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

Постановка задачі:

$$f(x) \to min \tag{1}$$

за умови

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

Функції $f(x), g_1(x), \ldots, g_s(x)$ визначені та диференційовні на множині X і $s < n$.

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

Алгоритм методу множників Лагранжа для задачі умовної оптимізації з обмеженнями у формі рівностей

Крок 1. Будуємо функцію Лагранжа

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

та складаємо систему

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

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

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

Приклад. Знайти екстремуми функції $F(x,y) = x^2 + 2y^2$ на множині, яка задана рівнянням $x^2 + y^2 - 2y = 0$.

Розв'язання. Для розв’язання задачі застосуємо метод множників Лагранжа для задачі умовної оптимізації з обмеженнями у формі рівностей.

Запишемо задачу в стандартному вигляді

$$x^2 + 2y^2 \to extr,$$

$$x^2 + y^2 - 2y = 0.$$

Складемо для неї функцію Лагранжа

$$L(x,y,\lambda) = x^2 + 2y^2 + \lambda\left(x^2 + y^2 - 2y\right)$$

і обчислимо частинні похідні

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

$$\frac{\partial L}{\partial y}(x,y,\lambda)=4y+2\lambda(y-1),$$

$$\frac{\partial L}{\partial \lambda}(x,y,\lambda)=x^2+y^2-2y.$$

Після цього потрібно розв'язати систему нелінійних алгебричних рівнянь:

$$\begin{cases} 2x(1+\lambda)=0 \\ (2+\lambda)y=\lambda \\ x^2+y^2-2y=0 \end{cases}.$$

<u>Перший випадок</u>. Нехай з першого рівняння системи $x=0$. Маємо

$$y=\frac{\lambda}{2+\lambda},\ y(y-2)=0.$$

Повинна виконуватися умова $\lambda\neq-2$. Тоді

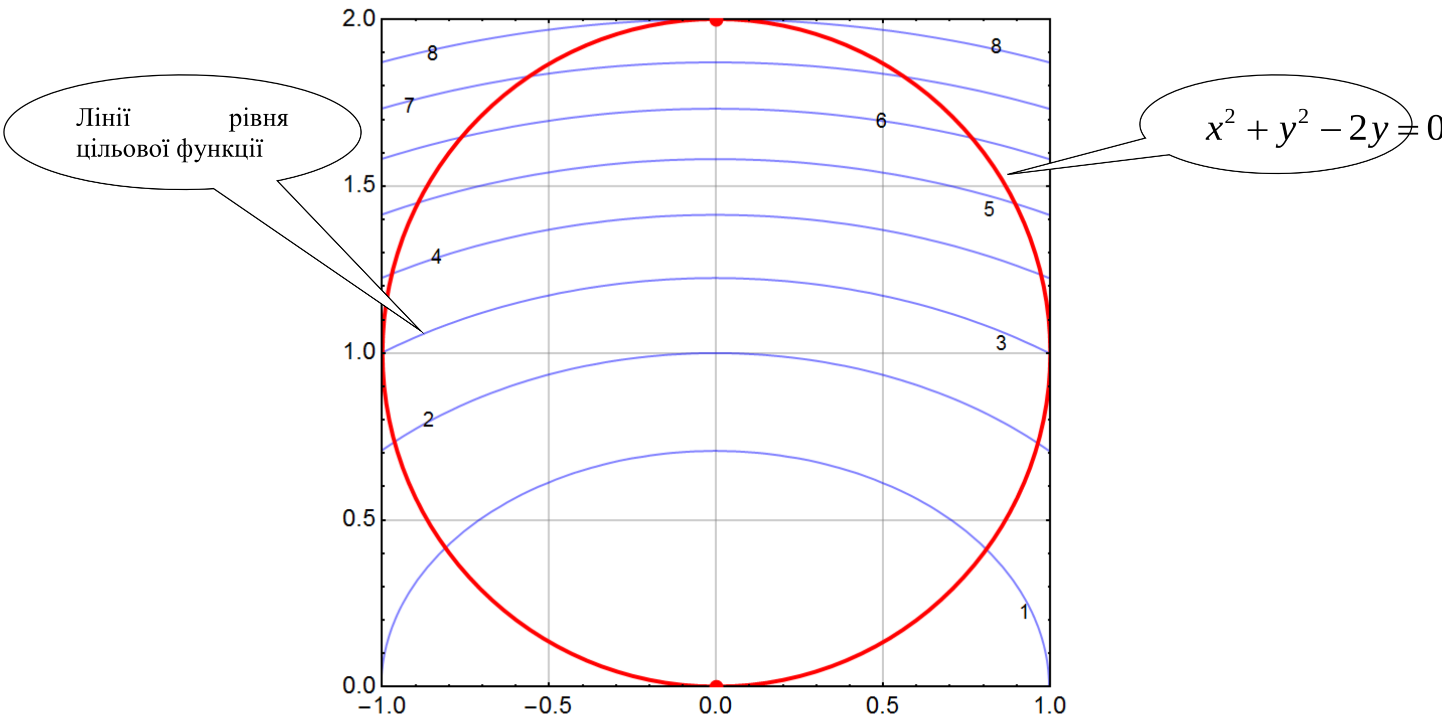
а) $y=0$, $\lambda=0$. Отже, стаціонарна точка є $M_1(0,0)$, $F(0,0)=0$.

б) $y=2$, $\lambda=-4$. Отже, стаціонарна точка є $M_2(0,2)$, $F(0,2)=8$.

<u>Другий випадок</u>. Нехай за першим рівнянням системи $1+\lambda=0$, тобто $\lambda=-1$. Маємо $y=-1$, $x^2+3=0$. Останнє рівняння дійсних коренів не має. Нових стаціонарних точок не знайдено.

За умовою задачі екстремуми цільової функції потрібно знайти на множині, що задається рівнянням $x^2+y^2-2y=0$, тобто на колі. Оскільки коло є компактна множина, а цільова функція є неперервна, то за теоремою Вейерштрасса цільова функція досягає своїх максимуму і мінімуму.

Теорема Вейерштраса. Нехай X – замкнена, обмежена множина в E^n, $f(x)$ – неперервна функція на множині X. Тоді існує точка глобального мінімуму функції $f(x)$ на множині X.



Для дослідження отриманих стаціонарних точок застосуємо критерій Сильвестра. Для функції $F(x,y)$

$$F'(x,y)=\begin{pmatrix}2x_1\\4x_2\end{pmatrix},\ F''(x,y)=\begin{pmatrix}2&0\\0&4\end{pmatrix},\ \Delta_1=2>0,\ \Delta_2=8>0.$$

Максимум цільової функції досягається в точці $(x,y)=(0,2)$ і дорівнює він 8, а мінімум досягається в точці $(x,y)=(0,0)$ і дорівнює 0.

Деякі теоретичні відомості з опуклого аналізу

Непорожня множина $X\subset E^n$ називається **опуклою**, якщо

$$\forall y,z\in X \text{ точки } x_\lambda=\lambda\, y+(1-\lambda)\,z\in X \text{ для всіх } \lambda,\quad 0\le\lambda\le1,$$

тобто множина X разом з будь-якими своїми точками y,z містить відрізок, який їх з’єднує.

Функція $f(x)$, яка визначена на опуклій множині $X\subset E^n$ називається ***опуклою,*** якщо

$$f\left(\lambda\, x_1+(1-\lambda)\,x_2\right)\le\lambda\, f(x_1)+(1-\lambda)\,f(x_2) \qquad (1)$$

для усіх $x_1,x_2\in X$ та $\lambda\in[0,1]$.

Геометричний сенс нерівності (1): функція $f(x)$ опукла на опуклій множині X, якщо для будь-яких точок $x_1,x_2\in X$ графік функції лежить не вище хорди, що з'єднує точки $(x_1,f(x_1))$ та $(x_2,f(x_2))$.

Приклад. З'ясуємо, за яких значень параметрів α та β функція є опуклою

$$f(x_1,x_2)=5x_1^2+\beta x_2^2+\alpha x_1+3x_2.$$

Складемо матрицю других похідних цієї функції та з умов її додатної визначеності отримаємо обмеження на параметри.

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 10x_1 + \alpha,$$

$$\frac{\partial f}{\partial x_2}(x_1, x_2) = 2\beta x_2 + 3,$$

$$\frac{\partial^2 f}{\partial x_1^2} = 10, \ \frac{\partial^2 f}{\partial x_2^2} = 2\beta, \ \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0, \ \frac{\partial^2 f}{\partial x_2 \partial x_1} = 0.$$

$$f''(x) = \begin{pmatrix} 10 & 0 \\ 0 & 2\beta \end{pmatrix}.$$

Матриця других похідних буде додатно визначеною, якщо $\beta > 0$ при будь-якому α.

Отже, функція є опуклою, якщо $\beta > 0$, α – будь-яке.

Приклад. З'ясуємо, за яких значень параметру α функція є опуклою

$$f(x_1, x_2) = \alpha x_1^2 x_2^2 + \left(x_1^2 + x_2^2\right)^2.$$

Складемо матрицю других похідних цієї функції та з умов її додатної визначеності отримаємо обмеження на параметри.

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 2(2+\alpha) x_1 x_2^2 + 4x_1^3,$$

$$\frac{\partial f}{\partial x_2}(x_1, x_2) = 2(2+\alpha) x_1^2 x_2 + 4x_2^3,$$

$$\frac{\partial^2 f}{\partial x_1^2} = 2(2+\alpha) x_2^2 + 12x_1^2, \ \frac{\partial^2 f}{\partial x_2^2} = 2(2+\alpha) x_1^2 + 12x_2^2,$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = 4(2+\alpha) x_1 x_2.$$

$$f''(x) = \begin{pmatrix} 2(2+\alpha) x_2^2 + 12x_1^2 & 4(2+\alpha) x_1 x_2 \\ 4(2+\alpha) x_1 x_2 & 2(2+\alpha) x_1^2 + 12x_2^2 \end{pmatrix}.$$

Функція буде опуклою, якщо

$$(2+\alpha) x_2^2 + 6x_1^2 > 0,$$

$$\left((2+\alpha) x_2^2 + 6x_1^2\right)\left((2+\alpha) x_1^2 + 6x_2^2\right) - 4(2+\alpha)^2 x_1^2 x_2^2 > 0.$$

З першої нерівності випливає, що $\alpha \geq -2$.

Наступна нерівність дає, поклавши $2 + \alpha = t$

$$2t\left(x_1^4 + x_2^4\right) + 12x_1^2 x_2^2 - t^2 x_1^2 x_2^2 > 0.$$

Якщо $t = 0$, то нерівність виконується.

При $t > 0$ розділимо нерівність на $2t$:

$$x_1^4 + x_2^4 - \frac{t^2 - 12}{2t} x_1^2 x_2^2 > 0.$$

Тоді нерівність виконуватиметься для будь-яких x_1 та x_2 при $\frac{t^2 - 12}{2t} \le 2$.

$$\frac{t^2 - 12}{2t} \le 2, \; t^2 - 4t - 12 \le 0, \; t^2 - 4t - 12 = 0, \; t_1 = -2, \; t_2 = 6.$$

Звідси одразу отримуємо, що $0 < t \le 6$.

Отже, $-2 \le \alpha \le 4$.