Лекция 33

Дифференцируемость сложного отображения и инвариантность формы первого дифференциала

Прежде всего докажем простой факт о линейных отображениях.

Предложение 1. *Пусть* $L : \mathbb{R}^d \to \mathbb{R}^m$ *линейное отображение. Тогда существует такое* $C > 0$, *что при всех* $\boldsymbol{x} \in \mathbb{R}^d$

$$\|L\boldsymbol{x}\| \le C\|\boldsymbol{x}\|.$$

Доказательство. Пусть в пространстве $\mathbb{R}^d$ задан базис $\boldsymbol{e}_1,\ \boldsymbol{e}_2,\ ...,\ \boldsymbol{e}_d$, а в пространстве $\mathbb{R}^m$ – базис $\boldsymbol{e}'_1,\ \boldsymbol{e}'_2,\ ...,\ \boldsymbol{e}'_m$. Координаты вектора $\boldsymbol{x} \in \mathbb{R}^d$ обозначим $x_1,\ x_2,\ ...,\ x_d$. В силу неравенства треугольника и определения евклидовой нормы имеем

$$\|L\boldsymbol{x}\| \le \sum_{j=1}^{d} \|L\boldsymbol{e}_j\| \cdot |x_j| \le \max_{j=1,\ ...,\ d} |x_j| \cdot \sum_{j=1}^{d} \|L\boldsymbol{e}_j\| \le \|\boldsymbol{x}\| \cdot \sum_{j=1}^{d} \|L\boldsymbol{e}_j\|,$$

и остаётся положить $C = \sum\limits_{j=1}^{d} \|L\boldsymbol{e}_j\|$. □

Доказательство следующей теоремы такое же, как и в одномерном случае, но сама формула для производной сложной функции многих переменных будет обсуждаться после доказательства.

Теорема 1. *(Дифференцируемость сложного отображения).* *Пусть отображение* $f : \mathbb{R}^d \mapsto \mathbb{R}^m$ *дифференцируемо в точке* $\boldsymbol{a} \in \mathbb{R}^d$, *а отображение* $g : \mathbb{R}^m \mapsto \mathbb{R}^n$ *дифференцируемо в точке* $f(\boldsymbol{a}) \in \mathbb{R}^m$. *Тогда сложное отображение* $g \circ f : \mathbb{R}^d \mapsto \mathbb{R}^n$ *дифференцируемо в точке* $\boldsymbol{a}$ *и* $d(g \circ f)(\boldsymbol{a}) = dg(f(\boldsymbol{a})) \circ df(\boldsymbol{a})$.

Другими словами, в условиях теоремы *дифференциал композиции равен композиции дифференциалов.*

Доказательство. Так как f и g дифференцируемы в точках $\boldsymbol{a}$ и $f(\boldsymbol{a})$ соответственно, то

$$f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) = df(\boldsymbol{a})\boldsymbol{h} + \alpha(\boldsymbol{h})\|\boldsymbol{h}\|,$$

и

$$g(f(\boldsymbol{a})+\boldsymbol{q}) - g(f(\boldsymbol{a})) = dg(\boldsymbol{a})\boldsymbol{q} + \beta(\boldsymbol{q})\|\boldsymbol{q}\|,$$

где вектора $\boldsymbol{h} \in \mathbb{R}^d$, а $\boldsymbol{q} \in \mathbb{R}^m$ берутся из проколотых окрестностей нуля в своих пространствах. Доопределим функции $\alpha : \mathbb{R}^d \mapsto \mathbb{R}^m$ и $\beta : \mathbb{R}^m \mapsto \mathbb{R}^n$ при $\boldsymbol{h} = \boldsymbol{0}$ и $\boldsymbol{q} = \boldsymbol{0}$, полагая $\alpha(\boldsymbol{0}) = \boldsymbol{0}$ и $\beta(\boldsymbol{0}) = \boldsymbol{0}$ (отметим, что $\alpha(\boldsymbol{0}) \in \mathbb{R}^m,\ \beta(\boldsymbol{0}) \in \mathbb{R}^n$). Далее имеем

$$\begin{aligned} g(f(\boldsymbol{a}+\boldsymbol{h})) - g(f(\boldsymbol{a})) &= g(f(\boldsymbol{a}) + (f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}))) - g(f(\boldsymbol{a})) = \\ &= dg(f(\boldsymbol{a}))(f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a})) + \beta(f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}))\|f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a})\| = \\ &= dg(f(\boldsymbol{a}))(df(\boldsymbol{a})\boldsymbol{h} + \alpha(\boldsymbol{h})\|\boldsymbol{h}\|) + \beta(f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}))\|df(\boldsymbol{a})\boldsymbol{h} + \alpha(\boldsymbol{h})\|\boldsymbol{h}\|\| = \\ &= dg(f(\boldsymbol{a}))(df(\boldsymbol{a})\boldsymbol{h}) + \gamma(\boldsymbol{h})\|\boldsymbol{h}\|, \end{aligned}$$

где

$$\gamma(\boldsymbol{h}) = dg(f(\boldsymbol{a}))\alpha(\boldsymbol{h}) + \beta(f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}))\|df(\boldsymbol{a})\frac{\boldsymbol{h}}{\|\boldsymbol{h}\|} + \alpha(\boldsymbol{h})\|.$$

Если мы докажем, что $\|\gamma(\boldsymbol{h})\| \to 0$ при $\|\boldsymbol{h}\| \to 0$, то это будет значить, что для композиции выполняется определение дифференцируемости.

Из неравенства треугольника получим

$$\|\gamma(\boldsymbol{h})\| \leq \|dg(f(\boldsymbol{a}))\alpha(\boldsymbol{h})\| + \|\beta(f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}))\| \cdot \left(\left\|df(\boldsymbol{a})\frac{\boldsymbol{h}}{\|\boldsymbol{h}\|}\right\| + \|\alpha(\boldsymbol{h})\|\right).$$

Так как $df(\boldsymbol{a})$ и $dg(f(\boldsymbol{a}))$ – это линейные отображения, то по предыдущему предложению существуют такие константы A и B, что для любых векторов $\boldsymbol{h} \in \mathbb{R}^d$ и $\boldsymbol{q} \in \mathbb{R}^m$ выполнены неравенства $\|df(\boldsymbol{a})\boldsymbol{h}\| \leq A\|\boldsymbol{h}\|$ и $\|dg(f(\boldsymbol{a}))\boldsymbol{q}\| \leq B\|\boldsymbol{q}\|$. Тогда

$$0 \leq \|\gamma(\boldsymbol{h})\| \leq B\|\alpha(\boldsymbol{h})\| + \|\beta(f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}))\| \cdot (A + \|\alpha(\boldsymbol{h})\|) \to 0, \text{ если } \|\boldsymbol{h}\| \to 0.$$

Таким образом, для композиции отображений $g \circ f : \mathbb{R}^d \mapsto \mathbb{R}^n$ выполнено определение дифференцируемости. □

Теперь обсудим, как композиция дифференциалов будет выглядеть в координатах. Пусть в пространствах $\mathbb{R}^d$, $\mathbb{R}^m$, $\mathbb{R}^n$ фиксированы базисы. Выясним, какой вид имеет матрица линейного отображения $d(g \circ f)(\boldsymbol{a})$. Матрицы отображений $df(\boldsymbol{a})$ и $dg(f(\boldsymbol{a}))$ – это матрицы Якоби отображений f и g точках $\boldsymbol{a}$ и $f(\boldsymbol{a})$ соответственно. Тогда композиции линейных отображений $d(g \circ f)(\boldsymbol{a})$ соответствует произведение матриц Якоби (так как композиции линейных операторов при заданных базисах соответствует произведение матриц этих операторов в данных базисах):

$$J_{(g\circ f)}(\boldsymbol{a}) = \begin{pmatrix} \frac{\partial(g\circ f)_1}{\partial x_1}(\boldsymbol{a}) & \dots & \frac{\partial(g\circ f)_1}{\partial x_d}(\boldsymbol{a}) \\ \dots & \dots & \dots \\ \frac{\partial(g\circ f)_n}{\partial x_1}(\boldsymbol{a}) & \dots & \frac{\partial(g\circ f)_n}{\partial x_d}(\boldsymbol{a}) \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(f(\boldsymbol{a})) & \dots & \frac{\partial g_1}{\partial y_m}(f(\boldsymbol{a})) \\ \dots & \dots & \dots \\ \frac{\partial g_n}{\partial y_1}(f(\boldsymbol{a})) & \dots & \frac{\partial g_n}{\partial y_m}(f(\boldsymbol{a})) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{a}) & \dots & \frac{\partial f_1}{\partial x_d}(\boldsymbol{a}) \\ \dots & \dots & \dots \\ \frac{\partial f_m}{\partial x_1}(\boldsymbol{a}) & \dots & \frac{\partial f_m}{\partial x_d}(\boldsymbol{a}) \end{pmatrix}.$$

В частном случае $n = 1$ для функции $g : \mathbb{R}^m \mapsto \mathbb{R}$ и отображения $f : \mathbb{R}^d \mapsto \mathbb{R}^m$ имеем:

$$J_{(g\circ f)}(\boldsymbol{a}) = \begin{pmatrix} \frac{\partial(g\circ f)}{\partial x_1}(\boldsymbol{a}) & \dots & \frac{\partial(g\circ f)}{\partial x_d}(\boldsymbol{a}) \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\partial g}{\partial y_1}(f(\boldsymbol{a})) & \dots & \frac{\partial g}{\partial y_m}(f(\boldsymbol{a})) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{a}) & \dots & \frac{\partial f_1}{\partial x_d}(\boldsymbol{a}) \\ \dots & \dots & \dots \\ \frac{\partial f_m}{\partial x_1}(\boldsymbol{a}) & \dots & \frac{\partial f_m}{\partial x_d}(\boldsymbol{a}) \end{pmatrix}.$$

Таким образом, j-й элемент матрицы Якоби функции $(g \circ f) : \mathbb{R}^d \mapsto \mathbb{R}$ представляет собой частные производные этой функции по переменной x_j :

$$\frac{\partial(g\circ f)}{\partial x_j}(\boldsymbol{a}) = \frac{\partial g}{\partial y_1}(f(\boldsymbol{a})) \cdot \frac{\partial f_1}{\partial x_j}(\boldsymbol{a}) + \dots + \frac{\partial g}{\partial y_m}(f(\boldsymbol{a})) \cdot \frac{\partial f_m}{\partial x_j}(\boldsymbol{a}). \tag{1}$$

В ещё более простом частном случае рассмотрим функцию $g : \mathbb{R}^m \mapsto \mathbb{R}$ и отображение $f : \mathbb{R} \mapsto \mathbb{R}^m$. Отображение f состоит из m координатных функций $f_j : \mathbb{R} \to \mathbb{R}$, $j = 1, \dots, m$, то есть координатные функции являются функциями одной переменной, поэтому можно говорить об их производных в обычном смысле (то есть как в первом семестре). В функцию $u = g(y_1, \dots, y_m)$ вместо переменных подставим функции $f_j : \mathbb{R} \to \mathbb{R}$, $j = 1, \dots, m$. Тогда сложная функция

$$(g \circ f)(t) = g(f_1(t), \dots, f_m(t))$$

является функцией одной переменной. Если все функции f_j дифференцируемы (а это значит, отображение f дифференцируемо) в точке $a \in \mathbb{R}$, а функция g дифференцируема в точке $f(a)$, то производную $g \circ f : \mathbb{R} \to \mathbb{R}$ можно найти, пользуясь доказанной теоремой:

$$(g\circ f)'(a) = \begin{pmatrix} \frac{\partial g}{\partial y_1}(f(a)) & \dots & \frac{\partial g}{\partial y_m}(f(a)) \end{pmatrix} \cdot \begin{pmatrix} f_1'(a) \\ \dots \\ f_m'(a) \end{pmatrix} = \frac{\partial g}{\partial y_1}(f(a))\cdot f_1'(a) + \dots + \frac{\partial g}{\partial y_m}(f(a))\cdot f_m'(a). \tag{2}$$

Теперь, используя выражения для дифференциала функции в координатах, мы можем сформулировать следующее важное следствие.

Теорема 2. ***(Инвариантность формы первого дифференциала).*** *Пусть отображение $f : \mathbb{R}^d \to \mathbb{R}^m$ дифференцируемо в точке $\boldsymbol{a}$, а функция $g : \mathbb{R}^m \to \mathbb{R}$ дифференцируема в точке $\boldsymbol{y} = f(\boldsymbol{a})$. Если в выражение $dg(\boldsymbol{y}) = \sum\limits_{k=1}^{m} \frac{\partial g}{\partial y_k}(\boldsymbol{y})dy_k$ вместо независимых переменных y_k $(k = 1, ..., m)$ подставить функции $y_k = f_k(\boldsymbol{x})$, то при $\boldsymbol{x} = \boldsymbol{a}$ полученное выражение будет дифференциалом сложной функции $u(\boldsymbol{x}) = g(f(\boldsymbol{x}))$, то есть выражение в координатах для первого дифференциала не изменяется, если вместо независимых переменных подставить функции.*

Доказательство. Подставим в выражение для дифференциала сложного отображения формулы вида (1) для частных производных $g \circ f$ и сгруппируем элементы при каждой производной $\frac{\partial g}{\partial y_k}(f(\boldsymbol{a}))$ $(k = 1, ..., m)$, после чего свернем выражения в скобках при частных производных $\frac{\partial g}{\partial y_k}$, пользуясь тем, что каждое такое выражение – это запись в координатах полного дифференциала функции f_k :

$$\begin{aligned} d(g\circ f)(\boldsymbol{a}) &= \sum_{j=1}^{d} \frac{\partial (g\circ f)}{\partial x_j}(\boldsymbol{a})dx_j = \\ &= \sum_{j=1}^{d} \left(\frac{\partial g}{\partial y_1}(f(\boldsymbol{a}))\cdot \frac{\partial f_1}{\partial x_j}(\boldsymbol{a}) + \dots + \frac{\partial g}{\partial y_m}(f(\boldsymbol{a}))\cdot \frac{\partial f_m}{\partial x_j}(\boldsymbol{a}) \right) dx_j = \\ &= \sum_{k=1}^{m} \left(\frac{\partial g}{\partial y_k}(f(\boldsymbol{a}))\cdot \left(\sum_{j=1}^{d} \frac{\partial f_k}{\partial x_j}(\boldsymbol{a})dx_j \right) \right) = \sum_{k=1}^{m} \frac{\partial g}{\partial y_k}(f(\boldsymbol{a}))df_k\boldsymbol{a}. \end{aligned}$$

С другой стороны

$$dg(f(\boldsymbol{a})) = \sum_{k=1}^{m} \frac{\partial g}{\partial y_k}(f(\boldsymbol{a}))dy_k.$$

Таким образом, получили тот же вид для полного дифференциала, только вместо y_k подставлено $f_k(\boldsymbol{a})$. □

В следующем разделе мы приведём примеры применения доказанных теорем.

Правила дифференцирования

Теперь мы можем сформулировать правила дифференцирования, которые легко доказываются с помощью инвариантности формы первого дифференциала.

Предложение 2. ***(Правила дифференцирования).*** *Пусть функции $f : \mathbb{R}^d \mapsto \mathbb{R}$ и $g : \mathbb{R}^d \mapsto \mathbb{R}$ дифференцируемы в точке $\boldsymbol{x} \in \mathbb{R}^d$. Тогда функции*

$$\alpha f + \beta g : \mathbb{R}^d \mapsto \mathbb{R}\ (\alpha, \beta \in \mathbb{R}),\ fg : \mathbb{R}^d \mapsto \mathbb{R} \text{ и } f/g : \mathbb{R}^d \mapsto \mathbb{R}\ (\text{при } g(\boldsymbol{x}) \neq 0)$$

дифференцируемы в точке $\boldsymbol{x}$ и

$$d(\alpha f+\beta g)(\boldsymbol{x})=\alpha df(\boldsymbol{x})+\beta dg(\boldsymbol{x}),\ d(fg)(\boldsymbol{x})=f(\boldsymbol{x})dg(\boldsymbol{x})+g(\boldsymbol{x})df(\boldsymbol{x}),$$
$$d\left(\frac{f}{g}\right)(\boldsymbol{x})=\frac{g(\boldsymbol{x})df(\boldsymbol{x})-f(\boldsymbol{x})dg(x)}{g^2(\boldsymbol{x})}.$$

Доказательство. Докажем только случай произведения функций, так как остальные случаи доказываются аналогично. Пусть $\varphi:\mathbb{R}^2\mapsto\mathbb{R}$ – функция двух вещественных переменных и $\varphi(x,y)=xy$. Тогда $d\varphi(x,y)=xdy+ydx$. Если $x=f(\boldsymbol{x})$, а $y=g(\boldsymbol{x})$, то выражение для дифференциала примет вид

$$d\varphi(f(\boldsymbol{x}),g(\boldsymbol{x}))=f(\boldsymbol{x})dg(\boldsymbol{x})+g(\boldsymbol{x})df(\boldsymbol{x})$$

в силу инвариантности формы первого дифференциала. □

Приведём примеры применения доказанных утверждений.

Пример 1. ***1)*** *Пусть функция $g:\mathbb{R}^m\to\mathbb{R}$ дифференцируема в точке $\boldsymbol{b}$. Рассмотрим единичный вектор $\boldsymbol{v}$ пространства $\mathbb{R}^d$, координаты которого, как известно, являются проекциями на координатные оси, то есть $\boldsymbol{v}=\begin{pmatrix}\cos\alpha_1\\ \dots\\ \cos\alpha_m\end{pmatrix}$. Координаты точки $\boldsymbol{b}$ – это набор чисел $(b_1,\ \dots,\ b_m)$. Тогда производная по направлению $\frac{\partial g}{\partial \boldsymbol{v}}(\boldsymbol{b})$ в силу доказанного в предыдущих лекциях равна $dg(\boldsymbol{b})\boldsymbol{v}$.*

Можно, однако, заметить, что в силу определения производной по направлению как предела мы считаем производную функции

$$u(t)=g(b_1+t\cdot\cos\alpha_1,\ \dots,\ b_m+t\cdot\cos\alpha_m)$$

в точке $t=0$. Тогда мы можем воспользоваться формулой (2), рассматривая отображение $f(t)=\begin{pmatrix}b_1+t\cdot\cos\alpha_1\\ \dots\\ b_m+t\cdot\cos\alpha_m\end{pmatrix}$. Согласно этой формуле

$$\frac{\partial g}{\partial \boldsymbol{v}}(\boldsymbol{b})=\frac{\partial g}{\partial y_1}(f(0))\cdot(b_1+t\cdot\cos\alpha_1)'\dots+\frac{\partial g}{\partial y_m}(f(0))\cdot(b_m+t\cdot\cos\alpha_m)'=$$
$$=\frac{\partial g}{\partial y_1}(\boldsymbol{b})\cos\alpha_1+\dots+\frac{\partial g}{\partial y_m}(\boldsymbol{b})\cos\alpha_m.$$

Легко видеть, что правая часть как раз и равна $dg(\boldsymbol{b})\boldsymbol{v}$, то есть произведению матрицы дифференциала на вектор $\boldsymbol{v}$.

2) *Вычислим частные производные функции $u(x,y,z)=x^2+y^2+z^2+xyz$:*

$$du=d(x^2+y^2+z^2+xyz)=2xdx+2ydy+2zdz+yzdx+xzdy+yzdx=$$
$$=(2x+yz)dx+(2y+xz)dy+(2z+xy)dz.$$

При dx стоит частная производная $\frac{\partial u}{\partial x}$, то есть $u'_x=2x+yz$. Аналогично,

$$u'_y=2y+xz,\ u'_z=2z+xy.$$

3) *Рассмотрим функцию $g:\mathbb{R}^3\mapsto\mathbb{R}$ и отображение $f:\mathbb{R}^2\mapsto\mathbb{R}^3$. Пусть*

$$y_1=x_1+x_2,\ y_2=x_1x_2,\ y_3=x_2\sin x_1.$$

Тогда в силу инвариантности формы первого дифференциала

$$\begin{aligned}d(g\circ f) = g'_{y_1}dy_1 + g'_{y_2}dy_2 + g'_{y_3}dy_3 = \\ = g'_{y_1}(dx_1+dx_2) + g'_{y_2}(x_1dx_2+x_2dx_1) + g'_{y_3}(x_2\cos x_1 dx_1 + \sin x_1 dx_2) = \\ = (g'_{y_1} + x_2 g'_{y_2} + (x_2\cos x_1)g'_{y_3})dx_1 + (g'_{y_1} + x_1 g'_{y_2} + (\sin x_1)g'_{y_3})dx_2 = \\ = (g\circ f)'_{x_1}dx_1 + (g\circ f)'_{x_2}dx_2.\end{aligned}$$

Из последнего равенства мы и находим частные производные сложного отображения:

$$(g\circ f)'_{x_1} = g'_{y_1} + x_2 g'_{y_2} + (x_2\cos x_1)g'_{y_3},\ \ (g\circ f)'_{x_2} = g'_{y_1} + x_1 g'_{y_2} + (\sin x_1)g'_{y_3}.$$