

Міністерство освіти і науки України  
Донбаська державна машинобудівна академія (ДДМА)

**Л. В. Васильєва, Н. С. Грудкіна**

# **МАТЕМАТИЧНІ МЕТОДИ ДОСЛІДЖЕННЯ ОПЕРАЦІЙ**

**Курс лекцій**

Затверджено  
на засіданні вченої ради ДДМА  
Протокол № 9 від 27.04.2023

Краматорськ  
ДДМА  
2023

**Рецензенти:**

*Гончаров О. А.*, д-р фіз.-мат. наук, професор кафедри прикладної математики та моделювання складних систем Сумського державного університету;

*Чумак О. О.*, канд. пед. наук, доцент кафедри Донбаської національної академії будівництва і архітектури, м. Краматорськ.

**Васильєва, Л. В.**

В 19 Математичні методи дослідження операцій : курс лекцій / Л. В. Васильєва, Н. С. Грудкіна. – Краматорськ : ДДМА, 2023. – 112 с.

ISBN 978-617-7889-40-2.

Курс лекцій призначений для студентів, що починають роботу з математичними методами дослідження операцій. Міститься матеріал з основних розділів курсу «Математичні методи дослідження операцій» з урахуванням специфіки галузі знань 12 «Інформаційні технології». Наведено теоретичні відомості й алгоритми чисельного пошуку безумовного й умовного екстремумів функцій однієї та багатьох змінних; розглянуто методи послідовного, прямого, градієнтного й випадкового пошуків екстремумів функцій, а також розв'язання екстремальних задач методами лінійного та динамічного програмування.

**УДК 519.8 : 004.42**

© Л. В. Васильєва,  
Н. С. Грудкіна, 2023

© ДДМА, 2023

ISBN 978-617-7889-40-2

## ЗМІСТ

ВСТУП .....	6
1 КЛАСИФІКАЦІЯ ПРОЦЕДУР СИНТЕЗУ ТА ПОНЯТТЯ ЕФЕКТИВНОЇ ТОЧКИ Й ОБЛАСТІ КОМПРОМІСІВ .....	7
2 СПОСОБИ ЗГОРТАННЯ ВЕКТОРНОГО КРИТЕРІЮ .....	11
2.1 Частинні критерії .....	11
2.2 Мультиплікативні критерії .....	11
2.3 Адитивні критерії.....	12
2.4 Максимінні критерії .....	13
3 ПОСТАНОВКА ЗАДАЧІ ОПТИМІЗАЦІЇ ТА КЛАСИФІКАЦІЯ МЕТОДІВ ПОШУКУ ЕКСТРЕМУМУ .....	15
3.1 Постановка задачі оптимізації.....	15
3.2 Необхідна й достатня умова екстремуму .....	17
3.3 Класифікація методів пошуку екстремуму .....	19
4 МЕТОДИ ОДНОМІРНОЇ ОПТИМІЗАЦІЇ .....	21
4.1 Метод послідовного рівномірного перебирання (сканування).....	21
4.2 Метод дихотомії (половинного розподілу) .....	22
4.3 Метод золотого перетину.....	23
4.4 Метод квадратичної інтерполяції .....	25
5 МЕТОДИ БАГАТОМІРНОЇ ОПТИМІЗАЦІЇ .....	27
6 ГРАДІЄНТНІ МЕТОДИ.....	30
6.1 Градієнтні методи зі змінним кроком.....	30
6.2 Метод найшвидшого спуску.....	31
6.3 Особливості пошуку оптимуму для цільових функцій із яружним рельєфом поверхні відгуку .....	33
6.4 Обчислення градієнта цільової функції .....	35
6.5 Метод спряжених градієнтів (метод Флетчера-Рівса) .....	35
7 МЕТОДИ БЕЗУМОВНОЇ ОПТИМІЗАЦІЇ НУЛЬОВОГО ПОРЯДКУ .....	38
7.1 Метод покоординатного спуску (метод Гауса-Зейделя).....	38
7.2 Метод прямого пошуку (метод Хука-Дживса) .....	39

7.3 Симплексний метод .....	40
7.4 Метод Розенброка.....	43
7.5 Метод паралельних дотичних (метод Пауелла) .....	45
8 МЕТОДИ ВИПАДКОВОГО ПОШУКУ.....	47
8.1 Метод випадкового пошуку з перерахунком.....	47
8.2 Метод випадкового пошуку за найкращим випробуванням (пошук за статистичним градієнтом).....	48
9 МЕТОДИ БЕЗУМОВНОЇ ОПТИМІЗАЦІЇ ДРУГОГО ПОРЯДКУ .....	50
9.1 Метод Ньютона.....	50
9.2 Метод Ньютона–Рафсона .....	51
9.3 Обчислення матриці Гессе.....	52
9.4 Методи безумовної оптимізації з перемінною метрикою .....	53
10 МЕТОДИ УМОВНОЇ ОПТИМІЗАЦІЇ.....	57
10.1 Класифікація методів .....	57
10.2 Методи внутрішніх штрафних функцій .....	58
10.3 Методи зовнішніх штрафних функцій .....	59
10.4 Метод Фіакко й Мак-Корміка .....	60
10.5 Метод проєкції градієнту.....	61
10.6 Комплексний пошук Бокса .....	63
10.7 Метод Моррісона.....	65
10.8 Особливості пошуку оптимуму при використанні максимінних критеріїв .....	66
11 ЛІНІЙНЕ ПРОГРАМУВАННЯ.....	69
11.1 Постановка задачі лінійного програмування.....	69
11.2 Властивості розв’язків задачі ЛП .....	71
11.3 Графічний метод розв’язання задачі ЛП.....	72
11.4 Симплексний метод розв’язання задачі ЛП .....	76
11.5 Метод побудови опорних планів .....	77
11.6 Знаходження оптимального розв’язку .....	79
11.7 Алгоритм симплексного методу .....	80
11.8 Симплексний метод із штучним базисом .....	84
12 ТРАНСПОРТНА ЗАДАЧА ЛП .....	89
12.1 Постановка транспортної задачі .....	89

12.2 Побудова вихідного опорного плану.....	91
12.3 Метод потенціалів .....	92
13 МЕТОДИ ДИСКРЕТНОЇ ОПТИМІЗАЦІЇ.....	94
13.1 Постановка задачі дискретної оптимізації.....	94
13.2 Метод Гоморі .....	95
13.3 Метод гілок і меж .....	96
14 ДРОБОВО-ЛІНІЙНЕ ПРОГРАМУВАННЯ .....	98
14.1 Постановка задачі і властивості її розв'язку.....	98
14.2 Графічний метод розв'язання задачі ДрЛП .....	99
15 ДИНАМІЧНЕ ПРОГРАМУВАННЯ .....	104
15.1 Сутність динамічного програмування.....	104
15.2 Алгоритм методу динамічного програмування.....	105
15.3 Використання методу ДП у задачі про вибір траєкторії .....	106
СПИСОК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ.....	110

## ВСТУП

Дослідження операцій – теорія використання чисельних методів аналізу у процесі прийняття рішень в усіх галузях цілеспрямованої діяльності. Основний метод теорії дослідження операцій – метод математичного моделювання.

Лише окремі задачі ДО можна розв'язати аналітично та порівняно невелику кількість – чисельними методами «вручну». Основна частина цих задач вимагає використання обчислювальної техніки.

Спеціалісти в галузі ІТ повинні володіти сучасним математичним апаратом і засобами обчислювальної техніки для забезпечення навичок отримання кваліфікованого розв'язання задач оптимального проектування об'єктів і систем й оптимального керування ними. Вибір методу оптимізації для конкретної задачі значною мірою визначається характером відповідної математичної моделі об'єкта чи процесу, видом критерію ефективності (лінійний або нелінійний щодо параметрів, які оптимізуються) й обмежень (типу рівностей або нерівностей), кількістю параметрів та інших чинників. У кожному конкретному випадку можна з багатьох методів вибрати найбільш раціональний для даного класу задач.

Цей курс лекцій укладено для студентів закладів вищої освіти галузі знань 12 «Інформаційні технології» і містить теми із лінійного, нелінійного, динамічного програмування та чисельних методів оптимізації. Матеріали супроводжуються побудовою та алгоритмом розв'язування основних видів класичних задач із дослідження операцій, у кінці кожної теми наведені питання для самоперевірки.

Сподіваємось, що курс лекцій буде корисним для усіх, хто вивчає основи математичного моделювання виробничих, економічних та інших видів процесів людської діяльності та можливі шляхи розв'язання складних моделей із застосуванням комп'ютерних розрахунків.

# 1 КЛАСИФІКАЦІЯ ПРОЦЕДУР СИНТЕЗУ ТА ПОНЯТТЯ ЕФЕКТИВНОЇ ТОЧКИ Й ОБЛАСТІ КОМПРОМІСІВ

Проектування – це процес одержання описів, достатніх для виготовлення нового технічного об'єкта в заданих умовах. Процес проектування ділиться на стадії й етапи. Етап об'єднує виконання проектних процедур із створення описів, що відносять до одного аспекту або ієрархічного рівня.

При виконанні проектних процедур вирішуються задачі синтезу й аналізу описів. При розв'язанні задач синтезу визначається склад елементів і засіб їхнього зв'язку між собою, а при розв'язанні задач аналізу оцінюються властивості синтезованої структури.

Класифікація типових проектних процедур синтезу подана на рис. 1.1. Визначення структури об'єкта зводиться до знаходження переліку типів елементів, що складають об'єкт, і засобів зв'язку елементів між собою. Якщо серед варіантів структури об'єкта ведуться пошуки не будь-якого прийняттого варіанта, а найкращого у деякому сенсі, то таку задачу синтезу називають структурною оптимізацією.

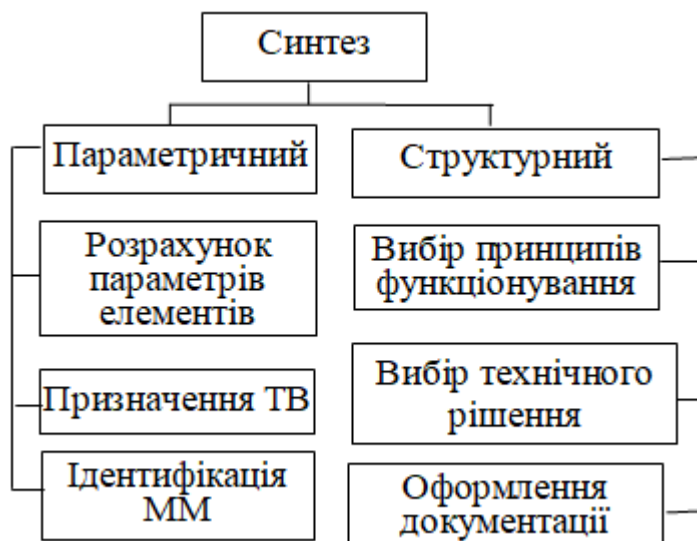


Рисунок 1.1 – Класифікація типових проектних процедур синтезу

Параметричний синтез полягає у визначенні числових значень параметрів елементів за заданих структурі й умовах працездатності, що впливають на вихідні параметри об'єкта.

Розрахунок же внутрішніх параметрів, оптимальних із позиції деякого критерію, при заданій структурі об'єкта, називають параметричною оптимізацією. Класифікація процедур параметричного синтезу показана на рис. 1.2.

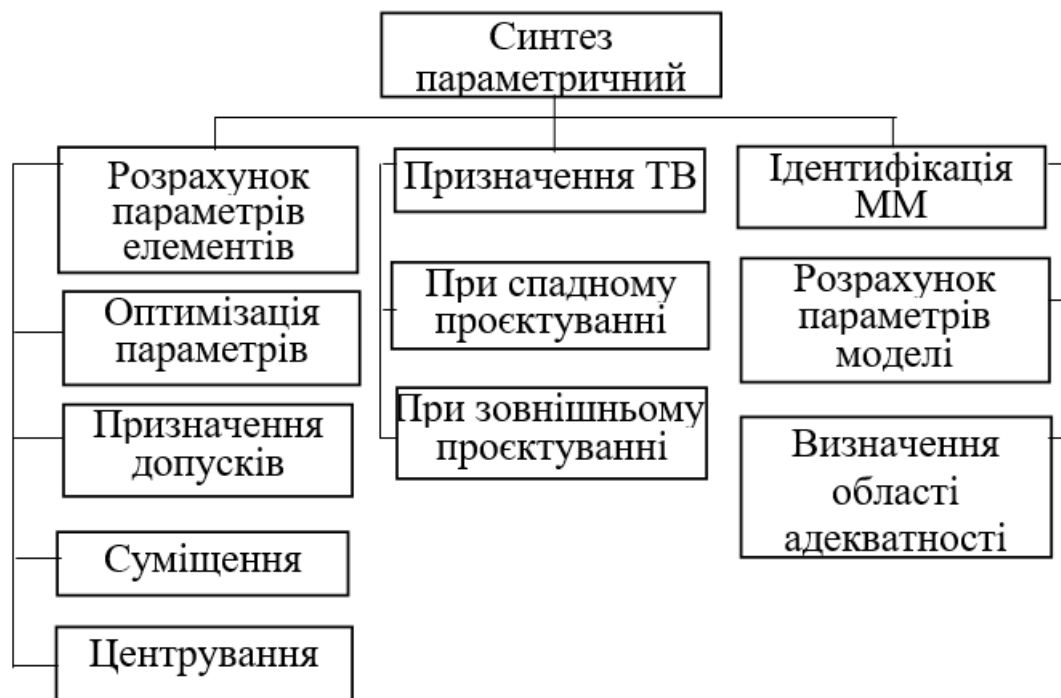


Рисунок 1.2 – Класифікація процедур параметричного синтезу

### *Поняття ефективної точки й області компромісів*

Під оптимізацією параметрів розуміють визначення таких номінальних значень внутрішніх параметрів  $X$  об'єкта проєктування, за яких функція  $f(X)$ , названа цільовою функцією, функцією якості, критерієм оптимальності, набуває екстремального (максимального або мінімального) значення.

Основна проблема постановки екстремальних задач полягає у формулюванні цільової функції. Це пов'язано з тим, що будь-який технічний об'єкт спочатку має векторний характер критеріїв оптимальності (багатокритеріальність), а поліпшення одного з вихідних параметрів, як правило,



призводить до погіршення іншого, тому що усі вихідні параметри є функціями тих самих керованих параметрів і не можуть змінюватися незалежно один від одного. Такі вихідні параметри називають конфліктними. Тому при оптимізації зазвичай неможливо домогтися одночасного поліпшення усіх вихідних параметрів.

Розв'язок задачі багатокритеріальної оптимізації (компромісний розв'язок)  $X^* \in D_X$  є ефективною точкою, якщо для неї справедлива нерівність:

$$Y(X^*) \leq Y(X), X \in D_X,$$

тобто будь-яка компонента

$$y_j(X^*) \leq y_j(X), j = \overline{1, m},$$

але хоча б для одного  $k$  знайдеться точка  $X \in D_X$ , у якій виконується стійка нерівність:

$$y_k(X^*) < y_k(X).$$

Множина всіх ефективних точок називається областю компромісів, або областю розв'язків, оптимальних за Парето.

Для визначення мінімуму за Парето необхідно перейти від задачі векторної оптимізації до задачі нелінійної оптимізації зі спеціально сконструйованою скалярною функцією цілі:

$$f(X) = f(y_1(X), y_2(X), \dots, y_m(X)),$$

тобто перейти до побудови комплексного критерію, при якому цільова функція тим або іншим засобом об'єднує вихідні параметри. Процес утворення глобального критерію  $f(X)$  для задачі багатокритеріальної оптимізації називається об'єднанням або згортанням векторного критерію.

У залежності від того, яким способом обираються й об'єднуються вихідні параметри в скалярну функцію цілі, розрізняють частинні, мультиплікативні, адитивні й інші критерії.

### **Питання для самоперевірки**

1. Опишіть етапи та стадії процесу проектування.
2. Вказати, що розуміють під структурною оптимізацією.
3. Наведіть класифікацію процедур параметричного синтезу.

4. Сформулюйте поняття ефективної точки й області компромісів.
5. Наведіть класифікацію типових проєктних процедур синтезу.
6. Вкажіть основні проблеми постановки екстремальних задач.
7. Розшифруйте поняття області компромісів.
8. Сформулюйте поняття ефективної точки.
9. Визначте поняття згортання векторного критерію.

## 2 СПОСОБИ ЗГОРТАННЯ ВЕКТОРНОГО КРИТЕРІЮ

### 2.1 Частинні критерії

Частинні критерії можуть застосовуватися у випадках, коли серед вихідних параметрів  $y_j(X)$  можна виділити один основний параметр  $y_k(X)$ , який найбільш повно визначає ефективність проєктованого об'єкта. Цей параметр приймається за цільову функцію  $f(X) = y_k(X)$ .

Тоді задача оптимізації стає типовою задачею нелінійного програмування.

Знайти:

$$\underset{X \in D_X}{\text{extremum}} f(X), \quad (2.1)$$

де  $D_X$  – область, що задається прямими обмеженнями на керовані параметри  $X_{min} \leq X \leq X_{max}$  й умовами працездатності, що можуть мати вигляд нерівностей  $y_j \leq TT_j$ ,  $y_j \geq TT_j$  і рівностей  $y_j = TT_j$ .

### 2.2 Мультиплікативні критерії

Розділимо вихідні параметри об'єкта на три групи за типом відповідних їм умов працездатності.

До першої групи віднесемо параметри  $y_j^+$ , що мають умову працездатності вигляду  $y_j^+ \geq TT_j$ , тобто ті параметри, для яких бажано максимальне збільшення.

До другої групи віднесемо параметри  $y_j^-$  з умовами працездатності виду  $y_j^- \leq TT_j$ . Для цих параметрів бажана мінімізація.

Третя група буде утворена параметрами  $y_j^{\bar{}}$  з умовами працездатності типу рівностей  $y_j^{\bar{}} = TT_j \pm \Delta y_j$ , де  $\Delta y_j$  – максимально припустиме відхилення  $y_j^{\bar{}}$  від  $TT_j$ .

Мультиплікативні критерії можуть застосовуватися в тих випадках, коли відсутні умови працездатності типу рівності і вихідні параметри не можуть набувати нульових значень. Тоді мультиплікативна цільова функція, яка мінімізується, буде мати вигляд:

$$f(X) = \prod_{j=1}^q y_j^-(X) / \prod_{j=q+1}^m y_j^+(X). \quad (2.2)$$

Цільова функція, що підлягає максимізації, має вигляд:

$$f(X) = \prod_{j=q+1}^m y_j^+(X) / \prod_{j=1}^q y_j^-(X). \quad (2.3)$$

### 2.3 Адитивні критерії

У адитивних критеріях цільова функція утворюється шляхом додавання перетворених вихідних параметрів. Перетворення здійснюється за допомогою введення нормувальних множників – вагових коефіцієнтів.

Для випадку мінімізації цільової функції згорання векторного критерію буде мати вигляд:

$$f(X) = \sum_{j=1}^q \omega_j y_j^-(X) - \sum_{j=q+1}^m \omega_j y_j^+(X), \quad (2.4)$$

де  $\omega_j$  – ваговий коефіцієнт, що визначає ступінь важливості  $j$ -ого вихідного параметра ( $\omega_j$  вибираються проєктувальником і в процесі оптимізації залишаються незмінними).

Якщо помножити  $y_j^+$  і відповідні їм технічні вимоги  $TT_j$  на  $(-1)$ , тобто привести всі обмеження до типу  $y_j \leq TT_j$ , то цільову функцію, яка мінімізується, можна записати у вигляді:

$$f(X) = \sum_{j=1}^m \omega_j y_j(X). \quad (2.5)$$

За відсутності інформації про важливість частинних критеріїв передбачається їхня рівноцінність. Тоді в якості цільової функції береться сума відносних відхилень частинних критеріїв від їхніх оптимальних значень:

$$f(X) = \sum_{j=1}^m \frac{y_j(X) - y_j(X^*)}{|y_j(X^*)|}. \quad (2.6)$$

Якщо всі або основні умови працездатності мають вид рівностей, то цільова функція, що мінімізується, може приймати вид

$$f(X) = \sum_{j=1}^m \omega_j (y_j(X) - TT_j)^2. \quad (2.7)$$

## 2.4 Максимінні критерії

Ці критерії дозволяють досягти варіанта найкращого забезпечення умов працездатності. Введемо кількісну оцінку ступені виконання  $j$ -ої умови працездатності, позначимо її через  $z_j$  і будемо називати запасом працездатності параметра  $y_j$ . Розрахунок запасу можна виконати різноманітними способами, наприклад,

$$z_j = \omega_j \left( \frac{TT_j - y_{j\text{ном}}}{\delta_j} - 1 \right), \quad (2.8)$$

де  $\omega_j$  – ваговий коефіцієнт (важливість, корисність);

$y_{j\text{ном}}$  – оцінка математичного очікування  $j$ -го вихідного параметра;

$\delta_j$  – величина, що характеризує розсіювання  $j$ -го вихідного параметра, ( $\delta_j \approx 3\sigma_j$ ).

Тут передбачається, що всі умови працездатності зведені до вигляду  $y_j \leq TT_j$ .

Якість функціонування технічної системи буде характеризуватися вектором вихідних параметрів  $Z = (z_1, z_2, \dots, z_m)$ . Цілі розрахунку збігаються з цілями збільшення запасів працездатності  $z_j$ , причому у першу чергу тих із  $z_j$ , що є найменшими. Звідси приходимо до цільової функції:

$$f(X) = \min_{1 \leq j \leq m} z_j(X), \quad (2.9)$$

яка називається функцією мінімуму.

Оскільки потрібна її максимізація, тобто треба знайти:

$$\max_{X \in D_X} f(X) = \max_{X \in D_X} \min_{1 \leq j \leq m} z_j(X), \quad (2.10)$$

то критерій оптимізації з цільовою функцією мінімуму називається максимінним критерієм.

### **Питання для самоперевірки**

1. Укажіть основні способи згортання векторного критерію.
2. Укажіть форму цільової функції у разі застосування частинних критеріїв.
3. Наведіть основні групи мультиплікативних критеріїв.
4. Наведіть побудову цільової функції при використанні адитивних критеріїв.
5. Наведіть можливі форми розрахунку запасу параметрів при застосуванні максимінних критеріїв.
6. Укажіть вигляд цільової функції для максимінних критеріїв.
7. Укажіть означення максимінного критерію.

## 3 ПОСТАНОВКА ЗАДАЧІ ОПТИМІЗАЦІЇ ТА КЛАСИФІКАЦІЯ МЕТОДІВ ПОШУКУ ЕКСТРЕМУМУ

### 3.1 Постановка задачі оптимізації

Сукупність параметрів об'єкта або системи, при яких забезпечується екстремум цільової функції (критерію оптимальності), називаються оптимальними параметрами.

Аргументами цієї функції є керовані параметри. Будемо позначати вектор керованих параметрів через  $X = (x_1, x_2, \dots, x_n)$ , цільову функцію через  $f(X)$ , а область її визначення через  $D_X$ .

Функція  $f(X)$  досягає локального мінімуму (максимуму) у точці  $X^*$ , якщо для всіх точок області  $D_X$ , що лежать в нескінченно малому околі точки  $X^*$ , має місце нерівність:

$$f(X^*) \leq f(X), \quad (3.1)$$

або

$$f(X^*) \geq f(X) \text{ відповідно.}$$

Глобальним екстремумом називається точка  $X^*$ , у якій цільова функція, що мінімізується (максимізується), має найменше (найбільше) значення серед усіх локальних екстремумів, тобто нерівність (3.1) справедлива для всіх  $X \in D_X$ .

Функція  $f(X)$  називається однокстремальною (унімодалною), якщо вона має один екстремум, і багатокстремальною, якщо вона має більше одного мінімуму (максимуму).

Задача знаходження максимуму будь-якої функції  $f(X)$  легко зводиться до знаходження мінімуму заміною функції  $f(X)$  на  $-f(X)$  у зв'язку з виконанням тотожності  $\max f(X) = -\min (-f(X))$ .

Тому надалі будемо говорити про мінімізацію функції  $f(X)$  і задачу знаходження значень оптимальних параметрів  $X^*$  будемо записувати так:

$$\min_{X \in D_X} f(X) = f(X^*) = f^*. \quad (3.2)$$

Якщо екстремум шукають у необмеженій області  $D_X$ , то його називають безумовним, а методи пошуку – методами безумовної оптимізації.

У задачах оптимізації часто присутні обмеження, що зменшують область  $D_X$ , і шуканий екстремум стає умовним. Розрізняють обмеження прямі й функціональні.

Прямі обмеження мають вигляд:

$$x_{i \min} \leq x_i \leq x_{i \max}, \quad i = \overline{1, n}, \quad (3.3)$$

або скорочено

$$X_{\min} \leq X \leq X_{\max}.$$

Наприклад, для багатьох об'єктів параметри елементів не можуть бути від'ємними (геометричні розміри, вага, концентрація домішок, тощо).

Функціональні обмеження являють собою умови працездатності вихідних параметрів, що не ввійшли до цільової функції, і мають вигляд нерівностей або рівностей

$$G(X) \geq 0, \quad H(X) = 0, \quad (3.4)$$

де  $G(X)$ ,  $H(X)$  – вектор-функції вигляду

$$G(X) = (g_1(X), g_2(X), \dots, g_l(X));$$

$$H(X) = (h_1(X), h_2(X), \dots, h_m(X)).$$

Прямі обмеження можна розглядати як окремий випадок функціональних.

Обмеження (3.5) формують допустиму область пошуку – область працездатності:

$$D_X = \{X / G(X) \geq 0, \quad H(X) = 0\}. \quad (3.5)$$

Будь-яка точка  $X \in D_X$  є допустимим розв'язком задачі оптимізації або планом. Допустимий розв'язок, що оптимізує цільову функцію, називається оптимальним розв'язком або оптимальним планом.

Отже, задачу параметричної оптимізації можна сформулювати в такий спосіб.

Знайти:

$$\min_{X \in D_X} f(X), \quad (3.6)$$

де  $D_X = \{X / g_j(X) \geq 0; j = \overline{1, l}, h_j(X) = 0, j = \overline{1, m}\}$ .



### 3.2 Необхідна й достатня умова екстремуму

Нехай функція  $f(X)$  задана аналітично і не менше ніж двічі диференційована за змінними  $x_i$ .

Для того щоб точка  $X^*$  була точкою безумовного локального екстремуму функції  $f(X)$ , необхідно, щоб усі її частинні похідні за елементами  $x_i$  у точці  $X^*$  дорівнювали нулю, тобто

$$\left. \frac{\partial f(X)}{\partial x_i} \right|_{X = X^*} = 0, \quad i = \overline{1, n}. \quad (3.7)$$

Необхідна умова екстремуму (3.7) називається умовою стаціонарності, її можна записати так:

$$\text{grad } f(X^*) = \nabla f(X^*) = 0, \quad (3.8)$$

де  $\nabla f(X)$  – це градієнт функції в точці  $X$ , тобто  $n$ -мірний вектор, компоненти якого дорівнюють частинним похідним функції  $f(X)$  за змінними  $x_i$ .

Характер екстремуму функції в стаціонарній точці визначається поведінням  $f(X)$  навколо  $X^*$ : чи є вона опуклою або увігнутою (рис. 3.1, а).

Нагадаємо, що множина  $D$  є опуклою, якщо для будь-яких двох точок  $X_1, X_2 \in D$  усі точки:

$$\lambda X_1 + (1 - \lambda) X_2 \in D, \quad \text{де } 0 \leq \lambda \leq 1.$$

Якщо  $f(\lambda X_1 + (1 - \lambda) X_2) \geq \lambda f(X_1) + (1 - \lambda) f(X_2)$ , то функція називається увігнутою (опуклою вгору). Геометрично це зображено на рис. 3.1, б.

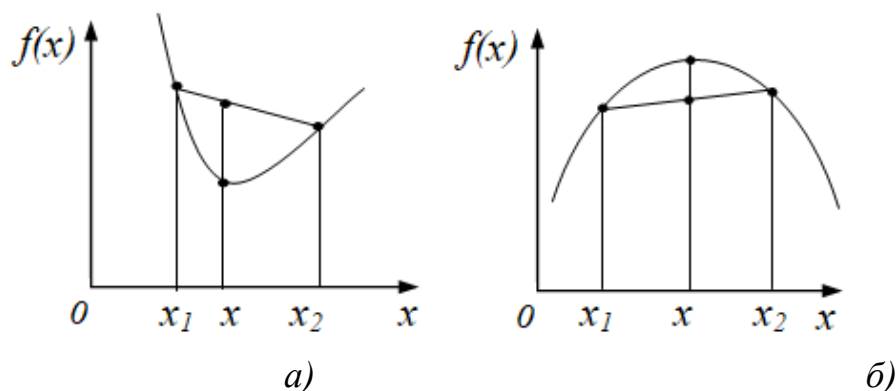


Рисунок 3.1 – Характер екстремуму функції в стаціонарній точці

Критерій опуклості й увігнутості функції  $n$  змінних демонструє наступна **теорема**.

**Теорема.** Диференційована функція  $f(X)$  строго увігнута навколо точки  $X_0$ , якщо виконані такі умови:

$$f_{11}(X_0) < 0, \quad \begin{vmatrix} f_{11}(X_0) & f_{12}(X_0) \\ f_{21}(X_0) & f_{22}(X_0) \end{vmatrix} > 0,$$

$$\begin{vmatrix} f_{11}(X_0) & f_{12}(X_0) & f_{13}(X_0) \\ f_{21}(X_0) & f_{22}(X_0) & f_{23}(X_0) \\ f_{31}(X_0) & f_{32}(X_0) & f_{33}(X_0) \end{vmatrix} < 0 \quad \text{і т. д.},$$

тобто знаки визначників чергуються, де

$$f_{ij}(X_0) = \frac{\partial^2 f(X_0)}{\partial x_i \partial x_j}.$$

Якщо усі визначники додатні, то функція  $f(X)$  строго опукла навколо точки  $X_0$ .

Таким чином, для того щоб у точці  $X^*$  досягався безумовний локальний мінімум (максимум), необхідно і достатньо рівності нулю всіх перших частинних похідних і строгої опуклості (увігнутості) функції навколо точки  $X^*$ .

На жаль, відсутні ознаки багатоекстремальних ситуацій, що не викликають труднощів при перевірці. Маючи координати екстремальної точки, не можна зробити ніяких висновків про локальний або глобальний характер екстремуму, не досліджуючи  $f(X)$  на усій області  $D_X$ . Винятком є задачі, де цільова функція є опуклою або увігнутою. Вона має один локальний мінімум (максимум), який і є глобальним, тобто опуклість або увігнутість функції є достатньою ознакою її унімодальності.

### 3.3 Класифікація методів пошуку екстремуму

Застосування методів диференціального числення для задач оптимізації не завжди дає позитивний результат. Якщо система рівнянь (3.7) є нелінійною, то розв'язати її аналітичним шляхом можна вкрай рідко. Потрібні чисельні методи. Крім того, часто цільова функція  $f(X)$  і обмеження  $G(X)$ ,  $H(X)$  визначаються неявно, тобто алгоритмічно. Тому потрібні методи пошукової оптимізації, у яких процедура пошуку оптимального розв'язку  $X^*$  пов'язана з проведенням проб у точках  $X_k$ ,  $k = 0, 1, 2, \dots$ . При цьому наступна точка визначається за допомогою рекурентних співвідношень:

$$X_{k+1} = R(X_k, f(X_k)), k = 0, 1, \dots \quad (3.9)$$

де  $k$  – номер ітерації;

$R$  – деякий оператор, що визначає метод пошуку.

**У залежності від типу шуканого екстремуму** розрізняють методи локальної і глобальної, умовної і безумовної оптимізації.

Методи, що дозволяють визначити екстремум тільки унімодальних функцій, називаються методами локальної оптимізації. Методи, що дозволяють визначити оптимальну точку багатоекстремальної функції, називаються методами глобальної оптимізації (наприклад, метод сканування). Методи, що дозволяють визначити оптимальну точку в умовах відсутності обмежень, називаються методами безумовної оптимізації, за наявності обмежень – методами умовної оптимізації.

Більшість методів пошуку відносять до методів безумовної оптимізації, але існують засоби зведення задач умовної оптимізації до безумовної.

**У залежності від порядку використовуваних похідних** цільової функції розрізняють методи безумовної оптимізації нульового, першого і другого порядків.

У методах нульового порядку (прямих методах) використовується інформація про значення функції. У методах першого порядку (градієнтних методах) використовується інформація про значення функції і її перших похідних. У методах другого порядку (методах Ньютонів) організація

пошуку екстремуму ведеться з урахуванням значень цільової функції, її перших і других похідних.

**У залежності від кількості параметрів** цільової функції розрізняють методи одномірного і багатомірного пошуку.

### **Питання для самоперевірки**

1. Сформулюйте поняття глобального екстремуму функції.
2. Наведіть поняття умовного та безумовного екстремуму функції.
3. Сформулюйте необхідну й достатню умови екстремуму.
4. Сформулюйте критерій опуклості й увігнутості функції  $n$  змінних.
5. Наведіть, за яким параметрами проводять класифікацію методів пошуку екстремуму.
6. Укажіть, які методи розрізняють в залежності від типу шуканого екстремуму.
7. Укажіть, яка інформація використовується у методах першого порядку (відносно похідних).
8. Укажіть, яким чином організовано пошук екстремуму при застосуванні методу Ньютона.

## 4 МЕТОДИ ОДНОМІРНОЇ ОПТИМІЗАЦІЇ

### 4.1 Метод послідовного рівномірного перебирання (сканування)

Для знаходження глобального екстремуму функції  $f(x)$  на відрізку  $[a, b]$  обирається крок  $h$  зміни аргументу  $x$  (рис. 4.1). Послідовно обчислюються значення  $y_i = f(x_i)$  у точках

$$x_1 = a, x_2 = x_1 + h, \dots, x_i = x_{i-1} + h, \dots, x_n = x_{n-1} + h,$$

де  $n = \left\lceil \frac{b-a}{h} \right\rceil + 1$ .

Крок  $h$  буде визначати точність визначення  $x^*$ .

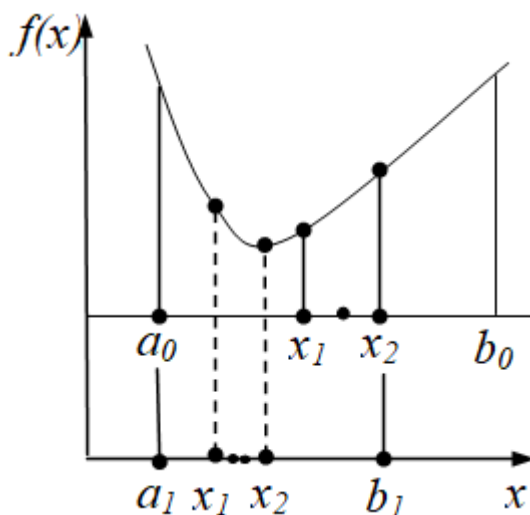


Рисунок 4.1 – Вибір кроку зміни аргументу для знаходження глобального екстремуму функції на відрізку

Знаходження найменшого значення функції виконується в циклі, у якому обчислюється поточне значення функції і порівнюється з найменшим із усіх попередніх значень  $f(x)$ . Якщо поточне значення  $y_i$  виявиться меншим за найменше  $y_{min}$  із попередніх значень, то його треба вважати новим найменшим значенням. У іншому випадку  $y_{min}$  залишається незмінним, тобто

$$y_{\min} = \begin{cases} y_i, & \text{якщо } y_i < y_{\min}; \\ y_{\min}, & \text{якщо } y_i \geq y_{\min}. \end{cases} \quad (4.1)$$

Для першого порівняння  $y_1$  із  $y_{\min}$  необхідно задати початкове значення  $y_{\min}$ .

Якщо  $f(x)$  є унімодальною, то метод можна модифікувати:

$$y_{\min} = \begin{cases} y_i, & \text{якщо } y_i < y_{\min}; \\ \text{кінець,} & \text{якщо } y_i \geq y_{\min}, \end{cases} \quad (4.2)$$

оскільки після досягнення мінімуму функція почне зростати й усі наступні  $y_i$  будуть більшими за  $y_{\min}$ .

## 4.2 Метод дихотомії (половинного розподілу)

У цьому методі проміжні значення  $x_1$  і  $x_2$  обираються на відстані  $\delta > 0$  від середини поточного інтервалу невизначеності (рис. 4.1).

$$\begin{cases} x_1 = \frac{a_k + b_k}{2} - \frac{\delta}{2}, \\ x_2 = \frac{a_k + b_k}{2} + \frac{\delta}{2}, \end{cases} \quad (4.3)$$

де  $\delta$  – параметр методу,

$a_k$  і  $b_k$  – межі інтервалу невизначеності на  $k$ -ому кроці пошуку.

Після обчислення значень  $f(x_1)$  і  $f(x_2)$  і їх порівняння, ці значення формують новий інтервал невизначеності:

$$\begin{cases} \text{якщо } f(x_1) \leq f(x_2), & \text{то } a_{k+1} = a_k, b_{k+1} = x_2; \\ \text{якщо } f(x_1) > f(x_2), & \text{то } a_{k+1} = x_1, b_{k+1} = b_k. \end{cases}$$

Потім знову обчислюють координати  $x_1$  і  $x_2$  і продовжують пошук. Після  $n$  обчислень функції ( $n=2k$ ) екстремум буде належати інтервалу

$$L = b_k - a_k = (b - a - \delta)/2^{n/2} + \delta. \quad (4.4)$$

Цей інтервал можна приймати за оцінку умови припинення пошуку, тобто пошук припиняється, коли  $l < \varepsilon$ . Тоді за екстремальне значення береться точка:

$$x^* = \frac{a_k + b_k}{2}. \quad (4.5)$$

**Зауваження.** Параметр  $\delta$  має бути меншим за  $\varepsilon$ . Можна покласти  $\delta = \varepsilon/2$ .

### 4.3 Метод золотого перетину

Відомо, що золотим перетином відрізка називається розподіл відрізка на дві нерівні частини так, що відношення довжини усього відрізка до довжини більшої частини дорівнює відношенню довжини більшої частини до довжини меншої частини, тобто

$$\frac{b-a}{b-x} = \frac{b-x}{x-a} \approx 1,618$$

Золотий перетин відрізка проводиться двома точками

$$\begin{cases} x_1 = a + \frac{3 - \sqrt{5}}{2} (b - a) \approx a + 0,382(b - a), \\ x_2 = a + \frac{\sqrt{5} - 1}{2} (b - a) \approx a + 0,618(b - a). \end{cases}$$

Ці точки розташовуються симетрично щодо центру інтервалу  $[a, b]$  з урахуванням значення коефіцієнта пропорційності  $\tau = 0,382$ . У свою чергу, точка  $x_1$  відповідає за золотий перетин відрізка  $[a, x_2]$ , а точка  $x_2$  – за золотий перетин відрізка  $[x_1, b]$  (рис. 4.2).

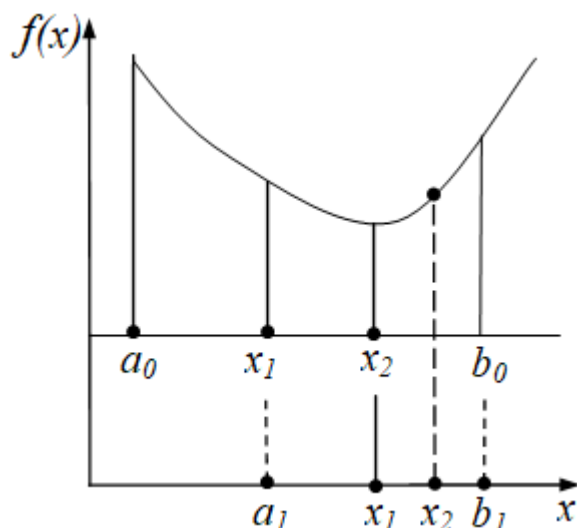


Рисунок 4.2 – Золотий перетин відрізка

Спираючись на цю властивість золотого перетину, можна запропонувати такий засіб мінімізації унімодальної на відрізку  $[a, b]$  функції. Проміжні значення  $x_1$  і  $x_2$  обираються симетрично щодо центру поточного інтервалу невизначеності в такий спосіб:

$$\begin{cases} x_1 = a_k + 0,382(b_k - a_k), \\ x_2 = b_k - 0,382(b_k - a_k). \end{cases} \quad (4.6)$$

Після обчислення в цих точках значення цільової функції, формують новий інтервал невизначеності (рис. б):

$$\begin{cases} \text{якщо } f(x_1) \leq f(x_2), & \text{то } a_{k+1} = a_k, b_{k+1} = x_2; \\ \text{якщо } f(x_1) > f(x_2), & \text{то } a_{k+1} = x_1, b_{k+1} = b_k. \end{cases}$$

Варто зауважити, що в методі золотого перетину немає необхідності у визначенні двох нових точок  $x_1^H$  і  $x_2^H$  згідно з формулою (4.6). Одна з них буде визначена на попередньому кроці.

Так, якщо  $f(x_1) \leq f(x_2)$ , тоді:

$$\begin{cases} x_1^H = a_k + 0,382(b_k - a_k), \\ x_2^H = x_1. \end{cases}$$

Після  $k$  ітерацій екстремум буде належати інтервалу довжини:

$$L = b_k - a_k = 0,618^k(b-a), \quad (4.7)$$



який приймається за оцінку умови припинення пошуку. Пошук припиняється, якщо  $l = b_k - a_k < \varepsilon$ . Тоді:

$$x^* = (a_k + b_k)/2.$$

Порівняння методів дихотомії і золотого перетину дозволяє зробити висновок, що досягнення однієї і тієї ж точності  $\varepsilon$  методом золотого перетину вимагає меншої кількості обчислень функції, оскільки на одному кроці методу дихотомії потрібно обчислити два значення функції, а в методі золотого перетину – одне.

#### 4.4 Метод квадратичної інтерполяції

У цьому методі за шуканий екстремум заданої унімодальної функції береться екстремум апроксимуючої її квадратичної функції.

Спочатку відшукується інтервал зміни  $x$ , якому повинна належати оптимальна точка  $x^*$ . Найпростіший спосіб визначення такого інтервалу полягає в послідовному обчисленні значень функції в точках:

$$x_0, x_0 + h, x_0 + 2h, x_0 + 4h, \dots$$

доти, поки чергове значення функції не стане більше за попереднє.

Знаючи інтервал, за його двома граничними точками і внутрішньою точкою можна побудувати інтерполяційний поліном:

$$\varphi(x) = a_0 + a_1x + a_2x^2$$

і задача знаходження  $x^*$  перетвориться на задачу мінімізації побудованого полінома  $\varphi(x)$ .

У результаті одержимо оцінку  $x^* = -a_1/(2a_2)$ . Якщо виявиться, що  $f(x^*)$  перевищує якийсь з трьох обчислених на інтервалі значень функції, то оцінка  $x^*$  вважається незадовільною і варто зменшити інтервал, на якому шукається мінімум.

Якщо три точки розташовані рівномірно і  $x_c$  – центральна, а  $x_a = x_c - h$ ,  $x_b = x_c + h$ , тоді маємо

$$x^* = x_c + \frac{h[f(x_a) - f(x_b)]}{2[f(x_a) - 2f(x_c) + f(x_b)]}. \quad (4.8)$$

### **Питання для самоперевірки**

1. Вкажіть основні методи одномірної оптимізації.
2. Вкажіть алгоритм застосування методу послідовного рівномірного перебору (сканування).
3. Наведіть алгоритм та основні розрахункові формули застосування методу дихотомії (половинного розподілу).
4. Наведіть основні розрахункові формули застосування методу золотого перетину.
5. Сформулюйте висновки порівняння методів дихотомії і золотого перетину.
6. Вкажіть алгоритм застосування квадратичної інтерполяції.
7. Вкажіть форму інтерполяційного поліному для методу квадратичної інтерполяції.

## 5 МЕТОДИ БАГАТОМІРНОЇ ОПТИМІЗАЦІЇ

При застосування методів багатомірної оптимізації будують таку послідовність точок

$$X_0, X_1, X_2, \dots, \text{ у якої } f(X_0) > f(X_1) > f(X_2) > \dots$$

Перед початком пошуку обирається деяка вихідна точка  $X_0 \in D_X$ . У якості  $X_0$  може бути обрана будь-яка точка, проте краще, якщо  $X_0$  розташована не занадто далеко від точки мінімуму. Потім необхідно обрати напрямок (вектор) прямування і величину кроку уздовж цього напрямку. З попередньої точки  $X_k$  перехід у точку  $X_{k+1}$  здійснюється за формулою:

$$X_{k+1} = X_k + \gamma_k P_k, k = 0, 1, \dots \quad (5.1)$$

де  $P_k$  – напрямок прямування;  $\gamma_k$  – довжина кроку.

Схема алгоритму пошуку екстремуму показана на рис. 5.1. На етапах 2 і 3 обирається напрямок пошуку й обчислюються координати нової точки. Далі в новій точці  $X$  обчислюється значення цільової функції  $f(X)$ . Крок закінчується перевіркою умов припинення пошуку, і в залежності від ступеня їхнього виконання пошук буде продовжено або припинено.

При аналізі різноманітних методів багатовимірної оптимізації корисні геометричні уявлення.

Будемо називати послідовність точок  $X_k$ , з'єднаних відрізками прямих, траєкторією пошуку. Геометричне уявлення функцій дається у вигляді поверхонь відгуку, а поверхні відгуку на рисунках представляються у вигляді сукупності ліній рівня. Лінія рівня має рівняння  $f(X) = a = const$  і є геометричним місцем точок простору керованих параметрів  $x_i$ , у яких значення цільової функції дорівнює  $a$ .

Геометричні уявлення зручно використовувати для відображення характерних рис функцій двох змінних  $f(X) = f(x_1, x_2)$ , наприклад таких, як яружність, багатоекстремальність, наявність сідлових точок (рис. 5.2).

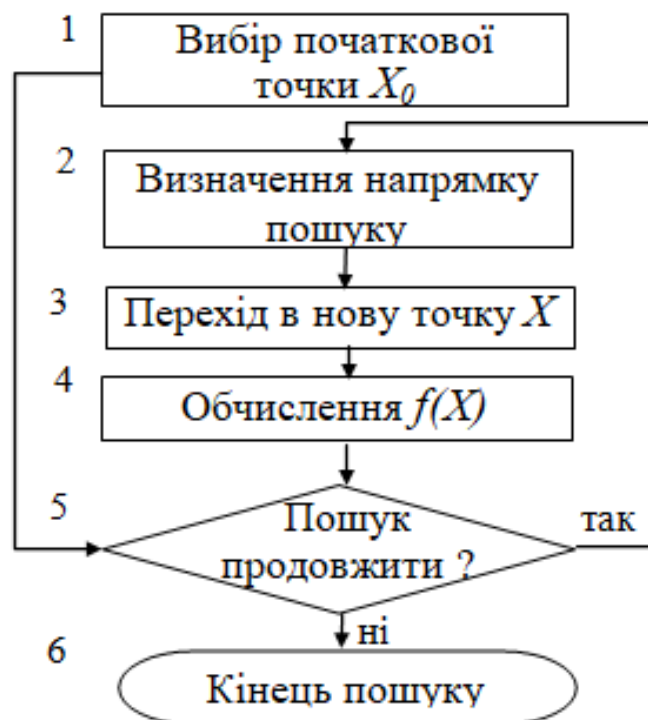


Рисунок 5.1 – Схема алгоритму пошуку екстремуму

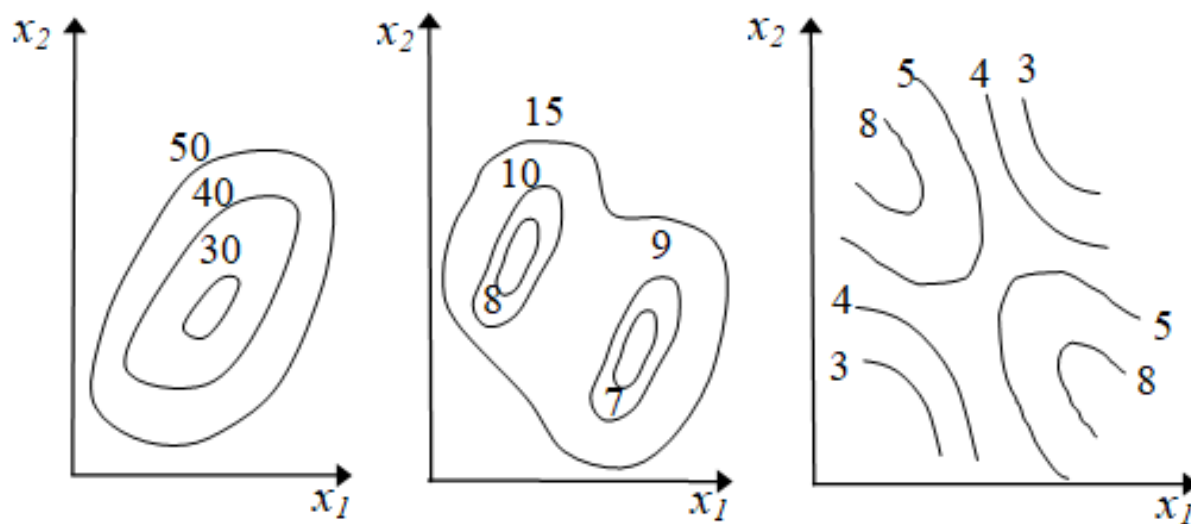


Рисунок 5.2 – Геометричне відображення характерних рис функцій двох змінних

### **Питання для самоперевірки**

1. Вкажіть, яким чином задається лінія рівня функції.
2. Сформулюйте поняття унімодальної функції.
3. Наведіть схему алгоритму пошуку екстремуму при багатомірній оптимізації.
4. Надайте геометричну інтерпретацію яружності, багатоекстремальності, наявності сідлових точок.

## 6 ГРАДІЄНТНІ МЕТОДИ

### 6.1 Градієнтні методи зі змінним кроком

Відомо, що градієнт  $\nabla f(X)$  ортогональний до поверхні відгуку цільової функції  $f(X)$  у точці  $X$ , і його напрямок збігається з локальним напрямком найшвидшого зростання функції.

Вектор  $-\nabla f(X)$ , протилежний градієнту, називається антиградієнтом і спрямований у бік найбільшого зменшення функції. У точці екстремуму градієнт дорівнює нулю. Тоді, обравши в якості напрямку спуску  $P_k$  антиградієнт функції в точці  $X_k$ , одержимо ітераційний процес вигляду:

$$X_{k+1} = X_k - \gamma_k \nabla f(X_k), \quad \gamma_k > 0, \quad k = 0, 1, \dots \quad (6.1)$$

У координатній формі це запишеться таким чином:

$$x_{ik+1} = x_{ik} - \gamma_k \nabla_i f(X_k), \quad i = \overline{1, n}.$$

Найчастіше в градієнтних методах у якості вектора  $P_k$  обирається вектор одиничної довжини, що має градієнтний напрямок:

$$P_k = -\nabla f(X_k) / \|\nabla f(X_k)\|,$$

де  $\|\nabla f(X_k)\|$  – норма вектора-градієнту, що дорівнює:

$$\sqrt{\sum_{i=1}^n \nabla_i^2 f(X_k)}.$$

Тоді одержимо таку формулу переходу з точки  $X_k$  у точку  $X_{k+1}$ :

$$X_{k+1} = X_k - \gamma_k \frac{\nabla f(X_k)}{\|\nabla f(X_k)\|}. \quad (6.2)$$

#### Алгоритм методу

**Етап 1.** Вибір початкової точки  $X_0$  і початкового кроку  $\gamma$ ,  $k = 0$ . Обчислення  $f(X_k)$ .

**Етап 2.** Обчислення градієнту функції  $\nabla f(X_k)$ .

**Етап 3.** Обчислення координат нової точки

$$X_{k+1} = X_k - \gamma_k \frac{\nabla f(X_k)}{\|\nabla f(X_k)\|}.$$

**Етап 4.** Якщо в новій точці  $f(X_{k+1}) < f(X_k)$ , то переходити до етапу 5, інакше – до етапу 6.

**Етап 5.** Прийняти  $k = k + 1$  і перейти до етапу 2.

**Етап 6.** Зменшити крок,  $\gamma = \gamma/2$ . Якщо при цьому крок більше граничного розміру  $\gamma > \gamma_{min}$ , то пошук продовжити з етапу 3 зі зменшеним кроком.

**Етап 7.** Покласти  $X^* = X_k$ ,  $f^* = f(X^*)$ .

Траєкторія пошуку мінімуму функції градієнтним методом наведена на рис. 6.1, а.

## 6.2 Метод найшвидшого спуску

Особливістю методу найшвидшого спуску є прямування з оптимальним кроком, розрахованим за допомогою одномірної мінімізації функції за  $\gamma$  уздовж антиградієнтного напрямку.

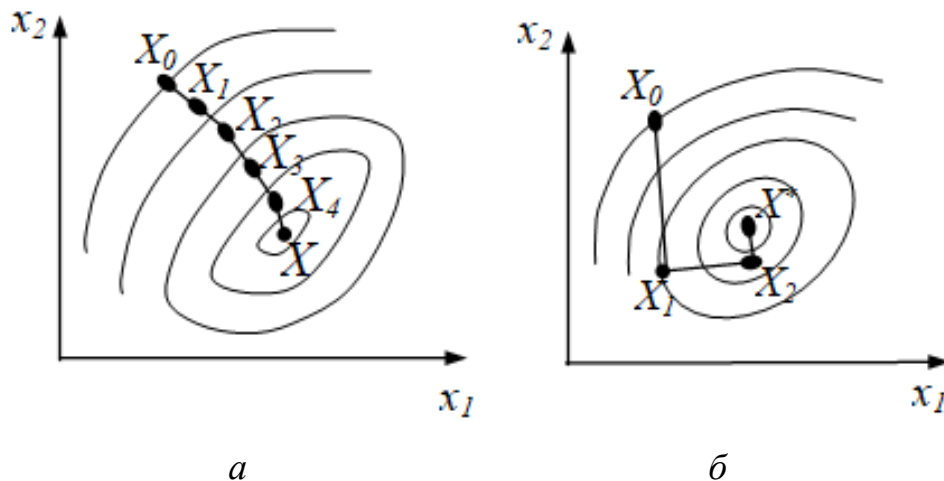


Рисунок 6.1 – Траєкторія пошуку мінімуму функції

На кожній ітерації величина кроку  $\gamma_k$  обирається з умови мінімуму функції  $f(X)$  у напрямку спуска, тобто:

$$f(X_k - \gamma_k \nabla f(X_k)) = \min_{\gamma > 0} f(X_k - \gamma \nabla f(X_k)) \quad (6.3)$$

або

$$f\left(X_k - \gamma_k \frac{\nabla f(X_k)}{\|\nabla f(X_k)\|}\right) = \min_{\gamma > 0} f\left(X_k - \gamma \frac{\nabla f(X_k)}{\|\nabla f(X_k)\|}\right).$$

Необхідна умова мінімуму функції  $f(X)$  у напрямку антиградієнта має вигляд:

$$\frac{df(\gamma)}{d\gamma} = (\nabla f(X_{k+1}), \nabla f(X_k)) = 0. \quad (6.4)$$

Згідно зі співвідношенням (6.4) маємо властивість ортогональності градієнтів у двох сусідніх точках. Тому до оптимального розв'язку веде зигзагоподібний шлях із поворотом під прямим кутом у кожній точці (рис. 6.1, б).

**Алгоритм методу** містить такі етапи.

**Етап 1.** Вибір початкової точки  $X_0$ ,  $k = 0$ .

**Етап 2.** Обчислення градієнту функції  $\nabla f(X_k)$ .

**Етап 3.** Знаходження одним із методів одномірної оптимізації оптимального кроку  $\gamma_k$  уздовж антиградієнтного напрямку.

**Етап 4.** Перевірка умови припинення пошуку. Пошук припиняється при виконанні однієї з таких умов:

- 1) при виконанні критерію завершення ітераційного процесу  $\gamma_k < \gamma_{min}$ ;
- 2) при значеннях норми вектора-градієнту  $\|\nabla f(X_k)\| < \nabla_{min}$ ;
- 3) при використанні заданого ресурсу кількості ітерацій  $k > k_{max}$  або кількості обчислень функції.

Якщо умови припинення пошуку виконуються, то виконується перехід до етапу 7, інакше – до етапу 5.

**Етап 5.** Обчислення координат нової точки:

$$X_{k+1} = X_k - \gamma_k \frac{\nabla f(X_k)}{\|\nabla f(X_k)\|}.$$

**Етап 6.** Прийняти  $k = k + 1$  і перейти до етапу 2.

**Етап 7.** Покласти  $X^* = X_k$ ,  $f^* = f(X^*)$ .



### 6.3 Особливості пошуку оптимуму для цільових функцій із яружним рельєфом поверхні відгуку

Розглянемо функцію двох змінних:

$$f(x_1, x_2) = (x_1 - 2)^2 + (x_2 - 3)^2 + 6.$$

Складемо матрицю других похідних функції (матрицю Гессе) і обчислимо визначники:

$$\frac{\partial f}{\partial x_1} = 2(x_1 - 2); \quad \frac{\partial f}{\partial x_2} = 2x_2 - 6;$$

$$\frac{\partial^2 f}{\partial x_1^2} = 2; \quad \frac{\partial^2 f}{\partial x_2^2} = 2; \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0;$$

$$H = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}; \quad f_{11} > 0, \quad \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} > 0.$$

Оскільки обидва визначники додатні, то функція є усюди опуклою і має глобальний мінімум у точці (2,3). Лінії рівня цієї функції являють собою кола.

Знайдемо власні числа матриці Гессе, тобто розв'яжемо рівняння:

$$\begin{vmatrix} 2 - \lambda & 0 \\ 0 & 2 - \lambda \end{vmatrix} = (2 - \lambda)^2 = 0; \quad \lambda_1 = \lambda_2 = 2$$

Тоді число обумовленості  $z = \lambda_{\max} / \lambda_{\min} = 2/2 = 1$ .

Антиградієнтний напрямок у будь-якій точці  $X$  є напрямком прямування до мінімуму, тоді для найшвидшого спуску достатньо обрати оптимальну величину кроку, щоб вирішити задачу за один крок. Траєкторія спуску показана на рис. 6.2, а.

Візьмемо тепер іншу функцію і проведемо аналогічні обчислення:

$$f(x_1, x_2) = 4(x_1 - 5)^2 + 36(x_2 - 3)^2 + 8;$$

$$\nabla_1 f = 8(x_1 - 5); \quad \nabla_2 f = 72(x_2 - 3);$$

$$H = \begin{pmatrix} 8 & 0 \\ 0 & 72 \end{pmatrix}; \quad f_{11} > 0; \quad \begin{vmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{vmatrix} > 0.$$

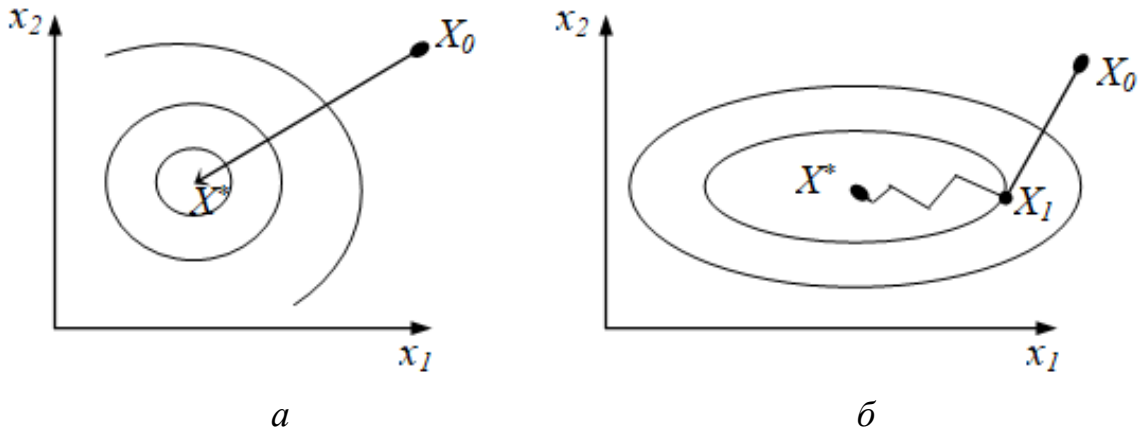


Рисунок 6.2 – Траєкторія спуску

Функція також є усюди опуклою, має глобальний мінімум у точці  $(5, 3)$ . Лінії рівня являють собою еліпси. Знайдемо власні числа і число обумовленості  $z$ :

$$\begin{vmatrix} 8 - \lambda & 0 \\ 0 & 72 - \lambda \end{vmatrix} = 0;$$

$$\lambda_1 = 8; \quad \lambda_2 = 72;$$

$$z = \lambda_{\max} / \lambda_{\min} = 72/8 = 9.$$

Лінії рівня будуть витягнуті уздовж осі  $ox_1$  у  $\sqrt{z} = 3$  раз. Траєкторія найшвидшого спуску стає зигзагоподібною, треба зробити декілька кроків для досягнення околу мінімуму (рис. 6.2, б). Функція уздовж осі  $ox_1$  змінюється набагато повільніше. Це явище називається ефектом ярів.

Напрямок антиградієнта в таких випадках істотно відхиляється від напрямку до точки мінімуму, що уповільнює збіжність градієнтного методу.

Таким чином, для сильно яружних функцій із числом обумовленості  $z \gg 1$  градієнтні методи мають малу швидкість збіжності і необхідно використовувати інші більш ефективні методи (метод спряжених градієнтів, метод Розенброка).

#### 6.4 Обчислення градієнта цільової функції

У більшості практичних задач градієнт цільової функції, що входить до алгоритму оптимізації, важко, а іноді навіть неможливо, одержати в явній формі від параметрів. Тому для обчислення градієнта використовується алгоритм чисельного методу:

$$\nabla_i f = \frac{f_{i+} - f_{i-}}{2a_i}, \quad i = \overline{1, n}, \quad (6.5)$$

де  $f_{i+} = f(x_1, \dots, x_{i-1}, x_i + a_i, x_{i+1}, \dots, x_n)$ ;  $f_{i-} = f(x_1, \dots, x_{i-1}, x_i - a_i, x_{i+1}, \dots, x_n)$ ;

$a_i$  – достатньо мала додатна величина, що використовується для наближеного обчислення градієнта,  $a_i = 10^{-4} - 10^{-6}$ .

#### 6.5 Метод спряжених градієнтів (метод Флетчера–Рівса)

Цей метод належить до групи методів спряжених напрямків. У цих методах кроки ітераційної процедури мінімізації цільової функції виконуються в спряжених напрямках.

Два  $n$ -мірних вектори  $X$  і  $Y$  називаються спряженими стосовно матриці  $H$  ( $H$ -спряженими), якщо скалярний добуток  $(X, HY) = 0$ . Тут  $H$  – симетрична додатно визначена матриця порядку  $n$ .

Методи спряжених напрямків мають у порівнянні з градієнтними методами більш високу швидкість збіжності. Наприклад, мінімум додатно визначеної квадратичної функції  $n$  змінних може бути знайдений максимум за  $n$  кроків.

Метод Флетчера–Рівса перетворює на кожному кроці антиградієнт  $-\nabla f(X_k)$  у напрямок  $P_k$ ,  $H$  – сполучений з раніше знайденими напрямками  $P_0, P_1, \dots, P_{k-1}$ . Тут  $H$  – матриця Гессе функції  $f(X)$ .

Відповідно до методу напрямки  $P_k$  обчислюються за формулами:

$$\begin{cases} P_0 = -\nabla f(X_0); \\ P_k = -\nabla f(X_k) + \beta_k P_{k-1}, \quad k \geq 1. \end{cases} \quad (6.6)$$

Коефіцієнт  $\beta_k$  обирається таким, щоб напрямки  $P_k$  і  $P_{k-1}$  були  $H$ -спряженими, тобто  $(P_k, H P_{k-1}) = 0$ . Ця умова буде виконана, якщо:

$$\beta_k = \frac{(\nabla f(X_k), \nabla f(X_k))}{(\nabla f(X_{k-1}), \nabla f(X_{k-1}))} = \frac{\|\nabla f(X_k)\|^2}{\|\nabla f(X_{k-1})\|^2}. \quad (6.7)$$

Ітераційна процедура має вид  $X_{k+1} = X_k + \gamma_k P_k$ .

Величина кроку  $\gamma_k$  обирається з умови мінімуму функції  $f(X)$  за  $\gamma$  у напрямку спуску, тобто в результаті розв'язання задачі одновимірної мінімізації.

$$f(X_k + \gamma_k P_k) = \min_{\gamma > 0} f(X_k + \gamma P_k). \quad (6.8)$$

**Алгоритм методу спряжених градієнтів складається з таких етапів.**

**Етап 1.** Обчислення в точці  $X_0$  вектора  $P_0 = -\nabla f(X_0)$ ,  $k = 0$ .

**Етап 2.** Знаходження мінімуму  $f(X)$  одним із методів одновимірного пошуку в напрямку  $P_k (k \geq 0)$ , тобто визначення кроку  $\gamma_k$  і точки  $X_{k+1} = X_k + \gamma_k P_k$ .

**Етап 3.** Обчислення величин  $f(X_{k+1})$  і  $\nabla f(X_{k+1})$ .

**Етап 4.** Якщо  $\|\nabla f(X_{k+1})\| < \varepsilon$ , то переходити до етапу 7.

**Етап 5.** Визначення нового напрямку  $P_{k+1}$  із співвідношення:

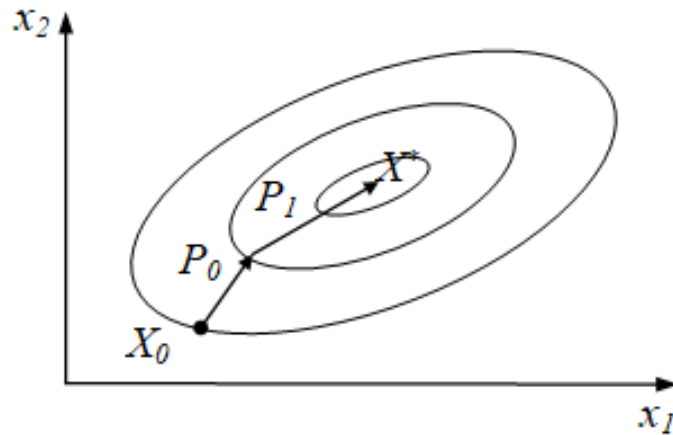
$$P_{k+1} = -\nabla f(X_{k+1}) + \frac{\|\nabla f(X_{k+1})\|^2}{\|\nabla f(X_k)\|^2} P_k.$$

**Етап 6.** Покласти  $k = k + 1$  і переходити до етапу 2.

**Етап 7.** Покласти  $X^* = X_{k+1}$ ,  $f^* = f(X^*)$ .

При мінімізації неквадратичної функції після  $(n + 1)$ -го кроку процедури 1–5 циклічно повторюються з заміною  $X_0$  на  $X_{n+1}$ . У випадку мініміза-

ції квадратичної функції двох змінних методом сполучених градієнтів траєкторія пошуку показана на рис. 6.3.



*Рисунок 6.3 – Траєкторія пошуку мінімізації квадратичної функції двох змінних методом сполучених градієнтів*

### **Питання для самоперевірки**

1. Вкажіть основні особливості градієнтних методів відшукування екстремуму.
2. Опишіть алгоритм застосування градієнтних методів зі змінним кроком.
3. Вкажіть основні етапи застосування методу найшвидшого спуску.
4. Вкажіть, за який напрямок відповідає вектор антиградієнту.
5. Сформулюйте особливості пошуку оптимуму для цільових функцій із яружним рельєфом поверхні відгуку.
6. Наведіть чисельну форму обчислення градієнту.
7. Наведіть алгоритм методу спряжених градієнтів.

## 7 МЕТОДИ БЕЗУМОВНОЇ ОПТИМІЗАЦІЇ НУЛЬОВОГО ПОРЯДКУ

Застосування градієнтних методів у ряді випадків виявляється важкою або недоцільною процедурою. Тому одержали розвиток методи оптимізації нульового порядку, в яких напрямок мінімізації цілком визначається обчисленням значень тільки цільової функції.

### 7.1 Метод покоординатного спуску (метод Гауса–Зейделя)

При застосуванні методу покоординатного спуску напрямок чергового кроку обирається уздовж якоїсь однієї координатної осі, наприклад  $x_1$ . Прямування в цьому напрямку проводиться вбік зменшення  $f(X)$  і виконується доти, поки  $f(X)$  спадає. Далі прямування починається уздовж нової координатної осі, наприклад  $x_2$  і т.д.

Після спусків уздовж  $n$  осей проводиться нова ітерація. У випадку, коли ні за однією з осей неможливо переміщення з кроком  $\gamma > \gamma_{min}$  із зменшенням функції  $f(X)$ , пошук припиняється й отримана точка  $X$  приймається за екстремальну  $X^*$ . Приклад можливої траєкторії покоординатного спуску приведено на рис. 7.1.

Наведемо основні етапи застосування методу покоординатного спуску.

**Алгоритм методу** стосовно до  $k$ -ої ітерації ( $k \geq 0$ ).

**Етап 1.** Знаходження мінімуму  $f(X)$  із точки  $X_k$  одним із методів одномірною пошуку в напрямку першої координатної осі, тобто в напрямку вектора  $l_1 = (1, 0, \dots, 0)$ . При цьому визначається точка:

$$X_{k1} = X_k + \gamma_1 l_1, \quad (7.1)$$

де  $\gamma_1$  знаходиться з умови  $f(X_k + \gamma_1 l_1) = \min_{\gamma} f(X_k + \gamma l_1)$ .

**Етап 2.** Послідовна одномірна мінімізація  $f(X)$  у напрямку інших координатних осей:

$$X_{k2} = X_{k1} + \gamma_2 l_2,$$

...

$$X_{kn} = X_{kn-1} + \gamma_n l_n.$$

У результаті одержимо точку:

$$X_{k+1} = X_k + \sum_{i=1}^n \gamma_i l_i = X_{kn}. \quad (7.2)$$

**Етап 3.** Перевірка умови завершення пошуку  $|\gamma_{max}| < \gamma_{min}$ . Якщо умова виконана, переходимо до етапу 5, інакше – до етапу 4.

**Етап 4.** Збільшити  $k$ :  $k = k + 1$ . Перейти до етапу 1.

**Етап 5.** Покласти  $X^* = X_{k+1}$ ,  $f^* = f(X^*)$ .

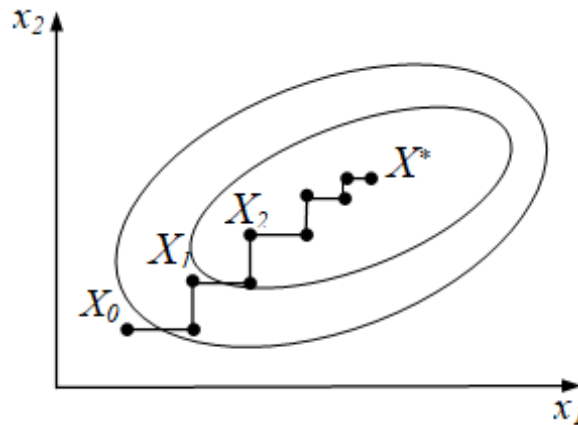


Рисунок 7.1 – Приклад можливої траєкторії покоординатного спуску

## 7.2 Метод прямого пошуку (метод Хука–Дживса)

**Алгоритм методу** прямого пошуку включає такі етапи.

**Етап 1.** Задають деяку початкову точку  $X_0$ , вектор зміни координат  $\Delta X$  у процесі дослідження околу точки  $X$ , найменше допустиме значення  $\varepsilon$  компонент вектора  $\Delta X$ .

**Етап 2.** Приймають, що  $X_0$  є базисною точкою  $X_B$  і обчислюють значення  $f(X_B)$ .

**Етап 3.** Циклічно змінюють кожну координату  $x_{iB}$  ( $i = \overline{1, n}$ ) базисної точки  $X_B$  на величину  $\Delta x_i$ , тобто обчислюють значення функцій у точках:

$$X = (x_{1B}, x_{2B}, \dots, x_{iB} \pm \Delta x_i, \dots, x_{nB}). \quad (7.3)$$

Якщо  $f(X) < f(X_B)$ , то відповідна координата  $x_i$  набуває нового значення  $x_{iB} + \Delta x_i$  або  $x_{iB} - \Delta x_i$ . У протилежному випадку значення цієї координа-

ти залишається незмінним. Після зміни  $n$ -ої координати одержують точку  $X_1$ . Якщо  $f(X_1) < f(X_B)$ , то переходять до етапу 4, інакше – до етапу 8.

**Етап 4.** Здійснюють спуск із точки  $X_1$  у точку:

$$X_2 = X_1 + (X_1 - X_B) = 2X_1 - X_B \quad (7.4)$$

або в координатній формі:

$$x_{i2} = 2x_{i1} - x_{iB}, \quad i = \overline{1, n}$$

і обчислюють значення  $f(X_2)$ .

**Етап 5.** Як і на етапі 3 циклічно змінюють кожну координату точки  $X_2$  відповідно до формули (7.4), виконуючи порівняння значень  $f(X)$  із значенням  $f(X_2)$ . Після зміни останньої координати одержують точку  $X_3$ . Якщо  $f(X_3) < f(X_1)$ , то переходять до етапу 6, інакше – до етапу 7.

**Етап 6.** Приймають, що  $X_1$  є новою базисною точкою  $X_B$ , а  $X_3$  – точкою  $X_1$  і повертаються до етапу 4.

**Етап 7.** Приймають, що  $X_1$  є новою базисною точкою  $X_B$  і повертаються до етапу 3.

**Етап 8.** Якщо  $\Delta x_i < \varepsilon$ , то переходять до етапу 10, інакше – до етапу 9.

**Етап 9.** Зменшують значення  $\Delta x_i$  (наприклад, у два рази) і повертаються до етапу 3.

**Етап 10.** Приймають  $X^* = X_B, f^* = f(X^*)$ .

До переваг методу прямого пошуку можна віднести простоту програмування. Недолік цього методу, як і методу покоординатного спуску, полягає в тому, що у випадку сильно витягнутих, вигнутих ліній рівня цільової функції або ліній з гострими кутами, він може виявитися непридатним для знаходження точку мінімуму.

### 7.3 Симплексний метод

Сиплексом називається регулярний (правильний) багатогранник у  $n$ -мірному евклідовому просторі, що має  $n + 1$  вершину. Наприклад, для випадку двох змінних регулярним сиплексом є рівносторонній трикутник (рис. 7.2), для випадку трьох змінних – тетраедр.



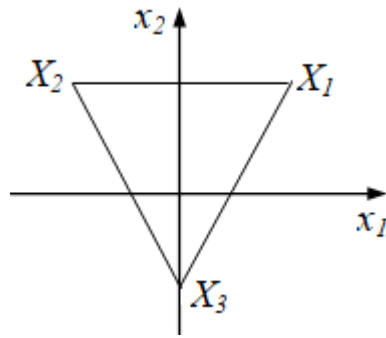


Рисунок 7.2 – Регулярний симплекс для випадку двох змінних

Координати вершин початкового регулярного симплекса можна визначити за допомогою матриці розміру  $(n + 1)n$ . Елемент матриці дорівнює  $j$ -ій координаті  $i$ -ї вершини симплекса.

$$\begin{matrix}
 & x_1 & x_2 & \dots & x_{n-1} & x_n \\
 X_1 & a_1 & a_2 & \dots & a_{n-1} & a_n \\
 X_2 & -R_1 & a_2 & \dots & a_{n-1} & a_n \\
 X_3 & 0 & -R_2 & \dots & a_{n-1} & a_n \\
 \dots & \dots & \dots & \dots & \dots & \dots \\
 X_n & 0 & 0 & \dots & -R_{n-1} & a_n \\
 X_{n+1} & 0 & 0 & \dots & 0 & -R_n
 \end{matrix} \quad , \quad (7.5)$$

де  $a_i = \sqrt{\frac{1}{2i(i+1)}}$ ;  $R_i = \sqrt{\frac{i}{2(i+1)}}$ ,  $i = \overline{1, n}$ .

Для випадку двох або трьох змінних матриця має вигляд (табл. 7.1), тобто симплекс розташований таким чином, що його центр ваги знаходиться в початку координат, довжина сторони симплекса дорівнює 1 (рис. 7.2).

Таблиця 7.1 – Матриця для симплекса

Точка	Координати точки		
	$x_1$	$x_2$	$x_3$
$X_1$	0.5	0.289	0.204
$X_2$	-0.5	0.289	0.204
$X_3$	0	-0.578	0.204
$X_4$	0	0	-0.612

Пошук мінімуму функції симплексним методом ведеться в такий спосіб. Встановлюються координати вершин початкового симплекса й у кожній із них обчислюється значення цільової функції  $f(X)$  (рис. 7.3).

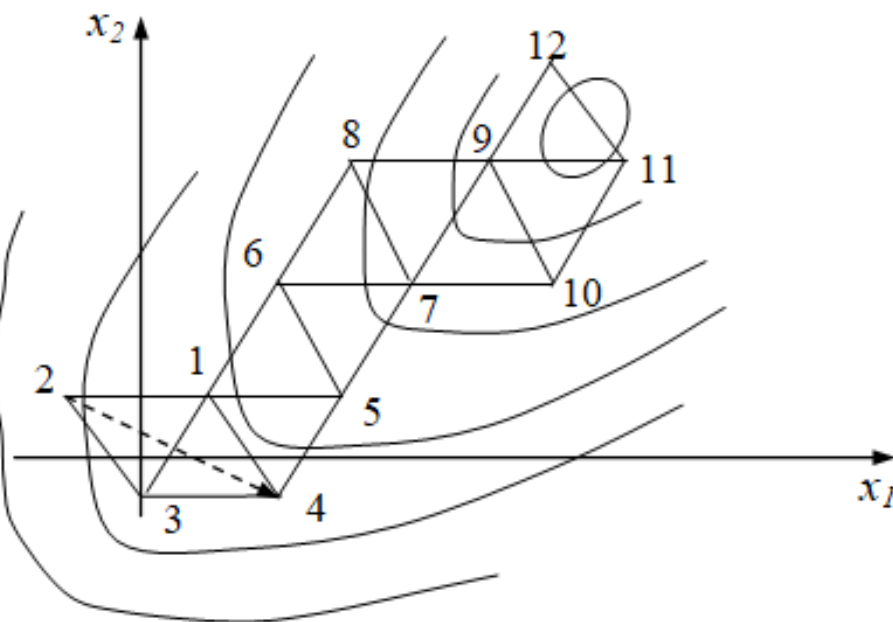


Рисунок 7.3 – Пошук мінімуму функції симплексним методом

Визначається вершина з найбільшим значенням функції  $f(X_{max})$ . Через цю вершину і центр протилежної грані симплекса проводиться пряма, на якій обирається нова вершина. Координати цієї точки визначаються за формулою:

$$X_H = 2X_C - X_{max} \quad (7.6)$$

або в координатній формі:

$$x_{in} = 2x_{ic} - x_{imax}, \quad i = \overline{1, n}.$$

Координати центру грані (центру ваги усіх вершин, за винятком  $X_{max}$ )  $X_C$  обчислюються за формулою:

$$x_{ic} = \frac{1}{n} \left( \sum_{j=1}^{n+1} x_{ij} - x_{imax} \right) \quad (7.7)$$

Потім вершина з найбільшим значенням  $f(X_{max})$  відкидається і новий симплекс будується на вершинах, що залишилися, і новій  $X_n$ .

Викладений метод використовує регулярні симплекси і схильний до зациклювання. Крім того, він має невелику точність, що обмежена розмірами симплекса, і погано працює в яружних ситуаціях.

Тому застосовується модифікація симплексного методу, що дозволяє усунути ці недоліки. Координати вершин початкового симплекса можуть бути визначені за формулами (7.5) або призначені довільно. На наступних етапах конфігурація симплекса змінюється, симплекс утрачає регулярність. Такий метод називається методом деформованого багатогранника або Нелдера–Міда.

## 7.4 Метод Розенброка

Метод Розенброка є розвитком методу покоординатного спуску. Основна ідея методу полягає в обертанні системи координат відповідно до зміни швидкості спадання цільової функції. Нові напрямки координатних осей визначаються таким чином, щоб одна з них відповідала напрямку найбільш швидкого спадання функції, а інші знаходяться з умови ортогональності. Ідея методу полягає в такому (рис. 7.4).

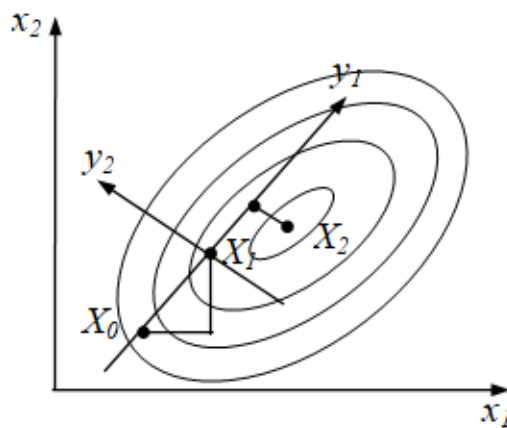


Рисунок 7.4 – Геометрична ілюстрація методу Розенброка

З початкової точки  $X_0$  здійснюють спуск за напрямками, паралельними координатним осям (покоординатний спуск). На наступній ітерації одна з осей повинна проходити в напрямку  $y_1 = X_1 - X_0$ , а інша в напрямку, орто-

гональному до  $y_1$ . Спуск уздовж цих осей призводить у точку  $X_2$ , що дає можливість побудувати новий вектор  $X_2 - X_1$ , і на його базі нову систему координат. У загальному випадку цей метод є найбільш ефективний при мінімізації яружних функцій, тому що результувальний напрямок пошуку буде розташований уздовж осі яру.

### Алгоритм методу Розенброка

**Етап 1.** Позначають через  $P_1, P_2, \dots, P_n$  напрямки координатних осей деякої точки  $X_k$  ( $k \geq 0$ ). Виконують одномірний пошук уздовж осі  $P_1$ , тобто знаходять оптимальний крок  $\gamma_1$  і точку  $X_{k1} = X_k + \gamma_1 P_1$ .

**Етап 2.** Виконують одномірну мінімізацію уздовж інших координатних осей:

$$X_{k2} = X_{k1} + \gamma_2 P_2,$$

....

$$X_{kn} = X_{kn-1} + \gamma_n P_n.$$

У результаті одержують точку:

$$X_{k+1} = X_{kn} = X_k + \sum_{i=1}^n \gamma_i P_i.$$

**Етап 3.** Обирають нові осі координат  $P_1, P_2, \dots, P_n$ . За першу ось приймають вектор  $P_1 = X_{k+1} - X_k$ . Інші осі будуються за допомогою процедури ортогоналізації Грама-Шмідта, тобто:

$$(P_i, P_j) = 0; i \neq j.$$

**Етап 4.** Перевіряють умову збіжності. Якщо  $\|X_{k+1} - X_k\| > \varepsilon$ , переходять до етапу 1.

**Етап 5.** Приймають  $X^* = X_{k+1}, f^* = f(X^*)$ .

## 7.5 Метод паралельних дотичних (метод Пауелла)

Цей метод використовує властивість квадратичної функції, яка полягає в тому, що будь-яка пряма, що проходить через точку мінімуму  $X^*$ , перетинає під рівними кутами дотичні до поверхонь рівного рівня функції в точках перетинання. Ідея методу полягає в наступному (рис. 7.5).

Обирається деяка точка  $X_0$  і виконується одномірний пошук уздовж довільного напрямку, що призводить у точку  $X_1$ . Потім із точки  $X_2$  здійснюється одномірний пошук уздовж прямої, паралельної  $X_0 - X_1$ . Отримана в результаті точка  $X_3$  разом із точкою  $X_1$  визначають напрямок  $X_1 - X_3$  одномірного пошуку, що призводить до точки  $X^*$ . У випадку квадратичної функції оптимальне значення знаходиться за  $n$  ітерацій.

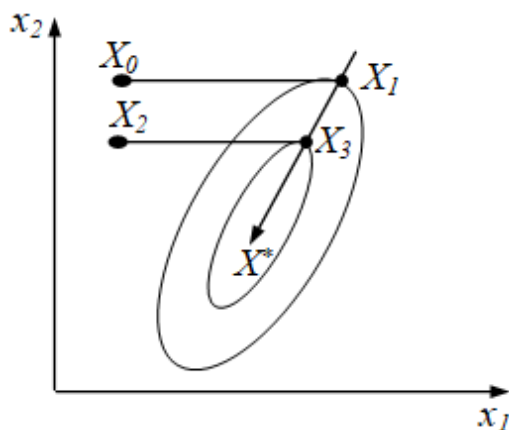


Рисунок 7.5 – Геометрична ілюстрація методу Пауелла

### Алгоритм методу Пауелла

**Етап 1.** З точки  $X_0$  виконують одномірний пошук уздовж  $n$ -ої осі, тобто знаходять оптимальний крок  $\gamma_n$  і точку  $X_1 = X_0 + \gamma_n P_n$ .

**Етап 2.** Виконують  $n$  покоординатних спусків уздовж ортогональних напрямків  $P_k$ ,  $k = \overline{1, n}$ . Величина кроку  $\gamma_k$  визначається з умови:

$$f(X_k + \gamma_k P_k) = \min f(X_k + \gamma P_k).$$

Отриманий крок  $\gamma$  визначає точку:

$$X_{k+1} = X_k + \gamma_k P_k.$$

Після  $n$  одномірних пошуків одержують точку  $X_{n+1}$ .

**Етап 3.** Здійснюють одномірний пошук уздовж напрямку:

$$P_{n+1} = X_{n+1} - X_1.$$

Знайдений крок  $\gamma_{n+1}$  дозволяє перейти в точку  $X_{n+2} = X_{n+1} + \gamma_{n+1} P_{n+1}$ .

**Етап 4.** Перевіряють умови закінчення пошуку. Якщо

$$\|X_{n+2} - X_{n+1}\| < \varepsilon,$$

переходять до етапу 6, інакше – до етапу 5.

**Етап 5.** Приймають  $X_0 = X_{n+2}$  і переходять до етапу 1.

**Етап 6.** Приймають  $X^* = X_{n+2}, f^* = f(X^*)$ .

### **Питання для самоперевірки**

1. Вкажіть основні методи безумовної оптимізації нульового порядку.
2. Сформулюйте основні етапи застосування методу покоординатного спуску.
3. Наведіть алгоритм застосування метод прямого пошуку (метод Хука-Дживса).
4. Сформулюйте основні етапи використання симплекс методу.
5. Наведіть геометричну інтерпретацію траєкторії спуску за методом Гауса-Зейделя.
6. Наведіть алгоритм методу паралельних дотичних (метод Пауелла).
7. Вкажіть основні етапи використання методу Розенброку.

## 8 МЕТОДИ ВИПАДКОВОГО ПОШУКУ

В основі всіх методів випадкового пошуку лежить внесення елементів випадку до процедури формування випробувальних точок, які використовуються для визначення напрямку пошуку.

### 8.1 Метод випадкового пошуку з перерахунком

Розглянемо алгоритм методу випадкового пошуку стосовно задачі умовної оптимізації за наявності прямих обмежень вигляду:

$$x_{i\min} \leq x_i \leq x_{i\max}, \quad i = \overline{1, n}.$$

Ітераційна процедура мінімізації цільової функції  $f(X)$  методом випадкового пошуку з перерахунком задається формулою:

$$X_{k+1} = X_k + \begin{cases} \gamma_k S \zeta, & \text{якщо } f(X_k + \gamma_k S \zeta) < f(X_k); \\ 0, & \text{в протилежному випадку,} \end{cases} \quad (8.1)$$

де  $\gamma_k$  – крок пошуку;

$S = (s_1, s_2, \dots, s_n)$  – вектор масштабних коефіцієнтів;

$\zeta = (\zeta_1, \zeta_2, \dots, \zeta_n)$  – випадковий  $n$ -мірний вектор із рівномірним законом розподілу в інтервалі  $[-1, 1]$ .

У координатній формі ітерація запишеться в такий спосіб:

$$x_{ik+1} = x_{ik} + \begin{cases} \gamma_k s_i \zeta_i, & \text{якщо } f(X_k + \gamma_k S \zeta) < f(X_k); \\ 0, & \text{в протилежному випадку.} \end{cases}$$

Масштабні коефіцієнти  $s_i$  визначаються співвідношенням:

$$s_i = \frac{x_{i\max} - x_{i\min}}{\max_i (x_{i\max} - x_{i\min})}, \quad i = \overline{1, n}. \quad (8.2)$$

Очевидно, усі  $s_i \leq 1$ .

Якщо якась із компонент  $x_i$  вектора  $X$  виходить за припустимі межі, тобто  $x_i > x_{imax}$  або  $x_i < x_{imin}$ , то цій компоненті присвоюється граничне значення:

$$x_i = \begin{cases} x_{i \max}, & \text{якщо } x_i > x_{i \max}, \\ x_{i \min}, & \text{якщо } x_i < x_{i \min}. \end{cases} \quad (8.3)$$

Завдяки цій умові забезпечується пошук уздовж межі області параметрів при порушенні обмежень.

Якщо при оптимізації пошукова точка цілком виходить за межі допустимої області, тобто  $X > X_{max}$  або  $X < X_{min}$ , то цей крок вважається невдалим, а значення усіх компонент вектора  $X$  покладаються рівними граничним. Якщо на якійсь ітерації маємо поспіль  $m$  невдалих спроб, то крок пошуку зменшують:

$$\gamma_{k+1} = \alpha \gamma_k, \quad (0 < \alpha < 1).$$

Якщо в процесі пошуку на якійсь ітерації виявиться, що для всіх  $m$  пробних кроків функція не спадає, то крок пошуку також зменшують:

$$\gamma_{k+1} = \alpha \gamma_k.$$

Пошук припиняється при виконанні умови  $\gamma_k < \gamma_{min}$ .

## 8.2 Метод випадкового пошуку за найкращим випробуванням (пошук за статистичним градієнтом)

Алгоритм методу стосовно  $k$ -ої ітерації ( $k \geq 0$ ).

**Етап 1.** З точки  $X_k$  виконують  $m$  випадкових випробувальних кроків:

$$\Delta X_j = \gamma_k \zeta_j, \quad j = \overline{1, m}, \quad (8.4)$$

де  $\gamma_k$  – крок пошуку;

$S$  – вектор масштабних коефіцієнтів;

$\zeta_j$  –  $n$ -мірний випадковий вектор на  $j$ -ому кроці.

Обчислюють значення цільової функції для всіх випробувальних кроків  $f_j(X_k + \Delta X_j)$ ,  $j = \overline{1, m}$ .



**Етап 2.** З усіх спроб обирають найкращу, що призводить до найбільшого зменшення цільової функції, тобто визначають  $\Delta X_n$  з умови:

$$f(X_k + \Delta X_n) = \min_j f_j(X_k + \Delta X_j). \quad (8.5)$$

Якщо  $f(X_k + \Delta X_n) < f(X_k)$ , переходимо до етапу 3, інакше – до етапу 5.

**Етап 3.** У напрямку найкращого випробування здійснюють робочий крок:

$$X_{k+1} = X_k + \Delta X_n. \quad (8.6)$$

**Етап 4.** Приймають  $k = k + 1$  і повертаються до етапу 1.

**Етап 5.** Зменшують крок пошуку  $\gamma_k = \alpha \gamma_k$ , ( $0 < \alpha < 1$ ).

**Етап 6.** Якщо  $\gamma_k > \gamma_{min}$ , повертаються до етапу 1.

**Етап 7.** Приймають  $X^* = X_k$ ,  $f^* = f(X^*)$ .

Із збільшенням кількості випробувань напрямок найкращого випробування наближається до найкращого – антиградієнтного. Проте, при використанні випадкового пошуку необхідності в обчисленні градієнта функції немає, що скорочує витрати машинного часу.

### **Питання для самоконтролю**

1. Вкажіть основні положення методів випадкового пошуку.
2. Наведіть алгоритм й обчислювані формули методу випадкового пошуку з перерахунком.
3. Сформулюйте основні етапи використання методу випадкового пошуку за найкращим випробуванням.
4. Наведіть основні розрахункові формули пошуку за статистичним градієнтом.
5. Наведіть висновки порівняння методів випадкового пошуку.

## 9 МЕТОДИ БЕЗУМОВНОЇ ОПТИМІЗАЦІЇ ДРУГОГО ПОРЯДКУ

### 9.1 Метод Ньютона

Якщо за напрямком спуску обрати у вигляді:

$$P_k = -H^{-1}(X_k) \nabla f(X_k),$$

де  $H^{-1}(X_k)$  – обернена матриця для матриці Гессе, то одержимо ітераційну процедуру виду:

$$X_{k+1} = X_k - H^{-1}(X_k) \nabla f(X_k), \quad (9.1)$$

у якій крок пошуку дорівнює 1.

Одержаний метод мінімізації називається методом Ньютона. Якщо матриця Гессе функції  $f(X)$  додатно визначена, то послідовність  $\{X_k\}$  збігається до точки локального мінімуму  $X^*$ .

Якщо  $f(X)$  – квадратична й опукла, тобто

$$f(X) = b_0 + (B, X) + 1/2(X, HX),$$

тоді  $H(X) = H = const$ ;  $\nabla f(X) = B + HX$ .

Звідси маємо:

$$X_1 = X_0 - H^{-1}(X_0) \nabla f(X_0) = X_0 - H^{-1}(B + HX_0) = -H^{-1}B.$$

Отже, для квадратичної опуклої функції метод Ньютона дозволяє знайти її мінімум, незалежно від початкового наближення  $X_0$  і ступеня яружності, за один крок. Якщо ж  $f(X)$  є неквадратичною, але опуклою, тоді метод Ньютона гарантує її монотонне спадання від ітерації до ітерації.

Істотним недоліком методу Ньютона є залежність збіжності від початкової точки  $X_0$  для неопуклих функцій. Якщо  $X_0$  знаходиться далеко від точки мінімуму, то метод може давати розходження.

## 9.2 Метод Ньютона–Рафсона

Збіжність методу Ньютона, незалежно від початкового наближення, забезпечується вибором не тільки напрямку спуску  $P_k$ , але і величиною кроку  $\gamma_k$  уздовж цього напрямку. Відповідний алгоритм називають методом Ньютона з регулюванням кроку або методом Ньютона–Рафсона. Ітераційна процедура визначається виразом:

$$X_{k+1} = X_k - \gamma_k H^{-1}(X_k) \nabla f(X_k). \quad (9.2)$$

Величина кроку  $\gamma_k$  обирається з умови мінімуму  $f(X)$  за  $\gamma$  у напрямку спуску, тобто в результаті розв'язання задачі одновимірної мінімізації:

$$f(X_k - \gamma_k H^{-1}(X_k) \nabla f(X_k)) = \min_{\gamma > 0} f(X_k - \gamma H^{-1}(X_k) \nabla f(X_k)).$$

Внаслідок накопичення помилок у процесі розрахунку матриця Гессе на деякій ітерації може виявитися від'ємно визначеною або її неможна буде обернути, тоді вважають, що  $H^{-1}(X_k) = E$ . Ітерація при цьому буде здійснюватися за методом найшвидшого спуску, тобто  $P_k = -\nabla f(X_k)$ .

### Алгоритм методу

**Етап 1.** Вибір початкової точки  $X_0$ ,  $k = 0$ , обчислення градієнту  $\nabla f(X_k)$ .

**Етап 2.** Обчислення матриці Гессе  $H(X_k)$  й оберненої матриці  $H^{-1}(X_k)$ .

**Етап 3.** Обчислення напрямку спуску:

$$P_k = -H^{-1}(X_k) \nabla f(X_k).$$

**Етап 4.** Знаходження одним із методів одновимірної мінімізації оптимального кроку  $\gamma_k$  уздовж напрямку  $P_k$ .

**Етап 5.** Обчислення координат нової точки  $X_{k+1} = X_k + \gamma_k P_k$  і градієнту  $\nabla f(X_{k+1})$ .

**Етап 6.** Перевірка умови завершення пошуку. Якщо:

$$\|X_{k+1} - X_k\| < \varepsilon \text{ і } \|\nabla f(X_{k+1})\| < \nabla_{min},$$

переходимо до етапу 8, інакше – до етапу 7.

**Етап 7.** Приймають  $k = k + 1$ , переходять до етапу 2.

**Етап 8.** Приймають  $X^* = X_{k+1}$ ;  $f^* = f(X^*)$ .

Кількість обчислень на кожній ітерації методу Ньютона–Рафсона більша, ніж при застосуванні градієнтних методів. Проте для одержання однакової точності розв’язку методом Ньютона–Рафсона потрібно в десятки разів менше ітерацій. Але в деяких задачах трудомісткість обчислень матриці Гессе потребує великих витрат машинного часу. У цих випадках доцільно застосовувати комбінацію градієнтного методу і методу Ньютона–Рафсона. На початку процесу, коли  $X_0$  знаходиться далеко від точки мінімуму  $X^*$ , можна застосувати градієнтний метод. Далі при зменшенні швидкості збіжності градієнтного методу необхідно перейти до методу Ньютона–Рафсона.

### 9.3 Обчислення матриці Гессе

Для побудови матриці Гессе необхідно обчислити другі частинні похідні функції  $f(X)$ :

$$\frac{\partial^2 f(X)}{\partial x_i \partial x_j}, \quad i, j = \overline{1, n}.$$

У більшості практичних задач одержати частинні похідні функції в явному виді важко. Тому частинні похідні заміняють кінцевими різницями вигляду:

$$\begin{aligned} \frac{\partial^2 f(X)}{\partial x_i^2} &= \frac{f_{i+} - 2f + f_{i-}}{a_i^2}; \\ \frac{\partial^2 f(X)}{\partial x_i \partial x_j} &= \frac{f_{i+j+} - f_{i+j-} - f_{i-j+} + f_{i-j-}}{4a_i a_j}, \end{aligned} \quad (9.3)$$

де  $f = f(x_1, x_2, \dots, x_n)$ ;

$$f_{i\pm} = f(x_1, \dots, x_{i-1}, x_i \pm a_i, x_{i+1}, \dots, x_n);$$

$$f_{i\pm j\pm} = f(x_1, \dots, x_i \pm a_i, \dots, x_j \pm a_j, \dots, x_n);$$

$a_i, a_j$  – кроки сітки апроксимації за  $x_i, x_j$ .

Обернена матриця Гессе  $H^{-1}(X)$  визначається, виходячи зі співвідношення  $H^{-1}(X)H(X) = E$ . У результаті утворюється  $n$  систем лінійних алгебраїчних рівнянь з  $n$  невідомими, які розв'язуються методом Гауса.

#### 9.4 Методи безумовної оптимізації з перемінною метрикою

Методи оптимізації другого порядку потребують обчислення й обертання матриці Гессе  $H(X)$ . На практиці обчислення й обертання матриць вимагає багато часу. У той же час швидкість збіжності методів другого порядку дуже висока, тому розробляються методи оптимізації, що за швидкістю збіжності близькі до методів другого порядку, але є істотно менше трудомісткими. До їхнього числа відносяться методи перемінної метрики, що не потребують обчислення других похідних цільової функції.

Ітераційна процедура методів перемінної метрики визначається формулою:

$$X_{k+1} = X_k - \gamma_k D(X_k) \nabla f(X_k), \quad (9.4)$$

де  $\gamma_k$  – величина кроку в напрямку спуска;

$P_k = -D(X_k) \nabla f(X_k)$  – напрямок спуску;

$D(X_k)$  – апроксимація оберненої матриці  $H^{-1}(X_k)$ .

На початку ітераційного процесу при прямуванні з точки  $X_0$  за  $D(X_0)$  обирається одинична матриця  $E$ . На кожному такому кроці матриця  $H^{-1}(X_{k+1})$  апроксимується, використовуючи інформацію, отриману на попередньому кроці:

$$H^{-1}(X_{k+1}) \approx D(X_{k+1}) = D(X_k) + \Delta D(X_k),$$

де  $\Delta D(X_k)$  – матриця поправок, яка обирається таким чином, щоб  $D(X_{k+1})$  була додатно визначеною матрицею. У такому випадку методи перемінної матриці збігаються.

Для квадратичних функцій після  $n$  кроків матриці  $D(X_n)$  і  $H^{-1}(X_n)$  збігаються, для неквадратичних:  $D(X_k) \rightarrow H^{-1}(X_k)$ .

$$k \rightarrow \infty$$

Для обчислення матриці поправок  $\Delta D(X_k)$  використовуються різноманітні співвідношення, які дають різні методи.

Введемо позначення:

$$\Delta X_k = X_{k+1} - X_k; \Delta Y_k = \nabla f(X_{k+1}) - \nabla f(X_k), D_k = D(X_k).$$

Матриці поправок  $\Delta D(X_k)$  визначаються:

1) у проєктивному методі Ньютона–Рафсона (НР):

$$\Delta D(X_k) = \frac{(D_k \Delta Y_k)(D_k \Delta Y_k)^T}{(\Delta Y_k)^T D_k \Delta Y_k}; \quad (9.5)$$

2) у методі Давидона–Флетчера–Пауелла (ДФП):

$$\Delta D(X_k) = \frac{\Delta X_k (\Delta X_k)^T}{(\Delta X_k)^T \Delta Y_k} - \frac{(D_k \Delta Y_k)(D_k \Delta Y_k)^T}{(\Delta Y_k)^T D_k \Delta Y_k}; \quad (9.6)$$

3) у методі Бroyдена (Б):

$$\Delta D(X_k) = \frac{(\Delta X_k - D_k \Delta Y_k)(\Delta X_k - D_k \Delta Y_k)^T}{(\Delta X_k - D_k \Delta Y_k)^T \Delta Y_k}; \quad (9.7)$$

4) у методі Пірсона 2 (П2):

$$\Delta D(x_k) = \frac{(\Delta x_k - D_k \Delta y_k)(\Delta x_k)^T}{(\Delta x_k)^T \Delta y_k}; \quad (9.8)$$

5) у методі Пірсона 3 (П3):

$$\Delta D(X_k) = \frac{(\Delta X_k - D_k \Delta Y_k)(D_k \Delta Y_k)^T}{(\Delta Y_k)^T D_k \Delta Y_k}; \quad (9.9)$$

6) у методі Пройдена – Флетчера – Гольдфарба–Шано (БФГШ):

$$\Delta D(X_k) = \left[ 1 + \frac{(\Delta Y_k)^T D_k \Delta Y_k}{(\Delta X_k)^T \Delta Y_k} \right] \cdot \frac{\Delta X_k (\Delta X_k)^T}{(\Delta X_k)^T \Delta Y_k} - \frac{\Delta X_k (\Delta Y_k)^T D_k + D_k \Delta Y_k (\Delta X_k)^T}{(\Delta X_k)^T \Delta Y_k}. \quad (9.10)$$

У методах перемінної метрики величина кроку  $\gamma_k$  визначається шляхом одномірної мінімізації функції  $f(X)$  за  $\gamma$  у напрямку спуска, тобто:

$$f(X_k - \gamma_k D(X_k) \nabla f(X_k)) = \min_{\gamma > 0} f(X_k - \gamma D(X_k) \nabla f(X_k)).$$

### Алгоритм методів

**Етап 1.** Вибір початкової точки  $X_0$ ,  $k = 0$ , обчислення  $\nabla f(X_k)$ ,  $D(X_k) = E$ .

**Етап 2.** Обчислення напрямку спуска  $P_k = -D(X_k) \nabla f(X_k)$ .

**Етап 3.** Знаходження одним із методів одномірної мінімізації розміру кроку  $\gamma_k$  уздовж напрямку  $P_k$ .

**Етап 4.** Обчислення координат точки  $X_{k+1} = X_k + \gamma_k P_k$  і значення градієнту  $\nabla f(X_{k+1})$ .

**Етап 5.** Обчислення значень  $\Delta X_k$ ,  $\Delta Y_k$ .

**Етап 6.** Перевірка умов припинення пошуку. Якщо

$$\|\Delta X_k\| = \|X_{k+1} - X_k\| < \varepsilon \text{ і } \|\nabla f(X_{k+1})\| < \nabla_{min},$$

переходимо до етапу 10, інакше – до етапу 7.

**Етап 7.** Обчислення  $\Delta D(X_k)$  відповідно до методу.

**Етап 8.** Перетворення матриці  $D(X_k)$  на матрицю  $D(X_{k+1})$  за формулою:

$$D(X_{k+1}) = D(X_k) + \Delta D(X_k).$$

**Етап 9.** Приймають  $k = k + 1$ , переходять до етапу 2.

**Етап 10.** Приймають  $X^* = X_{k+1}$ ,  $f^* = f(X^*)$ .

При практичних розрахунках матриця  $D(X_k)$  через кожні  $k = n$  кроків береться рівною  $E$ . Протягом наступних ітерацій матриця  $D(X_k)$  буде невинродженою і додатно визначеною. Методи перемінної метрики, як і методи другого порядку, найбільше доцільно застосовувати в окрузі точки мінімуму.

### Питання для самоперевірки

1. Вкажіть основні методи безумовної оптимізації другого порядку.
2. Наведіть алгоритм методу Ньютона та його недоліки.
3. Вкажіть основні етапи застосування методу Ньютона–Рафсона.
4. Запишіть матрицю Гессе.

5. Наведіть форму матриці поправок у проєктивному методі Ньютона–Рафсона.
6. Запишіть форму матриці поправок у методі Давидона–Флетчера–Пауелла.
7. Наведіть форму матриці поправок у методі Бройдена.
8. Запишіть співвідношення для матриць поправок для методів Пірсона.



## 10 МЕТОДИ УМОВНОЇ ОПТИМІЗАЦІЇ

### 10.1 Класифікація методів

До цього ми вирішували задачу безумовної оптимізації. Повернемося до загальної постановки задачі параметричної оптимізації.

Знайти

$$\min_{X \in D_X} f(X), \quad (10.1)$$

де  $D_X = \{X/g_j(X) \geq 0, j = \overline{1, l},$

$H_j(X) = 0; j = \overline{1, m} \}$ , причому цільова функція й обмеження є нелінійними функціями керованих параметрів. Задача умовної оптимізації (10.1) є задачею нелінійного програмування (НП).

Чисельні методи розв'язання задач нелінійного програмування можна розділити на прямі й непрямі. Прямі методи оперують безпосередньо з вихідною задачею оптимізації (метод проєкції градієнта, метод Бокса). Непрямі методи перетворюють вихідну задачу НП (13.1) з обмеженнями на послідовність задач безумовної оптимізації шляхом урахування обмежень за допомогою штрафних функцій.

При цьому мінімізується допоміжна функція:

$$F(X, \tau) = f(X) + \Phi(X, \tau), \quad (10.2)$$

де  $\Phi(X, \tau)$  – функція штрафу (бар'єрна функція);

$\tau$  – коефіцієнт штрафу.

У залежності від виду функції  $\Phi(X, \tau)$  розрізняють методи внутрішніх і зовнішніх штрафних функцій.

## 10.2 Методи внутрішніх штрафних функцій

Використання методів внутрішніх штрафних функцій (методів внутрішньої точки) дозволяє утримувати траєкторію пошуку всередині допустимої області. Методи застосовуються для розв'язання задач мінімізації з обмеженнями-нерівностями, тобто:

$$D_X = \{X/g_j(X) \geq 0, j = \overline{1, l}\}. \quad (10.3)$$

Крім того, при їхньому використанні початкова точка  $X_0$  повинна належати середині області  $D_X$ , тобто  $X_0 \in D_X, X_0 \notin dD_X$ .

Функціями штрафу можуть бути обрані:

$$\Phi_1(X, \tau) = -\tau \sum_{j=1}^l \ln g_j(X); \quad (10.4)$$

$$\Phi_2(X, \tau) = \tau \sum_{j=1}^l \frac{1}{g_j(X)}. \quad (10.5)$$

З наближенням  $X$  до межі області  $D_X$  значення  $\Phi_1$  і  $\Phi_2$  прямують до нескінченності. Це явище ототожнюється з наявністю непереборного бар'єру на межі області (рис. 10.1).

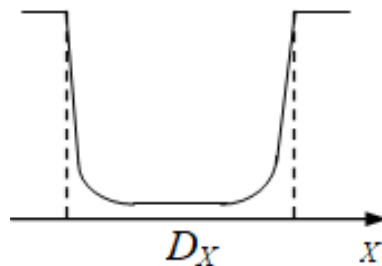


Рисунок 10.1 – Метод внутрішніх штрафних функцій

Вважають, що:

$$\Phi_2(X, \tau) = \begin{cases} \infty, & \text{якщо } X \notin D_X; \\ 0, & \text{якщо } X \in D_X, X \notin dD_X, \tau \rightarrow 0; \\ \infty, & \text{якщо } X \in D_X, X \rightarrow dD_X. \end{cases}$$

При використанні функцій  $\Phi_1$  і  $\Phi_2$  розв'язують послідовність задач безумовної оптимізації функції  $F(X, \tau)$  із значеннями  $\tau$ , що монотонно зменшуються на кожному етапі.

### 10.3 Методи зовнішніх штрафних функцій

У методах зовнішніх штрафних функцій (методах зовнішньої точки) функціями штрафу для обмежень типу (10.3) можуть бути обрані:

$$\Phi_3(X, \tau) = \frac{1}{\tau} \sum_{j=1}^l (\min[g_j(X), 0])^2, \quad (10.6)$$

$$\Phi_4(X, \tau) = \frac{1}{\tau} \sum_{j=1}^l \left( \frac{g_j(X) - |g_j(X)|}{2} \right)^2. \quad (10.7)$$

При використанні цих функцій не потрібно, щоб початкова точка  $X_0$  належала допустимій області  $D_X$ . Функції ведуть себе таким чином (рис. 10.2):

$$\Phi_3(X, \tau) = \begin{cases} 0, & \text{якщо } X \in D_X; \\ \infty, & \text{якщо } X \notin D_X. \end{cases}$$

Із зменшенням  $\tau$  (зростанням  $1/\tau$ ) будуть зближатися умовний мінімум  $f(X)$  і безумовний мінімум  $F(X, \tau)$ . При  $\tau \rightarrow 0$  ці мінімуми зливаються.

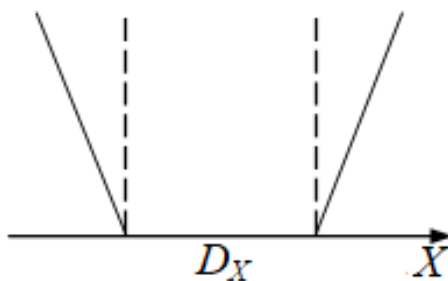


Рисунок 10.2 – Метод зовнішніх штрафних функцій

Варто зауважити, що якщо умовний мінімум функції  $f(X)$  лежить усередині області  $D_X$ , то він одночасно є і безумовним, тому що для  $X \in D_X$  завжди  $\min [g_j(X), 0] = 0$ .

## 10.4 Метод Фіакко й Мак-Корміка

Для задач із допустимою областю  $D_X$  у загальній постановці (10.1) використовуються комбіновані функції штрафу, наприклад:

$$\Phi_5(X, \tau) = \tau \sum_{j=1}^l \frac{1}{g_j(X)} + \frac{1}{\tau} \sum_{j=1}^m h_j^2(X), \quad (10.8)$$

$$\Phi_6(X, \tau) = \frac{1}{\tau} \sum_{j=1}^l (\min[g_j(X), 0])^2 + \frac{1}{\tau} \sum_{j=1}^m h_j^2(X). \quad (10.9)$$

Тоді допоміжна цільова функція буде мати вигляд:

$$F(X, \tau) = f(X) + \begin{cases} \Phi_5(X, \tau); \\ \Phi_6(X, \tau). \end{cases} \quad (10.10)$$

**Алгоритм методу** містить такі етапи:

**Етап 1.** Вибір початкової точки  $X_0$ , у випадку використання штрафної функції  $\Phi_5(X, \tau)$  для точки  $X_0$  повинна виконуватися нерівність  $g_j(X_0) > 0$ ,  $j = \overline{1, l}$ .

**Етап 2.** Приймаємо  $k = 1$ . Встановлення початкового значення коефіцієнта штрафу  $\tau_1$ . Якщо використовується штрафна функція  $\Phi_5$ , то можна покласти  $\tau_1 = 1$ , якщо  $\Phi_6$ , то  $\tau_1 = 0,01 \div 0,001$ .

**Етап 3.** Знаходження одним із відомих методів безумовної оптимізації мінімуму функції  $F(X, \tau)$ , тобто:

$$\min_X F(X, \tau_k) = F(X_k, \tau_k).$$

**Етап 4.** Перевірка умови завершення пошуку для отриманого на етапі 3 значення  $X_k$ . За критерій завершення пошуку приймається:

$$1) \|X_k - X_{k-1}\| < \varepsilon;$$

$$2) |f(X_k) - f(X_{k-1})| < \delta;$$

$$3) |F(X_k, \tau_k) - F(X_{k-1}, \tau_{k-1})| < \delta,$$

де  $\varepsilon$  і  $\delta$  – задані числа, які залежать від необхідної точності обчислень. Краще використовувати умови 1) і 2).

Якщо умови 1) і 2) виконуються, то пошук завершується, переходимо до етапу 7. У протилежному випадку виконують наступний етап.

**Етап 5.** Зменшення значення  $\tau_k$  у  $r$  раз:  $\tau_{k+1} = \tau_k/r$ . Звичайно приймають  $r = 5 \div 10$ .

**Етап 6.** Приймають  $k = k + 1$ , переходять до етапу 3.

**Етап 7.** Приймають  $X^* = X_k$ ,  $f^* = f(X^*)$ .

Загальним недоліком методів штрафних функцій є складність допоміжної функції  $F(X, \tau)$ , яка має яружну структуру. Причому ступінь яружності збільшується зі зменшенням параметра  $\tau$ . У той же час методи легко програмуються для ЕОМ і значно простіші, ніж прямі методи пошукової оптимізації.

## 10.5 Метод проєкції градієнту

Цей метод в основному застосовується при розв'язанні задач умовної оптимізації з обмеженнями типу рівностей, тобто задач вигляду:

$$\min_{X \in D_X} f(X), D_X = \{X / h_j(X) = 0, j = \overline{1, m}\}. \quad (10.11)$$

Умовний мінімум цільової функції знаходиться на гіперповерхні обмежень, тому що тільки тут виконуються умови  $H(X) = 0$ .

**Алгоритм методу проєкції градієнта** складається з двох етапів.

**Етап 1.** Повернення на гіперповерхню обмежень із поточної точки пошуку  $X_k$ , якщо ця точка вийшла за припустимі межі порушення обмежень  $\Delta H$  (рис. 10.3):  $\|H(X_k)\| \geq \Delta H$ . Таке переміщення відбувається із точки  $X_k$  за нормаллю до гіперповерхні обмежень (на рисунку 10.3 з точки  $X_0$  у точку  $X_1$ ).

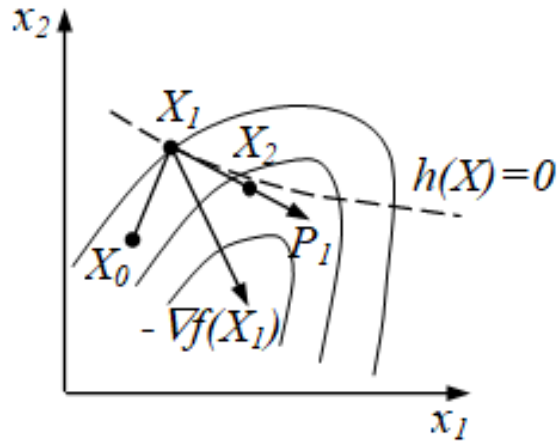


Рисунок 10.3 – Метод проєкції градієнта

Вектор приросту керованих параметрів визначається за формулою:

$$\Delta X_{k+1} = -D^k (D_k^k)^{-1} H(X_k),$$

де  $D_k$  – матриця розміром  $m \times n$ , рядками якої є градієнти функцій-обмежень  $h_j(X)$ , у точці  $X_k$

$$D_k = \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \dots & \frac{\partial h_1}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial h_m}{\partial x_1} & \dots & \frac{\partial h_m}{\partial x_n} \end{pmatrix}.$$

Таким чином, на етапі 1 одержимо  $X_{k+1} = X_k + \Delta X_{k+1}$ .

**Етап 2.** Переміщення убік умовного екстремуму уздовж гіперплощини, дотичної в точці  $X_k$  гіперповерхні обмежень. У площині двох параметрів це буде дотична до лінії обмеження. Напрямок пошуку мінімуму визначається за напрямком проєкції антиградієнта функції  $f(X)$  на цю гіперплощину:

$$P_k = -A_k \nabla f(X_k),$$

де  $A_k = E - D^k (D_k^k)^{-1} D_k$  – проєктувальна матриця.

Якщо величина кроку  $\gamma_k$  визначена, то нова точка  $X_{k+1}$  визначається згідно з формулою:

$$X_{k+1} = X_k + \gamma_k P_k = X_k - \gamma_k A_k \nabla f(X_k).$$

На рисунку 10.3 цей етап здійснюється з точки  $X_1$  у точку  $X_2$  уздовж напрямку  $P_1$ . Крок уздовж лінії обмеження призводить до його порушення, тому наступним кроком буде етап спуска на лінію обмежень і т. д.

Напрямок проєкції антиградієнта є найкращим локальним напрямком прямування до умовного мінімуму.

Модифікації викладеного методу дозволили одержати сімейство проєктивних методів або градієнтних методів із великою довжиною кроку. Зокрема, розроблені методи, що розв'язують задачу мінімізації при обмеженнях-нерівностях  $G(X) \geq 0$ .

## 10.6 Комплексний пошук Бокса

Цей метод є модифікацією методу деформованого багатогранника.

Він призначений для розв'язання задачі НП з обмеженнями-нерівностями:

$$\begin{aligned} \min_{X \in D_X} f(X), \quad D_X = \{X / g_j(X) \geq 0, j = \overline{1, l}, \\ x_{i \min} \leq x_i \leq x_{i \max}, i = \overline{1, n}\}. \end{aligned} \quad (10.12)$$

Для мінімізації  $f(X)$  у  $n$ -мірному просторі будують многогранники, що містять  $q \geq n + 1$  вершин (рекомендується  $q = 2n$ ), такі многогранники називають **комплексами**, що й обумовило назву методу.

Введемо такі позначення:

$X_k^s = (x_{1k}^s, x_{2k}^s, \dots, x_{nk}^s)$ ,  $s = \overline{1, q}$ ,  $k = 0, 1, \dots$  –  $s$ -та вершина комплексу на  $k$ -му етапі пошуку,

$X_k^r$  – вершина з максимальним значенням функції, тобто:

$$f(X_k^r) = \max\{f(X_k^1), \dots, f(X_k^q)\};$$

$X_k^c$  – центр ваги усіх вершин, за винятком  $X_k^r$ .

Координати центру ваги обчислюються за формулою:

$$x_{ik}^c = \frac{1}{q-1} \left( \sum_{s=1}^q x_{ik}^s - x_{ik}^r \right), \quad i = \overline{1, n}. \quad (10.13)$$

**Алгоритм** комплексного пошуку

**Етап 1.** Обирається деяка допустима точка  $X_0^l$  за першу вершину початкового комплексу. Координати інших  $q - 1$  вершин комплексу визначаються співвідношенням:

$$x_{io}^j = x_{i \min} + \zeta_i (x_{i \max} - x_{i \min}), \quad i = \overline{1, n}, \quad j = 1, q - 1,$$

де  $\zeta_i$  – випадкове число, рівномірно розподілене в інтервалі  $[0, 1]$ .

Отримані таким шляхом точки задовольняють прямим обмеженням, проте функціональні обмеження можуть не виконуватися.

**Етап 2.** У випадку порушення функціональних обмежень недопустима точка замінюється новою, яка лежить посередині відрізка, що з'єднує недопустиму точку з центром ваги обраних допустимих вершин. Дана операція повторюється доти, поки не виповнюються всі обмеження задачі.

**Етап 3.** Визначається вершина  $X_k^r$  з найбільшим значенням функції, а також координати центру ваги  $X_k^c$  інших вершин комплексу.

**Етап 4.** Вершина  $X_k^r$  відбивається щодо центру ваги  $X_k^c$  в нову точку  $X_k^H$ . Точка  $X_k^H$  визначається за формулою:

$$X_k^H = (\alpha + 1)X_k^c + \alpha X_k^r, \quad (10.14)$$

де  $\alpha > 0$  – коефіцієнт ( $\alpha = 1, 3$ ).

**Етап 5.** Якщо  $f(X_k^H) > f(X_k^r)$ , то нова вершина стає гіршою вершиною комплексу. У цьому випадку  $\alpha$  зменшують у два рази.

Якщо в результаті відбитку порушується якесь обмеження, то коефіцієнт  $\alpha$  зменшують щораз удвічі, поки  $X_k^H$  не стане допустимою.

**Етап 6.** Якщо значення функції мало змінюється протягом п'ятих послідовних ітерацій, тобто:

$$\left| f(X_{k+1}^c) - f(X_k^c) \right| < \varepsilon, \quad k = \overline{1, 5},$$

то розрахунки завершують і центр ваги комплексу вважають розв'язком задачі  $X^* = X_{k+1}^c$ . У протилежному випадку повертаються до етапу 3.



## 10.7 Метод Моррісона

Метод призначений для розв'язання задачі НП з обмеженнями-нерівностями. Складемо функцію:

$$F(X, \eta) = (f(X) - \eta)^2 + \sum_{j=1}^l [\min(0, g_j(X))]^2. \quad (10.15)$$

Величина  $\eta$  визначається за алгоритмом:

$$\begin{cases} \eta_{k+1} = \eta_k + \sqrt{F(X_k, \eta_k)}; \\ F(X_k, \eta_k) = \min_X F(X, \eta_k). \end{cases} \quad (13.16)$$

Передбачається, що можна задати початкове значення  $\eta_0$  таке, що  $\eta_0 \leq f(X^*)$ , де  $f(X^*)$  – нижня оцінка задачі умовної мінімізації.

Розв'язання припиняється, якщо виконується одна з умов:

- 1)  $\|X_{k+1} - X_k\| < \varepsilon$ ;
  - 2)  $f(X_k) - \eta_k < \varepsilon(1 + |\eta_k|)$ ;
  - 3)  $\eta_{k+1} - \eta_k < \varepsilon(1 + |\eta_k|)$ .
- (10.17)

У модифікованому методі Моррісона параметр  $\eta$  визначається за алгоритмом:

$$\eta_{k+1} = \eta_k + F(X_k, \eta_k) / (f(X_k) - \eta_k). \quad (10.18)$$

У цьому випадку забезпечується більша швидкість збіжності.

Для забезпечення роботи обох методів необхідним є виконання умови  $\eta_k \leq f(X^*)$ . Для цього потрібно проводити безумовну оптимізацію методами, що гарантують високу точність або задавати більш високу точність  $\varepsilon$ .

Різниця методів штрафних функцій і Моррісона полягає в тому, що в методі Моррісона зміна параметра штрафу  $\eta$  відбувається автоматично, а в методі штрафних функцій необхідно спеціально задавати послідовність  $\tau_k$ .

## 10.8 Особливості пошуку оптимуму при використанні максимінних критеріїв

Максимінна постановка задач параметричної оптимізації – одна з найбільш перспективних при проектуванні технічних об'єктів. Це пояснюється тим, що в результаті оптимізації поліпшуються запаси працездатності практично усіх вихідних параметрів об'єкта. Проте розв'язання задачі в такій постановці має свої особливості. Насамперед, це стосується умов формування цільової функції, яка має вигляд:

$$f(X) = \min_{1 \leq j \leq m} z_j(X), \quad (10.19)$$

де  $z_j(X)$  – запас працездатності параметра  $y_j(X)$ , що обчислюється за формулою:

$$z_j = \omega_j \left( \frac{TT_j - y_{j \text{ ном}}}{\delta_j} - 1 \right). \quad (10.20)$$

Функція  $f(X)$  підлягає максимізації. Ця функція не є гладкою, у результаті чого більшість методів першого й другого порядків для пошуку максимуму  $f(X)$  безпосередньо не застосовні.

Функція мінімуму в процесі пошуку формується в такий спосіб. У кожній точці на траєкторії пошуку знаходяться вихідні параметри  $y_j$  об'єкта, який оптимізується, і за ними визначаються оцінки  $z_j$ . Серед усіх оцінок у точці  $X_k$  виділяють мінімальну, наприклад, це буде оцінка  $z_p$ . У цьому випадку  $f(X) = z_p(X)$ .

До моменту зміни оцінки  $z_p$  на другу, гіперповерхня відгуку цільової функції передбачається достатньо гладкою, що дозволяє організувати процедуру максимізації  $f(X)$  одним із методів безумовної оптимізації.

Відбувається поліпшення оцінки  $z_p$ . Проте в силу конфліктності вихідних параметрів поліпшення оцінки  $z_p$  призведе до погіршення інших оцінок. Тому на деякому етапі пошуку мінімальною стане інша оцінка, наприклад  $z_q$ . Така зміна мінімальних оцінок, може спостерігатися в різних умовах (рис. 10.4).

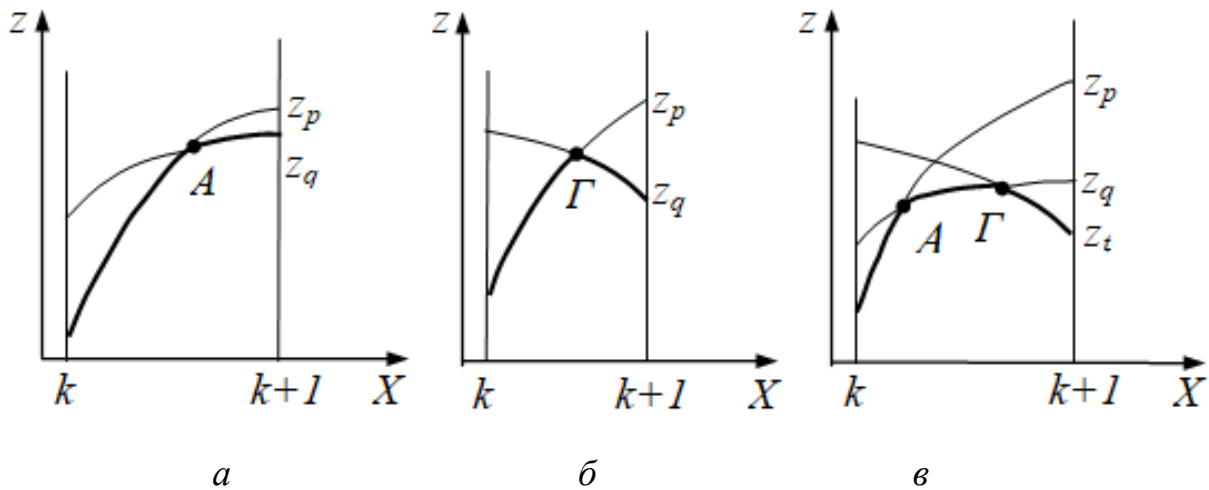


Рисунок 10.4 – Пошук оптимальності при використанні максиміальних критеріїв

На рисунку 10.4, а) наведений випадок, коли на  $(k + 1)$ -ому кроці пошуку цільовою буде фігурувати функція  $f(X) = z_q(X)$ . При цьому оцінки  $z_p$  і  $z_q$  мають тенденцію до збільшення, і зміну цільової функції можна трактувати як перехід з однієї гіперповерхні пошуку на другу. Тут застосовні методи безумовної оптимізації.

Аналіз поведінки оцінок на рисунку 10.4, б) показує, що обмежитися простою зміною оцінок у цільовій функції неможливо. Шуканий максимум повинен знаходитися на гіперповерхні перетинів двох гіперповерхонь  $z_p(X)$  і  $z_q(X)$ . Ця гіперповерхня перетину є гребенем гіперповерхні цільової функції  $f(X)$ . Тоді  $\Gamma$  – точка гребеня функції мінімуму.

У цьому випадку безумовний пошук доцільно замінити на умовний, вводячи обмеження типу рівності  $z_p(X) = z_q(X)$ , а в якості цільової функції можна розглядати попередню (стару) оцінку, тобто  $f(X) = z_p(X)$ .

У ситуації на рисунку 10.4, в) гребінь формується перетином гіперповерхонь нових оцінок  $z_q(X)$  і  $z_t(X)$ . Тому в алгоритмі пошуку проводиться зміна цільової функції  $f(X) = z_q(X)$  і вводиться обмеження  $z_q(X) - z_t(X) = 0$ . На рисунку 10.5 показані лінії рівня двох функцій  $z_1(X)$  і  $z_2(X)$  і точками виділена лінія  $z_1(X) = z_2(X)$ , що є гребенем цільової функції  $f(X)$ . По обидві сторони від гребеня значення  $f(X)$  зменшуються.

Отже, аналіз умов формування цільової функції вказує на те, що функція мінімуму  $f(X)$  не може бути гладкою. У точках гребенів  $f(X)$  не є диференційованою, тому більшість розглянутих методів безпосередньо не застосовні.

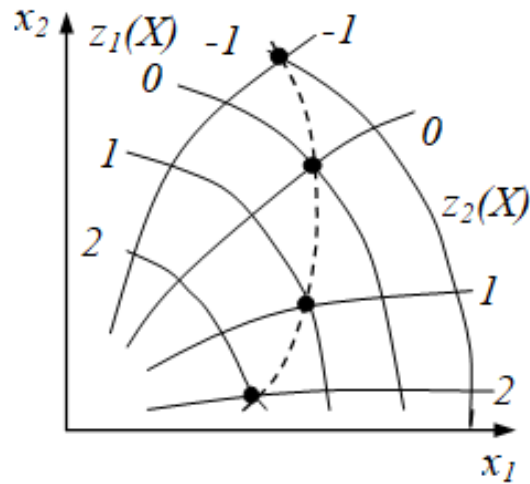


Рисунок 10.5 – Лінії рівня двох функцій

Один із найбільш ефективних алгоритмів розв’язання задач у максиміній постановці заснований на застосуванні методу проєкції градієнту. Використання цього методу базується на можливості формулювання рівняння гребенів у вигляді рівності конфліктних оцінок  $z_p(X) = z_q(X)$ , що розглядаються як обмеження задачі.

Наприклад, у випадку б) умовна оптимізація буде здійснюватися відповідно до методу проєкції градієнту цільової функції  $f(X) = z_p(X)$  на гіперповерхню обмежень  $z_p(X) - z_q(X) = 0$ .

### Питання для самоперевірки

1. Вкажіть основні методи умовної оптимізації.
2. Наведіть основні відмінності у методах внутрішніх і зовнішніх штрафних функцій.
3. Вкажіть, для яких задач є доцільним застосування комплексного пошуку Бокса.
4. Вкажіть особливості застосування й основні розрахункові формули методу Фіакко і Мак-Корміка.
5. Вкажіть основні етапи застосування методу проєкції градієнта.
6. Наведіть алгоритм використання методу Моррісона.
7. Вкажіть особливості пошуку оптимуму при використанні максиміальних критеріїв.

# 11 ЛІНІЙНЕ ПРОГРАМУВАННЯ

## 11.1 Постановка задачі лінійного програмування

Дуже широкий клас задач параметричної оптимізації складають задачі, у яких цільова функція виражається лінійно через керовані параметри, а обмеження записуються у вигляді лінійних рівностей або нерівностей.

Обчислення екстремуму лінійної функції (форми) при наявності лінійних обмежень на параметри, які оптимізуються, складає предмет лінійного програмування (ЛП).

Розглянемо задачу про суміші.

**Задача 1.** Для виготовлення заданого складу сплаву із свинцю, цинку й олова використовується сировина у виді п'ятьох матеріалів із тих же компонентів, що відрізняються складом і вартістю одного кілограму (табл. 11.1).

Таблиця 11.1 – Вхідні дані

Компоненти \ Матеріал	Склад, %				
	1	2	3	4	5
Свинець	10	10	40	60	30
Цинк	10	30	50	30	20
Олово	80	60	10	10	50
Вартість 1 кг, грн	4	4,5	5,8	6,2	7

Визначити, скільки потрібно взяти матеріалу кожного виду, щоб виготовити з мінімальною собівартістю новий сплав, що містить 50 % свинцю, не менше 20 % цинку і не більше 40 % олова.

### Побудова математичної моделі

Сформулюємо математичну модель задачі. Необхідно знайти значення змінних  $x_1, x_2, x_3, x_4, x_5$ , тобто план змішування вихідної сировини, за якого досягає мінімуму лінійна функція:

$$Z = 4x_1 + 4,5x_2 + 5,8x_3 + 6,2x_4 + 7x_5,$$

що являє собою загальну вартість суміші за таких лінійних обмежень:

$$\begin{cases} 10x_1 + 10x_2 + 40x_3 + 60x_4 + 30x_5 = 50, \\ 10x_1 + 30x_2 + 50x_3 + 30x_4 + 20x_5 \geq 20, \\ 80x_1 + 60x_2 + 10x_3 + 10x_4 + 50x_5 \leq 40, \\ x_1 + x_2 + x_3 + x_4 + x_5 = 1, \\ x_j \geq 0, \quad j = \overline{1,5}. \end{cases}$$

Умова невід'ємності змінних є природною та впливає з визначення цих змінних і підлягає урахуванню при розв'язанні цієї задачі.

**У загальному випадку задача ЛП формулюється так:** знайти значення змінних  $x_1, x_2, \dots, x_n$ , тобто вектор  $X = (x_1, x_2, \dots, x_n)$ , що мінімізують або максимізують лінійну функцію:

$$Z = c_1x_1 + c_2x_2 + \dots + c_nx_n = \sum_{j=1}^n c_jx_j \quad (11.1)$$

за таких обмежень

$$\begin{cases} \sum_{j=1}^n a_{ij}x_j \otimes b_i, \quad i = \overline{1,m}; \\ x_j \geq 0, \quad j = \overline{1,p}, \quad (p \leq n), \end{cases} \quad (11.2)$$

де  $\otimes$  означає один із можливих знаків відношення  $\leq, =, \geq$ .

Як відомо, задачу максимізації завжди можна звести до задачі мінімізації в силу співвідношення  $\max Z = -\min(-Z)$ .

Якщо на знак деякої змінної  $x_j$  обмеження відсутнє, то її можна записати як різницю двох невід'ємних змінних:

$$x_j' \geq 0 \text{ і } x_j'' \geq 0; \quad x_j = x_j' - x_j''.$$

Якщо в обмеженнях (11.2) є нерівності  $\leq$  або  $\geq$ , то шляхом введення додаткової або вільної змінної з коефіцієнтом  $+1$  або  $-1$  у ліву частину нерівності їх можна перетворити на рівність. Це дозволяє привести будь-яку задачу ЛП до канонічної форми.

**Визначення 1.** Канонічною або стандартною формою задачі ЛП називається задача мінімізації лінійної функції за умови невід'ємності всіх змінних і основних обмеженнях-рівностях.

Математична модель лінійної задачі в канонічній формі має вигляд:

$$Z = \sum_{j=1}^n c_j x_j \rightarrow \min ;$$

$$\begin{cases} \sum_{j=1}^n a_{ij} x_j = b_i, & i = \overline{1, m}; \\ x_j \geq 0, & j = \overline{1, n}. \end{cases} \quad (11.3)$$

Введемо такі вектори:

$$C = (c_1, c_2, \dots, c_n); X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}; A_0 = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}; A_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Тоді задачу (11.3) можна записати у векторній формі:

$$Z = CX \rightarrow \min, \sum_{j=1}^n A_j x_j = A_0, x_j \geq 0, j = \overline{1, n} \quad (11.4)$$

або в матричній формі:

$$Z = CX \rightarrow \min, AX = X_0, X \geq 0.$$

## 11.2 Властивості розв'язків задачі ЛП

**Визначення 2.** Допустимим розв'язком (планом) задачі ЛП називається вектор  $X = (x_1, x_2, \dots, x_n)$ , який задовольняє всім її обмеженням.

**Визначення 3.** План  $X = (x_1, \dots, x_n)$  задачі ЛП називається опорним (базисним), якщо система векторів умов  $A_j$ , що входять до розкладання вектору  $A_0 = \sum_{j=1}^n A_j x_j$  з додатними коефіцієнтами  $x_j$ , є лінійно незалежною.

Оскільки вектори  $A_j \in m$ -мірними, то кількість додатних компонент опорного плану не може перевищувати  $m$ .

**Визначення 4.** Опорний план називається не виродженим, якщо він містить рівно  $m$  додатних компонент. У іншому випадку план є виродженим.

**Визначення 5.** План  $X^* = (x_1^*, x_2^*, \dots, x_n^*)$  задачі ЛП, що оптимізує цільову функцію, називається оптимальним.

**Визначення 6.** Базисом опорного плану називається система з  $m$  лінійно незалежних векторів  $A_j$ , які відповідають додатним компонентам опорного плану.

**Визначення 7.** Компоненти опорного плану, які відповідають векторам його базису, називаються базисними, інші – небазисними змінними.

Розв’язання задачі ЛП засновано на таких твердженнях.

1. Множина планів  $L$  задачі ЛП опукла і являє собою опуклу багатогранну множину.

2. Множина планів задачі ЛП завжди містить у собі скінченну кількість граничних точок.

3. Цільова функція задачі ЛП досягає екстремуму в граничній точці.

4. Якщо лінійна функція досягає екстремуму в декількох крайніх точках, то вона досягає такого ж значення в будь-якій точці, що є опуклою комбінацією зазначених граничних точок  $X = X_1\lambda + (1 - \lambda) X_2$ .

5. Кожній вершині багатогранника (граничній точці) відповідає опорний план задачі ЛП і, навпаки, кожному опорному плану відповідає вершина багатогранника, тобто оптимальний план збігається з однією із вершин багатогранника.

Тому в процесі розв’язання задачі ЛП достатньо досліджувати тільки опорні плани (граничні точки), верхня межа яких обмежена числом  $C_n^m$ .

### 11.3 Графічний метод розв’язання задачі ЛП

Розв’язання задачі ЛП (11.1)–(11.2) полягає у визначенні екстремуму лінійної функції на деякій опуклій множині. Якщо ця множина є обмеженою, тоді вона являє собою опуклий багатогранник.

Для лінійної функції всі лінії рівня:

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



є паралельними і являють собою прямі лінії, площини або гіперплощини ( $n \geq 4$ ). Такі гіперплощини використовуються для знаходження в області  $L$  екстремальних точок цільової функції і називаються допустимими. Очевидно, екстремальними точками області  $L$  будуть ті точки, що відповідають граничним положенням допустимої гіперплощини (рис. 11.1). У кожній точці  $X$  градієнт лінійної функції  $\nabla Z = (c_1, c_2, \dots, c_n) = C$  сталий і збігається з вектором, складеним із коефіцієнтів даної функції  $Z$  при змінних і перпендикулярний лінії її рівня.

Для знаходження граничних положень допустимої гіперплощини потрібно побудувати довільну лінію рівня цільової функції (наприклад,  $Z = 0$ ) і, пересуваючи її паралельно самій собі в напрямку градієнту або антиградієнту, визначити екстремальні значення (рис. 11.1).

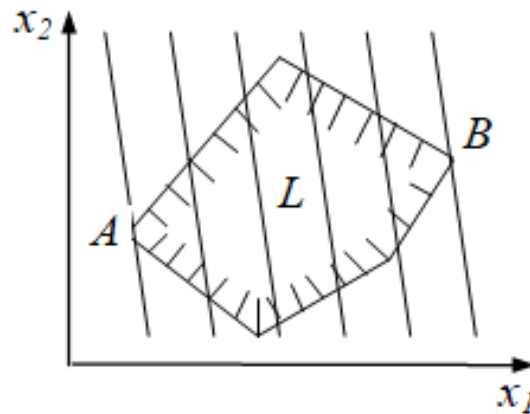


Рисунок 11.1 – Допустима гіперплощина

Можуть бути різноманітні варіанти взаємного розташування допустимої площини й многогранника. Допустима гіперплощина в одному із своїх граничних положень може виявитися паралельною ребру або грані (рис. 11.2, а), б)). Тоді екстремум лінійної форми досягається у всіх точках ребра або грані.

Якщо множина  $L$  є необмеженою, то цільова функція може виявитися необмеженою згори або знизу (рис. 11.2, б), в)). Може зустрітися випадок, коли обидва екстремуми не існують (рис. 11.2, з)).

Таким чином, графічний метод розв'язання задачі ЛП складається з таких етапів.

**Етап 1.** Побудувати область допустимих розв'язків  $L$ . Для цього спочатку необхідно побудувати область розв'язків кожного обмеження-нерівності задачі і знайти перетин цих областей.

**Етап 2.** Побудувати градієнт функції  $C = (c_1, \dots, c_n)$  і лінію рівня  $Z = const$ .

**Етап 3.** Пересуваючи лінію рівня паралельно самій собі, визначити її граничні положення. У тих точках  $L$ , на яких досягаються ці граничні положення, і буде екстремум цільової функції.

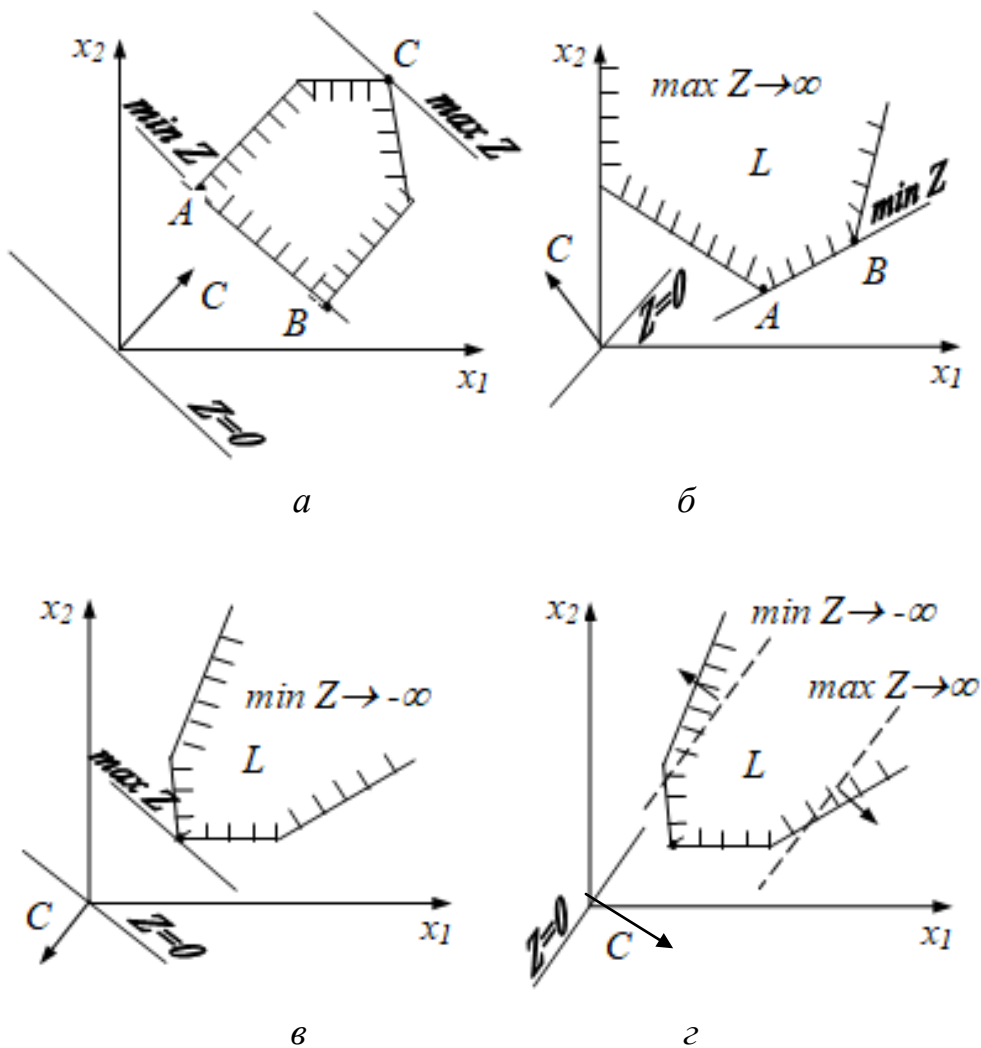


Рисунок 11.2 – Допустима гіперплощина

**Задача 2.** Знайти максимум і мінімум функції  $Z = 2x_1 + x_2$  за обмежень:

$$\begin{cases} -2x_1 + 3x_2 \leq 12; & \text{(I)} \\ 3x_1 + 2x_2 \leq 21; & \text{(II)} \\ x_1 - x_2 \leq 2; & \text{(III)} \\ x_1 + x_2 \geq 2; & \text{(IV)} \\ x_1 \geq 0, x_2 \geq 0. \end{cases}$$

Будуємо області  $L$  допустимих рішень задачі. Кожне з обмежень-нерівностей визначає півплощину, перетин яких визначає багатокутник (рис. 11.3). Для знаходження екстремумів будуємо допустиму пряму, прирівнюючи цільову функцію до нуля,  $Z = 2x_1 + x_2 = 0$ .

Будуємо градієнт цільової функції  $C = (2, 1)$ .

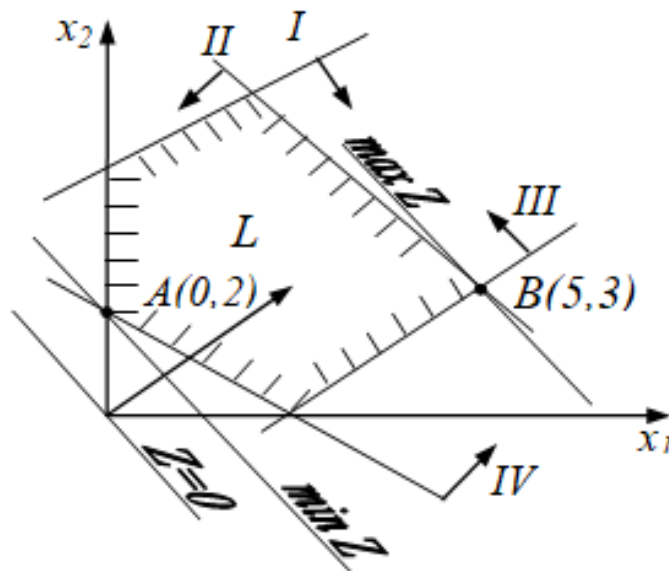


Рисунок 11.3 – Область  $L$  допустимих розв'язків задачі

Максимальне значення функція  $Z$  приймає у вершині  $B$ .

Координати точки знайдемо, як перетин відповідних прямих.

Таким чином, маємо:

$$B(5, 3) = X_{max} = (5, 3), \max Z = 13 \text{ та } A(0, 2) = X_{min} = (0, 2), \min Z = 2.$$

**Відповідь:**  $\max Z = Z(5, 3) = 13$ ,  $\min Z = Z(0, 2) = 2$ .

**Зауваження.** У випадку обмеженості  $L$  можна знайти координати всіх крайніх точок і порівняти в них значення лінійної форми.

## 11.4 Симплексний метод розв'язання задачі ЛП

Цей метод інакше називається методом послідовного поліпшення плану. Перед розв'язанням задачі ЛП симплексним методом її необхідно записати в канонічній формі:

$$Z = \sum_{j=1}^n c_j x_j \rightarrow \min,$$
$$\begin{cases} \sum_{j=1}^n a_{ij} x_j = b_i, & i = \overline{1, m}; \\ x_j \geq 0, & j = \overline{1, n}; \end{cases} \quad \text{або} \quad \begin{cases} \sum_{j=1}^n A_j x_j = A_0; \\ x_j \geq 0; & j = \overline{1, n}. \end{cases} \quad (11.5)$$

**Симплексний метод** складається з трьох основних етапів.

**Етап 1.** Визначення однієї з вершин багатогранника  $L$ , тобто вихідного опорного плану.

**Етап 2.** Перевірка опорного плану на оптимальність. Якщо план неоптимальний, переходять до наступного етапу.

**Етап 3.** Упорядковане перебирання вершин  $L$ , тобто опорних планів, шляхом викреслювання з базису визначеної змінної і запровадження замість неї нової. Повернення до етапу 2.

Кількість кроків у симплекс-методі складає не більше  $2m$ . На кожному кроці будується новий опорний план, якому відповідає менше значення лінійної форми, при цьому прямування до оптимальної вершини здійснюється за сусідніми вершинами області  $L$ .

У результаті багаторазового повторення етапів 2 і 3 буде або отримано оптимальний розв'язок, або виявлена необмеженість цільової функції, або встановлена несумісність умов ( $L = \emptyset$ ) і, відповідно, нерозв'язність задачі.

Процедура перерахунку коефіцієнтів при переході до нового базису формалізована і зводиться до заповнення стандартних симплексних таблиць.

## 11.5 Метод побудови опорних планів

Нехай  $X_0 = (x_1, \dots, x_n)$  опорний план канонічної задачі ЛП, якому відповідає  $m$ -мірний базис  $B_x = (A_{s1}, A_{s2}, \dots, A_{sm})$ .

Розв'язку  $X_0$  відповідає рівність:

$$\sum_{i=1}^m x_{si} A_{si} = A_0, \quad (11.6)$$

де усі  $x_{si} > 0$ .

Припустимо, що задача ЛП має більше, ніж один опорний план.

Спробуємо одержати, виходячи з  $X_0$ , новий опорний план.

Візьмемо довільний вектор умов  $A_j$  і розкладемо його за базисом  $B_x$

$$A_j = \sum_{i=1}^m x_{ij} A_{si}, \quad j = \overline{1, n}. \quad (11.7)$$

Припустимо, що для деякого вектора  $A_k$ , що не входить до вихідного базису  $B_x$ , хоча б один із коефіцієнтів  $x_{ik} > 0$ . Тоді:

$$A_k = \sum_{i=1}^m x_{ik} A_{si}.$$

Помножимо останній вираз на деяке невід'ємне число  $t$  і результат віднімемо з (11.7). Тоді одержимо:

$$\sum_{i=1}^m (x_{si} - tx_{ik}) A_{si} + tA_k = A_0. \quad (11.8)$$

Це розкладання вектора  $A_0$  визначає вектор  $X_1 = (x_{s1} - tx_{1k}, \dots, x_{sm} - tx_{mk}, t)$ , який у випадку невід'ємності його компонент є допустимим розв'язком задачі. Необхідно обрати  $t > 0$  таким чином, щоб  $x_{si} - tx_{ik} \geq 0$  для всіх  $i = \overline{1, m}$ .

Якщо для деякого  $i$  коефіцієнт  $x_{ik} \leq 0$ , то при будь-якому  $t > 0$  різниця буде додатною величиною. Тому розглянемо тільки  $x_{ik} > 0$ .

З нерівності  $x_{si} - tx_{ik} \geq 0$  маємо  $x_{s1} \geq tx_{ik}$ , тобто  $0 < t < \frac{x_{si}}{x_{ik}}$ .

Тоді за будь-яких значень  $t$ , що задовольняють нерівності:

$$0 < t < \min_i \frac{x_{si}}{x_{ik}},$$

вектор  $X_l$  буде допустимим розв'язком задачі ЛП, що містить  $m + 1$  додатну компоненту, але не є базисним. Щоб одержати новий допустимий базисний план, необхідно принаймні одну компоненту вектора  $X_l$  перетворити на нуль. Для цього треба за  $t$  обрати вираз:

$$t = \min_{x_{ik} > 0} \frac{x_{si}}{x_{ik}}.$$

Відношення  $x_{si}/x_{ik}$  називається симплексним. Нехай:

$$t = \min_{x_{ik} > 0} \frac{x_{si}}{x_{ik}} = \frac{x_{sr}}{x_{rk}}. \quad (11.9)$$

Для того щоб одержати базис опорного плану  $X_l$ , необхідно ввести в базис  $B_x$  вектор  $A_k$  і вивести з базису вектор  $A_{sr}$ . При цьому компоненти нового опорного плану  $X_l$  визначаються зі співвідношень:

$$\begin{cases} x'_{si} = x_{si} - \frac{x_{sr}}{x_{rk}} x_{ik}, & i \neq r; \\ x'_r = \frac{x_{sr}}{x_{rk}}, & i = r. \end{cases} \quad (11.10)$$

Новому допустимому базисному розв'язку  $X_l$  відповідає розкладання вектора правих частин, аналогічне (11.6):

$$A_0 = \sum_{i \neq r} x'_{si} \cdot A_{si} + x'_r A_k.$$

Аналогічно можна одержати розкладання для вектора  $A_j$ :

$$A_j = \sum_{i \neq r} x'_{ij} A_{si} + x'_{nj} A_k, \quad j = \overline{1, n}.$$

При цьому коефіцієнти розкладання можна обчислити за формулою:

$$\begin{cases} x'_{ij} = x_{ij} - \frac{x_{ri}}{x_{rk}} x_{ik}, & i = \overline{1, m}; \quad i \neq r_j; \quad j = \overline{1, n}; \\ x'_{rj} = \frac{x_{rj}}{x_{rk}}, & i = r; \quad j = \overline{1, n}. \end{cases} \quad (11.11)$$

Таким чином, процес одержання нових опорних планів складається з таких етапів.

**Етап 1.** Вибір вектора  $A_k$ , який вводиться до базису  $B_x$  відомого опорного плану.

**Етап 2.** Визначення вектора  $A_{sr}$ , виведеного з вихідного базису.

**Етап 3.** Використання (11.9) і формул повного перетворення (11.10), (11.11) для забезпечення відповідності одержаного базису новому опорному плану.

## 11.6 Знаходження оптимального розв'язку

Отже, нами викладений метод одержання допустимих базисних розв'язків (опорних планів). Але бажано вміти в якості наступного обирати такий опорний план, для якого значення цільової функції стає ближчим до оптимального, тобто зменшується.

Припустимо, що кожний опорний план є не виродженим. Нехай  $X_0 = (x_{s1}, \dots, x_{sm})$  – деякий опорний план, якому відповідає розкладання вектора правих частин  $A_0$  у вигляді:

$$A_0 = \sum_{i=1}^m x_{si} A_{si} .$$

і значення цільової функції:

$$Z_0 = \sum_{i=1}^m c_{si} x_{si} , \quad (11.12)$$

де усі  $x_{si} > 0$ . Розкладання вектора  $A_j$  за векторами базису  $A_{si}$  має вигляд (11.7):

$$A_j = \sum_{i=1}^m x_{ij} A_{si}, j = \overline{1, n} .$$

Йому відповідає значення лінійної форми:

$$Z_j = \sum_{i=1}^m c_{si} x_{ij}, \quad j = \overline{1, n}. \quad (11.13)$$

Складемо різницю:

$$\Delta_j = Z_j - c_j = \sum_{i=1}^m c_{si} x_{ij} - c_j, \quad j = \overline{1, n}. \quad (11.14)$$

Величину  $\Delta_j$  називаються оцінками векторів умов  $A_j$  щодо даного базису.

Помножимо (11.13) на  $t > 0$  і віднімемо результат із (11.12). Після перетворення одержимо:

$$Z_0 - t(Z_j - c_j) = \sum_{i=1}^m (x_{si} - tx_{ij}) c_{si} + tc_j. \quad (14.15)$$

Оскільки усі  $x_{si} > 0$ , то знайдеться таке  $t > 0$ , при якому коефіцієнти розкладання в (11.15) визначають новий опорний план.

Тому значення лінійної форми буде залежати від різниці  $\Delta_j = Z_j - c_j$ . Якщо  $\Delta_j > 0$ , то значення лінійної форми зменшується; якщо  $\Delta_j < 0$  – збільшується; якщо  $\Delta_j = 0$ , лінійна форма не змінюється.

Звідси має місце така **ознака оптимальності**:

1. Опорний план  $X^*$  є оптимальним розв'язком задачі ЛП, якщо  $\Delta_j \leq 0$  для всіх  $j = \overline{1, n}$ .

2. Якщо для деякого  $j$  величина  $\Delta_j > 0$  і усі відповідні цьому індексу  $x_{ij} \leq 0$ ,  $i = \overline{1, m}$ , тоді лінійна форма є необмеженою на множині планів задачі, і задача ЛП вважається нерозв'язною.

3. Якщо для деяких індексів  $j$  величина  $\Delta_j > 0$  і для кожного такого  $j$  принаймні одне з чисел  $x_{ij} > 0$ , то існує новий опорний план, для якого значення лінійної форми зменшується, що є більш близьким до оптимального.

## 11.7 Алгоритм симплексного методу

Перед розв'язанням задачі ЛП симплексним методом її необхідно представити в канонічній формі. Розв'язання починається з відомого опорного плану, якому відповідає одиничний базис. Тому задачу треба перетворити так, щоб її матриця основних обмежень у канонічній формі запису



містила принаймні  $m$  різноманітних одиничних векторів, із яких можна скласти одиничну матрицю  $m$ -го порядку.

Таким чином, якщо  $X_0 = (x_{s1}, \dots, x_{sm}, 0, \dots, 0)$  – вихідний опорний план задачі з одиничним базисом  $B_x = (A_{s1}, \dots, A_{sm})$ , то модель задачі має вигляд:

$$Z = \sum_{j=1}^n c_j x_j \rightarrow \min ;$$

$$\begin{cases} x_{si} + \sum_{j \neq si} a_{ij} x_j = b_i; & i = \overline{1, m}; \\ x_j \geq 0; & j = \overline{1, n}; & b_i \geq 0; & i = \overline{1, m}. \end{cases} \quad (11.16)$$

З (11.16) випливає, що в опорному плані  $X_0$  базисні змінні  $x_{si} = b_i$ ,  $i = \overline{1, m}$ . Тому алгоритм симплексного методу – це процес обчислення оптимального розв’язку за вихідним опорним планом  $X_0 = (x_{s1} = b_1, \dots, x_{sm} = b_m)$ .

Для дослідження  $X_0$  на оптимальність необхідно розкласти усі вектори  $A_j$  за векторами базису  $A_{si}$  і підрахувати різницю  $\Delta_j = Z_j - c_j$ .

Позначимо через  $X_j$  вектор, складений із компонентів  $A_j$  у даному базисі, тобто  $X_j = (x_{1j}, x_{2j}, \dots, x_{mj})^T$ . Тоді, очевидно, що  $x_{ij} = a_{ij}$ .

Корисно всю інформацію записувати у симплексну таблицю. Записавши вихідну інформацію про задачу, одержимо вихідну або нульову симплексну таблицю (табл. 11.2).

Таблиця 11.2 – Симплексна таблиця

$i$	$B_x$	$C_B$	$A_0$	$c_1$	...	$c_{s1}$	...	$c_{sr}$	...	$c_k$	...	$c_{sm}$	...	$c_m$	
				$A_1$	...	$A_{s1}$	...	$A_{sr}$	...	$A_k$	...	$A_{sm}$	...	$A_m$	
1	$A_{s1}$	$c_{s1}$	$x_{s1}$	$x_{11}$	...	1	...	0	...	$x_{1k}$	...	0	...	$x_{1n}$	$t$
2	$A_{s2}$	$c_{s2}$	$x_{s2}$	$x_{12}$	...	0	...	0	...	$x_{2k}$	...	0	...	$x_{2n}$	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
$r$	$A_{sr}$	$c_{sr}$	$x_{sr}$	$x_{r1}$	...	0	...	1	...	$x_{rk}$	...	0	...	$x_{rn}$	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
$m$	$A_{sm}$	$c_{sm}$	$x_{sm}$	$x_{m1}$	...	0	...	0	...	$x_{mk}$	...	1	...	$x_{mn}$	
$m+1$	-	-	$Z_0$	$\Delta_1$	...	$\Delta_{s1}=0$	...	$\Delta_{sr}=0$	...	$\Delta_k$	...	$\Delta_{sm}=0$	...	$\Delta_n$	



Таблиця заповнюється в такому порядку: нумеруються рядки (стовпчик  $i$ ), виписуються коефіцієнти цільової функції (рядок  $C$ ), виписується

базис (стовпчик  $B_x$ ), виписуються коефіцієнти цільової функції, що відповідають векторам базису (стовпчик  $C_B$ ), виписуються компоненти опорного плану (стовпчик  $A_0$ ), виписується матриця системи (стовпчики  $A_1, A_2, \dots, A_n$ ). У такий спосіб заповнюються перші  $m$  рядків таблиці.

У стовпчик  $A_0$  ( $m + 1$ )-го рядка записується значення лінійної форми для наявного опорного плану. Це значення обчислюється за формулою:

$$Z_0 = C_B A_0 = \sum_{i=1}^m c_{si} x_{si}. \quad (11.17)$$

У наступні клітини ( $m + 1$ )-го рядка записуються значення оцінок векторів умов  $A_j$ , (різниці  $\Delta_j$ ), що обчислюються за формулою:

$$\Delta_j = Z_j - c_j = C_B A_j - c_j = \sum_{i=1}^m c_{si} x_{ij} - c_j. \quad (11.18)$$

Очевидно, базисні різниці  $\Delta_{si} = 0$ .

Для перевірки опорного плану на оптимальність переглядаємо ( $m + 1$ )-ий рядок. При цьому можуть зустрітися такі випадки:

- 1) усі різниці  $\Delta_j \leq 0$ . Тоді опорний план  $X_0$  є оптимальним і  $\min Z = Z_0$ ;
- 2) для деякого  $j$  величина  $\Delta_j > 0$  і усі відповідні коефіцієнти стовпчика  $x_{ij} \leq 0, i = \overline{1, m}$ . Тоді лінійна форма є необмеженою на множині планів. У обох випадках процес обчислення закінчується;

3) для деяких індексів  $j \in \Delta_j > 0$  і для кожного такого  $j$  принаймні одне з чисел  $x_{ij} > 0$ . Це свідчить про те, що опорний план не є оптимальним і його можна поліпшити введенням до базису вектора, що відповідає одній із таких різниць. Звичайно до базису вводиться той вектор, якому відповідає  $\max \Delta_j$ . Якщо таких різниць декілька, то береться різниця з меншим індексом.

Нехай  $\max_j \Delta_j = \Delta_k$ . Тоді до базису варто ввести вектор  $A_k$  на місце вектора, якому відповідає мінімальне значення:

$$t = \min_{x_{ik} \geq 0} \frac{x_{si}}{x_{ik}} = x_{sr} / x_{rk}. \quad (11.19)$$

Тоді з базису виключають вектор  $A_{sr}$ . Новий базис буде складатися з векторів  $A_{s1}, A_{s2}, \dots, A_{sr-1}, A_{sr+1}, \dots, A_{sm}, A_k$ .

Число  $x_{rk}$  називається провідним елементом симплексної таблиці. Стовпчик  $k$  і рядок  $r$  називають також провідними (напрямними, визначальними).

При цьому всі елементи нової симплексної таблиці при переході до нового опорного плану перетворюються за такими рекурентними формулами:

$$x'_{ij} = \begin{cases} x_{ij} - \frac{x_{rj}}{x_{rk}} x_{ir}, i \neq r; \\ \frac{x_{rj}}{x_{rk}}, i = r, i = \overline{1, m+1}, j = \overline{0, n}. \end{cases} \quad (11.20)$$

Формули (11.20) дозволяють перетворити симплексну таблицю в такий спосіб. Напрямний рядок  $r$  ділять на провідний елемент  $x_{rk}$ . У новій симплексній таблиці на місці провідного елемента утворюється одиниця:  $x'_{rk} = 1$ , а на місці провідного стовпчика повинен утворитися одиничний вектор з одиницею в напрямному рядку. Інші рядки перетворюються так: із того рядка вихідної таблиці, що необхідно перетворити, віднімаємо перетворену напрямну, помножену на той елемент рядка, на місці якого варто одержати нуль.

У результаті перетворення нульової симплексної таблиці одержуємо наступну таблицю, у якій буде новий опорний план. Знову переглядаємо  $(m + 1)$ -ий рядок. Якщо усі  $\Delta_j \leq 0$ , то опорний план оптимальний. У протилежному випадку переходимо до нового опорного плану. Процес продовжується до тих пір, поки прийдемо або до оптимального плану, або переконаємося в необмеженості лінійної форми. На рис. 11.4 зображена блок-схема окремої ітерації розв'язання задачі ЛП симплексним методом.

Кожна ітерація складається з двох етапів. Перший етап полягає в перевірці досліджуваного опорного плану на оптимальність. Другий етап проводиться, якщо план не виявився оптимальним і при цьому не виявлена неможливість розв'язання задачі.

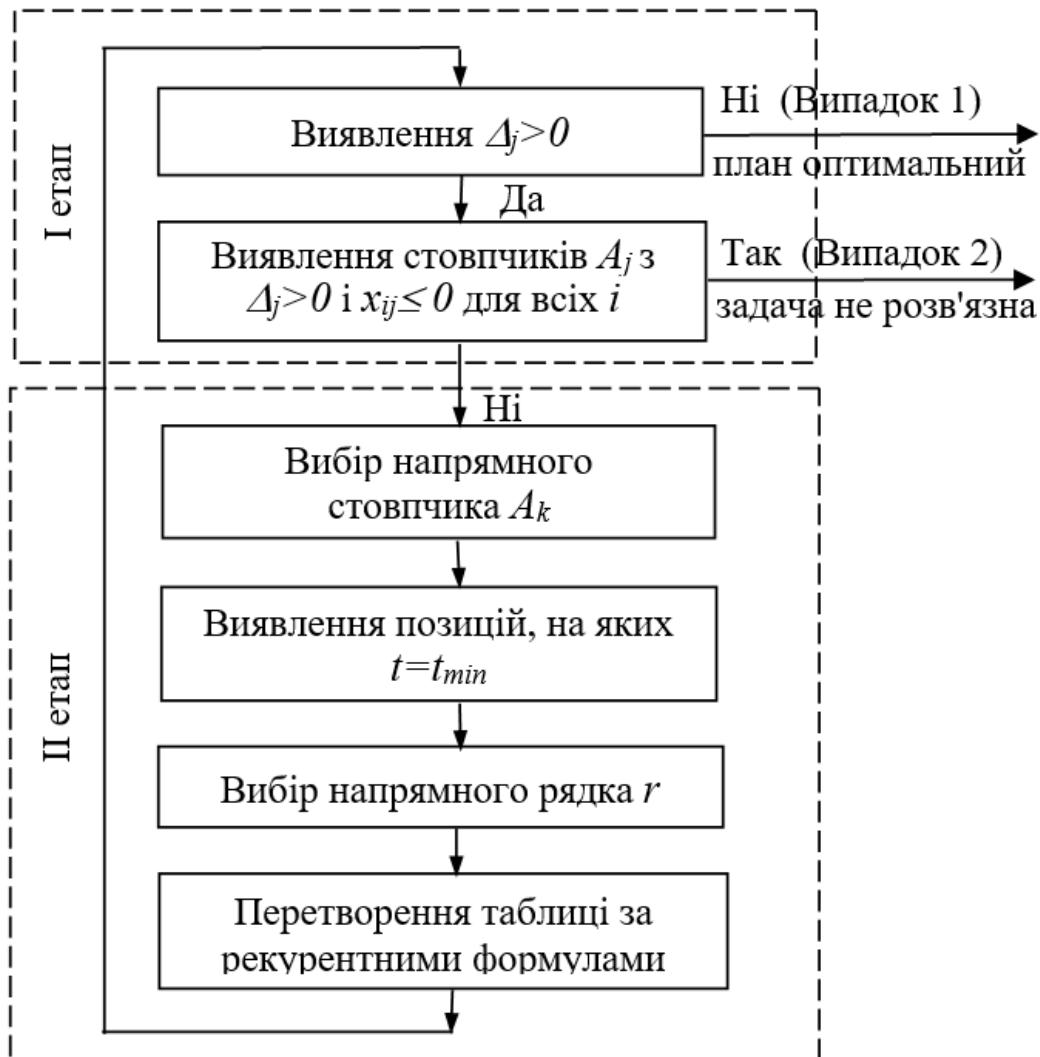


Рисунок 11.4 – Блок-схема окремої ітерації розв'язання задачі ЛП симплексним методом

### 11.8 Симплексний метод із штучним базисом

З алгоритму симплексного методу випливає, що розв'язання задачі ЛП симплексним методом зручно починати з такого опорного плану, якому відповідає базис, що складається з одиничних векторів. Для цього вектори  $A_1, \dots, A_n$ , які утворюють матрицю  $A$ , повинні містити принаймні  $m$  одиничних векторів, із яких можна скласти одиничну матрицю порядку  $m$ . Проте на практиці матриця  $A$  не завжди містить  $m$  одиничних векторів.

**Задача 3.** Потрібно мінімізувати функцію  $Z = -2x_1 + x_2 - x_3$  за обмежень:

$$\begin{cases} x_1 - x_2 + 2x_3 \leq 8, \\ 2x_1 - 2x_2 - x_3 = 4, \\ -x_1 + x_2 + x_3 \geq 1, \\ x_j \geq 0, j = \overline{1,3}. \end{cases}$$

### Розв'язання

За допомогою додаткових змінних  $x_4, x_5$  приведемо обмеження до канонічного вигляду:

$$\begin{cases} x_1 - x_2 + 2x_3 + x_4 = 8, \\ 2x_1 + 2x_2 - x_3 = 4, \\ -x_1 + x_2 + x_3 - x_5 = 1, \\ x_j \geq 0, j = \overline{1,5}. \end{cases}$$

Матриця  $A$  системи рівнянь набуває вигляду:

$$A = \begin{pmatrix} A_1 & A_2 & A_3 & A_4 & A_5 \\ 1 & -1 & 2 & 1 & 0 \\ 2 & 2 & -1 & 0 & 0 \\ -1 & 1 & 1 & 0 & -1 \end{pmatrix}.$$

Ця матриця містить тільки один одиничний вектор  $A_4$ . Побудуємо такі обмеження, щоб матриця  $A$  містила вектори:

$$A_6 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ і } A_7 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Для цього до лівих частин другого і третього рівнянь введемо із коефіцієнтами  $+1$  невід'ємні змінні  $x_6, x_7$ :

$$\begin{cases} x_1 - x_2 + 2x_3 + x_4 = 8, \\ 2x_1 + 2x_2 - x_3 + x_6 = 4, \\ -x_1 + x_2 + x_3 - x_5 + x_7 = 1, \end{cases}$$

$$x_{ij} \geq 0, j = \overline{1,7}.$$

Змінні  $x_6, x_7$  називаються штучними або фіктивними. Ці змінні вводяться в цільову функцію з великим додатним коефіцієнтом  $M$ , значення якого більше будь-якого скінченного числа, порівнюваного з  $M$  в процесі розв'язання.

У даному прикладі цільова функція допоміжної задачі буде такою:

$$Z' = -2x_1 + x_2 - x_3 + Mx_6 + Mx_7 \rightarrow \min.$$

Отриману задачу називають  $M$ -задачею або розширеною задачею стосовно вихідної.

У загальному випадку  $M$ -задачу для вихідної задачі:

$$\begin{cases} Z = CX \rightarrow \min, \\ AX = A_0, \quad X \geq 0, \quad A_0 \geq 0 \end{cases} \quad (11.21)$$

можна записати у вигляді:

$$\begin{cases} Z' = CX + M \sum_{j=n+1}^{n+k} x_j = \widehat{C}\widehat{X} \rightarrow \min; \\ \widehat{A}\widehat{X} = A_0, \quad \widehat{X} \geq 0, \quad A_0 \geq 0, \end{cases} \quad (11.22)$$

де  $\widehat{A} = (A, A_{n+1}, \dots, A_{n+k})$ ;  $A_{n+i}$  – штучні одиничні вектори;

$$\widehat{C} = (c_1, c_2, \dots, c_n, M, \dots, M); \quad \widehat{X} = (x_1, x_2, \dots, x_n, x_{n+1}, \dots, x_{n+k}), \quad k \leq m.$$

Зв'язок оптимальних розв'язків вихідної задачі і відповідної їй  $M$ -задачі дає наступна теорема.

**Теорема.** Якщо в оптимальному плані  $\widehat{X}^* = (x_1^*, x_2^*, \dots, x_n^*, x_{n+1}^*, \dots, x_{n+k}^*)$  розширеної задачі, значення штучних змінних  $x_{n+i}^* = 0, i = \overline{1, k}$ , то  $X^* = (x_1^*, \dots, x_n^*)$  є оптимальним планом вихідної задачі.  $M$ -метод дозволяє обчислити оптимальний план будь-якої розв'язної задачі без попереднього визначення початкового опорного плану. При цьому:

1) якщо в процесі розв'язання  $M$ -задачі буде отриманий оптимальний план, для якого  $x_{n+i} = 0$  для  $i = \overline{1, k}$ , то перші  $n$  його компонентів визначають розв'язок вихідної задачі;

2) якщо принаймні одна  $x_{n+i} > 0$  за будь-якого як завгодно великого  $M$ , то вихідна задача не має жодного плану, тобто умови задачі несумісні;

3) якщо в процесі аналізу  $M$ -задачі виявляється її нерозв'язаність, то нерозв'язаною є і вихідна задача.

Продовжимо розв'язання прикладу.

Заповнимо нульову й наступні симплекс-таблиці (табл. 11.3). При цьому в  $(m + 1)$ -ий рядок записуємо вільні члени, а в  $(m + 2)$ -й рядок – коефіцієнти при  $M$  у  $Z_0$  і  $\Delta_j$ . Вектор, який вводиться у базис, зв'язується з найбільшим додатним коефіцієнтом при  $M$  у різницях  $\Delta_j$ .

Таблиця 11.3 – Симплекс-таблиці

$i$	$B_X$	$C_B$	$A_0$	-2	1	-1	0	0	$M$	$M$	$N$
				$A_1$	$A_2$	$A_3$	$A_4$	$A_5$	$A_6$	$A_7$	
1	$A_4$	0	8	1	-1	2	1	0	0	0	0
2	$A_6$	$M$	4	2	2	-1	0	0	1	0	
3	$A_7$	$M$	1	-1	<span style="border: 1px solid black;">1</span>	1	0	-1	0	1	
4			0	2	-1	1	0	0	0	0	
5		$*M$	5	1	<span style="border: 1px solid black;">3</span>	0	0	-1	0	0	
1	$A_4$	0	9	0	0	3	1	-1	0		1
2	$A_6$	$M$	2	<span style="border: 1px solid black;">4</span>	0	-3	0	2	1		
3	$A_2$	1	1	-1	1	1	0	-1	0		
4			1	1	0	2	0	-1	0		
5		$*M$	2	<span style="border: 1px solid black;">4</span>	0	-3	0	2	0		
1	$A_4$	0	9	0	0	<span style="border: 1px solid black;">3</span>	1	-1			2
2	$A_1$	-2	1/2	1	0	-3/4	0	1/2			
3	$A_2$	1	3/2	0	1	1/4	0	-1/2			
4			1/2	0	0	<span style="border: 1px solid black;">11/4</span>	0	-3/2			
1	$A_3$	-1	3	0	0	1	1/3	-1/3			3
2	$A_1$	-2	11/4	1	0	0	1/4	1/4			
3	$A_2$	1	3/4	0	1	0	-1/12	-8/12			
4			-31/4	0	0	0	-11/12	-7/12			

У наступних рядках таблиці записується розв'язок розширеної задачі. У останньому рядку усі  $\Delta_j \leq 0$ , оптимальний розв'язок набуває вигляду:

$$X^* = (11/4, 3/4, 3), \quad \min Z = -31/4.$$

**Відповідь:**  $X^* = (11/4, 3/4, 3), \quad \min Z = -31/4.$

Варто зауважити, що штучні вектори  $A_6, A_7$ , які виключаються з базису, не бажано вводити в наступні базиси, тому надалі стовпчики даних векторів не заповнюються.

### Питання для самоперевірки

1. Зробіть постановку математичної моделі задачі лінійного програмування.
2. Запишіть канонічний вигляд задачі лінійного програмування.
3. Вкажіть рекурентні формули перерахунку симплексної таблиці.
4. Запишіть блок-схему ітерації симплексного методу.
5. Сформулюйте ознаку оптимальності опорного плану.
6. Вкажіть особливості використання симплексного методу із штучним базисом.
7. Сформулюйте теорему зв'язку оптимальних розв'язків вихідної задачі і відповідної їй  $M$ -задачі.



## 12 ТРАНСПОРТНА ЗАДАЧА ЛП

### 12.1 Постановка транспортної задачі

Розглянутий вище симплексний метод дозволяє розв'язувати будь-які задачі ЛП. Проте існують класи задач, що можуть бути розв'язані більш простими методами. Однією з таких задач ЛП є транспортна задача.

**Задача формулюється** таким чином. Нехай у  $m$  пунктах  $A_1, \dots, A_m$  випускається деякий однорідний продукт. Причому обсяг виробництва в пункті  $A_i$  складає  $a_i$  одиниць. Даний продукт споживається в  $n$  пунктах  $B_1, \dots, B_n$ , й обсяг споживання в пункті  $B_j$  складає  $b_j$  одиниць. Відомі транспортні витрати  $c_{ij}$  на перевезення одиниці продукту з пункту  $A_i$  у пункт  $B_j$ . Завдання полягає у визначенні такого плану перевезень, за якого потреби всіх споживачів цілком задоволені, весь продукт із пунктів виробництва вивезений і сумарні транспортні витрати є мінімальними.

**Математичну модель Т-задачі** можна представити у такому вигляді. Нехай  $x_{ij}$  – кількість продукту, перевезеного з пункту  $A_i$  у пункт  $B_j$ . Потрібно визначити план перевезень  $X = (x_{11}, \dots, x_{ij}, \dots, x_{mn})$ , який мінімізує сумарні транспортні витрати:

$$Z = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min, \quad (12.1)$$

задовольняє умовам:

$$\begin{cases} \sum_{j=1}^n x_{ij} = a_i, i = \overline{1, m}; \\ \sum_{i=1}^m x_{ij} = b_j, j = \overline{1, n}; \end{cases} \quad (12.2)$$

$$x_{ij} \geq 0, \quad i = \overline{1, m}, \quad j = \overline{1, n}.$$

Умови (12.2) гарантують повний вивіз продукту і повне задоволення попиту. Якщо:

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j, \quad (12.3)$$

тоді математична модель Т-задачі називається закритою. Це гарантує її можливість розв'язання.

У випадку порушення умови (12.3) модель Т-задачі є відкритою. Такі задачі вирішують шляхом зведення їх до закритих шляхом введення додаткового стовпчика або додаткового рядка. Т-задача являє собою задачу ЛП із  $m \times n$  змінними  $x_{ij}$  і  $m + n$  обмежень-рівностей. Матрицю  $X = (x_{ij})_{n \times m}$  називають планом перевезень, а матрицю  $C = (c_{ij})_{m \times n}$  – матрицею транспортних витрат або тарифів. Оскільки обмеження (12.2) залежні в силу виконання умови (12.3), то кожний опорний план містить  $m + n - 1$  базисних змінних. При цьому, якщо кількість додатних базисних змінних дорівнює  $m + n - 1$ , то план є не виродженим, якщо менше – план є виродженим.

Інформацію про транспортну задачу зручно подавати в таблиці Т-задачі (табл. 12.1).

Таблиця 12.1 – Інформація про транспортну задачу

Постачальник	Споживач				Обсяг виробництва $a_i$
	$B_1$	$B_2$	...	$B_n$	
$A_1$	$c_{11}$ $x_{11}$	$c_{12}$ $x_{12}$	...	$c_{1n}$ $x_{1n}$	$a_1$
$A_2$	$c_{21}$ $x_{21}$	$c_{22}$ $x_{22}$	...	$c_{2n}$ $x_{2n}$	$a_2$
...	...	...	...	...	...
$A_m$	$c_{m1}$ $x_{m1}$	$c_{m2}$ $x_{m2}$	...	$c_{mn}$ $x_{mn}$	$a_m$
Обсяг споживання $b_j$	$b_1$	$b_2$	...	$b_n$	$\sum$

## 12.2 Побудова вихідного опорного плану

**Метод «північно-західного кута».** Таблиця заповнюється, починаючи з лівого верхнього (північно-західного) кута, рухаючись далі за рядком праворуч або за стовпчиком вниз. У клітину (1, 1) заноситься менше з чисел  $a_1$  і  $b_1$ , тобто  $x_{11} = \min(a_1, b_1)$ .

Якщо  $a_1 > b_1$ , то  $x_{11} = b_1$ ,  $x_{i1} = 0$ ,  $i = \overline{2, m}$ , тобто перший стовпчик вважають закритим для заповнення інших клітин. Рухаючись далі за першим рядком, записуємо в сусідню клітину (1,2)  $x_{12} = \min(a_1 - b_1, b_2)$  і т. д. Якщо ж  $b_1 > a_1$ , то  $x_{11} = a_1$ ,  $x_{1j} = 0$ ,  $j = \overline{2, n}$ ;  $x_{21} = \min(a_2, b_1 - a_2)$  і т. д.

Після заповнення другої клітини (1,2) або (2,1), переходять до заповнення наступної клітини або за рядком, або за стовпчиком. Цей процес продовжується до повного вичерпання вантажу в постачальників або повного задоволення споживачів. Останньою заповненою клітиною буде  $(m, n)$ . План, отриманий за методом північно-західного кута, буде містити не більше  $m + n - 1$  перевезень.

**Задача 4.** У табл. 12.2 приведений план, побудований за методом північно-західного кута.

Таблиця 12.2 – План, побудований за методом північно-західного кута

$j \backslash i$	$B_1$	$B_2$	$B_3$	$B_4$	$a_i$
$A_1$	5 40	4 20	1	2	60
$A_2$	4	2 5	6 20	3 15	40
$A_3$	7	3	3	4 35	35
$b_j$	40	25	20	50	135

Для цієї задачі  $n + m - 1 = 6$ ;  $Z = 595$ .

**Метод мінімального елемента.** Заповнення таблиці починається з клітини з найменшим тарифом  $c_{ij}$ . Далі з рядків, що залишилися, або стовпчиків таблиці знову обирають найменший тариф, і процес розподілу запасів продовжується, тобто послідовність заповнення клітин визначається за величиною  $c_{ij}$ . Якщо  $\min c_{ij} = c_{lk}$ , то  $x_{lk} = \min(a_i, b_k)$ .

План, отриманий за методом мінімального елемента, для розглянутого приклада такий:

$$X_0 = \begin{pmatrix} 0 & 0 & 20 & 40 \\ 5 & 25 & 0 & 10 \\ 35 & 0 & 0 & 0 \end{pmatrix}, Z = 445 < 595.$$

### 12.3 Метод потенціалів

**Алгоритм методу потенціалів** включає такі етапи.

**Етап 1.** Побудувати вихідний опорний план  $X_0$  методом північно-західного кута або мінімального елемента.

**Етап 2.** Кожному постачальнику  $A_i$  ставиться у відповідність деяке число  $u_i$ , кожному споживачу  $B_j$  – число  $v_j$ . Ці числа називаються потенціалами. Обчислюють потенціали, розв'язуючи систему:

$$u_i + v_j = c_{ij}$$

для заповнених (базисних) клітин. Оскільки всіх потенціалів  $m + n$ , а зайнятих клітин  $m + n - 1$ , то одній невідомій надають довільне значення, найчастіше, нуль. Інші  $m + n - 1$  потенціалів визначаються однозначно.

**Етап 3.** Перевірка опорного плану на оптимальність. Для цього необхідно обчислити для вільних клітин різниці  $\gamma_{ij} = c_{ij} - (u_i + v_j)$ . Якщо усі  $\gamma_{ij} \geq 0$ , то план є оптимальним.

**Етап 4.** Якщо є клітини з  $\gamma_{ij} < 0$ , то переходять до нового опорного плану з меншим значенням цільової функції. До зайнятих клітин вводять клітину, для якої оцінка  $\gamma_{ij}$  мінімальна.

Нехай  $\min_{\gamma_{ij} < 0} \gamma_{ij} = \gamma_{lk}$ . Тоді клітина  $(l, k)$  вводиться до плану (рис. 12.1).

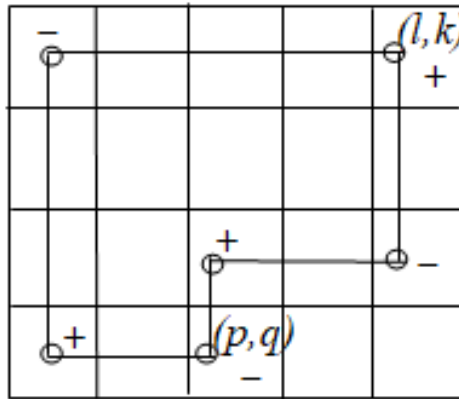


Рисунок 12.1 – Клітина  $(l, k)$  вводиться до плану

**Етап 5.** Побудувати цикл перерахунку з початковою клітиною  $(l, k)$  (рис. 12.1). Вихідна клітина циклу позначається знаком «+», а інші, чергуючись, знаками «-» і «+». Клітини циклу, позначені знаком «+», утворять додатний півланцюг, а знаком «-» – від'ємний. Серед усіх клітин від'ємного півланцюга знаходять клітину, що містить найменше значення  $x_{ij} = t$ . Нехай це клітина  $(p, q)$ . Тоді виконують перерахунок:

$$x_{ij}' = \begin{cases} x_{ij} + t, & \text{для "+"}; \\ x_{ij} - t, & \text{для "-"}; \\ x_{ij}, & \text{для інших.} \end{cases}$$

Одержують новий поліпшений план, із якого клітина  $(p, q)$  виведена і введена клітина  $(l, k)$ , що містить  $x_{lk} = t$ . Переходять до етапу 2.

### Питання самоперевірки

1. Постановка математичної моделі транспортної задачі.
2. Указати умову, за якої маємо транспортну задачу закритого типу.
3. Указати основні методи побудови початкового плану.
4. Сформулювати умову перевірки на оптимальність отриманого плану перевезень.
5. Укажіть алгоритм побудови початкового плану перевезень методом мінімального елемента.
6. Укажіть алгоритм побудови початкового плану перевезень методом північно-західного кута.
7. Наведіть основні розрахункові формули методу потенціалів.

## 13 МЕТОДИ ДИСКРЕТНОЇ ОПТИМІЗАЦІЇ

### 13.1 Постановка задачі дискретної оптимізації

Розглянемо таку задачу оптимізації.

Знайти

$$\min_{X \in D_X} f(X), \quad (13.1)$$

де  $D_X = \{X \in R \mid g_j(X) \geq 0, j = \overline{1, l}; h_j(X) = 0, j = \overline{1, m}\}$ .

Якщо множина  $R$  є скінченною або зліченною, тоді умова  $X \in R$  – це умова дискретності й дана задача є задачею дискретного програмування. Зазвичай умова дискретності розділена за окремими змінними  $x_j \in R_j, j = \overline{1, n}$ , де  $R_j$  – скінченна або зліченна множина. Якщо вводяться обмеження, що  $x_j$  – ціле число, то маємо задачу цілочислового програмування, що є окремим випадком дискретного програмування.

У задачах дискретного програмування область допустимих розв'язків є неопуклою і незв'язною. Тому є неможливим для її розв'язання застосування стандартних методів, що використовуються при заміні дискретної задачі її неперервним аналогом із подальшим округленням розв'язку до найближчого цілочислового.

Розглянемо таку задачу ЛП:

$$x_1 - 3x_2 + 3x_3 \rightarrow \max;$$

$$\begin{cases} 2x_1 + x_2 - x_3 \leq 4, \\ 4x_1 - 3x_2 \leq 2, \\ -3x_1 + 2x_2 + x_3 \leq 3. \end{cases}$$

$$x_j \geq 0 - \text{цілі}, j = \overline{1, 3}.$$

Ігноруючи умову цілочисельності, симплекс-метод дає оптимальний план  $X^* = (0,5; 0; 4,5); Z^* = 14$ .

Округлення компонентів цього плану не дає допустимого розв'язку. Шуканий цілочисловий розв'язок  $X^* = (2; 2; 5); Z^* = 11$ .

## 13.2 Метод Гоморі

Задача цілочислового лінійного програмування (ЦЛП) ставиться аналогічно звичайній задачі ЛП:

$$Z = \sum_{j=1}^n c_j x_j \rightarrow \min;$$
$$\begin{cases} \sum_{j=1}^n a_{ij} x_j = b_i, & i = \overline{1, m}; \\ x_j \geq 0, & j = \overline{1, m}; \end{cases} \quad (13.2)$$
$$x_j - \text{цiле}, j = \overline{1, p} \leq n.$$

Метод Гоморі (метод відсікання, метод січних площин) приводить розв'язання задачі (13.2) до розв'язання ряду спеціально побудованих задач ЛП, кожна з яких утворюється з вихідної шляхом додавання до її умов додаткового обмеження (перетину) вигляду:

$$\sum \alpha_j x_j \geq \alpha_0. \quad (13.3)$$

При цьому перетин будується таким чином, що:

- 1) будь-який цілочисловий розв'язок вихідної задачі йому задовольняє;
- 2) нецілочисловий оптимальний вихідної задачі йому не задовольняє.

Такий перетин називається **правильним**. Геометрично це означає провести перетин (гіперплощину) у многограннику  $L$ , що відкидав би нецілочислову оптимальну вершину, не торкаючись при цьому цілочислових.

Приєднуючи це лінійне обмеження до обмежень неперервної задачі ЛП, симплекс-методом знаходиться її новий і т. д. Після декількох таких ітерацій потрібна цілочислова точка буде на межі множини допустимих розв'язків й один з оптимальних розв'язків стає цілочисловим.

Розв'язання задачі (13.2) алгоритмом Гоморі виконується в такому порядку.

**Етап 1.** Розв'язується вихідна задача ЛП без накладання умови цілочисельності.

**Етап 2.** Якщо оптимальний розв'язок є цілочисловим, то процес завершується. Якщо умова цілочисельності не виконана, то для базисної змінної, що має найбільшу дробову частину, будується правильний перетин.

**Етап 3.** Утворення розширеної задачі шляхом додавання правильного перетину до системи лінійних обмежень задачі ЛП. Повернення до етапу 1.

Через скінченну кількість ітерацій одержимо цілочисловий оптимальний план.

### 13.3 Метод гілок і меж

Метод гілок і меж відносять до групи комбінаторних методів дискретного програмування. Ідея цих методів полягає в тому, що на кожному кроці пошуку замість оцінювання окремого варіанта оцінюється група варіантів. Якщо результат такого оцінювання негативний, то з подальшого розгляду виключається не окремий варіант, як у методі повного перебору, а група варіантів, що істотно зменшує кількість кроків пошуку. Продемонструємо застосування методу гілок і меж для задачі ЦЛП (13.2). Як і в методі січних площин, процес починається з розв'язання неперервної задачі ЛП. Якщо при цьому отриманий оптимальний план  $X_0$  не задовольняє умові цілочисельності, то значення цільової функції  $Z_0 = Z(X_0)$  дає нижню оцінку для шуканого розв'язку.

Нехай деяка змінна  $x_k$  ( $1 \leq k \leq p$ ) не одержала в плані  $X_0$  цілочислового розв'язку. Тоді множина  $L$  розбивається на дві підмножини введенням додаткових обмежень:

$$x_k \leq [x_k], \quad x_k \geq [x_k] + 1. \quad (13.4)$$

Знову розв'язують задачу ЛП (13.2) за додаткових умов (13.4). Обирають план  $X_1$  із мінімальною оцінкою  $Z_1 = Z(X_1)$ . Якщо цей план задовольняє умові цілочисельності, то він буде оптимальним планом задачі ЦЛП. В іншому випадку продовжують розгалуження. При цьому щораз для чергового розгалуження обирають підмножину з найменшою оцінкою. Процес продовжується доти, поки продовження розгалуження стає неможливим.



Геометрично метод гілок і меж для задачі ЦЛП інтерпретується таким чином. Гіперплощина, обумовлена цільовою функцією задачі, вдавлюється усередину багатогранника планів відповідної неперервної задачі ЛП до зустрічі з найближчою цілочисловою точкою цього багатогранника.

### **Питання для самоперевірки**

1. Зробіть постановку задачі дискретної оптимізації.
2. Зробіть постановку задачі цілочислового програмування.
3. Укажіть алгоритм використання методу Гоморі (метод відсікання).
4. Наведіть геометричне тлумачення побудови правильного перетину.
5. Сформулюйте особливості й основні етапи використання методу гілок і меж.
6. Наведіть геометричну інтерпретацію використання методу гілок і меж для задачі ЦЛП.

## 14 ДРОБОВО-ЛІНІЙНЕ ПРОГРАМУВАННЯ

### 14.1 Постановка задачі й властивості її розв'язку

При плануванні виробництва велику роль грають, наприклад, собівартість продукції, рентабельність виробництва, середня вартість однієї тонни продукції й інші питомі показники. Частина з цих показників треба мінімізувати, інші – максимізувати.

Математичні моделі, що описують ці питомі показники, як правило, мають дробово-лінійну структуру.

Задача дробово-лінійного програмування (ДрЛП) формулюється таким чином.

Знайти мінімум дробово-лінійної функції:

$$Z = \frac{\sum_{j=1}^n c_1 j x_j}{\sum_{j=1}^n c_2 j x_j} = \frac{Z_1(X)}{Z_2(X)} \quad (14.1)$$

за лінійних обмежень

$$\begin{cases} \sum_{j=1}^n a_{ij} x_j = b_i, & i = \overline{1, m}, \\ x_j \geq 0, & j = \overline{1, n}. \end{cases} \quad (14.2)$$

Це канонічна форма задачі ДрЛП. Будь-яка інша задача ДрЛП може бути приведена до вигляду (14.1) – (14.2).

Як і в задачі ЛП, множина допустимих розв'язків  $L$  задачі ДрЛП є опуклою, що має кінцеву кількість граничних точок. Знаменник  $Z_2(X)$  – лінійна і неперервна функція, тому для зручності будемо вважати, що  $Z_2(X) > 0$  в області  $L$ . В іншому випадку вважатимемо, що знак мінус належить до чисельника.

Для задачі ДрЛП зберігаються за аналогією з задачею ЛП всі основні поняття: допустимий, базисний, оптимальний розв'язок. Методи розв'язання задач ДрЛП засновані на твердженнях, аналогічних для задач ЛП.

## 14.2 Графічний метод розв'язання задачі ДрЛП

Розглянемо задачу знаходження екстремуму дробово-лінійної функції:

$$Z = \frac{c_{11}x_1 + c_{12}x_2}{c_{21}x_1 + c_{22}x_2} \quad (14.3)$$

за лінійних обмежень:

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 \leq b_i, & i = \overline{1, m}, \\ x_1, x_2 \geq 0. \end{cases} \quad (14.4)$$

На площині  $x_1 O x_2$  будемо множини  $L$  допустимих розв'язків. Для з'ясування, в якій із граничних точок  $L$  досягається екстремум функції, виразимо з (14.3) змінну  $x_2$  через  $x_1$ .

$$x_2 = \frac{c_{11} - zc_{21}}{zc_{22} - c_{12}} x_1 = kx_1.$$

Тут кутовий коефіцієнт  $k$  залежить від  $Z$ . Для встановлення характеру цієї залежності візьмемо похідну за  $Z$ :

$$\frac{dk}{dz} = \frac{-c_{21}(zc_{22} - c_{12}) - c_{22}(c_{11} - zc_{21})}{(zc_{22} - c_{12})^2} = \frac{c_{12}c_{21} - c_{11}c_{22}}{(zc_{22} - c_{12})^2}.$$

Оскільки знаменник дробу завжди додатний, а чисельник не залежить від  $Z$ , то похідна зберігає знак або  $+$ , або  $-$ . Отже, при збільшенні  $Z$  коефіцієнт  $k$  або збільшується, або зменшується. Справедливо й обернене твердження: при збільшенні  $k$  функція  $Z$  або зростає, або спадає.

Тому, з'ясувавши напрямок обертання для зростання  $Z$ , поворотом прямої  $x_2 = kx_1$  навколо початку координат знаходимо ті вершини багатогранника  $L$ , у яких досягається екстремум дробово-лінійної функції.

При розв'язанні задач графічним методом можуть зустрітися такі випадки:

- 1) область  $L$  є обмеженою. Мінімум і максимум функції  $Z$  досягаються (рис. 14.1, а));
- 2) область  $L$  є необмеженою, але обидва екстремуми функції  $Z$  існують (рис. 14.1, б));

3) область  $L$  є необмеженою і один екстремум не досягається. Допустима пряма  $x_2 = kx_1$  зайняла положення, паралельне одному із ребер многогранника (рис. 14.1, в)). У цьому випадку отриманий так називаний асимптотичний мінімум;

4) область  $L$  є необмеженою й обидва екстремуми асимптотичні (рис. 14.1, г)).

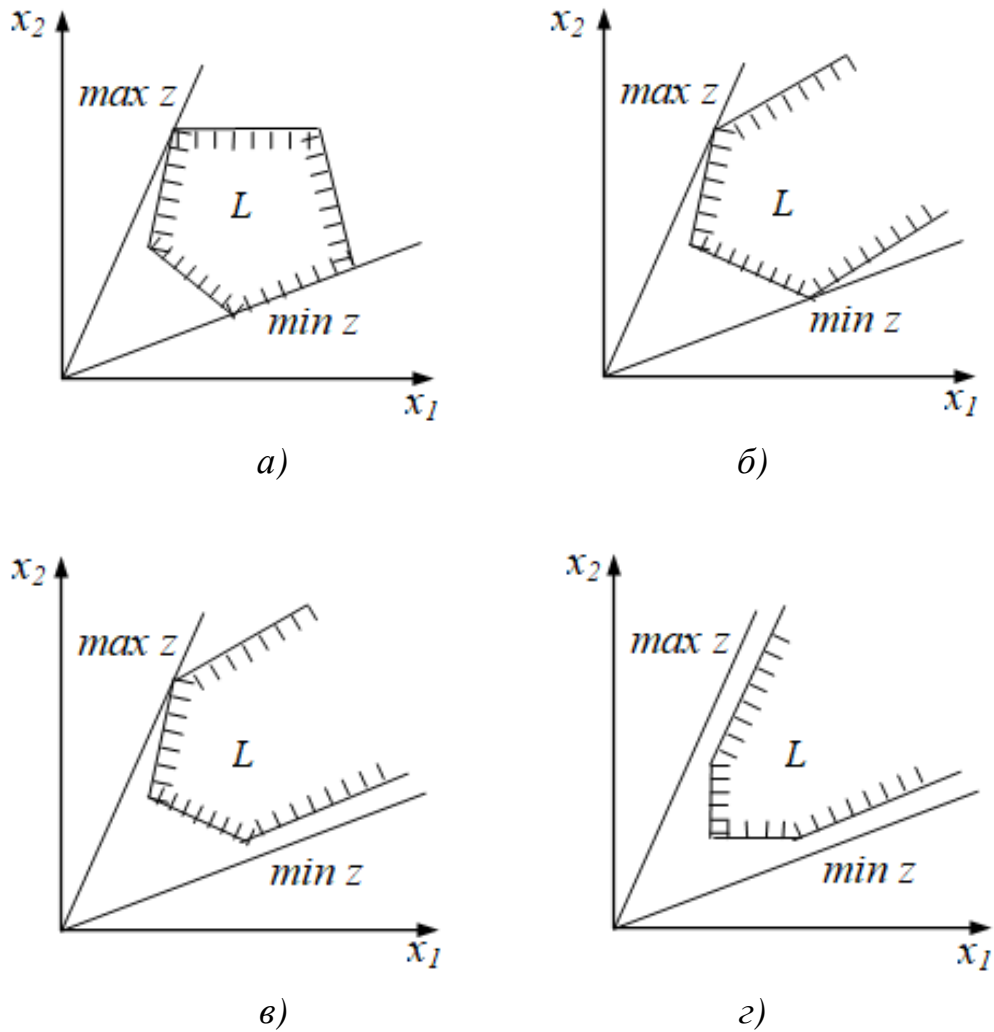


Рисунок 14.1 – Розв’язання задачі графічним методом

Нехай тепер дробово-лінійна функція задачі буде неоднорідною:

$$Z = \frac{c_{11}x_1 + c_{12}x_2 + c_{10}}{c_{21}x_1 + c_{22}x_2 + c_{20}}. \quad (14.5)$$

У цьому випадку багатогранник  $L$  будується аналогічно, а допустима пряма робить поворот навколо деякої точки  $A(x_1^0, x_2^0)$ , координати якої знаходяться таким чином:

а) зробимо заміну змінних:

$$\begin{cases} x_1 = y_1 + x_1^0, \\ x_2 = y_2 + x_2^0; \end{cases}$$

б) підставимо ці формули в (14.5), отримаємо:

$$Z = \frac{c_{11}y_1 + c_{12}y_2 + c_{11}x_1^0 + c_{12}x_2^0 + c_{10}}{c_{21}y_1 + c_{22}y_2 + c_{21}x_1^0 + c_{22}x_2^0 + c_{20}}.$$

Координати точки  $A$  підберемо таким чином, щоб неоднорідна функція в нових змінних  $y_1, y_2$  стала однорідною, тобто таким чином, щоб:

$$\begin{cases} c_{11}x_1^0 + c_{12}x_2^0 + c_{10} = 0; \\ c_{21}x_1^0 + c_{22}x_2^0 + c_{20} = 0. \end{cases}$$

На практиці в обох випадках можна не з'ясувати характер зміни кутового коефіцієнта допустимої прямої від  $Z$ , а обчислити просто значення дробово-лінійної функції в потрібних крайніх точках області  $L$ . Серед отриманих чисел обирають найбільше й найменше, що дає максимум і мінімум функції, а відповідні їм граничні точки визначають оптимальні розв'язки.

**Задача 6.** Знайти максимум і мінімум функції:

$$Z = \frac{3x_1 + 4x_2 + 9}{2x_1 + x_2 + 1}$$

за умов:

$$\begin{cases} -3x_1 + 4x_2 \leq 13; \\ x_1 + x_2 \geq 5; \\ -7x_1 + 5x_2 \geq 0; \\ x_1, x_2 \geq 0. \end{cases}$$

### Розв'язання

Будуємо багатогранник  $L$  і пряму  $PQ$ :  $2x_1 + x_2 + 1 = 0$  (рис. 14.2).

Пряма  $PQ$  не перетинає область  $L$ , отже, знаменник функції  $Z$  в області  $L$  не дорівнює нулю. Знаходимо координати точки  $A(x_1^0, x_2^0)$  із системи рівнянь:

$$\begin{cases} 3x_1 + 4x_2 + 9 = 0; \\ 2x_1 + x_2 + 1 = 0. \end{cases}$$

Одержуємо  $x_1^0 = 1$ ,  $x_2^0 = -3$ .

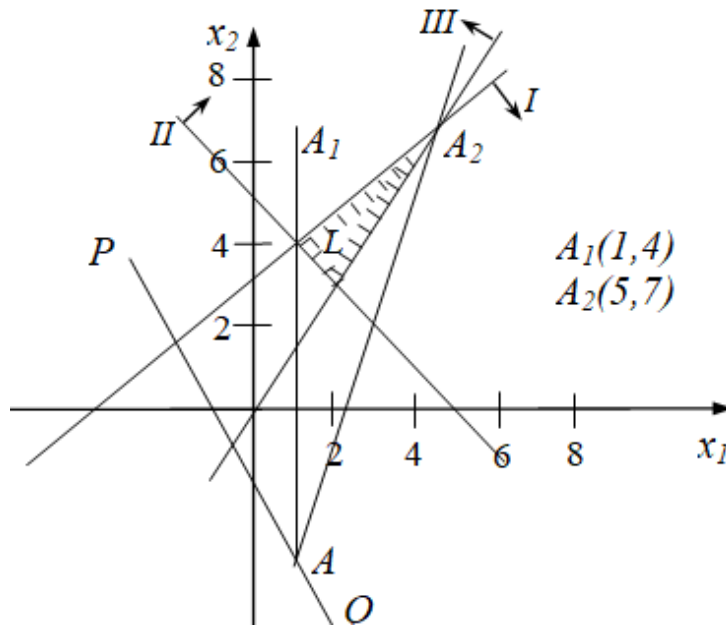


Рисунок 14.2 – Графічний розв'язок до задачі 6

Позначаємо на рисунку 14.2 точку  $A$  і, обертаючи допустиму пряму навколо цієї точки, знаходимо її граничні положення (точки  $A_1, A_2$ ). Обчислимо значення  $Z$  у цих точках:

$$Z(A_1) = \frac{3 \cdot 1 + 4 \cdot 4 + 9}{2 \cdot 1 + 4 + 1} = \frac{28}{7} = 4;$$

$$Z(A_2) = \frac{3 \cdot 5 + 4 \cdot 7 + 9}{2 \cdot 5 + 7 + 1} = \frac{52}{18} = 2 \frac{8}{9}.$$

Оскільки  $Z(A_1) > Z(A_2)$ , то  $\max Z = Z(A_1) = 4$ ,  $X_{\max} = (1, 4)$ ,  
 $\min Z = Z(A_2) = 2\frac{8}{9}$ ,  $X_{\min} = (5, 7)$ .

**Відповідь:**  $\max Z = Z(1, 4) = 4$ ,  $\min Z = Z(5, 7) = 2\frac{8}{9}$ .

Більш складні задачі ДрЛП розв'язують шляхом їх зведення до задач ЛП.

### Питання для самоперевірки

1. Зробіть постановку задачі дробово-лінійного програмування.
2. Укажіть основні властивості розв'язку задачі дробово-лінійного програмування.
3. Сформулюйте основні етапи застосування графічного методу розв'язання задачі дробово-лінійного програмування.
4. Надайте графічну інтерпретацію розв'язання задачі у випадку, коли область  $L$  є необмеженою, але обидва екстремуми функції  $Z$  існують.
5. Надайте графічну інтерпретацію розв'язання задачі у випадку, коли область  $L$  є необмеженою й один екстремум не досягається.
6. Надайте графічну інтерпретацію розв'язання задачі у випадку, коли область  $L$  є необмеженою й обидва екстремуми асимптотичні.

## 15 ДИНАМІЧНЕ ПРОГРАМУВАННЯ

### 15.1 Сутність динамічного програмування

Динамічне програмування (ДП) – це обчислювальний метод для розв'язання задач визначеної структури, який виник і сформувався в п'ятдесятих роках двадцятого століття завдяки роботам Р. Беллмана.

Припустимо, що деяка керована система знаходиться в початковому стані  $S_0 \in DS_0$ . З часом завдяки здійсненню деякого керування  $X$  система переходить у кінцевий стан  $S_n \in DS_n$ . Процес переходу зі стану  $S_0$  у  $S_n$  розпадається на  $n$  етапів. Причому, якщо система знаходиться на  $i$ -ому етапі в стані  $S_i$ , то стан  $S_{i+1}$  визначається не тільки вектором стану  $S_i$ , але і рішенням (керуванням)  $X_{i+1}$ , що прийняте на  $(i+1)$ -ому етапі, тобто вектор станів наступного етапу можна представити у вигляді:

$$S_{i+1} = \varphi_{i+1}(S_i, X_{i+1}),$$

де  $X_{i+1} \in DX_{i+1}$ .

У результаті реалізації  $(i+1)$ -го кроку буде забезпечений визначений прибуток, що залежить від вихідного стану системи  $S_i$  і обраного керування  $X_{i+1}$ , рівний  $R_{i+1}(S_i, X_{i+1})$ . Тоді загальний прибуток за  $n$  кроків або цільову функцію можна представити у вигляді:

$$f(S) = \sum_{i=0}^{n-1} R_{i+1}(S_i, X_{i+1}). \quad (15.1)$$

Задача динамічного програмування полягає в тому, щоб із множини можливих розв'язків  $X \in D_X$  знайти такий розв'язок  $X^*$ , що дозволить перевести систему з початкового стану  $S_0$  у кінцевий  $S_n$  таким чином, щоб цільова функція (15.1) набула екстремального значення.

Розв'язання задачі методом динамічного програмування засновано на **принципі оптимальності Беллмана**: яким би не був стан системи перед черговим кроком, треба вибрати керування на цьому кроці так, щоб прибуток на даному кроці плюс оптимальний прибуток на всіх наступних кроках був екстремальним.



Звідси випливає, що оптимальну стратегію керування можна одержати, якщо спочатку знайти оптимальну стратегію керування на  $n$ -му кроці, потім на двох останніх кроках, потім на трьох останніх і т. д., тобто розв'язання задачі динамічного програмування проводиться від кінця до початку.

Математичний запис принципу оптимальності називається **функціональним рівнянням Беллмана** або рекурентним співвідношення. Для  $k$ -го етапу воно має вигляд:

$$f_k(S_{n-k}) = \underset{X_{n-k+1}}{ext} [R_{n-k+1}(S_{n-k}, X_{n-k+1}) + f_{k-1}(S_{n-k+1})], \quad (15.2)$$

де  $f_k(S_{n-k})$  – екстремальне значення цільової функції за  $k$  етапів, починаючи зі стану  $S_0$ .

## 15.2 Алгоритм методу динамічного програмування

Алгоритм методу містить такі етапи.

**Етап 1.** Записують функціональне рівняння для кінцевого стану ( $k=1$ ) у вигляді:

$$f_1(S_{n-1}) = \underset{X_n}{ext} [R_n(S_{n-1}, X_n) + f_0(S_n)].$$

Оскільки за межами кінцевого стану  $S_n$  процес не розглядається, то  $f_0(S_n) = 0$ . Тому рівняння буде мати вигляд:

$$f_1(S_{n-1}) = \underset{X_n}{ext} R_n(S_{n-1}, X_n).$$

Розглядають набір фіксованих станів  $S_{n-1} \in DS_{n-1}$ , розв'язків  $X_n \in DX_n$  і значень  $R_n$ , що відповідають їм. Серед розв'язків обирають таке  $X_n^*$ , що забезпечує  $ext R_n$ .

**Етап 2.** Збільшують  $k$  на одиницю і записують відповідне функціональне рівняння:

$$f_k(S_{n-k}) = \underset{X_{n-k+1}}{ext} [R_{n-k+1}(S_{n-k}, X_{n-k+1}) + f_{k-1}(S_{n-k+1})].$$

Для кожного можливого стану  $S_{n-k}$  знаходять значення  $R_{n-k+1}$  у залежності від розв'язку  $X_{n-k+1}$ . Потім порівнюють суми  $R_{n-k+1} + f_{k-1}$  і визначають

екстремальну суму для кожного стану  $S_{n-k}$  і відповідний умовно-оптимальний розв'язок  $X_{n-k+1}^*$ , тобто те значення, за якого функція  $f_n(S_{n-k}^*)$  приймає оптимальне значення.

**Етап 3.** Перевірити  $k$ . Якщо  $k < n$ , перейти до етапу 2.

**Етап 4.** Записати функціональне рівняння для  $k = n$ :

$$f_n(S_0) = \underset{X_1}{ext} [R_1(S_0, X_1) + f_{n-1}(S_1)].$$

Тут  $f_n(S_0)$  – екстремальне значення цільової функції за усі  $n$  етапів, починаючи зі стану  $S_0$ . На даному етапі припущень про можливі стани процесу не роблять, тому що початковий стан  $S_0$  відомий. Для цього стану варто визначити оптимальний розв'язок з урахуванням всіх отриманих умовно-оптимальних розв'язків другого етапу.

**Етап 5.** Проходять весь процес у прямому напрямку від  $S_0$  до  $S_n$  і визначають оптимальний розв'язок  $X^*$  для всього процесу, що надає цільовій функції  $f_n(S_0)$  екстремального значення.

### 15.3 Використання методу ДП у задачі про вибір траєкторії

Розглянемо таку задачу.

**Задача 7.** Потрібно перевезти вантаж із міста 1 у місто 10. Мережа доріг і вартість перевезення одиниці вантажу з міста  $s = \overline{1,9}$  у місто  $j = \overline{2,10}$  показана на рис. 15.1. Необхідно визначити маршрут доставки вантажу з міста 1 у місто 10, якому відповідають мінімальні сумарні витрати.

#### Розв'язання

Розіб'ємо всю множину вершин на підмножини  $\{1\}$ ,  $\{2, 3, 4\}$ ,  $\{5, 6, 7\}$ ,  $\{8, 9\}$ ,  $\{10\}$ . Процес розв'язання задачі розбивається на 4 етапи. Перенумеруємо етапи від кінцевої вершини до початкової і введемо позначення:  $k = \overline{1,4}$  – номер кроку;  $c_{sj}$  – вартість перевезення вантажу з міста  $s$  до міста  $j$ ;  $f_k(s)$  – мінімальні витрати на перевезення вантажу від міста  $s$  до кінцевого міста, якщо до кінцевого міста залишилося  $k$  кроків;  $J_k(s)$  – номер міста, через який потрібно їхати з міста  $s$ , щоб досягти  $f_k(s)$ .

Запишемо рівняння для першого етапу ( $k = 1$ ):

$$f_1(s) = \min_j [c_{sj} + f_0(10)] = \min_j c_{sj}.$$

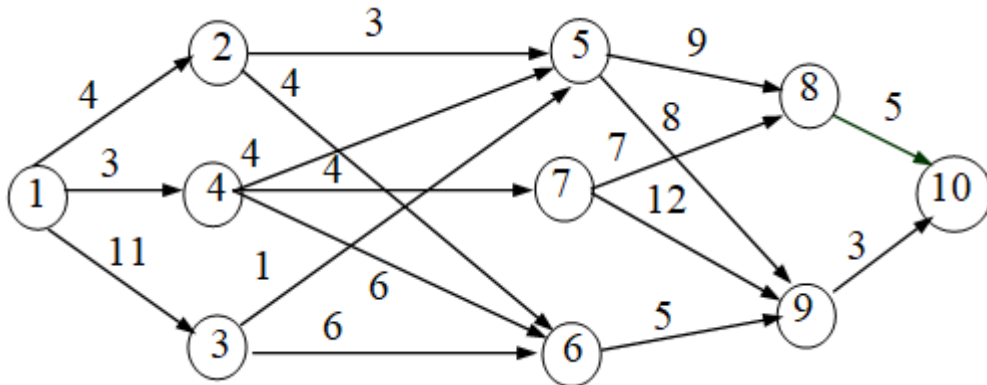


Рисунок 15.1 – Мережа доріг і вартість перевезення одиниці вантажу

У місто 10 вантаж може бути доставлений із міста 8 або 9. Тому обчислюємо

$$f_1(8) = c_{8,10} = 5; \quad s = 8, J_1(8) = 10;$$

$$f_1(9) = c_{9,10} = 3; \quad s = 9, J_1(9) = 10.$$

Ці дані можна сформуванати в табл. 15.1.

Таблиця 15.1 – План етапу 1

$j \backslash s$	10	$f_1(s)$	$J_1(s)$
8	5	5	10
9	3	3	10

Для другого етапу ( $k = 2$ ) функціональне рівняння має вигляд:

$$f_2(s) = \min_j [c_{sj} + f_1(j)].$$

Маємо:

$$f_2(5) = \min [c_{58} + f_1(8), c_{59} + f_1(9)] = \min [9 + 5, 8 + 3] = 11.$$

Одержуємо  $s = 5$ ,  $J_2(5) = 9$ , тобто умовно-оптимальний план проходить через місто 9. Аналогічно знаходяться інші значення й заповнюється табл. 15.2.

Таблиця 15.2 – Умовно-оптимальний план етапу 2

$j \backslash s$	8	9	$f_2(s)$	$J_2(s)$
5	9+5	8+3	11	9
6		5+3	8	9
7	7+5	12+3	12	8

Для третього етапу ( $k = 3$ ) функціональне рівняння має вигляд:

$$f_3(s) = \min_j [c_{sj} + f_2(j)].$$

Значення  $f_2(j)$  береться з табл. 15.2 і заповнюється табл. 15.3.

Таблиця 15.3 – Умовно-оптимальний план етапу 3

$j \backslash s$	5	6	7	$f_3(s)$	$J_3(s)$
2	3+11	4+8		12	6
3	1+11	6+8		12	5
4	4+11	6+8	4+12	14	6

Для останнього етапу:

$$f_4(s) = \min_j [c_{sj} + f_3(j)].$$

У результаті одержуємо табл. 15.4.

Таблиця 15.4 – Оптимальний план останнього етапу

$j \backslash s$	2	3	4	$f_4(s)$	$J_4(s)$
1	4+12	11+12	3+14	16	2

З останньої таблиці видно, що мінімальні витрати на перевезення вантажу складають  $f_4(1) = 16$  й оптимальний маршрут проходить через місто 2. Далі, рухаючись за таблицями, знаходимо оптимальний маршрут  $1 \rightarrow 2 \rightarrow 6 \rightarrow 9 \rightarrow 10$ .

**Відповідь:** оптимальний маршрут  $1 \rightarrow 2 \rightarrow 6 \rightarrow 9 \rightarrow 10$  забезпечує мінімальні витрати на перевезення вантажу у 16 ум. од.

### Питання для самоперевірки

1. Сформулюйте принцип оптимальності Беллмана.
2. Наведіть функціональне рівняння Беллмана.
3. Наведіть алгоритм методу динамічного програмування.
4. Опишіть основні етапи використання методу динамічного програмування у задачі про вибір траєкторії.

## СПИСОК РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

1. Катренко, А. В. Дослідження операцій [Текст] : підруч. / А. В. Катренко. – Л : Магнолія – 2006, 2009. – 352 с.
2. Григорків, В. С. Оптимізаційні методи та моделі : підручник / В. С. Григорків, М. В. Григорків. – Чернівці : Чернівецький нац. ун-т, 2016. – 400 с.
3. Таха Хэмди. Введение в исследование операций / Таха Хэмди. – М. : Вильямс. 2001. – 912 с.
4. Синеглазов, В. М. Математичні методи оптимізації : навч. посіб. / В. М. Синеглазов, О. А. Зеленков, Ш. І. Аскеров. ; Нац. Авіаційний ун-т. – К. : Освіта України, 2018. – Ч. 1. – 329 с.
5. Фон Нейман, Дж. Теория игр и экономическое поведение / Дж. Фон Нейман, Моргенштерн О. – М. : Наука, 1970. – 708 с.
6. Шиян, А. А. Теорія ігор: основи та застосування в економіці та менеджменті : навчальний посібник / А. А. Шиян // . – Вінниця : ВНТУ, 2009. – 164 с.
7. Бех, О. В. Математичне програмування [Текст] : навч. посіб. / О. В. Бех, Т. А. Городня, А. Ф. Щербак. – Л. : Магнолія – 2006, 2009. – 200 с.
8. Исследование операций в экономике: Учебное пособие для вузов / Кремер Н.Ш., Путко Б.А., Тришин И.Н., Фридман М.Н.; Под ред. проф. Н.Ш. Кремера // – М.: ЮНИТИ, 2003. – 407 с.
9. Snyman, J. A.; Wilke, D. N. (2018). Practical Mathematical Optimization: Basic Optimization Theory and Gradient-Based Algorithms (2nd ed.). Berlin: Springer. ISBN 978-3-319-77585-2.
10. Optimization Methods and Software. – Режим доступу: <https://www.tandfonline.com/toc/goms20/current>
11. Васильєва, Л. В. Математичні методи дослідження операцій : навч. посібник / Васильєва Л. В., Богдан М. П. – Краматорськ : ДДМА, 2018. – 144 с.
12. Васильєва, Л. В. Математичне моделювання в біотехнічних системах : навч. посіб. для студентів закладів вищої освіти спеціальності 122 «Комп'ютерні науки» / Васильєва Л. В., Тарасов О. Ф. – Житомир : Євро-Волинь, 2021. – 104 с.

13. Васильєва, Л. В. Використання комп'ютерних технологій при вирішенні задач лінійного програмування : навчальний посібник для студентів вищих навчальних закладів / Васильєва Л. В., Гетьман І. А. – Краматорськ : ДДМА, 2005. – 112 с.

14. Гончаров, О. А. Чисельні методи розв'язання прикладних задач : навч. посіб. / О. А. Гончаров, Л. В. Васильєва, А. М. Юнда. – Суми : СумДУ, 2020. – 142 с.

15. Тарасов, О. Ф. Сучасні методи проєктування програмних систем на основі ООП : навч. посіб. для студентів закладів вищої освіти спеціальності 122 «Комп'ютерні науки» / Тарасов, О. Ф., Васильєва Л. В. – Житомир : Євро-Волинь, 2021. – 68 с.

16. Шевцов, С. О. Розв'язання задач з економіки методами математичного аналізу : посібник до практичних занять і самостійної роботи / Шевцов С. О., Грудкіна Н. С. – Краматорськ : ДДМА, 2019. – 55 с.

*Навчальне видання*

**ВАСИЛЬЄВА Людмила Володимирівна,  
ГРУДКІНА Наталія Сергіївна**

# **МАТЕМАТИЧНІ МЕТОДИ ДОСЛІДЖЕННЯ ОПЕРАЦІЙ**

**Курс лекцій**

Редагування, комп'ютерне верстання

Я. О. Бершацька

51/2022. Формат 60 × 84/16. Ум. друк. арк. 6,47.  
Обл.-вид. арк. 5,35. Тираж 100 пр. Зам. № 5

Видавець і виготівник  
Донбаська державна машинобудівна академія  
84313, м. Краматорськ, вул. Академічна, 72.  
Свідоцтво суб'єкта видавничої справи  
ДК № 1633 від 24.12.2003