Лекция 35

Касательная плоскость и нормаль

Рассмотрим некоторые геометрические объекты, связанные с многомерными отображениями. Сразу отметим, что мы изучим только трёхмерный случай, однако все вводимые понятия могут быть обобщены на пространства и больших размерностей.

Перед изучением материала полезно повторить темы из аналитической геометрии об уравнении плоскости и уравнениях прямой в пространстве.

Поверхностью в $\mathbb{R}^3$ будем называть график любой функции $z = f(x, y)$, определённой и непрерывной в области (то есть связном открытом множестве) $\Omega \subset \mathbb{R}^2$, где под графиком понимается множество точек в $\mathbb{R}^3$ с координатами $(x, y, f(x, y)),\ (x, y) \in \Omega$.

Будет говорить, что поверхности $z = f_1(x, y)$ и $z = f_2(x, y)$ **являются касательными друг к другу в точке** (a, b, c), если $c = f_1(a, b) = f_2(a, b)$ и

$$f_1(x, y) - f_2(x, y) = o(\sqrt{(x-a)^2 + (y-b)^2})$$

при $\sqrt{(x-a)^2 + (y-b)^2} \to 0$.

Теорема 1. *(Теорема о касательной плоскости). Пусть функция $f : \mathbb{R}^2 \to \mathbb{R}$ дифференцируема в точке (a, b) и $z_0 = f(a, b)$. Тогда плоскость*

$$z - z_0 = \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b) \tag{1}$$

и поверхность $z = f(x, y)$ касаются друг друга в точке (a, b).

Доказательство. Обозначим вектор с координатами $\begin{pmatrix} x-a \\ y-b \end{pmatrix}$ через $\bar{h}$. Тогда уравнение (1) перепишется в виде $z = z_0 + df(a, b)\bar{h}$.

Обозначим правую часть этого уравнения через $g(x, y)$. В силу дифференцируемости f в точке (a, b) тогда получим для любой точки (x, y)

$$f(x, y) - g(x, y) = f(a, b) + df(a, b)\bar{h} + o(\|\bar{h}\|) - z_0 - df(a, b)\bar{h} = o(\|\bar{h}\|),$$

где $\|\bar{h}\| = \sqrt{(x-a)^2 + (y-b)^2}$. Стремление к нулю расстояния между точками (x, y) и (a, b) равносильно тому, что $x \to a,\ y \to b$.

Таким образом всё свелось к определению касательных поверхностей. □

Из последней теоремы мы видим, что *функция двух переменных дифференцируема в точке тогда и только тогда, когда у неё существует касательная плоскость в этой точке.*

Теперь напишем определение нормали к поверхности.

Определение 1. *Пусть функция $f : \mathbb{R}^2 \to \mathbb{R}$ дифференцируема в точке (a, b). Обозначим $z_0 = f(a, b)$. Тогда прямая, заданная каноническим уравнением*

$$\frac{x-a}{f'_x(a, b)} = \frac{y-b}{f'_y(a, b)} = \frac{z-z_0}{-1}$$

называется ***нормалью*** *к поверхности $z = f(x, y)$ в точке (a, b).*

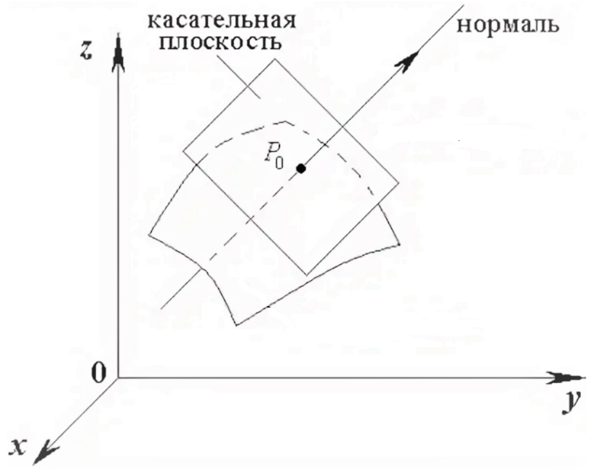

Рис. 1: Касательная плоскость и нормаль.

Мы видим, что нормаль – это прямая, перпендикулярная касательной плоскости и проходящая через точку касания.

Пример 1. *Найдём касательную плоскость и нормаль к* поверхности $z = x^2 - y^2$ *(она называется гиперболическим параболоидом или седлом) в точках* $(0, 0)$ *и* $(1, 2)$. *Имеем*

$$f'_x(0,0) = 2 \cdot 0 = 0 = f'_y(0,0), \ \ f'_x(1,2) = 2 \cdot 1 = 2 \ \text{и} \ f'_y(1,2) = -2 \cdot 2 = -4.$$

В точке $(0, 0)$ *значение* $z = 0$, *поэтому касательная плоскость в начале координат задаётся уравнением* $z = 0$, *а нормаль – уравнением* $\frac{x}{0} = \frac{y}{0} = \frac{z}{-1}$. *В точке* $(1, 2)$ $z = 1-4 = -3$,

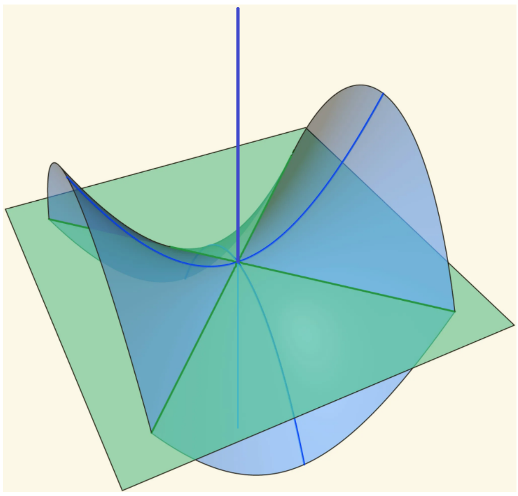

Рис. 2: Касательная плоскость и нормаль к гиперболическому параболоиду в $(0, 0)$.

поэтому касательная плоскость задаётся уравнением

$$z + 3 = 2(x - 1) - 4(y - 2) \ \Leftrightarrow \ 2x - 4y - z + 3 = 0,$$

а нормаль – системой уравнений

$$\frac{x-1}{2} = \frac{y-2}{-4} = \frac{z+3}{-1}.$$

Полезно записать уравнение касательной плоскости и нормали для поверхности, заданной неявно и параметрически.

Определение 2. *Пусть частная производная* $\frac{\partial f}{\partial x_j}(\boldsymbol{a})$ *функции* $f : \mathbb{R}^d \to \mathbb{R}$ *существует в окрестности* $U(\boldsymbol{a})$ *точки* $\boldsymbol{a}$ *и представляет собой функцию* $u = f'_{x_j}(\boldsymbol{x})$ $(\boldsymbol{x} \in U(\boldsymbol{a}))$, *имеющую частную производную по переменной* x_i *в точке* $\boldsymbol{a}$. *Тогда говорят, что функция* f *имеет частную производную второго порядка по переменным* x_j *и* x_i *в точке* $\boldsymbol{a}$ *и пишут*

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(\boldsymbol{a}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a})$$

или $(f'_{x_j})'_{x_i}(\boldsymbol{x}) = f''_{x_j x_i}(\boldsymbol{a})$.

Частные производные третьего и последующих порядков определяются аналогично. Частные производные высших порядков, взятые по разным переменным, называются *смешанными производными*. Например, при $j \neq i$ $f''_{x_i x_j}(\boldsymbol{a})$ является смешанной производной. В общем случае смешанные производные не равны, то есть, например, $f''_{x_i x_j}(\boldsymbol{a}) \neq f''_{x_j x_i}(\boldsymbol{a})$. Сейчас мы сформулируем теоремы, в которых будут указаны условия, достаточные для равенства смешанных производных второго порядка.

Теорема 2. *(*Теорема Шварца*). Если функция* $f : \mathbb{R}^d \to \mathbb{R}$ *в некоторой окрестности точки* $\boldsymbol{a}$ *имеет смешанные частные производные второго порядка* $f''_{x_i x_j}$ *и* $f''_{x_j x_i}$, *которые непрерывны в точке* $\boldsymbol{a}$, *то* $f''_{x_i x_j}(\boldsymbol{a}) = f''_{x_j x_i}(\boldsymbol{a})$.

Доказательство этой теоремы мы не приводим. Его можно также найти, например, в первом томе книги В. А. Зорича.

Упражнение 1. Показать, что $f''_{xy}(0,0) \neq f''_{yx}(0,0)$, если:

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

Какое условие теоремы Шварца нарушено?

Теорема 3. *(*Теорема Юнга*). Если функции* $f'_{x_1} : \mathbb{R}^2 \to \mathbb{R}$ *и* $f'_{x_2} : \mathbb{R}^2 \to \mathbb{R}$ *определены в некоторой окрестности точки* $\boldsymbol{a} = (a_1, a_2)$ *и дифференцируемы в самой точке* $\boldsymbol{a}$, *то* $f''_{x_1 x_2}(\boldsymbol{a}) = f''_{x_2 x_1}(\boldsymbol{a})$.

Доказательство. Пусть $g(x_1) = f(x_1, a_2 + h) - f(x_1, a_2)$. Тогда

$$g(a_1 + h) - g(a_1) = f(a_1 + h, a_2 + h) - f(a_1 + h, a_2) - f(a_1, a_2 + h) + f(a_1, a_2)$$

и по теореме Лагранжа

$$g(a_1 + h) - g(a_1) = hg'(a_1 + \theta_1 h) = h(f'_{x_1}(a_1 + \theta_1 h, a_2 + h) - f'_{x_1}(a_1 + \theta_1 h, a_2)).$$

Используя дифференцируемость функции f'_{x_1} в точке $\boldsymbol{a}$, получим равенства (убедитесь, что они получаются по определению дифференцируемости):

$$f'_{x_1}(a_1 + \theta_1 h, a_2 + h) - f'_{x_1}(a_1, a_2) = f''_{x_1 x_1}(\boldsymbol{a})\theta_1 h + f''_{x_1 x_2}(\boldsymbol{a})h + o(h)$$

и

$$f'_{x_1}(a_1 + \theta_1 h, a_2) - f'_{x_1}(a_1, a_2) = f''_{x_1 x_1}(\boldsymbol{a})\theta_1 h + o(h),$$

поэтому

$$\begin{aligned}
&g(a_1+h)-g(a_1)= \\
&\qquad = h(f'_{x_1}(a_1+\theta_1 h, a_2+h) - f'_{x_1}(a_1,a_2) - f'_{x_1}(a_1+\theta_1 h, a_2) + f'_{x_1}(a_1,a_2)) = \\
&\qquad\qquad = h^2 f''_{x_1x_2}(\boldsymbol{a}) + o(h^2).
\end{aligned}$$

Если же положить $p(x_2)=f(a_1+h,x_2)-f(a_1,x_2)$, то точно также получим, что

$$\begin{aligned}
&p(a_2+h)-p(a_2)= \\
&\qquad = f(a_1+h,a_2+h)-f(a_1+h,a_2)-f(a_1,a_2+h)+f(a_1,a_2)= \\
&\qquad\qquad = h^2 f''_{x_2x_1}(\boldsymbol{a}) + o(h^2).
\end{aligned}$$

Таким образом, $f''_{x_1x_2}(\boldsymbol{a})+o(1)=f''_{x_2x_1}(\boldsymbol{a})+o(1)$, откуда, переходя к пределу при $h\to 0$, получим $f''_{x_1x_2}(\boldsymbol{a})=f''_{x_2x_1}(\boldsymbol{a})$. □

Отметим, что теорема Юнга верна и для функций большего числа переменных. Для доказательства следует зафиксировать все переменные, кроме тех двух, равенство смешанных производных по которым мы доказываем, и применить уже доказанный вариант теоремы для функции двух переменных.

Также полезно отметить, что теоремы Юнга и Шварца не следуют друг из друга. В первой из них не требуется дифференцируемости частных производных, зато нужны существование и непрерывность в точке смешанных производных, а в теореме Юнга не требуется существование смешанных производных, зато нужна дифференцируемость первых частных производных в точке.

Дифференциалы высших порядков

Определение 3. *Функция $f:\mathbb{R}^d\to\mathbb{R}$ дважды дифференцируема в точке $\boldsymbol{a}\in\mathbb{R}^d$, если её частные производные первого порядка дифференцируемы в этой точке. Если все частные производные функции f порядка $(n-1)$ дифференцируемы в точке, то функция f называется n раз дифференцируемой в этой точке.*

Сразу заметим, что из определения дифференцируемости порядка n следует, что производные $(n-1)$-го порядка дифференцируемы, откуда по теореме Юнга следует, что **смешанные производные порядка $\boldsymbol{n}$ дифференцируемой $\boldsymbol{n}$ раз функции равны.**

Кроме того, из достаточных условий дифференцируемости, получаем следующее утверждение.

Предложение 1. *Пусть все частные производные функции f порядка n существуют в некоторой окрестности точки $\boldsymbol{a}$ и непрерывны в точке $\boldsymbol{a}$. Тогда функция f n раз дифференцируема в точке $\boldsymbol{a}$.*

Доказательство. Для доказательства используем индукцию по n. База – это достаточное условие дифференцируемости, где используется непрерывность частных производных. Если функция f дифференцируема $n-1$ раз, то она имеет все частные производные порядка $(n-1)$, а каждая из этих производных по условию имеет все частные производные в некоторой окрестности точки $\boldsymbol{a}$, непрерывные в точке $\boldsymbol{a}$. Тогда, снова по достаточному условию дифференцируемости, все эти частные производные дифференцируемы в точке $\boldsymbol{a}$, то есть выполнено определение дифференциала порядка n. □

Определим теперь второй дифференциал и изучим вопрос о том, как он выражается через частные производные.

Итак, пусть функция $f : \mathbb{R}^d \to \mathbb{R}$ дифференцируема и все частные производные f'_{x_j} $(j = 1, ..., d)$ дифференцируемы в точке $\boldsymbol{a} \in \mathbb{R}^d$. Если f дифференцируема в некоторой окрестности U точки $\boldsymbol{a}$, то при каждом фиксированном $\boldsymbol{h} \in \mathbb{R}^d$ из некоторой проколотой окрестности нуля и при каждом $\boldsymbol{x} \in U$ мы можем рассмотреть определённую на U функцию

$$g(\boldsymbol{x}) = df(\boldsymbol{x})\boldsymbol{h} = f'_{x_1}(\boldsymbol{x})h_1 + ... + f'_{x_d}(\boldsymbol{x})h_d.$$

Функция f дважды дифференцируема в точке $\boldsymbol{a}$, так как все частные производные дифференцируемы в этой точке. Тогда

$$dg(\boldsymbol{a})\boldsymbol{q} = d(df(\boldsymbol{a})\boldsymbol{h})\boldsymbol{q} = \sum_{j=1}^{d} f''_{x_1 x_j}(\boldsymbol{a})q_j h_1 + ... + \sum_{j=1}^{d} f''_{x_d x_j}(\boldsymbol{a})q_j h_d = \sum_{j,i=1}^{d} f''_{x_j x_i}(\boldsymbol{a})q_j h_i.$$

Полученная билинейная форма является симметричной, поэтому по ней можно построить квадратичную форму, полагая $\boldsymbol{q} = \boldsymbol{h}$:

$$d(df(\boldsymbol{a})\boldsymbol{h})\boldsymbol{h} = \sum_{j,i=1}^{d} f''_{x_j x_i}(\boldsymbol{a})dx_j(\boldsymbol{h})dx_i(\boldsymbol{h}) = \sum_{j,i=1}^{d} f''_{x_j x_i}(\boldsymbol{a})h_j h_i.$$

Это и есть второй дифференциал. Таким образом, *второй дифференциал записывается в виде следующей квадратичной формы:* $d^2 f(\boldsymbol{a}) = \sum\limits_{j,i=1}^{d} f''_{x_j x_i}(\boldsymbol{a})dx_j dx_i$.

Аналогично определяется *дифференциал порядка n в точке $\boldsymbol{a}$* :

$$d^n f(\boldsymbol{a}) = \sum_{1 \le j_1, ..., j_n \le d} \frac{\partial^n f}{\partial x_{j_1} ... \partial x_{j_n}}(\boldsymbol{a})dx_{j_1} ... dx_{j_n}.$$

Пример 2. *1) Найдём второй дифференциал функции $u(x, y, z) = x^2 + y^2 + z^2 + xyz$. Для этого вычислим частные производные функции u* :

$$u''_{xx} = (u'_x)'_x = (2x + yz)'_x = 2 = u''_{yy} = u''_{zz},\ u''_{xy} = z,\ u''_{xz} = y,\ u''_{yz} = x.$$

Так как функция u удовлетворяет всем условиям теоремы Юнга, то её смешанные производные равны. Таким образом,

$$d^2 u = 2dx^2 + 2dy^2 + 2dz^2 + 2zdxdy + 2ydxdz + 2xdydz.$$

2) Считая частные производные третьего порядка функции $u(x, y, z) = x^3 + y^3 + z^3 + xyz$, получим $d^3 u = 6(dx^3 + dy^3 + dz^3 + dxdydz)$.

Считая дифференциалы старших порядков, часто прибегают к упрощенной записи, суть которой поясним на примере вычисления дифференциалов функции двух переменных.

Для функции $u = f(x, y)$ имеем

$$du = f'_x dx + f'_y dy,\ d^2 u = f'''_{xx} dx^2 + 2f''_{xy} dxdy + f''_{yy} dy^2,$$
$$d^3 u = f'''_{xxx} dx^3 + 3f'''_{xxy} dx^2 dy + 3f'''_{yyx} dy^2 dx + f'''_{yyy} dy^2$$

и так далее. Мы видим, что коэффициенты при частных производных совпадают с биномиальными коэффициентами для степеней того же порядка, поэтому, используя индукцию, можно записать дифференциал порядка n в виде $d^n u = \left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)^n \cdot f$. Скобки раскрываются также, как и в биноме Ньютона, а запись $\frac{\partial^n}{\partial x_{j_1}...\partial x_{j_d}} dx_{j_1}...dx_{j_d} \cdot f$ равносильна взятию соответствующей частной производной, то есть $\frac{\partial^n}{\partial x_{j_1}...\partial x_{j_d}} dx_{j_1}...dx_{j_d} \cdot f = \frac{\partial^n f}{\partial x_{j_1}...\partial x_{j_d}} dx_{j_1}...dx_{j_d}$.

Аналогично можно записать дифференциал функции любого числа переменных, то есть если дана n раз дифференцируемая функция $f : \mathbb{R}^d \mapsto \mathbb{R}$, то

$$d^n f = \left(\frac{\partial}{\partial x_1}dx_1 + ...\frac{\partial}{\partial x_d}dx_d\right)^n \cdot f.$$

Формула Тейлора

В многомерном случае, как и в одномерном, справедливы формулы Тейлора с остатками в разных формах. Здесь ограничимся лишь остатком в форме Лагранжа.

Условимся, прежде всего, что под записью $d^n f(\boldsymbol{a})(\boldsymbol{h})$ понимается значение дифференциала порядка n на векторе $\boldsymbol{h}$, то есть

$$d^n f(\boldsymbol{a})(\boldsymbol{h}) = \sum_{1\le j_1,...,j_n \le d} \frac{\partial^n f}{\partial x_{j_1}...\partial x_{j_n}}(\boldsymbol{a}) dx_{j_1}(\boldsymbol{h})...dx_{j_n}(\boldsymbol{h}) = \sum_{1\le j_1,...,j_n \le d} \frac{\partial^n f}{\partial x_{j_1}...\partial x_{j_n}}(\boldsymbol{a}) h_{j_1} \cdot ... \cdot h_{j_n}.$$

Теорема 4. ***(Формула Тейлора с остаточным членом в форме Лагранжа).*** *Пусть функция* $f : \mathbb{R}^d \to \mathbb{R}$ *дифференцируема* $n+1$ *раз в некоторой окрестности* U *точки* $\boldsymbol{a}$. *Тогда для любой точки* $\boldsymbol{x} \in U$ *существует такая точка* $\boldsymbol{c} = \boldsymbol{a} + \theta(\boldsymbol{x} - \boldsymbol{a})$, *где* $0 < \theta < 1$, *что справедливо равенство*

$$f(\boldsymbol{x}) = f(\boldsymbol{a}) + df(\boldsymbol{a})\boldsymbol{h} + \frac{1}{2!}d^2 f(\boldsymbol{a})(\boldsymbol{h}) + ... + \frac{1}{n!}d^n f(\boldsymbol{a})(\boldsymbol{h}) + \frac{d^{n+1} f(\boldsymbol{c})}{(n+1)!}.$$

причём $\boldsymbol{h} = \boldsymbol{x} - \boldsymbol{a}$.

Доказательство. Доказательство сводится к аналогичному утверждению для функции одной переменной. Пусть $g(t) = f(\boldsymbol{a} + t(\boldsymbol{x} - \boldsymbol{a}))$. Эта функция $n+1$ раз дифференцируема в каждой точке отрезка $[0,1]$, так как функция f дифференцируема при $t \in [0,1]$ в каждой точке $\boldsymbol{a} + t(\boldsymbol{x} - \boldsymbol{a})$ по условию. Для функции $g : \mathbb{R} \mapsto \mathbb{R}$ можно записать формулу Тейлора с остаточным членом в форме Лагранжа при $t = 1$:

$$g(1) = g(0) + \frac{g'(0)}{1!} + ... + \frac{g^{(n)}(0)}{n!} + \frac{g^{(n+1)}(\theta)}{(n+1)!},$$

где $\theta \in (0,1)$ – некоторая постоянная.

Для функции g вычислим несколько производных в точке $t = 0$. Применяя все те же рассуждения, которые применялись при выводе формулы для производной по направлению, получим

$$g'(0) = \lim_{t\to 0} \frac{g(t) - g(0)}{t} = df(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a}).$$

Рассмотрим функцию

$$g'(\boldsymbol{a} + t(\boldsymbol{x} - \boldsymbol{a})) = df(\boldsymbol{a} + t(\boldsymbol{x} - \boldsymbol{a}))(\boldsymbol{x} - \boldsymbol{a}).$$

Для этой функции производная $g''(0)$ вычисляется также, то есть

$$g''(0) = d(df(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a}))(\boldsymbol{x} - \boldsymbol{a}) = d^2 f(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a}).$$

Далее можно рассуждать аналогично. Таким образом, справедливы равенства

$$g^{(j)}(0) = d^j f(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a}), \text{ где } j = 1, ..., n, \text{ а } g^{(n+1)}(\theta) = d^{n+1} f(\boldsymbol{c})(\boldsymbol{x} - \boldsymbol{a}),$$

а тогда при подстановке их в формулу Тейлора для g, получаем доказываемое равенство. □