Міністерство освіти і науки, молоді та спорту України
Державний вищий навчальний заклад
"Ужгородський національний університет"
Математичний факультет
Кафедра системного аналізу і теорії оптимізації

МЕТОДИЧНІ ВКАЗІВКИ

до практичних занять курсу
«МЕТОДИ ОПТИМІЗАЦІЇ»

Частина I

СКІНЧЕННОВИМІРНІ ЗАДАЧІ УМОВНОЇ ОПТИМІЗАЦІЇ

Ужгород – 2015

Гренджа В.І., Брила А.Ю., Ломага М.М. Методичні вказівки до практичних занять з курсу «Методи оптимізації». Ч. І. Скінченновимірні задачі умовної оптимізації.– Ужгород, 2015 .– 30с.

Розглядаються методи розв’язання скінченновимірних задач умовної оптимізації з обмеженнями рівностями (класична задача на умовний екстремум), скінченновимірних задач з обмеженнями рівностями і нерівностями та задач опуклого програмування. Наведено необхідні теоретичні відомості, загальні правила та приклади розв’язання деяких практичних задач. Запитання для самоконтролю разом із завданнями для самостійного виконання дозволяють більш глибоко засвоїти теоретичний матеріал.

Рецензенти: канд. фіз.-мат. наук, доц. Погоріляк О.О.,
канд. фіз.-мат. наук, доц. Юрченко Н.В.

Рекомендовано до друку Вченою радою математичного факультету ДВНЗ “Ужгородський національний університет” від 16 квітня 2015 року, протокол № 9.

Основні поняття та означення

При розв'язанні будь-якої задачі вибору необхідно визначити над якою множиною альтернатив потрібно здійснити вибір. Позначимо її X. Таку множину ще називають *множиною допустимих альтернатив* або ж *допустимою множиною*. Як правило, в скінченновимірних задачах умовної оптимізації допустима множина є власною підмножиною простору $\mathbf{R}^n$ і задається за допомогою обмежень (рівностей і нерівностей).

Звичайно, людина прагне завжди вибрати найкращу або непокращувану альтернативу $x^* \in X$, яку ще називають *оптимальною*. Часто вибір здійснюється шляхом попарного порівняння альтернатив за допомогою певним чином встановлених їх числових оцінок. Функція, яка кожній альтернативі ставить у відповідність її оцінку, називається *цільовою функцією*

$$f : X \to \mathbf{R}.$$

Задачу знаходження альтернативи $x^*_{\max} \in X$, такої, що $f(x^*_{\max}) \geq f(x), \forall x \in X$, називають *задачею максимізації* і записують так:

$$\max f(x),\ x \in X,$$

або ж

$$f(x) \to \max,\ x \in X.$$

Альтернативу $x^*_{\max}$ називають *оптимальним розв'язком* або розв'язком задачі, а допустиму альтернативу $x \in X$ – її *допустимим розв'язком*. Значення $f^*_{\max} = f(x^*_{\max})$ називають *максимумом функції* $f(x)$ *на множині* X, а $x^*_{\max}$ – *точкою глобального максимуму функції* $f(x)$ *на множині* X. Часто для її позначення використовують запис $x^*_{\max} \in \arg\max\limits_{x \in X} f(x)$.

Означення. Точка $x^*_{\max} \in X$ називається точкою *глобального максимуму* функції $f(x)$ на множині X, якщо $f(x)$ досягає в цій точці найбільшого значення, тобто

$$f(x^*_{\max}) \geq f(x),\ \forall x \in X.$$

Множину всіх точок глобального максимуму позначають $abs\max\limits_{x \in X} f(x)$.

Означення. Точка $x^*_{\max} \in X$ називається точкою *локального максимуму* функції $f(x)$ на множині X, якщо існує такий її ε-окіл

$$U_\varepsilon(x^*_{\max}) = \{x \in \mathbf{R}^n \mid \|x - x^*_{\max}\| < \varepsilon\},$$

що

$$\forall x \in X \cap U_\varepsilon(x^*_{\max}),\ f(x^*_{\max}) \geq f(x).$$

Множину точок локального максимуму позначають $loc\max\limits_{x \in X} f(x)$.

Задачу знаходження альтернативи $x^*_{\min} \in X$, такої, що $f(x^*_{\min}) \le f(x), \forall x \in X$, називають *задачею мінімізації* і записують так:

$$\min f(x), x \in X,$$

або ж

$$f(x) \to \min, x \in X.$$

Значення $f^*_{\min} = f(x^*_{\min})$ називають *мінімальним значенням функції* $f(x)$ *на множині* X, а $x^*_{\min}$ – *точкою глобального мінімуму*. Для точки глобального мінімуму часто використовують позначення $x^*_{\mathbf{min}} \in \arg\min_{x \in X} f(x)$.

Означення. Точка $x^*_{\min} \in X$ називається точкою *глобального мінімуму* функції $f(x)$ на множині X, якщо $f(x)$ досягає в цій точці найменшого значення, тобто

$$f(x^*_{\min}) \le f(x), \forall x \in X.$$

Множину точок глобального мінімуму позначають $abs \min_{x \in X} f(x)$.

Означення. Точка $x^*_{\min} \in X$ називається точкою *локального мінімуму* функції $f(x)$ на множині X, якщо існує такий її ε-окіл

$$U_\varepsilon(x^*_{\mathbf{min}}) = \{x \in \mathbf{R}^n \mid \|x - x^*_{\mathbf{min}}\| < \varepsilon\},$$

що

$$\forall x \in X \cap U_\varepsilon(x^*_{\mathbf{min}}),\ f(x^*_{\min}) \le f(x).$$

Множину точок локального мінімуму позначають $loc \min_{x \in X} f(x)$.

Відмітимо, що будь-яка задача мінімізації

$$\min f(x), x \in X,$$

еквівалентна задачі максимізації

$$\max\ (-f(x)), x \in X.$$

Якщо допустима множина X співпадає з простором $\mathbf{R}^n$, то задачу

$$f(x) \to extr, x \in \mathbf{R}^n,$$

називають *скінченновимірною задачею безумовної оптимізації*.

Якщо допустима множина X є власною підмножиною простору $\mathbf{R}^n$, то задачу

$$f(x) \to extr,\ x \in X \subset \mathbf{R}^n,$$

називають *скінченновимірною задачею умовної оптимізації*.

У роботі розглядаються скінченновимірні задачі умовної оптимізації та загальні правила їх розв'язування. У своїй більшості такі правила дозволяють виділити деяку підмножину допустимих точок, серед яких містяться розв'язки задачі, якщо вони існують. Ці підозрілі на екстремум цільової функції точки називають *критичними*. Виділити такі точки дозволяють, зокрема, *необхідні умови оптимальності* – умови, яким *повинна* задовольняти допустима точка, яка є точкою локального екстремуму цільової функції.

Дослідження характеру критичних точок, тобто виділення серед них таких точок, в яких дійсно досягається глобальний або локальний, строгий або нестрогий мінімум чи максимум цільової функції, здійснюється або на основі означень, або на основі достатніх умов оптимальності. *Достатні умови екстремуму* – це умови, з яких випливає, що критична точка, яка їм задовольняє, дійсно є точкою локального мінімуму або локального максимуму цільової функції.

Множина критичних точок може бути дещо ширшою за множину абсолютних і навіть локальних екстремумів. Однак, як правило, вона містить не дуже велику кількість точок. Тому, довівши, що розв'язки задачі існують, їх можна знайти тим чи іншим способом. Наприклад, їх можна знайти в результаті аналізу властивостей задачі або шляхом обчислення та порівняння значень цільової функції в усіх критичних точках і відбору з них точок з найменшим або з найбільшим значенням цільової функції, або на основі геометричного чи фізичного змісту задачі.

Очевидно, якщо існування розв'язку екстремальної задачі зрозуміло з її геометричного або фізичного змісту, або випливає з властивостей задачі, і при цьому існує тільки одна допустима критична точка, то вона і буде оптимальним розв'язком.

1. Задача з обмеженнями-рівностями. Класична задача на умовний екстремум

Класичною задачею на умовний екстремум або *гладкою скінченновимірною задачею з обмеженнями типу рівностей* називається задача:

$$f(x) \to extr, \tag{1}$$

$$g_i(x) = 0, \ i = \overline{1,m}, \quad x \in \mathbf{R}^n, \tag{2}$$

де функції $f(x)$, $g_i(x)$, $i = \overline{1,m}$ є диференційованими функціями змінних x_j, $j = \overline{1,n}$. Умови (2) прийнято називати *рівняннями зв'язку*.

Дана задача є задачею на умовний екстремум, тобто задачею знаходження екстремуму цільової функції $f(x)$ на множині $X = \left\{ x \in \mathbf{R}^n \,\middle|\, g_i(x) = 0, \ i = \overline{1,m} \right\}$.

1.1. Метод виключення змінних розв'язання класичної задачі на умовний екстремум

У тих випадках, коли систему (2) вдається перетворити до еквівалентного вигляду $x_i = \varphi_i(x_{m+1},...,x_n)$, $i = \overline{1,m}$, однозначно виразивши перші m змінних через всі інші, дану задачу на умовний екстремум можна звести до задачі безумовної оптимізації функції $n-m$ змінних $x_{m+1}, x_{m+2}, \dots, x_n$:

$$f\left(\varphi_1(x_{m+1},\dots,x_n),\dots,\varphi_m(x_{m+1},\dots,x_n),\, x_{m+1},\dots,x_n\right) \equiv \tilde{f}(x_{m+1},\dots,x_n) \to extr \quad (3)$$

Якщо $\left(x_{m+1}^*,\dots,x_n^*\right)$ – локальний (глобальний) розв'язок задачі (3), то, очевидно, локальним (глобальним) розв'язком задачі (1)-(2) є вектор $\left(\varphi_1(x_{m+1}^*,\dots,x_n^*),\dots,\varphi_m(x_{m+1}^*,\dots,x_n^*),\, x_{m+1}^*,\dots,x_n^*\right)$.

Описаний ***метод виключення змінних*** має обмежене застосування, так як явно виразити одну групу змінних через інші, як правило, складно або навіть неможливо – однозначні функції φ_i, $i=\overline{1,m}$, існують далеко не завжди.

Приклад 1. Знайти розв'язки задачі

$$xy^3 \to \max, \quad (4)$$
$$8x + 2y = 5, \quad (5)$$
$$x, y \in \mathbf{R}.$$

Розв'язання

Задача (4)-(5) є класичною задачею на умовний екстремум. Її ще можна переписати так:

$$f(x,y) = xy^3 \to \max,$$
$$g(x,y) = 8x + 2y - 5 = 0,$$
$$x, y \in \mathbf{R}.$$

Для розв'язання цієї задачі застосуємо метод виключення змінних. Виразимо змінну x через змінну y

$$x = \frac{1}{8}(5-2y) \quad (6)$$

і підставимо у цільову функцію f, в результаті одержимо функцію $\tilde{f}(y)$ залежну тільки від однієї змінної

$$\tilde{f}(y) = \frac{1}{8}(5-2y)y^3 = \frac{5}{8}y^3 - \frac{1}{4}y^4.$$

Таким чином задачу (4)-(5) зведено до задачі безумовної оптимізації

$$\tilde{f}(y) = \frac{5}{8}y^3 - \frac{1}{4}y^4 \to \max, \quad (7)$$

яку можна розв'язати використовуючи необхідні і достатні умови оптимальності.

Знайдемо критичні точки. У даному випадку це тільки стаціонарні точки, які одержуємо на основі необхідної умови оптимальності першого порядку

$$\tilde{f}'(y) = \frac{15}{8}y^2 - y^3 = 0.$$

Розв'язавши це рівняння одержимо стаціонарні точки: $\bar{y}_1 = 0$, $\bar{y}_2 = \dfrac{15}{8}$.

Для встановлення характеру стаціонарних точок використаємо достатню умову оптимальності другого порядку та умови оптимальності вищих порядків.

$$\tilde{f}''(y)=\frac{15}{4}y-3y^2.$$

$$\left.\begin{aligned}\tilde{f}''(\bar{y}_1)=0,\\ \tilde{f}'''(\bar{y}_1)=\frac{15}{8}\end{aligned}\right\}\Rightarrow \text{точка } \bar{y}_1 \text{ не є точкою екстремуму.}$$

$\tilde{f}''(\bar{y}_2)=\frac{-225}{64}<0\Rightarrow$ точка $\bar{y}_2$ є точкою локального максимуму. Оскільки точка $\bar{y}_2$ є єдиною точкою екстремуму і $\lim\limits_{|y|\to+\infty}\tilde{f}(y)=-\infty$, то ця точка є і точкою глобального максимуму.

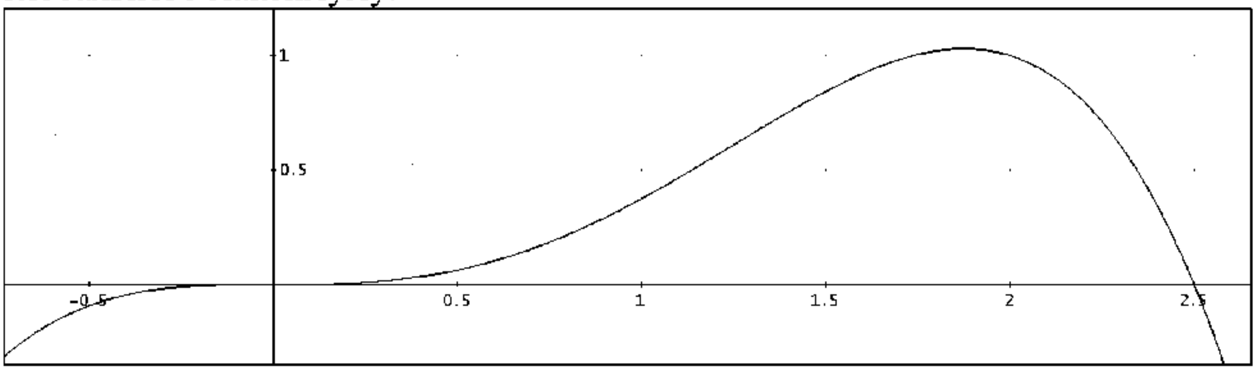

Рис 1. Графік функції $\tilde{f}(y)=\frac{5}{8}y^3-\frac{1}{4}y^4$

Повернемось до задачі (4)-(5). Оскільки точка $\bar{y}_2$ є глобальним розв'язком задачі (7), то точка $x^*_{\max}=(\bar{x}_2,\bar{y}_2)$, де згідно з (6) $\bar{x}_2=\frac{5}{8}-\frac{1}{4}\bar{y}_2=\frac{5}{32}$, є глобальним розв'язком задачі (4)-(5), $f(x^*_{\max})=f(\bar{x}_2,\bar{y}_2)\approx 1{,}03$.

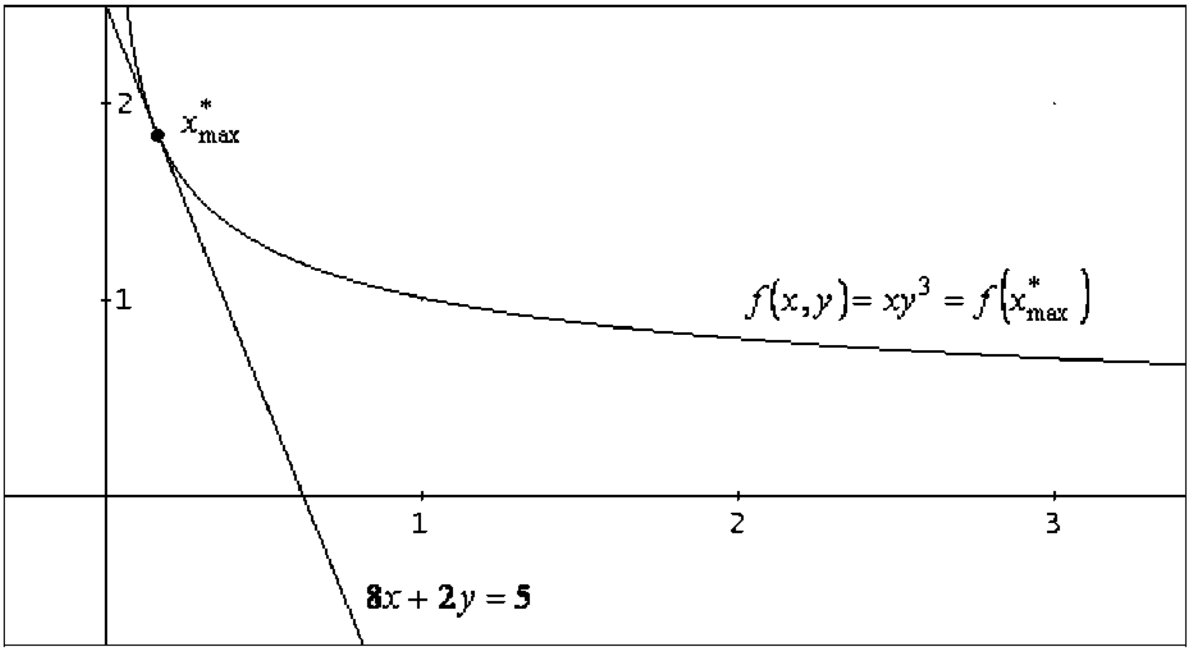

Рис 2. Графіки лінії обмежень та лінії рівня, що проходить через точку глобального максимуму $x^*_{\max}$

Відповідь: Точка $x^*_{\max} = \left(\dfrac{5}{32}, \dfrac{15}{8}\right)$ є точкою глобального максимуму, $f\left(x^*_{\max}\right) \approx 1{,}03$.

1.2. Метод невизначених множників Лагранжа

Більш загальний підхід до розв'язання задачі (1)-(2) дає ***метод невизначених множників Лагранжа***. При дослідженні цієї задачі даним методом важливу роль відіграє її *функція Лагранжа* $L(x, \lambda_0, \lambda)$,

$$L(x, \lambda_0, \lambda) = \lambda_0 f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) \iff L(x, \lambda_0, \lambda) = \lambda_0 f(x) + \langle \lambda, g(x) \rangle ,$$

де $x \in R^n$, $\lambda_0 \in R^1$, $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbf{R}^m$, $g(x) = \left(g_1(x), \ldots, g_m(x)\right)$. Частинні похідні цієї функції по координатам вектора x мають вигляд:

$$\frac{\partial L(x, \lambda_0, \lambda)}{\partial x_j} = \lambda_0 \frac{\partial f(x)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(x)}{\partial x_j}, \quad j = \overline{1, n}.$$

Складений із них вектор $L'_x(x, \lambda_0, \lambda)$ є градієнтом функції Лагранжа по x, тобто,

$$L'_x(x, \lambda_0, \lambda) = \left(\frac{\partial L(x, \lambda_0, \lambda)}{\partial x_1}, \ldots, \frac{\partial L(x, \lambda_0, \lambda)}{\partial x_n} \right) \iff$$

$$\iff L'_x(x, \lambda_0, \lambda) = \lambda_0 f'(x) + \sum_{i=1}^{m} \lambda_i g'_i(x),$$

де $f'(x)=\left(\frac{\partial f(x)}{\partial x_1},...,\frac{\partial f(x)}{\partial x_n}\right)$, $g_i'(x)=\left(\frac{\partial g_i(x)}{\partial x_1},...,\frac{\partial g_i(x)}{\partial x_n}\right)$, $i=\overline{1,m}$, – градієнти функцій $f(x)$ та $g_1(x),...,g_m(x)$ відповідно.

Метод або принцип Лагранжа розв'язання задачі (1)-(2) базується на тому, що її розв'язок може досягатись тільки в критичних точках задачі на безумовний екстремум

$$L(x,\lambda_0,\lambda)\to extr. \tag{8}$$

Тобто, принцип Лагранжа полягає у зведенні задачі на умовний екстремум (1)-(2) до задачі безумовної оптимізації (8) та у відшуканні і дослідженні її критичних точок.

На основі необхідних і достатніх умов локального мінімуму (максимуму) функції Лагранжа в задачі (8) одержуються необхідні і достатні умови локального екстремуму функції $f(x)$ в задачі (1)-(2). Ці умови задають теореми 1-3.

Теорема 1 (необхідна умова локального екстремуму першого порядку – принцип Лагранжа). *Нехай функції* $f(x)$, $g_i(x)$, $i=\overline{1,m}$, *неперервно-диференційовані в деякому околі точки* $x^*\in R^n$. *Якщо* x^* *– точка локального екстремуму в задачі* (1)-(2), *то існують одночасно не рівні нулю такі числа* λ_0^*, $\lambda_1^*,...,\lambda_m^*$, *що виконується умова стаціонарності по* x *функції Лагранжа* $L_x'(x^*,\lambda_0^*,\lambda^*)=0$, $\lambda^*=(\lambda_1^*,...,\lambda_m^*)$, *тобто*

$$\frac{\partial L(x^*,\lambda_0^*,\lambda^*)}{\partial x_j}=0,\ j=\overline{1,n},\Leftrightarrow \lambda_0^*\frac{\partial f(x^*)}{\partial x_j}+\sum_{i=1}^{m}\lambda_i^*\frac{\partial g_i(x^*)}{\partial x_j}=0,\ j=\overline{1,n}. \tag{9}$$

Якщо при цьому градієнти $g_1'(x^*),...,g_m'(x^*)$ *лінійно незалежні, то* $\lambda_0^*\neq 0$.

Умова (9) означає, що градієнти $f'(x^*)$, $g_1'(x^*),...,g_m'(x^*)$ лінійно залежні. Зокрема, якщо $m=1$, то $f'(x^*)$ і $g_1'(x^*)$ повинні бути колінеарними. Нагадаємо, що ці градієнти, якщо вони відмінні від нуля, направлені ортогонально до ліній рівня $L_{f(x^*)}=\left\{x\in\mathbf{R}^n \,\middle|\, f(x)=f(x^*)\right\}$ та $L_{g_1(x^*)}=\left\{x\in\mathbf{R}^n \,\middle|\, g_1(x)=g_1(x^*)=0\right\}$ відповідно, причому їх колінеарність означає, що розв'язок x^* повинен бути точкою дотику даних ліній.

З теореми 1 випливає, що розв'язками початкової задачі (1)-(2) можуть бути тільки ті точки $x^*\in R^n$, для яких існують такі одночасно не рівні нулю множники λ_0^*, $\lambda_1^*,...,\lambda_m^*$, що точка $(x^*,\lambda_0^*,\lambda^*)\in\mathbf{R}^{n+m+1}$ задовольняє $n+m$ рівнянням (2) і (9).

Довільна точка $x^*\in\mathbf{R}^n$, яка при деяких $\lambda_0^*\in\mathbf{R}^1$ і $\lambda^*\in\mathbf{R}^m$, одночасно не рівних нулю, задовольняє умовам (2) і (9) називається *умовно-стаціонароною* точкою задачі (1)-(2) і *стаціонарною* точкою функції Лагранжа. Тільки у таких точках може досягатись розв'язок цієї задачі. Числа $\lambda_0^*, \lambda_1^*,...,\lambda_m^*$ називаються *множниками Лагранжа*, що відповідають x^*.

Умови (2) та (9) утворюють систему $n+m$ рівнянь з $n+m+1$ невідомими. Легко бачити, якщо $(x^*, \lambda_0^*, \lambda^*)$ – розв'язок системи (2), (9), то $(x^*, \alpha\lambda_0^*, \alpha\lambda^*)$ при довільній константі α, $\alpha \neq 0$, також є її розв'язком, тобто, множники Лагранжа визначаються з системи з точністю до константи α. Тому у випадку, коли $\lambda_0^* \neq 0$, завжди можна вважати, що $\lambda_0^* = 1$. Для цього достатньо поділити всі множники Лагранжа λ_i^*, $i = \overline{1,m}$, на $\alpha \neq 0$. Отже, в системі (2), (9) завжди можна перейти до $n+m$ змінних, розглядаючи два випадки: 1) $\lambda_0^* = 0$ і 2) $\lambda_0^* = 1$.

Довільне додаткове припущення про задачу (1)-(2), яке забезпечує в теоремі 1 випадок $\lambda_0^* \neq 0$, називають *умовою регулярності*, а саму задачу *регулярною* (*невиродженою*). Зокрема, умови лінійної незалежності градієнтів $g_1'(x^*),...,g_m'(x^*)$ або лінійності функцій $g_1(x),...,g_m(x)$ є частинними прикладами умов регулярності. Для регулярної задачі (1)-(2) достатньо розглядати *регулярну функцію Лагранжа*

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

Перевірка умов регулярності, як правило, складна. Тому, зазвичай, систему рівнянь (2), (9) розв'язують, покладаючи в одному випадку $\lambda_0^* = 0$, а в другому – $\lambda_0^* = 1$. Якщо хоча б в одному з цих випадків система (2), (9) $n+m$ рівнянь з $n+m$ невідомими виявиться сумісною, то в результаті її розв'язання знайдемо всі підозрілі на екстремум умовно-стаціонарні точки $x^* \in X$ задачі (1)-(2) і відповідні їм множники Лагранжа $\lambda_0^*, \lambda_1^*,...,\lambda_m^*$.

Необхідна умова локальної оптимальності першого порядку не визначає характеру одержаних умовно-стаціонарних точок. Для цього потрібно провести їх додаткові дослідження на основі умов локальної оптимальності другого порядку.

Позначимо $L''_{xx}(x,\lambda_0,\lambda) = \left(\dfrac{\partial^2 L(x,\lambda_0,\lambda)}{\partial x_k \partial x_j} \right)_{k,j=\overline{1,n}}$ – матрицю других частинних похідних (гессіан) функції Лагранжа по координатам вектора x,

$$L''_{xx}(x,\lambda_0,\lambda) = \lambda_0 f''(x) + \sum_{i=1}^{m} \lambda_i g_i''(x),$$

де $f''(x) = \left(\dfrac{\partial^2 f(x)}{\partial x_k \partial x_j} \right)_{k,j=\overline{1,n}}$, $g_i'' = \left(\dfrac{\partial^2 g_i(x)}{\partial x_k \partial x_j} \right)_{k,j=\overline{1,n}}$, $i = \overline{1,m}$, – матриці других частинних похідних (гессіани) функцій $f(x)$ та $g_1(x),...,g_m(x)$ відповідно.

Теорема 2 (необхідна умова локального екстремуму другого порядку). *Нехай функції* $f(x)$, $g_i(x)$, $i=\overline{1,m}$, *два рази диференційовні в точці* $x^* \in R^n$ *і неперервно диференційовні в деякому околі цієї точки, причому градієнти* $g_1'(x^*),...,g_m'(x^*)$ *лінійно незалежні. Якщо* x^* *– точка локального мінімуму (максимуму) в задачі* (1)-(2), *то*

$$\left\langle h, L''_{xx}(x^*,\lambda_0^*,\lambda^*)\, h \right\rangle \ge 0 \quad \left(\left\langle h, L''_{xx}(x^*,\lambda_0^*,\lambda^*)\, h \right\rangle \le 0 \right) \tag{10}$$

при всіх $\lambda_0^* \in \mathbf{R}^1$, $\lambda^* \in \mathbf{R}^m$, *що задовольняють умову* (9), *і при всіх таких векторах* $h \in \mathbf{R}^n$, *що*

$$\left\langle g_i'(x^*), h \right\rangle = 0, \quad i=\overline{1,m}. \tag{11}$$

Нагадаємо, що умова (10) означає, що квадратична форма $\left\langle h, L''_{xx}(x^*,\lambda_0^*,\lambda^*)\, h \right\rangle$ є невід'ємно (недодатно) визначеною.

Теорема 3 (достатня умова локального екстремуму другого порядку). *Нехай функції* $f(x)$, $g_i(x)$, $i=\overline{1,m}$, *два рази диференційовані в точці* $x^* \in \mathbf{R}^n$, *яка задовольняє умові* (2), *тобто,* x^* *– допустима точка множини* X $(x^* \in X)$. *Якщо при деяких* $\lambda_0^* \in \mathbf{R}^1$, $\lambda^* \in \mathbf{R}^m$, *виконується умова* (9) *і, крім того,*

$$\left\langle h, L''_{xx}(x^*,\lambda_0^*,\lambda^*)\, h \right\rangle > 0 \quad \left(\left\langle h, L''_{xx}(x^*,\lambda_0^*,\lambda^*)\, h \right\rangle < 0 \right) \tag{12}$$

при всіх ненульових векторах $h \in \mathbf{R}^n$, *що задовольняють умову* (11). *Тоді* x^* *– точка строгого локального мінімуму (максимуму) цільової функції* $f(x)$ *в задачі* (1)-(2).

Умова (12) означає, що квадратична форма $\left\langle h, L''_{xx}(x^*,\lambda_0^*,\lambda^*)\, h \right\rangle$ додатньо (від'ємно) визначена.

У найпростіших випадках теореми 1-3 дозволяють розв'язати задачу (1)-(2) у явному вигляді. Сформулюємо ***правило невизначених множників Лагранжа розв'язування задач на умовний екстремум з обмеженнями типу рівностей*:**

1) скласти функцію Лагранжа $L(x,\lambda_0,\lambda)=\lambda_0 f(x)+\sum_{i=1}^{m}\lambda_i g_i(x)$.

Зауважимо, що попередньо задачу необхідно привести до вигляду (1)-(2);

2) виписати необхідну умову $L'_x(x,\lambda_0,\lambda)=0$ локального екстремуму першого порядку та умову $x \in X$ допустимості розв'язків задачі:

$$\begin{cases} \dfrac{\partial L(x.\lambda_0,\lambda)}{\partial x_j}=0, \ j=\overline{1,n}, \\ x \in X. \end{cases} \Leftrightarrow \begin{cases} \lambda_0 \dfrac{\partial f(x)}{\partial x_j}+\sum\limits_{i=1}^{m}\lambda_i \dfrac{\partial g_i(x)}{\partial x_j}=0, \ j=\overline{1,n}, \\ g_i(x)=0, i=\overline{1,m}. \end{cases} \tag{13}$$

Відмітимо, що умова $g_i(x)=0,\ i=\overline{1,m}$ рівносильна умові

$$L'_\lambda(x,\lambda_0,\lambda)=0 \Leftrightarrow \frac{\partial L}{\partial \lambda_i}(x,\lambda_0,\lambda)=0,\ i=\overline{1,m}.$$

3) відшукати умовно-стаціонарні точки задачі $x^* \in X$, які є стаціонарними точками функції Лагранжа $L(x,\lambda_0,\lambda)$, тобто, відшукати розв'язки системи (13), при умові, що не всі множники $\lambda_0,\ \lambda_1,\ldots,\lambda_m$ рівні нулю. При цьому, якщо виконуються умови регулярності задачі, то можна покласти $\lambda_0=1$ або $\lambda_0=\alpha=const\neq 0$. Якщо ж регулярність задачі не встановлена, то потрібно окремо розглянути випадки $\lambda_0=0$ і $\lambda_0\neq 0$. Якщо виявиться, що $\lambda_0\neq 0$, то можна покласти λ_0 рівним одиниці або довільній константі $\alpha\neq 0$;

4) на основі умов локальної оптимальності другого порядку встановити характер умовно стаціонарних точок;

5) відшукати розв'язки задачі серед точок локального екстремуму x^* або довести, що вони не існують;

6) у двовимірному випадку можна розглянути геометричну інтерпретацію задачі.

Описане правило складає основний зміст принципу Лагранжа дослідження задач з обмеженнями шляхом зведення початкової задачі до задачі без обмежень. Цей принцип застосовують не тільки до розв'язання задачі (1)-(2), а і до дослідження багатьох інших екстремальних задач з обмеженнями. При цьому слід мати на увазі, що на відміну від задачі з обмеженнями типу рівностей, в яких число λ_0 можна покладати рівним одиниці або довільній константі $\alpha\neq 0$, в задачах, в яких є нерівності і умови належності точок певним множинам (умови включення), знак λ_0 істотний, а саме: $\lambda_0>0$ в задачах на мінімум і $\lambda_0<0$ в задачах на максимум.

Проілюструємо застосування цього правила на прикладах. Спочатку розв'яжемо задачу, розв'язану у прикладі 1.

Приклад 2. Знайти розв'язки задачі

$$xy^3 \to \max,$$
$$8x+2y=5,$$
$$x,y\in \mathbf{R}.$$

Розв'язання

Дана задача є класичною задачею на умовний екстремум. Представимо її у вигляді задачі (1)-(2):

$$f(x,y)=xy^3 \to \max,$$
$$g(x,y)=8x+2y-5=0,$$
$$x,y\in \mathbf{R}.$$

Для розв'язання цієї задачі застосуємо правило невизначених множників Лагранжа.

1. Функція Лагранжа:
оскільки виконується умова регулярності то можемо записати регулярну функцію Лагранжа

$$L(x,y,\lambda)=f(x,y)+\lambda g(x,y)=xy^3+\lambda(8x+2y-5).$$

2. Необхідна умова локального екстремуму першого порядку (теорема 1):

$$L'_x(x,y,\lambda)=0 \;\Leftrightarrow\; \begin{cases} \dfrac{\partial L(x,y,\lambda)}{\partial x}\equiv y^3+8\lambda=0, \\ \dfrac{\partial L(x,y,\lambda)}{\partial y}\equiv 3xy^2+2\lambda=0; \end{cases}$$

$$x\in X \;\Leftrightarrow\; g(x,y)\equiv 8x+2y-5=0.$$

3. Пошук умовно-стаціонарних точок задачі:
шукаємо умовно-стаціонарні точки задачі, розв'язуючи систему рівнянь:

$$\begin{cases} L'_x(x,y,\lambda)=0, \\ g(x,y)=0. \end{cases} \Leftrightarrow \begin{cases} y^3+8\lambda=0, \\ 3xy^2+2\lambda=0, \\ 8x+2y-5=0. \end{cases} \Rightarrow \begin{cases} \lambda=-\dfrac{1}{8}y^3, \\ x=\dfrac{1}{8}(5-2y), \\ \dfrac{3}{8}(5-2y)y^2-\dfrac{2}{8}y^3=0. \end{cases}$$

Розв'язками системи є:

$$\left(x^{(1)},y^{(1)},\lambda^{(1)}\right)=\left(\frac{5}{8};0;0\right),$$

$$\left(x^{(2)},y^{(2)},\lambda^{(2)}\right)=\left(\frac{5}{8};0;0\right),$$

$$\left(x^{(3)},y^{(3)},\lambda^{(3)}\right)=\left(\frac{5}{32};\frac{15}{8};\frac{-3375}{4096}\right).$$

Отже, точки $z^{(1)}=z^{(2)}=\left(\frac{5}{8};0\right), z^{(3)}=\left(\frac{5}{32};\frac{15}{8}\right)$ – умовно стаціонарні точки задачі, а $\lambda^{(1)}=\lambda^{(2)}=0$ і $\lambda^{(3)}=-\frac{3375}{4096}$ – відповідні їм значення множника Лагранжа λ.

4. Дослідження характеру умовно-стаціонарних точок:

$$\frac{\partial^2 L(x,y,\lambda)}{\partial x^2}=0,\; \frac{\partial^2 L(x,y,\lambda)}{\partial y^2}=6xy,\quad \frac{\partial^2 L(x,y,\lambda)}{\partial x\partial y}=\frac{\partial^2 L(x,y,\lambda)}{\partial y\partial x}=3y^2,$$

$$L''_{xx}=\begin{pmatrix} 0 & 3y^2 \\ 3y^2 & 6xy \end{pmatrix},\quad g'(x)=(8;2).$$

Використовуючи умову (11) встановлюємо вигляд вектора h

$$\left\langle g'(x,y),h\right\rangle = 0 \quad \Leftrightarrow \quad (8;2)\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = 0 \quad \Leftrightarrow \quad 8h_1 + 2h_2 = 0 \quad \Rightarrow \quad h_2 = -4h_1 \ .$$

Отже, $h = (h_1;-4h_1)$.

Дослідимо точки $z^{(1)}$ і $z^{(2)}$:

$$\left\langle h, L''_{xx}\left(z^{(1)},\lambda^{(1)}\right)h\right\rangle = \left\langle h, L''_{xx}\left(z^{(2)},\lambda^{(2)}\right)h\right\rangle = (h_1;-4h_1)\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} h_1 \\ -4h_1 \end{pmatrix} = 0 \ \forall \ h \neq 0,$$

тобто достатня умова оптимальності не дозволяє встановити характер точок $z^{(1)}$ і $z^{(2)}$. Для дослідження цих точок розглянемо точки з їх ε-околу (враховуючи обмеження задачі можемо сказати, що точки з допустимої множини мають вигляд $z = \left(\dfrac{5-2y}{8};y\right)$).

$$\left.\begin{aligned} & f\left(x^{(1)}\right) = f\left(x^{(2)}\right) = f\left(\frac{5}{8};0\right) = 0, \\ & f\left(\frac{5-2\varepsilon}{8};\varepsilon\right) > 0, \\ & f\left(\frac{5-2(-\varepsilon)}{8};-\varepsilon\right) < 0. \end{aligned}\right\} \Rightarrow \text{точки } z^{(1)} \text{ і } z^{(2)} \text{ не є точками екстремуму,}$$

оскільки в околі цієї точки є точки, в яких функція приймає як більші так і менші значення.

Дослідимо точку $z^{(3)}$:

$$\left\langle h, L''_{xx}\left(z^{(3)},\lambda^{(3)}\right)h\right\rangle = (h_1;-4h_1)\begin{pmatrix} 0 & \dfrac{675}{64} \\ \dfrac{675}{64} & \dfrac{225}{128} \end{pmatrix}\begin{pmatrix} h_1 \\ -4h_1 \end{pmatrix} =$$

$$= \left(-\frac{675}{16}h_1;\frac{225}{64}h_1\right)\begin{pmatrix} h_1 \\ -4h_1 \end{pmatrix} = -\frac{225}{8}h_1^2 < 0 \ \ \forall h \neq 0 \Rightarrow z^{(3)} \in loc\max_{z\in X} f(z).$$

5. Дослідження існування глобальних розв'язків:

$$\lim_{\|z\|\to\infty} f(z) = -\infty,$$

тому згідно з наслідком теореми Вейерштрасса про найбільше і найменше значення функції $f(x,y)$ на множині X досягає найбільшого значення, тобто існує розв'язок задачі на максимум. Оскільки в задачі одержана єдина точка локального максимуму функції $f(x,y)$, то розв'язком задачі максимізації є точка $z^{(3)}$, тобто,

$$z^{(3)} \in abs\max_{z\in X} f(z), \ \ f^*_{\max} = f\left(z^{(3)}\right) \approx 1{,}03.$$

6. Графічний аналіз задачі:
множина допустимих розв'язків задачі є пряма. Лінії рівня цільової функції $f(x,y)$ – це сім'я кривих $f(x,y) \equiv xy^3 = \alpha, \alpha \in \mathbf{R}^1$. Глобальний мінімум функції $f(x,y)$ досягається в точці $z^{(3)} = \left(\frac{5}{32}; \frac{15}{8}\right)$, яка є точкою дотику прямої і максимальної лінії рівня $xy^3 = \alpha_{\min}$, $\alpha_{\min} = \frac{16875}{16384} = f^*_{\max}$.

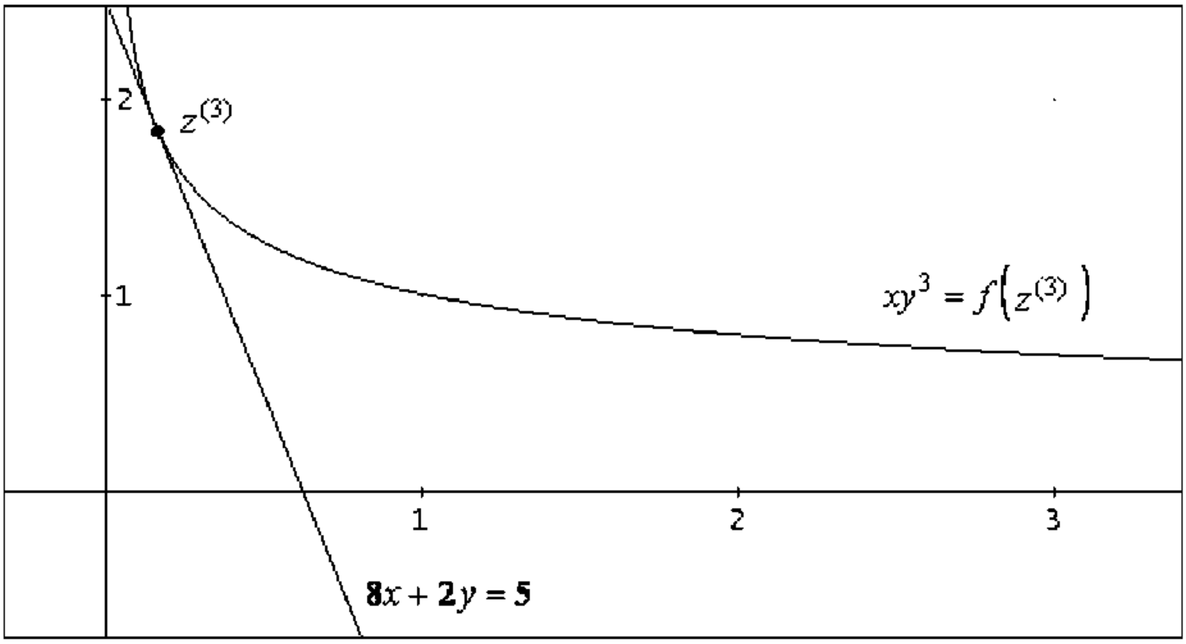

Рис. 3. Графічна інтерпретація задачі

Відповідь: $z^{(3)} \in abs\max\limits_{z \in X} f(z)$, $f^*_{\max} = f\left(z^{(3)}\right) \approx 1{,}03$.

Приклад 3. Знайти розв'язок і зробити графічний аналіз задачі

$$f(x_1, x_2) = 3x_1 + 4x_2 \to extr,$$
$$x_1^2 + x_2^2 = 25,$$
$$x, y \in \mathbf{R}.$$

Розв'язання

Дана задача є двовимірною гладкою класичною задачею на умовний екстремум. Застосовуємо правило розв'язування скінченновимірних гладких задач з обмеженнями типу рівностей.

1. Функція Лагранжа: $L(x, \lambda_0, \lambda_1) = \lambda_0(3x_1 + 4x_2) + \lambda_1(x_1^2 + x_2^2 - 25)$.

2. Необхідна умова локального екстремуму першого порядку (теорема 1):

$$L'_x(x, \lambda_0, \lambda_1) = 0 \iff \begin{cases} \dfrac{\partial L(x, \lambda_0, \lambda_1)}{\partial x_1} \equiv 3\lambda_0 + 2\lambda_1 x_1 = 0 \\ \dfrac{\partial L(x, \lambda_0, \lambda_1)}{\partial x_2} \equiv 4\lambda_0 + 2\lambda_1 x_2 = 0; \end{cases}$$

умова допустимості розв'язків задачі: $x \in X \iff g_1(x) \equiv x_1^2 + x_2^2 - 25 = 0$.

3. Пошук умовно-стаціонарних точок задачі (розглядаємо випадки $\lambda_0 = 0$ і $\lambda_0 \neq 0$):

1) нехай $\lambda_0 = 0 \quad \Rightarrow \quad \lambda_1 \neq 0$ (λ_0 і λ_1 одночасно не можуть приймати нульові значення); розв'язуємо систему рівнянь:

$$\begin{cases} L'_x(x,0,\lambda_1) = 0 \\ g_1(x) = 0 \end{cases} \Leftrightarrow \begin{cases} 3\lambda_0 + 2\lambda_1 x_1 = 0 \\ 4\lambda_0 + 2\lambda_1 x_2 = 0 \\ x_1^2 + x_2^2 - 25 = 0 \end{cases} \Rightarrow \begin{cases} 2\lambda_1 x_1 = 0 \\ 2\lambda_1 x_2 = 0 \\ x_1^2 + x_2^2 - 25 = 0 \end{cases} \Rightarrow (x_1, x_2, \lambda_1) \in \varnothing,$$

тобто, дана система рівнянь несумісна, а отже, у випадку $\lambda_0 = 0$ задача не має розв'язків;

2) нехай $\lambda_0 \neq 0 \Rightarrow \lambda_0 = const \neq 0$; покладаємо $\lambda_0 = 1$; шукаємо умовно-стаціонарні точки задачі, розв'язуючи систему рівнянь:

$$\begin{cases} L'_x(x,1,\lambda_1) = 0 \\ g_1(x) = 0 \end{cases} \Leftrightarrow \begin{cases} 3 + 2\lambda_1 x_1 = 0 \\ 4 + 2\lambda_1 x_2 = 0 \\ x_1^2 + x_2^2 = 25 \end{cases} \Rightarrow \begin{cases} \dfrac{x_1}{x_2} = \dfrac{3}{4} \\ \left(\dfrac{3}{4} x_2\right)^2 + x_2^2 = 25 \end{cases} \Rightarrow$$

$$\left(x_1^{(1)}, x_2^{(1)}, \lambda_1^{(1)}\right) = (-3, -4, 1/2),$$

$$\left(x_1^{(2)}, x_2^{(2)}, \lambda_1^{(2)}\right) = (3, 4, -1/2).$$

Отже, $x^{(1)} = (-3, -4)$, $x^{(2)} = (3, 4)$ – умовно стаціонарні точки задачі, а $\lambda_1^{(1)} = 1/2$ і $\lambda_1^{(2)} = -1/2$ – відповідні їм значення множника Лагранжа λ_1.

4. Дослідження характеру умовно-стаціонарних точок (теореми 2,3):

$$\frac{\partial^2 L(x,1,\lambda_1)}{\partial x_1^2} = \frac{\partial^2 L(x,1,\lambda_1)}{\partial x_2^2} = 2\lambda_1, \quad \frac{\partial^2 L(x,1,\lambda_1)}{\partial x_1 \partial x_2} = \frac{\partial^2 L(x,1,\lambda_1)}{\partial x_2 \partial x_1} = 0,$$

$$L''_{xx}(x,1,\lambda_1) = \begin{pmatrix} 2\lambda_1 & 0 \\ 0 & 2\lambda_1 \end{pmatrix}, \quad g'_1(x) = (2x_1, 2x_2);$$

$$\left\langle h, L''_{xx}(x,1,\lambda_1)\, h \right\rangle = (h_1, h_2) \begin{pmatrix} 2\lambda_1 & 0 \\ 0 & 2\lambda_1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = (2\lambda_1 h_1, 2\lambda_1 h_2) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = 2\lambda_1 (h_1^2 + h_2^2);$$

$$\left\langle g'_1(x), h \right\rangle = 0 \quad \Leftrightarrow \quad (2x_1, 2x_2) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = 0 \quad \Leftrightarrow \quad 2x_1 h_1 + 2x_2 h_2 = 0 \quad \Rightarrow \quad h_2 = -\frac{x_1}{x_2} h_1.$$

Отже, $h = (h_1, h_2) \neq (0,0)$, якщо $h_1 \neq 0$, а якщо $h_1 = 0$, то $h = (h_1, h_2) = (0,0)$;

$$\left\langle h, L''_{xx}\left(x^{(1)}, 1, \lambda_1^{(1)}\right) h \right\rangle = 2\lambda_1^{(1)} \left(h_1^2 + \left(-\frac{x_1^{(1)}}{x_2^{(1)}} h_1 \right)^2 \right) > 0 \quad \forall\ h \neq 0 \Rightarrow x^{(1)} \in loc \min_{x \in X} f(x),$$

$$\left\langle h, L''_{xx}\left(x^{(2)}, 1, \lambda_1^{(2)}\right) h \right\rangle = 2\lambda_1^{(2)} \left(h_1^2 + \left(-\frac{x_1^{(2)}}{x_2^{(2)}} h_1 \right)^2 \right) < 0 \quad \forall\ h \neq 0 \Rightarrow x^{(2)} \in loc \max_{x \in X} f(x),$$

5. Дослідження існування розв'язків:

функція $f(x)$ неперервна на обмеженій і замкненій множині X, тому згідно з теоремою Вейерштрасса про найбільше і найменше значення функції $f(x)$ на множині X досягає найбільшого і найменшого значень, тобто, існують розв'язки задач на максимум і на мінімум. Так як в задачі одержана єдина точка локального мінімуму і єдина точка локального максимуму функції

$f(x)$, то розв'язком задачі мінімізації є точка $x^{(1)}$, а розв'язком задачі максимізації – точка $x^{(2)}$, тобто,

$$x^{(1)} \in abs\min_{x\in X} f(x),\ f^*_{\min} = f\left(x^{(1)}\right) = -25,$$

$$x^{(2)} \in abs\max_{x\in X} f(x),\ f^*_{\max} = f\left(x^{(2)}\right) = 25\ .$$

Відмітимо, що внаслідок існування розв'язків задачі і наявності тільки двох умовно-стаціонарних точок, дослідження їх характеру можна було не проводити. Розв'язки задачі можна було знайти, порівнюючи значення функції $f(x)$ в умовно-стаціонарних точках $x^{(1)}$ і $x^{(2)}$:

$$f\left(x^{(1)}\right) = -9-16 = -25\ <\ f\left(x^{(2)}\right) = 9+16 = 25\quad \Rightarrow$$

$$\Rightarrow\quad x^{(1)} \in abs\min_{x\in X} f(x),\ f^*_{\min} = -25,\quad x^{(2)} \in abs\max_{x\in X} f(x),\ f^*_{\max} = 25\ .$$

6. Графічний аналіз задачі:
множина допустимих розв'язків задачі є колом з центром в точці (0,0) і радіусом $R=5$. Лінії рівня цільової функції $f(x)$ – це сім'я прямих $f(x) \equiv 3x_1 + 4x_2 = \alpha,\ \alpha \in R^1$. Глобальний мінімум функції $f(x)$ досягається в точці $x^{(1)} = (-3,-4)$, яка є точкою дотику до кола лінії мінімального рівня $3x_1 + 4x_2 = \alpha_{\min}$, $\alpha_{\min} = -25 = f^*_{\min}$. Глобальний максимум функції $f(x)$ досягається в точці $x^{(2)} = (3,4)$, яка є точкою дотику до кола лінії максимального рівня $3x_1 + 4x_2 = \alpha_{\max}$, $\alpha_{\max} = +25 = f^*_{\max}$.

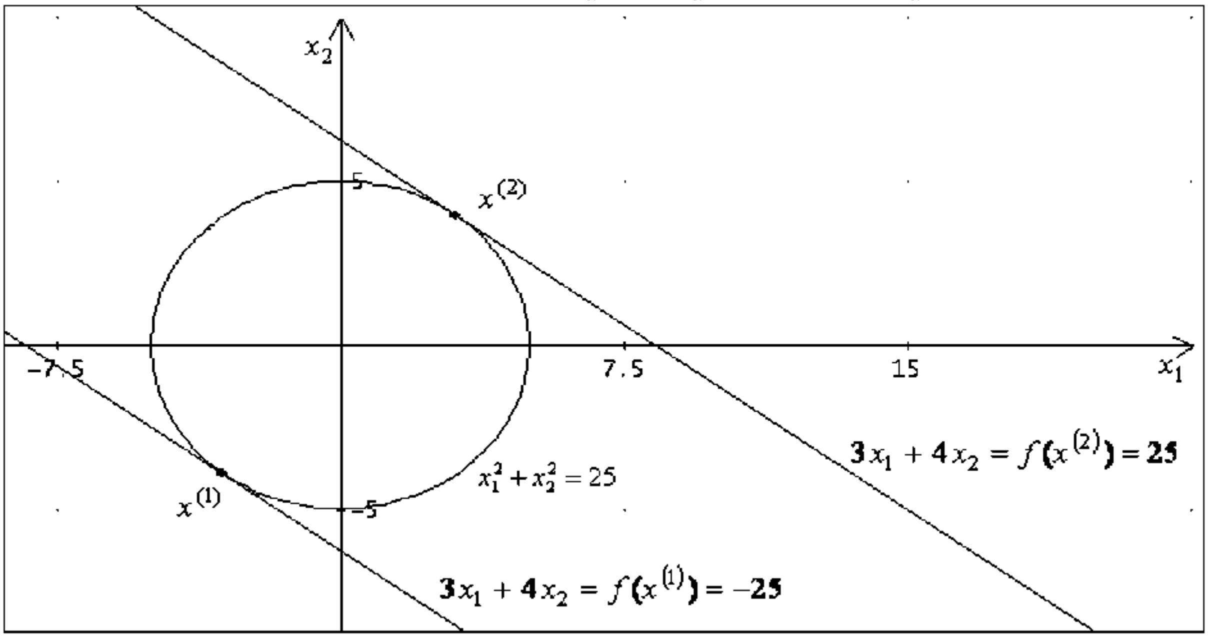

Рис. 4. Графічна інтерпретація задачі

***Відповідь*:** $x^{(1)} \in abs\min_{x\in X} f(x)$, $f^*_{\min} = f\left(x^{(1)}\right) = -25$, $x^{(2)} \in abs\max_{x\in X} f(x)$, $f^*_{\max} = f\left(x^{(2)}\right) = 25$.

2. Задача на умовний екстремум з обмеженнями рівностями і нерівностями

Задачею на умовний екстремум з рівностями і нерівностями називають задачу

$$f(x) \to extr, \tag{14}$$

$$g_i(x) \le 0, i = \overline{1,m}, \tag{15}$$

$$g_{m+k}(x) = 0, k = \overline{1,s}. \tag{16}$$

де функції f та $g_i, i = \overline{1, m+s}$, є диференційованими функціями n дійсних змінних.

Розв'язання даної задачі, як і класичної задачі на умовний екстремум, ґрунтується на використанні функції Лагранжа

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

Теорема 4 (теорема про невизначені множники Лагранжа). *Нехай x^* – точка локального мінімуму задачі* (14)-(16)*, а функції f та $g_i, i = \overline{1, m+s}$, є неперервно-диференційованими в деякому околі цієї точки. Тоді існують одночасно не рівні нулю множники Лагранжа $\lambda_0, \lambda_1, ..., \lambda_{m+s}$ такі, що для функції Лагранжа*

$$L(x, \lambda_0, \lambda_1, ..., \lambda_{m+s}) = \lambda_0 f(x) + \sum_{i=1}^{m+s} \lambda_i g_i(x)$$

виконуються умови:

1) стаціонарності по x

$$L'_x(x^*, \lambda_0, \lambda) = 0 \Leftrightarrow \frac{L(x^*, \lambda_0, \lambda)}{\partial x_j} = 0, j = \overline{1,n};$$

2) доповнюючої нежорсткості

$$\lambda_i g_i(x^*) = 0, i = \overline{1,m};$$

3) невід'ємності

$$\lambda_i \ge 0, i = \overline{0,m}.$$

Таким чином, ***правило невизначених множників Лагранжа розв'язування задач на умовний екстремум з рівностями та нерівностями таке***.

1. Скласти функцію Лагранжа

$$L(x, \lambda_0, \lambda_1, ..., \lambda_{m+s}) = \lambda_0 f(x) + \sum_{i=1}^{m+s} \lambda_i g_i(x)$$

2. Записати необхідні умови оптимальності (теорема 4).

3. Відшукати критичні точки, що задовольняють необхідні умови з множником Лагранжа $\lambda_0 = 0$ та $\lambda \ne 0$.

4. Відшукати розв'язок задачі серед усіх критичних точок або показати, що розв'язків немає.

Приклад 4. Знайти розв'язки задачі

$$f(x)=x_1^2+x_2^2+x_3^2 \to extr,$$
$$2x_1-x_2+x_3\le 5,$$
$$x_1+x_2+x_3=3.$$
$$x_1,x_2,x_3\in \mathbf{R}.$$

Розв'язання

Дана задача є задачею умовної оптимізації з обмеженнями рівностями і нерівностями. Записуємо її у вигляді

$$f(x)=x_1^2+x_2^2+x_3^2 \to extr,$$
$$g_1(x)=2x_1-x_2+x_3-5\le 0,$$
$$g_2(x)=x_1+x_2+x_3-3=0.$$

1. Складаємо функцію Лагранжа

$$L(x,\lambda_0,\lambda)=\lambda_0\left(x_1^2+x_2^2+x_3^2\right)+\lambda_1(2x_1-x_2+x_3-5)+\lambda_2(x_1+x_2+x_3-3).$$

2. Запишемо необхідні умови:

1) стаціонарності

$$\begin{cases}\dfrac{\partial L(x,\lambda_0,\lambda)}{\partial x_1}=0,\\ \dfrac{\partial L(x,\lambda_0,\lambda)}{\partial x_2}=0,\\ \dfrac{\partial L(x,\lambda_0,\lambda)}{\partial x_3}=0;\end{cases} \Leftrightarrow \begin{cases}2\lambda_0x_1+2\lambda_1+\lambda_2=0,\\ 2\lambda_0x_2-\lambda_1+\lambda_2=0,\\ 2\lambda_0x_3+\lambda_1+\lambda_2=0;\end{cases}$$

2) доповнюючої нежорсткості

$$\lambda_1(2x_1-x_2+x_3-5)=0;$$

3) невід'ємності

$$\lambda_0\ge 0,$$
$$\lambda_1\ge 0.$$

3. Знаходимо критичні точки: розв'язком системи

$$\begin{cases}2\lambda_0x_1+2\lambda_1+\lambda_2=0,\\ 2\lambda_0x_2-\lambda_1+\lambda_2=0,\\ 2\lambda_0x_3+\lambda_1+\lambda_2=0,\\ \lambda_1(2x_1-x_2+x_3-5)=0,\\ \lambda_0,\lambda_1\ge 0\end{cases}$$

є точка

$$\left(x^{(1)},\lambda_0^{(1)},\lambda_1^{(1)}\right)=\left(1,1,1,\frac{1}{2},0\right),$$

отже, точка $x^{(1)} = (1,1,1)$ є критичною точкою.

4. Дослідження існування розв'язків:

$$\lim_{\|x\|\to\infty} f(x) = \infty ,$$

тому за наслідком теореми Веєрштрасса про найбільше і найменше значення функції $f(x)$ досягає свого найменшого значення. Оскільки критична точка єдина, то розв'язком задачі мінімізації може бути лише вона.

Відповідь: $x^{(1)} \in \underset{x\in X}{abs\min} f(x)$, $f^*_{\min} = f\left(x^{(1)}\right) = 3$.

Питання для самоконтролю

1.Дайте означення точок локального і глобального екстремуму для задачі умовної оптимізації.

2. Запишіть загальну постановку класичної задачі на умовний екстремум з обмеженнями типу рівність.

3. Запишіть необхідну умову оптимальності для класичної задачі на умовний екстремум.

4. Запишіть достатню умову оптимальності для класичної задачі на умовний екстремум.

5. Запишіть загальну постановку задачі умовної оптимізації з обмеженнями рівностями і нерівностями.

6. Запишіть необхідну умову оптимальності для задачі умовної оптимізації з обмеженнями рівності і нерівності.

Завдання для самостійної роботи

1. Знайти розв'язки і зробити графічний аналіз задачі на умовний екстремум:

$$xy^3 \to extr ,$$
$$ax + by = c .$$

Значення коефіцієнтів a,b,c кожного з варіантів подано у таблиці

№	a	b	c	№	a	b	c
1	2	3	7	**11**	3	1	10
2	1	5	8	**12**	1	5	10
3	3	1	6	**13**	2	7	10
4	2	4	9	**14**	4	4	8
5	4	1	6	**15**	4	2	12
6	2	5	6	**16**	6	2	4
7	1	4	8	**17**	6	2	8
8	3	5	7	**18**	8	2	5
9	4	3	6	**19**	8	2	10
10	1	6	9	**20**	8	3	11

2. Знайти розв'язки і зробити графічний аналіз задачі на умовний екстремум:

$$f(x_1, x_2) = a\,x_1 + b\,x_2 \to extr ,$$
$$x_1^2 + x_2^2 = c^2 .$$

Значення коефіцієнтів a,b,c кожного з варіантів подано у таблиці

№	a	b	c	№	a	b	c	№	a	b	c
1	1	1	$\sqrt{2}$	**14**	-2	-2	$\sqrt{8}$	**27**	-3	3	$\sqrt{18}$
2	-1	-1	$\sqrt{2}$	**15**	-2	2	$\sqrt{8}$	**28**	3	-3	$\sqrt{18}$
3	-1	1	$\sqrt{2}$	**16**	2	-2	$\sqrt{8}$	**29**	-3	-4	5
4	1	-1	$\sqrt{2}$	**17**	2	3	$\sqrt{13}$	**30**	3	-4	5
5	1	2	$\sqrt{5}$	**18**	-2	-3	$\sqrt{13}$	**31**	-3	4	5
6	-1	-2	$\sqrt{5}$	**19**	2	-3	$\sqrt{13}$	**32**	5	12	13
7	-1	2	$\sqrt{5}$	**20**	-2	3	$\sqrt{13}$	**33**	-5	-12	13
8	1	-2	$\sqrt{5}$	**21**	3	2	$\sqrt{13}$	**34**	-5	12	13
9	2	1	$\sqrt{5}$	**22**	-3	-2	$\sqrt{13}$	**35**	5	-12	13
10	-2	-1	$-\sqrt{5}$	**23**	-3	2	$\sqrt{13}$	**36**	12	5	13
11	-2	1	$\sqrt{5}$	**24**	3	-2	$\sqrt{13}$	**37**	-12	-5	13
12	2	-1	$\sqrt{5}$	**25**	3	3	$\sqrt{18}$	**38**	-12	5	13
13	2	2	$\sqrt{8}$	**26**	-3	-3	$\sqrt{18}$	**39**	12	-5	13

3. Знайти розв’язки і зробити графічний аналіз задачі на умовний екстремум:

$$f(x_1,x_2)=a\,x_1^2+2x_1x_2+b\,x_2^2\to extr,$$
$$4x_1^2+cx_2^2=9.$$

Значення коефіцієнтів a,b,c кожного з варіантів подано у таблиці

№	a	b	c	№	a	b	c
1	1	1	3	**11**	5	8	6
2	3	4	5	**12**	5	3	2
3	1	1	2	**13**	9	3	1
4	1	2	6	**14**	9	7	3
5	1	1	1	**15**	3	6	7
6	3	2	2	**16**	7	4	2
7	6	4	3	**17**	13	4	1
8	3	5	6	**18**	7	11	6
9	7	9	5	**19**	9	5	2
10	5	2	1	**20**	7	13	7

4. Знайти розв’язки задачі на умовний екстремум:

$$x_1^2+x_2^2+x_3^2\to extr,$$
$$-2x_1+4x_2-x_3\le -1,$$
$$-3x_1+2x_2+ax_3=b,$$

Значення коефіцієнтів a,b кожного з варіантів подано у таблиці

№	a	b	№	a	b	№	a	b	№	a	b
1	2	1/2	**11**	3	3/2	**21**	5	1/2	**31**	6	3
2	2	1/3	**12**	3	4/3	**22**	5	1/4	**32**	6	4
3	3	1	**13**	4	1/2	**23**	5	3/4	**33**	6	5
4	3	2	**14**	4	3/4	**24**	5	3/2	**34**	6	5/2
5	4	1	**15**	4	3/5	**25**	6	1/2	**35**	6	5/3
6	4	2	**16**	**4**	3/2	**26**	6	1	**36**	6	5/4
7	4	3	**17**	5	1	**27**	6	3/4	**37**	8	2
8	2	2/3	**18**	5	2	**28**	6	3/2	**38**	8	3
9	2	4/3	**19**	5	3	**29**	6	4/3	**39**	8	4
10	3	1/2	**20**	5	4	**30**	6	2	**40**	2	1

3. Опукле програмування. Гладка основна задача мінімізації

Опуклим програмуванням називається розділ математики, в якому досліджуються задачі мінімізації опуклих функцій $f(x)$ на опуклих множинах X скінченновимірного простору $\mathbf{R}^n$, тобто, задачі

$$f(x) \to \min, \quad x \in X \subset \mathbf{R}^n \tag{17}$$

де $f(x)$ – опукла функція, визначена на опуклій множині $X \subset \mathbf{R}^n$. Такі задачі називають *опуклими задачами оптимізації*.

Нагадаємо основні поняття і факти, пов'язані з опуклими множинами і функціями.

Означення. Множина $X \subset \mathbf{R}^n$ називається *опуклою*, якщо $x(\alpha) = \alpha x^1 + (1-\alpha)x^2 \in X$ при всіх $x^1, x^2 \in X$, $\alpha \in [0,1]$. Іншими словами, множина $X \subset \mathbf{R}^n$ *опукла*, якщо разом з довільними двома своїми точками x^1 і x^2 вона містить і весь відрізок, що їх з'єднує, тобто, множину $[x^1, x^2] = \left\{ x \in \mathbf{R}^n \,\middle|\, x(\alpha) = \alpha x^1 + (1-\alpha)x^2,\ 0 \le \alpha \le 1 \right\}$.

Означення. Функція $f(x)$, визначена на опуклій множині $X \subset \mathbf{R}^n$, називається *опуклою* на X, якщо

$$f\left(\alpha x^1 + (1-\alpha)x^2\right) \le \alpha f(x^1) + (1-\alpha) f(x^2) \tag{18}$$

при всіх $x^1, x^2 \in X$, $\alpha \in [0,1]$. Якщо при всіх $x^1, x^2 \in X$, $x^1 \ne x^2$, і $\alpha \in (0,1)$ нерівність (18) виконується як строга, то функція $f(x)$ називається *строго опуклою* на X.

Якщо функція $f(x)$, $x \in \mathbf{R}^n$, два рази неперервно-диференційовна, то вона опукла тоді і тільки тоді, коли матриця $f''(x) = \left(\dfrac{\partial^2 f(x)}{\partial x_i \, \partial x_j} \right)_{i,j=\overline{1,n}}$, $x \in \mathbf{R}^n$, її других частинних похідних невід'ємно визначена.

В опуклій задачі оптимізації поняття локального і глобального розв'язків не розрізняються: довільний її локальний розв'язок є також глобальним. Тобто, можна говорити просто про розв'язок опуклої задачі.

Множина $X^* = \left\{ x^* \in \mathbf{R}^n \middle| f(x^*) = \min\limits_{x \in X} f(x) \right\}$ розв'язків опуклої задачі, якщо вона не порожня, є опуклою множиною. Якщо при цьому функція $f(x)$ строго опукла на X, то розв'язок задачі єдиний, тобто, X^* складається з однієї точки.

Основною задачею опуклого програмування прийнято називати задачу

$$f(x) \to \min, \quad g(x) \le 0, \; x \in Q, \tag{19}$$

в якій Q – опукла множина, $g(x) = (g_1(x), \ldots, g_m(x))$, $f(x)$ і компоненти $g_i(x)$, $i = \overline{1,m}$, вектор-функції $g(x)$ – опуклі функції.

Із властивостей опуклих множин і функцій випливає, що множини

$$\left\{ x \in \mathbf{R}^n \middle| \; g_i(x) \le 0 \right\}, i = \overline{1,m},$$

є опуклими, а множина

$$X = \left\{ x \in \mathbf{R}^n \middle| \; g(x) \le 0, \; x \in Q \right\}$$

допустимих розв'язків задачі (19) опукла як їх перетин з опуклою множиною Q. Важливим є випадок, коли множина X регулярна.

Означення. Говорять, що множина X допустимих розв'язків основної задачі опуклого програмування (19) *регулярна* (задовольняє умові Слейтера), якщо існує допустимий розв'язок $\tilde{x} \in X$, для якого виконується нерівність $g(\tilde{x}) < 0$. При цьому і саму задачу називають *регулярною.*

Критерій оптимальності для задачі (19) задає теорема Куна-Таккера. Він формулюється в термінах сідлової точки функції Лагранжа, яка для регулярної задачі (19) має вигляд

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x), \tag{20}$$

де $x \in Q$, $\lambda \ge 0$, $\lambda \in \mathbf{R}^m$.

Означення. Точка (x^*, λ^*), $x^* \in Q$, $\lambda^* \ge 0$, $\lambda^* \in \mathbf{R}^m$, називається *сідловою точкою* функції Лагранжа (20), якщо для всіх $x \in Q$, $\lambda \ge 0$, $\lambda \in \mathbf{R}^m$, виконується нерівність

$$L(x^*, \lambda) \le L(x^*, \lambda^*) \le L(x, \lambda^*) .$$

Теорема 5(Куна-Таккера). *Точка* x^* *є розв'язком основної задачі опуклого програмування* (19) *з регулярною множиною* X *допустимих розв'зків тоді і тільки тоді, коли існує невід'ємний вектор* $\lambda^* \in \mathbf{R}^m$ *такий, що* $(x^*, \lambda^*) \in \mathbf{R}^{n+m}$ *– сідлова точка функції Лагранжа* (20). *При цьому виконується умова доповнюючої нежорсткості*

$$\left\langle g(x^*), \lambda^* \right\rangle = 0 \quad \Leftrightarrow \quad \lambda_i^* g_i(x^*) = 0, \quad i = \overline{1,m} \; .$$

Вектор λ^* називається вектором Лагранжа, що відповідає оптимальному розв'язку x^*.

Згідно даної теореми для знаходження оптимального розв'язку x^* регулярної задачі (19) достатньо знайти сідлову точку $(x^*, \lambda^*) \in \mathbf{R}^{n+m}$ функції Лагранжа (20). Однак не для всякої задачі (19) її функція Лагранжа має сідлову точку.

Для *гладкої регулярної основної задачі опуклого програмування*, під якою розуміють регулярну задачу (19) з неперервно-диференційованими опуклими функціями $f(x)$ та $g_i(x)$, $i = \overline{1,m}$, і множиною $Q = \left\{ x \in \mathbf{R}^n \,\middle|\, x \ge 0 \right\}$, необхідні і достатні умови існування сідлової точки її функції Лагранжа $L(x,\lambda)$ мають вигляд:

$$x_j^* \ge 0\,, \quad \frac{\partial L(x^*, \lambda^*)}{\partial x_j} \ge 0\,, \quad x_j^* \frac{\partial L(x^*, \lambda^*)}{\partial x_j} = 0\,, \quad j = \overline{1,n}\,, \tag{21}$$

$$\lambda_i^* \ge 0\,, \quad \frac{\partial L(x^*, \lambda^*)}{\partial \lambda_i} \le 0\,, \quad \lambda_i^* \frac{\partial L(x^*, \lambda^*)}{\partial \lambda_i} = 0\,, \quad i = \overline{1,m}\,. \tag{22}$$

У найпростіших випадках теорема Куна-Таккера разом з цими умовами дозволяє знайти розв'язок x^* в явному вигляді. Сформулюємо ***правило розв'язування гладкої регулярної основної задачі опуклого програмування*:**

1) скласти функцію Лагранжа $L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$. Відзначимо, що попередньо задачу необхідно привести до вигляду (19);

2) виписати систему умов (21)-(22), які характеризують сідлову точку $(x^*, \lambda^*) \in \mathbf{R}^{n+m}$ побудованої функції Лагранжа $L(x,\lambda)$;

3) знайти розв'язок (x^*, λ^*) одержаної системи. Як правило, система (21)-(22) розв'язується складно. Для спрощення відшукання її розв'язку доцільно вважати, виходячи з логічного аналізу конкретної системи умов (21)-(22) або на основі геометричної інтерпретації і властивостей конкретної задачі, що всі або тільки певна частина координат x_j^*, $j = \overline{1,n}$, λ_i^*, $i = \overline{1,m}$, приймають додатні значення, а решта рівні 0. Якщо при деякому такому припущенні знайдемо розв'язок (x^*, λ^*) одержаної системи, то він є сідловою точкою функції Лагранжа, а його компонента x^* – розв'язком задачі. Якщо ж при зробленому припущенні виявиться, що система умов (21)-(22) несумісна, то слід розглядати всі інші можливі комбінації додатних і нульових значень вказаних змінних, поки або не буде знайдена сідлова точка (x^*, λ^*), або не буде встановлено, що задача нерозв'язна;

4) у двовимірному випадку можна розглянути геометричну інтерпретацію задачі.

Приклад 4. Розв'язати і зробити графічний аналіз екстремальної задачі:

$$f(x_1,x_2) = (x_1-6)^2 + (x_2-6)^2 \to \min,$$

$$x_1 + x_2 \le 6, \quad x_1 \ge 0,\ x_2 \ge 0 .$$

Розв'язання. Приводимо задачу до вигляду (19): $g_1(x) \equiv x_1 + x_2 - 6 \le 0$, причому в допустимій точці $\tilde{x} = (0,0)$ $g_1(\tilde{x}) \equiv -6 < 0$, $Q = \left\{x \in \mathbf{R}^2 \middle| x_1 \ge 0,\ x_2 \ge 0\right\}$, $f(x)$ та $g_1(x)$ – неперервно диференційовні опуклі функції, множина $X = \left\{x \in \mathbf{R}^2 \middle| g_1(x) \le 0,\ x \in Q\right\}$ допустимих розв'язків опукла і регулярна, а отже, дана задача – двовимірна гладка регулярна основна задача опуклого програмування. Застосовуємо правило розв'язування таких задач.

1) Функція Лагранжа: $L(x, \lambda_1) = (x_1-6)^2 + (x_2-6)^2 + \lambda_1(x_1 + x_2 - 6)$.

2) Умови (21)-(22):

$$x_1^* \ge 0,\ x_2^* \ge 0,$$

$$\frac{\partial L(x^*,\lambda^*)}{\partial x_1} \equiv 2(x_1^*-6) + \lambda_1^* \ge 0,$$

$$\frac{\partial L(x^*,\lambda^*)}{\partial x_2} \equiv 2(x_2^*-6) + \lambda_1^* \ge 0,$$

$$x_1^* \frac{\partial L(x^*,\lambda^*)}{\partial x_1} = 0 \Leftrightarrow \text{ або } x_1^* = 0, \text{ або } \frac{\partial L(x^*,\lambda^*)}{\partial x_1} \equiv 2(x_1^*-6) + \lambda_1^* = 0,$$

$$x_2^* \frac{\partial L(x^*,\lambda^*)}{\partial x_2} = 0 \Leftrightarrow \text{ або } x_2^* = 0, \text{ або } \frac{\partial L(x^*,\lambda^*)}{\partial x_2} \equiv 2(x_2^*-6) + \lambda_1^* = 0,$$

$$\lambda_1^* \ge 0, \quad \frac{\partial L(x^*,\lambda^*)}{\partial \lambda_1} \equiv x_1^* + x_2^* - 6 \le 0,$$

$$\lambda_1^* \frac{\partial L(x^*,\lambda^*)}{\partial \lambda_1} = 0 \Leftrightarrow \text{ або } \lambda_1^* = 0, \text{ або } \frac{\partial L(x^*,\lambda^*)}{\partial \lambda_1} \equiv x_1^* + x_2^* - 6 = 0.$$

3) Знаходження розв'язку (x^*, λ^*) одержаної системи умов: на основі логічного аналізу умов та геометричної інтерпретації задачі можна припустити, що $x_1^* > 0,\ x_2^* > 0,\ \lambda_1^* > 0$. При даному припущенні розв'язок (x^*, λ^*) знаходимо з системи:

$$\begin{cases} 2(x_1^* - 6) + \lambda_1^* = 0, \\ 2(x_2^* - 6) + \lambda_1^* = 0, \\ x_1^* + x_2^* - 6 = 0, \end{cases} \Rightarrow \begin{cases} x_1^* = (12 - \lambda_1^*)/2, \\ x_2^* = (12 - \lambda_1^*)/2, \\ (12 - \lambda_1^*)/2 + (12 - \lambda_1^*)/2 = 0, \end{cases} \Rightarrow$$

$\Rightarrow$ $(x^*, \lambda^*) = (x_1^*, x_2^*, \lambda_1^*) = (3,3,6)$ – сідлова точка функції Лагранжа $L(x, \lambda)$, $x^* = (x_1^*, x_2^*) = (3,3)$ – розв'язок задачі, $f^* = f(x^*) = 18$ – її значення.

4) Графічний аналіз задачі:

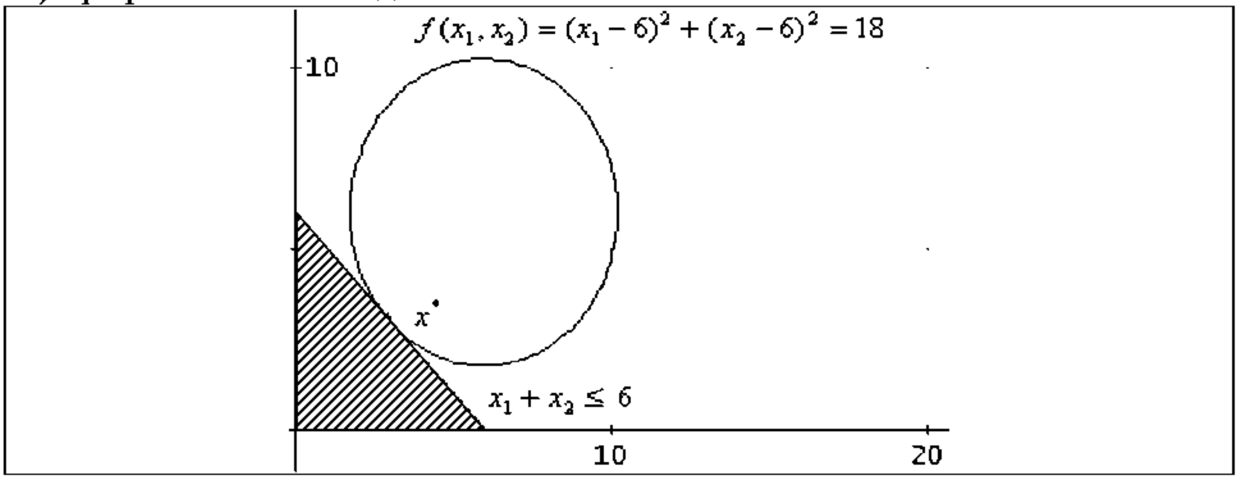

Рис. 5

множина допустимих розв'язків задачі – площина трикутника, розташованого в першому ортанті. Лінії рівня цільової функції $f(x)$ – кола $f(x_1, x_2) \equiv (x_1 - 6)^2 + (x_2 - 6)^2 = \alpha$, $\alpha > 0$, з центром в точці $(6,6)$ і радіусом $\sqrt{\alpha}$. Глобальний мінімум функції $f(x)$ досягається в точці $x^* = (x_1^*, x_2^*) = (3,3)$, яка є точкою дотику прямої $x_1 + x_2 = 6$ і кола $(x_1 - 6)^2 + (x_2 - 6)^2 = \alpha_{min}$, $\alpha_{min} = \sqrt{18} = f_{min}^*$.

Відповідь: $x^* = (3,3) \in abs \min\limits_{x \in X} f(x)$, $f_{min}^* = f(x^*) = 18$.

Питання для самоконтролю

1. Дайте означення опуклої множини.

2. Дайте означення опуклої функції.

3. Запишіть постановку задачі опуклого програмування.

4. Запишіть постановку основної задачі опуклого програмування.

5. Запишіть функцію Лагранжа для основної задачі опуклого програмування.

6. Дайте означення сідлової точки функції Лагранжа для основної задачі опуклого програмування.

7. Запишіть умови Куна-Таккера для гладкої основної задачі опуклого програмування.

Завдання для самостійної роботи

1. Розв'язати і зробити графічний аналіз екстремальної задачі:

$$f(x_1,x_2)=(x_1-a)^2+(x_2-a)^2\to\min,$$
$$x_1+x_2\le b,$$
$$x_1\ge 0,\ x_2\ge 0.$$

Значення коефіцієнтів a,b кожного з варіантів подано у таблиці

№	a	b	№	a	b	№	a	b	№	a	b
1	3	4	**11**	5	8	**21**	10	10	**31**	11	12
2	5	4	**12**	6	8	**22**	2	3	**32**	12	12
3	4	4	**13**	7	8	**23**	3	3	**33**	3	5
4	6	4	**14**	8	8	**24**	4	3	**34**	4	5
5	8	4	**15**	10	8	**25**	5	3	**35**	5	5
6	7	4	**16**	11	8	**26**	6	3	**36**	6	5
7	4	6	**17**	6	10	**27**	7	12	**37**	7	5
8	5	6	**18**	7	10	**28**	8	12	**38**	8	5
9	7	6	**19**	8	10	**29**	9	12	**39**	9	5
10	8	6	**20**	9	10	**30**	10	12	**40**	10	5

2. Знайти розв'язок задачі і дати її графічний аналіз:

$$f(x)=ax_1+bx_2\to\min,$$
$$(x_1-4)^2+(x_2-4)^2\le R,$$
$$x_1\ge 0, x_2\ge 0.$$

Значення коефіцієнтів a,b,R кожного з варіантів подано у таблиці

№	a	b	R	№	a	b	R	№	a	b	R
1	1	1	1	**11**	2	-1	5	**21**	3	1	10
2	-1	1	1	**12**	-2	-1	5	**22**	3	-1	10
3	1	-1	1	**13**	2	2	8	**23**	-3	-1	10
4	-1	-1	1	**14**	-2	2	8	**24**	-3	1	10
5	1	2	5	**15**	2	-2	8	**25**	3	2	13
6	-1	2	5	**16**	-2	-2	8	**26**	-3	2	13
7	1	-2	5	**17**	1	3	10	**27**	3	-2	13
8	-1	-2	5	**18**	-1	3	10	**28**	-3	-2	13
9	2	1	5	**19**	1	-3	10	**29**	2	3	13
10	-2	1	5	**20**	-1	-3	10	**30**	-2	3	13

Список використаної літератури

1. Алексеев В.М., Галеев Э.М., Тихомиров В.М. Сборник задач по оптимизации: Теория. Примеры. Задачи. – М: Наука, Глав. ред. физ.-мат. л-ры, 1984. – 288 с.
2. Галеев Э.М., Тихомиров В.М. Краткий курс теории экстремальных задач. – М.: Изд-во МГУ, 1989. – 204 с.
3. Ларин Р.М. Методы оптимизации. Примеры и задачи: Учеб. пособие/Р.М.Ларин, А.В.Плясунов, А.В.Пяткин. – Новосибирск: Новосиб. ун-т., 2003. – 115 с.
4. Моклячук М.П. Варіаційне числення. Екстремальні задачі. – Київ: 2003. – 380 с.
5. Сборник задач по математике для ВТУЗов. Ч.4. Методы оптимизации. Уравнения в частных производных. Интегральные уравнения: Учебное пособие /под ред. А.В. Ефимова – М.: Наука. Глав. ред. физ.-мат. л-ры, 1990. – 304 с.
6. Сухарев А.Г., Тимохов А.В., Федоров В.В. Курс методов оптимизации.– М.: Наука, Глав. ред. физ.-мат. л-ры, 1986. – 328 с.

ЗМІСТ

Відповідальний за випуск: завідувач кафедрою системного аналізу і теорії оптимізації к. ф.-м. н., доц. Кузка О.І.

Автори: к. ф.-м. н., доц. Гренджа В.І.,
к. ф.-м. н., доц. Брила А.Ю.,
викл. Ломага М.М.

Рецензенти: к.ф.-м.н., доц. Погоріляк О.О.,
к.ф.-м.н., доц.Юрченко Н.В.

МЕТОДИЧНІ ВКАЗІВКИ
до практичних робіт з курсу
«МЕТОДИ ОПТИМІЗАЦІЇ»

Частина І

СКІНЧЕННОВИМІРНІ ЗАДАЧІ УМОВНОЇ ОПТИМІЗАЦІЇ