ГЛАВА Функции нескольких переменных (продолжение)

§ Экстремумы функции некольких переменных

Пусть функция $f(\mathbf{x})$ определена при $\mathbf{x} = (x_1, \ldots, x_n)^T \in D \subset R^n$.

Определение. *1) Точка $\mathbf{a} = (a_1, \ldots, a_n)^T \in D$ называется точкой локального минимума функции $f(\mathbf{x})$, если:*

$$\exists \delta > 0: \ \forall x \in (U_\delta(\mathbf{a}) \cap D) \ \Rightarrow \ f(\mathbf{x}) \geq f(\mathbf{a}).$$

2) Точка $\mathbf{a} = (a_1, \ldots, a_n)^T \in D$ называется точкой локального максимума функции $f(\mathbf{x})$, если:

$$\exists \delta > 0: \ \forall x \in (U_\delta(\mathbf{a}) \cap D) \ \Rightarrow \ f(\mathbf{x}) \leq f(\mathbf{a}).$$

Здесь

$$U_\delta(\mathbf{a}) = \{\mathbf{x} \in R^n: \ \rho(\mathbf{x}, \mathbf{a}) < \delta\}$$

— δ-окрестность точки $\mathbf{a}$; $\rho(\cdot)$ — метрика в пространстве R^n.

Точки локального минимума и максимума называются точками локального экстремума.

Теорема (необходимые условия локального экстремума). *Пусть $\mathbf{a}$ — внутренняя точка области D, и функция $f(\mathbf{x})$ имеет частные производные по своим аргументам в этой точке. Тогда для того чтобы точка $\mathbf{a}$ являлась точкой локального экстремума функции $f(\mathbf{x})$ необходимо, чтобы выполнялись условия:*

$$f'_{x_i}(\mathbf{a}) = 0, \quad i = 1, \ldots, n.$$

Доказательство. Рассмотрим функцию

$$F(x_1) = f(x_1, a_1, \ldots, a_n).$$

Тогда точка $x_1 = a_1$ является точкой локального экстремума функции $F(x_1)$. По теореме Ферма имеем:

$$F'(a_1) = f'_{x_1}(\mathbf{a}) = 0.$$

Аналогично показывается равенство нулю в точке $\mathbf{a}$ остальных частных производных функции $f(\mathbf{x})$. Теорема доказана.

Замечание. Отметим, что необходимое условие локального экстремума: $f'_{x_i}(\mathbf{a}) = 0$, $i = 1, \ldots, n$, можно переписать в виде: $df(\mathbf{a}) = 0$, или $\vec{\nabla} f(\mathbf{a}) = \mathbf{0}$.

Замечание. Выписанные условия экстремума являются только необходимыми, но не достаточными. Например, пусть $f(x, y) = x^2 - y^2$. Тогда $f'_x(0,0) = 0$, $f'_y(0,0) = 0$, но тем не менее точка $(0,0)$ не является точкой локального экстремума функции $f(x,y)$. В самом деле, $f(\delta, 0) > 0$, $f(0, \delta) < 0$ для любого сколь угодно малого $\delta > 0$, т.е. в любой окрестности точки $(0,0)$ имеются как точки, в которых значение функции больше, чем $f(0,0) = 0$, так и точки, в которых меньше.

Далее точку $\mathbf{a}$, для которой $df(\mathbf{a}) = 0$, будем называть стационарной точкой функции $f(\mathbf{x})$, или точкой, "подозрительной" на экстремум.

Замечание. Если точка $\mathbf{a}$ не является внутренней точкой области D и/или в этой точке не существуют частные производные функции $f(\mathbf{x})$, то пользоваться сформулированной выше теоремой, вообще говоря, нельзя.

Далее пусть в точке $\mathbf{a}$ выполнены необходимые условия локального экстремума. Для установления достаточных условий дополнительно предположим, что функция

$f(\mathbf{x})$ имеет непрерывные частные производные до второго порядка включительно в окрестности точки $\mathbf{a}$. Тогда по формуле Тейлора для любой точки $\mathbf{x}$ из этой окрестности имеем:

$$f(\mathbf{x}) = f(\mathbf{a}) + df(\mathbf{a}) + R_2,$$

где

$$R_2 = \frac{1}{2}\left(\frac{\partial}{\partial x_1}dx_1 + \ldots + \frac{\partial}{\partial x_n}dx_n\right)^2 f(\tilde{\mathbf{x}}) = \frac{1}{2}\sum_{i,j=1}^{n} f''_{x_i x_j}(\tilde{\mathbf{x}})dx_i dx_j,$$

где $\tilde{\mathbf{x}}$ — некоторая точка, лежащая на отрезке, соединяющем точки $\mathbf{a}$ и $\mathbf{x}$. Выражение R_2 представляет собой квадратичную форму относительно $dx_1, \ldots, dx_n$:

$$R_2 = \frac{1}{2}(\mathbf{dx})^T \mathbf{S}(\tilde{\mathbf{x}})\mathbf{dx}.$$

Здесь $\mathbf{dx} = (dx_1, \ldots, dx_n)^T$; $\mathbf{S}(\tilde{\mathbf{x}}) = \{f''_{x_i x_j}(\tilde{\mathbf{x}})\}_{i,j=1}^n$ — матрица Гессе в точке $\tilde{\mathbf{x}}$.

Учитывая стационарность точки $\mathbf{a}$, получаем, что

$$\Delta f = f(\mathbf{x}) - f(\mathbf{a}) = R_2.$$

Таким образом, для того, чтобы точка $\mathbf{a}$ была точкой локального минимума (максимума), достаточно потребовать, чтобы выражение R_2 было положительным (отрицательным) при всех $\mathbf{x}$ из некоторой окрестности точки $\mathbf{a}$.

Для продолжения анализа рассмотрим некоторые результаты из теории квадратичных форм.

Квадратичной формой называется функция вида

$$\varphi(\mathbf{z}) = \mathbf{z}^T \mathbf{A}\mathbf{z} = \sum_{i,j=1}^{n} a_{ij} z_i z_j.$$

Здесь $\mathbf{z} = (z_1, \ldots, z_n)^T$; $\mathbf{A} = \{a_{ij}\}_{i,j=1}^n$ — симметричная матрица квадратичной формы.

Определение. *Квадратичная форма (матрица квадратичной формы) называется:*

1) положительно определенной, если $\varphi(\mathbf{z}) > 0$, $\forall \mathbf{z} \neq \mathbf{0}$*;*

2) отрицательно определенной, если $\varphi(\mathbf{z}) < 0$, $\forall \mathbf{z} \neq \mathbf{0}$*;*

3) знакопостоянной положительной, если $\varphi(\mathbf{z}) \geq 0$, $\forall \mathbf{z} \in R^n$*;*

4) знакопостоянной отрицательной, если $\varphi(\mathbf{z}) \leq 0$, $\forall \mathbf{z} \in R^n$*;*

5) знакопеременной, если $\exists \mathbf{z}_1, \mathbf{z}_2 \in R^n$*:* $\varphi(\mathbf{z}_1) > 0$, $\varphi(\mathbf{z}_2) < 0$.

Пример. Пусть $n = 2$. Тогда

1) $\varphi(x, y) = x^2 + y^2$ — положительно определенная квадратичная форма;

2) $\varphi(x, y) = -x^2 - y^2$ — отрицательно определенная квадратичная форма;

3) $\varphi(x, y) = (x - y)^2$ — знакопостоянная положительная квадратичная форма;

4) $\varphi(x, y) = -(x - y)^2$ — знакопостоянная отрицательная квадратичная форма;

5) $\varphi(x, y) = x^2 - y^2$ — знакопеременная квадратичная форма.

Теорема (критерий Сильвестра). *1) Для того, чтобы квадратичная форма была положительно определенной необходимо и достаточно, чтобы все левые верхние угловые миноры ее матрицы были положительными.*

2) Для того, чтобы квадратичная форма была отрицательно определенной необходимо и достаточно, чтобы левые верхние угловые миноры ее матрицы чередовали знак, начиная с "−".

Пример. Пусть

$$\varphi(z_1, z_2, z_3) = 3z_1^2 + 2z_2^2 + 4z_3^2 - z_1z_2 + 2z_1z_3 + 4z_2z_3.$$

Матрица этой квадратичной формы имеет вид:

$$\mathbf{A} = \begin{pmatrix} 3 & -1/2 & 1 \\ -1/2 & 2 & 2 \\ 1 & 2 & 4 \end{pmatrix}.$$

Вычислим ее левые верхние угловые миноры:

$$\Delta_1 = 3 > 0, \quad \Delta_2 = 23/4 > 0, \quad \Delta_3 = 7 > 0.$$

Согласно критерию Сильвестра, данная квадратичная форма — положительно определенная.

Пример. Пусть

$$\varphi(z_1, z_2, z_3) = -z_1^2 - z_2^2 - z_3^2 + z_1z_2.$$

Матрица этой квадратичной формы имеет вид:

$$\mathbf{A} = \begin{pmatrix} -1 & 1/2 & 0 \\ 1/2 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Вычислим ее левые верхние угловые миноры:

$$\Delta_1 = -1 < 0, \quad \Delta_2 = 3/4 > 0, \quad \Delta_3 = -3/4 < 0.$$

Согласно критерию Сильвестра, данная квадратичная форма — отрицательно определенная.

Вернемся теперь к рассмотрению экстремумов функции.

Теорема (достаточные условия локального экстремума). *Пусть* $\mathbf{a}$ *— внутренняя точка области* D*, функция* $f(\mathbf{x})$ *имеет непрерывные частные производные по своим аргументам в этой точке до второго порядка включительно, и* $df(\mathbf{a}) = 0$*. Тогда:*

1) если матрица $\mathbf{S}(\mathbf{a})$ *положительно определена, то* $\mathbf{a}$ *— точка локального минимума;*

2) если матрица $\mathbf{S}(\mathbf{a})$ *отрицательно определена, то* $\mathbf{a}$ *— точка локального максимума;*

3) если матрица $\mathbf{S}(\mathbf{a})$ *— знакопеременная, то* $\mathbf{a}$ *не является точкой локального экстремума;*

4) если матрица $\mathbf{S}(\mathbf{a})$ *— знакопостоянная, но не знакоопределенная, то точка* $\mathbf{a}$ *может как являться точкой локального экстремума, так и не являться.*

Для доказательства данной теоремы достаточно заметить, что, в силу непрерывности, если матрица Гессе положительно определена (отрицательно определена, знакопеременна) в точке $\mathbf{a}$, то тогда она будет таковой и во всех точках $\tilde{\mathbf{x}}$ из некоторой малой окрестности точки $\mathbf{a}$. Таким образом, в пунктах 1)–3) вопрос о достаточных условиях локального экстремума решается на основе проверки знака второго дифференциала $d^2f(\mathbf{a})$.

Для иллюстрации пункта 4) теоремы рассмотрим функции $f_1(x, y) = x^4 + y^4$ и $f_2(x, y) = x^3 + y^3$. В обоих случаях точка $(0, 0)$ — это стационарная точка, и матрицы Гессе в этой точке — нулевые (знакопостоянные). Однако, для функции $f_1(x, y)$ точка $(0, 0)$ является точкой экстремума (минимума), а для функции $f_2(x, y)$ — нет.

Если матрица $\mathbf{S}(\mathbf{a})$ — знакопостоянная, но не знакоопределенная, то для построения достаточных условий локального экстремума надо задействовать производные и дифференциалы порядка выше второго (см. формулу Тейлора):

$$\Delta f = f(\mathbf{x}) - f(\mathbf{a}) = df(\mathbf{a}) + \frac{1}{2!}d^2 f(\mathbf{a}) + \frac{1}{3!}d^3 f(\mathbf{a}) + \frac{1}{4!}d^4 f(\mathbf{a}) + \ldots .$$

Отметим, что хороших (простых) критериев проверки знакоопределенности однородных форм порядка выше 2, вообще говоря, нет.

Определение. *1) Точка $\mathbf{a} \in D$ называется точкой глобального минимума функции $f(\mathbf{x})$ в области D, если: $f(\mathbf{x}) \geq f(\mathbf{a})$, $\forall x \in D$.*

2) Точка $\mathbf{a} \in D$ называется точкой глобального максимума функции $f(\mathbf{x})$ в области D, если: $f(\mathbf{x}) \leq f(\mathbf{a})$, $\forall x \in D$.

Точки глобального минимума и максимума называются точками глобального экстремума. Глобальные экстремумы (если они существуют) могут достигаться либо в точках локального экстремума внутри области D, либо на границе области D. Заметим, что для проверки граничных точек области D на экстремальность сформулированные выше теоремы применять нельзя. В этом случае можно воспользоваться методами поиска условного экстремума (см. далее).

Пример. Найдем экстремумы функции

$$f(x, y) = x^3 - 3xy + 2y^2 - y + 1$$

в пространстве R^2.

Выписываем необходимые условия экстремума:

$$f'_x = 3x^2 - 3y = 0, \quad f'_y = -3x + 4y - 1 = 0.$$

Данная система уравнений имеет два решения: $(1, 1)$ и $(-\frac{1}{4}, \frac{1}{16})$.

Проверяем достаточные условия экстремума:

$$f''_{x^2} = 6x, \quad f''_{xy} = -3, \quad f''_{y^2} = 4.$$

Значит,

$$\mathbf{S}(x, y) = \begin{pmatrix} 6x & -3 \\ -3 & 4 \end{pmatrix}.$$

Матрица $\mathbf{S}(1, 1)$ положительно определена, следовательно, точка $(1, 1)$ — точка локального минимума; матрица $\mathbf{S}(-\frac{1}{4}, \frac{1}{16})$ — знакопеременная, следовательно, точка $(-\frac{1}{4}, \frac{1}{16})$ не является точкой локального экстремума.

Нетрудно заметить, что

$$\sup_{R^2} f(x, y) = +\infty, \quad \inf_{R^2} f(x, y) = -\infty,$$

т.е. глобальных экстремумов в области R^2 у функции нет.

Пример. Рассмотрим задачу: найти экстремумы функции

$$f(x, y) = (x - 1)^2 + (y - 2)^2 + 1$$

при $(x, y) \in D$, где

$$D: \; x^2 + y^2 \leq 9.$$

Ищем сначала локальные экстремумы внутри области D:

$$f'_x = 2(x-1) = 0, \quad f'_y = 2(y-2) = 0.$$

Получили одну стационарную точку: $(1,2) \in D$.

Имеем

$$f''_{x^2} = f''_{y^2} = 2, \quad f''_{xy} = 0.$$

Строим матрицу Гессе:

$$\mathbf{S}(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Она положительно определена при любых x и y, значит, $(1,2)$ — точка локального (нетрудно заметить, что и глобального) минимума (см. рисунок 1).

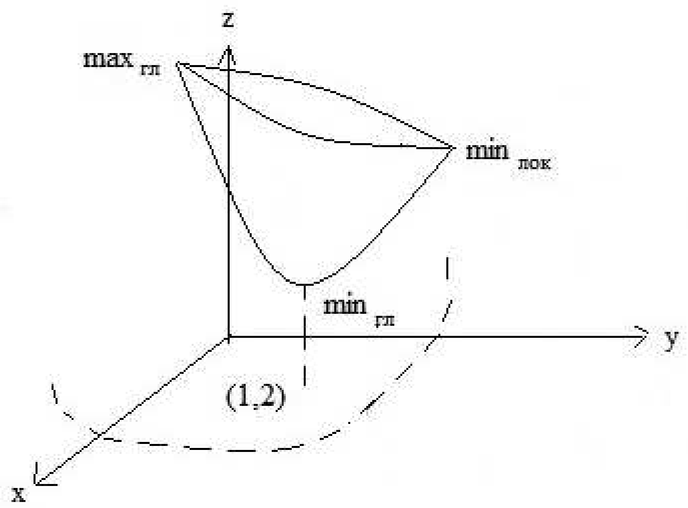

Рис. 1.

Глобальный максимум функции в области D достигается на границе этой области. Из уравнения границы:

$$x^2 + y^2 = 9$$

можно исключить одну из переменных, например,

$$y = \pm\sqrt{9-x^2}, \quad x \in [-3,3],$$

и свести задачу к поиску безусловного экстремума:

$$F(x) = f(x, \pm\sqrt{9-x^2}) = (x-1)^2 + (\pm\sqrt{9-x^2} - 2)^2 + 1 \to \text{extr}$$

при $x \in [-3,3]$.

Можно уравнение границы записать в параметрическом виде:

$$\begin{cases} x = 3\cos t, \\ y = 3\sin t, \end{cases} \qquad t \in [0, 2\pi].$$

Тогда снова получим одномерную задачу на поиск безусловного экстремума:

$$\tilde{F}(t) = f(3\cos t, 3\sin t) = (3\cos t - 1)^2 + (3\sin t - 2)^2 + 1 \to \text{extr}$$

при $t \in [0, 2\pi]$.

Решив полученную задачу в одной или другой форме, найдем глобальный максимум функции, достигаемый ею на границе области D. Также на границе области D имеется один локальный минимум.