувач повинен володіти знаннями з математичного аналізу, аналітичної геометрії, теорії матриць, диференціальних рівнянь, теорії ймовірностей та математичної статистики.

На результатах вивчення даної дисципліни базуються виконання магістерської дисертації та науково-дослідна практика, а також навчальні дисципліни для здобуття універсальних компетентностей дослідника (вибіркові).

На результатах вивчення даної дисципліни базуються математичні постановки задач та методи їх розв'язання, що досліджуються у дисертаціях здобувачів.

У посібнику матеріал викладається згідно силабусу і стосується вивчення: елементів теорії детермінованого лінійного та нелінійного програмування; методів чисельного пошуку екстремумів опуклих функцій на опуклих множинах; методів параметричного та цілочисельного програмування; термінології та визначення основних понять, символічного позначення основних операцій та їх змісту, що використовуються в теорії математичного програмування, теорії прийняття рішень, теорії ігор, теорії управління запасами і при імітаційному моделюванні, в обсязі достатнім для застосування базових знань, вмінь та навичок по основам побудови, ідентифікації і практичного використання оптимізаційних математичних моделей явищ, які мають місце в телекомунікаціях та радіотехніці.

Вивчення наведеного у посібнику матеріалу призводить до формування: фундаментальних теоретичних знань з математичного моделювання та оптимізації, які дозволяють виконувати аналіз та синтез складних телекомунікаційних систем та мереж, а також їх окремих елементів; прикладних практичних навичок застосування систем комп'ютерної математики типу MATLAB для побудови комп'ютерних математичних моделей та кількісного розв'язання оптимізаційних задач як на попередніх етапах проектування телекомунікаційних систем, пристроїв та засобів телекомунікацій, так і у реальному часі.

Навчальний матеріал у посібнику структурований за главами і підрозділами, у яких рельєфно виділено визначення, зауваження, формулювання теорем та їх доведення, приклади та висновки. Розділи є логічно завершеними, відносно самостійними, цілісними частинами навчальної дисципліни, засвоєння яких передбачає вивчення теоретичного матеріалу, відповіді на контрольні питання. проведення модульної контрольної роботи та аналіз результатів її виконання.

Посібник дозволяє студентам вивчити:

- термінологію, визначення, основні поняття, символічне позначення основних операцій та їх зміст;
- постановки задач математичного програмування, критерії та методи технічної та економічної оптимізації складних систем, що найбільш поширені в телекомунікаціях;

- математичні методи в задачах лінійного та нелінійного програмування та сформувані прикладні практичні навички стосовно фізичної інтерпретації результатів обчислювальних експериментів та їх візуалізації;
- інформаційні технології розв'язання класичних задач лінійного, нелінійного, дискретного (цілочисельного) програмування та дослідження операцій;
- прийоми формалізації задач структурно-функціонального аналізу, загальної стратегії їх розв'язання та системної оптимізації складних конструктивних елементів в телекомунікаціях.

В результаті вивчення викладеного у посібнику матеріалу студенти будуть вміти:

- виконувати аналітичні дослідження та комп'ютерне математичне моделювання складних систем та комплексів в телекомунікаціях, а також їх окремих компонентів;
- програмувати математичні моделі складних систем та комплексів в телекомунікаціях, а також їх окремих компонентів із використанням програмного забезпечення системи комп'ютерної математики MATLAB.

Навчальний посібник надає глибинні прикладні математичні знання і тим самим забезпечує достатній рівень математичної підготовки для написання магістерської дисертації.

Особистий вклад укладачів у розділі посібника.

Глави 1- 17: предметна фізична постановка завдання, приклади постановки та розв'язання задач у середовищі MATLAB – Лисенко О.І.; побудова аналітичних моделей – Алексєєва І.В.; приклади постановки та розв'язання задач у середовищі MATLAB – Тачиніна О.М.

Глави 18, 21 - 24: предметна фізична постановка завдання, побудова аналітичних моделей – Лисенко О.І.; приклади постановки та розв'язання задач у середовищі MATLAB – Явіся В.С.; імітаційне моделювання та обчислювальний експеримент у середовищі MATLAB + Simulink – Новіков В.І.; експериментальні дослідження на навчальному лабораторному обладнанні та із використанням реально функціонуючих телекомунікаційних систем – Валуйський С.В.

Глави 19 та 20: предметна фізична постановка завдання, побудова аналітичних моделей ; приклади постановки та розв'язання задач у середовищі MATLAB – Потьомкін М.М.

## **Розділ 1. Класичні математичні методи оптимізації**

### **Глава 1. Загальна методологія математичного програмування та дослідження операцій**

#### **1.1. Основні поняття та означення. Ефективність операцій**

**Означення 1.1.** *Операція* – будь-яка дія або система дій, об'єднаних єдиним задумом і спрямованих на досягнення певної мети.

**Означення 1.2.** Конкретну реалізацію послідовності дій у просторі і часі називають *сценарієм операції*.

**Означення 1.3.** Під *дослідженням операцій в автоматизованих системах на транспорті* розуміють застосування математичних методів для кількісного обґрунтування рішень стосовно розвитку автоматизованих комп'ютерно-інтегрованих технологій та підвищення ефективності їх використання.

Приклади дослідження операцій в автоматизованих системах на транспорті:

- система заходів щодо підвищення надійності автоматизованих систем на транспорті та технічних засобів автоматизації;
- система заходів зниження інформаційних ризиків (пошкодження інформації, знищення інформації, несанкціонований доступ до конфіденційної інформації);
- розміщення автоматизованих робочих місць;
- розміщення замовлень на виготовлення та технічних засобів автоматизації;
- створення системи гарантійного обслуговування автоматизованих систем на транспорті;
- оптимізація параметрів та структури автоматизованих систем на транспорті.

**Означення 1.4.** Характеристики дій (фактори операцій), які неможливо змінювати в умовах операцій, вважають *параметрами операцій*.

**Означення 1.5.** Характеристики дій, унаслідок змінювання яких виявляється позитивний результат, називають *змінними* або *керувальними параметрами*.

Отже, операціями завжди можна керувати. Будь-який вибір залежно від дослідника керувальних параметрів має назву розв'язок (рішення, план операції чи стратегія).

Розв'язки можуть бути такими, що задовольняють чи не задовольняють технічні та (або) економічні, та (або) екологічні, та (або) соціальні обмеження, і їх відповідно називають *допустимими* або *недопустимими*. Зрозуміло, що останніми нехтують, а серед перших обирають оптимальні, тобто ті, які у певному розумінні мають перевагу над іншими.

**Основна задача дослідження операцій** – попереднє кількісне обґрунтування оптимальних рішень. Під прийняттям рішення в технічних системах розуміють остаточне затвердження всіх дій, які входять до складу операції на основі запропонованого оптимального розв'язку та з урахуванням додаткової інформації. Існує поняття «особа, що приймає рішення» (ОПР).

Особа, що приймає рішення – це фізична особа або колектив фахівців. Для прийняття рішення ОПР використовує спеціальні методи й методики. Методи дослідження операцій є теоретичною основою процесу підготовки та прийняття рішення. Поряд з основною задачею (обґрунтування оптимальних розв'язків) під час дослідження операцій виконують додаткові задачі:

- порівняльний аналіз різних варіантів організації операцій;
- оцінювання впливу на результат операцій різних змін параметрів операцій;
- дослідження операцій для виявлення критичних елементів (критичних параметрів операцій), тобто таких складових операцій, порушення запланованого функціонування яких призводить до істотного погіршення остаточних результатів операцій.

Допоміжні задачі набувають особливого значення, коли дану операцію необхідно розглядати не ізольовано, а як елемент цілісної системи операцій. Дослідження цілісної системи операцій ґрунтується на так званому системному підході, який потребує комплексного врахування взаємної залежності та обумовленості складових елементів системи операцій. Для отримання раціональних рішень потрібно вміло використовувати як прийоми агрегування, тобто об'єднання кількох операцій, так і їх декомпозиції, тобто розбиття на окремі складові.

Припустімо, що нам вдалося виділити окрему операцію. Метою операції є отримання її найбільшої ефективності. Під ефективністю операції розуміють ступінь її налаштованості на виконання поставленої задачі. Що краще організована операція, то вона ефективніша.

**Означення 1.6.** *Критерієм оцінювання або показником ефективності, або цільовою функцією  $W$  називається функція фізичної величини чи сукупності фізичних величин, обчислення або вимірювання яких дозволяє кількісно оцінити результат виконання операції.*

Порівнюючи показники ефективності для різного складу дій, потрібно обрати найкращий спосіб для організації операції. Показник ефективності може бути заданий аналітично, алгоритмічно або його можна вимірювати. Конкретний вигляд показника ефективності, яким слід користуватися при числовій оцінці ефективності, залежить від спеціальних властивостей конкретної операції, її цільової спрямованості, а також від задачі дослідження. Ця задача може бути поставлена в різних формах: детермінованій або з урахуванням елементів невизначеності.

## **1.2. Математичні моделі операцій**

### **1.2.1. Загальні відомості про математичні моделі операцій**

Загальних способів побудови математичних моделей операцій не існує. У кожному випадку модель будують виходячи із цільової спрямованості операцій і задачі наукового дослідження з урахуванням необхідної точності рішення, а також з точністю, з якою відомі вихідні дані. Під час побудови математичної моделі спрощуване явище (в даному випадку операція) схематизується. Із безлічі факторів, що впливають на явище, виокремлюють порівняно невелику кількість найважливіших, а потім отриману схему опису явища подають у найбільш відповідних для неї математичних термінах (із використанням відповідного математичного апарату). У результаті встановлюються кількісні зв'язки між умовами операції, її параметрами та результатом операції, який оцінюється показником ефективності або сукупністю показників, якщо їх в даній задачі кілька.

Побудова математичної моделі – це мистецтво, яке вдосконалюється з досвідом. У складних випадках, коли немає єдиного рішення щодо математичної моделі операції, використовується так званий прийом

конкуренції моделей. Тобто, одне і те саме явище досліджується з використанням кількох математичних моделей. Після порівняння результатів теоретичних досліджень з експериментальними даними обирається найкраща за критерієм мінімального відхилення від експериментальних даних модель.

Для складних задач дослідження операцій характерною особливістю є корекція математичної моделі після кожного циклу дослідження. Побудова математичної моделі – найбільш важлива і відповідальна частина дослідження, яка потребує як глибоких математичних знань, так і знань предметної галузі, тобто суті модельованого явища.

Математичні моделі, які використовуються в задачах дослідження операцій, можна поділити на два класи:

- 1) аналітичні;
- 2) алгоритмічні.

Перші встановлюють формульні (аналітичні) залежності між показником ефективності, змінними і параметрами задачі. Зазвичай ці залежності записують у вигляді рівнянь або нерівностей: алгебричних, звичайних диференціальних або із частковими похідними. За допомогою аналітичних моделей вдається із задовільною точністю описати лише прості операції, де кількість елементів, що взаємодіють, порівняно невелика.

В операціях значного масштабу, де взаємодіють багато елементів, що перебувають під дією випадкових факторів, доцільно застосовувати алгоритмічні математичні моделі, які дозволяють виконувати імітаційне моделювання операцій. Суть цього моделювання полягає в тому, що процес розвитку операції імітується на комп'ютері з усіма випадковостями, які цю операцію супроводжують.

Найкращі результати дослідження операцій можна отримати у разі сумісного застосування аналітичних та імітаційних моделей: проста аналітична модель дозволяє визначити основні закономірності явища, з'ясувати напрям його розвитку. Подальше уточнення кількісних значень показника ефективності операції можливо досягти імітаційним моделюванням.

## 1.2.2. Детерміновані аналітичні моделі операцій

Розглянемо задачу дослідження операції в загальній постановці, тобто без врахування специфічних властивостей конкретної операції та її цілей. Припустимо, що необхідно виконати деяку операцію, тобто керований захід, на остаточний результат якого можливо вплинути обираючи змінні характеристики цієї операції  $x_1, x_2, \dots, x_n$ .

Ефективність операції оцінюється за допомогою чисельного показника  $W$ . Оптимальним розв'язком задачі вважаються ті значення  $x_1, x_2, \dots, x_n$ , при яких  $W$  досягає глобального максимуму. В літературі позначають:

$$M = \sup_{X \in G} W,$$

$X = [x_1, \dots, x_n]^T$ ,  $G$  – область допустимих розв'язків (ОДР).

В тому випадку коли необхідно досягти глобального мінімуму:

$$m = \inf_{X \in G} W.$$

Цю задачу можливо переформулювати в задачу пошуку глобального максимуму, розглядаючи замість показника ефективності  $W$  показник ефективності  $-W$ .

Якщо математична модель побудована, то це означає:

- 1) відомі фактори операції  $\alpha_1, \alpha_2, \dots, \alpha_m$ , на які дослідники впливати не можуть;
- 2) визначено фізичний зміст і математична формалізація факторів операції  $x_1, x_2, \dots, x_n$ , на які дослідники можуть впливати;
- 3) визначено спосіб отримання чисельного значення показника ефективності операції  $W$ , яке буде залежати від  $\alpha_1, \alpha_2, \dots, \alpha_m$  та  $x_1, x_2, \dots, x_n$ ;
- 4) з'ясовано ОДР  $G$ , тобто умови обмежень, що діють в даній операції

$$F_j: x_1, x_2, \dots, x_n \leq 0, = 0, \geq 0 \quad j = \overline{1, q}.$$

Загальна детермінована поставка задачі дослідження операції формулюється наступним чином: при заданих умовах (4) знайти такі

розв'язки (2), при яких показник ефективності (3) досягає глобального максимуму.

Специфічним для детермінованої задачі дослідження операції, порівняно із задачами пошуку екстремумів гладкої функції (допускається існування похідних та частинних мішаних похідних високих порядків), є наявність обмежень-нерівностей та недиференційованість, в деяких випадках, показника ефективності за елементами розв'язку  $x_1, x_2, \dots, x_n$ .

Загальних математичних методів знаходження екстремумів функцій при наявності довільних обмежень не існує, але коли показник ефективності та обмеження мають специфічні властивості, існують спеціальні методи розв'язання таких задач. Якщо показник ефективності  $W$  залежить від  $x_1, x_2, \dots, x_n$  лінійно і обмеження (4) мають вид лінійних алгебраїчних рівнянь або нерівностей, то глобальний максимум  $W$  знаходимо із використанням спеціального математичного апарату, що носить назву *лінійне програмування*. Якщо обмеження (4) та показник ефективності  $W$  мають інші властивості (наприклад, опуклі, сепарабельні, квадратичні), застосовується математичний апарат *опуклого або сепарабельного, або квадратичного програмування*. Якщо операція, за фізичним змістом операції, розділяється на декілька кроків чи станів, а показник ефективності  $W$  дорівнює сумі показників  $W_k$ , досягнутих на окремих етапах виконання операції, то для пошуку оптимального рішення можливо застосовувати метод *динамічного програмування*.

### ***Зауваження 1.1.***

1. Для розв'язання задачі пошуку екстремуму показника ефективності  $W$  зазвичай використовують числові методи, які попередньо перевіряють (тестують) під час розв'язання задач пошуку екстремумів функцій зі спеціальними властивостями (наприклад функції Розенброка, рис.1.1). Ці функції дають змогу з'ясувати досконалість числових методів з погляду уникнення зациклювань та спроможності обходити так звані «складки» та точки стаціонарності.

2. Показник ефективності може мати на ОДР  $G$  декілька локальних екстремумів (рис.1.2).

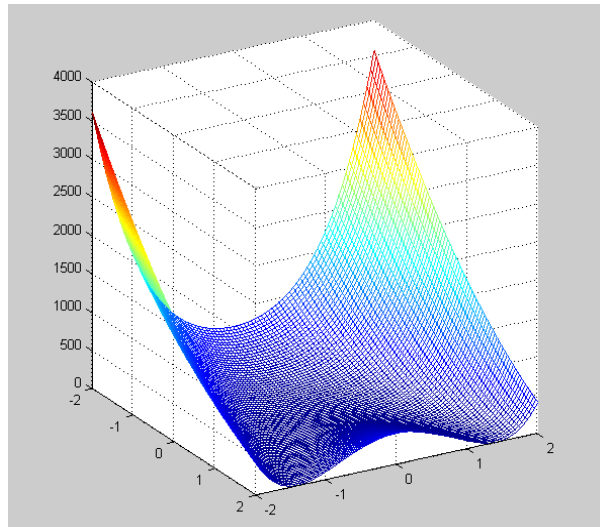


Рис. 1.1. Графічне зображення функції Розенброка

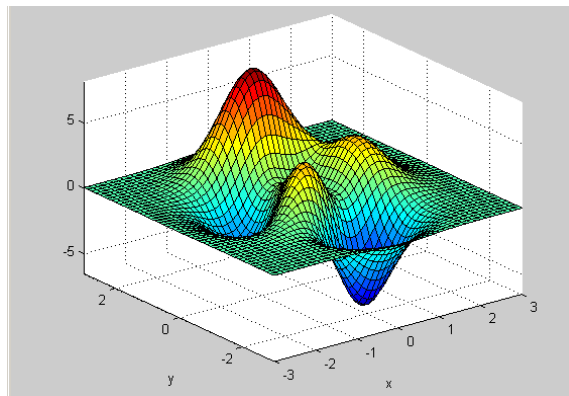


Рис. 1.2. Графічне зображення багато екстремального показника ефективності

### 1.2.3. Математичні моделі операцій з урахуванням невизначеності

Типовим для практики є випадок, коли не всі параметри операції, що не може змінювати дослідник, відомі, тобто є такі з них, які включають в себе елемент невизначеності. Наприклад, невизначеність параметрів каналів зв'язку може бути викликана змінами стану атмосфери (зміни погоди або сонячної активності) або дією штучних електромагнітних завад в телекомунікаційних системах спеціального призначення.

В умовах невизначеності до складу математичної моделі операції входять 3-и категорії факторів (параметрів) операції:

- 1)  $\alpha_1, \alpha_2, \dots, \alpha_m$ , які відомі заздалегідь і не можуть бути змінені;
- 2)  $Y_1, Y_2, \dots, Y_k$ , які є невідомими, і впливати на них дослідник не може;

3)  $x_1, x_2, \dots, x_n$ , які необхідно обрати для надання операції бажаних властивостей.

Від цих факторів залежить показник ефективності  $W$  та (або) функції, що описують ОДР  $G$ .

Якби фактори  $Y_1, Y_2, \dots, Y_k$  були відомі, то можливо було б обрати такі фактори  $x_1, x_2, \dots, x_n$ , при яких  $W$  досягав би глобального максимуму. Але яких саме значень  $y_1, y_2, \dots, y_k$  набудуть  $Y_1, Y_2, \dots, Y_k$  при реалізації операції невідомо і заздалегідь передбачити значення  $W$  теж буде неможливо.

**Постановка задачі дослідження операції в умовах невизначеності набуває вигляду:**

при заданих обмеженнях, фіксованих параметрах  $\alpha_1, \alpha_2, \dots, \alpha_m$ , із урахуванням невідомих факторів  $Y_1, Y_2, \dots, Y_k$  знайти такі елементи розв'язку  $x_1, x_2, \dots, x_n$ , які «по можливості» дозволяють показнику ефективності  $W$  досягти, як кажуть, «розумного» (раціонального) або вигідного значення.

Методи розв'язання задач дослідження операцій в такій постановці суттєво залежать від:

- природи факторів  $Y_1, Y_2, \dots, Y_k$ ;
- об'єму відомостей про математичну модель факторів  $Y_1, Y_2, \dots, Y_k$ .

Невизначеність  $Y_1, Y_2, \dots, Y_k$  можливо класифікувати наступним чином:

1)  $Y_1, Y_2, \dots, Y_k$  – випадкові величини або випадкові функції із відомими статистичними даними, які дозволяють ідентифікувати їх сумісну функцію розподілу або сумісну щільність ймовірності;

2)  $Y_1, Y_2, \dots, Y_k$  – невідомі фактори, які неможливо вивчати за допомогою статистичних методів, тому що відсутні відповідні статистичні дані, або явища, із якими пов'язані  $Y_1, \dots, Y_k$  взагалі не мають властивості статистичної стійкості;

3)  $Y_1, Y_2, \dots, Y_k$  – невідомі фактори, невизначеність яких пов'язана із діями «супротивника». Така ситуація є характерною для конкурентної боротьби на ринку телекомунікаційних послуг та в задачах захисту інформації.

**В першому випадку** математична модель задачі модифікується за рахунок зміни вигляду показника ефективності за допомогою одного з двох прийомів. Перший прийом: штучне зведення задачі з невизначеністю до детермінованої задачі дослідження операції шляхом заміни  $Y_1, Y_2, \dots, Y_k$ , що входять до складу показника ефективності, на їх відповідні математичні сподівання  $M[Y_1], \dots, M[Y_k]$ . Другий прийом: «оптимізація в середньому», коли в якості показника ефективності обирається математичне сподівання:

$$\bar{W} = M[W] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} W(y_1, \dots, y_k, x_1, \dots, x_n) f(y_1, \dots, y_k) dy_1 \cdot \dots \cdot dy_k \quad (1.1)$$

**В другому випадку**, базуючись на знанні діапазону значень, обчислюються так звані локально-оптимальні рішення: оптимальне рішення для фіксованих значень  $y_1, \dots, y_k$  із відомого діапазону  $y_1 \in [y_{min}; y_{max}], \dots, y_k \in [y_{min}; y_{max}]$  відповідно, які в подальшому аналізують і знаходять деякі значення  $x_1, x_2, \dots, x_n$ , які є компромісним розв'язком, що не є строго оптимальним ні для однієї з досліджуваних вище умов локальної оптимальності, але є прийнятними для розв'язання задачі в цілому. Таких прийнятних розумних розв'язків може бути декілька.

**В третьому випадку** для пошуку розв'язку використовується так звана теорія ігор – математична теорія конфліктних ситуацій. Вважаючи, що протидіюча нам у конфлікті сторона завжди прагне максимізувати наші збитки, то зрозумілою з нашого боку є стратегія (тобто таке рішення), яка мінімізує максимум нанесених збитків. Кажуть, що має місце так звана мінімаксна стратегія.

#### **1.2.4. Математичні моделі операцій з кількома показниками ефективності**

В практиці телекомунікаційних систем достатньо часто зустрічаються випадки, коли ефективність операції оцінюють не по одному, а по декількох показниках ефективності  $W_1, W_2, \dots, W_l$ , приміром, надійнісних, вартісних (економічних), технічних. Тобто, математична модель операції в задачі із декількома показниками ефективності, як в детермінованому випадку, так і у

випадку з невизначеністю, відрізняється від розглянутих тим, що замість скалярного показника ефективності  $W(x_1, \dots, x_n)$ , або  $W(Y_1, \dots, Y_k, x_1, \dots, x_n)$  використовується векторний показник ефективності:

$$\Omega = \left[ W_1(x_1, \dots, x_n), \dots, W_l(x_1, \dots, x_n) \right]^T$$

або

$$\Omega = \left[ W_1(Y_1, \dots, Y_k, x_1, \dots, x_n), \dots, W_l(Y_1, \dots, Y_k, x_1, \dots, x_n) \right]^T.$$

Бажаною є постановка задачі дослідження операцій, що вимагає пошуку такого розв'язку  $x_1, \dots, x_n$ , при якому усі складові векторного показника ефективності досягають свого екстремального значення. В загальному випадку такого розв'язку не існує, але використовуючи методи розв'язання багатокритеріальних задач, або, як ще кажуть, методи оптимізації за векторним критерієм, можливо побудувати процедури вилучення не конкурентно здатних варіантів рішення і тим самим звузити множину розв'язків до меж, в яких виконується подальший вибір раціонального рішення.

### Контрольні запитання та завдання

1. У чому полягає основна задача дослідження операцій.
2. Дайте визначення поняття «цільова функція».
3. Які математичні моделі використовуються в задачах дослідження операцій?
4. Сформулюйте загальну детерміновану поставку задачі дослідження операцій.
5. Поясніть, у яких випадках для дослідження операцій застосовують метод лінійного програмування, а у яких – метод динамічного програмування.

## **Глава 2. Структура методики дослідження операцій**

### **2.1. Складові методики дослідження операцій**

Складовими методики дослідження операцій є такі дії:

- 1) визначення цілей;
- 2) складання плану розроблення проекту, операції;
- 3) формулювання проблем;
- 4) побудова моделі;
- 5) розроблення обчислювального методу;
- 6) розроблення технічного завдання на програмування та налагодження програми;
- 7) збирання даних;
- 8) перевірка моделі;
- 9) реалізація результатів.

Розглянемо їх детальніше.

Першочергова мета будь-якого дослідження операцій полягає в тому, щоб з'ясувати, що прагне отримати керівник операції в результаті її проведення, тобто, які передбачувані результати проведення операції можна очікувати.

Цілі дослідження треба формулювати, виходячи із суті рішення або рішення, на яке орієнтована дана робота. Цілі не треба формулювати ані занадто вузько, ані занадто широко. Неправильне і неточне формулювання цілей може призвести дослідників до неправильного розв'язання поставленої задачі.

Другий етап дослідження полягає у складанні плану виконання проекту операції, тобто встановленню необхідних термінів завершення певних видів робіт. Це – одна з форм контролю за процесом розроблення проекту. Як документ, план розроблення проекту операції являє собою календарний графік виконання його етапів. Етапи можуть деталізуватися до рівня окремих завдань. Наприклад, етап розроблення обчислювального методу може мати такі завдання: розроблення методу розв'язання для кожної підмоделі задачі; опис і документальне оформлення методів розв'язання; перевірку запропонованих методів на обраних задачах невеликої розмірності; внесення уточнень та змін до методів розв'язання на підставі результатів пробних

розрахунків тощо. Складаючи план, треба також приділяти увагу розподілу робіт між окремими виконавцями.

Формулювання проблеми – третій етап дослідження. Він передбачає не тільки обговорення з керівником операції цілей дослідження, а й зібрання даних, які дають можливість уявити суть проблеми, що мало місце в минулому, чого треба очікувати в майбутньому, який характер співвідношень між змінними досліджуваної задачі. На основі цих результатів формулюється загальна схема побудови моделі і визначається напрям усієї подальшої роботи.

Одним із питань, пов'язаних із формулюванням проблеми, є визначення того, чи можна всю проблему подати у вигляді окремих підпроблем, щоб паралельно або послідовно дослідити їх незалежно одна від одної (декомпозиція). Друге питання пов'язане з визначенням міри деталізації розроблюваної моделі. Останнє залежить від обсягів виділених коштів, календарного плану розроблення проекту, мети дослідження.

Наступне питання стосується галузі застосування та розмірності розроблюваної моделі, визначення керованих і некерованих змінних, технологічних параметрів операції, показників ефективності, які дадуть змогу оцінити конкретні розв'язки розглянутої проблеми.

Четвертий етап дослідження пов'язаний із побудовою моделі. Вона відображає взаємозв'язок між керованими змінними, некерованими змінними, технологічними параметрами і показниками ефективності. Правильно побудована модель – основна умова успішного розроблення проекту операції.

Перед тим як розробляти модель, насамперед, треба з'ясувати питання про можливість використання тих чи інших показників і співвідношень у рамках моделі. Виокремлюють кілька різних типів співвідношень, які формують модель: співвідношення, які виходять із визначень; емпіричні співвідношення; нормативні співвідношення. Крім того, треба зібрати та ретельно проаналізувати великий обсяг даних.

На останньому етапі побудови моделі досліднику треба дати точне аналітичне або алгоритмічне формулювання досліджуваної проблеми.

Разом із роботою з побудови моделі потрібно обрати або розробити числовий метод розв'язання, що і є п'ятим етапом дослідження.

Для цього треба з'ясувати такі питання:

- чи треба використовувати імітаційне моделювання або будь-який з методів оптимізації:

- чи повинна модель враховувати випадковий характер деяких змінних або ж достатньо використовувати детермінований підхід;

- чи треба враховувати нелінійність певних співвідношень, чи достатньо обмежитися їх лінійною апроксимацією;

- чи можна використовувати наявні методи розв'язання, або треба розробити новий метод.

Отже, необхідно з'ясувати, які треба зробити припущення та який метод розробити, щоб застосування моделі було практично виправданим щодо використовуваних обчислювальних процедур. Цей етап також містить перевірку запропонованого числового методу з використанням модельних (тобто таких, точний розв'язок яких відомий) тестових задач, перевіряється можливість використання і коректність розроблених методів розв'язання.

Створення програм для ЕОМ передбачається на шостому етапі дослідження. Розроблення технічного завдання на програмування має виконуватися ретельно, що дасть змогу забезпечити якісніше документальне оформлення програм, результатом чого дослідження стає значною мірою орієнтованим на користувача, на задоволення його потреб. Треба звернути увагу на одну з робіт, які проводяться на цьому етапі, складання вхідних форм та вихідних (документів), їх обговорення та узгодження з керівництвом та виконавцями.

Вхідні форми – бази даних, які дають можливість забезпечити користувача інформацією, що швидко підготовлюється і легко оновлюється. Наглядні вихідні форми мають дати користувачу зрозумілу, добре підібрану й зручно розташовану інформацію.

Що стосується саме програмування і налагодження, то в багатьох випадках проект із дослідження операцій не потребує розроблення машинних програм, а у разі потреби завжди можна використати наявні програми.

На сьомому етапі здійснюється збирання та аналіз даних, які є необхідними для перевірки правильності моделі та практичного застосування результатів дослідження операцій, тоді як на попередніх етапах збирання даних пов'язане передусім з формулюванням проблеми та побудовою моделі. Тому проблема браку даних не є перешкодою до виконання продуктивних операційних досліджень, оскільки математична модель є засобом, який дає

змогу подолати труднощі отримання відповідних оцінок зведенням їх до простіших вимірювань.

Використання моделей, які розробляються при дослідженні операцій, допомагає в процесі прийняття рішень. Розв'язання проблеми зводиться до найпростіших вимірювань, встановлення вихід-них змінних і показників ефективності, які є функціями цих змінних. У цьому випадку може бути потрібно більше даних, але отримати їх значно простіше, а вимоги до їх точності будуть менш жорсткими.

Восьмий етап перевірки моделі охоплює дві фази: визначення способів перевірки і здійснення самої перевірки. На першій – обирають аналітичні й експериментальні методи перевірки несуперечності, чутливості, реалістичності та працездатності моделі. Для перевірки моделі будуть потрібні дані, отримані на попередньому етапі. Результати цієї роботи можуть призвести до необхідності перебудови моделі та, відповідно, до створення нових програм.

Найважливішим етапом, яким завершується операційне дослідження, є дев'ятий. Його можна розглядати як самостійну задачу. Отримані результати дослідження операцій треба подати рядом робочих процедур, які можна легко зрозуміти і застосувати діючій стороні.

## **2.2. Типові класи задач дослідження операцій**

За своєю змістовною постановкою множину задач дослідження операцій можна поділити на ряд класів, до яких зводиться більшість із них.

*Основні класи задач дослідження операцій:*

- управління запасами;
- розподіл ресурсів;
- ремонт та заміна обладнання;
- масове обслуговування;
- упорядкування та координація;
- вибір маршруту;
- пошук значення;
- змагальні;
- комбіновані.

Насправді задачі відповідних класів «виникають» одна з одної відповідно до того, як поширюється уявлення про досліджувану операцію.

Стисло охарактеризуємо перелічені класи задач.

*Задачі управління запасами.* Можна виокремити, принаймні, чотири основні причини, що призводять до потреби створення запасів:

- необхідність гарантувати безперебійність виробничого процесу;
- періодичність виробництва окремих матеріальних ресурсів у постачальників;
- особливості транспортування від постачальників до споживачів (невідповідність вантажопідйомності транспортних засобів і розмірів споживання);
- незбіжність ритму виробництва та постачання виробничих ресурсів із ритмом їх споживання.

Задача управління запасами в загальному випадку формулюється так. Існують певні запаси, витрати на зберігання яких є функцією їх величини. Відомі також витрати на доставляння ресурсів. Треба визначити оптимальний розмір постачання, частоту та терміни надходження ресурсів, щоб сумарні витрати були мінімальними. Критерієм оптимальності є сума витрат на зберігання і постачання ресурсів.

Задачі управління запасами можна класифікувати так:

- за кількістю періодів управління (поповнення запасів) – на одноперіодні та багатоперіодні;
- за характером поповнення запасів – із неперервною системою поповнення запасів (миттєве) і періодичні (із затримкою);
- за урахуванням попиту – на детерміновані й імовірнісні (статистичні);
- за кількістю типів ресурсів – на однопродуктові й багатопродуктові;
- за видом цільової функції – на задачі з пропорційними та непропорційними втратами.

*Задачі розподілу ресурсів* пов'язані з розподілом обмежених ресурсів за типами робіт, які треба виконати. При цьому можуть бути задані як роботи, так і ресурси, або тільки роботи. Перекидання, передавання ресурсів з одного типу робіт на інший, якщо не всі роботи можна виконати максимально ефективно через відсутність ресурсів, призводить до зменшення загальної

ефективності всіх типів робіт, взятих разом. Тому задачі розподілу ресурсів полягають у відшуканні такого розподілу ресурсів, за якого максимізується спільний прибуток або результат чи мінімізуються витрати. Задачі розподілу ресурсів дуже різноманітні за змістом і багато з них має спеціальну назву: транспортна задача; задача про призначення; задача про суміші; задача вибору оптимальних технологій тощо.

*Задачі ремонту та заміни обладнання* виникають у тих випадках, коли технічні характеристики працюючого обладнання погіршуються за рахунок старіння, спрацювання та з інших причин. Це призводить до необхідності заміни обладнання з метою зменшення сумарних витрат на експлуатацію або попередження виходу з ладу. У деяких випадках виникає потреба раціональної організації профілактичного обслуговування, тобто попереджувально-відновлювального ремонту.

Задачі цього класу класифікують:

1) за характером заміни обладнання:

- заміни обладнання на устаткування довгострокового використання;
- заміни обладнання на устаткування з метою попередження відмов;
- вибору оптимального плану попереджувального ремонту та технічного обслуговування;

2) за характером урахування витрат на обладнання – на дискретні та неперервні;

3) за виходом з ладу обладнання – на детерміновані та випадкові.

4) за стратегією заміни обладнання – на планові та змішані.

5) за часом урахування витрат на обладнання – з приведенням та без приведення витрат більш пізніх років випуску до розрахункового.

*Задачі масового обслуговування* умовно поділяють на задачі аналізу та задачі синтезу – оптимізації систем масового обслуговування. Задачі аналізу передбачають оцінювання ефективності функціонування систем за незмінних, наперед відомих вихідних характеристиках системи масового обслуговування: структури системи; дисципліни обслуговування; потоків вимог і законів розподілу часу їх обслуговування. Задачі синтезу спрямовані на пошук оптимальних параметрів системи масового обслуговування (вибір кількості каналів, їх послідовності, включення до роботи, пропускну

здатності) і характеристик функціонування (формування вхідного потоку вимог, вибір дисципліни обслуговування тощо).

*Задачі впорядкування та координації.* Задачі впорядкування пов'язані з визначенням оптимальної послідовності оброблення виробів, масивів інформації тощо. Задачі координації належать до комплексів операцій та складаються з певної сукупності окремих операцій, які повинні виконуватися з урахуванням часу у заданій послідовності. Це – задачі мережевого планування й управління. У цьому класі задач розглядають співвідношення між строками закінчення комплексу операцій та моментами початку всіх операцій комплексу.

*Задачі вибору маршруту* трапляються під час дослідження різноманітних процесів на транспорті, у системах зв'язку. Типова задача полягає у відшуканні найкращого маршруту, який пов'язує кілька пунктів. На допустимі маршрути може накладатися ряд обмежень, коли забороняється повернення до пройденого пункту.

Серед цих задач найбільш відомими є: задача вибору найкоротшого шляху між довільними пунктами; задача комівояжера; задача про максимальний потік.

*Задачі пошуку* полягають у відшуканні найкращого засобу отримання інформації, яка однозначно визначала б розв'язок. Критерієм у такій задачі є мінімум витрат двох видів: вартості отримання інформації й ціни помилки. У першому випадку йдеться про вартість вибірки, тобто вартість вибору спостережень, у другому – про помилки двох видів: помилки вибірки (виявлення того, що насправді є відсутнім) і помилки спостереження (пропускання того, що насправді має місце). Якщо на проведення пошуку виділені фіксовані ресурси, то що більше буде розмір вибірки, то менше обсяг ресурсів на кожне спостереження. Отже, за потреби зменшити помилки вибірки, як правило, зростає помилка спостереження і навпаки.

В обмеженій задачі пошуку обсяг ресурсів, які виділені на пошук, є заданим, і задача полягає у розробці плану пошуку, який мінімізує ціну помилки. У загальній задачі кількість ресурсів можна змінювати так, що її метою є мінімізація сумарних витрат ресурсів і ціни помилки.

*Змагальні задачі* – клас задач дослідження операцій, що виникають під час прийняття рішень в умовах конфліктів, незбігу інтересів осіб. Особливе місце в дослідженні проблем конфлікту займає вибір і порівняльний аналіз

можливих (допустимих) способів поведінки сторін, що дає підставу для ухвалення кожною стороною «розумних» рішень відносно своїх дій. Особи, які приймають рішення, повинні врахувати не тільки свої цілі, але й цілі, які переслідують інші учасники конфлікту. Відповідну інформацію вдається отримати не завжди, що створює додаткові труднощі як для дослідників, так і осіб, які приймають остаточне рішення.

*Комбіновані задачі містять кілька типових задач одночасно.*

### **Контрольні запитання та завдання**

1. Назвіть та поясніть основні складові методики дослідження операцій.
2. Наведіть та поясніть типові класи задач дослідження операцій.
3. Стисло охарактеризуйте основні класи задач дослідження операцій.
4. Надайте стисло характеристику задачі управління запасами.
5. У чому полягає суть задач ремонту та заміни обладнання?
6. Сформулюйте загальну постановку задачі масового обслуговування.

## Глава 3. Предмет математичного програмування

### 3.1. Загальна задача математичного програмування

Послідовність розв'язання задачі пошуку найкращих дій за заданим показником ефективності  $W$  (задача дослідження операцій) складається з наступних етапів:

- 1) якісна (неформальна, змістова, вербальна) постановка задачі дослідження операцій;
- 2) побудова математичної моделі;
- 3) математична постановка задачі;
- 4) розробка методу розв'язання задачі (метод оптимізації);
- 5) розробка методики та алгоритму реалізації запропонованого методу;
- 6) розробка комп'ютерної програми;
- 7) розрахунок (чисельний експеримент), імітаційне моделювання;
- 8) інтерпретація отриманого результату.

Загальною задачею математичного програмування є знаходження глобального екстремуму показника ефективності  $W$  на області допустимих значень  $G$ . Тобто математичне програмування дозволяє виконати 4) та 5) етапи задачі дослідження операцій.

Окрім терміну математичне програмування, ще використовують термін математична модель оптимізації, або математична модель розв'язання задач на екстремум. Набір змінних  $X = [x_1, x_2, \dots, x_n]^T$ , який задовольняє обмеженням  $G$ , називають *планом задачі математичного програмування*.

Система обмежень повинна бути сумісною, інакше множина планів буде порожньою множиною. Множина планів може бути як обмеженою, так і не обмеженою. План, що надає показнику ефективності оптимальне значення називається оптимальним, оптимальний розв'язок не завжди єдиний (рис. 3.1).

### 3.2. Класифікація задач математичного програмування

Задачі математичного програмування поділяються на класичні та некласичні.

Ознаками віднесення задач математичного програмування до класичних є вимоги:

- неперервності показника ефективності та обмежень по відношенню до керуючих змінних;
- існування за керуючими змінними перших та других частинних та мішаних частинних похідних від показника ефективності.

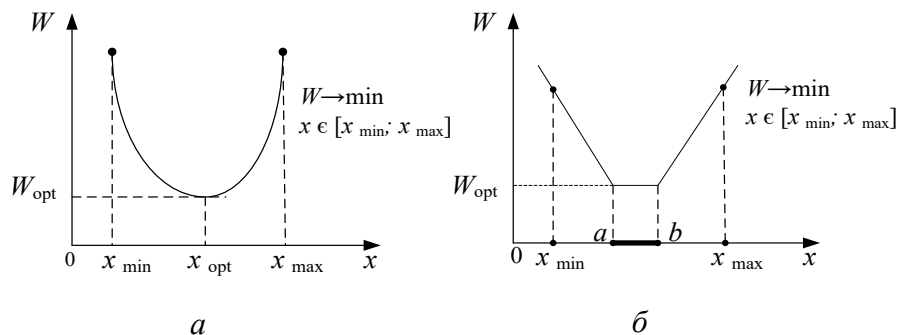


Рис. 3.1. Графічне зображення можливостей існування єдиного та безлічі оптимальних планів:

а – точка оптимуму єдина, б –  $x_{opt} \in [a; b]$  – континуум

Простою ознакою некласичності даних задач математичного програмування є вимога стосовно дискретності керуючих змінних.

*Класичні задачі* можливо поділити на 2 підкласи:

- 1) задачі пошуку безумовного екстремуму;
- 2) задачі пошуку умовного екстремуму (рис.3.2).

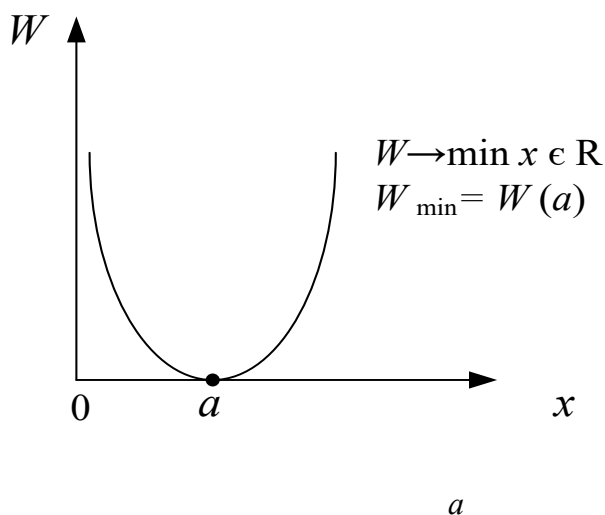
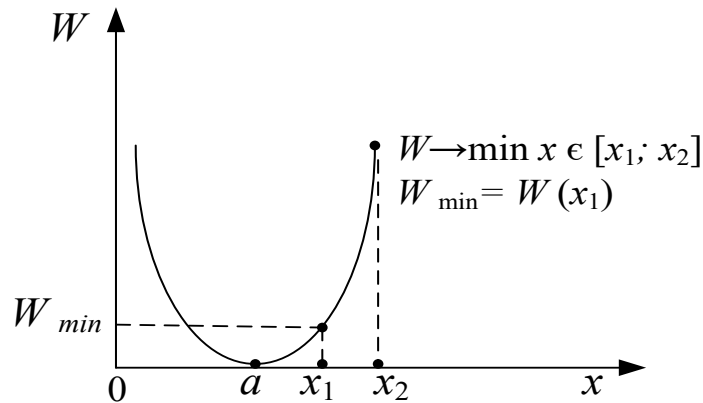


Рис. 3.2. Графічне зображення класифікації класичних задач математичного програмування: а – безумовний екстремум  $W_{min} = 0, x_{min} = a$ ;

б – умовний екстремум  $W_{min} = W(x_1), x_{min} = x_1$

Продовження рис. 3.2



б

Некласичні задачі математичного програмування поділяють на спеціальні та неспеціальні. До основних типів спеціальних задач математичного програмування відносяться:

### 1) Задача лінійного програмування(ЗЛП).

Загальна постановка ЗЛП має вигляд:

$$W = \sum_{j=1}^n c_j x_j \rightarrow \max_{X \in G}, \quad (3.1)$$

де  $X = [x_1, x_2, \dots, x_n]^T$ ,  $G$  – область допустимих розв'язків, яка задається обмеженнями

$$\sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = \overline{1, m}, x_j \in \mathbb{R}, \quad (3.2)$$

$$x_j \geq 0, \quad j = \overline{1, n}. \quad (3.3)$$

В записаних виразах  $c_j, a_{ij}, b_i, i = \overline{1, m}, j = \overline{1, n}$  – параметри,  $x_j, j = \overline{1, n}$  – керуючі змінні.

### 2) Задача квадратичного програмування(ЗКП).

Загальна постановка ЗКП відрізняється від задачі лінійного програмування формою запису показника ефективності:

$$W = \sum_{j=1}^n c_j x_j + \frac{1}{2} \sum_{k=1}^n \sum_{j=1}^n d_{k_j} x_k x_j \rightarrow \max_{X \in G}, \quad (3.4)$$

де  $d_{k_j}$   $k = \overline{1, n}, j = \overline{1, n}$  параметри.

### 3) Задача сепарабельного програмування(ЗСП).

Загальна постановка ЗСП відрізняється від задач лінійного програмування формою запису показника ефективності:

$$W = \sum_{j=1}^n f_j x_j \rightarrow \max_{X \in G} \text{ - адитивний показник ефективності}$$

або

$$W = \prod_{j=1}^n f_j x_j \rightarrow \max_{X \in G} \text{ - мультиплікативний показник ефективності}$$

де  $f_j x_j$   $j = \overline{1, n}$  – відомі функції.

### 4) Задача геометричного програмування(ЗГП).

Загальна постановка ЗГП набуває вигляду:

$$W = \sum_{k=1}^L c_k \left( \prod_{j=1}^n x_j^{\alpha_{k_j}} \right) \rightarrow \max_{X \in G}, \quad (3.5)$$

де  $G$  – область допустимих розв'язків, що задається нерівностями:

$$\sum_{k=1}^L c_{ik} \left( \prod_{j=1}^n x_j^{\alpha_{ik_j}} \right) \leq b_i, \quad i = \overline{1, m}, x_j \geq 0, j = \overline{1, n}, \quad (3.6)$$

де  $c_k, c_{ik}, \alpha_{k_j}, \alpha_{ik_j}, b_i$   $k = \overline{1, L}, i = \overline{1, m}, j = \overline{1, n}$  – параметри.

### 5) Задача опуклого програмування(ЗОП).

Загальна постановка ЗОП полягає в тому, що показник ефективності і обмеження є опуклими функціями.

Якщо всі точки відрізка АВ належать ОДР за будь-якого розташування А, В на межі ОДР, то ОДР – опукла (рис. 3.3, а). Показник ефективності вважається опуклим, якщо за будь-якого розташування точок А та В ордината точки С буде більша за значення показника ефективності в цій точці:  $Y_c \geq W x_c$  (рис. 3.3, б).

б) **Задача дискретного програмування(ЗДП)**, яка виникає у випадку, якщо будь-яку із попередніх задач доповнити умовою, що хоча б

одна із компонент  $X = [x_1, x_2, \dots, x_n]^T$  може приймати окремі значення, наприклад, цілочисельні.

7) **Задача стохастичного програмування (ЗСтП)**, яка виникає в будь-якому із попередніх випадків, коли всі або деякі, перелічені в пунктах 1, 2, 3, 4, параметри є випадковими величинами.

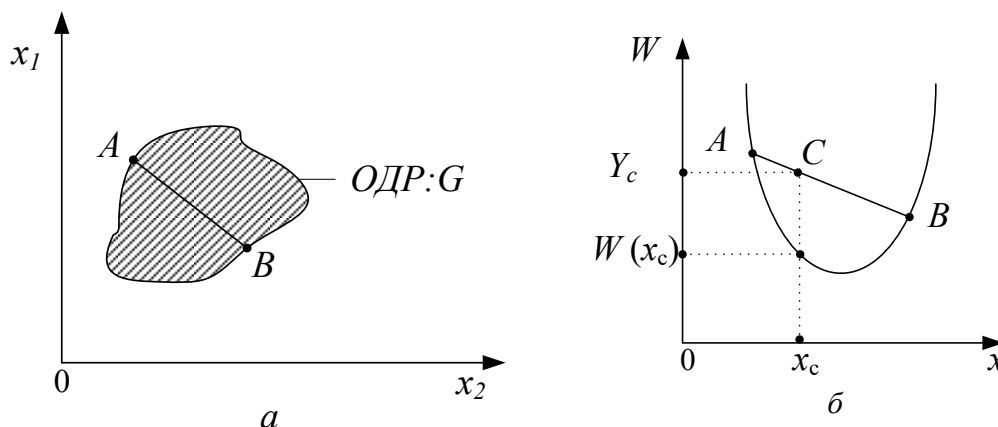


Рис. 3.3. Графічне зображення щодо пояснення змісту терміну «опукла область допустимих рішень» та «опуклий показник ефективності»:

$a$  – точки  $A, B$  належать межі області допустимих розв'язків (ОДР);

$б$  –  $W$  - показник ефективності, опуклий донизу

Усі інші типи задач математичного програмування належать до неспеціальних, некласичних задач математичного програмування.

### 3.3. Теорема про достатні умови глобального максимуму

**Означення 3.1.** В загальній задачі математичного програмування вектор змінних  $\hat{X}$  є *точкою глобального максимуму*, якщо він належить області існування планів задачі  $G$  і показник ефективності на ньому значення не менше, ніж в будь-якій допустимій точці:

$$\hat{X} \in G \text{ та } W \hat{X} \geq W X \quad \forall X \in G.$$

**Означення 3.2.** Вектор змінних є *точкою локального максимуму*, якщо він належить допустимій множині та на ньому досягається значення показника ефективності більше або рівне значенню показника ефективності в деякому малому околі цього вектора

$$\hat{X} \in G, W \hat{X} \geq W X \quad \forall X \in G \cap M_\epsilon \hat{X},$$

де  $M_\varepsilon \hat{X}$  –  $\varepsilon$ -окіл вектора  $\hat{X}$ , який в даному випадку є множиною точок, що задовольняють умові (рис.3.4, а)

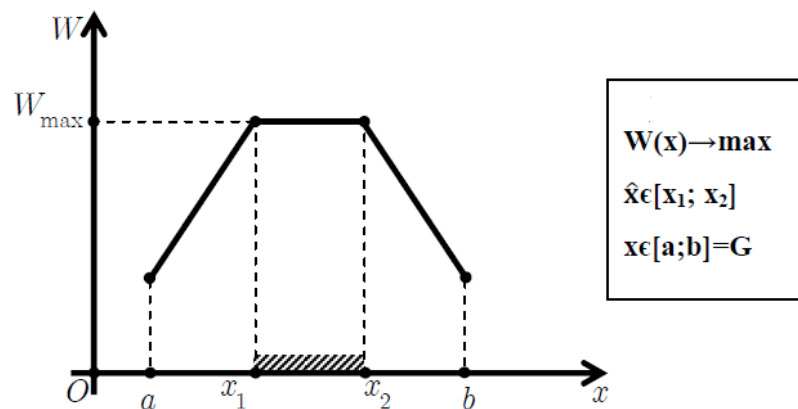
$$|X - \hat{X}| = \sqrt{\sum_{j=1}^n x_j - \hat{x}_j^2} \leq \varepsilon.$$

**Означення 3.3.** Якщо  $W \hat{X} > W X \forall X \in G$ , то *глобальний максимум* називається *строгим* або *сильним*.

**Означення 3.4.** Якщо  $W \hat{X} > W X \forall X \in G \cap M_\varepsilon \hat{X}$ , то *локальний максимум* називається *строгим* або *сильним*.

**Теорема 3.1.** Припустимо, що допустима множина  $G$  обмежена і замкнена (границя входить до складу ОДР  $G$ ), ще й опукла, а неперервний показник ефективності є опуклим догори на  $G$ . Тоді локальний максимум є глобальним. Якщо  $W X$  строго вигнута догори функція, то розв’язок задачі пошуку максимуму єдиний (існує єдиний глобальний максимум) (рис.3.4, б).

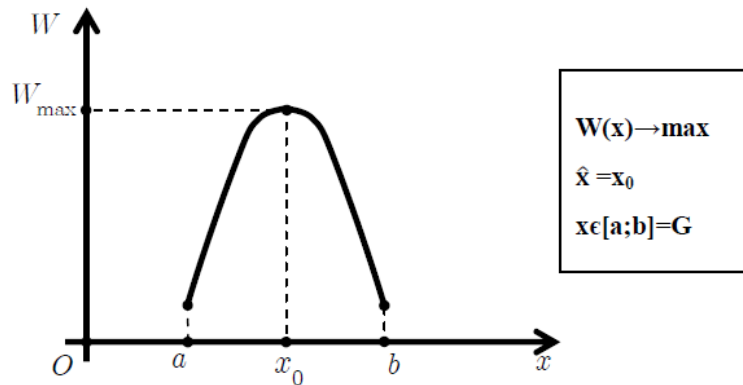
**Теорема 3.2.** Припустимо, що допоміжна множина  $G$  – компактна і опукла, тоді неперервний показник ефективності  $W X$  досягає глобального максимуму у внутрішній або граничній точці множини  $G$  (рис.3.4, в, г).



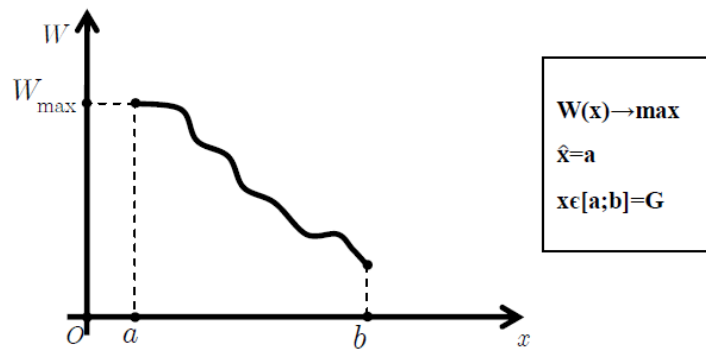
а

Рис. 3.4. Графічне зображення теорем: а, б – про глобальний максимум опуклої догори функції; в, г – теореми Вейєрштраса

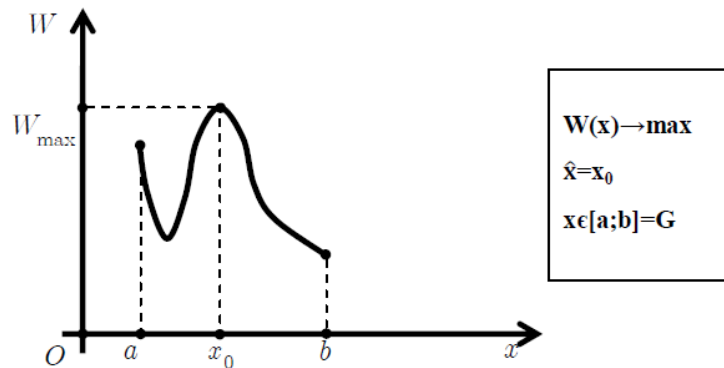
Продовження рис. 3.4



б



в



г

Головною специфічною ознакою некласичних задач математичного програмування є те, що вони мають яскраво означену прикладну спрямованість. При математичній постановці задач, які спрямовані на розв'язання конкретних практичних завдань, зазвичай виникає значна кількість специфічних обмежень, які класична теорія не враховує.

## Контрольні запитання та завдання

1. Наведіть та поясніть класифікацію задач математичного програмування.
2. У чому полягає суть задачі квадратичного програмування?
3. Поясніть суть задачі опуклого програмування.
4. У чому полягає суть задачі дискретного програмування?
5. Поясніть суть задачі стохастичного програмування.
6. Дайте визначення поняття «глобальний максимум».
7. Наведіть та поясніть достатні умови глобального максимуму.
8. У чому полягає суть теореми Вєрштраса?

## Глава 4. Типові задачі математичного програмування

### 4.1. Приклади типових задач математичного програмування

Кожна з наведених нижче задач має свою назву, що відображає її первинне походження, яке в багатьох випадках не має жодного стосунку у прямій постановці до автоматизованих систем керування на транспорті та засобів автоматизації, але прикладна привабливість цих задач для таких систем полягає в тому, що в математиці їх більшість може бути зведена до цих класичних прикладів. Типові приклади виконують як функцію прототипа-аналога, так і нагромаджувача знань з об'єктивно-орієнтованого способу розв'язання технічних задач в галузі автоматизації. Зводячи технічні задачі до прикладів, наведених нижче, отримаємо прийоми формалізації специфічних актуальних задач автоматизації та вже відпрацьований ефективний математичний та алгоритмічний інструмент розв'язання цих задач.

#### *Задача про розкрій*

На фірмі, що спеціалізується на виробництві корпусів для системних блоків, листи металу можуть розкрюватися декількома способами. Якщо лист розкroїти за  $j$ -м способом  $j = \overline{1, n}$ , то отримаємо  $a_{ij}$  корпусів  $i$ -го виду  $i = \overline{1, m}$ , при цьому величина відходів з одного листа дорівнює  $c_j$  м<sup>2</sup>.

Необхідно знайти, яку кількість листів металу необхідно розкрювати за кожним із способів для того, щоб отримати корпусів  $i$ -го виду не менше ніж  $b_i$  із мінімальною кількістю відходів.

Якщо через  $x_j$  позначити кількість листів металу, розкрюєних  $j$ -м способом, то математична модель та постановка задачі набуває вигляду:

$$W = \sum_{j=1}^n c_j x_j \rightarrow \min_{X \in G}, \quad (4.1)$$

де ОДР  $G$  являє собою умови стосовно виконання обмежень на виробництво корпусів  $j$ -го типу:

$$\sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = \overline{1, m} \quad (4.2)$$

та на  $x_j \quad j = \overline{1, n}$ , які випливають із фізичного змісту цих змінних:

$$x_j \geq 0, x_j \in \mathbb{N}_0, X = [x_1, \dots, x_n]^T \quad j = \overline{1, n} . \quad (4.3)$$

### ***Задача виробничого планування або задача оптимального використання ресурсів виробництва***

Для виробництва продукції  $j$ -го виду  $j = \overline{1, n}$  фірма має обмежені ресурси (виробничі приміщення, спеціалізовані прилади, оргтехніка, витратні матеріали, кількість фахівців, фінансові ресурси тощо)  $b_i \quad i = \overline{1, m}$ . Витрати ресурсів  $i$ -го виду на виготовлення одиниці продукції  $j$ -го виду дорівнюють  $a_{ij}$ .

Необхідно знайти скільки та якої продукції виробляти, щоб отримати максимальний прибуток. Вважаємо, що збут продукції кожного виду відбувається повністю.

Якщо позначити через  $x_j, P_j$  відповідно об'єм продукції  $j$ -го виду та прибуток від його реалізації, то математична модель постановки задачі набуває вигляду

$$W = \sum_{j=1}^n P_j x_j \rightarrow \max_{X \in G} \quad (4.4)$$

за умови виконання обмежень  $G$  по ресурсах:

$$\sum_{j=1}^n a_{ij} x_j \leq b_i \quad i = \overline{1, m} \quad (4.5)$$

із урахуванням фізичного змісту об'ємів виробництва

$$x_j \geq 0, x_j \in \mathbb{N}_0, X = [x_1, \dots, x_n]^T \quad j = \overline{1, n} . \quad (4.6)$$

### **Задача про суміші або задача про дієту**

Сучасний фітотерапевтичний центр виконує лікувальні роботи на основі препаратів природного походження. Кожен вид рослини характеризується відповідним вмістом лікувальних елементів. Відомо вміст  $i$ -го елемента в одиниці  $j$ -го рослинного препарату  $a_{ij}$   $i = \overline{1, m}$ ,  $j = \overline{1, n}$ , а також вартість  $c_j$  одиниці препарату  $j$ -го виду і добова мінімальна лікувальна доза  $i$ -го препарату  $b_i$ .

Необхідно скласти лікувально-профілактичний раціон мінімальної вартості. Якщо позначити через  $x_j$  кількість препарату  $j$ -го виду, що включено до складу раціону, то математична модель та постановка задачі набувають вигляду:

$$W = \sum_{j=1}^n c_j x_j \rightarrow \min_{X \in G} \quad (4.7)$$

за умови виконання обмежень  $G$  на добову дозу

$$\sum_{j=1}^n a_{ij} x_j \geq b_i \quad i = \overline{1, m} \quad (4.8)$$

та фізичний зміст кількості препарату

$$x_j \geq 0, x_j \in \mathbb{R} \quad j = \overline{1, n} \quad (4.9)$$

### **Задача про призначення**

Фірма, що займається гарантійним обслуговуванням та ремонтом комп'ютерної техніки має у своєму розпорядженні  $n$  ремонтних бригад. Відома продуктивність  $c_{ij}$  кожної  $i$ -ої бригади  $i = \overline{1, m}$  при виконанні  $j$ -ої роботи.

Потрібно так розподілити бригади за роботами, щоб досягти максимальної сумарної виробничої потужності. Зрозуміло, що  $i$ -та бригада може виконувати в деякому інтервалі часу лише одну задану роботу.

Якщо нас цікавить тільки заданий інтервал часу, то немає сенсу розглядати ті роботи, які протягом усього інтервалу часу не виконуються. Або розглядати ті бригади, яким у даному інтервалі часу роботи не

вистачило (табл. 4.1), тобто у процесі побудови математичної моделі задачі вважатимемо, що кількість робіт, які необхідно виконати, дорівнює кількості бригад [усі роботи обов'язково виконуються, усі бригади працюють].

Таблиця 4.1

**Пояснення щодо можливого робочого завантаження бригад**

Номер бригади	Номер роботи							
	1	2	3	4	5	6	7	8
1	0	1	0	0	0	0	0	0
2	0	0	0	0	0	1	0	0
3	1	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0

Математичну модель описаного явища можна подати у вигляді матриці (табл. 4.2).

Таблиця 4.2

**Побудова матриці математичної моделі**

Номер бригади	Номер роботи $j$					
	1	2	...	$n-2$	$n-1$	$n$
1	1	0	...	0	0	0
2	0	1	...	0	0	0
...	...	...	...	...	...	...
$n-2$	0	0	0	0	0	1
$n-1$	0	0	0	0	1	0
$n$	0	0	0	1	0	0

Відповідно до змісту задачі про призначення, в окремому стовпці, наприклад третьому, чи в окремому рядку, наприклад другому, може бути лише одна одиниця, яку інтерпретуємо як призначення другої бригади для виконання роботи третього типу.

Остаточна математична модель та постановка задачі набувають вигляду:

$$W = \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \max_{X \in G} \quad (4.10)$$

за умови виконання обмежень  $G$  :

$$x_{ij} \in [0,1], \sum_{i=1}^n x_{ij} = 1, \sum_{j=1}^n x_{ij} = 1, \quad (4.11)$$

де  $X = [x_{11}, x_{12}, \dots, x_{ij}, \dots, x_{nn}]^T$  – кількість компонент вектора дорівнює  $n^2$ .

### **Задача оптимального розміщення**

Великою телекомунікаційною компанією укладені договори на постачання обладнання замовникам у задані терміни, у певному асортименті та обсязі. Для виконання договірних зобов'язань керівництво компанії опрацьовує заходи щодо розширення виробництва на наявних підприємствах за рахунок їх реконструкцій, а також завдяки будівництву нових виробничих приміщень.

Визначити обсяги виробництва продукції на наявних підприємствах і тих, що реконструюються та будуються, а також обсяги постачання продукції від підприємств-постачальників до споживачів; і при цьому сумарні витрати на виробництво і доставлення продукції мають бути мінімальними. Побудову математичної моделі почнемо із введення та пояснення фізичного змісту таких позначень:

$i$  – вид продукції, яка виробляється  $i = \overline{1, m}$  ;

$j$  – номер підприємства, яке виробляє продукцію  $j = \overline{1, n}$  ;

$k$  – номер споживача продукції  $k = \overline{1, l}$  ;

$b_{ij}$  – об'єм продукції  $i$ -го виду, який виробляє  $j$ -те підприємство протягом заданого часу (виробнича потужність підприємства);

$c_{ij}$  – вартість виробництва продукції  $i$ -го виду, виробленої  $j$ -им підприємством;

$p_{ik}$  – об'єм поставки продукції  $i$ -го виду  $k$ -му споживачу згідно договорів;

$x_{ij}$  – об'єм виробництва продукції  $i$ -го виду на  $j$ -му підприємстві;

$x_{ijk}$  – об'єм поставки продукції  $i$ -го виду, виробленого  $j$ -им підприємством  $k$ -му споживачу;

$W$  – сумарні виробничі та транспортні витрати, тобто показник ефективності.

Математична модель і постановка задачі набувають вигляду:

$$W = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} + \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^l c_{ijk} x_{ijk} \rightarrow \min_{X \in G} \quad (4.12)$$

за умови виконання обмежень  $G$ , які накладаються на виробничі потужності кожного підприємства:

$$\sum_{k=1}^l x_{ijk} \leq b_{ij} \quad i = \overline{1, m}, j = \overline{1, n}, \quad (4.13)$$

на баланс виробництва і споживання продукції:

$$x_{ij} = \sum_{k=1}^l x_{ijk} \quad i = \overline{1, m}, j = \overline{1, n} \quad (4.14)$$

на задоволення запиту споживачів:

$$\sum_{j=1}^n x_{ijk} = p_{ik} \quad i = \overline{1, m}, k = \overline{1, l} \quad (4.15)$$

та невід'ємність об'єму поставок і виробництва продукції:

$$x_{ij} \geq 0, x_{ijk} \geq 0 \quad i = \overline{1, m}, j = \overline{1, n}, k = \overline{1, l}, \quad (4.16)$$

$$X = \left[ \overbrace{x_{11}, \dots, x_{ij}, \dots, x_{mn}}^{m \times n}, \overbrace{x_{111}, \dots, x_{mnl}}^{m \times n \times l} \right]^T.$$

### ***Задача про розподіл ресурсів***

В телекомунікаційній компанії є технічні і людські ресурси  $R_1, R_2, \dots, R_m$  (наприклад, кількість каналів та серверів на їх обслуговування, кількість бригад монтажу та ремонту обладнання, кількість бригад технологічної підтримки абонентів і т.д.) для надання телекомунікаційних послуг у кількості відповідно  $b_1, \dots, b_m$  одиниць. За допомогою цих ресурсів можливо надати послуги  $T_1, \dots, T_n$  (голосовий зв'язок, визначення координат мобільної станції, можливість прийому зображень у реальному часі, можливість забезпечити безпроводовий інтернет і т.д.). Для забезпечення

послуги  $T_j$  необхідно  $a_{ij}$  одиниць ресурсу  $R_i$   $i = \overline{1, m}; j = \overline{1, n}$ . Кожна одиниця ресурсу  $R_i$  коштує  $d_i$   $i = \overline{1, m}$ . Кожну послугу  $T_j$  можливо реалізувати за ціною  $c_j$   $j = \overline{1, n}$ . На кожний вид послуги є свій запит: відомо, що ринок телекомунікаційних послуг може скористатись не більше, ніж  $K_j$  одиницями послуги  $T_j$   $j = \overline{1, n}$ . Задача полягає в тому, щоб знайти які послуги і в якій кількості реалізувати для отримання максимального прибутку.

Побудуємо математичну модель задачі. Позначимо:

$x_1, x_2, \dots, x_n$  – кількісний вираз послуги відповідно  $T_1, T_2, \dots, T_n$ , які заплановано «виробити».

Умови запиту та фізичний зміст накладають на  $x_j$   $j = \overline{1, n}$  обмеження

$$0 \leq x_j \leq K_j \quad j = \overline{1, n}.$$

Обмеження виду

$$\sum_{j=1}^n a_{ij} x_j \leq b_i \quad i = \overline{1, m}$$

формалізують вимогу не перевищення наявного запасу відповідного ресурсу.

Обчислимо прибуток  $L$  в залежності від елементів рішення  $x_1, \dots, x_n$ .

Собівартість одиниці послуги  $T_j$  дорівнює

$$S_j = a_{1j} d_1 + a_{2j} d_2 + \dots + a_{nj} d_n = \sum_{i=1}^m a_{ij} d_i \quad j = \overline{1, n}$$

Чистий прибуток  $q_j$ , який буде отримано від реалізації одиниці послуги  $T_j$  дорівнює різниці між її ціною для продажу  $c_j$  та собівартістю  $S_j$ :

$$q_j = c_j - S_j \quad j = \overline{1, n}.$$

Загальний чистий прибуток від реалізації усіх послуг дорівнює значенню виразу:

$$W = \sum_{j=1}^n q_j x_j .$$

Математична модель і постановка задачі набувають вигляду:

$$W = \sum_{j=1}^n q_j x_j \rightarrow \max_{X \in G} \quad (4.17)$$

за умови виконання обмежень G:

$$\begin{aligned} 0 \leq x_j \leq K_j \quad j = \overline{1, n} , \\ \sum_{j=1}^n a_{ij} x_j \leq b_i \quad i = \overline{1, m} , \\ X = [x_1, \dots, x_n]^T . \end{aligned} \quad (4.18)$$

Цей запис означає: знайти такі невід'ємні та обмежені зверху значення  $x_j$   $j = \overline{1, n}$  , при яких задовольняються ресурсні обмеження та максимізується показник ефективності.

### *Задача про перевезення*

Описова постановка задачі полягає в тому, що фірма, яка працює за технологією інтернет-магазину має на всій території України  $m$  складських приміщень  $c_1, \dots, c_m$  та  $n$  пунктів споживання продукції  $\Pi_1, \dots, \Pi_n$  .

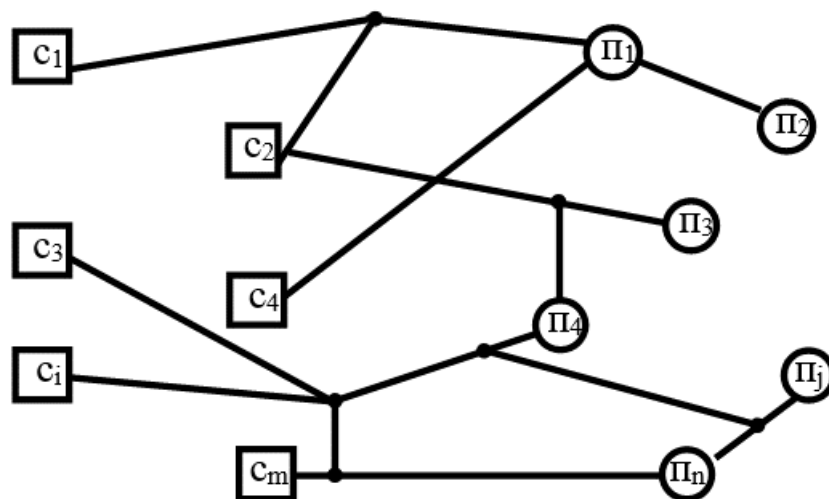


Рис. 4.1. Схема розташування складів  $c_i$   $i = \overline{1, m}$  , споживачів  $\Pi_j$   $j = \overline{1, n}$  та мережі транспортних сполучень

Суть задачі полягає у складанні плану перевезень із складів  $c_i$   $i = \overline{1, m}$  в пункти споживання  $\Pi_j$   $j = \overline{1, n}$  деякого телекомунікаційного обладнання. На складах  $c_i$   $i = \overline{1, m}$  є запаси цього обладнання у кількості  $a_i$  одиниць. Користувачі послуг  $\Pi_j$  інтернет-магазину надали заявок відповідно  $b_j$  одиниць обладнання. Заявки можливо виконати, якщо сума усіх заявок не перевищує суму усіх запасів  $\sum_{j=1}^n b_j \leq \sum_{i=1}^m a_i$ .

Склади пов'язані із пунктами споживання  $\Pi_j$   $j = \overline{1, n}$  мережею доріг із визначеними тарифами перевезення. Вартість перевезення одиниці товару зі складу  $c_i$  у пункт споживання  $\Pi_j$  дорівнює  $v_{ij}$   $i = \overline{1, m}; j = \overline{1, n}$ . Потрібно скласти план перевезень, тобто вказати з якого складу в які пункти і яку кількість обладнання необхідно спрямувати таким чином, щоб заявки були виконані, а загальні витрати на всі перевезення були мінімальні.

Побудуємо математичну модель задачі. Позначимо  $x_{ij}$  – кількість одиниць обладнання, яка спрямовується зі складу  $c_i$  у пункт  $\Pi_j$ .

Розв'язок (план перевезень) складається із  $n \times m$  чисел, які представимо у вигляді прямокутної матриці:

$$P = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} = x_{ij} \left| \begin{array}{l} i=\overline{1, m} \\ j=\overline{1, n} \end{array} \right.$$

Потрібно обрати такі невід'ємні значення змінних  $x_{ij}$ , що є елементами матриці перевезень  $P$ , щоб були виконані наступні умови:

1. Ємність складу перевищувати не можна (це означає, що загальна кількість обладнання, яке було взяте із кожного складу не повинна перевищувати його запасів на цьому складі):

$$\sum_{j=1}^n x_{ij} \leq a_i \quad i = \overline{1, m} .$$

2. Заявки, сформовані споживачами, повинні бути виконаними:

$$\sum_{i=1}^m x_{ij} = b_j \quad j = \overline{1, n} .$$

3. Загальна вартість перевезень повинна обчислюватись за формулою:

$$W = \sum_{i=1}^m \sum_{j=1}^n v_{ij} x_{ij} .$$

Постановка задачі формулюється наступним чином: знайти такий план перевезень  $P$ , щоб вартість  $W$  була мінімальною і всі обмеження були виконаними, за умови, коли сума усіх заявок дорівнює сумі усіх запасів

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

тобто з кожного складу буде вивезено усе, що на ньому є, і нерівність пункту 1 перетворюється у рівність:

$$\sum_{j=1}^n x_{ij} = a_i \quad i = \overline{1, m} .$$

Постановка задачі математичного програмування в цьому випадку набуває вигляду:

$$W = \sum_{i=1}^m \sum_{j=1}^n v_{ij} x_{ij} \rightarrow \min_{X \in G} \quad (4.19)$$

за умови виконання обмежень  $G$ :

$$\begin{aligned} x_{ij} &\geq 0 \quad i = \overline{1, m}; j = \overline{1, n} , \\ \sum_{j=1}^n x_{ij} &= a_i \quad i = \overline{1, m} , \\ \sum_{i=1}^m x_{ij} &= b_j \quad j = \overline{1, n} , \\ \sum_{j=1}^n b_j &= \sum_{i=1}^m a_i . \end{aligned} \quad (4.20)$$

Ця задача ще має назву «транспортна задача із правильним балансом».

### **Задача про виробництво складного обладнання.**

Протягом часу  $T$  планується виробництво складного телекомунікаційного обладнання, кожен компонент якого складається із  $n$  елементів  $E_1, E_2, \dots, E_n$ . Замовлення на виготовлення цих елементів можливо розмістити на  $m$  різних підприємствах  $\Pi_1, \Pi_2, \dots, \Pi_m$ . За інтервал часу  $T$  на підприємстві  $\Pi_i$  можливо виробити  $a_{ij}$  елементів типу  $E_j$   $i = \overline{1, m}; j = \overline{1, n}$ .

Зважаючи на те, що нас цікавить лише певний комплект елементів, який дозволяє виконати повну збірку відповідного пристрою, сформулюємо наступну мету операції: необхідно розподілити замовлення по підприємствах таким чином, щоб кількість повних комплектів, виготовлених за час  $T$  була максимальною. Розподілити замовлення означає визначити для кожного підприємства  $\Pi_i$   $i = \overline{1, m}$  частку часу  $x_{ij}$  від загального часу  $T$ , яку це підприємство буде працювати виробляючи елемент  $E_j$  ( $j = \overline{1, n}$ ).

Враховуючи, що

$$x_{ij} = \frac{\tau_{ij}}{T},$$

де  $\tau_{ij}$  — час, який повинно витратити підприємство  $\Pi_i$  на виробництво елемента  $E_j$ , а також той факт, що

$$\sum_{j=1}^n \tau_{ij} \leq T$$

отримаємо обмеження на змінну  $x_{ij}$ :

$$\sum_{j=1}^n x_{ij} \leq 1 \quad i = \overline{1, m}.$$

Підкреслимо, що при виконанні загальної задачі враховуються лише повні комплекти, що складаються із усіх елементів  $E_1, \dots, E_n$ . Тому визначимо кількість повних комплектів, які зможуть виробляти усі

підприємства разом за час  $T$ . Зрозуміло, що загальна кількість елементів типу  $E_j$  обчислюється за виразом:

$$N_j = \sum_{i=1}^m a_{ij} x_{ij} \quad j = \overline{1, n} .$$

Показник ефективності – це кількість повних комплектів, що дорівнює кількості елементів, яких вироблено найменше:

$$W = \min_{j \in \overline{1, \dots, n}} \sum_{i=1}^m a_{ij} x_{ij} .$$

Математична постановка задачі набуває вигляду:

$$W = \min_{j \in \overline{1, \dots, n}} \sum_{i=1}^m a_{ij} x_{ij} \rightarrow \max_{X \in G} . \quad (4.21)$$

при виконанні обмежень:

$$\begin{aligned} \sum_{j=1}^n x_{ij} &\leq 1 \quad i = \overline{1, m} , \\ x_{ij} &\geq 0 \quad i = \overline{1, m}; j = \overline{1, n} , \end{aligned} \quad (4.22)$$

де

$$X = [x_{11}, \dots, x_{ij}, \dots, x_{mn}]^T .$$

Із наведених прикладів видно, що більшість задач сформульовані із використанням лише лінійних функцій від елементів рішення як у показнику ефективності, так і в обмеженнях. Нагадаємо, що за цими елементами необхідно було знайти максимум або мінімум показника ефективності. Такі задачі в математичному програмуванні утворюють окремий клас задач, що отримали назву **задачі лінійного програмування**.

## Контрольні запитання та завдання

1. Наведіть та поясніть математичну модель та постановку задачі про розкрій.
2. Наведіть та поясніть математичну модель та постановку задачі про суміші.
3. Наведіть та поясніть математичну модель та постановку задачі про призначення
4. Наведіть та поясніть математичну модель та постановку задачі оптимального розміщення
5. Наведіть та поясніть математичну модель та постановку задачі про розподіл ресурсів
6. Наведіть та поясніть математичну модель та постановку задачі про перевезення
7. Наведіть та поясніть математичну модель та постановку задачі про виробництво складного обладнання

## Глава 5. Задачі лінійного програмування

### 5.1. Поняття про лінійне програмування

Найбільш дослідженим розділом математичного програмування є лінійне програмування.

*Лінійне програмування* вивчає важливу для практики задачу відшукування максимуму (мінімуму) лінійної функції за наявності обмежень у вигляді лінійних нерівностей або рівнянь.

Оскільки прикладна сторона математичного програмування – це оптимізація рішень взагалі і в техніці зокрема, то лінійне програмування ще називають лінійною оптимізацією. Алгоритми та програми, створені на базі методів лінійного програмування, що застосовуються в комп'ютерній техніці, утворюють так звані системи комп'ютерної математики. Ці системи є невід'ємними елементами сучасних інформаційних технологій пошуку оптимального розв'язку. Розглянемо ключовий елемент цієї технології, а саме – методи розв'язання задач лінійного програмування.

### 5.2. Форми запису задачі лінійного програмування

Загальна задача лінійного програмування полягає в пошуку максимуму (мінімуму) лінійної функції.

$$W = c_1x_1 + c_2x_2 + \dots + c_nx_n = \sum_{j=1}^n c_jx_j \rightarrow \max (\min)_{X \in G} \quad (5.1)$$

за умови використання обмежень:

$$a_{i1}x_1 + \dots + a_{in}x_n = \sum_{j=1}^n a_{ij}x_j \leq b_i \quad i = \overline{1, k}, \quad (5.2)$$

$$a_{i1}x_1 + \dots + a_{in}x_n = \sum_{j=1}^n a_{ij}x_j = b_i \quad i = \overline{k+1, m}, \quad (5.3)$$

$$x_j \geq 0 \quad i = \overline{1, n}, \quad (5.4)$$

які утворюють область допустимих розв'язків  $G$ , якій повинен задовольняти

вектор керуючих змінних  $X = [x_1, \dots, x_n]^T$ .

Нагадаємо, що функція  $W$  називається показником ефективності або цільовою функцією, або критерієм оптимальності, або лінійною формою. Сукупність значень невідомих керуючих змінних  $X = [x_1, \dots, x_n]^T$ , що задовольняють умовам (5.2), (5.3), (5.4) задачі (5.1) – (5.4) має назву розв'язку. Розв'язок  $X = [x_1, \dots, x_n]^T$  називається оптимальним, якщо він максимізує(мінімізує) значення показника ефективності.

Якщо в задачі (5.1) – (5.4) вилучити обмеження (5.3), то задача лінійного програмування буде називатися *симетричною задачею лінійного програмування*.

Якщо в задачі (5.1) – (5.4) вилучити обмеження (5.2), то отримана задача лінійного програмування буде називатися *канонічною або основною задачею лінійного програмування*.

Виконуючи формальні перетворення, можливо перейти від симетричної задачі лінійного програмування, до основної і навпаки. Перехід від нерівності до рівності виконується завдяки використанню допоміжних додатних змінних:

$$a_{i1}x_1 + \dots + a_{in}x_n \leq b_i \Leftrightarrow \begin{cases} a_{i1}x_1 + \dots + a_{in}x_n + y_i = b_i, & i = \overline{1, k} \\ y_i \geq 0, \end{cases}$$

або

$$a_{i1}x_1 + \dots + a_{in}x_n \geq b_i \Leftrightarrow \begin{cases} a_{i1}x_1 + \dots + a_{in}x_n - y_i = b_i, & i = \overline{1, k} \\ y_i \geq 0. \end{cases}$$

Перехід від рівняння до нерівності виконується шляхом заміни одного рівняння двома нерівностями.

$$a_{i1}x_1 + \dots + a_{in}x_n = b_i \Leftrightarrow \begin{cases} a_{i1}x_1 + \dots + a_{in}x_n \geq b_i, \\ a_{i1}x_1 + \dots + a_{in}x_n \leq b_i & i = \overline{k+1, m}. \end{cases}$$

Підкреслимо, що у випадку, коли за фізичним змістом задачі (5.1) – (5.4) деяка змінна  $x_s$  може мати будь який знак, то необхідно ввести в дослідження дві допоміжні невід'ємні змінні і цю змінну замінити їх різницею:

$$x_s \in \mathbb{R} \Leftrightarrow x_s = u_s - v_s, \quad u_s \geq 0, v_s \geq 0$$

### 5.3. Геометрична інтерпретація задач лінійного програмування

Поставимо задачу лінійного програмування у вигляді

$$W = c_1 x_1 + c_2 x_2 + \dots + c_n x_n \rightarrow \max(\min)_{X \in G} \quad (5.5)$$

за умови виконання обмежень  $G$  :

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n \leq b_1, \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n \leq b_m, \quad x_j \geq 0 \quad j = \overline{1, n} \end{cases} \quad (5.6)$$

Кожна із нерівностей системи (5.6) в Евклідовому просторі є півпростором із граничними гіперплощинами

$$\begin{aligned} a_{i1}x_1 + \dots + a_{in}x_n &= b_i \quad i = \overline{1, m}, \\ x_j &= 0 \quad j = \overline{1, n}. \end{aligned}$$

Припустимо, що система нерівностей (5.6) сумісна. Результат її розв'язання утворює область допустимих розв'язків, яка, як відомо із курсу лінійної алгебри, буде обмежена опуклим багатогранником, грані якого можуть співпадати із частинами гіперплощин, а вершини утворюються як точки перетину цих гіперплощин. Будь-яка внутрішня і гранична точка області допустимих розв'язків є допустимим розв'язком задачі лінійного програмування. Якщо прирівняти показник ефективності до нуля, то отримаємо рівняння гіперплощини в  $n$ -вимірному Евклідовому просторі

$$W = c_1 x_1 + c_2 x_2 + \dots + c_n x_n = 0,$$

яка проходить через початок координат і є перпендикулярною до вектора-градієнта  $\vec{C} = [c_1, \dots, c_n]^T$ , складеного із коефіцієнтів показника ефективності.

Напрямок вектора  $\vec{C}$  вказує напрям зростання функції  $W$ . Тому, для пошуку максимального значення функції  $W$  необхідно пересувати гіперплощину в напрямку вектора  $\vec{C}$  так, щоб область допустимих розв'язків і ця пересунута гіперплощина мали хоча б одну спільну точку.

На прикладі двовимірного Евклідового простору проілюструємо рисунками типові випадки, що трапляються при розв'язанні задач лінійного програмування (рис.5.1).

Вершини ОДР називають *опорними точками*, а розв'язки, що існують в цих вершинах – *опорними розв'язками*.

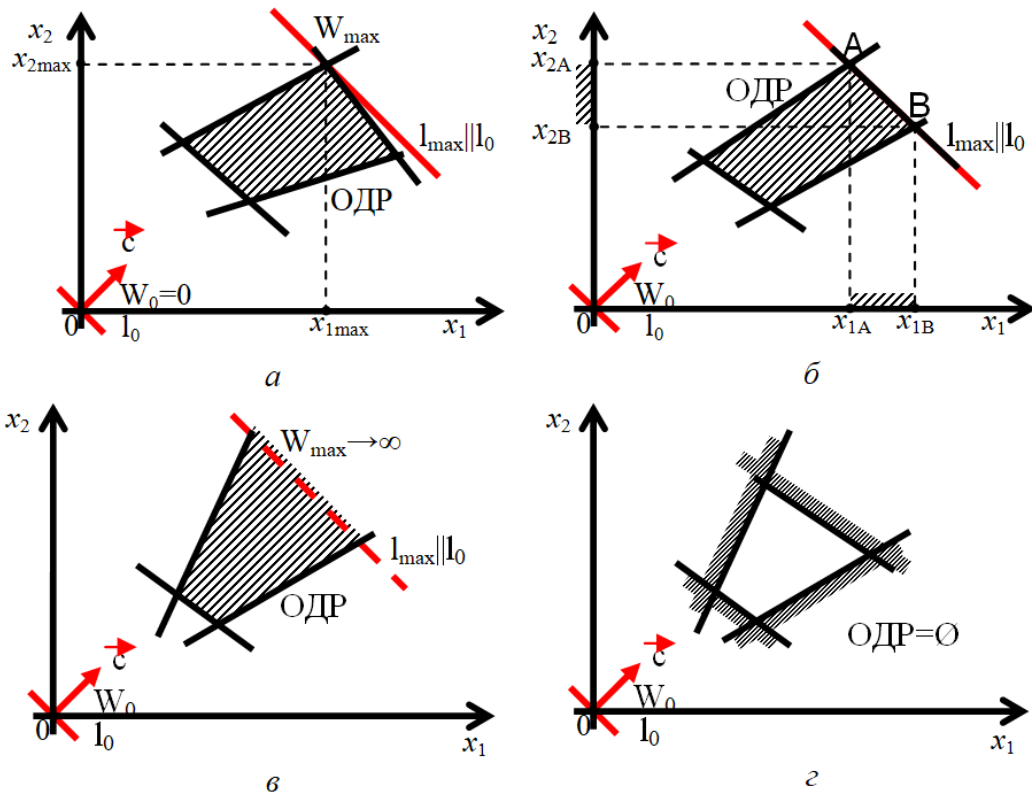


Рис. 5.1. Графічне зображення розташування основної площини

$$W = c_1x_1 + c_2x_2 + \dots + c_nx_n \text{ та ОДР:}$$

- а – задача лінійного програмування яка має єдиний розв'язок;
- б – задача лінійного програмування яка має безліч оптимальних розв'язків;
- в – задача лінійного програмування яка має абстрактний розв'язок (цей розв'язок не має фізичного змісту, тому що не враховано всі ресурсні обмеження, які не дозволяють досягти нескінченно великого значення показника ефективності);
- г – система обмежень несумісна  $G = \emptyset$  : задача лінійного програмування яка не має розв'язків.

#### 5.4. Графічний метод розв'язання задач лінійного програмування

**Приклад 5.1.** Розглянемо приклад виробничого планування, при  $n = 2$ . Припустимо, що телекомунікаційна фірма виготовляє вироби двох видів:

- А – супутник зв'язу,

- В – навігаційний супутник.
- Для їх виробництва використовуються ресурси:
- G – базові модулі сонячних батарей;
- D – функціонально необхідні елементи життєзабезпечення супутника;
- E – модулі системи орієнтації, стабілізації та навігації; в об'ємах відносно:  $G = 600$ ,  $D = 800$ ,  $E = 240$ .

Норми витрат ресурсів наведено в таблиці 5.1.

Таблиця 5.1

**Норми витрат ресурсів**

Норми витрат ресурсів	A	B
C	24	8
D	8	8
E	3	8

Прибуток від реалізації виробу A складає 40 у.о., B – 50 у.о.

Знайти об'єм виробництва, що забезпечує максимальний прибуток.

**Рішення**

□ **Побудова математичної моделі:**

Позначимо  $x_{1,2}$  – об'єми виробництва виробу A та B. Тоді прибуток фірми від реалізації  $x_1$  виробу A та  $x_2$  виробу B складатиме

$$W = 40x_1 + 50x_2.$$

Обмеження G (область допустимих розв'язків), пов'язані із ресурсами та фізичним змістом змінних, мають вигляд:

$$\begin{cases} 24x_1 + 8x_2 \leq 600; \\ 8x_1 + 8x_2 \leq 480; \\ 3x_1 + 8x_2 \leq 240; \\ x_1 \in N_0, x_2 \in N_0, \end{cases}$$

де  $N_0 = 0, 1, 2, \dots$ . Знехтуємо умовою цілочисельності розв'язків, а залишимо умову їх невід'ємності, тобто розв'яжемо так звану послаблену задачу.

**Математична постановка задачі:** необхідно максимізувати показник ефективності

$$W = 40x_1 + 50x_2 \rightarrow \max_{x_{1,2} \in G}$$

за умови виконання записаних обмежень  $G$ . Як бачимо, задача лінійного програмування сформульована як симетрична задача лінійного програмування.

Знайдемо область допустимих розв'язків цієї задачі. Для цього розв'яжемо нерівності відносно змінної  $x_2$ :

$$\begin{cases} x_2 \leq -3x_1 + 75; \\ x_2 \leq -x_1 + 60; \\ x_2 \leq -\frac{3}{8}x_1 + 30; \\ x_{1,2} \geq 0. \end{cases}$$

Границі:

$$\begin{aligned} x_2 &= -3x_1 + 75, \\ x_2 &= -x_1 + 60, \\ x_2 &= -\frac{3}{8}x_1 + 30, \\ x_1 &= 0, x_2 = 0. \end{aligned}$$

Виконаємо побудову області допустимих значень (рис.5.2) та знайдемо положення основної прямої

$$W = 40x_1 + 50x_2 = 0 \Rightarrow l_0 : x_2 = -\frac{4}{5}x_1.$$

Знайдемо  $W_{max}$ :

$$x_2 = -\frac{4}{5}x_1 + \left(\frac{W}{50}\right) = -\frac{4}{5}x_1 + W^*,$$

де  $W^*$  – відрізок, який пряма відтинає на осі ординат. Чим більше значення  $W^*$ , тим більше  $W$  тому, що  $W^* = \frac{W}{50}$ .

Знайдемо  $x_{1_{max}}, x_{2_{max}}$  як точку перетину границь:

$$\begin{cases} x_2 = -\frac{3}{8}x_1 + 30, \\ x_2 = -3x_1 + 75, \end{cases} \Leftrightarrow \begin{cases} x_1 = 17\frac{1}{7} \approx 17,14, \\ x_2 = 23\frac{4}{7} \approx 23,57. \end{cases}$$

$$W_{\max} = 40 \cdot 17,14 + 50 \cdot 23,57 = 1864,1.$$

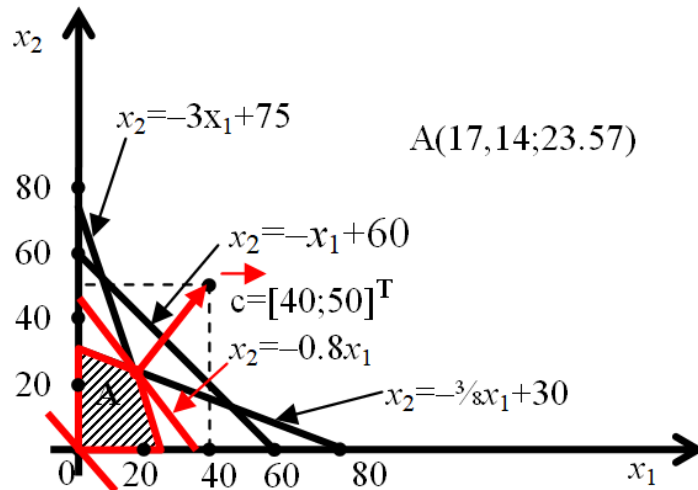


Рис. 5.2. Графік знаходження області допустимих значень

Обчислимо оптимальний розв'язок із урахуванням умови цілочисельності та невід'ємності змінних  $x_1$  та  $x_2$  (рис.5.3).

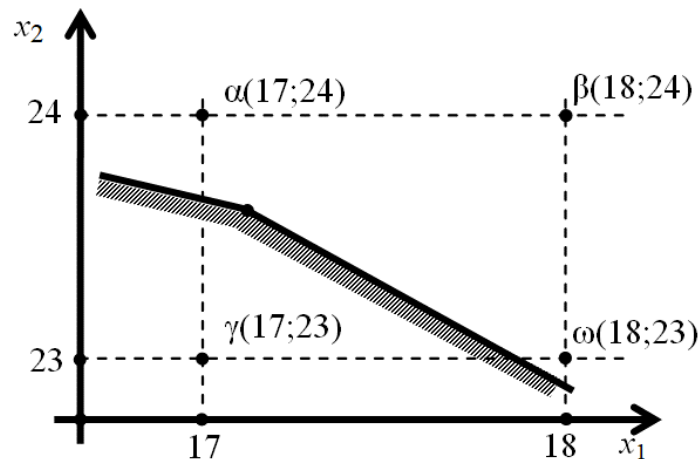


Рис. 5.3. Розташування оптимального розв'язку біля цілочислових значень

З рисунку 5.3 (із результатів обчислення можливості попадання  $\alpha, \beta, \gamma, \omega$  в область допустимих значень) отримаємо наступний результат: лише  $\gamma (17,23)$  належить області допустимих значень, тобто

$$W_{\max} = 40 \cdot 17 + 50 \cdot 23 = 1830$$

і досягається це значення критерію при випуску продукції: 17 одиниць типу  $A$  і 23 одиниць типу  $B$ . ■

**Приклад 5.2.** Знайти оптимальний розв'язок основної задачі лінійного програмування:

$$W = -5x_1 - 2x_2 - 18 \rightarrow \min_{X \in G},$$

де ОДР  $G$  задана наступними обмеженнями:

$$\begin{cases} x_1 - x_2 + x_3 = 4, \\ 2x_1 - x_2 - x_3 - x_4 = -5, \\ x_1 + x_2 - x_5 = -4, \\ x_2 + x_6 = 5, \\ 2x_1 - 2x_2 - x_6 + 2x_7 = 7, \end{cases}$$

при умові, що  $x_i \geq 0 \quad i = \overline{1,7}$ ,  $X = [x_1, \dots, x_7]^T \in \mathbb{R}^7$ .

### **Рішення**

□ Кількість керуючих змінних  $n = 7$ , а кількість обмежень-рівностей  $m = 5$ , тому  $n - m = 7 - 5 = 2$ . Отже, використаємо графічно-аналітичний метод розв'язання задачі лінійного програмування:

- побудуємо ОДР  $G$ ;
- графічним способом знайдемо точку, що відповідає мінімальному значенню критерію  $W$  (оптимальний розв'язок);
- обчислимо оптимальний розв'язок та мінімальне значення критерію.

1) Дослідження системи рівнянь на сумісність.

Із курсу лінійної алгебри відомо: для сумісності системи лінійних рівнянь необхідно і достатньо, щоб ранг матриці системи  $r_A$  дорівнював рангу розширеної матриці  $r_{AB}$ . Запишемо систему обмежень у матричному вигляді:

$$A \cdot X = B.$$

де

$$A = \begin{bmatrix} 1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 2 & -1 & -1 & -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 2 & -2 & 0 & 0 & 0 & -1 & 2 \end{bmatrix}, B = \begin{bmatrix} 4 \\ -5 \\ -4 \\ 5 \\ 7 \end{bmatrix}.$$

$A$  – матриця системи,  $B$  – матриця вільних коефіцієнтів,

$$[AB] = \begin{bmatrix} 4 \\ -5 \\ -4 \\ 5 \\ 7 \end{bmatrix} - \text{розширена матриця системи.}$$

Ранг матриці це найбільший порядок відмінного від нуля мінора:

$$\Delta_1 = 1 \neq 0,$$

$$\Delta_2 = \begin{vmatrix} 1 & -1 \\ 2 & -1 \end{vmatrix} \neq 0,$$

$$\Delta_3 = \begin{vmatrix} 1 & -1 & 1 \\ 2 & -1 & -1 \\ 1 & 1 & 0 \end{vmatrix} = 1 \cdot (-1) \cdot 0 + (-1) \cdot (-1) \cdot 1 + 2 \cdot 1 \cdot 1 -$$

$$-1 \cdot (-1) \cdot 1 - 1 \cdot 1 \cdot (-1) - 2 \cdot (-1) \cdot 0 = 5 \neq 0,$$

$$\Delta_4 = \begin{vmatrix} 1 & -1 & 1 & 0 \\ 2 & -1 & -1 & -1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & 1 & 0 \\ 2 & -1 & -1 \\ 1 & 0 & 0 \end{vmatrix} = -1 \cdot (-1) \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1 \neq 0,$$

$$\begin{aligned} \Delta_5 &= \begin{vmatrix} 1 & -1 & 1 & 0 & 0 \\ 2 & -1 & -1 & -1 & 0 \\ 1 & 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 \\ 2 & -2 & 0 & 0 & 0 \end{vmatrix} = -1 \cdot \begin{vmatrix} 1 & -1 & 1 & 0 \\ 2 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 2 & -2 & 0 & 0 \end{vmatrix} = \\ &= -1 \cdot (-1)^{2+4} \cdot -1 \cdot \begin{vmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 2 & -2 & 0 \end{vmatrix} = \\ &= \begin{vmatrix} 0 & 1 \\ 2 & -2 \end{vmatrix} = -2 \neq 0. \end{aligned}$$

Таким чином маємо, що

$$r_A = m = 5.$$

Визначників більшого порядку немає, тому що матриця  $A$  є прямокутною матрицею розміру  $5 \times 7$ . Зрозуміло, що розширена матриця системи матиме розмір  $5 \times 8$ , а це означає що найбільший порядок визначника розширеної матриці співпадає із найбільшим порядком визначника системи і дорівнює 5.

Висновок:  $r_A = r_{AB} = 5$ , система обмежень у формі рівностей сумісна.

2) Вибір вільних та базисних змінних.

Враховуючи той факт, що  $n = 7$ ,  $m = 5$ , тобто  $n - m = 2$ , можливо обрати дві змінні, наприклад,  $x_1$  та  $x_2$  в якості вільних і 5 змінних  $x_3, x_4, x_5, x_6, x_7$  в якості базисних. Тоді із системи обмежень знаходимо:

$$\begin{cases} x_3 = -x_1 + x_2 + 4 \geq 0, \\ x_4 = 3x_1 - 2x_2 + 1 \geq 0, \\ x_5 = x_1 + x_2 + 4 \geq 0, \\ x_6 = -x_2 + 5 \geq 0, \\ x_7 = \frac{1}{2}(-2x_1 + 2x_2 + x_6 + 7) = -x_1 + \frac{1}{2}x_2 + 6 \geq 0. \end{cases}$$

Враховуючи невід'ємність змінних будемо область допустимих розв'язків, аналітичний вираз для якої в залежності від змінних  $x_1$  та  $x_2$  має вигляд:

$$\begin{cases} x_2 \geq x_1 - 4, \\ x_2 \leq \frac{3}{2}x_1 + \frac{1}{2}, \\ x_2 \geq -x_1 - 4, \\ x_2 \leq 5, \\ x_2 \geq 2x_1 - 12, x_{1,2} \geq 0. \end{cases}$$

Знайдемо положення основної гіперплощини:

$$W = -5x_1 - 2x_2 - 18 \Leftrightarrow x_2 = -\frac{5}{2}x_1 - \underbrace{\frac{W + 18}{2}}_b;$$

$$\begin{cases} b = 0, \\ x_2 = -\frac{5}{2}x_1, \end{cases}$$

$\vec{c} = [-5, -2]$  – вектор градієнту.

Оптимальне (мінімальне) значення показника ефективності дорівнює:

$$W_{\min} = -70,5$$

і досягається це значення у точці  $x_{1\min}; x_{2\min} = 8,5; 5$  (рис.5.4). ■

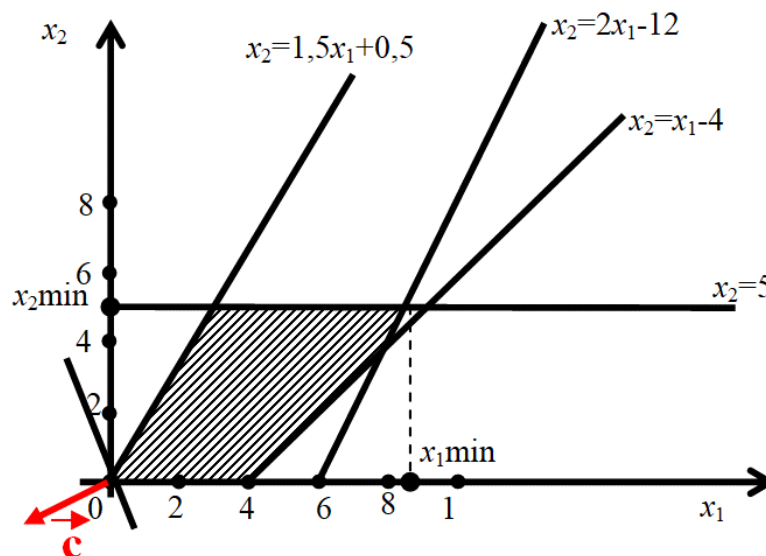


Рис. 5.4. Графічне зображення побудови області допустимих розв'язків

**Висновок.** Наведена у підрозд. 5.3 геометрична інтерпретація задачі лінійного програмування та приклади дають підстави стверджувати таке:

**Принцип 1.** Оптимальний розв'язок, якщо він існує, розташований в одній із «опорних точок». За аналітичною ознакою опорна точка – це така точка, в якій хоча б  $k = n - m$  ( $n$  - кількість невідомих,  $m$  - кількість сумісних обмежень рівностей) змінних перетворюються в 0. Геометрично опорна точка визначається як кутова точка ОДР  $G$ . Проілюструємо принцип 1 графічними образами (рис. 5.5) для випадку  $k = n - m = 2$ .

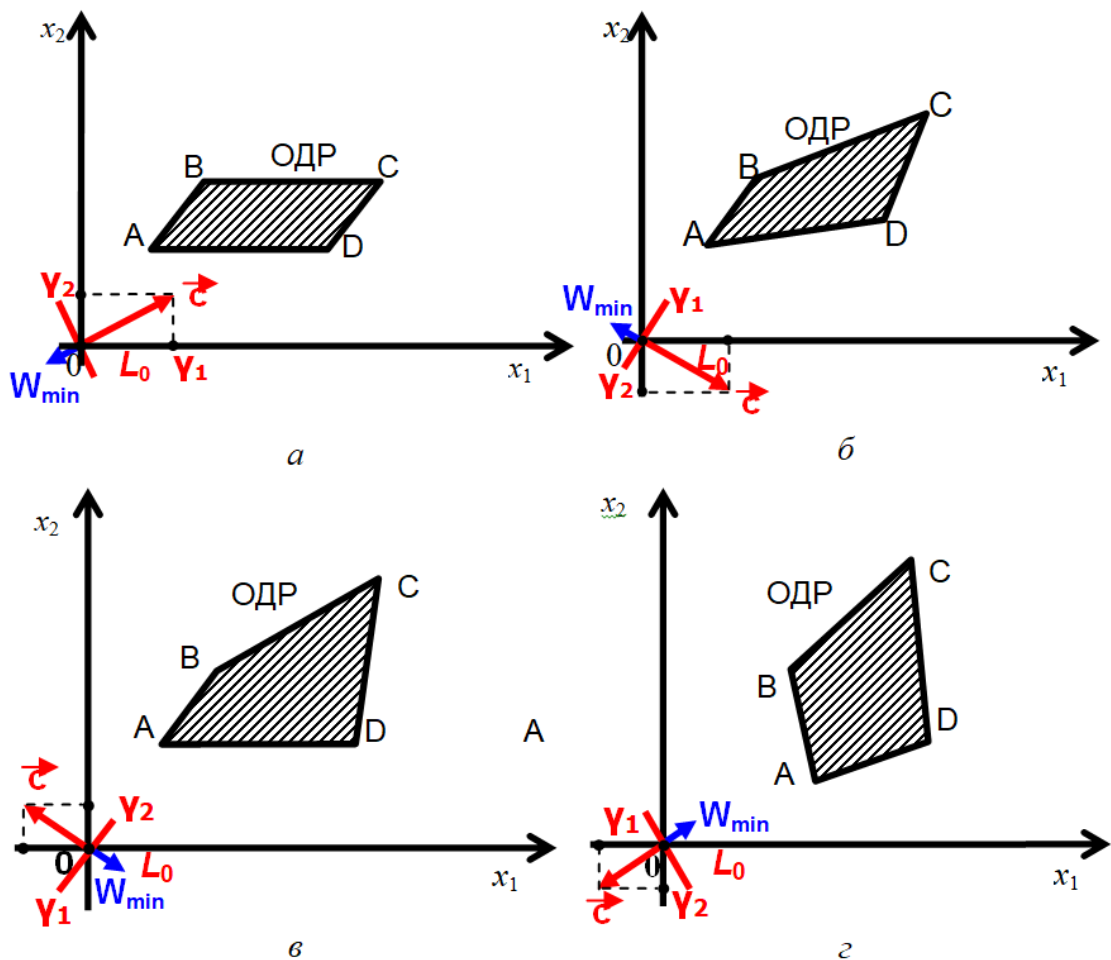


Рис. 5.5. Положення основної прямої та напрямку її пересування під час пошуку найменшого значення в точці, яка задовольняє область допустимих розв'язків:

$$a - \gamma_1, \gamma_2 > 0; \quad б - \gamma_1 > 0, \gamma_2 < 0; \quad в - \gamma_1 < 0, \gamma_2 > 0; \quad г - \gamma_1, \gamma_2 < 0$$

Зрозуміло, що у загальному випадку, коли  $k = 2$  всі змінні  $x_3, x_4, \dots, x_n$  можна виразити через  $x_1$  та  $x_2$ , що дозволяє критерій

$$W = c_1 x_1 + c_2 x_2 + c_3 x_3 + \dots + c_n x_n$$

записати у вигляді:

$$W = \gamma_1 x_1 + \gamma_2 x_2 + \gamma_0.$$

Припустимо, що необхідно знайти оптимальне значення критерію  $W$ .  
Якщо  $\gamma_0 = 0$ , то основна пряма  $L_0$ :

$$\gamma_1 x_1 + \gamma_2 x_2 = 0$$

проходить через початок координат.

Розглянемо за допомогою рис. 5.5  $a - z$  напрямки переміщення цієї прямої в площині двох змінних  $x_1, x_2$ , які дозволяють знайти найменше значення показника ефективності.

**Принцип 2.** Для пошуку оптимального розв'язку необхідно переходити від однієї опорної точки до іншої, пересуваючись у напрямі зменшення або збільшення показника ефективності залежно від вимог задачі лінійного програмування.

**Принцип 3.** Застосування геометричного методу розв'язання задачі лінійного програмування ускладнюється, коли  $n - m = 3$ . Геометричний метод зовсім втрачає придатність до розв'язання практичних задач при  $n - m > 3$ . Тому, у довільному випадку застосовуються чисельні методи пошуку екстремуму показника ефективності, які ґрунтуються на записаних вище принципах (1) та (2).

Універсальним методом розв'язання задач лінійного програмування є *симплекс-метод*.

### Контрольні запитання та завдання

1. У чому полягає суть лінійного програмування?
2. Наведіть та поясніть форми запису задачі лінійного програмування.
3. Наведіть геометричну інтерпретацію задачі лінійного програмування.
4. У чому полягає суть графічного методу розв'язання задач лінійного програмування?
5. Поясніть суть «опорних точок».

## Глава 6. Симплекс-метод розв'язання задачі лінійного програмування

### 6.1. Поняття про симплекс-метод

У назві основного методу розв'язання основної задачі лінійного програмування використовується термін "симплекс", що означає  $n$ -вимірний тетраedr або  $n$ -вимірний трикутник.

Симплекс-метод знаходження локального мінімуму будь-якої функції декількох змінних лінійної або нелінійної запропоновано Нелдером і Мідом. Цей метод хоча і містить у своїй назві слово «симплекс» не має нічого спільного із симплекс методом розв'язання задачі лінійного програмування. Суть методу Нелдера – Міда полягає у спеціальній процедурі обчислення координат вершин цього  $n$ -вимірного трикутника для наступної ітерації (наближення) в залежності від результату порівняння значень показника ефективності у вершинах  $n$ -вимірного трикутника, координати яких можуть бути обчислені у попередній ітерації. "Найгірша" вершина, в якій показник ефективності приймає найбільше значення, якщо відбувається пошук мінімального значення показника ефективності, відкидається і замінюється новою.

Координати нової вершини отримують, наприклад, наступним прийомом: «відображенням» старої вершини відносно прямої, що проходить через дві інші вершини. Окрім «відображення» для пошуку координат нової вершини використовуються так звані процедури "продовження", "стискання" або "скорочення". В результаті застосування означених прийомів та процедур значення показника ефективності у вершинах трикутників на кожній ітерації зменшується і при цьому зменшується «розмір» самого  $n$ -вимірного трикутника, стискаючись поступово до точки мінімального значення показника ефективності (рис. 6.1).

Розглянемо симплекс-метод розв'язання задачі лінійного програмування.



$$W = c_1x_1 + c_2x_2 + \dots + c_nx_n \rightarrow \min_{\substack{x_j \geq 0 \quad j=1, \overline{n} \\ y_i \geq 0 \quad i=1, \overline{m}}} , \quad (6.2)$$

який необхідно мінімізувати за невід'ємними змінними  $x_j \geq 0 \quad j = \overline{1, n}$  та  $y_i \geq 0 \quad i = \overline{1, m}$ , то отримаємо основну задачу лінійного програмування.

Отже, в результаті переходу до основної задачі лінійного програмування маємо:

1) Рівняння обмежень задані у формі, де базисні (залежні) змінні  $y_i \geq 0 \quad i = \overline{1, m}$  виражені через незалежні (вільні)  $x_j \geq 0 \quad j = \overline{1, n}$  змінні.

2) Загальна кількість змінних дорівнює « $n + m$ », де  $n$  – кількість початкових,  $m$  – додаткових змінних.

3) Показник ефективності явно залежить від початкових змінних. Вважаємо, що усі коефіцієнти при додаткових змінних показника ефективності дорівнюють 0.

**Приклад 6.1.** Дана задача лінійного програмування:

$$W = x_1 - 2x_2 - 3x_3 \rightarrow \min_{x_1, x_2, x_3 \geq 0}$$

за умови виконання обмежуючих нерівностей

$$\begin{cases} 2x_1 - x_2 + x_3 \leq 5, \\ x_3 - 2x_2 \leq -1, \\ x_5 - 2x_4 + x_1 \leq -1, \\ x_5 - x_1 \leq 0, \\ x_3 + 2x_1 + x_5 \geq 0. \end{cases}$$

Потрібно привести цю задачу до вигляду основної задачі лінійного програмування.

**Рішення**

□ 1. Приведемо обмеження нерівності до стандартної форми

$$\begin{cases} 2x_1 - x_2 + x_3 \leq 5, \\ x_3 - 2x_2 \leq -1, \\ x_5 - 2x_4 + x_1 \leq -1, \\ x_5 - x_1 \leq 0, \\ -x_3 - 2x_1 - x_5 \leq 0. \end{cases}$$

2. Використовуємо для отримання обмежень рівностей додаткові змінні

$$\begin{cases} y_1 = 5 - 2x_1 - x_2 + x_3, \\ y_2 = -1 - x_3 - 2x_2, \\ y_3 = -1 - x_5 - 2x_4 + x_1, \\ y_4 = 0 - (x_5 - x_1), \\ y_5 = 0 - (-x_3 - 2x_1 - x_5); \end{cases} \Leftrightarrow \begin{cases} y_1 = 5 - 2x_1 + x_2 - x_3, \\ y_2 = -1 - x_3 + 2x_2, \\ y_3 = -1 - x_5 + 2x_4 - x_1, \\ y_4 = -x_5 + x_1, \\ y_5 = x_3 + 2x_1 + x_5. \end{cases}$$

3. Постановка задачі набуває вигляду

$$W = x_1 - 2x_2 - x_3 \rightarrow \min_{\substack{x_{1,2,3} \geq 0 \\ y_{1,2,3} \geq 0}}$$

за умови виконання обмежень, наведених в п. 2. ■

### 6.3. Основні прийоми та способи симплекс-методу розв'язання задач лінійного програмування

Для розв'язання основної задачі лінійного програмування використовуємо наведені в підрозділі 5.4 принципи побудови оптимального розв'язку.

Прийоми та способи симплекс-методу розв'язання задачі лінійного програмування викладені у припущенні, що в задачі лінійного програмування використовується  $n$  змінних та  $m$  незалежних лінійних рівнянь-обмежень.

#### **Приєм 1. Вибір вільних змінних.**

Оберемо будь-які  $k$  змінних  $k = n - m$  ( $r = m$  - ранг системи обмежень рівностей) в якості вільних і представимо через них  $m$  базисних змінних, що залишилися. Припустимо, що у якості вільних обрані перші  $k$  змінних, тобто  $x_1, x_2, \dots, x_k$ , а інші  $m$  представимо через них:



## **Приём 2. Покращення опорного розв'язку.**

Якщо серед коефіцієнтів  $\gamma_1, \dots, \gamma_k$  є від'ємні, то збільшуючи ті змінні, при яких коефіцієнти від'ємні, досягаємо покращення розв'язку, тобто зменшення показника ефективності.

Припустимо, що є єдиний від'ємний коефіцієнт  $\gamma_1$ . Необхідно виконати дві дії:

1) Збільшити  $x_1$ , але так, щоб жодна із базисних змінних  $x_{k+1}, \dots, x_n$  не стала від'ємною, якщо деякий із коефіцієнтів  $\alpha_{k+11}, \dots, \alpha_{n1}$  при  $x_1$  від'ємний.

Збільшувати  $x_1$  можливо без обмежень, якщо всі коефіцієнти при  $x_1$  у виразах для обчислення базисних змінних додатні. Але в цьому випадку показник ефективності  $W$  прямує до  $-\infty$  при  $x_1 \rightarrow +\infty$ , тобто оптимального розв'язку, який має фізичний зміст, не існує. Існує абстрактний розв'язок. Розв'язання задачі припиняється, необхідно переформулювати постановку задачі.

2) Виключити  $x_1$  із списку вільних змінних і вставити у список базисних, а із списку базисних виключити ту змінну, припустимо  $x_L$ , яка першою досягне значення 0 при збільшенні  $x_1$ . Випишемо рівняння для  $x_L$ :

$$x_L = \alpha_{L1} \cdot x_1 + \dots + \alpha_{Lk} \cdot x_k + \beta_L$$

в якому покладемо, що  $x_2 = 0, \dots, x_k = 0, x_L = 0$ .

Тоді

$$x_1 = -\frac{\beta_L}{\alpha_{L1}} \geq 0$$

тому, що  $\alpha_{L1} < 0, \beta_L \geq 0$ .

Отримане значення  $x_1$  – це те значення, при якому  $x_L = 0$ . Взагалі першою досягне нуля та змінна із складу  $x_{k+1}, \dots, x_n$  для якої  $-\frac{\beta_L}{\alpha_{L1}}$  або  $\left| \frac{\beta_L}{\alpha_{L1}} \right|$  буде найменшим ( $L \in [k+1, n], L \in \mathbb{N}$ ). Припустимо, що це  $x_r, r \in [k+1, n], r \in \mathbb{N}$ .

### **Прийом 3. Зміна складу вільних змінних.**

Обираємо новий склад вільних змінних  $x_2, x_3, \dots, x_k, x_r$  та базисних  $x_1, x_{k+1}, \dots, x_{r-1}, x_{r+1}, \dots, x_n$ . Обчислимо нові базисні змінні через нові вільні та перевіряємо умову невід'ємності базисних змінних при нульових вільних змінних, тобто з'ясуємо чи є розв'язок опорним. Перевіряємо отримані опорні розв'язки на оптимальність, використовуючи прийоми 1 та 2.

**Висновок:** практична реалізація симплекс методу потребує розробки двох алгоритмів:

1. Визначення опорного розв'язку основної задачі лінійного програмування.
2. Визначення оптимального розв'язку основної задачі лінійного програмування.

**Приклад 6.2.** Пошук оптимального розв'язку шляхом поступового покращення результату із використанням трьох, описаних вище прийомів.

Постановка задачі:

$$W = 5x_1 - 2x_3 \rightarrow \min_{X \in G} \quad (6.3)$$

при умові виконання обмежень  $G$ :

$$\begin{cases} -5x_1 - x_2 + 2x_3 \leq 2, \\ -x_1 + x_3 + x_4 \leq 5, \\ 3x_1 + 5x_4 \leq 7, \\ x_i \geq 0, \quad i = 1, 2, 3, 4, \quad X = [x_1, x_2, x_3, x_4]^T \end{cases} \quad (6.4)$$

#### **Методика розв'язання задачі:**

- 1) виконати перехід до стандартного виду основної задачі лінійного програмування;
- 2) перевірити систему обмежень-рівностей на сумісність;
- 3) виконати обчислення оптимального розв'язку із використанням прийомів симплекс методу.

#### **Рішення**

- 1. Перехід до стандартного вигляду задачі лінійного програмування:

$$(6.4) \Leftrightarrow \begin{cases} -5x_1 - x_2 + 2x_3 + y_1 = 2; \\ -x_1 + x_3 + x_4 + y_2 = 5; \\ 3x_1 + 5x_4 + y_3 = 7; \\ x_{1,2,3,4} \geq 0, y_{1,2,3} \geq 0, \end{cases} \quad (6.5)$$

де (6.5) задає ОДР  $G$ .

2. Перевірка обмежень-рівностей на сумісність:

$$(6.5) \Leftrightarrow \begin{cases} -5x_1 - 1x_2 + 2x_3 + 0x_4 + y_1 = 2, \\ -1x_1 + 0x_2 + 1x_3 + 1x_4 + y_2 = 5, \\ -3x_1 + 0x_2 + 0x_3 + 5x_4 + y_3 = 7, \end{cases} \quad (6.6)$$

де

$$A = \begin{bmatrix} -5 & -1 & 2 & 0 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ -3 & 0 & 0 & 5 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 \\ 5 \\ 7 \end{bmatrix}.$$

Ранг матриці  $A : r_A = 3$ , що дорівнює  $m$  – кількості обмежень-рівностей. Отже, система обмежень сумісна та лінійно залежних обмежень немає.

Підкреслимо, що зазвичай сумісності та незалежності обмежень можливо досягти ще при побудові математичної моделі задачі, якщо враховувати фізичний зміст цих обмежень.

3. Обчислення оптимального розв'язку задачі лінійного програмування симплекс методом почнемо з вибору базисних та вільних змінних:

$$\begin{cases} y_1 = 5x_1 + x_2 - 2x_3 + 2, \\ y_2 = x_1 - x_3 - x_4 + 5, \\ y_3 = 3x_1 - 5x_4 + 7, \\ x_{1,2,3,4} \geq 0, \\ y_{1,2,3} \geq 0, \end{cases} \quad (6.7)$$

де система (6.7) задає ОДР  $G$ , на якій оптимізується критерій

$$W = 5x_1 - 2x_3 \rightarrow \min_{\left[ X^T \ Y^T \right]^T \in G},$$

$$X = [x_1, x_2, x_3, x_4]^T, Y = [y_1, y_2, y_3]^T.$$

Загальна кількість змінних в переформульованій задачі  $n = 7$ ,  $m = 3$  – кількість сумісних, лінійно незалежних обмежень-рівностей. Кількість вільних змінних  $k = 7 - 3 = 4$ .

**Приєм 1.**

Обираємо в якості вільних змінних  $x_{1,2,3,4}$  і припускаємо, що вони дорівнюють нулю. В результаті отримаємо розв'язок:

$$X_1 = [x_1 = 0; x_2 = 0; x_3 = 0; x_4 = 0; y_1 = 2; y_2 = 5; y_3 = 7],$$

який можливо вважати опорним.

При цьому  $W X_1 = 0$ .

**Приєм 2.**

2.1) Коефіцієнт при  $x_3$  виразу для обчислення показника ефективності від'ємний, тому за рахунок збільшення  $x_3$  можливо зменшити  $W$ . Обираємо в якості нової вільної змінної ту змінну (серед  $y_{1,2,3}$ ), для якої модуль відношення вільних коефіцієнтів до коефіцієнта при  $x_3$  найменший:

$$\left| \frac{2}{-2} \right| = 1 < \left| \frac{5}{-1} \right|.$$

Обираємо  $y_1$ . Таким чином, вільними змінними є  $x_1, x_2, x_4, y_1$ .

Обчислимо нові базисні змінні  $x_3, y_2, y_3$ :

$$\begin{cases} x_3 = -\frac{1}{2}y_1 + \frac{5}{2}x_1 + \frac{1}{2}x_2 + 1, \\ y_2 = \frac{1}{2}y_1 - \frac{3}{2x_1} - \frac{1}{2}x_2 - x_4 + 4, \\ y_3 = 3x_1 - 5x_4 + 7. \end{cases}$$

Перевіряємо знайдений розв'язок, чи є він опорним.

$$X_2 = [y_1 = 0; x_{1,2} = 0; x_3 = 1; y_2 = 4; y_3 = 7]$$

$X_2$  – опорний розв'язок. Спостерігається зменшення значення критерію:

$$W X_2 = y_1 - x_2 - 2 = -2.$$

2.2) Коефіцієнт при  $x_2$  від'ємний. Можливо зменшити значення показника ефективності завдяки збільшенню  $x_2$ . Аналогічно викладеному у 2.1) визначаємо нові вільні та базисні змінні:

$$\begin{cases} x_3 = -y_2 + x_1 - x_4 + 5, \\ x_2 = y_1 - 2y_2 - 3x_1 - 2x_4 + 8, \\ y_3 = 3x_1 - 5x_4 + 7, \end{cases}$$

$$X_3 = \left[ \underbrace{y_1 = 0; x_1 = 0; x_4 = 0; y_2 = 0}_{\text{вільні змінні}}; \underbrace{x_2 = 8; x_3 = 5; y_3 = 7}_{\text{базисні змінні}} \right],$$

$$W X_3 = 2y_2 + 3x_1 + 2x_4 - 10 = -10.$$

Всі коефіцієнти при вільних змінних у виразі для обчислення показника ефективності невід'ємні, тому досягати зменшення показника ефективності за рахунок додаткових вільних змінних неможливо. Досягнуто глобального мінімуму

$$W = -10$$

в точці  $x_1 = 0, x_4 = 0, y_1 = 0, y_2 = 0, x_2 = 8, x_3 = 5, y_3 = 7$ . ■

**Зауваження 6.1** У прикладах 6.1, 6.2 при застосуванні прийомів симплекс-методу отримали невід'ємні опорні розв'язки. У більшості практичних випадків трапляються ситуації, коли за обраних вільних змінних з нульовими значеннями отримують від'ємні значення базисних змінних, тобто не виконується умова невід'ємності змінних. Тому доцільно розглянути спеціальний алгоритм обчислення опорного розв'язку.

### Контрольні запитання та завдання

1. У чому полягає суть симплекс-методу?
2. Наведіть та поясніть процедуру зведення стандартної форми обмежень нерівностей до обмежень рівностей основної задачі лінійного програмування.
3. Поясніть основні прийоми та способи симплекс-методу розв'язання задач лінійного програмування.
4. Наведіть та поясніть методику розв'язання задач лінійного програмування на основі симплекс-методу.

5. Яка система обмежень вважається стандартною формою обмежень нерівностей?
6. У чому полягає суть методу Нелдера – Міда?
7. Наведіть приклади графічного зображення прийомів симплекс-методу Нелдера – Міда.

## Глава 7. Алгоритм пошуку опорних розв'язків основної задачі лінійного програмування

### 7.1. Табличний алгоритм заміни базисних змінних (стандартні таблиці)

Як було показано в гл. 6, прийоми реалізації симплекс-методу розв'язання задачі лінійного програмування ґрунтуються на багаторазових перерахуваннях системи рівнянь обмежень основної задачі лінійного програмування відносно нових вільних та базисних змінних. Для прискорення цих розрахунків розглянемо спеціальний табличний алгоритм. Припустімо, що потрібно виконати мінімізацію показника ефективності, представленого лінійною функцією

$$W = c_0 + c_1x_1 + c_2x_2 + c_3x_3 \rightarrow \min_{x_{1,2,3} \geq 0}$$

за умови виконання обмежень

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

Представимо цю задачу у вигляді основної задачі лінійного програмування у формі зручній для застосування стандартизованого табличного алгоритму:

$$W = c_0 - c_1x_1 - c_2x_2 - c_3x_3 = c_0 - \gamma_1x_1 + \gamma_2x_2 + \gamma_3x_3 \rightarrow \min_{x_{1,2,3} \geq 0}$$

$$\begin{cases} y_1 = b_1 - a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ y_2 = b_2 - a_{21}x_1 + a_{22}x_2 + a_{23}x_3, \\ y_3 = b_3 - a_{31}x_1 + a_{32}x_2 + a_{33}x_3, \end{cases}$$

де  $\gamma_1 = -c_1$ ,  $\gamma_2 = -c_2$ ,  $\gamma_3 = -c_3$ .

Складемо таблиці (див. табл. 7.1 та табл. 7.2), в яких позначено: ПЕ - показник ефективності; БЗ - базисні змінні; ВЕ - вільний елемент.

## Вихідне подання інформації

Показник ефективності (ПЕ), Базисні змінні (БЗ)	Вільний елемент (ВЕ)	Коефіцієнти при вільних змінних		
		$x_1$	$x_2$	$x_3$
$W$	$c_0$	$\gamma_1$	$\gamma_2$	$\gamma_3$
$y_1$	$b_1$	$a_{11}$	$a_{12}$	$a_{13}$
$y_2$	$b_2$	$a_{21}$	$a_{22}$	$a_{23}$
$y_3$	$b_3$	$a_{31}$	$a_{32}$	$a_{33}$

Припустимо, що необхідно виконати заміну базисної змінної  $y_3$  на  $x_2$  і, навпаки, вільної змінної  $x_2$  на  $y_3$ . В аналітичному вигляді така заміна змінних виконується із використанням рівняння:

$$x_2 = \frac{b_3}{a_{32}} - \left( \frac{a_{31}}{a_{32}} \cdot x_1 + \frac{1}{a_{32}} \cdot y_3 + \frac{a_{33}}{a_{32}} \cdot x_3 \right),$$

в якому елемент  $a_{32}$  називають *розв'язувальним елементом* і, відповідно, у стандартизованій таблиці рядок та стовпець називають *розв'язувальними рядком та стовпцем* (Табл. 7.2). Якщо вираз для  $x_2$  підставити у формули для обчислення  $W$ ,  $y_1$ ,  $y_2$ , то отримаємо таблицю перерахунку коефіцієнтів рівнянь обмежень для нових базисних змінних  $y_1, y_2, x_2$  та нових вільних змінних  $x_1, y_3, x_3$  (Табл. 7.2).

В загальному випадку алгоритм перетворення  $x_j$  в  $y_i$  і навпаки із використанням стандартної таблиці складається з наступних дій:

Формули перерахунку коефіцієнтів після заміни змінних

ПЕ БЗ	ВЕ	Коефіцієнти при вільних змінних		
		$x_1$	$y_3$	$x_3$
$W$	$c_0 + b_3 \left( -\frac{\gamma_2}{a_{32}} \right)$	$\gamma_1 + a_{31} \left( -\frac{\gamma_2}{a_{32}} \right)$	$-\frac{\gamma_2}{a_{32}}$	$\gamma_3 + a_{33} \left( -\frac{\gamma_2}{a_{32}} \right)$
$y_1$	$b_1 + b_3 \left( -\frac{a_{12}}{a_{32}} \right)$	$a_{11} + a_{31} \left( -\frac{a_{12}}{a_{32}} \right)$	$-\frac{a_{12}}{a_{32}}$	$a_{13} + a_{33} \left( -\frac{a_{12}}{a_{32}} \right)$
$y_2$	$b_2 + b_3 \left( -\frac{a_{22}}{a_{32}} \right)$	$a_{21} + a_{31} \left( -\frac{a_{22}}{a_{32}} \right)$	$-\frac{a_{22}}{a_{32}}$	$a_{23} + a_{33} \left( -\frac{a_{22}}{a_{32}} \right)$
$x_2$	$b_3 \left( \frac{1}{a_{32}} \right)$	$a_{31} \left( \frac{1}{a_{32}} \right)$	$\frac{1}{a_{32}}$	$a_{33} \left( \frac{1}{a_{32}} \right)$

1. Виділення в стандартній таблиці розв'язувального елемента  $a_{ij}$ .

Обчислення зворотної величини  $\lambda = \frac{1}{a_{ij}}$ . Для наочності обчислень

пропонується записати  $\lambda$  у правому нижньому куті фрагменту таблиці, де розташований розв'язувальний елемент.

2. Множення усіх елементів розв'язувального рядка (крім  $a_{ij}$ ) на  $\lambda$ .

Результат записується у відповідному фрагменті таблиці справа знизу.

3. Множення усіх елементів розв'язувального стовпця (окрім  $a_{ij}$ ) на  $-\lambda$ . Результат записується у відповідному фрагменті таблиці справа знизу.

4. Виділення у розв'язувальному рядку усіх попередніх, а у розв'язувальному стовпці – усіх нових елементів за виключенням самого розв'язувального елемента  $a_{ij}$ .

5. Обчислення добутку нового елемента розв'язувального стовпця на попередній елемент розв'язувального рядка. Для всіх попередніх елементів таблиці окрім тих, що розташовані у розв'язувальних рядку та стовпці. Запис добутку у відповідний фрагмент таблиці.

6. Формулювання остаточного результату:

- 1) заміна  $x_j$  на  $y_i$  і навпаки;
- 2) залишити у розв'язувальних стовпці та рядку лише нижні елементи фрагментів таблиці;
- 3) заміна інших елементів таблиці сумою попередніх чисел та результатів розрахунку згідно пункту 5.

**Приклад 7.1.** Необхідно виконати заміну змінних у наступній задачі лінійного програмування:

$$W = -x_1 - x_3 + 2 \rightarrow \min_{x_{1,2,3} \geq 0}$$

при виконанні умов:

$$\begin{cases} x_1 - x_2 + 2x_3 - 5 \geq 0, \\ 2x_1 - x_2 + 1 \geq 0, \\ 2x_2 - x_3 - 1 \geq 0. \end{cases}$$

### Рішення

□ Переформулюємо критерій та обмеження нерівності у вигляді:

$$W = 2 - (x_1 + x_3) \rightarrow \min_{\substack{x_{1,2,3} \geq 0 \\ y_{1,2,3} \geq 0}}$$

$$\begin{cases} y_1 = -5 - x_1 + x_2 - 2x_3, \\ y_2 = 1 - 2x_1 + x_2, \\ y_3 = -1 - 2x_2 + x_3. \end{cases}$$

Представимо вихідні дані у вигляді таблиці 7.3 та виконаємо заміну змінних, наприклад  $y_2$  на  $x_1$  у рядку і  $x_1$  на  $y_2$  у стовпчику, із використанням наведеного вище алгоритму перетворення  $x_j$  в  $y_i$  і навпаки.

Таблиця 7.3

		Вихідні дані							
ПЕ	БЗ	ВЕ		Коефіцієнти при вільних змінних					
				$x_1 \rightarrow y_2$	$x_2$	$x_3$			
$W$		2	1/2	1	1/2	0	1/2	-1	0
$y_1$		-5	-1/2	-1	-1/2	1	-1/2	-2	0
$y_2 \leftrightarrow x_1$		1	-1/2	-2	-1/2	1	-1/2	0	0
$y_3$		-1	0	0	0	-2	0	1	0

Результат заміни змінних наведено в табл.7.4.

Таблиця 7.4

Результат заміни змінних				
ПЕ БЗ	ВЕ	Коефіцієнти при вільних змінних		
		$y_2$	$x_2$	$x_3$
$W$	$5/2$	$1/2$	$1/2$	$1$
$y_1$	$-11/2$	$-1/2$	$1/2$	$-2$
$x_1$	$-1/2$	$-1/2$	$-1/2$	$0$
$y_3$	$-1$	$0$	$-2$	$1$

■

## 7.2. Алгоритм пошуку опорного розв'язку за допомогою стандартних таблиць

Основна задача лінійного програмування у стандартній формі має вигляд:

$$W = c_0 - \gamma_1 x_1 + \gamma_2 x_2 + \dots + \gamma_n x_n \rightarrow \min_{\substack{x_i \geq 0 \quad i=1, n \\ y_j \geq 0 \quad j=1, m}}$$

за умови виконання обмежень рівностей:

$$y_j = b_j - a_{j1} x_1 + a_{j2} x_2 + \dots + a_{jn} x_n, \quad j = \overline{1, m}$$

В кожній вершині області допустимих розв'язків розташовані опорні розв'язки, для яких характерним є те, що хоча б  $n$  змінних дорівнюють 0, а інші змінні невід'ємні.

Покладемо  $x_i = 0 \quad i = \overline{1, n}$ ,  $y_j = b_j \quad j = \overline{1, m}$ . Тоді ознакою того, що записаний вище розв'язок

$$\begin{cases} x_i = 0 & i = \overline{1, n}, \\ y_j = b_j & j = \overline{1, m} \end{cases}$$

є опорним, слід вважати невід'ємність вільних елементів  $b_j \geq 0 \quad j = \overline{1, m}$ .

Якщо ця умова не виконується, то необхідно перейти від розв'язку, що не задовольняє область допустимих розв'язків до опорного. Алгоритм пошуку опорного розв'язку повинен призводити на кожній ітерації або до зменшення кількості від'ємних вільних коефіцієнтів обмежень рівностей, які задають рівняння для пошуку базисних змінних, або зменшувати модуль цих від'ємних коефіцієнтів, і тим самим, поступово, наближати базисні змінні до границі області допустимих розв'язків.

Розглянемо один з таких алгоритмів, який використовує стандартні таблиці.

**Алгоритм пошуку опорного рішення за допомогою стандартних таблиць:**

**Приєм 1. Дослідити область допустимих розв'язків на існування.**

Знайти в стандартній таблиці від'ємний вільний елемент для будь-якої базисної змінної. Якщо в знайденому рядку немає від'ємних коефіцієнтів при вільних змінних, то це ознака відсутності розв'язку тому, що

$$y_j = b_j - a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n < 0$$

при  $b_j < 0$  та  $x_i \geq 0$ , а змінна  $y_j$  повинна бути невід'ємною. Виконати умову невід'ємності неможливо.

**Приєм 2. Пошук розв'язувального елемента.**

Припустимо, що у рядку із від'ємним вільним коефіцієнтом є від'ємний коефіцієнт при деякій вільній змінній. Розглядаючи стовпець з від'ємним коефіцієнтом при деякій вільній змінній, знаходимо в ньому елемент однакового з вільним коефіцієнтом знаку, але такий, щоб відношення вільного коефіцієнта до нього було мінімальним. Цей елемент і є розв'язувальним.

**Приєм 3. Виконуємо заміну змінних із застосуванням стандартних таблиць.**

Розглянемо на прикладі застосування алгоритму пошуку опорного рішення із використанням стандартних таблиць:

**Приклад 7.2.** Знайти, якщо він існує, опорний розв'язок основної задачі лінійного програмування із наступними рівняннями обмеженнями:

$$\begin{cases} y_1 = 1 - x_1 - 2x_2 + x_3, \\ y_2 = -5 - (-2x_1 + x_2 - x_3), \\ y_3 = 2 - x_1 + x_2, \\ y_4 = 1 - x_2 + x_3. \end{cases}$$

### Рішення

□ Запишемо вихідні дані у вигляді симплекс-таблиці (табл.7.5).

Таблиця 7.5

#### Вихідні дані

ПЕ БЗ	ВЕ		Коефіцієнти при вільних змінних					
			$x_1 \rightarrow y_3$		$x_2$		$x_3$	
$W$	1	2	-1	$\boxed{1}$	-2	1	1	0
$y_1$	-5	4	-2	$\boxed{2}$	1	2	-1	0
$y_3 \leftrightarrow x_1$	$\boxed{2}$	2	$\textcircled{1}$	1	$\boxed{1}$	1	$\boxed{0}$	0
$y_4$	1	0	0	$\boxed{0}$	-1	0	1	0

Обираємо другий рядок та перший стовбець, тому що у другому рядку є від'ємний коефіцієнт  $-2$  та від'ємний вільний коефіцієнт  $-5$ . Зупинимось на стовпчику із коефіцієнтом  $-2$  та обчислимо відношення  $\frac{-5}{-2} = 2,5 \geq \frac{2}{1} = 2$ , тобто в якості розв'язувального елемента вибираємо  $a_{31} = 1$ . Заповнимо табл.7.6, яка буде таблицею першої ітерації до опорного розв'язку.

Таблиця 7.6

#### Перша ітерація

БЗ	ВЕ		Коефіцієнти при вільних змінних					
			$y_3$		$x_2$		$x_3 \leftrightarrow y_2$	
$y_1$	3	-1	1	2	-1	3	1	$\boxed{1}$
$y_2 \leftrightarrow x_3$	$\boxed{-1}$	1	$\boxed{2}$	-2	$\boxed{3}$	-3	$\textcircled{-1}$	-1
$x_1$	2	0	1	0	1	0	0	$\boxed{0}$
$y_4$	1	-1	0	2	-1	3	1	$\boxed{1}$

Повторно застосовуємо алгоритм пошуку опорного розв'язку з використанням стандартних таблиць (табл. 7.6).

Заповнюємо табл. 7.7, яка буде таблицею другої ітерації до опорного розв'язку.

Таблиця 7.7

**Друга ітерація**

Базисні змінні	Вільний елемент	Коефіцієнти при вільних змінних		
		$y_3$	$x_2$	$y_2$
$y_1$	2	3	2	1
$x_3$	1	-2	-3	-1
$x_1$	2	1	1	0
$y_4$	0	2	2	1

У табл. 7.7 усі вільні елементи невід'ємні. Опорний розв'язок знайдено. ■

**Приклад 7.3.** Знайти, якщо він існує, опорний розв'язок задачі лінійного програмування із наступними рівняннями обмеженнями:

$$\begin{cases} y_1 = -4 - x_1 + 2x_2, \\ y_2 = -3 - x_1 - x_2 + x_3, \\ y_3 = -10 - (2x_1 - x_2 + x_3), \\ y_4 = -2 - x_1 + x_2. \end{cases}$$

**Рішення**

□ Обираємо, наприклад, перший рядок та перший стовпець із від'ємним коефіцієнтом (табл. 7.8).

Як на початку розв'язання задачі лінійного програмування, так і в процесі отримання кожної наступної ітерації контролюємо виникнення умови, коли вільний елемент будь-якої базисної змінної від'ємний, а серед коефіцієнтів при вільних змінних немає від'ємного.

**Вихідні дані**

БЗ	ВЕ		Коефіцієнти при вільних змінних					
			$x_1 \leftrightarrow y_4$		$x_2$		$x_3$	
$y_1$	-4	2	-1	-1	2	-1	0	0
$y_2$	-3	-2	1	1	-1	1	1	0
$y_3$	-10	-4	2	2	-1	2	1	0
$y_4$	2	2	-1	-1	1	-1	0	0

Так в рядках 2 та 3 маємо:

$$y_2 = -5 - y_4 + x_3 < 0 ,$$

$$y_3 = -14 - 2y_4 + x_2 + x_3 < 0 \quad \forall x_2, x_3 \geq 0, y_4 \geq 0.$$

**Перша ітерація**

БЗ	ВЕ	Коефіцієнти при вільних змінних		
		$y_4$	$x_2$	$x_3$
$y_1$	-2	-1	1	0
$y_2$	-5	1	0	1
$y_3$	-14	2	1	1
$x_1$	2	-1	-1	0

Тобто розв'язку, який належить ОДР за ознакою невід'ємності управляючих змінних, не існує. Отже, область допустимих розв'язків – порожня множина. Як бачимо, з'ясування питання про існування допустимих розв'язків відбувається в процесі пошуку опорного розв'язку. ■

## Контрольні запитання та завдання

1. У чому полягає суть опорних розв'язків задачі лінійного програмування?
2. Поясніть, у чому полягає суть табличного алгоритму заміни базисних змінних.
3. Розкрийте суть поняття розв'язувальний елемент.
4. Наведіть та поясніть алгоритм пошуку опорного розв'язку за допомогою стандартних таблиць.
5. Охарактеризуйте основні прийоми алгоритму пошуку опорного розв'язку за допомогою стандартних таблиць.

## Глава 8. Алгоритм пошуку оптимального розв'язку основної задачі лінійного програмування

### 8.1. Алгоритм пошуку оптимального розв'язку за допомогою стандартних таблиць

У гл. 6, 7 розглянуто основні складові (прийоми симплекс-методу) пошуку оптимального опорного розв'язку задачі лінійного програмування, тобто задачі знаходження мінімуму показника ефективності:

$$W = c_0 - \gamma_1 x_1 + \dots + \gamma_n x_n \rightarrow \min_{\substack{x_i \geq 0 \quad i=1, n \\ y_j \geq 0 \quad j=1, m}}$$

при виконанні рівнянь-обмежень

$$y_i = b_j - a_{j1} x_1 + \dots + a_{jn} x_n, \quad j = \overline{1, m}$$

Покажемо на прикладі, як у зручній формі табличного алгоритму заміни змінних виконувати оптимізацію, тобто розв'язувати основу задачу лінійного програмування в цілому.

#### **Приклад 8.1.** Розв'язати основну задачу лінійного програмування

$$W = 0 - -x_1 + 2x_2 + x_3 \rightarrow \min_{\substack{x_{1,2,3} \geq 0 \\ y_{1,2,3,4} \geq 0}}$$

при виконанні рівнянь-обмежень

$$\begin{cases} y_1 = 2 - (x_1 + x_2 - 2x_3), \\ y_2 = 1 - x_1 - x_2 + x_3, \\ y_3 = 5 - x_2 + x_3, \\ y_4 = 2 - (2x_1 - x_2). \end{cases}$$

#### **Рішення**

□ При  $x_1 = x_2 = x_3 = 0$ ,  $y_1 = 2, y_2 = 1, y_3 = 5, y_4 = 2$ , тобто опорний розв'язок існує, але він не є оптимальним, тому що коефіцієнти при  $x_2$  та  $x_3$ , у виразі для  $W$  – додатні. Це означає, що за рахунок збільшення  $x_2$  або  $x_3$  можливо зменшити значення показника ефективності. Перейдемо до застосування алгоритму стандартних таблиць (див. табл.8.1).

Вихідні дані

Показник ефективності, базисні змінні	Вільний елемент		Коефіцієнти при вільних змінних					
			$x_1$		$x_2$		$x_3 \leftrightarrow y_2$	
$W$	0	-1	-1	-1	2	1	1	$\boxed{-1}$
$y_1$	2	2	1	2	1	-2	-2	$\boxed{2}$
$y_2 \leftrightarrow x_3$	$\boxed{1}$	1	$\boxed{1}$	1	$\boxed{-1}$	-1	$\textcircled{1}$	1
$y_3$	5	-1	0	-1	1	1	1	$\boxed{-1}$
$y_4$	2	0	2	0	-1	0	0	$\boxed{0}$

В першому рядку табл.8.1, що відображає вихідні данні, показник ефективності має додатні коефіцієнти при вільних змінних  $x_2$  та  $x_3$ . Тому можливо обрати будь який із стовбців в якості розв'язувального. Нехай це буде стовбець з  $x_3$ , тобто  $x_3$  виключимо із складу вільних змінних. Розв'язувальний елемент в обраному стовбці повинен буди додатний. Таких коефіцієнтів два: один у рядку з  $y_2$  та інший у рядку з  $y_3$ .

Оберемо як розв'язувальний елемент той, для якого відношення вільного елемента до елемента обраного в якості розв'язувального найменше( табл. 8.1).

В таблиці отриманій після першої ітерації(див. табл. 8.2), у рядку «Показник ефективності», тільки при  $x_2$  коефіцієнт додатний, тому  $x_2$  необхідно вивести зі складу вільних змінних. В якості розв'язувального елемента вибираємо  $a_{32} = 2$ (табл. 8.2).

Таблиця 8.2

## Перша ітерація

Показник ефективності, базисні змінні	Вільний елемент		Коефіцієнти при вільних змінних					
			$x_1$		$x_2 \leftrightarrow y_3$		$y_2$	
$W$	-1	-6	-2	3/2	3	$\boxed{-3/2}$	-1	3/2
$y_1$	4	2	3	-1/2	-1	$\boxed{1/2}$	2	-1/2
$x_3$	1	2	1	-1/2	-1	$\boxed{1/2}$	1	-1/2
$y_3 \leftrightarrow x_2$	$\boxed{4}$	2	$\boxed{-1}$	-1/2	$\textcircled{2}$	1/2	$\boxed{-1}$	-1/2
$y_4$	2	2	2	-1/2	-1	$\boxed{1/2}$	0	-1/2

У подальшому для отримання другої та третьої ітерацій застосовуємо табличний алгоритм заміни базисних змінних (табл. 8.3, 8.4).

Таблиця 8.3

## Друга ітерація

Показник ефективності, базисні змінні	Вільний елемент		Коефіцієнти при вільних змінних					
			$x_1$		$y_3$		$y_2 \leftrightarrow y_1$	
$W$	-7	-2	-1/2	-5/6	-3/2	-1/6	1/2	$\boxed{-1/3}$
$y_1 \leftrightarrow y_2$	$\boxed{6}$	4	$\boxed{5/2}$	5/3	$\boxed{1/2}$	1/3	$\textcircled{3/2}$	2/3
$x_3$	3	-2	1/2	-5/6	1/2	-1/6	1/2	$\boxed{-1/3}$
$x_2$	2	2	-1/2	5/6	1/2	1/6	-1/2	$\boxed{-1/3}$
$y_4$	4	2	3/2	5/6	1/2	1/6	-1/2	$\boxed{-1/3}$

## Третя ітерація

Показник ефективності, базисні змінні	Вільний елемент	Коефіцієнти при вільних змінних		
		$x_1$	$y_3$	$y_1$
$W$	-9	-4/3	-5/3	-1/3
$y_2$	4	5/3	1/3	2/3
$x_3$	1	-1/3	1/3	-1/3
$x_2$	4	1/3	2/3	1/3
$y_4$	6	7/3	2/3	1/3

Як бачимо опорний розв'язок

$$x_1 = 0, x_2 = 4, x_3 = 1, y_1 = 0, y_2 = 4, y_3 = 0, y_4 = 6$$

є оптимальним, тому що всі коефіцієнти при вільних змінних у виразі для обчислення показника ефективності від'ємні. Остаточо отримаємо оптимальне(мінімальне) значення показника ефективності

$$W_{\min} = -9. \blacksquare$$

**Зауваження 8.1.** Якщо у кожному стовпці із додатнім коефіцієнтом у рядку показника ефективності не існує додатного елемента, то це означає, що не дивлячись на існування області допустимих розв'язків обмеженого розв'язку задача не має: при необмеженому збільшенні вільної змінної показник ефективності  $W \rightarrow -\infty$ .

**Алгоритм пошуку оптимального розв'язку основної задачі лінійного програмування з використанням симплекс-методу:**

1. Якщо всі вільні елементи (виключаючи рядок  $W$ ) у симплекс таблиці(стандартній таблиці) невід'ємні і при цьому у рядку  $W$  (без урахування вільного елемента) немає жодного додатного коефіцієнта, то оптимальний розв'язок досягнуто.

2. Якщо в рядку  $W$  серед коефіцієнтів вільних змінних є додатній, а у стовпці, що йому відповідає, немає жодного додатного коефіцієнту

(елементу) , то лінійна функція  $W$  не обмежена знизу і оптимального розв'язку не існує.

3. Якщо у стовпці є додатні коефіцієнти, то вільну змінну, яка відповідає цьому стовпцю необхідно поміняти місцями із тією базисною змінною, для якої відношення вільного елемента до обраного додатного коефіцієнту є найменшим.

**Зауваження 8.2.** При розв'язанні задачі лінійного програмування може виникнути ситуація коли при заміні змінних не відбувається зміни значення показника ефективності (вироджена задача лінійного програмування). Для припинення процедури пошуку  $W_{min}$  можливо використати ознаку того, що показник ефективності залишається незмінним після спеціально встановленої кількості циклів, або ознаку того, що показник ефективності вже неможливо зменшити, тому що немає додатних коефіцієнтів у рядку  $W$  симплекс таблиці.

**Зауваження 8.3.** У процесі розв'язання задачі лінійного програмування може виникнути зациклювання, суть якого полягає в тому, що після кількох замін змінних відбувається повернення до початкових значень опорного розв'язку та показника ефективності. Для подолання зациклювання потрібно після першого повернення до початкових умов оптимізації замінити розв'язувальний елемент.

**Приклад 8.2.** Розв'язати основну задачу лінійного програмування

$$W = 2x_1 - x_2 \rightarrow \min \begin{matrix} x_{1,2,3,4} \geq 0 \\ y_{1,2,3} \geq 0 \end{matrix}$$

при виконанні рівнянь-обмежень

$$\begin{cases} y_1 = x_1 - x_2, \\ y_2 = -x_2 + x_3 + 2, \\ y_3 = x_3 + x_4 + 1. \end{cases}$$

**Рішення**

□ 1. Перейдемо до стандартної форми задачі лінійного програмування:

$$W = 0 - -2x_1 + x_2 \rightarrow \min \begin{matrix} x_{1,2,3,4} \geq 0 \\ y_{1,2,3} \geq 0 \end{matrix}$$

$$\begin{cases} y_1 = 0 - -x_1 + x_2 , \\ y_2 = 2 - x_2 - x_3 , \\ y_3 = 1 - -x_3 - x_4 . \end{cases}$$

2. Із стандартної форми задачі лінійного програмування (див. п.1) випливає, що вільні коефіцієнти невід'ємні. Тобто при  $x_1=x_2=x_3=0$  має місце опорний розв'язок. Таким чином, опорний розв'язок шукати не потрібно. Переходимо до алгоритму покращення значень показника ефективності на опорних розв'язках, тобто до алгоритму пошуку оптимального розв'язку за допомогою стандартних таблиць(див. табл.. 8.5 та табл.. 8.6).

Таблиця 8.5

**Вихідні дані**

Показник ефективності, базисні змінні	Вільний елемент	Коефіцієнти при вільних змінних			
		$x_1$	$x_2 \leftrightarrow y_1$	$x_3$	$x_4$
$W$	0 0	-2 1	1 $\square$	0 0	0 0
$y_1 \leftrightarrow x_2$	$\square$ 0	$\square$ -1	$\bigcirc$ 1	$\square$ 0	$\square$ 0
$y_2$	2 0	0 1	1 $\square$	-1 0	0 0
$y_3$	1 0	0 0	0 $\square$	-1 0	-1 0

Таблиця 8.6

**Перша ітерація**

Показник ефективності, базисні змінні	Вільний елемент	Коефіцієнти при вільних змінних			
		$x_1$	$y_1$	$x_3$	$x_4$
$W$	0	-1	-1	0	0
$x_2$	0	-1	1	0	0
$y_2$	2	1	-1	-1	0
$y_3$	1	0	0	-1	-1

Отже:

1. Показник ефективності не змінився і дорівнює  $W = 0$ ;
2. У рядку із  $W$  у таблиці вже не існує додатних коефіцієнтів.

Тому робимо висновок, що оптимальний розв'язок досягнуто в точці  $x_1 = x_2 = x_3 = x_4 = y_1 = 0, x_2 = 2, y_3 = 1$ . ■

## 8.2. Розв'язання задачі лінійного програмування з використанням системи комп'ютерної математики Matlab

Постановка задачі:

$$W = f^T \cdot X \rightarrow \min_{X \in G},$$

де ОДР  $G$  задана системою нерівностей

$$\begin{cases} A \cdot X \leq b, \\ A_{eq} \cdot X = b_{eq}, \\ lb \leq X \leq Ub. \end{cases}$$

у якій прийняті наступні позначення:

$X$  – вектор змінних, за якими виконується оптимізація;

$f$  – матриця-стовпець коефіцієнтів показника ефективності;

$A$  – прямокутна матриця;

$b$  – матриця-стовпець вільних елементів(коефіцієнтів) у обмеженнях нерівностях;

$A_{eq}$  – прямокутна матриця;

$b_{eq}$  – матриця-стовпець вільних елементів(коефіцієнтів) у обмеженнях рівностях;

$lb, Ub$  – матриці-стовпці покомпонентних обмежень вектора  $X$ .

Методика застосування функції  $linprog()$  :

1. Виконати математичну постановку прикладної задачі;
2. Задати матриці  $f, A, b, A_{eq}, b_{eq}, lb, Ub$ ;
3. Звернення до функції розв'язок  $linprog()$  та отримання результату

$$[X, W] = linprog(f, A, b, A_{eq}, b_{eq}, lb, Ub)$$

4. Якщо якийсь із обмежень не задається, то замість цих аргументів при використанні функції *linprog()* задають порожні масиви у вигляді [...].

### Контрольні запитання та завдання

1. У чому полягає суть оптимального розв'язку задачі лінійного програмування?
2. Наведіть та поясніть алгоритм пошуку оптимального розв'язку основної задачі лінійного програмування з використанням симплекс-методу.
3. У якому випадку опорний розв'язок є оптимальним?
4. У яких випадках оптимальних розв'язків не існує?
5. Наведіть ознаку того, що показник ефективності вже не можна зменшити.
6. У чому полягає суть алгоритму пошуку оптимального розв'язку за допомогою стандартних таблиць?
7. Поясніть суть основної задачі лінійного програмування.
8. Наведіть дії, необхідні для подолання зациклювання у процесі розв'язання задачі лінійного програмування.
9. Поясніть призначення функції системи комп'ютерної математики Matlab *linprog()*.
10. Наведіть методику застосування функції *linprog()*.

## Глава 9. Транспортна задача лінійного програмування

### 9.1. Постановка транспортної задачі

Симплекс-метод – універсальний метод розв’язку задачі лінійного програмування, але існує клас задач лінійного програмування, які дозволяють отримати розв’язок простішими методами. Найбільш відомою є *транспортна задача*. Класична транспортна задача лінійного програмування формулюється наступним чином: існує  $m$  пунктів відправлення  $A_1, A_2, \dots, A_m$ , в яких зосереджено запаси деякого товару у кількості відповідно  $a_1, a_2, \dots, a_m$  одиниць. Крім того є  $n$  пунктів призначення  $B_1, B_2, \dots, B_n$ , що надали заявки на отримання товару відповідно у кількості  $b_1, b_2, \dots, b_n$  одиниць. Вважаємо, що сума усіх заявок дорівнює сумі усіх товарів:

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j. \quad (9.1)$$

Відомою є вартість  $c_{ij}$  – перевезення одиниці товару від кожного пункту відправлення  $A_i$  до кожного пункту призначення  $B_j$ . Таблиця вартості перевезення відома:

$$C = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn} \end{bmatrix}. \quad (9.2)$$

Необхідно скласти такий план перевезень, при якому всі заявки були б виконані і при цьому загальна вартість всіх перевезень була б мінімальною. При такій постановці задачі показником ефективності плану перевезень є вартість. Тому поставлену задачу називають транспортною задачею за критерієм вартості.

Дамо математичну постановку цієї задачі. Позначимо  $x_{ij}$   $x_{ij} \geq 0$  – кількість товару, який відправляється із  $i$ -го пункту відправлення  $A_i$  до  $j$ -го

пункту призначення  $B_j$   $i = \overline{1, m}, j = \overline{1, n}$ . На змінні  $x_{ij}$   $i = \overline{1, m}, j = \overline{1, n}$  накладаються наступні умови (обмеження) (рис. 9.1):

1. Сумарна кількість товару, що відвантажується з кожного пункту відправлення в усі пункти призначення, має дорівнювати запасу в даному пункті (рис. 9.1).

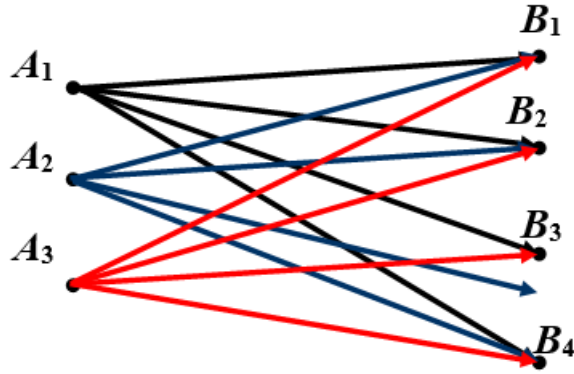


Рис. 9.1. Графічне зображення виникнення обмежень

Ця умова дає  $m$  рівнянь-обмежень:

$$\left\{ \begin{array}{l} \sum_{j=1}^n x_{1j} = a_1, \\ \dots\dots\dots \\ \sum_{j=1}^n x_{mj} = a_m. \end{array} \right. \quad (9.3)$$

2. Сумарна кількість товару, що поступає в кожен пункт призначення із усіх пунктів відправлення повинна дорівнювати замовленню, що надійшло від даного пункту. Ця умова дає  $n$  рівнянь-обмежень.

$$\left\{ \begin{array}{l} \sum_{i=1}^m x_{i1} = b_1, \\ \dots\dots\dots \\ \sum_{i=1}^m x_{in} = b_n. \end{array} \right. \quad (9.4)$$

3. Сумарна вартість всіх перевезень повинна бути мінімальною

$$W = c_{11}x_{11} + \dots + c_{mn}x_{mn} = \sum_{i=1}^m \sum_{j=1}^n c_{ij}x_{ij} \rightarrow \min_{x_{ij} \geq 0} \quad (9.5)$$

Показник ефективності (9.5), рівняння-обмеження (9.3). (9.4) є лінійними функціями відносно змінних  $x_{ij}$   $i = \overline{1, m}, j = \overline{1, n}$ , тобто задача про транспортні перевезення за вартісним критерієм поставлена як основна задача лінійного програмування. Як і будь яку задачу лінійного програмування, транспортну задачу лінійного програмування можливо розв'язати із використанням симплекс-методу розв'язання задачі лінійного програмування, але розмірність цієї задачі, яка дорівнює кількості змінних параметрів, які потрібно знайти, щоб задовольнити (9.5), дорівнює  $m \times n$ . Це означає, що при  $m = 10$  та  $n = 10$  (всього 10 складів та 10 пунктів прийому товару) отримаємо 100 змінних. Для досить простої транспортної задачі лінійного програмування різко зростає кількість невідомих і традиційний, розглянутий вище, симплекс-метод потребує значних обчислювальних потужностей.

Враховуючи той факт, що всі коефіцієнти при змінних в обмеженнях (9.3) та (9.4) дорівнюють 1 або 0 можливо спростити процедуру обчислення оптимального розв'язку. Перед початком викладення методу розв'язання транспортної задачі з'ясуємо особливості системи рівнянь обмежень (9.3) та (9.4).

Якщо скласти ліві і праві частини цих обмежень, то отримаємо рівняння: сума запасів дорівнює сумі замовлень. Тобто всього лінійно незалежних обмежень в (9.3) та (9.4) буде  $(m + n - 1)$ . Це означає, що ранг системи рівнянь-обмежень дорівнює  $r = m + n - 1$ . Тому можливо розв'язати ці рівняння відносно  $(m + n - 1)$  базисної змінної, вважаючи інші вільними змінними.

$$k = mn - (m + n - 1) = m(n - 1) - (n - 1) = (m - 1)(n - 1)$$

Раніше було показано, що в задачі лінійного програмування оптимальний розв'язок досягається в одній із вершин області допустимих розв'язків і при цьому хоча б  $k$  змінних перетворюються в 0. Це означає, що у випадку транспортної задачі  $(m - 1)(n - 1)$  змінних теж повинні дорівнювати 0 у вершині ОДР (опорній точці). При розв'язанні транспортної задачі зазвичай використовують наступну термінологію:

- перевезення  $x_{ij}$  – кількість одиниць вантажу, що перевозять із пункту відправлення  $A_i$  до пункту прийому  $B_j$ ;
- план перевезень  $x_{ij}$  – будь-яка сукупність значень  $x_{ij}$ ;
- допустимий план  $x_{ij \text{ д}}$  – план, який задовольняє балансовим умовам: усі заявки задоволені, усі запаси вичерпані, при цьому  $x_{ij} \geq 0$ ;
- опорний план  $x_{ij \text{ оп}}$  – допустимий план, в якому відмінні від 0 не більше, ніж  $r = m + n - 1$  базисних перевезень  $x_{ij}$ , а всі інші дорівнюють 0;
- оптимальний план  $x_{ij \text{ опт}}$  – опорний план, при якому досягається найменше значення показника ефективності (вартості перевезень).

Розглянемо методи розв'язання транспортної задачі. Вони не потребують застосування симплекс-таблиць, але використовують так звану транспортну таблицю (табл. 9.1).

Таблиця 9.1

Зразок транспортної таблиці

Пункт відправлення	Пункт призначення					Запаси
	$B_1$	$B_2$	$B_3$	...	$B_n$	
$A_1$	$c_{11}$	$c_{12}$	$c_{13}$	...	$c_{1n}$	$a_1$
$A_2$	$c_{21}$	$c_{22}$	$c_{23}$	...	$c_{2n}$	$a_2$
...	...	...	.....	...	...	...
$A_m$	$c_{m1}$	$c_{m2}$	$c_{m3}$	...	$c_{mn}$	$a_m$
$b_j$	$b_1$	$b_2$	$b_3$	...	$b_n$	$\sum b = \sum a$

До неї записують:

- пункти відправлення і пункти призначення;
- запаси, розташовані у пунктах відправлення;
- заявки, розташовані у пунктах призначення;

- вартість перевезень з кожного пункту відправлення в кожний пункт призначення одиниці вантажу.

Вартість перевезень  $c_{ij}$  розташуємо у правому верхньому куті кожного віконця, а всередині цього віконця будемо записувати перевезення  $x_{ij}$ . Пам'ятаючи визначення опорного плану, можливо стверджувати, що у  $r = m + n - 1$  віконці будуть записані числа, відмінні від 0. Ці віконця будемо називати базисними, а інші (пусті або нульові) – вільними. Таким чином, розв'язок транспортної задачі можливо знайти, якщо заповнити транспортну таблицю і при цьому виконати наступні умови:

- сума перевезень у кожному рядку транспортної таблиці дорівнює запасу даного пункту відправлення;
- сума перевезень у кожному стовпці транспортної таблиці дорівнює заявці пунктів прийому;
- загальна вартість перевезень повинна бути мінімальною.

Всі наступні дії, пов'язані із пошуком оптимального плану будуть використовувати зручну (наочну) форму представлення інформації про транспортну задачу у вигляді транспортної таблиці.

## 9.2. Табличний метод пошуку опорного плану (метод «північно-західного кута»)

Пошук розв'язку транспортної задачі, як і будь-якої задачі лінійного програмування, починається зі знаходження *опорного розв'язку* або *опорного плану*. На відміну від загального випадку основної задачі лінійного програмування з довільними обмеженнями і показником ефективності, розв'язок транспортної задачі завжди існує. Дійсно, із фізичних міркувань зрозуміло, що певне допустимий розв'язок (допустимий план) має існувати.

Пояснимо на конкретному прикладі основні прийоми методу пошуку опорного розв'язку, який отримав назву *метод «північно-західного кута»* (табл.9.2).

Припустимо, що умова транспортної задачі задана за допомогою транспортної таблиці. Необхідно знайти опорний розв'язок (побудувати опорний план). Розв'язання задачі починаємо із лівої верхньої позиції верхнього віконця із координатами 1;1, що відповідає образному

сприйняттю цієї позиції як «північно-західного» напрямку на географічній карті північної півкулі.

Таблиця 9.2

**Вихідні дані та приклад побудови опорного розв'язку**

Пункт відправлення	Пункт призначення					Запаси $a_i$
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	
$A_1$	10 18	8 27	5 3	6	9	48
$A_2$	6	7	8 30	6	5	30
$A_3$	8	7	10 9	8 12	7 6	27
$A_4$	7	5	4	6	8 20	20
$b_j$	18	27	42	12	26	125

За рахунок пункту відправлення  $A_1$  задовольняємо вимогу пункту призначення  $B_1$  і записуємо перевезення у віконце із координатами 1;1 . Потім задовольняємо запит пункту призначення  $B_2$  і далі записуємо залишок у віконце із координатами 1;3 . Недовиконання заявки пункту призначення  $B_3$  за рахунок пункту відправлення  $A_1$  доповнюємо поставками із  $A_2$  та  $A_3$  і т.д. Віконця таблиці, в яких стоять ненульові перевезення є базовими. Їх кількість повинна дорівнювати

$$r = m + n - 1 = 3 + 5 - 1 = 8 .$$

Вільні (порожні) означають 0 перевезень, тобто це вільні змінні, яких повинно бути щонайменше

$$k = (m - 1)(n - 1) = (4 - 1)(5 - 1) = 12 .$$

Таким чином, можна стверджувати, що опорний розв'язок знайдено. Обчислимо його вартість:

$$W = 18 \cdot 10 + 27 \cdot 8 + 3 \cdot 5 + 30 \cdot 8 + 9 \cdot 10 + 12 \cdot 8 + 6 \cdot 7 + 20 \cdot 8 = 1039 .$$

Зрозуміло, що знайдене значення  $W$  не є оптимальним, хоча б тому, що не використовувався алгоритм покращення опорного розв'язку. Перейдемо до алгоритму покращення опорного плану. Виконаємо так звану циклічну перестановку, наприклад, 18 одиниць вантажу із віконця з координатами

1;1 перемістимо у віконце із координатами 1;3 і далі через віконце 2;3 до 2;1, зберігаючи балансові співвідношення (див. табл.9.3).

Показник ефективності для транспортної таблиці набуває значення:

$$W = 18 \cdot 6 + 27 \cdot 8 + 21 \cdot 5 + 12 \cdot 8 + 9 \cdot 10 + 12 \cdot 8 + 6 \cdot 7 + 20 \cdot 8 = 913,$$

що на 126 одиниць менше показника ефективності, розрахованого для попереднього опорного плану. Наведений в табл. 9.3, як приклад, спосіб циклічної перестановки лежить в основі алгоритму оптимізації плану перевезень.

Таблиця 9.3

**Покращення опорного плану**

Пункт відправлення	Пункт призначення					Запаси $a_i$
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	
$A_1$	10	8 27	5 21	6	9	48
$A_2$	6 18	7	8 12	6	5	30
$A_3$	8	7	10 9	8 12	7 6	27
$A_4$	7	5	4	6	8 20	20
$b_j$	18	27	42	12	26	125

Перед тим, як перейти до алгоритму оптимізації, розглянемо так званий «вироджений» план транспортної задачі, коли нулю дорівнюють не лише вільні змінні, але й частина базисних (табл. 9.4).

Кількість базисних змінних згідно заданої транспортної таблиці повинна дорівнювати  $r = 4 + 5 - 1 = 8$ , але в отриманому за допомогою метода "північно-західного" кута плані перевезень ненульовими є 6 змінних, тобто виродженою є задача, в якій деякі з базисних змінних дорівнюють 0. Для побудови стійкого алгоритму обчислення оптимального розв'язку бажано на кожній ітерації мати всі базисні змінні ненульовими.

Приклад «виродженого» плану

Пункт відправлення	Пункт призначення					Запаси $a_i$
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	
$A_1$	10	10				20
$A_2$			20	10		30
$A_3$				25		25
$A_4$					20	20
$b_j$	10	10	20	35	20	95

З метою уникнення ситуації виникнення "виродженого" плану достатньо штучно на незначну величину  $\varepsilon$  змінити запаси і заявки (див. табл.9.5), а потім, після знаходження оптимального плану, покласти  $\varepsilon = 0$ .

Таблиця 9.5

Приклад уникнення «виродженого» плану

Пункт відправлення	Пункт призначення					Запаси $a_i$
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	
$A_1$	10	10	$\varepsilon$			$20 + \varepsilon$
$A_2$			$20 - \varepsilon$	$10 + \varepsilon$		30
$A_3$				$25 - \varepsilon$	$\varepsilon$	25
$A_4$					$20 - \varepsilon$	$20 - \varepsilon$
$b_j$	10	10	20	35	20	95

У табл. 9.5 маємо розрахункову кількість ненульових базисних змінних. Вона дорівнює 8. У табл. 9.3, 9.4 не вказана вартість перевезень одиниці товару  $c_{ij}$ , тому що автори ставили за мету розглянути лише приклад побудови опорного плану та прийоми уникнення появи «виродженого» плану.

### 9.3. Метод пошуку оптимального плану за допомогою циклічного перерахунку (метод циклічного перерахунку або розподільний метод)

Циклом у транспортній таблиці будемо називати кілька клітин, з'єднаних між собою уявною неперервною ламаною лінією, яка в кожній спеціально обраній клітині розвертається на  $90^\circ$  так, щоб утворити з початковою клітиною замкнену лінію. Так, наприклад, у табл. 9.6 наведено два цикли:

- Ц1 – з чотирма вершинами;
- Ц2 – з шістьма вершинами.

Кожен цикл має парну кількість вершин і сторін, тобто ланцюгів, що зв'язують ці вершини. Вершини позначимо «+» або «-» залежно від збільшення, або зменшення перевезень. Цикл з позначеними вершинами будемо називати *перенесеннями*. Перенести певну кількість одиниць вантажу зазначеним циклом означає збільшити перевезення, що містяться в додатних вершинах циклу, на деяку кількість одиниць вантажу, а перевезення, що містяться у від'ємних вершинах – зменшити на ту саму кількість. Зрозуміло, що за будь-якого циклічного перенесення, що зберігає невід'ємність перевезень, виконується закон збереження матерії, тобто зберігаються балансові умови, але вартість плану змінюється. Вартістю циклу називається зміна вартості перевезень під час переміщення одиниці вантажу зазначеним циклом.

Таблиця 9.6

Пояснення щодо методу циклічного перерахунку

Пункт відправлення	Пункт призначення					Запаси $a_i$
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	
$A_1$	$c_{11}$ -	$c_{12}$	$c_{13}$ +	$+ c_{14}$	$c_{15}$ -	$a_1$
$A_2$	$c_{21}$ +	$c_{22}$	$c_{23}$ -	$c_{24}$	$c_{25}$	$a_2$
$A_3$	$c_{31}$	$c_{32}$ +	$c_{33}$	$- c_{34}$	$c_{35}$	$a_3$
$A_4$	$c_{41}$	$c_{42}$ -	$c_{43}$	$c_{44}$	$c_{45}$ +	$a_4$
$b_j$	$b_1$	$b_2$	$b_3$	$b_n$	$b_n$	$\sum b = \sum a$

Кількісна вартість циклу дорівнює алгебраїчній сумі вартостей, що розташовані у вершинах циклу, із врахуванням знаку вершин. Наприклад, вартість Ц1 та Ц2 можливо обчислити:

$$V_{ц1} = c_{21} - c_{23} + c_{13} - c_{11},$$

$$V_{ц2} = c_{32} - c_{42} + c_{45} - c_{15} + c_{14} - c_{34}.$$

Якщо вартість циклу від'ємна, то переміщення по цьому циклу  $\Delta X$  одиниці вантажу призводить до зменшення загальної вартості перевезень на  $\Delta x \cdot V_{ц}$  одиниць.

Процедура зменшення загальної вартості припиняється тоді, коли циклів з від'ємною вартістю вже не буде. Тобто покращення плану перевезень бути не може. Оптимальний план досягнутий.

При покращенні плану перевезень за рахунок циклічних перенесень, як правило, користуються прийомами симплекс-методу: на кожній ітерації, присвяченій зміні циклу, замінюють одну базисну змінну вільною і навпаки. Загальна кількість базисних змінних залишається незмінною. Доведено, що для будь-якої вільної клітини транспортної таблиці завжди існує цикл і при тому єдиний, одна з вершин якого лежить у цій вільній клітині, а всі інші у базисних клітинах.

**Приклад 9.1.** Розглянемо комплексне застосування методу північно-західного кута та циклічного перерахунку для пошуку оптимального плану транспортної задачі, заданої за допомогою транспортної табл. 9.7.

Таблиця 9.7

**Вихідні дані прикладу 9.1**

Пункт відправлення	Пункт призначення				Запаси $a_i$
	$B_1$	$B_2$	$B_3$	$B_4$	
1	2	3	4	5	6
$A_1$	- 10	+ 7	6	8	31
	22 Ц1	9			
$A_2$	+ 5	- 6	5	4	48
		25	23		

Продовження таблиці 9.7

1	2	3	4	5	6
$A_3$	8	7	6	7	38
			18	20	
$b_j$	22	34	41	20	117

**Рішення**

□ 1. Складемо опорний план методом північно-західного кута. Кількість базисних змінних повинна дорівнювати

$$r = m + n - 1 = 6,$$

що відповідає плану, наведеному у таблиці 9.1. Задача «не вироджена».

Обчислимо значення показника ефективності

$$W = 22 \cdot 10 + 9 \cdot 7 + 25 \cdot 6 + 23 \cdot 5 + 18 \cdot 6 + 20 \cdot 7 = 796.$$

2. Покращимо план(див. табл. 9.8), замінивши вільну клітинку з координатами 2;1 .

Таблиця 9.8

**Покращений план за циклом Ц1**

Пункт відправлення	Пункт призначення				Запаси $a_i$
	$B_1$	$B_2$	$B_3$	$B_4$	
$A_1$	10	7	6	8	31
		31			
$A_2$	5	6	- 5	+ 4	48
	22	3	23	Ц2	
$A_3$	8	7	+ 6	- 7	38
			18	20	
$b_j$	22	34	41	20	117

Вартість циклу

$$V_{ц1} = 5 + 7 - 10 - 6 = -4$$

Покращимо план за рахунок переміщення 22 одиниць вантажу, інакше у 1;1 буде від'ємне число, чого неможливо допустити, бо від'ємних перевезень не може бути.

Покращений план(табл. 9.8) має показник ефективності

$$W_1 = 796 - 22 \cdot 4 = 708$$

Сформулюємо правило підбору циклу, що є чотирикутником:

- 1). Лише в одній вершині розташоване вільне віконце;
- 2). У вільному віконці завжди «+»;
- 3). Суми вартості по діагоналі, що включає вільне віконце менше за суму вартості іншої діагоналі.

3. Друга ітерація. Обираємо цикл Ц2(табл. 9.8) з вільним віконцем із координатами (2; 4):

$$V_{ц2} = 4 + 6 - 5 - 7 = -2$$

$$\Delta x \cdot V_{ц2} = 20 \cdot -2 = -40$$

$$W_2 = 708 - 40 = 668.$$

Результат перенесень за Ц2 представлено у табл. 9.9.

Таблиця 9.9

**Покращений план за циклом Ц2**

Пункт відправлення	Пункт призначення				Запаси $a_i$
	$B_1$	$B_2$	$B_3$	$B_4$	
$A_1$	10	7 31	6	8	31
$A_2$	5 22	6 3	5 3	4 20	48
$A_3$	8	7	6 38	7	38
$b_j$	22	34	41	20	117

План, наведений в табл. 9.9 є оптимальним, тому що всі цикли побудовані для вільних клітинок, мають додатну вартість. ■

Отже, розподільний (циклічний) метод полягає в безпосередньому пошуку вільних клітин транспортної таблиці з від'ємною вартістю циклу і в перенесенні перевезень за циклом.

**Приклад 9.2.** Знайти оптимальний план перевезень для транспортної задачі, заданої транспортною таблицею (табл. 9.10).

Таблиця 9.10

**Вихідні дані прикладу 9.2**

Пункт відправлення	Пункт призначення			Запаси $a_i$
	$B_1$	$B_2$	$B_3$	
$A_1$	- 10 20	7 20	+ 6 $\varepsilon$	$40+\varepsilon$
$A_2$	+ 5	6	+ 5 23	23
$A_3$	8	7	6 $20-\varepsilon$	$20-\varepsilon$
$b_j$	20	20	43	83

**Рішення**

□ 1. Будуємо опорний план методом північно-західного кута та перевіряємо його на невиродженість:

$$r = n + m - 1 = 3 + 3 - 1 = 5 > 4.$$

Опорний план вироджений.

2. Для уникнення виродженості використовуємо метод збурень запасів та заявок на величину  $\pm\varepsilon$ , відповідно у першому та третьому рядках.

3. Розподільчим методом покращуємо розв'язок, використовуючи цикл Ц1(див. табл. 9.11):

$$V_{\text{ц1}} = 5 + 6 - 10 - 5 = -4,$$

$$\Delta x \cdot V_{ц1} = 20 \cdot -4 = -80$$

Отримана транспортна таблиця(табл. 9.11) відображає оптимальний план перевезень. Циклів з від'ємною вартістю по відношенню до вільних клітинок немає. Показник ефективності дорівнює

$$W = 20 \cdot 5 + 20 \cdot 7 + 20 + \varepsilon \cdot 6 + 20 - \varepsilon \cdot 6 + 3 \cdot 5 \Big|_{\varepsilon=0} = 495.$$

Таблиця 9.11

**Покращення плану з використанням Ц1**

Пункт відправлення	Пункт призначення			Запаси $a_i$
	$B_1$	$B_2$	$B_3$	
$A_1$	10	7 20	6 $20 + \varepsilon$	$40 + \varepsilon$
$A_2$	5 20	6	5 3	23
$A_3$	8	7	6 $20 - \varepsilon$	$20 - \varepsilon$
$b_j$	20	20	43	83



**Контрольні запитання та завдання**

1. У чому полягає суть транспортної задачі лінійного програмування?
2. Наведіть та поясніть постановку транспортної задачі.
3. Розкрийте суть симплекс-методу розв'язання задачі лінійного програмування.
4. Наведіть математичну постановку транспортної задачі.
5. Поясніть процедуру пошуку розв'язку транспортної задачі методом «північно-західного кута».

## Глава 10. Метод потенціалів розв'язання транспортної задачі

### 10.1. Обґрунтування методу потенціалів

Розглянутий у главі 9 розподільчий метод розв'язання транспортної задачі має суттєвий недолік: необхідно знаходити цикли для всіх вільних віконць і обчислювати їх вартість. Для зменшення кількості обчислень за рахунок обчислення вартості лише тих циклів, де ця вартість від'ємна запропоновано *метод потенціалів*.

Ідея методу полягає в наступному. Припустимо, що задана транспортна задача у вигляді

$$W = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min \quad x_{ij} \geq 0 \quad (i = \overline{1, m}; j = \overline{1, n}) \quad (10.1)$$

за умови

$$\begin{cases} \sum_{j=1}^n x_{ij} = a_i, \\ \sum_{i=1}^m x_{ij} = b_j, \end{cases} \quad (10.2)$$

при цьому

$$\sum_{j=1}^n b_j = \sum_{i=1}^m a_i. \quad (10.3)$$

Уявимо, що кожен з пунктів відправлення  $A_i$  та призначення (прийому)  $B_j$  сплачують певні кошти у розмірі відповідно  $\alpha_i$  та  $\beta_j$ . Позначимо суму коштів  $\alpha_i + \beta_j = c_{ij}$ .

Підкреслимо, що  $\alpha_i, \beta_j, c_{ij}$  – уявні абстрактні вартості і платежі, тобто вони можуть бути як додатними, так і від'ємними або нульовими. Для скорочення запису позначимо всю сукупність платежів  $(\alpha_i, \beta_j)$ . Тоді матриця псевдовартостей матиме вигляд:

$$\begin{bmatrix} c_{11} = \alpha_1 + \beta_1 & c_{12} = \alpha_1 + \beta_2 & \dots & c_{1n} = \alpha_1 + \beta_n \\ c_{21} = \alpha_2 + \beta_1 & c_{22} = \alpha_2 + \beta_2 & \dots & c_{2n} = \alpha_2 + \beta_n \\ \dots & \dots & \dots & \dots \\ c_{m1} = \alpha_m + \beta_1 & c_{m2} = \alpha_m + \beta_2 & \dots & c_{mn} = \alpha_m + \beta_n \end{bmatrix}.$$

Доведемо теорему про платежі.

**Теорема 10.1.** Для заданої сукупності платежів  $(\alpha_i, \beta_j)$ , загальна псевдо-вартість перевезень не залежить від плану перевезень, а залежить лише від сукупності платежів

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

► Розглянемо вираз для обчислення псевдо вартості:

$$\begin{aligned} W &= \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} = \sum_{i=1}^m \sum_{j=1}^n (\alpha_i + \beta_j) x_{ij} = \sum_{i=1}^m \sum_{j=1}^n (\alpha_i) x_{ij} + \sum_{i=1}^m \sum_{j=1}^n (\beta_j) x_{ij} = \\ &= \sum_{i=1}^m \alpha_i \sum_{j=1}^n x_{ij} + \sum_{j=1}^n \beta_j \sum_{i=1}^m x_{ij} = \sum_{i=1}^m \alpha_i a_i + \sum_{j=1}^n \beta_j b_j. \end{aligned}$$

Як бачимо загальна псевдовартість перевезень не залежить від допустимого плану перевезень, тобто такого плану, що задовольняє системі балансових обмежень і умов  $x_{ij} \geq 0$ , а залежить лише від запасів  $a_i$ , заявок  $b_j$  та платежів  $(\alpha_i, \beta_j)$ , що і треба було довести. ◀

З'ясуємо зв'язок між псевдовартістю  $c_{ij}$  перевезення та її істинним значенням. Припустимо, що план перевезень  $(x_{ij})$  не вироджений: кількість базисних віконць з додатними перевезеннями у таблиці перевезень дорівнює  $m + n - 1$ . Покладемо, що в усіх базисних віконцях псевдовартості дорівнюють істинним вартостям

$$c_{ij} = \alpha_i + \beta_j = c_{ij} \quad \forall x_{ij} > 0.$$

У вільних віконцях  $x_{ij} = 0$  і при цьому співвідношення між  $c_{ij}$  та  $c_{ij}$  будь-яке. Це співвідношення є індикатором (показником) того, чи є план

перевезень, якому відповідає дана таблиця оптимальним, чи його можливо покращити.

Доведемо теорему.

**Теорема 10.2.** Якщо для всіх базисних віконць допустимого плану  $c_{ij} = c_{ij}$ , а для всіх вільних віконць допустимого плану  $c_{ij} \leq c_{ij}$ , то план є оптимальним.

Нагадаємо, що для базисних віконць перевезення повинні бути додатними, а для вільних віконць  $x_{ij} = 0$ .

► Визначимо вартість плану  $x_{ij}$  заданого умовами даної теореми та з урахуванням "теореми про платежі":

$$W = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} = \sum_{i=1}^m \alpha_i a_i + \sum_{j=1}^n \beta_j b_j = const.$$

Виконаємо заміну плану  $x_{ij} \rightarrow (x_{ij}^*)$ , в якому змінимо частково або повністю склад вільних або базисних змінних, але співвідношення між елементами матриці істинних вартостей та псевдовартостей залишилось незмінним, тобто таким, що відповідає умовам теореми. Це означає, що для базисних змінних  $(x_{ij}^*)$ , які теж були базисними у плані  $(x_{ij})$  виконується умова  $c_{ij} = c_{ij}$ , а для інших базисних змінних з плану виконується умова  $c_{ij} \geq c_{ij}$ .

В результаті отримаємо вираз для обчислення загальної вартості перевезень при новому плані  $(x_{ij}^*)$ :

$$W^* = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}^* \geq \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} = \sum_{i=1}^m \alpha_i a_i + \sum_{j=1}^n \beta_j b_j = const$$

що і треба було довести. ◀

**Зауваження 10.1.** Жодними змінами плану перевезень їх підсумкова вартість не може бути зменшена.

Прокоментуємо на прикладі зміст доведеної теореми. Припустимо, що  $m = n = 3$ . Тоді кількість базисних віконць повинні дорівнювати

$m + n - 1 = 5$ . Загальна вартість  $W$  при виконанні умов теореми обчислюється за формулою:

$$\begin{aligned}
 W &= \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \cdot \begin{bmatrix} x_{11} & x_{12} & 0 \\ 0 & x_{22} & x_{23} \\ x_{31} & 0 & 0 \end{bmatrix}^T \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \\
 &= \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \cdot \begin{bmatrix} x_{11} & 0 & x_{31} \\ x_{12} & x_{22} & 0 \\ 0 & x_{23} & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = c_{11}x_{11} + c_{12}x_{12} + c_{22}x_{22} + c_{23}x_{23} + c_{31}x_{31}.
 \end{aligned}$$

Припустимо, що план змінено, обчислимо вартість зміненого плану:

$$\begin{aligned}
 W^* &= \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \cdot \begin{bmatrix} 0 & x_{21}^* & 0 \\ x_{12}^* & 0 & 0 \\ x_{13}^* & x_{23}^* & x_{33}^* \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \\
 &= c_{12}x_{12}^* + c_{31}x_{31}^* + c_{21}x_{21}^* + c_{23}x_{23}^* + c_{33}x_{33}^* \geq \\
 &\geq c_{12}x_{12}^* + c_{31}x_{31}^* + c_{21}x_{21}^* + c_{23}x_{23}^* + c_{33}x_{33}^*.
 \end{aligned}$$

Доведена теорема виконується і для виродженого плану. Таким чином доведено, що ознакою оптимальності плану перевезень є виконання двох умов:

1.  $c_{ij} = c_{ij}$  для всіх базисних віконець;
2.  $c_{ij} \leq c_{ij}$  для всіх вільних віконець.

План для якого виконуються властивості 1 та 2 носить назву *потенціального*, а платежі, що йому відповідають  $(\alpha_i, \beta_j)$  носять назву *потенціалів пунктів*  $A_i$  та  $B_j$ .

Використовуючи вищезначені терміни, можливо переформулювати теорему коротше: будь-який потенціальний план є оптимальним.

## 10.2. Побудова потенціального плану

Для розв'язання транспортної задачі необхідно побудувати потенціальний план. Цей план можливо побудувати за рахунок послідовних наближень від початкового плану, знайденого, наприклад, методом північно-західного кута. Існує властивість платежів та псевдовартостей, яка полягає в

наступному: якою не була б система платежів  $(\alpha_i, \beta_j)$ , що задовольняє умові  $c_{ij} = c_{ij}$  для всіх базисних віконць, для кожного вільного віконця ціна циклу перерахунку дорівнює різниці між вартістю  $c_{ij}$  та псевдовартістю  $c_{ij}$ , що відноситься до даного віконця.

**Приклад 10.1.** Розглянемо транспортну задачу, не проставляючи запасів і заявок, наприклад, для випадку  $m = 5, n = 6$  (табл.10.1)

Таблиця 10.1

Приклад циклу перерахунку

Пункт відправлення	Пункт призначення					Запаси
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	
$A_1$	$c_{11}$	$c_{12}$	$c_{13}$	$c_{14}$	$c_{15}$	$c_{16}$
$A_2$	$c_{21}$	$c_{22}$	$c_{23}$	$c_{24}$	$c_{25}$	$c_{26}$
$A_3$	$c_{31}$	$c_{32}$	$c_{33}$	$c_{34}$	$c_{35}$	$c_{36}$
$A_4$	$c_{41}$	$c_{42}$	$c_{43}$	$c_{44}$	$c_{45}$	$c_{46}$
$A_5$	$c_{51}$	$c_{52}$	$c_{53}$	$c_{54}$	$c_{55}$	$c_{56}$

**Рішення**

□ Знайдемо

$$r = n + m - 1 = 5 + 6 - 1 = 10.$$

В таблиці для прикладу заштриховано 10 базисних віконць . Обираємо будь-яке вільне віконце, наприклад, 1, 5 і будемо цикл перерахунку так, щоб додатна вершина перебувала у цьому віконці, а всі інші в базисних. Ціна визначеного циклу дорівнює:

$$V_{15} = c_{15} - c_{35} + c_{33} - c_{23} + c_{22} - c_{12}.$$

Зважаючи на те, що для всіх базисних клітинок псевдовартості дорівнюють вартостям, отримуємо:

$$\begin{aligned} V_{15} &= c_{15} - \alpha_3 + \beta_5 + \alpha_3 + \beta_3 - \alpha_2 + \beta_3 + \alpha_2 + \beta_2 - \alpha_1 + \beta_2 = \\ &= c_{15} - \alpha_1 + \beta_5 = c_{15} - c_{15}. \end{aligned}$$

■

**Висновок.** У разі використання методу потенціалів для розв'язання транспортної задачі немає потреби визначати цикли з від'ємною ціною, що є найбільш трудомісткою процедурою розподільного методу.

### 10.3. Методика розв'язання транспортної задачі методом потенціалів

1. Виконати перше наближення до оптимального плану, наприклад, за допомогою метода північно-західного кута. Перевірити задачу на не виродженість, тобто обчислити кількість базисних клітинок, перевезення в яких повинні бути додатними. Якщо перше наближення продемонструвало виродженість транспортної задачі (кількість вільних віконць з нульовими перевезеннями перевищує  $m \cdot n - m + n - 1$ ), необхідно застосувати метод збурення заявок та запасів і перейти до не виродженої задачі.

2. Визначити для плану першого наближення платежі, виходячи із умови, що в будь-якому базисному віконці псевдовартості дорівнюють вартостям  $c_{ij} = \alpha_i + \beta_j = c_{ij}$ . Всього записаних вище рівнянь буде  $r = m + n - 1$ , а невідомих, що входять до їх складу  $\alpha_i$   $i = \overline{1, m}$  та  $\beta_j$   $j = \overline{1, n}$  -  $m + n$ . Тобто одну з невідомих можливо для простоти покласти нульовою, а всі інші обчислити так, щоб виконувалось записане вище рівняння.

3. Обчислити псевдовартості для всіх вільних віконць. Якщо виявляється, що псевдовартості вільних віконць не перевищують вартості цих віконць, то оптимальний план знайдено.

4. Якщо хоча б у одному з вільних віконць псевдовартість перевищує вартість  $c_{ij} > c_{ij}$ , то потрібно покращити план перевезень за рахунок

переміщення перевезень за циклом, який відповідає обраному вільному віконцю із від'ємною різницею  $0 > \tilde{c}_{ij} - c_{ij}$ .

5. Знову обчислити платежі та псевдовартості. Перерахунки завершуються лише тоді, коли буде знайдено оптимальний план.

**Зауваження 10.2.** Поняття платежів, вартостей, псевдоплатежів, псевдовартостей, мають наглядну економічну інтерпретацію: компанії групи  $A$ , які зберігають товари, та компанії групи  $B$ , що накопичують та реалізують товари, разом фінансують перевезення. Оптимальним є той план, коли компанії  $A$  та  $B$  за ці перевезення не переплачують.

**Приклад 10.2.** Розв'язати транспортну задачу, задану табл. 10.2.

Таблиця 10.2

**Вихідні дані**

Пункт відправлення	Пункт призначення					Запаси $a_i$	Платежі $\alpha_i$
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$		
$A_1$	10 10	8 8	6 9	3 6	4 5	25	0
	17 -	8Ц1 +					
$A_2$	8 5	6 6	4 4	3 3	2 8	32	-2
	+ 13 -	19					
$A_3$	9 9	7 7	5 5	4 4	3 3	40	-1
			22	14	4		
$A_4$	14 14	12 10	10 8	9 8	8 8	20	4
					20		
$b_j$	17	21	41	14	24	117	
Платежі $\beta_j$	10	8	6	5	4		

### Рішення

□ Знайдемо методом північно-західного кута перше наближення до оптимального плану(див. табл. 10.2). Кількість ненульових віконць дорівнює 8, що співпадає із контрольним числом

$$r = m + n - 1 = 4 + 5 - 1 = 8 .$$

**Висновок.** Транспортна задача є не виродженою.

Додамо до транспортної таблиці рядок та стовпець платежів відповідно  $\beta_j$  та  $\alpha_i$  (табл.10.2.). Псевдовартості запишемо ліворуч зверху в кожній клітині (табл.10.3.).

Таблиця 10.3

#### Перша ітерація

Пункт відправлення	Пункт призначення					Запаси $a_i$	Платежі $\alpha_i$
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$		
$A_1$	10 10 4 -	8 8 21	9 9	8 6	7 5	25	0
$A_2$	5 5 13 +	6 6	4 4 19	3 3	2 8	32	-5
$A_3$	6 9	4 7	5 5 22 +	4 4 14 -	3 3 4	40	-4
$A_4$	11 14	9 10	10 8	9 8	8 8 20	20	1
$b_j$	17	21	41	14	24	117	
Платежі $\beta_j$	10	8	9	8	7		

Враховуючи, що для базисних віконць  $c_{ij} = \alpha_i + \beta_j$ , знаходимо  $c_{11} = c_{11} = 10 = 0 + \beta_1$  (за припущенням  $\alpha_1 = 0$ ), тобто  $\beta_1 = 10$ . Аналогічно знаходимо  $\beta_2 = 8$ . Далі  $\alpha_2 + \beta_2 = 6 \rightarrow \alpha_2 = -2$  і т.д.(табл. 10.2). Виконаємо

перерахунок базисного віконця за циклом Ц1(табл. 10.2). Перенесемо за цим циклом 13 одиниць вантажу (більше неможливо, інакше у віконці 2,2 виникне від'ємне число і план не буде допустимим) і зменшимо вартість плану на  $13 \cdot 3 = 39$ .

Не всі псевдовартості у вільних віконцях ( див.табл. 10.3) задовольняють умові  $c_{ij} \leq c_{ij}$ . Тому, план, наведений у таблиці першої ітерації не є оптимальним. Спробуємо покращити цей план шляхом переведення у базисні змінні одну з вільних змінних, для якої у відповідному віконці  $c_{ij} > c_{ij}$ , наприклад, у віконці 1,4 . Ціна циклу для віконця 1,4 дорівнює  $v_{14} = 6 - 8 = -2$ .

Перенесемо по циклу Ц2 (табл. 10.3) 4 одиниці і занесемо результат у таблицю другої ітерації(див табл. 10.4).

Таблиця 10.4

**Друга ітерація**

Пункт відправлення	Пункт призначення					Запаси $a_i$	Платежі $\alpha_i$
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$		
$A_1$	8 10	8 8 <b>21</b>	7 9	6 6 <b>4</b>	5 5	25	0
$A_2$	5 5 <b>17</b> +	5 6	4 4 <b>15</b>	3 3	2 8	32	-3
$A_3$	6 9	6 7	5 5 <b>26</b>	4 4 <b>10</b>	3 3 <b>4</b> +	40	-2
$A_4$	11 14	11 10	10 8 +	9 ЦЗ 8 <b>20</b>	8 8	20	3
$b_j$	17	21	41	14	24	117	
Платежі $\beta_j$	8	8	7	6	5		

План другої ітерації теж не оптимальний. Обираємо віконце 4,3 і переносимо 20 одиниць перевезень по циклу ЦЗ. Остаточоно отримаємо таблицю третьої ітерації(див табл. 10.5), яка за всіма ознаками надає оптимальний план перевезень (усі псевдовартості не перевищують відповідних вартостей).

Значення показника ефективності дорівнює:

$$W = 21 \cdot 8 + 4 \cdot 6 + 17 \cdot 5 + 15 \cdot 4 + 6 \cdot 5 + 10 \cdot 4 + 24 \cdot 3 + 20 \cdot 8 = 639$$

Таблиця 10.5

**Третя ітерація**

Пункт відправлення	Пункт призначення					Запаси $a_i$	Платежі $\alpha_i$
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$		
$A_1$	8 10	8 8 21	7 9	6 6 4	5 5	25	0
$A_2$	5 5 17 +	5 6	4 4 15	3 3	2 8	32	-3
$A_3$	6 9	6 7	5 5 6 -	4 4 10	3 3 24 +	40	-2
$A_4$	9 14	11 10	8 8 20	7 8	6 8	20	1
$b_j$	17	21	41	14	24	117	
Платежі $\beta_j$	8	8	7	6	5		

Таким чином, фірма, яка володіє усіма складами  $A$  та всіма пунктами реалізації  $B$ , але не включає до свого складу перевізників, має заплатити мінімальну суму  $W_{min} = 639$ . Платежі  $\alpha_i$  та  $\beta_j$  мають умовний зміст і використані як математична абстракція. ■

**Приклад 10.3.** Розв'язати методом потенціалів вироджену транспортну задачу. Умови транспортної задачі задані вихідною транспортною таблицею 10.6,  $m = 3$ ,  $n = 3$ .

Таблиця 10.6

**Вихідні дані**

Пункт відправлення	Пункт призначення			Запаси $a_i$	Платежі $\alpha_i$
	$B_1$	$B_2$	$B_3$		
$A_1$	6    6 20	4    4 ε	3    2	20+ε	0
$A_2$	7    3 +	5Ц1 5-ε	5    4 2ε	25+ε	1
$A_3$	6    3	4    6	3    3 30-2ε	30-2ε	0
$b_j$	20	25	30	75	
Платежі $\beta_j$	6	4	3		

**Рішення**

□ Перевіримо:  $r = m + n - 1 = 3 + 3 - 1 = 5 > 3$  – транспортна задача вироджена. Скористаємось методом збурення запасів та заявок(див табл. 10.6).

Перерахунок за циклом Ц1 дозволяє заповнити таблицю першої ітерації(див. табл. 10.7), а за циклом Ц2 – таблицю другої ітерації(див.табл10.8).

$$W_{\min}|_{\epsilon=0} = 20 \cdot 4 + 20 \cdot 3 + 5 \cdot 5 + 30 \cdot 3 = 255$$

Підкреслимо, що у виродженій транспортній задачі при виборі Ц1 або Ц2 із іншими вільними віконцями остаточний, тобто оптимальний, план зміниться, але мінімальне значення показника ефективності залишиться незмінним:  $W = 255$ . ■

Таблиця 10.7

## Перша ітерація

Пункт відправлення	Пункт призначення						Запаси $a_i$	Платежі $\alpha_i$
	$B_1$		$B_2$		$B_3$			
$A_1$	6	6	4	4	3	2	20+ $\epsilon$	0
	20	-	20+ $\epsilon$	+	112	+		
$A_2$	3	3	5	5	4	4	25+ $\epsilon$	1
	20		5- $\epsilon$	+	2 $\epsilon$	-		
$A_3$	2	3	4	6	3	3	30-2 $\epsilon$	0
					30-2 $\epsilon$	-		
$b_j$	20		25		30		75	
Платежі $\beta_j$	2		4		3			

Таблиця 10.8

## Друга ітерація

Пункт відправлення	Пункт призначення						Запаси $a_i$	Платежі $\alpha_i$
	$B_1$		$B_2$		$B_3$			
$A_1$	2	6	4	4	2	2	20+ $\epsilon$	0
			20+ $\epsilon$	+	2 $\epsilon$			
$A_2$	3	3	5	5	3	4	25+ $\epsilon$	1
	20		5+ $\epsilon$					
$A_3$	3	3	5	6	3	3	30-2 $\epsilon$	1
					30-2 $\epsilon$	-		
$b_j$	20		25		30		75	
Платежі $\beta_j$	2		4		3			

**Висновок.** Особливістю вироджених транспортних задач є те, що однакових мінімальних значень показника ефективності можна досягнути на кількох різних планах. Тобто, для «вироджених» транспортних задач характерною ознакою є неоднозначність розв'язку.

### **Контрольні запитання та завдання**

1. Розкрийте суть транспортної задачі.
2. У чому полягає суть методу потенціалів розв'язання транспортної задачі?
3. Сформулюйте ознаки оптимальності плану перевезень.
4. Розкрийте суть поняття «потенціальний план».
5. Поясніть процедуру побудови потенціального плану.
6. Наведіть та поясніть методику розв'язання транспортної задачі методом потенціалів.

## Глава 11. Деякі окремі випадки транспортних задач

### 11.1. Транспортні задачі з неправильним балансом

Досі розглядалися транспортні задачі із виконанням умови: сума запасів дорівнює сумі заявок

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j. \quad (11.1)$$

Ця класична транспортна задача має назву *транспортної задачі з правильним балансом*. Порушення умови правильного балансу перетворює класичну транспортну задачу в *транспортну задачу із неправильним балансом*. Порушення балансу поділяють на 2 типи:

1. Сума запасів у пунктах відправлення перевищує суму заявок

$$\sum_{i=1}^m a_i > \sum_{j=1}^n b_j, \quad (11.2)$$

тобто має місце *транспортна задача із надмірністю запасів*.

2. Сума заявок перевищує наявні запаси

$$\sum_{i=1}^m a_i < \sum_{j=1}^n b_j \quad (11.3)$$

тобто маємо *транспортну задачу із надмірністю заявок*.

У випадку надмірності запасу, постановка транспортної задачі набуває вигляду: знайти такий план перевезень  $x_{ij}$ , при якому усі заявки будуть виконані, а загальна вартість перевезень набуває мінімального значення:

$$W = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min_{x_{ij} \geq 0} \quad (11.4)$$

при виконанні системи обмежень

$$\begin{cases} \sum_{j=1}^n x_{ij} < a_i & i = \overline{1, m_1}, \\ \sum_{j=1}^n x_{ij} = a_i & i = \overline{m_1 + 1, m}, \\ \sum_{i=1}^m x_{ij} = b_j & j = \overline{1, n}, \end{cases} \quad (11.5)$$

та умов (11.2).

Як бачимо, частина або всі рівняння-обмеження перетворюються в нерівності-обмеження. Розв'язок цієї транспортної задачі можливо отримати або класичним симплекс-методом або за допомогою методів розв'язання транспортних задач, розглянутих вище (розподільчий метод, метод потенціалів).

Перед тим, як скористатись відомими методами розв'язання транспортної задачі, зведемо транспортну задачу із надмірністю запасів, до класичної транспортної задачі із правильним балансом.

Для цього використаємо уявний пункт призначення (фіктивний, хибний)  $B_\Phi$ , якому належить заявка, що якраз дорівнює перевищенню запасів над заявками:

$$b_\Phi = \sum_{i=1}^m a_i - \sum_{j=1}^n b_j$$

Покладемо вартості перевезень від всіх пунктів відправлення до фіктивного пункту призначення  $B_\Phi$  рівними 0, тобто  $c_{i\Phi} = 0 \quad i = \overline{1, m}$ .

За фізичним змістом  $x_{i\Phi}$  – означає кількість вантажу, який залишається невідправленим з  $i$ -го пункту відправлення.

**Висновок.** За рахунок фіктивного пункту призначення  $B_\Phi$  із заявкою  $b_\Phi$ , всі нерівності-обмеження у транспортній задачі із надмірністю запасів перетворено у рівняння-обмеження, тобто задача зведена до класичної транспортної задачі із правильним балансом.

### ***Транспортна задача з надмірністю запасів***

У цьому випадку виконується нерівність (11.3):

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

тобто запасів недостатньо для задоволення усіх заявок. Необхідно скласти план таких перевезень, при якому усі запаси будуть вивезені і при цьому вартість перевезень буде мінімальною. Для розв'язання введемо фіктивний пункт відправлення  $A_\Phi$ , із фіктивним уявним запасом  $a_\Phi$ , який дорівнює самій кількості вантажу, якої не вистачає для задоволення усіх заявок

$$a_{\Phi} = \sum_{j=1}^n b_j - \sum_{i=1}^m a_i$$

При цьому покладемо, що усі вартості перевезень із фіктивного пункту відправлення, до будь якого пункту призначення, дорівнюють 0:

$$c_{\Phi j} = 0 \quad j = \overline{1, n}$$

Складові плану перевезень  $x_{\Phi j}$  ( $j = \overline{1, n}$ ) показують величину недопостачання вантажу у відповідний пункт призначення.

**Висновок.** За допомогою фіктивного пункту відправлення транспортну задачу з надмірністю замовлень зведено до транспортної задачі з правильним балансом.

**Зауваження 11.1.** Окрім прийому зведення транспортної задачі із неправильним балансом до транспортної задачі із правильним балансом за допомогою фіктивних пунктів відправлень та пунктів прийому, можливо використати умови «нормування», наприклад домножити усі заявки на коефіцієнт

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

для транспортної задачі із надмірністю заявок або на коефіцієнт

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

для транспортних задач із надмірністю запасів.

**Приклад 11.1.** Розв'язати транспортну задачу з надмірністю запасів (табл. 11.1).

## Вихідні дані

Пункт відправлення	Пункт призначення				Запаси $a_i$	Платежі $\alpha_i$
	$B_1$	$B_2$	$B_3$	$B_\Phi$		
$A_1$	5 5 18	7 7 21	6 6 11	1 0 0	50	0
$A_2$	4 6	6 6	5 5 22Ц1	0 0 18	40	-1
$A_3$	4 8	6 4 +	5 5 -	0 0 20 -	20	-1
$b_j$	18	21	33	38	110	
Платежі, $\beta_j$	5	7	6	1		

**Рішення**

□ Перевіряємо балансове співвідношення

$$\left. \begin{aligned} a_1 + a_2 + a_3 &= 50 + 40 + 20 = 110, \\ b_1 + b_2 + b_3 &= 18 + 21 + 23 = 72, \end{aligned} \right\} \Rightarrow \sum_{i=1}^3 a_i > \sum_{j=1}^3 b_j.$$

Має місце надмірність запасів.

Використовуємо фіктивний пункт призначення  $B_\Phi$  із заявкою  $b_\Phi = 110 - 72 = 38$ . В результаті використання  $B_\Phi$  вдалося перейти до транспортної задачі із правильним балансом.

Розв'язуємо класичну транспортну задачу методом потенціалів. Перевіряємо задачу на невиродженість:

$$r = n + m - 1 = 4 + 3 - 1 = 6.$$

Кількість ненульових базисних змінних, отриманих методом північно-західного кута дорівнює 6 і співпадає із  $r$ . Тобто транспортна задача (табл. 11.1) є невиродженою. Виконаємо заміну базисних і вільних віконць за циклами Ц1 (табл. 11.1), Ц2 (табл. 11.2) та Ц3 (табл. 11.3).

Таблиця 11.2

## Перше наближення

Пункт відправлення	Пункт призначення				Запаси $a_i$	Платежі $\alpha_i$
	$B_1$	$B_2$	$B_3$	$B_\phi$		
$A_1$	5 5 <b>18</b>	7 7 1 -	6 6 31 -	1 0 +	50	0
$A_2$	4 6	6 6	5 Ц2 2 +	5 0 38 -	40	-1
$A_3$	2 8	4 4 20	3 5 -	-2 0	20	-3
$b_j$	18	21	33	38	110	
Платежі $\beta_j$	5	7	6	1		

Таблиця 11.3

## Друге наближення

Пункт відправлення	Пункт призначення				Запас и $a_i$	Платежі $\alpha_i$
	$B_1$	$B_2$	$B_3$	$B_\phi$		
$A_1$	5 5 <b>18</b>	7 7 1 -	5 6	0 0 31 +	50	0
$A_2$	5 6	7 +6	5 Ц3 33	5 0 31	40	0
$A_3$	2 8	4 4 20	2 5	-3 0	20	-3
$b_j$	18	21	33	38	110	
Платежі $\beta_j$	5	7	5	0		

Заповнимо таблиці першого (табл. 11.2), другого (табл. 11.3) та третього (табл. 11.4) наближення відповідно.

## Третє наближення

Пункт відправлення	Пункт призначення				Запаси $a_i$	Платежі $\alpha_i$
	$B_1$	$B_2$	$B_3$	$B_\phi$		
$A_1$	5 5 18	6 7	5 6	0 0 32	50	0
$A_2$	5 6	6 6 1	5 5 33	0 0 6	40	0
$A_3$	3 8	4 4 20	3 5	-2 0	20	-2
$b_j$	18	21	33	38	110	
Платежі $\beta_j$	5	6	5	0		

**Висновок:**

1. Третє наближення дає оптимальний план, із якого випливає, що  $W_{min} = 341$ .
2. У пункті відправлення  $A_1$  залишилось невідправлених 32, у  $A_2$  – 6, а в  $A_3$  – 0. ■

**11.2. Розв'язок транспортної задачі за критерієм часу**

Досі транспортна задача ставилась як задача, у якій необхідно було мінімізувати вартість перевезень. Але в багатьох випадках практики важливу роль відіграє тривалість часу  $T$ , протягом якого усі перевезення буде завершено. Так, наприклад: при ліквідації наслідків надзвичайних ситуацій необхідно мінімізувати як час надходження інформації про стихійне лихо, так і час до надання допомоги; при багатопроекторній обробці інформації необхідно розподілити обчислювальні ресурси, щоб мінімізувати тривалість обчислень.

Найкращим слід вважати такий план перевезень  $x_{ij}$ , при якому час закінчення усіх перевезень мінімізується

$$T \rightarrow \min_{\substack{x_{ij} \geq 0 \\ i=1, m, j=1, n}} \quad (11.6)$$

Транспортна задача, у якій оптимальним вважається план із мінімальним часом усіх перевезень, носить назву *транспортна задача за критерієм часу*.

Математична постановка транспортної задачі за критерієм часу стосовно балансових обмежень виконується так само, як і класична транспортна задача, із тією різницею, що критерій, тобто показник ефективності, має вигляд:

$$T = \max t_i \rightarrow \min_{x_{ij} > 0} \quad (11.7)$$

і виконуються балансові обмеження:

$$\sum_{i=1}^m x_{ij} = b_j \quad (j = \overline{1, n}), \quad (11.8)$$

$$\sum_{j=1}^n x_{ij} = a_i \quad (i = \overline{1, m}). \quad (11.9)$$

**Зауваження 11.2.** Використання строгого обмеження  $x_{ij} > 0$  показує, що максимальний інтервал часу обирається не взагалі, а лише з тих віконць транспортної таблиці, в яких перевезення строго більше нуля (строго додатні, тобто реально виконуються). Для розв'язку транспортної задачі за критерієм часу теж зручно використовувати транспортну таблицю, у якій в правому верхньому куті буде записано  $t_{ij}$  – час перевезення вантажу із пункту відправлення  $i$  в пункт призначення  $j$  (замість  $c_{ij}$ ).

Загалом транспортна задача за критерієм часу не відноситься до задач лінійного програмування тому, що показник ефективності  $T$  є нелінійною функцією змінних  $x_{ij}$ . Але зручна форма транспортної таблиці, яка використовується для розв'язку класичної транспортної задачі, дозволяє виконати обчислення без порушень балансових обмежень, тобто побудувати алгоритм, який ґрунтується на використанні допустимого плану.

Метод обчислень, що мінімізує час перевезень  $T$ , побудований із використанням транспортної таблиці і технології циклічних перенесень, які дозволяють не порушити балансові обмеження, отримав назву *методу*

заборонених віконць. "Забороненими" вважаються віконця без перевезень і при цьому із найбільшим часом перевезень  $t_{ij}$ . "Заборона" вказує, із якого допустимого віконця потрібно "виштовхнути" перевезення в інше віконце із меншим часом перевезень. Покращення плану перевезень припиняється тоді, коли найбільший час перевезень вже зменшити неможливо. Ознака: всі віконця заблоковані перевезеннями або заборонені до перевезень.

**Приклад 11.2.** Вихідні дані транспортної задачі наведені в табл. 11.5. Необхідно мінімізувати час перевезень.

Таблиця 11.5

Вихідні дані

Пункт відправлення	Пункт призначення					Запаси $a_i$
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	
$A_1$	10	8	5 25	6	7	25
$A_2$	5 21	6 Ці	6 +	6 2	9 11	34
$A_3$	4 +	8 14	7 -	8 13	5 15	42
$A_4$	11	4 23	5	8	9	23
$b_j$	21	37	40	11	15	124

### Рішення

□ 1. Викреслюємо віконця із координатами 1,1 , 2,5 , 4,1 та 4,5 , як віконця із значним часом перевезень.

2. Будь-яким способом заповнюємо транспортну таблицю допустимим планом:

1) Якщо транспортну таблицю заповнити будь-яким планом не звертаючи уваги на баланс записів та заявок, але лише виконувати умову невід'ємності перевезень, то отримаємо план перевезень;

2) Якщо транспортну таблицю, додатково до невід'ємності, заповнити такими значеннями  $x_{ij}$ , які враховують умови виконання балансу записів та заявок, то отримаємо допустимий план перевезень;

3) Якщо вимоги п.2. доповнити вимогою додатності  $r = m + n - 1$  базових перевезень та нульовими перевезеннями в інших віконцях таблиці, то отримаємо опорний план перевезення;

Підкреслимо ще раз, що *оптимальний план перевезень для транспортної задачі за критерієм вартості* досягається лише на опорному плані. Для транспортної задачі за критерієм часу оптимальний план перевезень не обов'язково досягається на опорному, він може досягатися на будь-якому допустимому плані перевезень. Виконаємо перенесення перевезень за циклами Ц1 (див. табл. 11.6) та Ц2 (див. табл. 11.7).

Таблиця 11.6

**Перша ітерація**

Пункт відправлення	Пункт призначення					Запаси $a_i$
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	
$A_1$	10	8	5	6	5	25
$A_2$	7	6	6	Ц2 6	9	34
$A_3$	1	8	2	8	-5	42
$A_4$	11	4	5	8	9	23
$b_j$	21	37	40	11	15	124

Таблиця 11.7

**Друга ітерація**

Пункт відправлення	Пункт призначення					Запаси $a_i$
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	
$A_1$	10	8	5	6	5	25
$A_2$	5	6	6	6	9	34
$A_3$	4	8	5	8	-5	42
$A_4$	11	4	5	8	9	23
$b_j$	21	37	40	11	15	124

Отже,  $T_{min} = 6$ . ■

**Приклад 11.3.** Скласти план перевезень, якщо вихідні дані транспортної задачі задані табл. 11.8.

**Рішення**

□

Таблиця 11.8

**Вихідні дані**

Пункт відправлення	Пункт призначення					Запаси $a_i$
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	
$A_1$	1 15	3 5	4	2	5	20
$A_2$	5	2 15	3 15	1	2	30
$A_3$	8	7	6 10	10	4 25	35
$A_4$	3	4	3	6 25	5 15	40
$b_j$	15	20	25	25	40	125

Таблиця 11.9

**Остаточний результат**

Пункт відправлення	Пункт призначення					Запаси $a_i$
	$B_1$	$B_2$	$B_3$	$B_4$	$B_5$	
$A_1$	1 15	3 5	4	2 10	5	20
$A_2$	5	2 10	3	1 15	2 5	30
$A_3$	8	7	6	10	4 35	35
$A_4$	3 5	4 10	3 25	6	5	40
$b_j$	15	20	25	25	40	125

Отже,  $T_{min} = 4$ . ■

## Контрольні запитання та завдання

1. У чому полягає суть транспортної задачі?
2. Порушення яких умов перетворює класичну транспортну задачу в транспортну задачу із неправильним балансом?
3. Які методи застосовують для розв'язання транспортної задачі з неправильним балансом?
4. У чому полягає суть транспортної задачі за критерієм часу. Наведіть приклади.
5. Наведіть та поясніть математичну постановку транспортної задачі за критерієм часу.

## Глава 12. Післяоптимізаційний аналіз задачі лінійного програмування

### 12.1. Ідея аналізу параметричної чутливості

Післяоптимізаційний аналіз або, як ще кажуть, аналіз чуттєвості полягає у зв'язуванні впливу структурних, параметричних та структурно-параметричних змін у математичній моделі (математичній постановці) задачі на отриманий оптимальний розв'язок для тієї постановки задачі лінійного програмування, яка вважається вихідною.

Розглянемо приклад 5.1, в якому було з'ясовано, що для отримання максимального прибутку необхідно випустити продукцію типу  $A$  (супутники зв'язку)  $\hat{x}_1 = 17\frac{1}{7} \approx 17$  та типу  $B$  (навігаційні супутники)  $\hat{x}_2 = 23\frac{4}{7} \approx 23$ .

Оптимальний розв'язок було отримано за умови, що вартість виробів  $A$  та  $B$  складає відповідно 40 та 50 умовних одиниць. У зв'язку зі змінами, що відбуваються в світовій економіці, керівнику фірми (особі, що приймає рішення) важливо знати, як вплине зміна вартості продукції  $A$  та  $B$  (питомий прибуток) на запланований випуск продукції  $\hat{x}_{1,2}$  (див. рис.12.1).

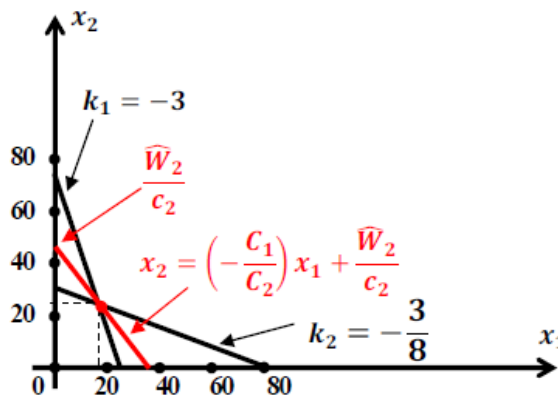


Рис. 12.1. Графік визначення меж зміни співвідношення між  $c_1$  та  $c_2$  у виразі для обчислення показника ефективності, за яких оптимальний розв'язок  $(x_1, x_2)$  залишається незмінним

Як бачимо, при умові

$$-3 < -\frac{c_1}{c_2} < -\frac{3}{8} \Leftrightarrow 3 > \frac{c_1}{c_2} > \frac{3}{8} \Leftrightarrow 3 > \frac{c_1}{c_2} > 0,375$$

оптимальний план випуску продукції не зміниться, тому що «основна пряма»

$$W = c_1x_1 + c_2x_2$$

проходитиме через кутову точку із координатами  $\left( \hat{x}_1 = 17\frac{1}{7}, \hat{x}_2 = 23\frac{4}{7} \right)$ , що

забезпечує максимальне значення показника ефективності. При  $\frac{c_1}{c_2} = \frac{3}{8}$  або

$\frac{c_1}{c_2} = 3$  має місце невизначеність.

Якщо  $\frac{c_1}{c_2} < \frac{3}{8}$ , то оптимальною точкою буде 0,30, а якщо  $\frac{c_1}{c_2} > 3$ , то

оптимальною точкою буде 25,0, що з фізичної точки зору означає відповідно недоцільність випуску супутників зв'язку або навігаційних супутників. При зміні співвідношення між  $c_1$  та  $c_2$  відбувається зміна оптимального значення показника ефективності.

Зрозуміло, що взагалі зміна будь-якого параметра математичної моделі задачі лінійного програмування (об'єм ресурсів та їх вартість), норми витрат ресурсів на виготовлення одиниці продукції впливатиме на оптимальне рішення і значення показника ефективності. Кількісне дослідження цих змін і виконується в процесі аналізу чутливості математичної моделі задачі лінійного програмування. Наведений простий приклад аналізу чутливості оптимального рішення задачі лінійного програмування та значення показника ефективності до зміни параметрів математичної моделі має ілюстративний характер і для використання у багатовимірному випадку ( $n \geq 3$ , де  $n$  – кількість змінних) стає громіздким.

Для дослідження впливу властивостей задачі лінійного програмування на її оптимальне за вихідною постановкою рішення в загальному багатовимірному випадку використовують спеціальним чином переформульовану вихідну задачу лінійного програмування, яка отримала назву *двоїста (спряжена) задача лінійного програмування*, або використовують так зване *параметричне програмування*.

## 12.2. Ідея фізичного змісту побудови математичної моделі двоїстої задачі лінійного програмування

У прикладі 5.1 було поставлено за мету отримати максимальний прибуток від виробництва виробів  $A$  та  $B$ , але нічого не було сказано про вартість ресурсу, який використовувався для створення цих виробів, тобто нічого не було сказано про витрати, які необхідно мінімізувати, але таким чином, щоб не зменшити питомий прибуток (собівартість виробленої продукції).

Припустимо, що одиниця ресурсу  $C$ ,  $D$ ,  $E$  коштує відповідно  $u_1$ ,  $u_2$ ,  $u_3$ . Тоді вартість виробів можливо обчислити за виразами відповідно  $24u_1 + 8u_2 + 3u_3$  і  $8u_1 + 8u_2 + 8u_3$ .

Задача лінійного програмування із змінними  $u_{1,2,3}$  формулюється так: якою повинна бути вартість кожного окремого ресурсу для того, щоб не зменшуючи вартості одиниці виробу, досягти мінімуму сумарної вартості ресурсів. Математична постановка задачі в цьому випадку набуває вигляду:

$$W = 600u_1 + 480u_2 + 240u_3 \rightarrow \min_{u_{1,2,3} \geq 0}$$

за умов

$$\begin{cases} 24u_1 + 8u_2 + 3u_3 \geq 40, \\ 8u_1 + 8u_2 + 8u_3 \geq 50. \end{cases}$$

Записана задача лінійного програмування має назву двоїстої по відношенню до задачі прикладу 5.1.

## 12.3. Загальна постановка і правила побудови двоїстої задачі

Кожній задачі лінійного програмування можливо поставити у відповідність задачу, яку називають двоїстою до неї. Припустимо, що загальна задача лінійного програмування (вихідна, або, як ще кажуть, пряма задача) задана у вигляді:

$$W = \sum_{j=1}^n c_j x_j \rightarrow \max_{x_j \geq 0} \quad (12.1)$$

за умови виконання обмежень

$$\sum_{j=1}^n a_{ij} x_j \leq b_i \quad i = \overline{1, m} . \quad (12.2)$$

Двоїста (спряжена до неї) задача має вигляд:

$$\tilde{W} = \sum_{i=1}^m b_i u_i \rightarrow \min_{u_i \geq 0} \quad j = \overline{1, m} , \quad (12.3)$$

$$\sum_{i=1}^m a_{ij} u_i \geq c_j \quad j = \overline{1, n} . \quad (12.4)$$

Задача (12.3), (12.4) – двоїста до задачі (12.1), (12.2) – будується за наступними правилами:

1. Виконується впорядкування вихідної задачі до виду (12.3), (12.4), тобто якщо цільова функція максимізується, то нерівності-обмеження повинні бути приведені до вигляду « $\leq$ », а якщо мінімізується, то до вигляду « $\geq$ ». Досягти виконання потрібної орієнтації знаку обмежень можливо множенням обох його частин на  $-1$ .

2. Якщо вихідна задача є задачею максимізації, то двоїста буде задачею мінімізації. При цьому вектор, який утворено із коефіцієнтів при невідомих в показнику ефективності вихідної задачі співпадає із вектором констант в правих частинах системи обмежень двоїстої задачі, і, навпаки, коефіцієнти при невідомих в показнику ефективності двоїстої задачі є відповідними правими частинами системи обмежень вихідної задачі.

3. Кожній змінній  $u_i$  двоїстої задачі відповідає  $i$ -те обмеження вихідної задачі і навпаки, кожній змінній  $x_j$  вихідної задачі відповідає  $j$ -те обмеження двоїстої задачі.

4. Матриця коефіцієнтів двоїстої задачі може бути отримана транспонуванням матриці  $A = \|a_{ij}\|_{m \times n}$ , складеної із коефіцієнтів при невідомих в системі обмежень вихідної задачі.

Задачі (12.1), (12.2) та (12.3), (12.4) утворюють симетричну пару взаємно двоїстих задач. Якщо використати позначення:

$$C = c_1, c_2, \dots, c_n ,$$

$$B = b_1, b_2, \dots, b_m \quad T,$$

$$A = \left\| a_{ij} \right\|_{\substack{i=1, \dots, m \\ j=1, \dots, n}} - \text{прямокутна матриця розміром } m \times n,$$

$$X = x_1, x_2, \dots, x_n \quad T,$$

$$U = u_1, u_2, \dots, u_m \quad T,$$

$$Q_k = 0, \dots, 0 \quad T - \text{вектор-стовпець нулів розмірності } k;$$

то пряму і двоїсту задачі лінійного програмування можливо записати у вигляді:

$$\begin{cases} W = C \cdot X \rightarrow \max_{X \geq o_n} \\ A \cdot X \leq B \end{cases} \quad - \text{пряма,}$$

$$\begin{cases} \tilde{W} = B^T \cdot U \rightarrow \min_{U \geq o_m} \\ A^T \cdot U \geq C^T \end{cases} - \text{двоїста.}$$

Розглянемо двоїсту задачу лінійного програмування (12.3), (12.4) як вихідну задачу лінійного програмування і, скориставшись правилами переходу до двоїстої задачі, отримаємо пряму задачу лінійного програмування (10.1), (10.2):

$$\begin{cases} -\tilde{W} = -B^T \cdot U \rightarrow \max, \\ -A^T \cdot U \geq -C^T. \end{cases} \quad U \geq o_m$$

І далі, після перетворення, можемо записати:

$$\begin{cases} -W = -C \cdot X \rightarrow \min, \\ -A \cdot X \geq -B; \end{cases} \quad X \geq o_n \Leftrightarrow \begin{cases} W = C \cdot X \rightarrow \max, \\ A \cdot X \leq B. \end{cases} \quad X \geq o_n$$

### **Висновок:**

1. Зведення задачі лінійного програмування до двоїстої задачі лінійного програмування є прямою задачею.

2. У системі «пряма–двоїста» обидві задачі лінійного програмування рівноправні. Кожну з них можна розглядати як пряму, тоді інша вважається двоїстою до неї.

**Приклад 12.1.** Задана задача лінійного програмування:

$$\begin{cases} W = x_1 - 4x_2 - 3x_3 \rightarrow \min ; \\ 0 \leq x_{1,2,3} \\ 3x_1 + 4x_2 + x_3 \leq 7, \\ x_1 + 2x_2 + x_3 = 6, \\ x_3 \geq 4. \end{cases}$$

Необхідно побудувати двоїсту задачу лінійного програмування.

**Рішення**

□ Приведемо задану задачу лінійного програмування до вигляду (12.1), (12.2).

$$\begin{cases} -W = -x_1 + 4x_2 + 3x_3 \rightarrow \max, \\ x_{1,2,3} \geq 0 \\ 3x_1 + 4x_2 + x_3 \leq 7, \\ x_1 + 2x_2 + x_3 \leq 6, \\ -x_1 - 2x_2 - x_3 \leq -6, \\ -x_3 \leq -4, \end{cases} \Leftrightarrow \begin{cases} -W = -1 & 4 & 3 \cdot x_1 & x_2 & x_3 \xrightarrow{T} \max, \\ x_{1,2,3} \geq 0 \\ \begin{bmatrix} 3 & 4 & 1 \\ 1 & 2 & 1 \\ -1 & -2 & -1 \\ 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \leq \begin{bmatrix} 7 \\ 6 \\ -6 \\ -4 \end{bmatrix} \end{cases}$$

Двоїста ЗЛП набуває вигляду:

$$\tilde{W} = \begin{bmatrix} 7 & 6 & -6 & -4 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \rightarrow \min_{u_{1,2,3,4} \geq 0}$$

$$\begin{bmatrix} 3 & 1 & -1 & 0 \\ 4 & 2 & -2 & 0 \\ 1 & 1 & -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \geq \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} \Leftrightarrow \begin{cases} \tilde{W} = 7u_1 + 6u_2 - 6u_3 - 4u_4 \rightarrow \min_{u_{1,2,3,4} \geq 0}; \\ 3u_1 + u_2 - u_3 \geq -1, \\ 4u_1 + 2u_2 - 2u_3 \geq 4, \\ u_1 + u_2 - u_3 - 4u_4 \geq 3. \end{cases}$$

**Приклад 12.2.** Пряма задача лінійного програмування задана у вигляді:

$$\begin{cases} W = 6x_1 + 10x_2 \rightarrow \max_{x_1 \geq 0}, \\ 5x_1 + 3x_2 \geq 10, \\ x_1 - x_2 \leq 4. \end{cases}$$

Побудувати двоїсту задачу.

**Рішення**

□ Приведемо пряму задачу до стандартного вигляду. Для цього позначимо

$$x = x_2 - x_3 \in \mathbb{R}, \quad x_{2,3} \geq 0$$

Тоді:

$$\begin{cases} W = 6x_1 + 10x_2 - 10x_3 \rightarrow \max_{x_{1,2,3} \geq 0}; \\ -5x_1 - 3x_2 + 3x_3 \leq -10, \\ x_1 - x_2 + x_3 \leq 4, \end{cases} \Leftrightarrow \begin{cases} W = [6 \quad 10 \quad -10] \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \rightarrow \max_{x_{1,2,3} \geq 0}; \\ \begin{bmatrix} -5 & -3 & 3 \\ 1 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \leq \begin{bmatrix} -10 \\ 4 \end{bmatrix}. \end{cases}$$

Двоїста задача набуває вигляду:

$$\begin{cases} W = [-10 \quad 4] \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \rightarrow \min_{u_{1,2} \geq 0}, \\ \begin{bmatrix} -5 & 1 \\ -3 & -1 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \geq \begin{bmatrix} 6 \\ 10 \\ -10 \end{bmatrix}, \end{cases} \Leftrightarrow \begin{cases} \tilde{W} = -10u_1 + 4u_2 \rightarrow \min_{u_{1,2} \geq 0}, \\ -5u_1 + u_2 \geq 6, \\ -3u_1 - u_2 \geq 10, \\ 3u_1 + u_2 \geq -10, \end{cases} \Leftrightarrow \begin{cases} \tilde{W} = -10u_1 + 4u_2 \rightarrow \min_{u_{1,2} \geq 0}, \\ -5u_1 + u_2 \geq 6, \\ -3u_1 - u_2 = 10. \end{cases}$$

**Приклад 12.3.** Побудувати задачу двоїсту до заданої задачі лінійного програмування

$$\begin{cases} W = x_1 + x_2 + x_3 \rightarrow \min, \\ x_{123} \geq 0, \\ x_1 + 2x_2 + x_3 \geq 9, \\ 2x_1 + x_3 \geq 4. \end{cases}$$

**Рішення**

□

$$\begin{cases} W = [111] \cdot [x_1 \ x_2 \ x_3]^T \rightarrow \min, \\ x_{123} \geq 0, \\ \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \geq \begin{bmatrix} 9 \\ 4 \end{bmatrix}. \end{cases}$$

$$\Rightarrow \begin{cases} \tilde{W} = [9 \ 4] \cdot [u_1 \ u_2]^T, \\ \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \leq \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \end{cases} \Leftrightarrow \begin{cases} \tilde{W} = 9u_1 + 4u_2 \rightarrow \max; \\ x_{123} \geq 0, \\ u_1 + 2u_2 \leq 1, \\ 2u_1 \leq 1, \\ u_1 + u_2 \leq 1. \end{cases}$$

■

## 12.4. Основні теореми двоїстості

**Теорема 12.1.** Якщо одна із двоїстих задач має оптимальний розв'язок  $\hat{X} = [\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n]^T$ , то і інша має оптимальний розв'язок  $\hat{U} = [\hat{u}_1, \hat{u}_2, \dots, \hat{u}_m]^T$ , при цьому екстремальні значення показника ефективності обох задач

$$\text{співпадають } \sum_{j=1}^n c_j \hat{x}_j = \sum_{i=1}^m b_i \hat{u}_i.$$

**Зауваження 12.1.** Якщо показник ефективності однієї із задач необмежений, то інша задача не має розв'язків.

**Наслідок 12.1.**

1. Для існування розв'язку однієї з двоїстих задач, необхідно і достатньо, щоб інша мала хоча б один розв'язок;

2. Для того щоб  $X = [x_1, x_2, \dots, x_n]^T$  та  $U = [u_1, u_2, \dots, u_m]^T$  були оптимальними розв'язками пари двоїстих задач, необхідно і достатньо, щоб виконувалась рівність:

$$\sum_{j=1}^n c_j x_j = \sum_{i=1}^m b_i u_i$$

**Теорема 12.2.** Якщо будь яка змінна  $\hat{x}_j$ ,  $j = \overline{1, n}$  оптимального розв'язку прямої задачі додатна, то  $j$ -те обмеження двоїстої задачі на оптимальному розв'язку перетворюється у рівність. Якщо оптимальне рівняння прямої задачі перетворює будь яке  $i$ -те обмеження у строгу нерівність, то в оптимальному розв'язку двоїстої задачі змінна  $\hat{u}_i$  дорівнює 0.

**Зауваження 12.2.** Ця теорема ще носить назву «умова доповняльної нежорсткості» і формально записується у вигляді:

$$\begin{cases} \left( \sum_{i=1}^m a_{ij} \hat{u}_i - c_j \right) \cdot \hat{x}_j = 0, & j = \overline{1, n}, \\ \left( \sum_{j=1}^n a_{ij} \hat{x}_j - b_i \right) \cdot \hat{u}_i = 0, & i = \overline{1, m}. \end{cases}$$

**Теорема 12.3.** В оптимальному розв'язку двоїстої задачі значення змінних чисельно дорівнюють частинним похідним  $\frac{\partial W_{max}}{\partial b_i}$ , для прямої задачі:

$$\hat{u}_i = \frac{\partial W_{max}}{\partial b_i}, \quad i = \overline{1, m},$$

де  $\Delta W_{max}$  – зміна максимального значення показника ефективності  $W$ , при одному і тому самому двоїстому плані, викликано малою зміною  $b_i$  прямої задачі.

► Покажемо, що при змінненні правих частин  $b_i$ ,  $i = \overline{1, m}$  обмежень прямої задачі невідомі двоїстої задачі можливо інтерпретувати як оцінки

впливу цих змінних на оптимальне значення показника ефективності прямої задачі. Позначимо

$$\Delta B = [\Delta b_1, \dots, \Delta b_m]^T,$$

де  $\Delta b_i$  – приріст  $i$ -ої правої частини прямої задачі.

Розглянемо дві двоїсті задачі:

1) Пряма:

$$\begin{cases} W = C \cdot X_1 \rightarrow \max, \\ X_1 \geq 0_n \\ A \cdot X_1 \leq B \end{cases}$$

та двоїста:

$$\begin{cases} \tilde{W} = B^T \cdot U_1 \rightarrow \min, \\ U_1 \geq 0_m \\ A^T \cdot U_1 \geq C^T. \end{cases}$$

2) Пряма:

$$\begin{cases} W = C \cdot X_2 \rightarrow \max, \\ X_2 \geq 0_n \\ A \cdot X_2 \leq B + \Delta B \end{cases}$$

де

$$\Delta B = \begin{bmatrix} \Delta b_1 \\ \vdots \\ \Delta b_m \end{bmatrix} \rightarrow 0 \Leftrightarrow \begin{bmatrix} \Delta b_1 \rightarrow 0 \\ \vdots \\ \Delta b_m \rightarrow 0 \end{bmatrix}$$

та двоїста:

$$\begin{cases} \tilde{W} = (B + \Delta B)^T \cdot U_2 \rightarrow \min, \\ U_2 \geq 0_m \\ A^T \cdot U_2 \geq C^T. \end{cases}$$

Вимога, яка полягає в тому, що  $\Delta B \rightarrow 0$ , означає, що заміна  $B$  на  $B + \Delta B$  у двоїстій задачі не призводить до зміни її оптимального розв'язку  $\hat{U}_2 = \hat{U}_1$ . Тоді, відповідно до першої теореми двоїстості, можна записати для першої пари задач:

$$W_{\max_1} = \hat{W}_{\min_1} \Leftrightarrow C\hat{X}_1 = B^T\hat{U}_1$$

та для другої пари задач:

$$W_{\max_2} = \hat{W}_{\min_2} \Leftrightarrow C\hat{X}_2 = B + \Delta B^T \hat{U}_1.$$

Обчислимо приріст показника ефективності, прямої задачі

$$\begin{aligned} \Delta W_{\max} &= W_{\max_2} - W_{\max_1} = B + \Delta B^T \cdot \hat{U}_1 - B^T \hat{U}_1 = \\ &= B^T + \Delta B^T - B^T \hat{U}_1 = \Delta B^T \hat{U}_1 = [\Delta b_1, \dots, \Delta b_m] \cdot \begin{bmatrix} \hat{u}_1 \\ \vdots \\ \hat{u}_m \end{bmatrix} = \sum_{i=1}^m \Delta b_i \cdot \hat{u}_i. \blacktriangleleft \end{aligned}$$

Із доведеної теореми 12.3 випливає, що розв'язок двоїстої задачі  $u_i$  (двоїста оцінка) кількісно дорівнює приросту максимального значення показника ефективності прямої задачі при зміні правої частини  $i$ -го обмеження прямої задачі на одиницю.

### Контрольні запитання та завдання

1. У чому полягає суть післяоптимізаційного аналізу задачі лінійного програмування?
2. Поясніть необхідність проведення післяоптимізаційного аналізу задачі лінійного програмування.
3. У чому полягає суть двоїстої задачі лінійного програмування?
4. Наведіть та поясніть математичну постановку і правила побудови двоїстої задачі.

## Глава 13. Аналіз розв'язку задачі лінійного програмування

### 13.1. Аналіз розв'язку задачі лінійного програмування на основі двоїстих оцінок

Аналіз розв'язку задачі лінійного програмування

$$W = \sum_{j=1}^n c_j x_j \rightarrow \max_{x_j \geq 0, j=1, n} \quad (11.1)$$

за умови виконання обмежень

$$\sum_{j=1}^n a_{ij} x_j \leq b_i \quad i = \overline{1, m} \quad (11.2)$$

базується на дослідженні впливу варіації параметрів  $c_j$ ,  $a_{ij}$ ,  $b_i$  ( $i = \overline{1, m}$ ,  $j = \overline{1, n}$ ) на цей розв'язок. Такий аналіз можливо виконати із використанням наступних прийомів та способів:

- 1) аналізу внутрішньої структури кожного із отриманих розв'язків із використанням властивостей двоїстих оцінок;
- 2) співставлення варіантів розв'язків, отриманих шляхом розв'язання прямої задачі, при різних значеннях її параметрів.

Перший напрямок використовує теореми двоїстості, другий напрямок пов'язаний із так званою *параметричною оптимізацією* (*параметричне програмування*), яка базується, здебільшого, на комп'ютерному обчислювальному експерименті.

#### ***Аналіз розв'язків задачі лінійного програмування на основі двоїстих оцінок***

Двоїсті оцінки слід розглядати як:

1. *Показники дефіцитності(коштовності) ресурсів і продукції.*

Як впливає із теореми 12.2.  $\hat{u}_i$  є оцінкою  $i$ -го ресурсу. Чим більше значення оцінки, тим вище дефіцитність ресурсу. Для недефіцитного ресурсу  $\hat{u}_i = 0$ .

2. *Показник впливу обмежень на значення показника ефективності.*

В теоремі 12.3 було показано, що  $\hat{u}_i = \frac{\partial W_{\max}}{\partial b_i}$ ,  $i = \overline{1, m}$ . При

незначному прирості  $\Delta b_i$ , оцінка  $\hat{u}_i$  є точною мірою впливу обмеження на показник ефективності. Для практики важливо знайти граничні значення правих частин системи обмежень-нерівностей прямої задачі (нижньої та верхньої границь), при яких величини оцінок  $\hat{u}_i$ ,  $i = \overline{1, m}$  залишаються незмінними.

### 3. Показники доцільності виробництва окремих видів продукції.

Ця властивість випливає із теореми 12.2. Головний зміст цієї властивості полягає в тому, що в оптимальний план виробництва доцільно включати лише виробництво тієї  $j$ -ої продукції, для якої витрати на придбання і використання сировини не перевищують прибуток:

$$\sum_{i=1}^m a_{ij} \hat{u}_i < c_j \quad j = \overline{1, n}$$

### 4. Інструмент співставлення сумарних умовних витрат та результатів.

Ця властивість випливає із теореми 12.1, в якій встановлено зв'язок між екстремальними значеннями показника ефективності прямої та двоїстої задач.

Перейдемо до детального розгляду чотирьох записаних вище властивостей. З економічної інтерпретації двоїстих задач випливає, що рівність показника ефективності при оптимальних планах означає, що оцінка всіх затрат виробництва повинна дорівнювати оцінці вартості продукту, який виготовлено.

**Означення 13.1.** Допустимим інтервалом стійкості двоїстих оцінок по відношенню до  $i$ -го обмеження називають інтервал вигляду

$$\left[ b_i - \Delta b_i^H, b_i + \Delta b_i^B \right], \quad i = \overline{1, m},$$

де  $\Delta b_i^H$  – нижня границя зменшення;  $\Delta b_i^B$  – верхня границя збільшення.

Якщо

$$b_i + \Delta b_i \in \left[ b_i - \Delta b_i^H, b_i + \Delta b_i^B \right],$$

де  $\Delta b_i$  – зміна  $i$ -го ресурсу, то вплив  $\Delta b_i$  на змінення величини прибутку, обчисленого в результаті оптимізації прямої задачі лінійного програмування, можливо оцінити за виразом:

$$\Delta \hat{W}_i = \hat{u}_i \cdot \Delta b_i \quad i = \overline{1, m} . \quad (11.3)$$

При цьому сумарний вплив змінення усіх  $\Delta b_i$  обчислюється як сума окремих впливів.

Обчислити  $\Delta b_i^H$  та  $\Delta b_i^B$  можливо за формулами:

$$\Delta b_i^H = \min_{j=\overline{1, m_1}} \left\{ \frac{\hat{x}_j}{d_{ij}} \right\}, \quad d_{ij} > 0 \quad (11.4)$$

та

$$\Delta b_i^B = \left| \max_{j=\overline{1, m_1}} \left\{ \frac{\hat{x}_j}{d_{ij}} \right\} \right|, \quad d_{ij} < 0, \quad (11.5)$$

де  $\|d_{ij}\| = \|a_{ij}\|^{-1}$ ,  $i = \overline{1, m_1}$ ,  $j = \overline{1, m_1}$ ,  $m_1$  – кількість базисних змінних,  $a_{ij}$  – елемент матриці  $A$ .

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1m_1} & | & a_{1m_1+1} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ a_{m_1 1} & \cdots & a_{m_1 m_1} & | & a_{m_1 m_1+1} & \cdots & a_{m_1 n} \end{bmatrix} = \begin{bmatrix} | & a_{1m_1+1} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ | & a_{m_1 m_1+1} & \cdots & a_{m_1 n} \end{bmatrix},$$

де  $A_1 = \begin{bmatrix} a_{11} & \cdots & a_{1m_1} \\ \vdots & \ddots & \vdots \\ a_{m_1 1} & \cdots & a_{m_1 m_1} \end{bmatrix}$  – квадратна матриця, що складається із

коефіцієнтів при базисних змінних ( $m_1 \leq m$ , де  $m_1$  – кількість лінійно незалежних обмежень в прямій задачі лінійного програмування)

Доцільність включення до плану випуску нових видів продукції оцінюється параметром:

$$\Delta_j = \sum_{i=1}^m a_{ij} \hat{u}_i - c_j \quad j = \overline{1, n} . \quad (11.6)$$

Якщо  $\Delta_j < 0$ , то даний вид продукції після його введення до плану покращує план.

Якщо  $\Delta_j > 0$ , то включення продукту до плану випуску недоцільно.

Зрозуміло, що приріст прибутку за рахунок збільшення деякого ресурсу на величину  $\Delta b_i$  дорівнює

$$\Delta \hat{W}_i = \Delta b_i \cdot \hat{u}_i \quad (11.7)$$

і при цьому витрати на придбання ресурсу складають

$$\Delta V_i = \Delta b_i \cdot v_i, \quad (11.8)$$

де  $v_i$  – вартість одиниці  $i$ -го ресурсу.

**Висновок.** : Захід із придбанням ресурсу слід вважати доцільним, якщо  $\Delta \hat{W}_i - \Delta V_i > 0$ .

**Приклад 13.1.** Для виготовлення чотирьох видів продукції А, Б, В, Г використовуються три види ресурсів  $R_1, R_2, R_3$ . Наявність ресурсу, норми її витрат на виготовлення одиниці продукції, питомий прибуток наведені в таблиці 13.1.

Таблиця 13.1

Вихідні дані прямої задачі лінійного програмування

Вид ресурсу	Наявність ресурсу	Норми витрат на виготовлення одиниці продукції			
		А	Б	В	Г
$R_1$	240	2	1	1	3
$R_2$	60	1	0	2	1
$R_3$	300	1	2	1	0
Питомий прибуток, $C_j$	-	4	2	3	5

Завдання полягає в тому, щоб розв'язати пряму задачу лінійного програмування і виконати післяоптимізаційний аналіз.

Основні етапи виконання завдання полягають у наступному:

1. Знайти оптимальний розв'язок прямої та двоїстої задач  $\hat{X}, \hat{U}$ .

2. Визначити зміни максимального прибутку при зміні ресурсу:

$$R_1 \text{ на } \Delta b_1 = -10,$$

$$R_2 \text{ на } \Delta b_2 = 60,$$

$$R_3 \text{ на } \Delta b_3 = 30.$$

3. Оцінити доцільність введення у план випуску п'ятого виду продукції Д, якщо норми витрат ресурсів  $R_1, R_2, R_3$  на одиницю випуску складатимуть відповідно 2, 4, 2, а питомий прибуток  $-C_5 = 15$ .

4. Оцінити доцільність придбання 100 одиниць ресурсу  $R_3$  за ціною  $V_3 = 0,5$ .

### **Рішення**

□ Розв'язати пряму та двоїсту задачу лінійного програмування і виконати післяоптимізаційний аналіз.

1. Постановка та розв'язок прямої та двоїстої задач лінійного програмування:

$$\begin{cases} W = 4x_1 + 2x_2 + 3x_3 + 5x_4 = [4 \ 2 \ 3 \ 5] \cdot X \rightarrow \max_{x_{1,2,3,4} \geq 0}, \\ \begin{bmatrix} 2 & 1 & 1 & 3 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 1 & 0 \end{bmatrix} \cdot X \leq \begin{bmatrix} 240 \\ 60 \\ 300 \end{bmatrix}, \end{cases}$$

де  $X = [x_1, x_2, x_3, x_4]^T$ .

2. Двоїста задача лінійного програмування

$$\begin{cases} \tilde{W} = [240 \ 60 \ 300] \cdot U \rightarrow \min_{u_{1,2,3} \geq 0}, \\ \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 1 \\ 3 & 1 & 0 \end{bmatrix} \cdot U \geq \begin{bmatrix} 4 \\ 2 \\ 3 \\ 5 \end{bmatrix}, \end{cases}$$

де  $U = [u_1, u_2, u_3]^T$ .

3. Результат обчислення оптимального рішення за допомогою функцій Matlab набуває вигляду:

- Прямої задачі лінійного програмування

$$\hat{X} = [\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4]^T = [60 \quad 120 \quad 0 \quad 0]^T,$$

$$\hat{W} = 480.$$

- Двоїстої задачі лінійного програмування

$$\hat{U} = [\hat{u}_1, \hat{u}_2, \hat{u}_3]^T = [1,6 \quad 0,6 \quad 0,2]^T,$$

$$\hat{W} = 480.$$

4. Аналіз результатів розв'язку прямої задачі лінійного програмування:

а) найбільш дефіцитним є ресурс  $R_1$ ,  $\hat{u}_1 = 1,6$ , найменш дефіцитним є ресурс  $R_3$ ,  $\hat{u}_3 = 0,2$ ;

б) обчислимо зміну максимального прибутку при зміні ресурсу.

5. Знайдемо інтервали стійкості двоїстих оцінок, тобто межі, в яких вони точно вимірюють вплив зміни обмежень на показник ефективності прямої задачі лінійного програмування.

5.1. Визначення інтервалу стійкості оцінки по відношенню до обмеження  $b_1$ , пов'язаного з першим ресурсом.

Базисними невідомими, що входять до оптимального плану є  $x_1, x_2, x_3$ , тому матриця  $A_1$  набуває вигляду:

$$A_1 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 1 \end{bmatrix} = \|a_{ij}\|_{3 \times 3}.$$

Тоді  $A_1^{-1} = \|a_{ij}\|_{3 \times 3}^{-1} = \|d_{ij}\|_{3 \times 3}$ :

$$A_1^{-1} = \begin{bmatrix} 0,8 & -0,2 & -0,4 \\ -0,2 & -0,2 & 0,6 \\ -0,4 & 0,6 & 0,2 \end{bmatrix}.$$

Знаходимо

$$\Delta b_1^H = \min_{j=1,2,3} \left\{ \frac{60}{0,8} \right\} = 75, \quad d_{1j} > 0;$$

$$\Delta b_1^B = \left| \max_{j=1,2,3} \left\{ \frac{120}{-0,2}; \frac{0}{-0,4} \right\} \right| = 0, \quad d_{1j} < 0.$$

Остаточню інтервал стійкості оцінок по відношенню до першого обмеження набуває вигляду:

$$\left[ b_1 - \Delta b_1^H, b_1 + \Delta b_1^B \right] = \left[ 240 - 75, 240 + 0 \right] = \left[ 165, 240 \right].$$

5.2. Визначаємо інтервал стійкості по відношенню до ресурсу 2:

$$\Delta b_2^H = \min_{j=1,2,3} \left\{ \frac{0}{0,6} \right\} = 0, \quad d_{2j} > 0;$$

$$\Delta b_2^B = \left| \max_{j=1,2,3} \left\{ \frac{60}{-0,2}; \frac{120}{-0,2} \right\} \right| = 300, \quad d_{2j} < 0.$$

Остаточню:

$$\left[ b_2 - \Delta b_2^H, b_2 + \Delta b_2^B \right] = \left[ 60 - 0, 60 + 300 \right] = \left[ 60, 360 \right].$$

5.3. Визначаємо інтервал стійкості по відношенню до ресурсу 3:

$$\Delta b_3^H = \min_{j=1,2,3} \left\{ \frac{120}{0,6}; \frac{0}{0,2} \right\} = 0, \quad d_{3j} > 0;$$

$$\Delta b_3^B = \left| \max_{j=1,2,3} \left\{ \frac{60}{-0,4} \right\} \right| = 150, \quad d_{3j} < 0.$$

Остаточню:

$$\left[ b_3 - \Delta b_3^H, b_3 + \Delta b_3^B \right] = \left[ 300 - 0, 300 + 150 \right] = \left[ 300; 450 \right].$$

Перевіримо, чи перебуває ресурс в межах стійкості після його зміни згідно п.2. завдання.

$$R_1: b_1 = 240 - 10 = 230 \in [165, 240],$$

$$R_2: b_2 = 60 + 60 = 120 \in [60, 360],$$

$$R_3: b_3 = 300 + 30 = 330 \in [300, 450].$$

**Висновок 1.** Після зміни весь ресурс перебуває в межах стійкості.

6. Вплив зміни окремого ресурсу на максимальний прибуток, тобто  $\Delta \hat{W}_i$   $i = 1, 2, 3$  можливо обчислити за формулою

$$\Delta \hat{W}_i = \hat{u}_i \cdot \Delta b_i \quad i = 1, 2, 3$$

$$\Delta \hat{W}_1 = \hat{u}_1 \cdot \Delta b_1 = 1,6 \cdot -10 = -16,$$

$$\Delta \hat{W}_2 = \hat{u}_2 \cdot \Delta b_2 = 0,6 \cdot 60 = 36,$$

$$\Delta \hat{W}_3 = \hat{u}_3 \cdot \Delta b_3 = 0,2 \cdot 30 = 6.$$

Сумарний вплив знаходимо як суму окремих впливів:

$$\Delta \hat{W} = -16 + 36 + 6 = 26.$$

7. Оцінимо доцільність введення у план випуску п'ятого виду продукції Д. Для цього обчислимо характеристику  $\Delta_5$ :

$$\Delta_5 = \sum_{i=1}^3 a_{i5} \hat{u}_i - C_5 = 2 \cdot 1,6 + 4 \cdot 0,6 + 2 \cdot 0,2 - 15 = -9 < 0$$

**Висновок 2.** Прибуток перевищує витрати, тому введення у план випуску п'ятого виду продукції є вигідним.

8. Приріст ресурсу  $R_3$  на величину  $\Delta b_3 = 100$

$$b_3 = 300 + 100 = 400 \in [300, 450]$$

не виводить цей ресурс за межі стійкості, тому

$$\Delta \hat{W}_3 = \hat{u}_3 \cdot \Delta b_3 = 0,2 \cdot 100 = 20,$$

але витрати на придбання 100 одиниць ресурсу  $R_3$  складатимуть:

$$\Delta V_3 = \Delta b_3 \cdot V_3 = 100 \cdot 0,5 = 50 > 20,$$

тобто величина додаткового прибутку менше за витрати.

**Висновок 3.** Закупка ресурсу  $R_3$  на умовах п.4 недоцільна. ■

### 13.2. Аналіз розв'язку задачі лінійного програмування на основі використання параметричного програмування

Дослідження змінення розв'язку задачі лінійного програмування залежно від змінення параметрів математичної моделі задачі лінійного програмування є предметом *параметричного програмування* в задачі

лінійного програмування. Постановка загальної задачі параметричного програмування в задачі лінійного програмування формулюється таким чином: знайти на відрізку  $[\alpha; \beta]$  скінчену кількість інтервалів  $I_q \subset [\alpha; \beta]$ ,  $q = \overline{1, Q}$ ,  $Q \in \mathbb{N}$ , що містять значення параметра  $t \in I_q$ ,  $q = \overline{1, Q}$ , для яких оптимальний розв'язок задачі лінійного програмування виду

$$W = \sum_{j=1}^n c_j' + c_j'' \cdot t \cdot x_j \rightarrow \max_{x_j \geq 0, j=1, n} \quad (13.9)$$

за умови виконання обмежень

$$\sum_{j=1}^n a_{ij}' + a_{ij}'' \cdot t \cdot x_j \leq b_i' + b_i'' \cdot t, \quad i = \overline{1, m} \quad (13.10)$$

досягається в одній і тій самій кутовій точці (вершині) області допустимих розв'язків.

Економічний зміст задачі параметричного програмування в задачі лінійного програмування полягає в тому, що з плином часу, який можливо інтерпретувати як параметр  $t$ , змінюються як вартість  $c_j$   $t$  одиниці  $j$ -того виробленого товару (вартість на момент виготовлення плюс вартість, пов'язана із строком збереження), так і запаси  $b_i$   $t$   $i$ -того ресурсу (відбувається вихід із ладу за рахунок старіння та перебування довгий час у неробочому стані) та витрат ресурсу  $a_{ij}$   $t$  на виготовлення одиниці товару (наприклад, пов'язані з модернізацією виробництва).

Дамо геометричну інтерпретацію задачі параметричного програмування в задачі лінійного програмування. Розглянемо залежність від часу  $t$  лише коефіцієнтів показника ефективності. Задача параметричного програмування в задачі лінійного програмування набуває вигляду: для кожного  $t \in [\alpha; \beta]$  знайти свій вектор змінних, який максимізує показник ефективності:

$$W_t = \sum_{j=1}^n c_j' + c_j'' \cdot t \cdot x_j = C \cdot t \cdot X \rightarrow \max_{x_j \geq 0, j=1, n} \quad (13.11)$$

при виконанні умов:

$$\sum_{j=1}^n a_{ij} x_j \leq b'_i, \quad i = \overline{1, m}, \quad (13.12)$$

де  $C t = [c'_1 + c''_1 t, \dots, c'_n + c''_n t]$ ,  $X = [x_1, \dots, x_n]^T$ .

Припустимо, що область допустимих розв'язків є опуклим багатогранником. Рівнянню  $C t \cdot X = 0$  відповідає сім'я гіперплощин, що проходять через початок координат.

Якщо параметру  $t$  надати деяке значення  $t = t_1 \in [\alpha; \beta]$ , то гіперплощина займе фіксоване положення. Переміщення гіперплощини  $C t_1 \cdot X = W_{t_1}$  паралельно фіксованому положенню у бік збільшення  $W_{t_1}$  призведе до попадання цієї площини, наприклад, у вершину  $A$  області допустимих розв'язків (рис. 13.1).

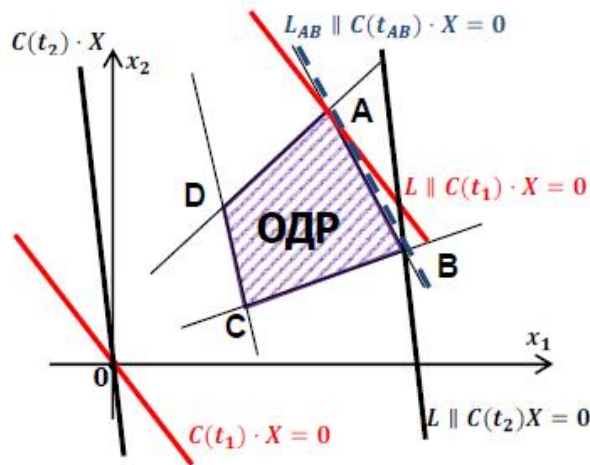


Рис. 13.1. Геометрична інтерпретація задачі параметричного програмування в задачі лінійного програмування

Змінення  $t$  в межах  $t \in t_{AB}, t_{AD}$  залишає незмінним оптимальний розв'язок  $\hat{X} = X_A$ .

Якщо

$$\begin{cases} t \in -\infty, t_{AB} \cap [\alpha; \beta] \\ t_{AB} > \alpha \end{cases}$$

тобто  $t \in [\alpha, t_{AB}]$ , то  $\hat{X} = X_B$  і т.д.

При  $t = t_{AB}$  задача лінійного програмування стає виродженою. Остаточний результат розв'язання задачі параметричного програмування в задачі лінійного програмування призведе до розбиття відрізка  $[\alpha; \beta]$  на кінцеве число інтервалів, в яких зміна  $t$  не призводить до зміни оптимального розв'язку, але в послідовно розташованих інтервалах ці розв'язки будуть різними. Методика розв'язання задачі параметричного програмування в задачі лінійного програмування із використанням стандартних процедур лінійного програмування складається з 2 етапів:

***I етап – обчислення оптимального розв'язку задачі лінійного програмування.***

Параметру  $t \in [\alpha; \beta]$  надається фіксоване значення. Задача параметричного програмування перетворюється в задачу лінійного програмування із фіксованими параметрами, яку розв'язують, наприклад, симплекс-методом.

***II етап – обчислення інтервалу сталості оптимального розв'язку.***

На цьому етапі знаходять інтервал  $t_H, t_B \subset [\alpha; \beta]$ , для якого оптимальний розв'язок, отриманий за допомогою процедур етапу 1 не змінюється. Знайдений інтервал виключають із  $[\alpha; \beta]$ , задають нове фіксоване значення для інтервалу і повторюють обчислення у відповідності з першим етапом. Розв'язок припиняється лише тоді, коли буде виконано розбиття всього відрізка  $[\alpha; \beta]$  на інтервали сталості оптимального розв'язку.

***Приклад 13.2.*** Визначити інтервали зміни параметра  $t \in [0, 8]$  і знайти значення змінних  $x_1, x_2$ , при яких лінійна функція

$$W_t = 4x_1 + 2 + t x_2$$

досягає максимуму в одній і тій самій вершині області допустимих розв'язків при виконанні системи обмежень:

$$\begin{cases} 2x_1 - 5x_2 \leq 10, \\ x_1 + x_2 \geq 5, \\ -x_1 + x_2 \leq 4, \\ 4x_1 + 5x_2 \leq 40, \quad x_{1,2} \geq 0. \end{cases}$$

### Ришення

□ 1. Побудова області допустимих розв'язків (ОДР) (рис.13.2):

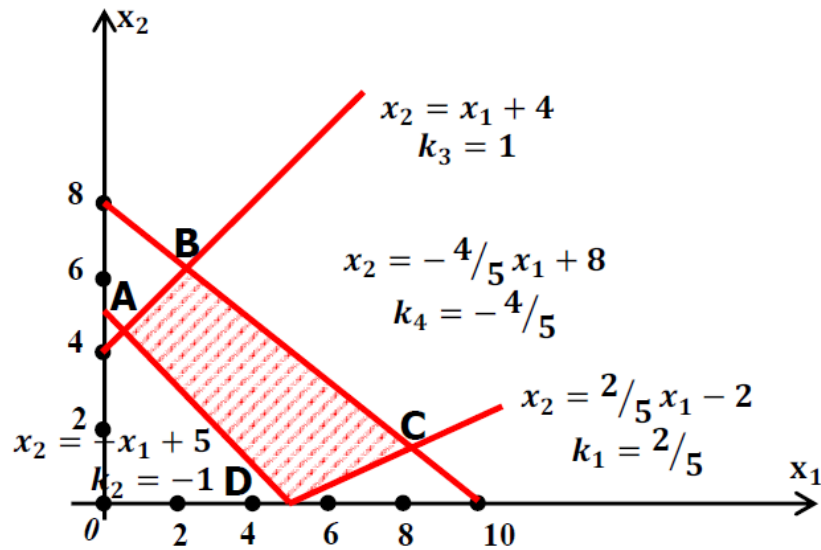


Рис. 13.2. Область допустимих розв'язків задачі 13.2

2. Побудова головної прямої (рис.13.3):

$$\begin{cases} W_t = 4x_1 + 2 + t \cdot x_2 \\ t \in [0,8] \end{cases} \Leftrightarrow \begin{cases} x_2 = \left(-\frac{4}{t+2}\right)x_1 + \left(\frac{W_t}{t+2}\right) = k_t \cdot x_1 + b_t \\ t \in [0,8] \end{cases}$$

$$k_t = -\frac{4}{t+2}, \quad b_t = \frac{W_t}{t+2}.$$

Якщо  $t \in [0,3]$ , то  $k_t \in \left[-2, -\frac{4}{5}\right]$ ,  $\hat{X} = X_C = \begin{bmatrix} x_{1C} \\ x_{2C} \end{bmatrix}$ .

Знайдемо координати точки  $C$ :

$$\begin{cases} x_2 = -\frac{4}{5}x_1 + 8 \\ x_2 = \frac{2}{5}x_1 - 2 \end{cases} \Leftrightarrow \begin{cases} x_2 = \frac{25}{3} \\ x_2 = \frac{4}{3} \end{cases} \Leftrightarrow \hat{X}_C = \begin{bmatrix} 25/3 \\ 4/3 \end{bmatrix}$$



## Рішення

□

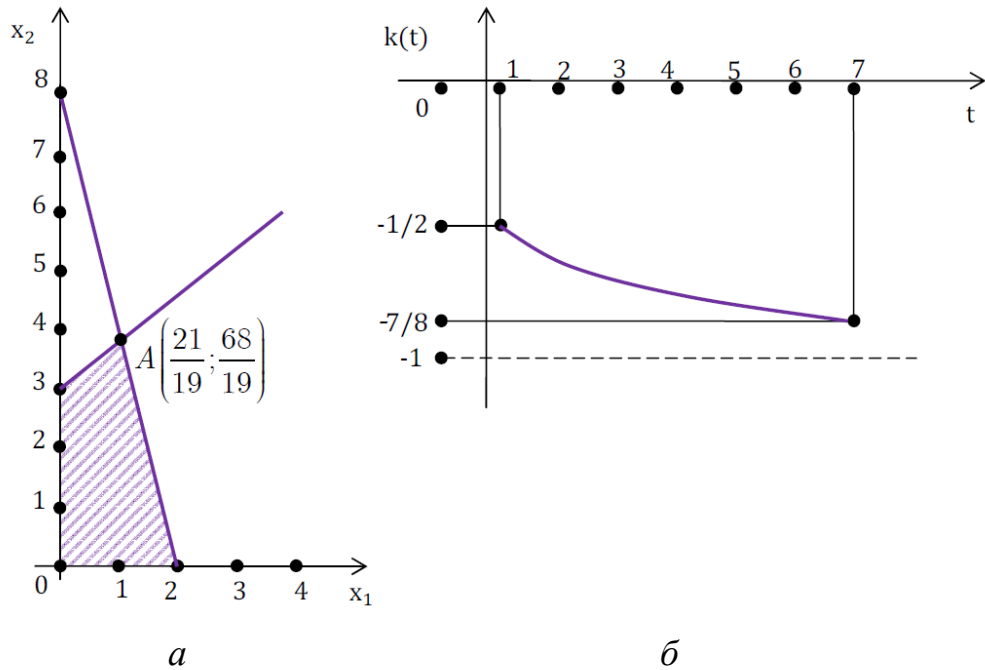


Рис. 13.4. Розв'язання задачі параметричного програмування:  
 а–область допустимих розв'язків задачі; б– графік залежності зміни кутового коефіцієнта  
 головної прямої від зміни параметра  $t$

$$\text{Відповідь: } t \in [1; 7] \rightarrow \hat{X}_A = \left[ \frac{21}{19}; \frac{68}{19} \right]^T . \blacksquare$$

### Контрольні запитання та завдання

1. У чому полягає суть задач лінійного програмування?
2. Які прийоми та способи використовують для аналізу розв'язку задач лінійного програмування?
3. Поясніть процедуру проведення аналізу розв'язку задач лінійного програмування на основі двоїстих оцінок.
4. У чому полягає суть параметричної оптимізації?
5. Наведіть та поясніть двоїсті оцінки, які використовують для аналізу розв'язків задач лінійного програмування.
6. Надайте визначення допустимого інтервалу стійкості двоїстих оцінок.
7. Наведіть та поясніть основні етапи розв'язку прямої задачі лінійного програмування і виконання післяоптимізаційного аналізу.

## Глава 14. Цілочислові задачі лінійної оптимізації

### 14.1. Постановка цілочислових задач лінійної оптимізації

Під час розв'язання більшості задач автоматизації необхідно, щоб величини (кількість каналів зв'язку, окремих агрегатів обладнання, обслуговуючого персоналу, корегувальних пристроїв тощо) обчислювались у цілих невід'ємних числах.

**Означення 14.1.** *Задачі цілочислового програмування* — різновид задач, які припускають, що шукані значення повинні бути цілими числами.

Задачі цілочислової оптимізації можуть бути як лінійними, так і нелінійними. Розглянемо цілочислові задачі лінійної оптимізації (лінійного програмування).

У цих задачах і показник ефективності, і обмеження є лінійними функціями своїх аргументів, але частина або всі ці аргументи можуть бути лише цілими та невід'ємними. Якщо вимога невід'ємної цілочисловості стосується лише частинних змінних, то така задача називається частковою цілочисловою.

Математична постановка цілочисельних задач лінійного програмування має вигляд:

$$\left\{ \begin{array}{l} W = \sum_{j=1}^n c_j x_j \rightarrow \max (\min) \\ \qquad \qquad \qquad x_j \in \mathbb{N}_0 \quad j=1, n \\ \sum_{j=1}^n a_{ij} x_j \leq = \geq b_i, \quad i = \overline{1, m} \\ \mathbb{N}_0 = 0, 1, 2, 3, \dots \quad \text{— розширений ряд} \\ \qquad \qquad \qquad \qquad \qquad \qquad \text{натуральних чисел} \end{array} \right. \quad (12.1)$$

### 14.2. Метод Гоморі розв'язання цілочислових задач лінійного програмування

Метод Гоморі будується на застосуванні симплекс-методу і методу відтинання. Головна ідея методу Гоморі полягає в такому.

На **етапі 1** знаходять оптимальний розв'язок задачі, яка була отримана з вихідної цілочислової задачі лінійного програмування, за умови, що вимога цілочисловості не враховувалась. Якщо отриманий розв'язок є цілочисловим,

то мета цілочислової задачі лінійного програмування досягнута і обчислення припиняється. Якщо розв'язок задачі лінійного програмування не є цілочисловим, то його не вважають розв'язком задачі лінійного програмування і відбувається перехід до етапу 2.

На **етапі 2** до умов задачі додаються обмеження, які відтинають від ОДР отриманий нецілочисловий розв'язок, але не відтинають жодного цілочислового розв'язку, що належить ОДР. Після цього відбувається повернення до етапу 1 і знову виконується розв'язання задачі лінійного програмування симплекс-методом, але з урахуванням нового обмеження. Якщо новий розв'язок теж не є цілочисловим, то відбувається перехід до етапу 2 і до складу обмеження включається ще одне додаткове обмеження та виконується повернення до етапу 1.

Вихід з двохетапної процедури відбувається тоді, коли буде знайдено цілочисловий розв'язок, що належить ОДР.

Для з'ясування специфічних особливостей методу Гоморі розглянемо його геометричне вирішення (рис.14.1).

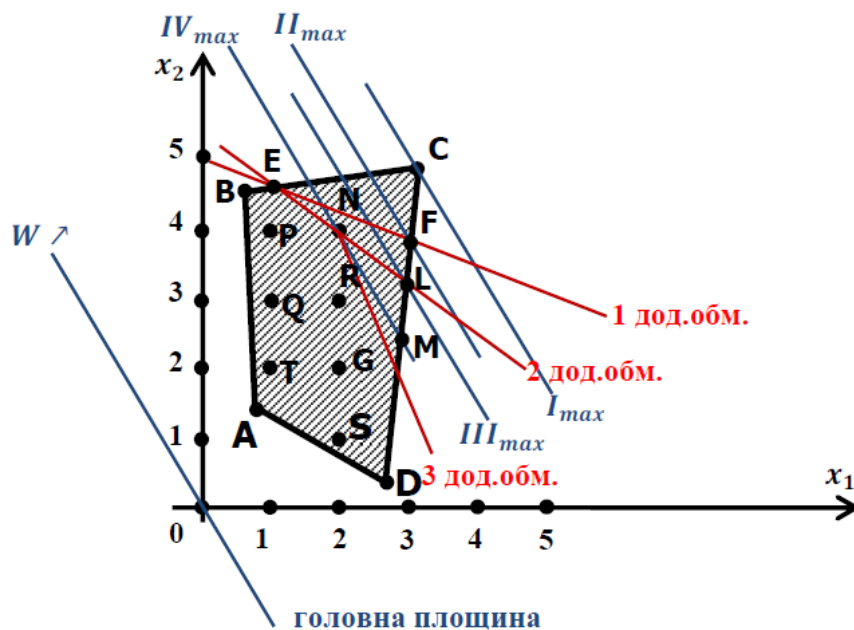


Рис. 14.1. Геометричне вирішення ідеї методу Гоморі

**Означення 14.2.** *Послабленою цілочисловою задачею лінійного програмування* називають задачу, у якій знято обмеження на цілочисловість.

Для послабленої цілочисельної задачі лінійного програмування (або просто задачі лінійного програмування) максимум досягається в точці  $C$ , але

$x_2$ , не є цілим. Точка  $C$  (дробовий розв'язок) відсікається додатковим обмеженням  $EF$  (перше додаткове обмеження), зберігаючи область допустимих рішень цілочисельної задачі лінійного програмування (точки  $P, N, Q, R, T, G, S$ ).

Будуємо додаткове обмеження 1 та переходимо до послабленої задачі що досягає максимуму в точці  $F$ , і при цьому оптимальне значення  $x_1$  буде дробовим. Будуємо додаткове обмеження 2 ( $ENL$ ) та переходимо до максимуму показника ефективності задачі лінійного програмування в точку  $L$ , в якій знову оптимальні  $x_1, x_2$  – дробові. Будуємо додаткове обмеження 3, яке в послабленій задачі лінійного програмування дає кутову точку  $N$ , в якій оптимальні  $x_1, x_2$  є цілочисельними і додатними. Вважаємо, що цілочисельна задача лінійного програмування розв'язана.

Як бачимо, додаткові обмеження кожного разу відтинали від області допустимих розв'язків послабленої цілочисельної задачі лінійного програмування (тобто області допустимих розв'язків задачі лінійного програмування) дробові розв'язки і зберігали область допустимих розв'язків цілочисельної задачі лінійного програмування.

Критерієм побудови "правильного" додаткового обмеження є те, що воно повинно відтинати оптимальну нецілочисельну точку і залишати незайнятими усі цілочисельні точки вихідної цілочисельної задачі лінійного програмування.

**Приклад 14.1.** Розв'язати цілочисельну задачу лінійного програмування методом Гоморі:

$$\begin{cases} W = 2x_1 + 4x_2 + 3x_3 \rightarrow \max_{x_{1,2,3} \in \mathbb{N}_0}; \\ 2x_1 + 3x_2 + x_3 \leq 20, \\ 9x_1 + 7x_2 + 10x_3 \leq 40. \end{cases}$$

### **Ришення**

□ **Етап 1.** Сформулюємо послаблену цілочислову задачу лінійного програмування у вигляді основної задачі лінійного програмування:

$$\begin{cases} -W = -2x_1 - 4x_2 - 3x_3 + 0x_4 + 0x_5 \rightarrow \min_{x_{1,2,3,4,5} \geq 0} \\ 2x_1 + 3x_2 + x_3 + x_4 + 0x_5 = 20, \\ 9x_1 + 7x_2 + 10x_3 + 0x_4 + x_5 = 40. \end{cases}$$

За допомогою функції **linprog** системи комп'ютерної математики MATLAB знайдемо оптимальний розв'язок:

```
>> f=[-2;-4;-3;0;0]
f = -2
    -4
    -3
     0
     0
>> A=[2 3 1 1 0
      9 7 10 0 1]
A = 2 3 1 1 0
    9 7 10 0 1
>> b=[20;40]
b = 20
    40
>> lb=[0;0;0;0;0]
lb = 0
     0
     0
     0
     0
>> [X,W]=linprog(f,[],[],A,b,lb,[])
Optimization terminated successfully.
X = 0.0000
    5.7143
    0.0000
    2.8571
    0.0000
W = -22.8571
```

**Висновок.** Серед базисних змінних  $x_1, x_2, x_3 \in$  дробовий розв'язок  $x_2$ .

Тому це рішення не можна вважати оптимальним розв'язком вихідної цілочисельної задачі лінійного програмування.

**Етап 2.** Виконаємо побудову першого додаткового обмеження:

$$\begin{aligned} x_2 &= \frac{40}{7} - \left( \frac{9}{7}x_1 + \frac{10}{7}x_3 + \frac{1}{7}x_5 \right) \Leftrightarrow \\ \Leftrightarrow x_2 &= \left[ \frac{40}{7} \right] + \left\{ \frac{40}{7} \right\} - \left( \left( \left[ \frac{9}{7} \right] + \left\{ \frac{9}{7} \right\} \right) x_1 + \left( \left[ \frac{10}{7} \right] + \left\{ \frac{10}{7} \right\} \right) x_3 + \left( \left[ \frac{1}{7} \right] + \left\{ \frac{1}{7} \right\} \right) x_5 \right) \Leftrightarrow \\ \Leftrightarrow - \left\{ \frac{40}{7} \right\} + \left\{ \frac{9}{7} \right\} x_1 + \left\{ \frac{10}{7} \right\} x_3 + \left\{ \frac{1}{7} \right\} x_5 &= \left[ \frac{40}{7} \right] - \left( \left[ \frac{9}{7} \right] x_1 + \left[ \frac{10}{7} \right] x_3 + \left[ \frac{1}{7} \right] x_5 \right) - x_2 \end{aligned}$$

Якщо  $x_{1,2,3}$  та  $x_{4,5}$  цілі, то і права частина записаного вище рівняння – ціле число. Тоді і ліва частина останнього рівняння теж буде цілим числом:

$$-\frac{5}{7} + \frac{2}{7}x_1 + \frac{3}{7}x_3 + \frac{1}{7}x_5 = x_6, \text{ яке задовольняє нерівності } x_6 \geq -\frac{5}{7}, x_6 - \text{ ціле}$$

невід'ємне число,  $x_6 \in 0, 1, 2, \dots$ . Нагадаємо, що  $x_{1,2,3,4,5}$  невід'ємні.

Записане означає, що додаткове обмеження у вигляді нової змінної  $x_6$ , не обмежено нічим, окрім загальної вимоги до невід'ємності та цілочисельності. З цього випливає, що нове обмеження може вплинути на дробовий розв'язок задачі лінійного програмування, залишивши цілочисельні незмінними.

Сформулюємо на прикладі  $x_6$  правило побудови додаткового обмеження:

$$\left( \frac{9}{7} - \left[ \frac{9}{7} \right] \right) x_1 + \left( \frac{10}{7} - \left[ \frac{10}{7} \right] \right) x_3 + 0x_4 + \left( \frac{1}{7} - \left[ \frac{1}{7} \right] \right) x_5 - x_6 = \frac{40}{7} - \left[ \frac{40}{7} \right].$$

Послаблену цілочислову задачу лінійного програмування із додатковим обмеженням розв'язуємо, використовуючи функції **linprog**:

```
>> f=[-2;-4;-3;0;0;0]
f = -2
    -4
    -3
     0
     0
```

```

0
>> A=[2 3 1 1 0 0
      9 7 10 0 1 0
      2/7 0 3/7 0 1/7 -1]
A = 2.0000  3.0000  1.0000  1.0000  0  0
     9.0000  7.0000 10.0000  0  1.0000  0
     0.2857  0  0.4286  0  0.1429 -1.0000
>> b=[20;40;5/7]
b = 20.0000
     40.0000
     0.7143
>> lb=[0;0;0;0;0;0]
lb = 0
     0
     0
     0
     0
     0
>> [X,W]=linprog(f,[],[],A,b,lb,[])
Optimization terminated successfully.
X =
  0.0000
  5.0000
  0.0000
  5.0000
  5.0000
  0.0000
W = -20.0000.

```

**Висновок:**  $x_1 = 0$ ,  $x_2 = 5$ ,  $x_3 = 0$ ,  $W = 20$  – оптимальний цілочисловий розв’язок цілочислової задачі лінійного програмування. ■

### 14.3. Метод гілок і границь розв'язання цілочислової задачі лінійного програмування

Ідея методу та прийоми його застосування полягають в тому, що спочатку в області допустимих рішень системи обмежень знаходять оптимальне рішення послабленої цілочисельної задачі лінійного програмування із використанням, наприклад, симплекс-методу. Для визначеності покладемо, що відбувається пошук максимуму показника ефективності. Якщо в отриманому розв'язку деякі базисні змінні є дробами, то обираємо будь-яку із цих змінних і будуємо два обмеження: в одному обмежена величина змінної менше або дорівнює розв'язку задачі лінійного програмування округленому до найближчого цілого числа знизу; а, в іншому, обмежена величина змінної більше або дорівнює розв'язку задачі лінійного програмування округленому до найближчого цілого числа зверху. Наприклад, при побудові додаткового обмеження за деякою  $i$ -ою змінною  $x_i = \frac{9}{2} \in [4; 5]$ , перше обмеження набуває вигляду  $x_i \leq 4$ , а друге обмеження набуває вигляду  $5 \leq x_i$ .

Першим та другим додатковим обмеженням виключаємо з області допустимих розв'язків вихідної цілочисельної задачі лінійного програмування проміжок  $4; 5$  із дробовими значеннями невідомої  $x_i$ . Цей проміжок поділяє область допустимих розв'язків на дві частини ОДР<sub>1</sub> та ОДР<sub>2</sub>, де нова ОДР<sub>1</sub> отримана завдяки врахуванню в обмеженні вихідної задачі додаткового обмеження  $x_i \leq 4$ , а ОДР<sub>2</sub> – врахуванням обмеження  $5 \leq x_i$ .

В результаті такого розподілу області допустимих значень отримаємо дві нові задачі (підзадачі) лінійного програмування. Якщо після їх розв'язання отримаємо значення невідомих, які знову не є цілочисельними, то обираємо для продовження обчислень, що наближає до цілочисельного розв'язку ту підзадачу, у якій значення показника ефективності більше і знову повторюємо процедуру формування додаткових обмежень і пов'язаних з ними підзадач.

Графічно - образне сприйняття процесу наближення до цілочисельного оптимального розв'язку і дало назву *метод гілок і границь*.

**Приклад 14.2.** Для ОДР<sub>1</sub> цілочисельної задачі лінійного програмування(ЦЗЛП) необхідно знайти максимум показника ефективності  $W_{max}$ .

**Ришення**

□ Побудуємо схему (рис.14.2).

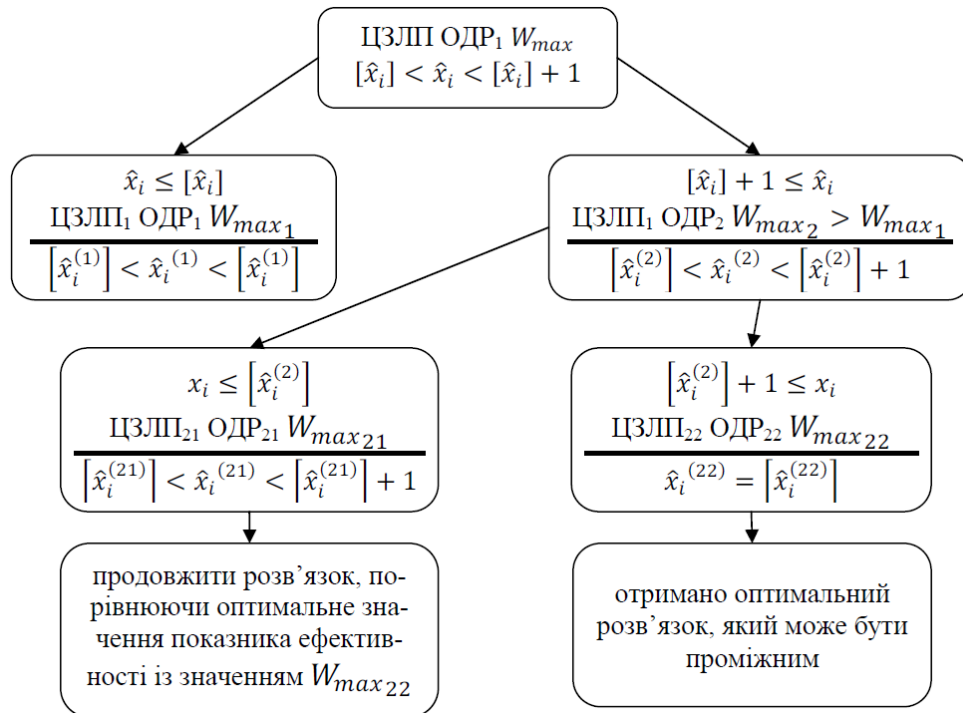


Рис. 14.2. Приклад структури методу гілок та границь за умови, що всі змінні, крім  $x_i$ , на кожній ітерації залишаються цілочисловими та невід’ємними: ЦЗЛП – цілочислова задача лінійного програмування

Процедура розгалуження припиняється тоді, коли буде знайдено цілочисловий розв’язок. Межами вважаються значення показника ефективності задачі і підзадач кожної гілки. На кожній ітерації подальшому розгалуженню (розподіленню на нові підзадачі) підлягає та гілка, на якій значення показника ефективності виявилось більшим. Тому окремі підзадачі (гілки, за якими отримано менші значення показника ефективності) можна відкинути, але можливі випадки з поверненням, коли в подальших ітераціях відбулося зменшення показника ефективності нижче раніше виявленого рівня на гілках, що вважалися неперспективними. Оскільки множина всіх розв’язків цілочислової задачі лінійного програмування скінченна, то після скінченної кількості ітерацій, оптимальний розв’язок буде знайдено. ■

**Приклад 14.3.** Знайти максимум показника ефективності методом гілок та границь із використанням функції **linprog** системи комп'ютерної математики MATLAB:

$$\begin{cases} W = 2x_1 + 4x_2 + 3x_3 \rightarrow \max_{x_{1,2,3} \in \mathbb{N}_0} ; \\ 2x_1 + 3x_2 + x_3 \leq 20, \\ 9x_1 + 7x_2 + 10x_3 \leq 40. \end{cases}$$

### *Ришення*

#### □ *Метод гілок і границь*

Вихідна послаблена цілочислова зображена як симетрична задача лінійного програмування у форматі

```
>> f=[-2;-4;-3]
>> A=[2 3 1
      9 7 10]
>> b=[20;40]
>> lb=[0;0;0]
>> [X,W]=linprog(f,A,b,[],[],lb,[])
Optimization terminated successfully.
X = 0.0000
    5.7143
    0.0000
W = -22.8571
```

#### *Гілка 1.*

```
>> A=[2 3 1
      9 7 10
      0 1 0]
>> b=[20;40;5]
>> [X,W]=linprog(f,A,b,[],[],lb,[])
Optimization terminated successfully.
X = 0.0000
    5.0000
    0.5000
W = -21.5000
```

#### *Гілка 1.1.*

```

>> Aeq=[0 0 1]
>> beq=[0]
>> [X,W]=linprog(f,A,b,Aeq,beq,lb,[])
Optimization terminated successfully.
X = 0.5556
    5.0000
    0
W = -21.1111

```

**Гілка 1.1.1.**

```

>> Aeq=[1 0 0
        0 0 1]
>> beq=[0;0]
>> [X,W]=linprog(f,A,b,Aeq,beq,lb,[])
Optimization terminated successfully.
X = 0
    5.0000
    0
W = -20.0000

```

**Гілка 2.**

```

>> A=[2 3 1
      9 7 10
      0 -1 0]
>> b=[20;40;-6]
>> [X,W]=linprog(f,A,b,[],[],lb,[])

```

Exiting: One or more of the residuals, duality gap, or total relative **error** has stalled:  
the primal appears to be infeasible (and the dual unbounded).  
(The dual residual < TolFun=1.00e-008.)

```

X = 0.0001
    5.8937
    0.0006
W = -23.5768

```

**Висновок.** Отримано оптимальний цілочисловий розв'язок:

$x_1=0$ ,  $x_2=5$ ,  $x_3=0$ ,  $W = 20$ . ■

## Контрольні запитання та завдання

1. Наведіть та поясніть постановку цілочислових задач лінійної оптимізації.
2. Дайте визначення задачі цілочислового програмування.
3. У чому полягає суть методу Гоморі розв'язання цілочислових задач лінійного програмування?
4. Наведіть та поясніть геометричну інтерпретацію методу Гоморі розв'язання цілочислових задач лінійного програмування.
5. Поясніть призначення функції *linprog* системи комп'ютерної математики MATLAB.

## Глава 15. Нелінійне програмування

### 15.1. Класичні умови екстремуму задачі нелінійного програмування

Загальна задача нелінійного програмування полягає в знаходженні екстремуму показника ефективності:

$$W = x_1, \dots, x_n \rightarrow \underset{x_j \in \mathbb{R}, j=1, n}{\text{extrem}} \quad (15.1)$$

на множині планів задачі, яка задається системою умов

$$g_i(x_1, \dots, x_n) \leq \geq b_i \quad i = \overline{1, m}, \quad (15.2)$$

де умова дискретності змінних та їх знаку теж враховується за допомогою функцій  $g_i(x_1, \dots, x_n)$ . Підкреслимо, що функції  $g_i(x_1, \dots, x_n)$   $i = \overline{1, m}$  та показник ефективності  $W(x_1, \dots, x_n)$  усі або частково є нелінійними.

До специфічних властивостей задачі нелінійного програмування слід віднести:

- 1) багатоекстремальність;
- 2) оптимальний план (план, на якому досягається, наприклад, глобальний максимум) може бути як внутрішньою, так і граничною точкою області допустимих розв'язків;
- 3) область допустимих розв'язків (ОДР) може не бути опуклою та зв'язаною, тобто, складається з декількох частин, що не перетинаються (рис.15.1);

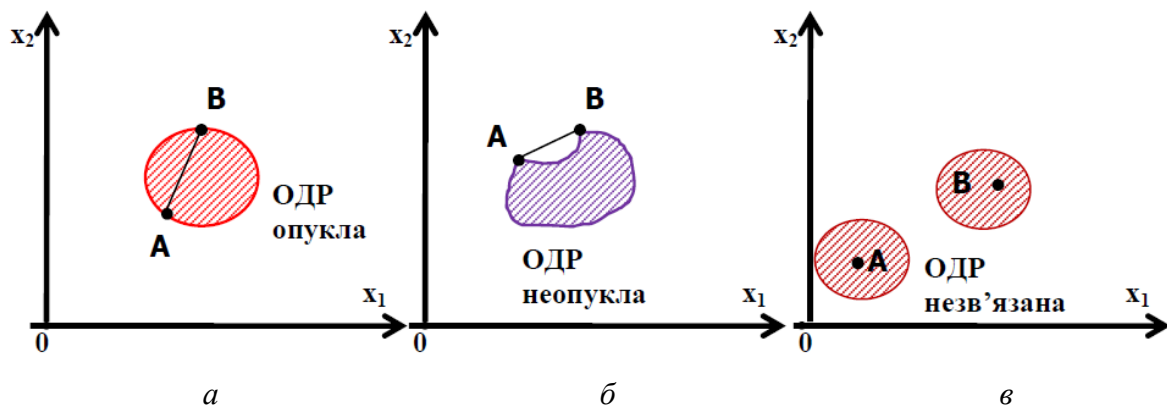


Рис. 15.1. Графічне пояснення особливостей ОДР

4) показник ефективності та обмеження можуть мати кутові точки та розриви.

Класична детермінована задача пошуку екстремуму, тобто максимуму або мінімуму деякої неперервної та кусково-диференційованої функції  $n$  змінних ставиться наступним чином: знайти всі значення вектора  $X^T = [x_1, \dots, x_n]$ , при яких функція  $W(X)$  досягає екстремуму, при виконанні ОДР  $G$ , яка задається за допомогою обмежень у формі рівнянь, тобто у обмеженнях (15.2) існує лише  $m$  обмежень-рівностей, і математично записується у вигляді:

$$W(X) \rightarrow \text{extrem.} \\ X \in G \subset \mathbb{R}^n$$

Різницю  $n - m$  називають *кількістю степенів свободи* задачі нелінійного програмування, а саму задачу нелінійного програмування із умовами (обмеженнями) називають *задачею нелінійного програмування на умовний екстремум*. Розглянемо застосування методу прямої підстановки та методу множників Лагранжа для розв'язання класичної задачі нелінійного програмування на умовний екстремум.

*Метод прямої підстановки*

Припустимо, що

$$X^T = [X_k^T, X_m^T],$$

де  $X_k^T = [x_1, \dots, x_k]$  –  $k$ -вимірний вектор вільних змінних,  $X_m^T = [x_{k+1}, \dots, x_n]$  –  $m$ -вимірний вектор базисних змінних ( $k + m = n$ ).

Тоді розв'язуючи у явному вигляді рівняння-обмеження відносно базисних змінних  $x_{k+i} = \varphi_i$ ,  $i = \overline{1, m}$  і підставляючи їх у вираз показника ефективності, отримаємо класичну безумовну задачу нелінійного програмування:

$$W(x_1, \dots, x_k, \varphi_{k+1}, \dots, \varphi_n) \rightarrow \text{extrem.} \\ [x_1, \dots, x_k] \in \mathbb{R}^k \quad (15.3)$$

Якщо  $k = 1$ , то показник ефективності буде залежати лише від однієї змінної. Припустимо, що необхідно знайти максимум показника

ефективності  $W_{max}$ . За визначенням, в точці  $\hat{x}_1$  досягається строгий локальний максимум, якщо  $W \hat{x}_1 > W \hat{x}_1 + \Delta x_1 \quad \forall \Delta x_1 \rightarrow 0$ .

Знайдемо диференціальні умови існування  $\hat{x}_1$  – точки локального максимуму. Розкладемо  $W x_1 = W \hat{x}_1 + \Delta x_1$  за умови  $\Delta x_1 \rightarrow 0$  в ряд Тейлора в околі точки  $\hat{x}_1$ :

$$\begin{aligned} \Delta W \hat{x}_1 &= W \hat{x}_1 + \Delta x_1 - W \hat{x}_1 = \\ &= \left. \frac{dW}{dx_1} x_1 \right|_{x_1=\hat{x}_1} \cdot \Delta x_1 + \frac{1}{2!} \left. \frac{d^2W}{dx_1^2} x_1 \right|_{x_1=\hat{x}_1} \cdot \Delta x_1^2 + o(\Delta x_1) < 0, \quad \forall \Delta x_1 \rightarrow 0. \end{aligned}$$

Із записаного виразу отримуємо достатні умови локального максимуму функції однієї змінної:

$$\begin{cases} \left. \frac{dW}{dx_1} x_1 \right|_{x_1=\hat{x}_1} = 0, \\ \left. \frac{d^2W}{dx_1^2} x_1 \right|_{x_1=\hat{x}_1} < 0. \end{cases} \quad (15.4)$$

**Приклад 15.1.** На космічній платформі з апаратурою системи космічного зв'язку розташовано контейнер з ядерною енергетичною установкою циліндричної форми. Для максимізації маси ядерного пального за заданою масою контейнера необхідно виготовити контейнер максимального об'єму за заданої площі поверхні (рис.15.2).

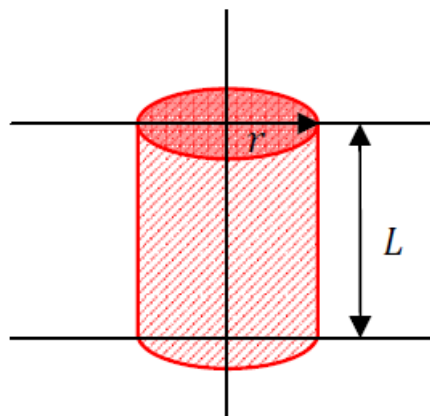


Рис. 15.2. Форма контейнера

## Ришення

□ Математична модель задачі:

$$\begin{cases} W(r, L) = \pi r^2 L; \\ S_0 = 2\pi r^2 + 2\pi r L. \end{cases}$$

Математична постановка задачі:

$$\begin{cases} W(r, L) = \pi r^2 L \rightarrow \max_{r, L}; \\ S_0 = 2\pi r^2 + 2\pi r L. \end{cases}$$

Розв'язання методом прямої підстановки:

$$L = \frac{1}{2\pi r} (S_0 - 2\pi r^2),$$

$$W(r) = \frac{S_0}{2} r - \pi r^3.$$

Необхідна умова екстремуму:

$$\frac{dW(r)}{dr} = 0.$$

Обчислюємо похідну:

$$\frac{S_0}{2} - 3\pi r^2 = 0 \Leftrightarrow r_{\max} = \sqrt{\frac{S_0}{6\pi}}.$$

За фізичним змістом мінімальний об'єм  $V_{\min} = 0$  при  $r = 0$  або  $L = 0$ , тобто знайдене значення  $r_{\max}$  відповідає максимальному об'єму.

Використаємо достатні умови для формального підтвердження цього факту:

$$\frac{d^2W(r)}{dr^2} = -6\pi r < 0, \quad r > 0.$$

Отже

$$\begin{cases} r_{\max} = r = \sqrt{\frac{S_0}{6\pi}}; \\ L_{\max} = \sqrt{\frac{2S_0}{3\pi}}. \end{cases} \blacksquare$$

**Приклад 15.2.** Знайти екстремум функції  $W(x_1, x_2)$ :

$$W(x_1, x_2) = \frac{1}{x_1 - 1^2 + x_2 - 1^2 + 1}.$$

**Ришення**

□ Необхідна умова екстремуму:

$$\frac{\partial W(x_1, x_2)}{\partial x_1} = \frac{2x_1 - 1}{x_1 - 1^2 + x_2 - 1^2 + 1} = 0,$$

$$\frac{\partial W(x_1, x_2)}{\partial x_2} = \frac{2x_2 - 1}{x_1 - 1^2 + x_2 - 1^2 + 1} = 0.$$

За геометричним змістом задачі зрозуміло, що в точці екстремуму  $\hat{x}_1 = \hat{x}_2 = 1$  функція  $W(x_1, x_2)$  приймає найбільше значення  $W(\hat{x}_1, \hat{x}_2) = 1$ .

Як видно з графіку (див. рис.15.3), функція  $W(x_1, x_2)$  має єдиний екстремум, що є глобальним максимумом:

$$\hat{x}_1 = \hat{x}_2 = 1,$$

$$W(\hat{x}_1, \hat{x}_2) = 1. \blacksquare$$

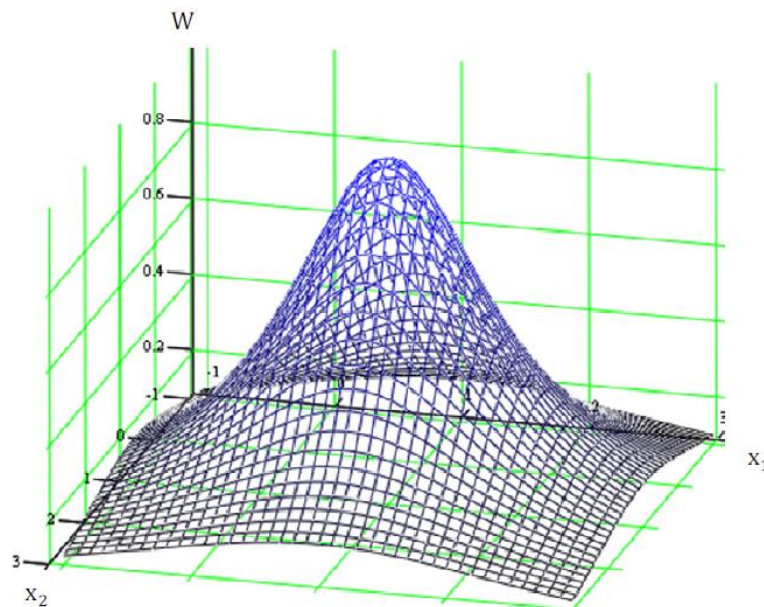


Рис. 15.3. Графічне зображення визначення типу екстремуму функції  $W(x_1, x_2)$

Покладемо  $k > 1, m < n, k + m = n$ , ( $k = n - m$  - кількість степенів свободи)  $k, n, m \in \mathbb{N}$ . За означенням,  $W X_k$  має в точці  $X_k$  строгий локальний максимум, якщо

$$W X_k > W \hat{X}_k + \Delta X_k \quad \forall \Delta x_i \rightarrow 0 \quad i = \overline{1, k},$$

$$X_k^T = [x_1, \dots, x_k], \quad \Delta X_k^T = [\Delta x_1, \dots, \Delta x_k].$$

Розкладання в ряд Тейлора скалярної функції векторного аргументу відносно  $X_k$  дає можливість обчислити приріст показника ефективності в малому околі точки максимуму цього показника ефективності, тобто  $\hat{X}_k$ :

$$\begin{aligned} \Delta W \hat{X}_k &= W \hat{X}_k + \Delta \hat{X}_k - W \hat{X}_k = \\ &= \left[ \frac{\partial W X_k}{\partial X_k} \right]_{X_k = \hat{X}_k} \cdot \Delta X_k + \frac{1}{2!} \Delta X_k^T \left[ \frac{\partial}{\partial X_k} \left( \frac{\partial W X_k}{\partial X_k} \right)^T \right]_{X_k = \hat{X}_k} \cdot \Delta X_k + o \|\Delta X_k\| \\ &\quad \forall \Delta X_k \rightarrow 0, \end{aligned}$$

де  $o \|\Delta X_k\|$  - доданки вищого порядку малості порівняно із

$$\|\Delta X_k\| = \sqrt{\Delta x_1^2 + \Delta x_2^2 + \dots + \Delta x_k^2}.$$

Виходячи із означення локального максимуму, маємо:

$$\Delta W \hat{X}_k < 0, \quad \forall \Delta X_k \rightarrow 0,$$

$$\left[ \frac{\partial W X_k}{\partial X_k} \right]_{X_k = \hat{X}_k}^T = \left[ \frac{\partial W}{\partial X_1}, \dots, \frac{\partial W}{\partial X_k} \right] = 0.$$

тобто  $\left. \frac{\partial W X_k}{\partial X_k} \right|_{X_k = \hat{X}_k}$  - градієнт показника ефективності в точці локального максимуму.

Квадратична форма:

$$\frac{1}{2!} \Delta X_k^T \left[ \frac{\partial}{\partial X_k} \left( \frac{\partial W}{\partial X_k} X_k \right)^T \right]_{X_k = \hat{X}_k} \cdot \Delta X_k < 0 \quad \forall \Delta X_k \rightarrow 0 \quad (15.5)$$

побудована із використанням квадратичної симетричної матриці

$$\left[ \frac{\partial}{\partial X_k} \left( \frac{\partial W}{\partial X_k} X_k \right)^T \right] = \begin{bmatrix} \frac{\partial^2 W}{\partial X_1^2} & \frac{\partial^2 W}{\partial X_1 \partial X_2} & \dots & \frac{\partial^2 W}{\partial X_1 \partial X_k} \\ \frac{\partial^2 W}{\partial X_2 \partial X_1} & \frac{\partial^2 W}{\partial X_2^2} & \dots & \frac{\partial^2 W}{\partial X_2 \partial X_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 W}{\partial X_k \partial X_1} & \frac{\partial^2 W}{\partial X_k \partial X_2} & \dots & \frac{\partial^2 W}{\partial X_k^2} \end{bmatrix}, \quad (15.6)$$

яка носить назву: *матриця Гессе*. Ця квадратична форма у випадку локального максимуму повинна бути від'ємно-визначеною, тобто:

$$\Delta X_k^T \left[ \frac{\partial}{\partial X_k} \left( \frac{\partial W}{\partial X_k} X_k \right)^T \right]_{X_k = \hat{X}_k} \cdot \Delta X_k < 0$$

Квадратична форма буде від'ємно-визначеною, коли матриця Гессе задовольняє спеціальній умові. Цю матрицю Гессе, яка задовольняє умові від'ємної визначеності квадратичної форми, теж називають від'ємно-визначеною.

Ознакою від'ємної визначеності квадратичної форми, побудованої із використанням квадратної симетричної матриці

$$Q = \begin{bmatrix} q_{11} & \dots & q_{1n} \\ \dots & \dots & \dots \\ q_{n1} & \dots & q_{nn} \end{bmatrix}$$

є : усі кутові діагональні визначники матриці  $Q$  непарного порядку повинні бути від'ємними, а парного – додатними, тобто,

$$q_{11} < 0,$$

$$\begin{vmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{vmatrix} > 0,$$

$$\begin{vmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{vmatrix} < 0$$

і т.д.

Якщо показник ефективності в точці  $\hat{X}_k$  має мінімум, то достатня умова мінімуму має вигляд:

$$\begin{cases} \left. \frac{\partial W}{\partial X_k} \right|_{X_k = \hat{X}_k} = 0, \\ \left\| \Delta_k X_k^T \left[ \frac{\partial}{\partial X_k} \left( \frac{\partial W}{\partial X_k} \right) \right]^T \right\|_{X_k = \hat{X}_k} > 0. \end{cases} \quad (15.7)$$

Ознакою додатної визначеності квадратичної форми є додатність усіх кутових визначників квадратичної симетричної матриці  $Q$ , на якій ця форма побудована.

**Приклад 15.3.** Перевірити знаковизначеність квадратичної форми, заданої симетричною матрицею

$$Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix}, X_k \in \mathbb{R}^3.$$

□

$$q_1 = 1 > 0, q_2 = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 2 - 1 = 1 > 0,$$

$$q_3 = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{vmatrix} = 6 + 1 + 1 - 2 - 1 - 3 = 2 > 0.$$

$$[\Delta x_1, \Delta x_2, \Delta x_3] \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{bmatrix} > 0 \quad \forall [\Delta x_1, \Delta x_2, \Delta x_3]^T \in \mathbb{R}^3 \setminus [0, 0, 0]^T.$$

**Висновок:** матриця  $Q$  додатно-визначена.

Якщо задану матрицю  $Q$  помножити на  $-1$ , то отримаємо:

$$\Delta X_3^T \begin{bmatrix} -1 & -1 & -1 \\ -1 & -2 & -1 \\ -1 & -1 & -3 \end{bmatrix} \cdot \Delta X_3 < 0,$$

де

$$\Delta X_3^T = [\Delta x_1, \Delta x_2, \Delta x_3]^T,$$

тобто отримана форма є від'ємно-визначеною.

Проілюструємо ознаку від'ємної визначеності квадратної симетричної матриці  $-Q$ :

$$-Q = \begin{bmatrix} -1 & -1 & -1 \\ -1 & -2 & -1 \\ -1 & -1 & -3 \end{bmatrix},$$

$$q_1 = -1 < 0, \quad q_2 = \begin{vmatrix} -1 & -1 \\ -1 & -2 \end{vmatrix} = 1 > 0,$$

$$q_3 = \begin{vmatrix} -1 & -1 & -1 \\ -1 & -2 & -1 \\ -1 & -1 & -3 \end{vmatrix} = -2 < 0$$

■

**Зауваження 15.1.** Окрема умова  $\frac{\partial W}{\partial X_k} X_k = 0$  має назву *необхідної умови екстремуму першого порядку*.

Умова

$$\begin{cases} \frac{\partial W}{\partial X_k} X_k = 0, \\ \frac{\partial}{\partial X_k} \left( \frac{\partial W}{\partial X_k} X_k \right)^T \leq 0 \end{cases}$$

має назву *необхідної умови максимуму другого порядку*. Матрицю Гессе в цьому випадку називають *від'ємно-напіввизначеною*.

Необхідна умова мінімуму другого порядку набуває вигляду:

$$\begin{cases} \frac{\partial W}{\partial X_k} = 0, \\ \frac{\partial}{\partial X_k} \left( \frac{\partial W}{\partial X_k} \right)^T \geq 0. \end{cases}$$

Матрицю Гессе в цьому випадку називають *додатно-напіввизначеною*.

Як бачимо, для розв'язання задачі пошуку екстремуму скалярного показника ефективності із векторним аргументом на основі записаних вище достатніх умов необхідно, щоб  $W(x_1 \dots x_k)$  мала частинні похідні до другого порядку включно за своїм аргументом, і при цьому існувала можливість отримання явного розв'язку системи обмежень, відносно базисних змінних. Але остання умова досить часто не виконується. Тому, за рахунок використання так званого принципу розширення вектору змінних та показника ефективності, було запропоновано спеціальний метод розв'язання класичної задачі нелінійного програмування на умовний екстремум, який отримав назву *метод множників Лагранжа*.

### *Метод множників Лагранжа*

Метод множників Лагранжа полягає в тому, що класичну задачу нелінійного програмування на умовний екстремум з обмеженнями-рівностями замінюють спрощеною умовною задачею із розширеним вектором змінних

$$\left[ X^T, \Lambda^T \right]^T$$

де  $X^T = [x_1, \dots, x_n]$  – вихідний вектор змінних,  $\Lambda^T = [\lambda_1, \dots, \lambda_m]$  – вектор додаткових змінних, який називають *вектором множників Лагранжа*, а також із розширеним показником ефективності

$$L(X, \Lambda) = W(X) + \Lambda^T (B - G(X)). \quad (15.8)$$

Таким чином можна записати розширений показник ефективності

$$L(X, \Lambda) = W(x_1, \dots, x_k) + \sum_{i=1}^m \lambda_i (b_i - g_i(x_1, \dots, x_n)) \quad (15.9)$$

який називають *функцією Лагранжа* або *лагранжіаном*.

Необхідна умова екстремуму в розширеній задачі нелінійного програмування має вигляд:

$$\left\{ \begin{array}{l} \frac{\partial L(X, \Lambda)}{\partial X} = \frac{\partial W(X)}{\partial X} + \left( -\frac{\partial G^T(X)}{\partial X} \right) \Lambda = 0 \Leftrightarrow \\ \Leftrightarrow \frac{\partial W(X)}{\partial x_j} - \sum_{i=1}^m \lambda_i \frac{\partial g_i(X)}{\partial x_j} = 0 \quad j = \overline{1, n}; \\ \frac{\partial L(X, \Lambda)}{\partial \Lambda} = B - G(X) = 0 \Leftrightarrow b_i - g_i(X) = 0 \quad i = \overline{1, m}. \end{array} \right. \quad (15.10)$$

Після обчислення розв'язків записаної системи рівнянь знаходимо точки екстремумів. В залежності від умови задачі виділяємо точки, в яких показник ефективності  $W(\hat{X})$  досягає максимуму або мінімуму і, в подальшому, виконуємо пошук глобального максимуму або мінімуму. В багатьох практичних випадках із фізичного змісту задачі зрозуміло, в якій точці досягається найбільше або найменше значення показника ефективності із урахуванням обмежень.

**Приклад 15.4.** Розв'яжемо приклад 15.1, використовуючи метод множників Лагранжа.

### **Рішення**

□ Запишемо функцію Лагранжа

$$L(r, l, \lambda) = \pi r^2 l + \lambda(S_0 - 2\pi r^2 - 2\pi r l).$$

Необхідна умова набуває вигляду:

$$\left\{ \begin{array}{l} \frac{\partial L}{\partial r} = 2\pi r l + \lambda(-4\pi r) - 2\pi l = 0, \end{array} \right. \quad (1)$$

$$\left\{ \begin{array}{l} \frac{\partial L}{\partial l} = \pi r^2 + \lambda(-2\pi r) = 0, \end{array} \right. \quad (2)$$

$$\left\{ \begin{array}{l} \frac{\partial L}{\partial \lambda} = S_0 - 2\pi r^2 - 2\pi r l = 0; \end{array} \right. \quad (3)$$

Розв'язання задачі (1)-(3):

$$2 \Rightarrow \lambda = \frac{r}{2},$$

$$1 \rightarrow rl - 2r^2 = 0 \rightarrow \begin{cases} r = \frac{l}{2}; \\ r = 0 - \text{не задовольняє фізичному змісту.} \end{cases}$$

$$3 \rightarrow S_0 - 6\pi r^2 = 0 \rightarrow r_{max} = \sqrt{\frac{S_0}{6\pi}}; l_{max} = \sqrt{\frac{2S_0}{3\pi}}; \lambda_{max} = \frac{1}{2} \sqrt{\frac{S_0}{6\pi}}. \blacksquare$$

Розглянемо на прикладі двовимірної задачі лінійного програмування фізичний (економічний) зміст множників Лагранжа:

$$\begin{cases} W(x_1, x_2) \rightarrow \max; \\ g(x_1, x_2) = b. \end{cases}$$

Припустимо, що умовний максимум досягається в точці  $\hat{X}^T = [\hat{x}_1, \hat{x}_2]$ .

Розглядаючи залежність  $\hat{X}$  від зміни параметра  $b$ , отримаємо залежності:

$$\hat{x}_1 = \hat{x}_1(b), \hat{x}_2 = \hat{x}_2(b),$$

$$W_{max} = W(\hat{x}_1, \hat{x}_2) \text{ , } \hat{x}_1, \hat{x}_2 \text{ - функції } b \text{ .}$$

Знайдемо похідну від оптимального значення показника ефективності по параметру  $b$ :

$$\frac{\partial \hat{W}}{\partial b} = \frac{\partial W}{\partial \hat{x}_1} \cdot \frac{\partial \hat{x}_1}{\partial b} + \frac{\partial W}{\partial \hat{x}_2} \cdot \frac{\partial \hat{x}_2}{\partial b} \text{ ,}$$

А також похідну по параметру  $b$  від обох частин обмеження-рівності:

$$\frac{\partial g}{\partial \hat{x}_1} \cdot \frac{\partial \hat{x}_1}{\partial b} + \frac{\partial g}{\partial \hat{x}_2} \cdot \frac{\partial \hat{x}_2}{\partial b} = 1.$$

Виходячи з того, що в заданій задачі нелінійного лінійного програмування функція Лагранжа (15.9) має вигляд:

$$L(x_1, x_2, \lambda) = W(x_1, x_2) + \lambda(b - g(x_1, x_2)) \text{ ,}$$

знаходимо частинні похідні в оптимальній точці:

$$\begin{cases} \frac{\partial L}{\partial \hat{x}_1} = \frac{\partial W}{\partial \hat{x}_1} - \hat{\lambda} \frac{\partial g}{\partial \hat{x}_1} = 0, \\ \frac{\partial L}{\partial \hat{x}_2} = \frac{\partial W}{\partial \hat{x}_2} - \hat{\lambda} \frac{\partial g}{\partial \hat{x}_2} = 0. \end{cases}$$

Підставляючи частинні похідні у вираз для обчислення похідної для оптимального значення показника ефективності, отримаємо:

$$\frac{\partial \hat{W}}{\partial b} = \frac{\hat{\lambda} \partial g}{\partial \hat{x}_1} \cdot \frac{\partial \hat{x}_1}{\partial b} + \frac{\hat{\lambda} \partial g}{\partial \hat{x}_2} \cdot \frac{\partial \hat{x}_2}{\partial b} = \hat{\lambda} \left( \frac{\partial g}{\partial \hat{x}_1} \cdot \frac{\partial \hat{x}_1}{\partial b} + \frac{\partial g}{\partial \hat{x}_2} \cdot \frac{\partial \hat{x}_2}{\partial b} \right) = \hat{\lambda}.$$

Узагальнюючи отриманий результат на випадок  $n$ -мірного вектора  $X$  та на випадок дії  $m$  обмежень  $m < n$ , отримаємо:

$$\left. \frac{\partial \hat{W}}{\partial b_i} \right|_{\substack{X=\hat{X} \\ \Lambda=\hat{\Lambda}}} = \hat{\lambda}_i \quad i = \overline{1, m}. \quad (15.11)$$

### **Фізичний(економічний) зміст**

Вважаючи, що  $W(X)$  – прибуток і при цьому виконується максимізація показника ефективності, або  $W(X)$  – вартість (витрати) виробництва і при цьому виконується мінімізація показника ефективності, за умови, що  $b_i$  – об'єм  $i$ -го ресурсу, робимо висновок про те, що  $i$ -й множник Лагранжа показує на скільки зміниться максимальний прибуток або мінімальна вартість виробництва при зміні  $i$ -го ресурсу на 1.

**Зауваження 15.3.** У більшості задач, присвячених системним дослідженням складних технічних систем, спеціально виділяють вектор стану, який можна інтерпретувати як вектор базисних змінних.

*Одноетапною процедурою прийняття рішення з обмеженнями у формі рівнянь називається результат розв'язання задачі нелінійного програмування, поставленої у вигляді:*

$$\begin{cases} W(X,U) \rightarrow \max_{X \in \mathbb{R}^n, U \in \mathbb{R}^m} (\min), \\ F(X,U) = 0. \end{cases} \quad (15.12)$$

де

- $X$  –  $n$ -вимірний вектор,
- $U$  –  $m$ -вимірний вектор,
- $W(X,U) : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$  – скалярний показник ефективності,
- $F(X,U) : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$  –  $n$ -вимірна векторна функція,
- $F^T(X,U) = [f_1(X,U), \dots, f_n(X,U)]$ .

Необхідні умови для пошуку оптимального розв'язку мають вигляд:

$$\left\{ \begin{array}{l} \frac{\partial L}{\partial X} \Big|_{\substack{X=\hat{X} \\ U=\hat{U} \\ \Lambda=\hat{\Lambda}}} = \frac{\partial W(X,U)}{\partial X} + \frac{\partial}{\partial X} F^T(X,U) \Lambda \Big|_{\substack{X=\hat{X} \\ U=\hat{U} \\ \Lambda=\hat{\Lambda}}} = 0; \\ \frac{\partial L}{\partial U} \Big|_{\substack{X=\hat{X} \\ U=\hat{U} \\ \Lambda=\hat{\Lambda}}} = \frac{\partial W(X,U)}{\partial U} + \frac{\partial}{\partial U} F^T(X,U) \Lambda \Big|_{\substack{X=\hat{X} \\ U=\hat{U} \\ \Lambda=\hat{\Lambda}}} = 0; \\ F(\hat{X}, \hat{U}) = 0, \end{array} \right. \quad (15.13)$$

де

$$L(X, U, \Lambda) = W(X, U) + \Lambda^T \cdot F(X, U),$$

$$\Lambda^T = [\lambda_1, \dots, \lambda_m] - \text{множник Лагранжа.}$$

**Приклад 15.5.** Дано:

$$W(x_1, x_2) = \frac{1}{x_1^2 - 1 + x_2^2 - 1 + 1}.$$

1) Знайти точку екстремуму та із використанням матриці Гессе з'ясувати тип екстремуму.

2) Розв'язати задачу

$$\begin{cases} W(x_1, x_2) \rightarrow \max(\min); \\ x_1^2 + x_2^2 = 1 \end{cases}$$

застосовуючи метод множників Лагранжа.

**Ришення**

□ 1. Враховуючи результат розв'язання прикладу 15.2 маємо

$$\hat{x}_1 = \hat{x}_2 = 1, W(\hat{x}_1, \hat{x}_2) = 1.$$

Тоді в точці екстремуму матриця Гессе та її діагональні визначники набувають вигляду:

$$Q = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}, q_1 = -2 < 0, q_2 = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = 4 > 0.$$

Отже, квадратична форма від'ємно визначена, тобто (1;1)– точка максимуму.

2. Побудуємо лагранжіан:

$$L(x_1, x_2, \lambda) = W(x_1, x_2) + \lambda(1 - x_1^2 - x_2^2)$$

та скористаємося необхідною умовою екстремуму

$$\begin{cases} \frac{\partial L}{\partial x_1} = 0, \\ \frac{\partial L}{\partial x_2} = 0, \\ \frac{\partial L}{\partial \lambda} = 0. \end{cases}$$

В результаті отримаємо

$$\begin{cases} x_{11} = x_{12} = -\frac{\sqrt{2}}{2}; \\ x_{21} = x_{22} = \frac{\sqrt{2}}{2}. \end{cases}$$

За геометричним змістом задачі зрозуміло, що

$$\begin{cases} x_{1\min} = x_{2\min} = -\frac{\sqrt{2}}{2}; \\ x_{1\max} = x_{2\max} = \frac{\sqrt{2}}{2}, \end{cases}$$

$$W_{\min}\left(-\frac{\sqrt{2}}{2}; -\frac{\sqrt{2}}{2}\right) = \frac{1}{\left(\frac{\sqrt{2}}{2} + 1\right)^2 + \left(\frac{\sqrt{2}}{2} + 1\right)^2 + 1},$$

$$W_{\max}\left(\frac{\sqrt{2}}{2}; \frac{\sqrt{2}}{2}\right) = \frac{1}{\left(\frac{\sqrt{2}}{2} - 1\right)^2 + \left(\frac{\sqrt{2}}{2} - 1\right)^2 + 1}.$$

■

## Контрольні запитання та завдання

1. У чому полягає суть задачі нелінійного програмування?
2. Наведіть та поясніть класичні умови екстремуму задачі нелінійного програмування.
3. Назвіть різновиди ОДР задачі нелінійного програмування.
4. У чому полягає суть методу прямої підстановки розв'язку класичної задачі нелінійного програмування на умовний екстремум?
5. Наведіть та поясніть необхідні умови екстремуму першого та другого порядку.
6. У чому полягає суть методу множників Лагранжа розв'язку класичної задачі нелінійного програмування на умовний екстремум?
7. Наведіть та поясніть необхідну умову екстремуму в розширеній задачі нелінійного програмування.

## Глава 16. Найпростіша задача нелінійного програмування в умовах невід'ємності змінних

У випадку найпростішої задачі нелінійного програмування з усіх обмежень загальної задачі нелінійного програмування залишається лише вимога невід'ємності змінних:

$$W(X) \rightarrow \max_{X \geq 0}. \quad (16.1)$$

На прикладі скалярного показника ефективності, що залежить від двох змінних, розглянемо умови існування його максимуму при невід'ємності змінних. Припустимо:

$$W = W(x_1, x_2) = \frac{1}{x_1 - a^2 + x_2 - b^2 + 1}.$$

Будемо позначати  $x_{1\max}$ ,  $x_{2\max}$  точки максимуму  $W(x_1, x_2)$ , коли обмежень немає,  $x_{1,2} \in R$ ,  $x_{1\max} = a$ ,  $x_{2\max} = b$  (див. рис.16.1).

Будемо позначати  $\hat{x}_1, \hat{x}_2$  координати точки максимуму  $W(x_1, x_2)$  в умовах, коли  $x_1 \geq 0, x_2 \geq 0$  (рис.16.1).

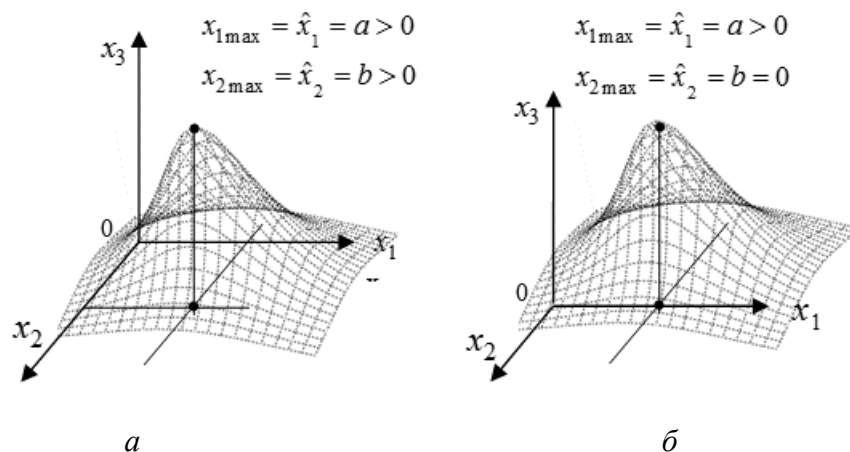
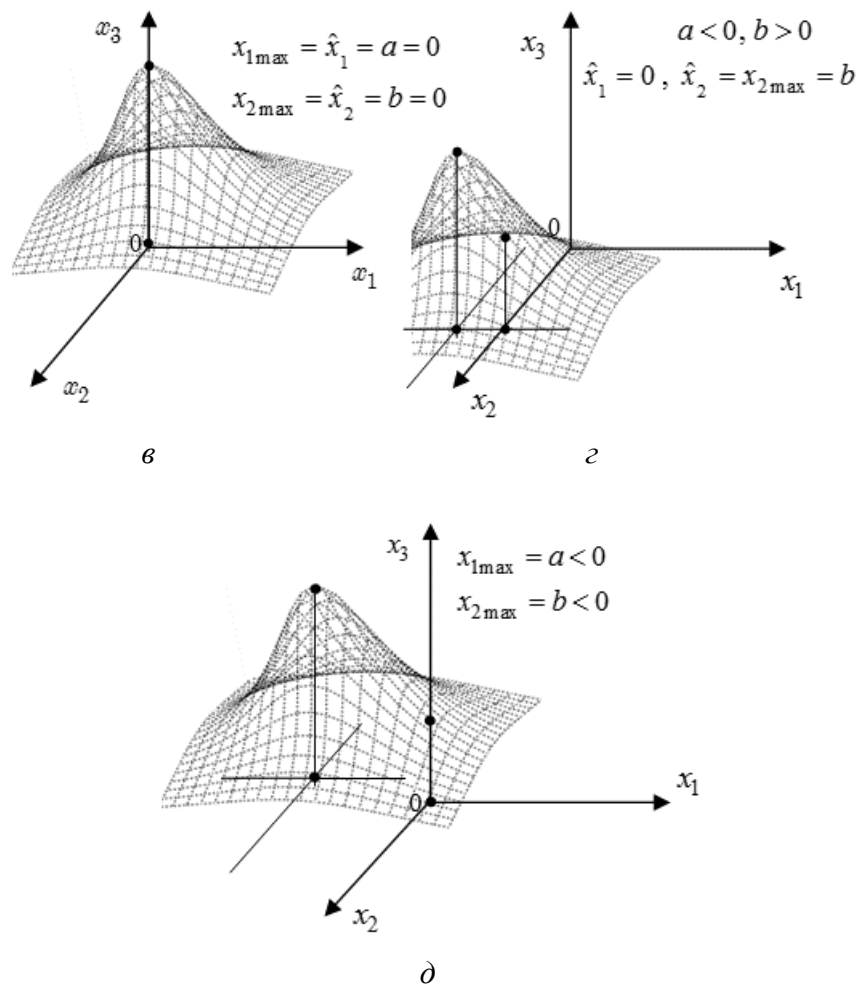


Рис. 16.1. Графічна ілюстрація стосовно пошуку необхідних умов існування найбільшого значення показника ефективності у точці  $\hat{x}_{1,2}$ , представленого у вигляді унімодальної двовимірної функції в умовах дії найпростіших обмежень  $x_{1,2} \geq 0$

Продовження рис. 16.1



Запишемо необхідні умови існування найбільшого значення показника ефективності в точці  $\hat{x}_{1,2}$ , зображеного у вигляді унімодальної двовимірної функції в умовах дії найпростіших обмежень  $x_{1,2} \geq 0$ :

– для рис. 16.1, а – виконується умова  $\frac{\partial W}{\partial x_i} = 0, i = 1, 2$ ;

– для рис. 16.1, б, д відповідно  $\frac{\partial W}{\partial x_1} < 0, \frac{\partial W}{\partial x_2} = 0$ ;  $\frac{\partial W}{\partial x_1} < 0, \frac{\partial W}{\partial x_2} < 0$ .

**Висновок** щодо необхідної умови існування максимуму:

$$\left\{ \begin{array}{l} \left. \frac{\partial W}{\partial x_1} \right|_{x_1=\hat{x}_1} \leq 0; \\ \left. \frac{\partial W}{\partial x_2} \right|_{x_2=\hat{x}_2} \leq 0; \\ \hat{x}_1 \geq 0; \\ \hat{x}_2 \geq 0; \\ \left. \frac{\partial W}{\partial x_1} \right|_{x_1=\hat{x}_1} \cdot \hat{x}_1 = 0; \\ \left. \frac{\partial W}{\partial x_2} \right|_{x_2=\hat{x}_2} \cdot \hat{x}_2 = 0, \end{array} \right. \rightarrow \left\{ \begin{array}{l} \left. \frac{\partial W(X)}{\partial X} \right|_{X=\hat{X}} \leq 0; \\ \hat{X} \geq 0; \\ \left( \left. \frac{\partial W}{\partial X} \right|_{X=\hat{X}} \right)^T \cdot \hat{X} = 0, \end{array} \right.$$

де  $X^T = [x_1, x_2]$ .

Якщо замість  $W$   $x_1, x_2$  із одним максимумом використовується показник ефективності із декількома максимумами (як кажуть, багатомодова функція), то записана умова зберігається і використовується як необхідна умова перевірки існування локального максимуму в умовах найпростіших обмежень-нерівностей (всі координати невід'ємні).

Для  $n$ -вимірному випадку, тобто коли  $X^T = [x_1, \dots, x_n]$ , можливо отримати строго аналітично ті ж самі співвідношення, що і для розглянутого двовимірному прикладу, як необхідну умову існування локального максимуму, якщо скористатись розкладанням у ряд Тейлора показника ефективності поблизу точки  $\hat{X}$  і виконати вимогу

$$\left\{ \begin{array}{l} \Delta W \hat{X} = W \hat{X} + \Delta X - W \hat{X} \leq 0 \quad \forall \Delta X \rightarrow 0_n; \\ X \geq 0_n. \end{array} \right.$$

Ці необхідні вимоги існування локального максимуму в  $n$ -вимірному випадку запишемо у вигляді:

$$1) \text{ якщо } \hat{x}_j = 0, \text{ то } \left. \frac{\partial W}{\partial x_j} \right|_{x=\hat{x}_j} \leq 0;$$

$$2) \text{ якщо } \hat{x}_j > 0, \text{ то } \left. \frac{\partial W}{\partial x_j} \right|_{x=\hat{x}_j} = 0.$$

**Зауваження 16.1.** В умовах 1), 2)  $j$  може приймати будь-яке значення від 1 до  $n$  як окремо, так і по групі координат.

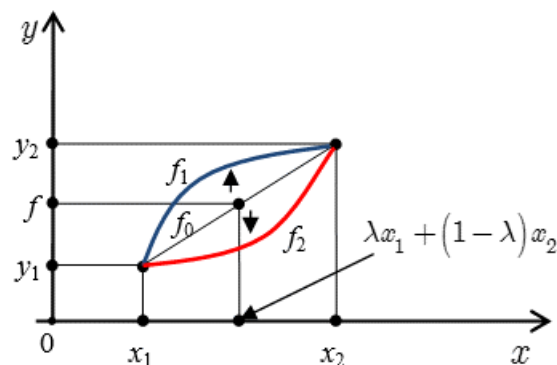
**Зауваження 16.2.** У випадку опуклої догори функції  $W(X)$  для будь-якого  $X \in \mathbb{R}^n$  отримаємо достатню умову існування глобального максимуму при  $X \geq 0$  у тому ж самому вигляді, який було записано вище у пунктах 1) та 2).

**Зауваження 16.3.** Визначення опуклості функції графічно зображено на рис. 16.2. Для векторного аргумента вважаємо функцію опуклою відповідно догори

$$f_1 \lambda x_1 + 1 - \lambda x_2 > \lambda f x_1 + 1 - \lambda f x_2 ,$$

$$f_2 \lambda x_1 + 1 - \lambda x_2 < \lambda f x_1 + 1 - \lambda f x_2 ,$$

де  $\lambda$  – скаляр.



$$f_0 \lambda x_1 + 1 - \lambda x_2 = \lambda f x_1 + 1 - \lambda f x_2 - \text{пряма};$$

$$f_1 \lambda x_1 + 1 - \lambda x_2 > \lambda f x_1 + 1 - \lambda f x_2 - \text{опукла догори};$$

$$f_2 \lambda x_1 + 1 - \lambda x_2 < \lambda f x_1 + 1 - \lambda f x_2 - \text{опукла донизу}$$

Рис. 16.2. Визначення характеру опуклості двовимірної функції

**Приклад 16.1.** Розв'язати графічно задачі:

$$1) \begin{cases} W(x_1, x_2) = x_1^2 - 5x_1 + x_2^2 - 8x_2 \rightarrow \max \min; \\ x_1, x_2 \geq 0; \\ x_1 + 2x_2 \geq 4; \\ \frac{5}{2}x_1 + 2x_2 \leq 20, \end{cases}$$

$$2) \begin{cases} W(x_1, x_2) = x_1^2 - 2x_1 + x_2^2 \rightarrow \max \min; \\ x_1, x_2 \geq 0; \\ x_1 \leq 10; \\ x_2 \leq 9; \\ x_1 + \frac{3}{2}x_2 \leq \frac{15}{2}; \\ x_1 \cdot x_2 \leq 6, \end{cases}$$

$$3) \begin{cases} W(x_1, x_2) = x_2 - x_1^2 + 4x_1 \rightarrow \max; \\ x_1, x_2 \geq 0; \\ x_1 \cdot x_2 \geq 6; \\ x_1^2 + x_2^2 \leq 49; \\ x_2 \leq 6. \end{cases}$$

**Ришення**

□ **Графічні розв'язки задач 1-3 подані на рис. 16.3–16.5:**

Задача 1.

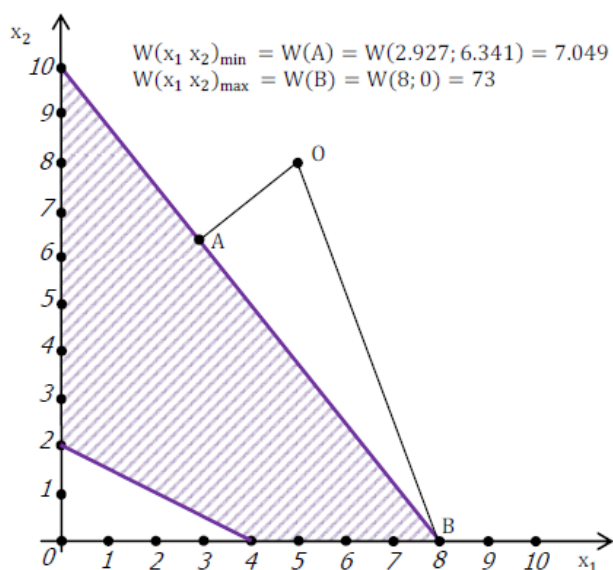


Рис. 16.3. Графічне зображення розв'язання задачі 1

Задача 2.

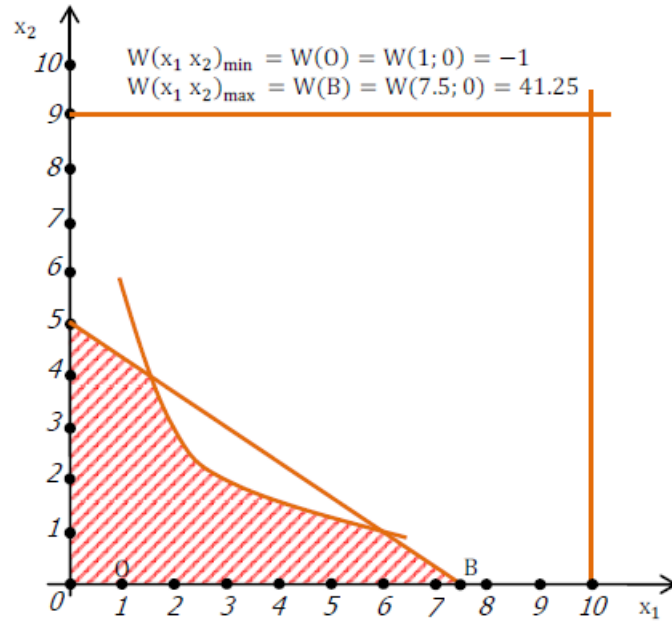


Рис. 16.4. Графічне зображення розв'язання задачі 2

Задача 3.

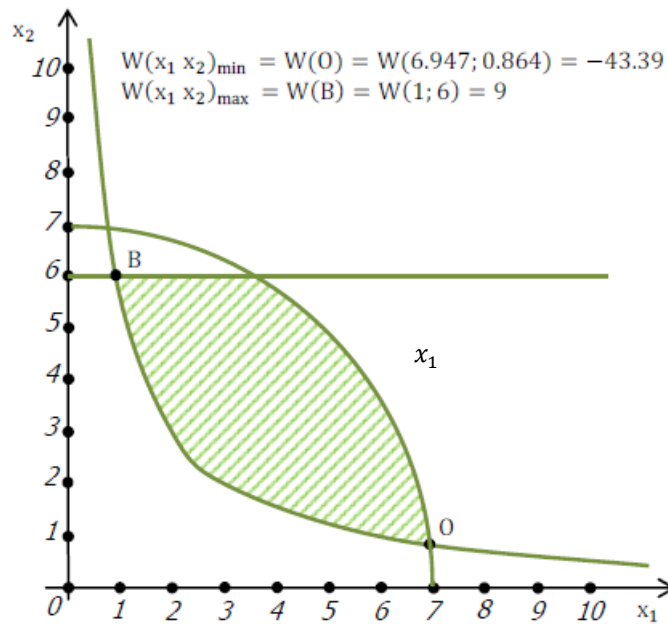


Рис. 16.5. Графічне зображення розв'язання задачі 3

■  
**Приклад 16.2.** Телекомунікаційна компанія надає користувачам послуги супутникового зв'язку ( $i = 1$ ) та безпроводового доступу до інтернету ( $i = 2$ ).

Знижка витрат на одиницю послуги при зростанні кількості користувачів обчислюються за формулою:

$$\Delta V_i = l_i \cdot x_i,$$

де

- $x_i$  – об'єм послуги  $i$ -го виду (кількість користувачів даної послуги);
- $V_i$  – витрати фірми на реалізацію  $i$ -го виду послуги при мінімальній сумарній кількості користувачів обох послуг;
- $l_i$  – коефіцієнти знижок питомих витрат при зростанні кількості користувачів.

Необхідно з'ясувати об'єм послуг кожного виду, при якому забезпечується мінімізація сумарних витрат при виконанні наступних обмежень:

- мінімальна сумарна кількість користувачів: 40;
- техніко-технологічний ресурс двох видів послуг складає відповідно: 160 одиниць; 210 одиниць;
- норми витрат першого ресурсу на обслуговування одного користувача супутникового зв'язку дорівнюють 2, а користувача безпроводового доступу до інтернету - 2,67;
- норми витрат другого ресурсу складають відповідно 3 і 2.

Необхідно побудувати математичну модель задачі, виконати математичну постановку задачі і розв'язати її графічним методом, якщо:

$$V_1 = 100 \text{ у.о.}, V_2 = 140 \text{ у.о.}, l_1 = l_2 = 1 \text{ у.о. на одного користувача.}$$

### ***Ришення***

□ 1. Математична модель:

$$\begin{cases} W(x_1, x_2) = 100 - x_1 x_1 + 140 - x_2 x_2; \\ x_1 + x_2 \geq 40; \\ 2x_1 + 2.67x_2 \leq 160; \\ 3x_1 + 2x_2 \leq 210; \\ x_{1,2} \geq 0. \end{cases}$$

2. Математична постановка:

$$W(x_1, x_2) \rightarrow \min_{x_{1,2} \in G},$$

$$x_1 - 50^2 + x_2 - 70^2 = 50^2 + 70^2 - W$$

де ОДР G задається обмеженнями-нерівностями:

$$\begin{cases} x_2 \geq -x_1 + 40; \\ x_2 \leq \frac{160 - 2x_1}{2,67}; \\ x_2 \leq -\frac{3}{2}x_1 + 105; \\ x_{1,2} \geq 0. \end{cases}$$

3. Графічний розв'язок(див. рис. 16.6).

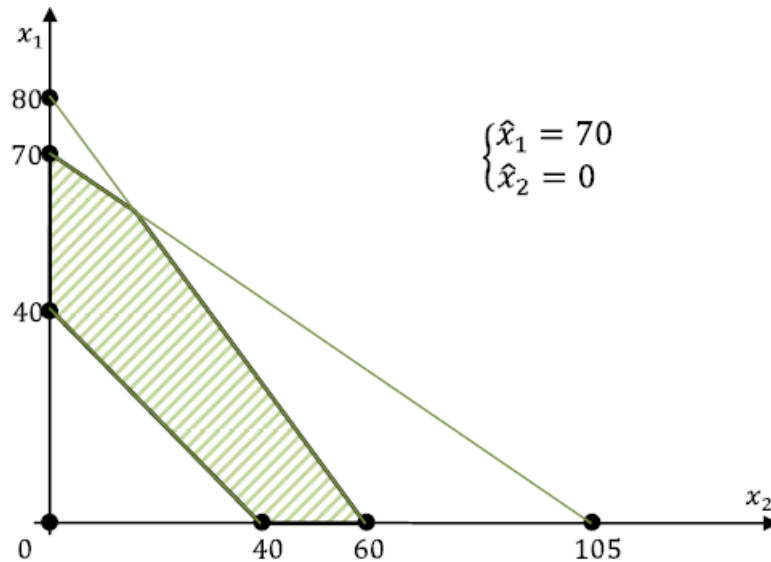


Рис. 16.6. Графічне зображення побудови області допустимих значень

Як бачимо

$$x_1 - 50^2 + x_2 - 70^2 = 50^2 + 70^2 - W = R^2,$$

тобто графічним образом виразу, який пов'язаний із показником ефективності  $W$ , є коло радіусу  $R^2 = 50^2 + 70^2 - W$ .

Чим менше  $W$ , тим більше  $R$ . Найбільше значення  $R$  досягається у точці

$$\begin{cases} \hat{x}_1 = 70 \\ \hat{x}_2 = 0 \end{cases}$$

## Контрольні запитання та завдання

1. У чому полягає суть задачі нелінійного програмування?
2. Наведіть та поясніть математичну постановку задачі нелінійного програмування в умовах невід'ємності змінних.
3. Наведіть та поясніть необхідні умови існування максимуму задачі нелінійного програмування.
4. Поясніть суть поняття «локальний максимум».
5. Назвіть та поясніть необхідні вимоги існування локального максимуму задачі нелінійного програмування.
6. Поясніть суть поняття «глобальний максимум».
7. Наведіть та поясніть достатні умови існування глобального максимуму задачі нелінійного програмування.

## Глава 17. Задачі опуклого та квадратичного програмування

### 17.1. Умови та теорема Куна-Таккера

Задача нелінійного програмування у некласичній постановці, тобто у постановці, коли в якості умов-обмежень у задачі нелінійного програмування використовуються обмеження-нерівності, формулюється наступним чином:

$$\begin{cases} W(X) \rightarrow \max_{x_i \geq 0, (i=1, n)}; \\ g_j X \leq b_j \quad (j = 1, m), \end{cases} \quad (17.1)$$

де  $X^T = [x_1, \dots, x_n]$ .

В якості показника ефективності розглядають опуклу догори функцію для всіх  $x_i \geq 0 \quad i = \overline{1, n}$ . Побудуємо функцію Лагранжа:

$$L(X, \Lambda) = W(X) + \Lambda^T \cdot (B - G X) \rightarrow \max_{\substack{x_i \geq 0, i=1, n \\ \lambda_j \geq 0, j=1, m}}, \quad (17.2)$$

де  $G^T X = [g_1 X, \dots, g_m X]$ ,  $B^T = [b_1, \dots, b_m]$ ,  $\Lambda^T = [\lambda_1, \dots, \lambda_m]$ .

Нам вдалося привести некласичну задачу нелінійного програмування до найпростішої некласичної задачі нелінійного програмування, тобто задачі нелінійного програмування, в якій на змінні накладена лише одна умова - умова їх невід'ємності.

Скористаємося результатом параграфу 16.1:

$$\begin{cases} \left. \frac{\partial L(X, \Lambda)}{\partial X} \right|_{X=\hat{X}} = \left( \frac{\partial W(X)}{\partial X} - \frac{\partial G^T X}{\partial X} \cdot \Lambda \right) \Big|_{X=\hat{X}} \leq 0; \\ \hat{X} \geq 0; \\ \left. \left( \frac{\partial W(X)}{\partial X} - \frac{\partial G^T X}{\partial X} \cdot \Lambda \right)^T \right|_{X=\hat{X}} \cdot \hat{X} = 0; \\ \left. \frac{\partial L(X, \Lambda)}{\partial \Lambda} \right|_{\Lambda=\hat{\Lambda}} = B - G \hat{X} \geq 0; \\ \hat{\Lambda} \geq 0; \\ B - G \hat{X}^T \hat{\Lambda} = 0, \end{cases}$$

$$\text{де } \frac{\partial G^T X}{\partial X} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_n}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial x_n} & \dots & \frac{\partial g_n}{\partial x_n} \end{bmatrix}.$$

Записані умови носять назву умов Куна-Таккера і є необхідними і достатніми умовами існування глобального максимуму задачі нелінійного програмування із опуклим догори показником ефективності при умові виконання обмежень, які утворюють опуклу множину допустимих розв'язків.

Задачі такого типу складають спеціальний розділ задач нелінійного програмування, який називається *опуклим програмуванням*.

Важлива властивість функції Лагранжа полягає в тому, що:

$$L(X, \hat{\Lambda}) \leq L(\hat{X}, \hat{\Lambda}) \leq L(\hat{X}, \Lambda). \quad (17.3)$$

Це означає, що точка  $(\hat{X}, \hat{\Lambda})$  є *сідловою* точкою функції Лагранжа (див. рис.17.1).

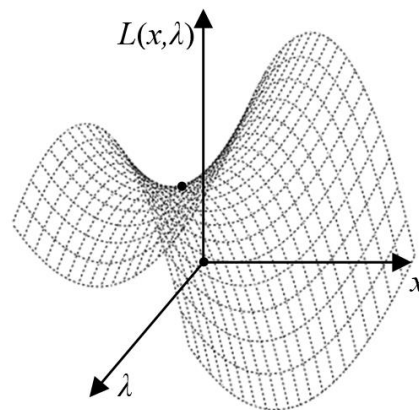


Рис. 17.1. Двовимірна функція Лагранжа із сідловою точкою

Тому умову Куна-Таккера можливо переписати у вигляді двоїстих задач:

$$\begin{cases} L(\hat{X}, \Lambda) = \max_{X \geq 0} L(X, \Lambda) \rightarrow \min_{\Lambda \geq 0}; \\ \frac{\partial L}{\partial \Lambda} \geq 0, \end{cases} \quad (17.4)$$

та

$$\begin{cases} L(X, \hat{\Lambda}) = \min_{\Lambda \geq 0} L(X, \Lambda) \rightarrow \max_{X \geq 0}; \\ \frac{\partial L}{\partial X} \leq 0. \end{cases} \quad (17.5)$$

Вперше властивість оптимальної точки  $(\hat{X}, \hat{\Lambda})$  як сідлової точки функції Лагранжа в задачах опуклого програмування, було доведено Куном і Таккером і сформульовано у формі теореми, яка в подальшому отримала назву теореми Куна-Таккера.

**Теорема 17.1.** Для того, щоб в задачі опуклого програмування деякий план  $\hat{X}$  був точкою глобального максимуму необхідно і достатньо існування вектора множників Лагранжа таких, що точка  $(\hat{X}, \hat{\Lambda})$  буде сідловою точкою функції Лагранжа:

$$L(X, \Lambda) = W(X) + \Lambda^T (B - G(X)).$$

**Зауваження 17.1.** Умови Куна-Таккера ще записують у більш детальному вигляді:

$$\hat{x}_j > 0 \rightarrow \left. \frac{\partial W}{\partial x_j} \right|_{X=\hat{X}} - \Lambda^T \cdot \left. \frac{\partial G}{\partial x_j} \right|_{X=\hat{X}} = 0,$$

$$\left. \frac{\partial W}{\partial x_j} \right|_{X=\hat{X}} - \Lambda^T \cdot \left. \frac{\partial G}{\partial x_j} \right|_{X=\hat{X}} < 0 \rightarrow \hat{x}_j = 0,$$

$$\left. \frac{\partial W}{\partial x_j} \right|_{X=\hat{X}} - \Lambda^T \cdot \left. \frac{\partial G}{\partial x_j} \right|_{X=\hat{X}} = 0 \rightarrow \hat{x}_j = 0 \text{ або } \hat{x}_j > 0,$$

$$\hat{x}_j = 0 \rightarrow \left. \frac{\partial W}{\partial x_j} \right|_{X=\hat{X}} - \Lambda^T \cdot \left. \frac{\partial G}{\partial x_j} \right|_{X=\hat{X}} = 0$$

або

$$\left. \frac{\partial W}{\partial x_j} \right|_{X=\hat{X}} - \Lambda^T \cdot \left. \frac{\partial G}{\partial x_j} \right|_{X=\hat{X}} < 0$$

і далі, відносно змінної  $\lambda_i$ :

$$\lambda_i > 0 \rightarrow g_i(\hat{X}) = b_i, \quad g_i(\hat{X}) < b_i \rightarrow \hat{\lambda}_i = 0,$$

$$\hat{\lambda}_i = 0 \rightarrow g_i \hat{X} = b_i \text{ або } g_i \hat{X} < b_i,$$

$$g_i \hat{X} = 0 \rightarrow \lambda_i = 0 \text{ або } \lambda_i > 0 \quad i = \overline{1, m}.$$

Умови Куна-Таккера, записані у детальному вигляді, називають умовами доповнювальної нежорсткості.

## 17.2. Квадратичне програмування

Спеціальним класом задач нелінійного програмування є так звані задачі квадратичного програмування, у яких показник ефективності складається з лінійної та квадратичної функцій, і при цьому всі обмеження-нерівності – лінійні.

$$\begin{cases} W \cdot X = G \cdot X + \frac{1}{2} X^T \cdot D \cdot X \rightarrow \max_{X \geq 0}; \\ AX \leq B, \end{cases} \quad (17.6)$$

де  $D$  – дійсна симетрична від'ємно-визначена матриця (тобто  $D$  складається із таких елементів, при яких  $X^T \cdot D \cdot X < 0 \quad \forall X \in \mathbb{R}^n \setminus \{0_n\}$ ).

Лінійні обмеження-нерівності утворюють опуклу область допустимих розв'язків. Показник ефективності складається із суми квадратичної форми

$$X^T \cdot D \cdot X = \sum_{k=1}^n \sum_{j=1}^n \alpha_{kj} x_k x_j < 0 \quad \forall X \in \mathbb{R}^n \setminus \{0_n\}$$

та лінійної форми:

$$C \cdot X = \sum_{j=1}^n c_j x_j$$

і є опуклою до гори функцією.

Це означає, що для розв'язання задачі квадратичного програмування можна застосувати методи та теореми опуклого програмування, зокрема теорему та умови Куна-Таккера. Функція Лагранжа для задачі квадратичного програмування має вигляд:

$$L(X, \Lambda) = GX + \frac{1}{2} X^T DX + \Lambda^T (B - AX), \quad (17.7)$$

що дає можливість отримати умови Куна-Таккера:

$$\left\{ \begin{array}{l} \left( \frac{\partial L}{\partial X} \right)^T \cdot X \Big|_{X=\hat{X}, \Lambda=\hat{\Lambda}} = C + D\hat{X} - \hat{\Lambda}^T A \hat{X}^T = 0; \\ \hat{X} \geq 0; \\ \frac{\partial L}{\partial \Lambda} \Big|_{X=\hat{X}, \Lambda=\hat{\Lambda}} = B - A\hat{X} \geq 0; \\ \frac{\partial L}{\partial X} \Big|_{X=\hat{X}, \Lambda=\hat{\Lambda}} = C + D\hat{X} - \hat{\Lambda}^T A \leq 0; \\ \Lambda^T \frac{\partial L}{\partial \Lambda} \Big|_{X=\hat{X}, \Lambda=\hat{\Lambda}} = \hat{\Lambda}^T B - A\hat{X} = 0; \\ \hat{\Lambda} \geq 0. \end{array} \right.$$

Розглянемо модифікацію щодо формулювання умов Куна-Таккера.  
Введемо вектори

$$V^T = v_1, \dots, v_n \geq 0,$$

$$\Omega^T = (\omega_1, \dots, \omega_m) \geq 0,$$

компоненти яких обираються, виходячи із умов:

1) якщо  $\frac{\partial L}{\partial x_j} < 0$ , то  $v_j > 0$ ,

2) якщо  $\frac{\partial L}{\partial x_j} = 0$ , то  $v_j = 0$ ;

3) якщо  $\frac{\partial L}{\partial \lambda_i} < 0$ , то  $\omega_i > 0$ ,

4) якщо  $\frac{\partial L}{\partial \lambda_i} = 0$ , то  $\omega_i = 0 \quad j = \overline{1, n}; i = \overline{1, m}$ .

Тоді умови Куна-Таккера можливо записати у вигляді:

$$\begin{cases} V^T \cdot \hat{X} = 0, \\ B - A\hat{X} - \Omega = 0, \\ C + D\hat{X} - \hat{\Lambda}^T A + V^T = 0, \\ \Lambda^T \cdot \Omega = 0, \\ \hat{X} \geq 0_n, \\ \hat{\Lambda} \geq 0_m, \\ V \geq 0_n, \\ \Omega \geq 0_m. \end{cases} \quad (17.8)$$

**Теорема 17.2.** *Необхідна та достатня умова оптимального розв'язку задачі квадратичного програмування.*

Вектор  $\hat{X}$  є розв'язком задачі квадратичного програмування тоді і лише тоді, коли існують такі вектори:

- $m$ -вимірний вектор  $\hat{\Lambda} \geq 0$  та  $\Omega \geq 0$ ,
- $n$ -вимірний вектор  $V \geq 0$ ,

при яких виконуються вищезазначені лінійні рівняння:

$$\begin{cases} B - A\hat{X} - \Omega = 0; \\ C + D\hat{X} - \hat{\Lambda}^T A + V^T = 0 \end{cases}$$

та нелінійні рівняння:

$$\begin{cases} V^T \cdot \hat{X} = 0; \\ \Lambda^T \cdot \Omega = 0, \end{cases}$$

що носять назву умови *доповнюючої нежорсткості*.

Для розв'язання системи лінійних рівнянь (17.8) можна застосувати методи розв'язання задач лінійного програмування, наприклад, симплекс-метод.

Серед методів розв'язання задач квадратичного програмування найбільш відомим є метод Франка-Вульфа.

В якості прикладу аналітичного пошуку розв'язків задачі квадратичного програмування розглянемо задачу квадратичного програмування як одноетапну процедуру прийняття рішення в наступній постановці:

$$\begin{cases} W(X, U) = \frac{1}{2} U^T R U + \frac{1}{2} X^T Q X = \frac{1}{2} \|U\|_R^2 + \frac{1}{2} \|X\|_Q^2 \rightarrow \min_{U \in \mathbb{R}^m}; \\ AX + BU + C = 0, \end{cases} \quad (17.9)$$

де

$A$  – невироджена матриця  $n \times n$ ,

$B$  – матриця розміру  $n \times m$ ;

$X \in E^n$  –  $n$ -вимірний евклідів простір,

$C \in E^n$ ,

$0_n^T = [0_1, \dots, 0_n]$ ,

$R, Q$  – симетричні додатньо-визначені матриці розміру відповідно  $m \times m, n \times n$ .

Функція Лагранжа може бути записана у наступному вигляді:

$$L(X, U, \Lambda) = \frac{1}{2} U^T R U + \frac{1}{2} X^T Q X + \Lambda^T (AX + BU + C). \quad (17.10)$$

Для мінімізації показника ефективності необхідно, щоб:

$$\begin{cases} \left. \frac{\partial L}{\partial X} \right|_{X=\hat{X}, U=\hat{U}, \Lambda=\hat{\Lambda}} = Q\hat{X} + A^T \hat{\Lambda} = 0; \\ \left. \frac{\partial L}{\partial U} \right|_{X=\hat{X}, U=\hat{U}, \Lambda=\hat{\Lambda}} = R\hat{U} + B^T \hat{\Lambda} = 0; \\ A\hat{X} + B\hat{U} + C = 0. \end{cases}$$

Із перших двох рівнянь знаходимо:

$$\hat{X} = -Q^{-1} A^T \Lambda,$$

$$\hat{U} = -R^{-1} B^T \Lambda$$

та підставляємо у третє рівняння (залежатиме лише від  $\Lambda$ ):

$$-A Q^{-1} A^T \hat{\Lambda} - B R^{-1} B^T \hat{\Lambda} + C = 0 \Leftrightarrow \hat{\Lambda} = (A Q^{-1} A^T + B R^{-1} B^T)^{-1} C$$

Підставимо  $\hat{\Lambda}$  у вирази для  $\hat{X}$  і  $\hat{U}$ :

$$\hat{X} = -Q^{-1} A^T (A Q^{-1} A^T + B R^{-1} B^T)^{-1} \cdot C,$$

$$\hat{U} = -R^{-1}B^T \left( A Q^{-1} A^T + B R^{-1} B^T \right)^{-1} \cdot C.$$

Достатня умова того, що в точці  $\hat{X}, \hat{U}$  показник ефективності досягає мінімуму, полягає в тому, що друга варіація показника ефективності поблизу цієї точки більша 0:

$$\delta^2 W_{X,U} \Big|_{\substack{X=\hat{X} \\ U=\hat{U}}} > 0,$$

де

$$\delta^2 W_{X,U} \Big|_{\substack{X=\hat{X} \\ U=\hat{U}}} = \frac{1}{2!} \begin{bmatrix} \Delta X^T & \Delta U^T \end{bmatrix} \cdot \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \cdot \begin{bmatrix} \Delta X \\ \Delta U \end{bmatrix} \quad \text{при} \quad \begin{matrix} \Delta X \rightarrow 0 \\ \Delta U \rightarrow 0 \end{matrix}.$$

Враховуючи той факт, що:

$$\begin{cases} A \cdot \Delta X + B \cdot \Delta U = 0. \\ \Delta X \rightarrow 0 \\ \Delta U \rightarrow 0 \end{cases} \rightarrow \Delta X = -A^{-1} B \Delta U,$$

отримаємо:

$$\begin{aligned} \delta^2 W_{X,U} \Big|_{\substack{X=\hat{X} \\ U=\hat{U}}} &= \frac{1}{2!} \Delta U^T \cdot R \cdot \Delta U + \frac{1}{2} \left( -A^{-1} \cdot B \cdot \Delta U \right)^T \cdot Q \cdot \left( -A^{-1} \cdot B \cdot \Delta U \right) = \\ &= \frac{1}{2} \Delta U^T \left( R + B^T \cdot A^{-T} \cdot Q \cdot A^{-1} \cdot B \right) \Delta U > 0. \end{aligned}$$

Тобто, матриця в дужках повинна задовольняти ознаці додатньої визначеності.

### Контрольні запитання та завдання

1. У чому полягає суть задачі нелінійного програмування в некласичній постановці?
2. Поясніть процедуру зведення некласичної задачі нелінійного програмування до найпростішої некласичної задачі нелінійного програмування.
3. Наведіть та поясніть необхідні і достатні умови існування глобального максимуму задачі нелінійного програмування.
4. Розкрийте суть поняття «сідлова точка функції Лагранжа».

5. Сформулюйте теорему Куна–Таккера та наведіть приклади її застосування.
6. Наведіть та поясніть умови доповнювальної нежорсткості Куна–Таккера.
7. У чому полягає суть задачі квадратичного програмування?
8. Які методи застосовують для розв'язання задачі квадратичного програмування?
9. Наведіть та поясніть необхідні та достатні умови оптимального розв'язку задачі квадратичного програмування.

## Глава 18. Огляд основних підходів до побудови числових методів розв'язання задач нелінійного програмування

### 18.1. Модельно-тестовий приклад

В якості модельно-тестового прикладу, який будемо використовувати для демонстрації змісту прийомів та способів, що складають зміст більшості чисельних методів розв'язання задач нелінійного програмування, розглянемо показник ефективності вигляду:

$$W(x_1, x_2) \Big|_{\substack{x_1=a \\ x_2=b}} = a^2 + ab + b^2 \rightarrow \min_{\substack{a \in [-3;3] \\ b \in [-3;3]}} \quad (18.1)$$

Використаємо необхідні і достатні умови пошуку екстремуму функції двох змінних і знайдемо при яких значеннях параметрів  $a$  та  $b$  показник ефективності досягає мінімуму:

1) необхідна умова:

$$\begin{cases} \frac{\partial W(a,b)}{\partial a} = 0, \\ \frac{\partial W(a,b)}{\partial b} = 0, \end{cases} \Rightarrow \begin{cases} 2a + b = 0, \\ 2b + a = 0, \end{cases} \Rightarrow \begin{cases} \hat{a} = 0, \\ \hat{b} = 0. \end{cases} \quad (18.2)$$

2) достатня умова існування мінімуму в точці  $0;0$ : матриця Гессе повинна бути додатно-визначеною:

$$\begin{bmatrix} \frac{\partial^2 W}{\partial a^2} & \frac{\partial^2 W}{\partial a \partial b} \\ \frac{\partial^2 W}{\partial a \partial b} & \frac{\partial^2 W}{\partial b^2} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} > 0, \quad (18.3)$$

$$\Delta_1 = q_{11} = 2 > 0,$$

$\Delta_2 = 2 \cdot 2 - 1 \cdot 1 = 3 > 0$ , де  $\Delta_{1,2}$  визначники, що розташовані на головній діагоналі матриці Гессе.

**Висновок:**

- 1) Підтверджено, що матриця Гессе додатно-визначена.
- 2) Показник ефективності досягає мінімуму і його значення дорівнює 0:

$$\begin{cases} \hat{a} = 0, \\ \hat{b} = 0, \end{cases} \quad \hat{W} \hat{a}, \hat{b} = 0.$$

## 18.2. Метод сканування (метод повного перебирання)

Зазвичай цей метод використовують для вибору початкових умов для роботи числового алгоритму або в задачах цілочислового програмування для уточнення цілочислових значень змінних, за якими розв'язується задача, розташованих біля екстремуму показника ефективності, який було знайдено в умовах послабленої цілочислової задачі.

Для реалізації методу сканування необхідно виконати такі дії:

- 1) задати крок дискретизації змінних, за якими виконується пошук екстремуму;
- 2) побудувати область дискретизації змінних, за якими виконується пошук екстремуму.

Для того щоб не пропустити жодної точки, використовуються різні способи їх обходу (способи сканування вузлів). Наприклад, на рис.18.1 показано рядковий та спіральний (сірий колір) способи сканування.

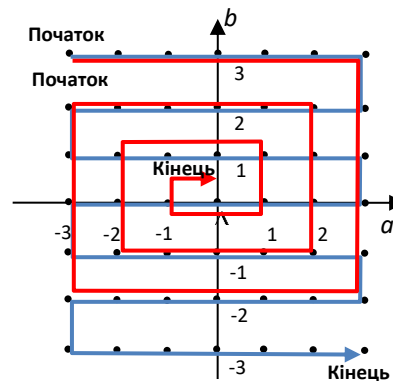


Рис. 18.1. Приклад побудови області дискретизації значень змінних, за якими виконується оптимізація показника ефективності  $W = a^2 + ab + b^2 \rightarrow \min \forall a, b \in [-3; 3]$  з вибором кроку дискретизації за параметрами  $\varepsilon_a = 1, \varepsilon_b = 1$

## 18.3. Метод Гаусса – Зейделя (метод покоординатного спуску)

Якщо показник ефективності  $W$  залежить від  $m$  змінних, то в цьому методі на кожному етапі(циклі) пошуку оптимальних значень параметрів припускають, що лише одна змінна впливає на значення показника ефективності, а інші залишаються незмінними (сталими). Після того, як буде

знайдено екстремальне значення показника ефективності за виділеною змінною, цю змінну вважають сталою із значенням, яке було знайдено при пошуку екстремуму показника ефективності за цією змінною. Покоординатний пошук екстремуму припиняється за відповідною умовою, яка може полягати в тому, що при повному виконанні етапів(циклів) покоординатної зміни за всіма змінними значення показника майже не змінюється.

Умови наближення до мінімуму показника ефективності за числової реалізації методу Гаусса–Зейделя показані на рис.18.2, де  $r$  – узагальнене позначення того параметру, від якого залежить показник ефективності і який на даному етапі (циклі) методу покоординатного спуску вважається змінюваним.

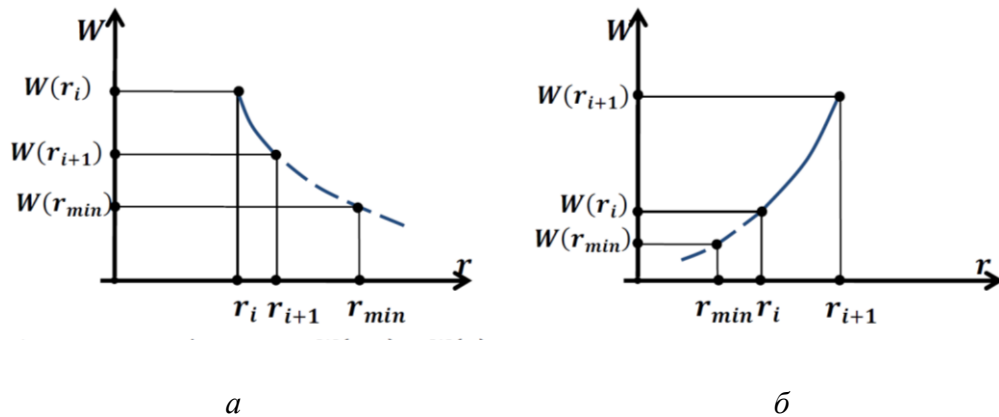


Рис. 18.2. Умови методів скалярної оптимізації (тобто пошуку екстремуму функції однієї змінної): а) якщо  $r_{i+1} > r_i$  і при цьому  $W(r_{i+1}) < W(r_i)$ , то точка мінімуму розташована праворуч від  $r_{i+1}$ ; б) якщо  $r_{i+1} > r_i$  і при цьому  $W(r_{i+1}) > W(r_i)$ , то точка мінімуму розташована ліворуч від  $r_i$

Існує велика кількість методів пошуку екстремуму функції однієї змінної, наприклад:

- 1) аналітичний (класичний) метод;
- 2) групи числових методів:
  - метод ділення відрізків навпіл;
  - метод золотого перерізу;
  - симетричні методи;
  - оптимальні методи;
  - метод ламаних;

- метод покриттів;
- метод дотичних;
- метод парабол тощо.

### 18.3.1. Класичний метод мінімізації функції однієї змінної у разі його застосування в методі Гаусса–Зейделя

Показник ефективності має вигляд:

$$W = a^2 + ab + b^2 \rightarrow \min ,$$

тобто вважається відомою залежність показника ефективності від змінних, за якими відбувається оптимізація.

**Етап I** (рис. 18.3)

$$\begin{cases} a_0 = 3, \\ b_0 = 3, \end{cases}$$

$$\begin{cases} b_0 = 3 = \text{const}, \\ a = \text{var}, \end{cases}$$

$$W(a, 3) = a^2 + 3a + 9,$$

$$\frac{dW}{da} = 2a + 3 = 0 \rightarrow a_1 = -\frac{3}{2}, b_1 = 3.$$

**Етап II**

$$\begin{cases} b_1 = \text{var}, \\ a_1 = -\frac{3}{2}, \end{cases}$$

$$W(-\frac{3}{2}, b) = \frac{9}{4} - \frac{3}{2}b + b^2,$$

$$\frac{dW}{db} = -\frac{3}{2} + 2b = 0 \rightarrow b_2 = \frac{3}{4} = 0,75, a_2 = -\frac{3}{2}.$$

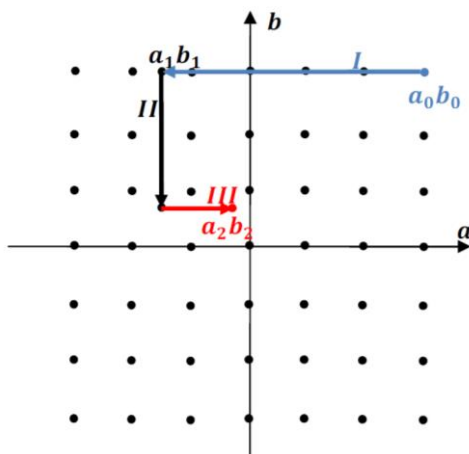


Рис. 18.3. Графічне зображення класичного методу пошуку екстремуму функції однієї змінної в методі Гаусса–Зейделя

### Етап III

$$\begin{cases} b_2 = 0,75, \\ a_2 = \text{var}, \end{cases}$$

$$W\left(a, \frac{3}{4}\right) = a^2 + \frac{3}{4}a + \frac{9}{16},$$

$$\frac{dW}{da} = 2a + \frac{3}{4} = 0 \rightarrow a_3 = -\frac{3}{8}$$

і т.д.

### 18.3.2. Метод екстраполяційного пошуку

Метод екстраполяційного пошуку ґрунтується на трьох припущеннях:

1) заздалегідь відоме бажане значення (область розташування бажаного значення показника ефективності):

$$W_{\epsilon} \in [W_0 - \epsilon_i; W_0 + \epsilon_i],$$

де  $W_0$  – певна величина, значення якої задають, базуючись на досвіді розв’язання подібних задач;  $\epsilon_i$  – допустима похибка обчислення ( $\epsilon_i > 0$ );  $i$  – номер етапу (циклу). Але при цьому невідома точка екстремуму показника ефективності, тобто значення змінних, за яких досягається  $W_{\epsilon}$ ;

2) значення показника ефективності може бути наближено представлено функціональною залежністю від змінних, за якими відбувається пошук екстремуму;

3) значення показника ефективності при заданому векторі змінних, за якими відбувається оптимізація, може бути виміряне або обчислене.

**Приклад 18.1.** Нехай залежність показника ефективності від змінного на даному циклі параметра є лінійною:

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1},$$

$$r_1 = r^i - \delta_i \rightarrow W r_1,$$

$$r_2 = r^i + \delta_i \rightarrow W r_2.$$

### Ришення

□ Тоді можемо записати:

$$\begin{aligned} \left( \frac{W(r^{i+1}) - W r_1}{W r_2 - W r_1} \right) &= \frac{r^{i+1} - r_1}{r_2 - r_1} \rightarrow r^{i+1} = \\ &= r_i - \delta_i + 2\delta_i \cdot \frac{W(r^{i+1}) - W r_1}{W r_2 - W r_1}, \end{aligned}$$

де беремо  $W(r^{i+1}) = W_0$ ,  $\delta_i > 0$ .

Після обчислення  $r^{i+1}$  вимірюємо або обчислюємо істинне значення показника ефективності  $W_{\text{іст}} r$  у точці  $r = r^{i+1}$ .

Якщо  $W_{\text{іст}}(r^{i+1}) \in [W_0 - \varepsilon_i; W_0 + \varepsilon_i]$ , то вважаємо, що  $r^{i+1} = r_{\text{ext}}$ .

Запам'ятовуємо це значення і переходимо до пошуку екстремуму показника ефективності за наступною координатою.

Якщо  $W_{\text{іст}}(r^{i+1}) \notin [W_0 - \varepsilon_i; W_0 + \varepsilon_i]$ , то повторюємо всі наведені вище

розрахунки за інших значень  $\delta_i$  та  $\varepsilon_i$ . ■

**Приклад 18.2.** Застосування методу екстраполяційного пошуку за умови квадратичної екстраполяції залежності показника ефективності від значень тієї змінної, за якою на даному циклі відбувається оптимізація.

Нехай  $W(r) = Ar^2 + Br + C$ ,  $A > 0$ .

**Ришення**

□ Якщо на  $i$ -му етапі виконати обчислення (вимірювання) значення показника ефективності, а саме:

$$W(r^i - \delta_i) = W_1,$$

$$W(r^i) = W_2,$$

$$W(r^i + \delta_i) = W_3,$$

то отримаємо систему трьох лінійних рівнянь для обчислення трьох невідомих параметрів  $A, B, C$ :

$$\begin{cases} A(r^i - \delta_i)^2 + B(r^i - \delta_i) + C = W_1, \\ A(r^i)^2 + B(r^i) + C = W_2, \\ A(r^i + \delta_i)^2 + B(r^i + \delta_i) = W_3. \end{cases}$$

Отже,

$$\begin{cases} A = \frac{\Delta A}{\Delta}, \\ B = \frac{\Delta B}{\Delta}, \\ C = \frac{\Delta C}{\Delta}, \end{cases}$$

$$\text{де } \Delta = \begin{vmatrix} (r^i - \delta_i)^2 & (r^i - \delta_i) & 1 \\ (r^i)^2 & (r^i) & 1 \\ (r^i + \delta_i)^2 & (r^i + \delta_i) & 1 \end{vmatrix}, \quad \Delta A = \begin{vmatrix} W_1 & (r^i - \delta_i) & 1 \\ W_2 & (r^i) & 1 \\ W_3 & (r^i + \delta_i) & 1 \end{vmatrix},$$

$$\Delta B = \begin{vmatrix} (r^i - \delta_i)^2 & W_1 & 1 \\ (r^i)^2 & W_2 & 1 \\ (r^i + \delta_i)^2 & W_3 & 1 \end{vmatrix}, \quad \Delta C = \begin{vmatrix} (r^i - \delta_i)^2 & (r^i - \delta_i) & W_1 \\ (r^i)^2 & (r^i) & W_2 \\ (r^i + \delta_i)^2 & (r^i + \delta_i) & W_3 \end{vmatrix}.$$

Мінімум квадратичної залежності досягається при  $r_{\text{ext}} = -B / 2A$ . ■

Знайдене значення змінної  $r_{\text{ext}}$ , за якою відбувається пошук екстремуму дає змогу обчислювати (вимірювати) істинне значення показника ефективності  $W_{\text{ICT}} r_{\text{ext}}$ .

Якщо  $W_{\text{ICT}} r_{\text{ext}} \in [W_0 - \varepsilon_i, W_0 + \varepsilon_i]$ , то вважаємо, що оптимізація за цією координатою завершена. Значення  $r_{\text{ext}}$  запам'ятовуємо і переходимо до пошуку екстремуму показника ефективності за такою координатою. Якщо  $W_{\text{ICT}} r_{\text{ext}}$  не належить заданій області, то повторюємо наведені вище обчислення, змінивши  $\delta_i$  або  $\varepsilon_i$ , або  $W_0$ .

#### 18.4. Метод градієнта

За умови використання методу градієнта на кожному етапі змінюються всі змінні, за якими відбувається пошук екстремуму. Припустимо, що таких змінних  $m$ . Показник ефективності має вигляд  $W = W(x_1, \dots, x_m)$ . Градієнтом диференційованої функції називається вектор, проекція якого на осі координат дорівнює  $\frac{\partial W}{\partial x_1}, \dots, \frac{\partial W}{\partial x_m}$  відповідно.

У методі градієнта використовується його властивість, яка полягає в тому, що модуль градієнта зменшується при наближенні до точки екстремуму (рис. 18.4):

$$\left| \frac{\partial W}{\partial X} \right|_{X_i} = |\text{grad } W|_i = \sqrt{\left( \frac{\partial W}{\partial x_1} \right)_i^2 + \dots + \left( \frac{\partial W}{\partial x_m} \right)_i^2} \rightarrow 0,$$

$$X_i^T = [x_1^{(i)}, \dots, x_m^{(i)}] \rightarrow X_{\text{ext}}.$$

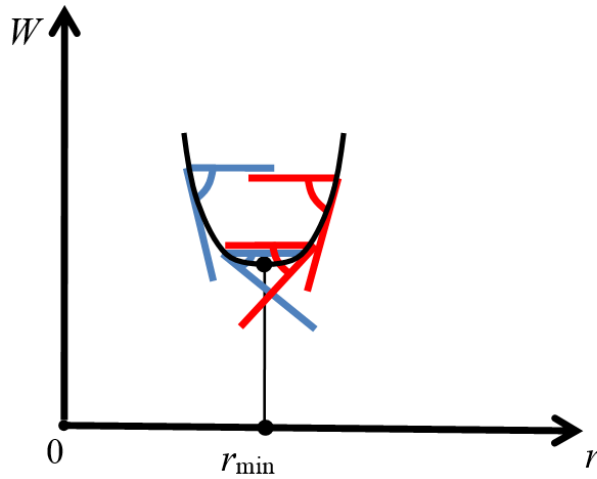


Рис. 18.4. Наближення модуля градієнта до 0 при наближенні аргумента до точки екстремуму показника ефективності

У загальному випадку, якщо невідома аналітична залежність показника ефективності від змінних, за якими відбувається оптимізація, значення градієнта в заданій точці  $X_i$  обчислюють наближено, тобто визначають субградієнт:

$$\begin{cases} \frac{\partial W}{\partial x_1^{(i)}} = \text{grad}_1^{(i)} W \cong \frac{1}{2\varepsilon_1} \left[ W \left( x_1^{(i)} + \varepsilon_1, x_2^{(i)} \dots x_m^{(i)} \right) - W \left( x_1^{(i)} - \varepsilon_1, x_2^{(i)} \dots x_m^{(i)} \right) \right], \\ \dots \\ \frac{\partial W}{\partial x_m^{(i)}} = \text{grad}_m^{(i)} W \cong \frac{1}{2\varepsilon_m} \left[ W \left( x_1^{(i)}, x_2^{(i)} \dots x_m^{(i)} + \varepsilon_m \right) - W \left( x_1^{(i)}, x_2^{(i)} \dots x_m^{(i)} - \varepsilon_m \right) \right], \end{cases}$$

де  $\varepsilon_j > 0$  ( $j = \overline{1, m}$ ) – пробний крок, використовуваний для зміни  $j$ -ї змінної на  $i$ -й ітерації.

Тобто:

$$\left( \overline{\text{grad}}^i W \right)^T = \left( \text{grad}_1^i W, \dots, \text{grad}_m^i W \right).$$

Для обчислення нового (наступного) наближення змінних, за якими відбувається пошук екстремуму, використовують вирази:

$$\begin{cases} x_1^{i+1} = x_1^i - \delta_1 \text{grad}_1^i W, \\ \dots \\ x_m^{i+1} = x_m^i - \delta_m \text{grad}_m^i W, \end{cases}$$

якщо відбувається пошук мінімального значення показника ефективності та

$$\begin{cases} x_1^{i+1} = x_1^i + \delta_1 \text{grad}_1^i W, \\ \dots \\ x_m^{i+1} = x_m^i + \delta_m \text{grad}_m^i W, \end{cases}$$

якщо відбувається пошук максимального значення показника ефективності, де  $\delta_j$  ( $j = \overline{1, m}$ ) – величина робочого кроку вздовж відповідної координати.

Отже, у методі градієнта процес пошуку екстремуму на кожній ітерації, тобто етапі (циклі) наближення до точки екстремуму, включає в себе дві обов'язкові операції (дії):

- пошук градієнта в поточній точці, коротка назва – пробний крок;
- пошук нового значення вектора змінних, коротка назва – робочий крок.

Графічне зображення траєкторії наближення до мінімального значення показника ефективності у разі використання методу градієнта показане на рис.18.5.

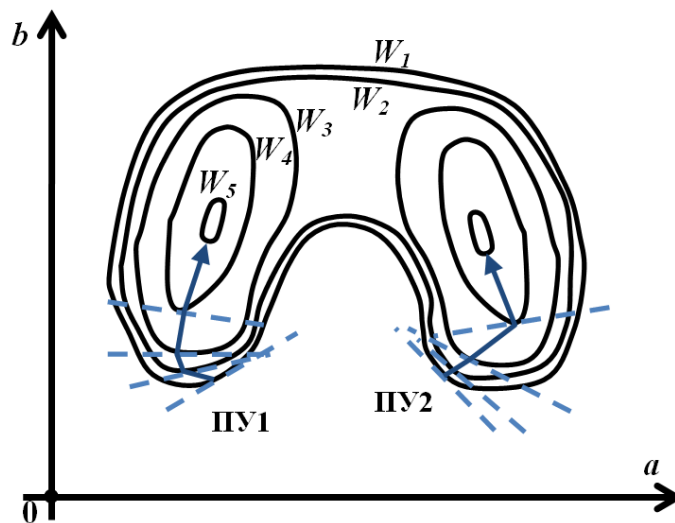


Рис. 18.5. Застосування методу градієнта для пошуку екстремуму багатомодової функції: ПУ – початкова умова

Метод градієнта дає змогу визначати локальний екстремум показника ефективності. Як бачимо з рис.18.5, результат пошуку локального екстремуму залежить від початкових умов.

**Приклад 18.3.** Припустімо, що показник ефективності має вигляд:

$$W = a^2 + ab + b^2 \rightarrow \min .$$

Розв'язати задачу методом градієнта (рис.18.6).

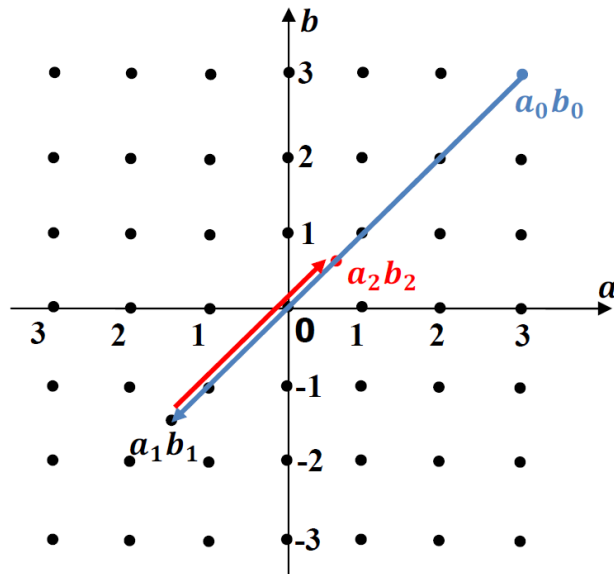


Рис. 18.6. Використання методу градієнта для пошуку екстремуму показника ефективності

### **Ришення**

□ 1. Перша ітерація. Початкові умови:  $a_0 = b_0 = 3$ .

1.1. Пробний крок:

$$\frac{\partial W}{\partial a} = 2a + b \Big|_{a_0; b_0} = 9;$$

$$\frac{\partial W}{\partial b} = 2b + a \Big|_{a_0; b_0} = 9.$$

1.2. Робочий крок:

$$\begin{cases} a_1 = a_0 - \delta \frac{\partial W}{\partial a} = -1,5; \\ b_1 = b_0 - \delta \frac{\partial W}{\partial b} = -1,5, \end{cases}$$

де  $\delta = 1/2$ .

2. Друга ітерація. Початкові умови:  $a_1 = b_1 = -1.5$ .

2.1. Пробний крок:

$$\frac{\partial W}{\partial a} = 2a + b \Big|_{a_1; b_1} = -4,5;$$

$$\frac{\partial W}{\partial b} = 2b + a \Big|_{a_1; b_1} = -4,5.$$

2.2. Робочий крок:

$$\begin{cases} a_2 = a_1 - \delta \frac{\partial W}{\partial a} = 0,75; \\ b_2 = b_1 - \delta \frac{\partial W}{\partial b} = 0,75. \end{cases}$$

Збіжність методу градієнта залежить як від початкових умов, так і від величини робочого кроку, тобто для підвищення ефективності методу градієнта (пошук локального екстремуму відбувається за меншу кількість кроків) необхідно цей метод модифікувати, надавши йому деяких оптимальних властивостей. ■

### 18.5. Метод найшвидшого спуску

Метод найшвидшого спуску – модифікація методу градієнта, суть якого полягає в тому, що після обчислення градієнта на  $i$ -й ітерації, переміщення вздовж цього градієнта відбувається від  $X^i \rightarrow X^{i+1}$  доти, поки зменшується значення показника ефективності, а не з фіксованим кроком, як це було в методі градієнта (рис. 18.7).

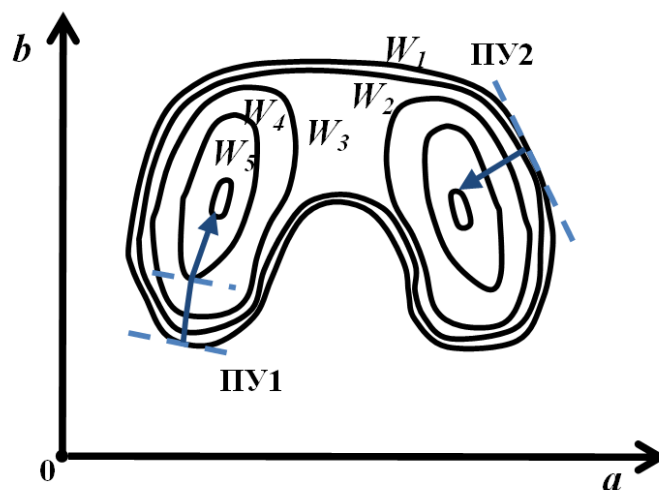


Рис. 18.7. Застосування методу найшвидшого спуску для пошуку екстремуму багатомодової функції

При використанні методу найшвидшого спуску зменшується кількість ітерацій і навіть можна досягти локального екстремуму за одну ітерацію за вдало обраних початкових умов (див. приклад 18.4 та рис. 18.8).

**Приклад 18.4.** Припустімо, що показник ефективності має вигляд:

$$W = a^2 + ab + b^2 \rightarrow \min .$$

Розв'язати задачу методом найшвидшого спуску.

1. Перша ітерація. Початкові умови:  $a_0 = b_0 = 3$ .

1.1 Пробний крок:

$$\frac{\partial W}{\partial a} = 2a + b \Big|_{a_0; b_0} = 9,$$

$$\frac{\partial W}{\partial b} = 2b + a \Big|_{a_0; b_0} = 9.$$

1.2. Робочий крок:

$$a_1 = a_0 - 9\delta = 3 - 9\delta;$$

$$b_1 = b_0 - 9\delta = 3 - 9\delta,$$

$$W(\delta) = 3(3 - 9\delta)^2,$$

$$\frac{\partial W}{\partial \delta} = -54(3 - 9\delta) = 0 \rightarrow \delta = \frac{1}{3}.$$

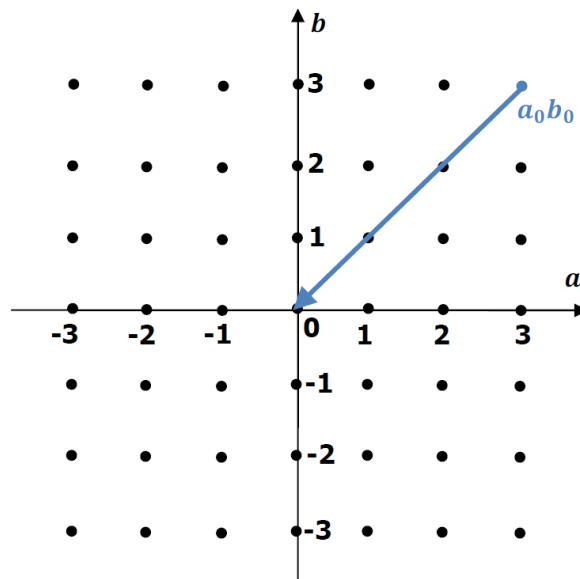


Рис. 18.8. Застосування методу найшвидшого спуску для пошуку мінімуму показника якості

Остаточно маємо:

$$a_1 = a_0 - 9\delta = 3 - 9\delta = 3 - 9 \cdot \frac{1}{3} = 0;$$

$$b_1 = b_0 - 9\delta = 3 - 9\delta = 3 - 9 \cdot \frac{1}{3} = 0.$$

## 18.6. Метод випадкового пошуку

В основу випадкового пошуку покладено метод випробувань і помилок, відповідно до якого вдало знайдене значення керувальних параметрів зберігається, а невдале – відкидається.

Метод випадкового пошуку використовують для пошуку гло-бальних і локальних екстремумів показника ефективності, причому як самостійно, так і комбіновано з методами регулярного (детермінованого) пошуку.

### Класифікація методів випадкового пошуку:

- 1) локальний випадковий пошук з поверненням;
- 2) локальний випадковий пошук за найкращою пробою;
- 3) локальний випадковий пошук за статистичним градієнтом;
- 4) глобальний випадковий пошук з незалежним вибором щільності розподілу пробних кроків;
- 5) генетичні алгоритми та еволюційне програмування.

Розглянемо як приклад основні елементи алгоритму локального випадкового пошуку (локального випадкового пошуку з поверненням, який на сьогодні є найпоширенішим). Алгоритм полягає в такому:

**Крок 1.** У випадковому напрямі здійснюється зміна значень параметрів, за якими відбувається пошук екстремуму:

$$X^{i+1} = X^i + \Delta X^{i+1},$$

де  $\Delta X^{i+1} = \delta E_{i+1}^T$ ;  $E_{i+1}^T = [e_1, \dots, e_n]_{i+1}$  – випадковий одиничний вектор  $e_1^2 + e_2^2 + \dots + e_n^2 = 1$  із заданою щільністю ймовірності, реалізація значень компонент якого отримана для обчислення  $\Delta X^{i+1}$ ;  $\delta$  – величина кроку за керувальними параметрами.

Якщо вимірювання (обчислення) показало, що показник ефективності в новій точці менший ніж значення показника ефективності у попередній точці (відбувається пошук мінімуму), тобто  $W(X^{i+1}) < W(X^i)$ , то значення показника ефективності  $W(X^{i+1})$  та  $X^{i+1}$  запам'ятовуються.

Якщо  $W\left(X^{i+1}\right) \geq W\left(X^i\right)$ , то виконується повернення до попередніх значень показника ефективності та керувальної змінної і знову обчислюється

$$X^{i+2} = X^i + \Delta X^{i+2},$$

де  $\Delta X^{i+2} = \delta E_{i+2}$ ,  $E_{i+2}^T = [e_1, \dots, e_n]_{i+2}$  – випадковий одиничний вектор, реалізація значень компонент якого отримана для обчислення  $\Delta X^{i+2}$ .

Обчислення виконується доти, поки показник ефективності стане меншим, ніж його попереднє значення. Останнє значення вектора керувальних змінних та значення показника ефективності запам'ятовується.

**Крок 2.** Виконується у випадковому напрямі від останнього значення вектора керувальних параметрів, яке було запам'ятоване наприкінці кроку 1.

У подальшому кроки 1 і 2 повторюються доти, поки не буде виконана умова зупинки алгоритму пошуку екстремального значення показника ефективності.

## 18.7. Застосування системи комп'ютерної математики MATLAB для розв'язка задач квадратичного та нелінійного програмування

*Квадратичне програмування: функція **quadprog***

Формат постановки задачі квадратичного програмування в MATLAB:

$$\begin{cases} W = \frac{1}{2} X^T \cdot H \cdot X + f \cdot X \rightarrow \min_{X \in \mathbb{R}^n}, \\ A \cdot X \leq b, \\ A_{eq} \cdot X = b_{eq}, \\ L_b \leq X \leq U_b. \end{cases}$$

Формат звернення до функції **quadprog**:

$[X, W, \text{flag}] = \mathbf{quadprog}(H, f, A, b, A_{eq}, b_{eq}, L_b, U_b)$ .

**Приклад 18.5.** Розглянемо задачу рівня бізнес-управління телекомунікаційною компанією. Перед ТОП-менеджерами компанії поставлено завдання щодо розподілу інноваційних ресурсів компанії на розвиток найприбутковіших технологій, тобто визначення так званого

інноваційного портфелю. Через невизначеність майбутнього запиту ринку цей портфель є ризикованим.

Для зменшення ризику втрати наявних інноваційних ресурсів необхідно виконати їх розбиття, наприклад, на чотири частини  $x_1, x_2, x_3, x_4$  для того, щоб розвивати, відповідно, чотири інноваційні технології  $A_1, A_2, A_3, A_4$ .

Позначимо

$$X^T = [x_1, x_2, x_3, x_4],$$

де  $x_i$  ( $i = \overline{1,4}$ ) відносні витрати (частина від загальної інноваційної суми):

$$x_1 + x_2 + x_3 + x_4 = 1, \quad x_{1,2,3,4} \geq 0.$$

Таким чином, інноваційний портфель визначається вектором  $X$  частинок від загальної суми інноваційних ресурсів. Ризик оцінюється як величина дисперсії доходу, що очікується. Уважаємо, що за результатом аналізу тенденцій розвитку ринку ІТ-технологій ідентифіковано вектор, що складається з математичних сподівань прибутковості чотирьох обраних для розвитку технологій

$$d^T = [d_1, d_2, d_3, d_4]$$

та матриці взаємних дисперсій  $V$ .

Тоді математичне сподівання прибутковості інноваційного портфелю можна обчислити так:

$$M_p = d^T X.$$

Припустімо, що:

$$d = [11.3; 13.2; 16.4; 17.4],$$

$$V = \begin{bmatrix} 102,0 & 27,1 & -52,3 & 66,5 \\ 27,1 & 148,8 & 42,1 & -66,4 \\ -52,3 & 42,1 & 246,5 & 56,9 \\ 66,5 & -66,4 & 56,9 & 272,3 \end{bmatrix}.$$

Дисперсія відхилення від прибутковості, що очікується ( $M_{p0} = 15$ ), обчислюється за формулою:

$$D = X^T V X.$$

Вираз для обчислення дисперсії  $D$  використовують для оцінки ризику інноваційного портфеля. Цей ризик потрібно мінімізувати за рахунок оптимального вибору компонент вектора  $X$  (задача Марковица).

Для наведеного прикладу матриці, що відповідають загальній постановці задачі, набувають вигляду:

$$H = V;$$

$$f - \text{відсутня, [ ]};$$

$$A - \text{відсутня, [ ]};$$

$$B - \text{відсутня, [ ]};$$

$$A_{eq} = \begin{bmatrix} d_1 & d_2 & d_3 & d_4 \\ 1 & 1 & 1 & 1 \end{bmatrix};$$

$$b_{eq} = \begin{bmatrix} 15 \\ 1 \end{bmatrix};$$

$$lb = [0, \dots, 0];$$

$$Ub - \text{відсутня, [ ]}.$$

Звернення до функції мінімізації квадратичного показника ефективності для нашого випадку:

$$[X, \min W, flag] = \text{quadprog}(V, [], [], [], A_{eq}, b_{eq}, lb, []),$$

де  $flag$  – символ якості виконання оптимізації: якщо  $flag$  дорівнює цілому додатному числу, то операція оптимізації виконана успішно; якщо  $flag = 0$ , то виконана максимальна кількість ітерацій без покращення результату (ознака зациклювання алгоритму); якщо  $flag$  дорівнює цілому від'ємному числу, то досягти екстремуму не вдалося – операція пошуку екстремуму не виконана.

Командне віконце MATLAB:

```
>> V=[102.0 27.1 -52.3 66.5
```

```
27.1 148.8 42.1 -66.4
```

```
-52.3 42.1 246.5 56.9
```

```
66.5 -66.4 56.9 272.3]
```

```
>> Aeq=[11.3 13.2 16.1 17.4
```

```
1 1 1 1]
```

```
>> beq=[15;1]
```

```
>> lb=[0;0;0;0]
```

```
>> [X,W,flag]=quadprog(V,[],[],[],Aeq,beq,lb,[])
```

$X = 0.0626$

0.4359

0.1439

0.3575

$\min W = 31.2725$

$\text{flag} = 1.$

*Нелінійне програмування: функція **fmincon***

Формат постановки задачі нелінійного програмування в MATLAB:

$$f(x) \rightarrow \min_{X \in \mathbb{R}^n},$$

$$\begin{cases} AX \leq b, \\ A_{eq} X = b_{eq}, \\ lb \leq X \leq Ub, \\ C X \leq 0, \\ C_{eq} X = 0. \end{cases}$$

Формат звернення до функції **fmincon**:

$$\begin{aligned} & [X, \min W, \text{flag}] = \\ & = \text{fmincon}(\text{fun}, X_0, A, b, A_{eq}, b_{eq}, lb, Ub, \text{nonlcon}, \text{options}, p_1, p_2), \end{aligned}$$

де  $X$  – вектор змінних, за яким відбувається оптимізація (керувальні змінні);  $W_{\min}$  – мінімальне значення показника ефективності;  $\text{flag}$  – символ якості виконання оптимізації (має той самий зміст, що й у випадку квадратичного програмування);  $\text{fun}$  – ім'я файлу-функції, яка обчислює значення показника ефективності (може залежати від кількох параметрів  $p_1$ ,  $p_2$ , значення яких передається за допомогою аргументів  $\text{fmincon}$ , що починаються з одинадцятої позиції у її вхідному списку; аргументом  $\text{fmincon}$  є вектор  $X$ ); вектори і матриці, які не використовуються, позначаються як [];  $X_0$  – початкове значення вектора  $X$ ;  $\text{nonlcon}$  – ім'я файлу-функції нелінійних обмежень (вхідним аргументом є вектор  $X$ , а вихідні аргументи – ліві частини обмежень);  $\text{options}$  – керувальна структура, яка задає властивості обчислювального алгоритму.

### **Приклад 18.6**

Постановка задачі

$$W=3x_1^2 + 2x_2^2 \rightarrow \min ,$$
$$x_1^2 + x_2^2 \leq 1 .$$

Розв'язання

Побудова файлу-функції: File\_New\_M-file

```
function fp1= myfunp1(x)
```

```
fp1=3·x(1)^2+2·x(2)^2;
```

Save as...myfunp1

Побудова файлу-функції: File\_New\_M-file

```
function [c, ceq]=myconp1(x)
```

```
c(1)=x(1)^2+x(2)^2-1;
```

```
ceq=[];
```

Save as...myconp1

```
>> [x, fp1, flag]=
```

```
=fmincon(@myfunp1,[0.7 0.7],[[],[],[],[],[],[],@myconp1)
```

```
x = 1.0e-004·(0.1895 -0.0235)
```

```
fp1 = 1.0882e-009
```

```
flag = 1
```

### **Приклад 18.7**

Постановка задачі

$$W=(x_1 - 5)^2 + (x_2 - 8)^2 \rightarrow \min(\max) ,$$

$$x_1 + 2x_2 \geq 4 ,$$

$$5/2x_1 + 2x_2 \leq 20 ,$$

$$x_{1,2} \geq 0 .$$

Розв'язання

Побудова файлу-функції: File\_New\_M-file

```
function fp2min= myfunp2min(x)
```

```
fp2min=(x(1)-5)^2+(x(2)-8)^2;
```

Save as...myfunp2min

```

>> A=[-1 -2 2.5 2]
>> b=[-4;20]
>> [x, fp2min, flag]=fmincon(@myfunp2min,[0 5],[A],[b],[[],[]],[0 0],[[]])
x = 2.9268 6.3415
fp2min = 7.0488
flag = 1
Побудова файлу-функції: File_New_M-file
function fp2max= myfunp2max(x)
fp2max=-((x(1)-5)^2+(x(2)-8)^2);
Save as...myfunp2max
>>[x, fp2max, flag]=fmincon(@myfunp2max,[0 5],[A],[b],[[],[]],[0 0],[[]])
x = 0 2
fp2max = -61
flag = 1

```

### **Приклад 18.8**

Постановка задачі

$$W=x_1^2 - 2x_1 + x_2^2 \rightarrow \min(\max),$$

$$x_1 + 3/2x_2 \leq 15/2,$$

$$0 \leq x_1 \leq 10,$$

$$0 \leq x_2 \leq 9,$$

$$x_1x_2 \leq 6.$$

Розв'язання

Побудова файлу-функції: File\_New\_M-file

```
function fp3min= myfunp3min(x)
```

```
fp3min=x(1)^2+x(2)^2-2·x(1);
```

Save as...myfunp3min

```
>> A=[1 1.5]
```

```
A = 1.0000 1.5000
```

```
>> b=[7.5]
```

```
b = 7.5000
```

Побудова файлу-функції: File\_New\_M-file

```
function [c, seq]=myconp3(x)
```

```

c(1)=x(1)·x(2)-6;
ceq=[];
Save as...myconp3
>> [x, fp3min, flag]=fmincon(@myfunp3min,[0 0],[A],[b],[[],[]],[0 0],[10 9],
@myconp3)
x = 1.0000      0
fp3min = -1
flag = 1
Побудова файлу-функції:File_New_M-file
function fp3max= myfunp3max(x)
fp3max=-(x(1)^2+x(2)^2-2·x(1));
Save as...myfunp3max
>>[x, fp3max, flag]=fmincon(@myfunp3max,[0 0],[A],[b],[[],[]],[0 0],[10 9],
@myconp3)
x = 1.0e-007·( 0   0.1000)
fp3max = -1.0000e-016
flag = 1
>> [x, fp3max, flag]=fmincon(@myfunp3max,[6 0],[A],[b],[[],[]],[0 0],[10 9],
@myconp3)
x = 7.5000      0
fp3max = -41.2500
flag = 1

```

### ***Приклад 18.9***

Постановка задачі

$$W = -x_1^2 + 4x_1 + x_2 \rightarrow \min (\max) ,$$

$$x_1 x_2 \geq 6,$$

$$x_1^2 + x_2^2 \leq 49 ,$$

$$0 \leq x_1 ,$$

$$0 \leq x_2 \leq 6.$$

Розв'язання

Побудова файл-функції:File\_New\_M-file

```
function fp4min= myfunp4min(x)
```

```

fp4min=-x(1)^2+4·x(1)+x(2);
Save as...myfunp4min
Побудова файл-функції:File_New_M-file
function [c, seq]=myconp4(x)
c(1)=-x(1)*x(2)+6;
c(2)=x(1)^2+x(2)^2-49;
seq=[];
Save as...myconp4
>> [x, fp4min, flag]=fmincon(@myfunp4min,[3 4],[],[],[],[0 0],[10 6],
@myconp4)
x = 6.9465    0.8637
fp4min = -19.6042
flag = 1
Побудова файл-функції:File_New_M-file
function fp4max= myfunp4max(x)
fp4max=-(-x(1)^2+4·x(1)+x(2));
Save as...myfunp4max
>>[x, fp4max, flag]=fmincon(@myfunp4max,[3 4],[],[],[0 0],[10 6],
@myconp4)
x = 2.0000    6.0000
fp4max = -10.0000
flag = 1

```

*Пошук глобального максимуму (мінімуму) поліноміальної функції (рис. 18.9, 18.10)*

```

Побудова файлу-функції: File_New_M-file
function fP=myfunP(x)
fP=4·sin(2·pi·x(1))·cos(1.5·pi·x(2))·(1-x(1)^2)·x(2)·(1-x(2));
Save as...myfunP
>> [x1,x2]=meshgrid(-1:0.05:1,0:0.05:1);
>> fP=4·sin(2·pi·x1)·cos(1.5·pi·x2)·(1-x1.^2)·x2·(1-x2);
>> mesh(x1,x2,fP)

```

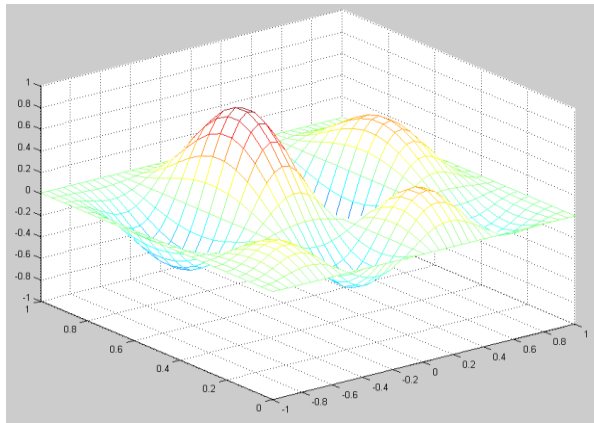


Рис. 18.9. Графічний образ полімодальної функції

```
>> [CMatr,h]=contour(x1,x2,fP);
>> clabel(CMatr,h)
>> grid on
```

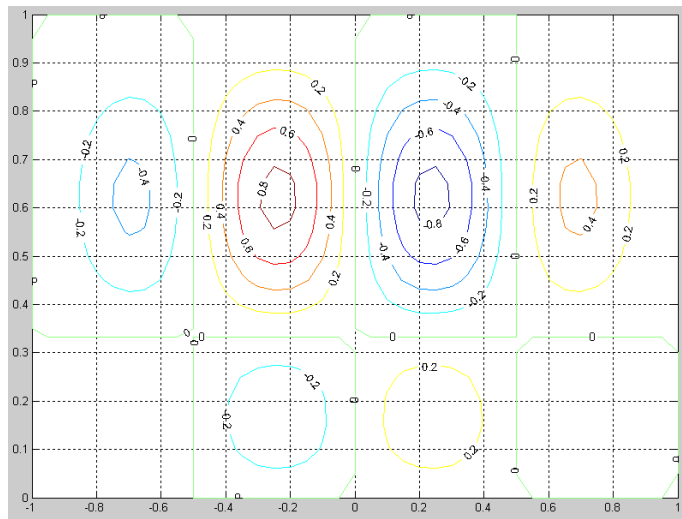


Рис. 18.10. Графічний образ ліній однакового рівня полімодальної функції

```
Початкові умови [-1; 0]
>> [X,fP,flag]=fmincon(@myfunP,-1;0,[],[],[],[],[-1;0],[1;1])
X =  -1
      0
fP = 0
flag = 1
Початкові умови [0.2;0.6]
>>[X,fP,flag]=fmincon(@myfunP,[0.2;0.6],[],[],[],[],[-1;0],[1;1])
X =  0.2373
      0.6211
```

```

fP = -0.8652
flag = 1
Початкові умови [-0.6;0.5]
>> [X,fP,flag]=fmincon(@myfunP,[-0.6;0.5],[[],[],[],[],[],[-1;0],[1;1])
X = -0.6874
      0.6211
fP = -0.4481
flag = 1

```

*Пошук глобального максимуму (мінімуму) функції Розенброка (рис.18.11, 18.12)*

```

Побудова файлу-функції: File_New_M-file
function fR=myfunR(x)
fR=100·(x(2)-x(1)^2)^2+(1-x(1))^2;
Save as...myfunR
>> [x1,x2]=meshgrid(-2:0.05:2,-2:0.05:2);
>> fR=100·(x2-x1.^2).^2+(1-x1).^2;
>> mesh(x1,x2,fR)

```

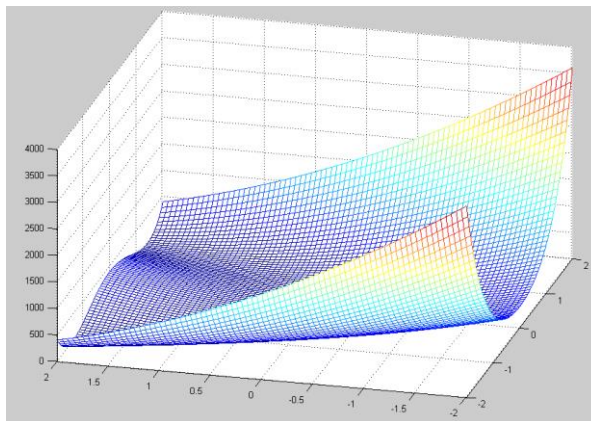


Рис. 18.11. Графічний образ функції Розенброка

```

>> [CMatr,h]=contour(x1,x2,fR);
>> clabel(CMatr,h)
>> grid on

>> p=[0.5 2 10 50 100 150 200 250 300 350 400];
>> [CMatr,h]=contour(x1,x2,fR,p);
>> clabel(CMatr,h)

```

```
>> grid on
>> [X,fR,flag]=fmincon(@myfunR,[-1;-1],[[],[],[],[],[-2;-2], [2;2])
X =  1.0000
      1.0001
fR = 9.7522e-009
flag = 1
```

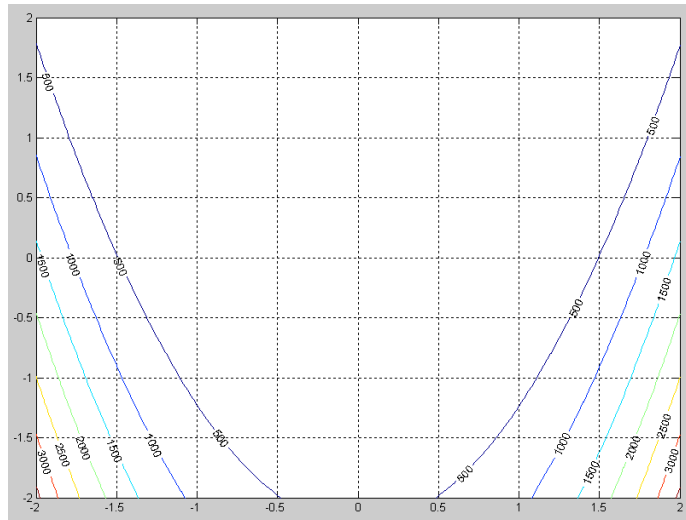


Рис. 18.12. Графічне зображення ліній рівного рівня функції Розенброка

*Пошук глобального максимуму (мінімуму) функції Вуда*

Побудова файлу-функції: File\_New\_M-file

```
function fB=myfunB(x)
```

```
f1=100·(x(2)-x(1)^2)^2+(1-x(1))^2+90·(x(4)-x(3)^2)^2+(1-x(3))^2;
```

```
f2=10.1·((x(2)-1)^2+(x(4)-1)^2)+19.8·(x(2)-1)·(x(4)-1);
```

```
fB=f1+f2;
```

Save as...myfunB

Початкові умови [3;3;3;3]

```
>> [X,fB,flag]=fmincon(@myfunB,[3;3;3;3],[[],[],[],[],[-3;-3;-3;-3],[3;3;3;3])
```

```
X =  1.0000
```

```
      1.0000
```

```
      1.0000
```

```
      1.0000
```

```
fB = 4.2696e-009
```

```
flag = 1
```

Початкові умови [-3;3;3;3]

```
>> [X,fB,flag]=fmincon(@myfunB,-3;3;3;3,[],[],[],[],[-3;-3;-3;-3],[3;3;3;3])
```

X = 1.0000

1.0000

1.0000

1.0000

fB = 3.4980e-010

flag = 1

### Контрольні запитання та завдання

1. Наведіть та поясніть основні методи розв'язання задач нелінійного програмування.
2. У чому полягає суть методу екстраполяційного пошуку?
3. Дайте визначення градієнта диференційованої функції.
4. Наведіть та поясніть геометричну інтерпретацію використання методу градієнта для пошуку екстремуму показника ефективності.
5. Наведіть та поясніть геометричну інтерпретацію застосування методу найшвидшого спуску для пошуку екстремуму багатомодової функції.
6. У чому полягає суть методу випадкового пошуку глобальних і локальних екстремумів показника ефективності?
7. Поясніть призначення функції quadprog системи комп'ютерної математики MATLAB.

## **Розділ 2. Еволюційні методи оптимізації та підходи до їх вдосконалення**

### **Глава 19. Еволюційні методи оптимізації**

Еволюційні методи оптимізації сформували окремий напрямок розвитку методів розв'язання оптимізаційних задач. До нього, зокрема, належать методи, в яких закладені принципи природних механізмів, що забезпечують ефективну адаптацію природних систем до навколишнього середовища. Основу цього напрямку сформували роботи Дж. Холанда, Л. Фогеля, І. Букатової, Л. Растрігіна та інших дослідників, які звернули увагу на процеси, які відбуваються в природі, надали їх формалізований опис та побудували деякі моделі еволюційного розвитку.

Необхідно зазначити, що еволюційний напрямок розв'язання задач оптимізації сформувався відносно недавно. Його появу пов'язують з публікацією Дж. Холандом в 1975 році роботи, присвяченої генетичним алгоритмам. З тих пір у світі спостерігається стрімке збільшення публікацій, присвячених застосуванню еволюційних методів для оптимізації складних систем, а також продовжуються роботи щодо створення нових та удосконалення існуючих методів цього класу.

Загальний інтерес до таких методів можна пояснити такими основними обставинами.

По-перше, стрімке зростання складності практичних задач, які підлягають розв'язанню, та жорсткі обмеження за часом, відведеним на отримання результатів. При цьому, внаслідок загального усвідомлення важливості цих факторів досить часто вважається за доцільне отримувати не оптимальні, а раціональні результати, які задовольняють практичним потребам.

По-друге, властивості, притаманні самим методам: універсальність, ефективність, відсутність спеціальних вимог до виду цільової функції і обмежень, а також можливість оптимізації складних систем з безпосереднім використанням моделей їх функціонування.

По-третє, досить висока обчислювальна спроможність загальнодоступної обчислювальної техніки та її постійне зростання, без якої ефективно застосування еволюційних методів оптимізації практично неможливе.

Водночас на сьогодні вважається, що еволюційні методи оптимізації можуть бути створені лише на основі аналізу процесів, які відбуваються в природі, тобто створення нового методу такого класу без деякого прообразу природного походження неможливе.

Окрім того, нині відсутній узгоджений погляд на сутність визначення „еволюційний метод” як загального поняття, та, відповідно, на сутність терміна „еволюційний метод оптимізації”. Для нівелювання такої невизначеності досить широко почали використовувати термін „біоінспіровані методи”, однак і він не охоплює усіх методів, які мають природні прототипи, та не дає однозначної відповіді на питання щодо фундаментальних механізмів, покладених у їх основу.

Така невизначеність у понятійному апараті та обов’язкова прив’язка до природних процесів не дають змоги надати відповідь на запитання, чим же еволюційні методи оптимізації принципово відрізняються від методів математичного програмування, хоча відповідь на нього може відкрити нові шляхи для вдосконалення існуючих та розроблення нових ефективних оптимізаційних методів.

Тобто нині існує методологічна проблема, пов’язана із розробленням теоретичних основ еволюційних методів оптимізації, визначенням їх зв’язку з фундаментальними науковими дисциплінами як підґрунтя для створення нових методів цього класу та удосконалення існуючих.

У цьому розділі розглянуті найбільш поширені сучасні методи оптимізації, які мають природні прототипи. При цьому наявність досить великої кількості варіантів реалізації окремих методів надає підстави розглядати наведені базові алгоритми як частину загальної методології розв’язання задач оптимізації.

Спільною рисою всіх розглядуваних методів є те, що вони призначені для розв’язання задач оптимізації без обмежень. Тому перед їх використанням розглядуване завдання необхідно звести до відповідної постановки оптимізаційної задачі, наприклад, з використанням методу штрафних функцій.

## 19.1. Генетичні алгоритми

Генетичні алгоритми (ГА) базуються на ідеї Ч. Дарвіна стосовно еволюційного розвитку біологічних видів, відповідно до якої у природі покращення ознак деякої популяції тварин від покоління до покоління відбувається завдяки трьом основним факторам:

- спадковості, коли нащадки успадковують свої характеристики від батьків;
- відбору, коли більше шансів вижити та залишити нащадків мають особини з кращими характеристиками;
- змінюваності, коли характеристики нащадків можуть змінюватись унаслідок впливу зовнішніх умов.

Для практичної реалізації впливу цих факторів користуються такими модельними уявленнями.

Відповідно до генетики будь-яка особина характеризується генотипом та фенотипом. Генетична інформація закодована в хромосомі, яка є послідовністю генів. Кожен ген може приймати деяке значення з множини припустимих. Тоді як генотип розглядається хромосома з визначеними значеннями генів.

Під фенотипом розуміють множину ознак особини, яку може оцінити зовнішній спостерігач. Загалом вважається, що кращий фенотип зумовлений кращим генотипом, тобто кращим генетичним матеріалом. Особини, які мають кращий фенотип, є кращими щодо їх пристосованості до виживання у визначених умовах та можливості залишити нащадків.

Окрім того, враховується, що під дією оточуючого середовища генотип особини може змінюватись, тобто деякі гени можуть поміняти своє поточне значення на інше (з множини припустимих).

Отже, спадковість забезпечується тим, що хромосома нащадка формується з частин хромосом батьків, відбір – тим, що в розмноженні беруть участь особини з кращим фенотипом, а змінюваність – тим, що окремі гени в хромосомі змінюють своє значення.

Таким чином, ГА оперують популяцією (множиною) особин, фенотип кожної з яких змінюється залежно від зміни її генотипу, а покращення фенотипів від покоління до покоління забезпечується за рахунок формування кращих генотипів.

До початку розрахунків визначають систему кодування змінних, яка дозволить подати сукупність незалежних змінних у вигляді хромосоми, перетворення якої за визначеними правилами дозволить генерувати множину розв'язків задачі оптимізації. Хромосома складається з генів, кожен з яких містить закодоване значення однієї незалежної змінної. Якщо кожен ген може приймати тільки одне з двох значень (0 або 1), тоді маємо діалельний варіант ГА, в іншому разі – поліалельний варіант.

Як характеристику фенотипу використовують функцію пристосованості (фітнес-функцію) – значення цільової функції, отриманої зведенням вихідної задачі оптимізації до вигляду, придатного для застосування ГА.

Для забезпечення спадковості та змінюваності використовують відповідні генетичні оператори, найпоширенішими з яких є кросовер (схрещування), мутація та інверсія.

При цьому кросовер формує хромосому нащадка шляхом комбінації ділянок хромосом батьків, мутація змінює значення випадково обраного гена шляхом випадкового вибору будь-якого іншого припустимого для цього гена значення, а інверсія міняє місцями значення двох випадково обраних генів.

Для діалельного ГА схема одноточкового кросовера наведена на рис. 19.1 (точка кросовера між другим та третім генами), мутації – на рис. 19.2 (мутує п'ятий ген), а інверсії – на рис. 19.3 (п'ятий та четвертий ген обмінюються значеннями).

Батьківська хромосома 1	0	0	1	1	0
Батьківська хромосома 2	1	1	0	0	0
Хромосома нащадка 1	0	0	0	0	0
Хромосома нащадка 2	1	1	1	1	0

Рис. 19.1. Схема одноточкового кросовера для діалельного ГА

Хромосома до мутації	1	1	1	1	0
Хромосома після мутації	1	1	1	1	1

Рис. 19.2. Схема мутації для діалельного ГА

Хромосома до інверсії	1	1	1	1	0
Хромосома після інверсії	1	1	1	0	1

Рис. 19.3. Схема інверсії для діалельного ГА

Для забезпечення відбору використовують правила, на основі яких окремі особини можуть бути обрані батьками для формування нащадка шляхом реалізації оператора кросовера (схрещування). Відповідно розрізняють види відбору: випадковий, турнірний, пропорційний тощо.

Випадковий відбір передбачає, що номери особин визначають з використанням рівномірного генератора випадкових чисел.

Турнірний відбір передбачає, що для визначення кожного з батьків організується турнір (змагання) між декількома випадково обраними особинами. Переможцем турніру вважається особина з найкращим значенням функції пристосованості.

Пропорційний відбір передбачає, що імовірність прийняти участь у формуванні нового покоління пропорційна перевищенню функцією пристосованості особини середнього значення пристосованості по популяції.

Схема розв'язання задачі оптимізації з використанням одного з варіантів ГА наведена на рис. 19.4. Особливістю реалізації цього варіанта ГА є те, що кожна особина може безпосередньо потрапити в нове покоління (не беручи участі в формуванні нащадка), а також використання лише одного з генетичних операторів.

На початку розрахунків створюють вихідну множину хромосом відповідно до визначеної кількості особин у популяції ( $M$ ) та обчислюють значення функції пристосованості для кожної з них.

Далі реалізуються два цикли: перший – за кількістю особин у популяції, а другий – за кількістю поколінь. При цьому перший цикл є вкладеним у другий.

Перший цикл реалізує формування нового покоління. На кожному кроці циклу пропорційним відбором з поточного покоління обирається деяка особина, яка (з деякою імовірністю) може бути або скопійована в нове покоління, або до неї може бути застосований один з генетичних операторів (кросовер, мутація або інверсія).

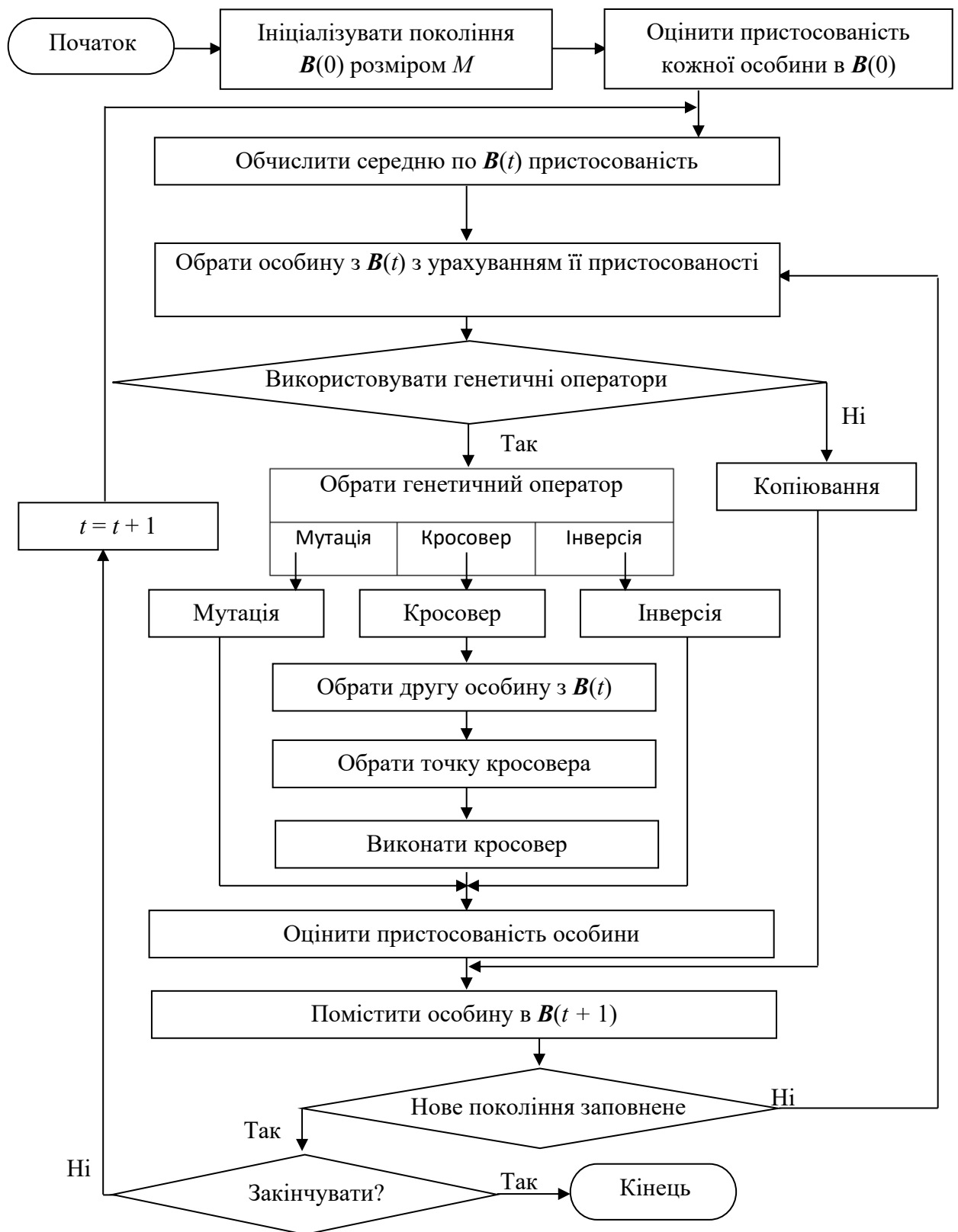


Рис. 19.4. Схема генетичного алгоритму (варіант за [6])

Через  $M$  кроків нове покоління буде сформоване. Хромосоми, які не узяли участі у формуванні нового покоління, виключаються з подальшого розгляду .

Робота алгоритму завершується або після досягнення кількістю поколінь наперед визначеного значення, або після стягування популяції у точку.

Зазвичай для ГА за розв'язок задачі оптимізації обирається найкраще значення функції пристосованості в останньому поколінні.

Необхідно зазначити, що нині запропонована досить велика кількість варіантів реалізації розглянутих генетичних операторів, наприклад, кросовер може бути одноточковим (як на рис. 19.1),  $n$ -точковим або рівномірним, а під час інверсії можуть мінятися не значення генів, а їх місце в хромосомі. У деяких випадках вважається за доцільне використання генетичних операторів, для яких немає аналогів у природі .

Окрім того, нині існує досить велика кількість варіантів формування нового покоління, зокрема широко застосовується принцип елітизму, відповідно до якого найкраща особина примусово переноситься до нового покоління. В такому разі за розв'язок задачі оптимізації приймають генотип елітної особини в останньому поколінні.

У загальному випадку, окрім наведених, можуть бути використані й інші ознаки завершення розрахунків, наприклад, кількість обчислень значень функції пристосованості, витрачений час на розрахунки тощо.

Нині ГА широко використовуються для розв'язання прикладних задач у багатьох галузях: для оптимізації багатоекстремальних функцій, налагодження параметрів нейронних мереж, синтезу архітектури обчислювальних систем , розв'язання багатоцільових задач теорії розкладів, оптимізації маршрутів перевезення вантажів, функціональної діагностики магістральних трубопроводів , проектування складних агрегатів літаків та металевих конструкцій, оптимального розкрою матеріалів, розподілу тактичних ракет проти групових цілей та неоднорідних засобів ураження по об'єктах противника тощо.

## 19.2. Метод LARES

В основу методу оптимізації LARES, що вперше був описаний у 2004 році як метод, призначений для розв'язання дискретних оптимізаційних задач, покладено відоме положення про те, що стабільний стан деякої речовини характеризується мінімумом її потенційної енергії, що відповідно потребує мінімізації енергії молекул, з яких вона складається.

Для реалізації цього положення з метою розв'язання задач оптимізації використовується концепція штучної хімічної реакції, яка відповідно до полягає в такому.

$V$  незалежних змінних задачі кодують набором дискретних змінних  $(x_j, j = 1, \dots, V)$ , які отримали назву „молекули”. Нехай  $\Omega_j$  є набором можливих значень змінної  $x_j$ ,  $\Omega_j = \{i_1, \dots, i_{M_j}\}$ , де  $M_j$  – загальна кількість можливих значень молекули  $j$ . Сукупність дискретних значень, що відповідають молекулі, називають „станами”. Після того, як вихідна задача закодована молекулярними змінними, єдиною інформацією, що використовується в алгоритмі, є функція енергії, яка відповідає цільовій функції. Розв'язком задачі оптимізації є нова „хімічна речовина”, яка складається з молекул, стани яких відповідають мінімуму функції енергії  $F$ .

Нехай  $Z$  є множиною всіх  $V$  молекул у системі, а  $x_j^g$  – поточний стан молекули  $j$ , якому відповідає поточне значення функції енергії,  $x_j^g = (x_1^g, \dots, x_V^g)$ . Зрозуміло, що в разі, коли значення функції енергії може бути поліпшено, тоді молекула  $j$  може бути переведена в новий стан  $x_j^t \neq x_j^g$ . Переведенням молекул з поточного стану  $x_j^g$  у стан  $x_j^t$  формують нову множину станів молекул, яким (у разі розв'язання задачі мінімізації) відповідає умова  $F(x^t) < F(x^g)$ .

Для пошуку такого кращого набору станів молекул використовується стратегія штучної хімічної реакції. Нехай  $L$ ,  $AR$ ,  $E$  та  $S$  є чотирма диз'юнктивними множинами, елементи яких – молекули-змінні. Нехай  $F$  – функція, що підлягає мінімізації, закодована в термінах молекул-змінних,  $F = F(x)$ . Нехай значенню найкращого поточного розв'язку (тобто мінімуму енергії) відповідає  $x_j^g = (x_1^g, \dots, x_V^g)$ .

Алгоритм методу LARES, схему якої наведено на рис. 19.5, полягає в перерозподілі молекул-змінних між множинами в такий спосіб, щоб забезпечити зменшення значення функції енергії.

Кожна множина асоціюється з деяким пристроєм хімічного реактора:  $L$  – з пристроєм завантаження;  $AR$  – з реактором активації;  $S$  – з пристроєм сепарації;  $E$  – з пристроєм відбору. Власне, сукупність позначень цих пристроїв і дала назву методу – LARES.

На першому етапі молекули-змінні ініціалізують у пристрої завантаження ( $L$ ) значеннями  $x_j^g$ . Частина молекул, відібрана з використанням рівномірного генератора випадкових чисел, надходить до реактора активації ( $AR$ ), у якому молекули переводять до нового стану з використанням правила: кожного разу, коли молекула-змінна  $x_j$  надходить до  $AR$ , їй (з використанням генератора рівномірного розподілу) призначається нове значення  $x_j^a$  (з множини припустимих), відмінне від поточного. При цьому  $x_j^a \neq x_j^g \in$  новим значенням, яке не змінюється доти, доки молекула належить множині  $AR$ .

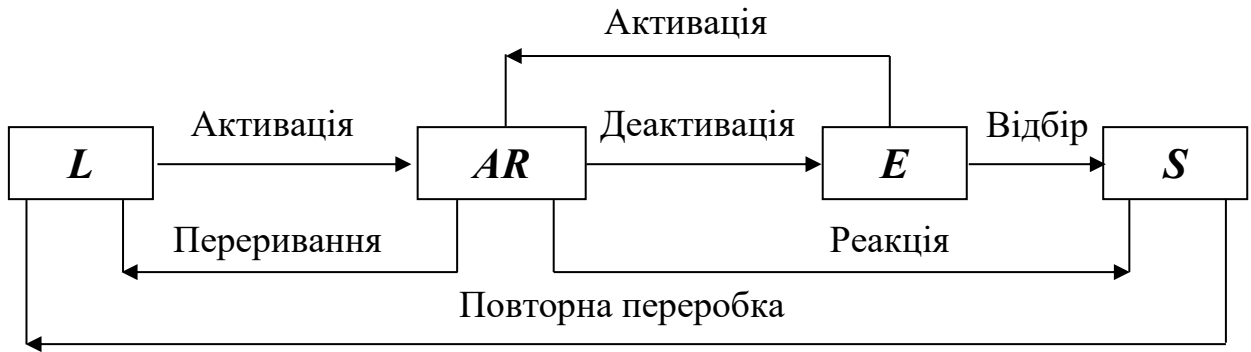


Рис. 19.5. Схема штучної хімічної реакції

Пробний вектор розв'язку задачі оптимізації  $x^t = (x^t_1, \dots, x^t_\nu)$  будується за формулою:

$$x_j^t = \begin{cases} x_j^g, & \text{якщо } x_j^t \notin AR, \\ x_j^a, & \text{якщо } x_j^t \in AR. \end{cases}$$

У разі, коли для вектора  $x^t_j$  поліпшення  $F$  не відбувається, молекули повертають до пристрою завантаження ( $L$ ) та здійснюють відбір іншої групи молекул з  $L$  для передавання до реактора активації  $AR$ .

У разі поліпшення значення функції енергії ті молекули, завдяки яким таке поліпшення досягнуте, передають до пристрою сепарації ( $S$ ).

Серед тих молекул, що залишились у реакторі активації, випадковим чином обирають деяку частину молекул, які передають до пристрою відбору ( $E$ ), де до них також застосовують операцію модифікації за правилом,

аналогічним правилу, використаному для пристрою активації. У разі поліпшення значення функції енергії ті молекули, завдяки яким таке поліпшення досягнуте, також передають до пристрою сепарації ( $S$ ).

Ті молекули, що залишились у  $E$ , повертають до реактора активації та здійснюють відбір іншої групи молекул для передавання до пристрою відбору. Процедуру передавання молекул з  $AR$  до  $E$  та в зворотному напрямку здійснюють задану кількість разів.

У разі, коли всі молекули переходять до пристрою сепарації, їх знову повертають до пристрою завантаження і повторюють алгоритм розрахунків. За розв'язок задачі оптимізації приймають вектор  $x^t_j = (x^t_1, \dots, x^t_V)$ , значення якого відповідають станам молекул у всіх пристроях, які вони отримали на момент виконання умови закінчення розрахунків.

Необхідно звернути увагу на те, що умови завершення розрахунків доцільно перевіряти перед виконанням кожного пересилання молекул між пристроями для виключення невиправданих витрат часу на розв'язання задачі оптимізації.

Конкретні умови завершення розрахунків розробляють з огляду на специфіку задачі оптимізації, що розв'язується. Додатково можуть використовуватись умови, засновані на граничній кількості розрахунків значення цільової функції та довжині “полиці”, яка характеризує сталість значення найкращого отриманого розв'язку.

Необхідно зазначити, що на відміну від ГА, який одночасно оперує з множиною розв'язків задачі оптимізації, метод LARES намагається покращити один розв'язок, елементи якого розподілені по пристроях штучного хімічного реактора.

Нині метод LARES достатньо ефективно використовується для оптимізації нестационарних систем, а також для розподілу неоднорідних ресурсів.

### **19.3. Метод моделювання відпалу**

Прообразом методу моделювання відпалу є широко розповсюджена в машинобудуванні технологія термічного оброблення кристалічних матеріалів – відпал, яка спрямована на виправлення дефектів їх структури, підвищення однорідності матеріалу та зняття в ньому внутрішніх напружень. Сутність

цієї технології полягає в нагріванні матеріалу до високої температури, його витримці при цій температурі протягом визначеного часу та подальшому охолодженні за спеціальним режимом. При цьому атоми з високо напружених ділянок мігрують по об'єму матеріалу, займаючи вільні місця в кристалічній решітці, що призводить до формування структури з малими дефектами та мінімізації внутрішньої енергії матеріалу.

Ідея перенесення механізму відпалу на розв'язання оптимізаційної задачі мінімізації була запропонована в 1983 році і полягає в тому, що покроковий процес оптимізації пов'язують з деякою температурою  $T$ . На кожному кроці здійснюють малу зміну стану об'єкта (значень незалежних змінних) та обчислюють відповідну зміну його енергії ( $\Delta E$ ), яка ототожнюється зі зміною значення цільової функції. Новий стан об'єкта приймається з імовірністю 1, якщо  $\Delta E < 0$  (тобто величина енергії зменшилась), та з імовірністю  $\exp(-\Delta E/T)$ , якщо  $\Delta E \geq 0$  (тобто величина енергії не змінилась або збільшилась). Під час розрахунків  $T$  поступово зменшують (у найпростішому випадку лінійно з коефіцієнтом  $\alpha < 1$ ), що призводить до поступового зменшення пошукової активності алгоритму оптимізації.

Блок-схему розглядуваного методу наведено на рис. 19.6, відповідно до якого реалізація ймовірнісного механізму в оптимізаційному алгоритмі на основі моделювання відпалу здійснюється в такий спосіб.

На початку розрахунків обирається вихідна точка, з якої буде здійснюватись пошук оптимуму та визначається початкове значення температури  $T$ .

Нове значення  $i$ -ї незалежної змінної генерується шляхом додавання до поточного значення  $x_i$  величини  $\Delta x_i$ . Для  $x_i + \Delta x_i$  розраховується значення цільової функції  $E(x_i + \Delta x_i)$  та величина її змінювання  $\Delta E$ .

Якщо відбулося покращення цільової функції  $E$ , тоді  $x_i + \Delta x_i$  приймається за поточне значення незалежної змінної і новий крок пошуку оптимуму буде здійснюватись з точки  $x_i + \Delta x_i$ .

Якщо ж покращання  $E$  не відбулося, тоді обирається випадкове число  $R$ , рівномірно розподілене в діапазоні  $[0, 1]$ , та перевіряється умова:

$$R \leq \exp(-\Delta E/T).$$

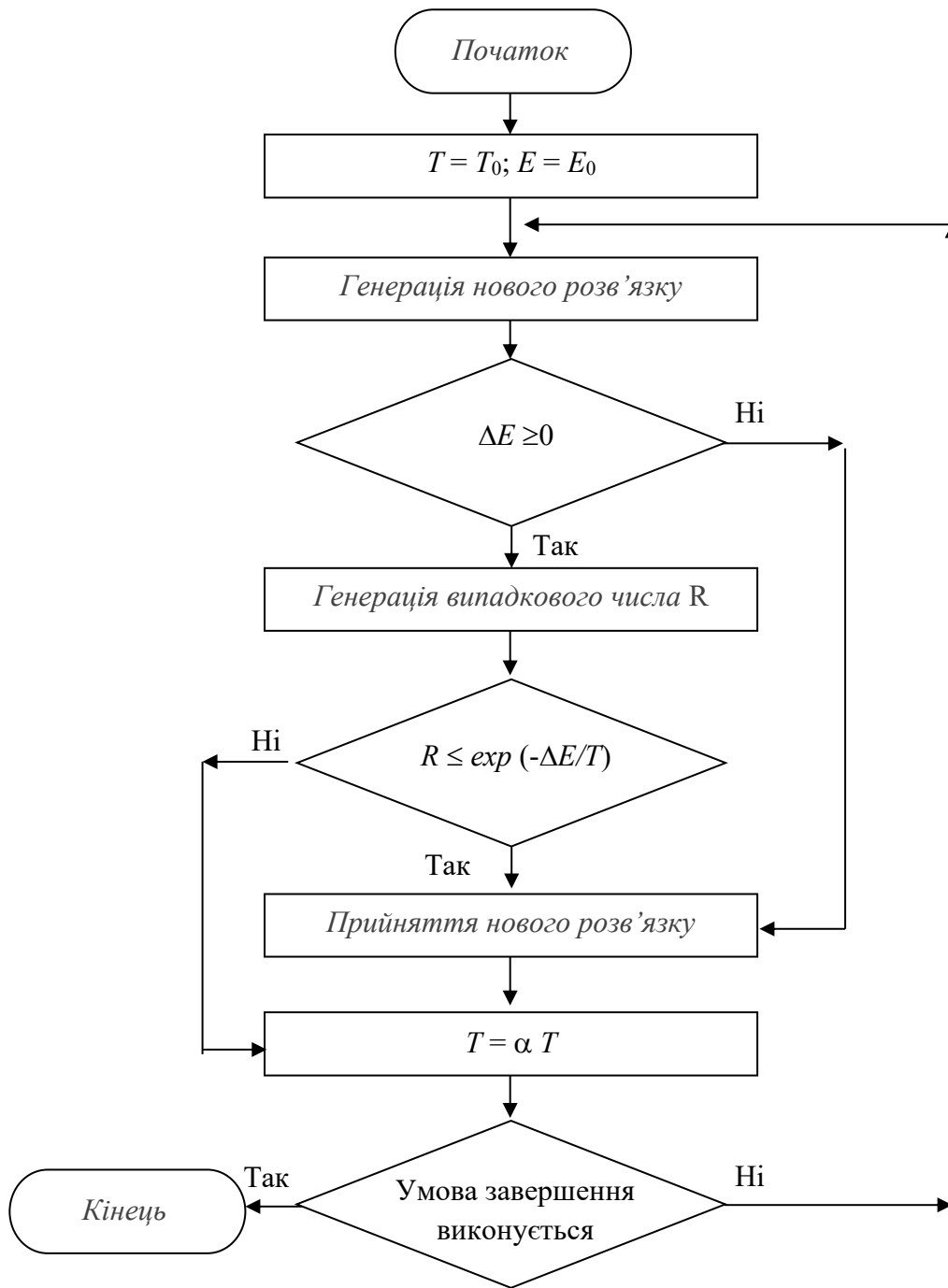


Рис. 19.6. Блок-схема алгоритму моделювання відпалу

Якщо ця умова виконується, тоді  $x_i + \Delta x_i$  приймається за новий вихідний стан для подальшого пошуку оптимуму. В іншому випадку вихідним вважається  $x_i$ .

Розв'язком задачі оптимізації вважається сукупність значень незалежних змінних, отриманих ними на час, коли температура  $T$  наближається до 0.

Необхідно зазначити, що такий алгоритм пошуку оптимуму дозволяє долати відносно неглибокі локальні мінімуми та підвищує збіжність до глобального оптимуму.

На відміну від ГА, метод моделювання відпалу, як і метод LARES, оперує одним розв'язком задачі оптимізації.

Як і в інших чисельних методах, ознаки закінчення розрахунків повинні обиратись з огляду на необхідність виключення невиправдано великих витрат часу на розв'язання задачі.

Доцільно звернути увагу на те, що існує досить велика кількість варіантів реалізації методу моделювання відпалу, які різняться використанням розподілом величини  $R$ , законами змінювання температури  $T$  тощо.

#### **19.4. Метод оптимізації колонією бджіл**

Перші варіанти методу оптимізації колонією бджіл були запропоновані в 2005 році. Цей метод моделює поведінку бджіл під час пошуку ними ділянок квітів з достатньою кількістю нектару. Цей метод належить до класу методів ройового інтелекту, що базуються на алгоритмічній реалізації таких природних механізмів поведінки соціальних істот, які забезпечують досягнення високої загальної ефективності функціонування соціуму за досить простих правил поведінки окремої істоти.

Для пошуку джерел їжі та її доставки до вулика в колонії бджіл існує розподіл функцій: первинний пошук джерел здійснюють розвідники, а транспортування їжі (нектару) забезпечують робочі бджоли. Окрім того, робочі бджоли оновлюють інформацію, надану розвідниками. Як первинна, так і оновлена інформація передається вільним (ще не зайнятим) бджолам і вони самостійно обирають те джерело їжі, з якого їм доцільно здійснювати фуражування.

Передавання інформації здійснюється під час специфічного бджолиного танцю, який може виконати на спеціальному майданчику кожна бджола, яка повернулася до вулика. Окремі елементи цього танцю містять інформацію про напрямок руху до джерела їжі, відстань до нього та його кормову цінність. Вільні бджоли спостерігають цей танок, і кожна з них

самостійно приймає рішення про доцільність здійснення фуражування з деякого джерела.

При цьому необхідно зазначити, що під час вибору джерела їжі кожна бджола враховує не тільки інформацію, отриману під час спостереження танцю, але й стан своїх крил. Справа в тому, що основною причиною вимирання робочих бджіл, тривалість життя яких влітку становить 1–2 місяці, є зношування їхніх крил. Молода бджола з „новими крилами” може здійснювати фуражування з віддалених джерел з повним навантаженням. Зношування крил спонукає її до обмежування відстані польотів та ваги транспортованого нектару.

Аналіз кожною бджолою інформації, здобутої нею, та інформації від інших бджіл призводить до того, що бджоли не концентруються на найкращому розвіданому джерелі їжі, а розподіляються групами по деякій площі. З одного боку, це забезпечує максимізацію кількості доставленої до вулика їжі всіма робочими бджолами за весь сезон, а з другої – забезпечує підвищення ефективності розвідки нових перспективних джерел їжі за рахунок додаткової інформації від робочих бджіл про околиці тих джерел, з яких вже здійснюється фуражування.

Якщо провести аналогію між цим механізмом та розв’язанням задачі оптимізації, то можна сказати, що поодинокі розвідники здійснюють випадковий глобальний пошук оптимуму, а робочі бджоли групами досліджують найбільш перспективні ділянки та концентруються на локальних оптимумах.

Базовий алгоритм оптимізації бджолиною колонією розглядає три типи бджіл:

- розвідників, які здійснюють випадковий пошук оптимуму в усьому просторі незалежних змінних,

$$x_{ij}^* = \text{random}(x_{i \max}; x_{i \min}), \quad (19.1)$$

де  $x_{ij}^*$  – нове значення  $i$ -ої незалежної змінної для  $j$ -ої бджоли-розвідника;

$x_{i \max}, x_{i \min}$  – діапазон змінювання  $i$ -ої незалежної змінної;

$\text{random}(x_{i \max}; x_{i \min})$  – випадкова величина, що має рівномірний розподіл у діапазоні  $x_{i \max} \dots x_{i \min}$ ;

- робочих бджіл, які закріплені за деяким джерелом їжі та здійснюють пошук оптимуму в околі цього джерела,

$$x_{ij}^* = x_{ij} + \text{random}(-1; 1) (x_{ij} - x_{ik}), \quad (19.2)$$

де  $x_{ij}^*$  – нове значення  $i$ -ї незалежної змінної для  $j$ -ї робочої бджоли;  
 $x_{ij}$  – поточне значення  $i$ -ї незалежної змінної для  $j$ -ї робочої бджоли;  
 $x_{ik}$  – поточне значення  $i$ -ї незалежної змінної для  $k$ -ї робочої бджоли;

- вільних робочих бджіл, які можуть бути завербовані робочими бджолами за результатами виконання ними танцю.

При цьому пошук оптимуму робочою бджолою в околі свого джерела їжі забезпечується тим, що своє поточне положення вона порівнює із положенням випадково обраної іншої робочої бджоли, тобто для формули (19.2)

$$k = \text{random}(1; N_{\text{роб}}), k \neq j, \quad (19.3)$$

де  $N_{\text{роб}}$  – загальна кількість робочих бджіл.

Під час вербування вільних бджіл оцінюється перспективність кожного джерела їжі, знайденого робочою бджолою, а кількість вільних бджіл  $N_{j \text{ вил}}$ , які вербуються  $j$ -ю робочою бджолою для дослідження  $j$ -го джерела, встановлюється пропорційно його привабливості  $\text{fitness}_j$ :

$$N_{j \text{ вил}} = N_{\text{вил}} \text{fitness}_j / \sum_{i=1}^{N_{\text{роб}}} \text{fitness}_i, \quad (19.4)$$

де  $N_{\text{вил}}$  – загальна кількість вільних бджіл.

Завербована вільна бджола модифікує своє положення за (19.2). Як робочі, так і завербовані вільні бджоли переходять у нове положення лише тоді, коли (для задачі максимізації)

$$\text{fitness}_j(x_{ij}^*) > \text{fitness}_j(x_{ij}). \quad (19.5)$$

Блок-схема базового варіанта алгоритму пошуку оптимуму колонією бджіл наведена на рис. 19.7.

За розв'язок задачі оптимізації приймається точка, якій відповідає максимальне значення  $\text{fitness}$ , знайдене всіма бджолами за всі цикли пошуку.

Надаючи загальну оцінку методу оптимізації колонією бджіл, необхідно зазначити, що його особливістю є спроба розділити процес пошуку оптимуму на два масштабні рівні: глобальний (що реалізується розвідниками) та локальний (що реалізується іншими бджолами). Окрім того, заслуговує на увагу механізм динамічного перерозподілу пошукових ресурсів, реалізований спрямуванням більшої кількості бджіл на ті ділянки, які здаються

більш перспективними з точки зору можливості розташування глобального оптимуму.

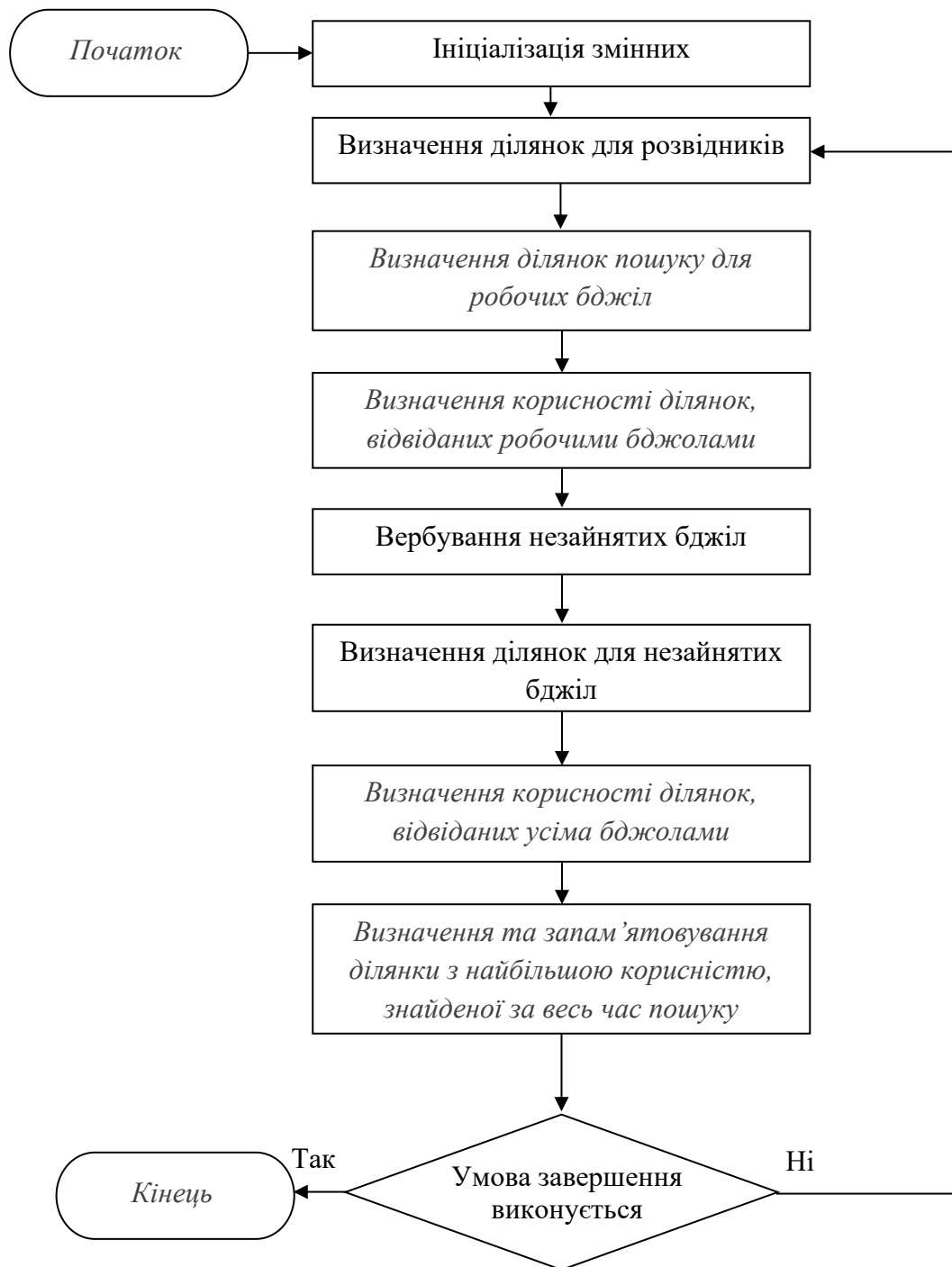


Рис. 19.7. Блок-схема методу оптимізації колонією бджіл (варіант)

Нині метод оптимізації бджолиною колонією досить широко використовується під час навчання нечітких систем та для розв'язання різних оптимізаційних задач. При цьому, запропоновані його модифікації, пов'язані

із перерозподілом кількості бджіл за спеціалізацією, використанням ідей методу моделювання відпалу тощо.

### 19.5. Метод оптимізації роєм частинок

Одним з перспективних методів дискретної оптимізації є метод PSO (particle swarm optimization – оптимізація роєм частинок), який моделює поведінку птахів у зграї (риб у косяку). Первісною основою методу PSO була імітація руху птахів у зграї (риб у косяку) з метою виявлення базових принципів, завдяки яким птахи літають (риби плавають) синхронно та вміють змінювати напрямок руху з перегрупуванням в оптимальні формації.

Відповідно до , у методі PSO особини, що називаються частинками, переміщуються в багатовимірному просторі рішень. Зміни координат частинок зумовлені природною соціально-психологічною тенденцією частинок конкурувати між собою. Тому, зміни в стані частинки залежать не тільки від її власного досвіду, але й від досвіду та знань її сусідів. У результаті таке моделювання спонукає частинки недетермінованим чином повертатися в оптимальні ділянки простору рішень.

Особини рою мають дуже просту поведінку: вони прагнуть перевершити досягнення сусідніх частинок та поліпшити власні. PSO-метод керує роєм частинок і кожна частинка уособлює потенційне рішення. За аналогією з генетичними алгоритмами рій можна трактувати як популяцію, а частинку – як особину.

Нехай  $x_{ij}(t)$  – позиція  $i$ -ї частинки в просторі пошуку в момент часу  $t$  ( $t$  – дискретні значення часу) по  $j$ -й координаті. Позиція частинки (за одиницю часу) змінюється додаванням швидкості  $v_{ij}(t)$  до координати поточної позиції за формулою:

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t), \quad (19.6)$$

де  $v_{ij}(t)$  – швидкість частинки  $i$  у вимірі  $j$  ( $j=1, \dots, n$ ) у момент часу  $t$ ;  $x_{ij}(t)$  – позиція частинки  $i$  у вимірі  $j$  у момент часу  $t$ ;  $n$  – кількість незалежних змінних.

Початкові координати частинок  $x_{ij}(0)$  визначають з використанням генератора випадкових чисел з припустимого діапазону варіювання незалежних змінних. Швидкість, що входить до формули (1.6), фактично

визначає ефективність оптимізаційного процесу та містить складові, які відповідають за обмін інформацією між частинками рою.

Власні знання частинки (когнітивні знання) вважаються пропорційними відстані від поточного положення частинки до точки, яка відповідає найкращому значенню цільової функції, знайденому нею за весь час її життя.

Знання рою (соціальні знання), які використовуються частинкою, вважаються пропорційними відстані від поточного положення частинки до точки, яка відповідає найкращому значенню цільової функції, знайденому сусідами частинки за весь час життя рою.

Існує два основних різновиди методу PSO: *gbest* та *lbest*, які різняться правилами формування множини сусідів частинки, що формують її соціальні знання. В різновиді *gbest* вважається, що всі частинки є сусідами, тобто соціальні знання пов'язуються з найкращим значенням цільової функції, знайденим усім роєм за весь час його життя. Тоді швидкість частинки розраховується за формулою:

$$v_{ij}(t+1) = v_{ij}(t) + c_1 [y_{ij}(t) - x_{ij}(t)] \text{ random} + c_2 [q_j(t) - x_{ij}(t)] \text{ random},$$

де  $c_1$ ,  $c_2$  – позитивні константи прискорення, які використовуються для варіювання ваг когнітивної та соціальної компонент швидкості частинки відповідно;

$y_{ij}(t)$  – найкраща за значенням цільової функції точка, знайдена частинкою за всі ітерації;

$q_j(t)$  – найкраща за значенням цільової функції точка, знайдена роєм за всі ітерації;

*random* – випадкова величина, що має рівномірний розподіл у діапазоні  $[0,1]$ .

У варіанті *lbest* вважається, що як сусіди можуть розглядатись частинки, що утворюють деякі геометричні конфігурації (трикутники, кола тощо). Необхідно зазначити, що більш ефективною для розрахунків швидкості частинок є формула

$$v_{ij}(t+1) = w v_{ij}(t) + c_1 [y_{ij}(t) - x_{ij}(t)] \text{ random} + c_2 [q_j(t) - x_{ij}(t)] \text{ random},$$

яка додатково враховує інерційність частинок  $w$ .

На основі розрахованих за (1.6) значень координат частинок ( $x_{ij}(t)$ ) для всіх частинок рою проводять розрахунки значень функції пристосованості

( $fitness_i$ ). Отримані значення  $fitness_i$  є основою для коригування когнітивних та соціальних знань.

На рис. 19.8 наведено блок-схему алгоритму методу PSO.

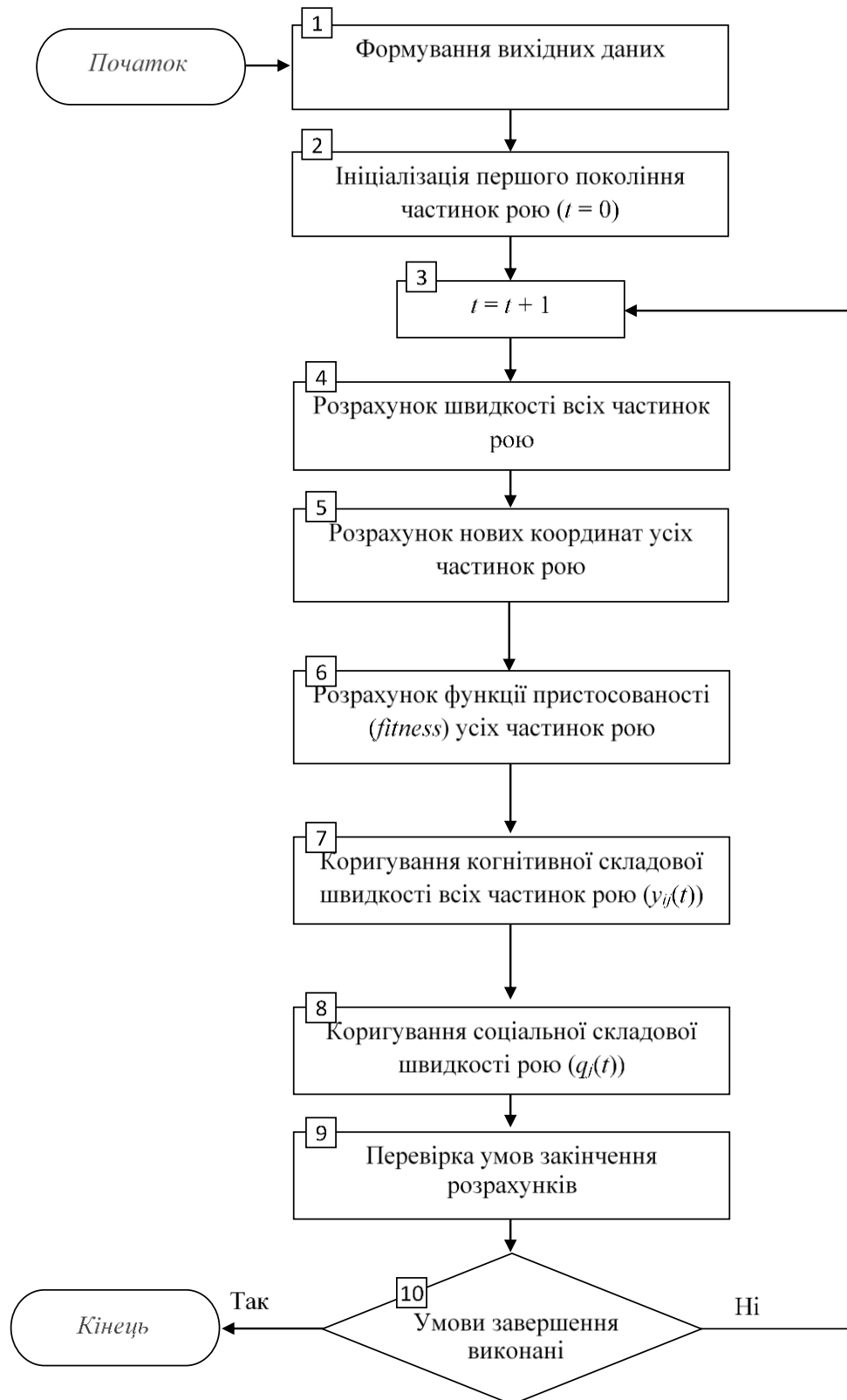


Рис. 19.8. Блок-схема алгоритму методу PSO (варіант)

За розв'язок задачі оптимізації приймають значення незалежних змінних у точці, яка відповідає найкращому значенню цільової функції, знайденому усім роєм за всі ітерації ( $q_j(t)$  на останній ітерації).

## 19.6. Метод оптимізації мурашиною колонією

Метод оптимізації мурашиною колонією базується на імітації природних механізмів самоорганізації, які мають місце в колоніях мурах під час пошуку та транспортування їжі до мурашника. Привабливість використання імітації цих природних механізмів пов'язана з тим, що мурахи завжди прокладають такий шлях до джерела їжі, який максимально наближений до найкоротшого, тобто фактично розв'язують задачу мінімізації.

Процес самоорганізації у такому разі забезпечується двома механізмами: механізмом модифікації зовнішнього середовища та механізмом вибору маршруту руху до джерела їжі.

Механізм модифікації зовнішнього середовища полягає в тому, що кожний мураха під час руху залишає за собою на землі феромон – досить стійку хімічну речовину зі специфічним запахом, який може сприйматись іншими мурахами протягом декількох діб. Чим більше мурах рухались одним шляхом, тим більшою буде концентрація феромону на ньому.

Механізм вибору маршруту руху до джерела їжі є імовірнісним та полягає в тому, що між декількома маршрутами мураха робить вибір з імовірністю, пропорційною концентрації феромону на них.

Якщо розглянути з позицій зазначених механізмів досить просту ситуацію, наведену на рис. 19.9, то видно, що перший маршрут є довшим за другий. У разі, коли два мурахи одночасно вийдуть з мурашника на пошук їжі за обома маршрутами, залишаючи феромонний слід, то швидше повернеться мураха, який рухався за другим маршрутом, що забезпечить концентрацію феромона на другому маршруті вдвічі більшу, ніж на першому.

Отже, для третього мурахи, який буде обирати шлях, імовірність вибору маршрутів (виходячи з концентрацій феромону на них) буде співвідноситись як 1:2, тобто він з імовірністю, близькою до 0,66, обере другий маршрут.

Очевидно, що збільшення кількості мурах, які обрали другий маршрут, призведе до того, що перший маршрут практично не використовуватиметься.

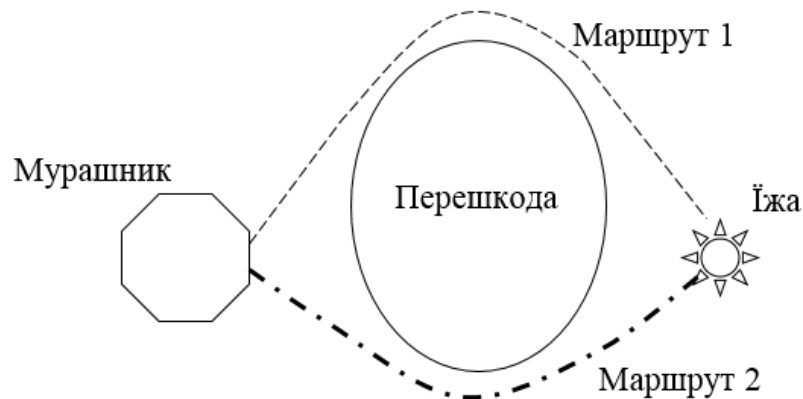


Рис. 19.9. Схема маршрутів руху мурах до джерела їжі

Розглянуті модельні уявлення наочно свідчать, що найбільш ефективним є використання методу оптимізації мурашиною колонією для розв'язання задач, які можуть бути зведені до графів, зокрема для задачі комівояжера, у якій міста розглядаються як вершини графа, а шляхи між містами – як його ребра.

Як відомо, задача комівояжера полягає в тому, щоб знайти найкоротший замкнений маршрут, який проходить через усі задані міста тільки один раз. Для розв'язання цієї задачі методом мурашиної колонії вважають, що мураха свій маршрут між містами обирає залежно від деяких обставин (переліку міст, у яких мураха ще не був, концентрації феромону на можливих ділянках руху та видимість міст). Після декількох ітерацій, які моделюють рух мурах, розв'язку задачі буде відповідати маршрут, на якому сконцентрована найбільша кількість феромону.

Перелік міст, у яких мураха ще не був, являє собою своєрідну пам'ять мурахи та гарантує, що він не буде заходити в кожне місто більше одного разу.

Концентрація феромону на ділянках між містами на початку розрахунків ( $\tau_0$ ) установлюється невеликою. Окрім того, визначають коефіцієнт швидкості випаровування феромону  $p \in [0, 1]$ . Тоді залежність для оновлення концентрації феромону на  $(t + 1)$  кроці алгоритму має вигляд:

$$\tau_{ij}(t+1) = (1-p)\tau_{ij}(t) + \sum_{k=1}^m \Delta\tau_{ijk}(t),$$

де  $t$  – номер ітерації;  $\tau_{ij}(t)$ ,

$\tau_{ij}(t+1)$  – концентрація феромону на ділянці між  $i$ -м та  $j$ -м містами на ітераціях  $t$  та  $t+1$  відповідно;

$m$  – кількість мурах у колонії;

$\Delta\tau_{ijk}$  – кількість феромону, залишена  $k$ -м мурахою на ділянці між  $i$ -м та  $j$ -м містами.

Кількість феромону, залишена  $k$ -м мурахою на ділянці між  $i$ -м та  $j$ -м містами, оновлюється після завершення маршруту відповідно до формули:

$$\Delta\tau_{ijk}(t) = \begin{cases} \frac{Q}{L_k(t)}, & \text{якщо } (i, j) \in T_k(t), \\ 0, & \text{якщо } (i, j) \notin T_k(t), \end{cases}$$

де  $T_k(t)$  – маршрут, який пройшов  $k$ -й мураха на ітерації  $t$ ;

$L_k(t)$  – довжина цього маршруту;

$Q$  – параметр, значення якого обирають одного порядку з довжиною оптимального маршруту.

Видимість міста асоціюється з бажанням мурахи відвідати деяке місто та обчислюється за формулою:

$$\eta_{ij} = 1 / D_{ij},$$

де  $D_{ij}$  – відстань між містами  $i$  та  $j$ .

Зрозуміло, що чим ближче місто до поточного розташування мурахи, тим більше його бажання відвідати це місто.

Імовірність переходу  $k$ -го мурахи з міста  $i$  до міста  $j$  на ітерації  $t$  визначається за формулою:

$$P_{ijk}(t) = \begin{cases} \frac{\tau_{ij}^\alpha(t)\eta_{ij}^\beta(t)}{\sum_{l \in J_{ik}} \tau_{il}^\alpha(t)\eta_{il}^\beta(t)}, & \text{якщо } j \in J_{ik}, \\ 0, & \text{якщо } j \notin J_{ik}, \end{cases}$$

де  $P_{ikj}$  – імовірність вибору  $k$ -м мурахою, який знаходиться в місті  $i$ , наступного  $j$ -го міста;

$\alpha, \beta$  – параметри алгоритму;

$J_{ik}$  – перелік міст, які необхідно відвідати  $k$ -му мураші, який знаходиться в місті  $i$ .

Параметри  $\alpha$  та  $\beta$  є настоюваними. Вони забезпечують змінювання співвідношення впливу концентрації феромону та видимості міста на імовірність вибору мурахою напрямку руху.

У разі, коли  $\alpha = 0$ , наступним буде обрано найближче місто. У разі, коли  $\beta = 0$ , наступним буде обрано місто, якому відповідає найбільша концентрація феромону.

Для визначення наступного міста, до якого буде рухатись мураха, використовується рулетковий відбір, відповідно до якого вважається, що колесо рулетки поділено на сектори, площі яких пропорційні розрахованій імовірності переходу в можливі міста, а згенероване випадкове число потраплянням в один із секторів визначає наступне місто. Зрозуміло, що частіше будуть обиратись ті міста, яким відповідають більші сектори рулетки.

За розв'язок задачі оптимізації приймається найкоротший маршрут, знайдений за весь час розрахунків.

Блок-схема методу мурашиної колонії наведена на рис. 19.10.

Необхідно зазначити, що в різних варіантах реалізації методу мурашиної колонії використовують різні варіанти початкового розміщення мурах. Вони можуть бути розподілені по різних містах, при цьому дозволяють або забороняють наявність декількох мурах в одному місті. Існують варіанти, в яких усі мурахи стартують з одного міста.

За результатами тестових розрахунків зазначається, що порівняно з точними методами, зокрема, динамічним програмуванням або методом гілок та границь, метод мурашиної колонії знаходить близькі до оптимуму розв'язки за значно менший час навіть для задач невеликої розмірності (вже для 20 міст). При цьому час оптимізації є поліноміальною функцією від розмірності задачі, тоді як для точних методів така залежність є експоненціальною.

Метод мурашиної колонії з незначними модифікаціями може бути використаний для розв'язання різних комбінаторних задач, зокрема квадратичної задачі про призначення, оптимізації маршрутів вантажних автомобілів, розфарбовування графа, а також для відбору супутників для

розрахунку координат у глобальній системі позиціонування. Окрім того, вони демонструють високу ефективність під час оптимізації нестационарних систем, наприклад, для трафіка в телекомунікаційних мережах.



Рис. 19.10. Блок-схема методу мурашиної колонії для розв'язання задачі комівояжера (варіант)

## 19.7. Метод оптимізації колонією бактерій

Метод оптимізації колонією бактерій базується на імітації природних механізмів самоорганізації, які мають місце в колоніях бактерій. При цьому розглядаються як особливості руху кожної окремої бактерії, так і взаємодія між ними, що разом забезпечують утворення досить сталих просторово-часових конфігурацій колонії у сприятливому середовищі.

Не дивлячись на те, що бактерії є найбільш чисельними істотами на Землі, більшість робіт, пов'язаних з використанням моделювання поведінки бактерій з метою розв'язання оптимізаційних задач, базуються на даних, накопичених біологами стосовно бактерії *E.Coli*. Це пояснюється тим, що вона є найбільш вивченою, а характер її поведінки у різних середовищах має чітку спрямованість на ефективний пошук місць харчування з максимальною концентрацією корисних речовин. При цьому дослідниками вважається, що бактерії намагаються максимізувати кількість енергії, отриманої ними в одиницю часу за рахунок їжі з урахуванням впливу хімічного складу середовища на їхнє взаємне розташування.

Бактерія *E.Coli* існує в кишечниках більшості ссавців. Особливості будови її рухового апарату та наявність чутливих рецепторів дозволяють бактерії дуже швидко реагувати на зміни в оточуючому середовищі практично без інерційного руху.

Бактерія *E.Coli* може здійснювати лише два типи рухів: прямолінійний рух у деякому напрямку та тривимірне хаотичне обертання, внаслідок якого вона отримує випадкову орієнтацію в просторі.

Існує досить багато причин, які спонукають бактерію до пересування, зокрема змінювання магнітного поля, температури, освітлення тощо. Однак з метою розв'язання задач оптимізації моделюють рух бактерій, зумовлений зміною хімічного складу оточуючого середовища. Він отримав назву хемотаксису. Вважається, що хемотаксис бактерії відповідає трьом основним принципам:

- у разі перебування в нейтральному середовищі бактерія здійснює пошукову активність шляхом послідовного використання поступального та обертального рухів;

- у разі перебування в сприятливому середовищі бактерія збільшує довжину поступального руху в напрямку збільшення концентрації їжі з метою досягнення ділянки з її максимальною концентрацією;

- у разі перебування в несприятливому середовищі або в разі руху в напрямку зменшення концентрації їжі бактерія намагається уникнути впливу цих факторів, змінюючи напрямок руху.

У сприятливих умовах бактерії можуть досить швидко розмножуватись, однак вони мають обмежений час життя. Звертає на себе увагу той факт, що хімічні речовини, які виділяють бактерії, можуть грати роль засобу комунікації між ними. Наприклад, у разі сприятливої ситуації бактерії можуть виділяти аттрактант – речовину, яка є сигналом для інших бактерій про знайдене джерело їжі, а виділення бактерією продуктів життєдіяльності грає роль репеленту – сигналу для сусідів щодо недоцільності наближення до неї на малі відстані.

Розглянуті модельні уявлення є базовими для алгоритмів, використовуваних у методі оптимізації колонією бактерій. Однак вони можуть доповнюватись та змінюватись, що є причиною існування низки варіантів реалізації цього оптимізаційного методу .

Найбільш розвинуті алгоритми розрахунків (рис. 19.11) передбачають наявність трьох циклів. При цьому зовнішній цикл (по  $l$ ) передбачає моделювання процесу “загибелі–розсіювання” бактерій; внутрішній цикл (по  $j$ ) призначений для моделювання хемотаксису; а в проміжковому циклі (по  $k$ ) розглядають репродукцію бактерій. При цьому розглядається задача мінімізації.

Моделювання процесу “загибелі–розсіювання” бактерій передбачає, що кожна бактерія з деякою імовірністю може бути перенесена в нову точку з випадково обраними координатами, тобто деяка бактерія гине, а замість неї у випадковій точці з’являється нова.

Процес репродукції бактерій передбачає, що для кожної  $i$ -ї бактерії визначають функцію здоров’я

$$J_{health}^i = \sum_{j=1}^{N_c} J(i, j, k, l),$$

де  $N_c$  – кількість кроків хемотаксису;

$J(i, j, k, l)$  – значення функції, яка мінімізується.

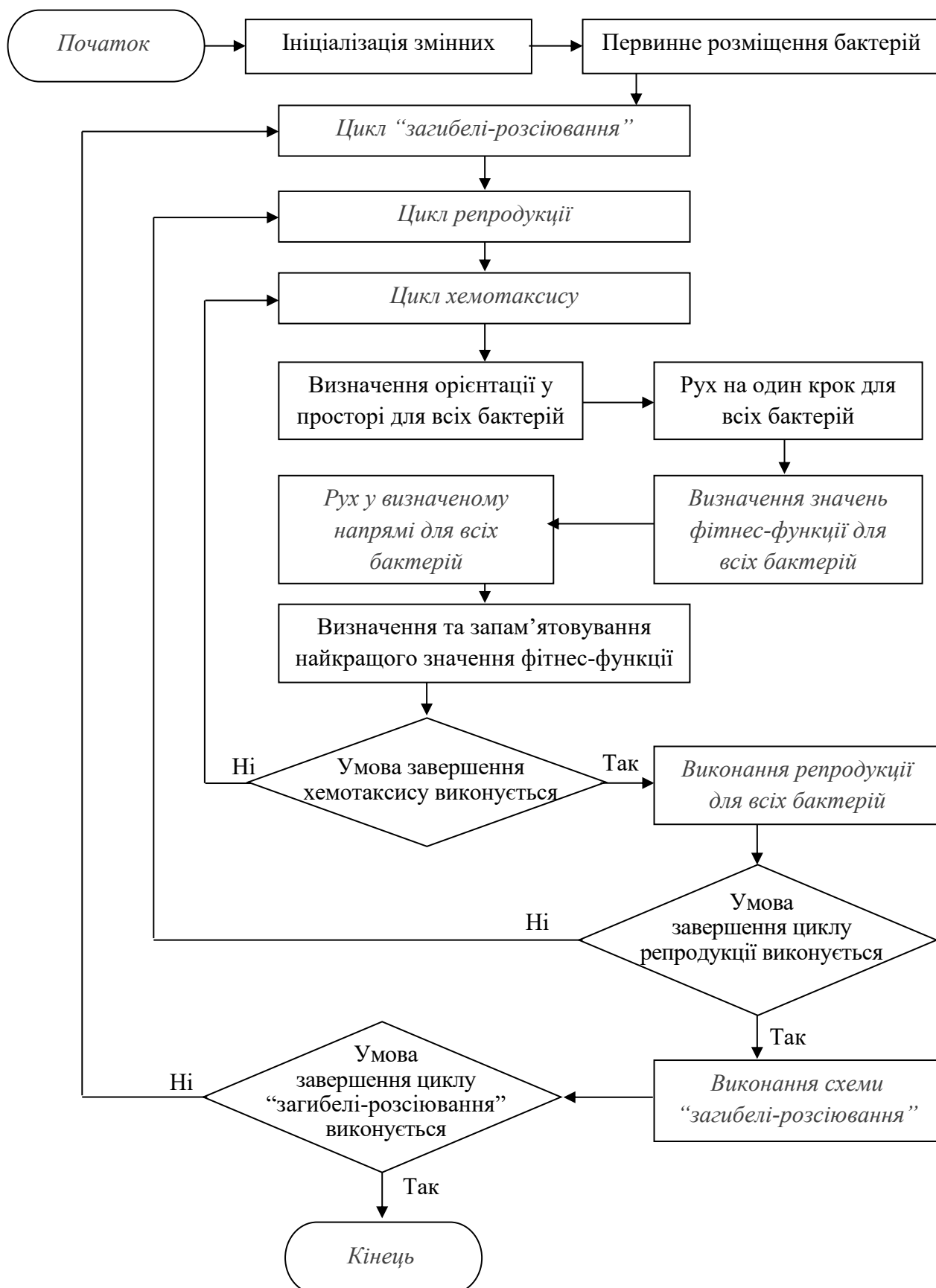


Рис. 19.11. Блок-схема методу оптимізації колонією бактерій (варіант)

За своїм змістом функція здоров'я характеризує успішність руху кожної бактерії в процесі оптимізації, при цьому менше значення функції здоров'я відповідає більш перспективній бактерії.

Усі бактерії впорядковують за збільшенням функції здоров'я, і гіршу половину з них виключають з подальшого розгляду. Кожну бактерію з тих, які залишились, дублюють, що забезпечує сталість чисельності колонії після завершення процесу репродукції.

Хемотаксис моделюють у такий спосіб.

З метою врахування обертання для кожної  $i$ -ї бактерії генерують вектор  $\Delta(i)$ , кількість елементів якого відповідає кількості незалежних змінних задачі оптимізації. При цьому кожний елемент вектора являє собою випадкову величину з рівномірним розподілом у інтервалі  $[-1, 1]$ .

Один крок переміщення бактерії для  $j+1$  циклу хемотаксису визначають за формулою:

$$\theta^i(j+1, k, l) = \theta^i(j, k, l) + zC(i) \frac{\Delta(i)}{\sqrt{\Delta^T(i)\Delta(i)}},$$

де  $\theta^i(j+1, k, l)$ ,  $\theta^i(j, k, l)$  – вектори координат  $i$ -ї бактерії на  $j$  та  $j+1$  кроках циклах хемотаксису відповідно;

$C(i)$  – вектор довжини переміщення  $i$ -ї бактерії за незалежними координатами;

$z$  – номер кроку в незмінному напрямі.

Якщо після одного переміщення бактерії ( $z = 1$ ) значення функції, яка оптимізується, покращилось, тоді рух у цьому напрямі продовжується або доти, поки значення функції не стане погіршуватись, або доти, поки не буде виконана визначена кількість кроків у незмінному напрямі.

Для забезпечення можливості врахування взаємного впливу бактерій як функція, яка фактично мінімізується, розглядається фітнес-функція  $J$  виду:

$$J = F + J_{cc},$$

де  $F$  – цільова функція вихідної задачі оптимізації,

$J_{cc}$  – функція, яка враховує профіль концентрації атрактантів та репелентів.

Профіль концентрації атрактантів та репелентів у деякій точці  $\theta$  оцінюється як

$$J_{cc}(\theta) = \sum_{i=1}^S \left[ -d_{attract} \exp(-w_{attract} \sum_{m=1}^p (\theta_m - \theta_m^i)^2) \right] + \sum_{i=1}^S \left[ h_{repel} \exp(-w_{repel} \sum_{m=1}^p (\theta_m - \theta_m^i)^2) \right],$$

де  $S$  – кількість бактерій у колонії;

$p$  – кількість незалежних змінних;

$\theta_m$  – вектор координат у точці  $\theta$ ;

$\theta^i$  – вектор координат  $i$ -ї бактерії;

$w_{attract}$ ,  $w_{repel}$  – показник швидкості дифузії атрактантів та репелентів відповідно;

$d_{attract}$ ,  $h_{repel}$  – рівень впливу атрактантів та репелентів на бактерію відповідно.

Величини  $w_{attract}$ ,  $w_{repel}$ ,  $d_{attract}$ ,  $h_{repel}$  є параметрами алгоритму, які підлягають підбору. Зазвичай вважають, що

$$d_{attract} = h_{repel}.$$

Тоді фітнес-функція для  $i$ -ї бактерії оцінюється як

$$J(\theta^i) = F(\theta^i) + J_{cc}(\theta^i).$$

Вважається [36; 37], що такий вигляд фітнес-функції краще моделює спроможність бактерії до пошуку їжі, уникнення середовища, несприятливого за хімічним складом, а також намагання колонії утворювати компактні конфігурації без надмірного наближення однієї бактерії до іншої.

За розв'язок задачі оптимізації приймають координати бактерії з найменшим значенням фітнес-функції, знайденим за весь час розрахунків.

Зрозуміло, що можуть бути реалізовані і спрощені варіанти розглядуваного методу (наприклад, без етапу “загибелі-розсіювання” або без урахування профілю концентрації атрактантів та репелентів), які на досить простих цільових функціях можуть забезпечити знаходження оптимуму.

Необхідно зазначити також, що оптимізація колонією бактерій відносно усіх розглянутих вище методів є найбільш складною з точки зору визначення комбінації значень усіх 10 настроюваних параметрів, необхідних для проведення розрахунків, яка б забезпечила збіжність методу до глобального оптимуму.

## Контрольні запитання та завдання

1. В чому полягає фізичний зміст генетичного алгоритму ?
2. Як математично реалізується генетичний алгоритм ?
3. В чому полягає фізичний зміст методу LARES?
4. Як алгоритмічно реалізується метод LARES?
5. В чому полягає фізичний зміст методу моделювання відпалу?
6. Як алгоритмічно реалізується метод моделювання відпалу?
7. В чому полягає фізичний зміст методу оптимізації колонією бджіл?
8. Як алгоритмічно реалізується метод оптимізації колонією бджіл?
9. В чому полягає фізичний зміст методу оптимізації роєм частинок?
10. Як алгоритмічно реалізується метод оптимізації роєм частинок?
11. В чому полягає фізичний зміст методу оптимізації мурашиною колонією?
12. Як алгоритмічно реалізується метод оптимізації мурашиною колонією?
13. В чому полягає фізичний зміст методу оптимізації колонією бактерій?
14. Як алгоритмічно реалізується метод оптимізації колонією бактерій?

## **Глава 20. Методичні підходи до вдосконалення еволюційних методів оптимізації та приклади їх використання**

### **20.1. Аналіз напрямків удосконалення еволюційних методів оптимізації**

Необхідно зазначити, що термодинаміка лише встановлює обов'язкові умови, за яких можливий перехід дисипативної системи від хаотичного до впорядкованого стану, та підкреслює те, що виведення системи з досягнутого впорядкованого стану потребує змінювання зовнішніх умов. Це має наслідком два висновки, які є принциповими для аналізу еволюційних методів оптимізації.

Першим висновком є те, що характер та показники впорядкованого стану системи залежать від характеристик первинного розподілу флуктуацій стану окремих її елементів, тобто в загальному випадку різний початковий розподіл флуктуацій може призвести до різних упорядкованих станів системи.

Другий висновок говорить про те, що досягнення впорядкованого стану не свідчить про досягнення глобального екстремуму деякої характеристики системи, тобто використання поняття „оптимізація” до процесів, що відбуваються в дисипативних системах, потребує достатньої обережності та попередніх уточнень.

Якщо в природних системах досягнення впорядкованого стану сприяє забезпеченню життєдіяльності біологічної системи (в селективному відборі – закріпленню корисних ознак, у фізичному процесі – створенню впорядкованого розподілу атомів у об'ємі матеріалу тощо), тобто досягненню корисної мети, то в еволюційних методах оптимізації швидке досягнення впорядкованого стану досить часто розглядається як невдала спроба розв'язку задачі.

Розглянемо причини такого принципово різного ставлення до факту досягнення впорядкованого стану стосовно природних систем та еволюційних методів оптимізації на прикладі моделі мурашника.

Нехай нам вдалося створити ситуацію, в якій буде штучно забезпечена велика первинна концентрація феромону на маршруті, що взагалі не пов'язаний з джерелом харчування. У цьому разі мурахи завжди будуть обирати цей маршрут, незважаючи на те, що його харчова корисність

практично відсутня. З погляду термодинаміки цей приклад ілюструє той факт, що первинний розподіл флуктуацій забезпечив формування впорядкованого стану, а з огляду на ефективність розв'язання задачі оптимізації, він свідчить про те, що невелике розташування первинних пошукових точок може привести еволюційний процес до впорядкованого стану, який не відповідає оптимальному розв'язку задачі.

У другому прикладі розглянемо ситуацію, в якій буде штучно підтримуватись актуальність джерела їжі, до якого вже склався сталий маршрут, тобто застабілізуємо зовнішні умови, які відповідають впорядкованому стану системи маршрутів. У цьому випадку оптимальний маршрут буде постійно характеризуватись підвищеною концентрацією феромону, і навіть виявлення нового, більш привабливого джерела їжі не змусить фуражирів обирати новий маршрут руху. Для стійкої переорієнтації фуражирів на новий маршрут необхідно забезпечити більшу концентрацію феромону на ньому порівняно зі старим, а виникнення такої ситуації є малоімовірним. З точки зору термодинаміки цей приклад відповідає ситуації сталості зовнішніх умов, що унеможливорює самостійний перехід системи в інший стан, а з огляду на ефективність застосування еволюційних методів оптимізації, це свідчить про те, що в разі переходу системи в процесі еволюції до впорядкованого стану, який не відповідає розв'язку задачі оптимізації, перейти до пошуку іншого розв'язку за тих самих параметрів алгоритму вже не вдасться.

На перший погляд, наведені приклади є лише ілюстрацією можливих варіантів руху мурах у штучно створених умовах, які не можуть бути відтворені в природі. Однак, реальність ще більш вражає: мова вже йде не про загальні поняття типу „корисно” або „некорисно”, а про виникнення впорядкованості, яка призводить до масової загибелі мурах.

Виникнення згубної для мурах упорядкованості пов'язана з таким відомим у термодинаміці дисипативних систем поняттям, як „точка біфуркації”. Особливість цієї точки полягає в тому, що вона є нестійкою: малі флуктуації станів елементів системи в цій точці можуть спрямувати еволюцію суттєво різними шляхами, причому після переходу через точку біфуркації на одну з гілок еволюції перехід на іншу гілку є практично неможливим унаслідок завершеності відбору елементами системи таких станів, які несумісні з іншими варіантами еволюції системи. Перехід до

іншого стану потребує кардинальної зміни або самої системи, або зовнішніх умов.

Стосовно мурах перехід через точку біфуркації та виникнення згубної впорядкованості ілюструється добре відомим явищем „смертельного млина”, сутність якого полягає в формуванні феромонної петлі (замкнутої ділянки маршруту) під час руху мурах-кочівників, яке неодноразово спостерігалось біологами. На початку в місці перетину випадкових маршрутів (точці біфуркації) існує незначна флуктуація феромонів. Якщо ватажок колони випадково надасть перевагу запаху на феромонній петлі та перейде на неї, тоді вся колона також перейде на петлю, підсилюючи концентрацію феромону. В результаті концентрація феромону на петлі досягає такого рівня, що звернути з неї мурахи вже не в змозі, причому кожний прохід колони петлею тільки збільшує концентрацію феромону. Утворюється впорядкований маршрут – замкнуте коло, рух по якому призводить до загибелі мурах від виснаження.

Виходячи з цих прикладів, можна стверджувати, що досягнення під час застосування еволюційних методів оптимізації кінцевого етапу еволюції – впорядкованого стану – може бути пов’язане з відсутністю корисного результату, адже сам факт переходу в такий стан не гарантує знаходження оптимуму вихідної задачі оптимізації, а розташування точки, в якій еволюція завершилась, суттєво залежить від розташування точок вихідної множини, яке і визначає початкову флуктуацію станів елементів термодинамічної системи. Окрім того, продовжити пошук оптимуму вже практично неможливо внаслідок сталості умов роботи еволюційного алгоритму.

Отже, ефективність еволюційних алгоритмів пов’язується з їхньою спроможністю щодо довготривалого підтримання режиму „регульованого хаосу”, в результаті чого забезпечується збалансованість хаотичної пошукової активності в просторі незалежних змінних та цілеспрямованого пошуку на вже визначених перспективних напрямках з можливістю переорієнтації на нові напрямки пошуку. Якщо повернутися до моделі мурашника, це означає, що з точки зору пошуку оптимуму найбільшу цінність має пошукова активність розвідників, а не впорядкований рух фуражирів, адже перехід мурах у режим транспортування їжі або „смертельного млина” фактично означає, що формування маршрутів вже завершилось. Єдиною надією на виявлення іншого оптимуму залишається

хаотичний рух залишків розвідників, але за сталості загальної кількості мурах (у моделі) їх частка настільки зменшується, що сподіватись на їхній успіх не доводиться.

Розглянуті особливості еволюційних методів оптимізації як основний шлях підвищення їх ефективності визначають підтримання режиму „регульованого хаосу” для запобігання швидкому переходу штучно створеної дисипативної системи до впорядкованого стану.

Зазвичай, самі еволюційні методи вже містять деяку кількість параметрів, необхідних для настроювання пошукової активності алгоритмів, зокрема для ГА ними, наприклад, є імовірності використання операторів мутації, інверсії тощо. Однак, досить часто можливостей одних параметрів настроювання недостатньо для впевненого підтримання режиму „регульованого хаосу”. Тому, більш глибинним шляхом є вдосконалення самих методів з метою запобігання їхній передчасній збіжності.

Для досягнення цієї мети існує чотири основних методичних підходи: більш повне врахування особливостей природних процесів; внесення в правила та розрахункові залежності в межах самого методу таких модифікацій, які не мають природних аналогів; перенесення перспективних ідей та підходів з одного еволюційного методу до іншого, а також комбінування методів.

У подальшому розглянемо приклади вдосконалення відомих еволюційних методів оптимізації з використанням деяких з визначених підходів.

## **20.2. Удосконалений генетичний алгоритм**

Необхідність удосконалення генетичних алгоритмів для підвищення їх пошукової активності зумовлена тим, що невдалий розв’язок задачі оптимізації вимагає проведення повного повторного циклу розрахунків. При цьому цілком втрачається вся корисна інформація (тобто популяція починає еволюціонувати від самого початку) і немає ніякої гарантії, що новий цикл розрахунків буде набагато кращим, ніж попередній.

Основна ідея вдосконалення генетичних алгоритмів полягає в тому, щоб у разі передчасної збіжності алгоритму оновити генетичний матеріал

популяції та примусити продовжити дослідження нових ділянок простору незалежних змінних без повторного його перезапуску.

З цією метою з визначених методичних підходів було обрано два напрямки: більш повне врахування особливостей природних процесів та внесення в правила та розрахункові залежності в межах самого методу таких модифікацій, які не мають природних аналогів.

Для більш повного врахування особливостей природних процесів було звернуто увагу на той факт, що в природі (на відміну від генетичних алгоритмів) у формуванні нового покоління можуть брати участь особини не тільки попереднього покоління, а й більш старших поколінь. Використання цього факту надає можливість оновити склад генетичного матеріалу в разі його збіднення.

Внесення в правила та розрахункові залежності в межах самого методу таких модифікацій, які не мають природних аналогів, було реалізоване за рахунок застосування до випадково обраної частини популяції нового генетичного оператора – оператора зсуву (рис. 19.12), причому з імовірністю 0,5 зсув здійснюється або праворуч, або ліворуч.

Як показник динаміки розвитку популяції обрана кількість поколінь, протягом якої екстремальне значення функції пристосованості залишається сталим ( $N\_Stal$ ).

Хромосома до зсуву	1	1	0	0	0
Хромосома після зсуву	0	1	1	0	0

Рис. 20.1. Схема оператора зсуву (праворуч)

Схема вдосконаленого ГА, який враховує вище викладені міркування, наведена на рис. 20.2 .

На першому кроці формується початкова популяція, еволюція якої за базовим ГА забезпечує знаходження деякого значення функції пристосованості.

Якщо функція пристосованості протягом визначеної кількості поколінь ( $N\_Turbo$ ) не покращується, вважається, що вона потрапила до локального оптимуму. Тоді на другому кроці оновлення популяції здійснюється із залученням для формування хромосоми нащадка одного з батьків з

попередніх поколінь (як номер покоління, так і особина в ньому визначаються випадково).

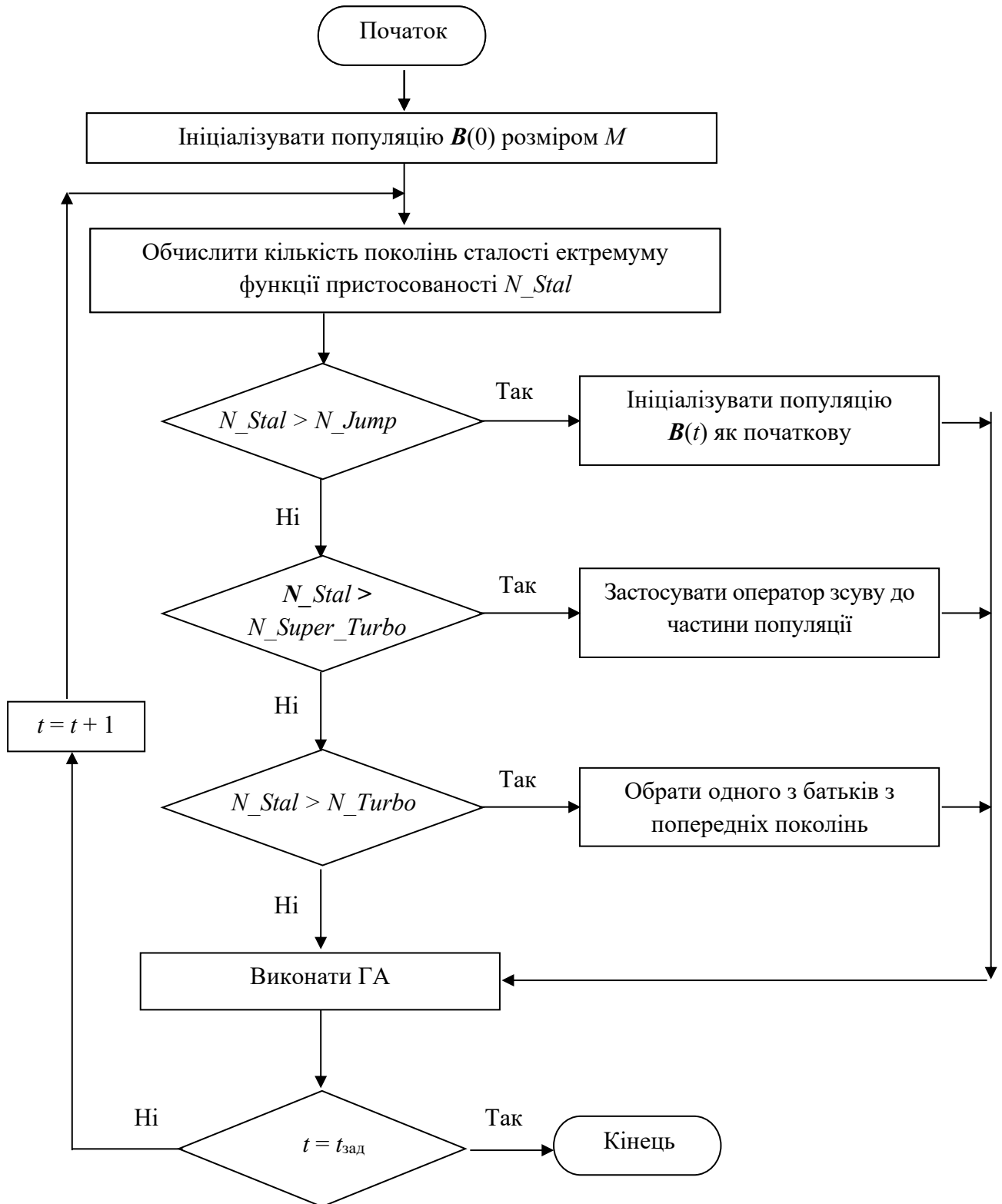


Рис. 20.2. Схема вдосконаленого генетичного алгоритму

Таке оновлення генетичного матеріалу відбувається протягом декількох поколінь, причому у випадку поліпшення значення функції пристосованості у подальшому здійснюється базовий ГА. Якщо позитивної динаміки протягом  $N\_Super\_Turbo$  поколінь досягти не вдалося, тоді виконується третій крок.

Якщо на другому кроці для оновлення генетичного матеріалу використовувалась еволюційна “вертикаль”, то на третьому кроці з цією самою метою використовується еволюційна “горизонталь”. Це досягається шляхом застосування до випадково обраної частини популяції оператора зсуву, що має наслідком перенос частини популяції у нову ділянку простору змінних.

Так само, як і другий крок, третій крок виконується протягом декількох поколінь. Якщо протягом  $N\_Jump$  поколінь функція пристосованості покращилась, то в подальшому здійснюється базовий ГА. У іншому разі переходимо до першого кроку, тобто популяція ініціалізується як початкова.

Під час реалізації зазначеного генетичного алгоритму доцільно використовувати елітизм, тобто примусове перенесення найкращої особини в наступне покоління, а також турнірний відбір для формування нової хромосоми, при цьому на другому кроці вдосконаленого алгоритму турнір для одного з батьків проводиться серед особин попереднього покоління.

Таким чином, запропонований варіант генетичних алгоритмів є свого роду надбудовою над загально використовуваними генетичними алгоритмами (зокрема стосовно розглянутого вище базового), що дозволяє в його межах реалізувати будь-яку комбінацію генетичних операторів.

### **20.3. Удосконалений метод оптимізації роєм частинок**

Передумови, які зумовлюють необхідність удосконалення методу оптимізації роєм частинок, аналогічні тим, які були розглянуті стосовно генетичних алгоритмів.

Під час удосконалення методу оптимізації роєм частинок були використані практично всі визначені методичні підходи, зокрема такі.

По-перше, метод оптимізації роєм частинок не враховує періодичне часткове оновлення популяції за рахунок загибелі та розмноження особин. Тому, першим напрямком його вдосконалення може бути врахування цього факту завдяки заміні деякої частини рою (імітація загибелі) новими

частинками, характеристики яких отримані на основі положень генетичних алгоритмів (імітація розмноження).

По-друге, поза увагою методу оптимізації роєм частинок залишився той факт, що деяка загальна інформація (наприклад, про місця зимівлі) може використовуватись декількома популяціями, ареали існування яких різняться за географією розташування. Тому, другим напрямком удосконалення цього методу може бути послідовне застосування декількох незалежних популяцій, які використовують найкращу точку як загальнодоступну інформацію.

На рис. 20.3 наведено схему вдосконаленого методу оптимізації роєм частинок, відповідно до якої розрахунки проводяться циклічно в декілька етапів.

На першому етапі (блок 1) здійснюють формування вихідних даних, до яких крім даних, необхідних для розрахунків за стандартним методом оптимізації роєм частинок, додатково належать: час життя одного рою ( $T$ ), кількість особин рою ( $p$ ), що підлягає оновленню в кожному поколінні, та загальна кількість роїв ( $K$ ).

Другим етапом (блок 2) здійснюють формування першого рою, при цьому під формуванням мається на увазі визначення для кожної частинки початкового положення в просторі незалежних змінних.

На третьому етапі (блок 3) для всіх частинок рою обчислюють значення функції пристосованості (*fitness*). При цьому, початкова координата кожної частинки вважається когнітивною складовою її швидкості, а найкраща серед усіх частинок когнітивна складова розглядається як початкове значення соціальної складової.

Четвертий етап (блоки 5–9) фактично повторює розрахунки, що відповідають стандартному методу оптимізації роєм частинок.

На п'ятому етапі (блок 10) здійснюють часткове оновлення поточної популяції з використанням генетичних алгоритмів. Для цього  $p$  разів використовується турнірний відбір двох батьків, з яких оператором кросовера формується нова особина. Отримана в такий спосіб особина заміщує в популяції найгіршу особину з тих, які брали участь у турнірах.

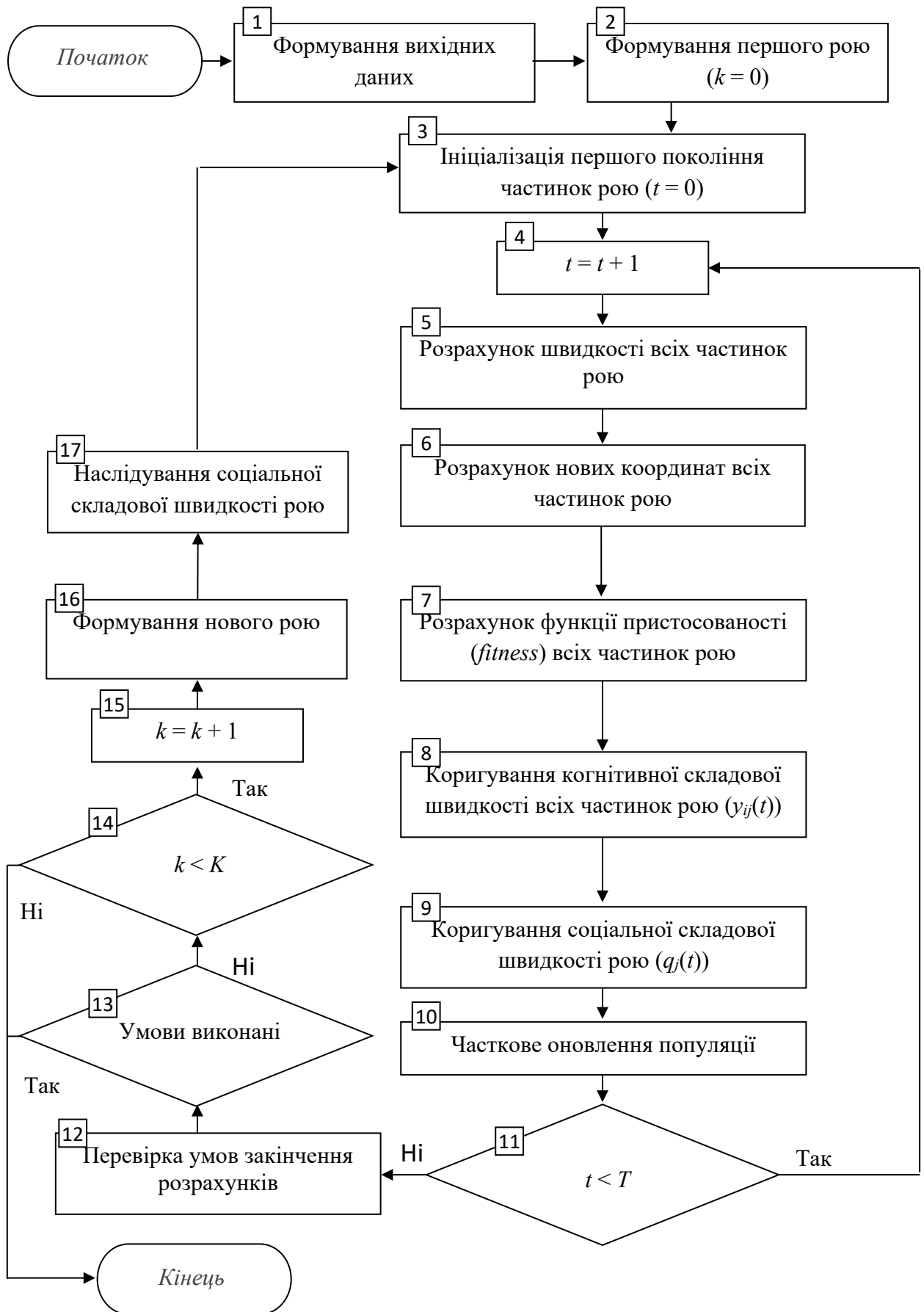


Рис. 20.3. Схема вдосконаленого методу оптимізації роєм частинок

Четвертий та п'ятий етапи виконують циклічно задану кількість разів, яка відповідає часу життя одного рою  $T$  (блок 11), а шостий етап (блоки 12 та 13) передбачає оцінювання умов завершення розрахунків.

У разі, коли лічильник кількості популяцій ( $k$ ) не досяг граничного значення ( $K$ ) (блок 15), переходять до сьомого етапу (блоки 16 та 17), на якому за аналогією з другим етапом (блок 2) формують новий рій та примусово встановлюють початкове значення його соціальної складової швидкості як значення соціальної складової швидкості останнього покоління попереднього рою.

Удосконалений метод був реалізований програмно та було проведено розв'язання тестової задачі (табл. 19.1).

Таблиця 20.1

**Результати розрахунків тестової задачі [7] за різними методами оптимізації**

Метод розв'язання	Збиток противника, од. важливості	Відносні результати
Максимального елемента	612,60	0,93
Удосконалений генетичний алгоритм	653,46	0,99
Метод оптимізації роєм частинок	618,61	0,94
Удосконалений метод оптимізації роєм частинок	658,43	1

Аналіз результатів тестових розрахунків (табл. 19.1) засвідчив, що запропоновані вдосконалення дозволили підвищити точність стандартного методу оптимізації роєм частинок до рівня, сумірного з удосконаленим ГА.

**20.4. Удосконалений метод оптимізації колонією бджіл**

Передумови, які зумовлюють необхідність удосконалення методу оптимізації колонією бджіл, аналогічні тим, які були розглянуті стосовно генетичних алгоритмів та методу оптимізації роєм частинок.

Під час розроблення вдосконаленого методу оптимізації колонією бджіл також були використані практично всі наведені вище методичні підходи.

Аналіз концептуальних положень, покладених в основу методу оптимізації колонією бджіл, показав, що існують деякі аспекти поведінки

бджіл під час пошуку джерел їжі, які можуть бути враховані додатково. Зокрема, бджоли-розвідники як один з орієнтирів використовують інтенсивність руху інших бджіл, тобто використовують інформацію від інших колоній.

Тому першим напрямком удосконалення методу оптимізації колонією бджіл може бути послідовне формування декількох колоній, які використовують найкращу точку попередньої колонії як загальнодоступну вихідну інформацію.

Окрім того, порівняння формул для розрахунку нових координат розміщення бджоли з формулою для визначення швидкості частинки за методом оптимізації роєм частинок показав, що вона є частковим випадком варіанта  $l_{best}$  (для  $w = c_1 = 0$ ), у якому соціальні знання частинки пов'язуються з сусідами.

Тому другим напрямком удосконалення методу оптимізації роєм бджіл може бути використання в ньому для розрахунку нових координат розміщення бджоли формул, запозичених з методу оптимізації роєм частинок.

Необхідно зазначити також, що ці формули забезпечують можливість урахування під час оптимізації не двох масштабів пошуку оптимуму, як це має місце в методі оптимізації колонією бджіл (глобального та локального), а трьох: глобального (розвідниками, які орієнтуються на найкращу точку поточної колонії), середнього (робочими бджолами, які орієнтуються на найкращі точки розвідників, які їх завербували) та локального (вільними бджолами, які орієнтуються на найкращу точку робочої бджоли, яка їх завербувала).

Тому третім напрямком удосконалення методу оптимізації роєм бджіл може бути використання трьох масштабів пошуку оптимуму. При цьому кожен розвідник вербує собі робочих бджіл, які, в свою чергу, вербують собі вільних. Найкраща соціальна складова швидкості передається розвідникам, які використовують її на наступному кроці пошуку.

Окремого розгляду потребує залежність, яка встановлює пропорційність між кількістю завербованих вільних бджіл та перспективністю  $j$ -го джерела їжі ( $fitness_j$ ) через відношення

$$fitness_j / \sum_{j=1}^{N_{роб}} fitness_j .$$

За такого використання величин  $fitness_j$  на вигляд  $fitness$ -функції, що підлягає оптимізації, накладається додаткове обмеження виду  $fitness_j > 0$ , порушення якого може призводити до від'ємних значень  $N_{j\text{ вил}}$  та/або діленню на 0. Забезпечення уникнення таких ситуацій суттєво ускладнює процес переходу від вихідної задачі оптимізації до  $fitness$ -функції.

Тому четвертий напрямок удосконалення методу оптимізації роєм бджіл передбачає введення до розгляду додаткової величини – корисності ділянок з джерелами їжі  $W_j$ , граничні значення для якої встановлюються, наприклад, у такий спосіб:

- для ділянки з найкращим значенням  $fitness$ -функції ( $fitness_{max}$ )  $W_{max} = 0,9$ ;

- для ділянки з найгіршим значенням  $fitness$ -функції ( $fitness_{min}$ )  $W_{min} = 0,1$ .

Тоді  $W_j$  обчислюється за формулою:

$$W_j = W_{min} + \frac{(W_{max} - W_{min})}{(fitness_{max} - fitness_{min})} fitness_j ,$$

а кількість завербованих вільних (або робочих) бджіл для  $j$ -ї ділянки визначається як

$$N_{j\text{ вил}} = N_{вил} \frac{W_j}{\sum_{j=1}^{N_{роб}} W_j} .$$

Схема вдосконаленого методу оптимізації роєм бджіл наведена на рис. 20.4, відповідно до якої розрахунки проводяться циклічно в декілька етапів.

На першому етапі (блок 1) здійснюють формування вихідних даних, які додатково до базового методу оптимізації роєм бджіл містять дані, необхідні для розрахунку швидкості відповідно до методу оптимізації роєм частинок, а також час життя одного рою ( $T$ ) та загальну кількість роїв ( $K$ ).

На другому етапі (блок 2) відбувається формування першого рою, що передбачає визначення для всіх бджіл координат вихідного положення.

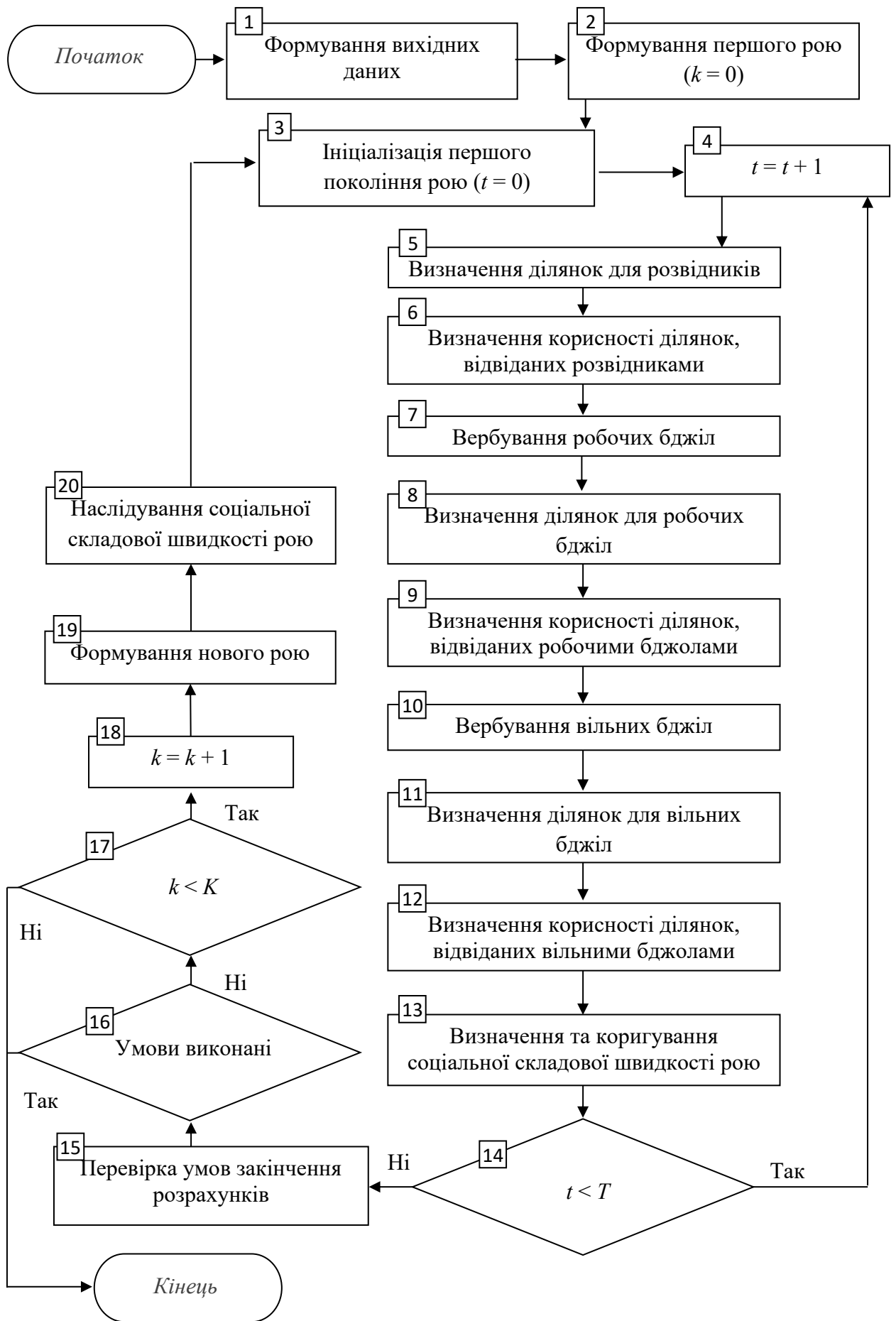


Рис. 20.4. Схема вдосконаленого методу оптимізації роєм бджіл

На третьому етапі (блок 3) для першого покоління рою ініціалізують значення швидкості всіх бджіл, обчислюють нові положення всіх бджіл та на основі значень *fitness*-функції для них визначають величину соціальної складової швидкості для розвідників.

Четвертий етап (блоки 4–14) передбачає циклічне повторення дій, пов'язаних з пошуком оптимуму одним роєм.

У блоці 5 визначають ділянки, які будуть відвідані розвідниками, а в блоці 6 оцінюється корисність цих ділянок та коригування (за необхідності) когнітивної складової швидкості кожного розвідника.

У блоці 7 відбувається вербування розвідниками робочих бджіл, при цьому кількість завербованих бджіл кожним розвідником пропорційна корисності відвіданої ним ділянки, а його когнітивна складова швидкості передається завербованим ним бджолам як соціальна складова їх швидкості.

У блоках 8 та 9 стосовно робочих бджіл здійснюються дії, аналогічні блокам 5 та 6 відповідно, а в блоці 10 відбувається вербування робочими бджолами вільних бджіл з відповідним передаванням їм як соціальної складової їх швидкості когнітивної складової швидкості відповідних робочих бджіл (аналогічно блоку 7).

У блоках 11 та 12 стосовно вільних бджіл здійснюються дії, аналогічні блокам 5 та 6 відповідно.

У блоці 13 визначається та (за необхідності) коригується значення соціальної складової швидкості усього рою, яке буде використане розвідниками наступного покоління або першого покоління нового рою.

Блоки 15–20 призначені для оцінювання умов завершення розрахунків та формування циклу розрахунків для декількох роїв.

Удосконалений метод оптимізації роєм бджіл було реалізовано програмно та проведено розв'язання тестової задачі з (табл. 20.2).

Аналіз результатів, наведених у табл. 20.2, свідчить, що точність удосконаленого методу оптимізації колонією бджіл підвищилась до рівня вдосконаленого методу оптимізації роєм частинок, при цьому час розрахунків порівняно з ним зменшився приблизно в 3–4 рази.

## Результати розрахунків задачі з [7] за різними методами оптимізації

Метод розв'язання	Збиток противника, од. важливості	Відносні результати	Час розрахунків
Максимального елемента	612,60	0,929	-
Удосконалений метод оптимізації роєм частинок	658,43	0,998	5 хв. 30 с
Оптимізації колонією бджіл	521,58	0,791	1 хв. 36 с
Удосконалений метод оптимізації колонією бджіл	659,39	1	1 хв. 38 с

## Контрольні запитання та завдання

1. Які ідеї покладено в основу аналізу напрямків удосконалення еволюційних методів оптимізації ?
2. В чому полягає фізичний зміст удосконаленого генетичного алгоритму?
3. Як математично реалізовується удосконалений генетичний алгоритм ?
4. В чому полягає фізичний зміст удосконаленого методу оптимізації колонією бджіл?
5. Як алгоритмічно реалізовується удосконалений метод оптимізації колонією бджіл?
6. В чому полягає фізичний зміст удосконаленого методу оптимізації роєм частинок?
7. Як алгоритмічно реалізовується удосконалений метод оптимізації роєм частинок?

### Розділ 3. Приклади практичного застосування математичних методів оптимізації в телекомунікаціях та радіотехніці

#### Глава 21. Моделі побудови та функціонування епізодичних радіомереж із БПЛА в зоні надзвичайної ситуації

У даній главі проведено аналіз особливостей побудови та функціонування епізодичних радіомереж на основі телекомунікаційних аероплатформ (ТА) в зоні надзвичайної ситуації в умовах швидкого і непередбачуваного переміщення мобільних абонентів (рятувальників). Крім того проведено аналіз задач управління мережею телекомунікаційних аероплатформ та методів підвищення пропускної здатності епізодичних радіомереж (ЕРМ) із застосуванням безпілотних літальних апаратів (БПЛА). Визначено ключову задачу дослідження та зроблено її постановку у загальному вигляді.

##### 21.1. Особливості побудови та функціонування епізодичних радіомереж на основі БПЛА

Приклади побудови архітектури із використанням БПЛА, що функціонують в якості телекомунікаційної аероплатформи (ТА), вказані нижче (рис. 21.1, рис. 21.2).

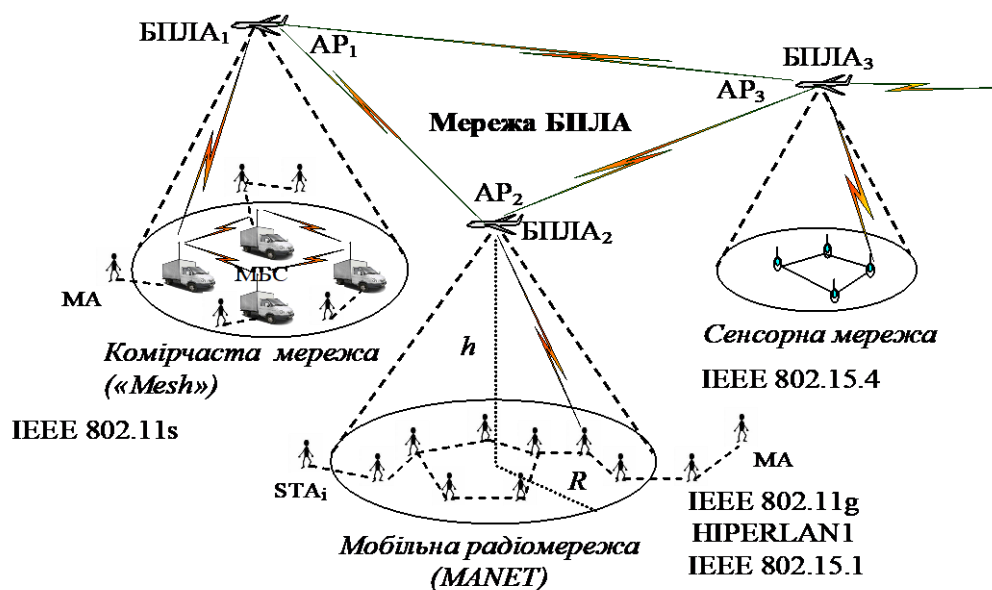


Рис. 21.1. Приклад архітектури дворівневої ЕРМ на основі БПЛА

Перший (наземний) рівень складають віддалені райони наземних вузлів, роль яких можуть грати мобільні абоненти (МА), мобільні базові станції (МБС) або сенсорні пристрої в залежності від виду ЕРМ (MANET, WMN або WSN).

Другий (повітряний) складає мережі БПЛА, що виконує роль опорної мережі для забезпечення зв'язності віддалених незв'язаних районів наземних вузлів (рис. 21.1) або підвищення зв'язності «проблемних» ділянок наземної мережі (наприклад, між вузлами 3 та 4, як показано на рис. 21.2).

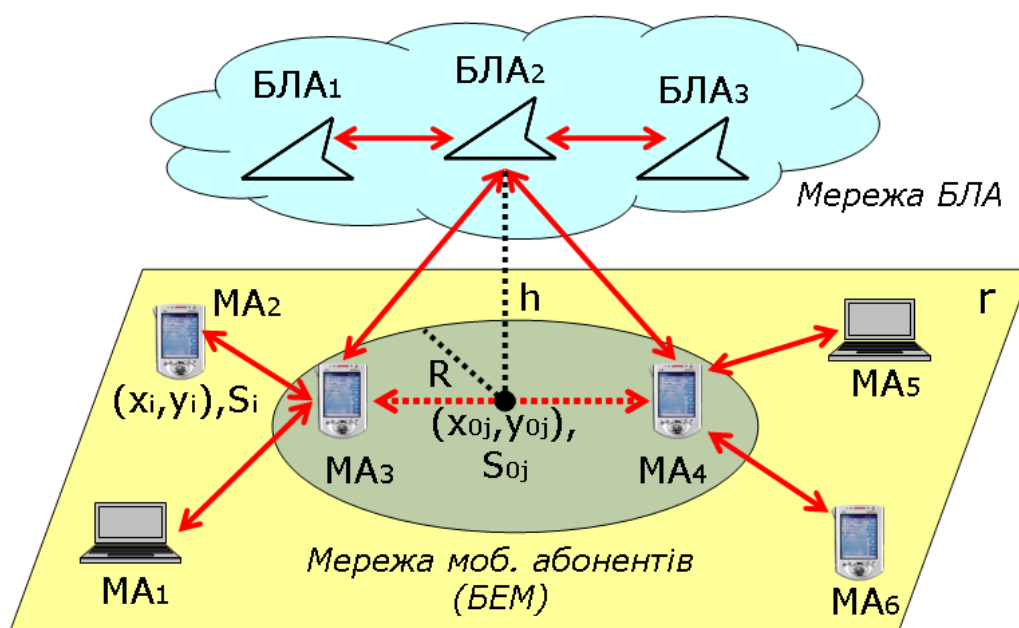


Рис. 21.2. Приклад застосування БПЛА для забезпечення зв'язності ЕРМ

Кожен БПЛА-вузол ЕРМ оснащується двома комплектами прийомо-передавальної апаратури та антенних систем (для зв'язку з наземними абонентами та для зв'язку між БПЛА), мережним процесором (маршрутизатором), буферним запам'ятовуючим пристроєм, GPS-навігатором. Виділяють 4 рівні функціональних можливостей БПЛА-ретрансляторів:

- БПЛА – шлюз;
- БПЛА – маршрутизатор;
- БПЛА – міст;
- БПЛА – комутатор.

Найбільш складну структуру має БПЛА – шлюз. Він забезпечує виконання на борту широкого переліку операцій з обробки сигналів та

повідомлень та може виступати в якості міжмережевого перетворювача. БПЛА – маршрутизатор та БПЛА – міст забезпечують обробку та ретрансляцію сигналів та повідомлень в межах однієї інформаційної мережі. При цьому БПЛА – маршрутизатор здійснює маршрутизацію повідомлень (пакетів), що передаються, у відповідності із збереженим на ньому альманасі мережі та користувачів. Найбільш просту структуру має БПЛА – комутатор, який виконує окремі операції обробки сигналів (підсилення, перенос в іншу обсть частотного спектру, корекцію крутизни фронтів). Найбільш раціональним варіантом побудови мережі, коли кожен БПЛА – вузол буде поєднувати усі вищевказані рівні функціональних можливостей.

Абонентами ЕРМ є стаціонарні та рухомі об'єкти (рис. 21.1):

Мобільні базові станції (МБС). Вони утворюють опорні мережі в різних географічних зонах (бойових дій, стихійних лих тощо), зв'язність між якими забезпечує мережа БПЛА.

Мобільні абоненти (МА). Мобільними абонентами можуть виступати члени пошуково-рятувальних бригад Державної служби надзвичайних ситуацій (ДСНС), особового складу військових підрозділів тощо.

Сенсорні прилади.

Усі рухомі та стаціонарні пристрої у свою чергу мають бути оснащені:

- двома комплектами прийомо-передавальної апаратури (для зв'язку між наземними абонентами та для зв'язку з БПЛА). Застосування двох окремих радіо інтерфейсів дозволяє розсилати маршрутні пакети, як через наземну мережу, так і через повітряну мережу БПЛА, та отримувати більшу кількість незалежних маршрутів між парою вузлів. Іншою перевагою застосування окремих радіо інтерфейсів є можливість використання різних частотних смуг для кожного інтерфейсу, що забезпечує кращу якість послуг (QoS) та баланс навантаження в мережі.

- мережним процесором (маршрутизатором), для здійснення ретрансляції пакетів, призначених іншим вузлам.

- буферним пристроєм (якщо вузли мережі підтримують режим енергозбереження, тобто управління потужністю випромінювання).

- засоби перетворення інформації у зручний для користувача вид (голос, дані, відео тощо), оскільки кожен наземний пристрій грає роль не тільки вузла (ретранслятора), а й кінцевого (термінального) устаткування.

В ЕРМ на основі БПЛА можуть бути використані наступні методи передачі інформації між наземними вузлами:

1) безпосередня передача повідомлення між наземними вузлами (без застосування БПЛА).

2) безпосередня ретрансляція повідомлення одержувачу через БПЛА – у випадку, якщо і джерело інформації, і одержувач знаходяться в межах зони видимості одного БПЛА-вузла;

3) зв'язок через проміжні наземні станції ретрансляції, що називають базовими, при якій кожне повідомлення, що поступило у БПЛА-вузол, обов'язково передається на базову станцію, що визначає подальший маршрут проходження повідомлення;

4) електронна пошта (ретрансляція з перенесенням в запам'ятовуючий пристрій БПЛА). В даному випадку повідомлення приймається БПЛА-вузлом від джерела інформації, зберігається в його запам'ятовуючому пристрої, та «скидається» споживачеві під час прольоту в зоні видимості;

5) зв'язок з використанням міжстанційних каналів між БПЛА.

Найбільш перспективним є поєднання 1,2 та 5-го методів. Воно дозволяє забезпечити глобальний зв'язок між наземними вузлами. При цьому можливі два варіанти організації каналів. У першому випадку інформація між двома наземними вузлами передається через мережу з багатократним перевипромінюванням БПЛА-вузлами по радіоканалах, що створюються тільки на час передачі інформації. У другому випадку між всіма вузлами мережі, які "бачать" один одного, організуються магістральні канали, що існують під час взаємної видимості, по яких, в міру необхідності, передається інформація. Кожний з цих варіантів має свої достоїнства і недоліки.

У першому випадку: простота реалізації, можливість використання одного комплексу прийомо-передавальної апаратури БПЛА-вузла і для обміну з абонентами, і для обміну з сусідніми вузлами. Економне використання ресурсу каналів поєднується з ускладненням режиму входження в зв'язок і труднощами при багатократній організації множинного доступу. Такий метод достатньо ефективний при малому та середньому трафіку (вхідному навантаженню мережі) і, як буде показано нижче, неефективний при великому трафіку.

У разі великого трафіку доцільна організація магістральних каналів. При цьому різке збільшення пропускної спроможності досягається шляхом

п'яти-шести кратного збільшення об'єму прийомо-передавальної апаратури на борту БПЛА-вузла, що у свою чергу приводить до необхідності усунення взаємного впливу магістральних каналів один на одного, а також взаємних перешкод пакетів усередині магістральних каналів.

*Особливості функціонування ЕРМ на основі БПЛА, перш за все динаміка та змінна топологія мережі (як наземних вузлів, так і мережі БПЛА), приводять при проектуванні до ряду проблем, до яких відносяться:*

- реалізація пакетної передачі інформаційних повідомлень;
- ретрансляція пакетів через проміжні вузли;
- організація множинного доступу абонентів до мережі;
- визначення маршрутів для передачі інформації по мережі;
- організація каналів по вибраному маршруту;
- управління топологією (місцеположенням) мережі БПЛА.

При *пакетній передачі інформації* кожне повідомлення розбивається на інформаційні пакети, що містять преамбулу та інформаційну частину. Преамбула складається з синхропосилки, що забезпечує підготовку приймального пристрою до прийому пакету (частотна, просторова і тимчасова синхронізація та автосупровід), і службової частини (у якій містяться відомості про номер пакету, адресата, маршрут проходження, види обробки та ін.). Не дивлячись на збільшення загального числа символів, що передаються, пакетний метод передачі інформації дозволяє підвищити пропускну здатність системи за рахунок одночасної організації декількох віртуальних ліній зв'язку, обслуговуючих одночасно велику кількість абонентів.

*Організація користування мережею* великою кількістю наземних вузлів та взаємодії БПЛА-вузлів з ними та один з одним здійснюється відповідно до правил, що називають протоколами множинного доступу (МД). Ці протоколи можна розділити на: детерміновані, випадкові та гідридні [1].

Перший клас протоколів упорядковує роботу абонентів та вузлів таким чином, що повністю усуває конфлікти, при яких два або більш абонентів (вузлів) одночасно передають повідомлення одному й тому ж вузлу. В таких протоколах розділення здійснюється шляхом статичного або динамічного закріплення частотного (FDMA), часового (TDMA), кодового (CDMA), просторового (SDMA) або гібридного (TDMA/CDMA, STDMA і т.п.) ресурсу

каналів за абонентами і вузлами та вимагає високої організації управління мережею.

Другий клас протоколів реалізується простіше і допускає виникнення конфліктів, тобто зіткнення пакетів. Історично першим вважається метод АЛОНА. Він ефективний у випадках невеликого завантаження мережі, коли втрати частини пропускної здатності системи через зіткнення пакетів виявляються меншими за втрати, що пов'язані з простоем ресурсу. При великому трафіку виникає ефект «нерегульованого перехрестя», тобто «пробки», коли вузол не в змозі ні прийняти, ні передати пакет далі. Частково ця проблема вирішується за рахунок тактування випадкового доступу, при якому всі абоненти і вузли починають випромінювати пакети в певний час з настанням чергового такту (S-ALOHA). Часовий інтервал між тактами покриває тривалість пакету і максимальний час його розповсюдження в просторі. Тоді при одночасному зверненні до вузла декількох абонентів буде прийнятий пакет від найближчого абонента, той, що поступив раніше з більш потужним сигналом, а решта пакетів буде відхилена.

Для зниження ймовірності зіткнення пакетів застосовуються методи доступу з контролем несучої (МДКН) або CSMA, тобто методи, що засновані на попередній перевірці стану радіоканалу. В простішому випадку перед передачею пакету вузол контролює стан каналу (наявність несучої або самої передачі). Якщо канал зайнятий, вузол відкладає передачу на більш пізній час. При звільненні каналу його передача може починатися різноманітними способами: відразу («жорсткий» МДКН), через випадковий інтервал часу («м'який»), з ймовірністю  $p$  ( $p$ -наполегливий) або з рандомізацією часу передачі на сегменти .

Для вирішення проблем «відкритого» та «схованого» терміналу запропонований ряд методів (протоколів) МД з контролем несучої та запобіганням зіткнень (CSMA/CA). Найбільш відомі – це MACA, MACA-BI, FAMA, MACAW, DBTMA та IEEE 802.11 DCF. При застосуванні перелічених методів канал тимчасово резервується на період передачі інформаційного пакету за допомогою діалогу коротких службових пакетів між відправником та одержувачем: запит відправника на передачу (RTS) та згода одержувача (CTS).

Загальна характеристика протоколів канального рівня, запропонованих для використання в ЕРМ, наведена в табл. 21.1.

Аналіз таблиці дозволяє зробити висновок, що в умовах високої мобільності вузлів, характерної для більшості ЕРМ, переважним є протокол IEEE 802.11, хоча він й далекий від досконалості. До того ж дослідження засвідчують про можливість застосування протоколу IEEE 802.11 не тільки в будівлі, а й зовні. При радіусі стільника 6 км функціонування МАС протоколу відповідає усім нормам стандарту 802.11.

Таблиця 21.1

**Основні характеристики існуючих протоколів каналного рівня  
безпроводових мереж**

Основні характеристики	Протокол каналного рівня			
	IEEE 802.11	IEEE 802.16	HiperLAN2	Bluetooth
Частота, ГГц	2.4 / 5.1	2 – 66	5.1	2.4
Відстань, м	До 500 / до 100	100 – 20000	До 250	10...100
Швидкість передачі в каналі, Мбіт/с	1 / 2 / 11 / 54	120	54	0.7...1
Метод доступу до каналу	DFWMAC (CSMA/CA)	OFDMA/TDMA / TDD	Polling / TDD	
Тип управління, організація мережі	Децентралізована, всі вузли одного рівня	Зонове (кластеризація мережі), централізоване управління ресурсами в кожній зоні (пікомережі)		
Мобільність вузлів	Легко реалізується на мережевому рівні	Ускладнена, викликає необхідність перебудови зон мережі		
Пропускна спроможність мережі	Обмежена взаємними завадами	Визначається конфігурацією мережі, час реконфігурації мережі – значний		

Окрім цього на рівні доступу (MAC) повинні вирішуватися і такі важливі задачі як синхронізація, управління енергозбереженням, роумінг, аутентифікація та шифрування та ін. Особливу увагу слід привернути до роумінгу, тобто процедурі входу наземного вузла у зону обслуговування БПЛА, перехід від однієї зони БПЛА до іншої, та переключення між зонами БПЛА при знаходженні вузла одночасно у декілька зонах обслуговування.

*Визначення маршруту* для кожного пакету є складною та відповідальною операцією і може здійснюватися як бортовими комплексами БПЛА, так і наземними вузлами. Відзначимо, що маршрутизатор БПЛА-вузла має ті ж самі функціональні властивості, що й маршрутизатор будь-якого наземного вузла. Це дозволяє використовувати існуючі протоколи маршрутизації, що розроблені для MANET, та, ще більш важливо, відмовитись від потреби допоміжного протоколу, що координує прибуття та відліт БПЛА з протоколом маршрутизації.

В теперішній час для використання в EPM запропонований ряд протоколів маршрутизації, які по способу побудови та підтримки маршруту класифікуються на дві основні групи :

- таблично-орієнтовні (DSDV, WRP, OLSR та ін.);
- зондові (DSR, AODV, TORA та ін.);
- гібридні (ZRP, R-зоновий та ін.)

При застосуванні таблично-орієнтовних методів кожен вузол постійно – на основі періодичного обміну вмістом маршрутних таблиць із сусідніми вузлами – будує маршрутну таблицю, що вміщає найкоротші шляхи (по критерію мінімальної вартості, де під вартістю може розумітися число ретрансляцій, час доставки, затрачена потужність передавача та ін.) до усіх інших вузлів.

При зондовій маршрутизації вузли формують маршрут передачі інформації по мірі необхідності («on-demand»), шляхом хвильової розсилки по мережі службових пакетів малого розміру (зондів-запитів) та збору зондів-відповідей, що вміщують інформацію про найкоротші маршрути (за вибраним критерієм) передачі інформації. Ця розсилка продовжується до тих пір, поки маршрут до необхідного вузла не буде встановлено, та знов відновлюється, коли встановлений маршрут втрачається. Такі протоколи ідеально підходять для застосування у EPM на основі БПЛА, оскільки

підтримують усі можливі випадки використання БПЛА для підвищення зв'язності наземних вузлів. Переміщення БПЛА у зазначене місце (наприклад, район «проблемної» гілки, як показано на рис. 21.2) еквівалентне введенню в мережу нового вузла зв'язку. При пошуку маршрутів між вузлами лівої (МА1, МА2, МА3) та правої (МА4, МА5, МА6) частини мережі шляхом розсилки зондів-запитів буде встановлено маршрути із використання БПЛА. Коли «проблемна» гілка мережі буде відновлена (тобто відбудеться злиття лівої та правої частини мережі), БПЛА може легко покинути заданий район баражування, а втрачені мережні маршрути завдяки «on-demand» природі можуть бути легко відтворені через проміжні наземні вузли.

*Організація каналу по маршруту передачі інформації* може здійснюватися двома способами: шляхом комутації каналів або шляхом комутації пакетів.

Перший спосіб полягає в створенні каналу «від кінця до кінця», з подальшою передачею повідомлення по всьому маршруту без розбиття на пакети. Недоліком такого способу є неможливість використання задіяних БПЛА-вузлів іншими абонентами, як під час прокладки маршруту, так і під час передачі повідомлення, а висока динаміка топології мережі швидко руйнує побудований канал.

Комутація пакетів має два основні різновиди [6]:

- комутація пакетів, при якій окремі пакети по мережі передаються незалежно по різних шляхах та «зшиваються» в повідомлення безпосередньо у абонента-одержувача інформації (це так званий датаграмний режим);

- комутація віртуальних каналів, при якій всі пакети одного повідомлення досягають одержувача по одному маршруту, визначуваному на етапі встановлення віртуального каналу.

Принципова відмінність комутації каналів від комутації віртуальних каналів полягає в наступному:

- у першому випадку канал недоступний для інших абонентів, навіть якщо в деякий момент часу по ньому нічого не передається

- у другому випадку має місце статистичне розділення каналів або поодинокі віртуальному каналу, або по окремих його ділянках можуть передаватися інформаційні пакети різних абонентів.

Комутація пакетів володіє рядом переваг. Це ефективне використання пропускної здатності мережі, висока надійність і живучість, пов'язані з

можливістю використання альтернативних маршрутів при виході з ладу окремих вузлів мережі або зіткненні пакетів. Проте реалізація переваг цього методу досягається тільки при відповідному управлінні інформаційними потоками, яке дозволяє забезпечити оптимізацію маршрутів і рівномірний розподіл навантаження на мережу.

При проектуванні цих систем характерне прагнення забезпечити широкі можливості обслуговування абонентів та велику кількість каналів, що одночасно діють, за рахунок значного ускладнення бортової апаратури БПЛА-ретрансляторів, що розробляється на основі нових інформаційних технологій. Так, наприклад, будуть вирішуватися завдання: передача будь-яких даних в цифровій формі – текст, зображення, мова, обмін інформацією між базами даних, між комп'ютерами, збір даних від датчиків контролю екологічної інформації, навколишнього середовища, технологічних систем контролю. При цьому забезпечується можливість використання достатньо простих терміналів, що не вимагають висококваліфікованого обслуговування, як стаціонарними, так і рухомими споживачами.

В цілому ЕРМ передачі інформації з використанням БПЛА характеризуються значно вищою складністю і рівнем організації, чим існуючі системи передачі інформації.

Основними особливостями ЕРМ на основі БПЛА, що існують з погляду їх аналізу і синтезу, є:

1. Складність структури мережі. Інформація передається пакетами через ретранслятори - вузли мережі. Радіоканали між джерелами інформації і БПЛА-вузлом, між БПЛА-вузлами та між БПЛА-вузлом і споживачем інформації самі є складними системами передачі інформації.

2. Динамічність топології мережі, тому безперечні переваги ЕРМ на основі БПЛА мають і зворотну сторону:

- нестійка зв'язність;
- складність маршрутизації;
- велика чутливість стану до дії перешкод;
- наявність доплерівських зсувів несучих частот;

3. Функціонування таких складних мереж не можливе без ефективної системи управління, складовою частиною якої є підсистема управління мережею БПЛА.

## **21.2. Аналіз задач управління мережею БПЛА в зоні надзвичайної ситуації**

За етапами задачі управління мережі БПЛА діляться на задачі планування, розгортання (організаційний спосіб реалізації) і оперативного управління (організаційно-технологічний спосіб).

*Етап планування* здійснюється центром управління ЕРМ на основі БПЛА. Змістом планування є (виходячи із прогнозованої обстановки та наявних ресурсів):

1. Планування топології мережі БПЛА (знаходження необхідної кількості БПЛА, визначення їхнього місця розташування або переміщення в просторі), що реалізує певну мету управління, виходячи з вимог до параметрів мережі та вимог щодо передачі трафіка.

2. Розподіл ресурсів (апаратурних, частотних, енергетичних, просторових) мережі БПЛА, вибір конкретних параметрів і режимів роботи технічних засобів (методів та алгоритмів управління).

*Етап розгортання* полягає в запуску заданої кількості БПЛА та управлінні їхнім польотом у задані райони баражування. При цьому задачі етапу розгортання (перепланування топології) мережі БПЛА можуть виконуватися й на етапі оперативного управління при значних змінах мережі (її ушкодження, введенні нових угруповань вузлів та ін.). Контроль за польотом БПЛА та роботою його бортових систем здійснюється із центра управління мережею.

*На етапі оперативного управління* за прийнятими критеріями ефективності постійно оцінюється стан мережі, і приймаються міри (відповідно до плану та реальної обстановки) по втриманню показників ефективності функціонування в заданих межах або їхній оптимізації. Задачі оперативного управління (на відмінність задач планування) вирішуються змішаним способом (централізовано/децентралізовано) у режимі реального часу, а за змістом багаторазово їх повторюють.

*По функціях* задачі управління мережі БПЛА діляться на дві групи:

Спеціальні задачі управління – визначення маршрутів польоту БПЛА та координація їхнього переміщення .

2. Універсальні задачі управління – характерні для будь-якої ЕРМ [9]: управління топологією, управління маршрутизацією, управління

навантаженням, управління радіоресурсом, управління енергоресурсом, управління безпекою. Однак, конкретна реалізація методів управління кожної функціональної підсистеми повинна враховувати призначення мережі БПЛА та особливості її архітектури (розмірність, мобільність, продуктивність й ін.).

Функціональна модель системи управління мережі БПЛА представлена на рис. 21.3.

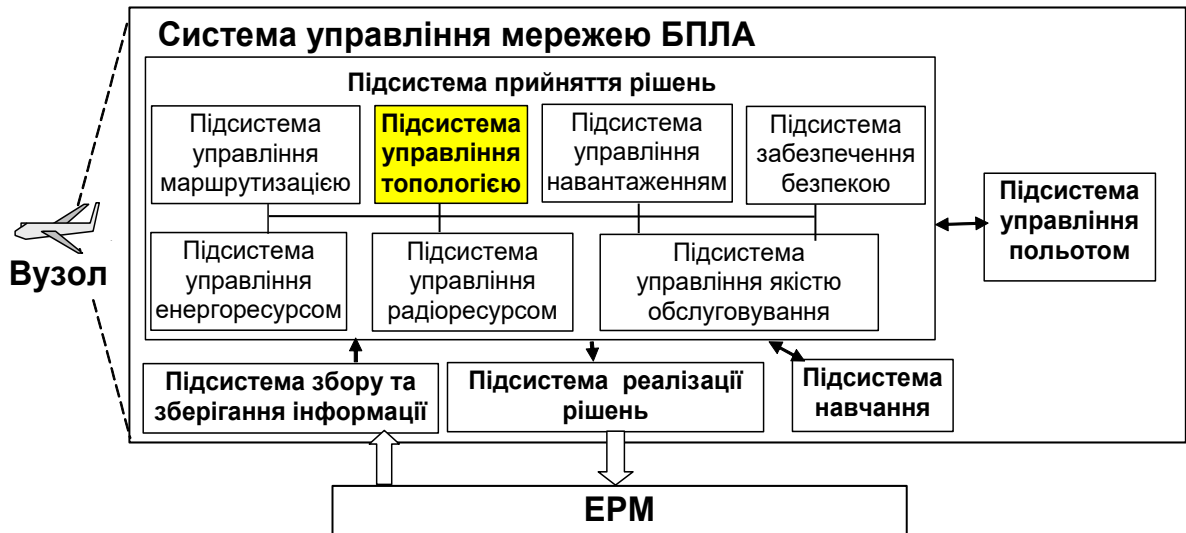


Рис. 21.3. Функціональна модель системи оперативного управління мережі БПЛА

Цілями системи управління мережі БПЛА можуть бути екстремум або підтримка (виступають як обмеження) заданих критеріїв ефективності функціонування мережі загалом (її зони або між окремою парою „відправник-адресат”)  $K_i = \{K_1, K_2, \dots, K_n\}$ :

–  $K_1$  – якість маршрутів всієї ЕРМ або її зони ( $K_1^1$  – довжина маршрутів передачі в зонах мережі,  $K_1^2$  – середній час доставки,  $K_1^3$  – пропускна здатність тощо);

–  $K_2$  – пропускна здатність всієї ЕРМ або її зони;

–  $K_3$  – ступінь покриття зонами БПЛА фіксованої кількості або виділених, пріоритетних наземних абонентів;

–  $K_4$  – структурна надійність (зв'язність);

–  $K_5$  – кількість БПЛА;

–  $K_6$  – безпека і т.д.

Наявність сукупності критеріїв ефективності обумовлює багатокритеріальний характер задач управління та значно ускладнює розробку формальних методів. Для рішення здійснення управління пропонується визначати головний критерій ефективності (виходячи з поточної ситуації), який підлягає оптимізації, а інші переводити в розряд обмежень. Для цього пропонується використати метод ієрархічного цільового динамічного оцінювання альтернатив.

Однією з головних задач управління мережею БПЛА є задача управління топологією (місцеположенням) мережі БПЛА. Класифікація задач управління топологією мережі БПЛА наведена на рис. 21.4.

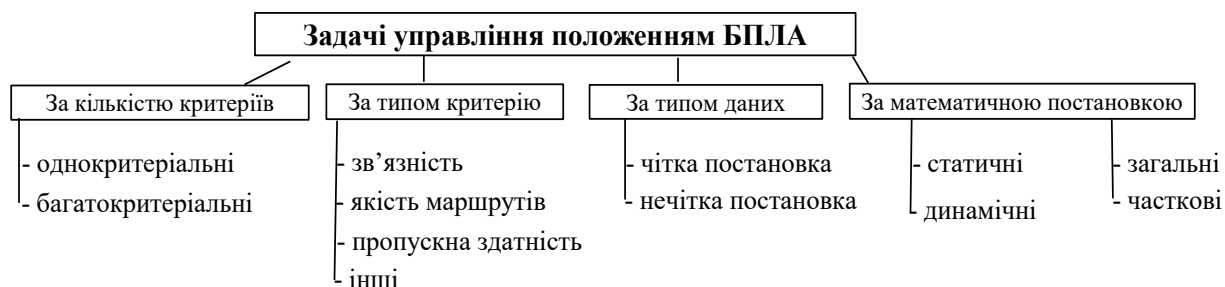


Рис. 21.4. Класифікація задач управління топологією (місцеположенням) мережі БПЛА

Першочерговою метою управління топологією (місцеположенням) мережі БПЛА є підвищення пропускної здатності мережі із одночасним забезпеченням структурної зв'язності та якості обслуговування мобільних абонентів. Тому в якості головного критерію ефективності управління місцеположенням БПЛА пропонується вибрати пропускну здатність мережі, а інші перевести в розряд обмежень.

Тоді управління місцеположенням телекомунікаційних аероплатформ (БПЛА) буде виконуватися за наступними принципами (етапами): планування (перепланування), розгортання, оперативне управління.

На етапі планування (перепланування) центром управління (ЦУ) здійснюється:

- збір інформації про початкову топологію  $((x_i, y_i)_t, v_i, X_{0k}, V_{0k})$  та функціонування мережі  $(P_i)$  та ввід вихідних даних  $(s^o, t_3^o, l^o), d^o, D^o, N, K, \Gamma$ ;

- розрахунок параметрів та тривалості структурної зв'язності, а також параметрів якості маршрутів (QoS), використовуючи удосконалені та існуючі математичні моделі, що розглядаються в наступних главах;

- аналіз виконання умов наявності структурної зв'язності та якості маршрутів;

- розрахунок пропускної здатності мережі  $S$ , згідно математичної моделі, що розглядається в главі 24.

Пошуку початкового (наступного) розміщення чергового БПЛА, що максимізує ПЗ мережі, при наявності можливих рішень, використовуючи запропонований алгоритм, що розглядаються в главі 24.

На *етапі розгортання* здійснюється:

- вивід (переміщення) БПЛА в точку початкового (наступного) розміщення.

- корегування маршрутів передачі даних та навантаження каналів.

На *етапі оперативного управління* здійснюється:

- адаптація роботи телекомунікаційних аероплатформ до реальних умов функціонування (рівня трафіку, створюваного кожним абонентом).

- періодична перевірка необхідності зміни БПЛА своєї позиції (при цьому усі вузли мережі вважаються фіксованими у заданий момент часу).

Дані принципи (етапи) реалізуються послідовно і оперативно для кожного БПЛА згідно методу, що розглядається в главі 24, в залежності від того, на скільки суттєво змінилася топологія мережі по відношенню до її попереднього стану (тобто практично від швидкості зміни топології мережі).

### **21.3. Аналіз методів підвищення пропускної здатності епізодичних радіомереж із застосуванням БПЛА**

Через низьку енергетику супутникових радіоліній та обмежену функціональність супутників зв'язку на їх зміну приходять нові телекомунікаційні системи на основі платформ повітряного базування, роль яких можуть виконувати безпілотні літальні апарати (БПЛА) [11]. Традиційно, акцент робився на відносно великі платформи, такі як *Global Hawk* та *Predator*, але останнім часом зростає увага до малих «мініатюрних БПЛА» (МБПЛА), що є більш дешевими, простими в експлуатації, легкими для транспортування, та не потребують злітно-посадочної смуги або

спеціальної пускової установки. Такі МБПЛА, оснащені засобами зв'язку, можуть більш детально досліджувати важкодоступні райони (стихійних лих, техногенних катастроф тощо) та підвищувати ефективність функціонування наземних епізодичних радіомереж, зокрема пропускну здатність мережі.

Недостатньо розв'язаною на сьогодні є задача оптимального оперативного розміщення множини телекомунікаційних аероплатформ (БПЛА) для підвищення пропускну здатності мережі із одночасним забезпеченням структурно-функціональної зв'язності мобільних абонентів.

Зв'язність мережі широко вивчається останнім часом, в тому числі і в контексті ЕРМ. Введення розосередженої мережі базових станцій (МБС) може значно підвищити зв'язність мережі. Деякі автори визначають критичну потужність наступним чином: що саме вузол має передавати в мережі, щоб забезпечувати зв'язність мережі із ймовірністю 1, коли кількість вузлів в мережі прямує до нескінченності. Існують розрахунки ймовірності того, що два вузла будуть з'єднані двох-стрибковим шляхом краще ніж напямую, а також, яким має бути діапазон передачі, щоб забезпечити зв'язність мережі з більшою ймовірністю. Доведено, що в мережі з  $N$  випадково розміщеними вузлами, якщо кожен вузол з'єднаний з менше ніж  $0,074 \log N$  сусідами, тоді мережа є асимптотично роз'єднаною з ймовірністю 1 при зростанні  $N$ ; якщо вузол зв'язаний з більш ніж  $5,1774 \log N$  сусідами, тоді мережа є асимптотично з'єднаною з ймовірністю 1 при зростанні  $N$ . Існують розрахунки того, як розмістити додаткові вузли так, щоб нарощена мережа була зв'язною і для того, щоб утворити  $k$ - зв'язний граф. Існують аналітичні вирази, що дають змогу визначення потрібної сукупності вузлів, що утворює, при заданій густині, майже без сумніву  $k$ - зв'язну мережу. Також запропоновані методи управління зв'язністю ЕРМ на основі управління потужністю вузлів та маршрутизацією повідомлень, а також управління діаграмою направленості антен вузлів.

Проблема підвищення пропускну здатності за допомогою БПЛА ще довго буде актуальною. Можна підвищити пропускну здатність ЕРМ на основі покриття БПЛА максимальної кількості наземних вузлів. Можна підвищити пропускну здатність мережі за допомогою одного БПЛА завдяки спеціальному маршруту його польоту.. Можна підвищити пропускну здатність мережі за допомогою управління навантаженням в вузлах мережі.

Вирішення проблеми оперативного управління інтелектуальними радіомережами започатковано в роботах проф. Романюка В.А. Проблеми передачі інформації через бортові системи космічних апаратів досліджувалися доц. Пашковим Д.П. В [8] надано огляд робіт, у яких викладено дослідження: залежності пропускну́ї здатності епізодичних радіомереж від зв'язності вузлів вивчається в роботі; кооперативного підходу управління групою БПЛА для утворення безпроводових мереж запропоновано; функціонування епізодичних мережна основі БПЛА із спрямованими антенами; проблема розміщення БПЛА, що функціонують у якості ретрансляторів, для вирішення задач повітряного спостереження вивчається в роботі; динамічного розміщення БПЛА- ретрансляторів для забезпечення зв'язку із мобільними абонентами; оптимізації пропускну́ї здатності багатоланкових епізодичних радіомереж; кооперативного управління мережею мініатюрних роторних БПЛА.

Як показує огляд літератури, запропоновані на сьогодні методи, що закладаються в систему управління топологією (місцеположенням) телекомунікаційних аероплатформ, вирішують лише часткові задачі забезпечення геометричної зв'язності роз'єднаних компонентів мережі, не враховуючи обмеження ємності каналних ресурсів, розподілу навантаження та обслуговування пакетів в вузлах мережі. Також більшість методів вирішують лише статичні задачі, не враховуючи характер мобільності абонентів та маневреність БПЛА, а отже підлягають вдосконаленню. Існуючі методи планування наземних радіомереж також не ефективні, оскільки мають велику складність та час обчислення, що не дозволяє БПЛА відпрацьовувати отримані рішення в режимі реального часу.

Отже, виникає актуальна наукова задача – розробка методу підвищення пропускну́ї здатності ЕРМ з управлінням положенням телекомунікаційних аероплатформ в умовах швидкого та непередбачуваного переміщення мобільних абонентів.

#### **21.4. Загальна постановка задачі підвищення пропускну́ї здатності епізодичних радіомереж із застосуванням БПЛА**

Для постановки загальної математичної задачі дослідження розглянемо функціональну модель ЕРМ із використанням вказано на рис. 21.5.

1-й рівень ЕРМ складає мережа мобільних абонентів (МА), що можуть довільним чином переміщуватися в деякому районі  $r$ . При наявності прямої видимості МА зв'язуються між собою через спільний широкомовний канал на частоті  $f_1$ , а при відсутності – застосовують принцип комутації (маршрутизації) пакетів через проміжні вузли. Таким чином абонентські термінали представляють собою багатофункціональні пристрої, що поєднують у собі прийомопередавач, модем, кодек, маршрутизатор та запам'ятовуючий пристрій, працюючи в одно частотному напівдуплексному режимі за принципом «store-and-forward» («прийняв – запам'ятав – передав»). Для ретрансляції повідомлень через ТА кожен МА має також другий комплект радіообладнання та пам'яті, працюючи у дуплексному двох частотному режимі на частотах  $f_2-f_3$ .

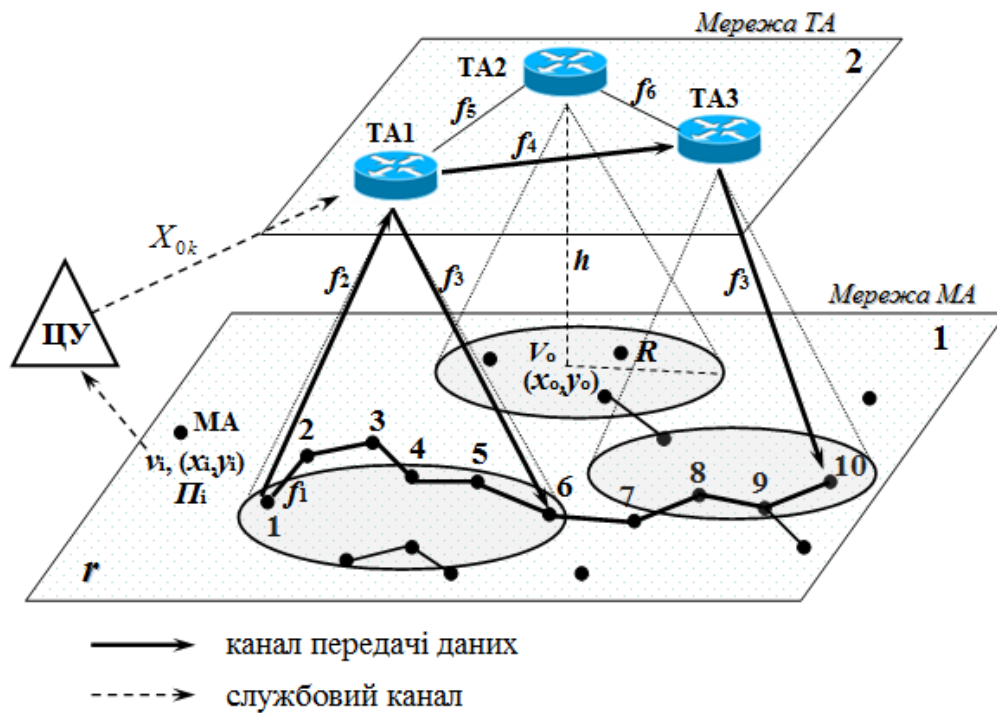


Рис. 21.5. Функціональна модель ЕРМ із використанням ТА

2-й рівень ЕРМ складає мережа телекомунікаційних аероплатформ на основі БПЛА, що баражують на висоті  $h$  по колу мінімального радіусу навколо проекції точки свого оптимального розміщення  $(x_{0k}, y_{0k}), k = \overline{1, K}$ , де  $K$  – кількість БПЛА в мережі, утворюючи стільники радіусом  $R$ . МА, що знаходяться в зоні радіо покриття БПЛА можуть ретранслювати пакети як в всередині стільника, так і в інші стільники, скорочуючи таким чином кількість ретрансляцій в довгих маршрутах. Тобто бортова апаратура БПЛА

також представляє собою складний багатофункціональний пристрій з окремими радіо інтерфейсами (для зв'язку з МА та між собою), що дозволяє виконувати маршрутизацію пакетів по їх адресній інформації, ретранслюючи їх в середині стільника або за його межі. Між стільникові з'єднання (ТА-ТА) працюють у дуплексному режимі із частотним ущільненням, використовуючи набір несучих частот, розподілених за стільниковим принципом, з окремим демодулятором на кожній. За допомогою єдиного передавача пакети направляються сусіднім ТА згідно наявним запитам в режимі розділення в часі. Вважається, що на борту кожного БПЛА є інформація про його місцеположення та розподіл частот по стільникам, що дозволяє при зміні його положення визначити яку з них слід використовувати в даний момент.

Таким чином можливі наступні варіанти маршрутів передачі даних між МА (наприклад, між МА1 та МА10), що позначені суцільними потовщеними стрілками на рис. 21.5:

- через мережу МА (МА1–МА2–...–МА10);
- через мережу БПЛА (МА1–ТА1–ТА3–МА10);
- змішаним шляхом (МА1–ТА1–МА6–...–МА10).

Таким маршрутам висуваються наступні *вимоги* до:

1) пропускної здатності (ПЗ) маршруту  $m_{ab}$  між заданою парою відправник  $a$  – адресат  $b$ :  $s(m_{ab}) \geq s^0$ ,  $a, b = \overline{1, N}$ ,  $m = \overline{1, M}$ , де  $N$  – кількість МА в мережі,  $M$  – кількість маршрутів в мережі,  $s^0$  – мінімальний допустимий рівень ПЗ маршруту;

2) затримки передачі (або кількості ретрансляцій) в маршруті між заданою парою відправник  $a$  – адресат  $b$ :  $t_z(m_{ab}) \leq t_z^0$  ( $l(m_{ab}) \leq l^0$ );

3) існування структурної зв'язності на всіх ділянках маршруту між заданою парою відправник  $a$  – адресат  $b$ :  $d_{ij} \leq d^0$  ( $D_{ik} \leq D^0$ )  $\forall ij \parallel ik \in m_{ab}$ ,  $i, j, a, b = \overline{1, N}$ ,  $k = \overline{1, K}$ , де  $d_{ij}, d^0$  – дальність між сусідніми МА та відповідне обмеження зверху, а  $D_{ik}, D^0$  – дальність між МА та БПЛА та відповідне обмеження зверху;

4) тривалості зв'язності кожної ділянки  $ij$  маршруту між заданою парою відправник  $a$  – адресат  $b$ :  $T_{z\epsilon ij} \geq T_{z\epsilon}^0$ , де  $T_{z\epsilon}^0$  – мінімальний час протягом якого

БПЛА може відпрацювати задане розміщення, встановити маршрут та здійснити передачу мінімальної кількості інформації.

Вибір маршруту здійснюється на основі функціонування одного з відомих методів маршрутизації. Для зручності управління топологією (місцеположенням) ТА краще використовувати таблично-орієнтовані методи (наприклад, OLSR), тоді кожен МА має власну маршрутну таблицю найкоротших шляхів  $P_i$  до усіх інших вузлів мережі.

Для здійснення управління топологією (місцеположенням) ТА в мережі використовується центр управління (ЦУ), винесений за межі району дислокації МА  $r$ . Використовуючи окремий службовий канал (показано на рис.1 пунктирною лінією), ЦУ через мережу БПЛА може легко зібрати вихідні дані про початкову топологію мережі в деякий момент часу  $t$ , а саме – координати та швидкість переміщення кожного МА  $(x_i, y_i)_t, \vec{v}_i, i = \overline{1, N}$  та попередньо виведених БПЛА  $X_{0k} = [x_{0k}, y_{0k}, z_{0k}], \vec{V}_{0k}, k = \overline{1, K}$ , а також дані про функціонування мережі  $P_i$  (існуючі маршрути та їх якість), та здійснити відповідне управлінське рішення (вивід нового БПЛА або переміщення одного з виведених в деяке нове положення простору  $X_{0k}$ ), що максимізує пропускну здатність мережі  $S$ . При цьому вважається, що на момент планування мережі ЦУ відома інформація про розподілення трафіку (навантаження)  $\gamma_{ab}$  між кожною парою відправник-адресат  $a-b$ , що задається матрицею тяжіння  $\Gamma = \|\gamma_{ab}(t)\|$ .

*Ціллю управління* положенням телекомунікаційних аероплатформ (БПЛА) є підвищення пропускну здатності ЕРМ із одночасним забезпеченням структурної зв'язності МА та якості маршрутів передачі даних між ними. При цьому під *пропускну здатністю мережі* будемо розуміти максимальне значення трафіку  $\gamma$ , яке мережа може обробити в одиницю часу при незмінній матриці розподілу трафіку  $\Gamma$ .

Підсумовуючи вищесказане, зведемо всі відомості про функціонування та ресурси в єдину таблицю вихідних даних (табл. 21.2).

## Вихідні дані

БПЛА	<p>Кількість на операцію: <math>K = 1 \dots 5</math></p> <p>Час баражування: <math>T_{\sigma j} \leq 4200, j = \overline{1, K}</math>.</p> <p>Висота баражування: <math>h_j \leq 3000 \text{ м}, j = \overline{1, K}</math>.</p> <p>Швидкість переміщення: <math>V_{0j} = 60 \dots 400 \text{ км/год}, j = \overline{1, K}</math></p> <p>Потужність передавача: <math>P_j = \text{const}, j = \overline{1, K}</math></p>
МА	<p>Кількість абонентів в мережі: <math>N &lt; 200</math></p> <p>Координати абонентів: <math>(x_i, y_i), i = \overline{1, N}</math></p> <p>Швидкість переміщення: <math>v_i = \text{const}, i = \overline{1, N}</math></p> <p>Потужність передавача: <math>p_i = \text{const}, i = \overline{1, N}</math></p> <p>Розмір району дислокації: <math>r = 10000 \times 10000 \text{ м}^2</math></p>
Частотне розподілення	<p>Канал МА-МА: напівдуплексний одно частотний (<math>f_1</math>).</p> <p>Канал МА-БПЛА-МА: дуплексний двох частотний (<math>f_2</math>-<math>f_3</math>).</p> <p>Канал БПЛА-БПЛА: Дуплексний багато частотний.</p> <p>Ширина смуги каналу: <math>\Delta f_1 = \dots = \Delta f_5 = 20 \text{ МГц}</math>.</p> <p>Діапазон частот: 2,4 / 5 ГГц.</p>
МД	<p>Тип МД: випадковий</p> <p>Протоколи МД: МДСЗ (МА-МА), АПР (МА-БПЛА-МА), FDMA (БПЛА-БПЛА).</p>
Маршрутизація	<p>Тип протоколу: таблично-орієнтовний (наприклад, OLSR);</p> <p>Алгоритм побудови маршрутів: Дейкстри.</p>
Трафік	<p>Тип: однорідний (Напр., VoIP+дані).</p> <p>Довжина пакету: <math>L=1024</math> біт.</p> <p>Швидкість передачі в каналах: <math>V=11 \text{ Мбіт/с}</math>.</p> <p>Матриця розподілу трафіку <math>\Gamma = \ \gamma_{ab}(t)\ </math></p>
Управління	<p>Тип: змішаний (централізовано-децентралізований)</p>
Допущення	<p>Буфери мають нескінченну довжину черги.</p> <p>Потік надходження пакетів підкоряється розподілу Пуассона.</p>

Використовуючи вище вказані вихідні дані, *загальну постановку задачі* можна сформулювати наступним чином: визначити розміщення  $X$  групи телекомунікаційних аероплатформ (БПЛА) для максимізації пропускної здатності мережі  $S$ , тобто

$$S = f(X) \rightarrow \max_{X \in \Omega}, \quad (21.1)$$

де  $\Omega$  – ОДЗ, що визначається вимогами до зв'язності та показників функціонування ЕРМ;

$$X = \begin{bmatrix} X_{01} \\ \dots \\ X_{0k} \end{bmatrix}, \text{ де } X_{01} = \begin{bmatrix} x_{01} \\ y_{01} \\ z_{01} \end{bmatrix}, \dots, X_{0k} = \begin{bmatrix} x_{0k} \\ y_{0k} \\ z_{0k} \end{bmatrix}, k = \overline{1, K}.$$

## 21.5. Висновки до глави 21

1. Аналіз функціонування ЕРМ на основі БПЛА показав, що вони відносяться до класу складних, багаторівневих, багатофункціональних, динамічних систем, що при розробці мають ряд проблем, пов'язаних з побудовою та підтримкою структури заданої зв'язності, забезпеченням якості обслуговування мобільних абонентів та підвищенням пропускної здатності мережі. Функціонування таких складних мереж не можливе без ефективної системи управління, складовою частиною якої є підсистема управління топологією (місцеположенням) мережі БПЛА.

2. Аналіз методів підвищення пропускної здатності ЕРМ із застосуванням ТА показав, що запропоновані на сьогодні методи, що закладаються в систему управління топологією (місцеположенням) телекомунікаційних аероплатформ, вирішують лише часткові задачі забезпечення геометричної зв'язності роз'єднаних компонентів мережі, не враховуючи обмеження ємності каналних ресурсів, розподілу навантаження та обслуговування пакетів в вузлах мережі. Також більшість методів вирішують лише статичні задачі, не враховуючи величину і характер мобільності абонентів та маневреність БПЛА, а отже підлягають вдосконаленню. Існуючі методи планування наземних радіомереж також не ефективні, оскільки мають велику складність та час обчислення, що не дозволяє БПЛА відпрацьовувати отримані рішення в режимі реального часу.

3. Таким чином, розробка методу підвищення пропускної здатності ЕРМ з управлінням положенням телекомунікаційних аероплатформ в умовах швидкого та непередбачуваного переміщення мобільних абонентів (рятувальників) є актуальною на сучасному етапі розвитку інтелектуальних радіомереж із самоорганізацією і має не тільки теоретичне, а й прикладне значення для забезпечення ефективного функціонування сучасних телекомунікаційних систем та мереж в надзвичайних ситуаціях.

### **Контрольні запитання та завдання**

1. В чому полягають особливості побудови та функціонування епізодичних радіомереж на основі БПЛА?
2. В чому полягає зміст аналізу задач управління мережею БПЛА в зоні надзвичайної ситуації ?
3. Наведіть основні результати аналізу методів підвищення пропускної здатності епізодичних радіомереж із застосуванням БПЛА.
4. В чому полягає фізичний зміст загальної постановки задачі підвищення пропускної здатності епізодичних радіомереж із застосуванням БПЛА ?
5. В чому полягає математична постановки задачі підвищення пропускної здатності епізодичних радіомереж із застосуванням БПЛА ?

## Глава 22. Математичні моделі оцінки зв'язності вузлів епізодичних радіомереж із БПЛА в зоні надзвичайної ситуації

Першою вимогою для підвищення пропускної здатності ЕРМ є забезпечення зв'язної структури мережі, тому даний розділ присвячено удосконаленню математичної моделі та розробці методики оцінки зв'язності вузлів ЕРМ із телекомунікаційними аероплатформи (БПЛА). Суть удосконалення моделі оцінки зв'язності вузлів ЕРМ із ТА полягає в поєднанні аналітичних моделі для оцінки достовірності передачі даних в радіоканалах з аналітичними моделями переміщення мобільних абонентів (рятувальників).

### 22.1. Вихідні дані та обмеження

*Вихідними даними* для оцінки структурної зв'язності вузлів ЕРМ є наступні параметри:  $(x_i, y_i)_t, i = \overline{1, N}$ ,  $X_{0k}, k = \overline{1, K}$  – координати МА та виведених БПЛА в момент часу  $t$ ;  $v_i = 2\text{м/с}$  або  $15\text{м/с}$  – середня швидкість переміщення МА (пішоходів або авто);  $h_{\max} = 3000\text{м}$  – максимальна висота баражування БПЛА;  $r_{\min} = 50\text{м}$  – мінімальний радіус баражування БПЛА; діапазон частот каналів МА-МА (МА-БПЛА) –  $2,4\text{ГГц}$ , каналів БПЛА-БПЛА –  $5\text{ГГц}$ ; ширина смуги кожного каналу  $\Delta f = 20\text{МГц}$ ; потужність передавача МА –  $p_i = \text{const}$  (однакова для всіх МА); потужність передавача БПЛА –  $p_k = \text{const}$  (однакова для всіх БПЛА); потужність шуму на вході приймача  $P_{u_i} = P_{u_k} = 10^{-10}\text{Вт}$  (однакова для всіх вузлів); коефіцієнт спрямованості бортової антени у напрямку МА –  $3\text{дБ}$ , усі інші антени – неспрямовані довжина пакету  $L=1024\text{біт}$ ; швидкість передачі  $V=11\text{Мбіт/с}$  (однакова для всіх каналів); модель радіоканалу МА-МА – Релея із незалежними завмираннями, усі інші канали – із адитивним білим гаусовим шумом (АБГШ).

*Обмеження* при оцінці безпосередньої зв'язності вузлів, що входять до складу маршрутів ЕРМ  $M = \{m_{ab}\}$ , можна записати наступним чином:  $\Omega_1$ :

$$\{d_{ij} \leq d^o, D_{ik} \leq D^o (R_{ik} \leq R^o), D_{kl} \leq D^o, T_{36 ij} \geq T_{36}^o\} \\ \forall ij, ik, kl \in m_{ab},$$

$$i, j = \overline{1, N}, k, l = \overline{1, K} \},$$

де  $d_{ij}, d^0$  – дальність між  $i$ -м та  $j$ -м МА та відповідне обмеження зверху, а  $D_{ik}, D^0$  – похила дальність між  $i$ -м МА та  $k$ -м БПЛА та відповідне обмеження зверху;  $R_{ik}, R^0$  – відстань між  $i$ -м МА та точкою надиру  $k$ -ого БПЛА та відповідно максимальний радіус зони стабільного покриття  $k$ -ого БПЛА;  $D_{kl}, D'$  – дальність між  $k$ -м та  $l$ -м БПЛА та відповідне обмеження зверху;  $T_{зв\ ij}, T_{зв}^0$  – тривалість зв'язності сусідніх вузлів та відповідне обмеження, що визначає мінімальний час протягом якого вузли можуть встановити зв'язок між собою та здійснити передачу мінімальної кількості інформації. Його можна визначити наступним чином:

$$T_{зв}^0 = T_{зб.\inf.} + T_{пош.\piиш.}^{\min} + T_{відпр.}^{\max} + T_{експл.}, \quad (22.1)$$

де  $T_{зб.\inf.}$  – час збору вихідної інформації  $((x_i, y_i)_t, v_i, P_i)$ ;  $T_{пош.\piиш.}^{\min}$  – час пошуку рішення (положення ТА) з мінімальним відхиленням від екстремального значення пропускної здатності;  $T_{відпр.}^{\max}$  – максимальний час відпрацювання БПЛА заданого рішення (визначається розмірами району дислокації МА та максимальною швидкістю переміщення БПЛА);  $T_{експл.}$  – мінімальний час експлуатації маршруту передачі даних у період його відносної стабільності.

## 22.2. Методика оцінки зв'язності вузлів епізодичних радіомереж із БПЛА

Таким чином *необхідно*: оцінити наявність та тривалість структурної зв'язності вузлів ЕРМ, а також можливість застосування обраного методу (системи) управління положенням телекомунікаційних аероплатформ (БПЛА).

*Крок 1.* Прогнозування середньої тривалості зв'язності  $T_{зв\ ij}$  вузлів ЕРМ на основі моделей переміщення МА, що розглядаються нижче.

*Крок 2.* Перевірка можливості відпрацювання БПЛА заданої цілі управління при заданій динаміці топології мережі (тривалості зв'язності), що визначається умовою  $T_{зв\ ij} > T_{зв\ ij}^0$ . Якщо умова виконується, тоді підвищення

пропускної здатності за допомогою зміни положення БПЛА можливе, крім того є можливість збільшити множину можливих рішень ( $T_{\text{пош.реш.}}$ ) для досягнення більшої точності, інакше – застосування методу не можливе.

*Крок 3.* Розрахунок дальності зв'язку від кожного МА до найближчого сусіда та центру зони покриття БПЛА.

*Крок 4.* Перевірка умов  $\Omega_1$  для кожного МА та перевірка цілісності мережі. Під цілісністю мережі розуміється наявність лише однієї компоненти зв'язності графу мережі. Перевірка цілісності мережі можливо шляхом побудови мінімального кістякового дерева (МКД) графу (наприклад, згідно алгоритму Пріма) та перевірка кожного ребра дерева на виконання умови  $\Omega_1$ . Якщо умови  $\Omega_1$  виконуються, то мережа є структурно зв'язаною на момент часу  $t$ , інакше – необхідне певне управлінське рішення (наприклад, вивід (переміщення) БПЛА).

### **22.3. Математична модель оцінки максимальної дальності радіозв'язку**

Максимальна дальність безпосереднього радіозв'язку пари вузлів мережі обмежується, по-перше, енергетикою радіолінії (потужністю передавача, чутливістю приймача, підсиленням антен, втрат у антенно-фідерному тракті, вільному просторі та ін), при якій відношення сигнал-шум в точці прийому не менше визначеної величини для забезпечення заданої ймовірності пакетної помилки (PER) при заданому виді сигнально-кодової конструкції.

По-друге, максимальна дальність безпосереднього радіозв'язку визначається ефективністю функціонування протоколу множинного доступу (МД) до спільного каналного ресурсу. При збільшенні протяжності радіолінії зростає кількість колізій при організації МД, а отже зменшується пропускна здатність та зростає затримка передачі. Тому максимальна дальність зв'язку може визначатися граничним значення пропускної здатності каналу  $s^0$  при заданій швидкості передачі  $V$ , розмірі пакету даних  $L$  та інтенсивності трафіку  $G$ .

Спочатку отримаємо аналітичну модель ймовірності пакетної помилки (PER) для каналів ERM.

Оскільки при передачі кодованого пакету довжини  $L$  використовується BPSK модуляція, тоді модель сигналу, що приймається, можна записати у наступному вигляді :

$$y_i = \rho_i x_i + n_i, i = 1, \dots, L, \quad (22.2)$$

де  $x_i, i = 1, \dots, L$  - передана BPSK послідовність ( $x_i = 1, -1$ );

$\rho_i$  - незалежні випадкові амплітуди каналу, що мають релеєвське розподілення;

$n_i$  - адитивний білий гаусів шум із нульовим середнім та дисперсією  $\sigma_n^2$ .

Будемо вважати, що кодований пакет був переданий з помилкою, якщо хоча б один біт цього пакету було змінено у процесі декодування. Виходячи з такого визначення, можна знайти вираз для ймовірності пакетної помилки:

$$PER = 1 - (1 - P_e)^L, \quad (22.3)$$

де  $L$  - довжина інформаційної послідовності;

$P_e$  - ймовірність помилки в одиницю часу.

Для теоретичної оцінки ймовірності  $P_e$  будемо використовувати наступний вираз [3]:

$$P_e \leq \sum_{d=d_f}^{\infty} a_d P(d), \quad (22.4)$$

де  $a_d$  - дистанційний спектр досліджуваного коду;

$d_f$  - вільна відстань коду;

$P(d)$  - ймовірність вибору помилкової послідовності з вагою  $d$ .

На практиці при обчисленнях за формулою (22.4) використовується кінцеве кількість доданків.

Дистанційні спектри різних кодів розраховані і для них складені таблиці [4]. Таким чином, для оцінки ймовірності помилки в одиницю часу достатньо знайти ймовірність вибору помилкової послідовності з вагою  $d$ . Вираз для ймовірності  $P(d)$  має наступний вигляд:

$$P(d) = \int_0^\infty \dots \int_0^\infty Q\left(\sqrt{\frac{R_c \sum_{i=1}^N \rho_i^2}{\sigma_n^2}}\right) \frac{\rho_1 \cdot \rho_2 \cdot \dots \cdot \rho_N}{\sigma^2} e^{-\frac{\sum_{i=1}^N \rho_i^2}{2\sigma^2}} d\rho_1 \dots d\rho_N, \quad (22.5)$$

де  $\sigma^2 = \langle |\rho_i|^2 \rangle / 2$ , а  $R_c$  – швидкість коду.

Для обчислення даного інтеграла необхідно використовувати апроксимацію Q-функції. У літературі, присвяченій аналізу завадостійкості згорткових кодів, застосовується апроксимація, побудована на основі границі Чернова, яка дає для релеєвського каналу з незалежними завмиранням оцінку, що відрізняється від експериментального результату на 0,8 дБ. Аналітичний вираз для ймовірності вибору помилкової послідовності з вагою  $d$  має вигляд:

$$P(d) = \frac{1}{2} \cdot \frac{\Gamma(d-1/2)}{\Gamma(d)} \cdot \frac{1}{\sqrt{2 \cdot R_c \cdot SNR}} \cdot \left( \frac{1}{1 + 2 \cdot R_c \cdot SNR} \right)^{d-1/2}, \quad (22.6)$$

де  $\Gamma(d)$  – гама-функція, а  $SNR$  – відношення сигнал/шум.

Для цифрового зв'язку  $SNR$  на вході приймача матиме наступний вигляд [76]:

$$SNR = \frac{P_c}{P_{ш}} \left( \frac{\Delta f}{V} \right), \quad (22.7)$$

де  $P_c$  – потужність сигналу на прийомі, Вт;

$P_{ш}$  – потужність шуму на прийомі (однакова для кожного вузла ЕРМ), Вт;

$\Delta f$  – ширина частотної смуги приймача (однакова для кожного вузла ЕРМ),

Гц;

$V$  – швидкість передачі в каналах ЕРМ, Мбіт/с.

Будемо вважати, що інтерференція від інших вузлів мережі відсутня.

У свою чергу потужність сигналу на прийомі можна визначити наступним чином:

$$P_c = \frac{P_t G_t G_r}{L_0}, \quad (22.8)$$

де  $P_t$  – потужність передавача (для МА вона позначається  $p_i$ , для БПЛА  $P_k$ ),

Вт;

$G_t, G_r$  – коефіцієнти спрямованості передавальної та приймальної антен;

$L_0 = \left(\frac{4\pi D}{\lambda}\right)^2$  – сумарні втрати у вільному просторі, де  $D$  – відстань між передавачем і приймачем (для каналу МА-МА вона позначається  $d_{ij}$ , для МА-БПЛА  $D_{ik}$ , для каналу БПЛА-БПЛА  $D_{kl}$ ), м;  $\lambda$  – довжина хвилі сигналу, м.

Будемо вважати, що згасанням у антенно-фідерному тракті можна знехтувати.

Якщо відомі координати вузлів (наприклад, через GPS), тоді відстань між вузлами  $i$  та  $j$  можна виразити наступним чином:

$$d_{ij} = \sqrt{|x_i - x_j|^2 + |y_i - y_j|^2} \quad (22.9)$$

Підставляючи (22.8) в (22.7), (22.7) в (22.6), (22.6) в (22.4) та (22.4) в (22.3) можемо отримати загальний вираз ймовірності пакетної помилки між наземними вузлами (МА)  $i$  та  $j$ :

$$PER_{ij} = 1 - \left( 1 - \sum_{d=d_f}^{\infty} a_d \cdot \frac{\Gamma(d-1/2)}{\Gamma(d)} \cdot \frac{1}{\sqrt{2R_c \cdot \frac{P_t G_t G_r \Delta f}{P_u V} \cdot \left(\frac{\lambda}{4\pi d_{ij}}\right)^2}} \right) \times \left( \frac{1}{1 + 2R_c \cdot \frac{P_t G_t G_r \Delta f}{P_u V} \cdot \left(\frac{\lambda}{4\pi d_{ij}}\right)^2} \right)^{d-1/2} \Bigg)^L, \quad i \neq j, i, j = \overline{1, N}. \quad (22.10)$$

У випадку каналу МА-БПЛА (БПЛА-БПЛА) (при когерентній модуляції BPSK в каналі з АБГШ та м'якому декодуванні Вітербі) ймовірність вибору помилкової послідовності з вагою  $d$  має наступний вигляд :

$$P(d) = Q(\sqrt{2 \cdot SNR \cdot R_c \cdot d}), \quad (22.11)$$

де  $Q(x)$  – гаусів інтеграл помилок.

Враховуючи, вище вказані співвідношення, аналогічним чином отримуємо вираз ймовірності пакетної помилки для каналів МА-БПЛА та БПЛА-БПЛА:

$$PER_{ik(kl)} = 1 - \left( 1 - \sum_{d=d_f}^{\infty} a_d Q \left( \sqrt{2R_c d \cdot \frac{P_t G_t G_r \Delta f}{P_u V} \cdot \left( \frac{\lambda}{4\pi D_{ik(kl)}} \right)^2} \right) \right)^L, \quad (22.12)$$

Якщо відомі координати проєкцій БПЛА-вузлів на земній поверхні  $x_{0k}, y_{0k}$  (наприклад, через GPS), тоді відстань між БПЛА та наземними вузлами (МА) можна виразити наступним чином:

$$D_{ik} = \sqrt{h^2 + (|x_{0k} - x_i|^2 + |y_{0k} - y_i|^2)}, \quad i = \overline{1, N}, \quad k = \overline{1, K}, \quad (22.13)$$

де  $h$  – висота розташування БПЛА над землею, м.

Використовуючи модуляцію BPSK в поєднанні зі загортковим кодом типу 133<sub>8</sub>,171<sub>8</sub> зі швидкістю  $\frac{1}{2}$ , для якого  $d_f = 10$ ,  $a_d = 36$  (усі інші вихідні дані вказані у підрозділі 22.1), отримуємо наступні залежності PER від максимальної дальності зв'язку для каналів МА-МА (рис. 22.1) та каналів МА-БПЛА (БПЛА-БПЛА) (рис. 22.2) при різних значеннях потужності передавача.

Максимальна дальність безпосереднього радіозв'язку також визначається ефективністю функціонування протоколу МД до спільного каналного ресурсу. Так при  $s^0=0,5$  максимальна протяжність радіолінії МА-МА  $d^o$  має становити не більше 600м для протоколу МД із сигналом «зайнято» (МДСЗ) (показано вертикальною лінією на рис. 22.1), МА-ТА  $D^o$  – не більше 3000м для адаптивного протоколу МД із резервуванням (АПР)  $n$  пакетів (показано вертикальною лінією на рис. 22.2), а на протяжність радіолінії ТА-ТА не має обмежень (окрім енергетики), оскільки протокол МД, що застосовується на ній, не залежить від нормованого часу розповсюдження пакету.

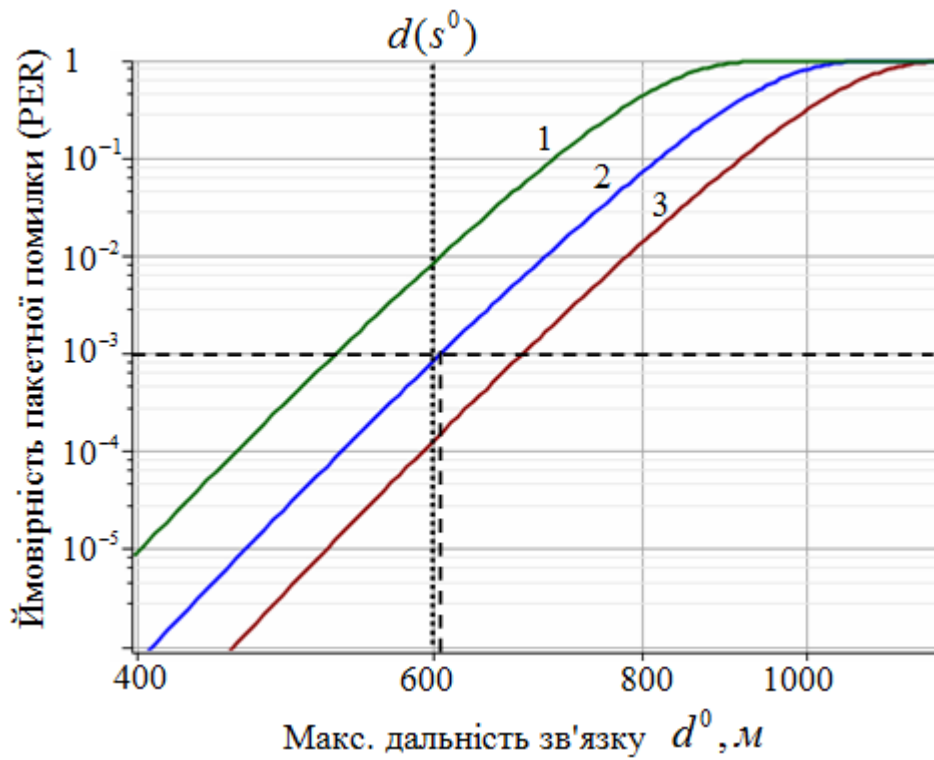


Рис. 22.1. Залежність PER від максимальної дальності зв'язку МА-МА для каналу із завмираннями згідно розподілу Релея при різних значеннях потужності передачі: 1 – 0,6Вт; 2 - 0,8Вт; 3 – 1Вт

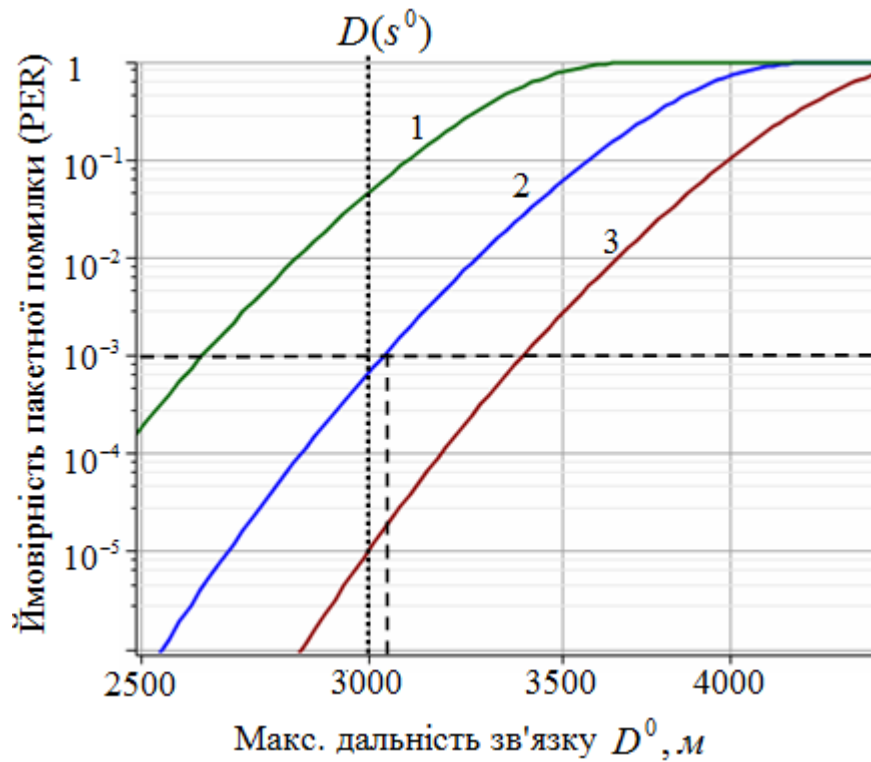


Рис. 22.2. Залежність PER від максимальної дальності зв'язку МА-БПЛА (БПЛА-БПЛА) для каналу АБГШ при різних значеннях потужності передачі: 1 – 3 Вт; 2 – 4 Вт; 3 – 5 Вт

Таким чином, щоб забезпечити задані вимоги до характеристик протоколів МД, потужність передавача слід обирати таким чином, щоб забезпечити необхідний рівень пакетної помилки на відстані, що визначається максимальним часом розповсюдження пакету в мережі. Так для забезпечення рівня PER  $10^{-3}$  на відстані 600 м потужність передачі МА має становити близько 0,8 Вт. У свою чергу для забезпечення того ж рівня PER на відстані 3000 м потужність передавачів БПЛА має становити близько 4 Вт.

Знаючи максимальну протяжність радіолінії МА-БПЛА, можна визначити максимальний радіус зони обслуговування, створюваної телекомунікаційною аероплатформою.

#### 22.4. Формування зони обслуговування БПЛА

Однією з визначальних характеристик функціонування ЕРМ є *зона покриття* БПЛА, що визначається як частина поверхні земного шару, у межах якої створюється необхідний для прийому наземною станцією (МА) рівень сигналу, що випромінюється БПЛА, а також забезпечується необхідний рівень сигналу від наземної станції на вході приймача БПЛА. Регламентом радіозв'язку визначено поняття *зони обслуговування* радіосистеми як частини поверхні Землі, на якій можуть розташовуватися наземні станції даної системи. На відміну від зони покриття у межах зони обслуговування не тільки забезпечується необхідний рівень сигналу на лініях між БПЛА та МА, але й дотримуються необхідні захисні інтервали по відношенню до сигналів інших систем, що заважають. Таким чином, зона обслуговування ніколи не виходить за межі зони покриття.

Для дослідження впливу мобільності БПЛА на формування зони обслуговування спочатку розрахуємо зону обслуговування при деякому фіксованому розміщенні платформи у просторі (рис. 22.3).

Якщо обмежити зону видимості БПЛА тією частиною земної поверхні, для якої БПЛА видно під кутом не менше деякого заданого кута місця  $\gamma$ , то будуть справедливі наступні співвідношення:

$$\begin{cases} (h + R_3) \sin \varphi = R_3 \cos \varphi; \\ \alpha + \varphi + \gamma = \pi/2, \end{cases} \quad (22.14)$$

де  $h$  – висота зависання БПЛА над поверхнею Землі;  $R_3$  – радіус Землі ( $\sim 6378$  км);  $\varphi$  – кут, що дорівнює половині максимального кутового розміру зони видимості відносно БПЛА;  $\alpha$  – кутова відстань границі зони обслуговування від точки надиру відносно центру Землі.

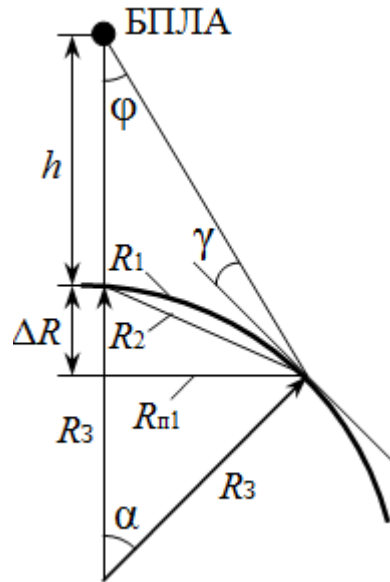


Рис. 22.3. Геометрія повздовжнього перетину зони покриття БПЛА

Тоді максимальна похила дальність від БПЛА до границі зони обслуговування визначається виразом:

$$D^2 = R_3^2 + (h + R_3)^2 - 2R_3(h + R_3)\cos\alpha, \quad (22.15)$$

а кривизна поверхні Землі під точкою зависання БПЛА дорівнює:  $\Delta R = R_3 \tan(\alpha/2)\sin\alpha$ .

Радіус зони покриття БПЛА може бути визначений через параметри кривизни поверхні зони покриття: довжину дуги, що відображає випуклість поверхні зони покриття,  $R_1 = \alpha R_3$ , де  $\alpha$  в радіанах; довжину хорди, що стягує цю дугу,  $R_2 = 2R_3 \sin(\alpha/2)$ ; відстань від точки зависання БПЛА до границі зони видимості  $R_{n1} = R_3 \sin\alpha$ . Відомо, що при  $h \leq 30$  км розходження між значеннями  $R_1$ ,  $R_2$  та  $R_{n1}$  не перевищує 0,1%, отже при розрахунку зони покриття можна користуватись будь-яким з вище вказаних параметрів. Також для збереження заданого інтервалу уразливості  $a$ , необхідного для безконфліктної роботи протоколу МД, максимальна похила дальність  $D$  не має перевищувати 5 км. Враховуючи ці припущення та вирази (22.14), (22.15)

розрахуємо параметри зони покриття в залежності від обраного кута місця  $\gamma$ .  
Результати розрахунків наведено у табл. 22.1.

Таблиця 22.1

**Розрахункові параметри зони обслуговування БПЛА**

Кут місця $\gamma$ , °	Висота розташування БПЛА $h$ , км	Діаметр 2R зони покриття, км
0	0	10
10	0,868	9,848
20	1,710	9,397
30	2,500	8,660
40	3,214	7,660
50	3,830	6,428
60	4,330	5,000
70	4,698	3,420
80	4,924	1,736

Зазвичай для наземної станції радіозв'язку кут місця лежить у межах від  $5^\circ$  до  $90^\circ$ . Для абонентських терміналів мереж на основі NARPs згідно рекомендації F.1569 мінімальний кут місця становить  $20^\circ$ . На практиці у межах щільно забудованого міста кут місця становить  $60-90^\circ$  [6]. Згідно даному припущенні розмір стільника одного мініатюрного БПЛА становитиме порядку 5-6 км у діаметрі.

Однак через мобільність БПЛА зона покриття зазнає певних змін. По-перше, через те, що БПЛА не може зависнути у заданій точці (якщо це не гелікоптер), а лише має змогу обертатися навколо неї по колу мінімального радіусу  $r_{\min}$ , що визначається конструкторськими даними платформи. По-друге, через неточність виводу БПЛА ( $\Delta x$ ,  $\Delta y$ ,  $\Delta z$ ), що визначається погодними умовами (зокрема швидкістю вітру) та похибкою системи позиціонування (наприклад, GPS). Це призведе відповідно до нестабільності зони покриття через відхилення діаграми спрямованості (ДС) антени аероплатформи (рис. 22.4а). У даному випадку зона стабільного прийому буде знаходитись у межах  $(0, R - \Delta'')$ , де  $\Delta'' = r_{\min} + \Delta x$ , а зона нестабільного прийому – у межах  $(R - \Delta''; R + \Delta'')$ . При  $R=3000\text{м}$ ,  $r_{\min}=200\text{м}$ ,  $\Delta x=100\text{м}$  зона

стабільного прийому  $2R_{ст}$  зменшиться на 600м, що становить 10% від зони покриття  $2R$  фіксованої платформи.

Для того, щоб збільшити зону стабільного прийому можна застосувати один з методів корекції ДС антени БПЛА [36], наприклад, пропорційний вихід країв ДС за зону стільника. У даному випадку зона обслуговування матиме форму еліпса через нахил кругової ДС, а зону стабільного прийому описуватиме мала вісь меншого еліпса  $b'$  при переміщені БПЛА по колу мінімального радіуса із максимальною нестабільністю  $\Delta x$  у бік центру зону обслуговування ( $\Delta' = r_{min} - \Delta x$ ) (рис. 22.4б). У свою чергу зону нестабільного прийому описуватиме велика вісь більшого еліпса  $a''$  при переміщені БПЛА по колу мінімального радіуса із максимальною нестабільністю  $\Delta x$  у протилежний бік від центру зону обслуговування ( $\Delta'' = r_{min} + \Delta x$ ).

Якщо представити головну пелюстку ДС антени БПЛА у вигляді конуса випромінювання з кутом  $\varphi$  (вершина конуса на БПЛА), нахиленого до Землі під кутом  $\theta$  (при вершині конуса), а поверхню землі – як січну площину в перетині такого конуса, то при використанні проекційних правил нарисної геометрії та апарату тригонометрії можна отримати діаметри зон стабільного та нестабільного прийому:

$$a'' = r''_1 + r''_2;$$

$$b' = (4 \tan^2(\varphi/2) \cdot (h/\cos\theta' + \sin\theta' \cdot (r'_2 - r'_1)/2)^2 - (r'_2 - r'_1)^2 \cos^2\theta')^{0,5}, \quad (3.16)$$

де  $r'_1 = h \cdot (\tan\theta^{(c)} - \tan(\theta^{(c)} - \varphi/2))$ ;  $r'_2 = h \cdot (\tan(\theta^{(c)} + \varphi/2) - \tan\theta^{(c)})$ ;

$$\theta^{(c)} = \frac{1}{2} \arctan\left(\frac{\Delta^{(c)}}{h} \frac{1 + \cos(\varphi) \sqrt{1 + \sin^2(\varphi) \cdot \Delta^{(c)2}/h^2}}{1 - \cos^2(\varphi) \cdot \Delta^{(c)2}/h^2}\right).$$

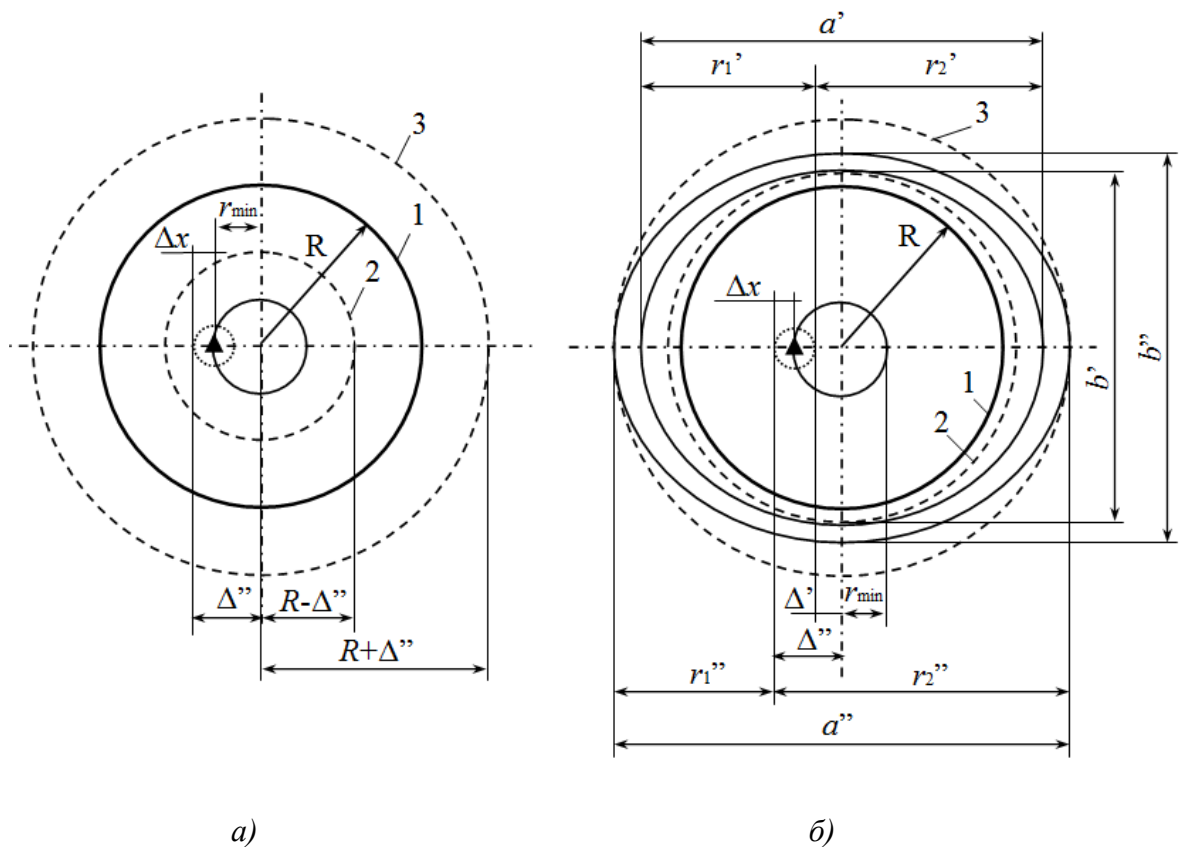


Рис. 22.4. Вид зверху заданої зони обслуговування (1) та зон стабільного (2) і нестабільного (3) покриття, що утворені без методів корекції положення ДС (а) та методом пропорційного виходу країв ДС за зону стільника (б)

Залежності приведені до  $2R$  діаметрів зон стабільного та нестабільного покриття від зміщення БПЛА  $\Delta$  при застосуванні методу пропорційного виходу країв ДС за зону стільника та без нього представлені на рис. 22.5. Аналізуючи дані залежності можемо бачити, що без застосування методів корекції положення ДС розмір зони стабільного покриття зменшується, а зони стабільного покриття відповідно збільшується, прямо пропорційно росту зміщення БПЛА  $\Delta$ . При застосуванні методу пропорційного виходу країв ДС за зону стільника розмір зони як стабільного, так і нестабільного покриття монотонно зростає із ростом зміщення БПЛА  $\Delta$ . При чому у даному випадку зона стабільного прийому завжди переважає задану фіксовану зону обслуговування, а розмір зони нестабільного прийому значно менше відповідних значень у випадку відсутності корекції положення ДС. При зміщенні БПЛА менш ніж на 500м ( $\Delta/2R \leq 0,1$ ) усі три зони будуть майже співпадати.

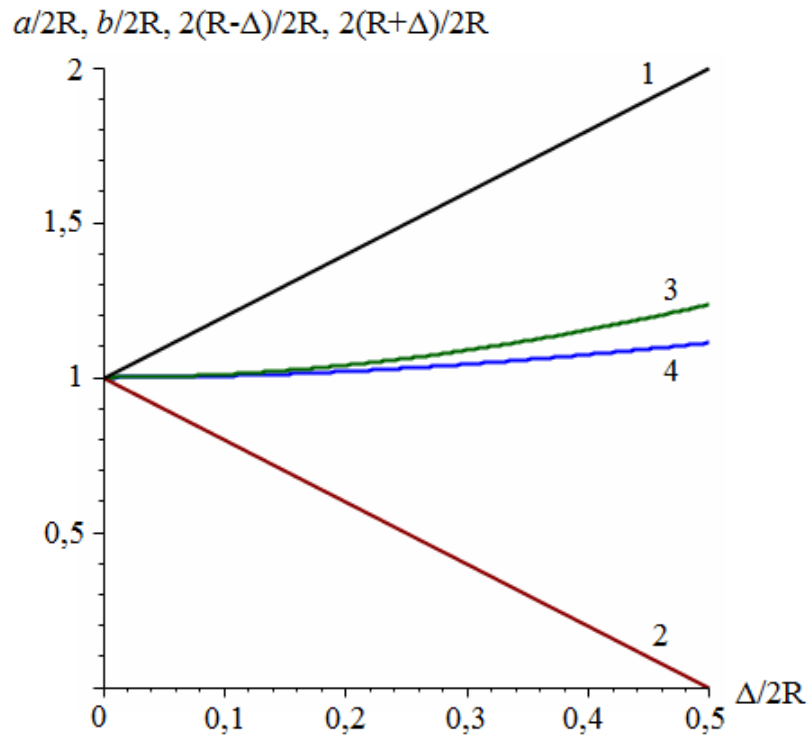


Рис. 22.5. Залежності приведених до  $2R$  діаметрів зон стабільного (2,4) та нестабільного (1,3) покриття від зміщення БПЛА  $\Delta$  при застосуванні методу пропорційного виходу країв ДС за зону стільника (3,4) та без нього (1,2) ( $2R=5\text{км}$ ,  $h=4,33\text{км}$ )

Без застосування методів корекції ДС бортової антени зону стабільного покриття можливо збільшити при застосуванні множини БПЛА, що синфазно рухаються з однаковим радіусом  $r$  та швидкістю  $V$  баражування на однаковій висоті ( $R=\text{const}$ ) із зміщенням позиції один відносно одного на кут однакової величини  $\omega=2\pi/N$ , де  $N=1,2,3,\dots$  – кількість БПЛА. Тоді зону стабільного покриття буде утворювати точка перетину двох кіл радіуса  $R$  сусідніх БПЛА (рис. 22.6). Зазначимо, що при цьому радіус зони нестабільного покриття залишиться незмінним  $R+r$ . Тоді із застосуванням тригонометричних перетворень радіус зони стабільного покриття  $L$  можна визначити наступним чином:

$$L = \begin{cases} R - r, & K < 2 \\ \sqrt{R^2 - \frac{1}{2}r^2(1 - \cos\frac{2\pi}{K})}, & K \geq 2 \end{cases} \quad (22.17)$$

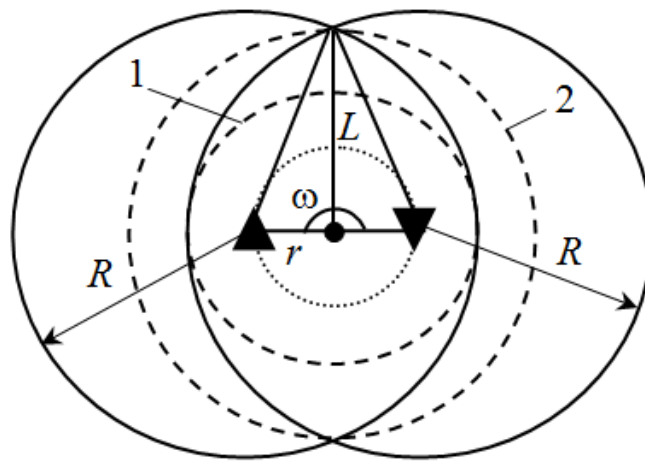


Рис. 22.6. Формування зони стабільного покриття із застосуванням одного БПЛА (1) та із застосуванням двох БПЛА (2)

Для різних значень радіуса баражування  $r$  (при  $R=2500\text{м}$ ) залежність розміру зони стабільного покриття від кількості БПЛА згідно (4) буде мати наступний вигляд (рис. 22.7).

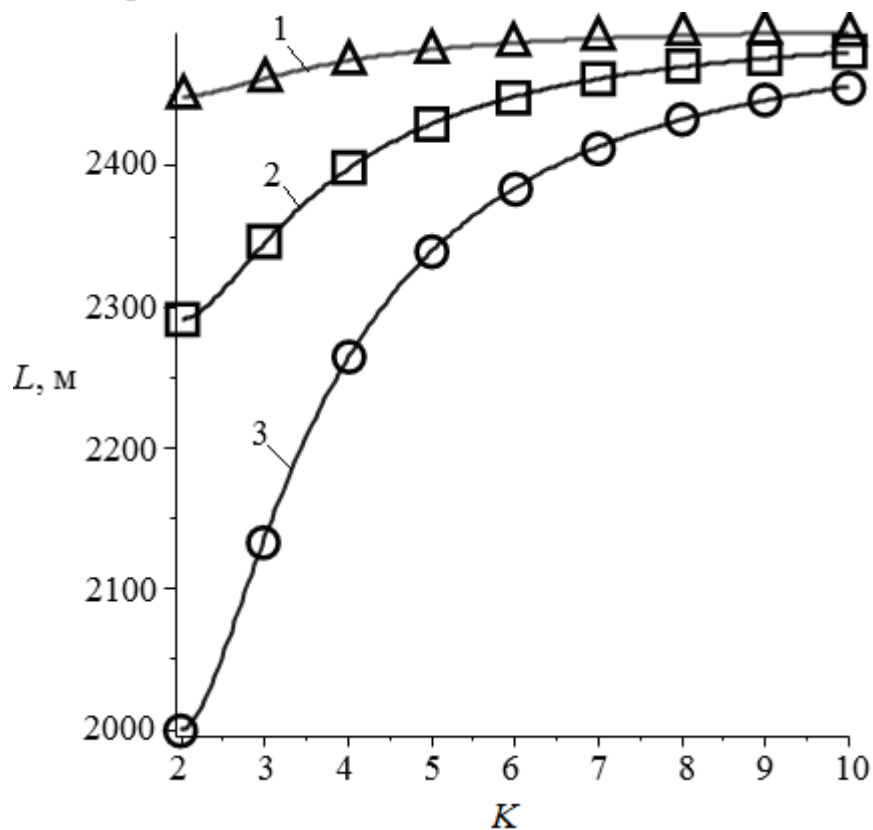


Рис. 22.7. Залежність розміру зони стабільного покриття від кількості БПЛА при різних значеннях радіуса баражування: 1 –  $r_1=500\text{м}$ , 2 –  $r_2=1000\text{м}$ , 3 –  $r_3=1500\text{м}$

Аналізуючи графіки на рис. 22.7 та вираз (22.17) можемо бачити, що найбільший вклад у збільшенні зони стабільного покриття вносить другий БПЛА (65% на випадок  $r=1000\text{м}$ ,  $R=2500\text{м}$ ), при застосуванні більшої кількості ретрансляторів приріст зменшується, а графіки асимптотично прямують до значення  $R$ . Також можемо бачити, що приріст  $L$  зменшується із зменшенням радіусу баражування  $r$ . Також, слід зазначити, що застосування декількох ретрансляторів при формуванні зони обслуговування одного стільника дозволяє збільшити надійність та інформаційну ємність мережі в  $K$ - кратному розмірі. Однак із збільшенням кількості ретрансляторів, що баражують по одному колу, збільшується також небезпека їх зіткнення, тому на практиці доцільніше буде застосовувати не більше 4 БПЛА.

## **22.5. Розрахунок тривалості зв'язності рятувальників в зоні надзвичайної ситуації**

Через свою мобільність абоненти ЕРМ не можуть постійно знаходитись у зоні радіо видимості один одного та зони покриття БПЛА, отже виникає практичний інтерес визначити часові рамки, в яких існує зв'язок між ними, тобто час життя з'єднання *LLT* (*Link life time*), протягом якого вони можуть здійснювати інформаційний обмін (побудувати маршрут та передати деяку кількість інформації). Вивчення цього питання започатковано у роботах [7 – 9], де запропоновано низку моделей мобільності вузлів для аналізу та моделювання ЕРМ. Однак запропоновані моделі охоплюють лише вузький набір варіантів переміщення МА, а отже підлягають вдосконаленню та доповненню.

ЕРМ передбачають довільне (за напрямком та швидкістю) переміщення вузлів. На практиці, якщо йде мова про мобільний зв'язок «пішохід-пішохід», у місті напрямок руху змінюється у відповідності із планом його забудови (частіше всього повороти на  $90^0$ ), а швидкість реально не перевищує 2 м/с, автомобілі з урахуванням обмежень швидкості у місті переміщуються з граничною швидкістю до 16 м/с.

### 22.5.1 Розрахунок на випадок прямої радіо видимості між вузлами

Розглянемо переміщення вузла  $B$  відносно вузла  $A$ , розташованого спочатку від нього на відстані  $d$ , зі швидкістю  $v$  під кутом  $\varphi$  (рис. 22.8а). Будемо рахувати, що *тривалість зв'язності* ( $T_{зв}$ ) – це проміжок часу, поки вузол  $B$  не опиниться в точці  $B'$ , що обмежена радіусом радіо видимості  $R$ . Тоді із використанням апарату тригонометрії отримаємо:

$$T_{зв} = \frac{d \cos \varphi + \sqrt{R^2 - d^2 \sin^2 \varphi}}{v}. \quad (22.18)$$

Для вихідних даних  $d = 100\text{м}$ ,  $v_1 = 2 \text{ м/с}$ ,  $v_2 = 15 \text{ м/с}$  та  $\varphi = 90^\circ$  отримаємо результати, приведені на рис. 3.8б. Згідно цим даним впливає, що тривалість зв'язку МА-МА ( $R = 500 \text{ м}$ ) може становити не більше 2-3 хвилин для пішоходів та не більше 0,5 хвилини для авто, а на випадок БПЛА-МА ( $R = 2500 \text{ м}$ ) становитиме приблизно 20-25 хв. для пішоходів та приблизно 6-7хв. для авто. Отже ефективність застосування БПЛА очевидна. Управляючи положенням БПЛА у просторі можна досягти іще більших значень тривалості зв'язності абонентів.

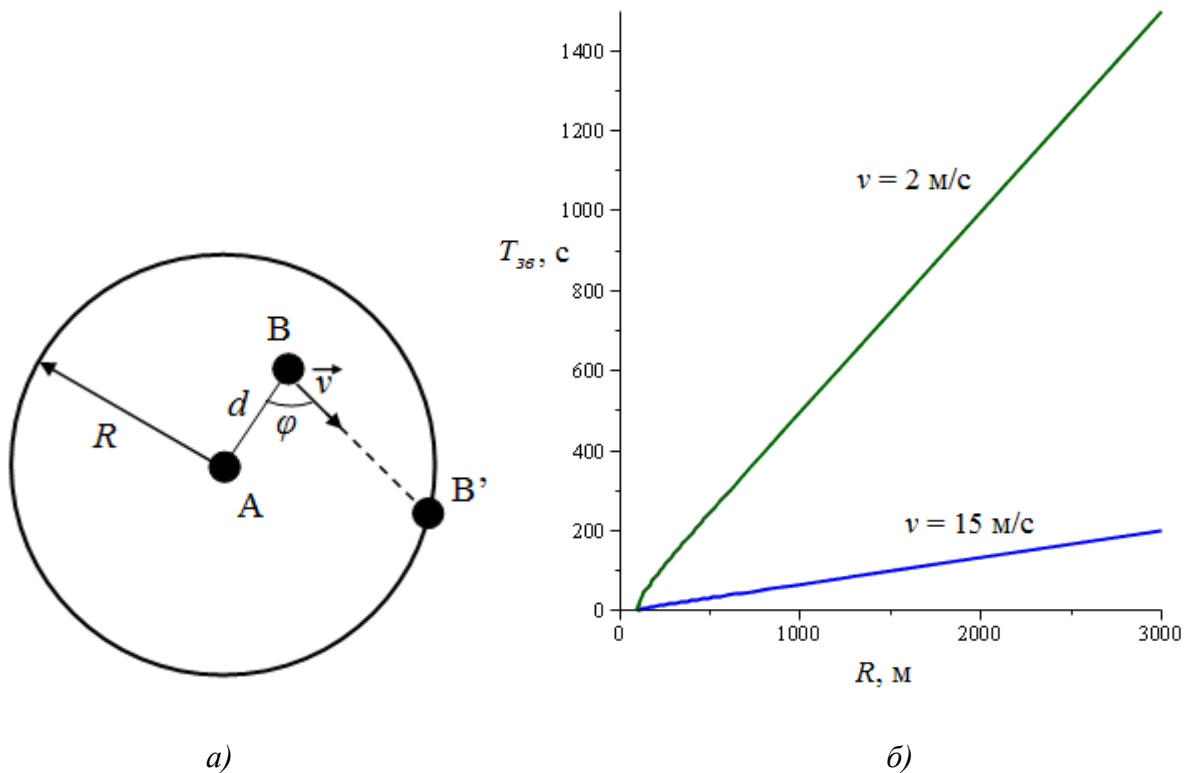


Рис. 22.8. Модель (а) та результати розрахунку тривалості зв'язності (б)

## 22.5.2 Розрахунок на випадок ретрансляції через проміжний вузол

Розглянемо модель, в якій вузол  $A$  може здійснювати зв'язок з вузлом  $C$  тільки за рахунок ретрансляції сигналу через вузол  $B$ . Вузол  $C$  знаходиться поза межами радіо видимості вузла  $A$ , при цьому  $C$  знаходиться в зоні видимості  $B$ . Для спрощення будемо рахувати, що радіус радіо видимості  $R$  для всіх вузлів однаковий.

Для здійснення зв'язку вузол  $A$  будує маршрут до  $C$ , і на це витрачається деякий час  $t$ . Оскільки  $B$  та  $C$  рухаються один відносно одного (та відносно  $A$  зі швидкостями  $v$  та  $w$  відповідно) (рис. 22.9а), наприклад, розходяться, тоді існує ймовірність відносно того, що за час  $t$  встановлення маршруту від  $A$  до  $C$  ретранслятор  $B$  та кінцевий вузол  $C$  розійдуться настільки, що передача повідомлень від  $A$  до  $C$  виявиться неможливою. Розрахуємо цей час.

Будемо вважати, що тривалість зв'язності ( $T_{зв}$ ) – це проміжок часу, поки вузол  $C$  не опиниться за межами видимості ретранслятора, тобто у точці  $C'$ . Тоді з використанням простих формул [10] отримаємо:

$$T_{зв} = \frac{\sqrt{AB^2 + AC^2 - 2AB \cdot AC \cdot \cos(\alpha + \beta + \gamma)}}{v^*}, \quad (22.19)$$

де  $v^*$  - відносна швидкість вузлів  $B$  та  $C$ , величина якої визначається напрямками та величинами векторів  $\vec{v}$  та  $\vec{w}$ .

Задаючи відстані між вузлами  $AB=200\text{м}$ ,  $AC=600\text{м}$ , отримаємо залежність тривалості зв'язності з однією ретрансляцією від відносної швидкості вузлів для різних значень кутів між вузлами (рис. 3.9б).

Аналізуючи рис. 22.9 можемо бачити, що значення тривалості зв'язності з однією ретрансляцією лежать у тих самих межах, що й у попередньому випадку (без ретрансляції), та визначаються розміщенням вузлів та напрямком й величиною швидкості їх переміщення.

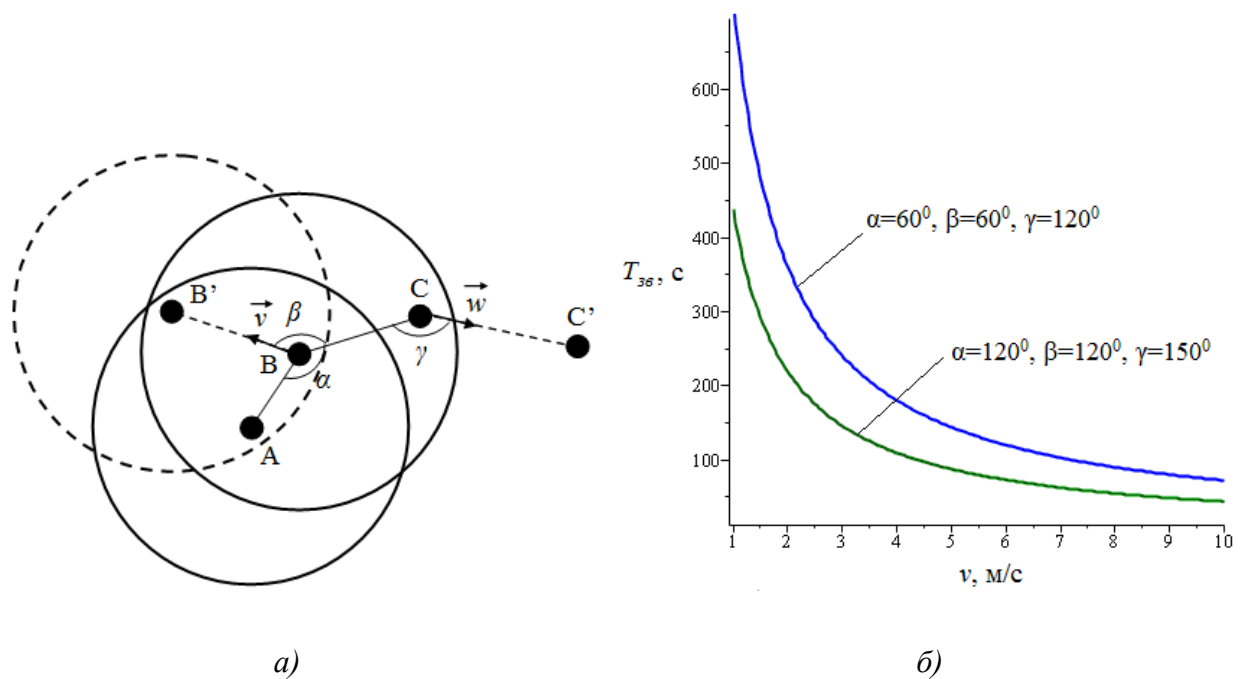


Рис. 22.9. Модель (а) та результати розрахунку тривалості зв'язності з однією ретрансляцією (б)

Тепер розглянемо поведінку фрагменту ЕРМ з 4-х вузлів. У початковому стані вузли ( $B, C, D$ ) знаходяться у межах радіо видимості вузла  $A$  (рис. 22.10а). З ними може здійснюватись обмін інформацією. Але через мобільність (показані вектори руху вузлів відносно  $A$ ) топологія мережі змінюється та через деякий час вузол  $B$  виходить з зони видимості  $A$ . Зв'язок з ним для вузла  $A$  на деякий час переривається (рис. 22.10б). Це продовжується до тих пір, поки  $B$  не потрапляє в зону радіо видимості вузла  $C$ , який, у свою чергу, буде служити ретранслятором для вузла  $A$ . Зауважимо, що протягом даного періоду часу вузли можуть змінювати напрямлення та швидкість переміщення. Тепер зв'язок між  $A$  і  $B$  здійснюється через ретранслятор  $C$  у два стрибки (рис. 22.10в). Можлива ситуація, коли вузол  $C$  не може більше служити ретранслятором (наприклад, розрядилася батарея) (рис. 22.10г). В цьому випадку алгоритм маршрутизації, адаптуючись до нових умов, створює інший маршрут – через БПЛА-ретранслятор (рис. 22.10д). В даному випадку можемо говорити про можливість зв'язку між вузлами  $A$  та  $B$  через БПЛА.

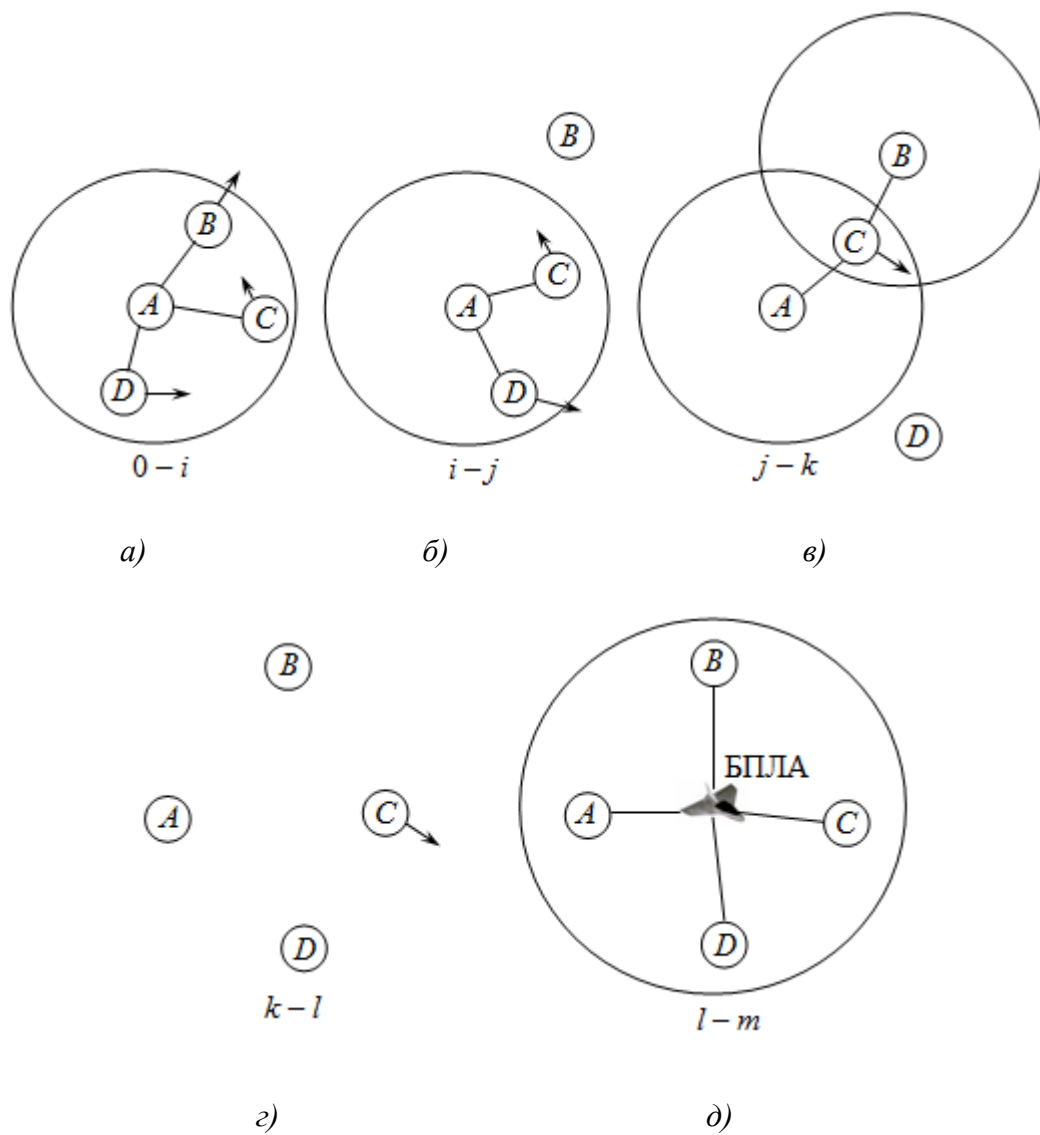


Рис. 22.10. Забезпечення зв'язності при переміщенні вузлів

Якщо відобразити вказану вище послідовність подій на схемі (рис. 22.11), то наочно побачимо порушення зв'язності між вузлами А та В на проміжках  $i-j$  та  $k-l$ .

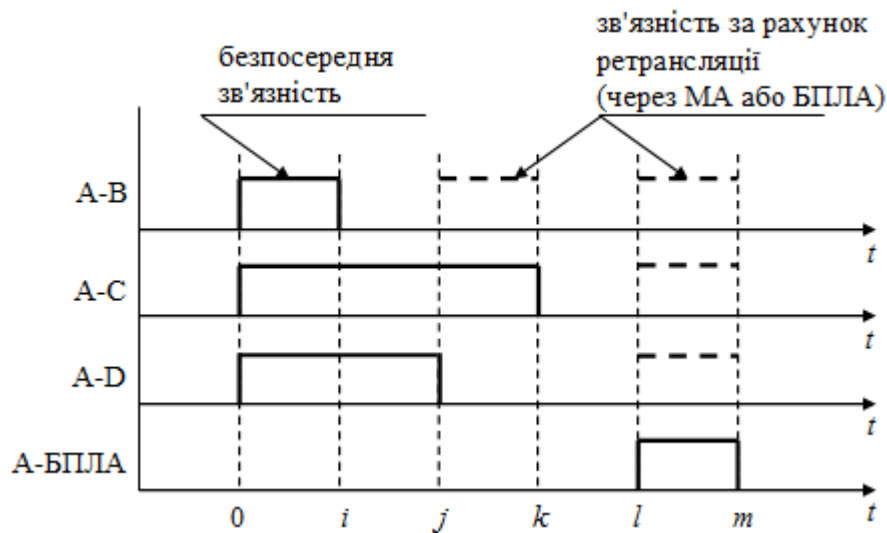


Рис. 22.11. Інтервали зв'язності

При передачі довгих повідомлень це може призвести до втрати інформації, якщо не вживати спеціальних заходів. Такими заходами можуть бути:

1. Дроблення повідомлень на пакети та відправка пакетів тільки після підтвердження наявності зв'язності та прогнозу на її підтримку протягом часу пересилки одного пакету. Для гарантії пересилки у вузла-відправника повинна існувати (або передбачатися) можливість «розширення» зони покриття за рахунок збільшення потужності передавача або зміни діаграми направленості антени [11], тобто вузол  $A$  «не відпускає»  $B$  протягом передачі хоча б одного пакету.

2. Передбачити наявність «резервних маршрутів», які автоматично включаються (м'яка естафетна передача) при перериванні основного.

3. Застосування алгоритмів локального відновлення маршруту на основі розрахунку координат вузлів в локальних координатах.

4. Архітектурні або алгоритмічні зміни, у тому числі, що порушують принцип однорангової мережі, тобто застосування мережі БПЛА у якості додаткових повітряних вузлів-ретрансляторів. Ці заходи будуть обговорюватися у главі 24.

### 22.5.3 Моделювання переміщення вузлів мережі

Використовуючи математичну модель, що розглядається у пункті 22.5.1, змоделюємо переміщення вузлів за різними сценаріями та порівняємо

тривалість зв'язності між вузлами ЕРМ при різних сценаріях мобільності рятувальників. Для моделювання переміщення вузлів пропонуються наступні сценарії мобільності:

1) «однакових напрямків» або «марш» ( $\varphi = const, v = const$ ). Згідно даного сценарію усі вузли в межах зони радіо видимості  $R$  рухаються прямолінійно та паралельно один одному в однаковому напрямку з однаковою швидкістю, як показано на рис. 22.12а.

2) «випадкових напрямків» або «різнобій» ( $\varphi = const, v = const$ ). Згідно даного сценарію усі вузли в межах зони радіо видимості  $R$  рухаються прямолінійно за різними випадковими напрямками з однаковою швидкістю, як показано на рис. 22.12б.

3) «випадкового блукання в полі» ( $\varphi = var, L = const, v = const$ ). Згідно даного сценарію усі вузли в межах зони радіо видимості  $R$  рухаються з однаковою швидкістю по ломаній траєкторії, випадково змінюючи свій напрямок через кожний постійний крок  $l \ll R$ , як показано на рис. 22.12в.

4) «випадкового блукання у місті» ( $\varphi = var(0^0, 90^0, 180^0, 270^0)$ ,  $L = const, v = const$ ). Згідно даного сценарію усі вузли в межах зони радіо видимості  $R$  рухаються з однаковою швидкістю по перпендикулярних прямих (вулицях) та випадковим чином змінюють свій напрямок на кожному перехресті через кожний постійний крок  $l \ll R$ , як показано на рис. 22.12г.

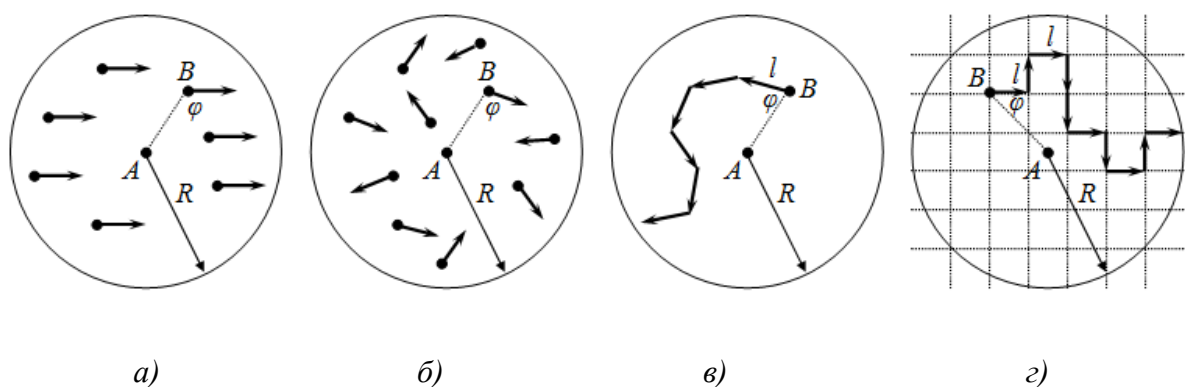


Рис. 22.12. Траєкторії руху вузлів ЕРМ в залежності від сценарію мобільності: а – «марш», б – «різнобій», в – «випадкове блукання в полі», г - «випадкове блукання у місті»

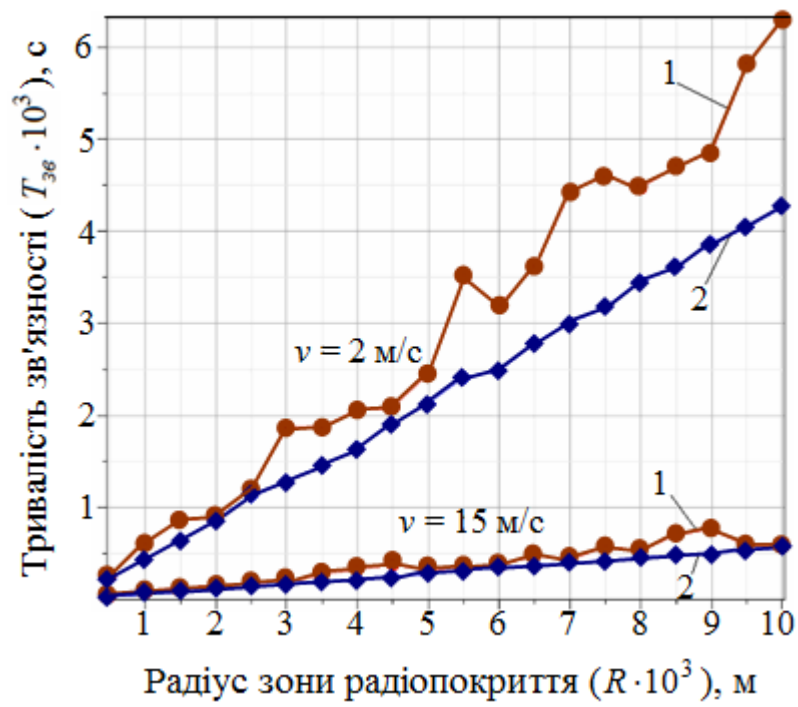
Координати усіх вузлів задаються випадково у межах зони радіо видимості  $R$  заданого фіксованого вузла (наприклад, БПЛА). Кількість вузлів

для кожного радіусу зони покриття пропорційна її площі, тобто  $R^2$ . Розрахуємо середній час, за який вузол досягає межі зони радіо видимості, в залежності від її радіусу при різних швидкостях та сценаріях переміщення. Моделювання виконувалось у програмному середовищі MAPLE. Результати моделювання приведені на рис. 22.13а,б.

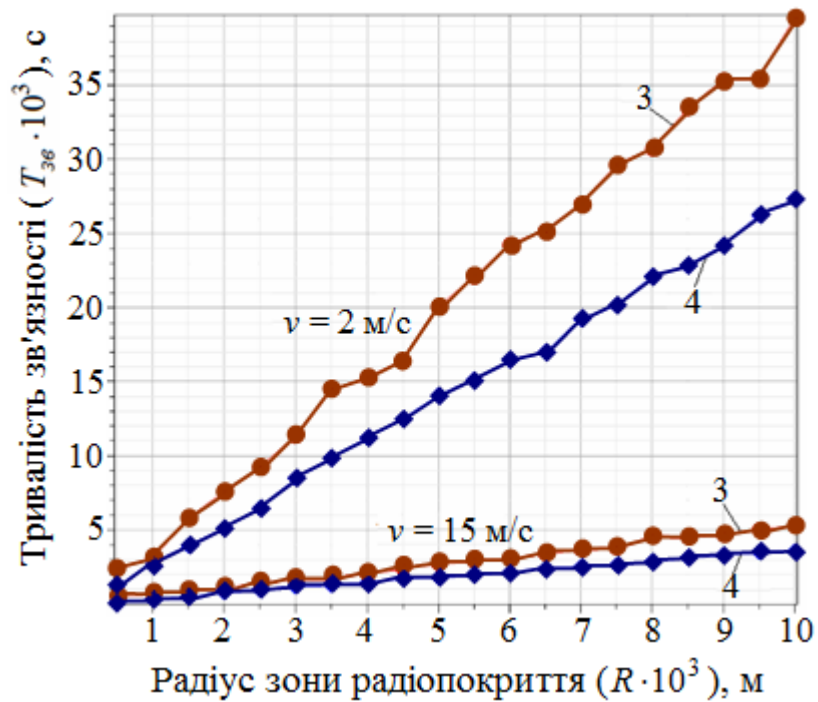
Аналізуючи рис. 22.13а,б, можемо бачити, що тривалість зв'язності ( $T_{зв}$ ) прямо пропорційна розміру зони покриття ( $R$ ) та обернено пропорційна швидкості переміщення вузлів ( $v$ ). На величину тривалості зв'язності також впливає тип (сценарій) мобільності вузлів. Найбільші значення тривалості зв'язності (при однакових  $R$  та  $v$ ) відповідають сценарію «випадкове блукання в полі» (крива 3), другі за величиною – «випадкове блукання у місті» (крива 4), треті за величиною – «марш» (крива 1), найменші – сценарію «різнобій» (крива 2). Так у випадку БПЛА ( $R=5000\text{м}$ ) середня тривалість зв'язку із пішоходами ( $v=2\text{м/с}$ ) буде відповідно – 20000, 14000, 2463, 2130 секунд, із автомобілями ( $v=15\text{м/с}$ ) відповідно – 2686, 1900, 327, 286 секунд. Знаючи ці дані можна визначити частоту відпрацювання нового положення БПЛА, що визначається при розробці методу підвищення пропускної здатності на основі управління положенням БПЛА, що розглядається в главі 24.

## 22.6. Висновки до глави 22

Показано, що максимальна дальність безпосереднього радіозв'язку пари вузлів мережі обмежується, по-перше, енергетикою радіолінії, при якій відношення сигнал-шум в точці прийому не менше визначеної величини для забезпечення заданої ймовірності пакетної помилки (PER) при заданому виді сигнально-кової конструкції. По-друге, максимальна дальність безпосереднього радіозв'язку визначається ефективністю функціонування протоколу МД до спільного каналного ресурсу.



a)



б)

Рис. 22.13. Залежність середнього часу досягнення МА границі зони покриття від її радіусу при різних швидкостях та сценаріях переміщення: 1 – «марш» (а), 2 – «різнобій» (а), 3 – «випадкове блукання в полі» (б), 4 - «випадкове блукання у місті» (б)

Так при мінімальній пропускній здатності каналів  $s^0=0,5$  максимальна протяжність радіолінії МА-МА  $d^o$  має становити не більше 600м для протоколу МДСЗ, МА-БПЛА  $D^o$  – не більше 3000м для адаптивного протоколу АПР, а на протяжність радіолінії БПЛА-БПЛА не має обмежень (окрім енергетики), оскільки протокол МД, що застосовується на ній, не залежить від нормованого часу розповсюдження пакету.

Показано, що знаючи максимальну похилу дальність радіолінії МА-БПЛА та допустимий кут місця абонентського терміналу, можна визначити максимальний радіус стільника, створюваного телекомунікаційною аероплатформою. Так при куті міста  $60^0-90^0$  у межах щільно забудованого міста він становитиме  $R \leq 1500$ м при висоті БПЛА  $2600 \leq h < 3000$ м. Але враховуючи маневрування БПЛА по колу мінімального радіусу та неточність його виводу, радіус зони стабільного покриття становитиме  $R^o \leq 1450$ м.

3. На основі простих моделей досліджено тривалість зв'язності мобільних вузлів ЕРМ (МА та БПЛА) в умовах прямої радіо видимості та з урахуванням ретрансляції. Показано, що тривалість зв'язності прямо пропорційна розміру зони покриття та обернено пропорційна швидкості переміщення вузлів. На величину тривалості зв'язності також впливає характер (сценарій) мобільності вузлів. Було виконано моделювання переміщення вузлів за 4-ма сценаріями: «марш», «різнобій», «випадкову блукання в полі» та «випадкове блукання в місті». Найбільші значення тривалості зв'язності відповідають третьому сценарію, а найменші – другому (при фіксованому радіусі зони покриття та швидкості переміщення вузлів). Так середня тривалість зв'язності з'єднання «БПЛА-пішохід» на випадок «різнобою» становитиме порядку 36 хвилин, а з'єднання «БПЛА-автомобіль» порядку 5 хвилин. Отримані результати можуть бути при розробці методу підвищення пропускної здатності на основі управління положенням БПЛА, що розглядається в главі 24.

### Контрольні запитання та завдання

1. Які вихідні дані потрібні для розв'язання задачі оцінки зв'язності вузлів епізодичних радіомереж із БПЛА в зоні надзвичайної ситуації ?
2. Які існують обмеження в задачі оцінки зв'язності вузлів епізодичних радіомереж із БПЛА в зоні надзвичайної ситуації ?
3. Сформулюйте етапи методики оцінки зв'язності вузлів епізодичних радіомереж із БПЛА.
4. В чому полягають особливості математичної моделі оцінки максимальної дальності радіозв'язку вузлів епізодичних радіомереж із БПЛА ?
5. За яким алгоритмом відбувається формування зони обслуговування БПЛА ?
6. Як виконується розрахунок тривалості зв'язності при умові прямої радіо видимості між вузлами ?
7. Як виконується розрахунок тривалості зв'язності при умові ретрансляції через проміжний вузол?
8. Які моделі переміщення вузлів мережі пропонується використовувати при імітаційному експерименті ?

## Глава 23. Математичні моделі оцінки показників функціонування епізодичних радіомереж із БПЛА в зоні надзвичайної ситуації

Задача підвищення пропускної здатності ЕРМ із ТА вирішується одночасно із забезпечення вимог до показників якості маршрутів передачі даних. Тому в даній главі досліджуються аналітичні моделі розрахунку трафіку в мережі, аналітичні моделі оцінки пропускної здатності та середньої затримки передачі в маршрутах мережі.

### 23.1. Вихідні дані, обмеження та допущення

Вихідними даними для оцінки якості маршрутів передачі даних в ЕРМ є наступні параметри:  $V=11$  Мбіт/с – швидкість передачі;  $L=1024$  біт – довжина пакету даних;  $\Pi = \|l(m_{ab})\|$ ,  $a, b = \overline{1, N}$  – маршрутна таблиця найкоротших шляхів, де  $l(m_{ab})$  – кількість ретрансляцій в маршруті;  $\Gamma = \|\gamma_{ab}(t)\|$  – матриця тяжіння, де  $\gamma_{ab}$  – середня інтенсивність потоку пакетів (на ЧНН), що надходить в маршрут  $m_{ab}$  між парою відправник-адресат  $a-b$ ;  $s^o = 5,5$  Мбіт/с – мінімально допустиме значення ПЗ маршруту  $m_{ab}$ ;  $t_3^o = 400$  мс – максимально допустиме значення затримки передачі в маршруті  $m_{ab}$  (для послуг реального часу та передачі даних з невисокою інтерактивністю згідно ІТУ Y.1541). Позначимо через  $\gamma$  повний трафік користувача, тобто

$$\gamma = \sum_{a=1}^N \sum_{b=1}^N \gamma_{ab}.$$

Обмеження при оцінці якості маршрутів в ЕРМ можна записати наступним чином:  $\Omega_2: \{s(m_{ab}) \geq s^o, t_3(m_{ab}) \leq t_3^o, a, b = \overline{1, N}\}$ .

Також було зроблено наступні допущення:

- 1) в мережі передається однорідний трафік (без пріоритетів);
- 2) потік надходження пакетів на обслуговування в вузлах мережі підкоряється пуасонівському розподіленню, тобто  $A(\tau) = 1 - \exp(-\lambda\tau)$ , де  $\lambda$  – інтенсивність надходження пакетів на обслуговування.

3) Розподілення часу обслуговування  $B(t)$  може бути довільним та визначається двома параметрами: інтенсивністю обслуговування  $\mu$  та дисперсією  $D$ .

4) спосіб обслуговування пакетів у вузлах мережі – з очікуванням без обмеження довжини черги.

5) дисципліна обслуговування пакетів в буфері – FIFO («перший прийшов – перший обслуговується»);

б) вважається, що виконується умова не зростання черги  $1/\mu < 1/\lambda$ .

### **23.2. Методика оцінки показників функціонування епізодичних радіомереж із БПЛА**

Таким чином *необхідно*: оцінити відповідність сукупності маршрутів передачі даних  $M = \{m_{ab}\}, a, b = \overline{1, N}$  вимогам до параметрів функціонування  $\Omega_2$ .

Методика вирішення даної задачі буде мати наступний *порядок оцінки*:

*Крок 1.* Розрахунок інтенсивності трафіку в каналах кожного маршруту  $m_{ab}$ , використовуючи моделі, що розглядаються в підрозділі 23.3.

*Крок 2.* Розрахунок пропускної здатності (ПЗ) та середньої затримки передачі в кожному маршруті  $m_{ab}$  (згідно матриць  $\Gamma$  та  $\Pi$ ), використовуючи аналітичні співвідношення, що розглядаються в підрозділі 23.4.

*Крок 3.* Перевірка умов  $\Omega_2$ : якщо  $s(m_{ab}) < s^o$  або  $t_s(m_{ab}) > t_s^o$ , тоді якість маршрутів в ЕРМ незадовільна, що потребує певного управлінського впливу (наприклад, виведення (переміщення) БПЛА).

### **23.3. Розрахунок інтенсивності трафіку в каналах епізодичних радіомереж із БПЛА**

На основі апарату теорії телекомунікаційних мереж та теорії масового обслуговування розроблено прості математичні моделі для розрахунку інтенсивності трафіку в каналах ЕРМ (рис. 23.1).

Використовуючи запропоновані моделі можемо легко розрахувати інтенсивність трафіку в кожному каналі ЕРМ, як суму інтенсивностей потоків між заданою парою відправник-адресат  $\gamma_{ab}$ , маршрути яких проходять через заданий канал.

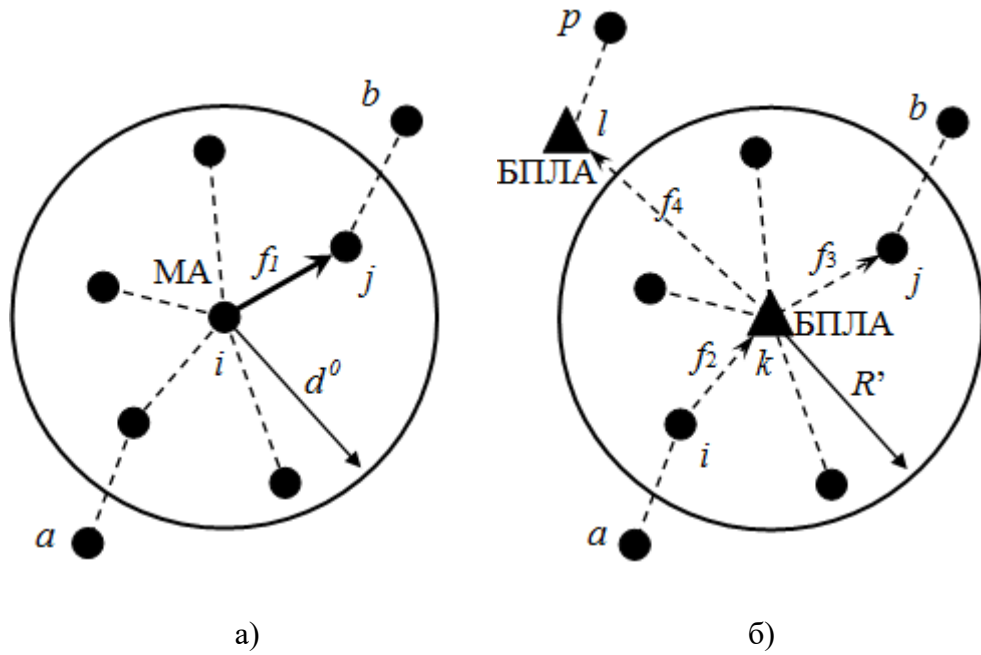


Рис. 23.1. Моделі для розрахунку інтенсивності трафіку в каналі МА-МА (а) та в каналах МА-БПЛА, БПЛА-БПЛА (б)

Тоді для широкоповного одночастотного каналу МА-МА інтенсивність трафіку можна визначити наступним чином:

$$\lambda_i^{(1)} = \sum_{j \in N_i} \sum_{a=1}^N \sum_{b=1}^N \gamma_{ab}, \quad a, b: i \in m_{ab}, \quad a \neq b \quad (23.1)$$

де  $N_i$  – множина сусідів  $i$ -ого МА, включаючи  $i$ .

Для спільного каналу МА-БПЛА та широкоповного каналу БПЛА-МА відповідно:

$$\lambda_{ik}^{(2)} = \sum_{i \in N_k} \sum_{a=1}^N \sum_{b=1}^N \gamma_{ab}, \quad a, b: ik \in m_{ab}; \quad (23.2)$$

$$\lambda_{kj}^{(3)} = \sum_{i \in N_k} \sum_{a=1}^N \sum_{b=1}^N \gamma_{ab}, \quad a, b: kj \in m_{ab}, \quad a \neq b,$$

де  $N_k$  – множина МА, що покриті  $k$ -м БПЛА.

Для дуплексного багаточастотного каналу БПЛА-БПЛА (в одному напрямку) відповідно:

$$\lambda_{kl}^{(4)} = \sum_{a=1}^N \sum_{b=1}^N \gamma_{ab}, \quad a, b: kl \in m_{ab}, \quad k \neq l, \quad a \neq b. \quad (23.3)$$

## 23.4. Дослідження пропускної здатності каналів епізодичних радіомереж із БПЛА

Основними характеристиками, що визначають ефективність функціонування протоколу множинного доступу до загального каналного ресурсу, є: середня швидкість передачі, середня затримка передачі та граничне значення трафіку, при якому досягається межа стійкої роботи мережі (межа стійкості). Ці параметри визначаються як функція системних параметрів, серед яких основними є інтенсивність трафіку, швидкість передачі, довжина пакету, геометричні розміри мережі або узагальнюючий їх інтервал уразливості.

*Середня швидкість передачі* представляє собою математичне очікування випадкової величини, що визначає кількість безконфліктно переданих пакетів за одиницю часу. В даній роботі за одиницю часу прийнято час передачі самого пакету  $T$ , тоді максимальним теоретичним значенням середньої швидкості передачі буде одиниця. Максимальне значення середньої швидкості передачі у літературі часто називають *пропускною здатністю* (іноді – ємністю) протоколу МД.

*Середня затримка передачі* – це нормований інтервал часу (в одиницях тривалості передачі пакету  $T$ ) з моменту появи на вході вузла запиту на передачу до завершення безконфліктної передачі пакету даних. При чому враховується тільки два фактори, що перешкоджають передачі: затримка доступу, тобто час до моменту отримання дозволу на передачу, та затримка повторної передачі, що зумовлена лише пошкодженням пакетів при їх інтерференції. Дана характеристика розраховується у наступному підрозділі.

Граничне значення трафіку характеризує область застосування протоколу МД у конкретних умовах, а також визначає кількість вузлів у зоні дії МА та БПЛА, що визначається також у подальших підрозділах.

Для аналізу вище вказаних характеристик та визначення їх потенційно можливих значень будемо використовувати аналітичні моделі, що будуються із використанням елементів теорії відновлення та припущення, що кількість абонентів нескінченно велика, кожен з яких генерує пакети з нескінченно малою швидкістю. Останнє є фактично припущенням про пуасонівський характер процесу надходження пакетів для передачі по радіоканалу. При цьому інтенсивність надходження пакетів вимірюється кількістю за час

передачі пакету  $T$  та позначається буквою  $G$ . Середня швидкість передачі  $S$  також вимірюється в кількості пакетів за час  $T$ , але що передані безконфліктно. Припускається також, що всі пакети мають однакову довжину  $L$ . Данні моделі оперують із середніми значеннями часових відрізків стану радіоканалу. Також будемо вважати, що службові пакети (квитанції), що підтверджують вдалий прийом пакету даних, мають нормовану довжину та передаються по окремому каналу безконфліктно.

Під інтенсивністю  $G$  надходження пакетів на передачу, що була вказана раніше, будемо розуміти сумарну інтенсивність, що включає як вхідний потік первинних пакетів  $Z_1$ , так і потік пакетів  $Z_2$ , що передаються повторно. Потік  $Z_1$  кожного вузла формують як власні пакети, так ті, що потрібно ретранслювати від інших вузлів (визначаються згідно заданої матриці тяжіння  $\Gamma$ ), а потік  $Z_2$  визначаються процедурними характеристиками протоколу МД та значеннями його системних параметрів, а саме інтервалом часу перед повторною передачею. На практиці для стійкої роботи мережі інтенсивність поточного трафіку не має перевищувати  $0,8g$ , де  $g$  – поріг стійкості, що визначається типом протоколу МД.

*Пропускна здатність маршруту* визначається мінімальною пропускною здатністю каналу, що входить до його складу, тобто

$$s(m_{ab}) = \min_{(i,j) \in m} \{s(c_{ij})\}. \quad (23.4)$$

Під *пропускною здатністю каналу* будемо розуміти середню швидкість передачі пакетів, тобто середню кількість безконфліктно переданих пакетів за інтервал часу (в даній роботі за одиницю часу прийнято час передачі самого пакету  $T = \mu$ ). Дана величина визначається протоколом множинного доступу, що використовуються в каналі. Для каналів МА-МА обрано протокол множинного доступу із сигналом «зайнято» (МДСЗ) для вирішення проблеми «скритих абонентів». Для каналів МА-БПЛА-МА обрано адаптивний протокол множинного із резервуванням пакетів (АПР), що дозволяє надсилати не один, а блоки  $n$  пакетів та управляти розміром блоку згідно інтенсивності трафіку, створюваного абонентами, на етапі оперативного управління. Для каналів БПЛА-БПЛА обрано частотне ущільнення згідно стільникового принципу в поєднанні з динамічним розподілом часового інтервалу всередині окремої комірки, тобто пакети

передаються на окремій частоті фактично без колізій і ПЗ таких каналів дорівнює 1.

Середню швидкість передачі каналу МА-МА, що працює на частоті  $f_1$  згідно протоколу МДСЗ, можна визначити за наступною формулою :

$$s_{ij}^{(1)} = \frac{G_{\Sigma}^{(1)}}{(\exp(2aG_{\Sigma}^{(1)}) - 1)((5a + \tau + 0,5)G_{\Sigma}^{(1)} + \exp(-2aG_{\Sigma}^{(1)})) + G_{\Sigma}^{(1)}(4a + \tau + 1) + 1}, \quad (23.5)$$

де  $a$ ,  $\tau$  – нормований максимальний час розповсюдження радіосигналу та сигналу «зайнято»;

$$G_{\Sigma}^{(1)} = \frac{\lambda_i^{(1)}}{\mu} = \lambda_i^{(1)}T - \text{сумарна інтенсивність трафіку, що надходить в канал МА-}$$

МА, нормована до тривалість передачі пакету  $T$ .

Підставляючи наступні вихідні дані:  $V=11\text{Мбіт/с}$ ,  $c=3*10^8\text{м/с}$ ,  $L=1024\text{біт}$  у вираз (23.5), матимемо наступні графіки залежності  $S$  на каналному рівні МА-МА (рис. 23.2 – рис. 23.4).

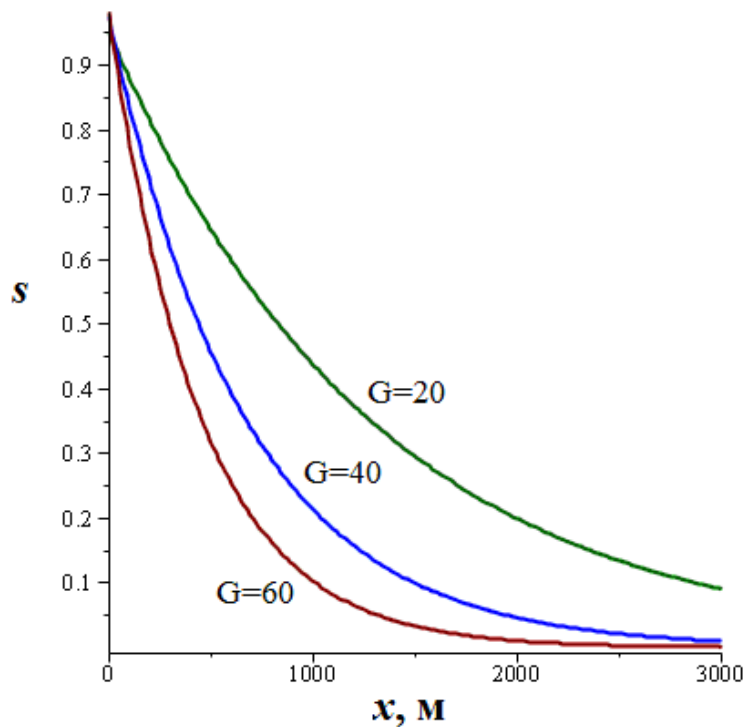


Рис. 23.2. Графіки залежності ПЗ каналу МА-МА від відстані між вузлами при різних значеннях інтенсивності трафіку

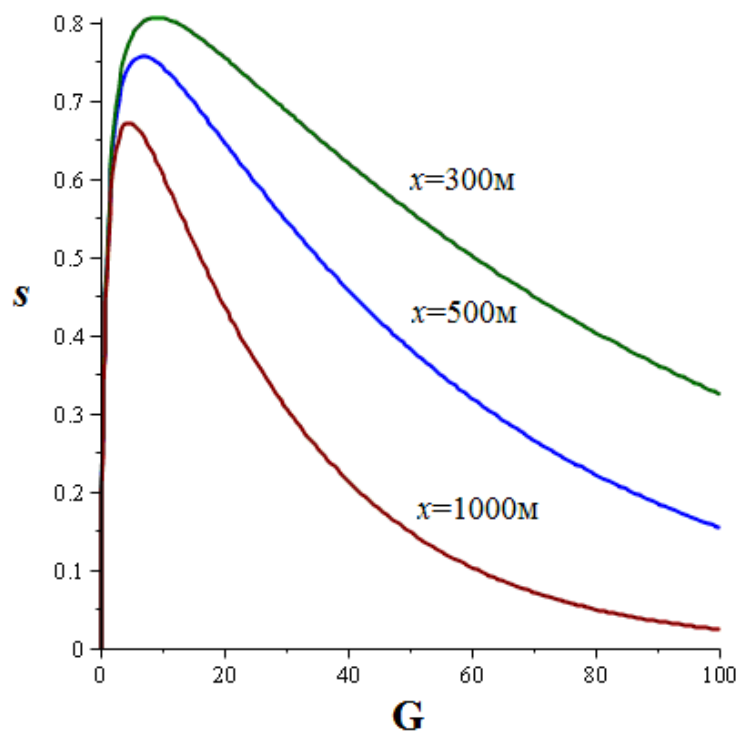


Рис. 23.3. Графіки залежності ПЗ каналу MA-MA від інтенсивності трафіку при різних значеннях відстані між вузлами

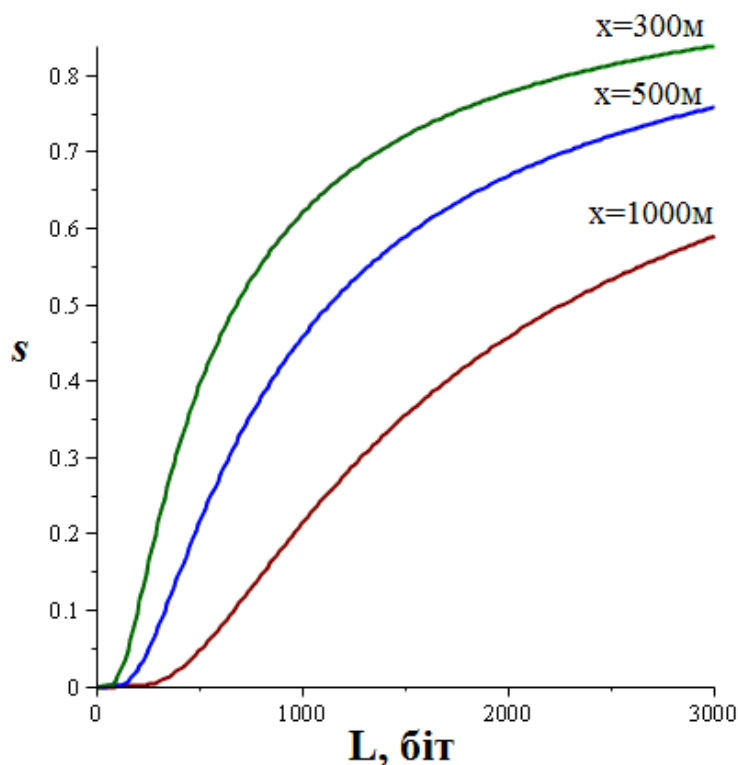


Рис. 23.4. Графіки залежності ПЗ каналу MA-MA від довжини пакету при різних значеннях відстані між вузлами (при  $G=40$ )

Аналізуючи графіки на рис. 23.2, можемо бачити, що середня швидкість передачі монотонно спадає при збільшенні максимальної відстані

між вузлами  $x$ . Це пов'язано з тим, що при збільшенні  $x$  зростає зона уразливості  $a$ , що спричиняє ріст кількості колізій при організації МД. Також згідно рис. 4.3 збільшення  $x$  призводить до зменшення межі стійкості  $g$  та діапазону стійкої роботи протоколу МД. Так при  $x=300\text{м}$  ( $a=0,011$ )  $g$  становить приблизно 9 ( $s_{\max} \approx 0,81$ ), при  $x=500\text{м}$  ( $a=0,0183$ )  $g \approx 6,8$  ( $s_{\max} \approx 0,76$ ), а при  $x=1000\text{м}$  ( $a=0,0366$ )  $g \approx 4,6$  ( $s_{\max} \approx 0,67$ ). Згідно рис. 4.4 проблему збільшення пропускної здатності при заданому  $x$  (або ж еквівалентно збільшення максимальної відстані між МА  $x$  при заданій  $s$ ) можна вирішити шляхом збільшення довжини пакету  $L$ , однак при цьому зростає ймовірність ураження пакета завадою та час доставки пакетів в мережі. Також цю проблему можна вирішити шляхом зменшення швидкості передачі  $V$  (тобто виділеної частотної смуги  $\Delta f$ ), але це – нераціонально, оскільки при цьому знижують інші основні характеристики мережі – пропускна здатність та коефіцієнт використання смуги частот. Отже, вибір максимальної відстані між вузлами (радіусу радіо покриття МА)  $x$  визначається мінімальним допустимим значенням пропускної здатності протоколу МД при заданих  $V$  та  $L$ .

Організація доступу до каналного ресурсу мережі БПЛА можливо декількома способами:

1. Організація  $n$ - частотного моно каналу, до якого підключені всі абоненти. Такий спосіб дозволяє кожен з переданих абонентом пакетів ретранслювати через мережу ретрансляторів в усіх стільниках мережі. Однак для уникнення самозбудження ланцюжка ретрансляторів, кожен з них після передачі пакету повинен мати період нечуттєвості, рівний  $2a+1$ , що значно збільшує затримку передачі кожного пакету. Окрім цього мобільність БПЛА значно ускладнює дотримання захисних інтервалів між стільниками, що працюють на однакових частотах. Ці проблеми можна вирішити шляхом застосування більшої кількості робочих частот та розробки ефективного методу розподілення радіочастотного ресурсу, адаптивного до переміщення БПЛА у просторі, що може бути предметом окремого дослідження.

2. Вище перелічені проблеми також можна вирішити за допомогою спеціальних процедур МД та пакетної маршрутизації, що можуть значно зменшити кількість ретрансляції та збільшити пропускну здатність мережі. Так завдяки адаптивним процедурам резервування пакетів та вирішення

конфліктів можливо забезпечити пропускну здатність МД  $S$  не нижче 0,9 при  $a=0,01$  у досить широкому діапазоні трафіку  $G$ . Інтервал уразливості  $a=0,01$  відповідає відстані між ретрансляторами порядку 5-15км (в залежності від системних параметрів), що цілком задовольняє для організації мережі на основі мініатюрних БПЛА. Більш того, ці процедури на відміну від звичайного МДСЗ дозволять швидко адаптуватися до змін вхідного трафіку, викликаного мобільністю МА та БПЛА.

Отже, застосовуючи, наприклад, адаптивний протокол випадкового МД із резервуванням (АПР), значення середньої швидкості передачі у каналах МА-БПЛА можна визначити наступним чином:

$$s_{ik}^{(2)} = \frac{GN \exp(-aG)}{1 + GB}, \quad (23.6)$$

де  $N$  – розмір блоку пакетів, що резервуються;

$B$  – середній час зайнятого стану каналу, який можна визначити наступним чином:

$$B = (N + b + 2a - r) \exp(-aG) + b + a + r, \quad (23.7)$$

де  $b$  – тривалість пакету резервування, що надсилає абонент, та пакету дозволу на передачу, що надсилає у відповідь ретранслятор;

$r$  – тривалість часу від надходження останнього конфліктного пакету, який можна визначити наступним чином:

$$r = a - \frac{1 - \exp(-aG)}{G} \quad (23.8)$$

Враховуючи, що  $G$  – це інтенсивність надходження блоків з  $N$  пакетів, пропускну здатність каналів МА-БПЛА, БПЛА-БПЛА можна записати так:

$$s_{ik}^{(2)} = \frac{GN \exp(-aG)}{1 + G \left( (N + b + a + \frac{1 - \exp(-aG)}{G}) \exp(-aG) + b + 2a - \frac{1 - \exp(-aG)}{G} \right)} \quad (23.9)$$

Підставляючи наступні вихідні дані:  $V=11$  Мбіт/с,  $c=3 \cdot 10^8$  м/с,  $L=1000$  біт у вираз (23.9), матимемо наступні графіки залежності  $S$  каналів МА-БПЛА та БПЛА-БПЛА (рис. 23.5 – рис. 23.9).

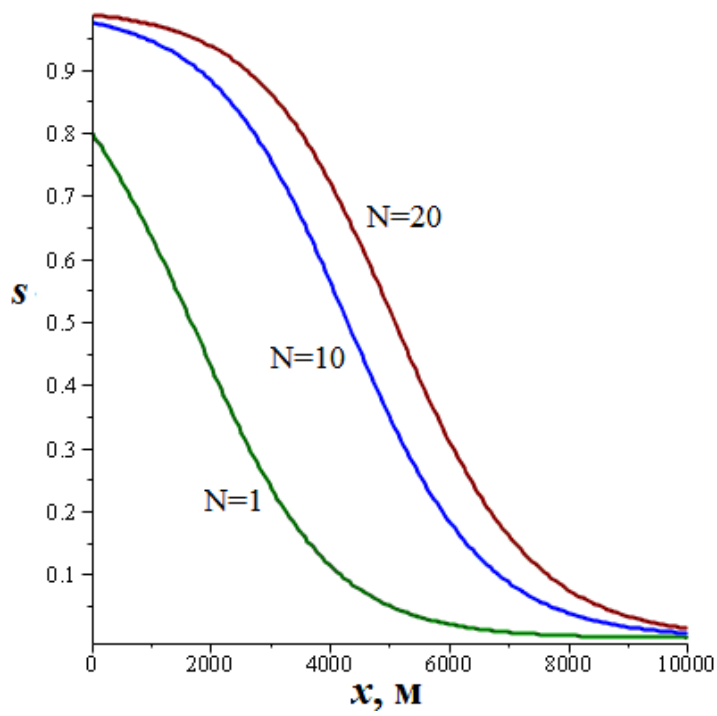


Рис. 23.5. Графіки залежності ПЗ каналу БПЛА-БПЛА (БПЛА-МА) від відстані між вузлами при різних значеннях розміру блоку резервування (при  $G=20$ )

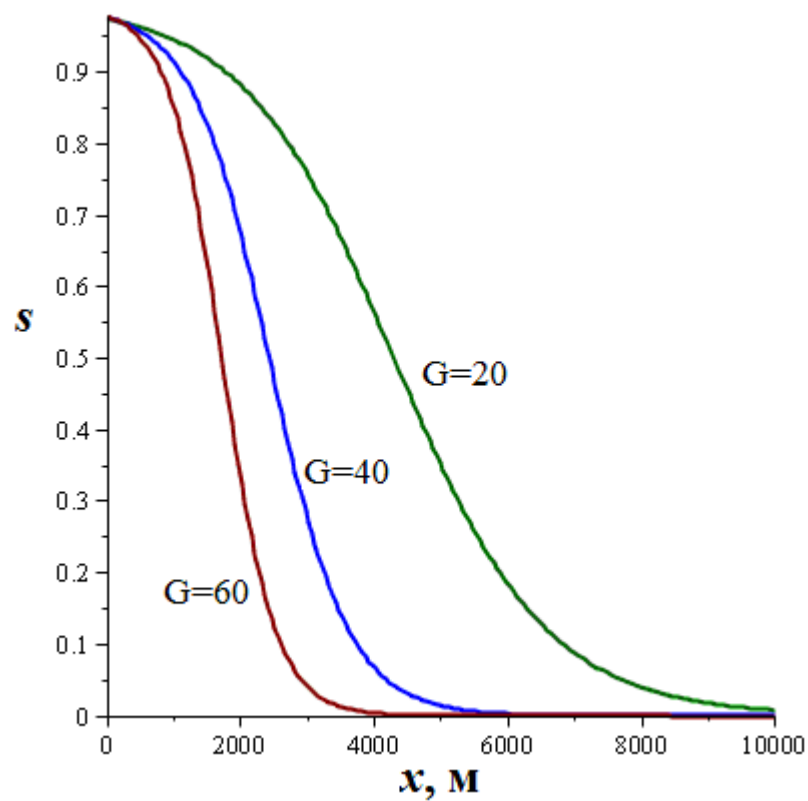


Рис. 23.6. Графіки залежності ПЗ каналу БПЛА-БПЛА (БПЛА-МА) від відстані між вузлами при різних значеннях інтенсивності трафіку (при  $N=10$ )

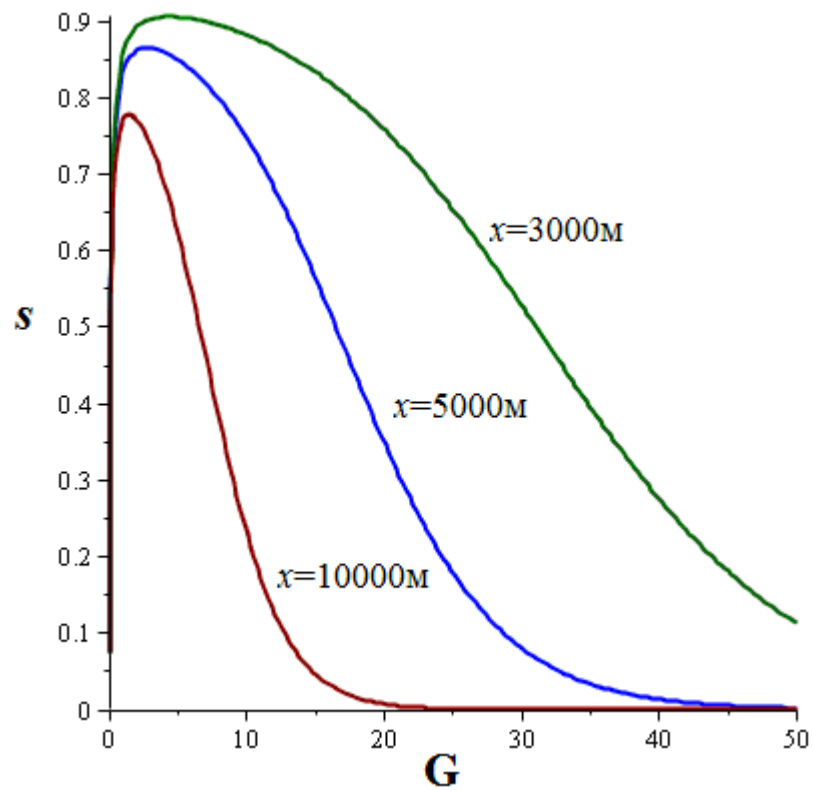


Рис. 23.7. Графіки залежності ПЗ каналу БПЛА-БПЛА (БПЛА-МА) від інтенсивності трафіку при різних значеннях відстані між вузлами (при  $N=10$ )

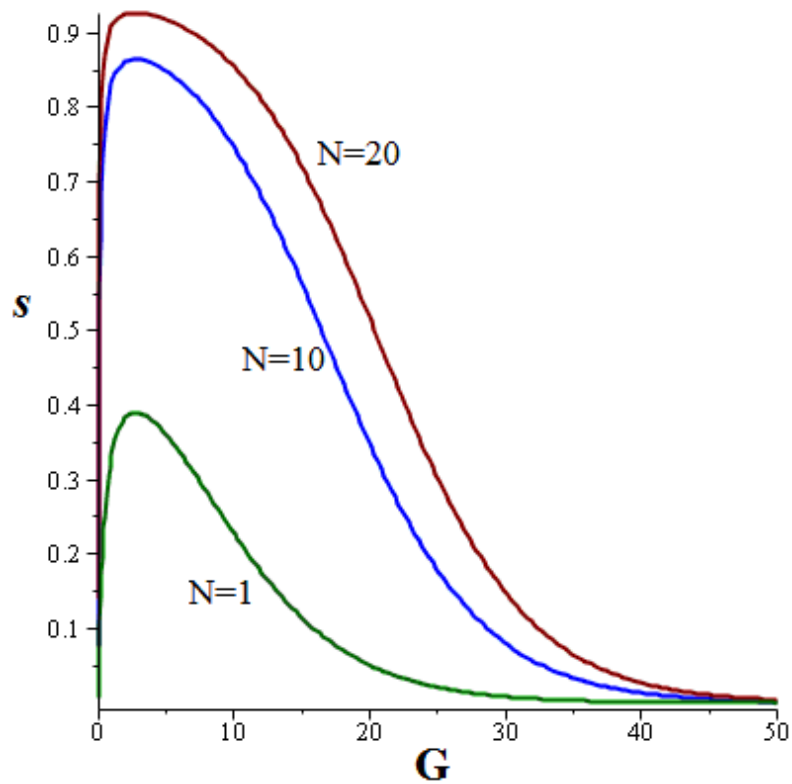


Рис. 23.8. Графіки залежності ПЗ каналу БПЛА-БПЛА (БПЛА-МА) від інтенсивності трафіку при різних значеннях розміру блоку резервування (при  $x=5000\text{m}$ )

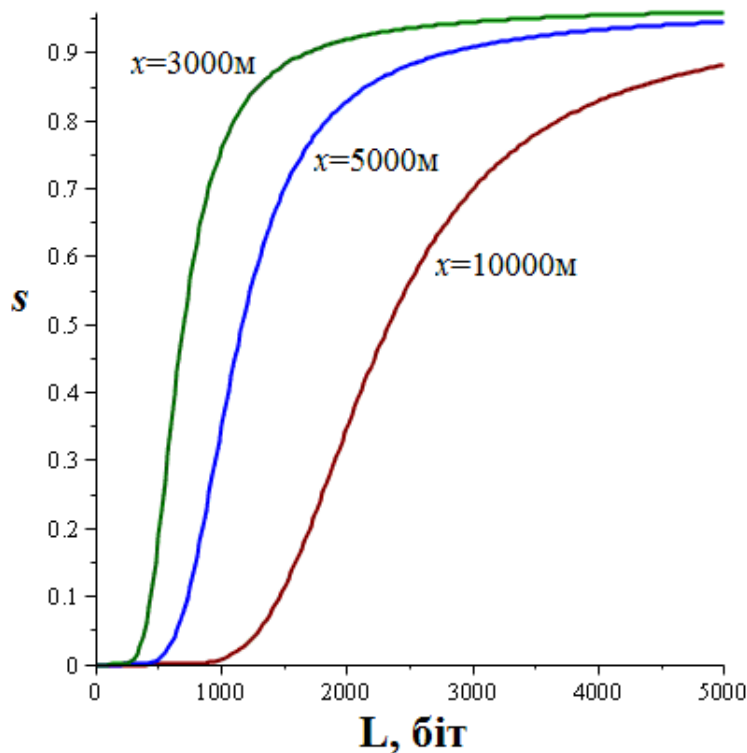


Рис. 23.9. Графіки залежності ПЗ каналу БПЛА-БПЛА (БПЛА-МА) від довжини пакету при різних значеннях відстані між вузлами (при  $N=10$ ,  $G=20$ )

Аналізуючи рис. 23.5, можемо бачити, що застосування блочного ( $N=10$ ,  $N=20$ ) резервування дає змогу значно розширити максимальну відстань між БПЛА та МА при заданій пропускну здатності на відміну від звичайного гнучкого протоколу МДСЗ ( $N=1$ ). Також згідно рис. 23.6 – рис. 23.9 можна бачити, що протокол АПР дозволяє значно розширити діапазон можливої інтенсивності трафіку та підвищити стійкість роботи мережі БПЛА при зміні вхідного навантаження.

Тоді використовуючи аналітичні моделі для обраних протоколів МД та вихідні дані, що вказані вище, отримано графіки залежності ПЗ каналів МА-МА та МА-БПЛА-МА від вхідного навантаження  $G$  (середньої кількості пакетів, що надійшли протягом часу передачі одного пакету  $T = \mu$ ) (рис. 23.10).

Аналізуючи графіки на рис. 23.10, можемо бачити, що канали МА-БПЛА-МА мають більшу пропускну здатність ніж канали МА-МА при однаковій інтенсивності трафіку (навантаженні) у діапазоні  $0 \dots 24$ , що визначається мінімально можливим рівнем ПЗ ( $s^0 = 0,5$ ). У такому випадку для підвищення ПЗ мережі доцільніше буде передавати дані через мережу

телекомунікаційних аероплатформ (БПЛА) при умові, що затримка передачі в маршруті не перевищує задану.

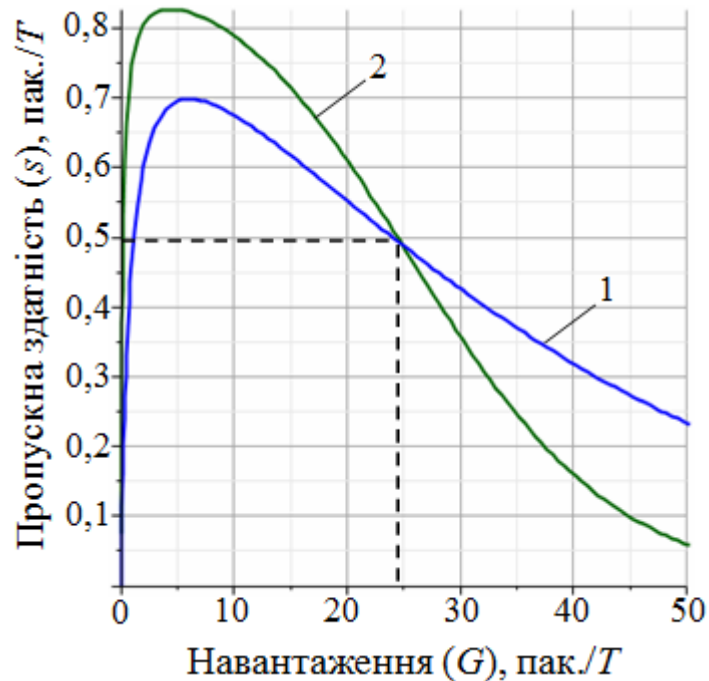


Рис. 23.10. Залежність пропускної здатності каналів ЕРМ від вхідного навантаження: 1 – канал МА-МА (МДС3), 2 – канал МА-БПЛА-МА (АПР, N=5)

### 23.5. Дослідження затримки передачі в каналах епізодичних радіомереж із БПЛА

Організація інформаційної взаємодії на мережному рівні між будь-якою парою абонентів мережі потребує наявності маршруту передачі даних заданої якості (QoS). Критерієм якості або метрикою маршруту (наприклад, при передачі голосового трафіку) може виступати кількість ретрансляцій або ж величина затримки передачі пакету із кінця в кінець через мережу проміжних вузлів-ретрансляторів. У загальному вигляді ця величина матиме наступний вигляд:  $D = T_{пак} + T_{дост} + T_{росп} + T_{обр} + T_{буф}$ , де  $T_{пак}$  – час пакетизації (формування пакету у вузлі-відправнику), що залежить від типу трафіку (алгоритму формування пакету, наприклад, для голосу – від типу кодеку);  $T_{дост}$  – середній час затримки доступу до каналу (для випадкового МД), що залежить від інтенсивності трафіку та пропускної спроможності каналу;  $T_{росп}$  – час розповсюдження сигналу в середовищі передачі (не залежить від типу трафіка);  $T_{обр}$  – час обробки пакету в проміжних вузлах мережі

(залежить від типу трафіку);  $T_{\text{буф}}$  – час затримки у буфері проміжних вузлах мережі (залежить від типу трафіку, дисципліни обслуговування, пріоритетності трафіку (SLA)). Для пакету IP середньої довжини 576 біт та при швидкості кодування 64кбіт/с час пакетизації становитиме 9 мс, чим можна знехтувати при грубій оцінці затримки передачі. Також будемо вважати, що пакети, які знаходяться на приймач, відразу потрапляють в буфер, а отже  $T_{\text{обр}}$  також дорівнює нулю.

Використовуючи апарат теорії масового обслуговування ЕРМ можна представити через сукупність блоків М/Г/1 (рис. 23.11).

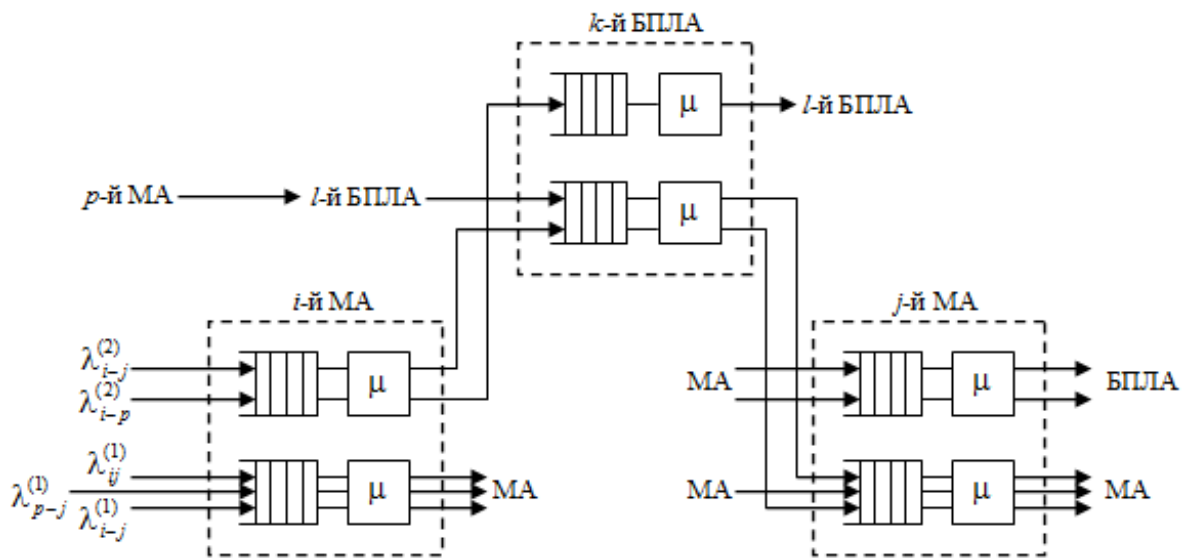


Рис. 23.11. ЕРМ, як мережа блоків типу М/Г/1

Тоді середній час перебування пакета в черзі буферу можна розрахувати за наступною формулою :

$$T_{\text{буф}} = T + W^{(z)} (1 + v^2) = \frac{1}{\mu} + \frac{G_{\Sigma}^{(z)}}{2\mu(1 - G_{\Sigma}^{(z)})} (1 + v^2) \quad (23.10)$$

де  $W^{(z)}$  – середній час затримки пакету в черзі при  $\mu = \text{const}$  ;

$v$  – коефіцієнт варіації  $\mu$  ;

$G_{\Sigma}^{(z)} = \frac{\lambda_{\Sigma}^{(z)}}{\mu} = \lambda_{\Sigma}^{(z)} T$  - сумарна інтенсивність трафіку, що надходить на частоті  $z$ .

Було проведено імітаційне моделювання середнього часу перебування пакету в блоці типу М/Г/1 (дані наведені у табл. 23.1), що підтверджують адекватність обраної аналітичної моделі (23.10).

При моделюванні застосовувався наступний алгоритм:

- 1) ПОЧАТОК.  $t_{приб} = 0; t_{кін} = 0; j=0; jj=0; N=0$ .
- 2)  $N=N+1$ .
- 3)  $t_{приб} = t_{приб} - \ln(Random)/\lambda$ .
- 4) Якщо  $t_{приб} > t_{кін}$ , то  $t_{поч} = t_{приб}$ , інакше  $t_{поч} = t_{кін}$ .
- 5)  $t_{кін} = t_{поч} + (1 + 2\sqrt{3} v(Random - 0,5))/\mu$ .
- 6)  $t = t_{кін} - t_{поч}$ .
- 7)  $j = j + t; jj = jj + t^2$ .
- 8) Якщо  $N < N_{max}$ , то перехід на крок 2, інакше на крок 9.
- 9)  $T_{буф} = j/N; \sigma_T = \sqrt{jj/N + T_{буф}^2}$ . КІНЕЦЬ.

В цьому алгоритмі використані наступні позначення:  $t_{приб}$  – момент прибуття пакету;  $t_{поч}$  – момент початку обробки пакету;  $t_{кін}$  – момент закінчення обробки пакету;  $N$  – число оброблених пакетів;  $t$  – час знаходження пакету в блоці;  $Random$  – генератор випадкових чисел в діапазоні (0, 1). Крок 3 алгоритму дозволяє задавати експоненціальний закон  $A(t)$  розподілення інтервалів часу між моментами надходження пакетів з інтенсивністю  $\lambda$ , а крок 5 – закон  $B(t)$  розподілення часу обслуговування з інтенсивністю  $\mu$  та коефіцієнтом варіації  $v$ .

Таблиця 23.1

**Результати моделювання середнього часу перебування пакету в блоці типу M/G/1**

Кіл-ть оброб. Пакетів	Інтен-ть надх. пакетів $\lambda$	Час обслуг. $1/\mu$	Коеф. варіації $v$	$T_{буф}$ (результати розрах.)	Результати імітаційного моделювання	
					$T_{буф}$	СКВ $\sigma_T$
10000	$1 \cdot 10^3$	$9,1 \cdot 10^{-5}$	0,5	$4,54 \cdot 10^{-6}$	$4,49 \cdot 10^{-6}$	1,258
10000	$1 \cdot 10^4$	$9,1 \cdot 10^{-5}$	0,5	$4,13 \cdot 10^{-4}$	$4,02 \cdot 10^{-4}$	1,412
10000	$1,0999 \cdot 10^4$	$9,1 \cdot 10^{-5}$	0,5	0,45	0,32	1,505

Тоді середній час затримки передачі пакету на ланці МА-МА матиме наступний спрощений вигляд:

$$D_{ij} = \left( \frac{G_{\Sigma}^{(1)}}{s_{ij}^{(1)}} - 1 \right) (4a + \tau + 0,5 + Y + X) + 4a + \tau + 1 + W_i^{(1)}, \quad (23.11)$$

де  $X$  – нормований середній час затримки повторної передачі;

$Y$  – нормований середній час початку передачі останнього конфлікуючого пакету.

У свою чергу середній час затримки передачі пакету на ланці БПЛА-БПЛА (МА-БПЛА) матиме наступний вигляд:

$$D_{ik} = \frac{1}{N} \left( \left( \frac{1}{P} - 1 \right) (X + (b + a + r)P_l) + N + a + W_i^{(2)} + W_k^{(3)} \right), \quad (23.12)$$

де  $P = \frac{\exp(-aG)}{1 + GB}$  - ймовірність вдалої передачі блоку пакетів;

$P_l = \frac{1}{1 + GB}$  - ймовірність вільного стану радіоканалу.

Оскільки канали БПЛА-БПЛА працюють з частотним розділенням, не заважаючи один одному, то будемо вважати, що максимальна затримка передачі в таких каналах однакова і визначається лише затримкою розповсюдження та затримкою в буфері проміжних вузлів, тобто  $D_{kl} = a + W_k^{(4)}$ .

Підставляючи наступні вихідні дані:  $V=11$ Мбіт/с,  $c=3 \cdot 10^8$ м/с,  $L=1000$ біт,  $b=0,1$ ,  $X=2$  у вирази (23.11), (23.12), матимемо наступні графіки залежності середньої затримки передачі однієї ланки МА-МА (при  $x=600$ м) та однієї ланки БПЛА-МА (при  $x=3000$ м,  $N=10,20,30$ ) від інтенсивності трафіку (рис. 23.12).

Аналізуючи рис. 23.12 можемо бачити, що при інтенсивності трафіку менше ніж 30 менший час затримки має ланка мережі із застосуванням БПЛА, аніж ланка МА-МА, при чому граничне значення інтенсивності трафіку тим більше, чим більше розмір блоку пакетів в протоколі АПР. Також слід зазначити, що при однаковій затримці передачі ланка із БПЛА має шестикратний вииграш у відстані. Отже для мінімізації затримки при передачі інформації на великі відстані доцільніше передавати інформацію через мережу БПЛА, а на коткі – через мережу МА. Тоді виникає практичний інтерес визначення граничної кількості ретрансляцій через МА, при якій необхідно переключатися на мережу БПЛА, що й буде розглянуто далі.

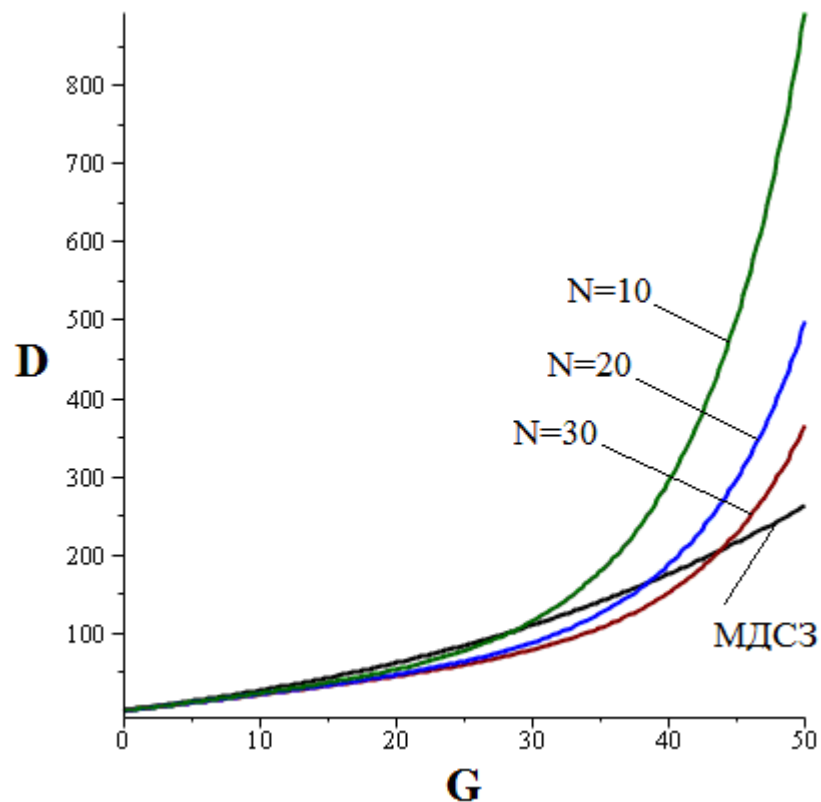


Рис. 23.12. Графіки залежності середньої затримки передачі однієї ланки МА-МА (при  $x=600\text{м}$ ) та однієї ланки МА-БПЛА-МА (при  $x=3000\text{м}$ ,  $N=10,20,30$ ) від інтенсивності трафіку

Якщо вважати, що маршрут передачі інформації складається з  $l$  ланок та кожна ланка має однакову середню інтенсивність трафіку, тоді загальна затримка пакету «із кінця в кінець» через мережі МА та мережі БПЛА становитиме відповідно:

$$t_3(m_{ab}) = D_{ij} \cdot l; \quad t_3(m_{ab}) = D_{ik} + (a + W_k)l. \quad (23.13)$$

Згідно отриманим аналітичним моделям обраних протоколів МД (23.13) було розраховано залежність середньої затримки передачі в маршруті від кількості ретрансляцій в маршруті при однаковому навантаженні в каналах (рис. 23.13).

Можемо бачити, що у даному випадку максимальна кількість ретрансляцій при передачі через мережу МА, що визначається заданою величиною затримки ( $t_3^0 = 400\text{мс}$ ), має становити не більше 14. Для скорочення кількості ретрансляцій в маршрутах доцільніше передавати пакети через мережу БПЛА. При чому всередині стільника вигідніше ретранслювати пакети при кількості ланок більше ніж 3, а через

міжплатформні лінії зв'язку – більше ніж 4. Це може бути використане при пошуку оптимального (найкоротшого) маршруту між заданою парою відправник-адресат.

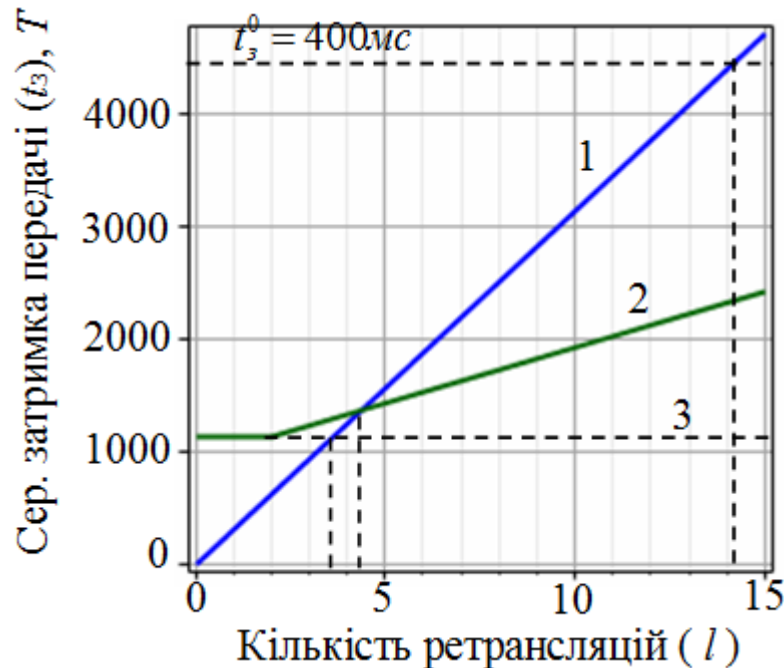


Рис. 23.13. Залежність середньої затримки в маршруті від кількості ретрансляцій для різних варіантів передачі: 1 – через мережу МА, 2 – через мережу БПЛА, 3 – через один БПЛА

### 23.6. Визначення кількості абонентів в стільниках МА та ТА

Популяція мережі (число абонентів  $n$ ) може бути визначена наступним чином. Для обраного протоколу МД встановлюється залежність пропускної спроможності каналу  $s$  від трафіку  $G$  при заданому значенні  $a$ , що визначається максимальною відстанню  $x$ . Наприклад, для гнучного МДСЗ вказана залежність відображається виразом (23.5), а для АПР – виразом (4.6). Графік вказаних залежностей показані на рис. 23.3 та рис. 23.7 відповідно. Абсциса екстремума функції  $s(G)$  або деякого її граничного значення (наприклад, 0,5) визначає гранично допустимий трафік  $G_{\max}$  для заданного протоколу при  $a=a_{\max}$ . Виходячи з вимог стійкості режиму мережі, ставимо наступну вимогу  $\sum_{i=1}^n \overline{G}_i < G_{\max}$ , де  $\overline{G}_i$  - середній трафік, що генерує кожен абонент стільника мережі.

Задаючись  $\overline{G}_i$ , знаходимо верхню границю числа абонентів стільника мережі:

$$n_{\max} \leq G_{\max} / \overline{G}_i. \quad (23.14)$$

При цій кількості абонентів не виключено, однак, що середня затримка передачі пакетів  $D$  може мати значення, що перевищують допустимі. Тому скористуємось залежністю величини середньої затримки передачі від пропускної здатності протоколу МД  $D(S)$ , що визначається виразом (23.11) для МДСЗ та виразом (23.12) для АПР. Задаючись допустимим значенням  $D_{\text{доп}}$ , знаходимо величину оптимального значення пропускної здатності протоколу  $s_{\text{опт}}$  та відповідну величину трафіку  $G_{\text{опт}}$ . Тоді оптимальну кількість абонентів стільника мережі можна визначити наступним чином:

$$n_{\text{опт}} = G_{\text{опт}} / \overline{G}_i. \quad (23.15)$$

Використовуючи вирази (23.5) та (23.6) та обмеження ПЗ  $s^0=0,5$  отримаємо наступні граничні допустимі трафіки: для стільника МА трафік пакетів становитиме  $G'_{\max} \approx 35$  при  $x=600\text{м}$ , для стільника БПЛА трафік пакетів (при  $N=10$ ) становитиме  $G''_{\max} \approx 300$  при  $x=3000\text{м}$ . Тоді задаючись середнім трафіком пакетів одного абонента  $\overline{G}_i = 5$  та використовуючи вираз (23.15), максимальна кількість абонентів в зоні дії МА становитиме  $n'_{\max} \leq 7$ , а в зоні дії БПЛА – становитиме  $n''_{\max} \leq 60$ . Максимальна кількість МА в зоні дії одного БПЛА залежить від плану розподілу потоків між абонентами мережі, (тобто величини трафіку, що передається за межі стільника), але вочевидь, що не переважає декількох одиниць із міркувань щодо забезпечення заданої ЕМС та безаварійного переміщення БПЛА у просторі.

Зрозуміло, що наведений розрахунок носить суто наближений характер через те, що окрім припущення щодо пуасонівського характеру потоку пакетів, ми задаємося середнім та рівним значенням трафіку, генерованого усіма абонентами, що не виконується на практиці. Не дивлячись на це, наведений розрахунок залишається справедливим при умові неперевикнення усіма абонентами сумарної допустимої величини трафіку в мережі, що обмежується максимальною середньою затримкою передачі. Саме цей параметр, виміряний кожним абонентом мережі шляхом обліку часу отримання квитанцій на ряд переданих пакетів, може служити обмеженням на генерований їм трафік або визначати момент переключення передачі пакетів з мережі МА на мережу БПЛА та навпаки. Знаючи кількість

наземних абонентів, їх координати та трафік створюваний ними, БПЛА може знайти своє оптимальне розміщення у просторі, що дасть змогу не тільки забезпечити геометричну зв'язність абонентів, а й підвищити інформаційні характеристики мережі (такі як пропускна здатність) завдяки з'єднання віддалених абонентів та розвантаження перевантажених каналів в районах скупчення абонентів. Зазначений принцип буде враховано при розробці методу управління топологією мережі БПЛА в главі 24.

### 23.7. Висновки до глави 23

1. Розроблено прості математичні моделі для розрахунку інтенсивності трафіку в каналах ЕРМ з урахуванням особливостей функціонування мережі (МД, частотного розподілу), що дозволяють розраховувати основні показники функціонування ЕРМ, а саме пропускну здатність та затримку передачі в маршрутах мережі.

2. Показано, що пропускна здатність маршруту передачі даних визначається мінімальною пропускну здатністю каналу, що входить до його складу. Показано, що канали МА-БПЛА-МА мають більшу пропускну здатність ніж канали МА-МА при однаковій інтенсивності трафіку (навантаженні) у діапазоні  $0 \dots 24$ , що визначається мінімально можливим рівнем ПЗ ( $s^0 = 0,5$ ). У такому випадку для підвищення ПЗ мережі доцільніше буде передавати дані через мережу телекомунікаційних аероплатформ (БПЛА) при умові, що затримка передачі в маршруті не перевищує задану.

3. При необхідності передачі інформації на більші відстані, аніж дальність прямого з'єднання, необхідно застосовувати механізм ретрансляції (комутації) пакетів через проміжні вузли. Максимальна дальність зв'язку у даному випадку визначається заданою величиною затримки передачі «із кінця в кінець» для забезпечення необхідної QoS. Так для застосувань реального часу невисокої інтерактивності (VoIP, дані) максимальна кількість ретрансляцій у випадку передачі через мережу МА при  $t_3^0 = 400 \text{ мс}$  має становити не більше 14. Для скорочення кількості ретрансляцій в маршрутах доцільніше передавати пакети через мережу БПЛА. При чому всередині стільника вигідніше ретранслювати пакети при кількості ланок більше ніж 3, а через міжплатформні лінії зв'язку – більше ніж 4. Це може бути

використане при пошуку оптимального (найкоротшого) маршруту між заданою парою відправник-адресат.

### **Контрольні запитання та завдання**

1. В чому полягають основні етапи методики кількісної оцінки значень показників функціонування епізодичних радіомереж із БПЛА ?
2. Наведіть алгоритм розрахунку інтенсивності трафіку в каналах епізодичних радіомереж із БПЛА .
3. В чому полягає зміст дослідження пропускної здатності каналів епізодичних радіомереж із БПЛА ?
4. В чому полягає зміст дослідження затримки передачі в каналах епізодичних радіомереж із БПЛА ?
5. Наведіть алгоритм визначення кількості абонентів в стільниках епізодичних радіомереж із БПЛА .

## Глава 24. Метод підвищення пропускної здатності епізодичних радіомереж з управлінням положенням БПЛА в зоні надзвичайної ситуації

Суть розробленого методу полягає в систематизації розглянутих вище математичних моделей оцінки структурної зв'язності (розглядалися у главі 22) та показників якості обслуговування мобільних абонентів (розглядалися у главі 23), а також удосконаленого алгоритму пошуку квазіоптимального положення ТА, що розглядається в даному розділі, та поєднанні їх в єдину обчислювальну процедуру для досягнення близьких до екстремальних значень пропускної здатності мережі в режимі реального часу.

### 24.1. Постановка задачі у математичному вигляді

Представимо ЕРМ у вигляді стохастичного ненаправленого зваженого графу  $G(V,E)$ , що складається з множини вершин (МА та БПЛА)  $V = \{v_i\} \cup \{b_k\}$  та множини ребер  $E = \{(i,j) | d_{ij} \leq d^0\} \cup \{(i,k) | R_{ik} \leq R^0\}$ ,  $i, j = \overline{1, N}, k = \overline{1, K}$  (рис. 24.1), що визначають матрицю зв'язності  $C = \|c_{ij}\|$ , де  $c_{ij} = \{0,1\}$  - булева змінна. У якості ваги ребра може виступати протяжність відповідної радіолінії ( $d_{ij} (R_{ik})$ ).

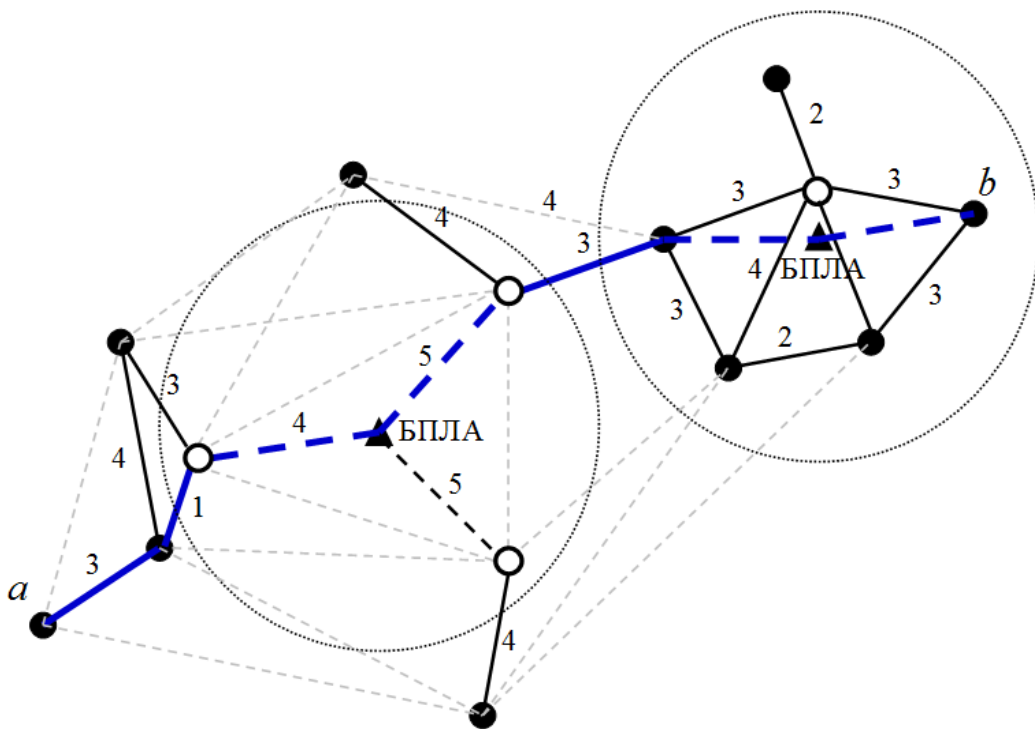


Рис. 24.1. Приклад графу  $G(V,E)$  ЕРМ, що складається з 14 наземних вузлів та двох БПЛА

Пропускна здатність мережі у даному випадку можна визначити, як сумарну пропускна здатність усіх маршрутів мережі, тобто

$$S(C) = \sum_{m=1}^M \sum_{a=1}^N \sum_{b=1}^N s(m_{ab}), a \neq b, \text{ де } s(m_{ab}) = \min_{(i,j) \in m} \{s(c_{ij})\}. \quad (24.1)$$

Нехай задані наступні *вихідні дані*:  $N$  – кількість МА;  $K$  – кількість БПЛА;  $r$  – розмір району дислокації МА, м<sup>2</sup>;  $(x_i, y_i)_t, i = \overline{1, N}$ ,  $(x_{0k}, y_{0k}), k = \overline{1, K}$  – координати МА та виведених БПЛА в момент часу  $t$ , м;  $d^0(R^0)$  – максимальна дальність зв'язку, м;  $s^0$  – мінімально допустиме значення ПЗ маршруту  $m_{ab}$ ;  $t_3^0$  – максимальна допустиме значення затримки передачі маршруту  $m_{ab}$ ;  $\Pi_i = \|l(m_{ab})\|, a, b = \overline{1, N}$  – маршрутна таблиця найкоротших шляхів, де  $l(m_{ab})$  – кількість ретрансляцій в маршруті;  $\Gamma = \|\gamma_{ab}(t)\|$  – матриця тяжіння.

Тоді можемо сформулювати наступну *задачу управління положенням телекомунікаційних аероплатформ ЕРМ*: знайти в режимі реального часу ( $T_{\text{пош.ршу.}} \ll T_{3\epsilon}$ ) координати положення БПЛА у просторі  $X_{0k}, k = \overline{1, K}$  (матрицю зв'язності  $C^*$ ), що забезпечує масимум пропускну здатності мережі  $S(C)$ :

$$C^* = \arg \max_{X_0 \in \Omega_{1,2,3}} S(C) = \arg \max_{X_0 \in \Omega_{1,2,3}} \sum_{a=1}^N \sum_{b=1}^N s(m_{ab}), a \neq b \quad (24.2)$$

при винонанні обмежень на множину управлінських дій та ресурси мережі

$$\Omega_1: \{d_{ij} \leq d^0, R_{ik} \leq R^0, D_{kl} \leq D', T_{3\epsilon ij} \geq T_{3\epsilon}^0 \forall ij, ik, kl \in m_{ab}, i, j = \overline{1, N}, k, l = \overline{1, K}\},$$

$$\Omega_2: \{s(m_{ab}) \geq s^0, t_3(m_{ab}) \leq t_3^0, a, b = \overline{1, N}\},$$

$$\Omega_3: \{2600 \leq h < 3000 \text{ м}, R \leq 1500 \text{ м}, N \leq 200, K \leq 10\}.$$

## 24.2. Опис методу

Для вирішення поставленої математичної задачі (5.2) запропонован *метод* на основі *управління положенням телекомунікаційних аероплатформ ЕРМ*, який представлений у вигляді схеми-алгоритму на рис. 24.2.

Запропонований метод складається з таких основних складових, поєднаних в єдину обчислювальну процедуру:

1. Удосконалена математична модель оцінки зв'язності вузлів ЕРМ із телекомунікаційними аероплатформи.
2. Уточнені математичні моделі оцінки показників функціонування ЕРМ із телекомунікаційними аероплатформи.
3. Удосконаленого алгоритму пошуку квазіоптимального положення телекомунікаційних аероплатформ, загальна схема якого показана на рис. 24.3.

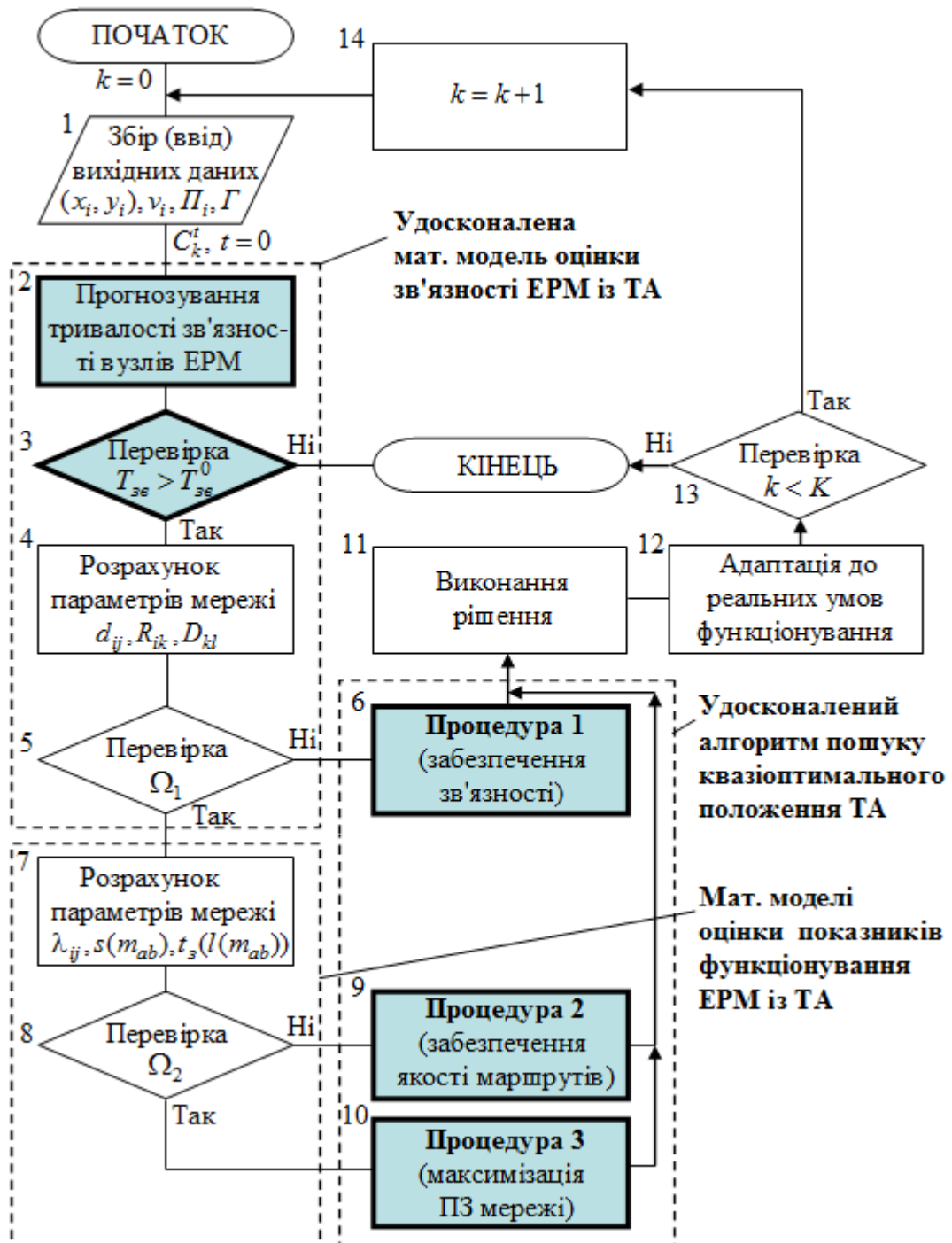


Рис. 24.2. Схема-алгоритм управління положенням телекомунікаційних аероплатформ ЕРМ для підвищення ПЗ мережі (продовження на рис. 24.3)

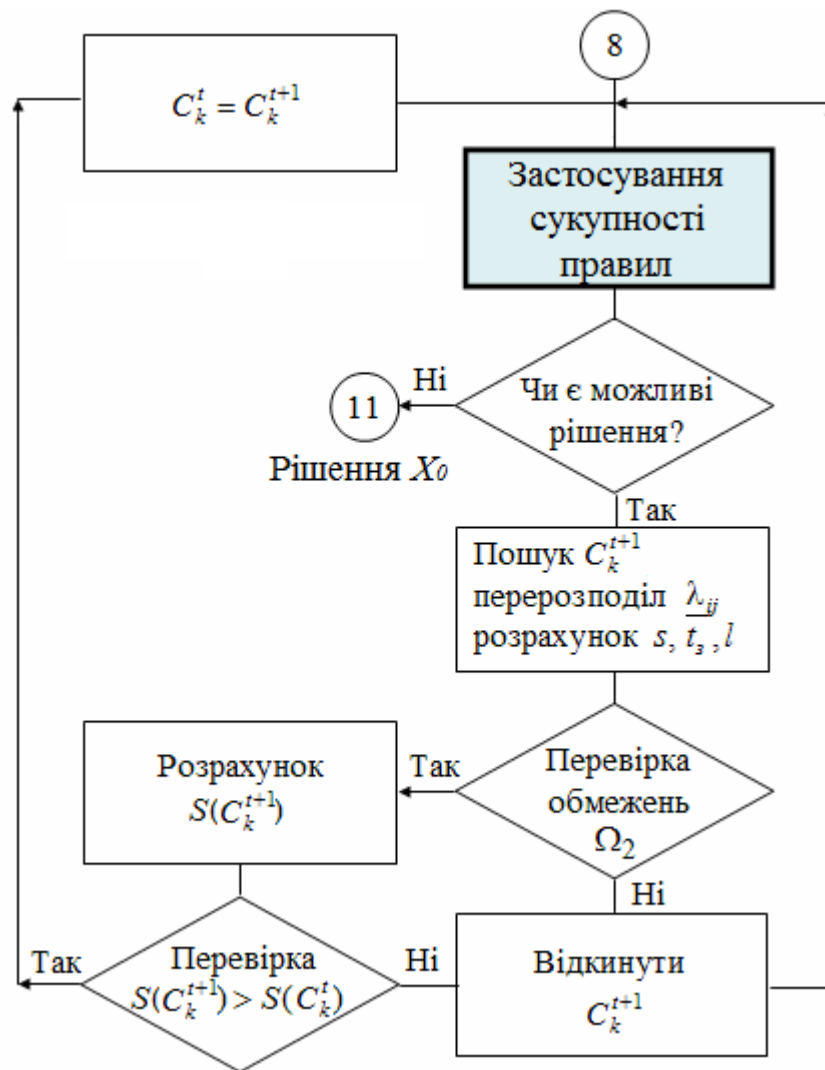


Рис. 24.3. Схема удосконаленого алгоритму пошуку квазіоптимального положення ТА із використанням сукупності правил (початок на рис. 24.2)

Цикл управління топологією (місцеположенням) мережі ТА включає в себе наступні кроки:

*Крок 1.* Збір інформації про початкову топологію мережі та ввід вихідних даних (на етапі планування) (блок 1):

- параметрів наземних вузлів:  $N, (x_i, y_i), S_i, i = \overline{1, N}$ , та виведених попередньо БПЛА:  $(x_{0j}, y_{0j}, z_{0j}), S_{0j}, j = \overline{1, K}$  (отримані через систему збору інформації по окремому каналу);

- допустимих значень параметрів:  $d^0, s^0, t_s^0 (l^0), D^0(R^0), r$ ;

- кількість БПЛА на операцію  $K$ ;

- параметри протоколу МД та протоколу маршрутизації.

Зазначені вище параметри та їх обмеження визначають початкову топологію мережі БЕМ без застосування БПЛА  $C_k, k = \overline{1, K}$ , де  $k=0$  – номер ітерації пошуку рішення (порядковий номер БПЛА).

*Крок 2.* Прогнозування тривалості зв'язності вузлів ЕРМ із ТА на основі розроблених моделей переміщення МА (блок 2) та перевірка умови  $T_{36\ ij} \geq T_{36}^0$  (блок 3), що визначає можливість БПЛА відпрацьовувати рішення до суттєвої зміни топології.

*Крок 3.* Аналіз наявності структурної зв'язності (блоки 3,4):

1. Розрахунок параметрів  $d_{ij}, D_{ik}, D_{kl}$ .

2. Перевірка виконання обмежень  $\Omega_1$  (блок 5). Якщо умови виконуються, тоді перевіряємо додаткову умову цілісності мережі  $k=1$ , інакше переходимо на крок 4.

*Крок 4.* Виконання процедури 1 забезпечення структурної зв'язності (блок 6), що розглядається нижче на основі спеціально розроблених правил.

У випадку наявності нових рішень, що забезпечують виконання умов  $\Omega_1$ , виконується вивід БПЛА у задану позицію (етап розгортання) (блок 11) та адаптація параметрів протоколу МД до реальних умов функціонування (блок 12) згідно процедур вказаних у [1] (етап оперативного управління).

Потім здійснюється перевірка наявності апаратного ресурсу (БПЛА) (блок 13). У випадку наявності ( $k < K$ ), переходимо на крок 1, інакше – КІНЕЦЬ.

*Крок 5.* Збір інформації про функціонування мережі (блок 1):

- матриці найкоротших маршрутів  $P_i$ ;

- матриці тяжіння  $G$ .

Збір зазначеної інформації може виконуватися на етапі планування у випадку наявності виведених БПЛА (через канал зв'язку із центром управління) або на етапі розгортання шляхом «зчитування» даних з будь-якого наземного вузла ЕРМ.

*Крок 6.* Аналіз вимог до якості маршрутів передачі (блоки 7,8):

1. Розрахунок параметрів функціонування мережі  $\lambda, s(m_{ab}), \bar{t}_3(l(m_{ab}))$  згідно співвідношень вказаних у розділі 4.

2. Перевірка виконання обмежень  $\Omega_2$ . Якщо умови виконуються, тоді переходимо на крок 8, інакше переходимо на крок 7.

*Крок 7.* Виконання процедури 2 для забезпечення якості маршрутів (блоки 9, 11-14), що розглядається нижче на основі спеціально розроблених правил.

У випадку наявності нових рішень, що забезпечують виконання умов  $\Omega_1$ , виконується вивід БПЛА у задану позицію (етап розгортання) (блок 11) та адаптація параметрів протоколу МД до реальних умов функціонування (блок 12) згідно процедур вказаних у [1] (етап оперативного управління).

Потім здійснюється перевірка наявності апаратного ресурсу (БПЛА) (блок 13). У випадку наявності ( $k < K$ ), переходимо на крок 1, інакше – КІНЕЦЬ.

*Крок 8.* Виконання процедури 3 для підвищення пропускної здатності мережі (блоки 10-14), загальний алгоритм якої показано на рис. 24.2:

1. Пошук нових рішень (топології із застосуванням поточного БПЛА)  $C_k^{t+1}$ , що забезпечують максимізацію ПЗ. Оскільки це є *NP*- складна задача, для зменшення складності пошуку можна застосовувати сукупність правил, а також центроїдну або решітчасту ініціалізацію, що детально розглядаються нижче.

2. Побудова маршрутних таблиць  $\Pi_i(C_k^{t+1})$ , що визначаються матрицею тяжіння та прийнятим методом маршрутизації. Перерозподіл потоків  $\lambda_{ij}$  згідно  $\Pi_i(C_k^{t+1})$ . Розрахунок параметрів  $s(m_{ab}), \bar{t}_z(l(m_{ab}))$  для існуючих пар відправник-адресат.

3. Перевірка виконання умов  $\Omega_2$  для  $C_k^{t+1}$ . Якщо  $\Omega_2$  не виконуються або  $S(C_k^{t+1}) < S(C_k^t)$ , то відкинути  $C_k^t$ , інакше  $C_k^t = C_k^{t+1}$ , здійснюється вивід БПЛА у задану позицію (етап розгортання) та адаптація параметрів протоколу МД до реальних умов функціонування (етап оперативного управління).

4. Перевірка на наявність апаратного ресурсу (БПЛА). У випадку наявності ( $k < K$ ), переходимо на крок 1, інакше – КІНЕЦЬ.

У випадку виведення усіх БПЛА кожним з них періодично на етапі оперативного управління здійснюється відпрацювання розглянутої вище схеми для перевірки необхідності зміни своєї позиції (при цьому усі вузли мережі вважаються фіксованими у заданий момент часу). При чому період відпрацювання методу управління має бути достатньо великим, щоб

побудувати маршрути та передати по ним мінімальну кількість даних, та одночасно достатньо малим, щоб топологія мережі не змінилася значним чином. Частоту відпрацювання даного методу визначається згідно прогнозування тривалості зв'язності на основі моделей переміщення МА (рятувальників), що розглядаються в главі 22.

Також припускається, що за час збору вихідних даних, виконання розрахунків та виведення БПЛА у задану позицію, топологія мережі не зміниться значним чином.

### **24.3. Удосконалений алгоритм пошуку квазіоптимального положення БПЛА**

Задача пошуку оптимального розміщення БПЛА (матриці зв'язності  $C_k$ ) в (24.2) відноситься до класу  $NP$ - повних. Застосування для її вирішення класичних методів приводить до експонентної складності. Отримання точного рішення для мережі, яка налічує сотні (тисячі) вузлів потребує значних затрат часу. Тому для скорочення повного перебору варіантів розміщення БПЛА пропонується використовувати сукупність правил, поєднаних в базу знань, що змінюють зв'язність мережі для підвищення пропускної здатності мережі та скорочують час обчислень. Це дозволяє отримати в реальному часі близькі до оптимальних рішення та використовувати запропонований алгоритм для оперативного управління положенням БПЛА. Критерієм визначення моменту перестройки топології мережі може бути не виконання однієї з умов  $\Omega_1$  або  $\Omega_2$ .

Застосування сукупності правил включає в себе ітераційну процедуру, що складається з модулів АНАЛІЗ, БАЗА ЗНАНЬ, УПРАВЛІННЯ та працює у наступній послідовності (рис. 24.4):

1. Аналіз поточного стану мережі (розрахунок параметрів ЕРМ за допомогою аналітичних моделей, наведених у главах 22,23).
2. Перевірка параметрів мережі на множині правил. Якщо немає умов для виконання правил, то процедуру закінчити, інакше – застосувати правила з бази знань для зміни зв'язності мережі.
3. Обчислення параметрів модифікованої мережі. Якщо параметри покращились, то перейти до кроку 2, інакше – відкинути останню зміну зв'язності і перейти до кроку 2.

Модуль АНАЛІЗ призначений для знаходження пропускну́ї здатності мережі  $S$ , середньої затримки передачі пакетів  $t_s$  по маршрутах при заданому вхідному навантаженні  $\Gamma$ , маршрутних рішеннях  $\Pi$ , зв'язності  $C$  та протоколі доступу до каналу за допомогою математичних моделей глави 23.

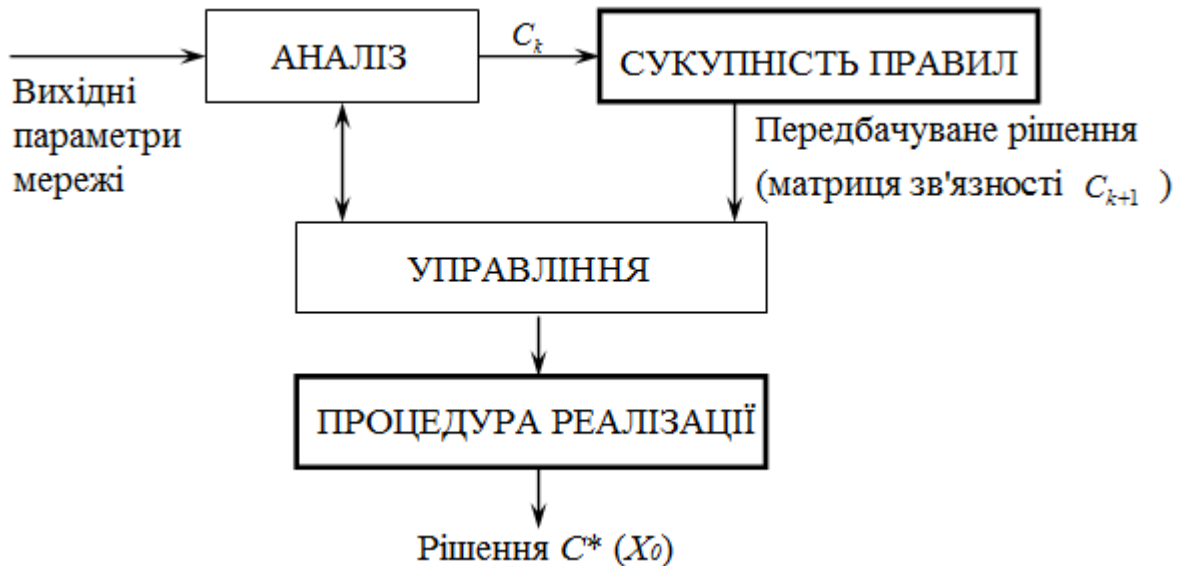


Рис. 24.4. Схема удосконаленого алгоритму пошуку квазіоптимального положення ТА із використанням сукупності правил

Модуль СУКУПНІСТЬ ПРАВИЛ включає в себе правила, що забезпечують напрямлений пошук зміни зв'язності, використовуючи параметри отримані з модуля АНАЛІЗ. Правила продукційного типу складаються з двох компонент: умова та дія [3]. Умова визначає можливість застосування правила у кількісному виразі, а дія використовується для задоволення умови та описує зміну зв'язності мережі, що призводить до системної або користувальницької оптимізації:

$$\text{Пр}_i: \alpha_i \rightarrow \beta_i, \quad (24.3)$$

де  $\text{Пр}_i$  –  $i$ -е правило;  $\alpha_i$  – умова;  $\beta_i$  – управлінська дія. Управлінська дія у даному випадку передбачає розміщення одного (або декількох) БПЛА у певних токах простору. Початкове розміщення може визначатися шляхом центроїдного або решітчастого способу ініціалізації, або ж їх модифікації, в залежності від обраної цілі управління (правила), що буде показано нижче.

Модуль УПРАВЛІННЯ визначає момент закінчення застосування правил та виконує роль інтерпретатора, працюючи циклічно. В кожному циклі він продивляється всі правила, щоб виявити ті, посилки яких

співпадають з відомими на даний момент фактами з пам'яті. Після вибору правило спрацьовує, його висновок вноситься у пам'ять, а потім цикл повторюється спочатку. В одному циклі може спрацювати тільки одне правило. Якщо декілька правил успішно зіставлені з фактами, то інтерпретатор на основі метаправил вибирає одне правило. Метаправила визначають перевагу правил в залежності від  $k$ -ї цілі управління мережею (МПр <sub>$i$</sub> : Пр <sub>$i$</sub>   $\rightarrow$   $w^k_i$ ). Те правило, що в більшій степені покращує потрібні параметри, буде мати більшу вагу  $w^k_i$ .

Ключовою особливістю в запропонованій процедурі є сукупність правил. Усі правила класифіковано на три групи:

- 1) для забезпечення вимог до зв'язності мережі ( $\Omega_1$ );
- 2) для забезпечення вимог до функціональних показників ( $\Omega_2$ );
- 3) для підвищення пропускнуої здатності мережі. Керованим параметром є взаємне розташування вузлів, а саме положення БПЛА відносно мобільних абонентів.

Запропоновані правила аналізують структуру графа мережі та намагаються створити таку структуру мережі, що буде мати найбільшу продуктивність, виявляючи перевантажені ділянки мережі та намагаючись пере направити потоки по маршрутам з меншою завантаженою. Всі правила, незалежно від цілей управління, визначають наступну дію – вивід або зміну позиції БПЛА у просторі, що призводить до зміни кількості вузлів у зоні покриття БПЛА.

Запропоновані правила аналізують структуру графа мережі та намагаються створити таку структуру мережі, що буде мати найбільшу продуктивність, виявляючи перевантажені ділянки мережі та намагаючись пере направити потоки по маршрутам з меншою завантаженою. Всі правила, незалежно від цілей управління, визначають наступну дію – вивід або зміну позиції БПЛА у просторі, що призводить до зміни кількості вузлів у зоні покриття БПЛА.

В результаті експериментальних досліджень структури мережі та застосування теорії графів були розроблені нові модифіковані правила та відповідні управлінські дії (щодо квазіоптимального розміщення БПЛА у просторі). Розглянемо їх детальніше.

**Правило №1.** *Якщо кількість компонент зв'язності (незв'язних підграфів) графу мережі  $k > 1$ , тоді телекомунікаційну аероплатформу (ТА) необхідно розмістити так, щоб з'єднати якомога більшу кількість компонент зв'язності.*

**Мета-правило №1.** *Якщо є декілька варіантів розміщення ТА, що поєднує однакову кількість компонент зв'язності, то обрати той, що покриває більшу кількість вузлів цих компонент.*

Для реалізації цього правила може бути використана процедура із використанням решітчастої ініціалізації (рис. 24.5), що передбачає перебір

варіантів розміщення БПЛА у кожному вузлі решітки та вибір такого що задовольняє правилу або мета-правилу.

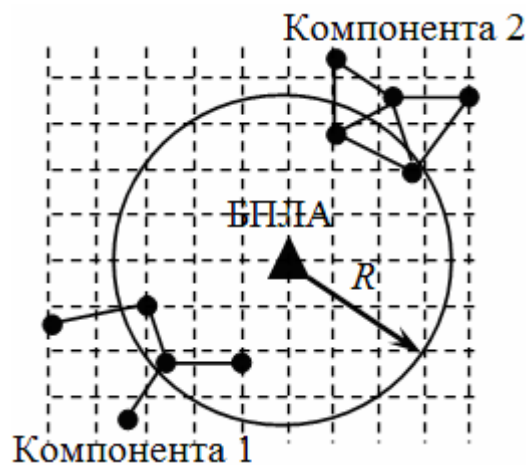


Рис. 24.5. Ілюстрація до реалізації правила №1 із застосуванням решітчастої ініціалізації

**Правило №2.** Якщо середня затримка передачі (кількість ретрансляцій) в деяких маршрутів більше необхідної, тоді ТА необхідно розмістити так, щоб зменшити середню затримку передачі (кількість ретрансляцій) в маршруті.

**Мета-правило №2.** Якщо є декілька варіантів розміщення ТА, що однаково зменшує затримку в маршрутах, то слід обрати той, що дозволяє максимізувати ПЗ маршрутів.

Для того, щоб зменшити середню затримку передачі (кількість ретрансляцій в маршруті) необхідно розмістити БПЛА так, щоб покрити пару заданих вузлів відправник-адресат у кращому випадку або ж пару найбільш віддалених (по кількості ретрансляцій) вузлів маршруту.

Для цього пропонується застосовувати таку просту процедуру, засновану на центроїдній ініціалізації (рис. 24.6):

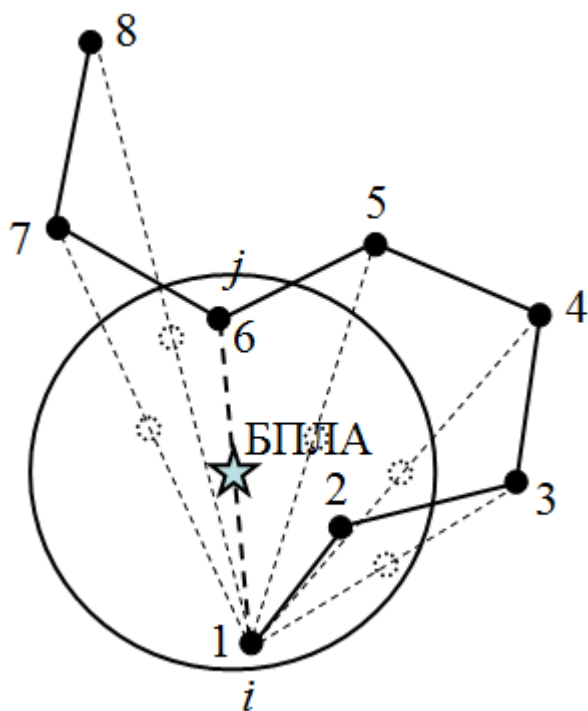


Рис. 24.6. Ілюстрація до реалізації правила №2 із застосуванням центроїдної ініціалізації

1. Визначити «проблемний» маршрут  $m_{ab}$  (маршрути) мережі, для якого не виконується умова  $\bar{t}_3 \leq t_3^0$  ( $l(m_{ab}) \leq l^0$ ), та кількість ретрансляцій  $l^0 = N - 1$ , що входять до нього, де  $N$  – кількість вузлів в маршруті.

2. Побудувати прямі відрізки, що з'єднують заданий вузол  $i$  маршруту (наприклад,  $i=1$ , як показано на рис. 24.6) зі всіма іншими вузлами  $j$  «проблемного» маршруту,  $i, j = \overline{1, N}$ .

3. Перевірити можливість покриття зоною обслуговування БПЛА радіусу  $R$  кожного відрізка  $ij$  (при розміщенні БПЛА в центрі кожного відрізка):  $r_{ij} \leq 2R$ .

4. Для всіх відрізків, для яких виконується попередня умова, розрахувати нову кількість ретрансляцій маршруту із використанням БПЛА по наступній формулі:

$$l_{ij}(m_{ab}) = \begin{cases} |N - j| + |1 - i| + 2, & i < j \\ |N - i| + |1 - j| + 2, & i > j \end{cases}, \quad a, b, i, j = \overline{1, N} \quad (24.4)$$

5. Серед усіх можливих відрізків вибрати той, що при розміщенні у його центрі БПЛА дозволить максимально скоротити кількість ретрансляцій в маршруті.

Обчислювальна складність процедури складатиме  $O(N^2)$ .

При застосуванні наведеної вище процедури БПЛА буде розміщений у центрі відрізка 1-6, що дозволить утворити новий маршрут 1-БПЛА-6-7-8, що має 4 ретрансляції замість 7-ми при відсутності БПЛА (маршрут 1-2-...-7-8).

У випадку коли є декілька рішень (щодо розміщення БПЛА), які дають однакову кількість ретрансляцій в маршруті, можна застосовувати певні метаправила направлені, наприклад, на досягнення більшої пропускної здатності маршруту.

**Правило №3.** Якщо необхідно підвищити пропускну здатність мережі, то ТА необхідно розмістити таким чином, щоб покрити максимальну кількість перевантажених вузлів мережі.

Перевантаженими вузли можуть бути у випадку, коли черга буферу постійно зростає через надвисоку інтенсивність надходження пакетів на обслуговування, та/або у випадку, коли через велике скупчення вузлів зростає кількість колізій при організації МД. У будь-якому випадку будемо вважати вузли перевантаженими, якщо пропускна здатність інцидентних їм каналів не задовольняє вимозі  $s(m_{ab}) \geq 0,5$ . Розміщення БПЛА над таким скупченням перевантажених вузлів (рис. 24.7) дозволить організувати нові маршрути передачі даних між вузлами (при функціонуванні одного з відомих протоколів маршрутизації), що дозволить вирівняти баланс навантаження в мережі та підвищити її пропускну здатність.

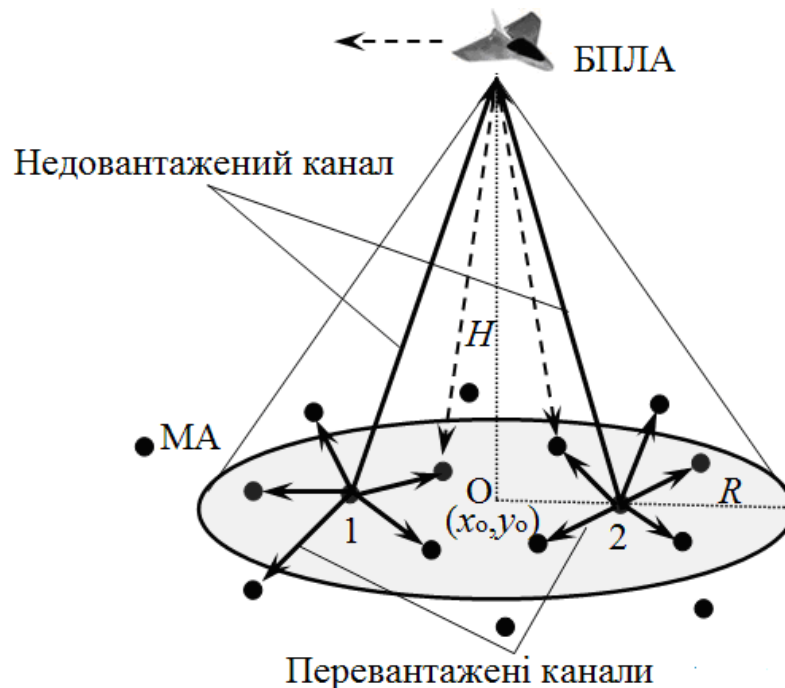


Рис. 24.7. Ілюстрація застосування правила №3 (точка О – центр мас вузлів 1 і 2)

Для реалізації цього правила може бути використана решітчаста ініціалізація у поєднанні з центроїдною:

1. Згідно решітчастої ініціалізації здійснюється пошук максимальної кількості перевантажених вузлів.

2. Згідно центроїдної ініціалізації здійснюється пошук квазіоптимального положення БПЛА в центрі мас перевантажених вузлів. Центр мас групи перевантажених вузлів можна визначити так:

$$x_{0k} = \sum_{i=1}^{N^{nep}} \alpha_i x_i, \quad y_{0k} = \sum_{i=1}^{N^{nep}} \alpha_i y_i, \quad (24.5)$$

де  $\alpha_i = 1/N^{nep}$ ,  $N^{nep}$  - кількість вузлів групи перевантажених вузлів. Також зазначимо, що  $\sum_{i=1}^{N^{nep}} \alpha_i = 1$ .

Але ці правила не враховують ситуації, коли перенавантажений вузол може бути розвантажений за допомогою перенаправлення трафіку на недовантажений вузол по іншому маршруту з використанням БПЛА, тому потребують додаткового правила, яке буде включати в себе таку ситуацію.

**Правило №4.** Якщо вузол  $i$  перевантажений, то знайти найближчий до нього неперевантажений вузол  $j$  в зоні покриття БПЛА (при відсутності такого вузла перемістити БПЛА, щоб забезпечити покриття вузла не порушуючи якість функціонування), перенаправити трафік з перенавантаженого вузла  $i$  до  $j$  (рис.24.8).

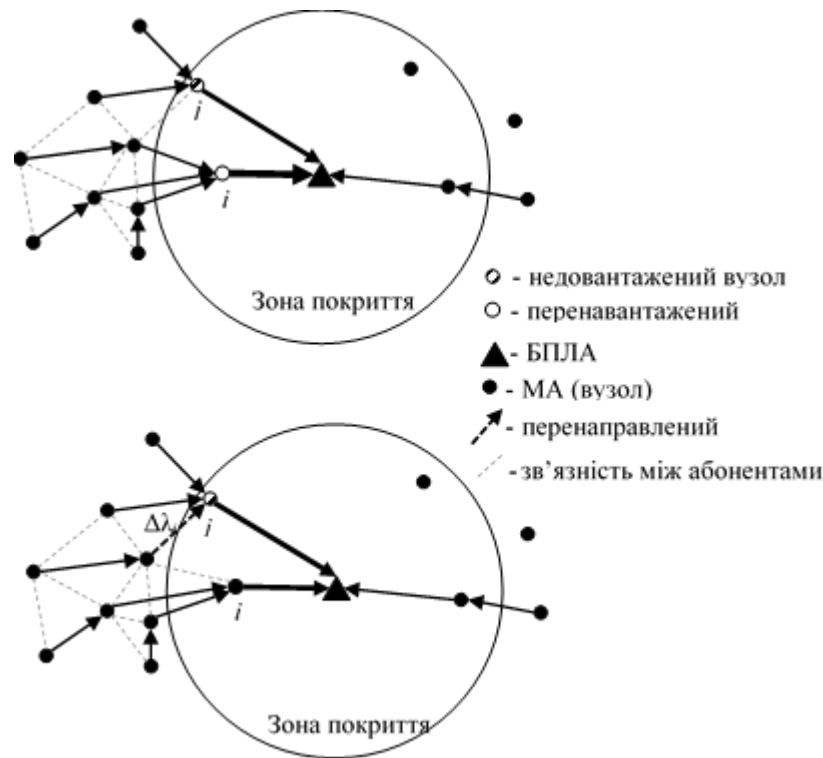


Рис. 24.8. Приклад перенаправлення трафіку між вузлами ЕРМ згідно правила №4.

Для реалізації правила необхідно:

1. Знайти перевантажений вузол  $I \lambda_i^{BX} \gg s_i$
2. На основі матриці зв'язності та матриці найкоротших маршрутів знайти найближчий недовантажений вузол  $j$ ;
3. Якщо вузол  $j$  знаходиться в зоні покриття цього БПЛА, то перехід на п.4, інакше пошук рішення на переміщення БПЛА для покриття вузла  $j$ .
4. Пошук рішення по перенаправленню частини трафіку  $\Delta\lambda_i = \lambda_i^{BX} - s_i$  від вузла  $I$  до  $j$ .

5. Якщо п.4 не приведе до перевантаження вузла  $j$ , рішення по перенаправленню трафіку знайдено, перехід на п.6.

Інакше на цей вузол перенаправляється частина трафіку  $\Delta\lambda_j = s_j - \lambda_j^{BX}$ , здійснюється пошук ще одного нового недовантаженого вузла  $p$ , на який перенаправляється  $\Delta\lambda_p = \Delta\lambda_i - \Delta\lambda_j$ , перехід до п.3.

6. Кінець.

#### 24.4. Оцінка ефективності запропонованого методу

Моделювання функціонування ЕРМ із ТА здійснювалося на базі комп'ютерного середовища Maple. Для цього застосовувалися наступні

вихідні дані:  $N=140$  – кількість МА;  $K=5$  – кількість БПЛА;  $r=10000 \times 10000 \text{ м}^2$  – розмір району дислокації МА;  $d^0=600 \text{ м}$  – максимальна дальність зв'язку МА-МА;  $R'=1500 \text{ м}$  – максимальна дальність зв'язку МА-БПЛА; алгоритм пошуку найкоротших шляхів – Дейкстри;  $V=11 \text{ Мбіт/с}$  – швидкість передачі;  $L=1024$  біт – довжина пакету;  $v_i=15 \text{ м/с}$ ,  $i=\overline{1, N}$  – середня швидкість переміщення МА. Усі інші данні повторюють вихідні дані попередніх розділів.

Для того, щоб показати ефективність запропонованого методу будемо розглядати декілька стратегій управління (СУ) положенням БПЛА:

СУ1: базова стратегія управління (покриття максимальної кількості вузлів).

СУ2: стратегія управління на основі методу повного перебору.

СУ3: стратегія управління на основі запропонованих правил (покриття максимальної кількості перевантажених вузлів).

*Порядок оцінки запропонованого методу наступний:*

1. Задання топології мережі випадковим чином та розміщення БПЛА згідно СУ1, СУ2 та СУ3 (рис. 24.9).

2. Розрахунок пропускної здатності мережі  $S(C)$  для кожної СУ. Визначення виграшу СУ2 та СУ3 відносно СУ1 та достовірності результатів кожної відносно методу повного перебору (рис. 24.10 – рис. 24.12).

3. Розрахунок часу пошуку рішення для запропонованого методу (СУ3) та методу повного перебору (СУ2) та порівняння значень із тривалістю зв'язності МА.

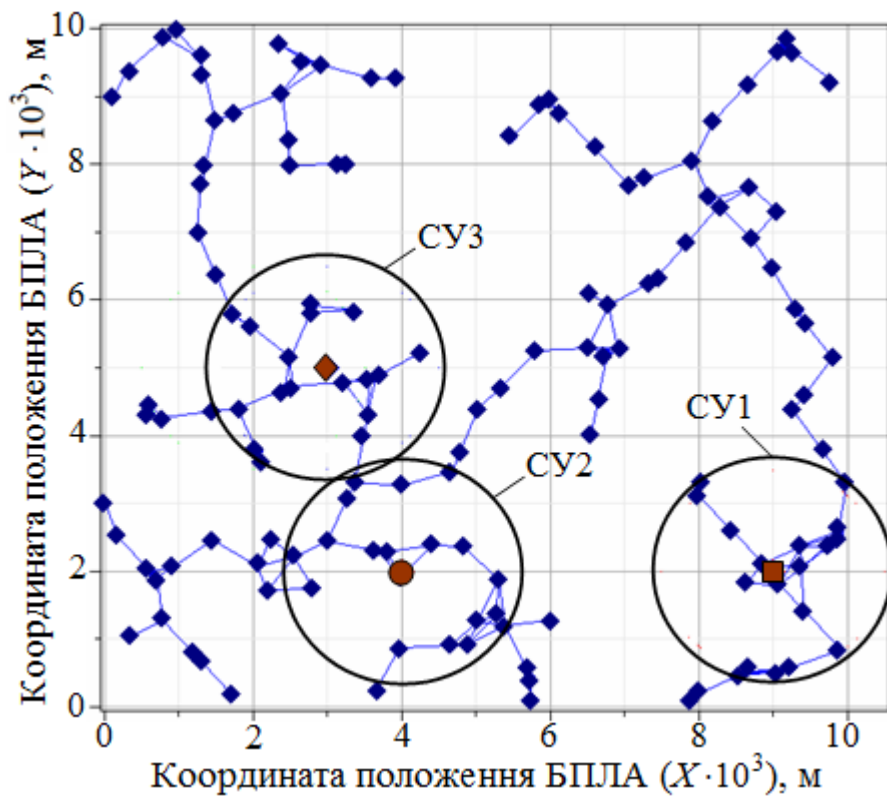


Рис. 24.9. Початкова топологія мережі із розміщенням БПЛА згідно СУ1, СУ2 та СУ3

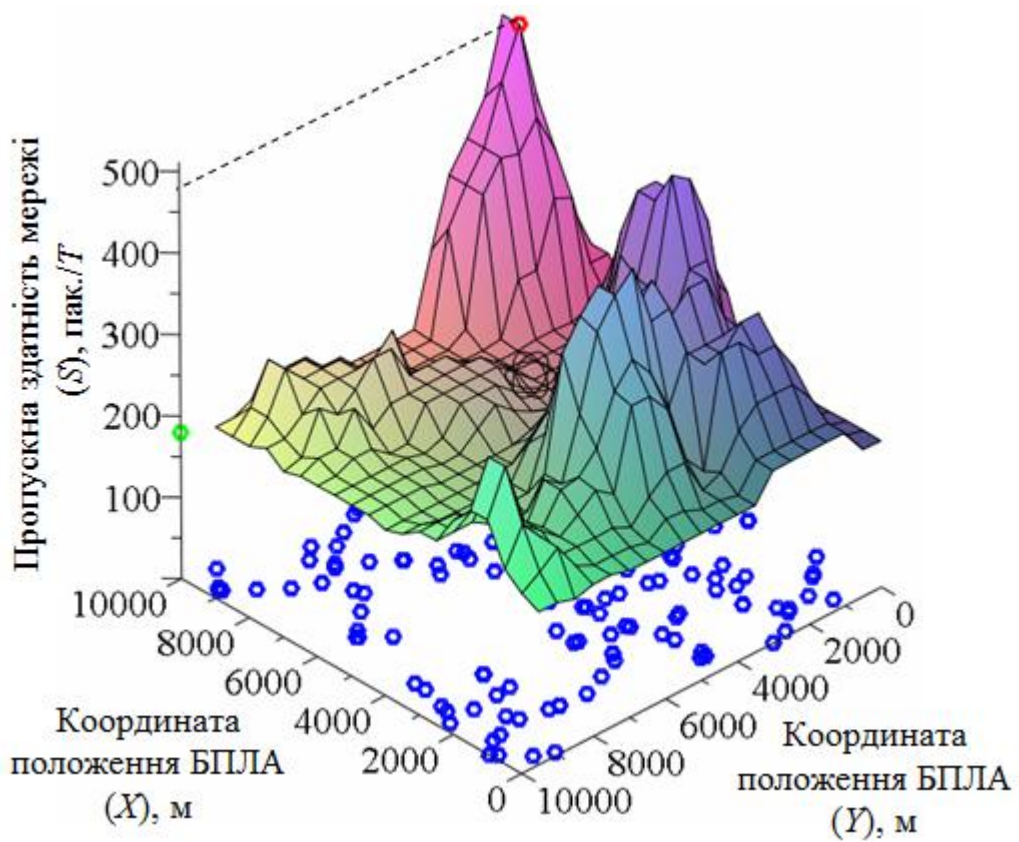


Рис. 24.10. Залежність ПЗ мережі від координат розміщення одного БПЛА (згідно СУ3)

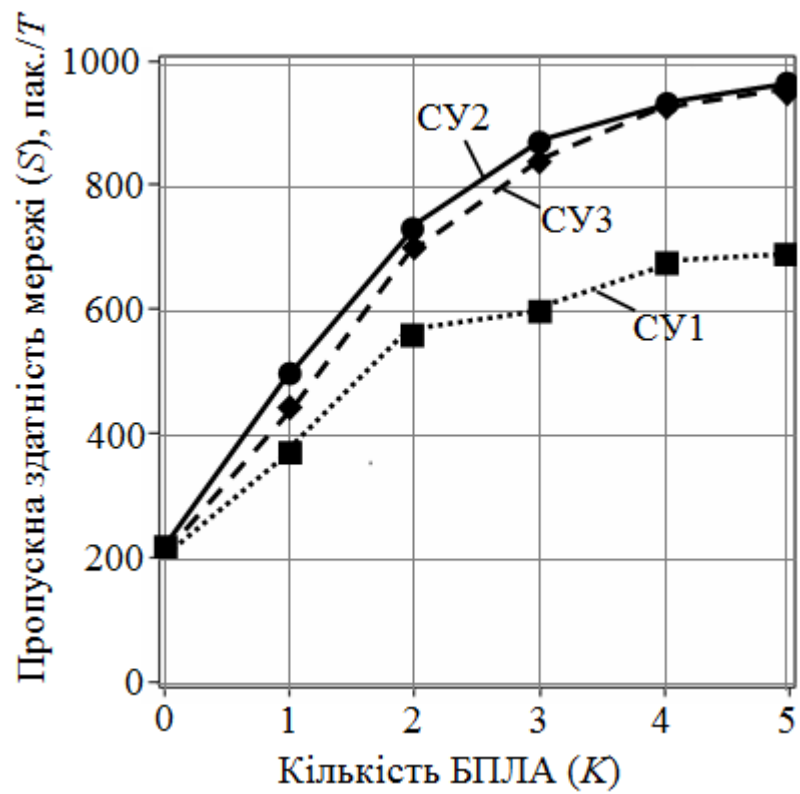


Рис. 24.11. Залежність ПЗ мережі від кількості виведених БПЛА згідно СУ1, СУ2, СУ3

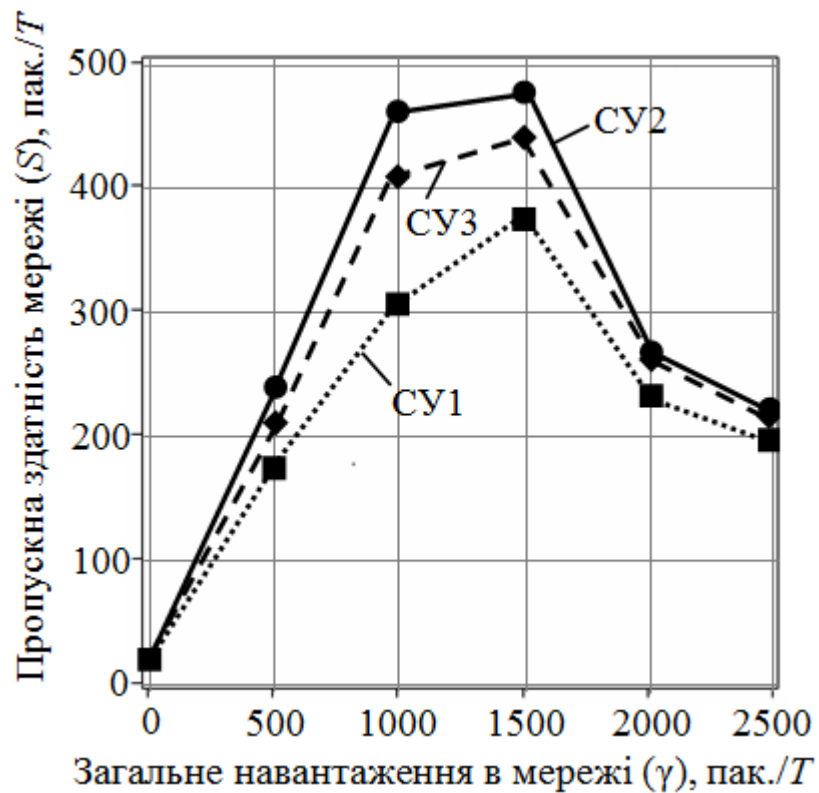


Рис. 24.12. Залежність ПЗ мережі від загального навантаження в мережі при різних стратегіях управління положенням БПЛА

Аналізуючи графіки залежності ПЗ мережі від загального навантаження в мережі при різних стратегіях управління можемо бачити, що в діапазоні трафіку 500...2000 пакетів протягом нормованого часу передачі пакету  $T$  запропонована стратегія управління положенням БПЛА (СУ3) значно переважають базову стратегію (СУ1), а в інших діапазонах усі СУ майже однаково неефективні.

Отже проведена оцінка ефективності запропонованого методу дозволяє зробити наступні висновки:

1. Середній вигравш запропонованого методу (СУ3 відносно СУ1) при випадковій генерації 100 початкових топологій становить 15-20%.

2. Середнє відхилення значення ПЗ (при СУ3) відносно методу повного перебору (СУ2) становить 5-7%.

3. Середній час отримання рішення для СУ3 становить 24с на відміну від десятків хвилин для СУ2, що при середній тривалості зв'язності між МА 348с, дозволяє виконувати управління положенням БПЛА в режимі реального часу.

#### **24.5. Висновки до глави 24**

1. Розроблено метод підвищення пропускної здатності епізодичних радіомереж з управлінням положенням телекомунікаційних аероплатформ. Суть розробленого методу полягає в систематизації математичних моделей оцінки структурної зв'язності та показників якості обслуговування мобільних абонентів, а також удосконаленого алгоритму пошуку квазіоптимального положення ТА, та поєднанні їх в єдину обчислювальну процедуру для досягнення близьких до екстремальних значень пропускної здатності мережі в режимі реального часу.

2. Задача пошуку оптимального розміщення БПЛА (матриці зв'язності  $C_k$ ) відноситься до класу  $NP$ - повних. Застосування для її вирішення класичних методів приводить до експонентної складності. Отримання точного рішення для мережі, яка налічує сотні (тисячі) вузлів потребує значних затрат часу. Тому для скорочення повного перебору варіантів розміщення БПЛА використовувати удосконалений алгоритм пошуку квазіоптимального положення ТА із застосуванням сукупності правил, поєднаних в базу знань, що змінюють зв'язність мережі для підвищення

пропускної здатності мережі та скорочують час обчислень. Це дозволяє отримати в реальному часі близькі до оптимальних рішення та використовувати запропонований алгоритм для оперативного управління положенням БПЛА.

3. Проведена оцінка ефективності запропонованого методу. Показано, що середній виграш запропонованого методу (відносно базової стратегії) становить 15-20%. При цьому середнє відхилення значення пропускної здатності відносно методу повного перебору становить 5-7%. Середній час отримання рішення згідно розробленого методу становить 24с на відміну від десятків хвилин методу повного перебору, що при середній тривалості зв'язності між МА 348с, дозволяє виконувати управління положенням БПЛА в режимі реального часу.

### **Контрольні запитання та завдання**

1. Виконайте математичну постановку задачі підвищення пропускної здатності епізодичних радіомереж з управлінням положенням БПЛА ?
2. Наведіть основні етапи методу підвищення пропускної здатності епізодичних радіомереж з управлінням положенням БПЛА .
3. В чому полягає удосконалення алгоритму пошуку квазіоптимального положення БПЛА ?
4. Як виконати імітаційний експеримент для перевірки ефективності запропонованого методу та (чи) алгоритму ?

## Навчально-методичні матеріали

### Основна література

1. Лисенко О.І., Тачиніна О.М., Алексеєва І. В. «Математичні методи моделювання та оптимізації. Частина 1. Математичне програмування та дослідження операцій: підручник» – К.: НАУ, 2017. – 212 с. ISBN 978-966-932-063-6 .
2. Лисенко О.І., Алексеєва І.В. Дослідження операцій. Конспект лекцій. — К: НТУУ «КПІ», 2016. – 196 с.
3. HANDONG UNITWIN FELLOWSHIP (Republic of Korea). Course [S084-Ukraine] Mathematical programming and operations research in telecommunications, 14 Lectures. Fall 2017. <https://www.hufocw.org/Course/263>
4. Математичні основи теорії телекомунікаційних систем: підручник / В.В. Поповський, С.О. Сабурова, В.Ф. Олійник та ін.: За загальною редакцією В.В. Поповського. – Харків: ТОВ. «Компанія СМІТ», 2006. – 564 с.
5. Ларіонов Ю.І., Левикін В.М., Хажмурадов М.А. Дослідження операцій в інформаційних системах.-Харків.: Компанія СМІТ, 2005.-364 с.
6. Xin-She Yang. Optimization Techniques and Applications with Examples. Hoboken, New Jersey: JohnWiley & Sons, 2018. – 364 p.
7. Романченко І.С., Борисюк С.Л., Потьомкін М.М. Еволюційні методи оптимізації та їх використання у військовій галузі досліджень: монографія. – Ж., ПП “Рута”, 2015. – 127 с. ISBN 978-617-581-234-1
8. Лисенко О.І., Романченко І.С., Чумаченко С.М., Данилюк С.Л., Новіков В.І., Тачинінв О.М., Кірчу П.І., Валуйський С.В. Моделі застосування інформаційно-телекомунікаційних технологій на основі безпілотних авіаційних комплексів у надзвичайних ситуаціях. – К.: НАУ, 2016. – 332 с.
9. Цифрові системи автоматичного керування для телекомунікацій. Комп’ютерний практикум [Електронний ресурс]: навч. посіб. для здобувачів ступеня бакалавра за освіт. програмою «Інженерія та програмування інфокомунікацій» спец. 172 Електронні комунікації та радіотехніка / КПІ ім. Ігоря Сікорського ; уклад.: О. І. Лисенко, О. М. Тачиніна, В. С. Явіся, В. І. Новіков, І. О. Сушин. – Електрон. текст. дані (1 файл: 5,67 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2024. – 204 с. – Назва з екрана. <https://ela.kpi.ua/handle/123456789/67807> (Гриф надано Методичною радою

КПІ ім. Ігоря Сікорського (протокол № 8 від 20.06.2024 р.) за поданням вченої ради Навчально-наукового інституту телекомунікаційних систем (протокол № 5 від 27.05.2024 р.)

10. Досягнення в телекомунікаціях 2019 / за наук. ред. М.Ю.Ільченка, С.О.Кравчука: монографія. - Київ: Інститут обдарованої дитини НАПН України, 2019.- 336 с. Рекомендовано до друку ВР КПІ ім.І.Сікорського (прот.№10 від 04.11.2019 р.) ISBN 978-617-7734-12-2

#### **Додаткова література**

11. Глоба Л.С., Дяденко О.М., Пилипенко А.Ю., Скулиш М.А. Математичні методи аналізу та керування телекомунікаційними мережами. К.: Інститут обдарованої дитини НАПН України, 2017. – 234 с.

12. Глоба Л.С. Розробка інформаційних ресурсів та систем. Підручник у 2 т. – К.: НТУУ „КПІ”, 2014. Т.1. -376 с.

13. Глоба Л.С. Розробка інформаційних ресурсів та систем. Підручник у 2 т. – К.: НТУУ „КПІ”, 2015. Т.2. -376 с.