Розділ 4
НЕЛІНІЙНЕ ПРОГРАМУВАННЯ

4.1. Постановка задачі нелінійного програмування

Загальна задача нелінійного програмування (ЗНП) визначається як задача знаходження мінімуму або максимуму функції $f(x_1, x_2, ..., x_n)$ на множині $G, G \subset R^n$, яка визначається системою обмежень

$$\begin{cases} g_i(x_1, x_2, ..., x_n) \le b_i, i = \overline{1,k}, \\ g_i(x_1, x_2, ..., x_n) = b_i, i = \overline{k+1,m}, \end{cases}$$
$$x_j \in R, j = \overline{1,n},$$

де хоча б одна з функцій $f, g_i, i = \overline{1,m}$ є нелінійною.

Таким чином, загальна постановка ЗНП має вигляд

$$f(x_1, x_2, ..., x_n) \to \max(\min), \tag{4.1}$$
$$\begin{cases} g_i(x_1, x_2, ..., x_n) \le b_i, i = \overline{1,k}, \\ g_i(x_1, x_2, ..., x_n) = b_i, i = \overline{k+1,m}, \end{cases} \tag{4.2}$$
$$x_j \in R, j = \overline{1,n}. \tag{4.3}$$

Умови (4.3) можуть бути замінені умовами невід'ємності, цілочисловості деяких змінних залежно від потреб практичних задач.

Якщо умови (4.2), (4.3) відсутні, то така задача називається *задачею безумовної оптимізації*.

Аналогічно до ЗЛП (задачі лінійного програмування) область G, яка визначається системою обмежень (4.2), вираз (4.3) називається *областю допустимих розв'язків* (ОДР), сукупність значень змінних $X = (x_1, x_2, \ldots, x_n)$, які задовольняють систему обмежень (4.2), (4.3) - *допустимим планом ЗНП*, а функція $f(x_1, x_2, ..., x_n)$ - *цільовою функцією*.

Оптимальним планом $X^* = \left(x_1^*, x_2^*, \ldots, x_n^*\right)$ *або точкою глобального максимуму (мінімуму)* називають допустимий план, на якому досягається максимум або мінімум функції f.

Отримане значення $f\left(X^*\right)$ називають *глобальним максимумом (мінімумом)*.

Зауважимо, що в теорії нелінійного програмування існує також поняття локального максимуму (мінімуму).

Точка $X^0 = \left(x_1^0, x_2^0, \ldots, x_n^0\right)$ називається точкою *локального максимуму (мінімуму)*, якщо існує такий проколотий окіл точки X^0, який належить ОДР і для будь-якої точки $X = (x_1, x_2, \ldots, x_n)$ цього околу виконується нерівність $f(X) < f\left(X^0\right)$ ($f(X) > f\left(X^0\right)$), при цьому число $f\left(X^0\right)$ називається *локальним максимумом (мінімумом)* функції $f(x_1, x_2, ..., x_n)$.

У ЗНП цільова функція може досягати свого екстремального (максимального або мінімального) значення як на границі ОДР, так і всередині неї, на відміну від задач лінійного програмування. Сама ж ОДР може бути неопуклою або, у разі опуклості, мати нескінчену кількість кутових точок (рис. 4.1).

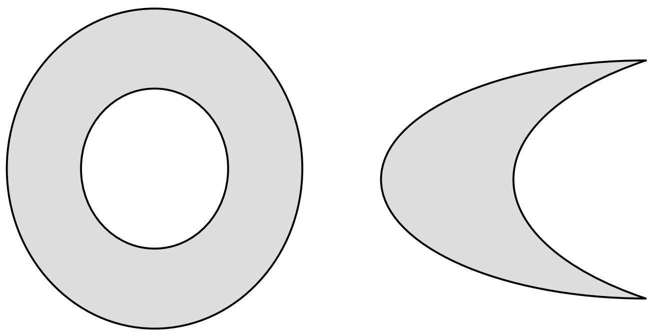

Рис. 4.1. Приклади неопуклих областей

Ряд практичних задач пов’язано зі знаходженням оптимального розв’язку за наявності деякої кількості обмежень на змінні, що суттєво ускладнює процедуру пошуку оптимуму. При цьому може порушуватися навіть основна умова, відповідно до якої екстремум повинен досягатися в критичних точках, тобто точках, у яких частинні похідні першого порядку або дорівнюють нулю, або не існують.

Наприклад, безумовний мінімум функції однієї змінної $y=(x+3)^2$ досягається в стаціонарній точці $x=-3$ (точку, в якій похідна дорівнює нулю, називають *стаціонарною*). Але, якщо потрібно знайти мінімум цієї функції з урахуванням додаткового обмеження $x \geq 0$, то буде знайдений умовний мінімум, якому відповідає точка $x=0$, яка не є стаціонарною, оскільки $y'(0)=2(x+3)\big|_{x=0}=6$.

Для ЗНП не існує універсального методу розв’язання. У зв’язку з цим при розв’язанні таких задач застосовують спеціальні методи, які залежать від виду цільової функції (4.1) і системи обмежень (4.2). До таких методів належать графічний метод і метод множників Лагранжа, градієнтні методи, опукле та квадратичне програмування. ЗНП, на змінні яких не накладено додаткових умов (4.2), тобто задачі пошуку безумовного екстремуму функції, розв’язують методами класичного математичного аналізу.

Нелінійне програмування поділяють на два основних типи задач:

- задачі опуклого програмування;
- задачі неопуклого програмування.

Опукле програмування характеризується тим, що цільова функція є опуклою для задачі пошуку мінімуму (увігнутою при пошуку максимуму), а обмеження задають опуклу ОДР. Всі інші типи задач належать до задач неопуклого програмування.

У свою чергу, в опуклому програмуванні виділяють окремий клас задач – задачі квадратичного програмування, цільова функція яких є поліномом другого порядку, а ОДР задається системою лінійних обмежень.

4.2. Графічний метод розв’язання ЗНП

Графічний метод розв’язання ЗНП можна застосовувати у випадку, коли функції $f, g_i, i=\overline{1,m}$ залежать від двох змінних:

$$f(x_1, x_2) \to \max(\min), \tag{4.4}$$

$$\begin{cases} g_i(x_1,x_2) \le b_i, i=\overline{1,k}; \\ g_i(x_1,x_2) = b_i, i=\overline{k+1,m}; \end{cases} \tag{4.5}$$

$$x_1, x_2 \in R. \tag{4.6}$$

Алгоритм графічного розв'язання ЗНП:

1. Будують ОДР, яка визначається умовами (4.5)-(4.6). Якщо ця область порожня, то задача розв'язків не має.

2. Будують лінії рівня $f(x_1,x_2)=C$ при різних значеннях параметра *C*.

3. Визначають напрям зростання для задачі максимізації (або спадання для задачі мінімізації) ліній рівня цільової функції.

4. Визначають точку X^* з ОДР, через яку проходить лінія рівня з найбільшим значенням параметра *C* для задачі пошуку максимуму функції (4.4) (або з найменшим значенням параметра *C* у разі пошуку мінімуму). Якщо цільова функція необмежена зверху (для задачі максимізації) або знизу (для задачі мінімізації) на ОДР, то задача розв'язків немає.

5. Визначають координати знайденої точки $X^* = (x_1^*, x_2^*)$ і обчислюють значення цільової функції $f^* = f(X^*)$.

Приклад 1. За допомогою графічного методу знайти екстремальні значення функції

$$f(x_1,x_2) = x_1^2 + x_2^2$$

при обмеженнях

$$\begin{cases} x_1 + x_2 \ge 1, \\ 2x_1 + 3x_2 \le 6, \\ x_1 \le 2, \\ x_1, x_2 \ge 0. \end{cases}$$

Розв'язання. ***Знаходження мінімуму функції.***

1. Будуємо ОДР (рис. 4.2).

2. Лінії рівня $f(x_1,x_2) = x_1^2 + x_2^2 = C$ - це сім'я концентричних кіл з центром у початку координат і радіусом $R = \sqrt{C}$.

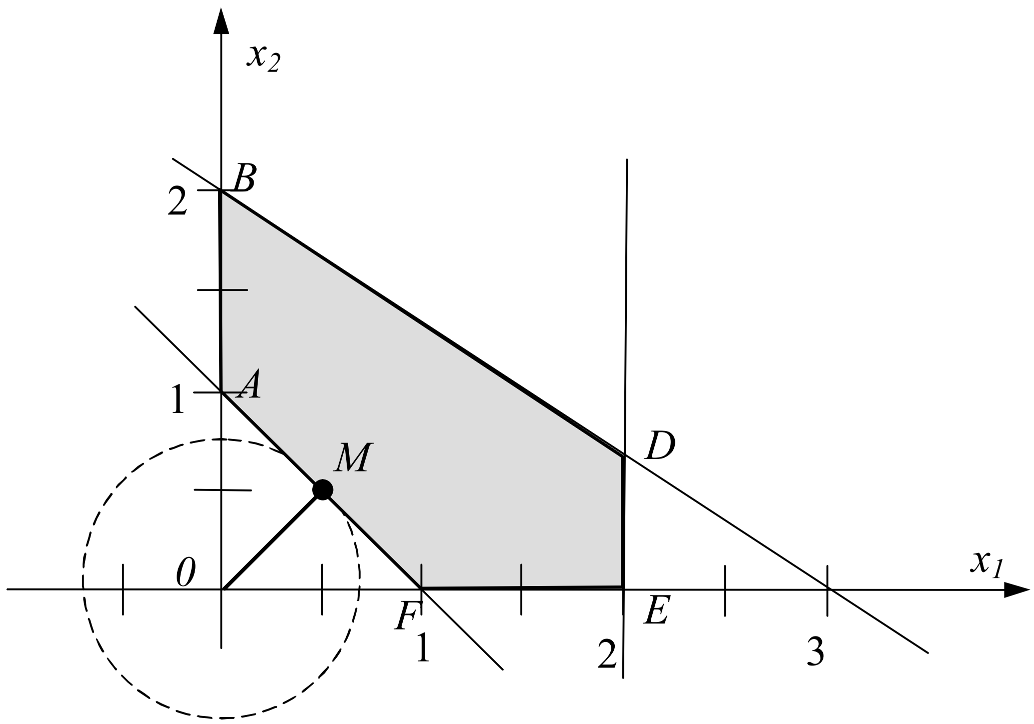

Рис. 4.2. Знаходження мінімуму функції прикладу 1

3. Визначаємо точку, яка може бути мінімумом. Для цього будуємо кола різних радіусів з центром у початку координат і відмічаємо точку, що належить ОДР і відповідає колу з найменшим радіусом. Очевидно, що це точка *M*.

4. Вектор-градієнт функції $grad\, f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right) = (2x_1, 2x_2)$ в точці (x_1, x_2) перпендикулярний до лінії рівня $f(x_1, x_2) = x_1^2 + x_2^2 = C$, що проходить через цю точку. Пряма $1 \cdot x_1 + 1 \cdot x_2 = 1$, яка є дотичною до кола $x_1^2 + x_2^2 = C$, має нормальний вектор $\vec{N} = (1{,}1)$. Оскільки вектори $grad\, f = (2x_1, 2x_2)$ і $\vec{N} = (1{,}1)$ колінеарні, то

$$\frac{2x_1}{1} = \frac{2x_2}{1}.$$

Звідси маємо рівняння прямої *OM*: $x_1 = x_2$. Розв’язуючи систему рівнянь, отримаємо координати точки *M*:

$$\begin{cases} x_1 + x_2 = 1, \\ x_1 = x_2 \end{cases} \Rightarrow x_1 = \frac{1}{2}, \quad x_2 = \frac{1}{2},$$

тобто $X^*\left(x_1^*, x_2^*\right) = \left(\frac{1}{2}, \frac{1}{2}\right)$.

5. Знайдемо мінімальне значення цільової функції:

$$\min f\left(x_1, x_2\right) = f^* = f\left(x_1^*, x_2^*\right) = f\left(\frac{1}{2}, \frac{1}{2}\right) = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1}{2}.$$

Відповідь: $f^* = \min f = f\left(\frac{1}{2}, \frac{1}{2}\right) = \frac{1}{2}$.

Знаходження максимуму функції.

За рис. 4.3 коло найбільшого радіусу перетинає область допустимих розв'язків у точці D, яка є точкою перетину прямих:

$$\begin{cases} 2x_1 + 3x_2 = 6 \\ x_1 = 2 \end{cases} \Rightarrow x_1 = 2, \quad x_2 = \frac{2}{3},$$

тобто $X^*\left(x_1^*, x_2^*\right) = \left(2, \frac{2}{3}\right)$.

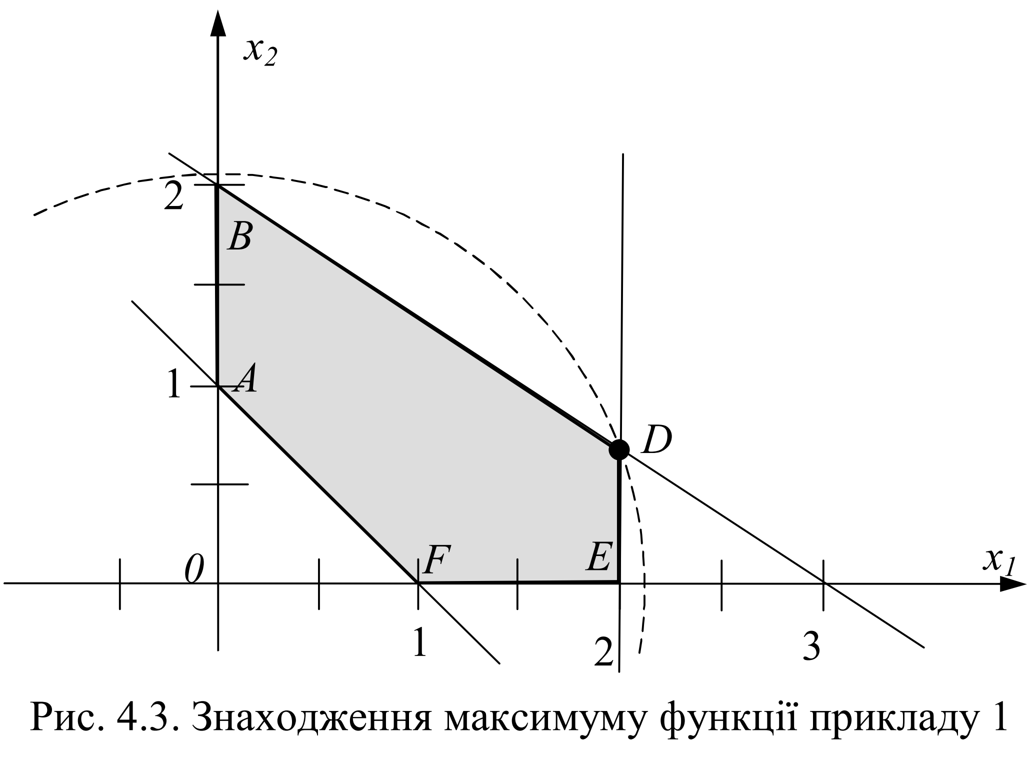

Рис. 4.3. Знаходження максимуму функції прикладу 1

Відповідь: $\max f = f^* = f\left(2, \frac{2}{3}\right) = 2^2 + \left(\frac{2}{3}\right)^2 = \frac{40}{9}$.

Приклад 2. За допомогою графічного методу знайти екстремальні значення функції

$$f(x_1, x_2) = x_1^2 - 6x_1 + x_2^2 - 8x_2$$

при обмеженнях

$$\begin{cases} x_1 x_2 \ge 7, \\ x_1 + x_2 \le 8, \\ x_1, x_2 \ge 0. \end{cases}$$

Розв'язання. Лініями рівня для функції $f(x_1, x_2) = x_1^2 - 6x_1 + x_2^2 - 8x_2$ є сім'я кіл $(x_1 - 3)^2 + (x_2 - 4)^2 = C$ з центром у точці $O'(3,4)$. З рис. 4.4 видно, що свого мінімуму цільова функція досягає в точці $O'(3,4)$, $\min f = f(3,4) = 3^2 - 18 + 4^2 - 32 = -25$.

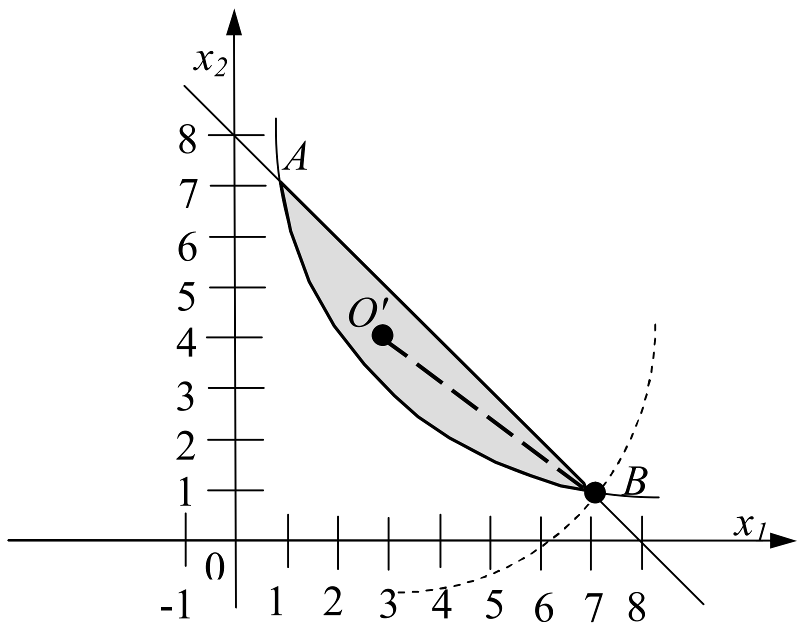

Рис. 4.4. Графічне розв'язання прикладу 2

Як бачимо, точка, у якій досягається мінімальне значення цільової функції знаходиться всередині ОДР.

Максимальне значення досягається в точці *B*, координати якої знаходяться як розв'язок системи рівнянь

$$\begin{cases} x_1 x_2 = 7, \\ x_1 + x_2 = 8, \end{cases} \Rightarrow x_1 = 7,\ x_2 = 1.$$

Відповідь: $\max f = f(7{,}1) = 7^2 - 42 + 1^2 - 8 = 0$, $\min f = f(3{,}4) = -25$.

4.3. Задачі нелінійного програмування без обмежень

Розглянемо задачу знаходження екстремуму функції багатьох змінних $f(x_1, x_2, ..., x_n)$.

Необхідні умови екстремуму функції.

Якщо функція $f(x_1, x_2, ..., x_n)$ має в точці $X^\circ = \left(x_1^\circ, x_2^\circ, ..., x_n^\circ\right)$ екстремум, то частинні похідні першого порядку по змінних $x_i, i = \overline{1,n}$ дорівнюють нулю або не існують.

Точка $X^\circ = \left(x_1^\circ, x_2^\circ, ..., x_n^\circ\right)$, у якій частинні похідні

$$\begin{cases} \dfrac{\partial f(x_1, x_2, ..., x_n)}{\partial x_1} = 0, \\ \dfrac{\partial f(x_1, x_2, ..., x_n)}{\partial x_2} = 0, \\ \dots \\ \dfrac{\partial f(x_1, x_2, ..., x_n)}{\partial x_n} = 0 \end{cases} \qquad (4.7)$$

дорівнюють нулю, називається *стаціонарною точкою функції* $f(x_1, x_2, ..., x_n)$.

Стаціонарні точки і точки, в яких частинні похідні не існують, називаються *критичними точками функції* $f(x_1, x_2, ..., x_n)$.

Для перевірки характеру отриманих точок необхідно використовувати достатні умови.

Нехай $X^\circ = \left(x_1^\circ, x_2^\circ, ..., x_n^\circ\right)$ – стаціонарна точка функції $f(x_1, x_2, ..., x_n)$, причому ця функція двічі диференційована в деякому околі точки X° і всі її частинні похідні другого порядку неперервні в цій точці.

Введемо квадратну матрицю розмірністю $n \times n$, елементами якої є частинні похідні другого порядку $a_{ij} = \dfrac{\partial^2 f(X)}{\partial x_i \, \partial x_j}$, $i = \overline{1,n}$, $j = \overline{1,n}$, тобто

$$H_f(X) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \, \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \, \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \, \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \, \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial^2 f}{\partial x_n \, \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \, \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{pmatrix}. \tag{4.8}$$

Матриця (4.8) називається *матрицею Гессе функції* f, а її визначник – *гессіаном функції* f.

Кутовими мінорами M_i, $i = \overline{1,n}$ квадратної матриці називаються мінори

$$M_1 = \frac{\partial^2 f}{\partial x_1^2}, \; M_2 = \begin{vmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \, \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_2 \, \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} \end{vmatrix}, \ldots,$$

$$M_n = \begin{vmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \, \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \, \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \, \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \, \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial^2 f}{\partial x_n \, \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \, \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{vmatrix}, \tag{4.9}$$

тобто мінори порядку $\overline{1,n}$ матриці (4.8), розташовані в лівому верхньому куті.

Достатні умови екстремуму функції.

Якщо всі кутові мінори M_i , $i=\overline{1,n}$ (4.9) матриці (4.8) додатні в точці X°, то в цій точці функція $f(x_1,x_2,...,x_n)$ має *мінімум*.

Якщо ж всі головні мінори гессіана функції f парного порядку додатні, а непарного порядку – від'ємні в точці X°, тобто $M_1<0, M_2>0, M_3<0,...,$ то в цій точці функція має *максимум*.

У випадку функції двох змінних $f(x_1,x_2)$ достатні умови екстремуму можна сформулювати так:

1) якщо $M_2>0$ в точці X°, то функція $f(x_1,x_2)$ має в цій точці екстремум, причому при $M_1>0$ – мінімум, а при $M_1<0$ – максимум;

2) якщо $M_2<0$ в точці X°, то в цій точці функція немає екстремуму;

3) якщо $M_2=0$, то потрібно додаткове дослідження.

Приклад 3. Знайти екстремум функції $f=x_1^3+x_2^3-9x_1x_2+6$.

Розв'язання. Знайдемо частинні похідні першого порядку і складемо систему (4.7):

$$\begin{cases} \dfrac{\partial f}{\partial x_1}=3x_1^2-9x_2=0, \\ \dfrac{\partial f}{\partial x_2}=3x_2^2-9x_1=0, \end{cases} \Rightarrow \begin{cases} 3x_1^2-9x_2=0, \\ 3x_2^2-9x_1=0. \end{cases}$$

Розв'язуючи цю систему, отримаємо дві стаціонарні точки $X_1^\circ=(0,0)$ і $X_2^\circ=(3,3)$. Оскільки

$$\frac{\partial^2 f}{\partial x_1^2}=6x_1,\quad \frac{\partial^2 f}{\partial x_2^2}=6x_2,\quad \frac{\partial^2 f}{\partial x_1\partial x_2}=\frac{\partial^2 f}{\partial x_2\partial x_1}=-9,$$

то матриця Гессе має вигляд

$$H_f(X)=\begin{pmatrix} 6x_1 & -9 \\ -9 & 6x_2 \end{pmatrix}.$$

У точці $X_1^\circ=(0,0)$: $H_f\left(X_1^\circ\right)=\begin{pmatrix} 0 & -9 \\ -9 & 0 \end{pmatrix}$, тобто $M_2=\begin{vmatrix} 0 & -9 \\ -9 & 0 \end{vmatrix}=-81$, отже в точці $X_1^\circ=(0,0)$ функція не має екстремуму.

У точці $X_1^\circ=(3,3)$: $H_f\left(X_1^\circ\right)=\begin{pmatrix} 18 & -9 \\ -9 & 18 \end{pmatrix}$, тобто $M_2=\begin{vmatrix} 18 & -9 \\ -9 & 18 \end{vmatrix}=$ $=324-81=263>0$, отже в точці $X_1^\circ=(3,3)$ функція має мінімум, оскільки $M_1=18>0$.

Відповідь: $\min f=f(3,3)=3^3+3^3-9\cdot3\cdot3+6=-21$.

Приклад 4. Знайти екстремум функції

$$f=-x_1^2-x_2^2-3x_3^2+x_1x_3+2x_1+5x_2+10x_3.$$

Розв'язання. Використовуючи необхідні умови екстремуму

$$\begin{cases} \dfrac{\partial f}{\partial x_1}=-2x_1+x_3+2=0, \\ \dfrac{\partial f}{\partial x_2}=-2x_2+5=0, \\ \dfrac{\partial f}{\partial x_3}=-6x_3+x_1+10=0, \end{cases} \Rightarrow \begin{cases} -2x_1+x_3+2=0, \\ -2x_2+5=0, \\ -6x_3+x_2+10=0, \end{cases}$$

маємо одну стаціонарну точку $X^\circ=\left(2,\dfrac{5}{2},2\right)$. Гессіан для функції f:

$$H_f(X)=\begin{pmatrix} -2 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & -6 \end{pmatrix}.$$

Обчислимо кутові мінори:

$$M_1 = -2 < 0, M_2 = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = 4 > 0,\ M_3 = \begin{vmatrix} -2 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & -6 \end{vmatrix} = -22 < 0.$$

Оскільки знаки мінорів чергуються, причому $M_1 < 0$, то точка $X^{\circ} = \left(2, \frac{5}{2}, 2\right)$ – точка максимуму функції.

Відповідь: $\max f = f\left(2, \frac{5}{2}, 2\right) = -4 - \frac{25}{4} - 12 + 4 + 4 + \frac{25}{2} + 20 = \frac{73}{4}$.

4.4. Задачі нелінійного програмування з обмеженнями-рівностями. Метод множників Лагранжа

Розглянемо задачу такого виду:

$$f(x_1, x_2, ..., x_n) \to \max(\min), \tag{4.10}$$

$$g_i(x_1, x_2, ..., x_n) = b_i, i = \overline{1, m}; \tag{4.11}$$

$$x_j \in R, j = \overline{1, n}, \tag{4.12}$$

де функції $f(x_1, x_2, ..., x_n)$, $g_i(x_1, x_2, ..., x_n)$, $i = \overline{1, m}$ неперервні разом зі своїми частинними похідними по $x_j, j = \overline{1, n}$.

Задача (4.10)-(4.12) є задачею знаходження умовного екстремуму.

Ця задача в багатьох випадках може бути розв'язана як задача безумовної оптимізації, що була отримана шляхом виключення з цільової функції m незалежних змінних за допомогою рівностей (4.11). Наявність обмежень-рівностей фактично дозволяє зменшити розмірність початкової задачі з n до $n - m$. Для ілюстрації розглянемо приклад.

Приклад 5. Знайти екстремум функції $f = x_1 x_2^2 x_3^3$, якщо $x_1 + 3x_2 - 4x_3 = 12$.

Розв'язання. Виключивши змінну x_1 з обмеження, отримаємо задачу знаходження безумовного екстремуму функції двох змінних:

$$f = (12 - 3x_2 + 4x_3)x_2^2 x_3^3 \to extr .$$

Метод виключення змінних може бути застосований лише у випадках, коли рівняння-обмеження можна розв'язати відносно деякого конкретного набору незалежних змінних або коли ця процедура нескладна.

У загальному випадку при розв'язанні ЗНП доцільно використовувати *метод множників Лагранжа*, який зводить задачу умовного екстремуму (4.10)-(4.12) до задачі безумовного екстремуму для функції

$$\begin{aligned} &L(x_1, x_2, ..., x_n, \lambda_1, ..., \lambda_m) = \\ &= f(x_1, x_2, ..., x_n) + \sum_{i=1}^{m} \lambda_i (b_i - g_i(x_1, x_2, ..., x_n)), \end{aligned} \tag{4.13}$$

де $\lambda_i, i = \overline{1,m}$ – невідомі параметри, які називаються *множниками Лагранжа*, а сама функція виду (4.13) називається *функцією Лагранжа*.

Необхідну умову екстремуму для функції Лагранжа (4.13) записують у вигляді системи рівнянь, складених з частинних похідних функції Лагранжа (4.13) по змінним $x_j, \lambda_i, i = \overline{1,m}, j = \overline{1,n}$:

$$\begin{cases} \dfrac{\partial L}{\partial x_j} = \dfrac{\partial f(x_1, x_2, ..., x_n)}{\partial x_j} - \sum\limits_{i=1}^{m} \dfrac{\lambda_i \, \partial g_i(x_1, x_2, ..., x_n)}{\partial x_j} = 0, \; j = \overline{1,n}, \\ \dfrac{\partial L}{\partial \lambda_i} = b_i - g_i(x_1, x_2, ..., x_n) = 0, \; i = \overline{1,m}. \end{cases} \tag{4.14}$$

Достатні умови існування умовного екстремуму в точці (X°, Λ°), яка є розв'язком системи (4.14), визначають за знаком другого диференціала $d^2 L(X, \Lambda)$:

$$d^2L(X,\Lambda) = \sum_{i,j=1}^{n} \frac{\partial^2 L(x_1, x_2, \ldots, x_n)}{\partial x_i \, \partial x_j} \, dx_i \, dx_j \, .$$

Функція (4.10) У стаціонарній точці (X°, Λ°) має умовний максимум, якщо в ній $d^2L(X,\Lambda) < 0$, і умовний мінімум, якщо $d^2L(X,\Lambda) > 0$.

Можна дослідити і іншим шляхом, розглянувши таку матрицю Гессе для функції Лагранжа (4.13):

$$H_L = \begin{array}{c} \begin{array}{cc} \overbrace{\qquad\qquad\qquad\qquad\qquad}^{m\ldots\text{стовпців}} & \qquad \overbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}^{n \ \ \text{стовпців}} \end{array} \\ \begin{pmatrix} 0 & 0 & \ldots & 0 & \dfrac{\partial g_1}{\partial x_1} & \dfrac{\partial g_1}{\partial x_2} & \ldots & \dfrac{\partial g_1}{\partial x_n} \\ 0 & 0 & \ldots & 0 & \dfrac{\partial g_2}{\partial x_1} & \dfrac{\partial g_2}{\partial x_2} & \ldots & \dfrac{\partial g_2}{\partial x_n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & \dfrac{\partial g_m}{\partial x_1} & \dfrac{\partial g_m}{\partial x_2} & \ldots & \dfrac{\partial g_m}{\partial x_n} \\ \dfrac{\partial g_1}{\partial x_1} & \dfrac{\partial g_2}{\partial x_1} & \ldots & \dfrac{\partial g_m}{\partial x_1} & \dfrac{\partial^2 L}{\partial x_1^2} & \dfrac{\partial^2 L}{\partial x_1 \, \partial x_2} & \ldots & \dfrac{\partial^2 L}{\partial x_1 \, \partial x_n} \\ \dfrac{\partial g_1}{\partial x_2} & \dfrac{\partial g_2}{\partial x_2} & \ldots & \dfrac{\partial g_m}{\partial x_2} & \dfrac{\partial^2 L}{\partial x_2 \, \partial x_1} & \dfrac{\partial^2 L}{\partial x_2^2} & \ldots & \dfrac{\partial^2 L}{\partial x_2 \, \partial x_n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \dfrac{\partial g_1}{\partial x_n} & \dfrac{\partial g_2}{\partial x_n} & \ldots & \dfrac{\partial g_m}{\partial x_n} & \dfrac{\partial^2 L}{\partial x_n \, \partial x_1} & \dfrac{\partial^2 L}{\partial x_n \, \partial x_2} & \ldots & \dfrac{\partial^2 L}{\partial x_n^2} \end{pmatrix} \end{array} \tag{4.15}$$

тоді:

1) точка (X°, Λ°) є точкою *умовного мінімуму*, якщо знаки кутових мінорів $M_{m+1}, M_{m+2}, \ldots, M_{m+1}, M_{m+n}$ матриці H_L співпадають зі знаком $(-1)^m$;

2) точка (X°, Λ°) є точкою *умовного максимуму*, якщо знаки кутових мінорів $M_{m+1}, M_{m+2}, \ldots, M_{m+1}, M_{m+n}$ матриці H_L чергуються, причому знак мінора M_{m+1} співпадає зі знаком $(-1)^{m+1}$.

У випадку, коли задача має лише одне обмеження, це правило спрощується: якщо визначник $\det H_L$ матриці (4.15) в точці $\left(X^{\circ}, \Lambda^{\circ}\right)$ від'ємний, то функція має в цій точці *умовний мінімум*, якщо додатний – *умовний максимум*.

Алгоритм методу множників Лагранжа для ЗНП:

1. Скласти функцію Лагранжа (4.13) для задачі (4.10)-(4.12).
2. Скласти систему (4.14).
3. Розв'язати систему (4.14) і визначити точки $\left(X^{\circ}, \Lambda^{\circ}\right)=\left(x_1^{\circ}, x_2^{\circ}, \ldots, x_n^{\circ}, \lambda_1^{\circ}, \ldots, \lambda_m^{\circ}\right)$, у яких функція Лагранжа може мати екстремум.
4. Перевірити кожну отриману точку $\left(X^{\circ}, \Lambda^{\circ}\right)$ на оптимальність.

Іноді перевірку оптимальності отриманої точки $\left(X^{\circ}, \Lambda^{\circ}\right)$ можна здійснити за допомогою гессіана, складеного для функції f, бо $b_i - g_i(x_1, x_2, \ldots, x_n) = 0, i = \overline{1, m}$.

У загальному випадку розв'язання системи (4.13) є досить складним через її нелінійність.

Приклад 6. Знайти екстремум функції $f = x_1^2 + (x_2 + 2)^2$, якщо $3x_1 + 2x_2 = 6$.

Розв'язання. Складемо функцію Лагранжа, враховуючи, що $g(x) = 6 - 3x_1 - 2x_2$:

$$L = x_1^2 + (x_2 + 2)^2 + \lambda(6 - 3x_1 - 2x_2).$$

Використовуючи систему рівнянь (4.14)

$$\begin{cases} \dfrac{\partial L}{\partial x_1} = 2x_1 - 3\lambda = 0, \\ \dfrac{\partial L}{\partial x_2} = 2(x_2 + 2) - 2\lambda = 0, \\ \dfrac{\partial L}{\partial \lambda} = 6 - 3x_1 - 2x_2 = 0, \end{cases}$$

знаходимо стаціонарну точку функції Лагранжа: $x_1 = \frac{30}{13}, x_2 = -\frac{6}{13}, \lambda = \frac{20}{13}$. Характер оптимальності отриманої точки визначимо за допомогою матриці Гессе для функції f. Маємо

$$\frac{\partial f}{\partial x_1} = 2x_1,\ \frac{\partial f}{\partial x_2} = 2(x_2 + 2),$$

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

тобто $H_f(X,\Lambda) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$.

Оскільки $M_1 = 2 > 0,\ M_2 = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4 > 0$, то в точці $X^* = \left(\frac{30}{13}, -\frac{6}{13}\right)$ функція $f = x_1^2 + (x_2 + 2)^2$ має мінімум, який дорівнює $\min f = \left(\frac{30}{13}\right)^2 + \left(-\frac{6}{13} + 2\right)^2 = \frac{1300}{169}$.

Відповідь: $\min f = f\left(\frac{30}{13}, -\frac{6}{13}\right) = \frac{1300}{169}$.

Приклад 7. Знайти екстремум функції $f = 2x_1 + 4x_2$, якщо $x_1^2 + x_2^2 = 20$.

Розв'язання. Функція Лагранжа має вигляд

$$L = 2x_1 + 4x_2 + \lambda\left(20 - x_1^2 - x_2^2\right).$$

Для знаходження стаціонарних точок запишемо систему (4.14):

$$\begin{cases} \frac{\partial L}{\partial x_1} = 2 - 2\lambda x_1 = 0, \\ \frac{\partial L}{\partial x_2} = 4 - 2\lambda x_2 = 0, \\ \frac{\partial L}{\partial \lambda} = 20 - x_1^2 - x_2^2 = 0. \end{cases}$$

Система має два розв'язки

$$\begin{cases} x_1 = 2, \\ x_2 = 4, \\ \lambda_1 = \dfrac{1}{2} \end{cases} \quad \text{і} \quad \begin{cases} x_1 = -2, \\ x_2 = -4, \\ \lambda_1 = -\dfrac{1}{2}. \end{cases}$$

З'ясуємо характер екстремуму в кожній з точок $\left(X^1, \Lambda^1\right) = \left(2, 4, \frac{1}{2}\right)$ та $\left(X^2, \Lambda^2\right) = \left(-2, -4, -\frac{1}{2}\right)$. Побудуємо матрицю (4.15), враховуючи, що $g(x_1, x_2) = \left(20 - x_1^2 - x_2^2\right)$:

$$H_L = \begin{pmatrix} 0 & g'_{x_1} & g'_{x_2} \\ g'_{x_1} & \dfrac{\partial^2 L}{\partial x_1^2} & \dfrac{\partial^2 L}{\partial x_1 \, \partial x_2} \\ g'_{x_2} & \dfrac{\partial^2 L}{\partial x_2 \, \partial x_1} & \dfrac{\partial^2 L}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 0 & -2x_1 & -2x_2 \\ -2x_1 & -2\lambda & 0 \\ -2x_2 & 0 & -2\lambda \end{pmatrix}$$

і знайдемо її визначник

$$\det H_L = \begin{vmatrix} 0 & -2x_1 & -2x_2 \\ -2x_1 & -2\lambda & 0 \\ -2x_2 & 0 & -2\lambda \end{vmatrix} = -8 \cdot \begin{vmatrix} 0 & x_1 & x_2 \\ x_1 & \lambda & 0 \\ x_2 & 0 & \lambda \end{vmatrix}.$$

У точці $\left(X^1, \Lambda^1\right) = \left(2, 4, \frac{1}{2}\right)$ отримаємо

$$\det H_L\left(X^1, \Lambda^1\right) = -8 \cdot \begin{vmatrix} 0 & 2 & 4 \\ 2 & \dfrac{1}{2} & 0 \\ 4 & 0 & \dfrac{1}{2} \end{vmatrix} = 80 > 0,$$

тому функція $f = 2x_1 + 4x_2$ має в цій точці умовний максимум, $\max f = f(2,4) = 20$.

Аналогічно для точки $\left(X^2, \Lambda^2\right) = \left(-2,-4,-\frac{1}{2}\right)$ маємо

$$\det H_L\left(X^2, \Lambda^2\right) = -8 \cdot \begin{vmatrix} 0 & -2 & -4 \\ -2 & -\frac{1}{2} & 0 \\ -4 & 0 & -\frac{1}{2} \end{vmatrix} = 8 \cdot \begin{vmatrix} 0 & 2 & 4 \\ 2 & \frac{1}{2} & 0 \\ 4 & 0 & \frac{1}{2} \end{vmatrix} = -80 < 0,$$

тобто точка $(-2,-4)$ є умовним мінімумом функції, $\min f = f(-2,-4) = -20$.

Оптимальність у стаціонарних точках можна також визначити за допомогою знака другого диференціала функції Лагранжа:

$$d^2 L = -2\lambda dx_1^2 + 2 \cdot 0 dx_1 dx_2 - 2\lambda dx_2^2 = -2\lambda\left(dx_1^2 + dx_2^2\right).$$

Оскільки $dx_1^2 + dx_2^2 > 0$, то знак другого диференціала залежить від знака λ. Таким чином, при $\lambda = \frac{1}{2}$ отримаємо $d^2 L = -\left(dx_1^2 + dx_2^2\right) < 0$, тобто в точці $(2,4)$ функція має максимум. Аналогічно, при $\lambda = -\frac{1}{2}$, $d^2 L = dx_1^2 + dx_2^2 > 0$ і в точці $(-2,-4)$ отримаємо умовний мінімум функції $f = 2x_1 + 4x_2$.

Відповідь: $\max f = f(2,4) = 20$, $\min f = f(-2,-4) = -20$.

Приклад 8. Знайти екстремум функції $f = 5x_1^2 + x_2^3 - 2x_1 x_2$, якщо $2x_1 + x_2 = 0$.

Розв'язання. Складемо функцію Лагранжа:

$$L = 5x_1^2 + x_2^3 - 2x_1 x_2 + \lambda\left(-2x_1 - x_2\right).$$

З необхідних умов екстремуму

$$\begin{cases} \dfrac{\partial L}{\partial x_1} = 10x_1 - 2x_2 - 2\lambda = 0, \\ \dfrac{\partial L}{\partial x_2} = 3x_2^2 - 2x_1 - \lambda = 0, \\ \dfrac{\partial L}{\partial \lambda} = -2x_1 - x_2 = 0 \end{cases}$$

знайдемо дві стаціонарні точки $\left(X^1, \Lambda^1\right) = (0,0,0)$ і $\left(X^2, \Lambda^2\right) = \left(\frac{3}{4}, -\frac{3}{2}, \frac{21}{4}\right)$. Для з'ясування характеру оптимальності отриманих точок обчислимо другий диференціал функції Лагранжа. Оскільки $\frac{\partial^2 L}{\partial x_1^2} = 10$, $\frac{\partial^2 L}{\partial x_2^2} = 6x_2$, $\frac{\partial^2 L}{\partial x_1 \partial x_2} = \frac{\partial^2 L}{\partial x_2 \partial x_1} = -2$, то

$$d^2 L = 10dx_1^2 + 2 \cdot (-2)dx_1 dx_2 + 6x_2 dx_2^2 = 10dx_1^2 - 4dx_1 dx_2 + 6x_2 dx_2^2.$$

З обмеження $2x_1 + x_2 = 0$ маємо $x_2 = -2x_1, dx_2 = -2dx_1$, тоді

$$d^2 L = 10dx_1^2 - 4dx_1\left(-2dx_1\right) + 6x_2\left(-2dx_1\right)^2 = 10dx_1^2 + 8dx_1^2 + 24x_2 dx_1^2 =$$
$$= \left(18 + 24x_2\right)dx_1^2.$$

Оскільки $d^2 L\Big|_{(0,0,0)} = 18dx_1^2 > 0$, то $X^1(0,0)$ є точкою умовного мінімуму функції $f = 5x_1^2 + x_2^2 - 2x_1 x_2$, $\min f = f(0,0) = 0$.

Аналогічно, $d^2 L\Big|_{\left(\frac{3}{4}, -\frac{3}{2}, \frac{21}{4}\right)} = -18dx_1^2 < 0$, то $X^1\left(\frac{3}{4}, -\frac{3}{2}\right)$ - точка умовного максимуму, $\max f = f\left(\frac{3}{4}, -\frac{3}{2}\right) = \frac{27}{16}$.

Відповідь: $\min f = f(0,0) = 0$, $\max f = f\left(\frac{3}{4}, -\frac{3}{2}\right) = \frac{27}{16}$.

Можна дослідити оптимальність кожної з отриманих точок, використовуючи матрицю (4.15), визначник якої дорівнює

$$\det H_L = \begin{vmatrix} 0 & -2 & -1 \\ -2 & 10 & -2 \\ -1 & -2 & 6x_2 \end{vmatrix} = -24x_2 - 18 .$$

Також можна було звести задачу знаходження умовного екстремуму функції двох змінних до знаходження екстремуму функції однієї змінної, дослідження якої відомо з курсу диференціального числення функції однієї змінної. Для цього потрібно було виразити одну зі змінних, наприклад x_2, з обмеження $2x_1 + x_2 = 0$ і підставити отриманий вираз у функцію $f = 5x_1^2 + x_2^3 - 2x_1x_2$.

4.5. Задачі опуклого та квадратичного програмування

4.5.1. Опуклі функції

Функція $f(X) = f(x_1, \ldots x_n)$ називається *опуклою* на деякій опуклій множині *G*, якщо для будь-яких точок $X_1 \in G$ та $X_2 \in G$, $G \subset R^n$ виконується нерівність

$$f(\lambda X_1 + (1-\lambda)X_2) \le \lambda f(X_1) + (1-\lambda) f(X_2) \tag{4.16}$$

для довільного $\lambda \in [0,1]$.

Графічний приклад опуклої функції на площині зображений на рис. 4.5.

Функція $f(X)$ називається *увігнутою* на деякій опуклій множині *G*, якщо для будь-яких точок $X_1 \in G$ та $X_2 \in G$, $G \subset R^n$ виконується нерівність

$$f(\lambda X_1 + (1-\lambda)X_2) \ge \lambda f(X_1) + (1-\lambda) f(X_2) \tag{4.17}$$

для довільного $\lambda \in [0,1]$.

Якщо в співвідношеннях (4.16), (4.17) для $\lambda \in (0,1)$ мають місце строгі нерівності, тоді функція називається *строго опуклою* (*строго увігнутою*).

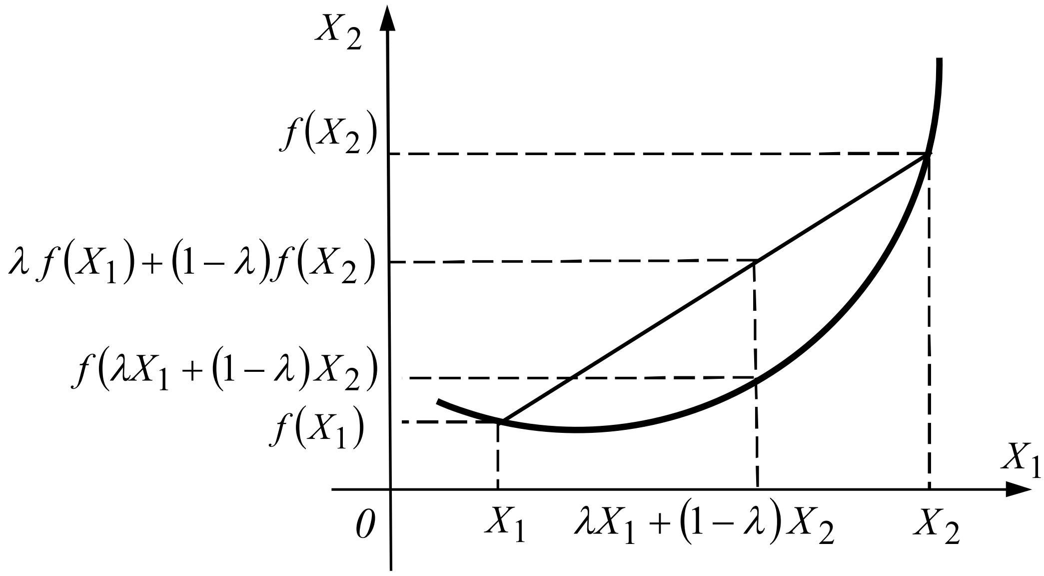

Рис. 4.5. Опукла функція

На рис. 4.6 наведений графічний приклад увігнутої функції однієї змінної.

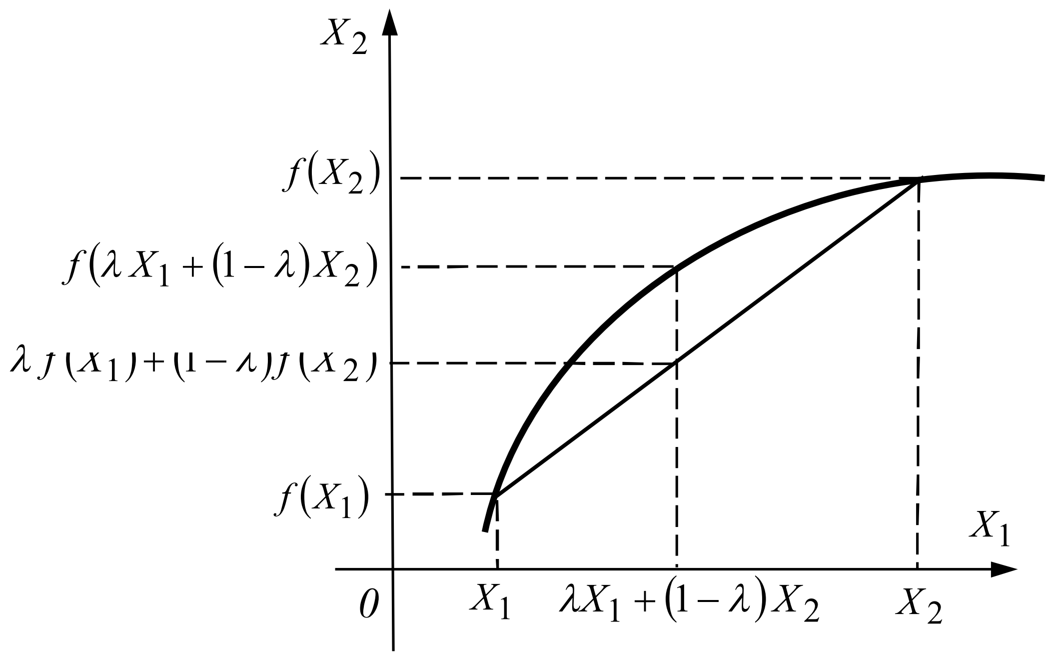

Рис. 4.6. Увігнута функція

Отже, опуклість (увігнутість) функції означає, що графік цієї функції лежить нижче (вище) від січної, яка з'єднує будь-які дві її точки.

Лінійна комбінація $f(X)=\sum_{i=1}^{n}\lambda_i f_i(X),\ i=\overline{1,n}$ опуклих (увігнутих) на деякій опуклій множині G функцій $f_i(X)$ також опукла (увігнута) на цій множині, якщо $\lambda_i \ge 0,\ i=\overline{1,n}$.

Очевидно, що якщо функція $f(X)$ – опукла, то функція «- $f(X)$» - увігнута, і навпаки.

Лінійна функція $f=c_1x_1+c_2x_2+...+c_nx_n$ є одночасно й опуклою й увігнутою функцією в просторі R^n, оскільки для будь-яких точок $X_1 \in R^n$ та $X_2 \in R^n$ і $\lambda \in [0,1]$ виконується рівність

$$f(\lambda X_1+(1-\lambda)X_2)=\lambda f(X_1)+(1-\lambda)f(X_2).$$

Дослідити функцію на опуклість або увігнутість можна за допомогою наведених визначень (4.16) і (4.17), але на практиці це, зазвичай, складно. Розглянемо критерії опуклості функції.

Теорема 1. Нехай функція $f(X)$ визначена на опуклій множині *G*, $G \subset R^n$ і має на цій множині неперервні частинні похідні другого порядку. Тоді $f(X)$ – строго опукла функція тоді і тільки тоді, коли всі кутові мінори M_i, $i=\overline{1,n}$ гессіана функції f – невід’ємні в усіх точках *G*.

Теорема 2. Нехай функція $f(X)$ визначена на опуклій множині *G*, $G \subset R^n$ і має на цій множині неперервні частинні похідні другого порядку. Тоді $f(X)$ – строго увігнута функція тоді і тільки тоді, коли всі кутові мінори гессіана функції f парного порядку додатні, а непарного порядку – від’ємні в усіх точках *G*, тобто $M_1<0, M_2>0, M_3<0,...$.

Приклад 9. Довести, що функція

$$f(x_1,x_2,x_3)=2x_1^2+3x_2^2+5x_3^2-4x_1x_2-6x_2x_3$$

є опуклою на всій області визначення.

Розв’язання. Знайдемо частинні похідні другого порядку:

$$\frac{\partial f(X)}{\partial x_1}=4x_1-4x_2;\qquad \frac{\partial f(X)}{\partial x_2}=6x_2-4x_1-6x_3;\qquad \frac{\partial f(X)}{\partial x_3}=10x_3-6x_2;$$

$$\frac{\partial^2 f(X)}{\partial x_1^2}=4;\qquad \frac{\partial^2 f(X)}{\partial x_2^2}=6;\qquad \frac{\partial^2 f(X)}{\partial x_3^2}=10;$$

$$\frac{\partial^2 f(X)}{\partial x_1\partial x_2}=\frac{\partial^2 f(X)}{\partial x_2\partial x_1}=-4;$$

$$\frac{\partial^2 f(X)}{\partial x_2\partial x_3}=\frac{\partial^2 f(X)}{\partial x_3\partial x_2}=-6;\qquad \frac{\partial^2 f(X)}{\partial x_1\partial x_3}=\frac{\partial^2 f(X)}{\partial x_3\partial x_1}=0$$

і складемо матрицю Гессе:

$$H_f(X)=\begin{pmatrix}4 & -4 & 0\\ -4 & 6 & -6\\ 0 & -6 & 10\end{pmatrix}.$$

Обчислимо кутові мінори:

$$M_1=4>0,\qquad M_2=\begin{vmatrix}4 & -4\\ -4 & 6\end{vmatrix}=8>0,\qquad M_3=\begin{vmatrix}4 & -4 & 0\\ -4 & 6 & -6\\ 0 & -6 & 10\end{vmatrix}=256>0.$$

Відповідь: Оскільки всі головні мінори гессіана додатні, то функція $f(x_1,x_2,x_3)=2x_1^2+3x_2^2+5x_3^2-4x_1x_2-6x_2x_3$ є опуклою на всій області визначення.

Зауваження. Для функції однієї змінної теореми 1, 2 набувають такого вигляду.

Теорема 3. Нехай функція двічі диференційована на (a,b), тоді:

а) якщо $f''(x)\ge 0, x\in(a,b)$, то $f(x)$ - є опуклою на (a,b);

б) якщо $f''(x)\le 0, x\in(a,b)$, то $f(x)$ - є увігнутою на (a,b).

Для опуклих (увігнутих) функцій справедливі такі теореми. Саме ці твердження гарантують успіх при розв’язанні задач нелінійного програмування.

Теорема 4. Нехай $f(X)$ – опукла функція, яка визначена на опуклій множині $G,\ G \subset R^n$. Тоді будь-який її локальний мінімум (максимум) збігається з глобальним мінімумом (максимумом).

Наслідки:

1. Якщо глобальний екстремум досягається у двох різних точках, то він досягається в будь-якій точці відрізка, який з'єднує ці точки.
2. Якщо $f(X)$ – строго опукла (увігнута) на множині *G,* то її глобальний мінімум (максимум) – єдиний.

4.5.2. Опукле програмування

Розглянемо задачу нелінійного програмування

$$f(x_1, x_2, ..., x_n) \to \max, \qquad (4.18)$$

$$g_i(x_1, x_2, ..., x_n) \le b_i,\ i = \overline{1,m}, \qquad (4.19)$$

$$x_j \ge 0,\ j = \overline{1,n}, \qquad (4.20)$$

де цільова функція (4.18) увігнута, а функції $g_i(x_1, x_2, ..., x_n),\ i = \overline{1,m}$ є опуклими.

Для задачі мінімізації цільова функція повинна бути опуклою.

Як відомо, якщо ОДР визначається опуклими функціями, то вона теж є опуклою.

У разі, коли система обмежень задачі містить рівності, то кожну рівність необхідно представити у вигляді двох нерівностей, а обмеження вигляду "≥", потрібно помножити на "-1", щоб звести задачу до потрібного вигляду (4.18) - (4.20) (п. 2.1.1).

Якщо на деяку змінну $x_j,\ j = \overline{1,n}$ не накладено умову невід-ємності, то потрібно представити її у вигляді $x_j = x'_j - x''_j,\ x'_j, x''_j \ge 0$.

Особливість опуклого програмування полягає в тому, що локальний і глобальний екстремуми тут обов'язково збігаються (теорема 4), що гарантує знаходження оптимального розв'язку задачі.

Кажуть, що множина допустимих розв'язків задовольняє *умові регулярності Слейтера (умова Слейтера)*, якщо існує, принаймні, одна точка X°, яка належить ОДР, така, що $g_i\left(X^\circ\right) < b_i, i = \overline{1,m}$.

Функцією Лагранжа задачі опуклого програмування (4.18) - (4.20) називається функція

$$L(x_1, x_2, ..., x_n, \lambda_1, ..., \lambda_m) = f(x_1, x_2, ..., x_n) + \sum_{i=1}^{m} \lambda_i \left(b_i - g_i(x_1, x_2, ..., x_n)\right),$$

де $\Lambda = \left(\lambda_1, ..., \lambda_m\right)$ – *вектор множників Лагранжа*.

Точка $\left(X^*, \Lambda^*\right) = \left(x_1^*, x_2^*, ..., x_n^*, \lambda_1^*, ..., \lambda_m^*\right)$ називається *сідловою точкою функції Лагранжа*, якщо

$$L\left(X, \Lambda^*\right) \le L\left(X^*, \Lambda^*\right) \le L\left(X^*, \Lambda\right) \tag{4.21}$$

для довільних $x_j \ge 0, j = \overline{1,n}, \lambda_i \ge 0, i = \overline{1,m}$.

Вираз (4.21) можна переписати у вигляді

$$\max_{x_j \ge 0} L\left(X, \Lambda^*\right) = L\left(X^*, \Lambda^*\right) = \min_{\lambda_i \ge 0} L\left(X^*, \Lambda\right), \; j = \overline{1,n}, \; i = \overline{1,m}.$$

Важливе місце в теорії опуклого програмування займають умови Куна-Таккера, які названі на честь Гарольда Куна і Альберта Таккера. Вони узагальнили метод множників Лагранжа, який використовується лише для знаходження оптимуму задач з обмеженнями-рівностями, на випадок, коли серед обмежень можуть бути як рівності, так і нерівності.

Умови Куна-Таккера визначають необхідну умову оптимальності, але при деяких додаткових умовах відносно цільової функції вони є і достатніми умовами оптимальності.

Теорема 5 (теорема Куна-Таккера). Для задачі опуклого програмування (4.18)-(4.20), множина допустимих розв'язків якої задовольняє умову регулярності, $X^* = \left(x_1^*, x_2^*, ..., x_n^*\right)$ є оптимальним планом тоді і тільки тоді, коли існує такий вектор $\Lambda^* = \left(\lambda_1^*, ..., \lambda_m^*\right)$, що $\left(X^*, \Lambda^*\right)$ є сідловою точкою функції Лагранжа.

Цю теорему ще називають *теоремою про сідлову точку*, оскільки вона встановлює зв'язок між оптимальним планом і сідловою точкою функції Лагранжа.

Якщо функції $f(X)$ і $g_i(X),\ i = \overline{1,m}$ неперервно диференційовані, то умова (4.21) еквівалентна таким *локальним умовам Куна-Таккера*:

$$\frac{\partial L\left(X^*, \Lambda^*\right)}{\partial x_j} \le 0,\ j = \overline{1,n}, \tag{4.22}$$

$$x_j^* \frac{\partial L\left(X^*, \Lambda^*\right)}{\partial x_j} = 0,\ x_j^* \ge 0,\ j = \overline{1,n}, \tag{4.23}$$

$$\frac{\partial L\left(X^*, \Lambda^*\right)}{\partial \lambda_i} \ge 0,\ i = \overline{1,m}, \tag{4.24}$$

$$\lambda_i^* \frac{\partial L\left(X^*, \Lambda^*\right)}{\partial \lambda_i} = 0,\ \lambda_i^* \ge 0, i = \overline{1,m}. \tag{4.25}$$

4.5.3. Квадратичне програмування

Нижче буде розглянуто задачу квадратичного програмування, що задовольняє всі сформульовані в п. 4.5.2 вимоги, які дозволяють записати необхідні та достатні умови для сідлової точки функції Лагранжа у вигляді (4.22)-(4.25).

Функція $F = \sum_{k=1}^{n} \sum_{j=1}^{n} c_{kj} x_k x_j$ називається *квадратичною формою* відносно змінних $x_1, ..., x_n$.

Квадратична форма F називається *додатно визначеною (від'ємно визначеною)*, якщо $F(X) > 0$ ($F(X) < 0$) для всіх значень $X = \left(x_1, ..., x_n\right)$, крім $X = 0$.

Якщо для довільних $X = \left(x_1, ..., x_n\right)$ $F(X) \ge 0$ ($F(X) \le 0$) і, крім того, існує такий набір змінних $X' = \left(x_1', ..., x_n'\right)$, де не всі $x_j',\ j = \overline{1,n}$ одночасно дорівнюють нулю, виконується рівність $F(X') = 0$, то така квадратична форма називається додатно напіввизначеною (від'ємно напіввизначеною).

Квадратична форма є опуклою функцією, якщо вона додатно напіввизначена, і увігнутою функцією, якщо вона від'ємно напіввизначена.

Тип квадратичної форми можна встановити, якщо її можна подати в канонічній формі

$$F = \sum_{j=1}^{n} \lambda_j x'_j \qquad (4.26)$$

за допомогою лінійних перетворень.

Очевидно, що

1) якщо $\lambda_j > 0, j = \overline{1,n}$, то квадратична форма F є додатно визначеною;

2) якщо $\lambda_j < 0, j = \overline{1,n}$, то F – від'ємно визначена;

3) якщо серед $\lambda_j, j = \overline{1,n}$ є як додатні, так і від'ємні числа, то квадратична форма F є невизначеною.

Простіше визначити вид квадратичної форми за допомогою кутових мінорів $M_i, i = \overline{1,n}$ (п. 4.3) матриці, що складена з коефіцієнтів квадратичної форми

$$C = \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \dots & \dots & \dots \\ c_{n1} & \dots & c_{nn} \end{pmatrix}. \qquad (4.27)$$

При цьому вважається, що матриця С – симетрична, тобто $c_{ij} = c_{ji}$, $i, j = \overline{1,n}$.

Теорема 6. Якщо для квадратичної форми $F = \sum_{k=1}^{n} \sum_{j=1}^{n} c_{kj} x_k x_j$ всі кутові мінори $M_i, i = \overline{1,n}$ відмінні від нуля, то її можна звести до вигляду виразу (4.26), де

$$\lambda_i = \frac{M_i}{M_{i-1}}, M_0 = 1, i = \overline{1,n}.$$

Таким чином, знаки коефіцієнтів $\lambda_i, i=\overline{1,n}$ визначаються знаками визначників $M_i, i=\overline{1,n}$.

Наслідки:

1. Якщо всі кутові мінори $M_i, i=\overline{1,n}$ додатні, то квадратична форма додатно визначена.

2. Якщо знаки кутових мінорів $M_i, i=\overline{1,n}$ чергуються, причому $M_1<0$, то квадратична форма від'ємно визначена.

3. Якщо ранг r матриці C менший за n і якщо її перші r визначників додатні, а решта дорівнює нулю, то квадратична форма додатно напіввизначена.

4. Якщо ранг r матриці C менший за n і знаки її перших r визначників чергуються, причому $M_1<0$, а решта дорівнює нулю, то квадратична форма від'ємно напіввизначена.

5. Квадратична форма є невизначеною, якщо кутові мінори мають як додатні, так від'ємні значення, причому знаки кутових мінорів не чергуються.

Задача нелінійного програмування

$$f=\sum_{k=1}^{n}\sum_{j=1}^{n}c_{kj}x_kx_j+\sum_{j=1}^{n}d_jx_j\to\max(\min), \tag{4.28}$$

$$\sum_{j=1}^{n}a_{ij}x_j\le b_i,\ i=\overline{1,m}, \tag{4.29}$$

$$x_j\ge 0,\ j=\overline{1,n}, \tag{4.30}$$

де $\sum_{k=1}^{n}\sum_{j=1}^{n}c_{kj}x_kx_j$ додатно напіввизначена (від'ємно напіввизначена) квадратична форма, називається *задачею квадратичного програмування*.

Функція Лагранжа для задачі квадратичного програмування (4.28)-(4.30) має вигляд

$$L=\sum_{k=1}^{n}\sum_{j=1}^{n}c_{kj}x_kx_j+\sum_{j=1}^{n}d_jx_j+\sum_{i=1}^{m}\lambda_i\left(b_i-\sum_{j=1}^{n}a_{ij}x_j\right). \tag{4.31}$$

Якщо функція L має сідлову точку $\left(X^*, \Lambda^*\right)$, то в цій точці виконуються локальні умови Куна-Таккера (4.22)-(4.25).

Після введення балансуючих змінних $v_j, w_i,\ j=\overline{1,n},\ i=\overline{1,m}$ отримаємо систему рівнянь:

$$\frac{\partial L\left(X^*, \Lambda^*\right)}{\partial x_j} + v_j = 0, \qquad (4.32)$$

$$\frac{\partial L\left(X^*, \Lambda^*\right)}{\partial \lambda_i} - w_i = 0, \qquad (4.33)$$

$$x_j^* v_j = 0, \qquad (4.34)$$

$$\lambda_i^* w_i = 0, \qquad (4.35)$$

$$\lambda_i^* \ge 0, w_i \ge 0, x_j^* \ge 0, v_j \ge 0,\ i=\overline{1,m},\ j=\overline{1,n}. \qquad (4.36)$$

Таким чином, для знаходження розв’язку задачі квадратичного програмування (4.28)-(4.30) необхідно знати невід’ємний розв’язок системи лінійних рівнянь (4.32)-(4.33), який задовольняє умови (4.34)-(4.36).

Це можна зробити, наприклад, за допомогою методу штучного базису (п. 2.6), застосованого для знаходження максимального значення функції $F = -\sum\limits_{k=1}^{n} My_k$ при умовах (4.32), (4.33), (4.36) з урахуванням (4.34)-(4.35), де $y_k, k=\overline{1,n}$ – штучні змінні, які введено в рівняння (4.32), тобто необхідно знайти розв’язок задачі лінійного програмування

$$F = -\sum_{k=1}^{n} My_k \to \max, \qquad (4.37)$$

$$\frac{\partial L\left(X^*, \Lambda^*\right)}{\partial x_j} + v_j + y_j = 0, \qquad (4.38)$$

$$\frac{\partial L\left(X^*, \Lambda^*\right)}{\partial \lambda_i} - w_i = 0, \qquad (4.39)$$

$$\lambda_i^* \ge 0, w_i \ge 0,\ i = \overline{1,m},$$
$$x_j^* \ge 0, v_j \ge 0\ \ j = \overline{1,n},$$

який задовольняє умови

$$x_j^* v_j = 0, \quad \lambda_i^* w_i = 0. \qquad (4.40)$$

Зауваження. Якщо ОДР задана лише лінійними нерівностями, то теорема Куна-Таккера буде вірною і без умови регулярності.

Алгоритм розв'язання задачі квадратичного програмування

1. Скласти функцію Лагранжа.
2. Записати необхідні та достатні умови існування сідлової точки функції Лагранжа у вигляді виразів (4.22)-(4.25).
3. Розв'язати методом штучного базису ЗЛП (4.37)-(4.39).
4. Для знайденого розв'язку перевірити умови (4.40).
5. У разі їх виконання записати оптимальний розв'язок початкової задачі і знайти значення цільової функції. Якщо умови не виконані, то знайти інший розв'язок задачі (4.37)-(4.39) і повернутися до кроку 4.

Використовуючи метод штучного базису з урахуванням умов (4.40), після скінченної кількості кроків або отримують оптимальний розв'язок, або доводять, що його не існує.

Приклад 10. Дана задача квадратичного програмування:

1) $f(x_1, x_2) = (x_1 - 10)^2 + (x_2 + 1)^2 \to \min,$

$$\begin{cases} 2x_1 + x_2 \ge 2, \\ -x_1 + x_2 \le 1, \\ 3x_1 - 2x_2 \le 6, \\ x_1, x_2 \ge 0; \end{cases}$$

2) $g(x_1,x_2)=-(x_1-1)^2-(x_2+1)^2 \to \max,$

$$\begin{cases} 2x_1+x_2 \ge 2, \\ -x_1+x_2 \le 1, \\ 3x_1-2x_2 \le 6, \\ x_1,x_2 \ge 0; \end{cases}$$

3) $z(x_1,x_2)=2(x_1-4)^2+(x_2-7)^2 \to \min,$

$$\begin{cases} 2x_1+x_2 \ge 2, \\ -x_1+x_2 \le 1, \\ 3x_1-2x_2 \le 6, \\ x_1,x_2 \ge 0. \end{cases}$$

Потрібно записати функцію Лагранжа даної задачі і знайти її сідлову точку, використовуючи розв'язок задачі, отриманий графічно.

Розв'язання. 1) Сім'єю ліній рівня функції $f(x_1,x_2)=(x_1-10)^2+(x_2+1)^2$ буде сім'я концентричних кіл $(x_1-10)^2+(x_2+1)^2=C$ з центром в точці $0_1(10,-1)$. Застосування графічного методу розв'язання ЗНП (див п. 4.2) дозволяє знайти точку $E(4,3)$, яка є підозрілою на екстремум (рис. 4.7).

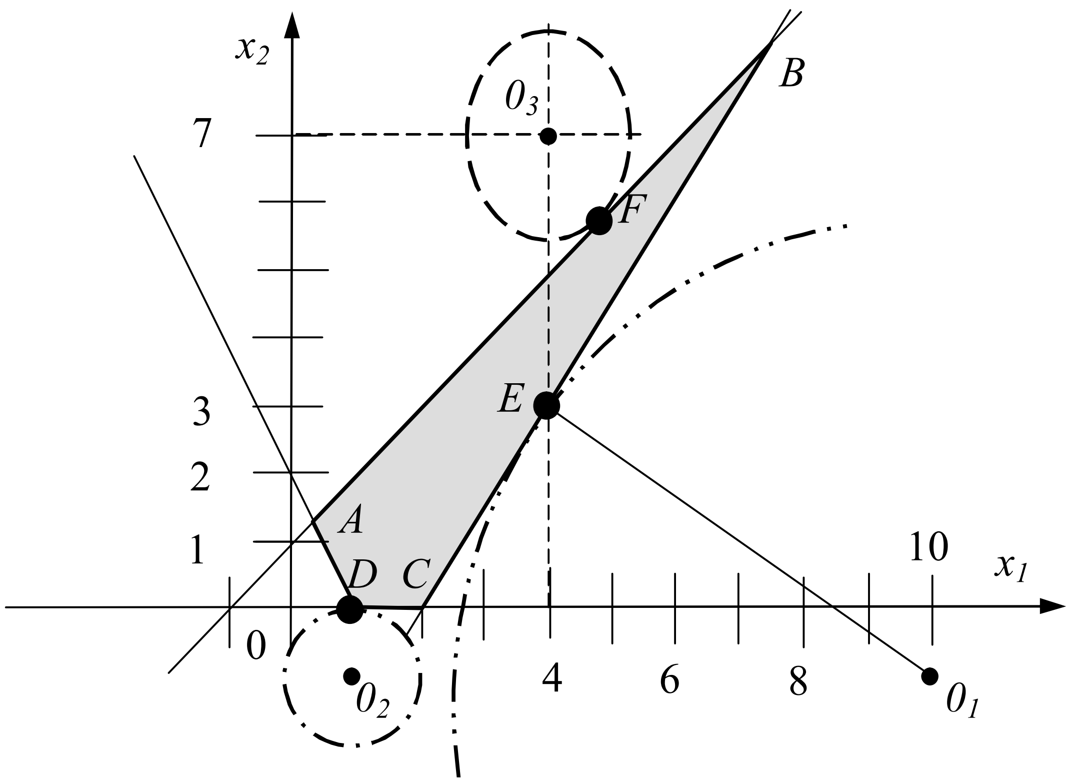

Рис. 4.7. Графічне розв'язання прикладу 10

Необхідно перевірити виконання умов Куна-Таккера для того, щоб довести, що в цій точці дійсно досягається мінімум функції. Зведемо поставлену задачу до вигляду (4.18)-(4.20):

$$\tilde{f}(x_1, x_2) = -f(x_1, x_2) = -(x_1 - 10)^2 - (x_2 + 1)^2 \to \max,$$

$$\begin{cases} 2x_1 + x_2 - 2 \ge 0, \\ x_1 - x_2 + 1 \ge 0, \\ -3x_1 + 2x_2 + 6 \ge 0, \\ x_1, x_2 \ge 0 \end{cases}$$

і складемо функцію Лагранжа:

$$L(X, \Lambda) = -(x_1 - 10)^2 - (x_2 + 1)^2 + \lambda_1(2x_1 + x_2 - 2) + \lambda_2(x_1 - x_2 + 1) + \\ + \lambda_3(-3x_1 + 2x_2 + 6).$$

Частинні похідні цієї функції

$$\frac{\partial L}{\partial x_1} = -2(x_1 - 10) + 2\lambda_1 + \lambda_2 - 3\lambda_3, \qquad \frac{\partial L}{\partial x_2} = -2(x_2 + 1) + \lambda_1 - \lambda_2 + 2\lambda_3,$$

$$\frac{\partial L}{\partial \lambda_1} = 2x_1 + x_2 - 2, \qquad \frac{\partial L}{\partial \lambda_2} = x_1 - x_2 + 1, \qquad \frac{\partial L}{\partial \lambda_3} = -3x_1 + 2x_2 + 6.$$

Використовуючи умови (4.23) і (4.25) та отриману графічним методом точку $E(4,3)$, запишемо систему рівнянь:

$$\begin{cases} x_1^* \dfrac{\partial L(X^*, \Lambda^*)}{\partial x_1} = 4\left(-2(4 - 10) + 2\lambda_1^* + \lambda_2^* - 3\lambda_3^*\right) = 4\left(12 + 2\lambda_1^* + \lambda_2^* - 3\lambda_3^*\right) = 0, \\ x_2^* \dfrac{\partial L(X^*, \Lambda^*)}{\partial x_2} = 3\left(-2(3 + 1) + \lambda_1^* - \lambda_2^* + 2\lambda_3^*\right) = 3\left(-8 + \lambda_1^* - \lambda_2^* + 2\lambda_3^*\right) = 0, \\ \lambda_1^* \dfrac{\partial L(X^*, \Lambda^*)}{\partial \lambda_1} = \lambda_1^*(2 \cdot 4 + 3 - 2) = 9\lambda_1^* = 0, \\ \lambda_2^* \dfrac{\partial L(X^*, \Lambda^*)}{\partial \lambda_2} = \lambda_2^*(4 - 3 + 1) = 2\lambda_2^* = 0, \\ \lambda_3^* \dfrac{\partial L(X^*, \Lambda^*)}{\partial \lambda_3} = \lambda_3^*(-3 \cdot 4 + 2 \cdot 3 + 6) = 0 \cdot \lambda_3^* = 0, \end{cases}$$

або

$$\begin{cases} 12 + 2\lambda_1^* + \lambda_2^* - 3\lambda_3^* = 0, \\ -8 + \lambda_1^* - \lambda_2^* + 2\lambda_3^* = 0, \\ 9\lambda_1^* = 0, \\ 2\lambda_2^* = 0. \end{cases}$$

Отже, $\lambda_1^* = \lambda_2^* = 0, \lambda_3^* = 4$. Таким чином, сідлова точка функції Лагранжа $\left(X^*, \Lambda^*\right) = (4,3,0,0,4)$. Перевіримо умови сідлової точки (4.21).

$$L\left(\tilde{O}, \Lambda^*\right) = -(x_1 - 10)^2 - (x_2 + 1)^2 + 0 \cdot (2x_1 + x_2 - 2) + 0 \cdot (x_1 - x_2 + 1) +$$
$$+ 4(-3x_1 + 2\tilde{o}_2 + 6) = -x_1^2 + 20\tilde{o}_1 - 100 - \tilde{o}_2^2 - 2\tilde{o}_2 - 1 - 12x_1 + 8\tilde{o}_2 + 24 =$$
$$= -x_1^2 + 8\tilde{o}_1 - \tilde{o}_2^2 + 6\tilde{o}_2 - 77 = -(x_1 - 4)^2 - (x_2 - 3)^2 - 52 =$$
$$= -52 - \left[(x_1 - 4)^2 + (x_2 - 3)^2\right] \le -52,$$

$$L\left(X^*, \Lambda^*\right) = -(4 - 10)^2 - (3 + 1)^2 + 0 \cdot (2 \cdot 4 + 3 - 2) + 0 \cdot (4 - 3 + 1) +$$
$$+ 4(-3 \cdot 4 + 2 \cdot 3 + 6) = -52,$$

$$L(X^*, \Lambda) = -(4 - 10)^2 - (3 + 1)^2 + \lambda_1^* \cdot (2 \cdot 4 + 3 - 2) + \lambda_2^* \cdot (4 - 3 + 1) +$$
$$+ \lambda_3^*(-3 \cdot 4 + 2 \cdot 3 + 6) = -52 + \left[9\lambda_1^* + 8\lambda_2^*\right] \ge -52,$$

тобто $L\left(X, \Lambda^*\right) \le L\left(X^*, \Lambda^*\right) \le L\left(X^*, \Lambda\right)$. Умова сідлової точки виконана, що вказує на те, що точка (4,3) є точкою, у якій функція $\tilde{f}(x_1, x_2) = -(x_1 - 10)^2 - (x_2 + 1)^2$ досягає максимального значення, а відповідно функція $f(x_1, x_2) = (x_1 - 10)^2 + (x_2 + 1)^2$ свого мінімального значення.

2) Згідно з рис. 4.7 функції $g(x_1, x_2) = -(x_1 - 1)^2 - (x_2 + 1)^2$ досягає свого максимального значення в точці $D(1,0)$. Функція Лагранжа

$$L(X,\Lambda)=-(x_1-1)^2-(x_2+1)^2+\lambda_1(2x_1+x_2-2)+\lambda_2(x_1-x_2+1)+$$
$$+\lambda_3(-3x_1+2x_2+6)$$

має сідлову точку $\left(X^*,\Lambda^*\right)=(1,0,0,0,0)$, для якої виконуються умови (4.21). Дійсно

$$L\left(\tilde{O},\Lambda^*\right)=-(x_1-1)^2-(x_2+1)^2+0\cdot(2x_1+x_2-2)+0\cdot(x_1-x_2+1)+$$
$$+0\cdot(-3x_1+2\tilde{o}_2+6)=-(x_1-1)^2-\tilde{o}_2^2-2\tilde{o}_2-1=-1-\left[(x_1-1)^2+\tilde{o}_2^2+2\tilde{o}_2\right]\le -1,$$
($\tilde{o}_2\ge 0\Rightarrow\ (x_1-1)^2+\tilde{o}_2^2+2\tilde{o}_2\ge 0$),

$$L\left(X^*,\Lambda^*\right)=-(1-1)^2-(0+1)^2+0\cdot(2\cdot 1+0-2)+0\cdot(1-1+1)+$$
$$+0\cdot(-3\cdot 1+2\cdot 0+6)=-1,$$

$$L(X^*,\Lambda)=-(1-1)^2-(0+1)^2+\lambda_1^*\cdot(2\cdot 1+0-2)+\lambda_2^*\cdot(1-1+1)+$$
$$+\lambda_3^*(-3\cdot 1+2\cdot 0+6)=-1+\left[\lambda_2^*+3\lambda_3^*\right]\ge -1.$$

3) Оскільки $N(-1,1)$ – нормальний вектор прямої $-x_1+x_2=1$, а $grad\, z=(4(x_1-4),2(x_2-7))$ – вектор-градієнт цільової функції $z(x_1,x_2)=2(x_1-4)^2+(x_2-7)^2$, то координати точки F знаходимо, розв'язавши систему рівнянь

$$\begin{cases}\dfrac{4(x_1-4)}{-1}=\dfrac{2(x_2-7)}{1},\\ -x_1+x_2=1,\end{cases}\quad x_1^*=\frac{14}{3},\ \ x_2^*=\frac{17}{3}.$$

Точка $\left(X^*,\Lambda^*\right)=\left(\frac{14}{3},\frac{17}{3},0,\frac{8}{3},0\right)$ є сідловою точкою функції Лагранжа

$$L(X,\Lambda)=-2(x_1-4)^2-(x_2-7)^2+\lambda_1(2x_1+x_2-2)+\lambda_2(x_1-x_2+1)+$$
$$+\lambda_3(-3x_1+2x_2+6),$$

оскільки

$$L\left(\tilde{O},\Lambda^*\right)=-2(x_1-4)^2-(x_2-7)^2+0\cdot(2x_1+x_2-2)+\frac{8}{3}\cdot(x_1-x_2+1)+$$

$$+0\cdot(-3x_1+2\tilde{o}_2+6)=-2x_1^2+16\tilde{o}_1-32-\tilde{o}_2^2+14\tilde{o}_2-49+\frac{8}{3}x_1-\frac{8}{3}\tilde{o}_2+\frac{8}{3}=$$

$$=-\frac{8}{3}-2\left(x_1-\frac{14}{3}\right)^2-\left(x_2-\frac{17}{3}\right)^2-\frac{8}{3}-\left[2\left(x_1-\frac{14}{3}\right)^2+\left(x_2-\frac{17}{3}\right)^2\right]\le-\frac{8}{3},$$

$$L\left(X^*,\Lambda^*\right)=-\frac{8}{3},$$

$$L(\tilde{O}^*,\Lambda)=-2\left(\frac{14}{3}-4\right)^2-\left(\frac{17}{3}-7\right)^2+\lambda_1^*\cdot\left(2\cdot\frac{14}{3}+\frac{17}{3}-2\right)+\lambda_2^*\cdot\left(\frac{14}{3}-\frac{17}{3}+1\right)+$$

$$+\lambda_3^*\cdot\left(-3\cdot\frac{14}{3}+2\cdot\frac{17}{3}+6\right)=-\frac{8}{3}+\left[13\lambda_1^*+\frac{10}{3}\lambda_3^*\right]\ge-\frac{8}{3}.$$

Відповідь: 1) $\min f=-\widetilde{f}(4,3)=f(4,3)=52$;

2) $\max g=g(1,0)=-1$;

3) $\min z=z\left(\frac{14}{3},\frac{17}{3}\right)=-\frac{8}{3}$;

Приклад 11. Знайти максимальне значення функції $f=-x_1^2+6x_1-x_2^2+8x_2$ при додаткових обмеженнях

$$\begin{cases}4x_1+5x_2\le 20,\\ -2x_1+3x_2\le 1,\\ x_1,x_2\ge 0.\end{cases}$$

Розв'язання. Це задача квадратичного програмування. Складемо функцію Лагранжа:

$$L(X,\Lambda)=-x_1^2+6x_1-x_2^2+8x_2+\lambda_1(20-4x_1-5x_2)+\lambda_2(1+2x_1-3x_2)$$

і запишемо для неї умови Куна-Таккера:

$$\begin{cases} \dfrac{\partial L}{\partial x_1} = -2x_1 + 6 - 4\lambda_1 + 2\lambda_2 \le 0, \\ \dfrac{\partial L}{\partial x_2} = -2x_2 + 8 - 5\lambda_1 - 3\lambda_2 \le 0, \end{cases} \tag{4.41}$$

$$\begin{cases} \dfrac{\partial L}{\partial \lambda_1} = 20 - 4x_1 - 5x_2 \ge 0, \\ \dfrac{\partial L}{\partial \lambda_2} = 1 + 2x_1 - 3x_2 \ge 0, \end{cases} \tag{4.42}$$

$$\begin{cases} x_1 \dfrac{\partial L}{\partial x_1} = x_1\left(-2x_1 + 6 - 4\lambda_1 + 2\lambda_2\right) = 0, \\ x_2 \dfrac{\partial L}{\partial x_2} = x_2\left(-2x_2 + 8 - 5\lambda_1 - 3\lambda_2\right) = 0, \end{cases} \tag{4.43}$$

$$\begin{cases} \lambda_1 \dfrac{\partial L}{\partial \lambda_1} = \lambda_1\left(20 - 4x_1 - 5x_2\right) = 0, \\ \lambda_2 \dfrac{\partial L}{\partial \lambda_2} = \lambda_2\left(1 + 2x_1 - 3x_2\right) = 0. \end{cases} \tag{4.44}$$

Перепишемо систему нерівностей (4.41)-(4.42) у вигляді

$$\begin{cases} 2x_1 + 4\lambda_1 - 2\lambda_2 \ge 6, \\ 2x_2 + 5\lambda_1 + 3\lambda_2 \ge 8, \\ 4x_1 + 5x_2 \le 20, \\ -2x_1 + 3x_2 \le 1. \end{cases} \tag{4.45}$$

Після введення додаткових невід'ємних змінних v_1, v_2, w_1, w_2 отримаємо систему рівнянь

$$\begin{cases} 2x_1 + 4\lambda_1 - 2\lambda_2 - v_1 = 6, \\ 2x_2 + 5\lambda_1 + 3\lambda_2 - v_2 = 8, \\ 4x_1 + 5x_2 + w_1 = 20, \\ -2x_1 + 3x_2 + w_2 = 1. \end{cases} \tag{4.46}$$

Тоді система нерівностей (4.43)-(4.44) набуде вигляду

$$\begin{cases} x_1 v_1 = 0, \\ x_2 v_2 = 0, \\ \lambda_1 w_1 = 0, \\ \lambda_2 w_2 = 0. \end{cases} \qquad (4.47)$$

Для знаходження розв'язку використаємо метод штучного базису (п. 2.6). У перше і друге рівняння системи (4.46), які відповідають обмеженням задачі, введемо штучні змінні y_1, y_2 і розглянемо задачу лінійного програмування

$$F = -My_1 - My_2 \to \max,$$

$$\begin{cases} 2x_1 + 4\lambda_1 - 2\lambda_2 - v_1 + y_1 = 6, \\ 2x_2 + 5\lambda_1 + 3\lambda_2 - v_2 + y_2 = 8, \\ 4x_1 + 5x_2 + w_1 = 20, \\ -2x_1 + 3x_2 + w_2 = 1, \end{cases}$$

$$x_1, x_2, v_1, v_2, \lambda_1, \lambda_2, w_1, w_2, y_1, y_2 \ge 0 .$$

Розв'язання проведемо за допомогою симплексних таблиць (табл. 4.1-4.3).

Таблиця 4.1

Початкова таблиця

Приклад 11		Вільна змінна						Вільний член
		$-x_1$	$-x_2$	$-\lambda_1$	$-\lambda_2$	$-v_1$	$-v_2$	*1*
Базисна змінна	y_1	2	0	4	-2	-1	0	6
	y_2	0	2	5	3	0	-1	8
	w_1	4	5	0	0	0	0	20
	w_2	-2	3	0	0	0	0	1
М-рядок		-2	-2	-9	-1	1	1	-14

Таблиця 4.2

Ітерація 1

Після обміну $\lambda_1 \leftrightarrow y_1$		Вільна змінна						Вільний член
		$-x_1$	$-x_2$	$-y_1$	$-\lambda_2$	$-v_1$	$-v_2$	1
Базисна змінна	λ_1	1/2	0	1/4	-1/2	-1/4	0	3/2
	y_2	-5/2	2	-5/4	11/2	5/4	-1	1/2
	w_1	4	5	0	0	0	0	20
	w_2	-2	3	0	0	0	0	1
М-рядок		5/2	-2	9/4	-11/2	-5/4	1	-1/2

Таблиця 4.3

Ітерація 2

Після обміну $\lambda_2 \leftrightarrow y_2$		Вільна змінна						Вільний член
		$-x_1$	$-x_2$	$-y_1$	$-y_2$	$-v_1$	$-v_2$	1
Базисна змінна	λ_1	3/11	2/11	3/22	1/11	-3/22	-1/11	17/11
	λ_2	-5/11	4/11	-5/22	2/11	5/22	-2/11	1/11
	w_1	4	5	0	0	0	0	20
	w_2	-2	3	0	0	0	0	1
М-рядок		0	0	1	1	0	0	0

Отже, отримали оптимальний план $x_1^* = x_2^* = 0$, $\lambda_1^* = \dfrac{17}{11}, \lambda_2^* = \dfrac{1}{11}$, $v_1 = v_2 = 0$, $w_1 = 20, w_2 = 1$. Перевіримо виконання умов (4.47):

$$\begin{cases} x_1^* v_1 = 0, \\ x_2^* v_2 = 0, \\ \lambda_1 w_1 = \dfrac{17}{11} \cdot 20 \neq 0, \\ \lambda_2 w_2 = \dfrac{1}{11} \cdot 1 \neq 0. \end{cases}$$

Умови (4.47) не виконуються, тому знайдемо інший розв’язок задачі, для чого оберемо в якості провідного будь-який стовпець з нульовою оцінкою в М-рядку, наприклад той, який відповідає вільній змінній x_1 (табл. 4.4).

Таблиця 4.4

Ітерація 3

Після обміну $x_1 \leftrightarrow w_1$		Вільна змінна						Вільний член
		$-w_1$	$-x_2$	$-y_1$	$-y_2$	$-v_1$	$-v_2$	1
Базисна змінна	λ_1	-3/44	-7/44	3/22	1/11	-3/22	-1/11	2/11
	λ_2	5/44	41/44	-5/22	2/11	5/22	-2/11	26/11
	x_1	1/4	5/4	0	0	0	0	5
	w_2	-1/2	11/2	0	0	0	0	11
М-рядок		0	0	1	1	0	0	0

Отриманий план $x_1^* = 5, x_2^* = 0$, $\lambda_1^* = \frac{2}{11}, \lambda_2^* = \frac{26}{11}$, $v_1 = v_2 = 0$, $w_1 = 0, w_2 = 11$ знов не задовольняє умови (4.47), бо $\lambda_2 w_2 = \frac{26}{11} \cdot 11 \neq 0$.

Продовжуючи обчислення, приходимо до табл. 4.5.

Таблиця 4.5

Ітерація 4

Після обміну $x_2 \leftrightarrow w_2$		Вільна змінна						Вільний член
		$-w_1$	$-w_2$	$-y_1$	$-y_2$	$-v_1$	$-v_2$	1
Базисна змінна	λ_1							1/2
	λ_2							1/2
	x_1							5/2
	x_2							2
М-рядок		0	0	1	1	0	0	0

Отримали розв'язок $x_1^* = 2{,}5$, $x_2^* = 2$, $\lambda_1^* = \lambda_2^* = \frac{1}{2}$, $v_1 = v_2 = 0$, $w_1 = w_2 = 0$. Легко перевірити, що цей розв'язок задовольняє умови (4.47). Звідси випливає, що точка $(X^*,\Lambda^*)=\left(\frac{5}{2},2,\frac{1}{2},\frac{1}{2}\right)$ є сідловою точкою функції Лагранжа.

Відповідь: $\max f = f\left(\frac{5}{2},2\right)=-\frac{25}{4}+6\cdot\frac{5}{2}-4+16=\frac{83}{4}$.

Приклад 12. Знайти максимальне значення функції $f=-2x_1^2+12x_1-x_2^2+6x_2$ при додаткових обмеженнях

$$\begin{cases} 2x_1+3x_2\le 12, \\ x_1-x_2\ge 1, \\ x_1,x_2\ge 0. \end{cases} \tag{4.48}$$

Розв'язання. Складемо функцію Лагранжа:

$$L(X,\Lambda)=-2x_1^2+12x_1-x_2^2+6x_2+\lambda_1(12-2x_1-3x_2)+\lambda_2(x_1-x_2-1).$$

Знаходимо частинні похідні:

$$\frac{\partial L}{\partial x_1}=-4x_1+12-2\lambda_1+\lambda_2,\ \frac{\partial L}{\partial x_2}=-2x_2+6-3\lambda_1-\lambda_2,$$
$$\frac{\partial L}{\partial \lambda_1}=12-2x_1-3x_2,\ \frac{\partial L}{\partial \lambda_2}=x_1-x_2-1,$$

і виписуємо систему рівнянь (4.23)-(4.25):

$$\begin{cases} x_1(-4x_1+12-2\lambda_1+\lambda_2)=0, \\ x_2(-2x_2+6-3\lambda_1-\lambda_2)=0, \\ \lambda_1(12-2x_1-3x_2)=0, \\ \lambda_2(x_1-x_2-1)=0. \end{cases}$$

При розв'язанні цієї системи потрібно розглянути можливі випадки значень невідомих:

1) $x_1>0, x_2>0$;
2) $x_1>0, x_2=0$;

3) $x_1 = 0, x_2 > 0$;

4) $x_1 = 0, x_2 = 0$.

1. У випадку, коли $x_1 > 0, x_2 > 0$, умови Куна-Таккера набувають вигляду

$$\begin{cases} -4x_1 + 12 - 2\lambda_1 + \lambda_2 = 0, \\ -2x_2 + 6 - 3\lambda_1 - \lambda_2 = 0, \\ \lambda_1(12 - 2x_1 - 3x_2) = 0, \\ \lambda_2(x_1 - x_2 - 1) = 0. \end{cases}$$

Звідси отримаємо такі розв'язки:

а) $x_1 = 3, x_2 = 3, \lambda_1 = \lambda_2 = 0$;

б) $x_1 = \frac{10}{3}, x_2 = \frac{7}{3}, \lambda_1 = 0, \lambda_2 = \frac{4}{3}$;

в) $x_1 = \frac{30}{11}, x_2 = \frac{24}{11}, \lambda_1 = \frac{6}{11}, \lambda_2 = 0$;

г) $x_1 = 3, x_2 = 2, \lambda_1 = \frac{2}{5}, \lambda_2 = \frac{4}{5}$.

Необхідно перевірити, що знайдені точки належать ОДР.

Перші три розв'язки (3,3), $\left(\frac{10}{3}, \frac{7}{3}\right)$ і $\left(\frac{30}{11}, \frac{24}{11}\right)$ не задовольняють обмеження задачі (4.48). Отже, маємо лише одну точку $(X, \Lambda) = \left(3, 2, \frac{2}{5}, \frac{4}{5}\right)$, $f(3,2) = 26$.

2. Якщо $x_1 > 0, x_2 = 0$, то за виразами (4.23) і (4.25) необхідно знайти розв'язок системи рівнянь

$$\begin{cases} -4x_1 + 12 - 2\lambda_1 + \lambda_2 = 0, \\ \lambda_1(12 - 2x_1) = 0, \\ \lambda_2(x_1 - 1) = 0, \end{cases}$$

який задовольняє нерівність $6 - 3\lambda_1 - \lambda_2 \le 0$. Серед отриманих розв'язків системи $(3,0,0,0)$, $(6,0,-6,0)$ і $(1,0,0,-8)$ тільки перший задовольняє цю нерівність і обмеженням (4.48). Таким чином, маємо ще одну точку $(X,\Lambda)=(3{,}0{,}0{,}0)$, $f(3{,}0)=18$.

3. Випадки, коли $x_1=0, x_2>0$ і $x_1=0, x_2=0$, можна взагалі не розглядати, бо вони суперечать другому обмеженню в умові задачі ($x_1-x_2\ge 1$).

Порівнюючи отримані значення функції, бачимо, що свого найбільшого значення вона досягає в точці X=(3,2).

Перевіримо виконання умов сідлової точки для $(X,\Lambda)=\left(3{,}2{,}\frac{2}{5},\frac{4}{5}\right)$:

$$L\left(X,\Lambda^*\right)=-2x_1^2+12x_1-x_2^2+6x_2+\frac{2}{5}(12-2x_1-3x_2)+\frac{4}{5}(x_1-x_2-1)=$$

$$=26-2(x_1+3)^2-(x_2-2)^2=26-\left[2(x_1+3)^2+(x_2-2)^2\right]\le 26,$$

$$L\left(X^*,\Lambda^*\right)=-2\cdot 9+12\cdot 3-4+12+\frac{2}{5}\cdot(12-6-6)+\frac{4}{5}\cdot(3-2-1)=26,$$

$$L\left(X^*,\Lambda\right)=26+0\cdot\lambda_1^*+0\cdot\lambda_2^*=26.$$

Як бачимо, умова (4.21) виконана.

Відповідь: $\max f=f(3{,}2)=26$.

Приклад 13. Для виробництва двох видів продукції А і В підприємство закуповує сировину в кількості 18 од. за ціною 180 грош. од. З одиниці сировини можна виготовити одиницю продукції А, затративши на переробку 4 грош. од., або одиницю продукції В, затративши 10 грош. од. Дохід від реалізації a одиниць продукції першого типу складає $a(30-a)$ грош. од., другого типу $a(60-2a)$ грош. од. Визначити стратегію підприємства, яка дозволить максимізувати прибуток.

Розв'язання. Введемо наступні позначення:

x_1 - кількість одиниць продукції А;

x_2 - кількість одиниць продукції В.

Дохід від реалізації виготовленої продукції складає $x_1(30-x_1)+x_2(60-2x_2)$ грош. од. Затрати на виробництво: $4x_1+10x_2+\frac{180}{18}(x_1+x_2)=14x_1+20x_2$. Таким чином, прибуток підприємства дорівнює

$$x_1(30-x_1)+x_2(60-2x_2)-14x_1-20x_2=-x_1^2+16x_1-2x_2^2+40x_2.$$

Отже, математична модель задачі має вигляд

$$f=-x_1^2+16x_1-2x_2^2+40x_2 \to \max,$$
$$\begin{cases} x_1+x_2 \le 18, \\ x_1, x_2 \ge 0. \end{cases}$$

Розв’язуючи задачу, наприклад графічним методом, отримуємо розв’язок $\max f = f(8{,}10)=264$.

Відповідь: Максимальний прибуток у розмірі 264 грош. од. буде отриманий, якщо буде виготовлено 8 од. продукції першого типу і 10 од. продукції другого типу.

4.6. Дробово-лінійне програмування

Задачею дробово-лінійного програмування (ЗДЛП) називається задача

$$f=\frac{c_1x_1+\ldots+c_nx_n+c_0}{d_1x_1+\ldots+d_nx_n+d_0} \to \max(\min), \tag{4.49}$$

$$\sum_{j=1}^{n} a_{ij}x_j \le b_i,\ i=\overline{1,m}, \tag{4.50}$$

$$x_j \ge 0,\ j=\overline{1,n}, \tag{4.51}$$

де $d_1x_1+\ldots+d_nx_n+d_0 \ne 0$ на ОДР.

4.6.1. Графічний метод розв’язання задачі дробово-лінійного програмування

Якщо ЗДЛП містить дві змінні, її можна розв’язати *графічним методом.*

Розглянемо таку задачу