Математический анализ, Коллоквиум 4

Балюк Игорь
@lodthe, GitHub
Материалы предоставил Егор Косов.

Дата изменения: 2020.09.01 в 20:42

Содержание

Исходный код предоставил Егор Косов. В данном файле я попытался исправить опечатки и облегчить некоторые моменты для понимания.

Оригинальный список вопросов

1 Метрические и нормированные пространства. Скалярное произведение и евклидово пространство. Неравенство Коши-Буняковского. Сходимость в метрических пространствах, открытые и замкнутые множества, предельные точки. Открытость открытого шара. Эквивалентное описание замкнутых множеств.

Оригинальный конспект.

1.1 Метрические и нормированные пространства.

Определение. Пусть X — множество. Функция $d : X \times X \to [0; +\infty)$ называется метрикой, если

1. $d(x, y) = 0 \iff x = y$;
2. $d(x, y) = d(y, x) \forall x, y \in X$;
3. $d(x, z) \leqslant d(x, y) + d(y, z) \forall x, y, z \in X$.

Пара (X, d) называется метрическим пространством.

Определение. Пусть X — линейное ($=$ векторное) пространство. Функция $\|\cdot\| : X \to [0; +\infty)$ называется нормой, если:

1. $\|x\| = 0 \iff x = 0$;
2. $\|\lambda x\| = |\lambda| \|x\|, \forall x \in X$;
3. $\|x + y\| \leqslant \|x\| + \|y\| \forall x, y \in X$.

Пара $(X, \|\cdot\|)$ называется нормированным пространством.

Нормой является привычной нам длина вектора. Аналогично метрике, мы будем часто работать с Евклидовой нормой: пусть $x \in \mathbb{R}^n$, тогда $\|x\| = \sqrt{x_1^2 + \dots + x_n^2}$.

Всякое нормированное пространство является метрическим с метрикой $d(x, y) = \|x - y\|$.

1.2 Скалярное произведение и евклидово пространство.

Определение. Пусть X — линейное ($=$ векторное) пространство. Функция $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{R}$ называется скалярным произведением, если для всех $x, y, z \in X$ и всех $a, b \in \mathbb{R}$ выполнены следующие условия:

1. $\langle x, x \rangle \geqslant 0$ и $\langle x, x \rangle = 0 \iff x = 0$;
2. $\langle x, y \rangle = \langle y, x \rangle$;
3. $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$.

Линейное пространство X со скалярным произведением называется Евклидовым.

Мы будем часто работать со следующим скалярным произведением: пусть $x, y \in \mathbb{R}^n$, тогда $\langle x, y \rangle = x_1 \cdot y_1 + \dots + x_n \cdot y_n$.

1.3 Неравенство Коши-Буняковского

Лемма. (Неравенство Коши-Буняковского) Пусть $\langle \cdot, \cdot \rangle$ скалярное произведение на линейном пространстве X, тогда $\forall x, y \in X$

$$|\langle x, y \rangle| \leqslant \sqrt{\langle x, x \rangle} \cdot \sqrt{\langle y, y \rangle}.$$

Доказательство. Заметим, что для $\lambda \in \mathbb{R}$ выполнено

$$0 \leqslant \langle x + \lambda y, x + \lambda y \rangle = \lambda^2 \langle y, y \rangle + 2\lambda \langle x, y \rangle + \langle x, x \rangle.$$

Не ограничивая общности, считаем, что $\langle y, y \rangle > 0$ (иначе y — нулевой вектор, доказательство тривиально). Это означает, что ветви параболы смотрят вверх. Но парабола лежит не ниже оси Ox, поэтому дискриминант этого трехчлена неположителен, т.е. $4|\langle x, y \rangle|^2 - 4\langle y, y \rangle \langle x, x \rangle \leqslant 0$. Откуда получаем:

$$|\langle x, y \rangle|^2 \leqslant \langle y, y \rangle \langle x, x \rangle \implies |\langle x, y \rangle| \leqslant \sqrt{\langle x, x \rangle} \cdot \sqrt{\langle y, y \rangle}.$$

∎

Следствие. На евклидовом пространстве функция $\|x\| := \sqrt{\langle x, x\rangle}$ является нормой.

Доказательство. Первые два свойства следуют из определения скалярного произведения. Неравенство треугольника следует из неравенства Коши-Буняковского:

$$\|x+y\|^2 = \langle x+y, x+y\rangle \leqslant \|x\|^2 + 2\cdot|\langle x,y\rangle| + \|y\|^2 \leqslant \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

■

Пример. На линейном пространстве $\mathbb{R}^k$ всех упорядоченных наборов $(x_1, \ldots, x_k)$ задано скалярное произведение $\langle x, y\rangle := \sum\limits_{j=1}^{k} x_j y_j$. Тем самым, на $\mathbb{R}^k$ задана естественная евклидова метрика $\|x - y\| := \sqrt{|x_1 - y_1|^2 + \cdots + |x_k - y_k|^2}$.

1.4 Сходимость в метрических пространствах, открытые и замкнутые множества, предельные точки.

Определение. Пусть (X, d) метрическое пространство.

1. Множество
$$B_r(x_0) := \{x \in X \mid d(x, x_0) < r\}$$
называется **открытым шаром** радиуса r.

2. Множество
$$\overline{B_r}(x_0) := \{x \in X \mid d(x, x_0) \leqslant r\}$$
называется **замкнутым шаром** радиуса r.

3. Последовательность точек $x_n \in X$ называется **сходящейся к точке** x, если для всякого $\varepsilon > 0$ найдется такой номер $N(\varepsilon)$, что $d(x, x_n) < \varepsilon$ для каждого $n \geqslant N(\varepsilon)$.

4. Последовательность точек $x_n \in X$ называется **фундаментальной**, если для всякого $\varepsilon > 0$ найдется такой номер $N(\varepsilon)$, что $d(x_k, x_n) < \varepsilon$ для всех $k, n \geqslant N(\varepsilon)$.

5. Точка x называется **предельной** для множества $M \subset X$, если для всякого $\varepsilon > 0$ выполнено $B_\varepsilon(x) \cap (M \setminus \{x\}) \neq \varnothing$.

6. Множество $U \subset X$ называется **открытым**, если для всякого $x \in U$ найдется такое $\varepsilon > 0$, что $B_\varepsilon(x) \subset U$.

7. Множество $F \subset X$ называется **замкнутым**, если множество $X \setminus F$ открыто.

1.5 Открытость открытого шара. Эквивалентное описание замкнутых множеств.

Лемма. Пусть (X, d) метрическое пространство. Тогда

1. если $x_n \to x, y_n \to y$, то $d(x_n, y_n) \to d(x, y)$;

2. предел сходящейся последовательности единственный;

3. любой открытый шар является открытым множеством;

4. множество F замкнуто тогда и только тогда, когда множество F содержит все свои предельные точки.

Доказательство.

1. Следует из оценки
$$|d(x_n, y_n) - d(x, y)| \leqslant |d(x_n, y_n) - d(x_n, y)| + |d(x_n, y) - d(x, y)| \leqslant d(y_n, y) + d(x_n, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$
Для всякого $\varepsilon > 0$, для каждого $n \geqslant N$, для некоторого N.

2. Следует из пункта 1). Предположим обратное. Пусть $\{x_n\} \to a$ и $\{x_n\} \to b$, при этом $a \neq b$. Возьмем некоторые непересекающиеся окрестности $U = U(a)$ и $V = V(b)$ точек a и b соответственно. Согласно определению предела вне окрестности U точки , в частности в окрестности V точки b, содержится лишь конечное число членов последовательности $\{x_n\}$. Однако точка b также является ее пределом, и потому в ее окрестности V должны находиться бесконечно много членов последовательности $\{x_n\}$, начиная с некоторого номера, а следовательно, получилось противоречие.

3.

Определение. Множество

$$B_r(x_0) := \{x \in X \mid d(x, x_0) < r\}$$

называется **открытым шаром** радиуса r.

Если $x \in B_r(x_0)$, то по неравенству треугольника $B_\varepsilon(x) \subset B_r(x_0)$ при $\varepsilon < r - d(x, x_0)$. Докажем это.

Заметим, что для для всех $t \in B_\varepsilon(x)$ мы имеем $d(x_0, t) \leqslant d(x_0, x) + d(x, t) < d(x_0, x) + (r - d(x, x_0)) = r$. Отсюда получили, что $t \in B_r(x_0)$.

4. Множество F замкнуто тогда и только тогда, когда $\forall x \notin F\ \exists \varepsilon > 0 :\ B_\varepsilon \cap F = \varnothing \iff$ всякая точка $x \notin F$ — не предельная для F.

■

2 Полные метрические пространства, полнота $\mathbb{R}^k$. Непрерывные отображения в метрических пространствах: определения, доказательства их эквивалентностей, основные свойства.

2.1 Полные метрические пространства, полнота $\mathbb{R}^k$.

Определение. Метрическое пространство называется полным, если каждая фундаментальная последовательность в нем сходится.

Замечание. На $\mathbb{R}^k$ справедливы соотношения

$$\max_{1 \leqslant j \leqslant k} |x_j| \leqslant \|x\| \leqslant \sqrt{k} \cdot \max_{1 \leqslant j \leqslant k} |x_j|$$

для вектора $x = (x_1, \ldots, x_k)$.

Теорема. Для любого конечномерного пространства $\mathbb{R}^m$ последовательность $\{x_n\} \xrightarrow{n\to\infty} a$ тогда и только тогда, когда для всякого i $\{(x_n)_i\} \xrightarrow{n\to\infty} a_i$.

Доказательство.

- **Необходимость**. По определению

$$\forall \varepsilon > 0\ \exists N(\varepsilon) > 0 :\ \forall n \geqslant N(\varepsilon) \implies \sum_{i=1}^{m} ((x_n)_i - a_i)^2 < \varepsilon^2 \implies |(x_n)_i - a_i| < \varepsilon\ \forall i,$$

а это означает, что $\{(x_n)_i\} \xrightarrow{n\to\infty} a_i$.

- **Достаточность**. По определению для всякого i и всякого $\varepsilon > 0$ существует $N_i(\varepsilon) :\ \forall n \geqslant N_i(\varepsilon) \implies |(x_n)_i - a_i| < \dfrac{\varepsilon}{\sqrt{m}}$. Но если выбрать $N = \max\{N_i \mid 1 \leqslant i \leqslant m\}$, то

$$\forall n \geqslant N \implies \sum_{i=1}^{m} ((x_n)_i - a_i)^2 < m \cdot \frac{\varepsilon^2}{m} = \varepsilon^2 \iff d(x_n, a) < \varepsilon.$$

■

Пример. Пространство $\mathbb{R}^k$ со стандартной евклидовой метрикой полное. Действительно. если последовательность векторов $x_n \in \mathbb{R}^k$ фундаментальна, то фундаментальны и последовательности координат $\{(x_n)_j\}_{j=1}^{\infty}$ для всякого $j \in \{1, \ldots, k\}$.

Тем самым, у j-ой координаты есть предел x_j для каждого $j \in \{1, \ldots, k\}$. То есть $|(x_n)_j - x_j| \to 0$. Значит, $x_n \to x := (x_1, \ldots, x_k)$. В доказательстве можно сослаться на одномерный случай.

Пример. Пусть $X = [0; \pi/2)$. Пространство X не является полным с метрикой $d_1(x, y) = |x - y|$, но является полным с метрикой $d_2(x, y) = |\operatorname{tg} x - \operatorname{tg} y|$.

Доказательство. Докажем, что пространство X не является полным с метрикой $d_1(x, y) = |x - y|$. Возьмем следующую последовательность:

$$x_1 = \frac{\pi}{2} - 1, \quad x_2 = \frac{\pi}{2} - \frac{1}{2}, \quad \ldots, \quad x_n = \frac{\pi}{2} - \frac{1}{n}, \quad \ldots$$

Тогда данная последовательность является фундаментальной, так как $|x_n - x_m| = \left|\dfrac{1}{n} - \dfrac{1}{m}\right| \leqslant \dfrac{1}{\min\{n, m\}}$, но она сходится к $\dfrac{\pi}{2}$, а оно не лежит в X. ■

2.2 Непрерывные отображения в метрических пространствах: определения, доказательства их эквивалентностей, основные свойства.

Определение. Пусть (X, d_X) и (Y, d_Y) — два метрических пространства. Отображение $f : X \to Y$ называется непрерывным в точке $x_0 \in X$, если для всякой последовательности $x_n \to x_0$ выполнено $f(x_n) \to f(x_0)$.

Лемма. Пусть (X, d_X) и (Y, d_Y) — два метрических пространства.

1. Отображение $f : X \to Y$ является непрерывным в точке $x \in X$ тогда и только тогда, когда для всякого $\varepsilon > 0$ найдется $\delta > 0$ такое, что $d_Y(f(x), f(x_0)) < \varepsilon$, если $d_X(x, x_0) < \delta$.
2. Отображение $f : X \to Y$ является непрерывным в каждой точке $x \in X$ тогда и только тогда, когда прообраз каждого открытого множества в Y будет открытым множеством в X (такие отображения будем называть просто непрерывными).

Доказательство.

1. Отображение f разрывно в точке $x_0 \iff$ найдется последовательность $x_n \to x_0$, для которой $f(x_n)$ не сходится к $f(x_0) \iff$ найдется число $\varepsilon > 0$ и последовательность $x'_n \to x_0$, для которой $d_Y(f(x'_n), f(x_0)) \geqslant \varepsilon \iff$ найдется такое число $\varepsilon > 0$, что для произвольного $\delta > 0$ существует $x_\delta \in B_\delta(x_0)$, для которого $d_Y(f(x_\delta), f(x_0)) \geqslant \varepsilon$.
2. Если прообраз любого открытого множества открыт, то для произвольного $\varepsilon > 0$ найдется такое $\delta > 0$, что $f^{-1}(B_\varepsilon(f(x_0))) \supset B_\delta(x_0)$, и значит отображение f непрерывно в точке x_0. Наоборот: пусть U — открыто в Y и $x_0 \in f^{-1}(U)$. Тогда в силу открытости найдется $\varepsilon > 0$, для которого $B_\varepsilon(f(x_0)) \subset U$. Из-за непрерывности в точке x_0 найдется такое $\delta > 0$, что $f^{-1}(B_\varepsilon(f(x_0))) \supset B_\delta(x_0)$, что дает открытость множества $f^{-1}(U)$.

■

Предложение. Пусть $f : X \to Y$ непрерывна в точке $a \in X$, $g : Y \to Z$ непрерывна в точке $f(a) \in Y$. Тогда композиция $g \circ f : X \to Z$ непрерывна в точке a.

Доказательство. Следует из определения непрерывности. ■

Следствие. Пусть $f, g : \mathbb{R}^k \to \mathbb{R}^m$ — непрерывные в точке a функции. Тогда $f + g$ и $f \cdot g$ — непрерывны в точке a.

Доказательство. Следует из свойства пределов:

$$\lim_{x \to x_0} (f(x) + g(x)) = \lim_{x \to x_0} f(x) + \lim_{x \to x_0} g(x) = f(x_0) + g(x_0) = (f + g)(x_0).$$

■

Определение. Пусть (X, d_X) и (Y, d_Y) — метрические пространства и пусть x_0 — предельная точка в X. Скажем, что предел функции $f : X \to Y$ в точке x_0 равен y_0, если функция g, определенная соотношением $g(x) = f(x)$ при $x \neq x_0$ и $g(x_0) = y_0$ иначе, непрерывна в точке x_0.

3 Компакты в метрических пространствах: определение и основные свойства. Образ компакта при непрерывном отображении. Критерий компактности в $\mathbb{R}^k$. Свойства непрерывных на компакте функций.

3.1 Компакты в метрических пространствах: определение и основные свойства. Образ компакта при непрерывном отображении.

Определение. Множество K в метрическом пространстве называется компактным тогда и только тогда, когда из произвольной последовательности $\{x_n\}_{n=1}^{\infty} \subset K$ можно выделить сходящуюся подпоследовательность $x_{n_k} \to x \in K$.

Лемма. Пусть K — компакт. Тогда

1. K — ограниченное множество;
2. K — замкнутое множество;
3. образ K при непрерывном отображении компактен.

Доказательство.

1. Зафиксируем произвольную точку $x_0 \in K$. Если K — неограниченное множество, то найдется последовательность $x_n \in K$, $d(x_n, x_0) \to \infty$. Переходя к подпоследовательности, имеем $x_{n_k} \to x$, $d(x_{n_k}, x_0) \to d(x, x_0)$. Противоречие.

2. Если $x_n \in K$, $x_n \to x_0$, то переходя к подпоследовательности $x_{n_k} \to x \in K$, в силу единственности предела $x_0 = x \in K$.

3. Рассмотрим последовательность $\{f(x_n)\}_{n=1}^{\infty}$, $x_n \in K$. Переходя к подпоследовательности имеем $x_{n_k} \to x \in K$. Так как f — непрерывное отображение, то $f(x_{n_k}) \to f(x) \in f(K)$.

■

3.2 Критерий компактности в $\mathbb{R}^k$.

Предложение. Множество K в $\mathbb{R}^k$ со стандартной евклидовой метрикой компактно тогда и только тогда, когда оно замкнуто и ограничено.

Доказательство. Необходимость этого условия следует из предыдущей леммы. Проверим достаточность. Пусть множество K — замкнуто и ограничено, и пусть $x_n \in K$. В силу ограниченности K ограниченными будут и все координаты $(x_n)_j$ последовательности x_n. Тогда найдется сходящаяся подпоследовательность первых координат $(x_{n_m})_1$. Далее, из последовательности $(x_{n_m})_2$ можно также извлечь сходящуюся подпоследовательность. Повторяя процедуру, получим подпоследовательность x'_n, у которой каждая координата сходится, то есть $(x'_n)_j \to x_j$ для некоторого x_j. Тем самым, $x'_n \to x = (x_1, \dots, x_k)$. В силу замкнутости K, вектор $x \in K$. ■

3.3 Свойства непрерывных на компакте функций.

Следствие. Пусть K — компакт, $f : K \to \mathbb{R}$ — непрерывная функция. Тогда образ $f(K)$ — ограниченное множество и найдутся точки $x_m, x_M \in K$, для которых $f(x_m) = \inf\limits_{x \in K} f(x), f(x_M) = \sup\limits_{x \in K} f(x)$.

4 Дифференцируемость отображений из $\mathbb{R}^k$ в $\mathbb{R}^m$, дифференциал. Непрерывность дифференцируемых отображений. Производная вдоль вектора и ее связь с дифференциалом. Частные производные.

4.1 Дифференцируемость отображений из $\mathbb{R}^k$ в $\mathbb{R}^m$, дифференциал.

Определение. Отображение $f : \mathbb{R}^k \to \mathbb{R}^m$ называется дифференцируемым в точке x, если для каждого $h \in \mathbb{R}^k$

$$f(x+h) = f(x) + Lh + \alpha(h) \|h\|,$$

где $L : \mathbb{R}^k \to \mathbb{R}^m$ — линейное отображение, $\lim\limits_{\|h\| \to 0} \|\alpha(h)\| = 0$. Линейное отображение L называют дифференциалом f в точке x и обозначают df.

Замечание. Напомним, что отображение $L : \mathbb{R}^k \to \mathbb{R}^m$ называется линейным, если

$$L(a_1 h_1 + a_2 h_2) = a_1 L h_1 + a_2 L h_2$$

для произвольных векторов $h_1, h_2, \in \mathbb{R}^k$ и произвольных чисел $a_1, a_2 \in \mathbb{R}$.

Если в $\mathbb{R}^k$ фиксирован базис $e := \{e_1, \dots, e_k\}$, а в $\mathbb{R}^m$ фиксирован $e' := \{e'_1, \dots, e'_m\}$, то линейное отображение L представимо в виде $L(h) = L(e_1)h_1 + \dots + L(e_k)h_k$, где $h = (h_1, \dots, h_k)$ в базисе e, а векторы $L(e_j) = (a_{1,j}, \dots, a_{m,j})$ в базисе e'.

В частности, каждое линейное отображение при фиксированных базисах e и e' в $\mathbb{R}^k$ и $\mathbb{R}^m$ соответственно записывается с помощью матрицы $A = (a_{ij})$. Кроме того,

$$\|Lh\| \leqslant (\|L(e_1)\| + \dots + \|L(e_k)\|) \cdot \max_{1 \leqslant j \leqslant k} |h_j| \leqslant C \|h\|$$

и каждое линейное отображение непрерывно на $\mathbb{R}^k$.

4.2 Непрерывность дифференцируемых отображений.

Следствие. Если отображение $f : \mathbb{R}^k \to \mathbb{R}^m$ дифференцируемо в точке x, то оно непрерывно в точке x.

Доказательство. Действительно, $\|f(x+h) - f(x)\| = \|df(h) + \alpha(h) \|h\|\| \leqslant C \|h\|$ (перенесли $f(x)$ влево, взяли норму от обеих частей) при h из некоторой окрестности нуля. ■

Замечание. Так как дифференцируемость f в точке x равносильна тому, что

$$\lim_{h\to 0}\frac{\|f(x+h)-f(x)-Lh\|}{\|h\|}=0,$$

и так как сходимость по норме равносильна покоординатной сходимости, то при фиксированном базисе $e' := \{e'_1,\dots,e'_m\}$ в $\mathbb{R}^m$ дифференцируемость отображения f равносильна дифференцируемости каждой координаты f_j в точке x. В этом случае $Lh=(L_1h,\dots,L_mh)$ в базисе e', где $L_j=df_j$ — дифференциал j-ой координаты.

Лемма. Если отображение $f:\mathbb{R}^k\to\mathbb{R}^m$ дифференцируемо в точке x, то для каждого вектора $h\in\mathbb{R}^k$ функция $t\to f(x+th)$ дифференцируема в точке 0 и $\dfrac{d}{dt}f(x+th)\Big|_{t=0}=df(h)$.

Доказательство. По определению

$$f(x+th)-f(x)=t\,df(h)+\alpha(th)\cdot|t|\,\|h\|\,.$$

Разделив на t и перейдя к пределу при $t\to 0$, получаем требуемое соотношение.

$$\lim_{t\to 0}\frac{f(x+th)-f(x)}{t}=\lim_{t\to 0}(df(h)+\alpha(th)\,\|h\|)=df(h).$$

Мы явно нашли df, поэтому функция дифференцируема по определению. ∎

4.3 Производная вдоль вектора и ее связь с дифференциалом.

Определение. Производная $\dfrac{\partial f}{\partial h}(x):=\dfrac{d}{dt}f(x+th)\Big|_{t=0}$ называется производной вдоль вектора h и может существовать и в случае, когда сама функция f не дифференцируема в точке x.

Как мы уже поняли, для дифференцируемости отображения достаточно исследовать дифференцируемость его координат, то есть дифференцируемость функции $f:\mathbb{R}^k\to\mathbb{R}$. Зафиксировав базис $e:=\{e_1,\dots,e_k\}$ в $\mathbb{R}^k$, условие дифференцируемости в точке $x=(x_1,\dots,x_k)$ переписывается в виде

$$f(x_1+h_1,\dots,x_k+h_k)=f(x_1,\dots,x_k)+c_1h_1+\dots+c_kh_k+\bar{\bar{o}}(\|h\|),$$

то есть $df(h)=c_1h_1+\dots+c_kh_k$. Из уже доказанного ясно, что $\dfrac{\partial f}{\partial e_j}(x)=df(e_j)=c_j$.

4.4 Частные производные.

Определение. Частной производной $\dfrac{\partial f}{\partial x_j}$ функции $f:\mathbb{R}^k\to\mathbb{R}$ в точке $x=(x_1,\dots,x_k)$ называется производная вдоль вектора e_j, то есть

$$\frac{\partial f}{\partial x_j}(x)=\frac{d}{dt}f(x_1,\dots,x_{j-1},t,x_{j+1},\dots,x_k)\Big|_{t=x_j}.$$

Замечание. При фиксированно базисе $e=\{e_1,\dots,e_k\}$ в $\mathbb{R}^k$ линейные функционалы $dx_1,\dots,dx_k$ оказываются сопряженным базисом к e. То есть $dx_i(e_j)=\delta_{i,j}$. Таким образом, $df=\dfrac{\partial f}{\partial x_1}\,dx_1+\dots+\dfrac{\partial f}{\partial x_k}\,dx_k$.

Замечание. В случае отображения $f:\mathbb{R}^k\to\mathbb{R}^m$ при фиксированных базисах e и e' в $\mathbb{R}^k$ и в $\mathbb{R}^m$ соответственно, компоненты матрицы дифференциала df имеют вид $a_{i,j}=\dfrac{\partial f_i}{\partial x_j}(x)$, то есть по строкам написаны градиенты $\nabla f_i(x)$.

5 Градиент функции и матрица Якоби отображения. Градиент, как направление наибольшего роста функции. Достаточное условие дифференцируемости функции в точке.

5.1 Градиент функции и матрица Якоби отображения.

Определение. При фиксированных базисах e в $\mathbb{R}^k$ и e' в $\mathbb{R}^m$ матрицу, соответствующую линейному отображению df, называют матрицей Якоби отображения f в точке x и обозначают $J_f(x)$.

Определение. *Градиентом* функции f называется вектор $\nabla f:=\left(\dfrac{\partial f}{\partial x_1},\dots,\dfrac{\partial f}{\partial x_n}\right)$.

5.2 Градиент, как направление наибольшего роста функции.

Лемма. Если f дифференцируема в точке x и $df \neq 0$, то наибольшее значение производной вдоль единичного вектора v (т.е. $\|v\| = 1$) достигается на векторе $\|\nabla f(x)\|^{-1}\nabla f(x)$.

Доказательство. Так как $\dfrac{\partial f}{\partial v}(x) = df(v) = \langle \nabla f(x), v\rangle$, то по неравенству Коши–Буняковского $\left|\dfrac{\partial f}{\partial v}(x)\right| \leqslant \|\nabla f(x)\|\|v\| = \|\nabla f(x)\|$. Если $v = \|\nabla f(x)\|^{-1}\nabla f(x)$, то в неравенстве достигается равенство. ■

Заметим, что наличия частных производных в точке недостаточно для дифференцируемости функции в этой точке.

Пример. Пусть

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

Функция f разрывна в нуле, а значит не дифференцируема, но в точке $(0,0)$ существуют обе частных производных. Действительно, если $x = r\cos\varphi, y = r\sin\varphi$, то функция $f(x,y) = \sin 2\varphi$. Таким образом, $f(x,y)$ в любой окрестности точки $(0,0)$ принимает значения из $[-1;1]$, но $\dfrac{\partial f}{\partial x}(0,0) = \dfrac{d}{dx}f(x,0) = 0$. Аналогично $\dfrac{\partial f}{\partial y}(0,0) = 0$.

5.3 Достаточное условие дифференцируемости функции в точке.

В следующей теореме сформулировано достаточное условие дифференцируемости.

Теорема. Если все частные производные $\dfrac{\partial f}{\partial x_j}$ существуют в окрестности точки x_0 и непрерывны в этой точке, то f — дифференцируема в точке x_0.

Доказательство. Для сокращения выкладок докажем теорему в случае $k = 2$. Заметим, что по теореме Лагранжа

$$\begin{aligned} f(x_1+h_1, x_2+h_2) - f(x_1,x_2) &= f(x_1+h_1,x_2+h_2) - f(x_1,x_2+h_2) + f(x_1,x_2+h_2) - f(x_1,x_2) \\ &= \frac{\partial f}{\partial x_1}(\xi_1, x_2+h_2)h_1 + \frac{\partial f}{\partial x_2}(x_1,\xi_2)h_2, \end{aligned}$$

где ξ_1 принадлежит интервалу с концами x_1, x_1+h_1, а ξ_2 — с концами x_2, x_2+h_2. Запишем теперь последнюю сумму в виде

$$\frac{\partial f}{\partial x_1}(x_1,x_2)h_1 + \frac{\partial f}{\partial x_2}(x_1,x_2)h_2 + \alpha(h)\|h\|,$$

где

$$\alpha(h) = \Big(\frac{\partial f}{\partial x_1}(\xi_1, x_2+h_2) - \frac{\partial f}{\partial x_1}(x_1,x_2)\Big)\frac{h_1}{\|h\|} + \Big(\frac{\partial f}{\partial x_2}(x_1,\xi_2) - \frac{\partial f}{\partial x_2}(x_1,x_2)\Big)\frac{h_2}{\|h\|}.$$

При малой $\|h\|$ выражения в скобках будут малы в силу непрерывности частных производных. Тем самым, $\lim\limits_{h\to 0}\|\alpha(h)\| = 0$, потому что $|h_1| \leqslant \|h\|$ и $|h_2| \leqslant \|h\|$ $\left(\text{потому что } \|h\| = \sqrt{h_1^2+h_2^2}\right)$. ■

6 Частные производные высоких порядков. Теоремы Шварца (б/д) и Юнга. Дифференциалы высоких порядков.

6.1 Частные производные высоких порядков.

Определение. Пусть $f\colon \mathbb{R}^k \to \mathbb{R}$ и предположим, что в некоторой окрестности $B_r(x_0)$ точки x_0 существует частная производная $\dfrac{\partial f}{\partial x_j}$. Если функция $x \mapsto \dfrac{\partial f}{\partial x_j}(x)$ в точке x_0 имеет частную производную по переменной x_i, то эта частная производная $\dfrac{\partial}{\partial x_i}\Big(\dfrac{\partial f}{\partial x_j}\Big)(x_0)$ называется *частной производной второго порядка* по переменным x_j и x_i и обозначается $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(x_0)$.

Замечание. Заметим, что частные производные $\dfrac{\partial^2 f}{\partial x_i\partial x_j}$ и $\dfrac{\partial^2 f}{\partial x_j\partial x_i}$ являются разными объектами и, вообще говоря, не совпадают (пример будет в рамках семинарских задач).

6.2 Теоремы Шварца и Юнга.

О совпадении смешанных частных производных позволяют судить следующие две теоремы, которые мы для простоты сформулируем в двумерном случае (общий случай, по сути, ничем не отличается).

Теорема 1 (Шварц)**.** *Пусть смешанные частные производные* $\frac{\partial^2 f}{\partial x \partial y}$ *и* $\frac{\partial^2 f}{\partial y \partial x}$ *существуют в окрестности точки* (x_0, y_0) *и непрерывны в этой точке. Тогда их значения в точке* (x_0, y_0) *совпадают.*

Теорема 2 (Юнг)**.** *Пусть* f *— дифференцируема в окрестности точки* (x_0, y_0)*, а ее частные производные* $\frac{\partial f}{\partial x}$ *и* $\frac{\partial f}{\partial y}$ *дифференцируемы в точке* (x_0, y_0)*. Тогда смешанные частные производные* $\frac{\partial^2 f}{\partial x \partial y}$ *и* $\frac{\partial^2 f}{\partial y \partial x}$ *в точке* (x_0, y_0) *совпадают.*

Приведем доказательство второй из этих теорем.

Доказательство. Не ограничивая общности, будем считать, что $(x_0, y_0) = (0, 0)$. Рассмотрим функцию

$$F(t, t) = f(t, t) - f(0, t) - f(t, 0) + f(0, 0).$$

Применяя теорему Лагранжа к функции $g(u) = f(t, u) - f(0, u)$, получаем

$$F(t, t) = g(t) - g(0) = g'(\xi)t = \Big(\frac{\partial f}{\partial y}(t, \xi) - \frac{\partial f}{\partial y}(0, \xi)\Big)t.$$

Дифференцируемость $\frac{\partial f}{\partial y}$ в точке $(0, 0)$ означает, что

$$\frac{\partial f}{\partial y}(x, y) = \frac{\partial f}{\partial y}(0, 0) + \frac{\partial^2 f}{\partial x \partial y}(0, 0)x + \frac{\partial^2 f}{\partial y^2}(0, 0)y + \bar{\bar{o}}(\sqrt{x^2 + y^2}).$$

Таким образом,

$$\frac{\partial f}{\partial y}(t, \xi) = \frac{\partial f}{\partial y}(0, 0) + \frac{\partial^2 f}{\partial x \partial y}(0, 0)t + \frac{\partial^2 f}{\partial y^2}(0, 0)\xi + \bar{\bar{o}}(\sqrt{t^2 + \xi^2}).$$

и

$$\frac{\partial f}{\partial y}(0, \xi) = \frac{\partial f}{\partial y}(0, 0) + \frac{\partial^2 f}{\partial y^2}(0, 0)\xi + \bar{\bar{o}}(\xi).$$

Т.к. $\xi \leqslant t$, то $\bar{\bar{o}}(\sqrt{t^2 + \xi^2}) = \bar{\bar{o}}(t)$ и $\bar{\bar{o}}(\xi) = \bar{\bar{o}}(t)$. Таким образом,

$$F(t, t) = \frac{\partial^2 f}{\partial x \partial y}(0, 0)t^2 + \bar{\bar{o}}(t^2).$$

Аналогично,

$$F(t, t) = \frac{\partial^2 f}{\partial y \partial x}(0, 0)t^2 + \bar{\bar{o}}(t^2).$$

Приравняв полученные выражения, поделив на t^2 и устремив t к нулю, получаем

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = \frac{\partial^2 f}{\partial y \partial x}(0, 0).$$

■

6.3 Дифференциалы высоких порядков.

Предположим, что $f\colon \mathbb{R}^k \to \mathbb{R}$ — дифференцируема в окрестности точки a и предположим, что ее частные производные $\frac{\partial f}{\partial x_j}$ дифференцируемы в точке a. Тогда при каждом $h \in \mathbb{R}^k$ возникает функция $x \mapsto df\big|_x(h) = \frac{\partial f}{\partial x_1}(x)h_1 + \ldots + \frac{\partial f}{\partial x_k}(x)h_k$, дифференцируемая в точке a.

Ее дифференциал $d(df(h))\big|_a(q) = \left(\sum\limits_{j=1}^{k} \frac{\partial^2 f}{\partial x_j \partial x_1}(a)q_j\right) h_1 + \ldots + \left(\sum\limits_{j=1}^{k} \frac{\partial^2 f}{\partial x_j \partial x_k}(a)q_j\right) h_k.$

То есть получена билинейная форма $d(df(h))\big|_a(q) = \sum\limits_{i,j=1}^{k} \frac{\partial^2 f}{\partial x_j \partial x_i}(a)q_j h_i$. Эта билинейная форма оказывается симметричной по теореме Юнга, а т.к. симметричная билинейная форма однозначно задается своей квадратичной формой $d(df(h))\big|_a(h) = \sum\limits_{i,j=1}^{k} \frac{\partial^2 f}{\partial x_j \partial x_i}(a)h_j h_i = \sum\limits_{i,j=1}^{k} \frac{\partial^2 f}{\partial x_j \partial x_i}(a)dx_j(h)dx_i(h)$, то эту квадратичную форму $d^2 f := \sum\limits_{i,j=1}^{k} \frac{\partial^2 f}{\partial x_j \partial x_i}(a)dx_j dx_i$ и называют **вторым дифференциалом** функции f.

Аналогично определяется дифференциал n-го порядка:

Определение. Если f — n раз дифференцируема в точке a, то

$$d^n f\big|_a := \sum_{1\leqslant j_1,\dots,j_n\leqslant k} \frac{\partial^n f}{\partial x_{j_1}\dots\partial x_{j_n}}(a)\, dx_{j_1}\dots dx_{j_n}.$$

Последняя запись означает лишь то, что при вычислении n-го дифференциала на векторе $h\in\mathbb{R}^k$ надо воспользоваться линейностью, а $[dx_{j_1}\dots dx_{j_n}](h) := dx_{j_1}(h)\dots dx_{j_n}(h) = h_{j_1}\dots h_{j_n}$.

7 Дифференциал суммы и произведения. Дифференциал обратного отображения.

7.1 Дифференциал суммы и произведения.

Теорема. Пусть функции $f,g\colon\mathbb{R}^k\to\mathbb{R}$ дифференцируемы в некоторой точке x. Тогда, для произвольных чисел $a,b\in\mathbb{R}$, функции $af+bg$ и fg дифференцируемы в точке x и $d(af+bg)=a\,df+b\,dg$ и $d(fg)=f\,dg+g\,df$.

Доказательство. Заметим, что

$$\begin{aligned}(af+bg)(x+h)-(af+bg)(x) &= a(f(x+h)-f(x))+b(g(x+h)-g(x))\\ &= a(df(h)+\bar{\bar{o}}(\|h\|))+b(dg(h)+\bar{\bar{o}}(\|h\|)) = a\,df(h)+b\,dg(h)+\bar{\bar{o}}(\|h\|).\end{aligned}$$

Таким образом, $d(af+bg)=a\,df+b\,dg$.

Для доказательства второго равенства заметим, что

$$\begin{aligned}(fg)(x+h)-(fg)(x) &= (f(x+h)-f(x))g(x+h)+f(x)(g(x+h)-g(x))\\ &= (df(h)+\bar{\bar{o}}(\|h\|))(g(x)+\bar{\bar{o}}(1))+f(x)(dg(h)+\bar{\bar{o}}(\|h\|))\\ &= g(x)\,df(h)+f(x)\,dg(h)+(df(h))\bar{\bar{o}}(1)+g(x)\bar{\bar{o}}(\|h\|)+f(x)\bar{\bar{o}}(\|h\|)+\bar{\bar{o}}(\|h\|).\end{aligned}$$

Мы использовали непрерывность функции g, т.е. $g(x+h)-g(x)=\bar{\bar{o}}(1)$ при $\|h\|\to 0$, в силу ее дифференцируемости в точке x.

Так как df — линейное отображение, то для некоторого числа $C>0$ выполнено $|df(h)|\leqslant C\|h\|$, а значит $(df(h))\bar{\bar{o}}(1)=\bar{\bar{o}}(\|h\|)$. Т.к. $f(x)$ и $g(x)$ просто числа, то $g(x)\bar{\bar{o}}(\|h\|)+f(x)\bar{\bar{o}}(\|h\|)=\bar{\bar{o}}(\|h\|)$. Таким образом, теорема доказана. ■

7.2 Дифференциал обратного отображения.

Теорема. Пусть $f\colon\mathbb{R}^k\to\mathbb{R}^k$ — есть непрерывная биекция между окрестностями $U(a)$ и $V(f(a))$, причем обратное отображение $f^{-1}\colon V(f(a))\to U(a)$ также непрерывно (т.е. f — гомеоморфизм между $U(a)$ и $V(f(a))$).

Предположим, что f — дифференцируемо в точке a и df — обратимое линейное отображение. Тогда f^{-1} — дифференцируемо в точке $f(a)$ и $df^{-1}\big|_{f(a)}=\big(df\big|_a\big)^{-1}$.

Доказательство. Нам нужно проверить, что

$$\lim_{\|q\|\to 0}\frac{\|f^{-1}(f(a)+q)-f^{-1}(f(a))-(df)^{-1}(q)\|}{\|q\|}=0.$$

Пусть $h=f^{-1}(f(a)+q)-f^{-1}(f(a))=f^{-1}(f(a)+q)-a$, тогда $q=f(a+h)-f(a)$ и $\|q\|\to 0$ тогда и только тогда, когда $\|h\|\to 0$.

Так как f — дифференцируемо в точке a, то

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

где $\lim\limits_{\|h\|\to 0}\|\alpha(h)\|=0$.

Таким образом,

$$\lim_{\|q\|\to 0}\frac{\|f^{-1}(f(a)+q)-f^{-1}(f(a))-(df)^{-1}(q)\|}{\|q\|}=\lim_{\|h\|\to 0}\frac{\big\|h-(df)^{-1}(df(h)+\alpha(h)\|h\|)\big\|}{\big\|df(h)+\alpha(h)\|h\|\big\|}.$$

Числитель в последнем выражении равен $\|h\|\big\|(df)^{-1}(\alpha(h))\big\|$. Для линейного отображения $(df)^{-1}$ найдется число $C>0$, для которого $\|(df)^{-1}(p)\|\leqslant C\|p\|, \forall p\in\mathbb{R}^k$.

Отсюда, подставив $p=df(h)$, получаем $C^{-1}\|h\|\leqslant\|\,df(h)\|$. Тем самым

$$\big\|df(h)+\alpha(h)\|h\|\big\|\geqslant\|\,df(h)\|-\|h\|\|\alpha(h)\|\geqslant\|h\|\big(C^{-1}-\|\alpha(h)\|\big).$$

Таким образом,

$$\frac{\big\|h-(df)^{-1}(df(h)+\alpha(h)\|h\|)\big\|}{\big\|df(h)+\alpha(h)\|h\|\big\|}\leqslant\frac{C\|h\|\|\alpha(h)\|}{\|h\|\big(C^{-1}-\|\alpha(h)\|\big)}=\frac{C\|\alpha(h)\|}{\big(C^{-1}-\|\alpha(h)\|\big)}\to 0$$

при $\|h\|\to 0$. ■

Замечание. Отметим, что матрица обратного линейного отображения есть обратная матрица к матрице исходного линейного отображения. Тем самым, матрица Якоби обратного отображения $J_{f^{-1}}(y)$ является обратной к матрице Якоби исходного отображения, т.е. равна $\big(J_f(f^{-1}(y))\big)^{-1}$.

8 Дифференциал композиции. Матрица Якоби композиции, правило вычисления частной производной сложной функции, инвариантность первого дифференциала.

8.1 Дифференциал композиции.

Теорема. Пусть $f\colon\mathbb{R}^k\to\mathbb{R}^m$, $g\colon\mathbb{R}^m\to\mathbb{R}^n$, причем отображение f дифференцируемо в точке a, отображение g дифференцируемо в точке $f(a)$. Тогда отображение $g\circ f$ дифференцируемо в точке a и $d(g\circ f)\big|_a=dg\big|_{f(a)}\circ df\big|_a$.

Замечание. Поясним запись $d(g\circ f)\big|_a=dg\big|_{f(a)}\circ df\big|_a$. Здесь $df\big|_a\colon\mathbb{R}^k\to\mathbb{R}^m$ есть линейное отображение и $dg\big|_{f(a)}\colon\mathbb{R}^m\to\mathbb{R}^n$ есть линейное отображение. Тогда их композиция $dg\big|_{f(a)}\circ df\big|_a\colon\mathbb{R}^k\to\mathbb{R}^n$ есть линейное отображение, действующее по правилу

$$dg\big|_{f(a)}\circ df\big|_a(h)=dg\big|_{f(a)}(df\big|_a(h)).$$

Доказательство. По условию $f(a+h)-f(a)=df(h)+\alpha(h)\|h\|$, где $\lim\limits_{\|h\|\to 0}\|\alpha(h)\|=0$ и $g(f(a)+q)-g(f(a))=dg(q)+\beta(q)\|q\|$, где $\lim\limits_{\|q\|\to 0}\|\beta(q)\|=0$. Мы также доопределим α и β в точке нуль нулем (т.е. считаем $\alpha(0)=0$ и $\beta(0)=0$). Тогда

$$\begin{aligned}g\big(f(a+h)\big)-g\big(f(a)\big)&=g\big(f(a)+[f(a+h)-f(a)]\big)-g\big(f(a)\big)\\&=dg[f(a+h)-f(a)]+\beta\big(f(a+h)-f(a)\big)\|f(a+h)-f(a)\|\\&=dg[df(h)+\alpha(h)\|h\|]+\beta\big(f(a+h)-f(a)\big)\|\,df(h)+\alpha(h)\|h\|\|.\end{aligned}$$

Тем самым,

$$g\big(f(a+h)\big)-g\big(f(a)\big)=dg[df(h)]+\gamma(h)\|h\|,$$

где

$$\begin{aligned}\|\gamma(h)\|&=\Big\|dg[\alpha(h)]+\beta\big(f(a+h)-f(a)\big)\|\,df(h/\|h\|)+\alpha(h)\|\Big\|\\&\leqslant\|dg[\alpha(h)]\|+\big\|\beta\big(f(a+h)-f(a)\big)\big\|(\|\,df(h/\|h\|)\|+\|\alpha(h)\|).\end{aligned}$$

Напомним, что для линейных отображений dg и df существуют такие постоянные A и B, что $\|\,df(h)\|\leqslant A\|h\|$ и $\|\,dg(q)\|\leqslant B\|q\|$, поэтому $\|\,df(h/\|h\|)\|+\|\alpha(h)\|\leqslant A+\|\alpha(h)\|$ и $\|\,dg[\alpha(h)]\|\leqslant B\|\alpha(h)\|$.

Так как $\big\|\beta\big(f(a+h)-f(a)\big)\big\|\to 0$ при $\|h\|\to 0$, получаем, что $\lim\limits_{\|h\|\to 0}\|\gamma(h)\|=0$. ■

8.2 Матрица Якоби композиции, правило вычисления частной производной сложной функции, инвариантность первого дифференциала.

Замечание. При фиксированных базисах $e=\{e_1,\dots,e_k\}$, $e'=\{e'_1,\dots,e'_m\}$, $e''=\{e''_1,\dots,e''_n\}$ в $\mathbb{R}^k$, $\mathbb{R}^m$ и $\mathbb{R}^n$ соответственно, матрица композиции линейных отображений есть произведение матриц этих линейных отображений.

Таким образом, в нашем случае для композиции функций $g\circ f$, где $f\colon\mathbb{R}^k\to\mathbb{R}^m$ и $g\colon\mathbb{R}^m\to\mathbb{R}^n$, по предыдущей теореме выполнено

$$\begin{pmatrix}\dfrac{\partial(g\circ f)_1}{\partial y_1}(a) & \dots & \dfrac{\partial(g\circ f)_1}{\partial y_k}(a)\\ \dots & & \\ \dfrac{\partial(g\circ f)_n}{\partial y_1}(a) & \dots & \dfrac{\partial(g\circ f)_n}{\partial y_k}(a)\end{pmatrix}=\begin{pmatrix}\dfrac{\partial g_1}{\partial x_1}(f(a)) & \dots & \dfrac{\partial g_1}{\partial x_m}(f(a))\\ \dots & & \\ \dfrac{\partial g_n}{\partial x_1}(f(a)) & \dots & \dfrac{\partial g_n}{\partial x_m}(f(a))\end{pmatrix}\begin{pmatrix}\dfrac{\partial f_1}{\partial y_1}(a) & \dots & \dfrac{\partial f_1}{\partial y_k}(a)\\ \dots & & \\ \dfrac{\partial f_m}{\partial y_1}(a) & \dots & \dfrac{\partial f_m}{\partial y_k}(a)\end{pmatrix}$$

В частности, в случае, когда $n=1$, для функции $g(x_1,\dots,x_m)$ и отображения

$$f(y_1,\dots,y_k)=(f_1(y_1,\dots,y_k),\dots,f_m(y_1,\dots,y_k)),$$

выполнено:

$$\begin{pmatrix} \frac{\partial(g\circ f)}{\partial y_1}(a) & \dots & \frac{\partial(g\circ f)}{\partial y_k}(a) \end{pmatrix} = \begin{pmatrix} \frac{\partial g}{\partial x_1}(f(a)) & \dots & \frac{\partial g}{\partial x_m}(f(a)) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(a) & \dots & \frac{\partial f_1}{\partial y_k}(a) \\ \dots & & \\ \frac{\partial f_m}{\partial y_1}(a) & \dots & \frac{\partial f_m}{\partial y_k}(a) \end{pmatrix}$$

Отсюда, во-первых получаем правило вычисления частной производной сложной функции:

$$\frac{\partial(g\circ f)}{\partial y_j}(a) = \frac{\partial g}{\partial x_1}(f(a))\frac{\partial f_1}{\partial y_j}(a) + \ldots + \frac{\partial g}{\partial x_m}(f(a))\frac{\partial f_m}{\partial y_j}(a)$$

. Во-вторых, получаем следующее свойство инвариантности первого дифференциала: для дифференциала выполнено равенство $dg = \frac{\partial g}{\partial x_1}\,dx_1 + \ldots + \frac{\partial g}{\partial x_m}\,dx_m$, где нам не важно, являются ли $dx_1, \ldots, dx_m$ — дифференциалами независимых переменных или же являются дифференциалами некоторых функций $x_j = f_j(y_1, \ldots, y_k)$.

Пример. Пусть $f(x,y) = \varphi(u,v,w)$, где $u = xy, v = x+y, w = x-y$. Тогда

$$\begin{aligned} df &= \frac{\partial\varphi}{\partial u}\,du + \frac{\partial\varphi}{\partial v}\,dv + \frac{\partial\varphi}{\partial w}\,dw = \frac{\partial\varphi}{\partial u}\,d(xy) + \frac{\partial\varphi}{\partial v}\,d(x+y) + \frac{\partial\varphi}{\partial w}\,d(x-y) \\ &= \frac{\partial\varphi}{\partial u}(x\,dy + y\,dx) + \frac{\partial\varphi}{\partial v}(dx+dy) + \frac{\partial\varphi}{\partial w}(dx-dy). \end{aligned}$$

В частности, $\frac{\partial f}{\partial x} = y\frac{\partial\varphi}{\partial u}(xy, x+y, x-y) + \frac{\partial\varphi}{\partial v}(xy, x+y, x-y) + \frac{\partial\varphi}{\partial w}(xy, x+y, x-y)$ и $\frac{\partial f}{\partial y} = x\frac{\partial\varphi}{\partial u}(xy, x+y, x-y) + \frac{\partial\varphi}{\partial v}(xy, x+y, x-y) - \frac{\partial\varphi}{\partial w}(xy, x+y, x-y)$.

9 Теорема о неявной функции: постановка вопроса, формулировка общей теоремы и доказательство в случае функции двух переменных.

9.1 Теорема о неявной функции: постановка вопроса, формулировка общей теоремы.

Пусть в $\mathbb{R}^2$ у нас имеется соотношение $F(x,y) = 0$. Нам бы хотелось понять при каких условиях данное уравнение возможно разрешить относительно y в виде явной зависимости $y = f(x)$.

Рассмотрим например $F(x,y) = x^2 + y^2 - 1$. Тогда уравнение $F(x,y) = 0$ задает обычную окружность и все решения данного уравнения относительно y имеют вид $y = \pm\sqrt{1-x^2}$. Ясно, что произвольный выбор знаков в разных точках x будет давать бесконечно много решений данного уравнения.

В тоже время в малой окрестности произвольной точки (x_0, y_0) на окружности (кроме $x_0 = \pm 1$) кривая $F(x,y) = 0$ единственным образом представима в виде графика непрерывной функции $y = f(x)$. В окрестности же точек $(\pm 1, 0)$ никакая дуга окружности не может быть представлена в виде графика функции $y = f(x)$. Зато эти дуги в окрестности точек $(\pm 1, 0)$ хорошо расположены относительно оси y и могут быть представлены в виде графика $x = g(y)$.

Чем же обусловлена такая особенность точек $(\pm 1, 0)$ в случае окружности? Заметим, что локально функция $F(x,y)$ представима в виде $F(x,y) = \frac{\partial F}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial F}{\partial y}(x_0,y_0)(y-y_0) + \bar{\bar{o}}(\sqrt{|x-x_0|^2 + |y-y_0|^2})$.

Таким образом, пренебрегая малыми более высокого порядка, наше уравнение $F(x,y) = 0$ в окрестности точки (x_0,y_0) похоже на линейное уравнение $\frac{\partial F}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial F}{\partial y}(x_0,y_0)(y-y_0) = 0$, которое в свою очередь разрешимо относительно y только в случае $\frac{\partial F}{\partial y}(x_0,y_0) \neq 0$.

В частности, в случае окружности как раз $\frac{\partial F}{\partial y}(\pm 1, 0) = 0$. Из данного эвристического рассуждения возникает гипотеза, что уравнение $F(x,y) = 0$ разрешимо относительно переменной y в некоторой окрестности данной точки (x_0,y_0), если производная $\frac{\partial F}{\partial y}(x_0,y_0)$ отлична от нуля. Именно это мы и докажем в следующей теореме уже в строго сформулированном виде.

Для сокращения всех записей будем использовать обозначение $F'_y(x,y) := \frac{\partial F}{\partial y}(x,y)$.

Теорема. Пусть $F\colon \mathbb{R}^2 \to \mathbb{R}$ — определена и непрерывно дифференцируема (т.е. частные производные непрерывно зависят от точки) в некоторой окрестности U точки $(a,b) \in \mathbb{R}^2$. Пусть 1) $F(a,b) = 0$ и 2) $F'_y(a,b) \neq 0$.

Тогда найдутся промежутки $I_x = (a-\alpha, a+\alpha)$ и $I_y = (b-\beta, b+\beta)$ и непрерывно дифференцируемая функция $f\colon I_x \to I_y$, для которых $I_x \times I_y \subset U$ и для каждой точки $(x,y) \in I_x \times I_y$ выполнено $F(x,y) = 0 \Leftrightarrow y = f(x)$.

Кроме того, $f'(x) = -\frac{F'_x(x, f(x))}{F'_y(x, f(x))}$.

9.2 Теорема о неявной функции: доказательство в случае функции двух переменных.

Доказательство.

1. Для определенности считаем, что $F'_y(a,b) > 0$. Так как производные функции F непрерывны в U, то в малой окрестности $\{(x,y)\colon \sqrt{|x-a|^2+|y-b|^2} < 2\beta\}$ точки (a,b) также выполнено $F'_y(x,y) > 0$.

   Так как $F'_y(a,y) > 0$ на отрезке $[b-\beta, b+\beta]$, то функция $y \mapsto F(a,y)$ монотонно возрастает на этом отрезке, откуда
$$F(a,b-\beta) < F(a,b) = 0 < F(a,b+\beta).$$

   Так как F непрерывна в U, то найдется такое число $\alpha < \beta$, что $F(x,b-\beta) < 0 < F(x,b+\beta)$ при $x \in (a-\alpha, a+\alpha)$.

   При каждом $x \in (a-\alpha, a+\alpha)$ рассмотрим функцию $y \mapsto F(x,y)$, заданную на отрезке $[b-\beta, b+\beta]$. Рассматриваемая функция есть непрерывная строго возрастающая функция на отрезке, причем на концах отрезка данная функция принимает значения разных знаков. Поэтому при каждом $x \in (a-\alpha, a+\alpha)$ существует единственная точка $y = f(x)$, для которой $F(x, f(x)) = 0$. Тем самым построена окрестность точки (a,b) вида $I_x \times I_y$ в которой построено единственное решение уравнения $F(x,y) = 0$ относительно переменной y.

2. Проверим непрерывность построенного решения в точке a. Ясно, что $f(a) = b$ в силу единственности нуля у функции $y \mapsto F(a,y)$ на I_y. Пусть теперь фиксировано некоторое $\varepsilon \in (0,\beta)$. Повторяя рассуждения первой части для интервала $(b-\varepsilon, b+\varepsilon)$ найдем интервал $(a-\delta, a+\delta)$ с $\delta < \alpha$ и функцию $\tilde{f}\colon (a-\delta, a+\delta) \to (b-\varepsilon, b+\varepsilon)$ для которых $F(x,y) = 0$ при $(x,y) \in (a-\delta, a+\delta) \times (b-\varepsilon, b+\varepsilon) \Leftrightarrow y = \tilde{f}(x), x \in (a-\delta, a+\delta)$.

   Так как $(a-\delta, a+\delta) \subset I_x$ и $(b-\varepsilon, b+\varepsilon) \subset I_y$, то в силу единственности решения f в $I_x \times I_y$ получаем, что $f(x) = \tilde{f}(x)$ при $x \in (a-\delta, a+\delta)$. Это означает, что $|f(x) - b| < \varepsilon$ при $|x-a| < \delta$.

   Теперь проверим непрерывность f в произвольной точке $x \in I_x$. Для этого просто примем за начальную точку построения произвольную точку (x,y) с $x \in I_x, y \in I_y$ и повторим рассуждение выше.

3. Докажем непрерывную дифференцируемость f на I_x и докажем формулу для вычисления производной. Пусть $x \in I_x$ и рассмотрим достаточно малое Δx, для которого $x + \Delta x \in I_x$. Пусть $y = f(x) \in I_y$ и $\Delta y = f(x+\Delta x) - f(x)$.

   Применим теорему Лагранжа к функции $t \mapsto F(x + t\Delta x, y + t\Delta y)$:
$$0 = F(x+\Delta x, y+\Delta y) - F(x,y) = F'_x(x+\xi\Delta x, y+\xi\Delta y)\Delta x + F'_y(x+\xi\Delta x, y+\xi\Delta y)\Delta y,$$
   где $\xi \in (0,1)$. Т.к. $F'_y \neq 0$ в $I_x \times I_y$, то
$$\frac{\Delta y}{\Delta x} = -\frac{F'_x(x+\xi\Delta x, y+\xi\Delta y)}{F'_y(x+\xi\Delta x, y+\xi\Delta y)}.$$

   В силу непрерывности f при $\Delta x \to 0$ выполнено, что и $\Delta y \to 0$, поэтому, в силу непрерывности производных функции F в $I_x \times I_y$, получается, что $f'(x) = -\dfrac{F'_x(x,y)}{F'_y(x,y)}$, где $y = f(x)$. Из формулы следует и непрерывность производной.

■

Аналогично доказывается следующий многомерный аналог предыдущей теоремы.

Теорема. Пусть $F\colon \mathbb{R}^{k+1} \to \mathbb{R}$ — определена и непрерывно дифференцируема (то есть частные производные непрерывно зависят от точки) в некоторой окрестности U точки $(a,b) = (a_1, \ldots, a_k, b) \in \mathbb{R}^{k+1}$.

Пусть 1) $F(a,b) = 0$ и 2) $F'_y(a,b) \neq 0$. Найдутся $I_x = (a_1 - \alpha_1, a_1 + \alpha_1) \times \ldots \times (a_k - \alpha_k, a_k + \alpha_k)$ и $I_y = (b-\beta, b+\beta)$ и непрерывно дифференцируемая функция $f\colon I_x \to I_y$, для которых $I_x \times I_y \subset U$ и для каждой точки $(x,y) \in I_x \times I_y$ выполнено $F(x,y) = 0 \Leftrightarrow y = f(x)$.

Кроме того, $\dfrac{\partial f}{\partial x_j}(x) = -\dfrac{F'_{x_j}(x, f(x))}{F'_y(x, f(x))}$.

Доказательство. Существование $I_x \times I_y$ и функции f, а также ее непрерывность, дословно повторяют рассуждение из предыдущей теоремы.

Если теперь фиксировать все переменные, кроме x_j и y, мы попадем в ситуацию предыдущей теоремы, откуда следует формула для вычисления частной производной. Из формулы следует непрерывность этой частной производной, а значит и непрерывная дифференцируемость f. ■

Замечание. Отметим, что формула для подсчет производной берется из дифференцирования тождества $F(x, f(x)) = 0$. Действительно, $\dfrac{\partial F}{\partial x_1}dx_1 + \ldots + \dfrac{\partial F}{\partial x_k} + \dfrac{\partial F}{\partial y}df = 0$, откуда выражая df и получаем нужную нам формулу.

10 Многомерная формула Тейлора.

10.1 Многомерная формула тейлора.

Лемма. Пусть функция f m раз дифференцируема в окрестности точки $a \in \mathbb{R}^k$. Рассмотрим функцию $\varphi(t) := f(a+th)$. Тогда φ m раз дифференцируема в окрестности точки нуль и $\varphi^{(m)}(t) = d^m f\big|_{a+th}(h)$.

Доказательство. Утверждение доказывается по индукции. База $m = 1$:

$$\varphi'(t) = \frac{\partial f}{\partial x_1}(a+th)h_1 + \ldots + \frac{\partial f}{\partial x_n}(a+th)h_n = df\big|_{a+th}(h).$$

Индуктивный переход:

$$\begin{aligned}\varphi^{(m+1)}(t) &= \frac{d}{dt}\Big[\sum_{j_1,\ldots,j_m} \frac{\partial^m f}{\partial x_{j_1}\ldots\partial x_{j_m}}(a+th)h_{j_1}\cdot\ldots\cdot h_{j_m}\Big] \\ &= \sum_{j_1,\ldots,j_m}\Big[\sum_{j=1}^{k}\frac{\partial^{m+1} f}{\partial x_j\partial x_{j_1}\ldots\partial x_{j_m}}(a+th)h_j\Big]h_{j_1}\cdot\ldots\cdot h_{j_m} = d^{m+1}f\big|_{a+th}(h).\end{aligned}$$

■

Теорема. Пусть функция f m раз непрерывно дифференцируема в окрестности точки a Тогда справедлива следующая формула Тейлора:

$$f(a+h) = f(a) + df\big|_a(h) + \frac{1}{2!}\,d^2 f\big|_a(h) + \ldots + \frac{1}{m!}\,d^{(m)} f\big|_a(h) + \bar{\bar{o}}(\|h\|^m).$$

Доказательство. Запишем для функции $\varphi(t) := f(a+th)$ формулу Тейлора с остаточным членом в интегральной форме:

$$\begin{aligned}&f(a+h) = \varphi(1) = \\ &\quad\varphi(0) + \varphi'(0)(1-0) + \frac{1}{2!}\varphi''(0)(1-0)^2 + \ldots + \frac{1}{(m-1)!}\varphi^{(m-1)}(0)(1-0)^{m-1} + \frac{1}{(m-1)!}\int\limits_0^1(1-t)^{m-1}\varphi^{(m)}(t)\,dt \\ &\quad= f(a) + df\big|_a(h) + \frac{1}{2!}\,d^2f\big|_a(h) + \ldots + \frac{1}{(m-1)!}\,d^{m-1}f\big|_a(h) + \frac{1}{m!}d^m f\big|_a(h) + R_m(h),\end{aligned}$$

где

$$R_m(h) := \frac{1}{(m-1)!}\int\limits_0^1 (1-t)^{m-1}\varphi^{(m)}(t)\,dt - \frac{1}{m!}\varphi^{(m)}(0) = \frac{1}{(m-1)!}\int\limits_0^1(1-t)^{m-1}\big(\varphi^{(m)}(t) - \varphi^{(m)}(0)\big)\,dt.$$

Отсюда

$$\begin{aligned}\frac{|R_m(h)|}{\|h\|^m} &\leqslant \frac{1}{m!}\frac{1}{\|h\|^m}\sup_{[0,1]}\big|\varphi^{(m)}(t) - \varphi^{(m)}(0)\big| \\ &\leqslant \frac{1}{m!}\sup_{[0,1]}\sum_{j_1,\ldots,j_m}\Big|\frac{\partial^m f}{\partial x_{j_1}\ldots\partial x_{j_m}}(a+th) - \frac{\partial^m f}{\partial x_{j_1}\ldots\partial x_{j_m}}(a)\Big|\frac{|h_{j_1}|}{\|h\|}\cdot\ldots\cdot\frac{|h_{j_m}|}{\|h\|} \to 0\end{aligned}$$

при $\|h\| \to 0$ в силу непрерывности частных производных m-го поряка. ■

11 Локальный экстремум: необходимое условие и достаточное условие.

11.1 Определение точки локального экстремума.

Определение. Точка a называется точкой *локального минимума (максимума)* функции f если всех точек x из некоторой окрестности $U(a)$ точки a выполнено $f(x) \geqslant f(a)$ $(f(x) \leqslant f(a))$.

Если для всех точек $x \in U(a), x \neq a$, выполнено $f(x) > f(a)$ $(f(x) < f(a))$, то точка a называется *точкой строгого локального минимума (максимума)*.

Точки локального минимума и максимума называются точками *локального экстремума.*

11.2 Необходимое условие локального экстремума.

Теорема (Необходимое условие локального экстремума)**.** Пусть a — точка локального экстремума функции f и предположим, что f дифференцируема в точке a. Тогда $df\big|_a = 0$ (или, что тоже самое, $\dfrac{\partial f}{\partial x_j}(a) = 0\,\forall j$).

Доказательство. Зафиксируем вектор h и функцию $\varphi(t) := f(a + th)$. Так как a — точка локального экстремума функции f, 0 — точка локального экстремума функции φ. Из одномерного случая известно, что $\varphi'(0) = 0$. Но $\varphi'(t) = df\big|_{a+th}(h)$, поэтому

$$df\big|_a(h) = \varphi'(0) = 0.$$

■

11.3 Достаточное условие локального экстремума.

Теорема (Достаточное условие локального экстремума)**.** Пусть f дважды непрерывно дифференцируема в окрестности точки a и предположим, что в точке a выполнено необходимое условие локального экстремума: $df\big|_a(h) = 0\,\forall h$.

Тогда

1. если $d^2f\big|_a(h) > 0\,\forall h \neq 0$, то a — точка строгого локального минимума;
2. если $d^2f\big|_a(h) < 0\,\forall h \neq 0$, то a — точка строгого локального максимума.

Доказательство. Докажем пункт 1), пункт 2) получается рассмотрением функции $-f$. По формуле Тейлора

$$f(a+h) - f(a) = \frac{1}{2}\,d^2f\big|_a(h) + \bar{\bar{o}}(\|h\|^2) = \|h\|^2\big(\frac{1}{2}\,d^2f\big|_a(\|h\|^{-1}h) + \bar{\bar{o}}(1)\big).$$

Заметим, что квадратичная функция $d^2f\big|_a(q) = \sum\limits_{i,j} \dfrac{\partial^2 f}{\partial x_i \partial x_j}(a) q_i q_j$ непрерывна (как функция аргумента q). Единичная сфера $\{q\colon \|q\| = 1\}$ — замкнутое и ограниченное множество, а значит компакт. Поэтому непрерывная функция $d^2f\big|_a(q)$ достигает на сфере своего минимума:

$$\min_{\|q\|=1} d^2f\big|_a(q) = m = d^2f\big|_a(q_0) > 0.$$

Поэтому

$$f(a+h) - f(a) \geqslant \|h\|^2\big(\frac{m}{2} + \bar{\bar{o}}(1)\big).$$

Существует такое δ, что при $\|h\| < \delta$ выполнено $|\bar{\bar{o}}(1)| < \dfrac{m}{4}$. Поэтому при $\|h\| < \delta$

$$f(a+h) - f(a) \geqslant \|h\|^2\big(\frac{m}{2} - \frac{m}{4}\big) = \frac{m\|h\|^2}{2} > 0.$$

■

Замечание. Отметим, что $d^2f\big|_a(h)$ — есть квадратичная форма, заданная матрицей

$$\begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(a) & \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \dots & \dfrac{\partial^2 f}{\partial x_1 \partial x_k}(a) \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \dfrac{\partial^2 f}{\partial x_2^2}(a) & \dots & \dfrac{\partial^2 f}{\partial x_2 \partial x_k}(a) \\ \dots & \dots & \dots & \end{pmatrix}$$

Предположения из предыдущей теоремы $d^2f\big|_a(h) > 0$ или $d^2f\big|_a(h) < 0$ $\forall h \neq 0$ означают положительную или отрицательную определенность квадратичной формы. Как известно из курса линейной алгебры, за положительную или отрицательную определенность квадратичной формы отвечает *критерий Сильвестра*:

1) все угловые миноры матрицы квадратичной формы $d^2f\big|_a$ положительны $\leftrightarrow$ $d^2f\big|_a$ — положительно определена (т.е. $d^2f\big|_a(h) > 0$ при каждом $h \neq 0$);

2) угловые миноры матрицы квадратичной формы $d^2f\big|_a$ начинаются с отрицательного, а затем чередуют знаки $\leftrightarrow$ $d^2f\big|_a$ — отрицательно определена (т.е. $d^2f\big|_a(h) < 0$ при каждом $h \neq 0$).

12 График функции. Касательная плоскость и касательное пространство к графику функции. Описание касательного пространства, как множества скоростей кривых, проходящих через данную точку.

12.1 График функции.

Для начала рассмотрим график функции $z = f(x, y)$, где $f: G \to \mathbb{R}$ — непрерывно дифференцируемая функция, и G — некоторая область в $\mathbb{R}^2$ (открытый круг, открытый прямоугольник).

График функции — это множество

$$\Gamma_f := \{(x, y, z) \in \mathbb{R}^3 \colon z - f(x, y) = 0,\ (x, y) \in G\} \subset \mathbb{R}^3.$$

Так как график f запараметризован парами чисел (x, y), то его естественно считать двумерной поверхностью в $\mathbb{R}^3$.

Так как f — дифференцируемая функция, то в окрестности любой точки (x_0, y_0, z_0) справедливо равенство

$$f(x, y) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + \bar{\bar{o}}\big(\sqrt{(x - x_0)^2 + (y - y_0)^2}\big).$$

То есть расстояние от точки $(x, y, f(x, y)) \in \Gamma_f$ до точки плоскости

$$z - z_0 = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$$

есть $\bar{\bar{o}}\big(\sqrt{(x - x_0)^2 + (y - y_0)^2}\big)$.

12.2 Касательная плоскость и касательное пространство к графику функции.

Плоскость

$$z - z_0 = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$$

естественно назвать касательной плоскостью к графику Γ_f в точке (x_0, y_0, z_0).

Линейное подпространство

$$\Big\{h = (h_x, h_y, h_z) \in \mathbb{R}^3 \colon h_z = \frac{\partial f}{\partial x}(x_0, y_0)h_x + \frac{\partial f}{\partial y}(x_0, y_0)h_y\Big\}$$

будем называть касательным пространством к Γ_f в точке (x_0, y_0, z_0) и обозначать $T_{(x_0, y_0, z_0)}\Gamma_f$.

Из сказанного ранее ясно, что точка (x, y, z) принадлежит касательной плоскости к графику функции f в точке (x_0, y_0, z_0) тогда и только тогда, когда вектор

$$(x - x_0, y - y_0, z - z_0) \in T_{(x_0, y_0, z_0)}\Gamma_f.$$

12.3 Описание касательного пространства, как множества скоростей кривых, проходящих через данную точку.

Определение. Кривой в $\mathbb{R}^k$ будем называть непрерывно дифференцируемое отображение $\gamma\colon (a, b) \to \mathbb{R}^k$.

Если $\gamma(t) = (\gamma_1(t), \dots, \gamma_k(t))$, то вектор $\dot{\gamma}(t_0) = (\dot{\gamma_1}(t_0), \dots, \dot{\gamma_k}(t_0))$ называют вектором скорости кривой γ в точке t_0.

Для касательной плоскости к графику функции существует инвариантное (относительно выбора базиса) описание.

Предложение. Вектор $h \in T_{(x_0, y_0, z_0)}\Gamma_f$ тогда и только тогда, когда найдется кривая $\gamma\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^3$ для которой $\gamma(0) = (x_0, y_0, z_0)$, $\gamma(t) \in \Gamma_f$ $\forall t \in (-1, 1)$ и $h = \dot{\gamma}(0)$.

Доказательство. Пусть γ — кривая из условия.

Тогда она имеет вид $\gamma(t) = (\gamma_x(t), \gamma_y(t), \gamma_z(t))$, причем $\gamma_z(t) = f(\gamma_x(t), \gamma_y(t))$. Тогда вектор

$$\dot{\gamma}(0) = \Big(\dot{\gamma_x}(0), \dot{\gamma_y}(0), \frac{\partial f}{\partial x}(\gamma_x(0), \gamma_y(0))\dot{\gamma_x}(0) + \frac{\partial f}{\partial y}(\gamma_x(0), \gamma_y(0))\dot{\gamma_y}(0)\Big) \in T_{(x_0, y_0, z_0)}\Gamma_f.$$

Наоборот, пусть $h \in T_{(x_0, y_0, z_0)}\Gamma_f$.

Рассмотрим кривую

$$\gamma(t) = \big(x_0 + th_x, y_0 + th_y, f(x_0 + th_x, y_0 + th_y)\big).$$

Тогда

$$\dot{\gamma}(0) = \Big(h_x, h_y, \frac{\partial f}{\partial x}(x_0, y_0)h_x + \frac{\partial f}{\partial y}(x_0, y_0)h_y\Big) = (h_x, h_y, h_z).$$

■

13 Поверхность в $\mathbb{R}^k$ и касательное пространство к ней. Описание касательного пространства к поверхности, заданной системой уравнений (доказательство в случае одного уравнения).

13.1 Поверхность в $\mathbb{R}^k$ и касательное пространство к ней.

Теперь мы можем дать общее определение $(k-1)$-мерной поверхности в $\mathbb{R}^k$ и касательного пространства к ней.

Определение. Подмножество $M \subset \mathbb{R}^k$ называется $(k-1)$-мерной поверхностью, если для каждой точки $a = (a_1, \ldots, a_k) \in M$ найдется номер j, окрестности

$$I = (a_1 - \alpha_1, a_1 + \alpha_1) \times \ldots \times (a_{j-1} - \alpha_{j-1}, a_{j-1} + \alpha_{j-1}) \times (a_{j+1} - \alpha_{j+1}, a_{j+1} + \alpha_{j+1}) \times \ldots \times (a_k - \alpha_k, a_k + \alpha_k)$$

и $J = (a_j - \alpha_j, a_j + \alpha_j)$ и непрерывно дифференцируемая функция $f\colon I \to J$, для которых для каждой точки $b \in I \times J$ выполнено $b \in M \Leftrightarrow x_j = f(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_k)$.

То есть $(k-1)$-мерной поверхностью называют такое подмножество $M \subseteq \mathbb{R}^k$, что у каждой точки $a \in M$ есть окрестность, в которой M совпадает с графиком некоторой функции относительной одной из координатных гиперплоскостей.

Пример. Пусть $F\colon \mathbb{R}^k \to \mathbb{R}$, причем для каждой точки $a \in \mathbb{R}^k$, в которой $F(a) = 0$, выполнено условие $\mathrm{rk}\nabla F = 1$. Тогда по теореме о неявной функции $\{a \in \mathbb{R}^k\colon F(a) = 0\}$ является $(k-1)$-мерной поверхностью в $\mathbb{R}^k$.

Определение. Касательным пространством $T_a M$ к поверхности M в точке $a \in M$ называется линейное пространство, состоящее из векторов скоростей кривых на M, проходящих через точку a. Касательной плоскостью называется плоскость $a + T_a M$.

13.2 Описание касательного пространства к поверхности, заданной системой уравнений (доказательство в случае одного уравнения).

Из определения $(k-1)$-мерной поверхности ясно, что касательное пространство действительно есть $(k-1)$-мерное линейное подпространство в $\mathbb{R}^k$.

Предложение. Пусть M задана уравнением $F(x) = 0$ и $\mathrm{rk}\nabla F(x) = 1$ $\forall x \in M$. Тогда $h \in T_a M \Leftrightarrow \langle \nabla F(a), h\rangle = dF\big|_a(h) = 0$.

Доказательство. Пусть γ — кривая на M, проходящая через точку a при $t = 0$ (т.е. $\gamma(0) = a$). Тогда $F(\gamma(t)) = 0$ и

$$\langle \nabla F(a), \dot\gamma(0)\rangle = \frac{\partial F}{\partial x_1}(\gamma(0))\dot{\gamma_1}(0) + \ldots + \frac{\partial F}{\partial x_k}(\gamma(0))\dot{\gamma_k}(0) = \frac{d}{dt}[F(\gamma(t))] = 0.$$

Значит $T_a M$ входит в пространство решений уравнения $\langle \nabla F(a), h\rangle = 0$. С другой стороны, т.к. $\mathrm{rk}\nabla F(A) = 1$, то пространство решений данного уравнения $(k-1)$-мерное, поэтому пространство $T_a M$ совпадает с пространством решений уравнения $\langle \nabla F(a), h\rangle = 0$. ■

В конце отметим, что аналогичным образом можно было определить график отображения

$$f(x_1, \ldots, x_k) = \big(f_1(x_1, \ldots, x_m), \ldots, f_{k-m}(x_1, \ldots, x_m)\big)$$

и назвать m-мерной поверхностью в $\mathbb{R}^k$ множество M, у каждой точки которого есть окрестность в которой M совпадает с графиком некоторого такого отображения.

Например, при таком определении по теореме о неявном отображении m-мерной поверхностью будет множество решений системы уравнений $F_1(x) = 0, \ldots, F_{k-m}(x) = 0$, при условии, что $\mathrm{rk}J_F = k - m$, где $F = (F_1, \ldots, F_{k-m})$. Определение касательного пространства остается тем же самым, а в случае поверхности, заданной системой уравнений касательное пространство задается системой линейных уравнений $\langle \nabla F_1(a), h\rangle = 0, \ldots, \langle \nabla F_{k-m}(a), h\rangle = 0$.

14 Формулировки теорем о неявном отображении и обратной функции. Параметрически заданные поверхности. Описание касательного пространства к поверхности, заданной параметрически.

14.1 Формулировка теоремы о неявном отображении.

Пусть мы рассматриваем систему уравнений

$$\begin{cases} F_1(x_1, \ldots, x_k, y_1, \ldots, y_n) = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ F_n(x_1, \ldots, x_k, y_1, \ldots, y_n) = 0. \end{cases}$$

Обозначим $x = (x_1, \dots, x_k)$, $y = (y_1, \dots, y_n)$, $F(x, y) = (F_1(x, y), \dots, F_n(x, y))$. По аналогии с предыдущей лекцией нам бы хотелось научиться локально, то есть в окрестности некоторой точки $(a, b) = (a_1, \dots, a_k, b_1, \dots, b_n)$, для которой $F(a, b) = 0$, научиться решать данную систему относительно переменных $y_1, \dots, y_n$, то есть находить функции $f_1, \dots, f_n$, для которых система

$$\begin{cases} y_1 = f_1(x_1, \dots, x_k), \\ \dots\dots\dots\dots\dots\dots \\ y_n = f_n(x_1, \dots, x_k) \end{cases}$$

равносильна исходной системе в рассматриваемой окрестности.

Приведем теперь строгую формулировку теоремы:

Теорема (О неявном отображении)**.** Пусть отображение $F\colon \mathbb{R}^{k+n} \to \mathbb{R}^n$ — определено и непрерывно дифференцируемо в некоторой окрестности U точки $(a, b) = (a_1, \dots, a_k, b_1, \dots, b_n) \in \mathbb{R}^{k+n}$. Пусть

1. $F(a, b) = 0$;
2. матрица $F'_y(a, b) := \begin{pmatrix} \dfrac{\partial F_1}{\partial y_1}(a, b) & \dots & \dfrac{\partial F_1}{\partial y_n}(a, b) \\ \dots & \dots & \dots \\ \dfrac{\partial F_n}{\partial y_1}(a, b) & \dots & \dfrac{\partial F_n}{\partial y_n}(a, b) \end{pmatrix}$ — обратима.

Тогда найдутся

$$I_x = (a_1 - \alpha_1, a_1 + \alpha_1) \times \ldots \times (a_k - \alpha_k, a_k + \alpha_k) \text{ и } I_y = (b_1 - \beta_1, b_1 + \beta_1) \times \ldots \times (b_n - \beta_n, b_n + \beta_n)$$

и непрерывно дифференцируемое отображение $f = (f_1, \dots, f_n)\colon I_x \to I_y$, для которых $I_x \times I_y \subset U$ и для каждой точки $(x, y) \in I_x \times I_y$ выполнено

$$F(x, y) = 0 \Leftrightarrow y = f(x).$$

Следствием предыдущей теоремы является следующая теорема об обратной функции.

Теорема (Об обратной функции)**.** Пусть отображение $f\colon \mathbb{R}^k \to \mathbb{R}^k$ — определено и непрерывно дифференцируемо в некоторой окрестности G точки $a \in \mathbb{R}^k$. Пусть $df\big|_a$ — обратимое линейное отображение. Тогда найдутся такие окрестности $U(a)$ и $V(f(a))$, что f есть диффеоморфизм этих окрестностей, то есть f — биекция окрестностей $U(a)$ и $V(f(a))$ и отображения f и f^{-1} непрерывно дифференцируемы на $U(a)$ и $V(f(a))$ соответственно.

14.2 Параметрически заданные поверхности.

Из определения $(k-1)$-мерной поверхности ясно, что касательное пространство действительно есть $(k-1)$-мерное линейное подпространство в $\mathbb{R}^k$.

Предложение. Пусть G — область в $\mathbb{R}^{k-1}$ и пусть $f\colon G \to \mathbb{R}^k$ — непрерывно дифференцируемое отображение, причем $\mathrm{rk} J_f(z) = k - 1$ $\forall z \in G$. Тогда для каждой точки $z_0 \in G$ найдется окрестность $U(z_0)$, для которой образ $f\big(U(z_0)\big)$ есть $(k-1)$-мерная поверхность в $\mathbb{R}^k$.

Кроме того, касательное пространство к данной поверхности в точке $a = f(z_0)$ совпадает с линейной оболочкой векторов $\dfrac{\partial}{\partial z_1} f(z_0), \dots, \dfrac{\partial}{\partial z_{k-1}} f(z_0)$.

Доказательство. Без ограничения общности считаем, что $\det J_{f_1, \dots, f_{k-1}}(z_0) \neq 0$. По теореме об обратной функции в некоторой окрестности $U(z_0)$ точки z_0 система

$$\begin{cases} x_1 = f_1(z_1, \dots, z_{k-1}), \\ \dots\dots\dots\dots\dots\dots \\ x_{k-1} = f_{k-1}(z_1, \dots, z_{k-1}) \end{cases}$$

равносильна системе

$$\begin{cases} z_1 = \varphi_1(x_1, \dots, x_{k-1}), \\ \dots\dots\dots\dots\dots\dots \\ z_{k-1} = \varphi_{k-1}(x_1, \dots, x_{k-1}) \end{cases}$$

для некоторых непрерывно дифференцируемых функции φ_j.

Отсюда, для этой окрестности образ $f(U(z_0))$ совпадает с графиком функции

$$x_k = f_k\big(\varphi_1(x_1, \dots, x_{k-1}), \dots, \varphi(x_1, \dots, x_{k-1})\big).$$

Перейдем к доказательству утверждения про касательное пространство. Заметим, что в силу условия $\mathrm{rk}J_f = k-1$ размерность указанной линейной оболочки равна $k-1$. Пусть γ — кривая на построенной поверхности M, имеющая вид $\gamma(t) = f(z(t))$, где $z(t)$ — кривая в G, $z(0) = z_0$. Тогда

$$\dot\gamma(0) = \frac{\partial}{\partial z_1} f(z(0))\dot z_1(0) + \ldots + \frac{\partial}{\partial z_{k-1}} f(z(0))\dot z_{k-1}(0),$$

что является элементом линейно оболочки векторов $\dfrac{\partial}{\partial z_1} f(z_0), \ldots, \dfrac{\partial}{\partial z_{k-1}} f(z_0)$. ■

14.3 Описание касательного пространства к поверхности, заданной параметрически.

Если поверхность $M \subset \mathbb{R}^3$ задана параметрически $x = x(u,v), y = y(u,v), z = z(u,v)$, то касательная плоскость задается уравнением

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ x'_u(u_0,v_0) & y'_u(u_0,v_0) & z'_u(u_0,v_0) \\ x'_v(u_0,v_0) & y'_v(u_0,v_0) & z'_v(u_0,v_0) \end{vmatrix} = 0.$$

15 Условный экстремум и метод множителей Лагранжа. Достаточное условие локального экстремума.

15.1 Условный экстремум и метод множителей Лагранжа.

Пусть $f\colon \mathbb{R}^k \to \mathbb{R}$ — непрерывно дифференцируемая функция и пусть $F\colon \mathbb{R}^k \to \mathbb{R}$ — также непрерывно дифференцируемая функция, $\mathrm{rk}\nabla F(x) = 1$ при $F(x) = 0$. Предположим, что мы хотим найти точки экстремума функции f при условии, что $F(x) = 0$. Тем самым мы ищем точки экстремума функции f на поверхности $\{x\colon F(x) = 0\}$.

Определение. Пусть M — поверхность в $\mathbb{R}^k$ и пусть f — функция в $\mathbb{R}^k$.

Точка $a \in M$ называется точкой *условного локального минимума (максимума)*, если для некоторой окрестности $U(a)$ точки a выполнено $f(b) \geq f(a)$ $(f(b) \leqslant f(a))$ $\forall\, b \in M \cap U(a)$. Если неравенство при $b \neq a$ строгое, то a называется точкой *строгого условного локального минимума (максимума)*.

Предложение (Необходимое условие условного локального экстремума)**.** Если a — точка условного локального экстремума, то $\nabla f(a) \perp T_a M$.

Доказательство. Пусть $h \in T_a M$, тогда $h = \dot\gamma(0)$ для некоторой кривой γ на M, $\gamma(0) = a$. Одномерная функция $f(\gamma(t))$ имеет в точке нуль локальный экстремум, поэтому

$$\langle \nabla f(a), \dot\gamma(0)\rangle = \frac{\partial f}{\partial x_1}(\gamma(0))\dot\gamma_1(0) + \ldots + \frac{\partial f}{\partial x_k}(\gamma(0))\dot\gamma_k(0) = \frac{d}{dt} f(\gamma(t))\Big|_{t=0} = 0.$$

■

В частности, в случае, когда M задано уравнением $F(x) = 0$, получаем, что в точке условного локального экстремума $\nabla f(a) \perp h$ $\forall h\colon h \perp \nabla F(a)$. Отсюда следует, что $\nabla f(a)$ и $\nabla F(a)$ пропорциональны, то есть существует число $\lambda\colon \nabla f(a) = \lambda \nabla F(a)$.

В случае, когда поверхность задана системой уравнений $F_1(x) = 0, \ldots, F_{k-m}(x) = 0$ (условный экстремум при нескольких ограничениях), условие $\nabla f(a) \perp T_a M$ в точке условного локального экстремума равносильно тому, что $\nabla f(a)$ лежит в линейной оболочке векторов $\nabla F_1(a), \ldots, \nabla F_{k-m}(a)$, то есть найдутся числа $\lambda_1, \ldots, \lambda_{k-m}$, для которых $\nabla f(a) = \lambda_1 \nabla F_1(a) + \ldots + \lambda_{k-m} \nabla F_{k-m}(a)$.

Заметим, что условие $\nabla f(a) = \lambda \nabla F(a)$ $(\nabla f(a) = \lambda_1 \nabla F_1(a) + \ldots + \lambda_{k-m} \nabla F_{k-m}(a))$ равносильно тому, что у функции

$$L_\lambda(x) := f(x) - \lambda F(x) \quad \big(L_{\lambda_1,\ldots,\lambda_{k-m}}(x) = f(x) - \lambda_1 F_1(x) - \ldots - \lambda_{k-m} F_{k-m}(x)\big)$$

в точке a дифференциал обращается в нуль $dL_\lambda\big|_a = 0$ (то есть все частные производные обращаются в нуль). Функцию $L_\lambda(x)$ называют **функцией Лагранжа**. Для поиска кандидатов в точки условного локального экстремума используют следующий метод множителей Лагранжа: по функции Лагранжа составляется система уравнений относительно переменных $a = (a_1, \ldots a_k)$ и λ

$$\begin{cases} \dfrac{\partial}{\partial x_1} L_\lambda(a) = 0 \\ \ldots\ldots\ldots\ldots\ldots \\ \dfrac{\partial}{\partial x_k} L_\lambda(a) = 0 \\ F(x) = 0, \end{cases}$$

где в случае нескольких условий $\lambda = (\lambda_1, \ldots, \lambda_{k-m})$, $F = (F_1, \ldots, F_{k-m})$.

15.2 Достаточное условие локального экстремума.

Теорема (Достаточное условие условного локального экстремума)**.** Пусть в точке $a \in M$ выполнено необходимое условие условного локального экстрмума, т.е. при некотором λ $dL_\lambda\big|_a = 0$. Тогда

1. если $d^2 L_\lambda\big|_a(h) > 0 \, \forall h \neq 0, h \in T_a M$, то a — точка строгого локального минимума;
2. если $d^2 L_\lambda\big|_a(h) < 0 \, \forall h \neq 0, h \in T_a M$, то a — точка строгого локального максимума.

Доказательство. Доказательство проведем в случае, когда M — $(k-1)$-мерная поверхность в $\mathbb{R}^k$ (одно условие). Заметим, что $L_\lambda(x) = f(x)$ на M. По определению в некоторой окрестности точки a поверхность M совпадает с графиком некоторой функции относительно одной из координатных осей. Без ограничения общности, считаем, что это график функции $x_k = \varphi(x_1, \ldots, x_{k-1})$. Тогда для функции

$$g(x_1, \ldots, x_{k-1}) := L_\lambda\big(x_1, \ldots, x_{k-1}, \varphi(x_1, \ldots, x_{k-1})\big)$$

точка $\tilde{a} := (a_1, \ldots, a_{k-1})$ является точкой обычного локального экстремума. Заметим, что

$$dg = \sum_{j=1}^{k-1} \Big[\frac{\partial L_\lambda}{\partial x_j} + \frac{\partial L_\lambda}{\partial x_k} \frac{\partial \varphi}{\partial x_j}\Big] \, dx_j$$

и, так как $\dfrac{\partial L_\lambda}{\partial x_k}(a) = 0$,

$$d^2 g\big|_{\tilde{a}} = \sum_{i,j=1}^{k-1} \frac{\partial^2 L_\lambda}{\partial x_i \partial x_j}(a) \, dx_i \, dx_j + 2 \sum_{i,j=1}^{k-1} \frac{\partial^2 L_\lambda}{\partial x_i \partial x_k}(a) \frac{\partial \varphi}{\partial x_j}(\tilde{a}) \, dx_i \, dx_j + \sum_{i,j=1}^{k-1} \frac{\partial^2 L_\lambda}{\partial^2 x_k}(a) \frac{\partial \varphi}{\partial x_i}(\tilde{a}) \frac{\partial \varphi}{\partial x_j}(\tilde{a}) \, dx_i \, dx_j.$$

Таким образом,

$$d^2 g\big|_{\tilde{a}}(q) = d^2 L_\lambda\big|_a(h),$$

где $h = \left(q_1, \ldots, q_{k-1}, \displaystyle\sum_{j=1}^{k-1} \frac{\partial \varphi}{\partial x_j}(\tilde{a}) q_j\right)$. Остается применить достаточное условие локального экстремума. ∎