Лекция 18. Высшие производные

Посмотрим повнимательнее на вторую производную.

Пусть $f\colon U \to \mathbb{R}$ — функция класса C^2, где $U \subset \mathbb{R}^n$ — открытое множество. Отображение Df ставит в соответствие каждой точке $p \in U$ линейное отображение из $\mathbb{R}^n$ в $\mathbb{R}$, причем, если $(e_1, \ldots, e_n)$ — стандартный базис в $\mathbb{R}^n$, отображение $Df(p)$, как мы знаем, переводит e_i в $(\partial f/\partial x_i)(p)$. Поскольку линейные отображения из $\mathbb{R}^n$ в $\mathbb{R}$ также образуют n-мерное векторное пространство, можно считать, что Df — это отображение из U в n-мерное пространство, действующее по формуле

$$p \mapsto \Big(\frac{\partial f}{\partial x_1}(p), \ldots, \frac{\partial f}{\partial x_1}(p)\Big).$$

Поскольку f — функция класса C^2, отображение Df также дифференцируемо. Его производная в точке $p \in U$ — это линейное отображение двух n-мерных пространств, которое задается матрицей (b_{ij}), где $b_{ij} = \dfrac{\partial}{\partial x_j}\Big(\dfrac{\partial f}{\partial x_i}\Big)(p)$. (В силу предложения 17.12 функция принадлежит классу C^2 тогда и только тогда, когда все эти «вторые частные производные» существуют и непрерывны всюду на U.) В этом месте нас ожидает приятный сюрприз: матрица (b_{ij}) всегда симметрична!

Предложение 18.1. *Пусть $f\colon U \to \mathbb{R}$ — функция класса C^2, где $U \subset \mathbb{R}^n$ — открытое множество. Тогда для любых $i, j \in [1; n]$ на U выполнено тождество* $\dfrac{\partial}{\partial x_j}\Big(\dfrac{\partial f}{\partial x_i}\Big) = \dfrac{\partial}{\partial x_i}\Big(\dfrac{\partial f}{\partial x_j}\Big)$.

Доказательство. Зафиксировав значения всех переменных, кроме x_i и x_j, можно считать, что $U \subset \mathbb{R}^2$ и $f\colon U \to \mathbb{R}$ — функция двух переменных (мы их будем обозначать x и y). Частную производную функции f по x (соотв. по y) будем обозначать f_x (соотв. f_y) — так часто обозначают частные производные в случае, когда переменные обозначены разными буквами.

Заметим, что для доказательства тождества $(f_x)_y = (f_y)_x$ нам достаточно показать, что интегралы от этих функций по любому квадрату, содержащемуся в U, совпадают. В самом деле, если бы значения этих двух непрерывных функций различались в какой-то точке U, то на некотором квадрате, содержащем эту точку, разность $(f_x)_y - (f_y)_x$ была бы строго положительна или строго отрицательна, так что интегралы от $(f_x)_y$ и $(f_y)_x$ по этому квадрату были бы различны.

Пусть теперь $\Delta = [a; a+h] \times [b; b+h]$ — произвольный квадрат, содержащийся в U. Покажем, что интегралы по Δ от каждой из функций $(f_x)_y$ и $(f_y)_x$ равны $f(a+h, b+h) - f(a+h, b) - f(a, b+h) + f(a, b)$. Вот как это делается для $(f_x)_y$:

$$
\begin{aligned}
\int_\Delta (f_x)_y \, dxdy &= \int\limits_a^{a+h} \left(\int\limits_b^{b+h} (f_x)_y \, dy \right) dx = \int\limits_a^{a+h} (f_x(x, b+h) - f_x(x, b)) \, dx \\
&= \int\limits_a^{a+h} f_x(x, b+h) \, dx - \int\limits_a^{a+h} f_x(x, b) \, dx \\
&= (f(a+h, b+h) - f(a, b+h)) - (f(a+h, b) - f(a, b)) \\
&= f(a+h, a+h) - f(a+h, a) - f(a, a+h) + f(a, b).
\end{aligned}
$$

Если вычислить $(f_y)_x$, интегрируя сначала по dx, а затем по dy, получится, очевидно, тот же результат. Это завершает доказательство предложения. $\square$

Для функций одного переменного высшие производные участвуют в формуле Тейлора. Для функций нескольких переменных формула Тейлора также имеет место и даже, при подходящем выборе обозначений, записывается почти так же. С обозначений мы и начнем.

Зафиксируем натуральное число n («число переменных»). *Мультииндексом* называется последовательность n целых неотрицательных чисел. Если $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ — мультииндекс, то полагают $|\alpha| = \alpha_1 + \cdots + \alpha_n$, $\alpha! = \alpha_1! \ldots \alpha_n!$, и $x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}$; если еще f — функция класса $C^{|\alpha|}$, то полагают $D^\alpha f = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \ldots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} f$ (иными словами, то, что получится, если функцию f продифференцировать α_1 раз по x_1, α_2 раз по x_2 и т. д.; если α состоит из одних нулей, то $D^\alpha f = f$). Например, «полиномиальная теорема» записывается так: $(x_1 + \cdots + x_n)^m = \sum_{|\alpha|=m} \dfrac{m!}{\alpha!} x^\alpha$.

Теорема 18.2 (формула Тейлора). *Пусть f — функция класса C^r на открытом множестве $U \subset \mathbb{R}^n$. Тогда для всякой $a \in U$ имеем*

$$
f(a+h) = \sum_{|\alpha| \leqslant r} \frac{D^\alpha f(a)}{\alpha!} h^\alpha + R(a, h),
$$

где $|R(a, h)| \leqslant \varphi(a, h) \|h\|^r$ и $\lim_{h \to 0} \varphi(a, h) = 0$. Более того, у всякой точки a существует такая окрестность $V \subset U$, что $\varphi(a, h)$ стремится к нулю при $h \to 0$ равномерно по $a \in V$.

(На жаргоне это звучит так: $R(a, h) = o(\|h\|^r)$, причем o локально равномерно по a.)

Доказательство. Для всякой функции $\psi\colon U \to \mathbb{R}$ положим $\omega_\psi(\varepsilon) = \sup\limits_{\|x_1 - x_2\| \leqslant \varepsilon} \|\psi(x_1) - \psi(x_2)\|$. Если ψ равномерно непрерывна, то $\lim_{\varepsilon \to 0} \omega_\psi(\varepsilon) = 0$.

Лемма 18.3. *Пусть $f\colon (p, q) \to \mathbb{R}$ — функция класса C^r. Тогда для всякого $a \in (p, q)$ имеем*

$$\left| f(a + h) - \sum_{k=0}^{r} \frac{d^k f}{dx^k}(a) h^k \right| \leqslant \omega_{f^{(r)}}(|h|) |h|^r$$

Доказательство леммы. Положим $f_1(t) = f(a + t) - \sum_{k=0}^{r} \frac{d^k f}{dx^k}(a) t^k$. Тогда $d^s f / dt^s(a) = 0$ при $0 \leqslant s \leqslant r$; так как $f_1^{(r)}(t) = f^{(r)}(a + t) - f^{(r)}(a)$, имеем $\omega_{f^{(r)}}(\varepsilon) = \omega_{f_1^{(r)}}(\varepsilon)$.

Применим теперь r-кратно к функции f_1 на отрезке $[0; h]$ формулу Лагранжа. Получим

$$\begin{aligned} f_1(h) = f_1(h) - f_1(0) = h f_1'(\theta_1 h) = h \cdot \theta_1 h f_1''(\theta_2 \theta_1 h) = \ldots = \\ = h^r \theta_1 \ldots \theta_{r-1} f_1^{(r)}(\theta_1 \ldots \theta_r h), \end{aligned}$$

где $0 < \theta_i < 1$. Поскольку

$$|f_1^{(r)}(\theta_1 \ldots \theta_r h)| = |f_1^{(r)}(\theta_1 \ldots \theta_r h) - f_1^{(r)}(0)| \leqslant \omega_{f_1^{(r)}}(|h|) = \omega_{f^{(r)}}(|h|),$$

лемма доказана. □

Лемма 18.4.

$$\frac{d^l f(t)}{dt^l} = \sum_{|\alpha| = l} \frac{l!}{\alpha!} h^\alpha D^\alpha f(a + th).$$

Доказательство леммы. Ввиду следствия 17.10 имеем

$$\frac{d}{dt} f(a + th) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a + th). \tag{18.1}$$

Иными словами, однократное дифференцирование $f(a + th)$ по t равносильно применению к f оператора $h_1 \frac{\partial}{\partial x_1} + \ldots + h_n \frac{\partial}{\partial x_n}$ (с последующей подстановкой $a+th$ в f). Поскольку отдельные слагаемые $h_i\frac{\partial}{\partial x_i}$ коммутируют друг с другом (h_i являются константами и тем самым выносятся за производныые, а частные дифференцирования по различным переменным коммутируют ввиду предложения 18.1), при применении r-й степени этого оператора можно раскрыть скобки и привести подобные так же, как при воведении в r-ю степень суммы n переменных. Ввиду «полиномиальной теоремы» при этом получается именно то, что утверждалось. $\square$

Положим теперь

$$\omega(\varepsilon) = Nl! \cdot \max_{|\alpha|=r} \omega_{D^\alpha f}(\varepsilon),$$

где $N = \binom{r+n-1}{n-1}$ — количество мультииндексов степени r. Применяя лемму 18.3 к функции $t \mapsto f(a+th)$ на отрезке $[0;1]$, получим, с учетом леммы 18.4, что

$$\left| f(a+h) - \sum_{|\alpha|\leqslant r} \frac{D^\alpha f(a)}{\alpha!} h^\alpha \right| \leqslant \omega(|h|)\|h\|^r,$$

где $\|h\| = \max|h_i|$. Осталось заметить, что у всякой точки $a \in U$ существует окрестность V с компактным замыканием, содержащимся в U; на этой окрестности все функции $D^\alpha f$ при $|\alpha| \leqslant r$ являются равномерно непрерывными, поэтому если определять ω только на V, то $\lim_{\varepsilon\to 0}\omega(\varepsilon) = 0$, что и доказывает теорему. $\square$