УМОВНІ ТА АБСОЛЮТНІ ЕКСТРЕМУМИ

Доведемо спочатку допоміжні теореми, що мають велике значення в аналізі й застосуваннях. Перша з них стосується неявної векторної функції $(y_1, y_2, ..., y_n) = f(x_1, ..., x_m)$, яка визначається з векторного рівняння (фактично системи n рівнянь) $F(x_1, ..., x_m, y_1, ..., y_n) = 0$.

Наприклад, потрібно розв'язати систему

$$\begin{cases} x_1^2 + x_2^2 = e^{x_1+y_1+y_2} \\ y_1^2 - y_2^2 = \sin(x_1 - y_1) \end{cases}$$

в околі точки $(x_1, x_2) = (0, 1)$ при значеннях $(y_1, y_2) = (0, 0)$. Систему можна записати, як

$$F(x_1, x_2, y_1, y_2) = 0,$$

де

$$F(x_1, x_2, y_1, y_2) = \begin{pmatrix} x_1^2 + x_2^2 - e^{x_1+y_1+y_2} \\ y_1^2 - y_2^2 - \sin(x_1 - y_1) \end{pmatrix}.$$

Загальний результат щодо можливості розв'язання подібних систем містить така теорема.

Теорема 4. **(Про неявну функцію).** Нехай $F : A \to \mathbb{R}^n$, $A \subset \mathbb{R}^{m+n}$ – відкрита множина, $(x^0, y^0) = (x_1^0, ..., x_m^0, y_1^0, ..., y_n^0) \in A$ і виконуються умови:

1) $F(x_1^0, ..., x_m^0, y_1^0, ..., y_n^0) = 0$;
2) $F \in C^1(A, \mathbb{R}^n)$;
3) $\frac{\partial(F_1, F_2, ..., F_n)}{\partial(y_1, y_2, ..., y_n)}(x^0, y^0) = det F'_y(x^0, y^0) \neq 0$.

Тоді існує куля $B(x^0, r) \subset \mathbb{R}^m$ і єдина функція $f \in C^1(B(x^0, r), \mathbb{R}^n)$, що задовольняє умови:

а) $f(x^0) = y^0$;
б) $\forall x \in B(x^0, r) \ : \ F(x, f(x)) = 0$;
в) $\forall x \in B(x^0, r) \ : \ f'(x) = -(F'_y(x, f(x))^{-1} F'_x(x, f(x))$,

де F'_x – похідна відображення F при фіксованому y.

Ідея доведення. В околі заданої точки можливе лінійне наближення функції

$$F(x, y) = F(x_1, ..., x_m, y_1, ..., y_n) = F(x_1^0, ..., x_m^0, y_1^0, ..., y_n^0) +$$
$$+ F'_{x_1}(x^0, y^0) \cdot (x_1 - x_1^0) + F'_{x_2}(x^0, y^0) \cdot (x_2 - x_2^0) + ... + F'_{x_m}(x^0, y^0) \cdot (x_m - x_m^0) +$$
$$+ F'_{y_1}(x^0, y^0) \cdot (y_1 - y_1^0) + ... + F'_{y_n}(x^0, y^0) \cdot (y_n - y_n^0) + r_2(x, y) =$$
$$= F'_x(x^0, y^0) \cdot (x - x^0) + F'_y(x^0, y^0) \cdot (y - y^0) + r_2(x, y).$$

Завдяки умові 2) функція диференційовна і похибка r_2 мала в околі точки (x^0, y^0). Завдяки умові 3) можна домножити на обернену матрицю:

$$0 = (F'_y(x^0, y^0))^{-1} F(x, y) =$$
$$= (F'_y(x^0, y^0))^{-1} F'_x(x^0, y^0) \cdot (x - x^0) + (y - y^0) + (F'_y(x^0, y^0))^{-1} r_2(x, y).$$

Звідси

$$y = y_0 - (F'_y(x^0, y^0))^{-1} F'_x(x^0, y^0) \cdot (x - x^0) - (F'_y(x^0, y^0))^{-1} r_2(x, y).$$

Це рівняння в просторі неперервних функцій $C(U)$, підібравши множину визначення U, можна записати у вигляді $y = f(y)$ і, завдяки малості правої частини, показати, що f – відображення стиску. Тоді на малій множині визначення U функція $y(x)$ визначиться однозначно і може бути порахована методом послідовних наближень.

Частинний випадок попередньої проблеми - потрібно з системи виразити (x_1, x_2) через (y_1, y_2) :

$$\begin{cases} y_1 = x_1^2 + x_2^2 - e^{x_1} \\ y_2 = \sin(1 + x_1 - x_2) \end{cases}$$

в околі точки $(x_1, x_2) = (0, 1)$ при значеннях $(y_1, y_2) = (0, 0)$.

Її розв'язок пропонує така теорема.

Теорема 5. **(Про обернену функцію).** Нехай $f : A \to \mathbb{R}^m,\ A \subset \mathbb{R}^m$ – відкрита множина, $x^0 \in A,\ y^0 = f(x^0)$ і виконуються умови:

1) $f \in C^1(A, \mathbb{R}^n)$;

3) $\frac{\partial(f_1,f_2,\ldots,f_m)}{\partial(x_1,x_2,\ldots,x_m)}(x^0) = det f'(x^0) \neq 0$.

Тоді існує куля $B(y^0, r) \subset \mathbb{R}^m$ і відкрита множина $U(x_0) \subset A$, що містить точку x_0 такі, що:

а) $f : U(x^0) \to B(y^0, r)$ – гомеоморфізм;

б) обернена функція $g = f^{-1} \in C^1(B(y^0, r), U(x^0))$;

в) $\forall y \in B(y^0, r)\ :\ g'(y) = (f'(g(y)))^{-1}$.

Доведення. Розглянемо відображення $F(x, y) = y - f(x),\ x \in A,\ y \in f(A)$. До нього застосуємо попередню теорему, помінявши місцями x і y. Всі її умови виконані. Тому існує функція $g : B(y_0, r) \to \mathbb{R}^m$ така, що:

а) $g(x^0) = y^0$;

б) $\forall y \in B(y^0, r)\ :\ f(g(y)) = y$;

в) $\forall y \in B(y^0, r)\ :\ f'(x) = (f'(g(y))^{-1}$.

Покладемо $U(x_0) := g(B(y^0, r))$. Тоді функція g обернена до функції $f : U(x_0) \to B(y_0, r)$. Отже, f – гомеоморфізм. Оскільки f – неперервна, то $U(x_0)$ – відкрита множина, як прообраз відкритої кулі.

На практиці нерідко потрібно знайти екстремуми функції за певних умов. Наприклад, найменше значення функції

$$f(a,b) = a + b, a > 0,\ b > 0,$$

за умови $a \cdot b = 1$ (найменше значення суми чисел з відомим добутком), чи найменше значення функції

$$f(x,y,z) = x^2 + y^2 + z^2$$

за умов $x+2y+3z = 1, 3x+2y+z = 1$ (найближча до початку координат точка прямої в просторі). Наведемо загальну теорію для розв'язання подібних задач.

Нехай $U \subset \mathbb{R}^m$ – відкрита множина,

$$f : U \to \mathbb{R},\ s < m,\ \varphi_i : U \to \mathbb{R},\ 1 \leq i \leq s,$$

$$M = \{x \in U \mid \varphi_i(x) = 0,\ 1 \leq i \leq s\}.$$

Рівняння $\varphi_1(x) = 0, ..., \varphi_s(x) = 0$, якими задається множина M, називають **рівняннями зв'язку.**

Означення 1. Нехай $x^0 \in M$. Функція f має в точці x^0 за умов $\varphi_1(x) = 0, ..., \varphi_s(x) = 0$

1) **локальний умовний максимум**, якщо

$$\exists r > 0 \; \forall x \in B(x^0, r) \cap M \; : \; f(x) \leq f(x^0).$$

2) **локальний умовний мінімум**, якщо

$$\exists r > 0 \; \forall x \in B(x^0, r) \cap M \; : \; f(x) \geq f(x^0).$$

3) **строгий локальний умовний максимум**, якщо

$$\exists r > 0 \; \forall x \in B(x^0, r) \cap M, \; x \neq x^0 \; : \; f(x) < f(x^0).$$

4) **строгий локальний умовний мінімум**, якщо

$$\exists r > 0 \; \forall x \in B(x^0, r) \cap M, \; x \neq x^0 \; : \; f(x) > f(x^0).$$

Теорема 1. **(Необхідна умова умовного екстремуму).** Нехай $s < m$, $f \in C^1(U)$, $\varphi_j \in C^1(U)$, $j = \overline{1, s}$, і матриця $(\frac{\partial \varphi_j}{\partial x_k}(x^0))_{j=1}^{s}{}_{k=1}^{m}$ має ранг s, x^0 – точка локального умовного екстремума функції f за обмежень $\varphi_1(x) = 0, ..., \varphi_s(x) = 0$.

Тоді існують дійсні числа $\lambda_1, ..., \lambda_s$ такі, що x^0 є точкою, підозрілою на екстремум функції

$$L = f + \lambda_1 \varphi_1 + ... + \lambda_s \varphi_s.$$

Функцію L називають **функцією Лагранжа,** а числа $\lambda_1, ..., \lambda_s$ – **множниками Лагранжа.**

Ідея доведення. Для визначеності вважатимемо, що

$$\frac{\partial(\varphi_1, ..., \varphi_s)}{\partial(x_{m-s+1}, ..., x_m)}(\overline{x}^0) \neq 0$$

(інакше змінні можна поміняти місцями).

Позначимо $\vec{x} := (x_1, ..., x_{m-s})$, $\vec{y} := (x_{m-s+1}, x_{m-s+2}, ..., x_m)$; $\vec{\varphi}(\vec{x}, \vec{y}) = (\varphi_1(\vec{x}, \vec{y}), \varphi_2(\vec{x}, \vec{y}), ..., \varphi_s(\vec{x}, \vec{y}))$. Тоді рівняння зв'язку можна записати так: $\vec{\varphi}(\vec{x}, \vec{y}) = 0$.

За допомогою теореми про неявну функцію в околі заданої точки можна виразити звідси $\vec{y}$ через $\vec{x} : \vec{y} = g(\vec{x})$. Підставимо в функцію, що досліджується на екстремум:

$$F(\vec{x}) = f(\vec{x}, \vec{y}) = f(\vec{x}, g(\vec{x})).$$

Отриманий вираз має екстремум в досліджуваній точці. Прирівняємо частинні похідні отриманого виразу до нуля:

$$F'_{x_k}(\vec{x}_0) = f'_{x_k}(\vec{x}^0, g(\vec{x}^0)) + f'_y(\vec{x}^0, g(\vec{x}^0)) \cdot g'_{x_k}(\vec{x}^0) = 0, \ 1 \leq k \leq m - s$$

(останній добуток – це скалярний добуток двох векторів). З теореми про неявну функцію

$$g'_{x_k}(\vec{x}^0) = -(\vec{\varphi}'_y(\vec{x}^0, \vec{y}^0))^{-1} \vec{\varphi}'_{x_k}(\vec{x}^0, \vec{y}^0).$$

Отже,

$$f'_{x_k}(\vec{x}^0, \vec{y}^0) - f'_y(\vec{x}^0, \vec{y}^0) \cdot (\vec{\varphi}'_y(\vec{x}^0, \vec{y}^0))^{-1} \vec{\varphi}'_{x_k}(\vec{x}^0, \vec{y}^0) = 0, \ 1 \leq k \leq m - s.$$

Якщо вектор $-f'_y(\vec{x}^0, \vec{y}^0) \cdot (\vec{\varphi}'_y(\vec{x}^0, \vec{y}^0))^{-1}$ позначити через λ, отримаємо

$$f'_{x_k}(\vec{x}^0, \vec{y}^0) + \lambda \cdot \vec{\varphi}'_{x_k}(\vec{x}^0, \vec{y}^0) = 0, \ 1 \leq k \leq m - s,$$

або розгорнуто

$$f'_{x_k}(\vec{x}^0, \vec{y}^0) + \sum_{p=1}^{s} \lambda_p \cdot (\varphi_p)'_{x_k}(\vec{x}^0, \vec{y}^0) = 0, \ 1 \leq k \leq m - s,$$

тобто частинна похідна функції Лагранжа по змінній x_k рівна 0.

Крім того, записавши означення

$$\lambda = -f'_y(\vec{x}^0, \vec{y}^0) \cdot (\vec{\varphi}'_y(\vec{x}^0, \vec{y}^0))^{-1},$$

ми можемо перетворити його:

$$f'_y(\vec{x}^0, \vec{y}^0) + \vec{\varphi}'_y(\vec{x}^0, \vec{y}^0) \cdot \lambda = \vec{0},$$

тобто частинні похідні функції Лагранжа по змінніх y рівні 0.

Теорема 2. Нехай виконуються умови:

1) $f \in C^2(U),\ \varphi \in C^2(U),\ 1 \le i \le s$;

2) $\overline{x}^0 \in M$ – критична точка функції $L = f + \sum\limits_{i=1}^{s} \lambda_i \varphi_i$ для деякого набору $\lambda_1, ..., \lambda_s \in \mathbb{R}$;

3) ранг матриці $\left(\frac{\partial \varphi_i(\overline{x}^0)}{\partial x_j}\right)_{i=1}^{s}\ {}_{j=1}^{m}$ рівний s. Для визначеності вважатимемо, що

$$\frac{\partial(\varphi_1, ..., \varphi_s)}{\partial(x_{m-s+1}, ..., x_m)}(\overline{x}^0) \neq 0.$$

Визначимо

$$Q(a_1, ..., a_{m-s}) = \sum_{i,j=1}^{m} \frac{\partial^2 L(\overline{x}^0)}{\partial x_i \partial x_j} a_i a_j,$$

де $a_{s-m+1}, ..., a_m$ визначені з рівнянь

$$\sum_{j=1}^{m} \frac{\partial \varphi_i(\overline{x}^0)}{\partial x_j} a_j = 0,\ 1 \le i \le s. \tag{1}$$

Тоді якщо Q – додатно визначена на $\mathbb{R}^{m-s}$, то $\overline{x}^0$ – точка строгого відносного локального мінімума, якщо Q – від'ємно визначена на $\mathbb{R}^{m-s}$, то $\overline{x}^0$ – точка строгого відносного локального максимума.

Ідея доведення. З рівняння $\vec{\varphi}(\vec{x}, \vec{y}) = 0$. виразити другий аргумент через перший за допомогою теореми про неявну функцію, підставити в функцію Лагранжа і переконатися, що її друга похідна є додатно або від'ємно визначеною.

Приклади. 1. Знайдемо екстремуми функції $f(x_1, x_2) = x_1 + x_2 + 1$ за умови $x_1^2 + x_2^2 = 1$.

Маємо $\varphi_1(x_1, x_2) = x_1^2 + x_2^2 - 1$. Складемо функцію Лагранжа

$$L(x_1, x_2, \lambda_1) = x_1 + x_2 + 1 + \lambda_1(x_1^2 + x_2^2 - 1).$$

Шукаємо точки, підозрілі на екстремум:

$$\begin{cases} \frac{\partial L}{\partial x_1} = 1 + 2\lambda_1 x_1 = 0 \\ \frac{\partial L}{\partial x_2} = 1 + 2\lambda_1 x_2 = 0 \\ \quad x_1^2 + x_2^2 = 1 \end{cases}$$

Ясно, що $\lambda_1 \neq 0$ і $x_1 = x_2 = -\frac{1}{2\lambda_1}$. Підставляємо в третє рівняння:

$$\frac{1}{4\lambda_1^2} + \frac{1}{4\lambda_1^2} = 1,$$

звідки $\lambda_1 = \pm\frac{1}{\sqrt{2}}$. Отже, є дві точки, підозрілих на екстремум:

$$(x_1, x_2) = (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}),\ \lambda = \frac{1}{\sqrt{2}};$$

$$(x_1, x_2) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}),\ \lambda = -\frac{1}{\sqrt{2}}.$$

Перевіримо достатні умови. З диференціала функції зв'язку

$$d\phi_1(x, a) = 2x_1 a_1 + 2x_2 a_2 = 0$$

виражаємо один з приростів і підставляємо в функцію

$$Q = d^2 L(x, a) = 2\lambda_1 a_1^2 + 2\lambda_1 a_2^2.$$

Для першої точки:

$$a_2 = -a_1, \quad Q = 2\sqrt{2}a_1^2 > 0,\ a_1 \neq 0,$$

тому це точка строгого умовного локального мінімума. Для другої точки:

$$a_2 = -a_1, \quad Q = -2\sqrt{2}a_1^2 < 0,\ a_1 \neq 0,$$

тому це точка строгого умовного локального максимума.

Зауваження. Користуючись цією теорією, можна знаходити найбільші і найменші значення на множині, досліджуючи спочатку внутрішні точки, а потім границю.

Приклад. Знайти найменше і найбільше значення функції $f(x_1, x_2) = x_1^4 + x_2^4 + 3$ на множині $\{(x_1, x_2) \mid x_1^2 + x_2^2 \leq 9\}$.

Оскільки функція неперервна і множина компактна, то за теоремою Вейєрштрасса мінімум і максимум обов'язково досягаються. Знайдемо підозрілі на екстремум точки.

I. Розглянемо внутрішні точки, для яких $x_1^2+x_2^2 < 9$. Необхідні умови:

$$\begin{cases} \frac{\partial f}{\partial x_1} = 4x_1^3 = 0 \\ \frac{\partial f}{\partial x_2} = 4x_2^3 = 0 \end{cases}$$

Отже, маємо точку $(0, 0)$.

II. Точки на межі $x_1^2 + x_2^2 = 9$. Тоді маємо умовний екстремум, записуємо функцію Лагранжа

$$L(x_1, x_2, \lambda_1) = x_1^4 + x_2^4 + \lambda_1(x_1^2 + x_2^2 - 9).$$

Необхідні умови:

$$\begin{cases} \frac{\partial L}{\partial x_1} = 4x_1^3 + 2\lambda_1 x_1 = 0 \\ \frac{\partial L}{\partial x_2} = 4x_2^3 + 2\lambda_1 x_2 = 0 \\ \quad x_1^2 + x_2^2 = 9 \end{cases}$$

Якщо $x_1 = 0$, то $x_2 = \pm 3$. Якщо $x_2 = 0$, то $x_1 = \pm 3$. Якщо ж $x_1 \neq 0,\ x_2 \neq 0$, то $x_1^2 = x_2^2 = -\frac{1}{2}\lambda_1$, звідки $x_1^2 = x_2^2 = \frac{9}{2}$, звідки отримаємо ще 4 підозрілі на екстремум точки $(\pm\frac{3}{\sqrt{2}}, \pm\frac{3}{\sqrt{2}})$.

Тепер порівнюємо значення функції:

$$f(0, 0) = 0,\ f(0, \pm 3) = f(\pm 3, 0) = 81, (\pm\frac{3}{\sqrt{2}}, \pm\frac{3}{\sqrt{2}}) = \frac{81}{2}.$$

Отже в точках $(0, \pm 3), (\pm 3, 0)$ функція набуває найбільшого значення, а в точці $(0, 0)$ найменшого.