Э. И. Зверович

ВЕЩЕСТВЕННЫЙ И КОМПЛЕКСНЫЙ АНАЛИЗ

Учебное пособие
в шести частях

Часть 3
Дифференциальное исчисление функций векторного аргумента

Минск

БГУ

2004

В этом томе излагается теоретический материал, который преподается студентам математических специальностей университетов во втором — третьем семестрах.

Глава 14

ПРЕДЕЛЫ И НЕПРЕРЫВНОСТЬ ОТОБРАЖЕНИЙ ИЗ $\mathbb{R}^n$ В $\mathbb{R}^m$

§ 1. Некоторые свойства евклидовых пространств $\mathbb{R}^n$ и их подмножеств

1. Норма и скалярное произведение в $\mathbb{R}^n$

Пусть $\mathbb{R}$ — множество всех вещественных чисел.

Определение 1. *Евклидово n-мерное пространство $\mathbb{R}^n$ определим как множество всех n-членных последовательностей вещественных чисел, т. е.*

$$\mathbb{R}^n := \left\{ \boldsymbol{x} = (x_1\,, \ldots\,, x_n) \;\middle|\; x_1 \in \mathbb{R}, \ldots\,, x_n \in \mathbb{R} \right\}. \tag{14.1}$$

Элементы множества $\mathbb{R}^n$ часто называются его *точками.* Пространства $\mathbb{R} = \mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^3$ называются соответственно *координатной прямой, координатной плоскостью* и *координатным пространством.* Каждое из этих трех пространств допускает наглядное и широко используемое геометрическое изображение (рис. 1).

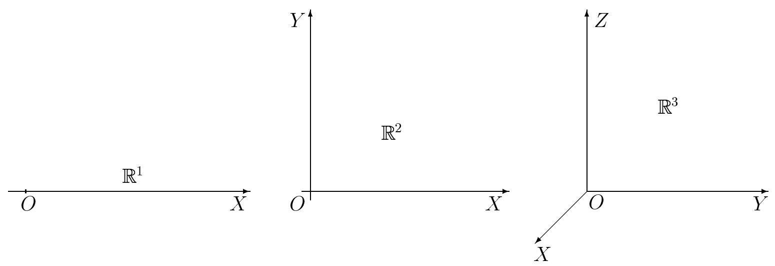

Рис. 1. Координатные пространства малых размерностей

Если $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, то числа $x_1, \ldots, x_n$ называются *координатами* точки $\boldsymbol{x}$. Из определения 1 следует, что равенство $\boldsymbol{x} = \boldsymbol{y}$, где $\boldsymbol{y} = (y_1, \ldots, y_n)$, равносильно системе n равенств соответствующих координат, т. е.

$$\boldsymbol{x} = \boldsymbol{y} \quad \Longleftrightarrow \quad \begin{cases} x_1 = y_1\,, \\ \cdots\cdots\,, \\ x_n = y_n\,. \end{cases} \tag{14.2}$$

Точки из $\mathbb{R}^n$ часто называются также *векторами*, ибо пространство $\mathbb{R}^n$, наделенное операциями

$$\begin{aligned} \boldsymbol{x} + \boldsymbol{y} &:= (x_1 + y_1, \ldots, x_n + y_n)\,, \\ \lambda \cdot \boldsymbol{x} &:= (\lambda x_1, \ldots, \lambda x_n)\,, \quad \text{где} \quad \lambda \in \mathbb{R}, \end{aligned} \tag{14.3}$$

действительно есть линейное (векторное) пространство (над полем $\mathbb{R}$ и n-мерное).

Кроме операций (14.3) введем еще *скалярное произведение* векторов из $\mathbb{R}^n$, т. е. отображение из $\mathbb{R}^n \times \mathbb{R}^n$ в $\mathbb{R}$ (будем обозначать его одним из символов $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ или $\boldsymbol{x} \cdot \boldsymbol{y}$, или $\boldsymbol{x}\,\boldsymbol{y}$), определяемое равенством

$$\boldsymbol{x}\,\boldsymbol{y} = \boldsymbol{x} \cdot \boldsymbol{y} = \langle \boldsymbol{x}, \boldsymbol{y} \rangle := \sum_{k=1}^{n} x_k y_k\,.$$

В векторном пространстве $\mathbb{R}^n$ вводится понятие длины (евклидовой нормы) вектора $\boldsymbol{x} = (x_1, \ldots, x_n)$ следующим образом:

$$|\boldsymbol{x}| := \sqrt{(x_1)^2 + \cdots + (x_n)^2}\,. \tag{14.4}$$

При $n = 1$, т. е. в $\mathbb{R}$, понятие длины совпадает с понятием модуля (абсолютной величины). Изучим некоторые связи между введенными понятиями.

Теорема 1. *Если* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$, *то имеют место следующие соотношения:*

(a) $|\boldsymbol{x}| \geqslant 0$ *причем* $|\boldsymbol{x}| = 0 \iff \boldsymbol{x} = \boldsymbol{0}$*;*

(b) $|\boldsymbol{x}\,\boldsymbol{y}| \leqslant |\boldsymbol{x}| \cdot |\boldsymbol{y}|$ (неравенство Коши — Буняковского)*; здесь равенство равносильно линейной зависимости векторов* $\boldsymbol{x}$ *и* $\boldsymbol{y}$*;*

(c) $\big||\boldsymbol{x}| - |\boldsymbol{y}|\big| \leqslant |\boldsymbol{x} + \boldsymbol{y}| \leqslant |\boldsymbol{x}| + |\boldsymbol{y}|$;

(d) $|\lambda \boldsymbol{x}| = |\lambda| \cdot |\boldsymbol{x}|$.

◀ (a) Неравенство $\sqrt{x_1^2+\cdots+x_n^2} \geqslant 0$ очевидно. Далее имеем

$$\sqrt{x_1^2+\cdots+x_n^2}=0 \iff x_1^2+\cdots+x_n^2=0 \iff x_1=\cdots=x_n=0\,,$$

т. е. $\boldsymbol{x}=(0\,,0\ldots,0)=\mathbf{0}$ (нуль-вектор).

(b) Предположим сначала, что векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ линейно зависимы, т. е. либо $\boldsymbol{y}=\mathbf{0}$, либо $\boldsymbol{x}=\lambda\boldsymbol{y}$ при некотором $\lambda\in\mathbb{R}$. В первом случае неравенство Коши — Буняковского очевидно, так как обращается в равенство $0=0$. Во втором случае имеем

$$|\,\boldsymbol{x}\,\boldsymbol{y}\,| = \left|\sum_{k=1}^{n} x_k y_k\right| = \left|\sum_{k=1}^{n}\lambda y_k^2\right| = \sqrt{\sum_{k=1}^{n}(\lambda y_k)^2}\cdot\sqrt{\sum_{k=1}^{n} y_k^2} = |\,\boldsymbol{x}\,|\cdot|\,\boldsymbol{y}\,|\,.$$

Предположим теперь, что векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ линейно независимы, т. е. что при любом $\lambda\in\mathbb{R}$ имеет место неравнство $\boldsymbol{x}-\lambda\boldsymbol{y}\neq\mathbf{0}$. Тогда в силу (a) имеем

$$0<|\,\lambda\boldsymbol{y}-\boldsymbol{x}\,|^2=|\,\boldsymbol{y}\,|^2\cdot\lambda^2+2\boldsymbol{x}\,\boldsymbol{y}\cdot\lambda+|\,\boldsymbol{x}\,|^2\,.$$

Получен квадратный трехчлен относительно λ, положительный при всех $\lambda\in\mathbb{R}$. Как известно, положительность трехчлена при всех λ равносильна отрицательности его дискриминанта, т. е. выполняется неравенство $\langle\boldsymbol{x}\,,\boldsymbol{y}\rangle^2-|\,\boldsymbol{x}\,|^2\cdot|\,\boldsymbol{y}\,|^2<0$, равносильное неравенству Коши — Буняковского.

(c) Используя неравенство Коши — Буняковского, имеем

$$|\boldsymbol{x}+y|^2=\sum_{k=1}^{n}(x_k+y_k)^2=\sum_{k=1}^{n}x_k^2+2\sum_{k=1}^{n}x_k y_k+\sum_{k=1}^{n}y_k^2\leqslant$$
$$\leqslant|\boldsymbol{x}|^2+2|\boldsymbol{x}|\cdot|\boldsymbol{y}|+|\boldsymbol{y}|^2=(|\boldsymbol{x}|+|\boldsymbol{y}|)^2.$$

Отсюда находим $|\,\boldsymbol{x}+\boldsymbol{y}\,|^2\leqslant(|\,\boldsymbol{x}\,|+|\,\boldsymbol{y}\,|)^2$, что равносильно правому неравенству (c). Левое неравенство (c) является следствием правого.

(d) Имеем $|\,\lambda\boldsymbol{x}\,|=\sqrt{\sum\limits_{k=1}^{n}(\lambda x_k)^2}=|\lambda|\cdot|\,\boldsymbol{x}\,|\,.$ ▶

Изучим теперь свойства скалярного произведения векторов как функции его сомножителей.

Теорема 2. *Если*[1] $\boldsymbol{x}$, $\boldsymbol{x}^1$, $\boldsymbol{x}^2$, $\boldsymbol{y}$, $\boldsymbol{y}^1$, $\boldsymbol{y}^2 \in \mathbb{R}^n$ *и* $\lambda \in \mathbb{R}$, *то*

(a) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$ (симметричность);

(b) $\left.\begin{array}{l} \langle \lambda\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \lambda\boldsymbol{y} \rangle = \lambda \cdot \langle \boldsymbol{x}, \boldsymbol{y} \rangle\,, \\ \langle \boldsymbol{x}^1 + \boldsymbol{x}^2, \boldsymbol{y} \rangle = \langle \boldsymbol{x}^1, \boldsymbol{y} \rangle + \langle \boldsymbol{x}^2, \boldsymbol{y} \rangle\,, \\ \langle \boldsymbol{x}, \boldsymbol{y}^1 + \boldsymbol{y}^2 \rangle = \langle \boldsymbol{x}, \boldsymbol{y}^1 \rangle + \langle \boldsymbol{x}, \boldsymbol{y}^2 \rangle \end{array}\right\}$ (билинейность);

(c) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geqslant 0$; $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0 \iff \boldsymbol{x} = \boldsymbol{0}$;

(d) $|\boldsymbol{x}| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$;

(e) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \dfrac{1}{4} \cdot \left(|\boldsymbol{x} + \boldsymbol{y}|^2 - |\boldsymbol{x} - \boldsymbol{y}|^2\right)$.

◀ (a) Имеем

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{k=1}^{n} x_k y_k = \sum_{k=1}^{n} y_k x_k = \langle \boldsymbol{y}, \boldsymbol{x} \rangle\,.$$

(b) Имеем

$$\langle \lambda\boldsymbol{x}, \boldsymbol{y} \rangle = \\ = \sum_{k=1}^{n} \lambda x_k y_k = \sum_{k=1}^{n} x_k \cdot \lambda y_k = \langle \boldsymbol{x}, \lambda\boldsymbol{y} \rangle = \lambda \sum_{k=1}^{n} x_k y_k = \lambda \cdot \langle \boldsymbol{x}, \boldsymbol{y} \rangle\,;$$

$$\langle \boldsymbol{x}^1 + \boldsymbol{x}^2, \boldsymbol{y} \rangle = \\ = \sum_{k=1}^{n} (x_k^1 + x_k^2) y_k = \sum_{k=1}^{n} x_k^1 y_k + \sum_{k=1}^{n} x_k^2 y_k = \langle \boldsymbol{x}^1, \boldsymbol{y} \rangle + \langle \boldsymbol{x}^2, \boldsymbol{y} \rangle\,;$$

$$\langle \boldsymbol{x}, \boldsymbol{y}^1 + \boldsymbol{y}^2 \rangle = \\ = \sum_{k=1}^{n} x_k (y_k^1 + y_k^2) = \sum_{k=1}^{n} x_k y_k^1 + \sum_{k=1}^{n} x_k y_k^2 = \langle \boldsymbol{x}, \boldsymbol{y}^1 \rangle + \langle \boldsymbol{x}, \boldsymbol{y}^2 \rangle\,.$$

(c) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \sum\limits_{k=1}^{n} x_k^2 \geqslant 0$; $\sum\limits_{k=1}^{n} x_k^2 = 0 \iff x_1 = \cdots = x_n = 0$.

[1] Обращаю вимание читателя на то, что во всех приводимых ниже формулах верхние индексы у векторов (но не всегда у скаляров!) — это именно индексы, а не показатели степеней.

(d) $|\boldsymbol{x}| = \sqrt{\sum\limits_{k=1}^{n} x_k^2} = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$.

(e) Имеем

$$|\boldsymbol{x}+\boldsymbol{y}|^2 - |\boldsymbol{x}-\boldsymbol{y}|^2 = \langle \boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}\rangle - \langle \boldsymbol{x}-\boldsymbol{y}, \boldsymbol{x}-\boldsymbol{y}\rangle =$$
$$= \langle \boldsymbol{x},\boldsymbol{x}\rangle + \langle \boldsymbol{x},\boldsymbol{y}\rangle + \langle \boldsymbol{y},\boldsymbol{x}\rangle + \langle \boldsymbol{y},\boldsymbol{y}\rangle -$$
$$- \langle \boldsymbol{x},\boldsymbol{x}\rangle - \langle \boldsymbol{y},\boldsymbol{y}\rangle + \langle \boldsymbol{x},\boldsymbol{y}\rangle + \langle \boldsymbol{y},\boldsymbol{x}\rangle = 4\langle \boldsymbol{x},\boldsymbol{y}\rangle . \ \blacktriangleright$$

Определение 2. *Стандартным базисом в* $\mathbb{R}^n$ *называется следующая последовательность векторов*

$$\begin{aligned} \boldsymbol{e}_1 &:= (1, 0, \ldots, 0), \\ \boldsymbol{e}_2 &:= (0, 1, \ldots, 0), \\ &\ldots\ldots\ldots\ldots\ldots\ldots, \\ \boldsymbol{e}_n &:= (0, 0, \ldots, 1). \end{aligned}$$

Каждый вектор $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ допускает единственное представление в виде суммы

$$\boldsymbol{x} = (x_1, 0, \ldots, 0) + (0, x_2, \ldots, 0) + \ldots + (0, 0, \ldots, x_n) =$$
$$= x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_n\boldsymbol{e}_n .$$

Отметим, что для векторов стандартного базиса в $\mathbb{R}^3$ чаще используются другие обозначения

$$\boldsymbol{i} = (1, 0, 0); \qquad \boldsymbol{j} = (0, 1, 0); \qquad \boldsymbol{k} = (0, 0, 1).$$

В этих обозначениях каждый вектор $(x, y, z) \in \mathbb{R}^3$ однозначно представляется в виде

$$(x, y, z) = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k} .$$

Очевидно, что система векторов $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ — *ортонормированная*, т. е. $\langle \boldsymbol{e}_\nu, \boldsymbol{e}_\mu \rangle = 0$ при $\mu \neq \nu$, и $|\boldsymbol{e}_1| = \cdots = |\boldsymbol{e}_n| = 1$.

2. Полнота пространства $\mathbb{R}^n$. Метрика в $\mathbb{R}^n$

Определение 3. *Последовательность*[2]

$$\boldsymbol{x}^1\,,\ \boldsymbol{x}^2\,,\ \ldots\,,\ \boldsymbol{x}^k\,,\ \ldots \tag{14.5}$$

векторов из $\mathbb{R}^n$ *называется сходящейся к вектору* $\boldsymbol{a} \in \mathbb{R}$, *если выполняется условие*

$$\forall \varepsilon > 0 \quad \exists k_\varepsilon \in \mathbb{N} \quad \forall k \geqslant k_\varepsilon \ : \quad |\boldsymbol{x}^k - \boldsymbol{a}| \leqslant \varepsilon\,. \tag{14.6}$$

Обозначается этот факт обычным образом $\lim\limits_{k\to\infty} \boldsymbol{x}^k = \boldsymbol{a}$.

Теорема 3. *Сходимость последовательности векторов* $\left(\boldsymbol{x}^k\right)_{k=1}^{\infty}$ *к вектору* $\boldsymbol{a} = (a_1\,,\ \ldots\,,\ a_n) \in \mathbb{R}^n$ *равносильна покоординатной сходимости, т. е.*

$$\lim_{k\to\infty} \boldsymbol{x}^k = \boldsymbol{a} \quad \Longleftrightarrow \quad \begin{cases} \lim\limits_{k\to\infty} x_1^k = a_1\,; \\ \cdots\cdots\cdots\cdots\,; \\ \lim\limits_{k\to\infty} x_n^k = a_n\,. \end{cases} \tag{14.7}$$

◀ Для любого вектора $\boldsymbol{x} = (x_1\,,\ \ldots\,,\ x_n)$ справедливы следующие неравенства

$$|x_j| \leqslant |\boldsymbol{x}| \leqslant \sum_{j=1}^{n} |x_j|\,. \tag{14.8}$$

В самом деле, разлагая вектор $\boldsymbol{x}$ по векторам стандартного базиса, имеем $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$. Отсюда находим

$$|x_j| = \sqrt{x_j^2} \leqslant \sqrt{\sum_{j=1}^{n} x_j^2} = |\boldsymbol{x}| = \left|\sum_{j=1}^{n} x_j\,\boldsymbol{e}_j\right| \leqslant \sum_{j=1}^{n} |x_j| \cdot |\boldsymbol{e}_j| = \sum_{j=1}^{n} |x_j|\,.$$

Тем самым неравенства (14.8) установлены.

Применяя неравенства (14.8) к вектору $\boldsymbol{x}^k - \boldsymbol{a}$, получим

$$\forall\ j = 1\,,\ \ldots\,,\ n\ :\quad |x_j^k - a_j| \leqslant |\boldsymbol{x}^k - \boldsymbol{a}| \leqslant \sum_{j=1}^{n} |x_j^k - a_j|\,. \tag{14.9}$$

[2]См. сноску на стр. 6

Предположим, что $\lim\limits_{k\to\infty} \boldsymbol{x}^k = \boldsymbol{a}$. Тогда выполнено условие (14.6), и при тех же значениях k из левого неравенства (14.9) имеем $|x_j^k - a_j| \leqslant \varepsilon$, т. е. $\lim\limits_{k\to\infty} x_j^k = a_j$, $j = 1, \ldots, n$.

Обратно, предположим, что имеет место покоординатная сходимость, т. е.

$$\lim_{k\to\infty} x_j^k = a_j\,, \quad \text{для всех} \quad j = 1\,, \ldots, n\,. \tag{14.10}$$

Задавая $\varepsilon > 0$, найдем номер $k_\varepsilon \in \mathbb{N}$ так, чтобы было

$$\forall k \geqslant k_\varepsilon \;\; \forall j = 1\,, \ldots, n \; : \; |x_j^k - a_j| \leqslant \frac{\varepsilon}{n}\,.$$

При тех же значениях k в силу правого неравенства(14.9) имеем

$$|\boldsymbol{x}^k - \boldsymbol{a}| \leqslant \sum_{j=1}^{n} |x_j^k - a_j| \leqslant \sum_{j=1}^{n} \frac{\varepsilon}{n} = \varepsilon\,.$$

Таким образом $\lim\limits_{k\to\infty} \boldsymbol{x}^k = \boldsymbol{a}$. ▶

Теорема 4 (критерий Коши). *Сходимость в $\mathbb{R}^n$ последовательности векторов $\left(\boldsymbol{x}^k\right)_{k=1}^{\infty}$ равносильна выполнению следующего условия:*

$$\forall \varepsilon > 0 \;\; \exists k_\varepsilon \in \mathbb{N} \;\; \forall k\,, j \geqslant k_\varepsilon \; : \; \left|\boldsymbol{x}^k - \boldsymbol{x}^j\right| \leqslant \varepsilon\,. \tag{14.11}$$

◀ Предположим, что условие (14.11) выполнено. В силу неравенства $\left|x_\nu^k - x_\nu^j\right| \leqslant |\boldsymbol{x}^k - \boldsymbol{x}^j|$ имеем

$$\forall\, \nu = 1\,, \ldots, n \;\; \forall\, \varepsilon > 0 \;\; \exists\, k_\varepsilon \in \mathbb{N} \;\; \forall\, k\,, j \geqslant k_\varepsilon \; : \; |x_\nu^k - x_\nu^j| \leqslant \varepsilon\,.$$

В силу критерия Коши в $\mathbb{R}$ сходятся последовательности координат $\left(x_\nu^k\right)_{k=1}^{\infty}$. Отсюда и из предыдущей теоремы следует сходимость последовательности векторов $(\boldsymbol{x}^k)_{k=1}^{\infty}$.

Обратно, предположим, что $\lim\limits_{k\to\infty} \boldsymbol{x}^k = a \in \mathbb{R}^n$. Тогда по предыдущей теореме имеем $\forall\, \nu = 1\,, \ldots, n \; : \; \lim\limits_{k\to\infty} x_\nu^k = a^k$. В силу критерия Коши в $\mathbb{R}$ имеем

$$\forall\, \varepsilon > 0 \;\; \exists\, k_\varepsilon \in \mathbb{N} \;\; \forall\, k\,, j \geqslant k_\varepsilon \; : \; |x_\nu^k - x_\nu^j| \leqslant \frac{\varepsilon}{n}\,.$$

При тех же значениях k и j из правого неравенства (14.9) находим

$$\left|\boldsymbol{x}^k - \boldsymbol{x}^j\right| \leqslant \sum_{\nu=1}^{n} \left|x_\nu^k - a_\nu^k\right| \leqslant \sum_{\nu=1}^{n} \frac{\varepsilon}{n} = \varepsilon\,,$$

т. е. выполнено (14.11). ▶

Определение 4. *Евклидовым расстоянием между точками* $\boldsymbol{x}\,,\ \boldsymbol{y} \in \mathbb{R}^n$ *называется отображение* $d\ :\ \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}_+$*, задаваемое равенством* $d(\boldsymbol{x}\,,\boldsymbol{y}\,) := |\,\boldsymbol{x} - \boldsymbol{y}\,|$.

Теорема 5. *Для любых* $\boldsymbol{x}\,,\ \boldsymbol{y}\,,\ \boldsymbol{z} \in \mathbb{R}^n$ *имеем*
(a) $(\boldsymbol{x}\,,\boldsymbol{y}\,) \geqslant 0$; $\quad d(\boldsymbol{x}\,,\boldsymbol{y}\,) = 0 \iff \boldsymbol{x} = \boldsymbol{y}$ (неотрицательность);
(b) $d(\boldsymbol{x}\,,\boldsymbol{y}\,) = d(\boldsymbol{y}\,,\boldsymbol{x}\,)$ (симметричность);
(c) $d(\boldsymbol{x}\,,\boldsymbol{y}\,) \leqslant d(\boldsymbol{x}\,,\boldsymbol{z}\,) + d(\boldsymbol{z}\,,\boldsymbol{y}\,)$ (неравенство треугольника).

◀ (a) Из теоремы 1 (a) имеем $|\,\boldsymbol{x} - \boldsymbol{y}\,| \geqslant 0$, причем

$$\boldsymbol{x} - \boldsymbol{y} = 0 \iff \boldsymbol{x} - \boldsymbol{y} = \boldsymbol{0}\,, \quad \text{т. е.} \quad \boldsymbol{x} = \boldsymbol{y}\,.$$

(b) Из теоремы 1 (d) при $\lambda = -1$ имеем

$$d(\boldsymbol{x}\,,\boldsymbol{y}\,) = |\,\boldsymbol{x} - \boldsymbol{y}\,| = |\,\boldsymbol{y} - \boldsymbol{x}\,| = d(\boldsymbol{y}\,,\boldsymbol{x}\,)\,.$$

(c) Из неравенства треугольника для нормы имеем

$$d(\boldsymbol{x},\boldsymbol{y}) = |\boldsymbol{x} - \boldsymbol{y}| = |(\boldsymbol{x} - \boldsymbol{z}) + (\boldsymbol{z} - \boldsymbol{y})| \leqslant \\ \leqslant |\boldsymbol{x} - \boldsymbol{z}| + |\boldsymbol{z} - \boldsymbol{y}| = d(\boldsymbol{x},\boldsymbol{z}) + d(\boldsymbol{z},\boldsymbol{y})\,.$$ ▶

Теорема 5 показывает, что пространство $\mathbb{R}^n$ можно рассматривать как метрическое пространство, беря в качестве функции расстояния евклидово расстояние. В силу Теоремы 4 это метрическое пространство является полным.

Как известно, всякое метрическое пространство можно превратить в топологическое, вводя топологию, порожденную метрикой. Таким образом, $\mathbb{R}^n$ можно рассматривать как топологическое пространство. Топологию в $\mathbb{R}^n$, порожденную евклидовой метрикой, будем называть *естественной топологией.* В естественной топологии множество $\mathcal{A} \subset \mathbb{R}^n$ открыто, если любая точка $\boldsymbol{a} \in \mathcal{A}$ содержится в $\mathcal{A}$ вместе с некоторым открытым n-мерным шаром $\left\{\boldsymbol{x} \in \mathbb{R} \;\middle|\; |\boldsymbol{x} - \boldsymbol{a}| < r\right\}$.

3. Компактность в $\mathbb{R}^n$

Определение 5. *Семейство множеств* $\{\mathcal{U}_\alpha \mid \alpha \in I\}$ *называется открытым покрытием множества* $\mathcal{A} \subset \mathbb{R}^n$, *если все* $\mathcal{U}_\alpha$ *открыты в* $\mathbb{R}^n$, *и* $\mathcal{A} \subset \bigcup\limits_{\alpha \in I} \mathcal{U}_\alpha$.

Определение 6. *Множество* $\mathcal{A} \subset \mathbb{R}^n$ *называется компактным, если любое его открытое покрытие содержит конечное подпокрытие.*

Примером компактного множества является любое конечное множество точек. Примером множества, не являющегося компактным, может служить все пространство $\mathbb{R}^n$. В самом деле, для каждого $n \in \mathbb{N}$ возьмем открытый шар $\mathcal{U}_n := \{\boldsymbol{x} \in \mathbb{R}^n \mid |\boldsymbol{x}| < n\}$. Семейство шаров $\{\mathcal{U}_n \mid n \in \mathbb{N}\}$ есть открытое покрытие пространства $\mathbb{R}^n$, не содержащее конечного подпокрытия. Ранее был установлен критерий компактности в пространстве $\mathbb{R}$. Аналогичный критерий имеет место и в $\mathbb{R}^n$. Наша ближайшая цель — установить этот критерий.

Определение 7. *Замкнутым n-мерным параллелепипедом (брусом) называется множество вида*

$$\Pi = [a_1\,, b_1] \times [a_2\,, b_2] \times \cdots \times [a_n\,, b_n]\,;$$

открытым n-мерным параллелепипедом называется множество вида

$$\Pi^o = (a_1\,, b_1) \times (a_2\,, b_2) \times \cdots \times (a_n\,, b_n),$$

где $a_k < b_k$, $k = 1\,, 2\,, \ldots\,, n$.

Можно показать, что замкнутый брус является замкнутым множеством в $\mathbb{R}^n$, а открытый — открытым.

Определение 8. *Множество* $\mathcal{A} \subset \mathbb{R}^n$ *называется ограниченным, если существует замкнутый n-мерный параллелепипед* Π *такой, что* $\mathcal{A} \subset \Pi$ *или, что равносильно,* $\exists K \in \mathbb{R}_+ \ \forall \boldsymbol{x} \in \mathcal{A} : |\boldsymbol{x}| \leqslant K$.

Эта равносильность связана с тем, что любой параллелепипед содержится в некотором шаре, который в свою очередь содержится в некотором параллепипеде (рис. 2).

Теорема 6. *Предположения:* $\boldsymbol{x}_0 \in \mathbb{R}^n$; $B \subset \mathbb{R}^n$; B *— компактное множество;* $\{\mathcal{W}_\alpha \subset \mathbb{R}^{n+m} \mid \alpha \in I\}$ *— открытое покрытие множества* $\{\boldsymbol{x}_0\} \times B$*. Заключения: множество* $\{\boldsymbol{x}_0\} \times B$ *— компактное; существует открытое множество* $\mathcal{U} \subset \mathbb{R}^n$*, такое, что* $\boldsymbol{x}_0 \in \mathcal{U}$*, и* $\mathcal{U} \times B$ *покрывается конечным подсемейством семейства* $\{\mathcal{W}_\alpha \mid \alpha \in I\}$.

◀ Так как множество $\mathcal{W}_\alpha \subset \mathbb{R}^{n+m}$ открыто, то для любой точки $(\boldsymbol{x}_0, \boldsymbol{y}) \in \{\boldsymbol{x}_0\} \times B$ существуют открытые брусы $\mathcal{U}_{\boldsymbol{y}} \subset \mathbb{R}^n$ и $\mathcal{V}_{\boldsymbol{y}} \subset \mathbb{R}^m$, такие, что $(\boldsymbol{x}_0, \boldsymbol{y}) \subset \mathcal{U}_{\boldsymbol{y}} \times \mathcal{V}_{\boldsymbol{y}} \subset \mathcal{W}_\alpha$ при некотором α. Семейство открытых брусов $\{\mathcal{V}_{\boldsymbol{y}} \mid \boldsymbol{y} \in B\}$ есть, очевидно, открытое покрытие множества B. Так как B компактно, то существует конечное подпокрытие $\{\mathcal{V}_{\boldsymbol{y}_1}, \mathcal{V}_{\boldsymbol{y}_2}, \ldots, \mathcal{V}_{\boldsymbol{y}_k}\}$ множества B. Беря для каждого $\nu = 1, 2, \ldots, k$ соответствующие α_ν так, чтобы было $\mathcal{U}_{\boldsymbol{y}_\nu} \times \mathcal{V}_{\boldsymbol{y}_\nu} \subset \mathcal{W}_{\alpha_\nu}$, получим

$$\{\boldsymbol{x}_0\} \times B \subset \bigcup_{\nu=1}^{k} \{\boldsymbol{x}_0\} \times \mathcal{V}_{\boldsymbol{y}_\nu} \subset \bigcup_{\nu=1}^{k} \mathcal{U}_{\boldsymbol{y}_\nu} \times \mathcal{V}_{\boldsymbol{y}_\nu} \subset \bigcup_{\nu=1}^{k} \mathcal{W}_{\alpha_\nu},$$

т. е. конечное семейство $\{\mathcal{W}_{\alpha_1}, \mathcal{W}_{\alpha_2}, \ldots, \mathcal{W}_{\alpha_k}\}$ покрывает множество $\{\boldsymbol{x}_0\} \times B$. Поэтому множество $\{\boldsymbol{x}_0\} \times B$ — компактное. Полагая $\mathcal{U} := \mathcal{U}_{\boldsymbol{y}_1} \bigcap \mathcal{U}_{\boldsymbol{y}_2} \bigcap \ldots \bigcap \mathcal{U}_{\boldsymbol{y}_k}$, заключаем, что множество $\mathcal{U}$ — открытое, причем

$$\mathcal{U} \times B \subset \bigcup_{\nu=1}^{k} \mathcal{U} \times \mathcal{V}_{\boldsymbol{y}_\nu} \subset \bigcup_{\nu=1}^{k} \mathcal{U}_{\boldsymbol{y}_\nu} \times \mathcal{V}_{\boldsymbol{y}_\nu} \subset \bigcup_{\nu=1}^{k} \mathcal{W}_{\alpha_\nu}.$$

Таким образом, конечное семейство $\{\mathcal{W}_{\alpha_1}, \mathcal{W}_{\alpha_2}, \ldots, \mathcal{W}_{\alpha_k}\}$ покрывает множество $\mathcal{U} \times B$. ▶

Теорема 7. *Если все множества* $A_1, \ldots, A_n$ *компактны, то их декартово произведение* $A_1 \times \cdots \times A_n$ *компактно. В частности, любой замкнутый параллепипед* $[a_1, b_1] \times \cdots \times [a_n, b_n]$ *компактен.*

◀ Рассмотрим сначала случай декартового произведения двух компактных множеств $A_1 \times A_2$, где $A_1 \subset \mathbb{R}^n$, $A_2 \subset \mathbb{R}^m$. Пусть $\{W_\alpha \mid \alpha \in I\}$ — открытое в $\mathbb{R}^{n+m}$ покрытие множества $A_1 \times A_2$. В силу теоремы 6 для любого $\boldsymbol{x} \in A_1$ существует открытое множество $U_{\boldsymbol{x}} \subset \mathbb{R}^n$, такое, что $\boldsymbol{x} \in U_{\boldsymbol{x}}$, а $U_{\boldsymbol{x}} \times A_2$ покрывается конечным числом множеств W_α. Семейство $\{U_{\boldsymbol{x}} \mid \boldsymbol{x} \in A_1\}$ есть открытое покрытие множества A_1. Так как A_1 компактно, то это покрытие содержит конечное подпокрытие $\{U\boldsymbol{x}_1, \ldots, U\boldsymbol{x}_k\}$, и значит, семейство $\{U\boldsymbol{x}_1 \times A_2, \ldots, U\boldsymbol{x}_k \times A_2\}$ есть конечное покрытие множества $A_1 \times A_2$. Так как каждое $U\boldsymbol{x}_\nu \times A_2$ покрывается конечным числом множеств W_α, то и $A_1 \times A_2$ покрывается конечным числом множеств W_α. Значит, $A_1 \times A_2$ — компактное множество. Предположим, что $A_1 \times \cdots \times A_{n-1}$ — компактное множество. Представляя $A_1 \times \cdots \times A_n$ в виде $A_1 \times \cdots \times A_n = (A_1 \times \cdots \times A_{n-1}) \times A_n$ и применяя доказанную часть теоремы, заключаем, что множество $A_1 \times \cdots \times A_n$ компактное.

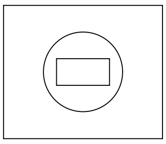

Рис. 2. К определению 8

Учитывая, что замкнутые интервалы $[a_\nu, b_\nu]$ в силу леммы Гейне — Бореля «о покрытиях» компактны, заключаем, что замкнутый брус $[a_1, b_1] \times \cdots \times [a_n, b_n]$ компактен как декартово произведение компактных множеств. ▶

Теорема 8 (критерий компактности в $\mathbb{R}^n$). *Множество $A \subset \mathbb{R}^n$ компактно тогда и только тогда, когда оно замкнуто и ограничено.*

◀ Предположим, что множество A замкнуто и ограничено. Тогда $\mathbb{R}^n \setminus A$ открытое, и существует замкнутый брус $\Pi \subset \mathbb{R}^n$, такой, что $A \subset \Pi$. Пусть $\{W_\alpha \mid \alpha \in I\}$ — открытое покрытие множества A. Семейство $\{W_\alpha \mid \alpha \in I\}$ вместе с открытым множеством $\mathbb{R}^n \setminus A$ есть открытое покрытие пространства $\mathbb{R}^n$, а значит, и бруса Π. Так как по теореме 7 брус Π компактен, то это его покрытие содержит конечное подпокрытие $\{W_{\alpha_1}, \ldots, W_{\alpha_k}, \mathbb{R}^n \setminus A\}$. Поскольку $(\mathbb{R}^n \setminus A) \bigcap A = \varnothing$, то семейство $\{W_{\alpha_1}, \ldots, W_{\alpha_k}\}$ есть конечное подпокрытие множества A. Таким образом, множество A компактное.

Обратно, пусть множество A — компактное. Предполагая, что оно не ограниченное, заключаем, что его открытое покрытие $\{U_k \mid k \in \mathbb{N}\}$, где $U_k = \{\boldsymbol{x} \in \mathbb{R}^n \mid |\boldsymbol{x}| < k\}$, не содержит конечного подпокрытия. В самом деле, объединение любого конечного семейства шаров $\{|\boldsymbol{x}| < k\}$ есть шар наибольшего радиуса, т. е. ограниченное множество.

Предполагая, что множество A не замкнутое, заключаем, что существует граничная точка $\boldsymbol{x}_0$ множества A, такая, что $\boldsymbol{x}_0 \notin A$. Полагая $U_k = \{\boldsymbol{x} \in \mathbb{R}^n \mid |\boldsymbol{x} - \boldsymbol{x}_0| > 1/k\}$, заключаем, что семейство $\{U_k \mid k \in \mathbb{N}\}$ есть открытое покрытие множества A, не содержащее конечного подпокрытия. В самом деле, никакое конечное семейство открытых множеств U_k не покрывает некоторую окрестность точки $\boldsymbol{x}_0$. А так как в любой окрестности точки $\boldsymbol{x}_0$ имеются точки множества A, то это конечное семейство не покрывает и множество A. Получено противоречие. ►

§ 2. Отображения из $\mathbb{R}^n$ в $\mathbb{R}^m$ и их пределы

1. Способы задания отображений из $\mathbb{R}^n$ в $\mathbb{R}^m$

Определение 9. *Отображениями из* $\mathbb{R}^n$ *в* $\mathbb{R}^m$ *условимся называть отображения следующего вида:*

$$f\,:\,\mathcal{A}\longrightarrow\mathbb{R}^m\,,\ \mathcal{A}\subset\mathbb{R}^m\,. \tag{14.12}$$

В зависимости от величин размерностей m и n отображения вида (14.12) можно подразделить следующим образом.

1) При $m=n=1$ отображение (14.12) есть скалярная функция одного скалярного аргумента. Такие функции до сих пор, в основном, нами и изучались.

2) В случае $n>1$, $m=1$ отображение (14.12) есть скалярная функция векторного аргумента (или вещественнозначная функция от n вещественных переменных).

3) В случае $n=1$, $m>1$ отображение (14.12) есть векторная функция скалярного аргумента.

4) При $n>1$, $m>1$ отображение (14.12) есть векторнозначная функция векторного аргумента.

Способы задания отображений из $\mathbb{R}^n$ в $\mathbb{R}^m$, вообще говоря, — те же, что и способы задания функций одного вещественного переменного. Именно: *аналитический* (явный, неявный, параметрический), *графический, табличный, компьютерный* и другие. Разумеется, все эти способы имеют свои особенности по сравнению со способами задания скалярных функций скалярного аргумента.

Явный аналитический способ задания отображения (14.12) предполагает наличие уравнения

$$\boldsymbol{y}=f(\boldsymbol{x})\,, \tag{14.13}$$

которое позволяет по каждому значению $\boldsymbol{x} \in \mathcal{A}$ вычислить соответствующее значение $\boldsymbol{y}$.

Желая переписать уравнение (14.13) в виде равносильной ему системы скалярных уравнений, введем понятие *отображения проектирования.*

Определение 10. *Отображением проектирования на* ν*-ю координату называется отображение* $\pi^\nu \ : \ \mathbb{R}^m \longrightarrow \mathbb{R}$*, для которого*

$$\pi^\nu \ : \ \boldsymbol{y} \longmapsto y_\nu \quad \text{\textit{или}} \quad \pi^\nu(\boldsymbol{y}) = y_\nu \,,$$

где $\boldsymbol{y} = (y_1 \,, \, \dots \,, \, y_m) \in \mathbb{R}^m$.

Действуя на обе части равенства (14.13) отображенями проектирования $\pi^1 \,, \, \dots \,, \, \pi^m$, получим

$$\begin{cases} y_1 = \pi^1(f(\boldsymbol{x}))\,, \\ \dots\dots\dots\dots\,, \\ y^m = \pi^m(f(\boldsymbol{x}))\,. \end{cases} \tag{14.14}$$

Определение 11. *Координатными функциями* f^ν *отображения* f *называются композиции* $f^\nu := \pi^\nu \circ f$, $\nu = 1 \,, \, \dots \,, \, m$, *где* π^ν *— отображения проектирования.*

Учитывая это определение и то, что $\boldsymbol{x} = (x_1 \,, \, \dots \,, \, x_n)$, перепишем систему (14.14) в следующем равносильном виде:

$$\begin{cases} y_1 = f^1(x_1 \,, \, \dots \,, \, x_n)\,, \\ \dots\dots\dots\dots\dots\dots\,, \\ y_m = f^m(x_1 \,, \, \dots \,, \, x_n)\,. \end{cases} \tag{14.15}$$

Пусть $\boldsymbol{e}'_1 \,, \, \dots \,, \, \boldsymbol{e}'_m$ — стандартный базис в $\mathbb{R}^m$. Учитывая, что

$$\boldsymbol{y} = (y_1 \,, \, \dots \,, \, y_m) = y_1 \boldsymbol{e}'_1 + \dots + y_m \boldsymbol{e}'_m = \pi^1(\boldsymbol{y})\boldsymbol{e}'_1 + \dots + \pi^m(\boldsymbol{y})\boldsymbol{e}'_m \,,$$

умножая первое равенство (14.14) на $\boldsymbol{e}'_1$, ... , умножая m-е равенство (14.14) на $\boldsymbol{e}'_m$ и складывая, получим равенство (14.13). Таким образом, все записи (14.13), (14.14), (14.15) эквивалентны.

Определение 12. *Графиком отображения* $f \ : \ \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^m$ *, называется множество* $\Gamma_f \subset \mathbb{R}^{m+n}$*, задаваемое равенством*

$$\Gamma_f := \left\{ (\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^{m+n} \ \middle| \ \boldsymbol{x} \in \mathcal{A} \,, \ \boldsymbol{y} = f(\boldsymbol{x}) \right\} \,. \tag{14.16}$$

Так как наглядное геометрическое изображение допускают только пространства $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^3$, то и график отображения из $\mathbb{R}^n$ в $\mathbb{R}^m$ допукает наглядное геометрическое изображение только при $2 \leqslant n + m \leqslant 3$. Кроме изученного ранее случая $n = m = 1$ наиболее часто используется наглядное геометрическое изображение в случае $n = 2$, $m = 1$. В этом случае мы имеем дело со *скалярной функцией двух вещественных переменных* $z = f(x, y)$, где $(x, y) \in \mathcal{A} \subset \mathbb{R}^2$, $z \in \mathbb{R}^1$. Графиком такой функции является множество

$$\Gamma_f = \left\{(x, y, z) \in \mathbb{R}^3 \;\middle|\; (x, y) \in \mathcal{A},\, z = f(x, y)\right\} . \tag{14.17}$$

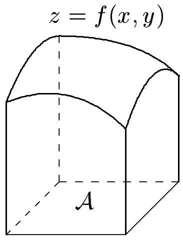

Рис. 3. График функции двух переменных

Если множество $\mathcal{A}$ — открытое и связное (т. е. *область*), а функция f — «достаточно хорошая», то график (14.17) можно изобразить в виде куска поверхности («крыши»), которая биективно проектируется на множество $\mathcal{A}$ (см. рис. 3).

Например, графиками функций

$$z = x^2 + y^2\,, \quad z = x \cdot y\,, \quad z = \sqrt{1 - x^2 - y^2}$$

являются соответствующие части поверхностей 2-го порядка (см. рис. 4).

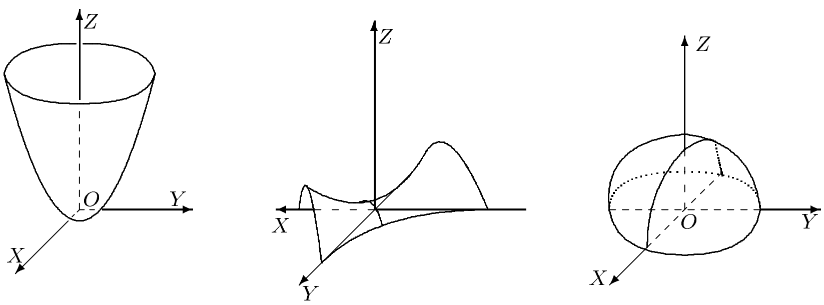

Рис. 4. Графики некоторых поверхностей 2-го порядка

2. Пределы отображений из $\mathbb{R}^n$ в $\mathbb{R}^m$

Пусть $f : \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, — отображение, и пусть $\boldsymbol{a} \in \mathbb{R}^n$ — точка прикосновения[3] множества $\mathcal{A}$.

Определение 13. *Вектор $\boldsymbol{b} \in \mathbb{R}^m$ называется пределом отображения f при $\boldsymbol{x} \to \boldsymbol{a}$, $\boldsymbol{x} \in \mathcal{A}$, если*

$$\forall \varepsilon \in \mathbb{R}_+ \; \exists \delta \in \mathbb{R}_+ \; \forall \boldsymbol{x} \in \mathcal{A} \; : \; |\boldsymbol{x} - \boldsymbol{a}| \leqslant \delta \implies |f(\boldsymbol{x}) - \boldsymbol{b}| \leqslant \varepsilon \, . \tag{14.18}$$

Записывают этот факт в виде $\lim\limits_{\substack{\boldsymbol{x} \to \boldsymbol{a}, \\ \boldsymbol{x} \in \mathcal{A}}} f(\boldsymbol{x}) = \boldsymbol{b}$ или $\lim\limits_{\boldsymbol{x} \to \boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{b}$, когда ясно, о каком множестве $\mathcal{A}$ идет речь. Определение 13 по форме не отличается от соответствующего определения предела функции одного переменного. Единственное отличие заключается в том, что в условии (14.18) символ $|\boldsymbol{z}|$ означает норму (длину) вектора $\boldsymbol{z}$ именно в том пространстве, которому принадлежит вектор $\boldsymbol{z}$.

Так же, как и в одномерном случае, условие (14.18) равносильно следующему условию:

$$\forall \mathcal{V}(\boldsymbol{b}) \; \exists \mathcal{U}(\boldsymbol{a}) \; : \; f(\mathcal{A} \cap \mathcal{U}(\boldsymbol{a})) \subset \mathcal{V}(\boldsymbol{b}) \, , \tag{14.19}$$

где $\mathcal{U}(\boldsymbol{a})$ и $\mathcal{V}(\boldsymbol{b})$ — окрестности точек $\boldsymbol{a}$ и $\boldsymbol{b}$ соответственно. В связи с этим определение предела отображения в форме (14.19) называется иногда определением «на языке окрестностей» в отличие от определения предела в форме (14.18), которое называется определением «на языке ε—δ». Напомним, что окрестности в (14.19) понимаются в смысле соответствующего пространства.

Замечания. 1. Из определения 13 вытекает, что *если существует предел $\lim f(\boldsymbol{x}) = \boldsymbol{b}$ при $\boldsymbol{x} \to \boldsymbol{a}$, $\boldsymbol{x} \in \mathcal{A}$, то существует и предел $\lim f(\boldsymbol{x}) = \boldsymbol{b}$ при $\boldsymbol{x} \to \boldsymbol{a}$, $\boldsymbol{x} \in \mathcal{B}$ для любого множества $\mathcal{B} \subset \mathcal{A}$, лишь бы $\boldsymbol{a}$ была точкой прикосновения множества $\mathcal{B}$.*

2. Из предыдущего замечания следует, что если для некоторого множества $\mathcal{B} \subset \mathcal{A}$ предел $\lim\limits_{\substack{\boldsymbol{x} \to \boldsymbol{a}, \\ \boldsymbol{x} \in \mathcal{B}}} f(\boldsymbol{x})$ не существует, то не существует и предел $\lim\limits_{\substack{\boldsymbol{x} \to \boldsymbol{a}, \\ \boldsymbol{x} \in \mathcal{A}}} f(\boldsymbol{x})$. Кроме того, если для двух подмножеств $\mathcal{B}_1$ и $\mathcal{B}_2$ множества $\mathcal{A}$ пределы $\lim\limits_{\substack{\boldsymbol{x} \to \boldsymbol{a}, \\ \boldsymbol{x} \in \mathcal{B}_1}} f(\boldsymbol{x})$ и $\lim\limits_{\substack{\boldsymbol{x} \to \boldsymbol{a}, \\ \boldsymbol{x} \in \mathcal{B}_2}} f(\boldsymbol{x})$, хотя и существуют, но различны, то предел $\lim\limits_{\substack{\boldsymbol{x} \to \boldsymbol{a}, \\ \boldsymbol{x} \in \mathcal{A}}} f(\boldsymbol{x})$ не сушествует.

---

[3] Это, напоминаю, означает, что $\mathcal{U}(\boldsymbol{a}) \cap \mathcal{A} \neq \varnothing$ для любой окрестности $\mathcal{U}(\boldsymbol{a}) \ni \boldsymbol{a}$.

Например, из существования предела $\lim\limits_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{b}$ следует существование того же самого предела «по любому направлению», т. е. предела $\lim\limits_{t\to+0} f(\boldsymbol{a}+t\boldsymbol{l})$ для любого ненулевого вектора $\boldsymbol{l}\in\mathbb{R}^n$. Отсюда вытекает, что если хотя бы по одному направлению $\boldsymbol{l}$ предел $\lim\limits_{t\to+0} f(\boldsymbol{a}+t\boldsymbol{l})$ не существует, или если пределы по двум различным направлениям $\boldsymbol{l}_1\neq\boldsymbol{l}_2$ различны, т. е. $\lim\limits_{t\to+0} f(\boldsymbol{a}+t\boldsymbol{l}_1)\neq\lim\limits_{t\to+0} f(\boldsymbol{a}+t\boldsymbol{l}_2)$, то предел $\lim\limits_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x})$ не существует.

Рассмотрим более конкретный пример. Пусть требуется установить, существует ли предел

$$\lim_{(x,y)\to(0,0)}\frac{xy}{x^2+y^2}\,. \tag{14.20}$$

Переходя к полярным координатам $x=r\cos\varphi$, $y=r\sin\varphi$, при фиксированном φ имеем

$$\lim_{(x,y)\to(0,0)}\frac{xy}{x^2+y^2}=\lim_{r\to+0}\cos\varphi\sin\varphi=\cos\varphi\sin\varphi\,.$$

Так как полученный ответ зависит от φ, то предел (14.20) не существует.

Вычисление пределов векторнозначных фукций векторного аргумента можно сводить к вычислению пределов скалярнозначных функций векторного аргумента на основании следующей теоремы.

Теорема 9. *Существование предела*

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x})=\boldsymbol{b}=(b_1\,,\,\dots\,,\,b_m)\in\mathbb{R}^m \tag{14.21}$$

векторнозначного отображения $f=(f^1\,,\,\dots\,,\,f^m)$ *равносильно существованию пределов*

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} f^k(\boldsymbol{x})=b_k\,,\ \ k=1\,,\,\dots\,,\,m\,, \tag{14.22}$$

всех координатных функций f^k *этого отображения.*

◄ При $k=1\,,\,\dots\,,\,m$ справедливы следующие неравенства:

$$\left|f^k(\boldsymbol{x})-b_k\right|\leqslant|f(\boldsymbol{x})-\boldsymbol{b}|\leqslant\sum_{k=1}^{m}\left|f^k(\boldsymbol{x})-b_k\right|\,. \tag{14.23}$$

Если существует предел (14.21), то

$$\forall\varepsilon\in\mathbb{R}_+\ \ \exists\delta\in\mathbb{R}_+\ \ \forall\boldsymbol{x}\in\mathcal{A}\ :\ \ |\boldsymbol{x}-\boldsymbol{a}|\leqslant\delta\Longrightarrow|f(\boldsymbol{x})-\boldsymbol{b}|\leqslant\varepsilon\,.$$

В силу левого неравенства (14.23) при тех же значениях переменного $\boldsymbol{x}$ будет $\left|f^k(\boldsymbol{x})-b_k\right| \leqslant \varepsilon$, т. е. $\lim\limits_{\boldsymbol{x}\to\boldsymbol{a}} f^k(\boldsymbol{x}) = b_k\,,\ k = 1\,,\,\ldots\,,\,m$. Обратно, если существуют пределы (14.22), то

$$\forall\varepsilon\in\mathbb{R}_+\ \exists\delta\in\mathbb{R}_+\ \forall\boldsymbol{x}\in\mathcal{A}\ :\ \ |\boldsymbol{x}-\boldsymbol{a}|\leqslant\delta \implies \left|f^k(\boldsymbol{x})-b_k\right|\leqslant\frac{\varepsilon}{m}\,. \tag{14.24}$$

Отсюда в силу правого неравенства (14.23) при тех же значениях переменного $\boldsymbol{x}$ имеем

$$|f(\boldsymbol{x})-\boldsymbol{b}|\leqslant\sum_{k=1}^{m}\left|f^k(\boldsymbol{x})-b_k\right|\leqslant\sum_{k=1}^{m}\frac{\varepsilon}{m}=\varepsilon\,.\ \blacktriangleright$$

Замечания. 1. Для скалярных функций векторного аргумента основные факты теории пределов и их обоснование практически не отличается от соответствующих фактов теории пределов скалярных функций скалярного аргумента. Перечислим, о каких фактах идет речь (конкретные формулировки и доказательства их оставляем читателю). Речь идет о следующих теоремах:

1) о пределе постоянной функции;

2) о единственности предела;

3) о финальной ограниченности функции, имеющей конечный предлел;

4) о связи предельного перехода с отношениями неравенства между функциями (в чистности, теорема о пределе «зажатой» функции);

5) о «бесконечно малых» и «бесконечно больших» функциях;

6) о пределе композиции функций;

7) критерий Коши существования предела функции.

2. Отметим, что так как пространства $\mathbb{R}^n$ при $n>1$ не являются линейно упорядоченными, то для функций нескольких переменных не имеет смысла понятие монотонности. Поэтому и теоремы о пределах монотонных функций на функции нескольких переменных, естественно, не распространяются.

3. Повторные пределы

Пусть $\boldsymbol{x} = (x_1\,, \ldots, x_n)$, $\boldsymbol{a} = (a_1\,, \ldots, a_n)$. В связи с тем, что имеет место равносильность

$$\boldsymbol{x} \to \boldsymbol{a} \quad \Longleftrightarrow \quad \begin{cases} x_1 \to a_1\,, \\ \ldots\ldots\,, \\ x_n \to a_n\,, \end{cases}$$

предел $\lim\limits_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x})$ иногда называют «кратным» пределом. При $n > 1$ наряду с кратными пределами можно рассматривать так называемые «повторные» пределы, например, следующие:

$$\lim_{x_n\to a_n}\left(\ldots\left(\lim_{x_2\to a_2}\left(\lim_{x_1\to a_1} f(x_1\,, x_2\,, \ldots, x_n)\right)\right)\ldots\right)$$

или

$$\lim_{x_1\to a_1}\left(\lim_{x_2\to a_2}\left(\ldots\left(\lim_{x_n\to a_n} f(x_1\,, x_2\,, \ldots, x_n)\right)\ldots\right)\right).$$

Появление самого́ понятия повторного предела — это новое качество, которое возникает при переходе от функций одного переменного к функциям нескольких переменных. Изучение простейших связей, существующих между кратными и повторными пределами, составляет содержание настоящего пункта. Ограничиваясь здесь только случаем функций двух переменных, рассмотрим сначала **примеры**.

1) Как было установлено выше, двойной предел $\lim\limits_{(x,y)\to(0,0)} \dfrac{xy}{x^2+y^2}$ не существует. Однако

$$\lim_{\substack{y\to 0,\\ y\neq 0}}\left(\lim_{\substack{x\to 0,\\ x\neq 0}} \frac{xy}{x^2+y^2}\right) = \lim_{\substack{y\to 0,\\ y\neq 0}} 0 = 0\,; \quad \lim_{\substack{x\to 0,\\ x\neq 0}}\left(\lim_{\substack{y\to 0,\\ y\neq 0}} \frac{xy}{x^2+y^2}\right) = \lim_{\substack{x\to 0,\\ x\neq 0}} 0 = 0\,.$$

Таким образом, двойной предел не существует, а оба повторных предела существуют и равны между собой.

2) Функция $f\ :\ (x\,,\,y) \longmapsto x\sin\dfrac{1}{y} + y\sin\dfrac{1}{x}$ определена на множестве $\mathcal{A} := \mathbb{R}^2 \setminus \{x\cdot y = 0\}$. В силу очевидных неравенств

$$0 \leqslant \left|x\sin\frac{1}{y} + y\sin\frac{1}{x}\right| \leqslant |x| + |y|$$

и учитывая, что

$$\lim_{(x,y)\to(0,0)} (|x|+|y|) = \lim_{r\to 0} r\cdot(|\cos\varphi|+|\sin\varphi|) = (|\cos\varphi|+|\sin\varphi|)\lim_{r\to 0} r = 0\,,$$

где $x = r\cos\varphi$, $y = r\sin\varphi$, имеем

$$\lim_{\substack{(x,y)\to(0,0),\\ (x,y)\in\mathcal{A}}} \left(x\sin\frac{1}{y} + y\sin\frac{1}{x}\right) = 0\,,$$

т. е. двойной предел существует. Посмотрим теперь повторные пределы. Так как предел $\lim\limits_{x\to 0}\sin\dfrac{1}{x}$ не существует, то и предел $\lim\limits_{x\to 0}\left(y\sin\dfrac{1}{x} + x\sin\dfrac{1}{y}\right)$ тоже не существует. Значит, не существует и повторный предел

$$\lim_{y\to 0}\left(\lim_{x\to 0}\left(y\sin\frac{1}{x} + x\sin\frac{1}{y}\right)\right)\,.$$

Так как $f(x\,,\,y) \equiv f(y\,,\,x)$, то и другой повторный предел не существует.

Эти примеры показывают, что из существования равных между собой повторных пределов не следует существование соответствующего кратного предела. С другой стороны, существование кратного предела отнюдь не гарантирует существования повторных пределов.

Однако при некоторых ограничениях связь между кратным и повторными пределами существует, как показывает следующая теорема.

Теорема 10 (о повторных пределах). *Предположим, что функция* $f\ :\ \mathcal{A}\longrightarrow\mathbb{R}$ *задана на множестве* $\mathcal{A}$*, представляющем собой окрестность точки* $(a\,,\,b)$*, из которой выброшены линии* $x = a$ *и* $y = b$*, и пусть существует двойной предел*

$$\lim f(x\,,\,y) = d\in\mathbb{R}\quad \textit{при}\quad (x,y)\to(a,b)\,,\,(x,y)\in\mathcal{A}\,.$$

Если

$$\forall\, y\neq b\ \ \exists \lim_{\substack{x\to a,\\ x\neq a}} f(x\,,\,y) = g(y)\in\mathbb{R}\,,\quad \textit{то}\quad \lim_{\substack{y\to b,\\ y\neq b}}\left(\lim_{\substack{x\to a,\\ x\neq a}} f(x\,,\,y) = d\right)\,.$$

◀ Зададим $\varepsilon\in\mathbb{R}_+$. В силу существования двойного предела существует $\delta\in\mathbb{R}_+$, такое, что

$$\forall\,(x\,,\,y)\in\mathcal{A}\ :\quad \left.\begin{array}{l}|x-a|\leqslant\delta\\ |y-b|\leqslant\delta\end{array}\right\}\quad\Longrightarrow\quad |f(x\,,\,y)-d|\leqslant\varepsilon\,.$$

Переходя здесь к пределу при $x \to a$, $x \neq a$, получим

$$0 < |y - b| \leqslant \delta \quad \Longrightarrow \quad |g(y) - d| \leqslant \varepsilon\,.$$

Последнее означает, что $\lim\limits_{\substack{y\to b,\\ y\neq b}} g(y) = d$ или, что равносильно,

$$\lim_{\substack{y\to b,\\ y\neq b}} \left(\lim_{\substack{x\to a,\\ x\neq a}} f(x\,,\, y) \right) = d. \quad \blacktriangleright$$

Пример. Функция $f\ :\ (x\,,\, y) \longmapsto x \sin\dfrac{1}{y}$ определена при $y \neq 0$. Так как $0 \leqslant \left|x \sin\dfrac{1}{y}\right| \leqslant |x|$, то для нее имеем $\lim\limits_{\substack{(x,y)\to(0,0),\\ y\neq 0}} x \sin\dfrac{1}{y} = 0$. Кроме того, при $y \neq 0$, очевидно, имеем $\lim\limits_{x\to 0} x \sin\dfrac{1}{y} = \sin\dfrac{1}{y} \cdot \lim\limits_{x\to 0} x = 0$. Значит,

$$\lim_{\substack{(x,y)\to(0,0),\\ y\neq 0}} x \sin\frac{1}{y} = \lim_{\substack{y\to 0,\\ y\neq 0}} \left(\lim_{x\to 0} x \sin\frac{1}{y} \right) = 0\,,$$

что согласуется с теоремой 10. Однако второй повторный предел, очевидно, не существует, что также не противоречит теореме 10.

§ 3. Непрерывность отображений из $\mathbb{R}^n$ в $\mathbb{R}^m$

1. Определение и локальные свойства

Понятие непрерывности отображений из $\mathbb{R}^n$ в $\mathbb{R}^m$ аналогично понятию непрерывности функций одного переменного. Понятие непрерывности отображения в точке может быть сведено к понятию его предела в этой точке.

Определение 14. *Отображение* $f\ :\ \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, *называется непрерывным в точке* $\boldsymbol{a} \in \mathcal{A}$, *если в* $\mathbb{R}^m$ *существует предел*

$$\lim_{\substack{\boldsymbol{x}\to\boldsymbol{a},\\ \boldsymbol{x}\in\mathcal{A}}} f(\boldsymbol{x}) \in \mathbb{R}^m. \tag{14.25}$$

В силу единственности предела непрерывное в точке $\boldsymbol{a} \in \mathcal{A}$ отображение f не может иметь иного предела, кроме $f(\boldsymbol{a})$, т. е.

$$\lim_{\substack{\boldsymbol{x}\to\boldsymbol{a},\\ \boldsymbol{x}\in\mathcal{A}}} f(\boldsymbol{x}) = f(\boldsymbol{a}).$$

Если $\boldsymbol{a}$ является изолированной точкой множества $\mathcal{A}$, то последнее равенство выполняется *для любого отображения* $f : \mathcal{A} \longrightarrow \mathbb{R}^m$. Поэтому *любое отображение* $f : \mathcal{A} \longrightarrow \mathbb{R}^m$ *является непрерывным в изолированной точке* $\boldsymbol{a} \in \mathcal{A}$. Отбрасывая случай изолированной точки как мало интересный, дадим равносильное определению 14 определение непрерывности отображения f *в предельной точке* $\boldsymbol{a} \in \mathcal{A}$.

Определение 15. *Отображение* $f : \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, *называется непрерывным в предельной точке* $\boldsymbol{a} \in \mathcal{A}$, *если выполняется равенство*

$$\lim_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in\mathcal{A}}} f(\boldsymbol{x}) = f(\boldsymbol{a}). \tag{14.26}$$

Перепишем условие непрерывности (14.26) «на языке окрестностей»

$$\forall\, \mathcal{V}(f(\boldsymbol{a}))\ \ \exists\,\mathcal{U}(\boldsymbol{a})\ :\ f\left(\dot{\mathcal{U}}(\boldsymbol{a}) \cap \mathcal{A}\right) \subset \mathcal{V}\,(f(\boldsymbol{a}))\,, \tag{14.27}$$

где обозначено $\dot{\mathcal{U}}(\boldsymbol{a}) := \mathcal{U} \setminus \{\boldsymbol{a}\}$ — проколотая окрестность точки $\boldsymbol{a}$. То же самое «на языке» $\varepsilon{-}\delta$ можно переписать следующим образом:

$$\forall\varepsilon \in \mathbb{R}_+\ \exists\delta \in \mathbb{R}_+\ :\ 0 < |x - a| \leqslant \delta \Longrightarrow |f(\boldsymbol{x}) - f(\boldsymbol{a})| \leqslant \varepsilon\,.$$

Отсюда вытекает еще одна равносильная форма определения непрерывности функции в предельной точке ее области определения: *непрерывность функции в точке равносильна тому, что бесконечно малому приращению ее аргумента в этой точке соответствует бесконечно малое приращение функции.*

Теорема 11 (о непрерывности композиции). *Пусть отображение* $f : \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, *непрерывно в точке* $\boldsymbol{a} \in \mathcal{A}$, *а отображение* $g : \mathcal{B} \longrightarrow \mathbb{R}^p$, $f(\mathcal{A}) \subset \mathcal{B} \subset \mathbb{R}^m$, *непрерывно в точке* $\boldsymbol{b} := f(\boldsymbol{a}) \in \mathcal{B}$ *Тогда композиция* $g \circ f : \mathcal{A} \longrightarrow \mathbb{R}^p$ *непрерывна в точке* $\boldsymbol{a}$.

◀ Вводя обозначение $\boldsymbol{c} := g(\boldsymbol{b})$, зададим окрестность $\mathcal{W}(\boldsymbol{c})$ точки $\boldsymbol{c} \in \mathbb{R}^p$. Так как отображение g непрерывно в точке $\boldsymbol{b}$, то

$$\forall\, \mathcal{W}(\boldsymbol{c}) \;\; \exists\, \mathcal{V}(\boldsymbol{b}) \; : \; g\,(\mathcal{V}(\boldsymbol{b}) \cap \mathcal{B}) \subset \mathcal{W}(\boldsymbol{c})\,. \tag{14.28}$$

Так как отображение f непрерывно в точке $\boldsymbol{a}$, то

$$\forall\, \mathcal{V}(\boldsymbol{b}) \;\; \exists\, \mathcal{U}(\boldsymbol{a}) \; : \; f\,(\mathcal{U}(\boldsymbol{a}) \cap \mathcal{A}) \subset \mathcal{V}(\boldsymbol{b})\,. \tag{14.29}$$

Считая, что в (14.28) и (14.29) окрестность $\mathcal{V}(\boldsymbol{b})$ — одна и та же, исключая ее, получим

$$\begin{aligned}(g \circ f)\,(\mathcal{U}(\boldsymbol{a}) \cap \mathcal{A}) &= g\,(f\,(\mathcal{U}(\boldsymbol{a}) \cap \mathcal{A})) = \\ &= g\,(f\,(\mathcal{U}(\boldsymbol{a}) \cap \mathcal{A}) \cap \mathcal{B}) \subset g\,(\mathcal{V}(\boldsymbol{b}) \cap \mathcal{B}) \subset \mathcal{W}(\boldsymbol{c})\,.\end{aligned}$$

Итак,

$$\forall\, \mathcal{W}(\boldsymbol{c}) \;\; \exists\, \mathcal{U}(\boldsymbol{a}) \; : \; (g \circ f)\,(\mathcal{U}(\boldsymbol{a}) \cap \mathcal{A}) \subset \mathcal{W}(\boldsymbol{c})\,,$$

т. е. композиция $g \circ f$ непрерывна в точке $\boldsymbol{a}$. ▶

Теорема 12. *Непрерывность отображения* $f \; : \; \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, *в точке* $\boldsymbol{a} \in \mathcal{A}$ *равносильна непрерывности в этой точке всех координатных функций этого отображения.*

◀ Утверждение теоремы вытекает из установленной ранее эквивалентности

$$\lim_{\boldsymbol{x} \to \boldsymbol{a}} f(\boldsymbol{x}) = f(\boldsymbol{a}) \quad \Longleftrightarrow \quad \begin{cases} \lim\limits_{\boldsymbol{x} \to \boldsymbol{a}} f^1(\boldsymbol{x}) = f^1(\boldsymbol{a})\,, \\ \dots\dots\dots\dots\dots\dots\,, \\ \lim\limits_{\boldsymbol{x} \to \boldsymbol{a}} f^m(\boldsymbol{x}) = f^m(\boldsymbol{a})\,, \end{cases}$$

связывающей предел отображения f с пределами его координатных функций $f^1, \dots, f^m$. ▶

Теорема 13. *Сумма, разность, произведение и частное (в тех точках, где знаменатель не равен нулю) непрерывных скалярных функций векторного аргумента являются непрерывными функциями.*

◀ Эта теорема является простым следствием теоремы о связи между пределами скалярных функций и арифметическими операциями над этими функциями. ▶

Теорема 14. *Если отображение* $f : \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, *непрерывно в точке* $\boldsymbol{a} \in \mathcal{A}$, *то оно финально ограничено при* $\boldsymbol{x} \to \boldsymbol{a}$, $\boldsymbol{x} \in \mathcal{A}$.

◀ Эта теорема является простым следствием теоремы о финальной ограниченности отображения, имеющего при $\boldsymbol{x} \to \boldsymbol{a}$, $\boldsymbol{x} \in \mathcal{A}$ предел, лежащий в $\mathbb{R}^m$. Напомним здесь только определение финальной ограниченности отображения f:

$$\exists\, \mathcal{U}(\boldsymbol{a}) \ \ \exists\, M \in \mathbb{R}_+ \ \ \forall\, \boldsymbol{x} \in \mathcal{U}(\boldsymbol{a}) \cap \mathcal{A} : \ \ |f(\boldsymbol{x})| \leqslant M\,,$$

где $\mathcal{U}(\boldsymbol{a})$ — окрестность точки $\boldsymbol{a}$. ▶

Теорема 15. *Если функция* $f : \mathcal{A} \longrightarrow \mathbb{R}$, $\mathcal{A} \subset \mathbb{R}^n$ *непрерывна в точке* $\boldsymbol{a} \in \mathcal{A}$, *и* $f(\boldsymbol{a}) \neq 0$, *то существует окрестность* $\mathcal{U}(\boldsymbol{a})$, *такая, что* $\forall\, \boldsymbol{x} \in \mathcal{U}(\boldsymbol{a}) \cap \mathcal{A}$ *число* $f(\boldsymbol{x})$ *имеет тот же знак, что и* $f(\boldsymbol{a})$.

◀ Так как $f(\boldsymbol{a}) \neq 0$, то существует интервал

$$\mathcal{V}\left(f(\boldsymbol{a})\right) := \left(f(\boldsymbol{a}) - \varepsilon\,,\, f(\boldsymbol{a}) + \varepsilon\right)\,,$$

не содержащий нуля. Значит, все числа этого интервала имеют тот же знак, что и $f(\boldsymbol{a})$. Так как функция f непрерывна в точке $\boldsymbol{a}$, то

$$\exists\, \mathcal{U}(a) \ : \ \ f\left(\mathcal{U}(\boldsymbol{a}) \cap \mathcal{A}\right) \subset \mathcal{V}\left(f(\boldsymbol{a})\right)$$

или, что равносильно, $\forall \boldsymbol{x} \in \mathcal{U}(\boldsymbol{a}) \cap \mathcal{A}$ будет: $f(\boldsymbol{x}) \in \mathcal{V}\left(f(\boldsymbol{a})\right)$, и потому $f(\boldsymbol{x})$ будет иметь тот же знак, что и $f(\boldsymbol{a})$. ▶

2. Важнейшие примеры непрерывных отображений

Сначала дадим следующее определение.

Определение 16. *Отображение* $f : \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$ *называется непрерывным на некотором множестве, если оно непрерывно в каждой точке этого множества.*

Теперь переходим к перечислению важнейших примеров.

1) *Постоянное отображение* $f(\boldsymbol{x}) \equiv \boldsymbol{c}$, где $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{c} \in \mathbb{R}^m$, непрерывно на $\mathbb{R}^n$, так как $\forall\, \boldsymbol{a} \in \mathbb{R}^n$: $\lim\limits_{\boldsymbol{x} \to \boldsymbol{a}} \boldsymbol{c} = \boldsymbol{c}$.

2) *Тождественное отображение* $\mathrm{Id} : \boldsymbol{x} \longmapsto \boldsymbol{x}$ непрерывно на $\mathbb{R}^n$, так как $\forall\, \boldsymbol{a} \in \mathbb{R}^n : \lim\limits_{\boldsymbol{x}\to\boldsymbol{a}} \boldsymbol{x} = \boldsymbol{a}$.

3) *Отображения проектирования* $\pi^k : \boldsymbol{x} \longmapsto x_k$, где $\boldsymbol{x} = (x_1, \ldots, x_n)$, $k = 1, \ldots, n$, непрерывны на $\mathbb{R}^n$ как координатные функции тождественного отображения.

4) *Линейная функция* $L : (x_1, \ldots, x_n) \longmapsto a_1x_1 + \ldots + a_nx_n$ непрерывна на $\mathbb{R}^n$ как сумма произведений непрерывных функций (постоянных функций и отображений проектирования).

5) *Линейным отображением (оператором)* называется всякое отображение, $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, которое можно представить в следующем виде:

$$L(\boldsymbol{x}) := \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{m1}x_n \end{pmatrix},$$

где $\boldsymbol{x} = (x_1, \ldots, x_n)$. Оно непрерывно на $\mathbb{R}^n$, так как линейны, а значит, и непрерывны все его координатные функции

$$L^k(\boldsymbol{x}) = \left(\pi^k \circ L\right)(\boldsymbol{x}) := a_{k1}x_1 + \ldots + a_{kn}x_n\,, \quad k = 1, \ldots, m\,.$$

6) *Целая рациональная функция,* т. е. всякая функция $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, представимая в виде конечной суммы

$$f(\boldsymbol{x}) = \sum_{k_1,\ldots,k_n} a_{k_1,\ldots,k_n}(x_1)^{k_1} \ldots (x_n)^{k_n}\,,$$

где $\boldsymbol{x} = (x_1, \ldots, x_n)$, а $a_{k_1 \ldots k_n}$ — числовые коэффициенты, непрерывна на $\mathbb{R}^n$ как сумма произведений непрерывных функций (постоянных функций и отображений проектирования).

7) *Дробная рациональная функция,* т. е. функция, представимая в виде отношения конечных сумм

$$f(\boldsymbol{x}) = \frac{\sum\limits_{k_1,\ldots,k_n} a_{k_1,\ldots,k_n}(x_1)^{k_1} \ldots (x_n)^{k_n}}{\sum\limits_{i_1,\ldots,i_n} b_{i_1,\ldots,i_n}(x_1)^{i_1} \ldots (x_n)^{i_n}},$$

представляет собой частное двух целых рациональных функций и потому непрерывна всюду на $\mathbb{R}^n$, кроме множества всех точек, в которых знаменатель равен нулю.

8) Функция *модуль* $\boldsymbol{x} \longmapsto |\boldsymbol{x}| := \sqrt{(x_1)^2 + \ldots + (x_n)^2}$ непрерывна на $\mathbb{R}^n$ как композиция непрерывных функций (целой рациональной функции и корня квадратного).

9) Функция *определитель* $\det : \mathbb{R}^{(n^2)} \longrightarrow \mathbb{R}$, задаваемая равенством

$$\det \begin{pmatrix} x_{11} & \ldots & x_{1n} \\ \ldots & \ldots & \ldots \\ x_{n1} & \ldots & x_{nn} \end{pmatrix} = \begin{vmatrix} x_{11} & \ldots & x_{1n} \\ \ldots & \ldots & \ldots \\ x_{n1} & \ldots & x_{nn} \end{vmatrix},$$

непрерывна на $\mathbb{R}^{(n^2)}$, так как определитель является целой рациональной функцией от его элементов $x_{\mu\nu}$.

10) Функция *скалярное произведение* $\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$, задаваемая равенством $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1 y_1 + \ldots + x_n y_n$ непрерывна как целая рациональная функция от координат векторов $\boldsymbol{x}$ и $\boldsymbol{y}$.

11) Отображение *векторное произведение* $\times : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, задаваемое равенством

$$\boldsymbol{x} \times \boldsymbol{y} := \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix},$$

непрерывно, так как его координатные функции являются целыми рациональными функциями от координат векторов $\boldsymbol{x} = (x_1, x_2, x_3)$ и $\boldsymbol{y} = (y_1, y_2, y_3)$

12) *Смешанное произведение*, т. е. функция из $\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$ в $\mathbb{R}$, значения которой вычисляются по формуле

$$\langle \boldsymbol{x}, \boldsymbol{y} \times \boldsymbol{z} \rangle = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix},$$

непрерывна как целая рациональная функция от координат перемножаемых векторов.

3. Глобальные свойства непрерывных отобржений

Сначала напомним понятие относительно открытого множества.

Определение 17. *Множество $\mathcal{K} \subset \mathbb{R}^n$ называется открытым относительно множества $\mathcal{A} \subset \mathbb{R}^n$, если его можно представить в виде $\mathcal{K} = \mathcal{A} \cap \mathcal{U}$, где $\mathcal{U} \subset \mathbb{R}^n$ — открытое множество.*

Теорема 16. *Нерперывность на множестве $\mathcal{A}$ отображения $f\ :\ \mathcal{A} \longrightarrow \mathbb{R}^n$, $\mathcal{A} \subset \mathbb{R}^m$, равносильна тому, что полный прообраз любого открытого множества $\mathcal{V} \subset \mathbb{R}^m$ открыт относительно $\mathcal{A}$.*

◀ Предположим сначала, что отображение f непрерывно на множестве $\mathcal{A}$. Пусть $\mathcal{V} \subset \mathbb{R}^n$ — открытое множество. Если $f^{-1}(\mathcal{V}) = \varnothing = \varnothing \cap \mathcal{A}$, то это множество открыто относительно $\mathcal{A}$, поскольку множество $\varnothing$ открыто в любой топологии. Предположим теперь, что $f^{-1}(\mathcal{V}) \neq \varnothing$. Тогда существует $\boldsymbol{a} \in \mathcal{A} \cap f^{-1}(\mathcal{V})$. Так как отображение f непрерывно в точке $\boldsymbol{a}$, то существует открытая окрестность $\mathcal{U}(\boldsymbol{a}) \ni \boldsymbol{a}$ такая, что $f(\mathcal{U}(\boldsymbol{a}) \cap \mathcal{A}) \subset \mathcal{V}$ или, что равносильно,

$$\{\boldsymbol{a}\} \subset \mathcal{U}(\boldsymbol{a}) \cap \mathcal{A} \subset f^{-1}(\mathcal{V})\ .$$

Объединяя эти три множества по всем $\boldsymbol{a} \in f^{-1}(\mathcal{V})$, получим

$$f^{-1}(\mathcal{V}) = \bigcup_{\boldsymbol{a} \in f^{-1}(\mathcal{V})} \{\boldsymbol{a}\} \subset \bigcup_{\boldsymbol{a} \in f^{-1}(\mathcal{V})} \mathcal{U}(\boldsymbol{a}) \cap \mathcal{A} \subset \bigcup_{\boldsymbol{a} \in f^{-1}(\mathcal{V})} f^{-1}(\mathcal{V}) = f^{-1}(\mathcal{V})\ .$$

Отсюда видно, что все включения на самом деле являются равенствами, и мы имеем $f^{-1}(\mathcal{V}) = \mathcal{A} \cap \mathcal{U}$, где множество $\mathcal{U} := \bigcup\limits_{\boldsymbol{a} \in f^{-1}(\mathcal{V})} \mathcal{U}(\boldsymbol{a})$ открыто как объединение открытых множеств.

Обратно, предположим, что полный прообраз открытого множества открыт относительно $\mathcal{A}$. Пусть $\boldsymbol{a} \in \mathcal{A}$ и $\boldsymbol{b} := f(\boldsymbol{a})$. Возьмем открытую окрестность $\mathcal{V}(\boldsymbol{b}) \ni \boldsymbol{b}$. По предположению ее полный прообраз можно представить в виде $f^{-1}(\mathcal{V}(\boldsymbol{b})) = \mathcal{A} \cap \mathcal{U}(\boldsymbol{a})$, где множество $\mathcal{U}(\boldsymbol{a}) \ni \boldsymbol{a}$ — открытое, т. е. открытая окрестность точки $\boldsymbol{a}$. Так как включение не исключает равенства, то полученное можно переписать в такой форме

$$\forall\, \mathcal{V}(\boldsymbol{b})\ \exists\, \mathcal{U}(\boldsymbol{a})\ :\ \ f(\mathcal{U}(\boldsymbol{a})) \subset \mathcal{V}(\boldsymbol{b})\,,$$

что совпадает с определением непрерывности «на языке окрестностей». ▶

Теорема 17. *Если отображение $f\ :\ \mathcal{A} \longrightarrow \mathbb{R}^m$ непрерывно на компактном множестве $\mathcal{A} \subset \mathbb{R}^n$, то образ $f(\mathcal{A}) \subset \mathbb{R}^m$ компактен.*

◀ Пусть $\left\{\mathcal{V}_\alpha \;\middle|\; \alpha \in I\right\}$ — открытое в $\mathbb{R}^m$ покрытие множества $f(\mathcal{A})$, т. е. $f(\mathcal{A}) \subset \bigcup\limits_{\alpha \in I} \mathcal{V}_\alpha$. Переходя здесь к полным прообразам, получим $\mathcal{A} \subset f^{-1}\left(\bigcup\limits_{\alpha \in I} \mathcal{V}_\alpha\right) = \bigcup\limits_{\alpha \in I} f^{-1}(\mathcal{V}_\alpha)$. Так как отображение

f непрерывно на $\mathcal{A}$, то по предыдущей теореме для любого $\alpha \in I$ существует открытое $\mathcal{U}_\alpha \subset \mathbb{R}^n$ и такое, что $f^{-1}(\mathcal{V}_\alpha) = \mathcal{A} \cap \mathcal{U}_\alpha$. Таким образом, $\mathcal{A} \subset \bigcup\limits_{\alpha \in I} \mathcal{A} \cap \mathcal{U}_\alpha \subset \bigcup\limits_{\alpha \in I} \mathcal{U}_\alpha$. Значит, семейство $\{\mathcal{U}_\alpha \mid \alpha \in I\}$ — открытое покрытие множества $\mathcal{A}$. Так как множество $\mathcal{A}$ — компактное, то это покрытие содержит конечное подпокрытие

$$\{\mathcal{U}_{\alpha_1}, \ldots, \mathcal{U}_{\alpha_k}\} \; ; \; \mathcal{A} \subset \bigcup_{\nu=1}^{k} \mathcal{U}_{\alpha_\nu} \, .$$

Переходя здесь к образам, получим

$$f(\mathcal{A}) \subset f\left(\bigcup_{\nu=1}^{k} \mathcal{U}_{\alpha_\nu}\right) \subset \bigcup_{\nu=1}^{k} f\left(\mathcal{U}_{\alpha_\nu} \cap \mathcal{A}\right) = \bigcup_{\nu=1}^{k} \mathcal{V}_{\alpha_\nu} \, .$$

Таким образом, $\{\mathcal{V}_{\alpha_1}, \ldots, \mathcal{V}_{\alpha_k}\}$ — конечное подпокрытие множества $f(\mathcal{A})$, и потому это множество компактное. ►

Теорема 18 (Вейерштрасс). *Если функция $f : \mathcal{A} \longrightarrow \mathbb{R}$, $\mathcal{A} \subset \mathbb{R}^n$, непрерывна на компактном множестве $\mathcal{A}$, то она ограничена, и на множестве $\mathcal{A}$ существуют точки, в которых она принимает свои минимальное и максимальное значения.*

◄ По предыдущей теореме множество $f(\mathcal{A}) \subset \mathbb{R}$ — компактное, а по критерию компактности оно ограниченное и замкнутое. Ограниченность множества $f(\mathcal{A})$ равносильна ограниченности функции f. Замкнутость его означает, что ему принадлежат все его граничные точки, в частности, точки $m := \inf f(\mathcal{A})$ и $M := \sup f(\mathcal{A})$. Так как $m, M \in f(\mathcal{A})$, то сущесвуют такие точки $\boldsymbol{x}_* \in \mathcal{A}$ и $\boldsymbol{x}^* \in \mathcal{A}$, что $f(\boldsymbol{x}_*) = m$, $f(\boldsymbol{x}^*) = M$. ►

Теорема 19. *Если отображение $f : \mathcal{A} \longrightarrow \mathbb{R}^m$ непрерывно на связном множестве $\mathcal{A} \subset \mathbb{R}^n$, то образ $f(\mathcal{A})$ связен.*

◄ Предположим противное, т. е. что множество $f(\mathcal{A})$ не является связным. Тогда существуют открытые в $\mathbb{R}^m$ множества $\mathcal{V}_1$ и $\mathcal{V}_2$ такие, что

$$\mathcal{V}_1 \cap f(\mathcal{A}) \neq \varnothing \, , \; \mathcal{V}_2 \cap f(\mathcal{A}) \neq \varnothing \, , \; \mathcal{V}_1 \cap \mathcal{V}_2 \cap f(\mathcal{A}) = \varnothing \, , \; \mathcal{V}_1 \cup \mathcal{V}_2 \cup f(\mathcal{A}) = f(\mathcal{A}) \, . \tag{14.30}$$

Так как отображение f непрерывно, то по теореме 16 существуют открытые в $\mathbb{R}^n$ множества $\mathcal{U}_1$ и $\mathcal{U}_2$ такие, что

$$f^{-1}(\mathcal{V}_1) = \mathcal{U}_1 \cap \mathcal{A} \, , \; f^{-1}(\mathcal{V}_2) = \mathcal{U}_2 \cap \mathcal{A} \, . \tag{14.31}$$

Учитывая эти равенства и переходя в (14.30) к полным прообразам, получим

$$\mathcal{U}_1 \cap \mathcal{A} \neq \varnothing\,,\ \mathcal{U}_2 \cap \mathcal{A} \neq \varnothing\,,\ \mathcal{U}_1 \cap \mathcal{U}_2 \cap \mathcal{A} = \varnothing\,,\ \mathcal{U}_1 \cup \mathcal{U}_2 \cup f(\mathcal{A}) = \mathcal{A}\,.$$

Эти соотношения означают, что множество $\mathcal{A}$ не является связным, и мы получили противоречие. ▶

Теорема 20. *Если функция $f\,:\,\mathcal{A} \longrightarrow \mathbb{R}$ непрерывна на связном множестве $\mathcal{A}$, то в некоторых точках множества $\mathcal{A}$ она принимает все значения, заключенные между* $\inf f(\mathcal{A})$ *и* $\sup f(\mathcal{A})$.

◀ На основании предыдущей теоремы заключаем, что множество $f(\mathcal{A})$ — связное. Так как $f(\mathcal{A}) \subset \mathbb{R}$, то в силу теоремы о классификации связных подмножеств числовой оси множество $f(\mathcal{A})$ является числовым промежутком. ▶

Определение 18. *Отображение $f\ :\ \mathcal{A} \longrightarrow \mathbb{R}^n$, $\mathcal{A} \subset \mathbb{R}^m$, называется равномерно непрерывным на множестве $\mathcal{A}$, если выполняется следующее условие:*

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall \boldsymbol{x}', \boldsymbol{x}'' : |\boldsymbol{x}' - \boldsymbol{x}''| \leqslant \delta \Longrightarrow |f(\boldsymbol{x}') - f(\boldsymbol{x}'')| \leqslant \varepsilon\,. \quad (14.32)$$

Теорема 21 (Кантор). *Если отображение $f\ :\ \mathcal{A} \longrightarrow \mathbb{R}^m$ непрерывно на компактном множестве $\mathcal{A} \subset \mathbb{R}^n$, то оно и равномерно непрерывно на $\mathcal{A}$.*

◀ Так как отображение f непрерывно на $\mathcal{A}$, то

$$\forall\, \boldsymbol{c} \in \mathcal{A}\ \ \forall\, \varepsilon > 0\ \ \exists\, \eta = \eta(\varepsilon, \boldsymbol{c}) > 0\ \ \forall\, \boldsymbol{x} \in \mathcal{A}\ :$$
$$|\boldsymbol{x} - \boldsymbol{c}| \leqslant \eta \implies |f(\boldsymbol{x}) - f(\boldsymbol{c})| \leqslant \frac{\varepsilon}{2}\,. \quad (14.33)$$

Обозначая $\mathcal{U}_\eta(\boldsymbol{c}) := \{|\boldsymbol{x} - \boldsymbol{c}| < \eta\}$ — открытый шар радиуса η с центром в точке $\boldsymbol{c}$, рассмотрим семейство шаров $\big\{\mathcal{U}_{\eta/2}(\boldsymbol{c})\ \big|\ \boldsymbol{c} \in \mathcal{A}\big\}$. Это семейство есть, очевидно, открытое покрытие множества $\mathcal{A}$. Так как множество $\mathcal{A}$ — компактное, то это покрытие содержит конечное подпокрытие $\{\mathcal{U}_1\,,\,\dots\,,\,\mathcal{U}_k\}$, где обозначено $\mathcal{U}_\nu := \mathcal{U}_{\eta_\nu/2}(\boldsymbol{c}_\nu)$. Положим $\delta := \dfrac{1}{4}\min\{\eta_1\,,\,\dots\,,\,\eta_k\}$, и пусть точки $\boldsymbol{x}', \boldsymbol{x}'' \in \mathcal{A}$ удовлетворяют неравенству $|\boldsymbol{x}' - \boldsymbol{x}''| \leqslant \delta$. При некотором ν будет $\boldsymbol{x}' \in \mathcal{U}_\nu$, и значит, $|\boldsymbol{x}' - \boldsymbol{c}_\nu| \leqslant \dfrac{\eta_\nu}{2}$. Поэтому из (14.33) следует, что

$|f(\boldsymbol{x}') - f(\boldsymbol{c}_\nu)| \leqslant \dfrac{\varepsilon}{2}$. Далее,

$$|\boldsymbol{x}'' - \boldsymbol{c}_\nu| \leqslant |\boldsymbol{x}'' - \boldsymbol{x}'| + |\boldsymbol{x}' - \boldsymbol{c}_\nu| \leqslant \delta + \frac{\eta_\nu}{2} \leqslant \eta_\nu\,,$$

и в силу (14.33) имеем $|f(\boldsymbol{x}'') - f(\boldsymbol{c}_\nu)| \leqslant \dfrac{\varepsilon}{2}$. И, наконец,

$$|f(\boldsymbol{x}') - f(\boldsymbol{x}'')| \leqslant |f(\boldsymbol{x}') - f(\boldsymbol{c}_\nu)| + |f(\boldsymbol{x}'') - f(\boldsymbol{c}_\nu)| \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,. \ \blacktriangleright$$

Глава 15

ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ ОТОБРАЖЕНИЙ ИЗ $\mathbb{R}^n$ В $\mathbb{R}^m$

§ 1. Дифференцируемые отображения и их основные свойства

1. Понятия дифференцируемости, производной, дифференциала

Напомним, что дифференцируемость функции f скалярного аргумента в точке x равносильна существованию в этой точке *конечной производной*

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=f'(x)\in\mathbb{R}. \qquad (15.1)$$

Для функций векторного аргумента (т. е. когда $\boldsymbol{x}\in\mathbb{R}^n$, $n>1$) левая часть равенства (15.1) теряет смысл, так как в этом случае в нее входила бы операция деления на вектор $\boldsymbol{h}\in\mathbb{R}^n$, которая при $n>1$ не имеет смысла. Поэтому, желая обобщить свойство дифференцируемости на отображения из $\mathbb{R}^n$ в $\mathbb{R}^m$, мы должны исходить из такого определения дифференцируемости функций одного переменного, которое не содержит операцию деления, например, из соотношения

$$f(x+h)-f(x)=L(h)+o(h)\quad\text{при}\quad h\to 0\,, \qquad (15.2)$$

где $L(h)=f'(x)\cdot h$ — линейная (т. е. однородная и аддитивная) функция от приращения h.

Перенося в (15.2) член $L(h)$ в левую часть, беря модули от обеих частей полученного равенства и деля на $|h|$, можно переписать равенство (15.2) в следующем равносильном виде:

$$\frac{|f(x+h)-f(x)-L(h)|}{|h|}=o(1)\quad\text{при}\quad h\to 0\,. \qquad (15.3)$$

Такая форма определения дифференцируемости уже допускает обобщение на отображения из $\mathbb{R}^n$ в $\mathbb{R}^m$, так как в (15.3) не может возникнуть деление на вектор, а под L можно понимать линейный оператор, не обязательно представимый в виде $L(h) = A \cdot h$.

Переходя непосредственно к определению дифференцируемости отображения $f\ :\ \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, будем предполагать, что множество $\mathcal{A}$ является *областью* (т. е. открытым и связным множеством), что упрощает ситуацию. Открытость позволяет отвлечься от изучения сложного вопроса дифференцируемости в граничных точках. Связность делает рассмотрения более наглядными, так как всегда можно вместо отображения f рассматривать его сужения на связные компоненты множества $\mathcal{A}$.

Определение 19. *Отображение $f\ :\ \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, называется дифференцируемым (по Фреше[1]) в точке $\boldsymbol{x} \in \mathcal{A}$, если существует линейное отображение (оператор) $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ такое, что выполняется равенство*

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{|f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x}) - L(\boldsymbol{h})|}{|\boldsymbol{h}|} = 0\,. \tag{15.4}$$

Определение 20. *Производной отображения f в точке $\boldsymbol{x} \in \mathcal{A}$ называется линейный оператор $L\ :\ \mathbb{R}^n \longrightarrow \mathbb{R}^m$, для которого выполняется равенство* (15.4).

Производную принято обозначать следующими символами:

$$f'(\boldsymbol{x}) = Df(\boldsymbol{x}) := L\,,$$

где L из (15.4).

Определение 21. *Дифференциалом отображения $\boldsymbol{y} = f(\boldsymbol{x})$ в точке $\boldsymbol{x}$ условимся называть его производную L в этой точке, заданную аналитическим способом $d\boldsymbol{y} = L(\boldsymbol{h})$.*

Обозначение $d\boldsymbol{y}$ для дифференциала будем называть *классическим.* Приведем другие, более точные обозначения для дифференциала

$$Df(\boldsymbol{x})(\boldsymbol{h}) = f'(\boldsymbol{x})(\boldsymbol{h}) = d\boldsymbol{y} := L(\boldsymbol{h})\,.$$

Левая часть этого равенства является наиболее точным обозначением для диффуренциала и читается так: *«дифференциал отображения f*

[1] *Фреше́* Морис Рене (1878—1973) — французский математик.

в точке $\boldsymbol{x}$ *на векторе* $\boldsymbol{h}$ *»,* подчеркивая тем самым его зависимость от приращеня $\boldsymbol{h}$, притом линейную.

Чтобы обосновать корректность определения 20, необходимо еще доказать *единственность*, т. е. что может существовать не более одного оператора L, для которого выполняется равенство (15.4).

Теорема 22. *Если отображение* $f : \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, *дифференцируемо в точке* $\boldsymbol{x} \in \mathcal{A}$, *то его производная определяется равенством* (15.4) *однозначно.*

◀ Пусть $L, M : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ — два линейных оператора, для которых выполняется равенство (15.4), т. е.

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{|f(\boldsymbol{x}+\boldsymbol{h})-f(\boldsymbol{x})-L(\boldsymbol{h})|}{|\boldsymbol{h}|} = \lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{|f(\boldsymbol{x}+\boldsymbol{h})-f(\boldsymbol{x})-M(\boldsymbol{h})|}{|\boldsymbol{h}|} = 0\,. \tag{15.5}$$

Надо показать, что $L = M$, т. е. что $\forall \boldsymbol{h} \in \mathbb{R}^n : L(\boldsymbol{h}) = M(\boldsymbol{h})$. Так как оба эти оператора — линейные, то $L(\boldsymbol{0}) = M(\boldsymbol{0}) = \boldsymbol{0}$. Пусть теперь $\boldsymbol{h} \neq \boldsymbol{0}$. В силу неравенства треугольника имеем

$$0 \leqslant \frac{|L(\boldsymbol{h})-M(\boldsymbol{h})|}{|\boldsymbol{h}|} \leqslant \frac{|L(\boldsymbol{h})-f(\boldsymbol{x}+\boldsymbol{h})+f(\boldsymbol{x})|}{|\boldsymbol{h}|} + \frac{|f(\boldsymbol{x}+\boldsymbol{h})-f(\boldsymbol{x})-M(\boldsymbol{h})|}{|\boldsymbol{h}|}\,.$$

Переходя здесь к пределу при $\boldsymbol{h} \to \boldsymbol{0}$ и учитывая равенства (15.5), получаем

$$\lim_{\substack{\boldsymbol{h}\to\boldsymbol{0},\\ \boldsymbol{h}\neq\boldsymbol{0}}} \frac{|L(\boldsymbol{h})-M(\boldsymbol{h})|}{|\boldsymbol{h}|} = 0\,. \tag{15.6}$$

Полагая здесь $\boldsymbol{h} = t\cdot\boldsymbol{l}$, где $t = |\boldsymbol{h}| > 0$, $|\boldsymbol{l}| = 1$, имеем

$$0 = \lim_{t\to+0} \frac{|L(t\boldsymbol{l})-M(t\boldsymbol{l})|}{|t\boldsymbol{l}|} = \lim_{t\to+0} \frac{t\,|L(\boldsymbol{l})-M(\boldsymbol{l})|}{t|\boldsymbol{l}|} = |L(\boldsymbol{l})-M(\boldsymbol{l})|\,.$$

Отсюда получаем тождество $L(\boldsymbol{l}) \equiv M(\boldsymbol{l})$, которое после умножения на t принимает вид $L(\boldsymbol{h}) \equiv M(\boldsymbol{h})$. Итак, $L = M$. ▶

Теорема 23. *Если отображение $f : \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, дифференцируемо в точке $\boldsymbol{x} \in \mathcal{A}$, то оно и непрерывно в этой точке.*

◀ Для дифференцируемого в точке $\boldsymbol{x}$ отображения f из равенства (15.4) вытекает следующее равенство

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} |f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x}) - L(\boldsymbol{h})| = 0\,. \tag{15.7}$$

Применяя неравенство треугольника, оценим модуль приращения отображения f

$$0 \leqslant |f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x})| \leqslant |f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x}) - L(\boldsymbol{h})| + |L(\boldsymbol{h})|\,. \tag{15.8}$$

Первое слагаемое в правой части этих неравенств бесконечно мало в силу (15.7), а второе — в силу непрерывности линейного оператора и того, что $L(\boldsymbol{0}) = \boldsymbol{0}$. Переходя к пределу в неравенствах (15.8) при $\boldsymbol{h} \to \boldsymbol{0}$, получаем $\lim\limits_{\boldsymbol{h}\to\boldsymbol{0}} |f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x})| = 0$, или, что равносильно, $\lim\limits_{\boldsymbol{h}\to\boldsymbol{0}} f(\boldsymbol{x}+\boldsymbol{h}) = f(\boldsymbol{x})$. ▶

2. Теоремы о дифференцируемости некоторых отображений

Определение 22. *Нормой линейного оператора $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ называется число $\|L\| := \sup\limits_{|\boldsymbol{x}|=1} |L(\boldsymbol{x})|$.*

Теорема 24. *Для любого линейного оператора $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ норма $\|L\|$ существует, причем $|L(\boldsymbol{x})| \leqslant \|L\| \cdot |\boldsymbol{x}|$ при любом $\boldsymbol{x} \in \mathbb{R}^n$.*

◀ Так как функция $\boldsymbol{x} \longmapsto |L(\boldsymbol{x})|$ — непрерывная, а сфера $|\boldsymbol{x}| = 1$ — компактная, то по теореме Вейерштрасса множество $\big\{L(\boldsymbol{x}) \;\big|\; |\boldsymbol{x}| = 1\big\}$ — ограниченное. Поэтому существует число $\|L\| := \sup\limits_{|\boldsymbol{x}|=1} |L(\boldsymbol{x})|$. Так как $\left|\dfrac{\boldsymbol{x}}{|\boldsymbol{x}|}\right| = 1$ при $\boldsymbol{x} \neq \boldsymbol{0}$, то

$$\left|L\left(\frac{\boldsymbol{x}}{|\boldsymbol{x}|}\right)\right| \leqslant \|L\|\,, \quad \text{откуда} \quad |L(\boldsymbol{x})| \leqslant \|L\| \cdot |\boldsymbol{x}|\,. \ \blacktriangleright$$

Теорема 25 (о дифференцируемости композиции). *Предположим, что отображение* $f : \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, *дифференцируемо в точке* $\boldsymbol{a} \in \mathcal{A}$, *а отображение* $g : \mathcal{B} \longrightarrow \mathbb{R}^p$, $\mathcal{B} \subset \mathbb{R}^m$, *дифференцируемо в точке* $\boldsymbol{b} := f(\boldsymbol{a}) \in f(\mathcal{A}) \subset \mathcal{B}$. *Тогда композиция* $g \circ f : \mathcal{A} \longrightarrow \mathbb{R}^p$ *дифференцируема в точке* $\boldsymbol{a}$, *причем*

$$(g \circ f)'(\boldsymbol{a}) = g'(\boldsymbol{b}) \circ f'(\boldsymbol{a}). \tag{15.9}$$

◀ Обозначая $L := f'(\boldsymbol{a})$, $M := g'(\boldsymbol{b})$, запишем условия дифференцируемости отображений f и g

$$f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) = L(\boldsymbol{h}) + \alpha(\boldsymbol{h}) \cdot |\boldsymbol{h}|, \tag{15.10}$$

$$g(\boldsymbol{b} + \boldsymbol{k}) - g(\boldsymbol{b}) = M(\boldsymbol{k}) + \beta(\boldsymbol{k}) \cdot |\boldsymbol{k}|, \tag{15.11}$$

где $\alpha(\boldsymbol{h}) \to \boldsymbol{0}$ при $\boldsymbol{h} \to \boldsymbol{0}$ и $\beta(\boldsymbol{k}) \to \boldsymbol{0}$ при $\boldsymbol{k} \to \boldsymbol{0}$, и можно положить $\alpha(\boldsymbol{0}) := \boldsymbol{0}$, $\beta(\boldsymbol{0}) := \boldsymbol{0}$.

Полагая в (15.11) $\boldsymbol{k} := f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a})$, заключаем, что $\boldsymbol{k} \to \boldsymbol{0}$ при $\boldsymbol{h} \to \boldsymbol{0}$ в силу непрервности отображения f. Используя выбранное значение приращения $\boldsymbol{k}$, а также равенства (15.10) и (15.11), имеем

$$\begin{aligned}
0 \leqslant |(g \circ f)(\boldsymbol{a} + \boldsymbol{h}) - (g \circ f)(\boldsymbol{a}) - (M \circ L)(\boldsymbol{h})| = \\
= |g\,(f(\boldsymbol{a} + \boldsymbol{h})) - g\,(f(\boldsymbol{a}))| - M\,(L(\boldsymbol{h})) = \\
= |g(\boldsymbol{b} + \boldsymbol{k}) - g(\boldsymbol{b}) - M(L(\boldsymbol{h}))| = |M(\boldsymbol{k}) + \beta(\boldsymbol{k}) \cdot \boldsymbol{k} - M\,(L(\boldsymbol{h}))| \leqslant \\
\leqslant |M(\boldsymbol{k} - L(\boldsymbol{h}))| + |\beta(\boldsymbol{k})| \cdot |\boldsymbol{k}| \leqslant \|M\| \cdot |f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) - L(\boldsymbol{a})| + \\
+ |\beta(\boldsymbol{k})| \cdot |f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a})| \leqslant \|M\| \cdot |\alpha(\boldsymbol{h})| \cdot |\boldsymbol{h}| + |\beta(\boldsymbol{k})| \cdot |L(\boldsymbol{h}) + \alpha(\boldsymbol{h}) \cdot |\boldsymbol{h}|| \leqslant \\
\leqslant [\|M\| \cdot |\alpha(\boldsymbol{h})| + \|L\| \cdot |\beta(\boldsymbol{k})| + |\alpha(\boldsymbol{h}) \cdot \beta(\boldsymbol{k})|] \cdot |\boldsymbol{h}|.
\end{aligned}$$

Выражение, находящееся в квадратных скобках правой части последнего неравенства, стемится к нулю при $\boldsymbol{h} \to \boldsymbol{0}$. Поэтому после деления на $|\boldsymbol{h}| \neq 0$ и перехода к пределу будем иметь

$$\lim_{\substack{\boldsymbol{h} \to \boldsymbol{0}, \\ \boldsymbol{h} \neq \boldsymbol{0}}} \frac{|(g \circ f)(\boldsymbol{a} + \boldsymbol{h}) - (g \circ f)(\boldsymbol{a}) - (M \circ L)(\boldsymbol{h})|}{|\boldsymbol{h}|} = 0,$$

т. е. $(g \circ f)'(\boldsymbol{a}) = M \circ L = g'(\boldsymbol{b}) \circ f'(\boldsymbol{a})$. ▶

Замечание. Формула (15.9) называется «цепным правилом». Применяя метод индукции, можно показать, что цепное правило справедливо для композиции любого конечного числа отображений. Например, в случае композиции трех отображений его можно записать в виде

$$(h \circ g \circ f)'(\boldsymbol{a}) = h'(\boldsymbol{c}) \circ g'(\boldsymbol{b}) \circ f'(\boldsymbol{a}) \tag{15.12}$$

или, что равносильно, в виде

$$D(h \circ g \circ f)(\boldsymbol{a}) = Dh(\boldsymbol{c}) \circ Dg(\boldsymbol{b}) \circ Df(\boldsymbol{a}), \tag{15.13}$$

где обозначено $\boldsymbol{b} := f(\boldsymbol{a})$, $\boldsymbol{c} := g(\boldsymbol{b})$.

Теорема 26. (a) *Если $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ — постоянное отображение, т. е. $f(\boldsymbol{x}) \equiv \boldsymbol{c}$, то оно дифференцируемо, причем*

$$f'(\boldsymbol{x}) \equiv O : \mathbb{R}^n \longrightarrow \mathbb{R}^m \quad \textit{(нуль-оператор).}$$

(b) *Если отображение $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ линейное, то оно и дифференцируемое, а его производная — постоянный оператор, равный $f'(\boldsymbol{x}) \equiv f$ или, что равносильно, $Df(\boldsymbol{x})(\boldsymbol{h}) \equiv f(\boldsymbol{h})$.*

(c) *Дифференцируемость отображения $f : \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, равносильна дифференцируемости всех его координатных функций f^k, причем*

$$Df(\boldsymbol{x})(\boldsymbol{h}) = \sum_{k=1}^{m} Df^k(\boldsymbol{x})(\boldsymbol{h})\boldsymbol{e}'_k, \tag{15.14}$$

где $\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_m$ — стандартный базис в $\mathbb{R}^m$.

◀ (a) Пусть $f(\boldsymbol{x}) \equiv \boldsymbol{c}$, и $O : \mathbb{R}^n \longrightarrow \{\boldsymbol{0}\}$ — нуль-оператор. Тогда имеем

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{|f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x}) - O(\boldsymbol{h})|}{|\boldsymbol{h}|} = \lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{|\boldsymbol{c} - \boldsymbol{c} - \boldsymbol{0}|}{|\boldsymbol{h}|} = \lim_{\boldsymbol{h}\to\boldsymbol{0}} 0 = 0,$$

т. е. $f'(\boldsymbol{x})(\boldsymbol{h}) = \boldsymbol{0}$ или $f' \equiv O$.

(b) Если $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ — линейный оператор, то имеем

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{|f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x}) - f(\boldsymbol{h})|}{|\boldsymbol{h}|} =$$

$$= \lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{|f(\boldsymbol{x}) + f(\boldsymbol{h}) - f(\boldsymbol{x}) - f(\boldsymbol{h})|}{|\boldsymbol{h}|} = \lim_{\boldsymbol{h}\to\boldsymbol{0}} 0 = 0,$$

т. е. $f'(\boldsymbol{x})(\boldsymbol{h}) = f(\boldsymbol{h})$ или $f'(\boldsymbol{x}) \equiv f$.

(c) Предположим, что отображение f дифференцируемо в точке $\boldsymbol{x}$. Так как отображения проектирования $\pi^1, \ldots, \pi^n$ — линейные, и значит, дифференцируемые, то координатные функции $f^k = \pi^k \circ f$ дифференцируемы как композиции дифференцируемых функций, и по цепному правилу имеем $Df^k = D(\pi^k \circ f) = D\pi^k \circ Df = \pi^k \circ Df$.

Обратно, предположим, что координатные функции $f^1, \ldots, f^m$ дифференцируемы в точке $\boldsymbol{x}$. Так как $f(\boldsymbol{x}) = \sum\limits_{k=1}^{m} f^k(\boldsymbol{x})\boldsymbol{e}'_k$, то имеем

$$0 \leqslant \frac{\left| f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x}) - \sum\limits_{k=1}^{m} Df^k(\boldsymbol{x})(\boldsymbol{h})\boldsymbol{e}'_k \right|}{|\boldsymbol{h}|} =$$

$$= \frac{1}{|\boldsymbol{h}|} \left| \sum_{k=1}^{m} \left[f^k(\boldsymbol{x}+\boldsymbol{h}) - f^k(\boldsymbol{x}) - Df^k(\boldsymbol{x})(\boldsymbol{h}) \right] \boldsymbol{e}'_k \right| \leqslant$$

$$\leqslant \sum_{k=1}^{m} \frac{\left| f^k(\boldsymbol{x}+\boldsymbol{h}) - f^k(\boldsymbol{x}) - Df^k(\boldsymbol{x})(\boldsymbol{h}) \right|}{|\boldsymbol{h}|} \to 0 \quad \text{при} \quad \boldsymbol{h} \to 0\,. \ \blacktriangleright$$

Теорема 27 (линейность операции дифференцирования). *Если отображения f, g : $\mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, дифференцируемы в точке $\boldsymbol{x} \in \mathcal{A}$, а λ, $\mu \in \mathbb{R}$, то отображение $\lambda f + \mu g$ дифференцируемо в точке $\boldsymbol{x}$, причем*

$$(\lambda f + \mu g)'(\boldsymbol{x}) = \lambda f'(\boldsymbol{x}) + \mu g'(\boldsymbol{x})\,. \tag{15.15}$$

◀ Для доказательства произведем следующую оценку:

$$0 \leqslant \frac{|(\lambda f + \mu g)(\boldsymbol{x}+\boldsymbol{h}) - (\lambda f + \mu g)(\boldsymbol{x}) - (\lambda f' + \mu g')(\boldsymbol{x})(\boldsymbol{h})|}{|\boldsymbol{h}|} =$$

$$= \frac{|\lambda\,[f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x}) - f'(\boldsymbol{x})(\boldsymbol{h})] + \mu\,[g(\boldsymbol{x}+\boldsymbol{h}) - g(\boldsymbol{x}) - g'(\boldsymbol{x})(\boldsymbol{h})]|}{|\boldsymbol{h}|} \leqslant$$

$$\leqslant |\lambda| \cdot \frac{|f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x}) - f'(\boldsymbol{x})(\boldsymbol{h})|}{|\boldsymbol{h}|} + |\mu| \cdot \frac{|g(\boldsymbol{x}+\boldsymbol{h}) - g(\boldsymbol{x}) - g'(\boldsymbol{x})(\boldsymbol{h})|}{|\boldsymbol{h}|} \to 0$$

при $\boldsymbol{h} \to \boldsymbol{0}$ в силу дифференцируемости отображений f и g в точке $\boldsymbol{x}$. Отсюда имеем

$$\lim_{\boldsymbol{h} \to \boldsymbol{0}} \frac{|(\lambda f + \mu g)(\boldsymbol{x}+\boldsymbol{h}) - (\lambda f + \mu g)(\boldsymbol{x}) - (\lambda f' + \mu g')(\boldsymbol{x})(\boldsymbol{h})|}{|\boldsymbol{h}|} = 0\,,$$

что равносильно равенству (15.15). ▶

Определение 23. *Отображение* $f : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \longrightarrow \mathbb{R}^m$ *от двух векторных аргументов* $\boldsymbol{x} \in \mathbb{R}^{n_1}$, $\boldsymbol{y} \in \mathbb{R}^{n_2}$ *называется билинейным, если оно линейно по каждому из своих аргументов при любом фиксированном значении другого аргумента. Нормой билинейного отображения назовем число*

$$\|f\| := \sup_{\substack{|\boldsymbol{x}|=1,\\ |\boldsymbol{y}|=1}} |f(\boldsymbol{x}, \boldsymbol{y})|\,. \tag{15.16}$$

Теорема 28. *Пусть* f *— билинейное отображение. Тогда:*

(a) *для любых* $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^{n_1}, \boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^{n_2}, \lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{R}$ *имеем*

$$\begin{gathered} f(\lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2\,, \mu_1\boldsymbol{y}_1 + \mu_2\boldsymbol{y}_2) = \lambda_1\mu_1 f(\boldsymbol{x}_1\,, \boldsymbol{y}_1) + \\ + \lambda_1\mu_2 f(\boldsymbol{x}_1\,, \boldsymbol{y}_2) + \lambda_2\mu_1 f(\boldsymbol{x}_2\,, \boldsymbol{y}_1) + \lambda_2\mu_2 f(\boldsymbol{x}_2\,, \boldsymbol{y}_2)\,; \end{gathered} \tag{15.17}$$

(b) *отображение* f *непрерывно на* $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$*;*

(c) *норма* $\|f\|$ *существует, причем* $|f(\boldsymbol{x}, \boldsymbol{y})| \leqslant |\boldsymbol{x}| \cdot |\boldsymbol{y}| \cdot \|f\|$.

◄ (a) Для доказательства равенства (15.17) достаточно применить свойство линейности сначала по одной переменной, затем по другой.

(b) Разлагая векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ по векторам стандартных базисов и используя свойство билинейности, имеем

$$f(\boldsymbol{x}\,,\, \boldsymbol{y}) = f\left(\sum_{\mu=1}^{n_1} x_\mu \boldsymbol{e}_\mu\,, \sum_{\nu=1}^{n_2} y_\nu \boldsymbol{e}'_\nu\right) = \sum_{\mu=1}^{n_1}\sum_{\nu=1}^{n_2} x_\mu y_\nu f(\boldsymbol{e}_\mu\,,\, \boldsymbol{e}'_\nu)\,.$$

Отсюда видно, что координатные функции билинейного отображения — целые рациональные функции от координат векторов $\boldsymbol{x}$, $\boldsymbol{y}$ и потому непрерывны. Значит, и f непрерывно на $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$.

(c) Так как множество $T := \{|\boldsymbol{x}| = 1\} \times \{|\boldsymbol{y}| = 1\}$ — компактное (как декартово произведение компактных множеств), а функция $T \longrightarrow \mathbb{R}_+ : (\boldsymbol{x}\,,\, \boldsymbol{y}) \longmapsto |f(\boldsymbol{x}\,,\, \boldsymbol{y})|$ — нпрерывная, то она ограничена на T и, значит, $\sup\limits_T |f(\boldsymbol{x}\,,\, \boldsymbol{y})| \in \mathbb{R}_+$. И, наконец, при $\boldsymbol{x} \neq \boldsymbol{0}$, $\boldsymbol{y} \neq \boldsymbol{0}$ имеем

$$|f(\boldsymbol{x}\,,\, \boldsymbol{y})| = |\boldsymbol{x}| \cdot |\boldsymbol{y}| \cdot \left|f\left(\frac{\boldsymbol{x}}{|\boldsymbol{x}|}\,, \frac{\boldsymbol{y}}{|\boldsymbol{y}|}\right)\right| \leqslant |\boldsymbol{x}| \cdot |\boldsymbol{y}| \cdot \|f\|\,. \ ►$$

Примеры билинейных отображений.

1) $f_1 : \mathbb{R}\times\mathbb{R} \longrightarrow \mathbb{R}$, где $f_1(x\,,\,y) := x\cdot y$ — произведение скаляров.

2) $f_2 : \mathbb{R}^n\times\mathbb{R}^n \longrightarrow \mathbb{R}$, где $f_2(\boldsymbol{x}\,,\,\boldsymbol{y}) := \langle \boldsymbol{x}\,,\,\boldsymbol{y}\rangle$ — скалярное произведение векторов из $\mathbb{R}^n$.

3) $f_3 : \mathbb{R}^3\times\mathbb{R}^3 \longrightarrow \mathbb{R}^3$, где $f_3(\boldsymbol{x}\,,\,\boldsymbol{y}) := \boldsymbol{x}\times\boldsymbol{y}$ — векторное произведение векторов из $\mathbb{R}^3$.

Замечание. Аналогично билинейному может быть введено понятие *трилинейного* и вообще любого n*-линейного (полилинейного)* отображения и его нормы. Не останавливаясь на соответствующих определениях, ограничимся лишь **примерами**.

1) Обозначим $\boldsymbol{x} := (x_1\,,\,x_2\,,\,x_3)$, $\boldsymbol{y} := (y_1\,,\,y_2\,,\,y_3)$, $\boldsymbol{z} := (z_1\,,\,z_2\,,\,z_3)$. Отображение $f : \mathbb{R}^3\times\mathbb{R}^3\times\mathbb{R}^3 \longrightarrow \mathbb{R}$, для которого

$$f(\boldsymbol{x}\,,\boldsymbol{y}\,,\boldsymbol{z}) = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix},$$

является трилинейным.

2) Обозначим $\boldsymbol{x}^k := (x_1^k\,,\,\ldots\,,\,x_n^k)$, $k = 1\,,\,\ldots\,,\,n$. Отображение

$$f : \underbrace{\mathbb{R}^n\times\ldots\times\mathbb{R}^n}_{n \text{ раз}} \longrightarrow \mathbb{R},$$

для которого

$$f(\boldsymbol{x}_1\,,\,\ldots\,,\,\boldsymbol{x}_n) = \begin{vmatrix} x_1^1 & \ldots & x_n^1 \\ \ldots & \ldots & \ldots \\ x_1^n & \ldots & x_n^n \end{vmatrix},$$

является n-линейным.

Теорема 29. *Если отображение $f : \mathbb{R}^{n_1}\times\mathbb{R}^{n_2} \longrightarrow \mathbb{R}^m$ — билинейное, то оно и дифференцируемое, причем*

$$Df(\boldsymbol{x}\,,\boldsymbol{y})(\boldsymbol{h}\,,\boldsymbol{k}) = f(\boldsymbol{h}\,,\boldsymbol{y}) + f(\boldsymbol{x}\,,\boldsymbol{k})\,. \tag{15.18}$$

◀ Используя свойство билинейности, имеем

$$0 \leqslant \frac{|f(\boldsymbol{x}+\boldsymbol{h},\boldsymbol{y}+\boldsymbol{k}) - f(\boldsymbol{x},\boldsymbol{y}) - f(\boldsymbol{h},\boldsymbol{y}) - f(\boldsymbol{x},\boldsymbol{k})|}{|(\boldsymbol{h},\boldsymbol{k})|} =$$
$$= \frac{|f(\boldsymbol{h},\boldsymbol{k})|}{|(\boldsymbol{h},\boldsymbol{k})|} \leqslant \frac{\|f\|\cdot|\boldsymbol{h}|\cdot|\boldsymbol{k}|}{\sqrt{|\boldsymbol{h}|^2+|\boldsymbol{k}|^2}} \leqslant \frac{\|f\|}{\sqrt{|\boldsymbol{h}|^2+|\boldsymbol{k}|^2}}\cdot\frac{|\boldsymbol{h}|^2+|\boldsymbol{k}|^2}{2} =$$
$$= \frac{\|f\|}{2}\cdot|(\boldsymbol{h},\boldsymbol{k})| \to 0$$

при $(\boldsymbol{h}, \boldsymbol{k}) \to (\mathbf{0}, \mathbf{0})$. Отсюда следует равенство

$$\lim_{(\boldsymbol{h},\boldsymbol{k})\to(\mathbf{0},\mathbf{0})} \frac{|f(\boldsymbol{x}+\boldsymbol{h}, \boldsymbol{y}+\boldsymbol{k}) - f(\boldsymbol{x}, \boldsymbol{y}) - f(\boldsymbol{h}, \boldsymbol{y}) - f(\boldsymbol{x}, \boldsymbol{k})|}{|(\boldsymbol{h}, \boldsymbol{k})|} = 0\,,$$

которое равносильно утверждению теоремы. ▶

В качестве **примеров**, в которых используется равенство (15.18), найдем дифференциалы тех билинейных отображений, примеры которых были приведены выше:

$$\begin{aligned}
&D(x\cdot y)(h,k) = h\cdot y + x\cdot k\,,\\
&D\langle \boldsymbol{x}, \boldsymbol{y}\rangle(\boldsymbol{h}, \boldsymbol{k}) = \langle \boldsymbol{h}, \boldsymbol{y}\rangle + \langle \boldsymbol{x}, \boldsymbol{k}\rangle\,,\\
&D\,[\boldsymbol{x}\times\boldsymbol{y}]\,(\boldsymbol{h}, \boldsymbol{k}) = \boldsymbol{h}\times\boldsymbol{y} + \boldsymbol{x}\times\boldsymbol{k}\,.
\end{aligned}$$

Отметим, что теорема, аналогичная теореме 29, справедлива для любых полилинейных отображений (а не только для билинейных).

Теорема 30. *Если функции* $f\,,\ g\ :\ \mathcal{A} \longrightarrow \mathbb{R}$, $\mathcal{A} \subset \mathbb{R}^n$, *дифференцируемы в точке* $\boldsymbol{x} \in \mathcal{A}$, *то их сумма, произведение, а при* $g(\boldsymbol{x}) \neq 0$ *и частное дифференцируемы в точке* $\boldsymbol{x}$, *причем*

$$(f+g)'(\boldsymbol{x}) = f'(\boldsymbol{x}) + g'(\boldsymbol{x})\,; \tag{15.19}$$

$$(f\cdot g)'(\boldsymbol{x}) = g(\boldsymbol{x})f'(\boldsymbol{x}) + f(\boldsymbol{x})g'(\boldsymbol{x})\,; \tag{15.20}$$

$$\left(\frac{f}{g}\right)'(\boldsymbol{x}) = \frac{g(\boldsymbol{x})f'(\boldsymbol{x}) - f(\boldsymbol{x})g'(\boldsymbol{x})}{[g(\boldsymbol{x})]^2}\,. \tag{15.21}$$

◀ Рассмотрим сначала частные случаи. Функция

$$s\ :\ (u,v) \longmapsto u+v$$

дифференцируема как линейная, причем $s' = s$. Функция $p\ :\ (u,v) \longmapsto u\cdot v$ дифференцируема как билинейная , причем $p'(u,v)(h,k) = hv + ku$. Докажем дифференцируемость частного $q\ :\ (u,v) \longmapsto \dfrac{u}{v}$ при $v \neq 0$ и равенство

$$q'(u,v)(h,k) = \frac{hv - ku}{v^2}\,. \tag{15.22}$$

С этой целью произведем следующие оценки:

$$0 \leqslant \frac{|q(u+h,v+k) - q(u,v) - q'(u,v)(h,k)|}{|(h,k)|} =$$
$$= \frac{\left|\dfrac{u+h}{v+k} - \dfrac{u}{v} - \dfrac{hv-ku}{v^2}\right|}{\sqrt{h^2+k^2}} = \frac{\left|\dfrac{vh-uk}{v(v+k)} - \dfrac{vh-uk}{v^2}\right|}{\sqrt{h^2+k^2}} =$$
$$= \frac{|vh-uk|}{\sqrt{h^2+k^2}} \cdot \left|\frac{1}{v}\left(\frac{1}{v+k} - \frac{1}{v}\right)\right| \leqslant \sqrt{u^2+v^2} \cdot \frac{|k|}{v(v+k)} \to 0$$

при $(h,k) \to (0,0)$.

Представляя сумму, произведение и частное функций в виде следующих композиций

$$f+g = s \circ (f,g)\,, \quad f \cdot g = p \circ (f,g)\,, \quad \frac{f}{g} = q \circ (f,g)\,,$$

заключаем, что они дифференцируемы как композиции дифференцируемых отображений. Производные их вычислим, применяя «цепное правило».

Для суммы имеем

$$D(f+g)(\boldsymbol{x}) = D\left(s \circ (f,g)\right)(\boldsymbol{x}) = Ds \circ D(f,g)(\boldsymbol{x}) =$$
$$= s \circ (Df(\boldsymbol{x})), Dg(\boldsymbol{x}) = s(Df(\boldsymbol{x}), Dg(\boldsymbol{x})) = Df(\boldsymbol{x}) + Dg(\boldsymbol{x})\,.$$

Для произведения имеем

$$D(f \cdot g)(\boldsymbol{x}) = D\left(p \circ (f,g)\right)(\boldsymbol{x}) = Dp(f(\boldsymbol{x}), g(\boldsymbol{x})) \circ D(f,g)(\boldsymbol{x}) =$$
$$= Dp(f(\boldsymbol{x}), g(\boldsymbol{x}))(Df(\boldsymbol{x}), Dg(\boldsymbol{x})) = g(\boldsymbol{x})Df(\boldsymbol{x}) + f(\boldsymbol{x})Dg(\boldsymbol{x})\,.$$

Для частного имеем

$$D\left(\frac{f}{g}\right)(\boldsymbol{x}) = D(q \circ (f,g))(\boldsymbol{x}) = Dq(f(\boldsymbol{x}), g(\boldsymbol{x})) \circ D(f,g)(\boldsymbol{x}) =$$
$$= Dq(f(\boldsymbol{x}), g(\boldsymbol{x}))(f(\boldsymbol{x}), g(\boldsymbol{x})) = \frac{g(\boldsymbol{x})Df(\boldsymbol{x}) - f(\boldsymbol{x})Dg(\boldsymbol{x})}{[g(\boldsymbol{x})]^2}\,. \ \blacktriangleright$$

Замечание. Теперь нам гарантирована дифференцируемость тех отображений $f : \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, координатные функции которых можно получить с помощью конечного числа арифметических операций и операций образования композиций из отображений проектирования

$\pi^1, \ldots, \pi^n$ и из всех тех функций, дифференцируемость которых была установлена ранее. Однако нахождение производных даже сравнительно простых функций может оказаться довольно сложным делом. Рассотрим в качестве примера функцию $f : (x, y) \longmapsto \sin(xy^2)$. Для вычисления ее дифференциала представим ее в виде композиции $f = \sin \circ p \circ (\pi^1, (\pi^2)^2)$ и применим «цепное правило». Тогда получим

$$\begin{aligned} Df(x,y)(h,k) &= (D\sin)(xy^2) \circ (Dp)(x,y^2) \circ (D\pi^1(x,y), D(\pi^2)^2(x,y))(h,k) = \\ &= \cos(xy^2) \cdot (Dp)(x,y^2) \circ (\pi^1(x,y)(h,k), 2\pi^2(x,y) \cdot \pi^2(x,y)(h,k)) = \\ &= \cos(xy^2) \cdot (Dp)(x,y^2)(\pi^1(h,k), 2y\pi^2(h,k)) = \\ &= \cos(xy^2) \cdot (Dp)(x,y^2)(h, 2yk) = (y^2h + 2xyk) \cdot \cos(xy^2)\,. \end{aligned}$$

Получилось довольно громоздкое вычисление. Возникает вопрос: можно ли подобные вычисления как-то упростить? Оказывается можно, и в следующем параграфе это будет сделано.

§ 2. Частные производные и матрица Якоби

1. Частные производные первого порядка

Пусть $u = f(\boldsymbol{x}) = f(x_1, \ldots, x_n)$ — скалярная функция векторного аргумента, определенная в области $\mathcal{A} \subset \mathbb{R}^n$. Рассматривая ее как функцию одного из переменных x_k при фиксированных значениях остальных переменных, мы приходим к понятию *частной производной* от f по x_k. Обозначая через $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ стандартный базис в $\mathbb{R}^n$, дадим следующее определение.

Определение 24. *Частной производной 1-го порядка от функции f по переменной x_k в точке $\boldsymbol{a} \in \mathcal{A}$ называется предел*

$$D_k f(\boldsymbol{a}) := \lim_{t \to 0} \frac{f(\boldsymbol{a} + t\boldsymbol{e}_k) - f(\boldsymbol{a})}{t}\,. \tag{15.23}$$

Приведем другие, всречающиеся в литературе, обозначения для частных производных

$$D_k f(\boldsymbol{a}) = f'_{x_k}(\boldsymbol{a}) = \frac{\partial f(\boldsymbol{a})}{\partial x_k} = \frac{\partial}{\partial x_k} f(a_1\,, \,\dots\,, \,a_n) :=$$
$$:= \lim_{h\to 0} \frac{f(\dots, a_k + h, \dots) - f(\dots, a_k, \dots)}{h}\,. \quad (15.24)$$

Здесь под знаком функции f многоточиями заменены значения тех переменных, которые не получают приращений (фиксируются). Отметим, что обозначение $\dfrac{\partial}{\partial x_k}$ (с круглой буквой ∂) применяется исключительно для обозначения частных производных функций одного переменного, обозначаемых символом $\dfrac{d}{dx}$. Производные от функций одного переменного принято называть *обыкновенными* производными, чтобы отличить их от частных производных. Учитывая, что частная производная $\dfrac{\partial f(\boldsymbol{a})}{\partial x_k}$ равна обыкновенной производной от функции $u = f(\dots\,, x_k\,, \dots)$ в точке $x_k = a_k$, легко объяснить геометрический смысл частной производной. Именно, частная производная $\dfrac{\partial f(\boldsymbol{a})}{\partial x_k}$ равна угловому каэффициенту касательной к графику функции $u = f(\dots\,, x_k\,, \dots)$ в точке $(a_k\,, f(\boldsymbol{a}))$.

Рассмотрим простые **примеры** на вычисление частных производных.

1) Для функции $f(x\,, y) = \sin(xy^2)$ имеем

$$\frac{\partial f(x\,, y)}{\partial x} = y^2\cos(xy^2)\,; \qquad \frac{\partial f(x\,, y)}{\partial y} = 2xy\cos(xy^2)\,.$$

2) Для функции $g(x\,, y) = x^y$ имеем

$$\frac{\partial g(x\,, y)}{\partial x} = \frac{\partial}{\partial x}x^y = yx^{y-1}\,;$$
$$\frac{\partial g(x\,, y)}{\partial y} = \frac{\partial}{\partial y}x^y = x^y\ln x\,.$$

После небольшой практики можно добиться такой же легкости при вычислении частных производных, как и при вычислении обыкновенных производных.

Отметим теперь применение частных производных для исследования локальных экстремумов функций векторного аргумента (определение локальных экстремумов функций векторного аргумента по форме не отличается от соответствующего определения для функций скалярного аргумента, и мы его здесь не приводим).

Теорема 31 (необходимое условие экстремума). *Если локальный экстремум скалярной функции $f : \mathcal{A} \longrightarrow \mathbb{R}$, определенной в области $\mathcal{A} \subset \mathbb{R}$, достигается в точке $\boldsymbol{a} \in \mathcal{A}$, и если существует конечная частная производная $D_k f(\boldsymbol{a})$, то $D_k f(\boldsymbol{a}) = 0$.*

◀ Обозначим, как обычно, $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\boldsymbol{a} = (a_1, \ldots, a_n)$. Фиксируя значения всех переменных x_j, $j \neq k$, кроме x_k, рассмотрим функцию $u = f(\ldots, x_k, \ldots)$ одного переменного x_k. Эта функция является сужением функции f на одномерное подмножество множества $\mathcal{A}$. Так как функция f имеет экстремум в точке $\boldsymbol{a}$, то и функция $u = f(\ldots, x_k, \ldots)$ имеет экстремум в соответствующей точке $x_k = a_k$. По необходимому условию экстремума функций одного переменного имеем $\dfrac{d}{dx_k} f(\ldots, a_k, \ldots) = 0$, что равносильно такому равенству $D_k f(\boldsymbol{a}) = 0$. ▶

Замечание. Доказанное необходимое условие экстремума не является достаточным. Так, в случае $n = 1$ функция $f : x \longmapsto x^3$ экстремумов не имеет, однако $f'(0) = 0$ (см. рис. 5).

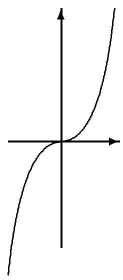

Рис. 5. График функции $x \longmapsto x^3$

В случае $n > 1$ недостаточность условий теоремы 31 может проявляться более эффектно. Рассмотрим, например, функцию

$$(x, y) \longmapsto y^2 - x^2,$$

график которой показан на рис. 6. Для нее имеем

$$\frac{\partial f(0,0)}{\partial x} = \frac{\partial f(0,0)}{\partial y} = 0,$$

однако она не имеет экстремума в точке $(0, 0)$. В самом деле, $f(0, 0) = 0$, но $f(0, y) = y^2 > 0$, $f(x, 0) = -x^2 < 0$ при $y \neq 0$ и $x \neq 0$ соответственно. Эти же неравенства показывают, что обе функции $f(0, y) = y^2$ и $f(x, 0) = -x^2$ имеют экстремумы в начале координат, причем первая имеет минимум, а вторая — максимум.

2. Частные производные высших порядков

Пусть скалярная функция $f\ :\ \mathcal{A}\longrightarrow\mathbb{R}$ определена в области $\mathcal{A}\subset\mathbb{R}^n$.

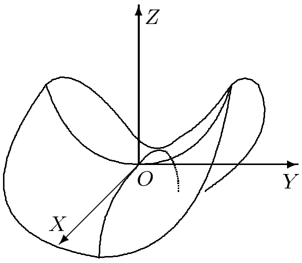

Рис. 6. Гиперболический параболоид

Предположим, что при любом $\boldsymbol{a}\in\mathcal{A}$ существует конечная частная производная $D_k f(\boldsymbol{x})$. Тем самым определяется функция

$$D_k f\ :\ \boldsymbol{x}\longmapsto D_k f(\boldsymbol{x})\,,$$

и можно ставить вопрос о частных производных этой новой функции, которые и называются *частными производными второго порядка от функции* f.

Определение 25. *Частные производные 2-го порядка от функции* f *в точке* $\boldsymbol{a}$ *определяются следующими равенствами*[2]:

$$f''_{x_k x_j}(\boldsymbol{a})=D_{kj}f(\boldsymbol{a})=\frac{\partial^2 f(\boldsymbol{a})}{\partial x_j\partial x_k}:=$$

$$:=D_j\left(D_k f\right)(\boldsymbol{a})=\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_k}\right)(\boldsymbol{a})\,,\quad \textit{где}\ \ k\,,\ j=1\,,\,\dots\,,\,n\,. \qquad (15.25)$$

Частные производные 3-го и более высоких порядков определяются аналогично, по индукции.

Так как в (15.25) индексы j и k могут независимо один от другого принимать значения $1\,,\,\dots\,,\,n$, всего может существовать n^2 различных частных производных 2=го порядка от функции f. Частные производные $\dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_k\partial x_k}=\dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_k^2}$ назваются *чистыми*, а частные произвоные $\dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_k\partial x_j}$ при $k\neq j$ — *смешанными*. Понятия чистых и смешанных производных имеют смысл и для частных производных более высоких порядков.

Образуем $n\times n$-матрицу из частных производных 2-го порядка от скалярной функции f в точке $\boldsymbol{a}$ (называемую также *матрицей*

[2]В первой строке формулы (15.25) приведены обозначения, а во второй — определения частных производных 2-го порядка.

Гессе[3], а ее определитель — *гессианом*):

$$\begin{pmatrix} \dfrac{\partial^2 f(\boldsymbol{a})}{\partial x_1^2} & \dfrac{\partial^2 f(\boldsymbol{a})}{\partial x_2 \partial x_1} & \dots & \dfrac{\partial^2 f(\boldsymbol{a})}{\partial x_n \partial x_1} \\ \dfrac{\partial^2 f(\boldsymbol{a})}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f(\boldsymbol{a})}{\partial x_2^2} & \dots & \dfrac{\partial^2 f(\boldsymbol{a})}{\partial x_n \partial x_2} \\ \dots & \dots & \dots & \dots \\ \dfrac{\partial^2 f(\boldsymbol{a})}{\partial x_1 \partial x_n} & \dfrac{\partial^2 f(\boldsymbol{a})}{\partial x_2 \partial x_n} & \dots & \dfrac{\partial^2 f(\boldsymbol{a})}{\partial x_n^2} \end{pmatrix}. \tag{15.26}$$

Важным для анализа является вопрос о том, симметрична ли эта матрица, т. е. выяснить условия, при выполнении которых имеет место равенство

$$\frac{\partial^2 f(\boldsymbol{a})}{\partial x_j \partial x_k} = \frac{\partial^2 f(\boldsymbol{a})}{\partial x_k \partial x_j}, \quad \text{где} \quad k \neq j\,. \tag{15.27}$$

Замечание. В общем случае равенство (15.27) выполняется не всегда. Рассмотрим, например, функцию

$$f(x\,,\,y) = \begin{cases} xy \cdot \dfrac{x^2 - y^2}{x^2 + y^2}\,, & (x\,,\,y) \neq (0\,,\,0)\,, \\ 0\,, & (x\,,\,y) = (0\,,\,0)\,. \end{cases}$$

Найдем ее частные производные $\dfrac{\partial f}{\partial x}$ и $\dfrac{\partial f}{\partial y}$. При $(x\,,\,y) \neq (0\,,\,0)$ имеем

$$\frac{\partial f(x\,,\,y)}{\partial x} = y\frac{x^2 - y^2}{x^2 + y^2} + xy\frac{2x(x^2 + y^2) - 2x(x^2 - y^2)}{(x^2 + y^2)^2} = \\ = y\frac{x^2 - y^2}{x^2 + y^2} + \frac{4x^2 y^3}{(x^2 + y^2)^2}\,.$$

при $(x\,,\,y) = (0\,,\,0)$ имеем

$$\frac{\partial f(0\,,\,0)}{\partial x} = \lim_{h \to 0} \frac{f(h\,,\,0) - f(0\,,\,0)}{h} = \lim_{h \to 0} 0 = 0\,.$$

[3] *Гессе* Людвиг Отто (1811—1874) немецкий математик.

Аналогично предыдущему находим

$$\frac{\partial f(x\,,\,y)}{\partial y}=x\frac{x^2-y^2}{(x^2+y^2)^2}-xy\frac{2y(x^2+y^2)+2y(x^2-y^2)}{(x^2+y^2)^2}=$$

$$=x\frac{x^2-y^2}{(x^2+y^2)^2}-\frac{4x^3y^2}{(x^2+y^2)^2}\quad\text{при}\quad (x\,,\,y)\neq(0\,,\,0)\,;$$

$$\frac{\partial f(0\,,\,0)}{\partial y}=\lim_{k\to 0}\frac{f(0\,,\,k)-f(0\,,\,0)}{k}=\lim_{k\to 0}0=0\,.$$

Теперь находим смешанные частные производные

$$\frac{\partial^2 f(0\,,\,0)}{\partial y\partial x}=\lim_{k\to 0}\frac{1}{k}\left(\frac{\partial f(0\,,\,k)}{\partial x}-\frac{\partial f(0\,,\,0)}{\partial x}\right)=\lim_{k\to 0}\frac{-k}{k}=-1\,;$$

$$\frac{\partial^2 f(0\,,\,0)}{\partial x\partial y}=\lim_{h\to 0}\frac{1}{h}\left(\frac{\partial f(h\,,\,0)}{\partial y}-\frac{\partial f(0\,,\,0)}{\partial y}\right)=\lim_{h\to 0}\frac{h}{h}=1\,.$$

Таким образом, имеем $\dfrac{\partial^2 f(0\,,\,0)}{\partial y\partial x}\neq\dfrac{\partial^2 f(0\,,\,0)}{\partial x\partial y}$, и значит, равенство (15.27) не выполняется. Следующая теорема дает достаточные условия, обеспечивающие выполнение равенства (15.27).

Теорема 32 (о смешанных производных). *Если смешанные частные производные $D_{kj}f$ и $D_{jk}f$ функции f существуют в некоторой окрестности точки $\boldsymbol{a}$ и непрерывны в точке $\boldsymbol{a}$, то $D_{kj}f(\boldsymbol{a})=D_{jk}f(\boldsymbol{a})$.*

◀ Так как при вычислении частных производных 2-го порядка фиксируются все координаты вектора $\boldsymbol{x}$, кроме двух, то с самого начала можно считать, что f — функция двух переменных x и y. Ее частные производные функции $\dfrac{\partial^2 f}{\partial y\partial x}$ и $\dfrac{\partial^2 f}{\partial x\partial y}$ пусть определены в некотором прямоугольнике Π с центром в точке $(a\,,\,b)$ и непрерывны в точке $(a\,,\,b)$. Чтобы доказать равенство

$$\frac{\partial^2 f(a\,,\,b)}{\partial y\partial x}=\frac{\partial^2 f(a\,,\,b)}{\partial x\partial y}\,,\tag{15.28}$$

рассмотрим вспомогательное выражение

$$W:=\frac{1}{hk}\left[f(a+h\,,\,b+k)-f(a+h\,,\,b)-f(a\,,\,b+k)+f(a\,,\,b)\right]\,,$$

где $h\neq 0\,,\;k\neq 0$ — достаточно малые числа. Функция

$$\varphi\;:\;x\longmapsto\frac{f(x\,,\,b+k)-f(x\,,\,b)}{k}$$

дифференцируема при $x \in [a - |h|\,, a + |h|]$, а ее производная равна

$$\varphi'(x) = \frac{1}{k}\left[\frac{\partial f(x\,, b + k)}{\partial x} - \frac{\partial f(x\,, b)}{\partial x}\right].$$

С помощью функции φ выражение W можно преобразовать, используя теорему Лагранжа «о конечных приращениях», следующим образом:

$$\begin{aligned} W = \frac{1}{h}\,[\varphi(a+h) - \varphi(a)] = \varphi'(a + \theta_1 h) = \\ = \frac{1}{k}\left[\frac{\partial f(a + \theta_1 h\,, b + k)}{\partial x} - \frac{\partial f(a + \theta_1 h\,, b}{\partial x}\right] = \\ = \frac{\partial^2 f(a + \theta_1 h\,, b + \theta_2 k)}{\partial y \partial x}\,, \quad \text{где} \quad \theta_1\,, \theta_2 \in [0\,, 1]\,. \end{aligned} \tag{15.29}$$

Вводя, далее, функцию

$$\psi\,.:\; y \longmapsto \frac{f(a+h\,, y) - f(a\,, y)}{h},$$

имеем

$$\psi'(y) = \frac{1}{h}\left[\frac{\partial f(a+h\,, y)}{\partial y} - \frac{\partial f(a\,, y)}{\partial y}\right].$$

С помощью функции ψ, опять используя теорему Лагранжа, преобразуем выражение W следующим образом:

$$\begin{aligned} W = \frac{1}{k}\,[\psi(b+k) - \psi(b)] = \psi'(b + \theta_3 k) = \\ = \frac{1}{h}\left[\frac{\partial f(a + h\,, b + \theta_3 k)}{\partial y} - \frac{\partial f(a\,, b + \theta_3 k)}{\partial y}\right] = \\ = \frac{\partial^2 f(a + \theta_4 h\,, b + \theta_3 k)}{\partial x \partial y}\,, \quad \text{где} \quad \theta_3\,, \theta_4 \in [0\,, 1]\,. \end{aligned} \tag{15.30}$$

Приравнивая между собой правые части равенств (15.29) и (15.30), имеем

$$\frac{\partial^2 f(a + \theta_1 h\,, b + \theta_2 k)}{\partial y \partial x} = \frac{\partial^2 f(a + \theta_4 h\,, b + \theta_3 k)}{\partial x \partial y}\,.$$

Переходя здесь к пределу при $(h\,, k) \to (0\,, 0)$ и используя непрерывность смешанных частных производных в точке $(a\,, b)$, получим равенство (15.28). ▶

3. Матрица Якоби

Производная в точке $\boldsymbol{a}$ отображения $f : \mathcal{A} \longrightarrow \mathbb{R}^m$ была определена как некоторый линейный оператор $f'(\boldsymbol{a}) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Из алгебры известно, что всякий линейный оператор из $\mathbb{R}^n$ в $\mathbb{R}^m$ можно задать матрицей размеров $m \times n$, если в этих пространствах зафиксировать базисы. Наша ближайшая цель — получить выражение для матрицы линейного оператора $f'(\boldsymbol{a})$, если в пространствах $\mathbb{R}^n$ и $\mathbb{R}^m$ зафиксированы стандартные базисы. Знание этой матрицы значительно облегчает вычисление производных.

Определение 26. *Матрицей Якоби*[4] *в точке* $\boldsymbol{a} \in \mathcal{A}$ *отображения* $f : \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, *называется матрица размеров* $m \times n$, *составленная из частных производных 1-го порядка координатных функций отображения* f *в точке* $\boldsymbol{a}$:

$$\big(D_k f^j(\boldsymbol{a})\big)_{\substack{j=1,\ldots,n\\ k=1,\ldots m}} = \begin{pmatrix} D_1 f^1(\boldsymbol{a}) & \ldots & D_n f^1(\boldsymbol{a}) \\ \ldots & \ldots & \ldots \\ D_1 f^m(\boldsymbol{a}) & \ldots & D_n f^m(\boldsymbol{a}) \end{pmatrix} \qquad (15.31)$$

В связи с этим определением удобно ввести понятие *градиента.*

Определение 27. *Градиентом скалярной функции* $f : \mathcal{A} \longrightarrow \mathbb{R}$, $\mathcal{A} \subset \mathbb{R}^n$, *в точке* $\boldsymbol{a}$ *называется* n*-мерный вектор (строка), координаты которого равны частным производным функции* f *в точке* $\boldsymbol{a}$, *т. е.*

$$\mathbf{grad}\, f(\boldsymbol{a}) := (D_1 f(\boldsymbol{a}), \ldots, D_n f(\boldsymbol{a}))\,. \qquad (15.32)$$

Сравнивая правые части равенств (15.31) и (15.32), видим, что градиент — это тот частный случай матрицы Якоби, когда она состоит из одной строки.

Пусть $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ — линейный оператор. Символом $[L]$ условимся обозначать его матрицу в стандартных базисах.

Теорема 33. *Предположим, что отображение* $f : \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, *дифференцируемо в точке* $\boldsymbol{a} \in \mathcal{A}$. *Тогда*:

(a) *в точке* $\boldsymbol{a}$ *существуют все частные производные всех координатных функций* f^k *отображения* f;

---

[4] *Якóби* Карл Густав Якоб (1804—1851) — немецкий математик.

(b) *матрица линейного оператора* $f'(\boldsymbol{a})$ *в стандартных базисах совпадает с матрицей Якоби, т. е.*

$$[f'(\boldsymbol{a})] = \begin{pmatrix} D_1 f^1(\boldsymbol{a}) & \dots & D_n f^1(\boldsymbol{a}) \\ \dots & \dots & \dots \\ D_1 f^m(\boldsymbol{a}) & \dots & D_n f^m(\boldsymbol{a}) \end{pmatrix}. \tag{15.33}$$

◀ Так как отображение f дифференцируемо в точке $\boldsymbol{a}$, то в этой точке дифференцируемы все его координатные функции f^j, причем

$$Df(\boldsymbol{a})(\boldsymbol{h}) = \sum_{k=1}^{m} Df^j(\boldsymbol{a})(\boldsymbol{h})\boldsymbol{e}_j^t, \tag{15.34}$$

где стандартный базис в $\mathbb{R}^m$ взят в виде:

$$\boldsymbol{e}_1^t = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \boldsymbol{e}_2^t = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ \dots, \ \boldsymbol{e}_m^t = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Здесь и в дальнейшем верхний индекс t над матрицами означает транспонирование. Далее, по определению дифференцируемости для координатных функций f_j имеем следующие равенства:

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{|f^j(\boldsymbol{a}+\boldsymbol{h}) - f^j(\boldsymbol{a}) - Df^j(\boldsymbol{a})(\boldsymbol{h})|}{|\boldsymbol{h}|} = 0.$$

Полагая здесь $\boldsymbol{h} = t\boldsymbol{e}_k$, где $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ — стандартный базис в $\mathbb{R}^n$, получим

$$\lim_{t\to 0} \left| \frac{f^j(\boldsymbol{a}+t\boldsymbol{e}_k) - f^j(\boldsymbol{a})}{t} - Df^j(\boldsymbol{a})(\boldsymbol{e}_k) \right| = 0.$$

Отсюда видно, что $D_k f^j(\boldsymbol{a}) = Df^j(\boldsymbol{a})(\boldsymbol{e}_k)$, где $k = 1, \dots, n$; $j = 1, \dots, m$, т. е. существуют все частные производные, и утверждение (a) доказано.

Для доказательства утверждения (b) положим $\boldsymbol{h} = \sum\limits_{k=1}^{n} h_k \boldsymbol{e}_k$ и преобразуем дифференциал $Df(\boldsymbol{a})(\boldsymbol{h})$:

$$Df(\boldsymbol{a})(\boldsymbol{h}) = \sum_{j=1}^{m} Df^j(\boldsymbol{a})(\boldsymbol{h}) \cdot \boldsymbol{e}_j^t = \sum_{j=1}^{m} Df^j(\boldsymbol{a}) \left(\sum_{k=1}^{n} h_k \boldsymbol{e}_k \right) \cdot \boldsymbol{e}_j^t =$$

$$= \sum_{j=1}^{m} \left(\sum_{k=1}^{n} D_k f^j(\boldsymbol{a}) h_k \right) \boldsymbol{e}_j^t = \begin{pmatrix} \sum\limits_{k=1}^{n} D_k f^1(\boldsymbol{a}) h_k \\ \vdots \\ \sum\limits_{k=1}^{n} D_k f^m(\boldsymbol{a}) h_k \end{pmatrix} =$$

$$= \begin{pmatrix} D_1 f^1(\boldsymbol{a}) & \dots & D_n f^1(\boldsymbol{a}) \\ \dots & \dots & \dots \\ D_1 f^m(\boldsymbol{a}) & \dots & D_n f^m(\boldsymbol{a}) \end{pmatrix} \cdot \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix},$$

где умножение матриц производится по стандартному правилу «строка на столбец». Итак, равенство (15.33) установлено. ▶

Замечание. Равенство

$$Df(\boldsymbol{a})(\boldsymbol{h}) = \begin{pmatrix} D_1 f^1(\boldsymbol{a}) & \dots & D_n f^1(\boldsymbol{a}) \\ \dots & \dots & \dots \\ D_1 f^m(\boldsymbol{a}) & \dots & D_n f^m(\boldsymbol{a}) \end{pmatrix} \cdot \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}, \tag{15.35}$$

дает представление дифференциала отображения f через частные производные и в связи с этим широко используется. В частном случае, когда f — скалярнгая функция, она упрощается и может быть записана в следующих формах:

$$f'(\boldsymbol{a})(\boldsymbol{h}) = D_1 f(\boldsymbol{a}) h_1 + \ldots + D_n f(\boldsymbol{a}) h_n = \langle \mathbf{grad} f(\boldsymbol{a}); \boldsymbol{h} \rangle. \tag{15.36}$$

Перепишем это равенство в классических обозначениях, полагая $f'(\boldsymbol{a})(\boldsymbol{h}) = df$, $h_k = dx_k$:

$$df(\boldsymbol{a}) = \frac{\partial f(\boldsymbol{a})}{\partial x_1} dx_1 + \ldots + \frac{\partial f(\boldsymbol{a})}{\partial x_n} dx_n .$$

Запись такого типа, хотя и часто встречается в литературе, является менее точной, чем (15.36).

Теорема 34. *Предположим, что отображение* $f : \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, *дифференцируемо в точке* $\boldsymbol{a} \in \mathcal{A}$, *а отображение* $g : \mathcal{B} \longrightarrow \mathbb{R}^p$, $\mathcal{B} \subset \mathbb{R}^m$, *дифференцируемо в точке* $\boldsymbol{b} := f(\boldsymbol{a}) \in f(\mathcal{A}) \subset$

$\subset \mathcal{B}$. *Тогда матрица Якоби композиции $g \circ f$ равна произведению (в том же порядке) матриц Якоби отображений f и g, вычисленных в соответствущих точках, т. е.*

$$\left[(g \circ f)'(\boldsymbol{a})\right] = [g'(\boldsymbol{b})] \cdot [f'(\boldsymbol{a})]\,. \tag{15.37}$$

◀ По теореме о производной композиции имеем

$$(g \circ f)'(\boldsymbol{a}) = g'(\boldsymbol{b}) \circ f'(\boldsymbol{a})\,. \tag{15.38}$$

Из алгебры известно, что матрица композиции линейных операторов равна произведению (в том же порядке) матриц операторов, образующих композицию. Таким образом, из (15.38) следует (15.37). ▶

Пример. Найти матрицу Якоби отображения $g \circ f \;:\; (x\,,\,y) \longmapsto (z_1\,,\,z_2)$, заданного следующими системами равенств:

$$g : \begin{cases} z_1 = u + v - w\,, \\ z_2 = u - v + w\,, \end{cases} \qquad f : \begin{cases} u = x + y\,, \\ v = xy\,, \\ w = \dfrac{x}{y}\,. \end{cases}$$

◀ По формуле (15.37) имеем

$$[(g \circ f)'(x\,,\,y)] = \begin{pmatrix} \dfrac{\partial z_1}{\partial u} & \dfrac{\partial z_1}{\partial v} & \dfrac{\partial z_1}{\partial w} \\ \dfrac{\partial z_2}{\partial u} & \dfrac{\partial z_2}{\partial v} & \dfrac{\partial z_2}{\partial w} \end{pmatrix} \cdot \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \\ \dfrac{\partial w}{\partial x} & \dfrac{\partial w}{\partial y} \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ y & x \\ \dfrac{1}{y} & -\dfrac{x}{y^2} \end{pmatrix} = \begin{pmatrix} 1 + y - \dfrac{1}{y} & 1 + x + \dfrac{x}{y^2} \\ 1 - y + \dfrac{1}{y} & 1 - x - \dfrac{x}{y^2} \end{pmatrix}. \;▶$$

4. Производные по вектору и по направлению

Пусть $\boldsymbol{v} \in \mathbb{R}^n$ — ненулевой вектор, а $f : \mathcal{A} \longrightarrow \mathbb{R}$, $\mathcal{A} \subset \mathbb{R}^n$ — скалярная функция векторного аргумента.

Определение 28. *Производной от функции f в точке $\boldsymbol{a}$ по вектору $\boldsymbol{v}$ называется предел*

$$\frac{\partial f(\boldsymbol{a})}{\partial \boldsymbol{v}} := \lim_{t\to+0} \frac{f(\boldsymbol{a}+t\boldsymbol{v}) - f(\boldsymbol{a})}{t}. \tag{15.39}$$

Считается, что производная по вектору существует, если этот предел существует и является числом.

Теорема 35. *Если функция f дифференцируема в точке $\boldsymbol{a}$, то ее производная по любому вектору $\boldsymbol{v} \neq \boldsymbol{0}$ существует, причем*

$$\frac{\partial f(\boldsymbol{a})}{\partial \boldsymbol{v}} = Df(\boldsymbol{a})(\boldsymbol{v}) = \langle \mathbf{grad} f(\boldsymbol{a}); \boldsymbol{v}\rangle. \tag{15.40}$$

◀ Так как функция f дифференцируема в точке $\boldsymbol{a}$, то имеем

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{|f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) - f'(\boldsymbol{a})(\boldsymbol{h})|}{|\boldsymbol{h}|} = 0.$$

Полагая здесь $\boldsymbol{h} = t\boldsymbol{v}$, $t > 0$ и умножая на $|\boldsymbol{v}|$, получим

$$\lim_{t\to+0} \left| \frac{f(\boldsymbol{a}+t\boldsymbol{v}) - f(\boldsymbol{a})}{t} - f'(\boldsymbol{a})(\boldsymbol{v}) \right| = 0,$$

что равносильно равенствам (15.40). ▶

Замечание. *Производной функции f по направлению $\boldsymbol{l}$* называется производная по вектору $\boldsymbol{l}$, длина которого равна 1. Из теоремы 35 вытекает существование производной по любому направлению $\boldsymbol{l} = (l_1, \dots, l_n)$ и равенство

$$\frac{\partial f(\boldsymbol{a})}{\partial \boldsymbol{l}} = \langle \mathbf{grad} f(\boldsymbol{a}); \boldsymbol{l}\rangle = \frac{\partial f(\boldsymbol{a})}{\partial x_1}\cdot l_1 + \ldots + \frac{\partial f(\boldsymbol{a})}{\partial x_n}\cdot l_n. \tag{15.41}$$

Отсюда, в частности, следует, что производные по направлениям векторов стандартного базиса равны соответствующим частным производным, т. е.

$$\frac{\partial f(\boldsymbol{a})}{\partial \boldsymbol{e}_k} = \frac{\partial f(\boldsymbol{a})}{\partial x_k}, \ k = 1, \dots, n.$$

Кроме того, *справедливо неравенство*

$$\left|\frac{\partial f(\boldsymbol{a})}{\partial \boldsymbol{l}}\right| \leqslant |\mathbf{grad} f(\boldsymbol{a})|. \tag{15.42}$$

◀ Действительно, применяя к правой части равенства (15.41) неравенство Коши — Буняковского и учитывая, что $|\boldsymbol{l}| = 1$, получим неравенство (15.42). Более того, равенство в этом неравенстве достигается тогда и только тогда, кагда векторы $\boldsymbol{l}$ и $\mathbf{grad} f(\boldsymbol{a})$ коллинеарны. ▶

5. Достаточные условия дифференцируемости

Дифференцируемость скалярной функции скалярного аргумента в точке равносильна существованию в этой точке конечной производной. В случае отображений из $\mathbb{R}^n$ в $\mathbb{R}^m$ ситуация усложняется. В силу теоремы 33 из дифференцируемости отображения в точке следует существование в этой точке всех его частных проихводных 1-го порядка в этой точке. Обратное, однако, неверно, т. е. из факта существования конечных частных производных (матрицы Якоби) в точке *не следует* дифференцируемость отображения в этой точке. Покажем это на примере.

Пример. Рассмотрим функцию

$$f \,:\, (x\,,\,y) \longmapsto f(x\,,\,y) := \begin{cases} \dfrac{x^2 y}{x^2+y^2} \quad \text{при} \quad (x\,,\,y) \neq (0\,,\,0)\,, \\ 0 \quad \text{при} \quad (x\,,\,y) = (0\,,\,0)\,. \end{cases}$$

Для нее имеем

$$\frac{\partial f(0\,,\,0)}{\partial x} = \lim_{h\to 0} \frac{f(h\,,\,0) - f(0\,,\,0)}{h} = \lim 0 = 0\,,$$

$$\frac{\partial f(0\,,\,0)}{\partial y} = \lim_{k\to 0} \frac{f(0\,,\,k) - f(0\,,\,0)}{k} = \lim 0 = 0\,.$$

Если бы данная функция была дифференцируемой в точке $(0\,,\,0)$, то ее производная была бы равна нуль-оператору, так как было бы

$$Df(0\,,\,0)(h\,,\,k) = \frac{\partial f(0\,,\,0)}{\partial x}\cdot h + \frac{\partial f(0\,,\,0)}{\partial y}\cdot k \equiv 0\,.$$

Кроме того, должно было бы выполняться равенство

$$\lim_{(h,k)\to(0,0)} \frac{|f(h\,,\,k) - f(0\,,\,0) - Df(0\,,\,0)(h\,,\,k)|}{\sqrt{h^2+k^2}} = 0\,,$$

равносильное следующему

$$\lim_{(h,k)\to(0,0)} \frac{|h^2 k|}{(h^2+k^2)^{3/2}} = 0\,.$$

Но этот последний предел на самом деле не существует, так как, переходя к полярным координатам $h = r\cos\varphi$, $k = r\sin\varphi$, можно преобразовать его к виду

$$\lim_{(h,k)\to(0,0)} \frac{|h^2 k|}{(h^2+k^2)^{3/2}} = \lim_{r\to+0} \cos^2\varphi\cdot|\sin\varphi| = \cos^2\varphi\cdot|\sin\varphi|\,,$$

где полученный результат зависит от направления φ. Получено противоречие.

Однако при некоторых ограничениях на частные производные функции дифференцируемость имеет место, как показывает следующая теорема.

Теорема 36. *Если все частные производные функции* $D_k f^j$ *отображения* $f : \mathcal{A} \longrightarrow \mathbb{R}^m$, $\mathcal{A} \subset \mathbb{R}^n$, *существуют в окрестности точки* $\boldsymbol{a} \in \mathcal{A}$ *и непрерывны в самóй точке* $\boldsymbol{a}$, *то отображение* f *дифференцируемо в этой точке.*

◀ Так как дифференцируемость векторнозначной функции равносильна дифференцируемости всех ее координатных функций, то достаточно с самого начала ограничиться только случаем скалярной функции векторного аргумента. Далее, доказательство для случая функций от n переменных в идейном плане не отличается от доказательства для случая функций от двух переменных. Различие между ними касается лишь технических деталей, поэтому ограничимся только случаем функции двух переменных $z = f(x\,,\,y)$, для которой условия теоремы пусть выполняются в окрестности точки $(a\,,\,b)$. В этом случае предполагаемый дифференциал должен иметь следующий вид:

$$Df(a\,,\,b)(h\,,\,k) = \frac{\partial f(a\,,\,b)}{\partial x}\cdot h + \frac{\partial f(a\,,\,b)}{\partial y}\cdot k\,. \qquad (15.43)$$

Используя теорему Лагранжа «о конечных приращениях», преобразуем приращение функции f:

$$\begin{gathered} f(a+h\,,\,b+k) - f(a\,,\,b) = \\ = [f(a+h\,,\,b+k) - f(a\,,\,b+k)] + [f(a\,,\,b+k) - f(a\,,\,b)] = \\ = \frac{\partial f(a+\theta_1 h\,,\,b+k)}{\partial x}\cdot h + \frac{\partial f(a\,,\,b+\theta_2 k)}{\partial y}\cdot k\,, \quad \text{где} \;\; \theta_1\,,\,\theta_2 \in [0\,,\,1]\,. \end{gathered}$$

Используя, далее, полученные выражения и неравенство Коши —

Буняковского, имеем

$$0 \leqslant \frac{\left| f(a+h\,,\, b+k) - f(a\,,\, b) - \left[\dfrac{\partial f(a\,,\, b)}{\partial x} \cdot h + \dfrac{\partial f(a\,,\, b)}{\partial y} \cdot k \right] \right|}{\sqrt{h^2+k^2}} =$$

$$= \left| \left[\frac{\partial f(a+\theta_1 h\,,\, b+k)}{\partial x} - \frac{\partial f(a\,,\, b)}{\partial x} \right] \cdot h + \right.$$

$$\left. + \left[\frac{\partial f(a\,,\, b+\theta_2 k)}{\partial y} - \frac{\partial f(a\,,\, b)}{\partial y} \right] \cdot k \right| \Big/ \sqrt{h^2+k^2} \leqslant$$

$$\leqslant \sqrt{\left[\frac{\partial f(a+\theta_1 h, b+k)}{\partial x} - \frac{\partial f(a,b)}{\partial x} \right]^2 + \left[\frac{\partial f(a, b+\theta_2 k)}{\partial y} - \frac{\partial f(a,b)}{\partial y} \right]^2} \to$$

$$\to 0 \quad \text{при} \quad (h\,,\, k) \to (0\,,\, 0)$$

в силу непрерывности частных производных в точке $(a\,,\, b)$. Таким образом, при $(h\,,\, k) \to (0\,,\, 0)$ имеем

$$\lim \frac{\left| f(a+h\,,\, b+k) - f(a\,,\, b) - \left[\dfrac{\partial f(a\,,\, b)}{\partial x} \cdot h + \dfrac{\partial f(a\,,\, b)}{\partial y} \cdot k \right] \right|}{\sqrt{h^2+k^2}} = 0\,,$$

т. е. функция f дифференцируема в точке $(a\,,\, b)$, а ее дифференциал вычисляется по формуле (15.43). ►

§ 3. Теоремы о «конечных приращениях», дифференциалы высших порядков и формула Тейлора для отображений из $\mathbb{R}^n$ в $\mathbb{R}^m$

1. Теоремы о «конечных приращениях»

Обобщения теоремы Лагранжа о «конечных приращениях» на случай отображений из $\mathbb{R}^n$ в $\mathbb{R}^m$ возможны, но они оказываются различными в зависимости от того, является ли данное отображение скалярнозначным или векторнозначным. Рассмотрение этого вопроса начнем с наиболее простого случая скалярной функции векторного аргумента. Предварительно введем понятие прямолинейного отрезка (или интервала), лежащего в $\mathbb{R}^n$.

Определение 29. *Прямой линией, лежащей в пространстве* $\mathbb{R}^n$, *называется множество всех точек* $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, *задаваемых уравнением*

$$\boldsymbol{x} = \boldsymbol{a} + t\boldsymbol{l} \quad \text{или системой} \quad \begin{cases} x_1 = a_1 + tl_1\,, \\ \ldots\ldots\ldots\ldots\,, \\ x_n = a_n + tl_n\,, \end{cases} \tag{15.44}$$

где $\boldsymbol{a}$ *и* $\boldsymbol{l} \neq \boldsymbol{0}$ *— фиксированные точки,* $t \in \mathbb{R}$ *— переменная, пробегающая всю числовую ось. Если переменная* t *пробегает отрезок или интервал, то его образ при отображении* (15.44) *называется соответственно прямолинейным отрезком или интервалом в* $\mathbb{R}^n$.

Условимся обозначать лежащие в $\mathbb{R}^n$ промежутки теми же символами, что и промежутки числовой оси.

Теорема 37 (первая теорема «о конечных приращениях»). *Пусть скалярная функция* $f : \mathcal{A} \longrightarrow \mathbb{R}$ *дифференцируема в области* $\mathcal{A} \subset \mathbb{R}^n$. *Если* $[\boldsymbol{a}, \boldsymbol{a} + \boldsymbol{h}] \subset \mathcal{A}$, *то существует точка* $\xi \in (\boldsymbol{a}, \boldsymbol{a} + \boldsymbol{h})$ *такая, что*

$$f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) = f'(\boldsymbol{\xi})(\boldsymbol{h})\,. \tag{15.45}$$

◀ Рассмотрим вспомогательную функцию $F : [0\,, 1] \longrightarrow \mathbb{R}$, определенную равенством $F(t) := f(\boldsymbol{a} + t\boldsymbol{h})$. Она дифференцируема как композиция дифференцируемых функций. Ее производную можно вычислить по формуле

$$F'(t) := \lim_{\tau \to 0} \frac{f(\boldsymbol{a} + (t+\tau)\boldsymbol{h}) - f(\boldsymbol{a} + t\boldsymbol{h})}{\tau} = f'(\boldsymbol{a} + t\boldsymbol{h})(\boldsymbol{h})\,,$$

так как определение этой производной совпадает с определением 28 производной от функции f по вектору $\boldsymbol{h}$ в точке $\boldsymbol{a} + t\boldsymbol{h}$. Применяя к функции F теорему Лагранжа «о конечных приращениях» функций скалярного аргумента, имеем $F(1) - F(0) = F'(\theta)$, где $\theta \in (0\,, 1)$ или, что равносильно, $f(\boldsymbol{a} + \boldsymbol{h}) - f(\boldsymbol{a}) = f'(\boldsymbol{a} + \theta\boldsymbol{h})(\boldsymbol{h})$. Полагая в этом равенстве $\xi := \boldsymbol{a} + \theta\boldsymbol{h}$, получим равенство (15.45). ▶

Приведем простое приложение этой теоремы.

Теорема 38. *Если все частные производные отображения $f : \mathcal{A} \longrightarrow \mathbb{R}^m$ тождественно равны нулю в области $\mathcal{A} \subset \mathbb{R}^n$, то $f(\boldsymbol{x}) \equiv \boldsymbol{c}$ — постоянное отображение.*

◀ Так как постоянство векторнозначного отображения равносильно постояноству всех его координатных функций, то, не ограничивая общности, будем считать, что $m = 1$, т. е. что f — скалярная функция. Так как все частные производные этой функции — постоянные, то они непрерывны в $\mathcal{A}$, откуда на основании теоремы 36 заключаем, что функция $f : \mathcal{A} \longrightarrow \mathbb{R}$ — дифференцируемая, причем $f'(\boldsymbol{x})(\boldsymbol{h}) \equiv 0$. Если $\boldsymbol{a} \in \mathcal{A}$ и $\{|\boldsymbol{x} - \boldsymbol{a}| < r\} \subset \mathcal{A}$, то ввиду теоремы 37 имеем $f(\boldsymbol{x}) - f(\boldsymbol{a}) = f'(\boldsymbol{\xi})(\boldsymbol{x} - \boldsymbol{a}) \equiv 0$, откуда $f(\boldsymbol{x}) \equiv f(\boldsymbol{a})$ при $|\boldsymbol{x} - \boldsymbol{a}| < r$. Таким образом, функция f — *локально постоянная* (т. е. постянная в окрестности каждой точки $\boldsymbol{a} \in \mathcal{A}$). Чтобы показать, что функция f — *глобально постоянная*, введем в рассмотрение два открытых множества:

$$U := \left\{ \boldsymbol{x} \in \mathcal{A} \;\middle|\; f(\boldsymbol{x}) = f(\boldsymbol{a}) \right\} \quad \text{и} \quad V := \left\{ \boldsymbol{x} \in \mathcal{A} \;\middle|\; f(\boldsymbol{x}) \neq f(\boldsymbol{a}) \right\} .$$

Открытость множества U вытекает из локального постоянства функции f, а открытость множества V — из локального свойства непрерывной функции f (теорема 15). Предполагая, что $V \neq \varnothing$, имеем

$$U \neq \varnothing\,,\ V \neq \varnothing\,,\ U \cap V = \varnothing\,,\ U \sqcup V = \mathcal{A}\,,$$

т. е. множество $\mathcal{A}$ не является связным, а в условии теоремы сказано, что $\mathcal{A}$ — область, т. е. связное множество. Полученное противоречие означает, что $V = \varnothing$, и значит, $U = \mathcal{A}$, т. е. функция f — глобально постоянная. ▶

Рассмотрим теперь **примеры**, показывающие, что теорема «о конечных приращениях» в форме равенства (15.45) перестает быть верной для векторнозначных отображений.

1) Вспомним геометрический эквивалент теоремы «о конечных приращениях» для скалярных функций (теорему Лагранжа): *среди внутренних точек графика дифференцируемой функции $f : [a\,, b] \longrightarrow \mathbb{R}$ существует точка M, в которой касательная параллельна хорде* (см. рис. 7). Здесь оказвывается существенным то, что весь этот график лежит в одной плос-

кости. Для пространственных кривых это свойство, вообще говоря, неверно.

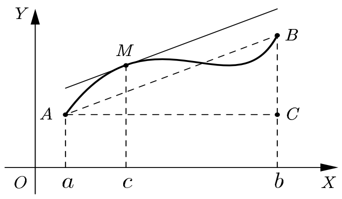

Рис. 7. К теореме Лагранжа

Возьмем, например, гладкий виток спирали, показанный на рис. 8. Его хорда AB —- вертикальная, а касательные во всех точках — «почти горизонтальные», и потому на нем нет точек, в которых касательная была бы параллельна хорде. С другой стороны, этот виток спирали является годографом[5] некоторой дифференцируемой вектор-функции $\boldsymbol{r} : [a\,,\,b] \longrightarrow \mathbb{R}^3$, и значит, для нее невозможно равенство типа (15.45).

2) Пусть вектор-функция $\boldsymbol{r} \;:\; [0\,,\,1] \longrightarrow \mathbb{R}^3$ задана своими координатными функциями

$$\boldsymbol{r} \;:\; \begin{cases} x = t\,, \\ y = t^2\,, \\ z = t^3\,, \end{cases} \qquad t \in [0\,,\,1]\,.$$

Предполагая, что существует $\theta \in [0\,,\,1]$ такое, что $\boldsymbol{r}(1) - \boldsymbol{r}(0) = \boldsymbol{r}'(\theta)$, получим

$$\begin{pmatrix}1\\1\\1\end{pmatrix} - \begin{pmatrix}0\\0\\0\end{pmatrix} = \begin{pmatrix}1\\2\theta\\3\theta^2\end{pmatrix}, \quad \text{откуда} \quad \begin{cases} 1 = 2\theta\,, \\ 1 = 3\theta^2\,. \end{cases}$$

Мы получили несовместную систему уравнений. Таким образом, предполагаемого значения θ не существует.

3) Пусть теперь отображение $f \;:\; \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ задано своими координатными функциями

$$f \;:\; \begin{cases} u = e^x \cos y\,, \\ v = e^x \sin y\,. \end{cases}$$

Предполагая, что для этого отображения справедливо равенство (15.45) и полагая $(a\,,\,b) = (0\,,\,0)$, $(h\,,\,k) = (0\,,\,2\pi)$, получим

$$f(0\,,\,2\pi) - f(0\,,\,0) = f'(\xi\,,\,\eta)(0\,,\,2\pi)\,.$$

Переписывая это равенство в векторно-матричной форме

$$\begin{pmatrix}0\\0\end{pmatrix} = \begin{pmatrix}\cos 2\pi\\ \sin 2\pi\end{pmatrix} - \begin{pmatrix}\cos 0\\ \sin 0\end{pmatrix} =$$
$$= \begin{pmatrix} e^{\xi}\cos\eta & -e^{\xi}\sin\eta \\ e^{\xi}\sin\eta & e^{\xi}\cos\eta \end{pmatrix} \begin{pmatrix}0\\2\pi\end{pmatrix} = 2\pi e^{\xi} \begin{pmatrix}-\sin\eta\\ \cos\eta\end{pmatrix},$$

[5] *Годóграфом* вектор-функции называется множество точек $\{\boldsymbol{r}(t) \mid t \in [a\,,\,b]\}$. Нельзя смешивать понятия годографа и графика, каковым в данном случае является множество пар $\{(t\,,\,\boldsymbol{r}(t)) \mid t \in [a\,,\,b]\}$. Образно говоря, если вектор-функция — это движение материальной точки, то годограф — траектория движения (т. е. дорога), а график — расписание движения.

приходим к противоречию, так как правая часть не равна нулевому вектору ни при каких значениях ξ и η.

Следующая теорема является обобщением теоремы 37 на случай векторнозначных отображений, но уже в форме *неравенства.*

Теорема 39 (вторая теорема «о конечных приращениях»). *Если отображение* $f : \mathcal{A} \longrightarrow \mathbb{R}^m$ *дифференцируемо в области* $\mathcal{A} \subset \mathbb{R}^n$, *и* $[\boldsymbol{a}, \boldsymbol{a}+\boldsymbol{h}] \subset \mathcal{A}$, *то*

$$|f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a})| \leqslant \sup_{\boldsymbol{\xi}} \|f'(\boldsymbol{\xi})\| \cdot |\boldsymbol{h}|, \tag{15.46}$$

где sup *берется по всем* $\boldsymbol{\xi} \in (\boldsymbol{a}, \boldsymbol{a}+\boldsymbol{h})$.

◀ Введем в рассмотрение вспомогательное отображение $F : [0, 1] \longrightarrow \mathbb{R}^m$, задаваемое равенством $F(t) := f(\boldsymbol{a} + t\boldsymbol{h})$, и вспомогательную функцию $\Phi : [0, 1] \longrightarrow \mathbb{R}$, задаваемую равенством

$$\Phi(t) := \langle F(t), F(1) - F(0)\rangle.$$

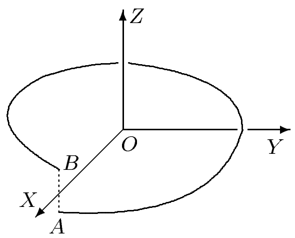

Рис. 8. Виток спирали

Оба эти отображения дифференцируемы как композиции дифференцируемых отображений, а их производные равны

$$F'(t) = f'(f(\boldsymbol{a} + t\boldsymbol{h}))(\boldsymbol{h}), \qquad \Phi'(t) = \langle F'(t), F(1) - F(0)\rangle.$$

Применяя к скалярной функции Φ теорему 37, заключаем, что при некотором $\theta \in (0, 1)$ выполняется равенство $\Phi(1) - \Phi(0) = \Phi'(\theta)$. Далее имеем

$$\begin{aligned}
|F(1) - F(0)|^2 &= \langle F(1) - F(0), F(1) - F(0)\rangle = \\
&= \langle F(1), F(1) - F(0)\rangle - \langle F(0), F(1) - F(0)\rangle = \Phi(1) - \Phi(0) = \\
&= \Phi'(\theta) = \langle F'(\theta), F(1) - F(0)\rangle \leqslant |F'(\theta)| \cdot |F(1) - F(0)| \leqslant \\
&\leqslant |F(1) - F(0)| \cdot \sup_{0<\theta<1} |F'(\theta)|.
\end{aligned}$$

Отсюда находим

$$|F(1) - F(0)|^2 \leqslant |F(1) - F(0)| \cdot \sup_{0<\theta<1} |F'(\theta)|. \tag{15.47}$$

Так как $F(1)-F(0)=f(\boldsymbol{a}+\boldsymbol{h})-f(\boldsymbol{a})$, то в случае $F(1)-F(0)=\mathbf{0}$ неравенство (15.46) очевидно. Если же $F(1)-F(0)\neq 0$, то, сокращая неравенство (15.47) на $|F(1)-F(0)|$, получаем

$$|F(1)-F(0)| \leqslant \sup_{0<\theta<1} |F'(\theta)|\,.$$

Полагая, далее, $\boldsymbol{\xi} := \boldsymbol{a}+\theta\boldsymbol{h} \in (\boldsymbol{a}\ ,\ \boldsymbol{a}+\boldsymbol{h})$, имеем

$$|f(\boldsymbol{a}+\boldsymbol{h})-f(\boldsymbol{a})| = |F(1)-F(0)| \leqslant \sup_{0<\theta<1} |F'(\theta)| =$$
$$= \sup_{0<\theta<1} |f'(\boldsymbol{a}+\theta\boldsymbol{h})(\boldsymbol{h})| = \sup_{\boldsymbol{\xi}} |f'(\xi)(\boldsymbol{h})| \leqslant \sup_{\boldsymbol{\xi}} \|f'(\boldsymbol{\xi})\|\cdot|\boldsymbol{h}|\,,$$

откуда непосредственно следует неравенство (15.46). ▶

2. Дифференциалы высших порядков

Пусть $f\,:\,\mathcal{A}\longrightarrow\mathbb{R}^m$ — отображение, дифференцируемое в области $\mathcal{A}\subset\mathbb{R}^n$. В этом случае в области $\mathcal{A}$ определено операторнозначное отображение

$$f'\ :\ \boldsymbol{x}\longmapsto f'(\boldsymbol{x})\,, \tag{15.48}$$

называемое *производным отображением 1-го порядка*, и можно ставить вопрос о дифференцируемости этого операторнозначного отображения. Производная отображения (15.48) в точке $\boldsymbol{x}$ называется *производной второго порядка отображения f в точке $\boldsymbol{x}$* и обозначается символом $f''(\boldsymbol{x})$. Эта производная есть линейный оператор, действующий из пространства $\mathbb{R}^n$ в пространство линейных операторов $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$, т. е. $f''(\boldsymbol{x})\ :\ \mathbb{R}^n\longrightarrow\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$. Производную $(p+1)$-го порядка можно определять аналогично, по индукции, как производную в точке $\boldsymbol{x}$ от производного отображения p-го порядка, т. е.

$$f^{(p+1)}(\boldsymbol{x}) := \left(f^{(p)}\right)'(\boldsymbol{x})\,. \tag{15.49}$$

Изучение производных высших порядков затруднено тем, что все они оказываются операторнозначными отображениями, а пространства линейных операторов, в которые они действуют, усложняются с увеличением порядков производных. С алгебраической точки зрения производная порядка p в точке $\boldsymbol{x}$ оказывается полилинейным

(p-линейным) отображением от приращений $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_p$ аргумента $\boldsymbol{x}$. Однако изучение алгебры полилинейных отображений программами университетского курса не предусмотрено, поэтому здесь производные высших порядков более подробно рассматривать не будем. Вместо них будем рассматривать *дифференциалы высших порядков*, более простые с точки зрения алгебры и более важные с точки зрения анализа. Поскольку дифференциалы любых порядков можно определять покоординатно, то всюду в этом параграфе будем рассматривать только скалярные функции $f : \mathcal{A} \longrightarrow \mathbb{R}$.

Между понятиями производной и дифференциала 1-го порядка по существу нет различия: производная — линейный оператор, а дифференциал — тот же оператор, но рассматриваемый как линейная форма

$$Df(\boldsymbol{x}) : \boldsymbol{h} \longmapsto Df(\boldsymbol{x})(\boldsymbol{h}).$$

Различие появляется, начиная со 2-го порядка. Определим сначала понятие дифференциала 2-го порядка и изучим некоторые его свойства. Предварительно введем обозначение $Df(\ )(\boldsymbol{h})$ для дифференциала, рассматриваемого как функция переменной $\boldsymbol{x}$ при фиксированном приращении $\boldsymbol{h}$, т. е. $Df(\ )(\boldsymbol{h}) : \boldsymbol{x} \longmapsto Df(\boldsymbol{x})(\boldsymbol{h})$.

Определение 30. *Дифференциалом 2-го порядка функции f в точке $\boldsymbol{x}$ на векторе $\boldsymbol{h}$ (дважды) называется квадратичная форма $\boldsymbol{h} \longmapsto D^2 f(\boldsymbol{x})(\boldsymbol{h})^2$, задаваемая равенством*

$$D^2 f(\boldsymbol{x})(\boldsymbol{h})^2 := D\,[Df(\ )(\boldsymbol{h})]\,(\boldsymbol{x})(\boldsymbol{h}). \tag{15.50}$$

Теорема 40. *Если все частные производные 2-го порядка функции f существуют в окрестности точки $\boldsymbol{x}$ и непрерывны в точке $\boldsymbol{x}$, то $D^2 f(\boldsymbol{x})(\boldsymbol{h})^2$ существует, причем матрица квадратичной формы* (15.50) *совпадает с матрицей Гессе* (15.26).

◀ Пусть $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\boldsymbol{h} = (h_1, \ldots, h_n)$. Функция

$$Df(\ )(\boldsymbol{h}) : \boldsymbol{x} \longmapsto Df(\boldsymbol{x})(\boldsymbol{h}) = \frac{\partial f(\boldsymbol{x})}{\partial x_1} \cdot h_1 + \ldots + \frac{\partial f(\boldsymbol{x})}{\partial x_n} \cdot h_n$$

при фиксированном $\boldsymbol{h}$ дифференцируема как линейная комбинация дифференцируемых функций. Поэтому выражение (15.50) можно представить в следующем виде:

$$D^2 f(\boldsymbol{x})(\boldsymbol{h})^2 := D\left[Df(\ )(\boldsymbol{h})\right](\boldsymbol{x})(\boldsymbol{h}) = D\left(\sum_{k=1}^{n} \frac{\partial f}{\partial x_k} \cdot h_k\right)(\boldsymbol{x})(\boldsymbol{h}) =$$

$$= \sum_{k=1}^{n} h_k \cdot D\left(\frac{\partial f}{\partial x_k}\right)(\boldsymbol{x})(\boldsymbol{h}) = \sum_{k=1}^{n}\sum_{j=1}^{n} h_k h_j \frac{\partial^2 f(\boldsymbol{x})}{\partial x_j \partial x_k}.$$

Отсюда видно, что матрицей квадратичной формы является матрица Гессе, если она симметрична. Но она и в самом деле симметрична в силу теоремы 32 (о смешанных производных). ▶

Для определения дифференциалов более высоких порядков напомним понятие формы степени p от n переменных.

Определение 31. *Формой степени $p \in \mathbb{N}$ от n переменных $h_1, \ldots, h_n$ называется однородная целая рациональная функция степени p от переменных $h_1, \ldots, h_n$, т. е. функция вида*

$$(h_1, \ldots h_n) \quad \longmapsto \sum_{i_1+\ldots+i_n=p} a_{i_1\ldots i_n} \cdot (h_1)^{i_1} \cdot \ldots \cdot (h_n)^{i_n},$$

где $a_{i_1\ldots i_n}$ — числовые коэффициенты, а индексы $i_1, \ldots, i_n$ — неотрицательные целые числа.

Например, линейную, квадратичную и кубичную формы можно представить соответственно в виде

$$\sum_{i=1}^{n} a_i h_i, \quad \sum_{i,k=1}^{n} a_{ik} h_i h_k, \quad \sum_{i,k,j=1}^{n} a_{ikj} h_i h_k h_j.$$

Естественно ожидать, что дифференциалы высших порядков можно определить по индукции как формы соответствующих степеней от приращения. Так оно на самом деле и есть.

Определение 32. *Дифференциалом порядка p функции f в точке $\boldsymbol{x}$ на векторе $\boldsymbol{h}$ (p раз) называется отображение $\boldsymbol{h} \longmapsto D^p f(\boldsymbol{x})(\boldsymbol{h})^p$, задаваемое равенством*

$$D^p f(\boldsymbol{x})(\boldsymbol{h})^p := D\left[D^{p-1} f(\ )(\boldsymbol{h})^{p-1}\right](\boldsymbol{x})(\boldsymbol{h}).$$

Теорема 41. *Если все частные производные p-го порядка функции f существуют в окрестности точки $\boldsymbol{x}$ и непрерывны в точке $\boldsymbol{x}$, то дифференциал $D^p f(\boldsymbol{x})(\boldsymbol{h})^p$ существует и является формой степени p от переменных $h_1, \ldots, h_n$.*

◀ При любых фиксированных значениях переменных $h_1, \ldots, h_n$ функция

$$D^{p-1}f(\ )(\boldsymbol{h})^{p-1} : \boldsymbol{x} \longmapsto D^{p-1}f(\boldsymbol{x})(\boldsymbol{h})^{p-1}$$

является линейной комбинацией частных производных порядка $p-1$ функции f с числовыми коэффициентами. Для этой функции выполняются, таким образом, условия теоремы 36. Значит, эта функция дифференцируема. Найдем и преобразуем ее дифференциал

$$\begin{aligned}
D^p f(\boldsymbol{x})(\boldsymbol{h})^p &:= D\left[D^{p-1}f(\ )(\boldsymbol{h})^{p-1}\right](\boldsymbol{x})(\boldsymbol{h}) = \\
&= D\left[\sum_{i_1,\ldots,i_{p-1}=1}^{n} \frac{\partial^{p-1} f}{\partial x_{i_{p-1}} \ldots \partial x_{i_1}} \cdot h_{i_{p-1}} \cdot \ldots \cdot h_{i_1}\right](\boldsymbol{x})(\boldsymbol{h}) = \\
&= \sum_{i_p=1}^{n} h_{i_p} \cdot \frac{\partial}{\partial x_{i_p}} \left[\sum_{i_1,\ldots,i_{p-1}=1}^{n} \frac{\partial^{p-1} f}{\partial x_{i_{p-1}} \ldots \partial x_{i_1}} \cdot h_{i_{p-1}} \cdot \ldots \cdot h_{i_1}\right](\boldsymbol{x}) = \\
&= \sum_{i_1,\ldots,i_p=1}^{n} \frac{\partial^p f(\boldsymbol{x})}{\partial x_{i_p} \ldots \partial x_{i_1}} \cdot h_{i_p} \cdot \ldots \cdot h_{i_1} .
\end{aligned}$$

Очевидно, что правая части последнего равенства является формой степени p от $h_1, \ldots, h_n$. ▶

Замечание. Учитывая зависимость между дифференциалом и производной по вектору, имеем

$$\begin{aligned}
&Df(\boldsymbol{x})(\boldsymbol{h}) = \frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{h}} = \left(h_1 \frac{\partial}{\partial x_1} + \ldots + h_n \frac{\partial}{\partial x_n}\right) f(\boldsymbol{x}); \\
&D^2 f(\boldsymbol{x})(\boldsymbol{h})^2 = \frac{\partial^2 f(\boldsymbol{x})}{\partial \boldsymbol{h}^2} = \left(h_1 \frac{\partial}{\partial x_1} + \ldots + h_n \frac{\partial}{\partial x_n}\right)^2 f(\boldsymbol{x}); \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots ; \\
&D^p f(\boldsymbol{x})(\boldsymbol{h})^p = \frac{\partial^p f(\boldsymbol{x})}{\partial \boldsymbol{h}^p} = \left(h_1 \frac{\partial}{\partial x_1} + \ldots + h_n \frac{\partial}{\partial x_n}\right)^p f(\boldsymbol{x}).
\end{aligned}$$

Если все частные производные порядка p непрерывны в точке $\boldsymbol{x}$, то в силу теоремы о смешанных производных операторы $\dfrac{\partial}{\partial x_i}$ и $\dfrac{\partial}{\partial x_k}$ перестановочны между собой. Поэтому со степенью оператора $\left(h_1 \dfrac{\partial}{\partial x_1} + \ldots + h_n \dfrac{\partial}{\partial x_n}\right)^p$ можно обращаться так, как со степенью обычного многочлена. Применив полиномиальную формулу[6], получим следующее представление дифференциала p-го порядка:

$$D^p f(\boldsymbol{x})(\boldsymbol{h})^p = \sum_{i_1+\ldots+i_n=p} \frac{p!}{i_1!\cdot\ldots\cdot i_n!}\cdot\frac{\partial^p f(\boldsymbol{x})}{(\partial x_1)^{i_1}\ldots(\partial x_n)^{i_n}}\cdot(h_1)^{i_1}\cdot\ldots\cdot(h_n)^{i_n}\,.$$

В частности, дифференциал p-го порядка функции двух переменных $z = f(x\,,\,y)$ можно вычислять по следующей формуле, для вывода которой используется аналог формулы бинома Ньютона:

$$D^p f(x\,,\,y)(h\,,\,k)^p = \sum_{i=0}^{p}\binom{p}{i}\cdot\frac{\partial^p f(x\,,\,y)}{(\partial x)^i(\partial y)^{p-i}}\cdot h^i k^{p-i}\,.$$

3. Формула Тейлора для функций векторного аргумента

Теорема 42. *Пусть функция* $f\,:\,\mathcal{A}\longrightarrow\mathbb{R}$ *имеет в точке* $\boldsymbol{x}_0$ *области* $\mathcal{A}\subset\mathbb{R}^n$ *дифференциалы до порядка* p *включительно. Если* $[\boldsymbol{x}_0\,,\,\boldsymbol{x}_0+\boldsymbol{h}]\subset\mathcal{A}$, *то справедливо равенство*

$$f(\boldsymbol{x}_0+\boldsymbol{h}) = f(\boldsymbol{x}_0)+Df(\boldsymbol{x}_0)(\boldsymbol{h})+\frac{D^2 f(\boldsymbol{x}_0)(\boldsymbol{h})^2}{2!}+\ldots+\frac{D^p f(\boldsymbol{x}_0)(\boldsymbol{h})^p}{p!}+ \\ + R_p(\boldsymbol{h})\,, \quad (15.51)$$

где $R_p(\boldsymbol{h}) = o(|\boldsymbol{h}|^p)$ *при* $\boldsymbol{h}\to\boldsymbol{0}$.

◄ Положим $\boldsymbol{h} = t\boldsymbol{l}$, где $t = |\boldsymbol{h}| > 0$, $|\boldsymbol{l}| = 1$. Рассмотрим вспомогательную функцию $\varphi\,:\,[0\,,1]\longrightarrow\mathbb{R}$, определенную равенством $\varphi(t) := f(\boldsymbol{x}_0+t\boldsymbol{h})$. Она p-кратно дифференцируема в точке $t=0$ как композиция p-кратно дифференцируемых функций, а ее последовательные производные равны

$$\varphi'(t) = Df(\boldsymbol{x}+t\boldsymbol{l})(\boldsymbol{l})\,,\quad \varphi''(t) = D^2 f(\boldsymbol{x}+t\boldsymbol{l})(\boldsymbol{l})^2\,,\ \ldots\,, \\ \varphi^{(p)}(t) = D^p f(\boldsymbol{x}+t\boldsymbol{l})(\boldsymbol{l})^p\,. \quad (15.52)$$

[6] *Полиномиальной формулой* называется обобщение формулы бинома Ньютона на случай степени суммы, в которой число слагаемых больше двух.

Для функции φ в окрестности точки $t = 0$ имеет место формула Маклорена

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \frac{\varphi''(0)}{2!}t^2 + \ldots + \frac{\varphi^{(p)}(0)}{p!}t^p + r_p(t)\,, \qquad (15.53)$$

где $r_p(t) = o(t^p)$ при $t \to 0$. Подставляя сюда вместо φ и ее производных их выражения через f, получим равенство (15.51), где обозначено $R_p(\boldsymbol{h}) := r_p(t) = r_p(|\boldsymbol{h}|) = o(|\boldsymbol{h}|)$ при $\boldsymbol{h} \to \boldsymbol{0}$. ►

Теорема 43. *Пусть функция* $f : \mathcal{A} \longrightarrow \mathbb{R}$ *имеет в области* $\mathcal{A} \subset \mathbb{R}^n$ *дифференциалы до порядка* $p+1$ *включительно. Если* $[\boldsymbol{x}_0\,,\ \boldsymbol{x}_0 + \boldsymbol{h}] \subset \mathcal{A}$, *то остаточный член формулы Тейлора* (15.51) *можно представить в форме Лагранжа*

$$R_p(\boldsymbol{h}) = \frac{D^{p+1}f(\boldsymbol{\xi})(\boldsymbol{h})^{p+1}}{(p+1)!} \qquad (15.54)$$

при некотором $\boldsymbol{\xi} \in (\boldsymbol{x}_0\,,\ \boldsymbol{x}_0 + \boldsymbol{h})$.

◄ В условиях теоремы введенная выше функция φ имеет конечные производные до порядка $p+1$ включительно, и согласно равенствам (15.52) имеем $\varphi^{(p+1)}(t) = D^{p+1}f(\boldsymbol{x} + t\boldsymbol{l})(\boldsymbol{l})^{p+1}$. Остаточный член формулы Маклорена (15.53) можно представить в форме Лагранжа

$$r_p(t) = \frac{\varphi^{p+1}(\theta)}{(p+1)!}t^{p+1}$$

при некотором $\theta \in (0\,,\,1)$. Полагая $\boldsymbol{\xi} := \boldsymbol{x}_0 + \theta\boldsymbol{h}$ и учитывая, что $\boldsymbol{h} = t\boldsymbol{l}$, имеем

$$R_p(\boldsymbol{h}) = r_p(t) = \frac{D^{p+1}f(\boldsymbol{x}_0 + \theta\boldsymbol{h})(\boldsymbol{l})^{p+1}}{(p+1)!}t^{p+1} = \frac{D^{p+1}f(\boldsymbol{\xi})(\boldsymbol{h})^{p+1}}{(p+1)!}\,. \ ►$$

Глава 16

НЕКОТОРЫЕ ПРИЛОЖЕНИЯ ДИФФЕРЕНЦИАЛЬНОГО ИСЧИСЛЕНИЯ ОТОБРАЖЕНИЙ ИЗ $\mathbb{R}^n$ В $\mathbb{R}^m$

§ 1. Локальные экстремумы функций векторного аргумента

1. Понятие локального экстремума и необходимые условия экстремума

Пусть $f\ :\ \mathcal{A} \longrightarrow \mathbb{R}$, $\mathcal{A} \subset \mathbb{R}^n$, — скалярная функция векторного аргумента. Здесь рассмотрим задачу нахождения внутренних локальных экстремумов функции f методами дифференциального исчисления. В связи с этим всюду будем предполагать, что $\mathcal{A}$ — область, а функция f дифференцируема в этой области. Напомним сначала понятие локального экстремума.

Определение 33. *Точка $\boldsymbol{a} \in \mathcal{A}$ называется точкой локального экстремума функции $f\ :\ \mathcal{A} \longrightarrow \mathbb{R}$, если существует проколотая окрестность $\dot{\mathcal{U}}(\boldsymbol{a})$ точки $\boldsymbol{a}$, такая, что $\forall\, \boldsymbol{x} \in \dot{\mathcal{U}}(\boldsymbol{a})$ выполняется хотя бы одно из следующих неравенств*:

(a) $f(\boldsymbol{x}) < f(\boldsymbol{a})$ (строгий локальный максимум);

(b) $f(\boldsymbol{x}) \leqslant f(\boldsymbol{a})$ (локальный максимум);

(c) $f(\boldsymbol{x}) > f(\boldsymbol{a})$ (строгий локальный минимум);

(d) $f(\boldsymbol{x}) \geqslant f(\boldsymbol{a})$ (локальный минимум).

Перенося в этих неравенствах $f(\boldsymbol{a})$ в левую часть, заключаем, что исследование функции f на экстремум сводится к исследованию знака приращения $\Delta f := f(\boldsymbol{x}) - f(\boldsymbol{a})$ в достаточно малой проколотой окрестности точки $\boldsymbol{a}$.

В предыдущей главе было показано, что необходимые условия локального экстремума дифференцируемой функции f в точке $\boldsymbol{a}$

заключаются в выполнении равенств

$$\frac{\partial f(\boldsymbol{a})}{\partial x_1} = \frac{\partial f(\boldsymbol{a})}{\partial x_2} = \ldots = \frac{\partial f(\boldsymbol{a})}{\partial x_n} = 0\,. \tag{16.1}$$

Точку $\boldsymbol{a} = (a_1\,, \ldots\,, a_n)$, в которой выполняются эти равенства, условимся называть *стационарной*. Так как обычно стационарные точки функции f заранее не известны, то для их нахождения необходимо решить систему n уравнений (16.1) относительно n неизвестных координат стационарной точки. Если стационарная точка $\boldsymbol{a}$ функции f найдена, то возникает вопрос: является ли она точкой экстремума, и если да, то какой именно? Иначе говоря, нужны достаточные условия экстремума.

2. Достаточные условия экстремума функции двух переменных

Здесь рассмотрим вопрос о достаточных условиях экстремума в наиболее простом и наиболее часто встречающемся случае функции двух переменных.

Теорема 44. *Пусть $(a\,, b)$ — стационарная точка дифференцируемой функции $z = f(x\,, y)$. Предположим, что существует второй дифференциал $D^2 f(x\,, y)(h\,, k)^2$, непрерывный в точке $(a\,, b)$. Введем следующие обозначения:*

$$A := \frac{\partial^2 f(a\,, b)}{\partial x^2}\,;\; B := \frac{\partial^2 f(a\,, b)}{\partial x \partial y} = \frac{\partial^2 f(a\,, b)}{\partial y \partial x}\,;\; C := \frac{\partial^2 f(a\,, b)}{\partial y^2}\,. \tag{16.2}$$

Если $A > 0$, $AC - B^2 > 0$, то в точке $(a\,, b)$ функция f имеет строгий локальный минимум.

Если $A < 0$, $AC - B^2 > 0$, то в точке $(a\,, b)$ функция f имеет строгий локальный максимум.

Если $AC - B^2 < 0$, то в точке $(a\,, b)$ функция f не имеет экстремума.

◄ Разложим функцию f в окрестности точки $(a\,, b)$ по формуле

Тейлора

$$f(a+h\,,\,b+k)=$$
$$=f(a\,,\,b)+Df(a\,,\,b)(h\,,\,k)+\frac{1}{2!}D^2f(a\,,\,b)(h\,,\,k)^2+r_2(h\,,\,k)\,. \quad (16.3)$$

Здесь $Df(a\,,\,b)(h\,,\,k)=h\dfrac{\partial f(a\,,\,b)}{\partial x}+k\dfrac{\partial f(a\,,\,b)}{\partial y}\equiv 0$, так как точка $(a\,,\,b)$ — стационарная. Учитывая, далее, обозначения (16.40), имеем

$$D^2f(a\,,\,b)(h\,,\,k)^2=h^2\frac{\partial^2 f(a\,,\,b)}{\partial x^2}+2hk\frac{\partial^2 f(a\,,\,b)}{\partial x\partial y}+k^2\frac{\partial^2 f(a\,,\,b)}{\partial y^2}=$$
$$=Ah^2+2Bhk+Ck^2. \quad (16.4)$$

Учитывая, далее, асимптотику остаточного члена $r_2(h\,,\,k)$, представим его в виде

$$r_2(h\,,\,k)=\rho^2\cdot\alpha(h\,,\,k)\,,\quad \text{где}\quad \rho:=\sqrt{h^2+k^2}\,,\quad \text{а}\quad \lim_{\rho\to+0}\alpha(h\,,\,k)=0\,.$$

Перенося в (16.3) член $f(a\,,\,b)$ в левую часть, представим приращение функции f в точке $(a\,,\,b)$ в следующем виде:

$$\Delta f(k\,,\,k):=f(a+h\,,\,b+k)-f(a\,,\,b)=\frac{Ah^2+2Bhk+Ck^2}{2}+\rho^2\alpha\,. \quad (16.5)$$

Для исследования знака приращения перейдем в окрестности точки $(a\,,\,b)$ к полярным координатам, полагая

$$h=\rho\cos\varphi\,,\quad k=\rho\sin\varphi\,,\quad (0\leqslant\varphi\leqslant 2\pi)\,.$$

Тогда получим

$$\Delta f(h\,,\,k)=\frac{\rho^2}{2}\cdot[\Lambda(\varphi)+2\alpha(h\,,\,k)]\,, \quad (16.6)$$

где обозначено

$$\Lambda(\varphi):=A\cos^2\varphi+2B\cos\varphi\sin\varphi+C\sin^2\varphi\,.$$

Так как предположено, что $A\neq 0$, то выражение $\Lambda(\varphi)$ можно преобразовать к следующему виду

$$\Lambda(\varphi)=\frac{(A\cos\varphi+B\sin\varphi)^2+(AC-B^2)\sin^2\varphi}{A}\,. \quad (16.7)$$

Предположим, что $AC - B^2 > 0$. Тогда из представления (16.7) очевидно, что непрерывная на отрезке $[0\,,\,2\pi]$ функция Λ нигде не обращается в нуль. Поэтому ее модуль ограничен снизу положительным числом, т. е.

$$\exists m \in \mathbb{R}_+ \quad \forall \varphi \in [0\,,\,2\pi] \;:\; |\Lambda(\varphi)| \geqslant m\,.$$

Так как $\alpha \to 0$ при $\rho \to +0$, то

$$\exists \rho_0 \in \mathbb{R}_+ \quad \forall \rho \in (0\,,\,\rho_0) \;:\; |2\alpha(h\,,\,k)| < m\,.$$

В силу этих неравенств при $\rho < \rho_0$ приращение (16.6) имеет тот же знак, что и $\Lambda(\varphi)$, т. е. тот же знак, что и A. Значит, при $A > 0$ оно положительно, и мы имеем минимум, а при $A < 0$ оно отрицательно, и мы имеем максимум.

Пусть теперь $AC - B^2 < 0$. Тогда слагаемые числителя в (16.7) имеют разные знаки, поэтому существуют два значения φ_1 и φ_2 такие, что $\Lambda(\varphi_1) > 0$, а $\Lambda(\varphi_2) < 0$. В качестве таких точек можно, например, взять корни уравнений $\sin\varphi = 0$ и $A\cos\varphi + B\sin\varphi = 0$ соответственно. Затем выбрать $\rho_0 \in \mathbb{R}_+$ так, чтобы было

$$\forall \rho \in (0\,,\,\rho_0) \;:\; |2\alpha(h\,,\,k)| < \min\left\{|\Lambda(\varphi_1)|\,,\,|\Lambda(\varphi_2)|\right\}\,.$$

При таком выборе ρ будут выполняться следующие неравенства:

$$\Delta f(\rho\cos\varphi_1\,,\,\rho\sin\varphi_1) > 0 \quad \text{и} \quad \Delta f(\rho\cos\varphi_2\,,\,\rho\sin\varphi_2) < 0\,.$$

Поскольку эти приращения имеют разные знаки, то функция f не имеет экстремума в точке $(a\,,\,b)$. ▶

Замечание. Теорема была доказана в предположении, что выполняются неравенства $A \neq 0$ и $AC - B^2 \neq 0$. Если хотя бы одно из эитх неравенств не выполняется, то требуется дополнительное исследование, поскольку здесь возможны самые разнообразные ситуации. Возьмем, например, следующие пять функций:

$$\begin{aligned}
&f_1(x\,,\,y) := x^4 + y^4\,;\\
&f_2(x\,,\,y) := -x^4 - y^4\,;\\
&f_3(x\,,\,y) := x^4\,;\\
&f_4(x\,,\,y) : -y^4\,;\\
&f_5(x\,,\,y) := x^4 - y^4\,.
\end{aligned}$$

У всех этих функций точка $(0\,,\,0)$ — стационарная, причем $A = B = C = 0$. Однако в этой точке:

функция f_1 имеет строгий локальный минимум;
функция f_2 имеет строгий локальный максимум;
функция f_3 имеет локальный минимум («жёлоб»);
функция f_4 имеет локальный максимум («хребет»);
функция f_5 не имеет экстремума («седло»).

3. Достаточные условия экстремума функции n переменных

Приведем здесь необходимые сведения из алгебры квадратичных форм, т. е. функций от $\boldsymbol{h} = (h_1\,,\,\ldots,h_n) \in \mathbb{R}^n$ вида

$$\boldsymbol{h} \longmapsto A(\boldsymbol{h})^2 := \sum_{k,j=1}^{n} a_{kj} h_k h_j\,, \quad \text{где} \quad a_{kj} = a_{jk}\,.$$

Определение 34. *квадратичная форма $A(\boldsymbol{h})^2$ называется:*
(a) *положительно определенной, если* $\forall\,\boldsymbol{h} \neq \boldsymbol{0}\ :\ A(\boldsymbol{h})^2 > 0$;
(b) *отрицательно определенной, если* $\forall\,\boldsymbol{h} \neq \boldsymbol{0}\ :\ A(\boldsymbol{h})^2 < 0$;
(c) *неопределенной, если* $\exists\,\boldsymbol{h}_1\ \exists\,\boldsymbol{h}_2\ :\ \begin{cases} A(\boldsymbol{h}_1)^2 > 0\,, \\ A(\boldsymbol{h}_2)^2 < 0\,. \end{cases}$

Для исследования вопроса об опредеоености квадратичной формы вводится в рассмртрение последовательность «главных миноров» ее матрицы $(a_{kj})_{k,j=1}^{n}$:

$$\Delta_1 := a_{11}\,,\ \Delta_2 := \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}\,, \ldots\,, \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}\,.$$

Из алгебры известна следующая теорема Д. Д. Сильвестра[1].

Теорема 45 (критерий Сильвестра). *Положительная определенность квадратичной формы* $\sum\limits_{k,j} a_{kj} h_k h_j$ *равносильна выполне-*

[1] *Сильвестр* Джеймс Джозеф (1814—1897) — английский математик.

нию следующих неравенств:

$$\Delta_1 > 0\,,\ \Delta_2 > 0\,,\ \ldots\,,\ \Delta_n > 0\,.$$

Отртицательная определенность квадратичной формы $\sum\limits_{k,j} a_{kj} h_k h_j$ *равносильна выполнению следующих неравенств:*

$$\Delta_1 < 0\,,\ \Delta_2 > 0\,,\ \ldots\,,\ (-1)^n \Delta_n > 0\,.$$

Критерий Сильвестра можно использовать для проверки выполнения следующего достаточного условия экстремума функции n переменных.

Теорема 46. *Пусть* $\boldsymbol{a}$ *— стационарная точка функции* $f\,:\,\mathcal{A} \longrightarrow \mathbb{R}$, $\mathcal{A} \subset \mathbb{R}^n$, *и пусть существует квадратичная форма* $A(\boldsymbol{h})^2 := D^2 f(\boldsymbol{a})(\boldsymbol{h})^2$. *Тогда*:

(a) *если квадратичная форма* $A(\boldsymbol{h})^2$ *положительно определенная, то в точке* a *функция* f *имеет строгий локальный минимум;*

(b) *если квадратичная форма* $A(\boldsymbol{h})^2$ *отрицательно определенная, то в точке* a *функция* f *имеет строгий локальный максимум;*

(c) *если квадратичная форма* $A(\boldsymbol{h})^2$ *неопределенная, то в точке* a *функция* f *не имеет экстремума.*

◄ Разложим функцию f по формуле Тейлора в окрестности точки $\boldsymbol{a}$:

$$f(\boldsymbol{a}+\boldsymbol{h}) = f(\boldsymbol{a}) + Df(\boldsymbol{a})(\boldsymbol{h}) + \frac{D^2 f(\boldsymbol{a})(\boldsymbol{h})^2}{2!} + r_2(\boldsymbol{h})\,. \tag{16.8}$$

Здесь $Df(\boldsymbol{a})(\boldsymbol{h}) \equiv 0$, так как точка $\boldsymbol{a}$ — стационарная. Далее, так как $r_2(\boldsymbol{h}) = o\big(|\boldsymbol{h}|^2\big)$ при $\boldsymbol{h} \to \boldsymbol{0}$, то $r_2(\boldsymbol{h}) = \rho^2 \cdot \alpha(\boldsymbol{h})$, где $\rho := |\boldsymbol{h}|$ и $\lim\limits_{\rho\to 0} \alpha(\boldsymbol{h}) = 0$. Таким образом, из равенства (16.8) вытекает следующее представление для приращения функции f:

$$\Delta f(\boldsymbol{h}) := f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) = \frac{\rho^2}{2} \cdot \big[D^2 f(\boldsymbol{a})(\boldsymbol{l})^2 + 2\alpha(\rho\boldsymbol{l})\big]\,, \tag{16.9}$$

где обозначено $\boldsymbol{h} = \rho\boldsymbol{l}$, $\rho > 0$, $\boldsymbol{l} = 1$.

Предположим, что квадратичная форма $D^2 f(\boldsymbol{a})(\boldsymbol{h})^2$ является положительно или отрицательно определенной. Тогда функция $\boldsymbol{l} \longmapsto \big|D^2 f(\boldsymbol{a})(\boldsymbol{l})^2\big|$ будет непрерывной и положительной на сфере

$\mathbb{S} := \{|\boldsymbol{l}\,| = 1\}$. Так как сфера $\mathbb{S}$ компактна, то эта функция ограничена снизу положительным числом, т. е.

$$\exists m \in \mathbb{R}_+ \quad \forall \boldsymbol{l} \in \mathbb{S} \ : \ \left|D^2 f(\boldsymbol{a})(\boldsymbol{l}\,)^2\right| \geqslant m\,.$$

Так как функция α бесконечно малая при $\rho \to 0$, то

$$\exists \rho_0 \in \mathbb{R}_+ \quad \forall \rho \in (0\,,\,\rho_0)\ ;\ |2\alpha(\rho \boldsymbol{l}\,)| < m\,.$$

Из двух последних неравенств и из (16.9) следует, что при $|\boldsymbol{h}\,| < \rho_0$ приращение $\Delta f(\boldsymbol{h}\,)$ имеет тот же знак, что и $D^2 f(\boldsymbol{a})(\boldsymbol{l}\,)^2$. Отсюда заключаем, что если квадратичная форма $D^2 f(\boldsymbol{a})(\boldsymbol{h}\,)^2$ положительно (отрицательно) определена, то функция f имеет в точке $\boldsymbol{a}$ минимум (максимум).

Предположим теперь, что форма $D^2 f(\boldsymbol{a})(\boldsymbol{h}\,)^2$ — неопределенная, т. е. существуют точки $\boldsymbol{h}_1 = \rho \boldsymbol{l}_1$ и $\boldsymbol{h}_2 = \rho \boldsymbol{l}_2$, в которых $D^2 f(\boldsymbol{a})(\boldsymbol{h}_1)^2 > 0$, $D^2 f(\boldsymbol{a})(\boldsymbol{h}_2)^2 < 0$. В этом случае имеем

$$\exists \rho_0 \in \mathbb{R}_+ \quad \forall \rho \in (0\,,\,\rho_0)\ :$$
$$|2\alpha(\rho\boldsymbol{l})| < \min\left\{|D^2 f(\boldsymbol{a})(\boldsymbol{l}_1)^2|\,,\,|D^2 f(\boldsymbol{a})(\boldsymbol{l}_2)^2|\right\}\,.$$

При таком выборе ρ будет $\Delta f(\boldsymbol{h}_1) > 0$ и $\Delta f(\boldsymbol{h}_2) < 0$. Значит, экстремума в точке $\boldsymbol{a}$ не будет. ▶

§ 2. Теоремы о неявных функциях

1. Неявные функции скалярного аргумента

С функциями, заданными неявно, мы уже всречались в главе 5 в связи с задачей вычисления производных от функций скалярного аргумента. Здесь рассмотрим их более подробно. Начнем рассмотрение с наиболее простого случая неявных функций скалярного аргумента.

Определение 35. *Пусть* $(x\,,\,y) \longmapsto F(x\,,\,y)$ *— функция двух переменных. Говорят, что функция* $y = y(x)$ *задана на множестве* X *неявно уравнением* $F(x\,,\,y) = 0$*, если* $\forall x \in X$ *выполняется равенство* $F\,(x\,,\,y(x)) \equiv 0$.

Основная проблема теории неявных функций — это *проблема суествования*, которая заключается в следующем. Дано уравнение $F(x\,,\,y)=0$, требуется узнать, существует ли функция $y=y(x)$, заданная неявно этим уравнением. Если существование такой функции установлено, то возникают проблемы ее единственности, непрерывности, дифференцируемости и т. д. В некоторых достаточно простых случаях существование неявной функции можно установить непосредственно, решив уравнение $F(x\,,\,y)=0$ относительно переменной y. Например, уравнение $x^2+y^2=1$ неявно задает функции $y=\pm\sqrt{1-x^2}$. Чаще, однако, встречаются более сложные случаи, когда нахождение переменной y из уравнения $F(x\,,\,y)=0$ затруднено и даже проблематично. Таковы, например, следующие уравнения:

$$x^3+y^3-3axy=0\,;$$
$$y-x-\varepsilon\sin y=0\,;$$
$$1+5\sin(x+y)-e^{\sqrt{xy}}=0\,;$$
$$y-x\operatorname{arctg}\frac{y}{x}=0\,;$$
$$\ln\sqrt{x^2+y^2}-\operatorname{arctg}\frac{y}{x}=0\,.$$

В связи с этим неявные функции обычно изучают непосредственно по уравнению $F(x\,,\,y)=0$, *не решая его.* Ниже рассматриваются проблемы существования, единственности, непрерывности и дифференцируемости функций, заданных неявно.

Теорема 47 (существование и единственность). *Предположим, что:*

(a) *функция* $F:\mathcal{A}\longrightarrow\mathbb{R}$, $\mathcal{A}\subset\mathbb{R}^2$, *непрерывна в прямоугольнике* $\big\{(x\,,\,y)\;\big|\;|x-x_0|\leqslant\Delta\,,\;|y-y_0|\leqslant\Delta'\big\}$ *с центром в точке* $(x_0\,,\,y_0)$*;*

(b) *выполняется равенство* $F(x_0\,,\,y_0)=0$*;*

(c) *при каждом* $x\in[x_0-\Delta\,,\,x_0+\Delta]$ *функция* $F(x\,,\,y)$ *строго монотонна по переменной* y.

Тогда существуют: число $\delta\in(0\,,\,\Delta]$ *и единственная функция* $y=y(x)$, *непрерывная при* $x\in(x-\delta\,,\,x+\delta)$ *и такая, что* $y_0=y(x_0)$ *и* $F(x\,,\,y(x))\equiv 0$.

◀ Предположим для определенности, что функция

$$y\longmapsto F(x_0\,,\,y)$$

строго возрастает. В силу возрастания справедливы неравенства $F(x_0\,,\,y_0-\Delta')<F(x_0\,,\,y_0)<F(x_0\,,\,y_0+\Delta')$, а так как $F(x_0\,,\,y_0)=0$, то $F(x_0\,,\,y_0-\Delta')<0<F(x_0\,,\,y_0+\Delta')$. Отсюда на основании локальных свойств непрерывных функций заключаем, что

$$\exists\delta\in(0\,,\,\Delta]\ \ \forall x\in(x_0-\delta\,,\,x_0+\delta)\ :\quad F(x\,,\,y_0-\Delta')<0<F(x\,,\,y_0+\Delta')\,. \tag{16.10}$$

Фиксируя произвольно $x\in(x_0-\delta\,,\,x_0+\delta)$, рассмотрим функцию $y\longmapsto F(x\,,\,y)$, непрерывную по y на отрезке $[y_0-\Delta'\,,\,y_0+\Delta']$. В силу неравенств (16.10) она принимает на концах этого отрезка значения разных знаков. Отсюда на основании теоремы Больцано — Коши заключаем, что существует точка $y(x)\in(y_0-\Delta'\,,\,y_0+\Delta')$, такая, что $F(x\,,\,y(x))=0$. Так как функция $y\longmapsto F(x\,,\,y)$ — *строго* возрастает, то точка $y(x)$ — единственная для каждого фиксированного значения x. Тем самым определяется функция $x\longmapsto y(x)$, притом единственная, со свойствами: $y(x_0)=y_0$ и $F(x\,,\,y(x))\equiv 0$ (см. рис. 9).

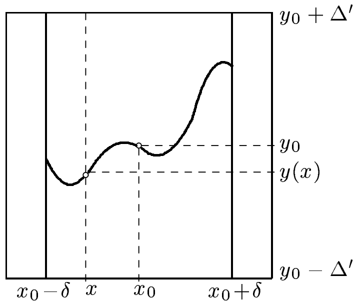

Рис. 9. К теореме 47

Осталось только показать, что функция

$$y=y(x)$$

непрерывна. С этой целью сформулируем полученный выше результат следующим образом:

$$\forall\Delta'\in\mathbb{R}_+\ \ \exists\delta\in\mathbb{R}_+\ :\quad y^{-1}\left((y_0-\Delta'\,,\,y_0+\Delta')\right)\supset(x_0-\delta\,,\,x_0+\delta)\,, \tag{16.11}$$

т. е. *полный прообраз любой окрестности точки y_0 содержит некоторую окрестность точки x_0.* Но это утвержение есть одна из форм определения непрерывности функции $y=y(x)$ в точке x_0, и значит, функция $y=y(x)$ непрерывна в точке x_0. Так как открытый интервал является окрестностью каждой своей точки, то утверждение (16.11) означает непрерывность функции $y=y(x)$ во всех точках $x\in(x_0-\delta\,,\,x_0+\delta)$ (а не только в точке x_0). ►

Замечание. Отметим, что теорема 47 имеет локальный характер. Это означает, что существование неявной функции гарантируется лишь в некоторой окрестности точки x_0.

Теорема 48 (дифференцируемость). *Предположим, что*:
(a) *функция* $F : \mathcal{A} \longrightarrow \mathbb{R}$, $\mathcal{A} \subset \mathbb{R}^2$, *и ее частные производные непрерывны в некоторой окрестности точки* $(x_0\,,\,y_0) \in \mathcal{A}$*;*
(b) *выполнены соотношения* $F(x_0\,,\,y_0) = 0$, $F'_y(x_0\,,\,y_0) \neq 0$;
Тогда существуют: число $\delta \in \mathbb{R}_+$ *и единственная функция* $y = y(x)$, *определенная при* $x \in (x_0 - \delta\,,\,x_0 + \delta)$ *и такая, что* $y_0 = y(x_0)$ *и* $F\,(x\,,\,y(x)) \equiv 0$. *Функция* $y = y(x)$ *дифференцируема, а для ее производной справедливо равенство*

$$y'(x) = -\frac{F'_x(x\,,\,y)}{F'_y(x\,,\,y)}\,. \tag{16.12}$$

◀ Предположим для определенности, что $F'_y(x_0\,,\,y_0) > 0$. Так как F'_y непрерывна,то в силу локальных свойств непрерывных функций существует прямоугольник

$$\Pi := \bigl\{(x\,,\,y)\ \big|\ |x - x_0| \leqslant \Delta\,,\,|y - y_0| \leqslant \Delta'\bigr\} \subset \mathcal{A}\,,$$

во всех точках которого выполняется неравенство $F'_y(x\,,\,y) > 0$. Из этого неравенства следует, что при $(x\,,\,y) \in \Pi$ функция F строго возрастает по переменной y. Таким образом, для функции F выполнены все условия предыдущей теоремы. Применяя ее, заключаем, что существуют: число $\delta \in (0\,,\,\Delta]$ и единственная функция $y = y(x)$, определенная и непрерывная на интервале $(x_0 - \delta\,,\,x_0 + \delta)$, и такая, что $y_0 = y(x_0)$ и $F(x\,,\,y(x)) \equiv 0$.

Желая установить дифференцируемость функции $y = y(x)$, зафиксируем $x \in (x_0 - \delta\,,\,x_0 + \delta)$ и возьмем $h \neq 0$ настолько малым, чтобы было $(x + h) \in (x_0 - \delta\,,\,x_0 + \delta)$. Используя, далее, теорему «о конечных приращениях», получим

$$\begin{aligned} 0 = F(x + h\,,\,y(x + h)) - F(x\,,\,y(x)) &= \\ = h \cdot F'_x(\xi\,,\,\eta) + [y(x + h) - y(x)] \cdot F'_y(\xi\,,\,\eta)\,, \end{aligned} \tag{16.13}$$

где $(\xi\,,\,\eta) \in [(x\,,\,y(x))\,;\,(x + h\,,\,y(x + h))]$. Из последнего соотношения следует, что $(\xi\,,\,\eta) \to (x\,,\,y(x))$ при $h \to 0$. Учитывая это, а также непрерывность частных производных F'_x и F'_y, заключаем, что правая часть равенства

$$\frac{y(x+h)-y(x)}{h} = -\frac{F'_x(\xi\,,\,\eta)}{F'_y(\xi\,,\,\eta)} \tag{16.14}$$

имеет предел, равный правой части равенства (16.12). Переходя в равенстве (16.14) к пределу при $h \to 0$, заключаем, что производная $y'(x)$ существует и что выполняется равенство (16.12). ▶

Замечания. 1. Чтобы найти значение производной $y'(x)$ (как функцию одного только переменного x), необходимо еще исключить переменную y из системы уравнений

$$F(x\,,\,y) = 0\,, \quad y'(x) = -\frac{F'_x(x\,,\,y)}{F'_y(x\,,\,y)}\,.$$

Методы такого исключения в классическом анализе, однако, не изучаются.

2. Отметим здесь другой, равносильный формуле (16.12), способ нахождения производной от функции $y = y(x)$, заданной неявно уравнением $F(x\,,\,y) = 0$.

Продифференцируем по переменной x тождество

$$F(x\,,\,y(x)) \equiv 0\,,$$

рассматривая его левую часть как композицию двух отображений

$$x \longmapsto \begin{pmatrix} x \\ y(x) \end{pmatrix} \longmapsto F(x\,,\,y(x))\,.$$

Тогда получим равенство нулю произведения соответствующих матриц Якоби:

$$\begin{pmatrix} \dfrac{\partial F(x\,,\,y)}{\partial x} & \dfrac{\partial F(x\,,\,y)}{\partial y} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ y'(x) \end{pmatrix} = 0\,.$$

Перемножая матрицы, получим равносильное формуле (16.12) равенство $F'_x + y' \cdot F'_y = 0$, из которого можно найти y'.

3. Можно показать, что при наличии у функции F достаточного количества непрерывных частных производных высших порядков неявная функция тоже будет иметь соответствующие производные высших порядков. Для их вычисления можно, например, диффренцировать тождество $F'_x + y' \cdot F'_y = 0$ по переменной x. Дифференцируя его и опуская ради краткости аргументы, получим равенство

$$F''_{xx} + y'F''_{xy} + y'F''_{yx} + (y')^2F''_{yy} + y''F'_y = 0\,,$$

из которого легко находится y''.

2. Неявные функции векторного аргумента

Излагаемый здесь материал является непосредственным обобщением материала, изложенного в предыдущем пункте.

Определение 36. *Пусть* $F : \mathcal{A} \longrightarrow \mathbb{R}$ *— скалярная функция векторного аргумента* $(\boldsymbol{x}\,, y) \in \mathcal{A} \subset \mathbb{R}^{n+1}$. *Говорят, что скалярная функция* $y = y(\boldsymbol{x}) \equiv y(x_1\,, \dots\,, x_n)$ *задана на множестве* $X \subset \mathbb{R}^n$ *неявно уравнением* $F(\boldsymbol{x}\,, y) = 0$, *если* $\forall \boldsymbol{x} \in X$ *имеем:* $F(\boldsymbol{x}\,, y(\boldsymbol{x})) \equiv 0$.

Теорема 49. *Предположим, что:*

(a) *функция* $(\boldsymbol{x}\,, y) \longmapsto F(\boldsymbol{x}\,, y)$ *непрерывна в* $(n+1)$*-мерном брусе*[2]

$$\left\{(\boldsymbol{x}\,, y) \in \mathbb{R}^{n+1} \;\middle|\; |x_1 - x_1^0| \leqslant \Delta_1\,, \dots\,, |x_n - x_n^0| \leqslant \Delta_n\,, |y - y_0| \leqslant \Delta'\right\}$$

с центром в точке $(\boldsymbol{x}^0\,, y^0) = (x_1^0\,, \dots\,, x_n^0\,, y^0)$;

(b) *выполняется равенство* $F(\boldsymbol{x}^0, y^0) = 0$.

(c) *при каждом* $\boldsymbol{x} = (x_1\,, \dots x_n)$ *из бруса*

$$\Pi_\Delta := \left\{\boldsymbol{x} \in \mathbb{R}^n \;\middle|\; |x_1 - x_1^0| \leqslant \Delta_1\,, \dots\,, |x_n - x_n^0| \leqslant \Delta_n\right\}$$

функция $y \longmapsto F(\boldsymbol{x}\,, y)$ *строго монотонна.*

Тогда существуют: числа $\delta_1 \in (0\,, \Delta_1]\,, \dots\,, \delta_n \in (0\,, \Delta_n]$ *и единственная функция* $y = y(\boldsymbol{x})$, *непрерывная в брусе*

$$\Pi_\delta := \left\{\boldsymbol{x} \in \mathbb{R}^n \;\middle|\; |x_1 - x_1^0| < \delta_1\,, \dots\,, |x_n - x_n^0| < \delta_n\right\} \tag{16.15}$$

и такая, что $y^0 = y(\boldsymbol{x}^0)$ *и* $F(\boldsymbol{x}\,, y(\boldsymbol{x})) \equiv 0$.

◄ Предположим для определенности, что функция

$$y \longmapsto F(\boldsymbol{x}^0, y)$$

строго возрастает. Тогда справедливы следующие неравенства:

$$F(\boldsymbol{x}^0, y^0 - \Delta') < F(\boldsymbol{x}^0, y^0) < F(\boldsymbol{x}^0, y^0 + \Delta')\,.$$

Учитывая, что $F(\boldsymbol{x}^0, y^0) = 0$, имеем такие неравенства:

$$F(\boldsymbol{x}^0, y^0 - \Delta') < 0 < F(\boldsymbol{x}^0, y^0 + \Delta')\,.$$

[2] Здесь и в дальнейшем краткости ради *брусом* будем называть прямоугольный параллелепипед соответствующей размерности.

Отсюда в силу локальных свойств непрерывных функций заключаем, что существует брус $\Pi_\delta \subset \Pi_\Delta$, такой. что

$$\forall \boldsymbol{x} \in \Pi_\delta \,:\, F(\boldsymbol{x}, y^0 - \Delta') < 0 < F(\boldsymbol{x}, y^0 + \Delta')\,.$$

Фиксируя любое значение $\boldsymbol{x} \in \Pi_\delta$, получаем, что непрерывная на отрезке $[y^0 - \Delta', y^0 + \Delta']$ функция $y \longmapsto F(\boldsymbol{x}\,, y)$ принимает на концах этого отрезка значения разных знаков. Отсюда на основании теоремы Больцано — Коши заключаем, что существует точка

$$y(\boldsymbol{x}) \in (y^0 - \Delta', y^0 + \Delta')\,,$$

в которой $F(\boldsymbol{x}\,, y(\boldsymbol{x})) = 0$. Эта точка — единственная, так как функция $y \longmapsto F(\boldsymbol{x}\,, y)$ — строго монотонная. Тем самым существование и единственность неявной функции установлены. Ее непрерывность в Π_δ вытекает из включения $y^{-1}((y^0 - \Delta', y^0 + \Delta')) \supset \Pi_\delta$ (т. е. полный прообраз интервала содержит брус). ►

Теорема 50. *Предположим, что:*

(a) *функция* $(\boldsymbol{x}\,, y) \longmapsto F(\boldsymbol{x}\,, y)$ *имеет непрерывные частные производные в некоторой окрестности точки* $(\boldsymbol{x}^0\,, y^0)$*;*

(b) $F(\boldsymbol{x}^0\,, y^0) = 0,\ F'_y(\boldsymbol{x}^0\,, y^0) \neq 0\,.$

Тогда существуют числа $\delta_1 > 0\,, \ldots, \delta_n > 0$ *и единственная функция* $y = y(\boldsymbol{x})$, *дифференцируемая в брусе* Π_δ *из* (16.15), *и такая, что* $y^0 = y(\boldsymbol{x}^0)$, $F(\boldsymbol{x}\,, y(\boldsymbol{x})) \equiv 0$, *причем*

$$\frac{\partial y(\boldsymbol{x})}{\partial x_k} = -\frac{F'_{x_k}(\boldsymbol{x}\,, y)}{F'_y(\boldsymbol{x}\,, y)} \quad \text{при} \quad k = 1\,, \ldots, n \quad \text{где} \quad \boldsymbol{x} = (x_1\,, \ldots, x_n)\,. \tag{16.16}$$

◄ Предположим для определенности, что $F'_y(\boldsymbol{x}^0\,, y^0) > 0$. Отсюда по локальному свойству непрерывных функций заключаем, что существует брус Π с центром в точке $(\boldsymbol{x}^0\,, y^0)$, во всех точках которого будет выполняться неравенство $F'_y(\boldsymbol{x}\,, y) > 0$. Из этого неравенства следует, что в брусе Π функция $y \longmapsto F(\boldsymbol{x}\,, y)$ строго возрастает. Таким образом, выполнены все условия предыдущей теоремы. Значит, выполнено и ее заключение. Остается только установить дифференциремость функции $y = y(\boldsymbol{x})$ и доказать равенства (16.16).

С этой целью зафиксируем точку $\boldsymbol{x} \in \Pi_\delta$ и зададим $t \in \mathbb{R}$ так, чтобы было $(\boldsymbol{x} + t\boldsymbol{e}_k) \in \Pi_\delta$, где $\boldsymbol{e}_1\,, \ldots, \boldsymbol{e}_n$ — стандартный базис в

$\mathbb{R}^n$. Применяя теорему о конечных приращениях, получим

$$0 = F\,(\boldsymbol{x} + t\boldsymbol{e}_k\,,\, y(\boldsymbol{x} + t\boldsymbol{e}_k)) - F(\boldsymbol{x}\,,\, y(\boldsymbol{x})) = \\ = tF'_{x_k}(\boldsymbol{\xi}\,,\,\eta) + [y(\boldsymbol{x} + t\boldsymbol{e}_k) - y(\boldsymbol{x})]\,F'_y(\boldsymbol{\xi}\,,\,\eta)\,, \quad (16.17)$$

где

$$(\boldsymbol{\xi}\,,\,\eta) \in [(\boldsymbol{x}\,,\,y(\boldsymbol{x})\,;\,(\boldsymbol{x} + t\boldsymbol{e}_k\,;\,y(\boldsymbol{x} + t\boldsymbol{e}_k))]\;.$$

Очевидно, что $(\boldsymbol{\xi}\,,\,\eta) \to (\boldsymbol{x}\,,\,y(\boldsymbol{x}))$ при $t \to 0$. Из равенства (16.17) находим

$$\frac{y(\boldsymbol{x} + t\boldsymbol{e}_k) - y(\boldsymbol{x})}{t} = -\frac{F'_{x_k}(\boldsymbol{\xi}\,,\,\eta)}{F'_y(\boldsymbol{\xi}\,,\,\eta)}\,. \quad (16.18)$$

В силу непрерывности частных производных правая часть этого равенства имеет конечный предел при $t \to 0$, и в пределе из равенства (16.18) получим равенство (16.16). Из равенства (16.16) следует, что частные производные функции $y = y(\boldsymbol{x})$ непрерывны как композиции непрерывных функций. Из непрерывности частных производных функции $y = y(\boldsymbol{x})$ следует ее дифференцируемость. ▶

§ 3. Принцип сжимающих отображений. Теоремы об обратном и неявном отображениях

1. Принцип сжимающих отображений

Здесь рассмотрим так называемый «принцип сжимающих отображений» — теорему С. Банаха[3], имеющую многочисленные приложения к проблемам существования и вычисления решений различных классов уравнений.

Определение 37. *Пусть M — метрическое пространство с функцией расстояния d. Отображение $f\,;\; M \longrightarrow M$ называется сжимающим, если*

$$\exists q \in (0\,,\,1) \;\; \forall x\,,\,y \in M \;:\; d(f(x)\,,\,f(y)) \leqslant q \cdot d(x\,,\,y)\,. \quad (16.19)$$

[3] *Банах* Стефан (1892—1945) — польский математик, один из создателей функционального анализа.

Условимся считать, что $x \to y \iff d(x\,,\,y) \to 0$. При таком понимании сходимости условие (16.19) означает, что сжимающее отображение непрерывно и, более того, равномерно непрерывно на M.

Теорема 51 (принцип сжимающих отображений). *Пусть M — полное метрическое пространство, а $f : M \longrightarrow M$ — сжимающее отображение. Тогда существует единственная неподвижная точка отображения f, т. е. такая точка $\widetilde{x} \in M$, что $\widetilde{x} = f(\widetilde{x})$.*

◀ Станем решать уравнение $x = f(x)$ методом последовательных приближений (итераций). С этой целью возьмем произвольную точку $x_0 \in M$ и образуем последовательность $(x_n)_{n=0}^{\infty}$, где $x_{n+1} = f(x_n)$ для всех $n \in \mathbb{N}$. Используя условие сжимаемости (16.19) и применяя индукцию, находим:

$$\begin{aligned}
&d(x_1\,,\,x_2) = d\,(f(x_0)\,,\,f(x_1)) \leqslant q\cdot d(x_0\,,\,x_1)\,;\\
&d(x_3\,,\,x_2) = d(f(x_2)\,,\,f(x_1)) \leqslant q\cdot d(x_2\,,\,x_1) \leqslant q^2\cdot d(x_1\,,\,x_0)\,;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\,;\\
&d(x_{n+1}\,,\,x_n) \leqslant q^n\cdot d(x_1\,,\,x_0)\,.
\end{aligned}$$

Используя, далее, неравенство треугольника, имеем

$$\begin{aligned}
&d(x_n\,,\,x_{n+p}) \leqslant\\
&\qquad \leqslant d(x_n\,,\,x_{n+1}) + d(x_{n+1}\,,\,x_{n+2}) + \ldots + d(x_{n+p-1}\,,\,x_{n+p}) \leqslant\\
&\qquad\qquad \leqslant \left(q^n + q^{n+1} + \ldots + q^{n+p-1}\right) d(x_1\,,\,x_0) < \frac{q^n}{1-q}d(x_1\,,\,x_0)\,.
\end{aligned}$$

Итак,

$$0 \leqslant d(x_n\,,\,x_{n+p}) < \frac{q^n}{1-q}d(x_1\,,\,x_0) \to 0 \ \text{ при } \ n \to \infty\,, \tag{16.20}$$

поскольку $0 < q < 1$. Таким образом, последовательность $(x_n)_{n=0}^{\infty}$ фундаментальна. Так как пространство M — полное, то существует предел $\lim\limits_{n\to\infty} x_n = \widetilde{x} \in M$. Переходя в равенстве $x_{n+1} = f(x_n)$ к пределу при $n \to \infty$ и используя непрерывность отображения f, получим

$$\widetilde{x} = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n) = f(\widetilde{x})\,,$$

т. е. $\widetilde{x} = f(\widetilde{x})$. Итак, точка $\widetilde{x}$ — неподвижная.

Осталось доказать единственность неподвижной точки. Предположим, что $\widehat{x} \in M$ — еще одна неподвижная точка отображения f. Используя условие сжимаемости, имеем

$$d(\widetilde{x}\,,\,\widehat{x}) = d(f(\widetilde{x})\,,\,f(\widehat{x})) \leqslant q \cdot d(\widetilde{x}\,,\,\widehat{x})\,,$$

т. е. $(1-q)\cdot d(\widetilde{x}\,,\,\widehat{x}) \leqslant 0$, откуда $d(\widetilde{x}\,,\,\widehat{x}) = 0$, и значит, $\widetilde{x} = \widehat{x}$. ▶

Замечания. 1. Отображение $f\ :\ M \longrightarrow M$ иногда интерпретируют как перемещение точки $x \in M$ в положение точки $f(x) \in M$. При такой интерпретации решения уравнения $x = f(x)$ оказываются *неподвижными точками* отображения f. Поэтому теоремы, гарантирующие существование решений уравнения $x = f(x)$, называют иногда *теоремами о неподвижных точках.* Принцип сжимающих отображений — одна из таких теорем.

2. Переходя к пределу в неравенстве (16.20) при $p \to \infty$, получим неравенство

$$d(x_0\,,\,\widetilde{x}) \leqslant \frac{q^n}{1-q} d(x_1\,,\,x_0)\,,$$

которое показывает, что «погрешность» n-го приближения x_n к решению уравнения $x = f(x)$ оценивается сверху n-ым членом геометрической прогрессии со знаменателем q.

2. Теорема об обратном отображении

Определение 38. *Якобианом дифференцируемого отображения* $f\ :\ \mathcal{A} \longrightarrow \mathbb{R}^n$, $\mathcal{A} \subset \mathbb{R}^n$, *называется определитель его матрицы Якоби, т. е. выражение*

$$\det[f'(\boldsymbol{x})] = \frac{D(f^1, \ldots, f^n)}{D(x_1, \ldots, x_n)} := \begin{vmatrix} \dfrac{\partial f^1(\boldsymbol{x})}{\partial x_1} & \ldots & \dfrac{\partial f^1(\boldsymbol{x})}{\partial x_n} \\ \ldots & \ldots & \ldots \\ \dfrac{\partial f^n(\boldsymbol{x})}{\partial x_1} & \ldots & \dfrac{\partial f^n(\boldsymbol{x})}{\partial x_n} \end{vmatrix}. \qquad (16.21)$$

Например, якобиан линейного отображения

$$\begin{cases} y_1 = a_{11}x_1 + \ldots + a_{1n}x_n\,, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\,, \\ y_n = a_{n1}x_1 + \ldots + a_{nn}x_n\,, \end{cases} \qquad (16.22)$$

совпадает с определителем его матрицы, т. е.

$$\frac{D(y_1,\ldots,y_n)}{D(x_1,\ldots,x_n)}=\begin{vmatrix}a_{11} & \ldots & a_{1n}\\ \ldots & \ldots & \ldots\\ a_{n1} & \ldots & a_{nn}\end{vmatrix} \tag{16.23}$$

В левых частях равенств (16.21) и (16.23) написаны различные применяемые обозначения для якобианов. Из алгебры известно, что если определитель (16.23) отличен от нуля, то для линейного отображения (16.22) существует единственное обратное отображение. Далеко идущее обобщение этого факта на произвольные непрерывно дифференцируемые отображения из $\mathbb{R}^n$ в $\mathbb{R}^n$ содержится в следующей теореме.

Теорема 52 (об обратном отображении). *Предположим, что отображение* $f\ :\ \mathcal{A}\longrightarrow\mathbb{R}^n$, $\mathcal{A}\subset\mathbb{R}^n$, *имеет непрерывную производную в некоторой окрестности точки* $\boldsymbol{a}\in\mathcal{A}$, *и пусть* $\det[f'(\boldsymbol{a})]\neq\neq 0$. *Тогда существуют окрестности* $\mathcal{U}$ *и* $\mathcal{V}$ *точек* $\boldsymbol{a}$ *и* $\boldsymbol{b}:=f(\boldsymbol{a})$ *соответственно, такие, что для сужения* $f\ :\ \mathcal{U}\longrightarrow\mathcal{V}$ *существует единственное обратное отображение* $f^{-1}\ :\ \mathcal{V}\longrightarrow\mathcal{U}$, *которое имеет непрерывную производную, равную*

$$\left(f^{-1}\right)'(\boldsymbol{y})=\left(f'(\boldsymbol{x})\right)^{-1},\quad \text{где}\quad \boldsymbol{y}=f(\boldsymbol{x})\,. \tag{16.24}$$

◀ Не ограничивая общности, будем считать, что $f(\mathbf{0})=\mathbf{0}$ и $f'(\mathbf{0})=\mathrm{I}$, где $\mathrm{I}\ :\ \mathbb{R}^n\longrightarrow\mathbb{R}^n$ — тождественный оператор. В противном случае мы бы могли вместо отображения f рассмотреть отображение

$$F\ :\ \boldsymbol{z}\longmapsto\left(f'(\boldsymbol{a})\right)^{-1}\left(f(\boldsymbol{z}+\boldsymbol{a})-f(\boldsymbol{a})\right)\,,$$

которое в окрестности точки $\boldsymbol{z}=\mathbf{0}$ всеми требуемыми свойствами обладает (т. е. $F(\mathbf{0})=\mathbf{0}$, $F'(\mathbf{0})=\mathrm{I}$).

Введем отображение $h\ :\ \boldsymbol{x}\longmapsto\boldsymbol{x}-f(\boldsymbol{x})$. В окрестности точки $\boldsymbol{x}=\mathbf{0}$ оно имеет непрерывную производную, причем $h(\mathbf{0})=\mathbf{0}$ (нуль-вектор) и $h'(\mathbf{0})=O$ (нуль-оператор). Из этих утверждений следует, что $\exists r\in\mathbb{R}_+$ такое, что в шаре $\mathbb{S}:=\{|\boldsymbol{x}|\leqslant r\}$ выполняются следующие неравенства:

$$\det[f'(\boldsymbol{x})]\neq 0\,,\quad \|f'(\boldsymbol{x})\|\leqslant 2\,,\quad \|h'(\boldsymbol{x})\|\leqslant\frac{1}{2}\,. \tag{16.25}$$

Отсюда и из теоремы о конечных приращениях при $\boldsymbol{x} \in \mathbb{S}$ имеем:

$$|h(\boldsymbol{x})| = |h(\boldsymbol{x}) - h(\boldsymbol{0})| \leqslant \sup_{\boldsymbol{\xi}} \|h'(\boldsymbol{\xi})\| \cdot |\boldsymbol{x}| \leqslant \frac{|\boldsymbol{x}|}{2} \leqslant \frac{r}{2}, \qquad (16.26)$$

где $\boldsymbol{\xi} \in [\boldsymbol{0}\,;\,\boldsymbol{x}]$. Из (16.26) следует, что h отображает шар $\mathbb{S}$ в шар $\left\{|\boldsymbol{x}| \leqslant \dfrac{r}{2}\right\}$. Рассмотрим теперь отображение $g\ :\ \boldsymbol{x} \longmapsto h(\boldsymbol{x}) + \boldsymbol{y}$. Если $|\boldsymbol{x}| \leqslant r$, $|\boldsymbol{y}| \leqslant \dfrac{r}{2}$, то в силу (16.26) имеем

$$|g(\boldsymbol{x})| = |h(\boldsymbol{x}) + \boldsymbol{y}| \leqslant |h(\boldsymbol{x})| + |\boldsymbol{y}| \leqslant \frac{r}{2} + \frac{r}{2} = r\,.$$

Таким образом, g отображает замкнутый шар $\mathbb{S}$ в себя. Далее, $\forall \boldsymbol{x}_1\,,\ \boldsymbol{x}_2 \in \mathbb{S}$ имеем

$$|g(\boldsymbol{x}_1) - g(\boldsymbol{x}_1)| = |h(\boldsymbol{x}_1) - h(\boldsymbol{x}_1)| \leqslant \sup_{\boldsymbol{\xi}} \|h'(\boldsymbol{\xi})\| \cdot |\boldsymbol{x}_1 - \boldsymbol{x}_2| \leqslant \frac{1}{2} \cdot |\boldsymbol{x}_1 - \boldsymbol{x}_2|\,,$$

где $\boldsymbol{\xi} \in [\boldsymbol{x}_1\,;\,\boldsymbol{x}_2] \subset \mathbb{S}$. Поэтому отображение $g\ :\ \mathbb{S} \longrightarrow \mathbb{S}$ — сжимающее с коэффициентом сжатия $q = \dfrac{1}{2}$. Так как метрическое пространство $\mathbb{S} \subset \mathbb{R}^n$ — полное, то для g выполняются все условия принципа сжимающих отображений. Применяя его, заключаем, что

$$\forall \boldsymbol{y} \in \mathcal{V}\ \ \exists!\, \boldsymbol{x} \in \mathbb{S}\ :\ \boldsymbol{x} = g(\boldsymbol{x})\,,$$

где обозначено $\mathcal{V} := \left\{|\boldsymbol{y}| \leqslant \dfrac{r}{2}\right\}$. Учитывая, что

$$\boldsymbol{x} = g(\boldsymbol{x}) \iff \boldsymbol{y} = f(\boldsymbol{x})\,,$$

имеем

$$\forall \boldsymbol{y} \in \mathcal{V}\ \ \exists!\, \boldsymbol{x} \in \mathbb{S}\ :\ \boldsymbol{y} = f(\boldsymbol{x})\,.$$

Таким обрзом, установлено существоване единственного отображения $f^{-1} : \mathcal{V} \longrightarrow \mathcal{U}$, обратного к сужению $f : \mathcal{U} \longrightarrow \mathcal{V}$, где обозначено $\mathcal{U} := f^{-1}(\mathcal{V}) \subset \mathbb{S}$. Осталось только исследовать его свойства.

С этой целью возьмем две пары точек $\boldsymbol{x}_1\,,\ \boldsymbol{x}_2 \in \mathcal{U}$ и $\boldsymbol{y}_1\,,\ \boldsymbol{y}_2 \in \mathcal{V}$, для которых

$$\boldsymbol{y}_1 = f(\boldsymbol{x}_1)\,,\ \ \boldsymbol{y}_2 = f(\boldsymbol{x}_2)\,,\ \ \boldsymbol{x}_1 = f^{-1}(\boldsymbol{y}_1)\,,\ \ \boldsymbol{x}_2 = f^{-1}(\boldsymbol{y}_2)\,.$$

Покажем, что они связаны неравенствами

$$\frac{1}{2} \cdot |\boldsymbol{x}_1 - \boldsymbol{x}_2| \leqslant |\boldsymbol{y}_1 - \boldsymbol{y}_2| \leqslant 2 \cdot |\boldsymbol{x}_1 - \boldsymbol{x}_2|. \tag{16.27}$$

Установим сначала правое неравенство

$$|\boldsymbol{y}_1 - \boldsymbol{y}_2| = |f(\boldsymbol{x}_1) - f(\boldsymbol{x}_2)| \leqslant \sup_{\boldsymbol{\xi}} \|f'(\boldsymbol{\xi})\| \cdot |\boldsymbol{x}_1 - \boldsymbol{x}_2| \leqslant 2 \cdot |\boldsymbol{x}_1 - \boldsymbol{x}_2|.$$

Произведем, далее, следующие оценки

$$|\boldsymbol{x}_1 - \boldsymbol{x}_2| = |g(\boldsymbol{x}_1) - g(\boldsymbol{x}_2)| = |h(\boldsymbol{x}_1) + \boldsymbol{y}_1 - h(\boldsymbol{x}_2) - \boldsymbol{y}_2| \leqslant \\ \leqslant |\boldsymbol{y}_1 - \boldsymbol{y}_2| + |h(\boldsymbol{x}_1) - h(\boldsymbol{x}_2)| \leqslant |\boldsymbol{y}_1 - \boldsymbol{y}_2| + \frac{1}{2} \cdot |\boldsymbol{x}_1 - \boldsymbol{x}_2|.$$

Перенося здесь последнее слагаемое в левую часть, получим левое неравенство (16.27). Неравенства (16.27) означают, что

$$\boldsymbol{x}_1 \to \boldsymbol{x}_2 \quad \Longleftrightarrow \quad \boldsymbol{y}_1 \to \boldsymbol{y}_2,$$

т. е. оба отображения f и f^{-1} непрерывны. Для установления дифференцируемости отображения f^{-1} произведем следующие преобразования и оценки:

$$|f^{-1}(\boldsymbol{y}_2) - f^{-1}(\boldsymbol{y}_1) - (f'(\boldsymbol{x}_1))^{-1} (\boldsymbol{y}_2 - \boldsymbol{y}_1)| = \\ = |\boldsymbol{x}_2 - \boldsymbol{x}_1 - (f'(\boldsymbol{x}_1))^{-1} (f(\boldsymbol{x}_2) - f(\boldsymbol{x}_1))| = \\ = \left|(f'(\boldsymbol{x}_1))^{-1} (f(\boldsymbol{x}_2) - f(\boldsymbol{x}_1) - f'(\boldsymbol{x}_1)(\boldsymbol{x}_2 - \boldsymbol{x}_1))\right| \leqslant \\ \leqslant \| (f'(\boldsymbol{x}_1))^{-1} \| \cdot |f(\boldsymbol{x}_2) - f(\boldsymbol{x}_1) - f'(\boldsymbol{x}_1)(\boldsymbol{x}_2 - \boldsymbol{x}_1)| = \\ = o(|\boldsymbol{x}_2 - \boldsymbol{x}_1|) = o(|\boldsymbol{y}_2 - \boldsymbol{y}_1|)$$

при $\boldsymbol{y}_2 \to \boldsymbol{y}_1$ в силу дифференцируемости отображения f и неравенств (16.27). Итак, отображение f^{-1} дифференцируемо, и выполняется равенство (16.24). Из этого равенства следует, что отображение $\left(f^{-1}\right)'$ непрерывно как композиция непрерывных отображений. ▶

3. Общая теорема о неявном отображении

Пусть $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, $\boldsymbol{y} = (y_1, \dots, y_m) \in \mathbb{R}^m$. Здесь будем использовать следующее обозначение:

$$(\boldsymbol{x}, \boldsymbol{y}) := (x_1, \dots, x_n, y_1, \dots, y_m) \in \mathbb{R}^{n+m}.$$

Кроме того, символы $f'_{\boldsymbol{x}}(\boldsymbol{x}, \boldsymbol{y})$ и $f'_{\boldsymbol{y}}(\boldsymbol{x}, \boldsymbol{y})$ будем использовать ниже для обозначения производной по одной векторной переменной при фиксированной другой.

Определение 39. *Пусть* $f : (\boldsymbol{x}, \boldsymbol{y}) \longmapsto f(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^m$ *— отображение из* $\mathbb{R}^{n+m}$ *в* $\mathbb{R}^m$. *Говорят, что отображение*

$$g : \mathcal{U} \longrightarrow \mathbb{R}^m,\ \mathcal{U} \subset \mathbb{R}^n$$

задано неявно уравнением $f(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0}$, *если*

$$\forall \boldsymbol{x} \in \mathcal{U} \ :\ f(\boldsymbol{x}, g(\boldsymbol{x})) \equiv \boldsymbol{0}.$$

Теорема 53. *Предположим, что:*

(a) *отображение* $f : G \longrightarrow \mathbb{R}^m$ *имеет непрерывную производную в области* $G \subset \mathbb{R}^{n+m}$;

(b) $f(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{0}$, $\det\left[f'_{\boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b})\right] \neq 0$,
где $(\boldsymbol{a}, \boldsymbol{b}) \in G$, $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$.

Тогда существуют: окрестность $\mathcal{U} \subset \mathbb{R}^n$ *точки* $\boldsymbol{a}$ *и единственное отображение* $g : \mathcal{U} \longrightarrow \mathbb{R}^m$, *такое, что* $\boldsymbol{b} = g(\boldsymbol{a})$ *и* $f(\boldsymbol{x}, g(\boldsymbol{x})) \equiv \boldsymbol{0}$. *Отображение* g *имеет непрерывную производную, равную*

$$g'(\boldsymbol{x}) = -(f'_{\boldsymbol{y}}(\boldsymbol{x}, \boldsymbol{y}))^{-1} \circ f'_{\boldsymbol{x}}(\boldsymbol{x}, \boldsymbol{y}). \tag{16.28}$$

◄ Построим вспомогательное отображение $F : G \longrightarrow \mathbb{R}^{n+m}$, $G \subset \mathbb{R}^{n+m}$, определив его для $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{y} \in \mathbb{R}^m$ следующим образом:

$$F(\boldsymbol{x}, \boldsymbol{y}) := (\boldsymbol{x}, f(\boldsymbol{x}, \boldsymbol{y})) = (x_1, \dots, x_n, f^1(\boldsymbol{x}, \boldsymbol{y}), \dots f^m(\boldsymbol{x}, \boldsymbol{y})).$$

Найдем его матрицу Якоби. Она имеет размеры $(m+n) \times (m+n)$ и вычисляется по формуле:

$$[F'] = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ \dfrac{\partial f^1}{\partial x_1} & \dfrac{\partial f^1}{\partial x_2} & \dots & \dfrac{\partial f^1}{\partial x_n} & \dfrac{\partial f^1}{\partial y_1} & \dots & \dfrac{\partial f^1}{\partial y_m} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dfrac{\partial f^m}{\partial x_1} & \dfrac{\partial f^m}{\partial x_2} & \dots & \dfrac{\partial f^m}{\partial x_n} & \dfrac{\partial f^m}{\partial y_1} & \dots & \dfrac{\partial f^m}{\partial y_m} \end{pmatrix}. \tag{16.29}$$

Так как все элементы этой матрицы непрерывны в G, то и производная F' непрерывна в G. Исходя из представления матрицы Якоби в виде (16.29), вычислим якобиан отображения F в точке $(\boldsymbol{a}, \boldsymbol{b})$:

$$\det[F'(\boldsymbol{a}, \boldsymbol{b})] = \begin{vmatrix} \dfrac{\partial f^1}{\partial y_1} & \dots & \dfrac{\partial f^1}{\partial y_m} \\ \dots & \dots & \dots \\ \dfrac{\partial f^m}{\partial y_1} & \dots & \dfrac{\partial f^m}{\partial y_m} \end{vmatrix} (\boldsymbol{a}, \boldsymbol{b}) = \det[f'_{\boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b})] \neq 0$$

в силу условия (b). Таким образом, для отобраажения F в окрестности точки $(\boldsymbol{a}, \boldsymbol{b})$ выполнены все условия предыдущей теоремы. Учитывая, что $F(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{a}, \boldsymbol{0})$ и применяя теорему 52, заключаем, что в некоторой окрестности $\widetilde{\mathcal{U}}$ точки $(\boldsymbol{a}, \boldsymbol{0})$ существует единственное непрерывно дифференцируемое отображение $F^{-1} : \widetilde{\mathcal{U}} \longrightarrow \widetilde{\mathcal{V}}$, обратное к сужению $F : \widetilde{\mathcal{V}} \longrightarrow \widetilde{\mathcal{U}} \subset \mathbb{R}^{m+n}$. Так как $F(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, f(\boldsymbol{x}, \boldsymbol{y}))$, то и $F^{-1}(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, \varphi(\boldsymbol{x}, \boldsymbol{y}))$, т. е. оба отображения F и F^{-1} «сохраняют» первые n координат. Далее, пользуясь произволом в выборе одной из окрестностей $\widetilde{\mathcal{U}}$ и $\widetilde{\mathcal{V}}$, выберем окрестность $\widetilde{\mathcal{U}}$ в виде декартова произведения $\widetilde{\mathcal{U}} = \mathcal{U} \times \mathcal{U}'$, где $\mathcal{U} \subset \mathbb{R}^n$ — открытая окрестность точки $\boldsymbol{a} \in \mathbb{R}^n$, а $\mathcal{U}' \subset \mathbb{R}^m$ — открытая окрестность точки $\boldsymbol{0} \in \mathbb{R}^m$. Полагая $g(\boldsymbol{x}) := \varphi(\boldsymbol{x}, \boldsymbol{0})$, получаем искомое неявное отображение $g : \mathcal{U} \longrightarrow \mathbb{R}^m$. Чтобы это показать, заметим, что $g(\boldsymbol{a}) = \boldsymbol{b}$ и что $F \circ F^{-1} = \mathrm{Id}$. Переписывая последнее равенство в виде $(\boldsymbol{x}; f(\boldsymbol{x}, \varphi(\boldsymbol{x}, \boldsymbol{y}))) \equiv (\boldsymbol{x}, \boldsymbol{y})$, получаем тождество $f(\boldsymbol{x}, \varphi(\boldsymbol{x}, \boldsymbol{y})) \equiv \boldsymbol{y}$. Полагая здесь $\boldsymbol{y} \equiv \boldsymbol{0}$, получаем требуемое тождество $f(\boldsymbol{x}, g(\boldsymbol{x})) \equiv \boldsymbol{0}$. Отображение g непрерывно дифференцируемо как композиция непрерывно дифференцируемых отображений

$g = \pi_1 \circ F^{-1} \circ \pi_2$, где π_1 и π_2 — некотрые отображения проектирования. И, наконец, равенство (16.28) можно доказать, дифференцируя тождество $f(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$, где $\boldsymbol{y} = g(\boldsymbol{x})$. Имеем

$$\begin{pmatrix} f'_{\boldsymbol{x}}(\boldsymbol{x},\boldsymbol{y}) & f'_{\boldsymbol{y}}(\boldsymbol{x},\boldsymbol{y}) \end{pmatrix} \circ \begin{pmatrix} \mathrm{I} \\ g'(\boldsymbol{x}) \end{pmatrix} = O\,.$$

Перемножая эти матрицы, находим

$$f'_{\boldsymbol{x}}(\boldsymbol{x}, \boldsymbol{y}) + f'_{\boldsymbol{y}}(\boldsymbol{x}, \boldsymbol{y}) \circ g'(\boldsymbol{x}) = O\,,$$

откуда легко следует равенство (16.28). ▶

§ 4. Условный экстремум

1. Постановка задачи и необходимый признак

В §1 этой главы была рассмотрена задача о нахождении локальных экстремумов функций векторного аргумента (или нескольких вещественных переменных). Если дополнительно предположить, что переменные не являются независимыми, а связаны некоторыми уравнениями, то возникает новый, более сложный, тип задач. Эти задачи называются задачами на *условный* или *относительный* экстремум.

В качестве конкретной задачи такого типа рассмотрим следующую задачу. *Среди всех треугольников заданного периметра* $2p$ *найти тот, площадь которого максимальна.* Желая формализовать постаноку этой задачи, обозначим через x, y, z длины сторон искомого треугольника. Вычисляя его площадь по известной формуле Герона[4], приходим к задаче нахожденя максимума функции

$$S = \sqrt{p(p-x)(p-y)(p-z)} \quad \text{при условии} \quad x + y + z = 2p\,.$$

В общем случае задачу на условный экстремум можно сформулировать следующим образом. *Найти локальные экстремумы функции*

$$u = f(x_1, \ldots, x_n, y_1, \ldots, y_m) \tag{16.30}$$

[4] *Герон* Александрийский (I век) — древнегреческий ученый.

от $(n+m)$ *скалярных переменных* $x_1, \dots, x_n, y_1, \dots, y_m$ *при условии, что эти переменные связаны между собой следующей системой скалярных уравнений:*

$$\begin{cases} \Phi^1(x_1, \dots, x_n, y_1, \dots, y_m) = 0\,, \\ \dots\dots\dots\dots\dots\dots\dots\dots\,, \\ \Phi^m(x_1, \dots, x_n, y_1, \dots, y_m) = 0\,. \end{cases} \tag{16.31}$$

Вводя векторные переменные $\boldsymbol{x} = (x_1\,, \dots, x_n) \in \mathbb{R}^n$, $\boldsymbol{y} = (y_1\,, \dots, y_m) \in \mathbb{R}^m$ и отображение

$$(\boldsymbol{x}, \boldsymbol{y}) \longmapsto \Phi(\boldsymbol{x}, \boldsymbol{y}) = \begin{pmatrix} \Phi^1(\boldsymbol{x}, \boldsymbol{y}) \\ \vdots \\ \Phi^m(\boldsymbol{x}, \boldsymbol{y}) \end{pmatrix}, \tag{16.32}$$

можно переписать задачу (16.30), (16.31) в виде задачи на отыскание локальных экстремумов функции $u = f(\boldsymbol{x}, \boldsymbol{y})$ векторного аргумента $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^n \times \mathbb{R}^m$ при условии $\Phi(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$. Предполагая, что отображение Φ определено в области $G \subset \mathbb{R}^n \times \mathbb{R}^m$, введем в рассмотрение множество

$$\Omega := \big\{ (\boldsymbol{x}, \boldsymbol{y}) \in G \;\big|\; \Phi(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0} \big\} \tag{16.33}$$

и определим понятие точек локального условного экстремума.

Определение 40. *Точка* $(\boldsymbol{a}, \boldsymbol{b}) \in \Omega$ *называется точкой локального условного экстремума функции* $u = f(\boldsymbol{x}, \boldsymbol{y})$ *при условии* $\Phi(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$, *если существует окрестность* $\mathcal{U} \subset \mathbb{R}^n \times \mathbb{R}^m$ *точки* $(\boldsymbol{a}, \boldsymbol{b})$ *такая, что* $\forall (\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{U} \cap \Omega$ *и* $(\boldsymbol{x}, \boldsymbol{y}) \neq (\boldsymbol{a}, \boldsymbol{b})$ *выполняется хотя бы одно из следующих неравенств:*

(a) $f(\boldsymbol{x}, \boldsymbol{y}) > f(\boldsymbol{a}, \boldsymbol{b})$ *(строгий локальный условный минимум);*

(b) $f(\boldsymbol{x}, \boldsymbol{y}) \geqslant f(\boldsymbol{a}, \boldsymbol{b})$ *(локальный условный минимум);*

(c) $f(\boldsymbol{x}, \boldsymbol{y}) < f(\boldsymbol{a}, \boldsymbol{b})$ *(строгий локальный условный максимум);*

(d) $f(\boldsymbol{x}, \boldsymbol{y}) \leqslant f(\boldsymbol{a}, \boldsymbol{b})$ *(локальный условный максимум).*

Исследование функции на условный экстремум будем проводить с использованием метода *множителей Лагранжа* при некоторых предположениях о дифференцируемости заданных отображений. Предварительно определим необходимые для этого понятия.

Определение 41. *Функция Лагранжа* $L = L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\lambda})$, *соответствующая задаче на экстремум функции* $u = f(\boldsymbol{x}, \boldsymbol{y})$ *при условии* $\Phi(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$, *определяется равенством*

$$L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\lambda}) := f(\boldsymbol{x}, \boldsymbol{y}) + \lambda_1 \Phi^1(\boldsymbol{x}, \boldsymbol{y}) + \ldots + \lambda_m \Phi^m(\boldsymbol{x}, \boldsymbol{y}). \tag{16.34}$$

Входящие сюда координаты вектора $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$ *называются множителями Лагранжа.*

Теорема 54 (необходимый признак условного экстремума). *Предположим, что в точке* $(\boldsymbol{a}, \boldsymbol{b}) \in \Omega$ *функция* $u = f(\boldsymbol{x}, \boldsymbol{y})$ *имеет локальный экстремум при условии* $\Phi(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$, *отображения* f *и* Φ *имеют непрерывные производные в некоторой окрестности точки* $(\boldsymbol{a}, \boldsymbol{b})$, *а* $\det\left[\Phi'_{\boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b})\right] \neq 0$. *Тогда существует* $\boldsymbol{\lambda}^0 \in \mathbb{R}^m$, *такое, что* $L'(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\lambda}^0) = \mathbf{0}$.

◀ Условия, наложенные в формулировке теоремы на отображение Φ, позволяют применить к уравнению $\Phi(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$ общую теорему о неявных отображениях. Применяя ее, заключаем, что существует окрестность $\mathcal{U} \subset \mathbb{R}^n$ точки $\boldsymbol{a}$ и единственное отображение $\boldsymbol{y} = y(\boldsymbol{x})$ из $\mathcal{U}$ в $\mathbb{R}^m$, имеющее непрерывную производную и такое, что $\boldsymbol{b} = y(\boldsymbol{a})$ и $\Phi(\boldsymbol{x}, y(\boldsymbol{x})) \equiv \mathbf{0}$. Последнее тождество означает, что $\forall \boldsymbol{x} \in \mathcal{U} : (\boldsymbol{x}, y(\boldsymbol{x})) \in \Omega$, и потому задача на экстремум функции $u = f(\boldsymbol{x}, \boldsymbol{y})$ при условии $\Phi(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$ равносильна при $\boldsymbol{x} \in \mathcal{U}$ задаче на экстремум функции $u = f(\boldsymbol{x}, y(\boldsymbol{x}))$ без дополнительных условий (безусловный). Применяя теорему о необходимом условии экстремума функции $u = f(\boldsymbol{x}, y(\boldsymbol{x}))$ в точке $\boldsymbol{x} = \boldsymbol{a}$, имеем:

$$f'_{\boldsymbol{x}}(\boldsymbol{a}, \boldsymbol{b}) + f'_{\boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b}) \circ y'(\boldsymbol{a}) = O. \tag{16.35}$$

Чтобы найти производную отображения $\boldsymbol{y} = y(\boldsymbol{x})$, заданного неявно уравнением $\Phi(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{0}$, запишем равенство

$$\Phi'_{\boldsymbol{x}}(\boldsymbol{a}, \boldsymbol{b}) + \Phi'_{\boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b}) \circ y'(\boldsymbol{a}) = O. \tag{16.36}$$

Из этого равенства находим

$$y'(\boldsymbol{a}) = -\left(\Phi'_{\boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b})\right)^{-1} \circ \Phi'_{\boldsymbol{x}}(\boldsymbol{a}, \boldsymbol{b}). \tag{16.37}$$

Подставляя это значение для $y'(\boldsymbol{a})$ в (16.35), получим

$$f'_{\boldsymbol{x}}(\boldsymbol{a}, \boldsymbol{b}) - f'_{\boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b}) \circ \left(\Phi'_{\boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b})\right)^{-1} \circ \Phi'_{\boldsymbol{x}}(\boldsymbol{a}, \boldsymbol{b}) = O. \tag{16.38}$$

Переходя здесь к матрицам, найдем

$$[f'_{\boldsymbol{x}}(\boldsymbol{a},\boldsymbol{b})] - \left[f'_{\boldsymbol{y}}(\boldsymbol{a},\boldsymbol{b})\right]\cdot\left[\Phi'_{\boldsymbol{y}}(\boldsymbol{a},\boldsymbol{b})\right]^{-1}\cdot[\Phi'_{\boldsymbol{x}}(\boldsymbol{a},\boldsymbol{b})] = \\ = (0\,,\ldots,\,0)\in\mathbb{R}^n\,. \quad (16.39)$$

Введем обозначение

$$\boldsymbol{\lambda}^0 = (\lambda_1^0,\ldots,\lambda_m^0) := -\left[f'_{\boldsymbol{y}}(\boldsymbol{a}\,,\boldsymbol{b})\right]\cdot\left[\Phi'_{\boldsymbol{y}}(\boldsymbol{a}\,,\boldsymbol{b})\right]^{-1}\in\mathbb{R}^m. \quad (16.40)$$

Тогда равенства (16.39) и (16.40) можно переписать соответственно в виде

$$D_{\boldsymbol{x}}\left(f+\lambda_1^0\Phi^1+\ldots+\lambda_m^0\Phi^m\right)(\boldsymbol{a}\,,\boldsymbol{b}) = \boldsymbol{0}\in\mathbb{R}^n, \quad (16.41)$$

$$D_{\boldsymbol{y}}\left(f+\lambda_1^0\Phi^1+\ldots+\lambda_m^0\Phi^m\right)(\boldsymbol{a}\,,\boldsymbol{b}) = \boldsymbol{0}\in\mathbb{R}^m, \quad (16.42)$$

а равенство $\Phi(\boldsymbol{a}\,,\boldsymbol{b}) = \boldsymbol{0}$ — в виде

$$D_{\boldsymbol{\lambda}}\left(f+\lambda_1^0\Phi^1+\ldots+\lambda_m^0\Phi^m\right)(\boldsymbol{a}\,,\boldsymbol{b}) = \boldsymbol{0}\in\mathbb{R}^m. \quad (16.43)$$

Система трех равенств (16.41) — (16.43) равносильна, очевидно, одному равенству $L'(\boldsymbol{a}\,,\boldsymbol{b}\,,\boldsymbol{\lambda}^0) = \boldsymbol{0}\in\mathbb{R}^{n+2m}$, где L — функция Лагранжа (16.34). ▶

Замечание. Переписывая равенство $L'(\boldsymbol{a}\,,\boldsymbol{b}\,,\boldsymbol{\lambda}^0) = \boldsymbol{0}\in\mathbb{R}^{n+2m}$ в координатах, получим следующую систему скалярных равенств:

$$\begin{cases} \dfrac{\partial}{\partial x_k}L(\boldsymbol{a},\boldsymbol{b},\boldsymbol{\lambda}^0) = 0\,,\ k = 1\,,\,\ldots\,,\,n\,, \\ \dfrac{\partial}{\partial y_k}L(\boldsymbol{a},\boldsymbol{b},\boldsymbol{\lambda}^0) = 0\,,\ k = 1\,,\,\ldots\,,\,m\,, \\ \Phi^k(\boldsymbol{a},\boldsymbol{b}) = 0\,,\ k = 1\,,\,\ldots\,,\,m\,. \end{cases} \quad (16.44)$$

Эти равенства представляют собой систему уравнений для нахождения стационарных точек задачи на условный экстремум.

Примеры. 1) Решим поставленную в п. 1 задачу нахождения треугольника заданного периметра $2p$, имеющего наибольшую площадь. Дело сводится к нахождению точек максимума функции $S^2 = (p-x)(p-y)(p-z)$ при условии $x+y+z = 2p$, причем $0\leqslant x,y,z\leqslant p$. Соответствуюая функция Лагранжа имеет вид

$$L = L(x,y,z,\lambda) = (p-x)(p-y)(p-z)+\lambda\cdot(x+y+z-2p)\,,$$

а система (16.44) — вид

$$\begin{cases} \dfrac{\partial L}{\partial x} = -(p-y)(p-z) + \lambda = 0\,, \\ \dfrac{\partial L}{\partial y} = -(p-x)(p-z) + \lambda = 0\,, \\ \dfrac{\partial L}{\partial z} = -(p-x)(p-y) + \lambda = 0\,, \\ x + y + z - 2p = 0\,. \end{cases}$$

Эта система переписывается в следующем равносильном виде:

$$\begin{cases} \lambda = \dfrac{S^2}{p-x} = \dfrac{S^2}{p-y} = \dfrac{S^2}{p-z}\,, \\ x + y + z = 2p\,. \end{cases}$$

Отсюда видно, что единственным решением системы является точка — $\left(\dfrac{2p}{3}, \dfrac{2p}{3}, \dfrac{2p}{3}\right)$, которая и является стационарной точкой. Поскольку S^2 — неотрицательная функция, определенная на компактном множестве (кубе) $[0\,,\,p]^3$, то в некоторой точке этого множества она принимает максимальное значение. Поскольку на краю куба $S^2 \equiv 0$, то максимум достигается во внутренней точке, а именно, в стационарной точке. Итак, искомый треугольник — правильный.

2) Исследовать на экстремум функцию $u = \dfrac{x_1 + \ldots + x_n}{n}$ при условии $x_1 \cdot \ldots \cdot x_n = c^n$, где c — положительная константа, а $x_1 > 0\,, \ldots\,, x_n > 0$.

◄ Составим функцию Лагранжа

$$L = \frac{x_1 + \ldots + x_n}{n} + \lambda \cdot (c^n - x_1 \cdot \ldots \cdot x_n)$$

и приравняем к нулю ее частные производные

$$\begin{cases} \dfrac{\partial L}{\partial x_1} = \dfrac{1}{n} - \lambda \cdot \dfrac{x_1 \cdot \ldots \cdot x_n}{x_1} = 0\,, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\,, \\ \dfrac{\partial L}{\partial x_n} = \dfrac{1}{n} - \lambda \cdot \dfrac{x_1 \cdot \ldots \cdot x_n}{x_n} = 0\,, \\ \dfrac{\partial L}{\partial \lambda} = c^n - x_1 \cdot \ldots \cdot x_n = 0\,. \end{cases}$$

Эта система легко решается, и ее решением является единственная стационарная точка $(x_1\,, \ldots\,, x_n) = (c\,, \ldots\,, c)$.

Так как $x_1 \cdot \ldots \cdot x_n = c^n$, то при $x_k \to +0$ должно быть $x_j \to +\infty$ при некотором $j \neq k$. Поэтому $\lim\limits_{x_k \to +0} \dfrac{x_1 + \ldots + x_n}{n} = +\infty$. Отсюда следует, что минимальное значение достигается во внутренней, а именно, в стационарной точке $(c, \ldots, c)$ и равно c. ►

Отсюда, между прочим, следует известное неравенство

$$\frac{x_1 + \ldots + x_n}{n} \geqslant \sqrt[n]{x_1 \cdot \ldots \cdot x_n},$$

причем равенство в нем достигается только при $x_1 = \ldots = x_n$.

2. Достаточный признак условного экстремума

Здесь будем пользоваться обозначениями из определений 40 и 41.

Теорема 55. *Пусть* $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\lambda}) \in \mathbb{R}^{n+2m}$ *— стационарная точка задачи на условный экстремум, для которой* $\det\left[\Phi'_{\boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b})\right] \neq 0$, *а функция Лагранжа* L *имеет в стационарной точке непрерывные частные производные 2-го порядка. Составим квадратичную форму от* $\boldsymbol{h} = (h_1, \ldots, h_n)$:

$$\left(\sum_{k=1}^{n} h_k \frac{\partial}{\partial x_k} + \sum_{j=1}^{m} Dy_j(\boldsymbol{a})(\boldsymbol{h}) \cdot \frac{\partial}{\partial y_j}\right)^2 L(\boldsymbol{a}, y(\boldsymbol{a}), \boldsymbol{\lambda}), \tag{16.45}$$

где

$$Dy(\boldsymbol{a})(\boldsymbol{h}) = \begin{pmatrix} Dy_1(\boldsymbol{a})(\boldsymbol{h}) \\ \vdots \\ Dy_m(\boldsymbol{a})(\boldsymbol{h}) \end{pmatrix} = -\left[\Phi'_{\boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b})\right]^{-1} \cdot \left[\Phi'_{\boldsymbol{x}}(\boldsymbol{a}, \boldsymbol{b})\right] \cdot \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}.$$

Если квадратичная форма положительно определенная, отрицательно определенная или неопределенная, то в точке $(\boldsymbol{a}, \boldsymbol{b})$ *будет соответственно: строгий локальный условный минимум, строгий локальный условный максимум, отсутствие экстремума.*

◄ В силу уравнений связи $\Phi(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0}$ функция Лагранжа $L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\lambda}) = f(\boldsymbol{x}, \boldsymbol{y}) + \langle \boldsymbol{\lambda}, \Phi(\boldsymbol{x}, \boldsymbol{y}) \rangle$ совпадает с функцией f на множестве Ω. В частности, обе эти функции на множестве Ω ведут себя одинаково с точки зрения экстремума. Отсюда, фиксируя стационарное значение множителей Лагранжа $\boldsymbol{\lambda}$, заключаем, что экстремальные свойства функции f зависят от свойств квадратичной формы $D^2L(\boldsymbol{a}, y(\boldsymbol{a}), \boldsymbol{\lambda})(\boldsymbol{h})^2$, где $\boldsymbol{y} = y(\boldsymbol{x})$ — отображение, заданное неявно уравнением связи. Имеем

$$DL(\boldsymbol{a}, y(\boldsymbol{a}), \boldsymbol{\lambda})(\boldsymbol{h}) = L'_{\boldsymbol{x}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\lambda})(\boldsymbol{h}) + L'_{\boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\lambda})(y'(\boldsymbol{a})(\boldsymbol{h}))\,.$$

Далее

$$\begin{aligned} D^2L(\boldsymbol{a}, y(\boldsymbol{a}), \boldsymbol{\lambda})(\boldsymbol{h})^2 = L''_{\boldsymbol{xx}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\lambda})(\boldsymbol{h})^2 + L''_{\boldsymbol{xy}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\lambda})(\boldsymbol{h})(y'(\boldsymbol{a})(\boldsymbol{h})) + \\ + L''_{\boldsymbol{yx}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\lambda})(y'(\boldsymbol{a})(\boldsymbol{h}))(\boldsymbol{h}) + L''_{\boldsymbol{yy}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\lambda})(y'(\boldsymbol{a})(\boldsymbol{h}))^2 + \\ + L'_{\boldsymbol{y}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\lambda})(y'(\boldsymbol{a})(\boldsymbol{h})^2\,. \end{aligned} \tag{16.46}$$

Здесь последнее слагаемое равно нулю, так как точка $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\lambda})$ — стационарная, а второе и третье слагаемые равны между собой в силу теоремы о смешанных производных. Учитывая это и переходя к координатам, преобразуем правую часть (16.46) к виду (16.45)

$$\begin{aligned} D^2L(\boldsymbol{a}, y(\boldsymbol{a}), \boldsymbol{\lambda})(\boldsymbol{h})^2 = L''_{\boldsymbol{xx}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\lambda})(\boldsymbol{h})^2 + L''_{\boldsymbol{xy}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\lambda})(\boldsymbol{h})(y'(\boldsymbol{a})(\boldsymbol{h})) + \\ + L''_{\boldsymbol{yx}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\lambda})(y'(\boldsymbol{a})(\boldsymbol{h}))(\boldsymbol{h}) + L''_{\boldsymbol{yy}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\lambda})(y'(\boldsymbol{a})(\boldsymbol{h}))^2 = \\ = \left(\left(\sum_{k=1}^{n} h_k \frac{\partial}{\partial x_k} \right)^2 + \sum_{j=1}^{m} y^j(\boldsymbol{a})(\boldsymbol{h}) \frac{\partial}{\partial y_j} \sum_{k=1}^{n} h_k \frac{\partial}{\partial x_k} + \right. \\ \left. + \sum_{k=1}^{n} h_k \frac{\partial}{\partial x_k} \sum_{j=1}^{m} y^j(\boldsymbol{a})(\boldsymbol{h}) \frac{\partial}{\partial y_j} + \left(\sum_{j=1}^{m} y^j(\boldsymbol{a})(\boldsymbol{h}) \frac{\partial}{\partial y_j} \right)^2 \right) L(\boldsymbol{a}, y(\boldsymbol{a}), \boldsymbol{\lambda}) = \\ = \left(\sum_{k=1}^{n} h_k \frac{\partial}{\partial x_k} + \sum_{j=1}^{m} Dy_j(\boldsymbol{a})(\boldsymbol{h}) \cdot \frac{\partial}{\partial y_j} \right)^2 L(\boldsymbol{a}, y(\boldsymbol{a}), \boldsymbol{\lambda})\,. \blacktriangleright \end{aligned}$$

§ 5. Замена переменных в дифференциальных выражениях

В теории дифференциальных уравнений (обыкновенных и с частными производными) часто возникают задачи вводить новые «неизвестные» функции и переходить к новым независимым переменным. Такого рода задачи называются задачами замены переменных и могут быть решены в рамках дифференциального исчисления функций скалярного и векторного аргументов. Здесь рассмотрим несколько типичных задач такого рода, предполагая, что все рассматриваемые функции имеют непрерывные производные всех требуемых порядков.

1. Замена независимой переменной

Рассмотрим выражение

$$W(x) = f\left(x\,,\, y\,,\, y'\,,\, \ldots\,,\, y^{(n)}\right)\,, \tag{16.47}$$

предполагая, что f — заданная функция от $(n+2)$-х переменных: x (независимая переменная), $y = y(x)\,,\, y'\,,\, \ldots\,,\, y^{(n)}$ — некоторая функция от x и ее производные функции до n-го порядка включительно. Пусть задана функция $x = \varphi(t)$, где t — новая независимая переменная. Требуется преобразовать выражение $W[\varphi(t)]$ к виду, аналогичному (16.47), где в качестве новой функции берется композиция $z = y \circ \varphi$, а переменная x, функция y и ее производные должны быть исключены.

Полагаяя в (16.47) $x = \varphi(t)$, получим

$$W[\varphi(t)] = f\left[\varphi(t)\,,\, (y \circ \varphi)(t)\,,\, (y' \circ \varphi)(t)\,,\, \ldots\,,\, (y^{(n)} \circ \varphi)(t)\right], \tag{16.48}$$

и остается только исключить композиции $y' \circ \varphi\,,\, \ldots\,,\, y^{(n)} \circ \varphi$. С этой целью продифференцируем композицию $z = y \circ \varphi$:

$z'(t) = y'(\varphi(t)) \cdot \varphi'(t) = (y' \circ \varphi)(t) \cdot \varphi'(t)$ и получим $(y' \circ \varphi)(t) = \dfrac{z'(t)}{\varphi'(t)}$.

Дифференцируя последнее равенство, имеем

$$(y' \circ \varphi)'(t) = (y'' \circ \varphi)(t) \cdot \varphi'(t) = \frac{d}{dt}\left(\frac{z'(t)}{\varphi'(t)}\right)\,,$$

откуда находим

$$(y'' \circ \varphi)(t) = \frac{1}{\varphi'(t)} \cdot \frac{d}{dt}\left(\frac{z'(t)}{\varphi'(t)}\right)\,.$$

Аналогично предыдущему получаем

$$(y''' \circ \varphi)(t) = \frac{1}{\varphi'(t)} \cdot \frac{d}{dt}\left(\frac{1}{\varphi'(t)} \cdot \frac{d}{dt}\left(\frac{z'(t)}{\varphi'(t)}\right)\right) = \left(\frac{1}{\varphi'(t)} \cdot \frac{d}{dt}\right)^2 \left(\frac{z'(t)}{\varphi'(t)}\right).$$

Применяя индукцию, имеем

$$(y^{(n)} \circ \varphi)(t) = \left(\frac{1}{\varphi'(t)} \cdot \frac{d}{dt}\right)^{n-1} \left(\frac{z'(t)}{\varphi'(t)}\right).$$

Подставляя найденные выражения для композиций в равенство (16.48), получим окончательно

$$W[\varphi(t)] = \\ = f\left(\varphi(t)\,,\ z(t)\,,\ \frac{z'(t)}{\varphi'(t)}\,,\ \ldots\,,\ \left(\frac{1}{\varphi'(t)}\frac{d}{dt}\right)^{n-1}\left(\frac{z'(t)}{\varphi'(t)}\right)\right). \quad (16.49)$$

Поскольку функция φ задана, то ее производные тоже можно считать известными, и поэтому выражение (16.49) можно привести к виду $W[\varphi(t)] = \widetilde{f}\left(t\,,\, z\,,\, z'\,,\, \ldots\,,\, z^{(n)}\right)$, где $\widetilde{f}$ — некоторая известная функция указанных аргументов.

2. Замена независимой переменной в случае ее неявного задания

Рассмотрим ту же задачу, предполагая, что зависимость $x = \varphi(t)$ между старой независимой переменной x и новой t задана неявно уравнением $\Phi(t\,,\ x) = 0$. Приводя выражение $W[\varphi(t)]$ к виду (16.49), замечаем, что задача будет решена, если удастся исключить переменные $x = \varphi(t)\,,\ \varphi',\ \ldots\,,\ \varphi^{(n)}$. Так как исключить переменную $x = \varphi(t)$ в общем случае мы не можем, то займемся исключением производных. С этой целью станем дифференцировать равенство $\Phi(t\,,\ x) = 0$ по переменной t:

ИМЕННОЙ УКАЗАТЕЛЬ

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

ЛИТЕРАТУРА

1. *Зорич В. А.* Математический анализ. М., 1997—1998. Ч. I—II.
2. *Толстов Г. П.* Элементы математического анализа. М., 1974. Т. I—II.
3. *Рудин У.* Основы математического анализа. М., 1966.
4. *Спивак М.* Математический анализ на многообразиях. Волгоград, 1996.
5. *Фихтенгольц Г. М.* Курс дифференциального и интегрального исчисления. СПб., 1997. Т. I—III.
6. *Фихтенгольц Г. М.* Основы математического анализа. М., 1968. Т. I—II.
7. *Ильин В. А., Садовничий В. А., Сендов Бл. Х.* Математический анализ. М., 1985.
8. *Никольский С. М.* Курс математического анализа. М., 1990—1991. Т. I—II.
9. *Кудрявцев Л. Д.* Курс математического анализа. М., 1988—1989. Т. 1—3.
10. *Демидович Б. П.* Сборник задач по курсу математического анализа. М., 1998.

ОГЛАВЛЕНИЕ