2

Дифференцирование

ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

Напомним, что функция $f\colon \mathbf{R} \to \mathbf{R}$ называется дифференцируемой в точке $a \in \mathbf{R}$, если существует такое число $f'(a)$, что

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = f'(a). \tag{1}$$

Это равенство, очевидно, теряет смысл в общем случае функции $f\colon \mathbf{R}^n \to \mathbf{R}^m$, однако ему можно придать форму, допускающую обобщение. Пусть $\lambda\colon \mathbf{R} \to \mathbf{R}$ — линейное отображение, определяемое формулой $\lambda(h) = f'(a) \cdot h$. Равенство (1) эквивалентно тогда равенству

$$\lim_{h\to 0} \frac{f(a+h)-f(a)-\lambda(h)}{h} = 0. \tag{2}$$

Последнее равенство часто интерпретируют так: $\lambda(h)+f(a)$ есть хорошая аппроксимация функции f вблизи точки a (см. задачу 2.9). Сосредоточим теперь наше внимание на линейном отображении λ и переформулируем определение дифференцируемости следующим образом.

Функция $f\colon \mathbf{R} \to \mathbf{R}$ дифференцируема в точке $a \in \mathbf{R}$, если существует такое линейное отображение $\lambda\colon \mathbf{R} \to \mathbf{R}$, что

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

В таком виде определение допускает простое обобщение на высшие размерности:

Функция $f\colon \mathbf{R}^n \to \mathbf{R}^m$ *дифференцируема* в точке $a \in \mathbf{R}^n$, если существует такое линейное отображение $\lambda\colon \mathbf{R}^n \to \mathbf{R}^m$, что

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

Заметим, что h — точка из $\mathbf{R}^n$, а $f(a+h)-f(a)-\lambda(h)$ — точка из $\mathbf{R}^m$, так что знаки нормы в этом определении существенны. Линейное отображение λ называется *производной* функции f в точке a и обозначается $Df(a)$.

2.1. Теорема. *Если функция $f\colon \mathbf{R}^n \to \mathbf{R}^m$ дифференцируема в точке $a \in \mathbf{R}^n$, то существует единственное линейное отображение $\lambda\colon \mathbf{R}^n \to \mathbf{R}^m$ такое, что*

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

Доказательство. Предположим, что линейное отображение $\mu\colon \mathbf{R}^n \to \mathbf{R}^m$ удовлетворяет условию

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

Полагая $d(h)=f(a+h)-f(a)$, получаем

$$\lim_{h\to 0} \frac{|\lambda(h)-\mu(h)|}{|h|} = \lim_{h\to 0} \frac{|\lambda(h)-d(h)+d(h)-\mu(h)|}{|h|} \leqslant$$

$$\leqslant \lim_{h\to 0} \frac{|\lambda(h)-d(h)|}{|h|} + \lim_{h\to 0} \frac{|d(h)-\mu(h)|}{|h|} = 0.$$

Но $tx \to 0$ при $t \to 0$ для всякого $x \in \mathbf{R}^n$. Поэтому для $x \neq 0$ имеем

$$0 = \lim_{t\to 0} \frac{|\lambda(tx)-\mu(tx)|}{|tx|} = \frac{|\lambda(x)-\mu(x)|}{|x|}.$$

Следовательно, $\lambda(x)=\mu(x)$. ∎

В дальнейшем мы обнаружим простой способ отыскания $Df(a)$. Пока же рассмотрим функцию $f\colon \mathbf{R}^2 \to \mathbf{R}$, определяемую формулой $f(x, y) = \sin x$. Покажем, что $Df(a, b) = \lambda$, где $\lambda(x, y) = (\cos a)\cdot x$. Для этого заметим, что

$$\lim_{(h,k)\to 0} \frac{|f(a+h, b+k)-f(a, b)-\lambda(h, k)|}{|(h, k)|} =$$

$$= \lim_{(h,k)\to 0} \frac{|\sin(a+h)-\sin a-(\cos a)\cdot h|}{|(h, k)|}.$$

Но так как $\sin'(a)=\cos a$, то

$$\lim_{h\to 0} \frac{|\sin(a+h)-\sin a-(\cos a)\cdot h|}{|h|} = 0.$$

А поскольку $|(h, k)| \geqslant |h|$, имеем также

$$\lim_{h \to 0} \frac{|\sin(a+h) - \sin a - (\cos a) \cdot h|}{|(h, k)|} = 0.$$

Часто удобно рассматривать матрицу отображения $Df(a)$: $\mathbf{R}^n \to \mathbf{R}^m$ относительно стандартных базисов в $\mathbf{R}^n$ и $\mathbf{R}^m$. Эта $(m \times n)$-матрица называется *матрицей Якоби* функции f в точке a и обозначается $f'(a)$. Если $f(x, y) = \sin x$, то $f'(a, b) = (\cos a, 0)$. Если f: $\mathbf{R} \to \mathbf{R}$, то $f'(a)$ есть (1×1)-матрица, единственным элементом которой служит число, обозначаемое в элементарном анализе $f'(a)$.

Производную $Df(a)$ можно определить и для функции f, заданной только на некотором открытом множестве, содержащем a. Но рассмотрение лишь функций, определенных на всем $\mathbf{R}^n$, упрощает формулировки теорем, не приводя по существу к потере общности.

Будем говорить, что функция f дифференцируема на множестве A, если f дифференцируема в каждой точке $a \in A$. Функция f: $A \to \mathbf{R}^m$ будет называться *дифференцируемой*, если ее можно продолжить до дифференцируемой функции на некотором открытом множестве, содержащем A.

Задачи

2.1*. Доказать, что если функция f: $\mathbf{R}^n \to \mathbf{R}^m$ дифференцируема в $a \in \mathbf{R}^n$, то она непрерывна в a. (Указание: использовать задачу 1.10.)

2.2. Функция f: $\mathbf{R}^2 \to \mathbf{R}$ *не зависит от второй переменной*, если для каждого $x \in \mathbf{R}$ имеем $f(x, y_1) = f(x, y_2)$ для всех $y_1, y_2 \in \mathbf{R}$. Показать, что f не зависит от второй переменной тогда и только тогда, когда существует такая функция g: $\mathbf{R} \to \mathbf{R}$, что $f(x, y) = g(x)$. Как выражается $f'(a, b)$ через g'?

2.3. Определить независимость функции f: $\mathbf{R}^2 \to \mathbf{R}$ от первой переменной и найти $f'(a, b)$ для таких f. Какие функции не зависят ни от первой, ни от второй переменной?

2.4. Пусть g — непрерывная функция на единичной окружности $\{x \in \mathbf{R}^2: |x| = 1\}$ и $g(0, 1) = g(1, 0) = 0$, $g(-x) = -g(x)$. Определим f: $\mathbf{R}^2 \to \mathbf{R}$ условиями

$$f(x) = \begin{cases} |x| \, g\left(\dfrac{x}{|x|}\right) & \text{при} \quad x \neq 0, \\ 0 & \text{при} \quad x = 0. \end{cases}$$

а) Показать, что функция h: $\mathbf{R} \to \mathbf{R}$, определенная равенством $h(t) = f(tx)$, где $x \in \mathbf{R}^2$, дифференцируема.

б) Показать, что f не дифференцируема в $(0, 0)$, кроме случая $g = 0$. (Указание: сначала показать, что $Df(0, 0) = 0$, рассматривая (h, k) с $k = 0$ и затем с $h = 0$.)

2.5. Пусть f: $\mathbf{R}^2 \to \mathbf{R}$ определена условиями

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

Показать, что f есть функция вида, рассмотренного в задаче 2.4, и потому f не дифференцируема в $(0, 0)$.

2.6. Пусть f: $\mathbf{R}^2 \to \mathbf{R}$ определяется равенством $f(x, y) = \sqrt{|xy|}$. Показать, что f не дифференцируема в $(0, 0)$.

2.7. Показать, что функция f: $\mathbf{R}^n \to \mathbf{R}$, удовлетворяющая условию $|f(x)| \leqslant |x|^2$, дифференцируема в нуле.

2.8. Пусть f: $\mathbf{R} \to \mathbf{R}^2$. Доказать, что f дифференцируема в a тогда и только тогда, когда f^1 и f^2 дифференцируемы в a, и что в этом случае

$$f'(a) = \begin{pmatrix} (f^1)'(a) \\ (f^2)'(a) \end{pmatrix}.$$

2.9. Функции f, g: $\mathbf{R} \to \mathbf{R}$ называются *равными с точностью до n-го порядка* в a, если

$$\lim_{h \to 0} \frac{f(a+h) - g(a+h)}{h^n} = 0.$$

а) Показать, что f дифференцируема в a тогда и только тогда, когда существует такая функция g вида $g(x) = a_0 + a_1(x - a)$, что f и g равны с точностью до первого порядка в a.

б) Пусть существуют $f'(a), \ldots, f^{(n)}(a)$. Показать, что f равна с точностью до n-го порядка в a функции g: $\mathbf{R} \to \mathbf{R}$, определенной равенством

$$g(x) = \sum_{i=1}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i.$$

(Указание: предел

$$\lim_{x \to a} \frac{f(x) - \sum\limits_{i=1}^{n-1} \dfrac{f^{(i)}(a)}{i!}(x-a)^i}{(x-a)^n}$$

можно вычислить с помощью правила Лопиталя.)

ОСНОВНЫЕ ТЕОРЕМЫ

2.2. Теорема. (Правило дифференцирования сложной функции.) *Если* $f\colon \mathbf{R}^n \to \mathbf{R}^m$ *дифференцируема в* a *и* $g\colon \mathbf{R}^m \to \mathbf{R}^p$ *дифференцируема в* $f(a)$, *то композиция* $g \circ f\colon \mathbf{R}^n \to \mathbf{R}^p$ *дифференцируема в* a *и*

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a).$$

Замечание. Это равенство можно записать также в форме

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a).$$

При $m = n = p = 1$ получаем известное правило дифференцирования сложных функций.

Доказательство. Пусть $b = f(a)$, $\lambda = Df(a)$ и $\mu = Dg(f(a))$.

Положим

$$\varphi(x) = f(x) - f(a) - \lambda(x - a), \tag{1}$$

$$\psi(y) = g(y) - g(b) - \mu(y - b), \tag{2}$$

$$\rho(x) = (g \circ f)(x) - (g \circ f)(a) - (\mu \circ \lambda)(x - a). \tag{3}$$

Тогда

$$\lim_{x \to a} \frac{|\varphi(x)|}{|x - a|} = 0, \tag{4}$$

$$\lim_{y \to b} \frac{|\psi(y)|}{|y - b|} = 0, \tag{5}$$

и мы должны показать, что

$$\lim_{x \to a} \frac{|\rho(x)|}{|x - a|} = 0.$$

Но

$$\rho(x) = g(f(x)) - g(b) - \mu(\lambda(x - a)) =$$

$$= g(f(x)) - g(b) - \mu(f(x) - f(a) - \varphi(x)) \text{ согласно (1)}$$

$$= [g(f(x)) - g(b) - \mu(f(x) - f(a))] + \mu(\varphi(x)) =$$

$$= \psi(f(x)) + \mu(\varphi(x)) \text{ согласно (2).}$$

Таким образом, достаточно доказать, что

$$\lim_{x\to a} \frac{|\psi(f(x))|}{|x-a|} = 0 \tag{6}$$

и

$$\lim_{x\to a} \frac{|\mu(\varphi(x))|}{|x-a|} = 0. \tag{7}$$

Но равенство (7) легко следует из (4) и задачи 1.10. Далее, в силу (5) для всякого $\varepsilon > 0$ существует такое $\delta > 0$, что

$$|\psi(f(x))| < \varepsilon |f(x) - b| \quad \text{при} \quad |f(x) - b| < \delta.$$

Последнее же неравенство справедливо при $|x - a| < \delta_1$ с надлежащим $\delta_1 > 0$. Тогда

$$|\psi(f(x))| < \varepsilon |f(x) - b| = \varepsilon |\varphi(x) + \lambda(x - a)| \leqslant$$
$$\leqslant \varepsilon |\varphi(x)| + \varepsilon M |x - a|$$

для некоторого M (см. задачу 1.10). А отсюда легко следует равенство (6). ▮

2.3. Теорема.

(1) *Если* $f\colon \mathbf{R}^n \to \mathbf{R}^m$ — *постоянная функция* (*т. е. существует такое* $y \in \mathbf{R}^m$, *что* $f(x) = y$ *для всех* $x \in \mathbf{R}^n$), *то*

$$Df(a) = 0$$

при всех $a \in \mathbf{R}^n$.

(2) *Если* $f\colon \mathbf{R}^n \to \mathbf{R}^m$ — *линейное отображение, то*

$$Df(a) = f$$

при всех $a \in \mathbf{R}^n$.

(3) *Функция* $f\colon \mathbf{R}^n \to \mathbf{R}^m$ *дифференцируема в* $a \in \mathbf{R}^n$ *тогда и только тогда, когда дифференцируема каждая координатная функция* f^i, *и*

$$Df(a) = (Df^1(a), \ldots, Df^m(a)).$$

Таким образом, $f'(a)$ *есть* $(m \times n)$-*матрица,* i-*й строкой которой служит* $(f^i)'(a)$.

(4) *Если* $s\colon \mathbf{R}^2 \to \mathbf{R}$ *определено равенством* $s(x, y) = x + y$, *то*

$$Ds(a, b) = s.$$

(5) *Если* $p\colon \mathbf{R}^2 \to \mathbf{R}$ *определено равенством* $p(x, y) = xy$, *то*

$$Dp(a, b)(x, y) = bx + ay,$$

так что

$$p'(a, b) = (b, a).$$

Доказательство.

(1) $\lim\limits_{h \to 0} \dfrac{|f(a+h) - f(a) - 0|}{|h|} = \lim\limits_{h \to 0} \dfrac{|y - y - 0|}{|h|} = 0.$

(2) $\lim\limits_{h \to 0} \dfrac{|f(a+h) - f(a) - f(h)|}{|h|} =$

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

(3) Если каждая координатная функция f^i дифференцируема в a и

$$\lambda = (Df^1(a), \ldots, Df^n(a)),$$

то

$$\begin{aligned} f(a+h) - f(a) - \lambda(h) &= \\ = (f^1(a+h) - f^1(a) - Df^1(a)(h), &\ldots \\ \ldots, f^n(a+h) - f^n(a) - Df^n(a)(h)) \end{aligned}$$

и потому

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} \leqslant$$

$$\leqslant \lim_{h \to 0} \sum_{i=0}^{n} \frac{|f^i(a+h) - f^i(a) - Df^i(a)(h)|}{|h|} = 0.$$

Если же, с другой стороны, f дифференцируема в a, то $f^i = \pi^i \circ f$ дифференцируема в a в силу (2) и теоремы 2.2.

(4) Следует из (2).

(5) Пусть $\lambda(x, y) = bx + ay$. Тогда

$$\lim_{(h, k) \to 0} \frac{|p(a+h, b+k) - p(a, b) - \lambda(h, k)|}{|(h, k)|} = \lim_{(h, k) \to 0} \frac{|hk|}{|(h, k)|}.$$

Но

$$|hk| \leqslant \begin{cases} |h|^2, & \text{если} \quad |k| \leqslant |h|, \\ |k|^2, & \text{если} \quad |h| \leqslant |k|. \end{cases}$$

Следовательно, $|hk| \leqslant |h|^2 + |k|^2$. Поэтому

$$\frac{|hk|}{|(h, k)|} \leqslant \frac{h^2+k^2}{\sqrt{h^2+k^2}} = \sqrt{h^2+k^2},$$

так что

$$\lim_{(h, k)\to 0} \frac{|hk|}{|(h, k)|} = 0. \blacksquare$$

2.4. Следствие. *Если $f, g: \mathbf{R}^n \to \mathbf{R}$ дифференцируемы в a, то*

$$D(f+g)(a) = Df(a) + Dg(a),$$
$$D(fg)(a) = g(a)\,Df(a) + f(a)\,Dg(a).$$

Если, кроме того, $g(a) \neq 0$, то

$$D\left(\frac{f}{g}\right)(a) = \frac{g(a)\,Df(a) - f(a)\,Dg(a)}{[g(a)]^2}.$$

Доказательство. Мы докажем лишь первое равенство, предоставив остальные читателю. Поскольку $f+g = s \circ (f, g)$, имеем

$$D(f+g)(a) = Ds(f(a), g(a)) \circ D(f, g)(a) =$$
$$= s \circ (Df(a), Dg(a)) = Df(a) + Dg(a). \blacksquare$$

Теперь нам гарантирована дифференцируемость тех функций $f: \mathbf{R}^n \to \mathbf{R}^m$, координатные функции которых получаются путем сложения, умножения, деления и композиции функций π^i (являющихся линейными отображениями), и тех функций, дифференцируемость которых установлена в элементарном анализе. Однако нахождение $Df(x)$ или $f'(x)$ может оказаться довольно сложным делом. Пусть, например, $f: \mathbf{R}^2 \to \mathbf{R}$ определена равенством $f(x, y) = \sin(xy^2)$. Так как $f = \sin \circ (\pi^1 \cdot [\pi^2]^2)$, то мы имеем

$$f'(a, b) = \sin'(ab^2) \cdot [b^2(\pi^1)'(a, b) + a([\pi^2]^2)'(a, b)] =$$
$$= \sin'(ab^2) \cdot [b^2(\pi^1)'(a, b) + 2ab(\pi^2)'(a, b)] =$$
$$= (\cos(ab^2)) \cdot [b^2(1, 0) + 2ab(0, 1)] =$$
$$= (b^2\cos(ab^2), 2ab\cos(ab^2)).$$

К счастью, мы найдем вскоре значительно более простой способ вычисления f'.

Задачи

2.10. Используя теоремы этого параграфа, найти f' для следующих функций:

а) $f(x, y, z) = x^y$,
б) $f(x, y, z) = (x^y, z)$,
в) $f(x, y) = \sin(x \sin y)$,
г) $f(x, y, z) = \sin(x \sin(y \sin z))$,
д) $f(x, y, z) = x^{y^z}$,
е) $f(x, y, z) = x^{y+z}$,
ж) $f(x, y, z) = (x + y)^z$.
з) $f(x, y) = \sin(xy)$,
и) $f(x, y) = [\sin(xy)]^{\cos 3}$,
к) $f(x, y) = (\sin(xy), \sin(x \sin y), x^y)$.

2.11. Найти f' в приводимых ниже примерах (где $g\colon \mathbf{R} \to \mathbf{R}$ непрерывно):

$$\text{а) } f(x, y) = \int_a^{x+y} g, \qquad \text{б) } f(x, y) = \int_a^{xy} g,$$

$$\text{в) } f(x, y, z) = \int_{x^y}^{\sin(x \sin(y \sin z))} g.$$

2.12. Функция $f\colon \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^p$ называется *билинейной*, если для любых $x, x_1, x_2 \in \mathbf{R}^n$, $y, y_1, y_2 \in \mathbf{R}^m$ и $a \in \mathbf{R}$

$$f(ax, y) = af(x, y) = f(x, ay),$$
$$f(x_1 + x_2, y) = f(x_1, y) + f(x_2, y),$$
$$f(x, y_1 + y_2) = f(x, y_1) + f(x, y_2).$$

а) Доказать, что если f билинейна, то

$$\lim_{(h, k) \to 0} \frac{|f(h, k)|}{|(h, k)|} = 0.$$

б) Доказать, что $Df(a, b)(x, y) = f(a, y) + f(x, b)$.

в) Показать, что формула для $Dp(a, b)$ в теореме 2.3 есть частный случай пункта б).

2.13. Определим [1]) $IP\colon \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ формулой $IP(x, y) = \langle x, y \rangle$.

а) Найти $D(IP)(a, b)$ и $(IP)'(a, b)$.

---

[1]) IP — сокращение от inner product (внутреннее произведение). — *Прим. перев.*

б) Показать, что если $f, g\colon \mathbf{R} \to \mathbf{R}^n$ дифференцируемы и $h\colon \mathbf{R} \to \mathbf{R}$ определено равенством $h(t) = \langle f(t), g(t)\rangle$, то

$$h'(a) = \langle f'(a)^{\mathrm{T}}, g(a)\rangle + \langle f(a), g'(a)^{\mathrm{T}}\rangle.$$

(Заметим, что $f'(a)$ есть $(n \times 1)$-матрица; транспонированная матрица $f'(a)^{\mathrm{T}}$ есть $(1 \times n)$-матрица, которую мы рассматриваем как элемент из $\mathbf{R}^n$.)

в) Показать, что если $f\colon \mathbf{R} \to \mathbf{R}^n$ дифференцируема и $|f(t)| = 1$ для всех t, то $\langle f'(t)^{\mathrm{T}}, f(t)\rangle = 0$.

г) Дать пример дифференцируемой функции $f\colon \mathbf{R} \to \mathbf{R}$, для которой функция $|f|$, определяемая равенством $|f|(t) = |f(t)|$, была бы недифференцируемой.

2.14. Пусть E_i $(i = 1, \ldots, k)$ — евклидовы пространства различных размерностей. Функция $f\colon E_1 \times \ldots \times E_k \to \mathbf{R}^p$ называется *полилинейной*, если при любом выборе точек $x_j \in E_j$, $j \neq i$, функция $g\colon E_i \to \mathbf{R}^p$, определяемая равенством $g(x) = f(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_k)$, линейна.

а) Показать, что если f — полилинейная функция и $i \neq j$, то для $h = (h_1, \ldots, h_k)$, где $h_l \in E_l$,

$$\lim_{h \to 0} \frac{|f(a_1, \ldots, h_i, \ldots, h_j, \ldots, a_k)|}{|h|} = 0.$$

(Указание: функция $g(x, y) = f(a_1, \ldots, x, \ldots, y, \ldots, a_k)$ билинейна.)

б) Доказать, что

$$Df(a_1, \ldots, a_k)(x_1, \ldots, x_k) = \sum_{i=1}^{k} f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_k).$$

2.15. Будем считать $(n \times n)$-матрицу точкой произведения $\mathbf{R}^n \times \ldots \times \mathbf{R}^n$ n экземпляров пространства $\mathbf{R}^n$, рассматривая каждую строку как элемент из $\mathbf{R}^n$.

а) Доказать, что $\det\colon \mathbf{R}^n \times \ldots \times \mathbf{R}^n \to \mathbf{R}$ дифференцируем и

$$D(\det)(a_1, \ldots, a_n)(x_1, \ldots, x_n) = \sum_{i=1}^{n} \det \begin{pmatrix} a_1 \\ \cdot \\ \cdot \\ \cdot \\ x_i \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix}.$$

б) Показать, что если a_{ij}: $\mathbf{R} \to \mathbf{R}$ дифференцируемы и $f(t) = \det(a_{ij}(t))$, то

$$f'(t) = \sum_{j=1}^{n} \det \begin{pmatrix} a_{11}(t) & \dots & a_{1n}(t) \\ \vdots & & \vdots \\ a'_{j1}(t) & \dots & a'_{jn}(t) \\ \vdots & & \vdots \\ a_{n1}(t) & \dots & a_{nn}(t) \end{pmatrix}.$$

в) Пусть $\det(a_{ij}(t)) \neq 0$ для всех t, функции $b_1, \dots, b_n$: $\mathbf{R} \to \mathbf{R}$ дифференцируемы, а функции $s_1, \dots, s_n$: $\mathbf{R} \to \mathbf{R}$ таковы, что $(s_1(t), \dots, s_n(t))$ при каждом t образует решение системы

$$\sum_{j=1}^{n} a_{ij} s_j(t) = b_j(t), \qquad i = 1, \dots, n.$$

Показать, что все s_i дифференцируемы, и найти $s'_i(t)$.

2.16. Предположим, что функция f: $\mathbf{R}^n \to \mathbf{R}^n$ дифференцируема и имеет дифференцируемую обратную f^{-1}: $\mathbf{R}^n \to \mathbf{R}^n$. Показать, что $(f^{-1})'(a) = [f'(f^{-1}(a))]^{-1}$. (Указание: учесть, что $(f \circ f^{-1})(x) = x$.)

ЧАСТНЫЕ ПРОИЗВОДНЫЕ

Приступим к задаче отыскания производных по одной из переменных. Пусть f: $\mathbf{R}^n \to \mathbf{R}$ и $a \in \mathbf{R}^n$. Предел

$$\lim_{h \to 0} \frac{f(a^1, \dots, a^i + h, \dots, a^n) - f(a^1, \dots, a^n)}{h},$$

если он существует, называется *i-й частной производной* функции f в точке a и обозначается $D_i f(a)$. Важно заметить, что $D_i f(a)$ есть обычная производная некоторой функции. А именно, $D_i f(a) = g'(a^i)$, где $g(x) = f(a^1, \dots, x, \dots, a^n)$. Это означает, что $D_i f(a)$ есть угловой коэффициент касательной в точке $(a, f(a))$ к кривой, полученной пересечением графика f с плоскостью $x^j = a^j$, $j \neq i$ (рис. 2.1). Это означает также, что вычисление $D_i f(a)$ есть задача, которую мы уже умеем решать. Если функция $f(x^1, \dots, x^n)$ задана формулой, в которую входят $x^1, \dots, x^n$, то $D_i f(x^1, \dots, x^n)$ находится дифференцированием функции, значение которой в x^i задается этой формулой, если считать все x^j при $j \neq i$ постоянными. Например, если $f(x, y) = \sin(xy^2)$, то $D_1 f(x, y) = y^2 \cos(xy)^2$ и $D_2 f(x, y) = 2xy \cos(xy^2)$. Если же $f(x, y) = x^y$, то $D_1 f(x, y) = yx^{y-1}$, а $D_2 f(x, y) = x^y \ln x$.

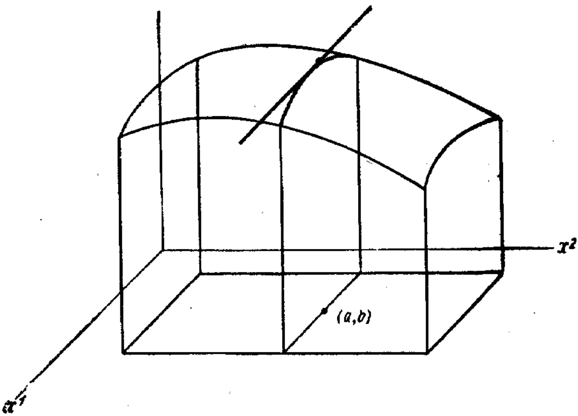

Рис. 2.1.

После небольшой практики (решив, например, задачи в конце этого параграфа) читатель достигнет той же легкости в вычислении $D_i f$, с какой он вычисляет обычные производные.

Если $D_i f(x)$ существует для всех $x \in \mathbf{R}^n$, то мы получаем функцию $D_i f$: $\mathbf{R}^n \to \mathbf{R}$. Ее j-я частная производная в точке x, т. е. $D_j(D_i f)(x)$, часто обозначается $D_{i,j} f(x)$. Заметим, что эта запись обращает порядок i и j. На самом деле порядок обычно неважен, поскольку для большинства функций (по поводу исключений см. задачи) $D_{i,j} f = D_{j,i} f$.

Существуют различные тонкие теоремы относительно условий, обеспечивающих это равенство; следующая теорема вполне достаточна. Мы сформулируем ее здесь, но доказательство отложим (задача 3.28).

2.5. Теорема. *Если $D_{i,j} f$ и $D_{j,i} f$ непрерывны на открытом множестве, содержащем a, то*

$$D_{i,j} f(a) = D_{j,i} f(a).$$

Функция $D_{i,j}f$ называется *(смешанной) частной производной второго порядка* функции f. (Смешанные) частные производные высших порядков определяются аналогично. Очевидно, что при надлежащих условиях теорему 2.5 можно использовать для доказательства равенства (смешанных) частных производных высших порядков. Если f имеет частные производные всех порядков, то порядок индексов

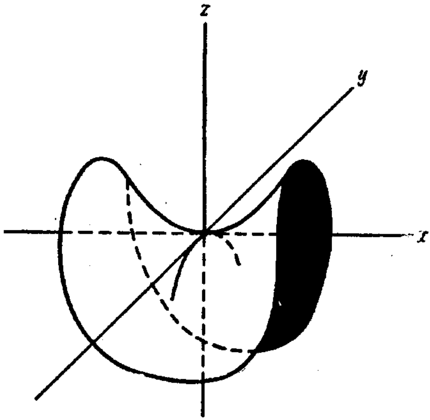

Рис. 2.2.

$i_1, \ldots, i_k$ в $D_{i_1, \ldots, i_k}f$ совершенно неважен. Такие функции называются функциями класса C^∞. В последующих главах часто будет удобно ограничиваться рассмотрением функций класса C^∞.

В следующем параграфе частные производные будут использованы для отыскания производных. Они имеют также другое важное применение — для отыскания максимумов функций.

2.6. Теорема. *Пусть $A \subset \mathbf{R}^n$. Если максимум (или минимум) $f\colon A \to \mathbf{R}$ достигается во внутренней точке a множества A и $D_i f(a)$ существует, то $D_i f(a) = 0$.*

Доказательство. Пусть $g_i(x) = f(x^1, \ldots, x, \ldots$ $\ldots, x^n)$. Очевидно g_i имеет максимум (или минимум) в a^i

и определена в открытом интервале, содержащем a^i. Следовательно, $D_i f(a) = g_i'(a^i) = 0$. ▮

Читатель помнит, что обращение теоремы 2.6 неверно даже при $n=1$ (если $f\colon \mathbf{R}\to\mathbf{R}$ определена равенством $f(x)=x^3$, то $f'(0)=0$, но 0 не является даже локальным максимумом или минимумом). При $n>1$ невозможность обращения теоремы 2.6 может проявляться довольно эффектным образом. Предположим, например, что $f\colon \mathbf{R}^2\to\mathbf{R}$ определена равенством $f(x, y)=x^2-y^2$ (рис. 2.2). Тогда $D_1 f(0, 0)=0$, ибо g_1 имеет минимум в 0, а $D_2 f(0, 0)=0$, ибо g_2 имеет максимум в 0. Очевидно $(0, 0)$ не является точкой ни относительного максимума, ни относительного минимума.

Если для отыскания максимума или минимума функции f на A используется теорема 2.6, то значения f в граничных точках должны быть рассмотрены особо — это нелегкое дело, поскольку границей A может быть все A. Один из способов справиться с этим указан в задаче 2.27; более же сильный метод, который часто используется, изложен в задаче 5.16.

Задачи

2.17. Найти частные производные следующих функций:

а) $f(x, y, z)=x^y$,
б) $f(x, y, z)=z$,
в) $f(x, y)=\sin(x\sin y)$,
г) $f(x, y, z)=\sin(x\sin(y\sin z))$,
д) $f(x, y, z)=x^{y^z}$,
е) $f(x, y, z)=x^{y+z}$,
ж) $f(x, y, z)=(x+y)^z$,
з) $f(x, y)=\sin(xy)$,
и) $f(x, y)=[\sin(xy)]^{\cos 3}$.

2.18. Найти частные производные следующих функций (где $g\colon \mathbf{R}\to\mathbf{R}$ непрерывна):

а) $f(x, y)=\int\limits_a^{x+y} g$,

б) $f(x, y)=\int\limits_y^{x} g$,

в) $f(x, y) = \int\limits_a^{xy} g$.

г) $f(x, y) = \int\limits_a^{\left(\int\limits_b^y g\right)} g$.

2.19. Найти $D_2 f(1, y)$ в случае, когда $f(x, y) = x^{x^{x^{x^y}}} + (\ln x)(\operatorname{arc\,tg}(\operatorname{arc\,tg}(\operatorname{arc\,tg}(\sin(\cos xy) - \ln(x+y)))))$. (Указание: это можно сделать очень просто).

2.20. Выразить частные производные f через производные функции g и h, если

а) $f(x, y) = g(x) h(y)$,
б) $f(x, y) = g(x)^{h(y)}$,
в) $f(x, y) = g(x)$,
г) $f(x, y) = g(y)$,
д) $f(x, y) = g(x+y)$.

2.21*. Пусть g_1, g_2: $\mathbf{R}^2 \to \mathbf{R}$ непрерывны и f: $\mathbf{R}^2 \to \mathbf{R}$ определена равенством

$$f(x, y) = \int\limits_0^x g_1(t, 0)\, dt + \int\limits_0^y g_2(x, t)\, dt.$$

а) Показать, что $D_2 f(x, y) = g_2(x, y)$.

б) Как следовало бы определить f, чтобы $D_2 f(x, y) = g_1(x, y)$?

в) Найти такую функцию f: $\mathbf{R}^2 \to \mathbf{R}$, у которой $D_1 f(x, y) = x$ и $D_2 f(x, y) = y$, и такую функцию, у которой $D_1 f(x, y) = y$ и $D_2 f(x, y) = x$.

2.22*. Показать, что функция f: $\mathbf{R}^2 \to \mathbf{R}$, у которой $D_2 f = 0$, не зависит от второй переменной. Если же $D_1 f = D_2 f = 0$, то f — постоянная.

2.23*. Пусть $A = \{(x, y): x < 0 \text{ или } x \geqslant 0 \text{ и } y \neq 0\}$.

а) Показать, что функция f: $A \to \mathbf{R}$, у которой $D_1 f = D_2 f = 0$, постоянна. (Указание: любые две точки в A можно соединить ломаной, каждое звено которой параллельно одной из осей.)

б) Найти такую функцию f: $A \to \mathbf{R}$, что $D_2 f = 0$, но f не независима от второй переменной.

2.24. Определим f: $\mathbf{R}^2 \to \mathbf{R}$ равенством

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

а) Показать, что $D_2 f(x, 0) = x$ для всех x и $D_1 f(0, y) = -y$ для всех y.

б) Показать, что $D_{1,2} f(0, 0) \neq D_{2,1} f(0, 0)$.

2.25*. Определим $f\colon \mathbf{R}^n \to \mathbf{R}$ равенством

$$f(x) = \begin{cases} e^{-x^{-2}} & \text{при} \quad x \neq 0, \\ 0 & \text{при} \quad x = 0. \end{cases}$$

Показать, что f есть функция класса C^∞ и $f^{(i)}(0) = 0$ для всех i. $\left(\text{Указание: предел } f'(0) = \lim\limits_{h \to 0} \dfrac{e^{-h^{-2}}}{h} = \lim\limits_{h \to 0} \dfrac{\frac{1}{h}}{e^{h^{-2}}}\right.$ можно вычислить по правилу Лопиталя; довольно легко вычисляется $f'(x)$ при $x \neq 0$, а тогда $f''(0) = \lim\limits_{h \to 0} f'(h)$ можно найти по правилу Лопиталя.)

2.26*. Пусть

$$f(x) = \begin{cases} e^{-(x-1)^{-2}} \cdot e^{-(x+1)^{-2}} & \text{при} \quad x \in (-1, 1), \\ 0 & \text{при} \quad x \notin (-1, 1). \end{cases}$$

а) Показать, что $f\colon \mathbf{R} \to \mathbf{R}$ есть функция класса C^∞, положительная на интервале $(-1, 1)$ и равная нулю во всех остальных точках.

б) Показать, что существует такая функция $g\colon \mathbf{R} \to [0, 1]$ класса C^∞, что $g(x) = 0$ при $x \leqslant 0$ и $g(x) = 1$ при $x \leqslant \varepsilon$. $\left(\text{Указание: положить } g(x) = \int\limits_0^x f \Big/ \int\limits_0^\varepsilon f\right.$, где f — функция класса C^∞, положительная на интервале $(0, \varepsilon)$ и равная нулю во всех остальных точках.)

в) Определим при заданном $a \in \mathbf{R}^n$ функцию $g\colon \mathbf{R}^n \to \mathbf{R}$ равенством

$$g(x) = f([x^1 - a^1]/\varepsilon) \cdot \ldots \cdot f([x^n - a^n]/\varepsilon).$$

Показать, что g — функция класса C^∞, положительная на $(a^1 - \varepsilon, a^1 + \varepsilon) \times \ldots \times (a^n - \varepsilon, a^n + \varepsilon)$ и равная нулю во всех остальных точках.

г) Показать, что для всякого открытого множества A и компакта $C \subset A$ существует такая неотрицательная функция $f\colon A \to \mathbf{R}$ класса C^∞, что $f(x) > 0$ для всех $x \in C$ и $f(x) = 0$ вне некоторого замкнутого множества, содержащегося в A.

д) Показать, что указанную в г) функцию $f\colon A \to [0, 1]$ можно выбрать так, чтобы $f(x) = 1$ для всех $x \in C$. (Указание: если функция f из г) такова, что $f(x) \geqslant \varepsilon$ при $x \in C$, то рассмотреть $g \circ f$, где g — функция из б).)

2.27. Определим g, h: $\{x \in \mathbf{R}^2: |x| \leqslant 1\} \to \mathbf{R}^3$ равенствами

$$g(x, y) = (x, y, \sqrt{1 - x^2 - y^2}),$$
$$h(x, y) = (x, y, -\sqrt{1 - x^2 - y^2}).$$

Показать, что максимум f на $\{x \in \mathbf{R}^3: |x| = 1\}$ есть либо максимум $f \circ g$, либо максимум $f \circ h$ на $\{x \in \mathbf{R}^2: |x| \leqslant 1\}$.

ПРОИЗВОДНЫЕ

Читатель, сравнивший задачи 2.10 и 2.17, вероятно уже догадался, что верно следующее утверждение.

2.7. Теорема. *Если f: $\mathbf{R}^n \to \mathbf{R}^m$ дифференцируема в a, то $D_j f^i(a)$ существует для всех i, j: $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, и $f'(a)$ есть $(m \times n)$-матрица $(D_j f^i(a))$.*

Доказательство. Предположим сначала, что $m = 1$, так что f: $\mathbf{R}^n \to \mathbf{R}$. Определим h: $\mathbf{R} \to \mathbf{R}^n$ равенством $h(x) = (a^1, \ldots, x, \ldots, a^n)$, где x стоит на j-м месте. Тогда $D_j f(a) = (f \circ h)'(a^j)$.

Но по теореме 2.2

$$(f \circ h)'(a^j) = f'(a) \cdot h'(a^j) = f'(a) \cdot \begin{array}{c} \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \\ \\ \leftarrow j\text{-е место.} \\ \\ \\ \\ \\ \end{array}$$

Так как $(f \circ h)'(a^j)$ имеет $D_j f(a)$ своим единственным элементом, то это показывает, что $D_j f(a)$ существует и является j-м элементом $(1 \times n)$-матрицы $f'(a)$.

Для произвольного m теорема следует теперь из теоремы 2.3, согласно которой каждая координатная функция f^i дифференцируема и i-й строкой матрицы $f'(a)$ служит $(f^i)'(a)$. ∎

В задачах приведено несколько примеров, показывающих, что обращение теоремы 2.7 неверно. Однако оно верно при одном дополнительном предположении.

2.8. Теорема. *Пусть функция f: $\mathbf{R}^n \to \mathbf{R}^m$ такова, что все частные производные $D_j f^i(x)$ существуют в не-*

котором открытом множестве, содержащем точку a, и непрерывны в a. Тогда f дифференцируема в a.

Доказательство. Как и при доказательстве теоремы 2.7, достаточно рассмотреть случай $m=1$, т. е. $f\colon \mathbf{R}^n \to \mathbf{R}$. Тогда

$$f(a+h)-f(a)=f(a^1+h^1, a^2, \ldots, a^n)-f(a^1, \ldots, a^n)+$$
$$+f(a^1+h^1, a^2+h^2, a^3, \ldots, a^n)-f(a^1+h^1, a^2, \ldots, a^n)+\ldots$$
$$\ldots+f(a^1+h^1, \ldots, a^n+h^n)-$$
$$-f\left(a^1+h^1, \ldots, a^{n-1}+h^{n-1}, a^n\right).$$

Напомним, что $D_1 f$ есть производная функции g, определяемой равенством $g(x)=f(x, a^2, \ldots, a^n)$. Применяя к g теорему о среднем, получаем

$$f(a^1+h^1, a^2, \ldots, a^n)-f(a^1, \ldots, a^n)=$$
$$=h^1 \cdot D_1 f(b_1, a^2, \ldots, a^n),$$

где b_1 — некоторое число, заключенное между a^1 и a^1+h^1.

Аналогично i-й член суммы равен

$$h^i D_i f(a^1+h^1, \ldots, a^{i-1}+h^{i-1}, b_i, \ldots, a^n)=h^i D_i f(c_i)$$

с некоторым c_i. Тогда

$$\lim_{h\to 0}\frac{\left|f(a+h)-f(a)-\sum_{i=1}^{n} D_i f(a)\cdot h^i\right|}{|h|}=$$
$$=\lim_{h\to 0}\frac{\left|\sum_{i=1}^{n}[D_i f(c_i)-D_i f(a)]\cdot h^i\right|}{|h|}\leqslant$$
$$\leqslant \lim_{h\to 0}\sum_{i=1}^{n}|D_i f(c_i)-D_i f(a)|=0,$$

поскольку функции $D_i f$ непрерывны в a. ▌

Хотя в доказательстве теоремы 2.7 было использовано правило дифференцирования сложной функции, легко можно было бы обойтись и без него. Поэтому после теоремы 2.8, дающей признак дифференцируемости функций, и теоремы 2.7, дающей выражения для их производных, это правило могло бы показаться даже излишним. Однако оно имеет чрезвычайно важное следствие, касающееся частных производных.

2.9. Теорема. *Пусть* $g_1, \ldots, g_m$: $\mathbf{R}^n \to \mathbf{R}$ *непрерывно дифференцируемы в* a *и* f: $\mathbf{R}^m \to \mathbf{R}$ *непрерывно дифференцируема в* $(g_1(a), \ldots, g_m(a))$. *Определим* F: $\mathbf{R}^n \to \mathbf{R}$ *равенством* $F(x) = f(g_1(x), \ldots, g_m(x))$. *Тогда*

$$D_iF(a) = \sum_{j=1}^{m} D_jf(g_1(a), \ldots, g_m(a))\, D_ig_j(a).$$

Доказательство. Функция F есть просто композиция $f \circ g$, где $g = (g_1, \ldots, g_m)$. Так как функции g_i непрерывно дифференцируемы в a, то из теоремы 2.8 следует, что g дифференцируема в a. Аналогично f дифференцируема в $(g_1(a), \ldots, g_m(a))$. Следовательно, в силу теоремы 2.2

$$F'(a) = f'(g(a)) \cdot g'(a) =$$

$$= (D_1f(g(a)), \ldots, D_mf(g(a))) \begin{pmatrix} D_1g_1(a) & \ldots & D_ng_1(a) \\ \vdots & & \vdots \\ D_1g_m(a) & \ldots & D_ng_m(a) \end{pmatrix}.$$

Но $D_iF(a)$ есть i-й член левой части этого равенства, в то время как $\sum_{j=1}^{m} D_jf(g_1(a), \ldots, g_m(a))\, D_ig_j(a)$ есть i-й член правой части. ∎

Теорему 2.9 также часто называют *правилом дифференцирования сложной функции*, но она слабее, чем теорема 2.2, поскольку g или f могут быть дифференцируемы и без того, чтобы g_i или f были непрерывно дифференцируемыми (см. задачу 2.32). Вычисления, опирающиеся на теорему 2.9, большей частью довольно просты. Некоторого ухищрения требует функция F: $\mathbf{R}^2 \to \mathbf{R}$, определенная равенством

$$F(x, y) = f(g(x, y), h(x), k(y)),$$

где h, k: $\mathbf{R} \to \mathbf{R}$. Чтобы применить теорему 2.9, определим $\bar{h}, \bar{k}$: $\mathbf{R}^2 \to \mathbf{R}$ равенствами

$$\bar{h}(x, y) = h(x), \quad \bar{k}(x, y) = k(y).$$

Тогда

$$D_1\bar{h}(x, y) = h'(x), \quad D_2\bar{h}(x, y) = 0,$$
$$D_1\bar{k}(x, y) = 0, \quad D_2\bar{k}(x, y) = k'(y)$$

и

$$F(x, y) = f(g(x, y), \bar{h}(x, y), \bar{k}(x, y)).$$

Положив $a = (g(x, y), h(x), k(y))$, получаем

$$D_1F(x, y) = D_2f(a) \cdot D_1g(x, y) + D_2f(a) \cdot h'(x),$$
$$D_2F(x, y) = D_1f(a) \cdot D_2g(x, y) + D_3f(a) \cdot k'(y).$$

Разумеется, нет необходимости действительно выписывать функции $\bar{h}$ и $\bar{k}$.

Задачи

2.28. Найти выражения для частных производных следующих функций:

а) $F(x, y) \quad = f(g(x)k(y), g(x) + k(y))$,
б) $F(x, y, z) = f(g(x+y), h(x+z))$,
в) $F(x, y, z) = f(x^y, y^z, z^x)$,
г) $F(x, y) \quad = f(x, g(x), h(x, y))$.

2.29. Пусть $f\colon \mathbf{R}^n \to \mathbf{R}$. Предел

$$\lim_{t \to 0} \frac{f(a+tx) - f(a)}{t},$$

если он существует, обозначается $D_xf(a)$ и называется *производной* функции f в точке a *по направлению* x.

а) Показать, что $D_{e_i}f(a) = D_if(a)$.

б) Показать, что $D_{tx}f(a) = tD_xf(a)$.

в) Показать, что если f дифференцируема в a, то $D_xf(a) = Df(a)(x)$ и потому $D_{x+y}f(a) = D_xf(a) + D_yf(a)$.

2.30. Пусть f определена, как в задаче 2.4. Показать, что $D_xf(0, 0)$ существует для всех x, однако если $g \neq 0$, то $D_{x+y}f(0, 0) \neq D_xf(0, 0) + D_yf(0, 0)$ для некоторых x и y.

2.31. Пусть $f\colon \mathbf{R}^2 \to \mathbf{R}$ определена, как в задаче 1.26. Показать, что $D_xf(0, 0)$ существует для всех x, хотя f даже не непрерывна в $(0, 0)$.

2.32. а) Пусть $f\colon \mathbf{R} \to \mathbf{R}$ определена условиями

$$f(x) = \begin{cases} x^2 \sin\dfrac{1}{x} & \text{при} \quad x \neq 0, \\ 0 & \text{при} \quad x = 0. \end{cases}$$

Показать, что f дифференцируема в 0, но f' разрывна в 0.

б) Пусть $f\colon \mathbf{R}^2 \to \mathbf{R}$ определена условиями

$$f(x, y) = \begin{cases} (x^2+y^2)\sin\dfrac{1}{\sqrt{x^2+y^2}} & \text{при} \quad (x, y) \neq 0, \\ 0 & \text{при} \quad (x, y) = 0. \end{cases}$$

Показать, что f дифференцируема в $(0, 0)$, но $D_i f$ разрывны в $(0, 0)$.

2.33. Показать, что непрерывность $D_i f^j$ в a можно исключить из условий теоремы 2.8.

2.34. Функция $f\colon \mathbf{R}^n \to \mathbf{R}$ называется *однородной* степени m, если $f(tx) = t^m f(x)$ для всех x. Показать, что если при этом f дифференцируема, то

$$\sum_{i=1}^{n} x^i D_i f(x) = m f(x).$$

(Указание: найти $g'(1)$, где $g(t) = f(tx)$.)

2.35. Доказать, что если $f\colon \mathbf{R}^n \to \mathbf{R}$ дифференцируема и $f(0) = 0$, то существуют такие $g_i\colon \mathbf{R}^n \to \mathbf{R}$, что

$$f(x) = \sum_{i=1}^{n} x^i g_i(x).$$

$\left(\text{Указание: если } h_x(t) = f(tx), \text{ то } f(x) = \int\limits_0^1 h_x'(t)\,dt.\right)$

ОБРАТНЫЕ ФУНКЦИИ

Предположим, что функция $f\colon \mathbf{R} \to \mathbf{R}$ непрерывно дифференцируема на открытом множестве, содержащем a, и $f'(a) \neq 0$. Если $f'(a) > 0$, то существует такой открытый интервал V, содержащий a, что $f'(x) > 0$ для всех $x \in V$, и аналогичное верно, когда $f'(a) < 0$. Таким образом, f возрастает (или убывает) на V, а потому взаимно однозначна и имеет обратную функцию f^{-1}, определенную на некотором открытом интервале W, содержащем $f(a)$. Кроме того, нетрудно показать, что f^{-1} дифференцируема и

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$$

для всех $y \in W$.

Аналогичные рассмотрения для высших размерностей значительно сложнее, но результат (теорема 2.11) весьма важен. Мы начнем с простой леммы.

2.10. Лемма. *Пусть* $A \subset \mathbf{R}^n$ — *параллелепипед и* $f: A \to \mathbf{R}^n$ *непрерывно дифференцируема. Если существует такое число* M, *что* $|D_j f^i(x)| \leqslant M$ *для всех* x *внутри* A, *то*

$$|f(x) - f(y)| \leqslant n^2 M |x - y|$$

для всех $x, y \in A$.

Доказательство. Имеем

$$f^i(y) - f^i(x) = \sum_{j=1}^{n} [f^i(y^1, \ldots, y^j, x^{j+1}, \ldots, x^n) - \\ - f^i(y^1, \ldots, y^{j-1}, x^j, \ldots, x^n)].$$

Применяя теорему о среднем, получаем

$$f^i(y^1, \ldots, y^j, x^{j+1}, \ldots, x^n) - f^i(y^1, \ldots, y^{j-1}, x^j, \ldots, x^n) = \\ = |y^j - x^j| \cdot D_j f^i(z_{ij})$$

с некоторым z_{ij}. Выражение в правой части не превосходит $M|y^j - x^j|$. Следовательно,

$$|f^i(y) - f^i(x)| \leqslant \sum_{j=1}^{n} |y^j - x^j| \cdot M \leqslant nM|y - x|,$$

поскольку $|y^j - x^j| \leqslant |y - x|$ для каждого j. Наконец,

$$|f(y) - f(x)| \leqslant \sum_{i=1}^{n} |f^i(y) - f^i(x)| \leqslant n^2 M \cdot |y - x|. \blacksquare$$

2.11. Теорема об обратной функции. *Предположим, что* $f: \mathbf{R}^n \to \mathbf{R}^n$ *непрерывно дифференцируема в некотором открытом множестве, содержащем* a, *и* $\det f'(a) \neq 0$. *Тогда существуют открытое множество* V, *содержащее* a, *и открытое множество* W, *содержащее* $f(a)$, *такие, что отображение* $f: V \to W$ *имеет непрерывное обратное отображение* $f^{-1}: W \to V$, *дифференцируемое и для всех* $y \in W$ *удовлетворяющее соотношению*

$$(f^{-1})'(y) = [f'(f^{-1}(y))]^{-1}.$$

Доказательство. Пусть λ — линейное отображение $Df(a)$. Оно невырожденно, поскольку $\det f'(a) \neq 0$.

Но $D(\lambda^{-1}\circ f)(a)=D(\lambda^{-1})(f(a))\circ Df(a)=\lambda^{-1}\circ Df(a)$ есть тождественное линейное отображение. Если теорема верна для $\lambda^{-1}\circ f$, то она очевидно верна и для f. Поэтому мы можем считать с самого начала, что λ — тождественное отображение. Если тогда $f(a+h)=f(a)$, то

$$\frac{|f(a+h)-f(a)-\lambda(h)|}{|h|}=\frac{|h|}{|h|}=1.$$

Но

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

Это означает, что равенство $f(x)=f(a)$ не может выполняться для значений x, произвольно близких к a, не равных a. Поэтому существует замкнутый параллелепипед U, содержащий a в качестве внутренней точки и такой, что

$$f(x)\neq f(a),\quad \text{если}\quad x\in U \quad \text{и}\quad x\neq a \tag{1}$$

Поскольку f непрерывно дифференцируема в открытом множестве, содержащем a, можно также считать, что

$$\det f'(x)\neq 0 \quad \text{для всех}\quad x\in U \tag{2}$$

и

$$|D_jf^i(x)-D_jf^i(a)|<\frac{1}{2n^2}\quad \text{для всех}\quad i,\ j\ \text{и}\ x\in U. \tag{3}$$

Заметим, что из (3) и леммы 2.10, примененной к $g(x)=f(x)-x$, вытекает, что

$$|f(x_1)-x_1-(f(x_2)-x_2)|\leqslant\frac{1}{2}|x_1-x_2|$$

для любых $x_1,\ x_2\in U$. Так как

$$|x_1-x_2|-|f(x_1)-f(x_2)|\leqslant$$
$$\leqslant|f(x_1)-x_1-(f(x_2)-x_2)|\leqslant\frac{1}{2}|x_1-x_2|,$$

то получаем

$$|x_1-x_2|\leqslant 2|f(x_1)-f(x_2)|\quad \text{для всех}\quad x_1,\ x_2\in U. \tag{4}$$

Далее, f отображает границу параллелепипеда U в компактное множество, не содержащее, согласно (1), $f(a)$ (рис. 2.3). Поэтому существует такое число $d>0$, что

$|f(a)-f(x)|\geqslant d$ для всех x, принадлежащих границе U. Пусть $W=\{y:|y-f(a)|<d/2\}$. Если $y\in W$ и x принадлежит границе U, то

$$|y-f(a)|<|y-f(x)|. \tag{5}$$

Покажем, что для всякого $y\in W$ существует единственное x внутри U, для которого $f(x)=y$. Для этого рассмотрим функцию $g\colon U\to\mathbf{R}$, определенную равенством

$$g(x)=|y-f(x)|^2=\sum_{i=1}^{n}(y^i-f^i(x))^2.$$

Эта функция непрерывна и потому имеет минимум на U. Если x принадлежит границе U, то $g(a)<g(x)$ в силу (5).

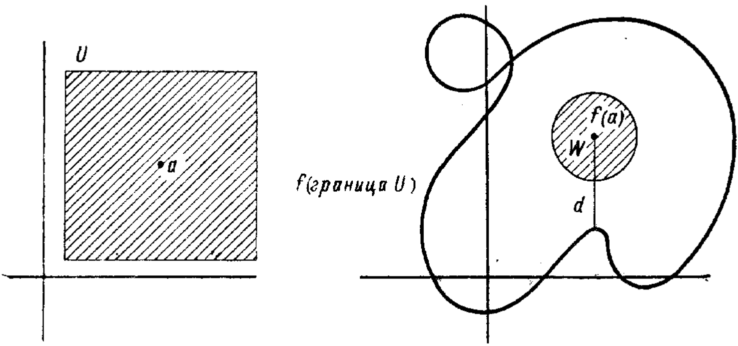

Рис. 2.3.

Следовательно, минимум g *не* достигается на границе U. Согласно теореме 2.6, тогда существует такая точка x внутри U, что $D_jg(x)=0$ для всех j, т. е.

$$\sum_{i=1}^{n}2(y^i-f^i(x))D_jf^i(x)=0 \quad \text{для всех } j.$$

Но в силу (2) матрица $(D_jf^i(x))$ имеет ненулевой определитель. Поэтому мы должны иметь $y^i-f^i(x)=0$ для всех i, т. е. $y=f(x)$. Тем самым доказано существование x. Единственность непосредственно следует из (4).

Обозначим через V пересечение внутренности U с $f^{-1}(W)$. Мы показали, что функция $f\colon V\to W$ имеет обратную f^{-1}: $W \to V$. Теперь (4) можно переписать в виде

$$|f^{-1}(y_1) - f^{-1}(y_2)| \leqslant 2|y_1 - y_2| \text{ для всех } y_1, y_2 \in W. \quad (6)$$

Это показывает, что f^{-1} непрерывна.

Осталось только доказать, что f^{-1} дифференцируема. Пусть $\mu = Df(x)$. Покажем, что f^{-1} дифференцируема в точке $y = f(x)$ и имеет в качестве производной μ^{-1}. Как и в доказательстве теоремы 2.2, для всех $x_1 \in V$ имеем

$$f(x_1) = f(x) + \mu(x_1 - x) + \varphi(x_1 - x),$$

где

$$\lim_{x_1 \to x} \frac{|\varphi(x_1 - x)|}{|x_1 - x|} = 0.$$

Поэтому

$$\mu^{-1}(f(x_1) - f(x)) = x_1 - x + \mu^{-1}(\varphi(x_1 - x)).$$

Так как каждое $y_1 \in W$ имеет вид $f(x_1)$, где $x_1 \in V$, то последнее равенство можно переписать так:

$$f^{-1}(y_1) = f^{-1}(y) + \mu^{-1}(y_1 - y) - \mu^{-1}\left(\varphi\left(f^{-1}(y_1) - f^{-1}(y)\right)\right),$$

и потому достаточно показать, что

$$\lim_{y_1 \to y} \frac{|\mu^{-1}(\varphi(f^{-1}(y_1) - f^{-1}(y)))|}{|y_1 - y|} = 0.$$

Следовательно (задача 1.10), достаточно убедиться в том, что

$$\lim_{y_1 \to y} \frac{|\varphi(f^{-1}(y_1) - f^{-1}(y))|}{|y_1 - y|} = 0.$$

Но

$$\frac{\varphi(f^{-1}(y_1) - f^{-1}(y))|}{|y_1 - y|} =$$

$$= \frac{|\varphi(f^{-1}(y_1) - f^{-1}(y))|}{|f^{-1}(y_1) - f^{-1}(y)|} \cdot \frac{|f^{-1}(y_1) - f^{-1}(y)|}{|y_1 - y|}.$$

Поскольку f^{-1} непрерывна, $f^{-1}(y_1) \to f^{-1}(y)$ при $y_1 \to y$. Поэтому первый множитель стремится к нулю. А так как в силу (6) второй множитель не превосходит 2, то произведение также стремится к 0. ▮

Следует заметить, что обратная функция f^{-1} может существовать даже если $\det f'(a)=0$. Например, если $f\colon \mathbf{R}\to\mathbf{R}$ определяется равенством $f(x)=x^3$, то $f'(0)=0$, но f имеет обратную функцию $f^{-1}(x)=\sqrt[3]{x}$. Все же одно можно сказать определенно: если $\det f'(a)=0$, то f^{-1} не может быть дифференцируема в $f(a)$. Чтобы доказать это, заметим, что $(f\circ f^{-1})(x)=x$. Если бы f^{-1} была дифференцируема в $f(a)$, то правило дифференцирования сложных функций дало бы $f'(a)\cdot(f^{-1})'(f(a))=I$ и, следовательно, $\det f'(a)\cdot\det(f^{-1})'(f(a))=1$, в противоречии с тем, что $\det f'(a)=0$.

Задачи

2.36 *. Пусть $A\subset\mathbf{R}^n$ — открытое множество и $f\colon A\to\mathbf{R}^n$ — такая взаимно однозначная функция, что $\det f'(x)\neq 0$ для всех x. Показать, что $f(A)$ — открытое множество и $f^{-1}\colon f(A)\to A$ дифференцируема. Показать также, что $f(B)$ открыто для всякого открытого множества $B\subset A$.

2.37. а) Пусть $f\colon \mathbf{R}^2\to\mathbf{R}$ — непрерывно дифференцируемая функция. Показать, что f *не* взаимно однозначна. (Указание: если, например, $D_1 f(x, y)\neq 0$ для всех (x, y) из некоторого открытого множества A, рассмотреть функцию $g\colon A\to\mathbf{R}^2$, определяемую равенством $g(x, y)=(f(x, y), y)$.)

б) Обобщить этот результат на случай непрерывно дифференцируемых $f\colon \mathbf{R}^n\to\mathbf{R}^m$ с $m<n$.

2.38. а) Пусть $f\colon \mathbf{R}\to\mathbf{R}$ такова, что $f'(a)\neq 0$ для всех $a\in\mathbf{R}$. Показать, что f взаимно однозначна (на всем $\mathbf{R}$).

б) Определим $f\colon \mathbf{R}^2\to\mathbf{R}^2$, положив $f(x, y)=(e^x\cos y, e^x\sin y)$. Показать, что f не взаимно однозначна, хотя $\det f'(x, y)\neq 0$ для всех (x, y).

2.39. Используя функцию $f\colon \mathbf{R}\to\mathbf{R}$, определенную условиями

$$f(x)=\begin{cases}\dfrac{x}{2}+x^2\sin\dfrac{1}{x} & \text{при } x\neq 0,\\ 0 & \text{при } x=0,\end{cases}$$

показать, что непрерывность производной нельзя исключить из предположений теоремы 2.11.

НЕЯВНЫЕ ФУНКЦИИ

Рассмотрим функцию f: $\mathbf{R}^2 \to \mathbf{R}$, определенную равенством $f(x, y) = x^2 + y^2 - 1$. Если точка (a, b) выбрана так, что $f(a, b) = 0$, причем $a \neq 1$, -1, то (рис. 2.4)

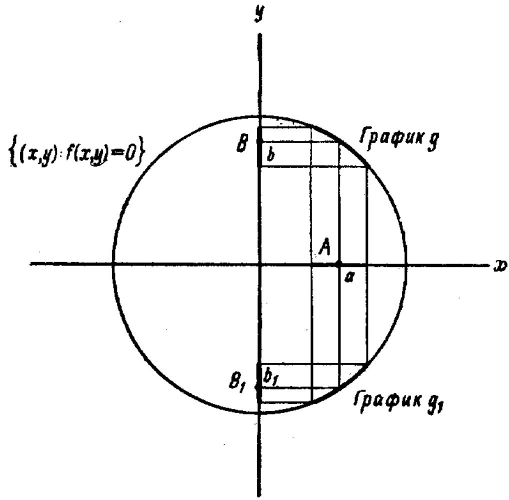

Рис. 2.4.

существуют такие открытые интервалы A и B, содержащие соответственно a и b, что для всякого $x \in A$ существует единственное $y \in B$, для которого $f(x, y) = 0$.

Поэтому можно определить функцию g: $A \to \mathbf{R}$ условиями $g(x) \in B$ и $f(x, g(x)) = 0$ (если $b > 0$, как изображено на рис. 2.4, то $g(x) = \sqrt{1 - x^2}$). Для рассматриваемой нами функции f имеется еще одно число b_1, при котором $f(a, b_1) = 0$. И здесь существует такой интервал B_1, содержащий b_1, что если $x \in A$, то $f(x, g_1(x)) = 0$ для однозначно определенного $g_1(x) \in B_1$ (здесь $g_1(x) = -\sqrt{1 - x^2}$). Обе функции g и g_1 дифференцируемы. Их называют *неявными* функциями, определенными уравнением $f(x, y) = 0$.

При $a = 1$ или -1 невозможно найти ни одной такой функции g, определенной в открытом интервале, содержащем a. Было бы желательно иметь простой критерий, позволяющий решать, когда вообще можно найти такую функцию. Более общим образом можно поставить вопрос так: пусть $f\colon \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$, причем $f(a^1, \ldots, a^n, b) = 0$; когда для каждого $(x^1, \ldots, x^n)$ вблизи $(a^1, \ldots, a^n)$ можно найти единственное y вблизи b, для которого бы $f(x^1, \ldots, x^n, y) = 0$? Еще более общим образом можно задаться вопросом об условиях разрешимости системы уравнений, зависящих от параметров $x^1, \ldots, x^n$, относительно m неизвестных: пусть

$$f_i\colon \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}, \qquad i = 1, \ldots, m,$$

причем

$$f_i(a^1, \ldots, a^n, b^1, \ldots, b^m) = 0, \qquad i = 1, \ldots m;$$

когда для каждого $(x^1, \ldots, x^n)$ вблизи $(a^1, \ldots, a^n)$ можно найти единственное $(y^1, \ldots, y^m)$ вблизи $(b^1, \ldots \ldots, b^m)$, удовлетворяющее уравнениям $f_i(x^1, \ldots, x^n, y^1, \ldots, y^m) = 0$, $i = 1, \ldots, m$? Ответ дает следующая теорема.

2.12. Теорема о неявной функции. *Предположим, что $f\colon \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^m$ непрерывно дифференцируема в некотором открытом множестве, содержащем (a, b), и $f(a, b) = 0$. Пусть M есть $(m \times m)$-матрица*

$$(D_{n+j} f^i(a)), \qquad 1 \leqslant i, j \leqslant m.$$

Тогда если $\det M \neq 0$, то существуют открытое множество $A \subset \mathbf{R}^n$, содержащее a, и открытое множество $B \subset \mathbf{R}^m$, содержащее b, со следующим свойством: для всякого $x \in A$ имеется единственное $g(x) \in B$, для которого $f(x, g(x)) = 0$. При этом функция $g\colon A \to B$ дифференцируема.

Доказательство. Определим $F\colon \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^n \times \mathbf{R}^m$ равенством $F(x, y) = (x, f(x, y))$. Тогда $\det F'(a, b) = \det M \neq 0$. В силу теоремы 2.11, существуют открытое множество $W \subset \mathbf{R}^n \times \mathbf{R}^m$, содержащее точку $F(a, b) = (a, 0)$, и содержащее точку (a, b) открытое множество в $\mathbf{R}^n \times \mathbf{R}^m$, которое можно считать имеющим вид $A \times B$, так что функция $F\colon A \times B \to W$ имеет дифференцируемую обратную $h\colon W \to A \times B$. Очевидно, h имеет вид $h(x, y) =$

$=(x,\ k(x,\ y))$, где k — некоторая дифференцируемая функция (поскольку F — функция такого вида). Пусть π: $\mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^m$ — функция, определенная равенством $\pi(x,\ y)=y$. Тогда $\pi \circ F = f$. Поэтому

$$f(x,\ k(x,\ y)) = f \circ h(x,\ y) =$$
$$= (\pi \circ F) \circ h(x,\ y) = \pi \circ (F \circ h)(x,\ y) = \pi(x,\ y) = y.$$

Таким образом, $f(x,\ k(x,\ 0))=0$, т. е. можно положить $g(x)=k(x,\ 0)$. ∎

Зная, что функция g дифференцируема, легко найти ее производную. Действительно, так как $f^i(x,\ g(x))=0$, то, применяя D_j к обеим частям этого равенства, получаем

$$0 = D_j f^i(x,\ g(x)) + \sum_{\alpha=1}^{m} D_{n+\alpha} f^i(x,\ g(x)) \cdot D_j g^{\alpha}(x),$$
$$i,\ j = 1,\ \ldots,\ m.$$

Поскольку $\det M \neq 0$, эта система уравнений относительно $D_j g^{\alpha}(x)$ разрешима. Ответ будет зависеть от значений $D_j f^i(x,\ g(x))$, а поэтому и от $g(x)$. Но это неизбежно, ибо функция g, вообще говоря, не единственна. Рассматривая функцию f: $\mathbf{R}^2 \to \mathbf{R}$, определяемую равенством $f(x,\ y)=x^2+y-1$, мы уже заметили, что уравнению $f(x,\ g(x))=0$ удовлетворяют две функции: $g(x)= \sqrt{1-x^2}$ и $g(x)=-\sqrt{1-x^2}$. Дифференцирование уравнения $f(x,\ g(x))=0$ дает здесь

$$D_1 f(x,\ g(x)) + D_2 f(x,\ g(x)) \cdot g'(x) = 0,$$

или $2x+2g(x)\cdot g'(x)=0$, т. е. $g'(x)=-x/g(x)$, а этому условию удовлетворяет и $g(x)=\sqrt{1-x^2}$ и $g(x)=-\sqrt{1-x^2}$. Несколько обобщив метод доказательства теоремы 2.12, мы получим результат, который будет иметь важное значение в гл. 5.

2.13. Теорема. *Пусть f: $\mathbf{R}^n \to \mathbf{R}^p$, $p \leqslant n$, непрерывно дифференцируема на открытом множестве, содержащем точку a. Если $f(a)=0$ и $(n \times p)$-матрица $(D_i f^j(a))$ имеет ранг p, то существуют открытое множество $A \subset \mathbf{R}^n$ и дифференцируемая функция*

h: $A \to \mathbf{R}^n$, имеющая дифференцируемую обратную, такие, что

$$f \circ h(x^1, \ldots, x^n) = (x^{n-p+1}, \ldots, x^n).$$

Доказательство. Мы можем рассматривать f как функцию из $\mathbf{R}^{n-p} \times \mathbf{R}^p$ в $\mathbf{R}^p$. Пусть M есть $(p \times p)$-матрица $(D_{n-p+i}f^j(a))$, $1 \leqslant i, j \leqslant p$. Если $\det M \neq 0$, то мы находимся точно в ситуации, рассмотренной при доказательстве теоремы 2.12, а там было показано, что существует такая функция h, что

$$f \circ h(x^1, \ldots, x^n) = (x^{n-p+1}, \ldots, x^n).$$

В общем же случае, поскольку $(D_i f^j(a))$ имеет ранг p, найдутся такие индексы $i_1 < \ldots < i_p$, что матрица $(D_i f^j(a))$, $1 \leqslant j \leqslant p$, $i = i_1, \ldots, i_p$, имеет ненулевой определитель. Пусть g: $\mathbf{R}^n \to \mathbf{R}^n$ — функция, переставляющая переменные x^i так, что

$$g(x^1, \ldots, x^n) = (\ldots, x^{i_1}, \ldots, x^{i_p}).$$

Тогда $f \circ q$ есть функция рассмотренного уже типа, так что $((f \circ g) \circ k)(x^1, \ldots, x^n) = (x^{n-p+1}, \ldots, x^n)$ при некоторой k. Искомой функцией будет $h = g \circ k$. ▌

Задачи

2.40. Решить задачу (2.15), используя теорему о неявной функции.

2.41. Пусть f: $\mathbf{R} \times \mathbf{R} \to \mathbf{R}$. Определим для каждого $x \in \mathbf{R}$ функцию g_x: $\mathbf{R} \to \mathbf{R}$ равенством $g_x(y) = f(x, y)$. Пусть при любом x существует единственное $y = c(x)$, для которого $g_x'(y) = 0$.

а) Показать, что если $D_{2,2}f(x, y) \neq 0$ для всех (x, y), то c дифференцируема и

$$c'(x) = -\frac{D_{2,1}f(x, c(x))}{D_{2,2}f(x, c(x))}.$$

(Указание: равенство $g_x'(y) = 0$ можно переписать в форме $D_2 f(x, y) = 0$.)

б) Показать, что если $c'(x)=0$, то существует y, для которого

$$D_{2,1}f(x, y)=0,$$
$$D_2 f(x, y)=0.$$

в) Пусть $f(x, y)=x(y\ln y-y)-y\ln x$. Найти

$$\max_{1/2\leqslant x\leqslant 2}\left(\min_{1/3\leqslant y\leqslant 1} f(x, y)\right).$$

ПО ПОВОДУ ОБОЗНАЧЕНИЙ

Этот параграф содержит краткое и не вполне беспристрастное обсуждение классических обозначений, связанных с частными производными. Приверженцы классических обозначений записывают частную производную $D_1 f(x, y, z)$ в виде

$$\frac{\partial f(x, y, z)}{\partial x} \quad \text{или} \quad \frac{\partial f}{\partial x} \quad \text{или}$$

$$\frac{\partial f}{\partial x}(x, y, z) \quad \text{или} \quad \frac{\partial}{\partial x} f(x, y, z)$$

или при помощи каких-либо других подходящих аналогичных символов. Такой способ записи ведет к тому, что вместо $D_1 f(u, v, w)$ пишут

$$\frac{\partial f}{\partial u}(u, v, w),$$

хотя могут (а для выражений типа $D_1 f(7, 3, 2)$ должны) использоваться символы вроде

$$\left.\frac{\partial f(x, y, z)}{\partial x}\right|_{(x, y, z)=(u, v, w)} \quad \text{или} \quad \frac{\partial f(x, y, z)}{\partial x}(u, v, w).$$

Аналогичные обозначения используются для $D_2 f$ и $D_3 f$. Производные высших порядков обозначаются символами типа

$$D_2 D_1 f(x, y, z)=\frac{\partial^2 f(x, y, z)}{\partial y\, \partial x}.$$

Для $f\colon \mathbf{R}\to\mathbf{R}$ символ ∂ автоматически заменяется первоначальным d, так что пишут $\frac{d\sin x}{dx}$, а не $\frac{\partial \sin x}{\partial x}$. Уже формулировка теоремы 2.2 потребовала бы в классической записи введения лишних букв.

Обычная запись вычисления $D_1(f\circ(g, h))$ такова: если $f(u, v)$ — некоторая функция и $u=g(x, y)$, $v=h(x, y)$, то

$$\frac{\partial f(g(x, y), h(x, y))}{\partial x}=\frac{\partial f(u, v)}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial f(u, v)}{\partial v}\frac{\partial v}{\partial x}.$$

[Символ $\partial u/\partial x$ означает $(\partial/\partial x)\, g(x, y)$, а $(\partial/\partial u) f(u, v)$ означает $D_1 f(u, v)=D_1 f(g(x, y), h(x, y))$.] Это равенство часто записывают просто в виде

$$\frac{\partial f}{\partial x}=\frac{\partial f}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial f}{\partial v}\frac{\partial v}{\partial x},$$

хотя f имеет в разных частях равенства неодинаковый смысл!

Обозначение df/dx, всегда представлявшееся преувеличенно соблазнительным, породило многие (обычно бессмысленные) определения для df и dx самих по себе, единственной целью которых было придать смысл равенству

$$df=\frac{df}{dx}\cdot dx.$$

Для $f\colon \mathbf{R}^2\to\mathbf{R}$, например, df *определяется* в классических курсах формулой

$$df=\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy$$

(что бы ни означали dx и dy).

Глава 4 содержит строгие определения, дающие возможность доказать вышеприведенные равенства. Действительно ли эти новые определения лучше классических — вопрос деликатный: пусть читатель судит о нем сам.