Здравствуйте!

Лекция №2

Частные производные

Пусть имеется функция n переменных $f(x_1, x_2, ..., x_n)$. Изменим значение i-й переменной с x_i на $x_i+\Delta x_i$. Величина

$$\Delta_i f = f(x_1, x_2, ..., x_i + \Delta x_i, ..., x_n) - f(x_1, x_2, ..., x_i, ..., x_n)$$

называется **частным приращением** функции $f(x)$ по i-й переменной.

Величина

$$\lim_{\Delta x_i \to 0} \frac{\Delta_i f}{\Delta x_i} = \frac{\partial f(x)}{\partial x_i} = f'_{x_i}(x_1, x_2, ..., x_n)$$

называется **частной производной** от функции $f(x)$ по переменной x_i и обозначается либо символом $\frac{\partial f(x)}{\partial x_i}$, либо символом $f'_{x_i}(x)$.

Отметим главное: при вычислении частной производной по какой-либо переменной **все остальные переменные выступают как константы.**

Вектор с компонентами

$$\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right) = \operatorname{grad} f$$

называется **градиентом** функции $f(x)$ и обозначается символом $\operatorname{grad} f$.

Полное приращение и дифференциал функции

Изменим теперь **все** переменные, заменяя x_1 на $x_1+\Delta x_1$, x_2 на $x_2+\Delta x_2$, …, x_n на $x_n+\Delta x_n$. Величина

$$\Delta f = f\,(x_1+\Delta x_1\,, x_2+\Delta x_2\,, ..., x_n+\Delta x_n\,) - f\,(x_1\,, x_2\,, ..., x_n\,)$$

называется **полным приращением** функции $f\,(x)$.

Определение 1. Если Δf представима в виде

$$\Delta f = \sum_{i=1}^{n} A_i \Delta x_i + o(\rho), \qquad \rho = \sqrt{\sum_{i=1}^{n} (\Delta x_i\,)^2}\,,$$

то функция $f\,(x)$ называется **дифференцируемой** в точке x, а комбинация

$$\sum_{i=1}^{n} A_i \Delta x_i = df\,(x)$$

называется **дифференциалом (**или, точнее, полным дифференциалом) функции $f\,(x)$ и обозначается $d\,f\,(x)$.

Определение 2. Если Δf представима в виде

$$\Delta f = \sum_{i=1}^{n} (A_i + \alpha_i) \Delta x_i,$$

где $\alpha_i \to 0$ при $\rho \to 0$, то функция $f(x)$ называется **дифференцируемой** в точке x.

Можно доказать, что эти два определения эквиваленты.

Теорема 1. Если $f(x)$ дифференцируема в точке x, то у нее в этой точке существуют все частные производные и $A_i = \partial f(x) / \partial x_i$

Доказательство. Дадим приращение Δx_i только одной переменной x_i, а остальные переменные оставим без изменения. Тогда

$$\Delta_i f = (A_i + \alpha_i)\Delta x_i,$$

где $\alpha_i \to 0$ при $\Delta x_i \to 0$. Деля на Δx_i

$$\frac{\Delta_i f}{\Delta x_i} = A_i + \alpha_i$$

и устремляя Δx_i к нулю, получим

$$\lim_{\Delta x_i \to 0} \frac{\Delta_i f}{\Delta x_i} = \frac{\partial f}{\partial x_i} = A_i,$$

что и требовалось доказать.

Если взять $f(x) = x_i$, то $\dfrac{\partial f}{\partial x_i} = 1$, а при $j \neq i$ $\dfrac{\partial f}{\partial x_j} = 0$. Тогда из общего выражения для дифференциала получим, что $dx_i = \Delta x_i$. Окончательно, выражение для дифференциала функции $f(x)$ приобретает вид

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i .$$

Если ввести вектор $dx = (dx_1, dx_2, ..., dx_n)$, то df можно представить в виде

$$df = (\mathrm{grad}\ f, dx).$$

Теорема 2. Если частные производные $\partial f / \partial x_i$ существуют не только в точке $x=(x_1, x_2, x_3, ..., x_n)$, но и в некоторой ее окрестности и **непрерывны** в точке x, то $f(x)$ дифференцируема в точке x.

Доказательство. Для простоты, окажем эту формулу только для функции двух переменных $f(x, y)$.

Итак, пусть дана $f(x, y)$. Тогда ее приращение может быть записано в виде

$$\Delta f(x, y) = f(x+\Delta x, y+\Delta y) - f(x, y) =$$

$$= [f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y)] + [f(x, y+\Delta y) - f(x, y)].$$

Заметим, что во второй скобке первые аргументы у $f(x, ...)$ одинаковы. Поэтому, применяя формулу Лагранжа, получим

$$f(x, y+\Delta y) - f(x, y) = f'_y(x, y+\theta\Delta y)\Delta y.$$

В силу непрерывности производной, можно записать

$$f'_y(x, y+\theta\Delta y) = f'_y(x, y) + \beta,$$

где $\beta \to 0$ при $\Delta y \to 0$. Поэтому окончательно

$$f(x, y+\Delta y) - f(x, y) = f'_y(x, y)\Delta y + \beta\Delta y.$$

Аналогично, в первой скобке у $f(\ldots, y+\Delta y)$ одинаковым является второй аргумент. Поэтому, по той же формуле Лагранжа

$$f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y) = f'_x(x+\theta\Delta x, y+\Delta y)\Delta x.$$

В силу непрерывности производной

$$f'_x(x+\theta\Delta x, y+\Delta y) = f'_x(x, y) + \alpha,$$

где $\alpha \to 0$ при $(\Delta x, \Delta y) \to (0,0)$. Поэтому

$$f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y) = f'_x(x, y)\Delta x + \alpha\Delta x.$$

Объединяя все вместе, получим, что

$$\Delta f(x, y) = \left(f'_x(x, y) + \alpha\right)\Delta x + \left(f'_y(x, y) + \beta\right)\Delta y,$$

где α и β стремятся к нулю при $(\Delta x, \Delta y) \to (0,0)$. Это и доказывает нашу теорему (см. определение 2).

Производные от сложной функции

Пусть $z = f(x_1, x_2, ..., x_n)$, а аргументы этой функции x_i сами являются функциями переменной $t = (t_1, t_2, ..., t_m)$, то есть

$$z = f(x_1, x_2, ..., x_n), \quad x_i = \varphi_i(t_1, t_2, ..., t_m).$$

Мы имеем, таким образом, дело со сложной функцией

$$z = f(\varphi_1(t), \varphi_2(t), ..., \varphi_n(t)).$$

Наша задача – научиться вычислять частные производны от z по t_k, то есть $\partial z / \partial t_k$.

Будем считать, что все частные производные $\partial f / \partial x_i$ и $\partial \varphi_i / \partial t_k$, $i = \overline{1, n}$ существуют и непрерывны. Отсюда будет следовать, что и $f(x)$ и все $\varphi_i(t)$ дифференцируемы.

Изменим только одну из компонент из переменной t, скажем, t_k, сделав ей приращение Δt_k. Но тогда уже **все** x_i изменятся на величины

$$\Delta x_i = \varphi_i(t_1, t_2, ..., t_k + \Delta t_k, ..., t_m) - \varphi_i(t_1, t_2, ..., t_k, ..., t_m).$$

В силу дифференцируемости $\varphi_i(t)$, можно записать

$$\Delta x_i = \left(\frac{\partial \varphi_i}{\partial t_k} + \alpha_i \right) \Delta t_k,$$

где все $\alpha_i \to 0$ при $\Delta t_k \to 0$. Заметим, что поэтому и все $\Delta x_i \to 0$ при $\Delta t_k \to 0$.

Но раз **все** x_i изменились на величину Δx_i, то z изменится на величину

$$\Delta z = f(x_1 + \Delta x_1, x_2 + \Delta x_2, \ldots, x_n + \Delta x_n) - f(x_1, x_2, \ldots, x_n),$$

и, в силу дифференцируемости $f(x)$, можно записать

$$\Delta z = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} + \beta_i \right) \Delta x_i,$$

где $\beta_i \to 0$ когда все $\Delta x_i \to 0$. Но так как при $\Delta t_k \to 0$ как раз все $\Delta x_i \to 0$, то можно сказать, что все $\beta_i \to 0$ когда $\Delta t_k \to 0$.

Подставляя сюда выражение для Δx_i, получим

$$\Delta z=\sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}+\beta_i\right)\left(\frac{\partial \varphi_i}{\partial t_k}+\alpha_i\right)\Delta t_k=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\frac{\partial \varphi_i}{\partial t_k}\Delta t_k+$$

$$+\sum_{i=1}^{n}\left(\beta_i\frac{\partial \varphi_i}{\partial t_k}+\alpha_i\frac{\partial f}{\partial x_i}+\beta_i\alpha_i\right)\Delta t_k.$$

Но так как при $\Delta t_k\to 0$ все α_i и β_i стремятся к нулю, то и величина

$$\gamma=\sum_{i=1}^{n}\left(\beta_i\frac{\partial \varphi_i}{\partial t_k}+\alpha_i\frac{\partial f}{\partial x_i}+\beta_i\alpha_i\right)$$

стремится к нулю при $\Delta t_k\to 0$.

Окончательно имеем

$$\Delta z=\Delta t_k\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\frac{\partial \varphi_i}{\partial t_k}+\gamma\Delta t_k,$$

откуда получаем

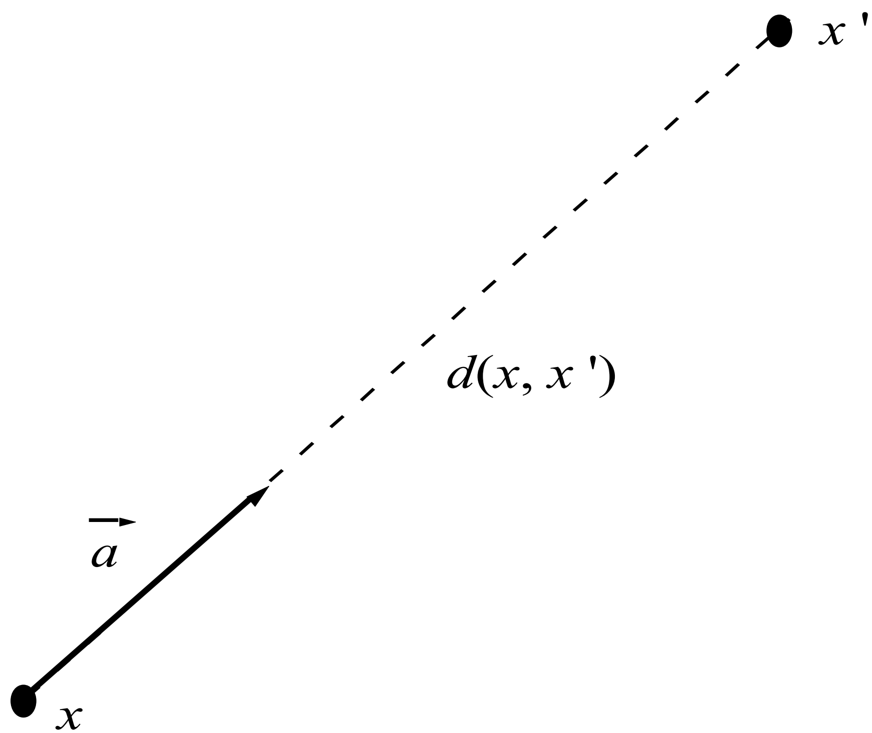
$$\frac{\partial z}{\partial t_k} = \lim_{\Delta t_k \to 0} \frac{\Delta z}{\Delta t_k} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial \varphi_i}{\partial t_k},$$

что и даёт формулу для вычисления производной от сложной функции от многих переменных

$$\frac{\partial z}{\partial t_k} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial \varphi_i}{\partial t_k} = \frac{\partial f}{\partial x_1} \cdot \frac{\partial \varphi_1}{\partial t_k} + \frac{\partial f}{\partial x_2} \cdot \frac{\partial \varphi_2}{\partial t_k} + \ldots + \frac{\partial f}{\partial x_n} \cdot \frac{\partial \varphi_n}{\partial t_k}.$$

Производная по направлению

Пусть задана функция $f(x)$, зависящая от n-мерной переменной $x=(x_1, x_2, ..., x_n)$ и пусть в нашем n-мерном пространстве задан вектор $\vec{a}=(a_1, a_2, ..., a_n)$ единичной длины, то есть $\sum_{i=1}^{n} a_i^2 = 1$.



Представим себе, что из точки x, двигаясь по вектору $\vec{a}$ (или в противоположном направлении) мы перешли в точку x'. При этом мы сместились на расстояние $d(x, x')$ (см. рис.).

Тогда **производной** от $f(x)$ по **направлению вектора** $\vec{a}$ называется величина

$$\lim_{x' \to x} \frac{f(x') - f(x)}{\pm d(x, x')} = \frac{\partial f}{\partial \vec{a}}.$$

Знак «+» перед $d(x, x')$ берется в том случае если мы двигались **по** направлению вектора $\vec{a}$, знак «–» – если мы двигались **против** вектора $\vec{a}$.

Выведем формулу для $\partial f / \partial \vec{a}$. Так как вектор $x' - x$ коллинеарен вектору $\vec{a}$, то их компоненты пропорциональны, то есть

$$\frac{x_1' - x_1}{a_1} = \frac{x_2' - x_2}{a_2} = \ldots = \frac{x_n' - x_n}{a_n} = t.$$

Обозначая это общее отношение через t, получим

$$x_i' = x_i + a_i t.$$

Находя $\pm d(x, x')$, получим

$$\pm d(x, x') = \pm \sqrt{\sum_{i=1}^{n} (x_i' - x_i)^2} = \pm \sqrt{t^2 \sum_{i=1}^{n} a_i^2} = t,$$

так как $\sum_{i=1}^{n} a_i^2 = 1$. Правильный знак перед $\pm d(x, x')$ учтен тем, что $\pm\sqrt{t^2}$ записан как t.

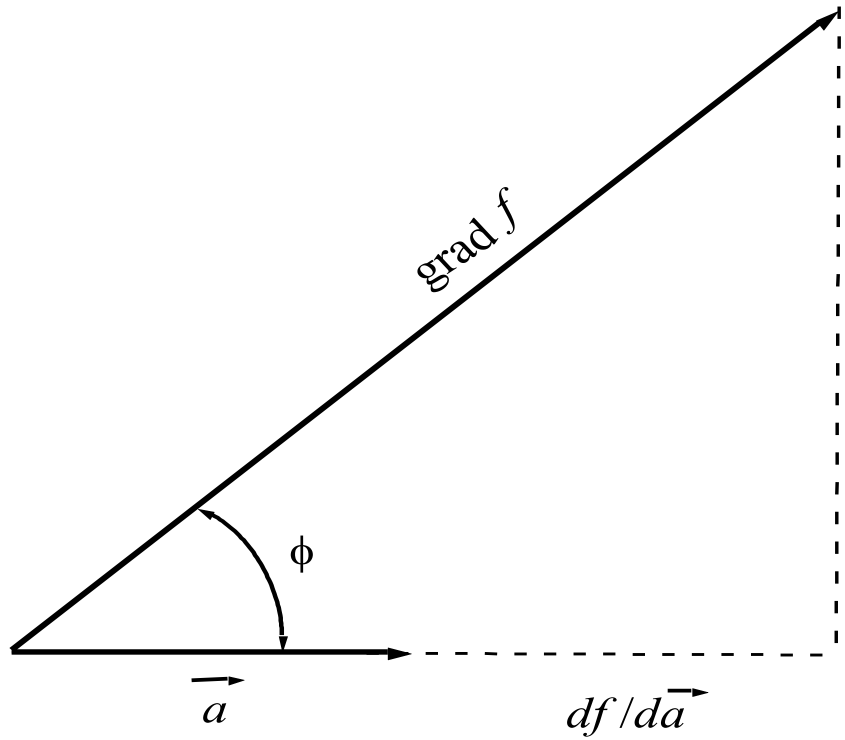
Тогда

$$\frac{\partial f}{\partial \vec{a}} = \lim_{t\to 0}\frac{f(x_1 + a_1t, x_2 + a_2t, ..., x_n + a_nt) - f(x_1, x_2, ..., x_n)}{t} =$$

$$= \frac{d}{dt} f(x_1 + a_1t, x_2 + a_2t, ..., x_n + a_nt)\Big|_{t=0}.$$

Используя формулу для производной от сложной функции и учитывая, что $d(x_i + a_it)/dt = a_i$, получим

$$\frac{\partial f}{\partial \vec{a}} = \sum_{i=1}^{n}\frac{\partial f(x_1 + a_1t, x_2 + a_2t, ..., x_n + a_nt)}{\partial x_i}\Bigg|_{t=0} a_i = \sum_{i=1}^{n}\frac{\partial f(x)}{\partial x_i}a_i$$



Вспоминая выражение для градиента, можно написать, что

$$\frac{\partial f}{\partial \vec{a}} = (\mathrm{grad}\ f, \vec{a}).$$

Таким образом, производная от функции $f(x)$ по какому-то направлению равна проекции градиента на это направление

Вспоминая, что $\|\vec{a}\| = 1$ и обозначая через φ угол между векторами $\vec{a}$ и grad f, можно записать, что

$$\frac{\partial f}{\partial \vec{a}} = \|\operatorname{grad} f\| \cos \varphi.$$

Так как $-1 \le \cos \varphi \le 1$, то

$\max \dfrac{\partial f}{\partial \vec{a}} = \|\operatorname{grad} f\|$ (получается при φ = 0),

$\min \dfrac{\partial f}{\partial \vec{a}} = -\|\operatorname{grad} f\|$ (получается при φ = π).

Глядя на эти формулы можно сказать следующее:
вектор grad f указывает нам, в каком направлении функция возрастает быстрее всего: функция быстрее всего возрастает при движении по направлению вектора градиента и быстрее всего убывает при движении против направления градиента.

Заметим, в заключение, что если вектор $\vec{a}$ имеет произвольную длину, то

$$\frac{\partial f}{\partial \vec{a}} = \frac{(\operatorname{grad} f, \vec{a})}{\|\vec{a}\|},$$

так как вектор $\vec{a}/\|\vec{a}\|$ имеет единичную длину и направлен в ту же сторону, что и вектор $\vec{a}$.