

## Методи оптимізації. Лекція 25.03.2022

### Класифікація задач оптимізації за деякими аспектами їхніх постановок

1. *За наявністю умов (обмежень):*
  - задачі безумовної оптимізації:  $f(x) \rightarrow \min, x \in E^n$ ;
  - задачі умовної оптимізації:  $f(x) \rightarrow \min, x \in X \subset E^n$ .
2. *За кількістю змінних:*
  - задачі одновимірної оптимізації: цільова функція залежить від однієї змінної,  $X \subset E^1$ ;
  - задачі скінечновимірної оптимізації, тобто  $n \geq 2, n < \infty$ ;
  - задачі нескінечновимірної оптимізації (мінімізації функціоналів).
3. *За виглядом допустимої області:*
  - задачі неперервної оптимізації;
  - задачі дискретної оптимізації.
4. *За наявністю випадкового фактору:*
  - детерміновані (немає випадкового фактору);
  - стохастичні (є випадковий фактор).
5. *За кількістю точок екстремуму:*
  - однократноекстремальні;
  - багатократноекстремальні.
6. *Із урахуванням часу:*
  - статичні (час не враховується);
  - динамічні.

Задача математичного програмування в загальному вигляді може бути сформульована так:

$$f(x) \rightarrow \min,$$

$$x \in X = \left\{ x \in E^n : x \in X_0, g_i(x) \leq b_i, i = \overline{1, m}, g_i(x) = b_i, i = \overline{m+1, p} \right\}.$$

Функції  $f(x)$  та  $g_i(x)$  вважаються опуклими функціями.

Тут  $X_0$  – множина простої структури. Такими множинами є: невід'ємний октант,  $n$ -мірна куля і  $n$ -мірний паралелепіпед та деякі інші.

На рисунку на площині показані:

а)  $X_0 = \{ x \in E^2, x = (x_1, x_2) : x_i \geq 0, i = 1, 2 \}$ . Перша чверть координатної площини або невід'ємний октант при  $n = 2$ ;

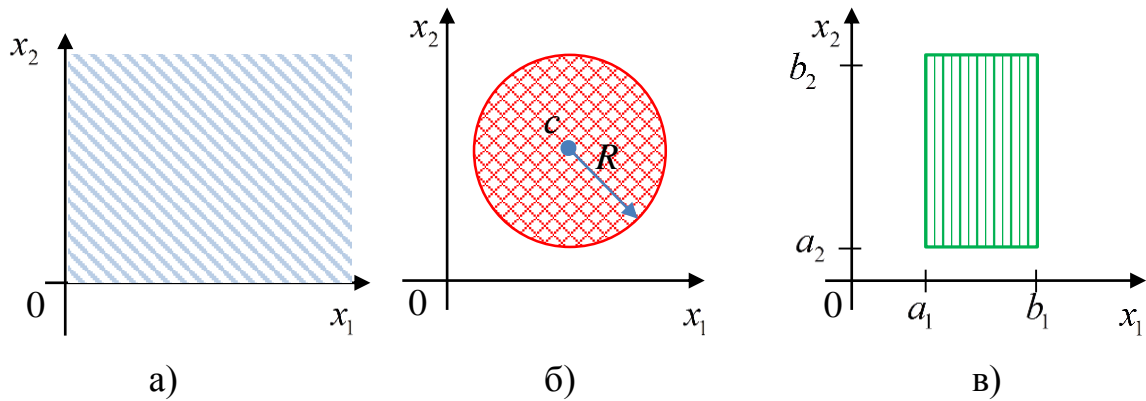
б) 2-вимірна куля – це коло.

$$X_0 = \left\{ x \in E^2, x = (x_1, x_2) : (x_1 - c_1)^2 + (x_2 - c_2)^2 \leq R^2 \right\},$$

де  $c = (c_1, c_2)$  – центр кола,  $R$  – радіус;

в)  $n$ -мірний паралелепіпед на площині – прямокутник.

$$X_0 = \{ x \in E^2, x = (x_1, x_2): a_i \leq x_i \leq b_i, i = 1, 2 \}.$$



Задачі математичного програмування діляться на

- задачі лінійного програмування (ЛП),
- задачі нелінійного програмування (НЛП).

Постановка задачі лінійного програмування.

$$f(x) = \sum_{j=1}^n c_j x_j \rightarrow \max(\min),$$

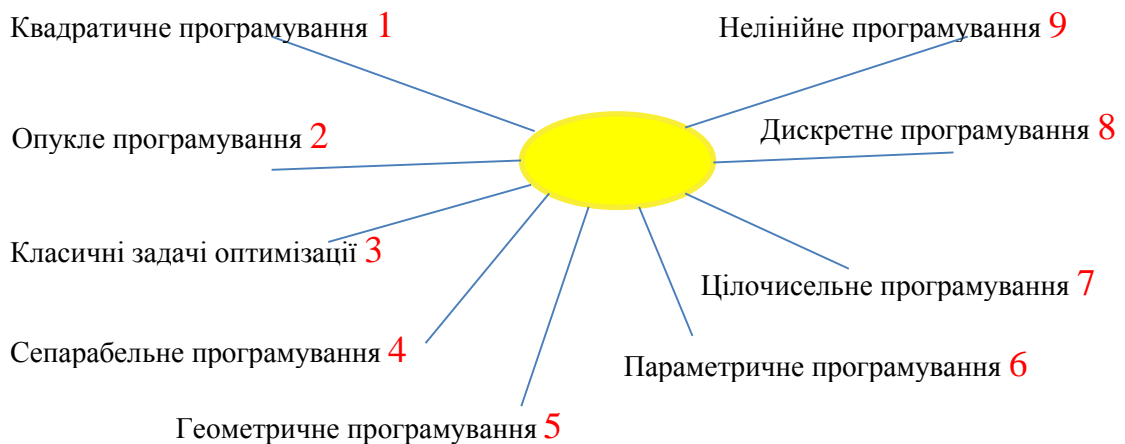
$$\sum_{i=1}^m a_{ij} x_j \leq b_i, \quad i = \overline{1, p},$$

$$\sum_{i=1}^m a_{ij} x_j = b_i, \quad i = \overline{p+1, m},$$

$$x_j \geq 0, \quad i = \overline{1, s}, \quad s \leq n.$$

Цільова функція і обмеження є лінійними. Основним методом розв'язання задач ЛП є симплекс-метод.

### Класифікація задач нелінійного програмування (НЛП)



### 1. Квадратичне програмування.

$$f(x) = (Ax, x) - (b, x) \rightarrow \min,$$

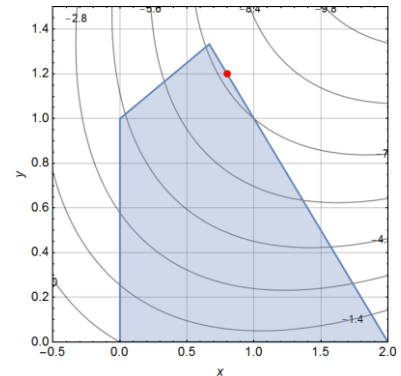
де  $A$  – симетрична додатно визначена матриця, функції  $g_i(x)$  – лінійні функції,  $b \in E^n$ .

Приклад.

$$f(x) = x_1^2 + 2x_2^2 - 2x_1x_2 - 2x_1 - 6x_2 \rightarrow \min,$$

$$\begin{cases} x_1 + x_2 \leq 2 \\ -x_1 + 2x_2 \leq 2 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0.$$



2. Опукле програмування. Функції  $f(x)$ ,  $g_i(x)$ ,  $i = \overline{1, m}$  – опуклі функції, визначені на опуклій множині.

3. Класичні задачі оптимізації. Функція  $f(x)$  є неперервно-диференційовною.

- задача безумовної оптимізації (без обмежень):

$$f(x) \rightarrow \min, x \in E^n. f(x) \in C^1(E^n).$$

Метод розв'язання:  $f'(x) = 0$ .

- задача умовної оптимізації:

$$f(x) \rightarrow \min, x \in X \subset E^n. f(x) \in C^1(X).$$

Метод розв'язання: метод множників Лагранжа.

4. Сепарабельне програмування. Функції  $f(x)$ ,  $g_i(x)$ ,  $i = \overline{1, m}$  – сепарабельні функції. Тобто функції вигляду:  $f(x) = \sum_{j=1}^n f_j(x_j)$ .

Приклад.

$$f(x) = x_1^2 + x_2^2 - 2x_1x_2 - 6x_1 - 8x_2 - 5x_3 \rightarrow \min,$$

$$\begin{cases} x_1 + x_2 + x_3 \leq 5 \\ x_1^2 + x_1x_2 - x_2 \leq 3 \end{cases}$$

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0.$$

5. Геометричне програмування. Функції  $f(x)$  и  $g_i(x)$ ,  $i = \overline{1, p}$  – ПОЗИНОМИ.

$$f(x) = \sum_{j=1}^n c_j \cdot (x_1)^{\alpha_{j1}} \cdot (x_2)^{\alpha_{j2}} \cdot \dots \cdot (x_n)^{\alpha_{jn}},$$

$$g_i(x) < 1, i = \overline{1, p},$$

$$c_j > 0, \alpha_{jk} - \text{дійсні числа.}$$

6. **Параметричне програмування.** Цільова функція і/або обмеження залежать від параметру.

Наприклад, параметрична задача лінійного програмування:

$$f(x) = \sum_{j=1}^n (c'_j + c''_j t) x_j \rightarrow \max,$$

$$\sum_{j=1}^m (a'_{ij} + a''_{ij} t) x_j = b'_i + b''_i t, \quad i = \overline{1, m},$$

$$x_j \geq 0, \quad j = \overline{1, n}, \quad t \in [\alpha, \beta].$$

7. **Цілочисельне програмування.** Змінні задачі вважаються цілочисельними.

8. **Дискретне програмування.** Множина допустимих розв'язків є дискретною множиною.

9. **Власне задачі НЛП,** тобто задачі які розв'язуються загальними методами розв'язання задач НЛП.

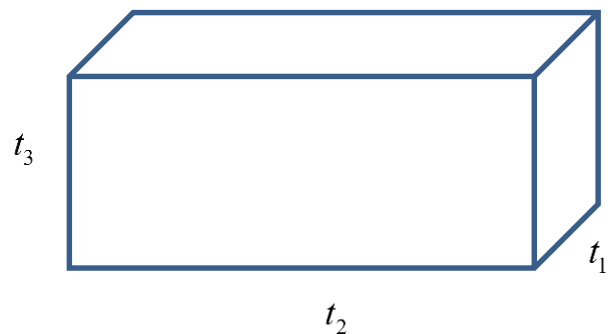
**Приклад.** Задача геометричного програмування.

Нехай потрібно переправити через річку  $400 \text{ м}^3$  гравію. Припустимо, що гравій вантажиться у відкритий ящик завдовжки  $t_1 \text{ м}$ , шириною  $t_2 \text{ м}$  і висотою  $t_3 \text{ м}$ . Бічні сторони і дно ящика виготовлені з матеріалу, вартість  $1 \text{ м}^2$  якого 10 грошових одиниць, а передня і задня стінки – із матеріалу по 20 грошових одиниць за  $\text{м}^2$ . Кожне перевезення ящика будь-якого розміру з одного берегу на інший і назад коштує 0,1 грошова одиниця, причому після його використання ящик не матиме залишкової вартості. Чому дорівнює мінімальна вартість транспортування  $400 \text{ м}^3$  гравію?

Очевидно, що при лінійних розмірах ящика  $t_1, t_2, t_3$  число рейсів, які потрібно виконати для перевезення  $400 \text{ м}^3$  гравію

становить  $\frac{400}{t_1 \cdot t_2 \cdot t_3}$ , а вартість

перевезень:  $0.1 \frac{400}{t_1 \cdot t_2 \cdot t_3}$ .



Загальна вартість матеріалу становить

$$40 \cdot t_2 \cdot t_3 + 20 \cdot t_1 \cdot t_3 + 10 \cdot t_1 \cdot t_2.$$

Отже, сумарна вартість перевезень з врахуванням вартості матеріалу становить:

$$g(t) = \frac{40}{t_1 \cdot t_2 \cdot t_3} + 40 \cdot t_2 \cdot t_3 + 20 \cdot t_1 \cdot t_3 + 10 \cdot t_1 \cdot t_2.$$

Природними обмеженнями будуть умови додатності змінних:

$$t_1 > 0, \quad t_2 > 0, \quad t_3 > 0.$$

Оптимальним розв'язком цієї задачі є  $t_1 = 2$ ,  $t_2 = 1$ ,  $t_3 = 0.5$ .  $g(t^*) = 100$ .

### Класичний метод пошуку екстремуму

Нехай  $f(x)$  є диференційовна у просторі  $E^n$ , тобто  $f(x) \in C^1(E^n)$ .

Тоді точками локального мінімуму або максимуму функції  $f(x)$  у просторі  $E^n$  можуть бути лише ті точки  $x$  у просторі, в яких  $f'(x) = 0$ , тобто

$$\frac{\partial f(x)}{\partial x_k} = 0, \quad k = \overline{1, n}. \quad (1)$$

Усі точки, які задовольняють системі (1) називають стаціонарними точками функції  $f(x)$ .

Сюди входять точки мінімуму, максимуму, а також точки, в яких функція не має ні мінімуму ні максимуму.

Залучаючи достатні умови або необхідні умови більш високого порядку, ніж перший, серед стаціонарних точок вибираємо точки мінімуму (максимуму).

У тих випадках, коли вдається виявити всі точки локального мінімуму (максимуму) функції  $f(x)$ , для визначення глобального мінімуму (максимуму) цієї функції у просторі  $E^n$  потрібно перебрати всі точки локального мінімуму (максимуму) та із них вибрати точку з найменшим (найбільшим) значенням функції.

**Приклад.** Знайти екстремуми функції

$$f(x, y) = -\frac{1}{9}(3x + 5y)^3 + \frac{3}{2}x^2 + \frac{5}{2}y^2 + 5xy - 2y.$$

Знайдемо похідні цієї функції

$$\frac{\partial f(x, y)}{\partial x} = -(3x + 5y)^2 + 3x + 5y,$$

$$\frac{\partial f(x, y)}{\partial y} = -\frac{5}{3}(3x + 5y)^2 + 5y + 5x - 2.$$

Знайдемо розв'язок системи

$$\begin{cases} -(3x + 5y)^2 + 3x + 5y = 0, \\ -\frac{5}{3}(3x + 5y)^2 + 5y + 5x - 2 = 0, \end{cases} \Rightarrow \begin{cases} (3x + 5y)[1 - 3x - 5y] = 0, \\ -\frac{5}{3}(3x + 5y)^2 + 5x + 5y - 2 = 0. \end{cases}$$

Ця система розпадається на дві системи

$$1) \begin{cases} 3x + 5y = 0, \\ 5x + 5y - 2 = 0. \end{cases} \Rightarrow \begin{cases} x = 1, \\ y = -\frac{3}{5}. \end{cases}$$

$$2) \begin{cases} 3x + 5y = 1, \\ 5x + 5y = \frac{11}{3}. \end{cases} \Rightarrow \begin{cases} x = \frac{4}{3}, \\ y = -\frac{3}{5}. \end{cases}$$

Точки  $M_0\left(1, -\frac{3}{5}\right)$  та  $M_1\left(\frac{4}{3}, -\frac{3}{5}\right)$  є стаціонарними точками.

Перевіримо критерій Сильвестра додатної визначеності матриці: матриця  $f''(x)$  **додатно визначена** в тому і тільки в тому випадку, коли всі її головні (кутові) мінори додатні.

Нагадаємо, що головними (кутовими) мінорами матриці  $A = \{a_{ij}\}$ ,  $i, j = \overline{1, n}$  називаються визначники:

$$\Delta_1 = a_{11}, \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \dots, \Delta_n = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

Матриця  $A = \{a_{ij}\}$ ,  $i, j = \overline{1, n}$  буде **від'ємно визначеною** в тому і тільки в тому випадку, коли мають місце такі знаки головних (кутових) мінорів:

$$\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \Delta_4 > 0, \dots$$

Обчислимо другі похідні функції  $f(x)$ :

$$\frac{\partial^2 f(x, y)}{\partial x^2} = -6(3x + 5y) + 3,$$

$$\frac{\partial^2 f(x, y)}{\partial x \partial y} = -10(3x + 5y) + 5.$$

$$\frac{\partial^2 f(x, y)}{\partial y \partial x} = -10(3x + 5y) + 5,$$

$$\frac{\partial^2 f(x, y)}{\partial y^2} = -\frac{50}{3}(3x + 5y) + 5.$$

$$f''(x, y)|_M = \begin{pmatrix} \left. \frac{\partial^2 f(x, y)}{\partial x^2} \right|_M & \left. \frac{\partial^2 f(x, y)}{\partial x \partial y} \right|_M \\ \left. \frac{\partial^2 f(x, y)}{\partial x \partial y} \right|_M & \left. \frac{\partial^2 f(x, y)}{\partial y^2} \right|_M \end{pmatrix}.$$

Обчислимо другі похідні функції  $f(x)$  в точці  $M_0\left(1, -\frac{3}{5}\right)$ :

$$f''(M_0) = \begin{pmatrix} 3 & 5 \\ 5 & 5 \end{pmatrix}, \Delta_1 = 3 > 0, \Delta_2 = \begin{vmatrix} 3 & 5 \\ 5 & 5 \end{vmatrix} = -10 < 0.$$

В точці  $M_0\left(1, -\frac{3}{5}\right)$  функція  $f(x, y)$  не має екстремуму.

Досліджуємо другу стаціонарну точку  $M_1\left(\frac{4}{3}, -\frac{3}{5}\right)$ .

Обчислимо другі похідні функції  $f(x, y)$  в точці  $M_1$ :

$$f''(M_1) = \begin{pmatrix} -3 & -5 \\ -5 & -\frac{35}{3} \end{pmatrix}, \Delta_1 = -3 < 0, \Delta_2 = \begin{vmatrix} -3 & -5 \\ -5 & -\frac{35}{3} \end{vmatrix} = 10 > 0.$$

Матриця других похідних  $f''(M_1)$  від'ємно визначена. Точка  $M_1\left(\frac{4}{3}, -\frac{3}{5}\right)$  є точка максимуму функції  $f(x, y)$ .

В класичному математичному аналізі традиційно розглядається така задача на умовний екстремум: знайти екстремуми функції  $f(x)$  за умови, що змінні  $x$  задовольняють обмеженням

$$g_1(x) = 0, \dots, g_s(x) = 0, x \in E^n.$$

Вважається, що  $f(x)$  та  $g_i(x)$ ,  $i = \overline{1, s}$  є визначеними та мають неперервні частинні похідні першого порядку у всьому просторі  $E^n$ .

Методами розв'язання такого класу задач є

- метод виключення частини змінних,
- метод множників Лагранжа.

### **Метод множників Лагранжа для задачі умовної оптимізації з обмеженнями у формі рівностей**

Нехай потрібно розв'язати таку задачу

$$f(x) \rightarrow \min \quad (1)$$

за умови

$$x \in X = \{x \in E^n : g_i(x) = 0, i = \overline{1, s}\}. \quad (2)$$

Припустимо, що функції  $f(x), g_1(x), \dots, g_s(x)$  визначені та диференційовні на множині  $X$  і  $s < n$ .

Основна ідея методу множників Лагранжа полягає у переході від задачі на умовний екстремум до задачі знаходження безумовного екстремуму спеціально побудованої функції Лагранжа з наступним застосуванням класичного методу пошуку екстремуму.

Введемо змінні  $\lambda_0, \lambda_1, \dots, \lambda_s$  і побудуємо функцію Лагранжа

$$L(x, \lambda) = \lambda_0 f(x) + \sum_{i=1}^s \lambda_i g_i(x). \quad (3)$$

Сформулюємо необхідну умову екстремуму для задачі (1),(2).

**Теорема (Ознака Лагранжа).** Для того, щоб вектор  $x_* = (x_1^*, \dots, x_n^*)$  був розв'язком задачі (1), (2), **необхідно** існування вектору  $\lambda^* = (\lambda_0^*, \lambda_1^*, \dots, \lambda_s^*)$ , такого що його компоненти  $\lambda_i^*$  ( $i = \overline{1, s}$ ) одночасно не дорівнюють нулеві і такі, що всі частинні похідні першого порядку функції Лагранжа  $L(x, \lambda)$  за змінними  $x_j$  ( $j = \overline{1, n}$ ) в точці  $x_*$  дорівнюють нулеві, тобто

$$\frac{\partial L(x_*, \lambda^*)}{\partial x_k} = \lambda_0^* \frac{\partial f(x_*)}{\partial x_k} + \sum_{i=1}^s \lambda_i^* \frac{\partial g_i(x_*)}{\partial x_k} = 0, \quad k = \overline{1, n}, \quad \lambda^* \neq 0. \quad (4)$$

Умова (4) має ясний геометричний сенс. Вона означає, що градієнти  $f'(x_*)$ ,  $g_1'(x_*)$ ,  $\dots$ ,  $g_s'(x_*)$  є лінійно залежними. У випадку, коли  $s = 1$ , вектори  $f'(x_*)$ ,  $g_1'(x_*)$  повинні бути колінеарними.

Якщо  $\lambda_0 \neq 0$ , то задачу (3) називають не виродженою (регулярною), при  $\lambda_0 = 0$  – виродженою. У багатьох практичних задачах  $\lambda_0 = 1$ .

Відзначимо, що вимірність задачі (1), (2) дорівнює  $n$ , а вимірність задачі (3) є  $n + s + 1$ .

### Алгоритм методу множників Лагранжа

**Крок 1.** Будуємо функцію Лагранжа

$$L(x, \lambda) = f(x) + \sum_{i=1}^s \lambda_i g_i(x)$$

та складаємо систему

$$\frac{\partial L(x, \lambda)}{\partial x_k} = 0, \quad k = \overline{1, n}, \quad \frac{\partial L(x, \lambda)}{\partial \lambda_i} = g_i(x) = 0, \quad i = \overline{1, s}.$$

**Крок 2.** Розв'язавши отриману систему алгебраїчних рівнянь якимось з відомих методів, отримаємо стаціонарні точки  $x$  функції Лагранжа.

**Крок 3.** Залучаючи необхідні умови більш високих порядків або достатні умови, або аналізуючи поведінку цільової функції в околі стаціонарних точок, виберемо серед стаціонарних точок точки мінімуму.

**Метод множників Лагранжа для задачі умовної оптимізації з обмеженнями у формі нерівностей**

Знайдемо розв'язок задачі

$$f(x) \rightarrow \min \quad (5)$$

за умов

$$x \in X = \left\{ x \in E^n : g_i(x) \leq 0, i = \overline{1, s} \right\}. \quad (6)$$

Допустимо, що функції  $f(x)$ ,  $g_1(x)$ ,  $\dots$ ,  $g_s(x)$  диференційовні та  $s < n$ .

У цьому випадку застосовується такий алгоритм:



### Алгоритм методу

**Крок 1.** Розглянемо задачу

$$f(x) \rightarrow \min, \quad (5)$$

$$g_i(x) < 0, \quad i = \overline{1, s}. \quad (8)$$

Задачу (5), (8) розв'язуємо класичним методом, тобто знаходимо стаціонарні точки, які є розв'язками системи алгебраїчних рівнянь

$$\frac{\partial f(x)}{\partial x_k} = 0, \quad k = \overline{1, n},$$

потім перевіряємо чи задовольняють ці точки умові (8). Якщо не задовольняють, виключаємо їх з подальшого розгляду, у протилежному випадку – досліджуємо чи є ці стаціонарні точки точками мінімуму цільової функції.

**Крок 2.** Розв'язуємо задачу

$$f(x) \rightarrow \min,$$

$$g_i(x) = 0, \quad i = \overline{1, s}$$

методом множників Лагранжа для обмежень у формі рівностей.

**Крок 3.** Серед всіх знайдених точок вибираємо точки мінімуму.

Алгоритм описаний.

**Приклад 1.** Знайти мінімум функції

$$f(x_1, x_2) = 4x_1^2 + x_2^2, \text{ при умові } x_1 + x_2 = 5.$$

Складаємо функцію Лагранжа

$$L(x, \lambda) = 4x_1^2 + x_2^2 + \lambda(x_1 + x_2 - 5).$$

Складаємо систему рівнянь

$$\frac{\partial L(x, \lambda)}{\partial x_1} = 8x_1 + \lambda = 0,$$

$$\frac{\partial L(x, \lambda)}{\partial x_2} = 2x_2 + \lambda = 0,$$

$$\frac{\partial L(x, \lambda)}{\partial \lambda} = x_1 + x_2 - 5 = 0.$$

Розв'язуємо отриману систему

$$\begin{cases} 8x_1 + \lambda = 0, \\ 2x_2 + \lambda = 0, \\ x_1 + x_2 - 5 = 0, \end{cases} \Rightarrow \begin{cases} x_1 = 1, \\ x_2 = 4, \\ \lambda = -8. \end{cases}$$

Для дослідження отриманої стаціонарної точки застосуємо критерій Сильвестра. Для функції  $f(x)$

$$f'(x) = \begin{pmatrix} 8x_1 \\ 2x_2 \end{pmatrix}, \quad f''(x) = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}, \quad \Delta_1 = 8 > 0, \quad \Delta_2 = 16 > 0.$$

Згідно з цим критерієм, маємо, що цільова функція в точці  $(1,4)$  має єдиний мінімум.

