Лекция 15.

Производная по направлению. Градиент функции.

Рассмотрим некоторое обобщение понятия частной производной функции многих переменных.

Пусть $M_0(x_0, y_0, z_0) \in \mathbb{R}^3$ - внутренняя точка области определения функции $f(x, y, z)$. Пусть задан вектор $e \in \mathbb{R}^3$, $\|e\| = 1$. Тогда $e = (\cos\alpha, \cos\beta, \cos\gamma)$, где α, β, γ – углы между вектором e и осями Ox, Oy и Oz соответственно.

Рассмотрим функцию $g(t) = f(x_0 + t\cos\alpha, y_0 + t\cos\beta, z_0 + t\cos\gamma)$, где $t \in \mathbb{R}$ – вещественный параметр.

Определение 1. *Производной функции* $f(x)$ *по направлению вектора* $e = (\cos\alpha, \cos\beta, \cos\gamma)$ в точке $M_0(x_0, y_0, z_0)$ называется производная сложной функции $g(t)$ в точке $t_0 = 0$, то есть число

$$\frac{\partial f}{\partial e}(M_0) = \lim_{t\to 0}\frac{f(M_0 + te) - f(M_0)}{t} = \lim_{t\to 0}\frac{g(t) - g(0)}{t} = g'(0).$$

Предположим, что функция $f(x)$ дифференцируема в точке M_0. Тогда функция $g(t)$ дифференцируема в нуле как сложная функция и справедлива формула:

$$\frac{\partial f}{\partial e}(M_0) = f'_x(M_0)\cdot\cos\alpha + f'_y(M_0)\cdot\cos\beta + f'_z(M_0)\cdot\cos\gamma.$$

Для случая функции двух переменных формула принимает вид:

$$\frac{\partial f}{\partial e}(M_0) = f'_x(M_0)\cdot\cos\alpha + f'_y(M_0)\cdot\sin\alpha,$$

где α - угол между вектором e и осью Ox.

Пусть теперь $x^0 = (x_1^0, \ldots, x_n^0)$ - фиксированная точка, внутренняя для области определения функции $f(x) = f(x_1, \ldots, x_n)$, и пусть задан вектор e, $e \in \mathbb{R}^n$ $\|e\| = 1$. В этом случае координаты вектора e равны его направляющим косинусам: $e = (\cos\alpha_1, \ldots, \cos\alpha_n)$, где α_i - угол между осью Ox_i и вектором e, $i = 1, \ldots, n$.

Рассмотрим функцию $g(t) = f(x^0 + te) = f(x_1^0 + t\cos\alpha_1, \ldots, x_n^0 + t\cos\alpha_n)$, где $t \in \mathbb{R}$ – вещественный параметр.

Определение 2. *Производной функции* $f(x)$ *по направлению* $e = (\cos\alpha_1, \ldots, \cos\alpha_n)$ в точке x^0 называется производная сложной функции $g(t)$ в точке $t_0 = 0$, то есть число

$$\frac{\partial f}{\partial e}(x^0) = \lim_{t\to 0}\frac{f(x^0 + te) - f(x^0)}{t} = \lim_{t\to 0}\frac{g(t) - g(0)}{t} = g'(0). \tag{1}$$

Замечание. Из наличия у функции производной по любому направлению в некоторой точке не следует, вообще говоря, ее дифференцируемость в этой точке. Например, у функции $f(x) = \begin{cases} \frac{x^4y^2}{x^8+y^4}, & x^2 + y^2 \neq 0, \\ 0, & x^2 + y^2 = 0 \end{cases}$ производная в точке $(0, 0)$ по любому направлению равна нулю, но функция не является даже непрерывной в этой точке (проверьте!)

Предположим, что функция $f(x)$ дифференцируема в точке x^0. Тогда функция $g(t)$ дифференцируема в точке $t_0 = 0$ как сложная функция, и производная (1) легко вычисляется по правилу дифференцирования сложной функции. Получаем формулу:

$$\frac{\partial f}{\partial e}(x^0) = f'_{x_1}(x^0)\cdot\cos\alpha_1 + \ldots + f'_{x_n}(x^0)\cdot\cos\alpha_n. \tag{2}$$

Из формулы (2) видно, что производная по направлению есть скалярное произведение вектора e и вектора частных производных функции $f(x)$. Вектор $grad\, f(x^0) = (f'_{x_1}(x^0),..., f'_{x_n}(x^0))$ называется *градиентом* функции $f(x)$ в точке x_0.

Таким образом, получаем равенство:

$$\frac{\partial f}{\partial e}(x^0) = (grad\ f(x^0), e). \tag{3}$$

Градиент часто представляют в виде: $grad\, f = \nabla f$ (читается: «набла *f*»), где $\nabla = \left(\frac{\partial}{\partial x_1},..., \frac{\partial}{\partial x_n} \right)$- так называемый *абстрактный вектор-оператор градиента*.

Отметим кстати, что формулу для дифференциала функции можно записать в следующем виде: $df = (grad f, \Delta x) = (\nabla f, \Delta x)$.

Что характеризует градиент функции? Какими свойствами обладает? Выясним это подробнее.

Лемма 1. Градиент функции (в данной точке) – это вектор, направление которого есть направление наибольшей скорости роста функции, а норма градиента равна этой наибольшей скорости роста.

Доказательство. Из формулы (3) получаем:

$$\frac{\partial f}{\partial e}(x^0) = (grad\ f(x^0), e) = \|grad\ f\| \cdot \|e\| \cdot \cos(grad\ f,^\wedge e).$$

Поскольку $\cos(grad\, f,^\wedge e) \le 1$, и достигает своего наибольшего значения 1, когда векторы сонаправлены, то легко видеть, что максимальное значение производной $\frac{\partial f}{\partial e}(x^0)$ будет в том и только в том случае, когда $e = \frac{grad\ f(x^0)}{\|grad\ f(x^0)\|}$, то есть когда вектор e совпадает с ортом градиента f (в точке x^0).

Какова же максимальная скорость роста функции f? Из (3), при $e = \frac{grad\ f(x^0)}{\|grad\ f(x^0)\|}$, получаем: $\frac{\partial f}{\partial e}(x^0) = (grad\ f(x^0), \frac{grad\ f(x^0)}{\|grad\ f(x^0)\|}) = \|grad\ f(x^0)\|$. Лемма доказана.

Замечание. Поскольку, в силу леммы 1, направление и норма градиента есть направление и величина максимальной скорости роста функции (в данной точке), то градиент $grad\, f(x)$ не зависит от выбора системы координат.

Рассмотрим теперь направление градиента функции по отношению к её *поверхности уровня*, то есть к геометрическому месту точек, определяемому уравнением: $P_c : f(x) = f(x_1,...,x_n) = c$, где c - некоторая константа.

Лемма 2. Градиент дифференцируемой в точке x^0 функции $f(x)$ ортогонален её поверхности уровня P_c, проходящей через точку x^0.

Доказательство. Пусть в малой окрестности точки x^0 взята произвольная точка $x, x \in P_c, x - x^0 = \Delta x \ne 0$. В силу дифференцируемости функции $f(x)$ в точке x^0, имеем: $0 = \Delta f = f(x^0 + \Delta x) - f(x^0) = (grad\, f(x^0), \Delta x) + \bar{o}(\|\Delta x\|)$. Разделив это равенство на $\|\Delta x\|$, получим:

$$0=\frac{\Delta f}{\|\Delta x\|}=f'_{x_1}(x^0)\frac{\Delta x_1}{\|\Delta x\|}+\ldots+f'_{x_n}(x^0)\frac{\Delta x_n}{\|\Delta x\|}+\frac{\bar{o}(\|\Delta x\|)}{\|\Delta x\|}=(grad\, f(x^0),\frac{\Delta x}{\|\Delta x\|})+\frac{\bar{o}(\|\Delta x\|)}{\|\Delta x\|}.$$

При переходе к пределу при $\Delta x\to 0$ в последнем соотношении вектор $\frac{\Delta x}{\|\Delta x\|}$ превращается в касательный вектор $e_{\text{кас.}}$ в точке x^0 к поверхности P_c, и получается равенство:

$(grad\, f(x_0),e_{\text{кас.}})=0$. Таким образом, $grad\, f(x^0)\perp e_{\text{кас.}}$. В силу произвольности точки $x\in P_c$ отсюда следует, что $grad\, f(x^0)\perp P_c$. Это завершает доказательство леммы.

Частные производные и дифференциалы высших порядков.

Если у функции $f(x)=f(x_1,\ldots,x_n)$ частная производная $\frac{\partial f}{\partial x_k}$ определена в некоторой области $D\subset\mathbb{R}^n$, то она также является функцией n переменных. Может случиться , что эта функция имеет частную производную по переменной x_i в некоторой внутренней точке $x^0=(x_1^0,\ldots,x_n^0)$ области D. Тогда эту производную $\frac{\partial}{\partial x_i}(\frac{\partial f}{\partial x_k})(x^0)=\frac{\partial^2 f}{\partial x_i\partial x_k}(x^0)$ называют *второй частной производной функции f сначала по переменной x_k, а затем по переменной x_i, в точке* $x^0=(x_1^0,\ldots,x_n^0)$ (то есть сначала производится дифференцирование по x_k, а затем по x_i). Если $x_k\neq x_i$, то частная производная второго порядка называется *смешанной.*

Далее, применяя такое же рассуждение ко второй частной производной, можно определить понятие третьей частной производной, и так деле. Основываясь на этом описании понятия второй частной производной, мы можем ввести следующее общее индуктивное определение:

Определение 3. Если у функции $f(x)=f(x_1,\ldots,x_n)$ определена частная производная $(n-1)$-го порядка $\frac{\partial^{n-1}f}{\partial x_{i_{n-1}}\ldots\partial x_{i_1}}(x)$ в некоторой области $D\subset\mathbb{R}^n$, и у неё существует частная производная по переменной x_{i_n} в точке $x^0=(x_1^0,\ldots,x_n^0)\in D$, то эта производная $\frac{\partial}{\partial x_{i_n}}(\frac{\partial^{n-1}f}{\partial x_{i_{n-1}}\ldots\partial x_{i_1}})(x^0)$ называется *частной производной n-го порядка функции $f(x)$ по переменным $x_{i_1},\ldots,x_{i_n}$* в точке x^0 и обозначается $\frac{\partial^n f}{\partial x_{i_n}\partial x_{i_{n-1}}\ldots\partial x_{i_1}}(x^0)$.

Аналогично частным производным первого порядка, существуют другие обозначения и для частных производных высших порядков. Производные $\frac{\partial^2 f}{\partial x_i\partial x_k}(x^0)$, $\frac{\partial^n f}{\partial x_{i_n}\partial x_{i_{n-1}}\ldots\partial x_{i_1}}(x^0)$ можно обозначать также $f''_{x_kx_i}$, $f^{(n)}_{x_{i_1}\ldots x_{i_n}}$ соответственно. Если среди переменных $x_{i_1},\ldots,x_{i_n}$ не все совпадают, то такая частная производная n- го порядка называется *смешанной.*

Пример 1. $f(x,y)=arctg\, xy$; $\quad\frac{\partial f}{\partial x}=\frac{y}{1+x^2y^2}$; $\frac{\partial f}{\partial y}=\frac{x}{1+x^2y^2}$;

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{x}{1+x^2y^2}\right) = \frac{1-x^2y^2}{\left(1+x^2y^2\right)^2} = \frac{\partial^2 f}{\partial y \partial x}$$ (Проверьте самостоятельно!)

Пример 2. Пусть $g(x,y) = \begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2}, & x^2+y^2>0 \\ 0, & x=y=0 \end{cases}$. Тогда

$$f'_x(x;y) = \begin{cases} y\dfrac{4x^2y^2+x^4-y^4}{\left(x^2+y^2\right)^2}, & x^2+y^2>0 \\ 0, & x=y=0 \end{cases}; \quad f'_y(x;y) = \begin{cases} x\dfrac{-4x^2y^2+x^4-y^4}{\left(x^2+y^2\right)^2}, & x^2+y^2>0 \\ 0, & x=y=0 \end{cases};$$

Следовательно, для смешанных частных производных второго порядка получаем:

$$f''_{xy}(0;0) = \lim_{y\to 0}\frac{f'_x(0;y)-f'_x(0;0)}{y} = -1; \qquad f''_{yx}(0;0) = \lim_{x\to 0}\frac{f'_y(x;0)-f'_y(0;0)}{x} = 1.$$

(Проверьте самостоятельно все вычисления!).

Рассмотрим теперь понятие n раз дифференцируемой функции, которое также вводится индуктивно.

Определение 4. Функция $f(x) = f(x_1,...,x_n)$ называется *дважды дифференцируемой в точке* $x^0 = (x_1^0,...,x_n^0)$, если она дифференцируема в некоторой окрестности этой точки, и все её частные производные дифференцируемы в точке x^0. Аналогично, если функция $f(x)$ $(n-1)$ раз $(n>1)$ дифференцируема в некоторой окрестности точки x^0, и все её частные производные $(n-1)$-го порядка дифференцируемы в точке x^0, то $f(x)$ называется n *раз дифференцируемой в точке* x^0.

Из определения 4 вытекает следующее достаточное условие для того, чтобы функция была n раз дифференцируема в данной точке.

Утверждение 1. Для того, чтобы функция $f(x)$ была n раз дифференцируема в данной точке $x^0 = (x_1^0,...,x_n^0)$, достаточно, чтобы она $(n-1)$ раз была дифференцируема в некоторой окрестности точки x^0, и все ее частные производные n-го порядка были непрерывны в самой точке x^0.

Доказательство. При $n=1$ получаем достаточное условие дифференцируемости функции в точке. Пусть $n>1$. Рассмотрим любую из производных $(n-1)$- го порядка функции $f(x)$. По условию все ее частные производные первого порядка непрерывны в точке x^0 (поскольку они являются производными n-го порядка самой функции $f(x)$). Согласно достаточному условию дифференцируемости, получаем, что любая производная порядка $(n-1)$ функции $f(x)$ дифференцируема в точке x^0. Кроме того, сама функция по условию $(n-1)$ раз дифференцируема в окрестности точки x^0. Следовательно, она n раз дифференцируема в точке x^0, что и требовалось доказать.

Рассмотрим теперь вопрос о том, зависят ли смешанные частные производные функции $f(x)$ по одному и тому же набору переменных от того, в каком порядке производится последовательное дифференцирование, и при каких условиях они совпадают. В приведённых выше примерах, как мы видели, в одном случае (пример 1) смешанные производные совпадают, а в другом (пример 2) различны. Сформулируем и докажем две теоремы о достаточных условиях равенства смешанных частных производных второго порядка.

Теорема 1 (Юнг). Если функция $f(x,y)$ дважды дифференцируема в точке (x_0,y_0), то $\dfrac{\partial^2 f}{\partial x \partial y}(x_0,y_0)=\dfrac{\partial^2 f}{\partial y \partial x}(x_0,y_0)$.

Доказательство. Условие теоремы означает, что частные производные функции $f(x,y)$ определены в некоторой окрестности, и дифференцируемы в самой этой точке. Пусть приращение h достаточно мало, так что точка (x_0+h, y_0+h) принадлежит указанной окрестности. Рассмотрим выражение

$$\Phi=\Phi(x_0,y_0;h)=f(x_0+h;y_0+h)-f(x_0+h,y_0)-f(x_0,y_0+h)+f(x_0,y_0),$$

которое можно представить следующими двумя способами. Во-первых, так:

$$\Phi=[f(x_0+h;y_0+h)-f(x_0+h,y_0)]-[f(x_0,y_0+h)-f(x_0,y_0)]=\varphi(x_0+h)-\varphi(x_0), \qquad (4)$$

где $\varphi(x)=f(x,y_0+h)-f(x,y_0)$. И во-вторых, так:

$$\Phi=[f(x_0+h;y_0+h)-f(x_0,y_0+h)]-[f(x_0+h,y_0)-f(x_0,y_0)]=\psi(y_0+h)-\psi(y_0), \qquad (5)$$

где $\psi(y)=f(x_0+h,y)-f(x_0,y)$.

Применяя теорему Лагранжа к дифференцируемой функции $\varphi(x)$ на интервале $(x_0;x_0+h)$, из (4) получаем:

$$\Phi=\varphi'_x(x_0+\theta h)\cdot h=[f'_x(x_0+\theta h,y_0+h)-f'_x(x_0+\theta h,y_0)]\cdot h=$$
$$=[f'_x(x_0+\theta h,y_0+h)-f'_x(x_0,y_0)]\cdot h-[f'_x(x_0+\theta h,y_0)-f'_x(x_0,y_0)]\cdot h,$$

где $\theta\in(0;1)$.

Далее, в последнем выражении в квадратных скобках стоят приращения дифференцируемой в точке (x_0,y_0) функции f'_x, которые можно представить следующим образом:

$$[f'_x(x_0+\theta h,y_0+h)-f'_x(x_0,y_0)]=f''_{xx}(x_0,y_0)\theta h+f''_{xy}(x_0,y_0)h+\alpha_1\theta h+\alpha_2 h;$$
$$[f'_x(x_0+\theta h,y_0)-f'_x(x_0,y_0)]=f''_{xx}(x_0,y_0)\theta h+\alpha_3 h,$$

где $\alpha_1,\alpha_2,\alpha_3$ - бесконечно малые при $h\to 0$. Подставляя полученные выражения в формулу для Ф, получаем:

$$\Phi=f''_{xx}(x_0,y_0)\theta h^2+f''_{xy}(x_0,y_0)h^2+(\alpha_1\theta h+\alpha_2 h)h-f''_{xx}(x_0,y_0)\theta h^2-\alpha_3 h^2.$$

Таким образом,

$$\Phi=f''_{xy}(x_0,y_0)h^2+\bar{o}(h^2), \quad h\to 0.$$

Совершенно аналогично, используя представление (5) , получаем:

$$\Phi=\psi'_y(y_0+\tilde{\theta}h)\cdot h=[f'_y(x_0+h,y_0+\tilde{\theta}h)-f'_y(x_0,y_0+\tilde{\theta}h)]\cdot h=$$
$$=[f'_y(x_0+h,y_0+\tilde{\theta}h)-f'_y(x_0,y_0)]\cdot h-[f'_y(x_0,y_0+\tilde{\theta}h)-f'_y(x_0,y_0)]\cdot h=$$
$$f''_{yx}(x_0,y_0)h^2+f''_{yy}(x_0,y_0)\tilde{\theta}h^2+(\beta_1 h+\beta_2\tilde{\theta}h)h-f''_{yy}(x_0,y_0)\tilde{\theta}h^2-\beta_3 h^2.$$

Поэтому, уничтожая слагаемые с противоположными знаками, имеем:

$$\Phi=f''_{yx}(x_0,y_0)h^2+\bar{o}(h^2), \quad h\to 0.$$

Поделив на h^2 и приравнивая правые части данных соотношений, получим:

$$f''_{xy}(x_0,y_0)+\bar{o}(1)=f''_{yx}(x_0,y_0)+\bar{o}(1),$$

откуда и следует искомое равенство: $f''_{yx}(x_0,y_0)=f''_{xy}(x_0,y_0)$, так как разность $f''_{yx}(x_0,y_0)-f''_{xy}(x_0,y_0)$ есть бесконечно малая величина. Теорема доказана.

Теорема 2 (Шварц). Если у функции $f(x,y)$ в некоторой окрестности точки (x_0,y_0) существуют частные производные $f'_x, f'_y, f''_{yx}, f''_{xy}$, причём производные f''_{yx}, f''_{xy} непрерывны в точке (x_0,y_0), то имеет место равенство: $f''_{yx}(x_0,y_0)=f''_{xy}(x_0,y_0)$.

Доказательство. Используем выражение Ф и некоторые выкладки из доказательства теоремы 1. Из условия теоремы о существовании частной производной f''_{xy}, применяя теорему Лагранжа, получаем:

$$\Phi = [\, f'_x(x_0+\theta h, y_0+h) - f'_x(x_0+\theta h, y_0)] \cdot h = f''_{xy}(x_0+\theta h, y_0+\theta_1 h)\cdot h^2,$$

где $\theta_1 \in (0;1)$.

С другой стороны, поменяв ролями x и y и снова применяя теорему Лагранжа, имеем:

$$\Phi = [\, f'_y(x_0+h, y_0+\tilde{\theta} h) - f'_y(x_0, y_0+\tilde{\theta} h)] \cdot h = f''_{yx}(x_0+\theta_2 h, y_0+\tilde{\theta} h)\cdot h^2,$$

где $\theta_2 \in (0;1)$.

Поделив на h^2 и приравнивая правые части полученных выражений, приходим к равенству:

$$f''_{yx}(x_0+\theta_2 h, y_0+\tilde{\theta} h) = f''_{xy}(x_0+\theta h, y_0+\theta_1 h).$$

При переходе к пределу при $h \to 0$, в силу условия непрерывности этих производных в точке (x_0, y_0), получаем искомое равенство: $f''_{yx}(x_0, y_0) = f''_{xy}(x_0, y_0)$. Теорема полностью доказана.

Из теоремы 1 выведем достаточное условие равенства смешанных производных высших порядков.

Теорема 3. Пусть функция $f(x)$ m раз ($m > 2$) дифференцируема в точке x^0. Тогда её частные производные m- го порядка не зависят от порядка последовательного выполнения операций дифференцирования.

Доказательство. Достаточно показать, что производная $\dfrac{\partial^m f}{\partial x_{i_m}\ldots\partial x_{i_{k+1}}\partial x_{i_k}\ldots\partial x_{i_1}}(x^0)$ не зависит от перестановки двух соседних операций дифференцирования, то есть доказать равенство:

$$\frac{\partial^m f}{\partial x_{i_m}\ldots\partial x_{i_{k+1}}\partial x_{i_k}\ldots\partial x_{i_1}}(x^0) = \frac{\partial^m f}{\partial x_{i_m}\ldots\partial x_{i_k}\partial x_{i_{k+1}}\ldots\partial x_{i_1}}(x^0).$$

С этой целью рассмотрим функцию $F(x) = \dfrac{\partial^{k-1} f}{\partial x_{i_{k-1}}\ldots\partial x_{i_1}}(x), 1 < k < m$. Из условия теоремы следует, что

1) при $1 < k < m-1$ функция $F(x)$ дважды дифференцируема в некоторой окрестности точки x^0;

2) при $k = m-1$ функция $F(x)$ дважды дифференцируема в точке x^0.

Но тогда, по теореме 1, её смешанные частные производные $\dfrac{\partial^2 F}{\partial x_{i_k}\partial x_{i_{k+1}}}, \dfrac{\partial^2 F}{\partial x_{i_{k+1}}\partial x_{i_k}}$ при $1 < k < m-1$тождественно совпадают в некоторой окрестности точки x^0, а при $k = m-1$ они совпадают в точке x^0. Это означает, что:

1) $\dfrac{\partial^{k+1} f}{\partial x_{i_{k+1}}\partial x_{i_k}\ldots\partial x_{i_1}} = \dfrac{\partial^{k+1} f}{\partial x_{i_k}\partial x_{i_{k+1}}\ldots\partial x_{i_1}}$ при $1 < k < m-1$ в некоторой окрестности точки x^0, откуда при дальнейшем дифференцировании по остальным переменным $x_{i_{k+2}},\ldots,x_{i_m}$ получается нужное равенство;

2) при $k = m-1$ равенство $\dfrac{\partial^{k+1} f}{\partial x_{i_{k+1}}\partial x_{i_k}\ldots\partial x_{i_1}} = \dfrac{\partial^{k+1} f}{\partial x_{i_k}\partial x_{i_{k+1}}\ldots\partial x_{i_1}}$ в точке x^0 совпадает с искомым равенством.

Теорема доказана.

Из теоремы 3 вытекает следующее простое утверждение.

Следствие. Если функция $f(x)$ m раз ($m \geq 2$) дифференцируема в точке x^0, то её частные производные m-го порядка можно записывать в следующей форме: $\frac{\partial^m f}{(\partial x_n)^{\alpha_n} \ldots (\partial x_1)^{\alpha_1}}$, где $0 \leq \alpha_j \leq m,$ $\alpha_1 + \ldots + \alpha_n = m.$