Содержание

1 Пространство $\mathbb{R}^n$ **2**
1.1 Метрическое пространство . . . 2
1.2 Типы точек и множеств в метрическом пространстве . . . 3
1.3 Нормированные линейные пространства . . . 7
1.4 Компактные множества . . . 9
1.5 Сходимость последовательности . . . 11

2 Предел и непрерывность отображения **15**
2.1 Предел . . . 15
2.2 Непрерывность отображения . . . 19

3 Многомерное дифференциальное исчисление **22**
3.1 Производная и дифференциал . . . 22
3.2 Правила дифференцирования . . . 25
3.3 Достаточное условие дифференцируемости . . . 27
3.4 Градиент и касательная плоскость . . . 29

4 Производные и дифференциалы высших порядков **31**
4.1 Частные производные высших порядков . . . 31
4.2 Дифференциалы высших порядков . . . 34
4.3 Формула Тейлора для функции многих переменных . . . 36

5 Экстремумы функции многих переменных **38**
5.1 Необходимое условие экстремума . . . 38
5.2 Достаточное условие экстремума функции n переменных . . . 39

6 Неявное отображение и обратное отображение **43**
6.1 Теорема Лагранжа о среднем . . . 43
6.2 Производная функции, заданной неявно . . . 44
6.3 Производная отображения, заданного неявно . . . 46
6.4 Обратимость отображения . . . 49

7 Условный экстремум **51**

Функции многих переменных

1 Пространство $\mathbb{R}^n$

Пространство, в котором будем работать – $\mathbb{R}^n$ – линейное пространство, состоящее из n-мерных векторов $x = (x_1, x_2, \dots, x_n)$, где $x_i \in \mathbb{R}$.

Но многие понятия данного раздела используют для произвольных пространств. И читателю полезно представлять себе максимально абстрактные пространства.

1.1 Метрическое пространство

Определение 1.1.1 *Пусть X – некоторое множество. Функция ρ: $X \times X \to \mathbb{R}$ называется метрикой (или расстоянием) на X, если $\forall x, y, z \in X$ выполнено: 1) $\rho(x, y) = 0 \Leftrightarrow x = y$;*

2) $\rho(x, y) = \rho(y, x)$ (аксиома симметрии);

3) $\rho(x, y) + \rho(y, z) \geqslant \rho(x, z)$ (аксиома треугольника).

При этом пара (X, ρ) называется метрическим пространством.

Заметим, что из аксиомы треугольника при $x = z$ следует неотрицательность расстояния: $\forall x, y \in X \ \rho(x, y) \geqslant 0$.

На одном и том же множестве X можно задать разные метрики, получив тем самым разные метрические пространства.

Если понятно, какая метрика задана, то часто метрическое пространство (X, ρ) обозначают также, как и множество, на котором оно задано – X.

Пример 1.1.1 $X = \mathbb{R}$, $\rho(x, y) = |x - y|$.

Пример 1.1.2 $X = \mathbb{R}^n$, $x = (x_1, ..., x_n) \in \mathbb{R}^n$,

$$\rho_2(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad \textit{Евклидова метрика, стандартна для } \mathbb{R}^n$$

Пример 1.1.3 $X = \mathbb{R}^n$, $p \geqslant 1$,

$$\rho_p(x, y) = \sqrt[p]{\sum_{i=1}^{n} |x_i - y_i|^p}$$

При $p = 1$ – Манхеттеновское расстояние (расстояние городских кварталов).

При $p = +\infty$: $\rho(x, y) = \max\limits_{i=1..n} |x_i - y_i|$ – расстояние Чебышёва.

Для доказательства того, что в последнем примере функция ρ действительно задает метрику, воспользуемся неравенством Минковского:

$$\left(\sum_{i=1}^{n}|a_i+b_i|^p\right)^{1/p} \leqslant \left(\sum_{i=1}^{n}|a_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n}|b_i|^p\right)^{1/p}$$

и проверим выполнение неравенства треугольника:

$$\rho(x,z) = \sqrt[p]{\sum_{i=1}^{n}|x_i-z_i|^p} = \sqrt[p]{\sum_{i=1}^{n}|(x_i-y_i)+(y_i-z_i)|^p} \leqslant$$

при $a_i = x_i - y_i$, $b_i = y_i - z_i$

$$\leqslant \sqrt[p]{\sum_{i=1}^{n}|x_i-y_i|^p} + \sqrt[p]{\sum_{i=1}^{n}|y_i-z_i|^p} = \rho(x,y)+\rho(y,z).$$

Пример 1.1.4 *Дискретная метрика (для любого X)* $\rho(x,y) = \begin{cases} 1, & x \neq y; \\ 0, & x = y. \end{cases}$

Пример 1.1.5 *Для* $f, g \in C[a,b] = X$*:* $\rho(f,g) = \sup\limits_{x\in[a,b]} |f-g|$.

Пример 1.1.6 *Для* $f, g \in C[a,b] = X$*:* $\rho(f,g) = \sqrt[p]{\int\limits_a^b |f-g|^p dx}$.

Заметим, что в последнем примере нельзя взять $X = R[a,b]$, так как расстояние между функциями, отличающимися в одной точке (а значит, различными), будет равно нулю, то есть не выполняется первая аксиома метрики.

1.2 Типы точек и множеств в метрическом пространстве

В этом пункте будем предполагать, что (X,ρ) – произвольное метрическое пространство.

Определение 1.2.1 ***Открытым (замкнутым) шаром*** *с центром $a \in X$ и радиусом r ($r > 0$) называется множество*

$$B_r(a) = \{x \in X : \rho(x,a) < r\} \, (\bar{B}_r(a) = \{x \in X : \rho(x,a) \leqslant r\}),$$

сфера *с центром $a \in X$ и радиусом r ($r > 0$):*

$$S_r(a) = \{x \in X : \rho(x,a) = r\}\,.$$

Пример 1.2.1 *Нарисуйте сферу $S_1(0,0)$ в пространстве $\mathbb{R}^2$ с метриками ρ_2, ρ_1, $\rho_{+\infty}$ (см. Пример 1.1.3).*

Пусть $M \subset X$ – некоторое множество. По отношению к множеству M точку $x_0 \in X$ можно охарактеризовать следующим образом:

Определение 1.2.2

1. *Точка x_0 называется **внутренней** точкой множества M, если существует шар $B_r(x_0) \subset M$, то есть точка x_0 лежит в M вместе с некоторым открытым шаром.*

2. *Точка x_0 называется **внешней** точкой множества M, если она является внутренней для дополнения M^C.*

3. *Иначе точка x_0 называется **граничной** точкой множества M.*

Таким образом, внутренняя точка обязательно принадлежит множеству, внешняя – не принадлежит, а граничная точка – это такая точка, что в любом шаре с центром в этой точке есть точки как из данного множества, так и не принадлежащие ему.

Будем использовать следующие обозначения:

$$\operatorname{Int} M - \text{множество внутренних точек (внутренность) } M$$
$$\partial M - \text{множество граничных точек (граница) } M.$$

Лемма 1.2.1 *В метрическом пространстве $(\mathbb{R}^n, \rho)$ выполнено:*

1. $\partial \bar{B}_r(x_0) = \partial B_r(x_0) = \partial S_r(x_0) = S_r(x_0)$;

2. $\operatorname{Int} B_r(x_0) = \operatorname{Int} \bar{B}_r(x_0) = B_r(x_0)$;

3. $\operatorname{Int} S_r(x_0) = \emptyset$.

Доказательство. Доказательство основано на определениях. Проделайте самостоятельно (например, для $\mathbb{R}^2$). □

Теперь определим понятие открытого и замкнутого множества.

Определение 1.2.3 *Множество $G \subset X$ называется **открытым** (в X), если все его точки – внутренние. Пустое множество $\emptyset$ считается открытым по определению.*

То есть, G – открыто, если

$$\forall x \in G \ \exists B_r(x) \subset G,$$

другими словами, вместе с каждой своей точкой оно содержит и некоторый открытый шар с центром в этой точке.

Пример 1.2.2 *$\emptyset$ и X – открыты в X; интервал (a,b) – открыт, а отрезок $[a,b]$ не открыт в (R,ρ_1).*

Лемма 1.2.2 *Открытый шар есть открытое множество.*

Доказательство. Пусть $\xi \in B_r(x_0)$. Возьмём $r_\xi = \frac{1}{2}(r - \rho(\xi, x_0))$ и покажем, что $B_{r_\xi}(\xi) \subset B_r(x_0)$.

Пусть $y \in B_{r_\xi}(\xi)$. Тогда

$$\rho(y, x_0) \leqslant \rho(y, \xi) + \rho(\xi, x_0) < \frac{r - \rho(\xi, x_0)}{2} + \rho(\xi, x_0) = \frac{r + \rho(\xi, r_0)}{2} < \frac{2r}{2} = r$$

то есть, $y \in B_r(x_0)$. □

Определение 1.2.4 ***Окрестностью*** *$U(x_0)$ точки $x_0 \in X$ называется любое открытое множество, содержащее эту точку.*
Проколотой окрестностью *$\overset{o}{U}(x_0)$ называется разность окрестности и данной точки: $\overset{o}{U}(x_0) = U(x_0) \setminus \{x_0\}$.*
Эпсилон-окрестностью *точки x_0 называется открытый шар радиуса ε: $U_\varepsilon(x_0) := B_\varepsilon(x_0)$.*

Определение 1.2.5 *Множество $F \subset X$ называется* ***замкнутым*** *в X, если его дополнение $F^C = X \setminus F$ открыто в X.*

Пример 1.2.3 *$\emptyset$ и X – замкнуты в X; интервал (a,b) – не замкнут, а отрезок $[a,b]$ замкнут в (R,ρ_1).*

Заметим, что в $\mathbb{R}^n$ (с евклидовой метрикой) только два множества $\emptyset$ и X являются открытыми и замкнутыми одновременно.

Пример 1.2.4 *Пусть $X = (-\infty, 0) \cup (1, +\infty)$. Тогда $M = (-\infty, 0)$ – открыто и замкнуто в X. В англоязычной литературе такие множества называются clopen set.*

Лемма 1.2.3 *Замкнутый шар есть замкнутое множество.*

Доказательство. Докажите самостоятельно. □

В следующей Лемме все множества – подмножества X.

Лемма 1.2.4 (Свойства открытых и замкнутых множеств) *1. Если G_α, $\alpha \in A$ – открыты, тогда $\bigcup\limits_{\alpha \in A} G_\alpha$ – открыто.*

2. *Если* $G_1, \ldots, G_n$ *– открыты, тогда* $\bigcap\limits_{i=1}^{n} G_i$ *– открыто.*

3. *Если* F_α, $\alpha \in A$ *– замкнуты, тогда* $\bigcap\limits_{\alpha\in A}$ *– замкнуты.*

4. *Если* $F_1, \ldots, F_n$ *– замкнуты, тогда* $\bigcup\limits_{i=1}^{n} G_i$ *– замкнуто.*

5. *Если* G *– открыто, а* F *– замкнуто, то* $G \setminus F$ *– открыто, а* $F \setminus G$ *– замкнуто в* X*.*

Доказательство. 1. Пусть $x \in G = \bigcup\limits_{\alpha\in A} G_\alpha$. Тогда $x \in G_\alpha$ для некоторого $\alpha \in A$, а значит $\exists B(x) \subset G_\alpha \subset G$, откуда следует, что G – открыто.

2. Пусть $x \in F = \bigcap\limits_{i=1}^{n} G_i$. Тогда $x \in G_i$ $\forall i = 1..n$ и $\exists r_1, \ldots, r_n$: $B_{r_i}(x) \in G_i$. Тогда $B_r(x) \in G$, где $r = \min\limits_{i=1..n} r_i$.

3,4. Доказательство следует из доказанного и законов де Моргана.

5. Доказательство основано на равенствах:

$$G \setminus F = G \cap F^C, \qquad F \setminus G = F \cap G^C.$$

$\square$

Пример 1.2.5 *Пересечение открытых множеств может не быть открытым, а объединение замкнутых – замкнутым:*

$$\bigcap_{n\in\mathbb{N}} \left(0,\ 1 + \frac{1}{n}\right) = (0, 1], \qquad \bigcup_{n\in\mathbb{N}} \left[\frac{1}{n}, 2 - \frac{1}{n}\right] = (0, 2).$$

Очевидно, что сфера – замкнутое множество, так как представима в виде разности замкнутого и открытого шаров: $S_r(a) = \bar{B}_r(a) \setminus B_r(a)$.

Определение 1.2.6 *Точка* x_0 *называется* ***предельной точкой*** *множества* M*, если в любой её проколотой окрестности точки есть точки из множества* M*. Множество предельных точек обозначают* M'*.*

Предельная точка не обязательно принадлежит множеству. Понятно, что в любой окрестности предельной точки имеется бесконечно много точек данного множества.

Определение 1.2.7 *Если к множеству* M *добавить все его предельные точки, то полученное множество называется* ***замыканием*** *множества* M*. Замыкание обозначают* $\operatorname{cl} M$ *или* $\bar{M}$*.*

$$\operatorname{cl} M = M \cup M', \quad M' - \textit{мн-во предельных точек}.$$

Следующую теорему часто используют как другое определение замкнутого множества.

Теорема 1.2.1 (Критерий замкнутости множества) *Множество замкнуто тогда и только тогда, когда оно содержит все свои предельные точки.*

Другими словами, F – замкнуто $\Leftrightarrow F = \operatorname{cl} F$.

Доказательство. 1. Пусть F – замкнуто, $x \in F' \setminus F$. Тогда $x \in F^C$ и F^C – открыто, а значит некоторая окрестность $U(x_0)$ целиком лежит в F^C и не может содержать точек из F, что противоречит тому, что x_0 – предельная для F.

2. Пусть теперь $F = \operatorname{cl} F$. Докажем, что F^C – открыто. Пусть $x \in F^C$. Тогда $x \notin F = \operatorname{cl} F$, то есть точка x не предельная для F. Тогда существует шар $B_r(x) \subset F^C$, не содержащий точек F, откуда следует, что точка x – внутренняя для F^C. То есть F^C – открыто. □

Определение 1.2.8 *Точка $x \in M$ и не являющаяся предельной точкой множества M называется **изолироанной точкой** множества M.*

Для изолированной точки существует окрестность, не содержащая других точек из M. Каждая точка множества M является либо его предельной точкой, либо изолированной.

Еще несколько простых свойств произвольного множества $M \subset X$:

1. $\operatorname{cl} M$ – замкнуто;
2. $\operatorname{Int} M$ – открыто;
3. ∂M – замкнуто.

1.3 Нормированные линейные пространства

Здесь мы вспомним понятие нормированного пространства и свойства нормы.

Пространство является ***линейным*** (или ***векторным***), если в нем определены операции сложения элементов и умножение элемента на число (для нас вещественное). Эти операции должны удовлетворять аксиомам сложения и умножения, и их результаты должны лежать в этом же пространстве.

Нормой $\|\cdot\|$ называется отображение $X \to \mathbb{R}$, удовлетворяющее аксиомам нормы:

1. $\|x\| = 0 \Leftrightarrow x = 0$;
2. $\forall \lambda \in \mathbb{R} \quad \|\lambda x\| = |\lambda| \cdot \|x\|$;

3. $\forall x, y \in X \quad \|x+y\| \leqslant \|x\| + \|y\|$ – неравенство треугольника.

Линейное пространство, на котором задана норма называется **нормированным пространством**.

Заметим, что условие $\|x\| \geqslant 0$ следует из неравенства треугольника при $y = -x$.

Всякое нормированное пространство можно сделать метрическим, если ввести метрику:

$$\rho(x, y) = \|x - y\|.$$

Таким образом на нормированные пространства переносятся все понятия, имеющиеся для метрических пространств.

Отметим ещё одно свойство нормы:

$$\|x - y\| \geqslant \Big| \|x\| - \|y\| \Big|.$$

Приведем примеры стандартных норм (сравните с Примером 1.1.3):

Пример 1.3.1 *1. На множестве $\mathbb{R}$ естественная норма $\|x\| = |x|$.*

2. В пространстве $\mathbb{R}^n$ с элементами $x = (x_1, x_2, \ldots, x_n)$ можно ввести следующие нормы:

$$\|x\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}, \qquad (\textit{при } p = 2 \textit{ евклидова норма});$$

$$\|x\|_\infty = \max |x_i|.$$

3. $X = C[a, b]$:

$$\|f\| = \max |f(x)| \quad – \textit{ равномерная норма};$$

или

$$\|f\|_p = \sqrt[p]{\int_a^b |f(x)|^p dx} \quad – \textit{ интегральная норма.}$$

Ещё один пример рассмотрим более подробно.

О норме линейного оператора

Множество линейных операторов, действующих из $\mathbb{R}^m$ в $\mathbb{R}^n$ будем обозначать $L(\mathbb{R}^m, \mathbb{R}^n)$. Это множество является линейным пространством. Норма определяется следующим образом:

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|},$$

то есть, $\|A\|$ – это инфимум таких чисел C, для которых при всех $x \in \mathbb{R}^m$ верно неравенство

$$\|Ax\| \leqslant C\|x\|.$$

Почему указанный супремум существует, будет ясно из п.4 следующей леммы, описывающей свойства нормы линейного оператора.

Лемма 1.3.1 (Свойства нормы линейного оператора) *Для линейного оператора $A \in L(\mathbb{R}^m, \mathbb{R}^n)$ верны свойства:*

1. $\|A\| = \sup\limits_{\|x\|\leqslant 1} \|Ax\| = \sup\limits_{\|x\|=1} \|Ax\|$;
2. $\|Ax\| \leqslant \|A\| \cdot \|x\|$;
3. $\|AB\| \leqslant \|A\| \cdot \|B\|$;
4. $\|A\| \leqslant C_A$, *где* $C_A = \left(\sum_{i,j} a_{ij}^2\right)^{1/2}$, *где* $\{a_{ij}\}$ *– матрица оператора* A.

Доказательство. 1. Так как для любого $x \neq 0$ верно $\left\|\frac{x}{\|x\|}\right\| = 1$, то

$$\|A\| = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|.$$

Тогда требуемое равенство следует из неравенства

$$\|A\| = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} \geqslant \sup_{\|x\|\leqslant 1} \frac{\|Ax\|}{\|x\|} \geqslant \sup_{\|x\|\leqslant 1} \|Ax\| \geqslant \sup_{\|x\|=1} \|Ax\| = \|A\|.$$

2. Сразу следует из определения нормы линейного оператора.
3. Следует из цепочки неравенств: $\|ABx\| \leqslant \|A\| \cdot \|Bx\| \leqslant \|A\| \cdot \|B\| \cdot \|x\|$.
4. Воспользуемся неравенством Коши–Буняковского:

$$\|Ax\|^2 = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij}x_j\right)^2 \leqslant \sum_{i=1}^{n} \left(\sum_{j=1}^{m} a_{ij}^2 \sum_{j=1}^{m} x_j^2\right) = \left(\sum_{i,j} a_{ij}^2\right)^{1/2} \sum_{j=1}^{n} x_j^2 = C_A\|x\|.$$

$\square$

1.4 Компактные множества

Здесь везде (X, ρ) – метрическое пространство.

Определение 1.4.1 *Говорят, что система множеств E_α, $\alpha \in A$, образует* ***покрытие*** *множества X, если $X \subset \bigcup\limits_{\alpha\in A} E_\alpha$.*

Определение 1.4.2 *Множество $K \subset X$ называется компактным (или компактом), если из любого его покрытия множествами, открытыми в X, можно выделить конечное покрытие.*

Пример 1.4.1 *Отрезок $[a, b]$ – компакт в $\mathbb{R}$ (по лемме Бореля–Лебега: из любого покрытия отрезка интервалами можно выделить конечное покрытие), $[a, b)$ – не компакт в $\mathbb{R}$.*

Множество

$$\Pi = \{x \in \mathbb{R}^n : \ a_i \leqslant x_i \leqslant b_i\}$$

будем называть **n-мерным параллелепипедом** (или **брусом**).

Лемма 1.4.1 *Брус Π – компакт.*

Доказательство. Пусть существует покрытие бруса $\Pi_0 = \Pi$ открытыми множествами E_α, $\alpha \in A$ такое, что из них нельзя выделить конечное покрытие.

Поделим каждую сторону Π_0 пополам. Получим 2^n новых параллелепипедов. Хотя бы один из них не допускает конечного покрытия, пусть это параллелепипед Π_1. Продолжаем и т.д.

$$\Pi = \Pi_0 \supset \Pi_1 \supset \Pi_2 \supset ...\Pi_p \supset ...$$

$$\Pi_p = \{x \in \mathbb{R}^n : \ a_i^p \leqslant x_i \leqslant b_i^p\}, \quad b_i^p - a_i^p \to 0 \text{ при } p \to \infty.$$

Получаем систему вложенных отрезков по p: $I_p^i = [a_i^p, \ b_i^p]$. По теореме Кантора найдется

$$\exists \eta_i \in \bigcap_{p=0}^{\infty} I_p^i$$

$$\eta = (\eta_1, \eta_2, ..., \eta_n) \in \Pi_i \quad \forall i \in \mathbb{N} \cup \{0\} \quad \Rightarrow$$

$$\exists E_{\alpha_0} : \ \eta \in E_{\alpha_0}, \ E_{\alpha_0} \text{ – открыто} \quad \Rightarrow \quad B(\eta, r) \subset E_{\alpha_0} \quad \Rightarrow$$

$$\exists p_0 : \ \forall p > p_0 \ \Pi_p \subset B(\eta, r) \subset E_{\alpha_0} \quad \Rightarrow \text{ Противоречие с построением } \Rightarrow$$

Π – компакт. □

Теорема 1.4.1 (Критерий компактности в $\mathbb{R}^n$) *Множество компактно в $\mathbb{R}^n$ тогда и только тогда, когда оно замкнуто и ограничено в $\mathbb{R}^n$.*

Доказательство. 1. Пусть K – компакт в $\mathbb{R}^n$. Докажем замкнутость. Пусть $x \notin K$ – предельная точка K. Тогда $\forall y \in K$ найдем пару окрестностей $U_y(y)$, $O_y(x)$: $U_y(y) \cap O_y(x) = \emptyset$. Множество окрестностей $U_y(y)$, $y \in \mathbb{R}$ образует открытое покрытие K. Выделим из него конечное покрытие

$$U(y_1), ..., U(y_m); \quad K \subset \bigcup_{i=1}^{m} U(y_i),$$

которому соответствует конечный набор окрестностей $\{O_i(x)\}$. Рассмотрим

$$O(x) = \bigcap_{i=1}^{m} O_i(x) \text{ – окрестность } x,$$

причем $O(x)$ не пересекается с $\bigcup_{i=1}^{m} U(y_i) \supset K$, значит $O(x) \cap K = \emptyset \Rightarrow x$ – не предельная. Противоречие. И значит K замкнуто.

Докажем ограниченность. Ограниченность множества $K \subset \mathbb{R}^n$ равносильна $\exists B(0, r) \supset K$. Рассмотрим множество шаров $B(x, n)$, $n \in \mathbb{N}$.

$$\bigcup_{n \in \mathbb{N}} B(x, n) \supset K \quad \text{– открытое покрытие } K.$$

Можно выбрать конечное покрытие $B(x, n_1), ..., B(x, n_p)$. Тогда

$$K \subset B(x, \max\{n_1, ..., n_p\}).$$

2. Обратно. Пусть K – замкнуто и ограничено в $\mathbb{R}^n$ и G_α – открытое покрытие K.

Так как K ограничено, то найдётся брус Π, содержащий K: $K \subset \Pi$. Пусть $G = \mathbb{R}^n \setminus K$ – открыто. Тогда объединение $\bigcup\limits_{\alpha} G_\alpha \cup G$ образует открытое покрытие бруса Π. Так как брус компактен, то выделим из этого покрытия конечное, которое будет и покрытием K. Значит K – компакт. $\square$

1.5 Сходимость последовательности

Введем понятие расширенного пространства $\mathbb{R}^n$, дополнив его бесконечно удаленной точкой.

Определение 1.5.1 $\bar{\mathbb{R}}^n = \mathbb{R}^n \cup \{\infty\}$. *При этом эпсилон-окрестностью бесконечности называется*

$$U_\varepsilon(\infty) = \{x \in \mathbb{R}^n : \ \rho(x, 0) > 1/\varepsilon\}.$$

Заметим, что $\bar{\mathbb{R}}^1 \neq \bar{\mathbb{R}}$, так как в $\bar{\mathbb{R}}$ содержатся точки $+\infty$ и $-\infty$.

Определение 1.5.2 *Последовательностью x^k в $\mathbb{R}^n$ называется отображение $\mathbb{N} \to \mathbb{R}^n$.*

Обозначать будем так:

$$x^k = \left(x_1^k, x_2^k, ..., x_n^k\right).$$

Определение 1.5.3 *Последовательность называется ограниченной, если существует шар $B_r(0)$, содержащий все члены последовательности.*

Заметим, что для ограниченности последовательности достаточно наличие шара, содержащего члены последовательности начиная с некоторого номера.

Замечание 1.5.1 *Ограниченность x^k равносильна ограниченности всех x_i^k, $\forall i \in \{1, ..., n\}$.*

Определение 1.5.4 *Пусть $A \in \bar{\mathbb{R}}^n$. Говорят, что A – предел последовательности x^k, если*

$$\forall \varepsilon > 0 \; \exists k_0 \in N : \; \forall k > k_0 \;\Rightarrow\; x^k \in U_\varepsilon(A).$$

Если $A \in \mathbb{R}^n$, то $x^k \in U_\varepsilon(A) \Leftrightarrow \|x^k - A\| < \varepsilon$ или $\rho(x^k, A) < \varepsilon$.
Если $A = \infty$, то $x^k \in U_\varepsilon(A) \Leftrightarrow \|x^k\| > 1/\varepsilon$ или $\rho(x^k, 0) > 1/\varepsilon$.

Определение 1.5.5 *Если $\lim\limits_{k\to\infty} x^k = A \in \mathbb{R}^n$, то x^k называется сходящейся последовательностью.*

Замечание 1.5.2 *Сходимость последовательности зависит от введенной метрики и нормы. Так при одной метрике данная последовательность может сходиться, а при другой – расходиться.*

Замечание 1.5.3 *Далее, говоря про пространство $\mathbb{R}^n$, используем по умолчанию естественную норму:*

$$\|x\| := \|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}.$$

Свойства сходящихся последовательностей в $\mathbb{R}^n$:

1. Если существует предел последовательности в $\bar{\mathbb{R}}$, то он единственен.

2. Пусть $x^{(k)} \in \mathbb{R}^n$, $A = \{A_i\}_{i=1}^n \in \mathbb{R}^n$. Тогда

$$x^{(k)} \to A \Leftrightarrow \lim x_i^{(k)} = A_i, i = 1, \dots, n,$$

то есть, сходимость последовательности в $\mathbb{R}^n$ равносильна сходимостям в $\mathbb{R}^1$ последовательностей каждой координаты.

Доказательство. Пусть $x^k \to A = (A_1, ..., A_n)$. Тогда

$$0 \leqslant |x_i^k - A_i| \leqslant \sqrt{(x_i^k - A_i)^2 + ... + (x_n^k - A_n)^2} \longrightarrow 0,$$

откуда следует $x_i^k \to A_i$.

Обратно. Пусть теперь $x_i^k \to A_i$, тогда

$$0 \leqslant \sqrt{(x_i^k - A_i)^2 + ... + (x_n^k - A_n)^2} \leqslant \sqrt{n \cdot \max(x_i^k - A_i)^2} =$$
$$= \sqrt{n} \cdot \max |x_i^k - A_i| \longrightarrow 0.$$

□

3. Если последовательность сходится, то она ограничена.

4. Линейность. Пусть $x^k \to x$, $y^k \to y$ $(x, y \in \mathbb{R}^n)$, $\lambda_k \to \lambda$ $(\lambda_k, \lambda \in \mathbb{R})$. Тогда

   (a) $x^k + y^k \longrightarrow x + y$;

   (b) $\lambda_k \cdot x^k \longrightarrow \lambda x$;

   (c) $\alpha x^k + \beta y^k \longrightarrow \alpha x + \beta y$.

5. Пусть $x^k \to A \in \bar{\mathbb{R}}^n$, тогда любая её подпоследовательность $x^{k_p} \longrightarrow A$.

Свойство 1 доказывается также, как в $\mathbb{R}$. Доказательства Свойств 3–5 удобно провести, сославшись на Свойство 2.

Заметим, что Свойство 2 не выполнено для $A = \infty$. Например, $(n, 0) \to \infty \in \mathbb{R}^2$.

Теорема 1.5.1 (Больцано–Вейерштрасса) *Если последовательность ограничена, то существует её сходящаяся подпоследовательность. У неограниченной последовательности существует подпоследовательность, стремящаяся к бесконечности.*

Доказательство. Докажем для $\mathbb{R}^2$. Пусть $x^k = (x_1^k, x_2^k)$ – ограничена. Следовательно, x_1^k – ограничена (в $\mathbb{R}$) и по теореме Больцано–Вейерштрасса в ней можно выделить сходящуюся подпоследовательность $x_1^{k_p}$. Рассмотрим теперь последовательность $x_2^{k_p}$ (она ограничена, т.к. x_2^k ограничена) и выделим в ней сходящуюся подпоследовательность $x_2^{k_{p_t}}$. Получаем сходящуюся в $\mathbb{R}^2$ подпоследовательность исходной последовательности $(x_1^{k_{p_t}}, x_2^{k_{p_t}})$. $\square$

Определение 1.5.6 *Последовательность $x_n \in X$ в метрическом пространстве (X, ρ) называется* ***фундаментальной*** *(или* ***последовательностью Коши****), если*

$$\forall \varepsilon > 0 \exists N : \forall k, m > N \quad \rho\left(x^k, x^m\right) < \varepsilon.$$

Теорема 1.5.2 (Критерий Коши) *В $\mathbb{R}^n$ последовательность сходится тогда и только тогда, когда она фундаментальна.*

Доказательство. 1. Пусть $x^k \to A$. Тогда фундаментальность следует из неравенства:

$$\|x^k - x^m\| \leqslant \|x^k - A\| + \|x^m - A\|.$$

2. Пусть x^k – фундаментальна. Тогда каждая её координатная последовательность фундаментальна, т.к. $\|x_i^k - x_i^m\| \leqslant \|x^k - x^m\|$. И по критерию Коши в $\mathbb{R}$ все последовательности x_i^k $(i = 1, .., n)$ сходятся, следовательно, x^k сходится. $\square$

Замечание 1.5.4 *В любом метрическом пространстве утверждение неверно.*

Определение 1.5.7 *Метрическое пространство, в котором любая фундаментальная последовательность сходится, называется* ***полным****.*

Таким образом, пространство $\mathbb{R}^n$ – полное.

Пример 1.5.1 *Пространство $(\mathbb{Q}, \rho)$, $\rho(x, y) = |x - y|$ – не полное, так как существует фундаментальная последовательность рациональных чисел $x_n \in \mathbb{Q}$, сходящаяся к иррациональному числу (в $\mathbb{R}$) и, значит, не сходящаяся в $\mathbb{Q}$.*

2 Предел и непрерывность отображения

2.1 Предел

Сформулируем понятие предела для более общего случая: отображения из подмножества $\mathbb{R}^m$ в $\mathbb{R}^n$. Определение предела функции получается как частный случай при $n = 1$.

Рассмотрим отображение $f : \mathbb{R}^m \supset E \to \mathbb{R}^n$ (действующее из множества $E \subset \mathbb{R}^m$ в $\mathbb{R}^n$). Оно каждой точке $x = (x_1, ..., x_m) \in E \subset \mathbb{R}^m$ ставит в соответствие точку $y = (y_1, ..., y_n) \in \mathbb{R}^n$:

$$f : \begin{cases} y_1 = f_1(x_1, ..., x_m) \\ y_2 = f_2(x_1, ..., x_m) \\ ... \\ y_n = f_n(x_1, ..., x_m) \end{cases}$$

Определение 2.1.1 (По Коши) *Пусть $f : \mathbb{R}^m \supset E \to \mathbb{R}^n$, x_0 – предельная точка E. Говорят, что $A \in \bar{R}^n$ – предел отображения f при $x \to x_0$ (по Коши), если*

$$\forall V(A)\ \exists \overset{o}{U}(x_0) :\ \forall x \in \overset{o}{U} \cap E\ \Rightarrow\ f(x) \in V(A)$$

или

$$\forall \varepsilon > 0\ \exists \delta = \delta(\varepsilon) > 0 :\ \forall x \in E :\ 0 < \rho(x, x_0) < \delta\ \Rightarrow\ \rho(f(x), A) < \varepsilon.$$

Замечание 2.1.1 *Неравенства $0 < \rho(x, x_0) < \delta$ и $\rho(f(x), A) < \varepsilon$ равносильны неравенствам $0 < \|x - x_0\| < \delta$ и $\|f(x) - A\| < \varepsilon$, соответственно.*

Замечание 2.1.2 *В определении предела на языке ε-окрестностей функции расстояния ρ вообще говоря, разные (одно в R^m, другое в $\mathbb{R}^n$).*

Замечание 2.1.3 *Для случая $\lim\limits_{x \to \infty} f(x) = \infty$ можно записать*

$$\forall \varepsilon > 0\ \exists \delta = \delta(\varepsilon) > 0 :\ \forall x \in E :\ \rho(x, 0) > \frac{1}{\delta}\ \Rightarrow\ \rho(f(x), 0) > \frac{1}{\varepsilon}$$

или

$$\forall \varepsilon > 0\ \exists \delta = \delta(\varepsilon) > 0 :\ \forall x \in E :\ \|x\| > \frac{1}{\delta}\ \Rightarrow\ \|f(x)\| > \frac{1}{\varepsilon}.$$

Определение 2.1.2 (По Гейне) *Пусть $f : \mathbb{R}^m \supset E \to \mathbb{R}^n$ и $x_0 \in \bar{R}^m$ – предельная для E. Говорят, что $A \in \bar{R}^n$ – предел отображения f при $x \to x_0$ (по Гейне), если*

$$\forall x^k \in E,\ x^k \neq x_0,\ x^k \xrightarrow[k \to \infty]{} x_0\ \Rightarrow\ f(x^k) \xrightarrow[k \to \infty]{} A.$$

Теорема 2.1.1 *Определения предела по Коши и по Гейне эквивалентны.*

Доказательство. Доказывается аналогично случаю функции одной переменной. Полезно это проделать самостоятельно. $\square$

Лемма 2.1.1 *Пусть $f : \mathbb{R}^m \supset E \to \mathbb{R}^n$, x_0 – предельная точка для E. Тогда*

1. $f \xrightarrow[x\to x_0]{} A \in \mathbb{R}^n \;\Leftrightarrow\; f_i \xrightarrow[x\to x_0]{} A_i,\ i \in \{1, ..., n\}$ *(сходимость покоординатная);*
2. *Если* $\lim\limits_{x\to x_0} f(x) = A$ *и* $\lim\limits_{x\to x_0} f(x) = B$, *то* $A = B$ $(A, B \in \bar{R}^n)$ *(единственность предела);*
3. *Если* $\lim\limits_{x\to x_0} f(x) = A \in \mathbb{R}^n$, *то* $\exists U(x_0)$: f *– ограничена в* $E \cap U(x_0)$.

Доказательство. Из определения по Гейне. $\square$

Теорема 2.1.2 (Арифметические свойства) *Пусть $f, g : \mathbb{R}^m \supset E \to \mathbb{R}^n$, x_0 – предельная точка для E, $\lim\limits_{x\to x_0} f(x) = A$, $\lim\limits_{x\to x_0} g(x) = B$. Тогда*

1. $\lim\limits_{x\to x_0} (f + g) = A + B$;
2. *Пусть* $\lambda\colon E \to \mathbb{R}$, $\lambda(x) \xrightarrow[x\to x_0]{} \lambda$, *тогда* $\lim\limits_{x\to x_0} (\lambda(x) f(x)) = \lambda A$;
3. *Если* $n = 1$, *то* $\lim\limits_{x\to x_0} (f \cdot g) = AB$ *и* $\lim\limits_{x\to x_0} \frac{f(x)}{g(x)} = \frac{A}{B}$, $B \neq 0$.

Доказательство. Упражнение. $\square$

Пример.

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

Докажем, что $\lim\limits_{(x,y)\to(0,0)} f(x, y)$ не существует.

$$x_1^n = \left(0, \frac{1}{n}\right) \longrightarrow (0, 0), \quad \lim_{n\to\infty} f\left(0, \frac{1}{n}\right) = 0,$$

$$x_2^n = \left(\frac{1}{n}, \frac{1}{n}\right) \longrightarrow (0, 0), \quad \lim_{n\to\infty} f\left(\frac{1}{n}, \frac{1}{n}\right) = \frac{1}{2},$$

следуя определений по Гейне, двойной предел $\lim\limits_{(x,y)\to(0,0)} f(x, y)$ не существует.

Рассмотрим повторные пределы:

$$\lim_{x\to 0}\lim_{y\to 0} f(x,y) = 0, \quad \lim_{y\to 0}\lim_{x\to 0} f(x,y) = 0.$$

Пример.

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

Рассмотрим повторные пределы:

$$\lim_{x\to 0}\lim_{y\to 0} f(x,y) = 1, \quad \lim_{y\to 0}\lim_{x\to 0} f(x,y) = -1.$$

Двойной предел не существует, так как

$$f\left(0, \frac{1}{n}\right) \longrightarrow -1, \quad f\left(\frac{1}{n}, 0\right) \longrightarrow 1.$$

Пример.

$$f(x,y) = \begin{cases} x + y\cdot \sin\frac{1}{x}, & x \neq 0; \\ 0, & x = 0. \end{cases}$$

Рассмотрим повторные пределы:

$$\lim_{x\to 0}\lim_{y\to 0} f(x,y) = 0,$$

$$\lim_{y\to 0}\lim_{x\to 0} f(x,y) \text{ – не существует (кроме } y=0\text{)}.$$

Двойной предел:

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0 \text{ (произведение беск. малой на ограниченную).}$$

Пример.

$$f(x,y) = \begin{cases} \frac{x^2y}{x^4+y^2}, & x^2+y^2 \neq 0; \\ 0, & x^2+y^2 = 0. \end{cases}$$

$$f\left(\frac{1}{n}, 0\right) \longrightarrow 0, \quad f\left(\frac{1}{n}, \frac{1}{n^2}\right) \longrightarrow \frac{1}{2}.$$

Рассмотрим предел по направлению $x = \alpha t$, $y = \beta t$, $t \to 0+$:

$$f(\alpha t, \beta t) = \frac{\alpha^2\beta t^3}{\alpha^4t^4 + \beta^2t^2} \longrightarrow 0,$$

т.е. по любому направлению предел равен нулю, но двойной предел не существует, т.к.

$$f\left(\frac{1}{n}, \frac{1}{n^2}\right) \longrightarrow \frac{1}{2}.$$

Теорема 2.1.3 (О повторном пределе) *Пусть* $f\colon \mathbb{R}^2 \supset \overset{o}{U}(x_0, y_0) \to \mathbb{R}$, $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y) = A$ *и*

$$\exists \delta > 0: \ \forall y: \ 0 < |y - y_0| < \delta \ \ \exists \lim_{x\to x_0} f(x,y) = \varphi(y).$$

Тогда

$$\lim_{y\to y_0} \varphi(y) = A.$$

Доказательство. Пусть $\varepsilon > 0$. Тогда

$$\exists \delta_1 : \forall (x,y) \in \overset{o}{U}_{\delta_1}(x_0,y_0) \cap \overset{o}{U}(x_0,y_0) \ \Rightarrow \ |f(x,y) - A| < \varepsilon.$$

Возьмем $\delta_2 = \min\{\delta_1, \delta\}$. Тогда перейдем к пределу при $0 < |y - y_0| < \delta_2$:

$$\lim_{x\to x_0} |f(x,y) - A| = |\varphi(y) - A| \leqslant \varepsilon \ \Leftrightarrow \ \lim_{y\to y_0} \varphi(y) = A.$$

$\square$

Теорема 2.1.4 (О вычислении двойного предела в полярных координатах) *Пусть* $f(x,y)\colon \overset{o}{U}(x_0,y_0) \to \mathbb{R}$. *Если* $\exists \rho_0 > 0: \ \forall \varphi \in [0, 2\pi), \ \forall \rho \in (0, \rho_0) \ \ \Rightarrow$

$$\big|f(x_0 + \rho\cos\varphi, y_0 + \rho\sin\varphi) - A\big| \leqslant F(\rho),$$

где $F(\rho) \to 0$ *при* $\rho \to 0+$, *то*

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = A.$$

Доказательство. Пусть $\varepsilon > 0$. Тогда

$$\exists \delta > 0 : \forall \rho \in (0, \delta) \ \Rightarrow \ \big|f(x_0 + \rho\cos\varphi, y_0 + \rho\sin\varphi) - A\big| \leqslant F(\rho) < \varepsilon,$$

$$0 < \sqrt{(x - x_0)^2 + (y - y_0)^2} = \rho < \delta.$$

$\square$

Пример.

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

$$\left|\frac{x^2 y}{x^2 + y^2}\right| = \rho\cos^2\varphi |\sin\varphi| \leqslant \rho \to 0.$$

Теорема 2.1.5 (Критерий Коши) *Пусть* $f\colon \mathbb{R}^m \supset E \to \mathbb{R}^n$, $x_0 \in \bar{\mathbb{R}}^m$ – *предельная для* E. *Тогда*

$$\exists \lim_{x\to x_0} f(x) \in \mathbb{R}^n \quad \Leftrightarrow$$

$$\forall \varepsilon > 0 \ \exists \delta = \delta(\varepsilon) > 0 : \ \forall x, y \in E \cap \overset{o}{U}_\delta(x_0) : \Rightarrow \ \|f(x) - f(y)\| < \varepsilon.$$

Доказательство. Докажите самостоятельно. $\square$

2.2 Непрерывность отображения

Определение 2.2.1 *Пусть* $f : \mathbb{R}^m \supset E \to \mathbb{R}^n$*. Говорят, что* f *непрерывно в точке* $x_0 \in E$*, если*

$$\forall V(f(x_0))\ \exists U(x_0) :\ f(U(x_0)) \subset V(f(x_0)).$$

Для $x_0 \in E$ возможно два случая:

1. x_0 – предельная точка для E. Тогда непрерывность f в x_0 равносильна тому, что $\lim\limits_{x\to x_0} = f(x_0)$.
2. x_0 – изолированная точка E. Тогда f непрерывно в x_0 всегда, так как в достаточно маленькой окрестности x_0 нет других точек из E.

Теорема 2.2.1 (Локальные свойства непрерывных отображений) *Пусть* $f, g : \mathbb{R}^m \supset E \to \mathbb{R}^n$*,* f *и* g *непрерывны в точке* $x_0 \in E$*. Тогда*

1. $f + g$ *непрерывно в* x_0*;*
2. *пусть* λ*:* $E \to \mathbb{R}$ *и* λ *непрерывно в* x_0*, тогда* λf *непрерывно в* x_0*;*
3. f *ограничено в* $U(x_0)$*;*
4. *при* $n = 1$ $f \cdot g$ *непрерывно в* x_0*;*
5. *при* $n = 1$ f/g *непрерывно в* x_0*, если* $g(x_0) \neq 0$.

Доказательство. Для предельной точки доказательство непосредственно следует из локальных свойств предела. Для изолированной – предоставляется читателю в качестве упражнения. □

Теорема 2.2.2 (О непрерывности композиции) *Пусть* $f : \mathbb{R}^m \supset E_1 \to E_2 \subset \mathbb{R}^n$*,* g*:* $E_2 \to \mathbb{R}^k$*,* $x_0 \in E_1$*,* f *непрерывно в* x_0*,* g *непрерывно в* $f(x_0)$*. Тогда* $g \circ f$ *непрерывно в* x_0*.*

Доказательство. Так как g непрерывно в $f(x_0)$, то

$$\forall U(g(f(x_0)))\ \exists U(f(x_0)) :\ \forall x \in U(f(x_0)) \cap E_2\ \Rightarrow g(f(x)) \in U(g(f(x_0))).$$

Так как f непрерывно в x_0, то по окрестности $U(f(x_0))$

$$\exists U(x_0) :\ \forall x \in U(x_0) \cap E_1\ \Rightarrow\ f(x) \in U(f(x_0)),$$

это и означает непрерывность $g(f)$ в x_0. □

Замечание 2.2.1 *Если $f(x)$ непрерывно на X, то $f(x,y) = f(x)$ (при $y \in \mathbb{R}$) непрерывно на $X \times \mathbb{R}$.*

Пример. Функция

$$f(x,y) = 1 + e^{-xy} \cdot \log_2(1 + |x| + 4|y|)$$

непрерывна на $\mathbb{R}^2$, так как получается из непрерывных функций конечным числом арифметических операций и суперпозиций.

Определение 2.2.2 *Пусть $F \subset E \subset \mathbb{R}^n$.*

1. *Говорят, что $x_0 \in F$ является внутренней точкой для F в E, если*

$$\exists B(x_0, r) \subset \mathbb{R}^n: \ B(x_0, r) \cap E \subset F;$$

2. *F называется открытым в E, если все точки F внутренние в E;*

3. *F называется замкнутым в E, если $E \setminus F$ открыто в E.*

Пример. $E = [0,1] \subset \mathbb{R}$, $F = \left(\frac{1}{2}, 1\right]$. Точка 1 является внутренней точкой для F в E, а множество F открыто в E.

Определение 2.2.3 *Пусть $f: \mathbb{R}^m \supset E \to \mathbb{R}^n$ и f непрерывно в каждой точке $F \subset E$. Тогда говорят, что f непрерывно на F и пишут $f \in C(F)$.*

Теорема 2.2.3 (Критерий непрерывности) *Пусть $f: \mathbb{R}^m \supset E \to \mathbb{R}^n$. f непрерывно на E тогда и только тогда, когда прообраз любого открытого в $\mathbb{R}^n$ множества открыт в E.*

Доказательство. 1. Пусть $f \in C(E)$ и G – открыто в $\mathbb{R}^n$. Рассмотрим $F = f^{-1}(G)$ – не пусто. Пусть $x_0 \in F$ и $V(f(x_0))$ – окрестность точки $f(x_0)$ из G. Тогда

$$\exists U_V(x_0): \ \forall x \in U_V(x_0) \cap E \ \Rightarrow \ f(x) \in V(f(x_0)) \ \Rightarrow \ U_V(x_0) \cap E \subset f^{-1}(G).$$

2. Пусть прообраз любого открытого множества открыт в E и $x_0 \in E$:

$$\forall U(f(x_0)) \ \Rightarrow \ f^{-1}\Big(U(f(x_0))\Big) \text{ – открыт в } E$$

и является окрестностью точки x_0. □

Замечание 2.2.2 *Аналогичное утверждение верно для замкнутого множества.*

Теорема 2.2.4 *Пусть $f : \mathbb{R}^m \supset E \to \mathbb{R}^n$, $f \in C(E)$ и E – компакт. Тогда $f(E)$ – компакт.*

Другими словами, образ компакта при непрерывном отображении – компакт.
Доказательство. Пусть G_α, $\alpha \in A$ – открытое покрытие $f(E)$. Так как f непрерывно, то множества $f^{-1}(G_\alpha)$ открыты в E и образуют покрытие E. Выделим конечное покрытие: $E \subset \bigcup_{i=1}^n f^{-1}(G_{\alpha_i})$. Следовательно, $f(E) \subset \bigcup_{i=1}^n G_{\alpha_i}$, а значит, $f(E)$ – компакт. □

Замечание 2.2.3 *Прообраз компакта при непрерывном отображении не обязательно компакт. Например, непрерывное биективное отображение полуинтервала на окружность $[0, 2\pi) \to S_1(0)$. При этом обратное отображение не является непрерывным.*

Замечание 2.2.4 *Отрезок в $\mathbb{R}^n$ – компакт.*

Доказательство. Пусть $a, b \in \mathbb{R}^n$. Тогда

$$[a, b] = \{x : x = a + t(b - a), t \in [0, 1]\}$$

и функция $x(t)$ – непрерывна на компакте $[0, 1]$. □

Определение 2.2.4 *Множество $G \subset \mathbb{R}^n$ называется линейно-связным (связным), если для любых $a, b \in G$ существует непрерывное отображение (путь) γ с концами a и b и носителем в G.*

Определение 2.2.5 *Областью в $\mathbb{R}^n$ называется открытое связное множество.*

Определение 2.2.6 *Пусть $f : \mathbb{R}^m \supset E \to \mathbb{R}^n$. Говорят, что f равномерно непрерывно на $D \subset E$, если*

$$\forall \varepsilon > 0\ \exists \delta = \delta(\varepsilon) > 0 :\ \forall x_1, x_2 \in D :\ \|x_1 - x_2\| < \delta\ \Rightarrow\ \|f(x_1) - f(x_2)\| < \varepsilon.$$

Теорема 2.2.5 (Глобальные свойства непрерывных отображений) *Пусть $f : \mathbb{R}^m \supset E \to \mathbb{R}^n$ и $f \in C(E)$.*

1. ***Теорема Кантора.*** *Если E – компакт, то f равномерно непрерывно на E.*
2. ***Теорема Вейерштрасса.*** *Если $n = 1$ и E – компакт, то f достигает наибольшего и наименьшего значений.*
3. ***Теорема Больцано-Коши.*** *Если $n = 1$ и E – связно, то $\forall a, b \in E$ и $\forall \gamma$, лежащего между $f(a)$ и $f(b)$: $\exists c \in E$: $f(c) = \gamma$.*

Доказательство. Докажем Теорему Больцано–Коши. Пусть $\varphi : [\alpha, \beta] \subset \mathbb{R} \to E$, $\varphi(\alpha) = a$, $\varphi(\beta) = b$. $f(\varphi(t))$ непрерывна на $[\alpha, \beta]$ как композиция непрерывных функций. $f(\varphi(\alpha)) = f(a)$, $f(\varphi(\beta)) = f(b)$. Применим теорему Больцано–Коши для функции в $\mathbb{R}$. □

3 Многомерное дифференциальное исчисление

3.1 Производная и дифференциал

Определение 3.1.1 *Пусть* $f : \mathbb{R}^m \supset E \to \mathbb{R}^n$, x_0 *– внутренняя точка* E. *Если существует такой линейный оператор* $A_f \in L(\mathbb{R}^m, \mathbb{R}^n)$, *что*

$$f(x_0 + h) = f(x_0) + A_f h + o(\|h\|), \quad h \to 0,$$

то говорят, что f *дифференцируемо в точке* x_0.

Замечание 3.1.1 *В определении выше запись* $o(\|h\|)$ *означает функцию, представимую в виде* $\alpha(h)\|h\|$, *где* $\alpha(h) \to 0$ *при* $h \to 0$. *При этом значение* $\alpha(0)$ *может быть не определено. Будем полагать* $\alpha(0) = 0$, *тогда* α *непрерывна в нуле.*

Определение 3.1.2 *Линейный оператор* A_f *в определении выше называется производной отображения* f *в точке* x_0, *а величина* $A_f h$ *– дифференциалом* f *в точке* x_0:

$$f'(x_0) = A_f, \quad df(x_0, h) = A_f h.$$

Также будем использовать обозначения $A_f(x_0)$, $A(x_0)$, $A_f(x_0)h$ *и т.п.*

Пример. Функция двух переменных $f : \mathbb{R}^2 \to \mathbb{R}$ $f(x, y) = x^2 + xy$ в точке (x_0, y_0). Зададим приращение $h = (h_x, h_y)$ и рассмотрим

$$\begin{multline} f(x_0 + h_x, y_0 + h_y) = (x_0 + h_x)^2 + (x_0 + h_x)(y_0 + h_y) = \\ = f(x_0, y_0) + (2x_0 + y_0)h_x + x_0 h_y + h_x^2 + h_x h_y. \quad (1) \end{multline}$$

Здесь $A_f = (2x_0 + y_0,\ x_0)$ и

$$|h_x^2 + h_x h_y| \leqslant |h_x| \cdot \big(|h_x| + |h_y|\big) \leqslant 2(h_x^2 + h_y^2) = o(\|h\|).$$

Лемма 3.1.1 (Необходимое условие дифференцируемости) *Пусть* f *– дифференцируемо в точке* x_0. *Тогда* f *непрерывно в точке* x_0.

Доказательство. По определению имеем

$$f(x_0 + h) - f(x_0) = A_f h + \alpha(h)\|h\|,$$

и при $h \to 0$ оба слагаемых стремятся к 0, следовательно, $f(x_0+h) - f(x_0) \to 0$, что и означает непрерывность f в точке x_0. □

Определение 3.1.3 *Пусть $f : \mathbb{R}^m \supset E \to \mathbb{R}^n$, x_0 – внутренняя точка множества E. И пусть $e \in \mathbb{R}^m \setminus \{0\}$. Производной f по направлению e называется*

$$\frac{\partial f}{\partial e} = \lim_{t \to 0+} \frac{f(x_0 + t \cdot e_0) - f(x_0)}{t},$$

где e_0 – орт вектора e: $e_0 = e/\|e\|$.

Определение 3.1.4 *Пусть $f : \mathbb{R}^m \supset E \to \mathbb{R}^n$, x_0 – внутренняя точка множества E. Частной производной отображения f по переменной x_i в точке x_0 будем называть*

$$\frac{\partial f}{\partial x_i} = \lim_{t \to 0} \frac{f(x_0 + t \cdot e_i) - f(x_0)}{t},$$

где e_i – i-ый базисный орт пространства $\mathbb{R}^m$.

Замечание 3.1.2 *Частные производные не равны производным по направлениям соответствующих базисных ортов. Если (иногда такие определения удобны) в определении производной по направлению рассматривать двусторонни предел при $t \to 0$, то частные производные будут совпадать с производными по направлениям соответствующих ортов.*

Если $f = (f_1, ..., f_n)$, то частная производная по x_i – это вектор в $\mathbb{R}^n$.

$$\frac{\partial f}{\partial x_i} = \left(\frac{\partial f_1}{\partial x_i}, \frac{\partial f_2}{\partial x_i}, \ ..., \ \frac{\partial f_n}{\partial x_i} \right).$$

Пример 3.1.1 *Пусть $f(x, y) = x^y$: $\{(x, y) \in \mathbb{R}^2, x > 0\} \to \mathbb{R}$. Тогда*

$$\frac{\partial f}{\partial x} = y \cdot x^{y-1}, \quad \frac{\partial f}{\partial y} = x^y \cdot \ln x.$$

Теорема 3.1.1 (Необходимое условие дифференцируемости) *Пусть $f : \mathbb{R}^m \supset E \to \mathbb{R}^n$, x_0 – внутренняя точка множества E. Если f дифференцируемо в точке x_0, то для любого вектора $e \neq 0$ существует $\frac{\partial f}{\partial e}$, а также существуют все частные производные.*

Доказательство. Рассмотрим $h = t \cdot e$:

$$f(x_0 + h) - f(x_0) = Ah + \alpha(h) \cdot \|h\|, \quad h \to 0,$$

воспользуемся линейностью A,

$$f(x_0 + te) - f(x_0) = Ate + \alpha(te) \cdot \|te\| = t \cdot A(e) + \alpha(te) \cdot |t| \cdot \|e\|,$$

разделим на t:

$$\frac{f(x_0+te)-f(x_0)}{t} = Ae + \alpha(te)\cdot\|e\|\cdot\frac{|t|}{t} \to A(e), \quad t\to 0,$$

так как второе слагаемое стремится к нулю как произведение бесконечно малой $\alpha(te)$ на ограниченную $\|e\|\operatorname{sign} t$. $\square$

Величина Ae выражается матрицей

$$T = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_m} \end{pmatrix},$$

которую принято называть матрицей Якоби.

Другими словами, вектор приращения функции должен иметь вид

$$\Delta f = \begin{pmatrix} \Delta f_1 \\ \vdots \\ \Delta f_n \end{pmatrix} = \begin{pmatrix} A_{11} & \dots & A_{1m} \\ \dots & \dots & \dots \\ A_{n1} & \dots & A_{nm} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_m \end{pmatrix} + o\left(\|h\|\right).$$

Для отображения $f(x)=x$ имеем

$$x_i + h_i - x_i = 1\cdot h_i \;\Leftrightarrow\; h_i = dx_i(x_0,h), \quad i=1,...,m.$$

И тогда для произвольного дифференцируемого f пишут

$$df(x_0,h) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \cdots & \frac{\partial f_1}{\partial x_m}(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(x_0) & \cdots & \frac{\partial f_n}{\partial x_m}(x_0) \end{pmatrix} \begin{pmatrix} dx_1(x_0,h) \\ \vdots \\ dx_m(x_0,h) \end{pmatrix},$$

или короче:

$$df = f'dx = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_m} \end{pmatrix} \begin{pmatrix} dx_1 \\ \vdots \\ dx_m \end{pmatrix}.$$

3.2 Правила дифференцирования

Теорема 3.2.1 (Арифметические свойства) *Пусть $f, g\colon \mathbb{R}^m \supset E \to \mathbb{R}^n$, f, g дифференцируемы в точке $x_0 \in E$. Тогда*

1. *$\forall \lambda, \mu \in \mathbb{R}$: $\lambda f + \mu g$ дифференцируема в x_0, причем*
$$A_{\lambda f+\mu g} = \lambda \cdot A_f + \mu \cdot A_g.$$
2. *Пусть $\lambda\colon E \to \mathbb{R}$, дифференцируема в x_0. Тогда*
$$A_{\lambda f} = f \cdot A_\lambda + \lambda \cdot A_f.$$
3. *Пусть $\lambda\colon E \to \mathbb{R}$, дифференцируема в x_0 и $\lambda(x_0) \neq 0$. Тогда*
$$A_{f/\lambda} = \frac{f \cdot A_\lambda - \lambda \cdot A_f}{\lambda^2}.$$

Доказательство. 1. Имеем
$$f(x_0+h) - f(x_0) = A_f h + o(\|h\|), \quad g(x_0+h) - g(x_0) = A_g h + o(\|h\|),$$
тогда
$$(\lambda f+\mu g)(x_0+h)-(\lambda f+\mu g)(x_0) = \lambda(f(x_0+h)-f(x_0))+\mu(g(x_0+h)-g(x_0)) = \\ = (\lambda A_f + \mu A_g)h + (\lambda+\mu)o(\|h\|) = (\lambda A_f + \mu A_g)h + o(\|h\|).$$
2. Аналогично:
$$(\lambda f)(x_0+h) - (\lambda f)(x_0) = \lambda(x_0+h)f(x_0+h) - \lambda(x_0)f(x_0) = \\ = \Big(\lambda(x_0) + A_\lambda h + o(\|h\|)\Big)\Big(f(x_0) + A_f h + o(\|h\|)\Big) - \lambda(x_0)f(x_0) = \\ = \lambda(x_0)A_f h + A_\lambda h f(x_0) + o(\|h\|) = (\lambda(x_0)A_f + f(x_0)A_\lambda)h + o(\|h\|),$$
где для получения $o(\|h\|)$ мы воспользовались непрерывностью линейных операторов A_λ и A_f.
3. Докажите самостоятельно. □

Теорема 3.2.2 (Дифференцирование композиции) *Пусть $g\colon \mathbb{R}^m \supset E \to F \subset \mathbb{R}^n$, $f\colon F \to \mathbb{R}^k$, g – дифференцируема в точке x_0, f – дифференцируема в точке $g(x_0)$. Тогда $f \circ g$ дифференцируема в точке x_0 и $A_{f\circ g} = A_f \circ A_g$, то есть*
$$(f(g))'(x_0) = f'(g(x_0)) \cdot g'(x_0).$$

Доказательство. Запишем определения дифференцируемости

$$f(x_0+h)-f(x_0)=A_fh+\alpha(h)\cdot\|h\|,\quad \alpha(h)\xrightarrow[h\to 0]{}0,\ \alpha(0)=0,$$

$$g\left(f(x_0)+t\right)-g\left(f(x_0)\right)=A_gt+\beta(t)\cdot\|t\|,\quad \beta(t)\xrightarrow[t\to 0]{}0,\ \beta(0)=0.$$

Пусть $t=f(x_0+h)-f(x_0)$. Заметим, что при $h\to 0$ выполнено $t\to 0$. Тогда

$$g\left(f(x_0+h)\right)-g\left(f(x_0)\right)=A_g\left(A_fh+\alpha(h)\|h\|\right)+\beta(t)\cdot\|t\|=$$

подставим $\|t\|=\Big\|A_fh+\alpha(h)\|h\|\Big\|$

$$=A_g\cdot A_fh+A_g(\alpha(h))\cdot\|h\|+\beta(t)\cdot\Big\|A_fh+\alpha(h)\cdot\|h\|\Big\|.$$

Рассмотрим второе слагаемое и применим свойство ограниченности линейного оператора

$$\|A_g(\alpha(h))\cdot\|h\|\|\leqslant C_g\|\alpha(h)\|\cdot\|h\|=o(\|h\|),$$

где $C_g=\left(\sum_{i,j}a_{ij}^2\right)^{1/2}$ и $\{a_{ij}\}$ – матрица оператора A_g.

Для третьего слагаемого имеем аналогично

$$\Big\|\beta(t)\cdot\Big\|A_fh+\alpha(h)\cdot\|h\|\Big\|\ \Big\|=o(\|h\|),$$

и тогда $g\left(f(x_0+h)\right)-g\left(f(x_0)\right)=A_g\cdot A_fh+o(\|h\|)$. □

Следствие 3.2.3 (Инвариантность формы первого дифференциала) *Выражение для дифференциала отображения $df=A_fdx=f'dx$ не зависит от того, является ли x зависимой или независимой переменной, а также от того, независимы ли компоненты x_1, ..., x_m вектора x.*

Теперь зададимся вопросом о производной обратного отображения. Для его существования необходимо равенство размерностей $m=n$.

Теорема 3.2.4 (О производной обратного отображения) *Пусть $f\colon\mathbb{R}^m\supset E\to\mathbb{R}^m$, x_0 – внутренняя точка множества E, $f(x_0)$ – внутренняя для $f(E)$, f дифференцируемо в x_0 и имеет обратное отображение f^{-1} – непрерывное в $f(x_0)$, и оператор A_f обратим. Тогда*

$$A_{f^{-1}}=A_f^{-1}.$$

Доказательство. Пусть $y_0 = f(x_0)$. Зададим приращение h и

$$f(x_0+h) - f(x_0) = A_f h + \alpha(h)\cdot\|h\|, \quad \alpha(h) \xrightarrow[h\to 0]{} 0,\ \alpha(0)=0.$$

Зададим точке $f(x_0)$ приращение t и возьмём

$$h = f^{-1}\left(f(x_0)+t\right) - f^{-1}\left(f(x_0)\right).$$

В силу непрерывности f^{-1} имеем $t\to 0 \Leftrightarrow h \to 0$. Тогда

$$f(x_0+h) - f(x_0) = t = A_f h + \alpha(h)\cdot\|h\|$$

или

$$A_f h = t - \alpha(h)\cdot\|h\|.$$

Подействуем оператором A_f^{-1}:

$$h = A_f^{-1}t - A_f^{-1}\left[\alpha(h)\right]\cdot\|h\|.$$

Так как $\left\|A_f^{-1}[\alpha(h)]\right\| \leqslant \left\|A_f^{-1}\right\|\cdot\|\alpha(h)\| \to 0$, то $A_f^{-1}\left[\alpha(h)\right]\cdot\|h\| = o(\|t\|)$, и осталось доказать, что $\frac{\|h\|}{\|t\|}$ ограничено. Будем считать, что $\left\|A_f^{-1}[\alpha(h)]\right\| < \frac{1}{2}$ и рассмотрим

$$\frac{\|h\|}{\|t\|} = \frac{\left\|A_f^{-1}t - A_f^{-1}\left[\alpha(h)\right]\cdot\|h\|\right\|}{\|t\|} \leqslant \frac{\|A_f^{-1}t\|}{\|t\|} + \frac{\left\|A_f^{-1}\left[\alpha(h)\right]\right\|\cdot\|h\|}{\|t\|} \leqslant$$

$$\leqslant \|A_f^{-1}\| + \frac{1}{2}\frac{\|h\|}{\|t\|} \quad\Rightarrow\quad \frac{\|h\|}{\|t\|} \leqslant 2\|A_f^{-1}\|,$$

откуда следует, что $A_f^{-1}\left[\alpha(h)\right]\cdot\|h\| = o(\|t\|)$ и

$$h = f^{-1}\left(f(x_0)+t\right) - f^{-1}\left(f(x_0)\right) = A_f^{-1}t + o(\|t\|),$$

что и означает $A_{f^{-1}} = A_f^{-1}$. □

3.3 Достаточное условие дифференцируемости

Определение 3.3.1 *Будем говорить, что f дифференцируемо на E, если f дифференцируемо в каждой точке $x_0 \in E$.*

Теорема 3.3.1 (Достаточное условие дифференцируемости) *Пусть $f\colon \mathbb{R}^m \supset E \to \mathbb{R}^n$, $x_0 \in E$ – внутренняя точка множества E. Если все частные производные $\dfrac{\partial f}{\partial x_i}(x)$, $i = 1,\dots,m$ определены в окрестности точки x_0 и непрерывны в точке x_0, то функция $f(x)$ дифференцируема в точке x_0.*

Доказательство. Так как дифференцируемость отображения f равносильна дифференцируемости всех f_i, то докажем для случая $n = 1$. И пусть $m = 2$ (при $m > 2$ доказательство аналогично).

Пусть $\dfrac{\partial f}{\partial x}(x, y)$ и $\dfrac{\partial f}{\partial y}(x, y)$ определены в шаре $B_\delta(x_0, y_0)$ и непрерывны в точке (x_0, y_0).

Пусть $x = x_0 + \Delta x$, $y = y_0 + \Delta y$. Запишем полное приращение функции:

$$\Delta f\left(x_0, y_0\right) = f(x, y) - f\left(x_0, y_0\right) = \left(f(x, y) - f\left(x_0, y\right)\right) + \left(f\left(x_0, y\right) - f\left(x_0, y_0\right)\right).$$

Рассмотрим функцию $f(x, y)$ как функцию одной переменной x. Тогда по теореме Лагранжа (для функции одной переменной) найдется точка ξ, лежащая между x и x_0 такая, что

$$f(x, y) - f\left(x_0, y\right) = \frac{\partial f}{\partial x}(\xi, y)\left(x - x_0\right).$$

Так как $\dfrac{\partial f}{\partial x}$ непрерывна в точке (x_0, y_0), то

$$\frac{\partial f}{\partial x}(\xi, y) = \frac{\partial f}{\partial x}\left(x_0, y_0\right) + \alpha(\xi, y), \quad \alpha(\xi, y) \to 0 \quad \text{при } (x, y) \to \left(x_0, y_0\right).$$

Аналогично по переменной y получим

$$f\left(x_0, y\right) - f\left(x_0, y_0\right) = \frac{\partial f}{\partial y}\left(x_0, \psi\right)\left(y - y_0\right),$$

$$\frac{\partial f}{\partial y}\left(x_0, \psi\right) = \frac{\partial f}{\partial y}\left(x_0, y_0\right) + \beta(\psi), \quad \beta(\psi) \to 0 \quad \text{при } (x, y) \to \left(x_0, y_0\right).$$

Тогда приращение функции имеет вид

$$\Delta f\left(x_0, y_0\right) = \frac{\partial f}{\partial x}\left(x_0, y_0\right)\Delta x + \frac{\partial f}{\partial y}\left(x_0, y_0\right)\Delta y + \alpha(\xi, y)\Delta x + \beta(\psi)\Delta y.$$

Докажем, что $\alpha(\xi, y)\Delta x + \beta(\psi)\Delta y = o\left(\sqrt{\Delta x^2 + \Delta y^2}\right)$ при $\Delta x \to 0$, $\Delta y \to 0$. Имеем

$$\left|\frac{\alpha(\xi, y)\Delta x + \beta(\psi)\Delta y}{\sqrt{\Delta x^2 + \Delta y^2}}\right| \leqslant \left|\frac{\alpha(\xi, y)\Delta x}{\sqrt{\Delta x^2 + \Delta y^2}}\right| + \left|\frac{\beta(\psi)\Delta y}{\sqrt{\Delta x^2 + \Delta y^2}}\right| \leqslant |\alpha(\xi, y)| + |\beta(\psi)| \to 0.$$

$\square$

Замечание 3.3.1 *Функции, имеющие непрерывные частные производные в E (а значит и дифференцируемые в E) называют непрерывно-дифференцируемыми на E и обозначают $C^1(E)$.*

Замечание 3.3.2 *Непрерывность частных производных не является необходимым условием дифференцируемости функции.*

Например, функция

$$f(x, y)=\begin{cases}\left(x^2+y^2\right) \sin \frac{1}{\sqrt{x^2+y^2}}, & x^2+y^2>0; \\ 0, & x^2+y^2=0,\end{cases}$$

Дифференцируема в точке $(0,0)$, так как

$$\Delta f(0,0)=0 \cdot x+0 \cdot y+o\left(x^2+y^2\right) \quad \text{при} (x, y) \rightarrow(0,0),$$

но частная производная

$$\frac{\partial f}{\partial x}=2 x \sin \frac{1}{\sqrt{x^2+y^2}}-\frac{x}{\sqrt{x^2+y^2}} \cos \frac{1}{\sqrt{x^2+y^2}}$$

не имеет предела при $(x, y) \rightarrow(0,0)$ (доказать это можно, рассмотрев предел по множеству $y=0$ - он не существует), а значит, и не является непрерывной в точке $(0,0)$ функцией.

Пример 3.3.1 *Для функции $f(x, y)$ получить выражения для производных в полярных координатах, т.е. найти $\dfrac{\partial f}{\partial r}, \dfrac{\partial f}{\partial \varphi}$.*

Напомним формулы перехода в полярные координаты: $x=r \cos \varphi$, $r \sin \varphi$. Получаем

$$\frac{\partial f}{\partial r}=\frac{\partial f}{\partial x} \frac{\partial x}{\partial r}+\frac{\partial f}{\partial y} \frac{\partial y}{\partial r}=\frac{\partial f}{\partial x} \cos \varphi+\frac{\partial f}{\partial y} \sin \varphi=\frac{x}{\sqrt{x^2+y^2}} \frac{\partial f}{\partial x}+\frac{y}{\sqrt{x^2+y^2}} \frac{\partial f}{\partial y}.$$

$$\frac{\partial f}{\partial \varphi}=\frac{\partial f}{\partial x}(-r \sin \varphi)+\frac{\partial f}{\partial y} r \cos \varphi=-y \frac{\partial f}{\partial x}+x \frac{\partial f}{\partial y}.$$

3.4 Градиент и касательная плоскость

Рассмотрим вещественнозначную функцию $f: \mathbb{R}^m \subset E \rightarrow \mathbb{R}$, дифференцируемую во внутренней точке x_0 множества E.

Дифференциал $df(x_0, h)=f'(x_0) h$ является линейной функцией вектора приращения h, а значит найдется такой вектор $\xi \in \mathbb{R}^m$, что дифференциал выражается скалярным произведением вектора : $df(x_0, h)=\xi \cdot h$.

Определение 3.4.1 *Градиентом функции f в точке x_0 называется вектор* $\operatorname{grad} f(x_0)$ *такой, что*

$$df(x_0, h)=\operatorname{grad} f(x_0) \cdot h.$$

Так как в координатном представлении дифференциал имеет вид

$$df(x_0,h)=\frac{\partial f}{\partial x_1}(x_0)h_1+\frac{\partial f}{\partial x_2}(x_0)h_2+\ldots+\frac{\partial f}{\partial x_m}(x_0)h_m,$$

то градиент имеет вид

$$\operatorname{grad} f(x_0)=f'(x_0)=\Big(\frac{\partial f}{\partial x_1}(x_0),\frac{\partial f}{\partial x_2}(x_0),\ldots,\frac{\partial f}{\partial x_m}(x_0)\Big).$$

Свойства градиента:

1. Производная по направлению:

$$\frac{\partial f}{\partial l}=\operatorname{grad} f\cdot l_0=\mathrm{Pr}_l \operatorname{grad} f=df(l_0),$$

т.е. производная по направлению равна скалярному произведению градиента на орт направления или проекции градиента на вектор направления, что тоже самое, что и значение дифференциала на орте направления.

2. $\max\limits_{l}\dfrac{\partial f}{\partial l}=\dfrac{\partial f}{\partial(\operatorname{grad} f)}=|\operatorname{grad} f|.$
Т.е. в направлении градиента производная по направлению максимальна и равна норме градиента.

3. $\operatorname{grad} f(x_0)$ ортогонален любой гладкой кривой, лежащей на поверхности уровня $f(x)=C$ и проходящей через точку x_0.

Доказательство. 1. Так как f дифференцируема в точке x_0, то $f(x_0+h)-f(x_0)=f'(x_0)h+o(\|h\|)$. Применим для $h=tl_0$:

$$\frac{\partial f}{\partial l}(x_0)=\lim_{t\to 0+}\frac{f(x_0+tl_0)-f(x_0)}{t}=\lim_{t\to 0+}\frac{f'(x_0)(tl_0)+o(|t|)}{t}=$$
$$=f'(x_0)l_0=\operatorname{grad} f(x_0)\cdot l_0.$$

3. Пусть кривая $\gamma:[a,b]\to\mathbb{R}^m$ лежит на поверхности уровня $f(x)=C$ и задается дифференцируемой функцией $\gamma(t)=(\gamma_1(t),\ldots\gamma_m(t))$. Точке x_0 соответствует t_0: $\gamma(t_0)=x_0$, $f(x_0)=C$.

Тогда при всех $t\in[a,b]$ верно равенство $f(\gamma(t))=C$. Дифференцируя его по t как суперпозицию отображений, получим в точке $t=t_0$:

$$\Big(\frac{\partial f}{\partial x_1},\ldots,\frac{\partial f}{\partial x_m}\Big)\cdot\Big(\gamma_1',\ldots,\gamma_m'\Big)=0,$$

что и означает ортогональность векторов $\operatorname{grad} f$ и направляющего вектора γ' в точке t_0. $\square$

Рассмотрим поверхность уровня в $\mathbb{R}^3$, заданную равенством $F(x, y, z) = C$. Пусть точка $M_0 = (x_0, y_0, z_0, C)$ лежит на поверхности, т.е. $F(x_0, y_0, z_0) = C$.

Из свойства 3 следует, что если $\operatorname{grad} F(M_0) \neq 0$, то касательные ко всем гладким кривым, лежащим на поверхности уровня $F(x, y, z) = C$ и проходящим через точку M_0 имеют общую нормаль $\operatorname{grad} F(M_0)$, а значит, лежат в одной плоскости. Эта плоскость называется **касательной плоскостью** (гиперплоскостью) к поверхности $F(x, y, z) = C$ в точке M_0.

Прямая, проходящая через точку M_0 перпендикулярно касательной плоскости, называется **нормалью** к поверхности.

Тогда касательная плоскость в точке M_0 имеет нормаль $\operatorname{grad} F(x_0)$ и уравнение касательной плоскости в точке M_0 имеет вид:

$$\frac{\partial F}{\partial x}(M_0)(x - x_0) + \frac{\partial F}{\partial y}(M_0)(y - y_0) + \frac{\partial F}{\partial z}(M_0)(z - z_0) = 0.$$

Уравнение нормали:

$$\frac{x - x_0}{\frac{\partial F}{\partial x}(M_0)} = \frac{y - y_0}{\frac{\partial F}{\partial y}(M_0)} = \frac{z - z_0}{\frac{\partial F}{\partial z}(M_0)}.$$

4 Производные и дифференциалы высших порядков

4.1 Частные производные высших порядков

Рассмотрим функцию двух переменных $u = f(x, y) : \mathbb{R}^2 \supset E \to \mathbb{R}$. Пусть в окрестности точки (x, y) существуют частные производные $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$. Эти частные производные также являются функциями двух переменных от x и y. Если существуют частные производные от этих функций, то они называются частными производными второго порядка от функции f и обозначаются:

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

Другое обозначение:

$$f''_{xx} = \frac{\partial^2 f}{\partial x^2}, \quad f''_{yy} = \frac{\partial^2 f}{\partial y^2}, \quad f''_{xy} = (f'_x)'_y = \frac{\partial^2 f}{\partial y \partial x}, \quad f''_{yx} = \left(f'_y\right)'_x = \frac{\partial^2 f}{\partial x \partial y}.$$

Производные, взятые по разным переменным, называются **смешанными производными**.

Пример 4.1.1 $f(x, y) = x^3y^2 + xy^4$.

$$\frac{\partial f}{\partial x} = 3x^2y^2 + y^4, \quad \frac{\partial f}{\partial y} = 2x^3y + 4xy^3.$$

$$\frac{\partial^2 f}{\partial x^2} = \left(3x^2y^2 + y^4\right)'_x = 6xy^2, \quad \frac{\partial^2 f}{\partial y^2} = \left(2x^3y + 4xy^3\right)'_y = 2x^3 + 12xy^2,$$

$$\frac{\partial^2 f}{\partial y \partial x} = \left(3x^2y^2 + y^4\right)'_y = 6x^2y + 4y^3, \quad \frac{\partial^2 f}{\partial x \partial y} = \left(2x^3y + 4xy^3\right)'_x = 6x^2y + 4y^3.$$

Следует заметить, что в данном случае смешанные производные оказались равными: $\dfrac{\partial^2 f}{\partial y \partial x} = \dfrac{\partial^2 f}{\partial x \partial y}$. *Этот результат не случайный (см. теорему чуть ниже).*

Для функций большего числа переменных и для производных более высоких порядков определения аналогичны. Например, для функции $u = f(x, y, z)$

$$\frac{\partial^4 f}{\partial x \partial y^2 \partial z} = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}\left(\frac{\partial}{\partial y}\left(\frac{\partial}{\partial z} f\right)\right)\right)$$

или, что то же самое,

$$f^{(4)}_{zyyx} = \left(\left(\left(f'_z\right)'_y\right)'_y\right)'_x.$$

Теорема 4.1.1 (О равенстве смешанных производных) *Пусть функция* $f : \mathbb{R}^m \supset G \to \mathbb{R}$ *имеет в окрестности точки* x_0 *смешанные производные* $\frac{\partial^2 f}{\partial x_i \partial x_j}$ *и* $\frac{\partial^2 f}{\partial x_j \partial x_i}$. *Если эти производные непрерывны в точке* x_0, *то*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x_0).$$

Доказательство. Так как при вычислении смешанных производных по переменным x_i и x_j остальные переменные фиксируются, то можно сразу рассматривать функцию двух переменных $f(x, y)$ и точку $(x, y) \in G \subset \mathbb{R}^2$.

Зададим приращение: $h = (\Delta x, \Delta y)$. Рассмотрим величину

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

Введем функцию $\varphi(x) = f(x, y+\Delta y) - f(x,y)$. Тогда

$$\omega = \varphi(x+\Delta x) - \varphi(x)$$

Функция $\varphi(x)$ дифференцируема и $\varphi'(x) = \frac{\partial f}{\partial x}(x, y+\Delta y) - \frac{\partial f}{\partial x}(x,y)$.

По теореме Лагранжа для функции $\varphi(x)$ имеем

$$\exists x_1 \in (x, x+\Delta x): \quad \omega = \varphi'(x_1)\Delta x,$$

$$\omega = \left(\frac{\partial f}{\partial x}(x_1, y+\Delta y) - \frac{\partial f}{\partial x}(x_1, y)\right)\Delta x =$$

опять по теореме Лагранжа $\exists y_1 \in (y, y+\Delta y)$:

$$= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}(x_1, y_1)\right)\Delta x \Delta y = \frac{\partial^2 f}{\partial y \partial x}(x_1, y_1)\Delta x \Delta y.$$

Теперь перепишем ω в виде

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

и введем функцию $\psi(y) = f(x+\Delta x, y) - f(x,y)$. Тогда $\omega = \psi(y+\Delta y) - \psi(y)$. Используя дважды теорему Лагранжа, получим

$$\exists x_2 \in (x, x+\Delta x), y_2 \in (y, y+\Delta y): \quad \omega = \frac{\partial^2 f}{\partial x \partial y}(x_2, y_2)\Delta x \Delta y,$$

т.е.

$$\frac{\partial^2 f}{\partial y \partial x}(x_1, y_1) = \frac{\partial^2 f}{\partial x \partial y}(x_2, y_2).$$

Устремляя $\Delta x \to 0$, $\Delta y \to 0$ и учитывая непрерывность f''_{xy} и f''_{yx} получим

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

$\square$

Замечание 4.1.1 *Случай, рассмотренный в теореме легко обобщить на случай смешанных производных любого порядка. А именно, результат вычисления смешанной производной не зависит от порядка дифференцирования (в случае их непрерывности).*

Замечание 4.1.2 *Условие непрерывности для равенства смешанных производных обязательно.*

Пример 4.1.2 *Рассмотрим функцию*

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

Докажем, что $\dfrac{\partial^2 f}{\partial x\partial y}(0,0)\neq\dfrac{\partial^2 f}{\partial y\partial x}(0,0)$.

Для первых производных имеем:

$$\frac{\partial f}{\partial x}(0,0)=\frac{\partial f}{\partial y}(0,0)=0,(\textit{по определению})$$

при $x^2+y^2\neq 0$*:*

$$\frac{\partial f}{\partial x}=y\frac{x^2-y^2}{x^2+y^2}+xy\frac{2x\left(x^2+y^2\right)-2x\left(x^2-y^2\right)}{\left(x^2+y^2\right)^2}=y\frac{x^2-y^2}{x^2+y^2}+\frac{4x^2y^3}{\left(x^2+y^2\right)^2},$$

$$\frac{\partial f}{\partial y}=x\frac{x^2-y^2}{x^2+y^2}-\frac{4x^3y^2}{\left(x^2+y^2\right)^2}.$$

Тогда для смешанных производных:

$$\frac{\partial^2 f}{\partial x\partial y}(0,0)=\lim_{x\to 0}\frac{f'_y(x,0)-f'_y(0,0)}{x}=\lim_{x\to 0}\frac{x}{x}=1,$$

$$\frac{\partial^2 f}{\partial y\partial x}(0,0)=\lim_{y\to 0}\frac{f'_x(0,y)-f'_x(0,0)}{y}=\lim_{y\to 0}\frac{-y}{y}=-1,$$

т.е. $\dfrac{\partial^2 f}{\partial x\partial y}(0,0)\neq\dfrac{\partial^2 f}{\partial y\partial x}(0,0)$.

Определение 4.1.1 *Множество функций* $f:E\to\mathbb{R}$*, имеющих все частные производные вплоть до* k*-го порядка непрерывные на* E*, будем обозначать* $C^k(E)$*. При этом,* $C^1(E)$ *– множество непрерывно дифференцируемых функций.*

4.2 Дифференциалы высших порядков

Пусть $f:\mathbb{R}^m\subset E\to\mathbb{R}$ – дифференцируемая функция.

Для удобства записи, введем формальный дифференциальный оператор d:

$$d=\sum_{k=1}^{m}\frac{\partial}{\partial x_k}dx_k.$$

Тогда дифференциал функции f на векторе $h = (dx_1, ..., dx_m)$ можно записать так:

$$df = \Big(\frac{\partial}{\partial x_1}dx_1 + \frac{\partial}{\partial x_2}dx_2 + ... + \frac{\partial}{\partial x_m}dx_m\Big)f = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + ... + \frac{\partial f}{\partial x_m}dx_m.$$

Этот дифференциал является функцией точки и вектора приращений h.

Определение 4.2.1 *Дифференциалом второго порядка d^2f функции f называется $d^2f := d(df)$. Более того, дифференциал n-го порядка определяется индуктивно: $d^nf = d(d^{n-1}f)$.*

В операторном виде имеет место запись

$$d^nf = \left(\frac{\partial}{\partial x_1}dx_1 + \frac{\partial}{\partial x_m}dx_m\right)^n f.$$

При этом возведение в степень n происходит формально. Произведение операторов определяется как композиция: $\frac{\partial}{\partial x}\frac{\partial}{\partial y} = \frac{\partial^2}{\partial x \partial y}$.

Пример 4.2.1 *Для функции $f : \mathbb{R}^2 \to \mathbb{R}$ вида $u = f(x, y) \in C^2(\mathbb{R}^2)$ имеем*

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy.$$

Далее будем считать приращение (dx, dy) постоянным. Тогда

$$d^2f = d\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right)dx + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right)dy =$$
$$= \frac{\partial^2 f}{\partial x^2}(dx)^2 + \frac{\partial^2 f}{\partial x \partial y}dxdy + \frac{\partial^2 f}{\partial y \partial x}dydx + \frac{\partial^2 f}{\partial y^2}(dy)^2.$$

Пользуясь равенством смешанных производных и обозначением $(dx)^2 = dx^2$, $(dy)^2 = dy^2$, получим

$$d^2f = \frac{\partial^2 f}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial x \partial y}dxdy + \frac{\partial^2 f}{\partial y^2}dy^2.$$

Для дифференциала порядка n будет верна формула, аналогичная биному Ньютона:

$$d^nf = \sum_{k=0}^{n} C_n^k \frac{\partial^n f}{\partial x^{n-k}\partial y^k}dx^{n-k}dy^k.$$

Заметим, что дифференциал второго порядка функции $f(x_1, ..., x_m)$ является квадратичной формой относительно dx_1, ... dx_m.

Пример 4.2.2 *Найти d^3f, если $f(x,y) = x^3y + x^2y^3 + y^4$.*

Распишем, вначале, выражение для d^3f:

$$d^3f = \left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)^3 f =$$
$$= \frac{\partial^3 f}{\partial x^3}dx^3 + 3\frac{\partial^3 f}{\partial x^2 \partial y}dx^2dy + 3\frac{\partial^3 f}{\partial x \partial y^2}dxdy^2 + \frac{\partial^3 f}{\partial y^3}dy^3.$$

Найдем частные производные до третьего порядка:

$$\frac{\partial f}{\partial x} = 3x^2y + 2xy^3, \quad \frac{\partial f}{\partial y} = x^3 + 3x^2y^2 + 4y^3;$$

$$\frac{\partial^2 f}{\partial x^2} = 6xy + 2y^3, \quad \frac{\partial^2 f}{\partial y^2} = 6x^2y + 12y^2;$$

$$\frac{\partial^3 f}{\partial x^3} = 6y, \quad \frac{\partial^3 f}{\partial x^2 \partial y} = 6x + 6y^2, \quad \frac{\partial^3 f}{\partial x \partial y^2} = 12xy, \quad \frac{\partial^3 f}{\partial y^3} = 6x^2 + 24y.$$

Окончательно,

$$d^3f = 6ydx^3 + 18(x + y^2)dx^2dy + 36xydxdy^2 + (x^2 + 4y)dy^3.$$

4.3 Формула Тейлора для функции многих переменных

Вспомним и запишем в удобном виде формулу Тейлора для функции $f(x)$ одной переменной, имеющей производные вплоть до $(n+1)$-го порядка. В точке $x_0 + \Delta x$ имеем

$$f(x_0 + \Delta x) = f(x_0) + f'(x_0)\Delta x + \frac{1}{2!}f''(x_0)(\Delta x)^2 + \dots + \frac{1}{n!}f^(n)(x_0)(\Delta x)^n + R_n,$$

где остаток запишем в форме Лагранжа:

$$R_n = \frac{1}{(n+1)!}f^{(n+1)}(\xi)(\Delta x)^{n+1}, \quad \xi \in (x_0, x_0 + \Delta x).$$

Или можно записать через дифференциалы:

$$f(x_0 + \Delta x) = f(x_0) + df(x_0) + \frac{1}{2!}d^2f(x_0) + \dots + \frac{1}{n!}d^nf(x_0) + R_n.$$

Получим обобщение для случая функции $f : \mathbb{R}^m \subset E \to \mathbb{R}$.

Теорема 4.3.1 *Пусть $f : \mathbb{R}^m \subset E \to \mathbb{R}$, $f \in C^{n+1}(E)$, x_0 – внутренняя точка E. Тогда для h такого, что $x_0 + h \in E$ $\exists \theta \in (0,1)$ такая, что*

$$f(x_0+h) = f(x_0) + \sum_{k=1}^{m} \frac{d^k f(x_0,h)}{k!} + \frac{1}{(m+1)!} d^{m+1}\left(x_0+\theta h, h\right).$$

Доказательство. Рассмотрим функцию $\varphi(t) = f(x_0+th)$ одной переменной. Она дифференцируема на $t \in [0,1]$ и

$$\varphi'(t) = f'(x_0+th)h = df(x_0+th,h).$$

Вычисляя вторую производную, получим

$$\varphi''(t) = d^2 f(x_0+th,h).$$

Продолжая далее по индукции, получим

$$\varphi^{(n)}(t) = d^n f(x_0+th,h).$$

Запишем формулу Тейлора для функции $\varphi(t)$ в точке 0:

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \frac{\varphi''(0)}{2!}t^2 + \ldots + \frac{\varphi^{(n)}(0)}{n!}t^n + \frac{\varphi^{(n+1)}(\theta t)}{(n+1)!}t^{n+1},$$

где $\theta \in (0,1)$.

Имеем $\varphi(0) = f(x_0)$, $\varphi(1) = f(x_0+h)$. Подставляя в формулу Тейлора для $\varphi(t)$ значение $t = 1$, получаем

$$f(x) = f(x_0+h) = f(x_0) + \sum_{k=1}^{n} \frac{d^k f(x_0)}{k!} + \frac{1}{(n+1)!} d^{n+1}\left(x_0+\theta h\right).$$

$\square$

Замечание 4.3.1 *Для остатка можно записать форму Пеано:*

$$R_n = o\left(\|h\|^n\right), \quad h \to 0.$$

Пример 4.3.1 *Написать формулу Маклорена для функции $f(x,y) = \dfrac{\sin x}{\cos y}$ в точке $(0,0)$ с $o(\|h\|^3)$.*

1-ый способ. Найдем все частные производные в точке $(0,0)$ до третьего порядка включительно:

$$f'_x(0,0) = 1, \quad f'_y(0,0) = 0;$$

$$f''_{xx}(0,0) = 0, \quad f''_{xy}(0,0) = 0, \quad f''_{yy}(0,0) = 0;$$

$$f'''_{xxx}(0,0) = -1, \quad f'''_{xxy}(0,0) = 0, \quad f'''_{xyy}(0,0) = 1, \quad f'''_{yyy}(0,0) = 0.$$

Тогда

$$\frac{\sin x}{\cos y} = x - \frac{x^3}{6} + \frac{xy^2}{2} + o(\|h\|^3).$$

2-ой способ. Воспользуемся известными формулами Маклорена для функций одной переменной:

$$\frac{\sin x}{\cos y} = \frac{x - \frac{x^3}{6} + o(x^3)}{1 - \frac{y^2}{2} + o(y^3)} = \left(x - \frac{x^3}{6} + o(x^3)\right)\left(1 + \frac{y^2}{2} + o(y^3)\right) =$$

$$= x - \frac{x^3}{6} + \frac{xy^2}{2} + o(x^3) + xo(y^3) = x - \frac{x^3}{6} + \frac{xy^2}{2} + o(\|h\|^3).$$

В последнем действии мы воспользовались тем, что

$$o(x^3) = o\left(\|h\|^3\right), \quad o(y^3) = o\left(\|h\|^3\right).$$

5 Экстремумы функции многих переменных

5.1 Необходимое условие экстремума

Рассмотрим функцию $u = f(x) = f(x_1, \ldots, x_n)$ в окрестности точки $x^0 = \left(x_1^0, \ldots, x_n^0\right)$.

Напомним определение локального экстремума.

Определение 5.1.1 *Функция $f(x)$ имеет **локальный максимум (минимум)** в точке x^0, если существует окрестность $U(x^0)$ этой точки, что для $\forall x \in U(x^0)$ выполнено $f(x) \leqslant f(x^0)$ $\left(f(x) \geqslant f(x^0)\right)$.*

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

Если в определении взять проколотую окрестность точки x^0 и взять строгие неравенства: $f(x) < f(x^0)$ $\left(f(x) > f(x^0)\right)$, то получится ***строгий*** экстремум.

Теорема 5.1.1 (Необходимое условие экстремума) *Если функция $f(x)$, $x \in \mathbb{R}^n$ имеет экстремум в точке x^0, то любая ее частная производная первого порядка обращается в точке x^0 в ноль или не существует.*

□ Зафиксируем в точке x^0 все переменные функции f кроме x_1. Функция $g(t) = f(t, x_2^0, \ldots, x_n^0)$ имеет экстремум в точке x_1^0 и, следовательно, $g'(x_1^0) = 0$ или не существует. Но $g'(x_1^0) = \dfrac{\partial f}{\partial x_1}(x^0)$. Аналогично для производных по другим переменным. ■

Замечание 5.1.1 *Это необходимое условие не является достаточным.*

Точка, в которой функция дифференцируема, и все частные производные первого порядка обращаются в ноль (т.е. $df = 0$), называется ***стационарной точкой***.

Точка, в которой частные производные первого порядка обращаются в ноль или не существуют (т.е. df не существует), называется ***критической точкой***.

Пример. Функция $u = x^2 - y^2$ в точке $(0, 0)$:

$$u(0,0) = 0, \quad \frac{\partial z}{\partial x}(0,0) = 0, \quad \frac{\partial x}{\partial y}(0,0) = 0,$$

но в окрестности точки $(0, 0)$ при $x = 0 : u(0, y) = -y^2 < 0$, а при $y = 0$: $u(x, 0) = x^2 > 0$. Значит, в точке $(0, 0)$ экстремума нет.

5.2 Достаточное условие экстремума функции n переменных

Здесь нам потребуются понятия алгебры, касающиеся квадратичных форм.

Квадратичная форма

$$\Phi(\xi) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \xi_i \xi_j,$$

где $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$, $a_{ij} = a_{ji} \in R$, называется

а) ***положительно определенной***, если $\forall \xi \neq 0 \quad \Phi(\xi) > 0$;

б) ***отрицательно определенной***, если $\forall \xi \neq 0 \quad \Phi(\xi) < 0$;

в) ***неопределенной (знакопеременной)*** , если $\exists \xi_1, \xi_2 \ : \quad \Phi(\xi_1) > 0, \Phi(\xi_2) < 0$.

Существуют также квадратичные формы, не являющиеся ни одной из перечисленных. Например, если она принимает нулевое и положительные (отрицательные) значения. Такие формы называют *полуопределенными*.

Примеры.

1. $\Phi_1(\xi) = \xi_1^2 + 2\xi_1\xi_2 + 3\xi_2^2 = (\xi_1 + \xi_2)^2 + 2\xi_2^2$ – положительно определенная;

2. $\Phi_2(\xi) = -\xi_1^2 - 3\xi_2^2$ – отрицательно определенная;

3. $\Phi_3(\xi) = \xi_1^2 - \xi_2^2$ – неопределенная;

4. $\Phi_4(\xi) = \xi_1^2 + 2\xi_1\xi_2 + \xi_2^2 = (\xi_1 + \xi_2)^2$ – полуопределенная.

Квадратичная форма определяется симметричной матрицей $A = \{a_{ij}\}_{i=1,\dots,n}^{j=1,\dots,n}$.

Критерий Сильвестра положительной определенности квадратичной формы: Квадратичная форма положительно определена тогда и только тогда, когда все главные миноры ее матрицы положительны, т.е.

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0, \quad \dots .$$

Для отрицательной определенности квадратичной формы A необходимо и достаточно положительная определенность формы $-A$. Это означает, что знаки главных миноров будут чередоваться, начиная с первого $a_{11} < 0$.

Невырожденная квадратичная форма является неопределенной, если выполнено хотя бы одно из условий:

1. один из главных миноров равен нулю;

2. главный минор чётного порядка отрицателен;

3. два главных минора нечётного порядка имеют разные знаки.

Теорема 5.2.1 (Отделимость от нуля положительно опр. кв. формы) *Пусть $\Phi(\xi)$ – положительно определенная квадратичная форма. Тогда*

$$\exists C > 0 : \quad \forall \xi \quad \Phi(\xi) \geqslant C\|\xi\|^2,$$

где $\|\xi\| = \sqrt{\xi_1^2 + \dots + \xi_n^2}$.

Доказательство. Рассмотрим значения квадратичной формы $\Phi(\xi)$ на сфере $S = \{x : x_1^2 + \dots + x_n^2 = 1\}$. При $\xi \in S \quad \Phi(\xi) > 0$.

Так как S есть компакт в $\mathbb{R}^n$ (оно замкнуто и ограничено), то функция $\Phi(\xi)$ достигает на S свое наименьшее значение (т. Вейерштрасса), обозначим это значение C.

Следовательно, для $\forall \xi \in S \quad \Phi(\xi) \geqslant C$.

Если $\xi \notin S$ и $\xi \neq 0$, то точка $\dfrac{\xi}{\|\xi\|} \in S$ и тогда

$$\Phi\left(\frac{\xi}{\|\xi\|}\right) \geqslant C.$$

Теперь воспользуемся однородностью квадратичной формы (однородность означает, что $\forall t$: $\Phi(tx) = t^k\Phi(x)$, здесь $k = 2$):

$$\Phi\left(\frac{\xi}{\|\xi\|}\right) = \frac{1}{\|\xi\|^2}\Phi(\xi) \quad \Rightarrow \quad \Phi(\xi) \geqslant C\|\xi\|^2.$$

$\square$

Дифференциал второго порядка $d^2f(x^0)$ является квадратичной формой переменных $dx_1, \ldots, dx_n$.

Теорема 5.2.2 (Достаточное условие экстремума) *Пусть функция $f(x)$ имеет в окрестности точки $x^0 \in \mathbb{R}^n$ непрерывные частные производные второго порядка и $df(x^0) = 0$. Тогда*

1. *если $d^2f(x^0)$ – положительно определенная квадратичная форма, то в точке (x_0, y_0)* ***строгий минимум*** *$f(x)$;*
2. *если $d^2f(x^0)$ – отрицательно определенная квадратичная форма, то в точке (x_0, y_0)* ***строгий максимум*** *$f(x)$;*
3. *если $d^2f(x^0)$ – неопределенная квадратичная форма, то в точке (x_0, y_0)* ***экстремума нет****.*

Доказательство. Запишем формулу Тейлора в точке x^0:

$$\Delta f(x^0) = \frac{1}{2}d^2f(x^0) + o\left(\|h\|^2\right).$$

1. Пусть $d^2f(x^0)$ – положительно определенная квадратичная форма. Тогда в силу предыдущей теоремы, $\exists C > 0$, что

$$d^2f(x^0) \geqslant C\|h\|^2.$$

Тогда

$$\Delta f(x^0) \geqslant \frac{1}{2}C\|h\|^2 + o\left(\|h\|^2\right) = \frac{1}{2}C\|h\|^2\left(1 + \alpha(h)\right),$$

где $\alpha(h) \to 0$ при $h \to 0$.

Следовательно, в некоторой окрестности точки x^0 $(1 + \alpha(h) > 0$ и тогда $\Delta f(x^0) > 0$, что означает, что в точке (x_0, y_0) минимум.

2. Если $d^2f(x^0)$ – отрицательно определенная квадратичная форма, применим рассуждения предыдущего пункта к форме $-d^2f(x^0)$.

3. Пусть $\Phi(h) = d^2f(x^0)$ - неопределенная квадратичная форма. Тогда

$$\exists h', h'' : \quad \Phi(h') > 0, \quad \Phi(h'') < 0.$$

Тогда для любой окрестности $U(x^0)$ найдётся такое $t > 0$, что точки $x^0 + th'$ и $x^0 + th'' \in U(x^0)$ и

$$\Delta_1 f = f(x^0 + th') - f(x^0) > 0, \quad \Delta_2 f = f(x^0 + th'') - f(x^0) < 0,$$

что и означает, что в точке x^0 экстремума нет. □

Замечание 5.2.1 *Если квадратичная форма полуопределена, то возможно как наличие экстремума, так и его отсутствие. Например, функции $f(x, y) = x^2 + y^4$ и $g(x, y) = x^2 - y^4$ имеют в точке $(0, 0)$ второй дифференциал, равный $2dx^2$ (полуопределенная положительно квадратичная форма), но f имеет минимум в точке $(0, 0)$, а g не имеет экстремума в точке $(0, 0)$.*

Пример 5.2.1 *$f(x, y, z) = x^3+y^2+z^2+6xy-4z$. Исследовать на экстремум.*

Найдем стационарные точки:

$$\begin{cases} \dfrac{\partial u}{\partial x} = 3x^2 + 6y = 0, \\ \dfrac{\partial u}{\partial y} = 2y + 6x = 0, \\ \dfrac{\partial u}{\partial z} = 2z - 4 = 0. \end{cases}$$

Этой системе удовлетворяют две точки: $A(6, -18, 2)$ и $B(0, 0, 2)$. Найдем вторые производные:

$$\frac{\partial^2 u}{\partial x^2} = 6x, \quad \frac{\partial^2 u}{\partial x \partial y} = 6, \quad \frac{\partial^2 u}{\partial x \partial z} = 0,$$

$$\frac{\partial^2 u}{\partial y^2} = 2, \quad \frac{\partial^2 u}{\partial y \partial z} = 0, \quad \frac{\partial^2 u}{\partial z^2} = 2,$$

и составим матрицу квадратичной формы второго дифференциала:

$$\begin{pmatrix} 6x & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Для точки $A(6, -18, 2)$ получаем:

$$\begin{pmatrix} 36 & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

$$\Delta_1 = 36 > 0, \quad \Delta_2 = \begin{vmatrix} 36 & 6 \\ 6 & 2 \end{vmatrix} > 0, \quad \Delta_3 = \begin{vmatrix} 36 & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{vmatrix} > 0,$$

следовательно, в точке $A(6, -18, 2)$ - локальный минимум функции $u(x, y, z)$.

Теперь рассмотрим точку $B(0, 0, 2)$. В ней квадратичная форма имеет матрицу

$$\begin{pmatrix} 0 & 6 & 0 \\ 6 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Рассмотрим значения функции вблизи точки $B(0, 0, 2)$:

$$u(0, 0, 2) = -4,$$

при $x = \Delta x$, $y = \Delta y$, $z = 2 + \Delta z$ имеем

$$\Delta u = u(x, y, z) - u(0, 0, 2) = \Delta x^3 + \Delta y^2 + 6\Delta x \Delta y + \Delta z^2 = \Delta x(\Delta x^2 + 6\Delta y) + \Delta y^2 + \Delta z^2.$$

Возьмем $\Delta y = \Delta z = 0$. Тогда знак Δu будет совпадать со знаком Δx, то есть принимать в окрестности точки $B(0, 0, 2)$ и положительные и отрицательные значения. Следовательно, в точке $B(0, 0, 2)$ экстремума нет.

Замечание 5.2.2 *Матрица квадратичной формы, соответствующей дифференциалу второго порядка, называется **матрицей Гессе**, а ее определитель – **гессианом**.*

6 Неявное отображение и обратное отображение

6.1 Теорема Лагранжа о среднем

Теорема Лагранжа о среднем (или о конечном приращении) играет важную роль в математическом анализе. Обобщим ее на случай функций нескольких переменных.

Теорема 6.1.1 (Теорема Лагранжа о среднем) *Пусть $f : \mathbb{R}^m \supset G \to \mathbb{R}$, G – область и отрезок $[x, x + h] \subset G$. Если f непрерывна на $[x, x + h]$ и дифференцируема на $(x, x + h)$, то найдется точка $\xi \in (x, x + h)$ такая, что*

$$f(x + h) - f(x) = f'(\xi)h.$$

Замечание 6.1.1 *Множество, для которого отрезок, соединяющий любые две его точки, содержится в нем, называется* ***выпуклым***. *В условии теоремы можно требовать выпуклость области.*

Доказательство. Введем вспомогательную функцию

$$\varphi(t) = f(x + th), \quad t \in [0, 1].$$

Функция $\varphi(t)$ удовлетворяет теореме Лагранжа: непрерывна на $[0, 1]$, дифференцируема на $(0, 1)$ как суперпозиция непрерывных/дифференцируемых отображений. Тогда найдется $\theta \in (0, 1)$:

$$\varphi(1) - \varphi(0) = \varphi'(\theta),$$

откуда получаем $f(x + h) - f(x) = f'(x + \theta h)h$ и нашлось $\xi = x + \theta h$. $\square$

Следствие 6.1.2 *Если f дифференцируемо в области G и $df = 0$ в любой точке G, то $f \equiv \mathrm{const}$ в G.*

Замечание 6.1.2 *Для векторнозначных отображений теорема Лагранжа не верна. Например, для $f(x) = \begin{pmatrix} \sin x \\ \cos x \end{pmatrix}\colon \mathbb{R} \to \mathbb{R}^2$ имеем на отрезке $[0, 2\pi]$: $f(2\pi) - f(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, но при этом ни в одной точке $f'(x) = \begin{pmatrix} \cos x \\ -\sin x \end{pmatrix}$ не равно 0.*

6.2 Производная функции, заданной неявно

Напомним понятие неявно заданной функции для функции одной переменной. Уравнение $F(x, y) = 0$ в прямоугольнике $G = \{(x, y) : x_0 - a < x < x_0 + a, \quad y_0 - b < y < y_0 + b\}$ задает функцию $y = y(x)$, если для $\forall x \in (x_0 - a, x_0 + a)$ $\exists! y \in (y_0 - b, y_0 + b)$ такой, что $F(x, y) = 0$.

Теорема 6.2.1 (О неявной функции) *Пусть функция $F(x, y)\colon \mathbb{R}^2 \supset U(x_0, y_0) \to \mathbb{R}$ и выполнены следующие условия:*

1) $F \in C^1(U(x_0, y_0))$;

2) $F(x_0, y_0) = 0$;

3) $F'_y(x_0, y_0) \neq 0$.

Тогда существует прямоугольник $K = \{(x, y) : x_0 - a \leqslant x \leqslant x_0 + a, \quad y_0 - b \leqslant y \leqslant y_0 + b\}$, в котором уравнение $F(x, y) = 0$ задает y как неявную функцию от x: $y = f(x)$. При этом функция $f(x)$ непрерывно дифференцируема на $(x_0 - a, x_0 + a)$ и ее производная

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

Доказательство. 1) Докажем существование неявной функции.

Пусть $F'_y > 0$ (если меньше, то переобозначим $F(x,y) = -F(x,y)$). Так как F'_y непрерывна, то существует прямоугольник K_1:

$$K_1 = \{(x,y) : x_0 - a_1 \leqslant x \leqslant x_0 + a_1,\ y_0 - b_1 \leqslant y \leqslant y_0 + b_1\} : \quad F'_y(x,y) > 0.$$

Рассмотрим функцию $\psi(y) = F(x_0, y)$: $\psi'(y) = F'_y(x_0, y) > 0$, следовательно, $\psi(y)$ возрастает. $\psi(y_0) = F(x_0, y_0) = 0$ и найдётся такое b, что $\psi(y_0 + b) > 0$, $\psi(y_0 - b) < 0$. Тогда существует такое a, что для всех $x \in [x_0 - a, x_0 + a]$ выполнено $F(x, y_0 - b) < 0$ и $F(x, y_0 + b) > 0$.

Зафиксируем $x^* \in [x_0 - a, x_0 + a]$. Введём функцию $\varphi(y) = F(x^*, y)$ – непрерывна, $\varphi(y_0 - b) < 0$, $\varphi(y_0 + b) > 0$, $\varphi'(y) > 0$ и $\varphi(y)$ – возрастает. Отсюда следует, что найдётся единственный y^*, такой, что $\varphi(y^*) = F(x^*, y^*) = 0$. То есть, в прямоугольнике $K = \{(x,y) : x_0 - a \leqslant x \leqslant x_0 + a,\ y_0 - b \leqslant y \leqslant y_0 + b\}$ уравнение $F(x,y) = 0$ задает функцию $y = f(x)$.

2) Докажем дифференцируемость и формулу для производной.

Возьмём две точки (x,y) и $(x + \Delta x, y + \Delta y)$: $F(x,y) = F(x + \Delta x, y + \Delta y) = 0$. По теореме Лагранжа найдется точка $\xi \in \mathbb{R}^2$, лежащая на отрезке, соединяющем точки (x,y) и $(x + \Delta x, y + \Delta y)$ такая, что:

$$\Delta F = F(x + \Delta x, y + \Delta y) - F(x,y) = F'_x(\xi)\Delta x + F'_y(\xi)\Delta y = 0,$$

отсюда

$$\Delta y = -\frac{F'_x(\xi)}{F'_y(\xi)}\Delta x.$$

Так как $F'_y(x_0, y_0) > 0$, то $\exists \alpha > 0$: $F'_y(x,y) \geqslant \alpha$.

Так как F'_x непрерывна на компактном множестве, то $\exists \beta > 0$: $|F'_x(x,y)| \leqslant \beta$.

Тогда $|\Delta y| \leqslant \frac{\beta}{\alpha}|\Delta x| \to 0$ при $x \to 0$. То есть, $y = f(x)$ – непрерывна. Так как F'_x и F'_y непрерывны, то

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = -\lim_{\Delta x \to 0} \frac{F'_x(\xi)}{F'_y(\xi)} = -\frac{F'_x(x,y)}{F'_y(x,y)}.$$

□

Замечание 6.2.1 *Для функции n переменных $y = y(x_1, \ldots, x_n)$, заданной неявно уравнением*

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

будет выполнено

$$\frac{\partial y}{\partial x_i} = -\frac{F'_{x_i}(x,y)}{F'_y(x,y)}, \quad i = 1, \ldots, n.$$

Пример 6.2.1 $e^y - e^x + x^2 + y^2 = 0$. *Найти* y'_x.

1-ый способ. Воспользуемся доказанной теоремой.

$$F(x,y) = e^y - e^x + x^2 + y^2, \quad F'_x = -e^x + 2x, \quad F'_y = e^y + 2y.$$

$$\Rightarrow y'_x = \frac{e^x - 2x}{e^y + 2y}.$$

2-ой способ. Продифференцируем равенство $e^y - e^x + x^2 + y^2 = 0$ *по* x, *считая* y *функцией от* x*:*

$$e^y y'_x - e^x + 2x + 2yy'_x = 0.$$

Отсюда выразим y'_x*:*

$$y'_x = \frac{e^x - 2x}{e^y + 2y}.$$

6.3 Производная отображения, заданного неявно

Нам будет удобно использовать прямоугольные окрестности. Пусть $a = (a_1, ..., a_m) \in \mathbb{R}^m$. Прямоугольной окрестностью точки $x_0 = (x_1^0, ..., x_m^0) \in \mathbb{R}^m$ будем называть множество

$$I_a(x_0) = \{x \in \mathbb{R}^m : \ |x_i - x_i^0| < a_i, \ i = 1, .., m\},$$

которое можно представить декартовым произведением одномерных окрестностей:

$$I_a(x_0) = U_{a_1}(x_1^0) \times ... \times U_{a_m} x_m^0.$$

Пусть $x = (x_1, \dots, x_m) \in \mathbb{R}^m$, $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ и есть система уравнений

$$\begin{cases} F_1(x_1 ..., x_m, y_1, ..., y_n) = 0, \\ \dots \\ F_n(x_1 ..., x_m, y_1, ..., y_n) = 0, \end{cases} \quad \Leftrightarrow \quad F(x,y) = 0, \qquad (*)$$

где отображение $F = (F_1, ..., F_n)$: $\mathbb{R}^{m+n} \supset G \to \mathbb{R}^n$ непрерывно дифференцируемо и якобиан $\det F'_y \neq 0$ в G.

Зададимся вопросом, при каких условия эта система разрешима относительно функций y_1, ..., y_n?

Определение 6.3.1 *Будем говорить, что система (*) в окрестности точки* (x_0, y_0) *задает неявное отображение* $y = f(x) : \mathbb{R}^m \to \mathbb{R}^n$, *если для каждого* $x \in U(x_0)$ *найдется единственный* $y \in V(y_0)$ *такой, что* $F(x,y) = 0$.

Для краткости обозначим

$$F'_x(x,y) = \begin{pmatrix} (F_1)'_{x_1} & \dots & (F_1)'_{x_m} \\ \vdots & & \\ (F_n)'_{x_1} & \dots & (F_n)'_{x_m} \end{pmatrix}_{(x,y)}, \quad F'_y(x,y) = \begin{pmatrix} (F_1)'_{y_1} & \dots & (F_1)'_{y_n} \\ \vdots & & \\ (F_n)'_{y_1} & \dots & (F_n)'_{y_n} \end{pmatrix}_{(x,y)}.$$

Заметим, что матрица F'_y квадратная. А значит, она обратима тогда и только тогда, когда $\det F'_y \neq 0$.

Теорема 6.3.1 (О неявном отображении) *Пусть отображение* $F : \mathbb{R}^{m+n} \supset U(x_0, y_0) \to \mathbb{R}^n$ *удовлетворяет условиям:*

1) $F \in C^1(U(x_0, y_0))$*;*

2) $F(x_0, y_0) = 0$*;*

3) $\det F'_y(x_0, y_0) \neq 0$.

Тогда существует $(m+n)$*-мерная прямоугольная окрестность* $I(x_0, y_0) = I_\alpha(x_0) \times I_\beta(y_0) \subset U(x_0, y_0)$*, в которой система (*) задает неявно отображение* $y = f(x)$*, причем* $f \in C^1(I_\alpha(x_0))$ *и*

$$f'(x) = -\Big[F'_y(x, f(x))\Big]^{-1} \cdot F'_x(x, f(x)).$$

Доказательство. Воспользуемся методом математической индукции по числу уравнений n. При $n = 1$ утверждение выполнено (теорема о неявной функции 6.2.1).

Пусть утверждение выполнено для размерности $n-1$. Докажем ее выполнение для n.

Так как определитель n-го порядка $\det F'_y(x_0, y_0) \neq 0$, то хотя бы один из элементов последней строки отличен от нуля. Пусть это $(F_n)'_{y_n} \neq 0$.

Тогда по теореме о неявной функции (6.2.1), последнее уравнение $F_n(x, y_1, .., y_n)$ определяет в некоторой окрестности $\tilde{I}(x_0, y_0)$ функцию $y_n = \tilde{f}(x_1, .., x_m, y_1, .., y_{n-1})$ класса C^1 в соответствующей окрестности точки $x_0, y_1^0, ..., y_{n-1}^0)$.

Подставим теперь найденное y_n в первые $(n-1)$ уравнения системы (*). Получим систему из $n-1$ уравнения и обозначим:

$$\begin{cases} \Phi_1(x_1, .., x_m, y_1, .., y_n) := F_1(x_1..., x_m, y_1, ..., y_{n-1}, \tilde{f}(x, y_1, .., y_{n-1})) = 0, \\ \qquad \dots \\ \Phi_{n-1}(x_1, .., x_m, y_1, .., y_n) := F_n(x_1..., x_m, y_1, ..., y_{n-1}, \tilde{f}(x, y_1, .., y_{n-1})) = 0. \end{cases} \qquad (**)$$

Покажем, что для этой системы выполнено индукционное предположение. Имеем: Φ_i класса C^1 в соответствующей окрестности точки

$(x_0, y_1^0, .., y_{n-1}^0)$, а также $\Phi_i(x_0, y_1^0, .., y_{n-1}^0) = 0$. Рассмотрим

$$\frac{\partial \Phi_i}{\partial y_k} = \frac{\partial F_i}{\partial y_k} + \frac{\partial F_i}{\partial y_n}\frac{\partial \tilde{f}}{\partial y_k}, \quad i = 1, .., n-1$$

и докажем, что определитель, состоящий из $\frac{\partial \Phi_i}{\partial y_j}$ отличен от нуля. Положим

$$\Phi_n(x_1, ..., x_m, y_1, ..., y_{n-1}) = F_n(x_1..., x_m, y_1, ..., y_{n-1}, \tilde{f}(x, y_1, .., y_{n-1}) \equiv 0,$$

но тогда

$$\frac{\partial \Phi_n}{\partial y_k} = \frac{\partial F_n}{\partial y_k} + \frac{\partial F_n}{\partial y_n}\frac{\partial \tilde{f}}{\partial y_k} \equiv 0.$$

Тогда можем записать

$$\begin{pmatrix} (F_1)'_{y_1} + (F_1)'_{y_n}\tilde{f}'_{y_1} & \dots & (F_1)'_{y_{n-1}} + (F_1)'_{y_n}\tilde{f}'_{y_1} & (F_1)'_{y_n} \\ \vdots & & & \\ (F_n)'_{y_1} + (F_n)'_{y_n}\tilde{f}'_{y_1} & \dots & (F_n)'_{y_{n-1}} + (F_n)'_{y_n}\tilde{f}'_{y_{n-1}} & (F_n)'_{y_n} \end{pmatrix} =$$

$$= \begin{pmatrix} (\Phi_1)'_{y_1} & \dots & (\Phi_1)'_{y_{n-1}} & (F_1)'_{y_n} \\ \vdots & & & \\ (\Phi_{n-1})'_{y_1} & \dots & (\Phi_{n-1})'_{y_{n-1}} & (F_{n-1})'_{y_n} \\ 0 & \dots & 0 & (F_n)'_{y_n} \end{pmatrix}$$

По нашему предположению $(F_n)'_{y_n} \neq 0$ и его минор отличен от нуля, следовательно, в некоторой окрестности точки $x_0, y_1^0, ..., y_{n-1}^0$:

$$\begin{vmatrix} (\Phi_1)'_{y_1} & \dots & (\Phi_1)'_{y_{n-1}} \\ \vdots & & \\ (\Phi_{n-1})'_{y_1} & \dots & (\Phi_{n-1})'_{y_{n-1}} \end{vmatrix} \neq 0.$$

Тогда в силу индукционного предположения система (**) в некоторой окрестности точки $x_0, y_1^0, ..., y_{n-1}^0$ задает функции

$$y_i = f_i(x), i = 1, .., n-1.$$

Для y_n получаем

$$y_n = \tilde{f}(x, f_1(x), ..., f_{n-1}(x)) =: f_n(x).$$

Осталось доказать формулу для производной. Для найденного отображения f имеем в окрестности точки x_0:

$$F(x, f(x)) \equiv 0.$$

Дифференцируя это равенство, получим

$$F'_x(x,y) + F'_y(x,y)f'(x,y) = 0, \qquad \text{где } y = f(x),$$

откуда следует требуемое. $\square$

Пример 6.3.1 *Функции $u = u(x,y)$, $v = v(x,y)$ заданы системой уравнений*

$$\begin{cases} xu + yv - u^3 = 0, \\ x + y + u + v = 0. \end{cases}$$

Найти u'_x, u'_y, v'_x, v'_y в точке $A(x,y;u,v) = (1,0;1,-2)$.
Имеем

$$F_1 = xu + yv - u^3, \quad F_2 = x + y + u + v,$$

$$\begin{pmatrix} (F_1)'_u & (F_1)'_v \\ (F_2)'_u & (F_2)'_v \end{pmatrix} = \begin{pmatrix} x - 3u^2 & y \\ 1 & 1 \end{pmatrix}\Bigg|_A = \begin{pmatrix} -2 & 0 \\ 1 & 1 \end{pmatrix},$$

$$\begin{pmatrix} (F_1)'_x & (F_1)'_y \\ (F_2)'_x & (F_2)'_y \end{pmatrix} = \begin{pmatrix} u & v \\ 1 & 1 \end{pmatrix}\Bigg|_A = \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}.$$

Тогда

$$\begin{pmatrix} u'_x & u'_y \\ v'_x & v'_y \end{pmatrix}\Bigg|_A = -\begin{pmatrix} -2 & 0 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -1 \\ -\frac{3}{2} & 0 \end{pmatrix}.$$

6.4 Обратимость отображения

Определение 6.4.1 *Пусть $G \in \mathbb{R}^m$ – область, т.е. открытое связное множество. Отображение $f : \mathbb{R}^m \supset G \to \mathbb{R}^m$ называется регулярным в G, если $f \in C^1(G)$ и его якобиан $\det f' \neq 0$ в G.*

Теорема 6.4.1 (о локальной обратимости отображения) *Пусть $G \subset \mathbb{R}^m$ – открытое множество, отображение $f : G \to \mathbb{R}^m$ регулярно. Тогда в любой точке $x_0 \in G$ оно* ***локально регулярно обратимо****, т.е. найдутся такие окрестности $A(x_0) \subset G$ и $B(y_0) \subset f(G)$, $y_0 = f(x_0)$, что отображение $f : A(x_0) \to B(y_0)$ взаимно однозначно, и обратное отображение $f^{-1} : B(y_0) \to A(x_0)$ регулярно.*

Доказательство. Пусть отображение f задается системой

$$\begin{cases} y_1 - f_1(x) = 0, \\ \ldots \\ y_m - f_m(x) = 0. \end{cases} \qquad (*)$$

Обозначим $F_i(x, y) = y_i - f_i(x)$, $i = 1..m$, – непрерывно дифференцируемые функции в области G.

Рассмотрим якобиан

$$\begin{vmatrix} \frac{\partial F_1}{\partial x_1}(x) & \ldots & \frac{\partial F_1}{\partial x_m}(x) \\ \ldots & \ldots & \ldots \\ \frac{\partial F_m}{\partial x_1}(x) & \ldots & \frac{\partial F_m}{\partial x_m}(x) \end{vmatrix}_{(x^0, y^0)} = (-1)^m \begin{vmatrix} -\frac{\partial f_1}{\partial x_1}(x) & \ldots & -\frac{\partial f_1}{\partial x_m}(x) \\ \ldots & \ldots & \ldots \\ -\frac{\partial f_m}{\partial x_1}(x) & \ldots & -\frac{\partial f_m}{\partial x_m}(x) \end{vmatrix}_{(x^0, y^0)} \neq 0.$$

Значит для системы (*) выполнены условия теоремы о неявном отображении. Тогда существуют клеточные окрестности $K(x^0)$ и $Q(y^0)$, в которых система (*) определяет переменные $x_1, ..., x_m$ как неявные (непрерывно-дифференцируемые) функции переменных $y_1, ..., y_m$. Обозначим эти функции $x_i = \varphi_i(y)$, $i = 1..m$.

Регулярность обратного отображения следует из равенства
$f'(x_0) \cdot (f')^{-1}(x_0) = I$. □

Следствие 6.4.2 *Для регулярного отображения образ открытого множества есть открытое множество.*

Пример 6.4.1 *Отображение, задающее полярные координаты* $(\rho, \varphi) \xrightarrow{f} (x, y)$*:*

$$x = \rho \cos \varphi, \qquad \rho \in [0, +\infty]$$
$$y = \rho \sin \varphi, \qquad \varphi \in [0, 2\pi).$$

Найдем якобиан:

$$\det f' = \begin{vmatrix} \cos \varphi & -\rho \sin \varphi \\ \sin \varphi & \rho \cos \varphi \end{vmatrix} = \rho.$$

Следовательно, локально отображение f *регулярно обратимо в окрестности любой точки кроме* $(0, 0)$*. Также оно обратимо как отображение:*

$$(0, +\infty) \times (0, 2\pi) \to \mathbb{R}^2 \setminus \{(x, y) : (x, 0), x \geqslant 0\}.$$

Пример 6.4.2 *Отображение* $(\rho, \varphi, \theta) \xrightarrow{f} (x, y, z)$*, определяемое функциями*

$$x = \rho \cos \varphi \sin \theta, \qquad \rho \in [0, +\infty]$$
$$y = \rho \sin \varphi \sin \theta, \qquad \varphi \in [0, 2\pi)$$
$$z = \rho \cos \theta, \qquad \theta \in [0, \pi]$$

соответствует переходу от декартовых координат к сферическим. Найдем дифференциал и якобиан этого отображения.

Матрица Якоби:

$$J(\rho,\varphi,\theta)=\begin{pmatrix} x'_\rho & x'_\varphi & x'_\theta \\ y'_\rho & y'_\varphi & y'_\theta \\ z'_\rho & z'_\varphi & z'_\theta \end{pmatrix}=\begin{pmatrix} \cos\varphi\sin\theta & -\rho\sin\varphi\sin\theta & \rho\cos\varphi\cos\theta \\ \sin\varphi\sin\theta & \rho\cos\varphi\sin\theta & \rho\sin\varphi\cos\theta \\ \cos\theta & 0 & -\rho\sin\theta \end{pmatrix}.$$

Якобиан равен (используем разложение по первой строке):

$$|J|=-\rho^2\sin\theta.$$

Таким образом, при $\rho>0$ и $\theta\in(0,\pi)$ это отображение локально обратимо.

Также оно будет обратимым как отображение

$$(0,+\infty)\times(0,2\pi)\times(0,\pi)\to\mathbb{R}^3\setminus\{(x,y,z):(x,0,z),x\geqslant 0,z\in\mathbb{R}\}.$$

7 Условный экстремум

Пусть $f\colon \mathbb{R}^m\supset E\to\mathbb{R}$ и $\varphi(x)=(\varphi_1(x),...,\varphi_n(x))\colon \mathbb{R}^m\supset E\to\mathbb{R}^n$. При этом $m>n$. Ограничим аргумент функции f условием $\varphi(x)=0$ и обозначим

$$\Omega=\{x\in E:\varphi_i(x)=0, i\in\{1,...,n\}\}.$$

Определение 7.0.1 *Точка $x_0\in E$ называется точкой* ***условного максимума*** *функции $f(x)$ при выполнении условий $\varphi(x)=0$, если существует окрестность $U(x_0)$, что для $\forall x\in U(x_0)\cap\Omega$ выполнено $f(x_0)\geqslant f_0(x)$.*

Аналогичным образом определяется точка условного минимума и точки строгих условных экстремумов.

Прямой метод нахождения условного экстремума заключается в следующем.

Из системы связей $\varphi(x)=0$ выразим переменные $x_1,\ldots,x_n$ (или любые n переменных) $(x=(x_1,\ldots,x_n)$ через $x_{n+1},\ldots,x_m$ и подставим в функцию $f(x)$. Получим функцию от $m-n$ переменных $x_{n+1},\ldots,x_m$. Далее остается найти обычный (безусловный) экстремум этой функции.

Пример 7.0.1 *Найти экстремумы функции $f(x,y)=x^2+y^2$ при условии $x+y-1=0$.*

Выразим из уравнения связи y: $y=1-x$ и подставим в функцию:

$$f(x,y)=f(x,1-x)=x^2+(1-x)^2=2x^2-2x+1.$$

Этот квадратный трехчлен имеет минимум в точке $x_0 = \dfrac{1}{2}$*, что соответствует точке* $(x_0, y_0) = \left(\dfrac{1}{2}, \dfrac{1}{2}\right)$.

Заметим, что выражение каких-либо n переменных из уравнений связи часто бывает довольно сложно или вовсе невозможно. Для таких случаев применяют метод Лагранжа. Опишем его.

Теорема 7.0.1 (Необходимое условие условного экстремума) *Пусть* $f, \varphi_1, ..., \varphi_n \colon \mathbb{R}^m \supset E \to \mathbb{R}$, $f, \varphi_i \in C^1(E)$, $\operatorname{rang} \varphi'(x_0) = n$ $(m, n \in \mathbb{N}$, $m > n)$. *Если* x_0 *– точка условного экстремума* f *с уравнениями связи* $\varphi = 0$*, то существует* $\lambda \in \mathbb{R}^n$*:*

$$\operatorname{grad} f(x_0) = \sum_{i=1}^{n} \lambda_i \operatorname{grad} \varphi_i(x_0).$$

Доказательство. Условие $\operatorname{rang} \varphi'(x_0) = n$ означает, что в матрице Якоби $\varphi'(x_0)$ есть ненулевой минор максимального порядка n. Будем считать, что это самый правый минор, содержащий последние n столбцов. Обозначим переменные удобным образом.

Будем рассматривать $\mathbb{R}^m$ как $\mathbb{R}^m = \mathbb{R}^{m-n} \times \mathbb{R}^n$, где элементы пространства $\mathbb{R}^{m-n}$ назовем $x = (x_1, ..., x_{m-n}) \in \mathbb{R}^{m-n}$, а элементы $\mathbb{R}^n$ назовем $y = (y_1, ..., y_n)$. При этом точка $x_0 = (x^0, y^0)$, где $x^0 \in \mathbb{R}^{m-n}$, $y^0 \in \mathbb{R}^n$. При этом данная функция $f = f(x, y)$ и уравнения связи $\varphi(x, y) = 0$.

Тогда выполнены условия теоремы о неявном отображении и система уравнений $\varphi(x, y) = 0$ задает неявно y как функции от x, то есть в некоторой окрестности $U(x^0)$ существует отображение $\psi \colon U(x^0) \to \mathbb{R}^n$ непрерывно-дифференцируемое и такое, что уравнение $\varphi(x, y) = 0$ равносильно $y = \psi(x)$.

Введём функцию $F(x) = f(x, \psi(x))$, $x \in U(x^0)$. Она имеет экстремум в точке x^0. Тогда для неё выполняется необходимое условия экстремума

$$F'_x(x^0) = f'_x(x^0, \psi(x^0)) + f'_y(x^0, \psi(x^0)\psi'(x_0)) = 0 \tag{*}$$

Но также в окрестности $U(x^0)$ выполнено $\varphi(x, \psi(x)) \equiv 0$. Продифференцируем:

$$\varphi'_x(x^0, \psi(x^0)) + \varphi'_y(x^0, \psi(x^0)) \cdot \psi'(x^0) = 0$$

Умножим это равенство на $\lambda = (\lambda_1, ..., \lambda_n)$ слева (как на строку):

$$\lambda \cdot \varphi'_x(x_0) + \lambda \cdot \varphi'_y(x_0) \cdot \psi'(x_0) = 0 \tag{**}$$

Составим разность (*) и (**):

$$\Big(f'_x(x_0) - \lambda \cdot \varphi'_x(x_0)\Big) + \Big(f'_y(x_0) - \lambda \cdot \varphi'_y(x_0)\Big)\psi'(x^0) = 0$$

Выберем $\lambda = {f'}_y(x_0)\Big(\varphi'_y(x_0)\Big)^{-1}$. Тогда

$$ {f'}_x(x_0) = \lambda \cdot \varphi'_x(x_0), \quad {f'}_y(x_0) = \lambda \cdot \varphi'_y(x_0), $$

откуда и следует требуемое. □

Определение 7.0.2 *В условиях теоремы функция*

$$ L(x,\lambda) = f(x) - \sum_{i=1}^{n} \lambda_i \varphi_i(x) $$

называется функцией Лагранжа. "Хорошо" найденные λ_i называются множителями Лагранжа.

Теорема 7.0.2 (Достаточное условие условного экстремума) *Пусть $n, m \in \mathbb{N}$, $m > n$, E – открыто в $\mathbb{R}^m$, $f \in C^2(E)$, $\varphi \in C^2(E)$, $\varphi(x_0) = 0$, $x_0 \in E$, $\operatorname{rang}\varphi'(x_0) = n$. Пусть также $N = \{h \in \mathbb{R}^m : d\varphi(x_0,h) = 0\}$, L – функция Лагранжа с хорошо найденными λ. Тогда*

1. *Если $d^2L(x_0,h)$ положительно определена на N, то x_0 – точка условного минимума.*
2. *Если $d^2L(x_0,h)$ отрицательно определена на N, то x_0 – точка условного максимума.*
3. *Если $d^2L(x_0,h)$ принимает значения разных знаков на N, то x_0 не является точкой условного экстремума.*

Доказательство. Докажем для случая $m = 2$, $n = 1$. $x_0 = (x^0, y^0)$.

$$ {f'}_x(x_0) = \lambda \cdot {\varphi'}_x(x_0), \quad {f'}_y(x_0) = \lambda \cdot {\varphi'}_y(x_0). $$

Пусть ${\varphi'}_y(x_0) \neq 0$, тогда $\exists U(x_0) \subset \mathbb{R}^2$ и $V(x_0) \subset \mathbb{R}$ ψ: $V(x^0) \to \mathbb{R}$, $\psi \in C^2(V(x_0))$

$$ \Big\{(x,y) \in U(x_0) : \ \varphi(x,y) = 0\Big\} = \Big\{(x, \psi(x)), \ x \in V(x^0)\Big\} $$

Аналогично предыдущей теореме, введём функцию $F(x) = f(x, \psi(x))$ и

$$ F'(x) = f'_x(x, \psi(x)) + f'_y(x, \psi(x))\psi'(x) $$

И из $\varphi(x, \psi(x)) \equiv 0$ получаем

$$ \varphi'_x + \varphi'_y \cdot \psi' = 0. $$

Продифференцируем еще раз:

$F''(x) = f''_{xx}(x,\psi(x)) + 2f''_{xy}(x,\psi(x))\cdot\psi'(x) + f''_{yy}(x,\psi(x))\cdot(\psi'(x))^2 + f'_y(x,\psi(x))\cdot\psi''(x),$

$0 = \varphi''_{xx}(x,\psi(x)) + 2\varphi''_{xy}(x,\psi(x))\cdot\psi'_x + \varphi''_{yy}(x,\psi(x))\cdot(\psi'(x))^2 + \varphi'_y(x,\psi(x))\cdot\psi''(x).$

Умножим второе на λ, вычтем из первого и воспользуемся тем, что $f'_y - \lambda\varphi'_y = 0$ в точке y^0. Получим в точке x_0:

$$F''(x) = L''_{xx} + 2L''_{xy}\cdot\psi'_x + L''_{yy}\cdot\psi'^2 = 0$$

Пусть $h = (dx, dy) \in N$. Тогда имея равенство

$$\varphi'_x(x_0)dx + \varphi'_y(x_0)dy = 0$$

выразим dy:

$$dy = -\Big(\varphi'_y(x_0)\Big)^{-1}\cdot\varphi'_x(x_0)dx$$

$$dy = \psi'(x^0)dx.$$

Запишем дифференциал второго порядка функции F:

$$d^2F(x_0,h) = F''_{xx}(x_0)dx^2 = L''_{xx}\cdot dx^2 + 2L''_{xy}\cdot\psi'(x)\cdot dx^2 + L''_{yy}\cdot\psi'^2\cdot dx^2 =$$

$$= L''_{xx}\cdot dx^2 + 2L''_{xy}dxdy + L''_{yy}dy^2 = d^2L\Big(x_0,(h,\psi'(x_0)\cdot h)\Big).$$

Далее пользуемся достаточным условием экстремума (безусловного) для функции F. $\square$

Пример 7.0.2 *Найти экстремумы функции при данных уравнениях связи*

$$u = xyz, \quad x+y-z=3,\ x-y-z=8.$$

$$L(x,y,z,\lambda_1,\lambda_2) = xyz - \lambda_1(x+y-z-3) - \lambda_2(x-y-z-8)$$

$$\begin{cases} yz-\lambda_1-\lambda_2=0 \\ xz-\lambda_1+\lambda_2=0 \\ xy+\lambda_1+\lambda_2=0 \\ x+y-z-3=0 \\ x-y-z-8=0 \end{cases} \quad\Rightarrow\quad x=\frac{11}{4},\ y=-\frac{5}{2},\ z=-\frac{11}{4}$$

$$d^2L = 0dx^2+0dy^2+0dz^2+2zdxdy+2ydxdz+2xdydz = -\frac{11}{2}dxdy-5dxdz+\frac{11}{2}dydz$$

$$d\Phi(x_0)=0 \quad\Rightarrow\quad \begin{cases} dx+dy-dz=0 \\ dx-dy-dz=0 \end{cases} \quad\Rightarrow\quad \begin{cases} dx=dz \\ dy=0 \end{cases}$$

$$d^L\Big|_N = -5dx^2 < 0$$

следовательно, точка $\left(\frac{11}{4},-\frac{5}{2},-\frac{11}{4}\right)$ *– точка условного максимума.*

Пример 7.0.3 *Найти экстремумы функции* $f(x,y) = x^2 + 12xy + 2y^2$ *при условии* $4x^2 + y^2 = 25$.

Запишем матрицы Якоби для функции $f_1(x,y) = 4x^2 + y^2 - 25$:

$$\begin{pmatrix} 8x & 2y \end{pmatrix}.$$

Её ранг равен 1 и равен числу уравнений связи.

Составим функцию Лагранжа:

$$L(x,y,\lambda) = x^2 + 12xy + 2y^2 + \lambda(4x^2 + y^2 - 25).$$

Найдем ее стационарные точки:

$$\begin{cases} \dfrac{\partial L}{\partial x} = 2x + 12y + 8\lambda x = 0, \\ \dfrac{\partial L}{\partial y} = 4y + 12x + 2\lambda y = 0, \\ \qquad\qquad 4x^2 + y^2 = 25. \end{cases} \Leftrightarrow \begin{cases} (1+4\lambda)x + 6y = 0, \\ 6x + (2+\lambda)y = 0, \\ 4x^2 + y^2 = 25. \end{cases}$$

Первые два уравнения умеют ненулевое решение при условии

$$\begin{vmatrix} 1+4\lambda & 6 \\ 6 & 2+\lambda \end{vmatrix} = 0,$$

что дает $\lambda_1 = 2$, $\lambda_2 = -\frac{17}{4}$.

При $\lambda_1 = 2$ *получаем две точки* $A(2,-3)$ *и* $B(-2,3)$.

При $\lambda_2 = -\frac{17}{4}$ *получаем* $C(\frac{3}{2},4)$, $D(-\frac{3}{2},-4)$.

Для проверки достаточных условий, запишем второй дифференциал функции Лагранжа:

$$d^2_{xx}L = (2+8\lambda)dx^2 + 24dxdy + (4+2\lambda)dy^2.$$

Дифференцируя уравнение связи, получим

$$8xdx + 2ydy = 0.$$

В точке $A(2,-3)$ *при* $\lambda_1 = 2$ *имеем:*

$$16dx - 6dy = 0 \Leftrightarrow dy = \frac{8}{3}dx,$$

$$d^2_{xx}L(A) = 18dx^2 + 24dxdy + 8dy^2 = 2(3dx + 2dy)^2 = 2\left(3 + \frac{8}{3}\right)^2 dx^2 > 0.$$

Следовательно, в точке $A(2,-3)$ *условный минимум.*

Аналогичным образом получаем, в точке $B(-2,3)$ *- условный минимум, в точках* $C(\frac{3}{2},4)$ *и* $D(-\frac{3}{2},-4)$ *- условные максимумы.*