ГЛАВА 13

ЭКСТРЕМУМЫ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

§ 1. Безусловный экстремум

Определение. Пусть на множестве $X \subset \mathbb{R}^n$ задана функция $f : X \to \mathbb{R}$. Точка $x^0 \in X$ называется точкой *строгого локального минимума (максимума)* функции f на множестве X, если

$$\exists \delta > 0 : \forall x \in \overset{o}{U}_\delta(x^0) \bigcap X \hookrightarrow \ f(x^0) < (>) f(x).$$

Если здесь строгое неравенство заменить нестрогим, то получится определение *нестрогого локального минимума (максимума)*. Точки минимума и максимума называются точками *экстремума*.

Определение. Пусть x^0 – точка локального экстремума функции f на множестве X. Тогда если $x^0 \in \operatorname{int} X$, то x^0 называется точкой *безусловного локального экстремума* функции f; если $x^0 \in \partial X$, то x^0 называется точкой *условного локального экстремума* функции f.

Лемма 1. x^0 – точка строгого безусловного локального минимума функции f тогда и только тогда, когда

$$\exists \delta > 0 : \forall x \in \overset{o}{U}_\delta(x^0) \hookrightarrow \ f(x^0) < f(x).$$

Доказательство. Пусть x^0 – точка строго безусловного локального минимума функции f. Тогда $\exists \delta_1 > 0 : \forall x \in \overset{o}{U}_{\delta_1}(x^0) \bigcap X \hookrightarrow$ $\hookrightarrow \ f(x^0) < f(x)$. Поскольку $x^0 \in \operatorname{int} X$, то $\exists \delta_2 > 0 : U_{\delta_2}(x^0) \subset X$. Определим $\delta = \min\{\delta_1, \delta_2\}$. Тогда $\forall x \in U_\delta(x^0) \hookrightarrow \ f(x^0) < f(x)$. Обратное утверждение очевидно. □

Утверждения, аналогичные лемме 1, справедливы и для максимума, и для нестрогих экстремумов. Иными словами, в определении безусловного экстремума множество X указывать не нужно.

Теорема 1. (Необходимое условие экстремума.) Пусть функция $f(x)$ определена в окрестности точки $x^0 \in \mathbb{R}^n$ и дифференцируема в точке x^0. Если x^0 – точка безусловного локального экстремума функции f, то $\operatorname{grad} f(x^0) = \overline{0}$.

Доказательство. Поскольку компоненты вектора $\operatorname{grad} f(x^0) = \overline{0}$ равны частным производным $\frac{\partial f}{\partial x_i}(x^0)$, то достаточно доказать, что $\frac{\partial f}{\partial x_i}(x^0) = 0 \quad \forall i \in \{1, \ldots, n\}$. Зафиксируем произвольное $i \in \{1, \ldots, n\}$ и рассмотрим функцию одной переменной $\varphi(x_i) = f(x_1^0, \ldots, x_{i-1}^0, x_i, x_{i+1}^0, \ldots, x_n^0)$. Поскольку x^0 – точка локального экстремума функции f, то x_i^0 – точка локального экстремума функции $\varphi(x_i)$. В силу теоремы Ферма $\varphi'(x_i^0) = 0$. Следовательно, $\frac{\partial f}{\partial x_i}(x^0) = \varphi'(x_i^0) = 0$. $\square$

Определение. Если $\operatorname{grad} f(x^0) = \overline{0}$, то точка x^0 называется *стационарной точкой* функции f.

Из теоремы 1 следует, что точки экстремума функции являются ее стационарными точками. Обратное неверно. Например, для функции одной переменной $f(x) = x^3$ точка $x_0 = 0$ является стационарной точкой, но не является точкой экстремума.

Пусть функция $f(x)$ дважды непрерывно дифференцируема в окрестности точки $x^0 \in \mathbb{R}^n$ (т. е. частные производные функции f до второго порядка включительно непрерывны в окрестности точки x^0). Тогда справедлива следующая формула Тейлора (глава 6, § 11, теорема 2):

$$f(x) - f(x^0) = df(x^0) + \frac{1}{2} d^2 f(x^0) + o(|\Delta x|^2) \quad \text{при} \quad \Delta x \to 0.$$

Через $f'_x(x^0)$ обозначим строку частных производных первого порядка или, что то же самое, координатную строку вектора градиента:

$$f'_x(x_0) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(x^0) & \cdots & \dfrac{\partial f}{\partial x_n}(x^0) \end{pmatrix},$$

а через $f''_{xx}(x^0)$ – матрицу частных производных второго порядка:

$$f''_{xx}(x^0) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1}(x^0) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_1}(x^0) \\ \cdots & \cdots & \cdots \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(x^0) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n}(x^0) \end{pmatrix}.$$

Поскольку для дважды непрерывно дифференцируемой функции f частные производные второго порядка не зависят от порядка дифференцирования, то $f''_{xx}(x^0)$ – симметрическая матрица.

Полагая $dx = \Delta x = x - x^0 = \begin{pmatrix} x_1 - x_1^0 \\ \cdots \\ x_n - x_n^0 \end{pmatrix}$, $\quad \Delta f = f(x) - f(x^0)$, получаем равенства

$$df(x^0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x^0)\, dx_i = f'_x(x^0)\, dx,$$

$$d^2 f(x^0) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x^0)\, dx_i\, dx_j = (dx)^T\, f''_{xx}(x^0)\, dx. \qquad (1)$$

Из формулы (1) следует, что второй дифференциал функции f является квадратичной формой относительно вектора dx.

Напомним, что квадратичная форма $k(x) = x^T\, M\, x$ называется

1) положительно определенной, если $k(x) > 0 \quad \forall x \neq \overline{0}$;

2) отрицательно определенной, если $k(x) < 0 \quad \forall x \neq \overline{0}$;

3) знаконеопределенной, если $\exists x_1, x_2 \in \mathbb{R}^n :\ k(x_1) > 0,\ k(x_2) < 0$.

Легко видеть, что квадратичная форма $k(x) = x^T\, M\, x$ отрицательно определена, если квадратичная форма $-k(x) = x^T\,(-M)\,x$ положительно определена. Для проверки положительной и отрицательной определенности квадратичной формы удобно использовать критерий Сильвестра. Доказательство того, что квадратичная форма знаконеопределена, проводят по определению.

Лемма 2. Если квадратичная форма $k(x)$ положительно определена, то

$$\exists \lambda > 0 : \forall x \in \mathbb{R}^n \hookrightarrow\ k(x) \geq \lambda |x|^2.$$

Доказательство. Поскольку функция $k(x) = x^T\, M\, x$ непрерывна, а единичная сфера $S = \{x \in \mathbb{R}^n\ :\ |x| = 1\}$ ограничена и замкнута, т. е. является компактом, то существует $\min\limits_{x \in S} k(x) = \lambda$. Из положительной определенности квадратичной формы $k(x)$ следует, что $\lambda > 0$. Из определения минимума получаем, что $\forall \widetilde{x} \in S \hookrightarrow\ k(\widetilde{x}) \geq \lambda$.

Если $x = \overline{0}$, то $k(x) = 0$, и неравенство $k(x) \geq \lambda |x|^2$ выполняется.

Если $x \neq \overline{0}$, то $\widetilde{x} = \frac{x}{|x|} \in S$, следовательно, $k(\widetilde{x}) \geq \lambda$, $\quad k(x) = x^T\, M\, x = |x|^2\, \widetilde{x}^T\, M\, \widetilde{x} = |x|^2\, k(\widetilde{x}) \geq \lambda |x|^2$. $\square$

Теорема 2. (Достаточные условия экстремума.) Пусть функция $f(x)$ дважды непрерывно дифференцируема в окрестности точки x^0 и пусть x^0 – стационарная точка функции f. Тогда

1) если квадратичная форма $d^2f(x^0)$ положительно определена, то x^0 – точка строгого безусловного локального минимума функции f;

2) если квадратичная форма $d^2f(x^0)$ отрицательно определена, то x^0 – точка строгого безусловного локального максимума функции f;

3) если квадратичная форма $d^2f(x^0)$ знаконеопределена, то x^0 не является точкой безусловного локального экстремума функции f;

4) если квадратичная форма $d^2f(x^0)$ не является ни положительно, ни отрицательно определенной и не является знаконеопределенной, то x^0 может быть точкой локального экстремума, а может и не быть.

Доказательство. Поскольку x^0 – стационарная точка функции f, то $df(x^0) = \overline{0}$, следовательно,

$$\Delta f = \frac{1}{2}d^2f(x^0) + o(|\Delta x|^2) \quad \text{при} \quad \Delta x \to 0. \tag{2}$$

1) Пусть квадратичная форма $d^2f(x^0)$ положительно определена. В силу леммы 2 $\exists\lambda > 0 \ : \ \forall\Delta x \in \mathbb{R}^n \hookrightarrow \ d^2f(x^0) \geq \lambda|\Delta x|^2$. Отсюда и из формулы Тейлора (2) следует, что $\Delta f \geq \frac{\lambda}{2}|\Delta x|^2 + o(|\Delta x|^2)$ при $\Delta x \to 0$. По определению o-малого $\lim\limits_{\Delta x\to 0} \frac{o(|\Delta x|^2)}{|\Delta x|^2} = 0$, поэтому

$$\lim_{\Delta x\to 0} \frac{\frac{\lambda}{2}|\Delta x|^2 + o(|\Delta x|^2)}{|\Delta x|^2} = \frac{\lambda}{2} > 0.$$

Следовательно,

$$\exists\delta > 0 \ : \ \forall\Delta x \in U_\delta(\overline{0}) \hookrightarrow \ \frac{\lambda}{2}|\Delta x|^2 + o(|\Delta x|^2) > 0,$$

поэтому $\forall x \in U_\delta(x^0) \hookrightarrow \ f(x) - f(x^0) = \Delta f \geq \frac{\lambda}{2}|\Delta x|^2 + o(|\Delta x|^2) > 0$, а значит, x^0 – точка строгого безусловного локального минимума.

Пункт (2) сводится к пункту (1) заменой функции $f(x)$ на $-f(x)$.

3) Пусть квадратичная форма $d^2f(x^0) = (dx)^T\, f''_{xx}(x^0)\, dx$ знаконеопределена, т. е. $\exists\xi_1, \xi_2 \in \mathbb{R}^n$: $\xi_1^T\, f''_{xx}(x^0)\, \xi_1 > 0$, $\xi_2^T\, f''_{xx}(x^0)\, \xi_2 < 0$. Применяя формулу (2) для $\Delta x = t\xi_1$, получим $f(x^0 + t\xi_1) - f(x^0) = \Delta f = \frac{1}{2}(\Delta x)^T\, f''_{xx}(x^0)\, \Delta x + o(|\Delta x|^2) = \frac{1}{2}(\xi_1)^T\, f''_{xx}(x^0)\, \xi_1\, t^2 + o(t^2)$

при $t \to 0$. Поскольку

$$\lim_{t\to 0} \frac{1}{t^2}\left(\frac{1}{2}(\xi_1)^T f''_{xx}(x^0)\,\xi_1\,t^2 + o(t^2)\right) = \frac{1}{2}(\xi_1)^T f''_{xx}(x^0)\,\xi_1 > 0,$$

то

$$\exists \delta_1 > 0 :\ \forall t \in (0,\delta_1) \hookrightarrow\ \Delta f = f(x^0 + t\xi_1) - f(x^0) > 0.$$

Аналогично,

$$\exists \delta_2 > 0 :\ \forall t \in (0,\delta_2) \hookrightarrow\ \Delta f = f(x^0 + t\xi_2) - f(x^0) < 0.$$

Поэтому $\forall \delta > 0\quad \exists t_1 = \min\left\{\frac{\delta_1}{2}, \frac{\delta}{2|\xi_1|}\right\}$, $\exists t_2 = \min\left\{\frac{\delta_2}{2}, \frac{\delta}{2|\xi_2|}\right\}$ такие, что $x^1 = x^0 + t_1\xi_1 \in U_\delta(x^0)$, $x^2 = x^0 + t_2\xi_2 \in U_\delta(x^0)$, $f(x^1) - f(x^0) > 0$, $f(x^2) - f(x^0) < 0$. Следовательно, точка x^0 не является ни точкой локального минимума, ни точкой локального максимума функции f.

4) Пусть $f(x)$ – функция одной переменной $x \in \mathbb{R}$. Тогда в случае $f''(x_0) = 0$ квадратичная форма $d^2 f(x_0) = f''(x_0)\,(dx)^2$ не является положительно определенной, отрицательно определенной, а также не является знаконеопределенной.

Для функции $f(x) = x^4$ имеем $f''(0) = 0$, а точка $x_0 = 0$ является точкой минимума.

Для функции $f(x) = x^3$: $f''(0) = 0$, а точка $x_0 = 0$ не является точкой экстремума. $\square$

Задача 1. Пусть функция $f : \mathbb{R}^2 \to \mathbb{R}$ дважды непрерывно дифференцируема и $\det f''_{xx}(x^0) < 0$. Может ли функция f достигать локальный безусловный экстремум в точке x^0?

§ 2. Условный экстремум

Пусть в окрестности точки $x^0 \in \mathbb{R}^n$ заданы скалярная функция $f(x)$ и вектор-функция $g(x) = \begin{pmatrix} g_1(x) \\ \cdots \\ g_m(x) \end{pmatrix}$. Рассмотрим задачу отыскания экстремума функции $f(x)$ на множестве $X \subset \mathbb{R}^n$, заданном системой уравнений $g(x) = \overline{0}$:

$$f(x) \to ext\ :\ g(x) = \overline{0}. \tag{1}$$