Лекция 22 (05.05.20)
Дифференцируемость функции в точке.
Достаточное условие дифференцируемости.
Геометрический смысл дифференцируемости

Линейные отображения

Определение 1. Отображение (функция) f из линейного пространства L_1 в линейное пространство L_2 называется **линейным**, если из существования $f(a)$ следует, что для любого $\alpha \in \mathbb{R}$ (или $\alpha \in \mathbb{C}$, если L_1 и L_2 — линейные пространства над полем комплексных чисел $\mathbb{C}$) существует $f(\alpha a) = \alpha f(a)$ и из существования $f(a)$ и $f(b)$ следует существование $f(a+b) = f(a) + f(b)$.

Теорема 1. *Определенная на $\mathbb{R}^n$ действительнозначная функция φ линейна тогда и только тогда, когда существуют $\alpha_k \in \mathbb{R}$, $k = 1, \ldots, n$, что для любого $\vec{x} \in \mathbb{R}^n$ функция $\varphi(\vec{x}) = \varphi(x_1, \ldots, x_n) = \sum\limits_{k=1}^{n} \alpha_k x_k$. Если φ линейна, то она непрерывна и существует такое $C \geqslant 0$, что для любого $\vec{x} \in \mathbb{R}^n$ верна оценка $|\varphi(\vec{x})| \leqslant C\|\vec{x}\|$.*

▼ Если $\varphi(\vec{x}) = \varphi(x_1, \ldots, x_n) = \sum\limits_{k=1}^{n} \alpha_k x_k$, то, очевидно, φ — линейная функция. Если φ — линейная функция, то положим $\alpha_k = \varphi(\vec{e}_k)$ (где $\vec{e}_k$ — вектор, k-ая координата которого 1, а остальные 0), $k = 1, \ldots, n$. Тогда

$$\varphi(\vec{x}) = \varphi\left(\sum_{k=1}^{n} x_k \vec{e}_k\right) = \sum_{k=1}^{n} x_k \varphi(\vec{e}_k) = \sum_{k=1}^{n} \alpha_k x_k.$$

Ясно, что φ непрерывна и по неравенству Коши–Буняковского

$$|\varphi(\vec{x})| = \left|\sum_{k=1}^{n} \alpha_k x_k\right| \leqslant \sqrt{\sum_{k=1}^{n} \alpha_k^2} \cdot \sqrt{\sum_{k=1}^{n} x_k^2} = \sqrt{\sum_{k=1}^{n} \alpha_k^2}\|\vec{x}\|$$

и, значит, заключительное утверждение теоремы верно с $C = \sqrt{\sum\limits_{k=1}^{n} \alpha_k^2}$. ▲Ж

Дифференцируемость функции в точке

Определение 2. Функция f из пространства $\mathbb{R}^n$ в $\mathbb{R}$ называется **дифференцируемой в точке** $x_0 \in \mathbb{R}^n$, если f определена в окрестности этой точки и ее приращение $\Delta f = f(x_0 + \Delta x) - f(x_0) = f(x) - f(x_0)$ представляется в виде

$$\Delta f = f(x_0 + \Delta x) - f(x_0) = L(\Delta x) + o(\|\Delta x\|),$$

где L — линейная функция, отображающая $\mathbb{R}^n$ в $\mathbb{R}$, $o(\|\Delta x\|) = o(1) \cdot \|\Delta x\|$, $o(1) \to 0$ при $\Delta x \to 0$ (т.е. при $\|\Delta x\| \to 0$), а приращение аргумента Δx таково, что f определена в точке $x_0 + \Delta x$ (что выполняется при достаточно малых Δx).

Определение 3. Линейная функция L называется (**полным**) **дифференциалом** (или **производным отображением** или **полной производной**) функции f в точке x_0 и обозначается $\boldsymbol{d}f(\Delta x)\big|_{x_0} = L(\Delta x)$. Будут также использоваться обозначения $\boldsymbol{d}f(x_0, \Delta x)$, $\boldsymbol{d}f(\Delta x)$ (если понятно, в какой точке происходит дифференцирование) и $\boldsymbol{d}f(x_0)$.

Теорема 2. *Если функция f из пространства $\mathbb{R}^n$ в $\mathbb{R}$ дифференцируема в точке, то она в ней непрерывна.*

▼ Возьмем любую последовательность $\Delta x_k \to 0$. Из непрерывности линейной функции L следует, что $L(\Delta x_k) \to 0$. Функция $f(x_0 + \Delta x_k)$ определена для всех достаточно малых Δx_k, т.е. начиная с некоторого номера, и приращение функции

$$\Delta f = f(x_0 + \Delta x_k) - f(x_0) = L(\Delta x_k) + o(\|\Delta x_k\|) \xrightarrow[k\to\infty]{} 0,$$

значит (по определению предела по Гейне), $\lim\limits_{\Delta x\to 0} f(x_0+\Delta x) = f(x_0)$ и, следовательно, f непрерывна в точке x_0. ▲

Определение 4. Частной производной (первого порядка) функции f (из $\mathbb{R}^n$ в $\mathbb{R}$) по переменной x_k в точке $\vec{x}^0$ называется предел

$$\lim_{t\to 0} \frac{f(\vec{x}^0 + t\vec{e}_k) - f(\vec{x}^0)}{t} = \lim_{t\to 0} \frac{f(x_1^0, \ldots, x_{k-1}^0, x_k^0 + t, x_{k+1}^0, \ldots, x_n^0) - f(x_1^0, \ldots, x_n^0)}{t}.$$

Обозначение. Частную производную функции f по переменной x_k в точке $\vec{x}^0$ обозначают

$$f'_{x_k}(\vec{x}^0) \text{ или } f'_{x_k}\big|_{\vec{x}=\vec{x}^0} \text{ или } \frac{\partial f(\vec{x}^0)}{\partial x_k} \text{ или } \frac{\partial f(\vec{x})}{\partial x_k}\Big|_{\vec{x}=\vec{x}^0}.$$

Фактически это обычная производная в ситуации, когда функция f рассматривается как функция переменной x_k при фиксировании остальных переменных.

Теорема 3. *Если функция f дифференцируема в точке $\vec{x}^0$, то она имеет в этой точке все частные производные $\frac{\partial f(\vec{x}^0)}{\partial x_k}$, $k = 1, \ldots, n$, и ее дифференциал*

$$\boldsymbol{d}f\left(\overrightarrow{\Delta x}\right)\big|_{\vec{x}^0} = L\left(\overrightarrow{\Delta x}\right) = \sum_{k=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_k}\Delta x_k.$$

▼ Дифференциал — линейная функция приращения аргумента, $\boldsymbol{d}f(\overrightarrow{\Delta x}) = L(\overrightarrow{\Delta x}) = \sum\limits_{k=1}^{n} \alpha_k \Delta x_k$ (по теореме 1), значит, $\Delta f = \sum\limits_{k=1}^{n} \alpha_k \Delta x_k + o\left(\|\overrightarrow{\Delta x}\|\right)$. Используя последнее равенство получаем, что существует

$$\lim_{t\to 0} \frac{f(\vec{x}^0 + t\vec{e}_k) - f(\vec{x}^0)}{t} = \lim_{t\to 0} \frac{\alpha_k t + o(|t|)}{t} = \lim_{t\to 0}(\alpha_k + o(1)) = \alpha_k,$$

то есть существует $\frac{\partial f(\vec{x}^0)}{\partial x_k} = \alpha_k$, $k = 1, \ldots, n$, и

$$\boldsymbol{d}f(\vec{x}^0) = \sum_{k=1}^{n} \alpha_k \Delta x_k = \sum_{k=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_k}\Delta x_k. \blacktriangle$$

Замечание 1. Если дифференциал $\boldsymbol{d}f\left(\overrightarrow{\Delta x}\right)\big|_{\vec{x}^0}$, являющийся при фиксированном $\vec{x}^0$ функцией от $\overrightarrow{\Delta x}$, рассматривать как функцию от $\vec{x}$, $\overrightarrow{\Delta x} = \vec{x} - \vec{x}^0$, то его частная производная по переменной x_k в любой точке равна $\frac{\partial f(\vec{x}^0)}{\partial x_k}$.

Отметим, что существование частных производных в точке не гарантирует даже непрерывности, тем более дифференцируемости, функции. Например, функция двух переменных

$$f(x, y) = \operatorname{sign} xy = \begin{cases} 0, & \text{если } xy = 0, \\ 1, & \text{если } xy > 0, \\ -1, & \text{если } xy < 0, \end{cases}$$

имеет в точке $(0,0)$ равные нулю частные производные $\frac{\partial f(0,0)}{\partial x}$ и $\frac{\partial f(0,0)}{\partial y}$ и разрывна в $(0,0)$.

Достаточное условие дифференцируемости

Определение 5. Функция f (из $\mathbb{R}^n$ в $\mathbb{R}$) имеет в точке $\vec{x}^0$ **непрерывную частную производную** $\frac{\partial f}{\partial x_k}$, если эта частная производная существует в окрестности точки $\vec{x}^0$ и непрерывна в ней.

Теорема 4. *Если функция f имеет непрерывные частные производные (первого порядка) в точке $\vec{x}^0$ $\frac{\partial f(\vec{x})}{\partial x_k}$, $k = 1, \ldots, n$, то f дифференцируема в этой точке.*

▼ Проведем индукцию по n. При $n = 1$ утверждение верно (по теореме для одной переменной), причем достаточно просто существования производной в точке). Предположим, что оно верно при $n = m$ и докажем его при $n = m + 1$. Имеем равенство

$$\Delta f = f(\vec{x}) - f(\vec{x}^0) = f(x_1, \ldots, x_m, x_{m+1}) - f(x_1, \ldots, x_m, x_{m+1}^0) + \\ + f(x_1, \ldots, x_m, x_{m+1}^0) - f(x_1^0, \ldots, x_m^0, x_{m+1}^0).$$

По теореме Лагранжа

$$f(x_1, \ldots, x_m, x_{m+1}) - f(x_1, \ldots, x_m, x_{m+1}^0) = \\ = \frac{\partial}{\partial x_{m+1}} f(x_1, \ldots, x_m, x_{m+1}^0 + \theta \Delta x_{m+1}) \Delta x_{m+1},$$

где $0 < \theta < 1$. Так как частная производная $\frac{\partial f}{\partial x_{m+1}}$ непрерывна в точке $\vec{x}^0$, то

$$\frac{\partial}{\partial x_{m+1}} f(x_1, \ldots, x_m, x_{m+1}^0 + \theta \Delta x_{m+1}) = \frac{\partial f(\vec{x}^0)}{\partial x_{m+1}} + o(1),$$

где $o(1) \to 0$ при $\overrightarrow{\Delta x} \to 0$.

Рассматривая функцию f как функцию только первых m переменных при фиксированной $m + 1$-ой переменной, по индукционному предположению имеем:

$$f(x_1^0 + \Delta x_1, \ldots, x_m^0 + \Delta x_m, x_{m+1}^0) - f(x_1^0, \ldots, x_m^0, x_{m+1}^0) = \\ = \sum_{k=1}^{m} \frac{\partial f(\vec{x}^0)}{\partial x_k} \Delta x_k + o\left(\|(\overrightarrow{\Delta x})_m\|\right),$$

где $(\overrightarrow{\Delta x})_m = (\Delta x_1, \ldots, \Delta x_m, 0)$.

В итоге имеем:

$$\Delta f = f(\vec{x}^0 + \overrightarrow{\Delta x}) - f(\vec{x}^0) = \frac{\partial f(\vec{x}^0)}{\partial x_{m+1}} \Delta x_{m+1} + o(1) \cdot \Delta x_{m+1} + \\ + \sum_{k=1}^{m} \frac{\partial f(\vec{x}^0)}{\partial x_k} \Delta x_k + o\left(\|(\overrightarrow{\Delta x})_m\|\right) = \sum_{k=1}^{m+1} \frac{\partial f(\vec{x}^0)}{\partial x_k} \Delta x_k + o\left(\|\overrightarrow{\Delta x}\|\right).\blacktriangle$$

Замечание 1. В теореме можно отказаться от непрерывности частной производной $\frac{\partial f}{\partial x_1}$, достаточно ее существования в точке $\vec{x}^0$. Ведь при $n = 1$ такое утверждение верно, а далее из верности этого предположения для $n = m$ следуя доказательству теоремы получаем его верность для $n = m + 1$. Так как порядок переменных несущественен, то, значит, в формулировке теоремы можно отказаться от требования непрерывности любой одной частной производной $\frac{\partial f}{\partial x_k}$, достаточно ее существования в точке $\vec{x}^0$.

Геометрический смысл дифференцируемости

Рассмотрим поверхность S в пространстве $\mathbb{R}^{n+1}$, являющуюся графиком функции f, $x_{n+1} = f(x_1, \ldots, n)$.

Определение 6. Плоскость (гиперплоскость) Π, проходящая через точку поверхности S $\vec{x}^0 \in \mathbb{R}^{n+1}$, $x^0_{n+1} = f(x^0_1, \ldots, x^0_n)$, и определяемая уравнением

$$x_{n+1} - x^0_{n+1} = \sum_{k=1}^{n} \alpha_k (x_k - x^0_k),$$

называется **касательной плоскостью** к поверхности S в точке $\vec{x}^0 \in \mathbb{R}^{n+1}$, если угол между двумя прямыми, прямой из плоскости, проходящей через точки $\left(x^0_1, \ldots, x^0_n, x^0_{n+1}\right)$ и $\left(x_1, \ldots, x_n, x^0_{n+1} + \sum_{k=1}^{n} \alpha_k (x_k - x^0_k)\right)$, и секущей, проходящей через точки $\left(x^0_1, \ldots, x^0_n, x^0_{n+1}\right)$ и $(x_1, \ldots, x_n, f(x_1, \ldots, x_n))$, стремится к нулю при $(x_1, \ldots, x_n) \to (x^0_1, \ldots, x^0_n)$.

Теорема 5. *Функция f, определенная в окрестности точки $\vec{x}^0 \in \mathbb{R}^n$, дифференцируема в этой точке тогда и только тогда, когда определяемая ею поверхность $S \subset \mathbb{R}^{n+1}$ имеет в точке $(x^0_1, \ldots, x^0_n, f(x^0_1, \ldots, x^0_n))$ касательную плоскость. При этом последняя определяется уравнением*

$$x_{n+1} - f(\vec{x}^0) = \sum_{k=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_k} (x_k - x^0_k).$$

▼ Пусть α — угол между двумя вышеуказанными прямыми, прямой из плоскости и секущей, а β и $\beta + \alpha$ — соответственно углы между ними и прямой, проходящей через точки $(x^0_1, \ldots, x^0_n, x^0_{n+1})$ и $(x_1, \ldots, x_n, x^0_{n+1})$. Угол β не может приближаться сколь угодно близко к $\pm\frac{\pi}{2}$, поскольку

$$|\operatorname{tg}\beta| = \left| \frac{x_{n+1} - x^0_{n+1}}{\|\overrightarrow{\Delta x}\|} \right| = \left| \frac{\sum_{k=1}^{n} \alpha_k \Delta x_k}{\|\overrightarrow{\Delta x}\|} \right|,$$

где $\overrightarrow{\Delta x} = (\Delta x_1, \ldots, \Delta x_n) \in \mathbb{R}^n$, $\Delta x_k = x_k - x^0_k$, $k = 1, \ldots, n$. По теореме предыдущей лекции о линейной функции на n-мерном пространстве существует такая постоянная C, что $|\operatorname{tg}\beta| < C$. А так как

$$|\operatorname{tg}(\alpha + \beta) - \operatorname{tg}\beta| \cdot \left\|\overrightarrow{\Delta x}\right\| = \left| \frac{\sin\alpha}{\cos(\alpha + \beta) \cdot \cos\beta} \right| \cdot \left\|\overrightarrow{\Delta x}\right\| = \left| f(\vec{x}) - f(\vec{x}^0) - \sum_{k=1}^{n} \alpha_k \Delta x_k \right|,$$

где последнее выражение — расстояние между точками $(x_1, \ldots, x_n, f(\vec{x}^0) + \sum\limits_{k=1}^{n} \alpha_k \Delta x_k)$ и $(x_1, \ldots, x_n, f(\vec{x}))$, то угол α, $-\frac{\pi}{2} < \alpha < \frac{\pi}{2}$, стремится к нулю тогда и только тогда, когда указанное расстояние $\left| f(\vec{x}) - f(\vec{x}^0) - \sum\limits_{k=1}^{n} \alpha_k \Delta x_k \right|$ является $o(\|\overrightarrow{\Delta x}\|)$, что эквивалентно дифференцируемости f в точке $\vec{x}^0$. При этом $\sum\limits_{k=1}^{n} \alpha_k \Delta x_k)$ является дифференциалом функции f и, значит, $\alpha_k = \frac{\partial f(\vec{x}^0)}{\partial x_k}$ по предыдущей лекции. ▲

Лекция 23 (08.05.20)
Производная по направлению..
Градиент функции.
Правила дифференцирования

Производная по направлению

Определение 1. Производной по направлению $\vec{w} = (w_1, \ldots, w_n)$, $\|\vec{w}\| = 1$, функции f в точке $\vec{x}^0$ называется

$$\frac{\partial f(\vec{x}^0)}{\partial \vec{w}} = \lim_{t\to+0} \frac{f(\vec{x}^0 + t\vec{w}) - f(\vec{x}^0)}{t}.$$

Отметим, что если существует $\frac{\partial f(\vec{x}^0)}{\partial x_k}$, то производная по направлению $\vec{e}_k$ в точке $\vec{x}^0$ — это $\frac{\partial f(\vec{x}^0)}{\partial x_k}$, а по направлению $-\vec{e}_k$ — это $-\frac{\partial f(\vec{x}^0)}{\partial x_k}$. Производная по направлению $\vec{w}$ функции f в точке $\vec{x}^0$ — это правая производная в точке 0 функции $f(\vec{x}^0 + t\vec{w})$ как функции одной переменной t. Левая производная этой функции отличается от производной по направлению $-\vec{w}$ знаком.

Теорема 1. *Если функция f дифференцируема в точке $\vec{x}^0$, то f имеет в этой точке производную по любому направлению $\vec{w}$ и при этом*

$$\frac{\partial f(\vec{x}^0)}{\partial \vec{w}} = \sum_{k=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_k} w_k.$$

▼ Поскольку f дифференцируема в точке $\vec{x}^0$, то

$$f(\vec{x}^0 + t\vec{w}) - f(\vec{x}^0) = \sum_{k=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_k} \cdot t w_k + o(t\|\vec{w}\|)$$

и, значит,

$$\frac{\partial f(\vec{x}^0)}{\partial \vec{w}} = \lim_{t\to+0} \frac{f(\vec{x}^0 + t\vec{w}) - f(\vec{x}^0)}{t} = \sum_{k=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_k} w_k. \blacktriangle$$

Отметим, что существование производных по любому направлению в точке не гарантирует даже непрерывности, тем более дифференцируемости функции в этой точке.

Например, функция двух переменных

$$f(x, y) = \begin{cases} 1, & \text{если } x^2 + \left(y - \frac{1}{2}\right)^2 = \frac{1}{4},\ x \neq 0 \text{ и } y \neq 0, \\ 0 & \text{в остальных точках} \end{cases}$$

имеет в точке $(0, 0)$ производную по любому направлению равную нулю, но разрывна в ней.

Градиент функции

Определение 2. Вектор $\left(\frac{\partial f(\vec{x}^0)}{\partial x_1}, \ldots, \frac{\partial f(\vec{x}^0)}{\partial x_n}\right)$ называется **градиентом функции** f в точке $\vec{x}^0$ и обозначается $\operatorname{grad} f(\vec{x}^0)$.

Определение 3. Если $\vec{x}$ и $\vec{y}$ — вектора из $\mathbb{R}^n$, то их **скалярным произведением** называется величина $(\vec{x}, \vec{y}) = \sum\limits_{k=1}^{n} x_k y_k$.

В введенных обозначениях производная по направлению

$$\frac{\partial f(\vec{x}^0)}{\partial \vec{w}} = \left(\operatorname{grad} f(\vec{x}^0), \vec{w}\right).$$

Теорема 2. *Если функция f дифференцируема в точке $\vec{x}^0$, то длина* $\operatorname{grad} f(\vec{x}^0)$ *равна максимальной величине производной по направлению в точке $\vec{x}^0$ и, если* $\operatorname{grad} f(\vec{x}^0) \neq 0$, *то он направлен в ту же сторону, что и единственный вектор направления $\vec{w}$, вдоль которого производная $\frac{\partial f(\vec{x}^0)}{\partial \vec{w}}$ максимальна.*

▼ Из неравенства Коши-Буняковского следует, что

$$\frac{\partial f(\vec{x}^0)}{\partial \vec{w}} \leqslant \|\operatorname{grad} f(\vec{x}^0)\| \cdot \|\vec{w}\| = \|\operatorname{grad} f(\vec{x}^0)\|.$$

Если $\operatorname{grad} f(\vec{x}^0) \neq 0$, то возьмем $\vec{w} = \frac{\operatorname{grad} f(\vec{x}^0)}{\|\operatorname{grad} f(\vec{x}^0)\|}$ и тогда

$$\frac{\partial f(\vec{x}^0)}{\partial \vec{w}} = \left(\operatorname{grad} f(\vec{x}^0), \frac{\operatorname{grad} f(\vec{x}^0)}{\|\operatorname{grad} f(\vec{x}^0)\|}\right) = \|\operatorname{grad} f(\vec{x}^0)\|.$$

По тому же неравенству Коши-Буняковского равенство может достигаться лишь на векторе, пропорциональном $\operatorname{grad} f(\vec{x}^0)$, а так как вектор $-\frac{\operatorname{grad} f(\vec{x}^0)}{\|\operatorname{grad} f(\vec{x}^0)\|}$ не годится (по этому направлению $\frac{\partial f(\vec{x}^0)}{\partial \vec{w}}$ минимальна), то такое направление единственно. ▲

Правила дифференцирования

Теорема 3. *Если функции f и g дифференцируемы в точке $\vec{x}^0$, то дифференцируемы в ней и функции $f \pm g$, $f \cdot g$, а если $g(\vec{x}^0) \neq 0$, то и $\frac{f}{g}$ и при этом*

$$\begin{gathered}
\boldsymbol{d}(f \pm g)(\vec{x}^0) = \boldsymbol{d}f(\vec{x}^0) \pm \boldsymbol{d}g(\vec{x}^0), \\
\boldsymbol{d}(f \cdot g)(\vec{x}^0) = f(\vec{x}^0)\boldsymbol{d}g(\vec{x}^0) + g(\vec{x}^0)\boldsymbol{d}f(\vec{x}^0), \\
\boldsymbol{d}\left(\frac{f}{g}\right)(\vec{x}^0) = \frac{g(\vec{x}^0)\boldsymbol{d}f(\vec{x}^0) - f(\vec{x}^0)\boldsymbol{d}g(\vec{x}^0)}{g^2(\vec{x}^0)}.
\end{gathered}$$

▼ Поскольку

$$\begin{gathered}
\Delta(f \pm g) = \Delta f \pm \Delta g = \boldsymbol{d}f(\vec{x}^0) + o(\|\overrightarrow{\Delta x}\|) + \\
+ \boldsymbol{d}g(\vec{x}^0) + o(\|\overrightarrow{\Delta x}\|) = \boldsymbol{d}f(\vec{x}^0) \pm \boldsymbol{d}g(\vec{x}^0) + o(\|\overrightarrow{\Delta x}\|),
\end{gathered}$$

то $f \pm g$ дифференцируема в точке $\vec{x}^0$ и $\boldsymbol{d}(f \pm g)(\vec{x}^0) = \boldsymbol{d}f(\vec{x}^0) \pm \boldsymbol{d}g(\vec{x}^0)$.

Аналогично, учитывая непрерывность (в силу дифференцируемости) функции g в точке $\vec{x}^0$, получаем

$$\Delta(f\cdot g) = f(\vec{x}^0+\overrightarrow{\Delta x})g(\vec{x}^0+\overrightarrow{\Delta x}) - f(\vec{x}^0)g(\vec{x}^0) = g(\vec{x}^0+\overrightarrow{\Delta x})\big(f(\vec{x}^0+\overrightarrow{\Delta x}) - f(\vec{x}^0)) +$$
$$+ f(\vec{x}^0)(g(\vec{x}^0+\overrightarrow{\Delta x}) - g(\vec{x}^0)) = (g(\vec{x}^0)+o(1))\Delta f + f(\vec{x}^0)\Delta g =$$
$$= \big(g(\vec{x}^0)+o(1)\big)\Big(\boldsymbol{d}f(\vec{x}^0)+o(\|\overrightarrow{\Delta x}\|)\Big) + f(\vec{x}^0)\Big(\boldsymbol{d}g(\vec{x}^0)+o(\|\overrightarrow{\Delta x}\|)\Big) = g(\vec{x}^0)\boldsymbol{d}f(\vec{x}^0)+$$
$$+ f(\vec{x}^0)\boldsymbol{d}g(\vec{x}^0) + g(\vec{x}^0)\cdot o(\|\overrightarrow{\Delta x}\|) + o(1)\cdot \boldsymbol{d}f(\vec{x}^0) + o(1)\cdot o(\|\overrightarrow{\Delta x}\|) + f(\vec{x}^0)\cdot o(\|\overrightarrow{\Delta x}\|),$$

где последние четыре слагаемых — $o(\|\overrightarrow{\Delta x}\|)$. Для всех, кроме $o(1)\cdot\boldsymbol{d}f(\vec{x}^0)$, это очевидно. По предыдущей лекции $\boldsymbol{d}f(\vec{x}^0) = O(\|\overrightarrow{\Delta x}\|)$, а так как $o(1)\cdot O(\|\overrightarrow{\Delta x}\|) = o(\|\overrightarrow{\Delta x}\|)$, то и $o(1)\cdot\boldsymbol{d}f(\vec{x}^0) = o(\|\overrightarrow{\Delta x}\|)$. Значит, $f\cdot g$ дифференцируема в точке $\vec{x}^0$ и для ее дифференциала верно равенство $\boldsymbol{d}(f\cdot g)(\vec{x}^0) = f(\vec{x}^0)\boldsymbol{d}g(\vec{x}^0) + g(\vec{x}^0)\boldsymbol{d}f(\vec{x}^0)$.

Точно также устанавливается дифференцируемость частного.

$$\Delta\left(\frac{f}{g}\right) = \frac{f(\vec{x}^0+\overrightarrow{\Delta x})}{g(\vec{x}^0+\overrightarrow{\Delta x})} - \frac{f(\vec{x}^0)}{g(\vec{x}^0)} = \frac{f(\vec{x}^0+\overrightarrow{\Delta x})g(\vec{x}^0) - f(\vec{x}^0)g(\vec{x}^0+\overrightarrow{\Delta x})}{g(\vec{x}^0+\overrightarrow{\Delta x})g(\vec{x}^0)} =$$
$$= \frac{g(\vec{x}^0)\Delta f - f(\vec{x}^0)\Delta g}{g(\vec{x}^0+\overrightarrow{\Delta x})g(\vec{x}^0)} = \left(\frac{1}{g(\vec{x}^0)}+o(1)\right)\frac{1}{g(\vec{x}^0)}\cdot\Big(g(\vec{x}^0)\boldsymbol{d}f(\vec{x}^0) + g(\vec{x}^0)\cdot o(\|\overrightarrow{\Delta x}\|) -$$
$$- f(\vec{x}^0)\boldsymbol{d}g(\vec{x}^0) - f(\vec{x}^0)\cdot o(\|\overrightarrow{\Delta x}\|)\Big) = \frac{g(\vec{x}^0)\boldsymbol{d}f(\vec{x}^0) - f(\vec{x}^0)\boldsymbol{d}g(\vec{x}^0)}{g^2(\vec{x}^0)} + \frac{o(\|\overrightarrow{\Delta x}\|)}{g(\vec{x}^0)} -$$
$$- \frac{f(\vec{x}^0)}{g^2(\vec{x}^0)}\cdot o(\|\overrightarrow{\Delta x}\|) + o(1)\cdot\boldsymbol{d}f(\vec{x}^0) + o(1)\cdot o(\|\overrightarrow{\Delta x}\|) - o(1)\cdot\frac{f(\vec{x}^0)}{g(\vec{x}^0)}\cdot\boldsymbol{d}g(\vec{x}^0) -$$
$$- o(1)\cdot\frac{f(\vec{x}^0)}{g(\vec{x}^0)}\cdot o(\|\overrightarrow{\Delta x}\|),$$

где, учитывая, что по предыдущей лекции $\boldsymbol{d}f(\vec{x}^0) = O(\|\overrightarrow{\Delta x}\|)$ и $\boldsymbol{d}g(\vec{x}^0) = O(\|\overrightarrow{\Delta x}\|)$, получаем, что шесть последних слагаемых — $o(\|\overrightarrow{\Delta x}\|)$. Значит, $\frac{f}{g}$ дифференцируема в точке $\vec{x}^0$ и $\boldsymbol{d}(\frac{f}{g})(\vec{x}^0) = \frac{g(\vec{x}^0)\boldsymbol{d}f(\vec{x}^0) - f(\vec{x}^0)\boldsymbol{d}g(\vec{x}^0)}{g^2(\vec{x}^0)}$. ▲

Дифференцируемость сложной функции

Теорема 4. *Если функция f дифференцируема в точке $\vec{x}^0 \in \mathbb{R}^n$, а функции $x_k(\vec{t})$, $\vec{t}\in\mathbb{R}^m$, $k = 1,\dots,n$, дифференцируемы в точке $\vec{t}^0 \in \mathbb{R}^m$, причем $x_k(\vec{t}^0) = x_k^0$, $k = 1,\dots,n$, то тогда функция $f\left(\vec{x}(\vec{t})\right) = f\left(x_1(\vec{t}),\dots,x_n(\vec{t})\right)$ дифференцируема в точке $\vec{t}^0$ и*

$$\boldsymbol{d}f\left(\vec{x}(\vec{t}^0)\right) = \sum_{i=1}^{n}\frac{\partial f(\vec{x}^0)}{\partial x_i}dx_i(\vec{t}^0) = \sum_{i=1}^{n}\frac{\partial f(\vec{x}^0)}{\partial x_i}\sum_{j=1}^{m}\frac{\partial x_i(\vec{t}^0)}{\partial t_j}\Delta t_j.$$

▼ Из дифференцируемости $f(\vec{x})$ в точке $\vec{x}^0$ следует, что

$$\Delta f = \sum_{i=1}^{n}\frac{\partial f(\vec{x}^0)}{\partial x_i}\Delta x_i + o(\|\overrightarrow{\Delta x}\|) = \sum_{i=1}^{n}\frac{\partial f(\vec{x}^0)}{\partial x_i}\left(\sum_{j=1}^{m}\frac{\partial x_i(\vec{t}^0)}{\partial t_j}\Delta t_j + o(\|\overrightarrow{\Delta t}\|)\right) + o(\|\overrightarrow{\Delta x}\|),$$

и

$$\Delta x_i = \sum_{j=1}^{m} \frac{\partial x_i(\vec{t}^0)}{\partial t_j} \Delta t_j + o(\|\overrightarrow{\Delta t}\|) = O(\|\overrightarrow{\Delta t}\|),$$

а значит, и $\|\overrightarrow{\Delta x}\| = O(\|\overrightarrow{\Delta t}\|)$, откуда следует, что $o(\|\overrightarrow{\Delta x}\|) = o(1) \cdot \frac{\|\overrightarrow{\Delta x}\|}{\|\overrightarrow{\Delta t}\|} \cdot \|\overrightarrow{\Delta t}\| = o(1) \cdot \|\overrightarrow{\Delta t}\| = o(\|\overrightarrow{\Delta t}\|)$. Следовательно,

$$\Delta f = \sum_{i=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_i} \sum_{j=1}^{m} \frac{\partial x_i(\vec{t}^0)}{\partial t_j} \Delta t_j + o(\|\overrightarrow{\Delta t}\|),$$

где двойная сумма — линейная функция от $\overrightarrow{\Delta t}$ и, значит, дифференциал функции f в точке $\vec{t}^0$. ▲

Следствие 1. *Если функция f дифференцируема в точке $\vec{x}^0 \in \mathbb{R}^n$, а функции $x_k(t)$, $t \in \mathbb{R}$, $k = 1, \ldots, n$, дифференцируемы в точке $t^0 \in \mathbb{R}$, причем $x_k(t^0) = x_k^0$, $k = 1, \ldots, n$, то тогда функция одной переменной $f(\vec{x}(t)) = f(x_1(t), \ldots, x_n(t))$ дифференцируема в точке t^0 и*

$$\frac{df(\vec{x}(t^0))}{dt} = \sum_{i=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_i} \frac{dx_i(t^0)}{dt}.$$

▼ Следствие — частный случай предыдущей теоремы. ▲

Следствие 2. *Если функция f дифференцируема в точке $\vec{x}^0 \in \mathbb{R}^n$, а функции $x_i(\vec{t})$, $\vec{t} \in \mathbb{R}^m$, $i = 1, \ldots, n$, имеют частные производные $\frac{\partial x_i(\vec{t}^0)}{\partial t_j}$, $i = 1, \ldots, n$, в точке $\vec{t}^0 \in \mathbb{R}^m$, причем $x_i(\vec{t}^0) = x_i^0$, $i = 1, \ldots, n$, то тогда функция $f\left(\vec{x}(\vec{t})\right) = f\left(x_1(\vec{t}), \ldots, x_n(\vec{t})\right)$ имеет в точке $\vec{t}^0$ частную производную $\frac{\partial f\left(\vec{x}(\vec{t}^0)\right)}{\partial t_j}$ и*

$$\frac{\partial f\left(\vec{x}(\vec{t}^0)\right)}{\partial t_j} = \sum_{i=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_i} \frac{\partial x_i(\vec{t}^0)}{\partial t_j}.$$

▼ Это следствие вытекает из предыдущего. ▲

Инвариантность первого дифференциала

Следствие 3. *При предположениях последней теоремы вычисление дифференциала функции $f(\vec{x}(\vec{t}))$ прямым способом*

$$\boldsymbol{d}f = \sum_{j=1}^{m} \frac{\partial f(\vec{x}(\vec{t}))}{\partial t_j}\Big|_{\vec{t}^0} \Delta t_j,$$

где $\frac{\partial f(\vec{x}(\vec{t}))}{\partial t_j}\Big|_{\vec{t}^0} = \sum_{i=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_i} \frac{\partial x_i(\vec{t}^0)}{\partial t_j}$, или последовательным способом

$$\boldsymbol{d}f = \sum_{i=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_i} dx_i = \sum_{i=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_i} \left(\sum_{j=1}^{m} \frac{\partial x_i(\vec{t}^0)}{\partial t_j} \Delta t_j \right)$$

(ведь $\boldsymbol{d}x_i = \sum\limits_{j=1}^{m} \frac{\partial x_i(\vec{t}^{\,0})}{\partial t_j}\Delta t_j$, если $\boldsymbol{d}x_i$ рассматривать как дифференциал x_i, а не как приращение x_i) приводит к одинаковому результату.

Приведенное свойство называется **инвариантностью первого дифференциала**.

Лекция 24 (12.05.20)
Производные высших порядков. Теоремы о равенстве смешанных производных

Частные производные второго порядка

Определение 1. Если частная производная $\frac{\partial}{\partial x_i} f(\vec{x})$ существует в окрестности точки $\vec{x}^{\,0}$ и для нее как функции существует частная производная по x_j в точке $\vec{x}^{\,0}$

$$\frac{\partial}{\partial x_j}\left(\frac{\partial}{\partial x_i} f(\vec{x})\right)\big|_{\vec{x}^{\,0}} = \frac{\partial}{\partial x_j}\left(\frac{\partial}{\partial x_i} f(\vec{x}^{\,0})\right),$$

то эту производную называют **частной производной второго порядка** функции f по переменным x_i и x_j в точке $\vec{x}^0$.

Обозначение 1. Частную производную второго порядка функции f по переменным x_i и x_j в точке $\vec{x}^0$ обозначают

$$\frac{\partial^2}{\partial x_j \partial x_i} f(\vec{x})\big|_{\vec{x}=\vec{x}^{\,0}} \text{ или } \frac{\partial^2 f(\vec{x})}{\partial x_j \partial x_i}\big|_{\vec{x}=\vec{x}^{\,0}} \text{ или } \frac{\partial^2 f(\vec{x}^{\,0})}{\partial x_j \partial x_i} \text{ или } f''_{x_i,x_j}(\vec{x}^{\,0}).$$

В двух первых обозначениях вместо “$\vec{x} = \vec{x}^{\,0}$” пишут также “$\vec{x}^{\,0}$”.

Определение 2. Если x_j и x_i одна и та же переменная, т.е. $j = i$, то частную производную $\frac{\partial^2 f(\vec{x}^{\,0})}{\partial x_i \partial x_i}$ называют **чистой** частной производной и обозначают $\frac{\partial^2 f(\vec{x}^{\,0})}{(\partial x_i)^2}$ или, по традиции опуская скобки, $\frac{\partial^2 f(\vec{x}^{\,0})}{\partial x_i^2}$.

Определение 3. Если x_j и x_i разные переменные, т.е. $j \neq i$, то частную производную $\frac{\partial^2 f(\vec{x}^{\,0})}{\partial x_j \partial x_i}$ называют **смешанной** частной производной.

Частные производные высших порядков

Аналогично частной производной второго порядка определяются частные производные любого натурального порядка.

Определение 4. Если частная производная порядка m ($m \in \mathbb{N}$) функции f по переменным $x_{i_1}, \ldots, x_{i_m}$ $\frac{\partial^m f(\vec{x})}{\partial x_{i_m} \ldots \partial x_{i_1}}$ существует в окрестности точки $\vec{x}^{\,0}$ и для нее как функции существует частная производная по переменной $x_{i_{m+1}}$ в точке $\vec{x}^{\,0}$

$$\frac{\partial}{\partial x_{i_{m+1}}}\left(\frac{\partial^m f(\vec{x})}{\partial x_{i_m} \ldots \partial x_{i_1}}\right)\Bigg|_{\vec{x}=\vec{x}^{\,0}},$$

то эту производную называют **частной производной порядка** $m + 1$ функции f по переменным $x_{i_1}, \ldots, x_{i_m}, x_{i_{m+1}}$ в точке $\vec{x}^{\,0}$.

Обозначение 2. Частную производную порядка $m+1$ функции f по переменным $x_{i_1}, \dots, x_{i_m}, x_{i_{m+1}}$ в точке $\vec{x}^0$ обозначают

$$\frac{\partial^{m+1}}{\partial x_{i_{m+1}} \partial x_{i_m} \dots \partial x_{i_1}} f(\vec{x}) \Big|_{\vec{x}=\vec{x}^0} \text{ или } \frac{\partial^{m+1} f(\vec{x})}{\partial x_{i_{m+1}} \partial x_{i_m} \dots \partial x_{i_1}} \Big|_{\vec{x}^0} \text{ или}$$
$$\frac{\partial^{m+1} f(\vec{x}^0)}{\partial x_{i_{m+1}} \partial x_{i_m} \dots \partial x_{i_1}} \text{ или } f^{(m+1)}_{x_{i_1}, \dots, x_{i_m}, x_{i_{m+1}}}.$$

Определение 5. Если $x_{i_{m+1}}, x_{i_m}, \dots, x_{i_1}$ одна и та же переменная, т.е. $i_{m+1} = i_m = \dots = i_1$, то такую частную производную называют **чистой** и обозначают $\frac{\partial^{m+1} f(\vec{x}^0)}{(\partial x_{i_1})^{m+1}}$ или, по традиции опуская скобки, $\frac{\partial^{m+1} f(\vec{x}^0)}{\partial x_{i_1}^{m+1}}$.

Определение 6. Если среди переменных $x_{i_{m+1}}, x_{i_m}, \dots, x_{i_1}$ встречаются различные, то частную производную $\frac{\partial^{m+1} f(\vec{x}^0)}{\partial x_{i_{m+1}} \partial x_{i_m} \dots \partial x_{i_1}}$ называют **смешанной** частной производной.

Теоремы о равенстве смешанных производных

Теорема 1 (Шварца). *Пусть функция двух переменных $f(x,y)$ имеет в точке (x_0, y_0) непрерывные смешанные производные $\frac{\partial^2 f}{\partial y \partial x}$ и $\frac{\partial^2 f}{\partial x \partial y}$. Тогда они равны,*

$$\frac{\partial^2}{\partial x \partial y} f(x_0, y_0) = \frac{\partial^2}{\partial y \partial x} f(x_0, y_0).$$

▼ Будем брать Δx и Δy настолько малыми, что точка $(x_0 + \Delta x, y_0 + \Delta y)$ лежит в окрестности точки (x_0, y_0), где существуют первые частные производные и смешанные производные второго порядка. Рассмотрим выражение

$$\Delta^2 f = f(x_0 + \Delta x, y_0 + \Delta y) - -f(x_0, y_0 + \Delta y) - f(x_0 + \Delta x, y_0) + f(x_0, y_0),$$

называемое второй разностью. Его можно рассматривать как разность по переменной x от разности по переменной y,

$$\Delta^2 f = [f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0)] - [f(x_0, y_0 + \Delta y) - f(x_0, y_0)] =$$
$$= \frac{\partial}{\partial x} [f(x_0 + \theta \Delta x, y_0 + \Delta y) - f(x_0 + \theta \Delta x, y_0)] \Delta x =$$
$$= \left[\frac{\partial}{\partial x} f(x_0 + \theta \Delta x, y_0 + \Delta y) - \frac{\partial}{\partial x} f(x_0 + \theta \Delta x, y_0) \right] \Delta x = \Delta_y f'_x(x_0 + \theta \Delta x, y) \Delta x,$$

где $0 < \theta < 1$ по теореме Лагранжа. Еще раз применяя теорему Лагранжа, имеем

$$\Delta^2 f = \Delta_y f'_x(x_0 + \theta \Delta x, y) \Delta x = \left[\frac{\partial}{\partial x} f(x_0 + \theta \Delta x, y_0 + \Delta y) - \frac{\partial}{\partial x} f(x_0 + \theta \Delta x, y_0) \right] \Delta x =$$
$$= \frac{\partial}{\partial y} \left(\frac{\partial}{\partial x} f(x_0 + \theta \Delta x, y_0 + \vartheta \Delta y) \right) \cdot \Delta x \cdot \Delta y = \frac{\partial^2}{\partial y \partial x} f(x_0 + \theta \Delta x, y_0 + \vartheta \Delta y) \cdot \Delta x \cdot \Delta y,$$

где $0 < \vartheta < 1$, $0 < \theta < 1$.

В силу непрерывности смешанной производной $\frac{\partial^2 f}{\partial y \partial x}$ в точке (x_0, y_0) получаем, что при $\Delta x \to 0$, $\Delta y \to 0$ так, что ни одно из приращений не обращается в 0,

$$\frac{\Delta^2 f}{\Delta y\, \Delta x} \xrightarrow[\Delta x \to 0 \Delta y \to 0]{} \frac{\partial^2 f}{\partial y\, \partial x} f(x_0, y_0).$$

Аналогично, при $\Delta x \to 0$, $\Delta y \to 0$ так, что ни одно из приращений не обращается в 0,

$$\frac{\Delta^2 f}{\Delta y\, \Delta x} \xrightarrow[\Delta x \to 0 \Delta y \to 0]{} \frac{\partial^2 f}{\partial x\, \partial y} f(x_0, y_0),$$

что и доказывает теорему. ▲

В теореме Шварца можно заменить требование существования и непрерывности смешанной производной $\frac{\partial^2 f}{\partial x \partial y}$ (или $\frac{\partial^2 f}{\partial y \partial x}$) на требование существования непрерывной частной производной первого порядка $\frac{\partial f}{\partial y}$ (соответственно,$\frac{\partial f}{\partial y}$) в окрестности точки (x_0, y_0).

Следствие 4. *Если функция f (из $\mathbb{R}^n$ в $\mathbb{R}$) имеет в точке $\vec{x}^0$ непрерывные смешанные частные производные $\frac{\partial^2}{\partial x_j \partial x_i} f$ и $\frac{\partial^2}{\partial x_i \partial x_j} f$, то эти частные производные равны.*

▼ Действительно, полагая $x_k = x_k^0$ при $k \neq i, j$, т.е. фиксируя все переменные, кроме i-ой и j-ой, сводим данное утверждение к утверждению теоремы Шварца. ▲

Следствие 1. *Если у функции f (из $\mathbb{R}^n$ в $\mathbb{R}$) любые смешанные частные производные порядка m непрерывны в точке $\vec{x}^0$, а производные порядков меньше m непрерывны в ее окрестности, то смешанные частные производные до порядка m включительно по одинаковым наборам переменных (с учетом их кратности, но без учета их порядка) равны.*

▼ Действительно, при $m = 2$ это предыдущее следствие. А если утверждение верно при $m = r$, то оно верно и при $m = r + 1$. Ведь по предположению (индукционному) все смешанные частные производные до порядка r включительно по одинаковым наборам переменных равны в некоторой окрестности точки $\vec{x}^0$ (т.е. первые r (или менее) дифференцирований можно совершать в некоторой окрестности точки $\vec{x}^0$ в любом порядке), а по следствию 1 можно менять порядок r-ого и $r+1$-ого дифференцирования в точке $\vec{x}^0$. Значит, все $r + 1$ дифференцирований можно совершать в точке $\vec{x}^0$ в любом порядке — результат не меняется. ▲

Теорема 2 (Юнга). *Если функция двух переменных $f(x, y)$ дважды дифференцируема в точке (x_0, y_0) (т.е. дифференцируемы частные производные первого порядка), то в этой точке смешанные производные второго порядка равны,*

$$\frac{\partial^2}{\partial x \partial y} f(x_0, y_0) = \frac{\partial^2}{\partial y \partial x} f(x_0, y_0).$$

▼ Как и в доказательстве теоремы Шварца рассмотрим выражение

$$\Delta^2 f = [f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0)] - [f(x_0, y_0 + \Delta y) - f(x_0, y_0)] =$$
$$= \frac{\partial}{\partial x} [f(x_0 + \theta \Delta x, y_0 + \Delta y) - f(x_0 + \theta \Delta x, y_0)] \cdot \Delta x =$$
$$= \left[\frac{\partial}{\partial x} f(x_0 + \theta \Delta x, y_0 + \Delta y) - \frac{\partial}{\partial x} f(x_0 + \theta \Delta x, y_0) \right] \Delta x = \Delta_y f'_x(x_0 + \theta \Delta x, y) \Delta x,$$

где $0 < \theta < 1$ (по теореме Лагранжа). В силу дифференцируемости частной производной f'_x

$$\Delta^2 f = \Delta_y f'_x(x_0 + \theta\Delta x, y)\Delta x = [f'_x(x_0 + \theta\Delta x, y_0 + \Delta y) - f'_x(x_0 + \theta\Delta x, y_0)]\,\Delta x =$$
$$= f'_x(x_0, y_0) \cdot \Delta x + f''_{x,x} \cdot \theta(\Delta x)^2 + f''_{x,y}(x_0, y_0) \cdot \Delta x \cdot \Delta y + o(\|(\theta\Delta x, \Delta y)\|) \cdot \Delta x -$$
$$- f'_x(x_0, y_0) \cdot \Delta x - f''_{x,x} \cdot \theta(\Delta x)^2 - o(|\theta\Delta x|) \cdot \Delta x = f''_{x,y}(x_0, y_0) \cdot \Delta x \cdot \Delta y +$$
$$+ o(\|(\Delta x, \Delta y)\|) \cdot \Delta x$$

и, значит,

$$\lim_{\Delta y = \Delta x \to 0} \frac{\Delta^2 f}{\Delta y \cdot \Delta x} = f''_{x,y}(x_0, y_0).$$

Аналогично,

$$\lim_{\Delta y = \Delta x \to 0} \frac{\Delta^2 f}{\Delta x \cdot \Delta y} = f''_{y,x}(x_0, y_0),$$

что и доказывает теорему. ▲

Следствие 2. *Если функция f из $\mathbb{R}^n$ в $\mathbb{R}$ дважды дифференцируема в точке $\vec{x}^0$, то для любых переменных x_i и x_j*

$$\frac{\partial^2}{\partial x_i \partial x_j} f(\vec{x}^0) = \frac{\partial^2}{\partial x_j \partial x_i} f(\vec{x}^0).$$

▼ Действительно, полагая $x_k = x_k^0$ при $k \neq i, j$, т.е. фиксируя все переменные, кроме i-ой и j-ой, сводим данное утверждение к утверждению теоремы. ▲

Следствие 3. *Если функция f из $\mathbb{R}^n$ в $\mathbb{R}$ m раз дифференцируема в точке $\vec{x}^0$, то смешанные производные до порядка m включительно по одинаковым наборам переменных (с учетом их кратности, но без учета их порядка) равны.*

▼ Действительно, при $m = 2$ это предыдущее следствие. А если утверждение верно при $m = r$, то оно верно и при $m = r + 1$. Ведь все частные производные до порядка r включительно непрерывны в некоторой окрестности точки $\vec{x}^0$ и, значит, совпадают при дифференцировании по одинаковым наборам переменных, а по предыдущему следствию можно менять порядок r-ого и $r + 1$-ого дифференцирования. Значит, все $r + 1$ дифференцирований можно совершать в любом порядке — результат получается одинаковый. ▲

Лекция 25 (15.05.20)
Дифференциалы высших порядков. Формула Тейлора и локальные экстремумы

Кратная дифференцируемость функции

Зная определение дифференцируемости функции в точке дадим теперь по индукции определение кратной дифференцируемости функции в точке.

Определение 1. Функция f называется $m + 1$ **раз дифференцируемой в точке** $\vec{x}^0$, где $m \in \mathbb{N}$, если сама функция и все ее частные производные порядка k, $1 \leqslant k < m$, дифференцируемы в некоторой окрестности точки $\vec{x}^0$, а все ее частные производные порядка m дифференцируемы в точке $\vec{x}^0$.

Утверждение. Из теоремы о достаточном условии дифференцируемости следует, что для того, чтобы функция f была m раз дифференцируемой в точке $\vec{x}^0$ достаточно, чтобы все ее частные производные порядка k, $1 \leqslant k < m$, существовали и были непрерывны в некоторой окрестности точки $\vec{x}^0$, а все частные производные порядка m существовали и были непрерывны в точке $\vec{x}^0$.

Определение 2. Функция f называется m **раз непрерывно дифференцируемой** в точке $\vec{x}^0$, если f дифференцируема m раз в некоторой окрестности точки $\vec{x}^0$ и все ее частные производные порядка m непрерывны в точке $\vec{x}^0$.

Дифференциал второго порядка

Определение 3. Если первый дифференциал $\boldsymbol{d}f(\vec{x}, \overrightarrow{\Delta x})$ как функция от $\vec{x}$ при любом фиксированном $\overrightarrow{\Delta x}$ дифференцируем в точке $\vec{x}^0$, то выражение, являющееся дифференциалом от первого дифференциала при таком же приращении $\overrightarrow{\Delta x}$, что и фиксированное в первом дифференциале, называется **вторым дифференциалом**. Если f дважды дифференцируема в точке $\vec{x}^0$, то второй дифференциал

$$\boldsymbol{d}^2 f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = \boldsymbol{d}^2 f(\vec{x}^0, \overrightarrow{\Delta x}) = \sum_{j=1}^{n} \boldsymbol{d}\left(\frac{\partial f(\vec{x}^0)}{\partial x_j}\right)\big|_{\vec{x}^0} \Delta x_j =$$
$$= \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f(\vec{x}^0)}{\partial x_i \partial x_j} \Delta x_j \Delta x_i.$$

Вместо Δx_i и Δx_j часто используют обозначения dx_i и dx_j.

Замечание 1. Иногда под вторым дифференциалом понимается не квадратичная, а билинейная форма переменных $\overrightarrow{\Delta_1 x}$ и $\overrightarrow{\Delta_2 x}$

$$\boldsymbol{d}^2 f(\overrightarrow{\Delta_1 x}, \overrightarrow{\Delta_2 x})\big|_{\vec{x}^0} = \boldsymbol{d}^2 f(\vec{x}^0, \overrightarrow{\Delta_1 x}, \overrightarrow{\Delta_2 x}) =$$
$$= \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f(\vec{x}^0)}{\partial x_i \partial x_j} \Delta_1 x_j \Delta_2 x_i.$$

Но мы таким понятием второго дифференциала пользоваться не будем.

Дифференциалы высших порядков

Аналогично дифференциалу второго порядка определяются дифференциалы любого порядка.

Определение 4. Если дифференциал m-ого порядка $\boldsymbol{d}^m f(\vec{x}, \overrightarrow{\Delta x})$ как функция от $\vec{x}$ при любом фиксированном $\overrightarrow{\Delta x}$ дифференцируем в точке $\vec{x}^0$, то выражение, являющееся дифференциалом от m-ого дифференциала, взятое при таком же приращении $\overrightarrow{\Delta x}$, что и зафиксированное в m-ом дифференциале, называется $m+1$**-ым дифференциалом** функции f в точке $\vec{x}^0$. Если функция f $m+1$ раз дифференцируема в точке $\vec{x}^0$, то $m+1$-ый дифференциал

$$\boldsymbol{d}^{m+1} f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = \boldsymbol{d}^{m+1} f(\vec{x}^0, \overrightarrow{\Delta x}) =$$
$$= \sum_{i_{m+1}=1}^{n}\sum_{i_m=1}^{n}\cdots\sum_{i_1=1}^{n} \frac{\partial^{m+1} f(\vec{x}^0)}{\partial x_{i_{m+1}} \partial x_{i_m} \ldots \partial x_{i_1}} \Delta x_{i_1} \ldots \Delta x_{i_{m+1}}.$$

Замечание 2. Иногда под $m+1$-ым дифференциалом понимается не квадратичная, а полилинейная форма переменных $\overrightarrow{\Delta_1 x}$, $\overrightarrow{\Delta_2 x}, \ldots, \overrightarrow{\Delta_{m+1} x}$

$$\boldsymbol{d}^{m+1} f\left(\overrightarrow{\Delta_1 x}, \overrightarrow{\Delta_2 x}, \ldots, \overrightarrow{\Delta_{m+1} x}\right)\big|_{\vec{x}^0} =$$
$$= \sum_{i_{m+1}=1}^{n} \sum_{i_m=1}^{n} \cdots \sum_{i_1=1}^{n} \frac{\partial^{m+1} f(\vec{x}^0)}{\partial x_{i_{m+1}} \partial x_{i_m} \ldots \partial x_{i_1}} \Delta_1 x_{i_1} \ldots \Delta_{m+1} x_{i_{m+1}}.$$

Но мы таким понятием высших дифференциалов пользоваться не будем.

Формула Тейлора

Определение 5. Для функции f из $\mathbb{R}^n$ в $\mathbb{R}$ дифференцируемой в точке $\vec{x}^0$ не менее m раз равенство

$$f(\vec{x}) = f(\vec{x}^0 + \overrightarrow{\Delta x}) = \sum_{k=0}^{m} \frac{1}{k!} \boldsymbol{d}^k f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} + r_m(\overrightarrow{\Delta x})\big|_{\vec{x}^0},$$

где $d^0 f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = f(\vec{x}^0)$, называется **формулой Тейлора** в точке $\vec{x}^0$ функции f; $r_m(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ — **остаточный член** формулы Тейлора. В случае $\vec{x}^0 = \vec{0}$ написанное равенство иногда называют **формулой Маклорена.**

Теорема 1 (остаточный член в форме Лагранжа). *Если функция f из $\mathbb{R}^n$ в $\mathbb{R}$ m раз непрерывно дифференцируема на отрезке $[\vec{x}^0, \vec{x}]$ и $m+1$ раз дифференцируема на интервале $(\vec{x}^0, \vec{x})$ (получающимся из отрезка отбрасыванием его концов), то найдется такое θ, $0 < \theta < 1$, что*

$$f(\vec{x}) = f(\vec{x}^0 + \overrightarrow{\Delta x}) = \sum_{k=0}^{m} \frac{1}{k!} \boldsymbol{d}^k f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} + \frac{1}{(m+1)!} \boldsymbol{d}^{m+1} f(\overrightarrow{\Delta x})\big|_{\vec{x}^0 + \theta \overrightarrow{\Delta x}}$$

(т.е. $r_m(\vec{x}) = \frac{1}{(m+1)!} \boldsymbol{d}^{m+1} f(\overrightarrow{\Delta x})\big|_{\vec{x}^0 + \theta \overrightarrow{\Delta x}}$).

▼ Рассмотрим при фиксированном $\overrightarrow{\Delta x} = \vec{x} - \vec{x}^0$ функцию одного переменного t, $0 \leqslant t \leqslant 1$, $g(t) = f(\vec{x}^0 + t\overrightarrow{\Delta x})$. Тогда

$$\begin{aligned} g(0) &= f(\vec{x}^0), \\ g'(0) &= \boldsymbol{d} f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = \sum_{i=1}^{n} \frac{\partial f(\vec{x}^0)}{\partial x_i} \overrightarrow{\Delta x_i} \end{aligned}$$

(по следствию теоремы о дифференцировании сложной функции)

$$\begin{aligned} g''(0) &= \boldsymbol{d}^2 f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^2 f(\vec{x}^0)}{\partial x_j \, \partial x_i} \Delta x_i \Delta x_j, \\ &\vdots \qquad \vdots \\ g^{(m)}(0) &= \boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}, \\ g^{(m+1)}(t) &= \boldsymbol{d}^{m+1} f(\overrightarrow{\Delta x})\big|_{\vec{x}^0 + t\overrightarrow{\Delta x}}. \end{aligned}$$

По формуле Тейлора с остаточным членом в форме Лагранжа

$$f(\vec{x}^0+t\overrightarrow{\Delta x})=\sum_{k=1}^{m}\frac{1}{k!}g^{(k)}(0)t^k+\frac{1}{(m+1)!}g^{(m+1)}(\theta t)t^{m+1}=$$

$$=\sum_{k=1}^{m}\frac{1}{k!}\boldsymbol{d}^k f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}\cdot t^k+\frac{1}{(m+1)!}\boldsymbol{d}^{m+1} f(\overrightarrow{\Delta x})\big|_{\vec{x}^0+\theta t\overrightarrow{\Delta x}}\cdot t^{m+1}.$$

Теперь, полагая $t=1$, получаем искомую формулу. ▲

Следствие 1 (остаточный член в форме Пеано). *Если функция f из $\mathbb{R}^n$ в $\mathbb{R}$ m раз непрерывно дифференцируема в точке $\vec{x}^0$, то*

$$f(\vec{x})=f(\vec{x}^0+\overrightarrow{\Delta x})=\sum_{k=0}^{m}\frac{1}{k!}\boldsymbol{d}^k f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}+o(\|\overrightarrow{\Delta x}\|^m)$$

(т.е. $r_m(\vec{x})=o(\|\overrightarrow{\Delta x}\|^m)$).

▼ При достаточно малом $\overrightarrow{\Delta x}$ по теореме 1

$$f(\vec{x})=f(\vec{x}^0+\overrightarrow{\Delta x})=\sum_{k=0}^{m-1}\frac{1}{k!}\boldsymbol{d}^k f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}+\frac{1}{m!}\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0+\theta\overrightarrow{\Delta x}}.$$

Дифференциал

$$\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0+\theta\overrightarrow{\Delta x}}$$

является суммой чденов

$$\frac{\partial^m f(\vec{x}^0+\theta\overrightarrow{\Delta x})}{\partial x_{i_m}\ldots\partial x_{i_1}}\Delta x_{i_1}\ldots\Delta x_{i_m}.$$

Так как f m раз непрерывно дифференцируема в точке $\vec{x}^0$, то

$$\frac{\partial^m f(\vec{x}^0+\theta\overrightarrow{\Delta x})}{\partial x_{i_m}\ldots\partial x_{i_1}}=\frac{\partial^m f(\vec{x}^0)}{\partial x_{i_m}\ldots\partial x_{i_1}}+o(1).$$

Значит,

$$\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0+\theta\overrightarrow{\Delta x}}=\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}+o(\|\overrightarrow{\Delta x}\|^m)$$

и верность следствия установлена. ▲

Замечание 1. Формула Тейлора с остаточным членом в форме Пеано верна просто при дифференцируемости функции f m раз в точке $\vec{x}^0$, но доказать это сложнее.

Локальные экстремумы функций многих переменных

Определение 6. Функция f из $\mathbb{R}^n$ в $\mathbb{R}$ имеет в точке $\vec{x}^0$ **локальный максимум** (**локальный минимум**), если f определена в некоторой окрестности точки $\vec{x}^0$ и

$$\exists B'_\delta(\vec{x}^0)\ \forall\vec{x}\in B'_\delta(\vec{x}^0):\ f(\vec{x})\leqslant f(\vec{x}^0)\quad\big(f(\vec{x})\geqslant f(\vec{x}^0)\big).$$

При этом говорят, что $\vec{x}^0$ — **точка локального максимума** (**локального минимума**), а величина $f(\vec{x}^0)$ — **локальный максимум** (**локальный минимум**) функции f.

Если в приведенном определении заменить нестрогое неравенство строгим, то получится определение **строгого локального максимума** (**строгого локального минимума**). При этом говорят, что $\vec{x}^0$ — **точка строгого локального максимума** (**строгого локального минимума**), а величина $f(\vec{x}^0)$ — **строгий локальный максимум** (**строгий локальный минимум**) функции f.

Определение 7. Термин **локальный экстремум** означает локальный максимум или локальный минимум, а термин **строгий локальный экстремум** — строгий локальный максимум или строгий локальный минимум.

Условия локальных экстремумов

Теорема 3 (необходимое условие локального экстремума). *Если функция f из $\mathbb{R}^n$ в $\mathbb{R}$ имеет в точке $\vec{x}^0$ локальный экстремум, то все частные производные $\frac{\partial f(\vec{x}^0)}{\partial x_k}$, которые существуют, равны нулю: $\frac{\partial f(\vec{x}^0)}{\partial x_k} = 0$.*

▼ Полагая $x_i = x_i^0$ при $i \neq k$, т.е. фиксируя все переменные, кроме k-ой, сводим утверждение теоремы к теореме Ферма утверждающей, что если функция одного переменного дифференцируема в точке локального экстремума, то ее производная в этой точке равна нулю. ▼

Следствие 2. *Если функция f из $\mathbb{R}^n$ в $\mathbb{R}$ имеет в точке $\vec{x}^0$ локальный экстремум и дифференцируема в ней, то для любого $\overrightarrow{\Delta x}$ имеем $\boldsymbol{d}f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = 0$ и $grad\, f(\vec{x}^0) = \vec{0}$.*

Определение 8. Точка $\vec{x} \in \mathbb{R}^n$, в которой обращаются в нуль все частные производные функции f, называется **стационарной точкой** функции f.

Лекция 26 (19.05.20)
Условия локального ъкстремума.
Неявные функции

Для дифференцируемых функций поиск точек локальных экстремумов обычно сводится к определению стационарных точек и исследованию, какие из них являются точками локальных экстремумов.

Теорема 1 (об условиях локального экстремума). *Пусть функция f из $\mathbb{R}^n$ в $\mathbb{R}$ m раз непрерывно дифференцируема в точке $\vec{x}^0$ и $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ первый отличный от нуля дифференциал функции f в точке $\vec{x}^0$.*

Если $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ положительно определен (т.е. $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} > 0$ при $\overrightarrow{\Delta x} \neq 0$), то f имеет в точке $\vec{x}^0$ строгий локальный минимум; если f имеет в $\vec{x}^0$ локальный минимум, то $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ определен неотрицательно (т.е. $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} \geqslant 0$ при любом $\overrightarrow{\Delta x}$).

Если $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ отрицательно определен (т.е. $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} < 0$ при $\overrightarrow{\Delta x} \neq 0$), то f имеет в точке $\vec{x}^0$ строгий локальный максимум; если f имеет в $\vec{x}^0$ локальный максимум, то $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ определен неположительно (т.е. $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} \leqslant 0$ при любом $\overrightarrow{\Delta x}$).

Если $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ принимает значения разных знаков (как строго больше нуля, так и строго меньше), то f не имеет в точке $\vec{x}^0$ локального экстремума.

▼ Если $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} > 0$ при $\overrightarrow{\Delta x} \neq 0$), то, так как единичная сфера в $\mathbb{R}^n$ компакт, значение

$$\eta = \min_{\|\overrightarrow{\Delta x}\|=1} \frac{1}{m!}\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$$

достигается в некоторой точке сферы (по 2-ой теореме Вейерштрасса) и $\eta > 0$. По формуле Тейлора с остаточным членом в форме Пеано

$$\Delta f = f(\vec{x}^0 + \overrightarrow{\Delta x}) - f(\vec{x}^0) = \frac{1}{m!}\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} + o(1) \cdot \|\overrightarrow{\Delta x}\|^m,$$

а

$$\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = \boldsymbol{d}^m f\left(\frac{\overrightarrow{\Delta x}}{\|\overrightarrow{\Delta x}\|}\right)\big|_{\vec{x}^0} \cdot \left\|\overrightarrow{\Delta x}\right\|^m$$

(это очевидно для каждого из членов $\frac{\partial^m f(\vec{x}^0)}{\partial x_{i_m} \ldots \partial x_{i_1}} \cdot \Delta x_{i_1} \ldots \Delta x_{i_m}$ дифференциала), поэтому

$$\Delta f \geqslant (\eta + o(1)) \cdot \|\overrightarrow{\Delta x}\|^m.$$

Найдем такую проколотую δ-окрестность $B'_\delta(\vec{x}^0)$, что на ней $|o(1)| < \eta$. Тогда на $B'_\delta(\vec{x}^0)$

$$\Delta f \geqslant (\eta + o(1)) \cdot \|\overrightarrow{\Delta x}\|^m > 0$$

и, значит, в точке $\vec{x}^0$ строгий локальный минимум.

Если в точке $\vec{x}^0$ локальный минимум функции f, то в некоторой окрестности $\vec{x}^0$

$$\Delta f = \frac{1}{m!}\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} + o(1) \cdot \|\overrightarrow{\Delta x}\|^m \geqslant 0.$$

Значит,

$$\lim_{\alpha \to +0} \alpha^{-m} \left(\frac{1}{m!}\boldsymbol{d}^m f(\alpha\overrightarrow{\Delta x})\big|_{\vec{x}^0} + o(1) \cdot \|\alpha\overrightarrow{\Delta x}\|^m\right) =$$
$$= \frac{1}{m!}\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} + \lim_{\alpha \to +0} o(1) \cdot \|\overrightarrow{\Delta x}\|^m = \frac{1}{m!}\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} \geqslant 0.$$

для любого $\overrightarrow{\Delta x}$, т.е. дифференциал $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ определен неотрицательно.

Другие случаи, когда дифференциал $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ отрицательно определен или когда $\vec{x}^0$ — точка локального максимума функции f, рассматриваются аналогично. Они могут быть также сведены к уже изученным случаям рассмотрением функции $-f$.

И, наконец, если дифференциал $\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0}$ принимает значения разных знаков, то, в соответствии с предыдущими пунктами теоремы, не выполняется необходимое условие как локального минимума, так и локального максимума. ▲

Замечание. Отметим, что дифференциал нечетного порядка или нулевой или принимает значения разных знаков, так как $\boldsymbol{d}^m f(\overrightarrow{\Delta x}) = -\boldsymbol{d}^m f(-\overrightarrow{\Delta x}$ при нечетном m.

На практике наиболее часто встречается случай $m = 2$. В этом случае определение характера стационарной точки приводит к изучению квадратичной формы

$$\boldsymbol{d}^m f(\overrightarrow{\Delta x})\big|_{\vec{x}^0} = \sum_{j=1}^n \sum_{i=1}^n \frac{\partial^2 f(\vec{x}^0)}{\partial x_j \partial x_i} \Delta x_i \, \Delta x_j.$$

Для анализа квадратичной формы применяется известный в алгебре критерий Сильверста.

Утверждение. *Квадратичная форма* $\sum\limits_{j=1}^{n}\sum\limits_{i=1}^{n} a_{j,i}t_it_j$ *с симметричной матрицей* $(a_{j,i})_{j,i=1}^{n}$ *(т.е.* $a_{j,i}=a_{i,j}$*) положительно (отрицательно) определена тогда и только тогда, когда все главные миноры матрицы*

$$|a_{j,i}|_{j,i=1}^{k} > 0,\ k=1,\ldots,n,$$
$$\left((-1)^k\,|a_{j,i}|_{j,i=1}^{k} > 0,\ k=1,\ldots,n,\right).$$

Теорема о неявной функции

Определение 1. Если переменная y, являющаяся функцией аргументов $x_1, x_2, \ldots, x_n$, задается посредством функционального уравнения

$$F(y, x_1, \ldots, x_n) = 0,$$

то говорят, что y как функция $x_1, x_2, \ldots, x_n$ **задана неявно** или что y — **неявная функция**.

Естественно, возникают вопросы: при каких условиях функциональное уравнение $F(y, x_1, \ldots, x_n) = F(y, \vec{x}) = 0$, $\vec{x} \in \mathbb{R}^n$, однозначно определяет функцию $y(\vec{x}) = y(x_1, \ldots, x_n)$; при каких условиях $y(\vec{x})$ непрерывна и при каких дифференцируема. Ответы (окончательности которых не будем касаться) содержаться в следующей теореме.

Теорема 2 (о неявной функции). *Пусть* $F(y, \vec{x})$ *функция* $1+n$ *переменного* y, $\vec{x} = (x_1, \ldots, x_n)$, *непрерывна в некоторой окрестности точки* $(y^0, \vec{x}^{\,0}) \in \mathbb{R}^{1+n}$ *и имеет в этой окрестности непрерывную частную производную* F'_y. *Тогда если* $F(y^0, \vec{x}^{\,0}) = 0$, *а* $F'_y(y^0, \vec{x}^{\,0}) \neq 0$, *то*

$$\exists \varepsilon_0 > 0\ \forall \varepsilon \in (0, \varepsilon_0)\ \exists B_\delta(\vec{x}^{\,0}) \subset \mathbb{R}^n\ \forall x \in B_\delta(\vec{x}^{\,0})\ \exists!\, y \in B_\varepsilon(y^0):\ F(y, \vec{x}) = 0,$$

т.е. для любого достаточно малого числа $\varepsilon > 0$ *найдется такая окрестность* $B_\delta(\vec{x}^{\,0})$ *точки* $\vec{x}^{\,0}$, *что в пределах этой окрестности существует единственная функция* $y(\vec{x})$, *удовлетворяющая условию* $y(\vec{x}) \in B_\varepsilon(y^0)$ *и являющаяся решением уравнения* $F(y, \vec{x}) = 0$, *причем эта функция* $y(\vec{x})$ *непрерывна в некоторой окрестности точки* $\vec{x}^{\,0}$.

Если дополнительно потребовать дифференцируемость $F(y, \vec{x})$ *в точке* $(y^0, \vec{x}^{\,0})$ *(в окрестности точки* $(y^0, \vec{x}^{\,0})$*), то функция* $y(\vec{x})$ *будет дифференцируема в точке* $\vec{x}^{\,0}$ *(в некоторой окрестности точки* $\vec{x}^{\,0}$*) и при этом*

$$\boldsymbol{d}y(\overrightarrow{\Delta x})\big|_{\vec{x}} = \sum_{k=1}^{n} -\frac{F'_{x_k}(y, \vec{x})}{F'_y(y, \vec{x})}\Delta x_k,$$

где $y = y(\vec{x})$, *в точке* $\vec{x} = \vec{x}^{\,0}$ *(в некоторой окрестности точки* $\vec{x}^{\,0}$*).*

▼ Выберем число $\varepsilon_0 > 0$ таким, что в $2\varepsilon_0$-окрестности точки $(y^0, \vec{x}^{\,0})$ функция $F(y, \vec{x})$ непрерывна, а частная производная $F'_y(y, \vec{x})$ непрерывна и сохраняет знак (строго положительна, если $F'_y(y^0,\vec{x}^0) > 0$; строго отрицательна, если $F'_y(y^0,\vec{x}^0) > 0$). Так как $F(y,\vec{x}^0)$ как функция одной переменной y имеет производную одного знака на отрезке $[-\varepsilon_0,\varepsilon_0]$, то $F(y,\vec{x}^0)$ — строго монотонная функция от y на $[-\varepsilon_0,\varepsilon_0]$, $F(y^0,\vec{x}^0)=0$, значит, $F(y^0+\varepsilon,\vec{x}^0)\cdot F(y^0-\varepsilon,\vec{x}^0)<0$ для любого $\varepsilon\in(0,\varepsilon_0)$. Пользуясь непрерывностью $F(y,\vec{x}^0)$ найдем такую окрестность точки $\vec{x}^0$ $B_\delta(\vec{x}^0)$, $0<\delta<\varepsilon_0$, что для $\vec{x}\in B_\delta(\vec{x}^0)$

$$F(y^0+\varepsilon,\vec{x})\cdot F(y^0-\varepsilon,\vec{x})<0.$$

При этом

$$\overline{B}_\varepsilon(y^0)\times B_\delta(\vec{x}^0)=\{(y,\vec{x}):\ y\in\overline{B}_\varepsilon(y^0),\ \vec{x}\in B_\delta(\vec{x}^0)\}\subset$$
$$\subset\overline{B}_{\varepsilon_0}(y^0)\times B_{\varepsilon_0}(\vec{x}^0)=\{(y,\vec{x}):\ y\in\overline{B}_{\varepsilon_0}(y^0),\ \vec{x}\in B_{\varepsilon_0}(\vec{x}^0)\}\subset B_{2\varepsilon_0}\big((y^0,\vec{x}^0)\big)$$

(в силу неравенства треугольника). Теперь по теореме Больцано–Коши о промежуточном значении и в силу строгой монотонности $F(y,\vec{x})$ как функции одной переменной y при $\vec{x}\in B_\delta(\vec{x}^0)\subset B_{\varepsilon_0}(\vec{x})$ получаем, что

$$\forall x\in B_\delta(\vec{x}^0)\ \exists!\, y\in B_\varepsilon(y^0):\ F(y,\vec{x})=0.$$

Тем самым показано, что в пределах $B_\delta(\vec{x}^0)$ существует единственная функция $y(\vec{x})$, удовлетворяющая условию $y(\vec{x})\in B_\varepsilon(y^0)$, которая является решением уравнения $F(y,\vec{x})=0$. Так как при этом $y\left(B_\delta(\vec{x}^0)\right)\subset B_\varepsilon(y^0)$, то, значит, $y(\vec{x})$ непрерывна в точке $\vec{x}^0$, $y(\vec{x}^0)=y^0$.

У любой точки $(y(\vec{x}),\vec{x})$, где $\vec{x}\in B_\delta(\vec{x}^0)\subset B_{\varepsilon_0}(\vec{x}^0)$, $y(\vec{x})\in B_\varepsilon(y^0)\subset B_{\varepsilon_0}(\vec{x}^0)$, и значит, $(y(\vec{x}),\vec{x})\in B_{2\varepsilon_0}\left((y^0,\vec{x}^0)\right)$, есть окрестность, в которой $F(y,\vec{x})$ непрерывна и имеет непрерывную частную производную F'_y, $F(y(\vec{x}),\vec{x})=0$, $F'_y(y(\vec{x}),\vec{x})\neq 0$. Значит, для точки $(y(\vec{x}),\vec{x})$ выполняются те же условия, что и для точки $y^0,\vec{x}^0)$, поэтому используя уже доказанную непрерывность $y(\vec{x})$ в точке $\vec{x}^0$ делаем вывод, что $y(\vec{x})$ непрерывна в любой точке $\vec{x}\in B_\delta(\vec{x}^0)$.

Теперь перейдем к дифференцируемости. Если $F(y,\vec{x})$ дифференцируема в точке $(y^0,\vec{x}^0)$, то

$$\Delta F=F'_y(y^0,\vec{x}^0)\Delta y+\sum_{k=1}^n F'_{x_k}(y^0,\vec{x}^0)\Delta x_k+o(1)(|\Delta y|+\|\overrightarrow{\Delta x}\|),$$

причем фигурирующее тут $o(1)$ не больше того, которое присутствует в обычной записи последнего члена в виде $o(1)\cdot\|(\Delta y,\overrightarrow{\Delta x})\|$, так как $|\Delta y|+\|\overrightarrow{\Delta x}\|=\|(\Delta y,0)\|+\|(0,\overrightarrow{\Delta x})\|\geqslant\|(\Delta y,\overrightarrow{\Delta x})\|$ (по неравенству треугольника для норм), где $\|\overrightarrow{\Delta x}\|$ — норма в $\mathbb{R}^n$, а три последующие нормы — в $\mathbb{R}^{1+n}$.

Если взять $y=y(\vec{x})$, то $\Delta F=F(y(\vec{x}),\vec{x})-F(y^0,\vec{x}^0)=0-0=0$ и, значит,

$$F'_y(y^0,\vec{x}^0)\cdot\Delta y+\sum_{k=1}^n F'_{x_k}(y^0,\vec{x}^0)\cdot\Delta x_k+o(1)\cdot\Big(|\Delta y|+\|\overrightarrow{\Delta x}\|\Big)=0.$$

Как уже установлено, если $\overrightarrow{\Delta x}\to 0$, то $\Delta y=y(\vec{x})-y^0\to 0$ и, значит, $o(1)\to 0$ при $\overrightarrow{\Delta x}\to 0$. Будем брать приращение $\overrightarrow{\Delta x}$ столь малым, что $|o(1)|\leqslant\frac{1}{2}\left|F'_y(y^0,\vec{x}^0)\right|$. Тогда

$$\Delta y=\frac{-\sum_{k=1}^n F'_{x_k}(y^0,\vec{x}^0)\cdot\Delta x_k+o(1)\cdot\|\overrightarrow{\Delta x}\|}{F'_y(y^0,\vec{x}^0)+0(1)}=\sum_{k=1}^n O(1)\cdot\Delta x_k+o(1)\cdot\|\overrightarrow{\Delta x}\|=O(1)\cdot\|\overrightarrow{\Delta x}\|.$$

Следовательно, в предшествующем равенстве член

$$o(1)\cdot\left(|\Delta y|+\|\overrightarrow{\Delta x}\|\right)=o(1)\cdot\|\overrightarrow{\Delta x}\|,$$

где последнее $o(1)\to 0$ при $\overrightarrow{\Delta x}\to o$, т.е.

$$F'_y(y^0,\vec{x}^{\,0})\cdot\Delta y+\sum_{k=1}^{n}F'_{x_k}(y^0,\vec{x}^{\,0})\cdot\Delta x_k+o(1)\cdot\|\overrightarrow{\Delta x}\|=0.$$

Отсюда получаем, что

$$\Delta y=\sum_{k=1}^{n}-\frac{F'_{x_k}(y^0,\vec{x}^{\,0})}{F'_y(y^0,\vec{x}^{\,0})}\cdot\Delta x_k+o(\|\overrightarrow{\Delta x}\|,$$

т.е. $y(\vec{x})$ дифференцируема в точке $\vec{x}^{\,0}$ и

$$\boldsymbol{d}y(\overrightarrow{\Delta x})\big|_{\vec{x}^{\,0}}=\sum_{k=1}^{n}-\frac{F'_{x_k}(y^0,\vec{x}^{\,0})}{F'_y(y^0,\vec{x}^{\,0})}\cdot\Delta x_k.$$

Если $F(y,\vec{x})$ дифференцируема в окрестности точки $(y^0,\vec{x}^{\,0})$, то для каждой точки $(y(\vec{x}),\vec{x})$ из некоторой окрестности точки $(y^0,\vec{x}^{\,0})$ выполняются те же условия, что и для точки $(y^0,\vec{x}^{\,0})$, и, значит, $y(\vec{x})$ при это будет дифференцируема в точке $\vec{x}$. Отсюда и из непрерывности $y(\vec{x})$ следует, что существует такая окрестность точки $\vec{x}^{\,0}$, на которой $y(\vec{x})$ дифференцируема. ▲

Следствие 1. *При предположениях теоремы с дополнительным требованием дифференцируемости $F(y,\vec{x})$ в точке $(y^0,\vec{x}^{\,0})$ (в окрестности точки $(y^0,\vec{x}^{\,0})$) неявная функция $y(\vec{x})$ имеет в точке $\vec{x}^{\,0}$ (в некоторой окрестности точки $\vec{x}^{\,0}$) частные производные*

$$\frac{\partial y(\vec{x})}{\partial x_k}=-\frac{F'_{x_k}(y,\vec{x})}{F'_y(y,\vec{x})}$$

в точке $\vec{x}=\vec{x}^{\,0}$, $y=y(\vec{x})=y(\vec{x}^{\,0})$ (в окрестности точки $\vec{x}^{\,0}$, $y=y(\vec{x})$).

▼ Следствие сразу следует из вида дифференциала $\boldsymbol{d}y(\overrightarrow{\Delta x})$. ▲

Следствие 2. *Если $F(y,\vec{x})$ удовлетворяет предположениям теоремы и непрерывно дифференцируема в точке $(y^0,\vec{x}^{\,0})$ (в окрестности точки $(y^0,\vec{x}^{\,0})$), то неявная функция $y(\vec{x})$ непрерывно дифференцируема в точке $\vec{x}^{\,0}$ (в некоторой окрестности точки $\vec{x}^{\,0}$).*

▼ Действительно, по следствию 1 частная производная $\frac{\partial y(\vec{x})}{\partial x_k}$ непрерывна в точке $\vec{x}^{\,0}$ (в окрестности точки $\vec{x}^{\,0}$) как отношение двух непрерывных функций. ▲

ДОБАВЛЕНИЕ

Теорема о неявных функциях

Приведём для ознакомления только формулировку теоремы о неявных функциях. Доказывать её не будем.

Теорема 3 (о неявных функциях). *Пусть m функций $F_i(\vec{y},\vec{x})$, $i=1,\ldots,m$, от $m+n$ переменных $(\vec{y},\vec{x})$, $\vec{y}\in\mathbb{R}^m$, $\vec{x}\in\mathbb{R}^n$, непрерывны в некоторой окрестности точки $(\vec{y}^{\,0},\vec{x}^{\,0})\in\mathbb{R}^{m+n}$ и имеют в этой окрестности непрерывные частные производные $\frac{\partial F_i(\vec{y},\vec{x})}{\partial y_j}$, $i=1,\ldots,m$, $j=1,\ldots,m$. Тогда если в точке $(\vec{y}^{\,0},\vec{x}^{\,0})$ все функции обращаются в нуль, $F_i(\vec{y}^{\,0},\vec{x}^{\,0})=0$, $i=1,\ldots,m$, а определитель, называемый **определителем Якоби** или **якобианом**,*

$$\det\left(\frac{\partial F_i(\vec{y}^{\,0},\vec{x}^{\,0})}{\partial y_j}\right)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant m}}\neq 0,$$

то

$$\exists\,\varepsilon_j^0>0,\ j=1,\ldots,m,\ \forall\varepsilon_j\in(0,\varepsilon_j^0),\ j=1,\ldots,m,\ \exists B_\delta(\vec{x}^{\,0})\subset\mathbb{R}^n\ \forall\vec{x}\in B_\delta(\vec{x}^{\,0})$$
$$\exists!\,\vec{y}=(y_1,\ldots,y_m)\in\prod_{j=1}^m B_{\varepsilon_j}(y_j^0):\ F_i(\vec{y},\vec{x})=F_i(y_1,\ldots,y_m,\vec{x})=0,\ i=1,\ldots,m,$$

то есть для любых достаточно малых чисел $\varepsilon_j>0$, $j=1,\ldots,m$, найдется такая окрестность $B_\delta(\vec{x}^{\,0})$ точки $\vec{x}^{\,0}$, что в пределах этой окрестности существуют единственные m функций $y_j(\vec{x})$, $j=1,\ldots,m$, удовлетворяющие условиям $y_j(\vec{x})\in B_{\varepsilon_j}(y_j^0)$, $j=1,\ldots,m$, и являющиеся решением системы уравнений $F_i(\vec{y},\vec{x})=F_i(y_1,\ldots,y_m,\vec{x})=0$, $i=1,\ldots,m$, причем эти функции непрерывны в некоторой окрестности точки $\vec{x}^{\,0}$.

Если дополнительно потребовать дифференцируемость $F_i(\vec{y},\vec{x})$, $i=1,\ldots,m$, в точке $(\vec{y}^{\,0},\vec{x}^{\,0})$ (в некоторой ее окрестности), то функции $y_j(\vec{x})$, $j=1,\ldots,m$, будут дифференцируемы в точке $\vec{x}^{\,0}$ (в некоторой ее окрестности). Причем, если $F_i(\vec{y},\vec{x})$, $i=1,\ldots,m$, непрерывно дифференцируемы в точке $(\vec{y}^{\,0},\vec{x}^{\,0})$ (в некоторой ее окрестности), то функции $y_j(\vec{x})$, $j=1,\ldots,m$, будут непрерывно дифференцируемы в точке $\vec{x}^{\,0}$ (в некоторой ее окрестности).