Лекция 15. Высшие производные

Пусть $f\colon U \to \mathbb{R}$ — функция класса C^2, где $U \subset \mathbb{R}^n$ — открытое множество. Посмотрим повнимательнее на ее вторую производную.

Как мы выяснили в предыдущей лекции, 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 $Df(p)$ 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 (по крайней мере для вычислительных целей) 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, в соответствии с тем, что объяснялось на предыдущей лекции, 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 14.9 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! Именно, имеет место такое

Предложение 15.1. *Пусть $f\colon U \to \mathbb{R}$ — функция класса C^2, где $U \subset \mathbb{R}^n$ — открытое множество. Тогда и для любых целых $i, j \in [1; n]$ 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.

Начнем с простой леммы:

Лемма 15.2. *Пусть $f\colon U \to \mathbb{R}$ — непрерывная функция, где $U \subset \mathbb{R}$ — открытое множество. Если для всякого замкнутого прямоугольника $\Delta \subset U$ выполнено равенство $\int_\Delta f\,dxdy = 0$, то $f(x, y) = 0$ для всех $(x, y) \in U$.*

Доказательство леммы. Если найдется точка p, для которой $f(p) \neq 0$, то найдется содержащий ее замкнутый прямоугольник $\Delta \subset U$, на котором знак функции f постоянен. Пусть, для определенности, $f > 0$

всюду на Δ. Тогда существует такое число $m > 0$, что $f(q) \geqslant m$ для всех $q \in \Delta$; следовательно, интеграл f по Δ строго положителен, в противоречие с условием. $\square$

Пусть теперь $\Delta = [a; a+h] \times [b; b+k]$ — произвольный замкнутый прямоугольник, содержащийся в U. Ввиду доказанной леммы нам достаточно установить равенство

$$\int_\Delta (f_x)_y \, dxdy = \int_\Delta (f_y)_x \, dxdy \tag{15.1}$$

(из того, что f — класса C^2, вытекает, что $(f_x)_y$ и $(f_y)_x$ — непрерывные функции на U). И действительно, покажем, что обе части (15.1) равны $f(a+h, a+k) - f(a+h, a) - f(a, a+k) + f(a, b)$. Вот как это делается для левой части:

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

Если вычислить правую часть (15.1), интегрируя сначала по dx, а затем по dy, получится, очевидно, тот же результат. Это завершает доказательство равенства (15.1) и предложения. $\square$

Замечание 15.3. Симметрия частных производных обобщается (при надлежащем понимании) и на бесконечномерный случай, но с другим доказательством.

Из всего сказанного немедленно вытекает, что функция принадлежит классу C^r тогда и тогда, когда все ее частные производные порядка r существуют и непрерывны, а также что эти производные не зависят от того, в каком порядке производится дифференцирование.

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 (это будет «число переменных» — размерность пространства, на открытом подмножестве которого определена наша функция). *Мультииндексом* называется последовательность n целых неотрицательных чисел. Если $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ —

мультииндекс, то полагают $|\alpha| = \alpha_1 + \dots + \alpha_n$, $\alpha! = \alpha_1! \dots \alpha_n!$, и $x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}$; если еще f — функция класса $C^{|\alpha|}$, то полагают $D^\alpha f = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \dots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} f$ (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 $D^\alpha f = f$). Пример на эти обозначения: $(x_1 + \dots + x_n)^m = \sum_{|\alpha|=m} \dfrac{m!}{\alpha!} x^\alpha$ (подразумевается, что суммирование проводится по всем мультииндексам α, для которых $|\alpha| = m$). Теперь можно сформулировать

Предложение 15.4 (формула Тейлора). *Пусть f — функция класса C^r, определенная в окрестности точки $a = (a_1, \dots, a_n) \in \mathbb{R}^n$. Тогда*

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

при $x \to 0$.

Доказательство. Для достаточно близкого к началу координат $h \in \mathbb{R}^n$ положим $\varphi(t) = f(a + th)$. Очевидно, в некоторой окрестности нуля функция φ принадлежит классу C^r. Для начала вычислим производные от функции φ.

Лемма 15.5.
$$\frac{d^l \varphi(t)}{dt^l} = \sum_{|\alpha|=l} \frac{l!}{\alpha!} h^\alpha D^\alpha f(a + th).$$

Доказательство леммы. Ввиду следствия 14.7 имеем

$$d\psi/dt = \sum_{i=1}^n h_i \frac{\partial g}{\partial x_i}(a + th), \tag{15.2}$$

где g — гладкая функция, определенная в окрестности a, и $\psi(t) = g(a + th)$. Полагая в этом равенстве $g = f$, получаем утверждение леммы для $l = 1$. Далее будем рассуждать по индукции. Именно, пусть установлено равенство

$$\frac{d^{l-1} \varphi(t)}{dt^{l-1}} = \sum_{|\beta|=l-1} \frac{(l-1)!}{\beta!} h^\beta D^\beta f(a + th).$$

Продифференцируем его по t; при этом в правой части каждое слагаемое вида $\dfrac{(l-1)!}{\beta!} h^\beta D^\beta f(a + th)$ перейдет, в силу формулы (15.2), в

$$\sum_{i=1}^n \frac{(l-1)!}{\beta!} h_i h^\beta \frac{\partial D^\beta f}{\partial x_i}(a + th).$$

Стало быть,
$$\frac{d^l\varphi(t)}{dt^l} = \sum_{|\alpha|=l} c_\alpha h^\alpha D^\alpha f(a+th),$$
где для $\alpha = (\alpha_1, \ldots, \alpha_n)$ имеем
$$c_\alpha = \sum_{i=1}^{n} \frac{(l-1)!}{\alpha_1!\ldots\alpha_{i-1}!(\alpha_i-1)!\alpha_{i+1}!\ldots\alpha_n!} \tag{15.3}$$
(слагаемые, в которых в знаменателе под знаком факториала стоят отрицательные числа, считаем равными нулю).

Однако выражение в правой части равно $l!/(\alpha_1!\ldots\alpha_n!) = l!/\alpha!$. В этом можно убедиться несколькими разными способами (например, можно заметить, что такой же рекуррентной формуле удовлетворяют коэффициенты разложения $(x_1+\cdots+x_n)^l$). Это завершает индукцию и доказательство леммы. $\square$

Теперь вспомним, что по ходу доказательства одномерной формулы Тейлора с остаточным членом в форме Пеано (теорема 7.18) мы установили (для функции φ, принадлежащей классу C^r) неравенство
$$\left|\varphi(t) - \sum_{i=0}^{r} \frac{\varphi^{(i)}(c)}{i!}(t-c)^i\right| \leqslant M\cdot|t-c|^r,$$
где $M = \sup_{x\in[c;t]}|\varphi^{(r)}(x) - \varphi^{(r)}(c)|$ (напомним, что это доказывается r-кратным применением теоремы Лагранжа к выражению под модулем в левой части). Применив это неравенство к нашей φ, положив $c=0$ и $t=1$, получим, ввиду леммы, что
$$\left|f(a+h) - \sum_{|\alpha|\leqslant r}\frac{D^\alpha f(0)}{\alpha!}h^\alpha\right| \leqslant \sum_{|\alpha|\leqslant r}\frac{r!}{\alpha!}|h^\alpha| \sup_{0\leqslant t\leqslant 1}\left|D^\alpha f(a+th) - D^\alpha f(a)\right|. \tag{15.4}$$
Однако же $|h^\alpha| \leqslant \|h\|^{|\alpha|}$ (норма — максимум модулей координат), и
$$\lim_{h\to 0}\sup_{\substack{|\alpha|=r\\ 0\leqslant t\leqslant 1}}\left|D^\alpha f(a+th) - D^\alpha f(a)\right| = 0$$
в силу непрерывности всех частных производных порядка r. Стало быть, левая часть (15.4) не превосходит константы $\sum_{|\alpha|=r}(r!/\alpha!)$ (равной, собственно говоря, n^r, но сейчас это несущественно) на $\|h\|^r$ и на величину, стремящуюся к нулю при $h\to 0$. Это и означает, что «остаточный член» в многомерной формуле Тейлора есть $o(\|h\|^r)$, что и требовалось. $\square$

Вернемся теперь к началу лекции. Сейчас мы дадим более инвариантную формулировку свойства симметрии частных производных.

Пусть $f\colon U \to F$ — отображение класса C^2, где F — векторное пространство и U — открытое подмножество в векторном пространстве E (пространства E и F предполагаются конечномерными). Производная отображения f есть отображение $Df\colon U \to \mathrm{Hom}(E, F)$, а производная этой производной (существующая и непрерывная, так как f является C^2-гладким) есть отображение $D^2f\colon U \to \mathrm{Hom}(E, \mathrm{Hom}(E, F))$. Чтобы описать пространство, в которое бьет отображение D^2f, нам придется совершить небольшой экскурс в линейную алгебру.

Определение 15.6. Пусть E, F и G — векторные пространства (не обязательно конечномерные). Отображение $B\colon E \times F \to G$ называется *билинейным,* если $B(\lambda e_1 + \mu e_2, f) = \lambda B(e_1, f) + \mu B(e_2, f)$ и $B(e, \lambda f_1 + \mu f_2) = \lambda B(e, f_1) + \mu B(e, f_2)$ для любых $e, e_1, e_2 \in E$, $f, f_1, f_2 \in F$ и λ и μ из основного поля.

Иными словами, можно сказать, что отображение билинейно тогда и только тогда, когда оно линейно по каждому из своих аргументов (при фиксированном другом аргументе). Ясно, что билинейные отображения из $E \times F$ в G образуют векторное пространство, которое мы будем обозначать $\mathrm{Bil}(E, F; G)$ (общепринятого обозначения нет).

Основное свойство билинейных отображений сформулировано в следующем простом, но важном предложении.

Предложение 15.7. *Пусть E, F и G — векторные пространства. Тогда имеется естественный изоморфизм векторных пространств $\Phi\colon \mathrm{Bil}(E, F; G) \to \mathrm{Hom}(E, \mathrm{Hom}(F, G))$. Отображение Φ задается следующим образом: если B — билинейное отображение из $E \times F$ в G, то гомоморфизм $\Phi(B)$ переводит вектор $e \in E$ в гомоморфизм $f \mapsto B(e, f)$; обратное отображение Φ^{-1} переводит гомоморфизм $\varphi\colon E \to \mathrm{Hom}(F, G)$ в билинейное отображение $\Phi^{-1}(\varphi)\colon E \times F \to G$, действующее по правилу $(e, f) \mapsto (\varphi(e))(f)$.*

Словосочетание «естественный изоморфизм» означает, грубо говоря, что изоморфизм не зависит от произвольных выборов (например, мы не выбирали в пространствах базисов). Доказательство предложения (то есть проверка того, что отображения Φ и Φ^{-1} обладают всеми требуемыми свойствами) оставляется в качестве обязательного упражнения (это проще, чем разобрать написанное доказательство).

Вернемся к анализу. Инвариантная формулировка свойства симметрии частных производных выглядит так:

Предложение 15.8. *Пусть $\varphi\colon U \to F$ — отображение класса C^2, где F — конечномерное векторное пространство и U — открытое подмножество в конечномерном векторном пространстве E. Тогда для* *всякого $a \in U$ билинейное отображение $B_a : E \times E \to F$, соответствующее отображению $D^2\varphi(a) \colon E \to \mathrm{Hom}(E, F)$, является симметричным: $B_a(v_1, v_2) = B_a(v_2, v_1)$ для любых $v_1, v_2 \in E$.*

Доказательство. Пусть $E = \mathbb{R}^m$, $F = \mathbb{R}^n$, и отображение $f \colon U \to F$ действует по формуле

$$(x_1, \ldots, x_m) \mapsto (\varphi_1(x_1, \ldots, x_m), \ldots, \varphi_n(x_1, \ldots, x_m)).$$

Обозначим через $(e_1, \ldots, e_m)$ и $(f_1, \ldots, f_n)$ стандартные базисы в пространствах E и F соответственно. Как мы выяснили на предыдущей лекции,

$$Df(a)(e_j) = \sum_{i=1}^{n} \Bigl(\frac{\partial f_i}{\partial x_j}\Bigr)(a) \cdot f_i.$$

Выберем в пространстве $\mathrm{Hom}(E, F)$ базис, состоящий из гомоморфизмов h_{ij} $(1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m)$, где $h_{ij}(e_j) = f_i$ и $h_{ij}(e_k) = 0$ при $k \neq i$. В этих терминах можно сказать, что отображение Df действует по формуле

$$a \mapsto \sum_{i=1}^{n} \sum_{j=1}^{m} \Bigl(\frac{\partial f_i}{\partial x_j}\Bigr)(a) \cdot h_{ij}.$$

Соответственно, отображение $D^2 f(a)$, то есть производная Df в точке a, действует так:

$$D^2 f(a)(e_k) = \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{\partial}{\partial x_k} \Bigl(\frac{\partial \varphi_i}{\partial x_j}\Bigr)(a) \cdot h_{ij}.$$

Стало быть,

$$B_a(e_k, e_j) = (D^2 f(a)(e_k))(e_j) = \frac{\partial}{\partial x_k} \Bigl(\frac{\partial \varphi_i}{\partial x_j}\Bigr)(a) \cdot f_i;$$

аналогично,

$$B_a(e_j, e_k) = (D^2 f(a)(e_k))(e_j) = \frac{\partial}{\partial x_j} \Bigl(\frac{\partial \varphi_i}{\partial x_k}\Bigr)(a) \cdot f_i;$$

В силу того, что отображение F является C^2-гладким, таковы же и все φ_i, так что, по предложению 15.1, $\dfrac{\partial}{\partial x_k} \Bigl(\dfrac{\partial \varphi_i}{\partial x_j}\Bigr)(a) = \dfrac{\partial}{\partial x_k} \Bigl(\dfrac{\partial \varphi_i}{\partial x_k}\Bigr)(a)$. Стало быть, $B_a(e_k, e_j) = B_a(e_j, e_k)$ для любых j и k. Из того, что билинейное отображение симметрично на базисных векторах, немедленно следует, что оно симметрично и для любых пар векторов. $\square$