

Лекция 16

Функции многих переменных

Последовательности в $\mathbb{R}^n$.

ОПРЕДЕЛЕНИЕ 16.1. Пространством $\mathbb{R}^n$ мы будем называть множество точек $x = (x_1, \dots, x_n)$, где $x_i \in \mathbb{R}$ ($i = 1, \dots, n$), снабженное расстоянием, т. е. функцией

$$\rho(x, y) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}$$

Иногда вместо $\rho(x, y)$ будем писать $|x - y|$.

Основные свойства расстояния в $\mathbb{R}^n$:

- 1) $\rho(x, y) = 0 \iff x = y$ (т. е. $x_1 = y_1, \dots, x_n = y_n$);
- 2) $\rho(x, y) = \rho(y, x)$ для любых x и y ;
- 3) $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ для любых x, y и z (так называемое неравенство треугольника)

Отметим, что расстояние, удовлетворяющее свойствам 1)–3), можно задать и другими формулами, например

$$\tilde{\rho}(x, y) = |x_1 - y_1| + \dots + |x_n - y_n|.$$

Легко проверить, что все эти свойства выполнены.

ОПРЕДЕЛЕНИЕ 16.2. ε -окрестностью точки $a \in \mathbb{R}^n$ называется множество $O_\varepsilon(a) = \{x \in \mathbb{R}^n : \rho(x, a) < \varepsilon\}$ ($|x - a| < \varepsilon$).

ОПРЕДЕЛЕНИЕ 16.3. Рассмотрим последовательность $x^{(k)} \in \mathbb{R}^n$, $k = 1, 2, \dots$. Будем говорить, что $x^{(k)} \rightarrow a$ при $k \rightarrow \infty$ или $\lim_{k \rightarrow \infty} x^{(k)} = a$, если $|x^{(k)} - a| \rightarrow 0$ при $k \rightarrow \infty$, т. е.

$$\lim_{k \rightarrow \infty} \sqrt{(x_1^{(k)} - a_1)^2 + \dots + (x_n^{(k)} - a_n)^2} = 0.$$

ТЕОРЕМА 16.4 (об эквивалентности сходимости в $\mathbb{R}^n$ покоординатной сходимости). Последовательность $x^{(k)}$ сходится в $\mathbb{R}^n$ к a тогда и только тогда, когда $x_i^{(k)}$ сходится к a_i для каждого $i = 1, \dots, n$:

$$\lim_{k \rightarrow \infty} x^{(k)} = a \iff \lim_{k \rightarrow \infty} x_i^{(k)} = a_i \text{ для всех } i = 1, \dots, n.$$

ДОКАЗАТЕЛЬСТВО. Необходимость ($\Rightarrow$). Для любого $i = 1, \dots, n$ справедлива оценка

$$|x_i - y_i| \leq \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2} = \rho(x, y).$$

Поэтому

$$|x_i^{(k)} - a_i| \leq |x^{(k)} - a| \rightarrow 0 \quad (k \rightarrow \infty).$$

Достаточность ($\Leftarrow$). Пусть $x_1^{(k)} - a_1 \rightarrow 0, \dots, x_n^{(k)} - a_n \rightarrow 0$ при $k \rightarrow \infty$. Тогда $(x_1^{(k)} - a_1)^2 \rightarrow 0, \dots, (x_n^{(k)} - a_n)^2 \rightarrow 0$ при $k \rightarrow \infty$. Следовательно, $(x_1^{(k)} - a_1)^2 + \dots + (x_n^{(k)} - a_n)^2 \rightarrow 0$ при $k \rightarrow \infty$. Значит,

$$\sqrt{(x_1^{(k)} - a_1)^2 + \dots + (x_n^{(k)} - a_n)^2} \rightarrow 0 \text{ при } k \rightarrow \infty. \quad \square$$

ПРИМЕР 16.5. 1) В $\mathbb{R}^3$ последовательность $x^{(k)} = (\frac{1}{k}, \frac{k}{k+1}, (1 + \frac{1}{k})^k)$ сходится к $(0, 1, e)$.

2) В $\mathbb{R}^2$ последовательность $x^{(k)} = (\frac{1}{k}, k)$ не имеет предела, так как вторая координата $x_2^{(k)} = k$ не имеет (конечного) предела.

ОПРЕДЕЛЕНИЕ 16.6. Последовательность $x^{(k)}$ называется *последовательностью Коши* (или *последовательностью, сходящейся в себе*, или *фундаментальной*), если для любого $\varepsilon > 0$ существует такое $N = N(\varepsilon)$, что для всех $k, m \geq N$ справедлива оценка $\rho(x^{(k)}, x^{(m)}) < \varepsilon$.

ТЕОРЕМА 16.7 (о полноте $\mathbb{R}^n$). Любая фундаментальная последовательность в $\mathbb{R}^n$ сходится.

ДОКАЗАТЕЛЬСТВО. Поскольку $|x_i^{(k)} - x_i^{(m)}| \leq \rho(x_i^{(k)}, x_i^{(m)})$ (см. выше), для любого $i = 1, \dots, n$ имеем

$$\forall \varepsilon > 0 \quad \exists N : \quad |x_i^{(k)} - x_i^{(m)}| \leq \varepsilon \quad \forall k, m \geq N$$

Поэтому, каждая (числовая) последовательность координат $x_i^{(k)}$ фундаментальна, следовательно она сходится (материал прошлого семестра). Каждая координата сходится, следовательно (по теореме 18.6) сходится последовательность $x^{(k)}$. $\square$

ЗАМЕЧАНИЕ. Каждая сходящаяся последовательность в $\mathbb{R}^n$ является фундаментальной. Показывается это дословно так же, как для последовательностей в $\mathbb{R}$.

ОПРЕДЕЛЕНИЕ 16.8. Шаром радиуса r с центром в точке a называется множество

$$B_r(a) = \{x \in \mathbb{R}^n : \rho(x, a) \leq r\}$$

ОПРЕДЕЛЕНИЕ 16.9. Множество M называется *ограниченным* в $\mathbb{R}^n$, если существует шар $B_r(0)$, который целиком его содержит: $B_r \supset M$. Иными словами, если существует такое $r > 0$, что $\rho(x, 0) \leq r$ для всех точек $x \in M$.

ЛЕММА 16.10 (Больцано–Вейерштрасса). *Ограниченная в $\mathbb{R}^n$ последовательность $\{x^{(k)}\}$ имеет сходящуюся в $\mathbb{R}^n$ подпоследовательность $\{x^{(k_m)}\}$.*

ДОКАЗАТЕЛЬСТВО. Последовательность $\{x^{(k)}\}$ ограничена в $\mathbb{R}^n$. Это означает, что существует такое $r > 0$, что $\rho(x^{(k)}, 0) < r$. Отсюда, в частности, вытекает, что $|x_i^{(k)}| \leq r$ для любого $i = 1, \dots, n$.

Выберем из ограниченной числовой последовательности $\{x_1^{(k)}\}$ сходящуюся подпоследовательность $x_1^{(k_m)}$ (это можно сделать по лемме Больцано–Вейерштрасса из первого семестра). Эта последовательность $x_1^{(k_m)}$ сходится к некоторому числу a_1 .

Рассмотрим подпоследовательность $x^{(k_m)}$ в $\mathbb{R}^n$. Переобозначим эту последовательность по-старому, чтобы не было большого количества индексов: $\{x^{(k)}\}$. Эта последовательность по-прежнему обладает свойством $|x_i^{(k)}| \leq r$ ($i = 1, \dots, n$). Кроме того, ее первая координата сходится к a_1 .

Выберем теперь из ограниченной числовой последовательности $\{x_2^{(k)}\}$ сходящуюся подпоследовательность $x_2^{(k_m)}$. Эта последовательность $x_2^{(k_m)}$ сходится к некоторому числу a_2 . Рассмотрим подпоследовательность $\{x^{(k_m)}\}$ в $\mathbb{R}^n$. Она имеет первые две сходящиеся координаты.

Повторяя эту процедуру n раз, получаем в итоге подпоследовательность $\{x^{(k_m)}\}$, у которой все координаты $\{x_i^{(k_m)}\}$, $i = 1, \dots, n$, являются сходящимися последовательностями. В силу теоремы 16.4 эта последовательность $\{x^{(k_m)}\}$ сходится в $\mathbb{R}^n$. $\square$

ОПРЕДЕЛЕНИЕ 16.11. а) Точка $x \in M$ называется *внутренней точкой* множества M , если эта точка входит в множество M с некоторой своей ε -окрестностью: $O_\varepsilon(x) \subset M$.

б) Точка $x \notin M$ называется *внешней точкой* множества M , если некоторая ее ε -окрестность не имеет с этим множеством общих точек: $O_\varepsilon(x) \cap M = \emptyset$.

в) Точка x называется *граничной точкой* множества M , если она не является для этого множества ни внутренней, ни внешней. Иными словами, любая ее ε -окрестность содержит как точки, принадлежащие M , так и точки, не принадлежащие этому множеству.

Обозначения:

$\text{Int } M$ — множество внутренних точек множества M ;

$\text{Ext } M$ — множество внешних точек множества M ;

∂M — множество граничных точек множества M .

Из определения видно, что

$$\mathbb{R}^n = \text{Int } M + \text{Ext } M + \partial M.$$

Здесь знак «+» вместо знака объединения означает, что эти множества попарно не пересекаются.

ПРИМЕР 16.12. 1) Рассмотрим множество $M = \{0 \leq x_1 < 1, 0 \leq x_2 < 1\} \subset \mathbb{R}^2$.

Нетрудно видеть, что $(0, 1/2) \in \partial M$, $(1, 1/2) \in \partial M$, $(1/2, 1/2) \in \text{Int } M$, $(0, -1) \in \text{Ext } M$. Легко понять, что граница этого множества состоит из четырех отрезков.

2) В $\mathbb{R}$ рассмотрим множество $M = \mathbb{Q}$. Любая точка $x \in \mathbb{R}$ принадлежит границе множества $\mathbb{Q}$, поскольку любая окрестность $O_\varepsilon(x) = (x - \varepsilon, x + \varepsilon)$ содержит как рациональные числа, так и иррациональные (т.е. как принадлежащие $\mathbb{Q}$, так и не принадлежащие). Таким образом, у множества $M = \mathbb{Q}$ нет ни внутренних, ни внешних точек. Все точки для него являются граничными.

В $\mathbb{R}^2$ можно рассмотреть множество $\mathbb{Q}^2 = \{x = (q_1, q_2) \in \mathbb{R}^2 : q_1, q_2 \in \mathbb{Q}\}$, обладающее аналогичными свойствами.

3) Рассмотрим в качестве множества M отрезок $[0, 1]$ на оси Ox_1 в пространстве $\mathbb{R}^2$: $M = \{x \in \mathbb{R}^2 : x_1 \in [0, 1], x_2 = 0\}$. Любая точка этого отрезка не является внутренней (в любой ε -окрестности заведомо есть точки с ненулевой второй координатой). Поэтому любая точка этого множества M является граничной. Нетрудно видеть, что любая точка, не принадлежащая M , является внешней для M .

Отметим, что в $\mathbb{R}$ отрезок $[0, 1]$, разумеется, имеет внутренность $\text{Int } M = (0, 1)$; граница состоит из двух точек: $\partial M = \{0, 1\}$; остальные точки являются внешними: $\text{Ext } M = (-\infty, 0) \cup (1, +\infty)$.

ОПРЕДЕЛЕНИЕ 16.13. Множество M называется *открытым* в $\mathbb{R}^n$, если всего его точки являются внутренними: $M = \text{Int } M$.

Множество M называется *замкнутым* в $\mathbb{R}^n$, если все его граничные точки входят в M : $\partial M \subset M$. В этом случае $M = \text{Int } M + \partial M$.

ПРИМЕР 16.14. а) Окрестность $O_\varepsilon(a)$ является открытым множеством.

Действительно, пусть $x \in O_\varepsilon(a)$. Это означает, что $r = \rho(x, a) < \varepsilon$. Положим $\delta = (\varepsilon - r)/2$. Тогда $O_\delta(x) \subset O_\varepsilon(a)$. Действительно, если $y \in O_\delta(x)$, то

$$\rho(y, a) \leq \rho(y, x) + \rho(x, a) < \delta + r = \frac{\varepsilon - r}{2} + r = \frac{\varepsilon + r}{2} < \frac{\varepsilon + \varepsilon}{2} = \varepsilon.$$

б) Шар $B_r(a)$ является замкнутым множеством. Действительно, любая точка, не принадлежащая этому шару, является внешней для него: пусть $x \notin B_r(a)$, т.е. $\rho(x, a) > r$, тогда, выбрав $\delta = (\rho(x, a) - r)/2$, мы получим окрестность $O_\delta(x)$, не имеющую общих точек с $B_r(a)$, поскольку для любой точки $y \in O_\delta(x)$ справедливо неравенство

$$\rho(y, a) \geq \rho(x, a) - \rho(x, y) > \rho(x, a) - \frac{\rho(x, a) - r}{2} = \frac{\rho(x, a) + r}{2} > \frac{r + r}{2} = r.$$

в) Множество $M = \{(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1 < 1, 0 \leq x_2 < 1\}$ не является ни открытым, ни замкнутым (часть границы принадлежит множеству, а часть — нет).

г) Отрезок на плоскости $\mathbb{R}^2$ является замкнутым множеством, поскольку $\partial M \subset M$.

д) $\emptyset$ является открытым и замкнутым одновременно.

е) $\mathbb{R}^n$ является открытым и замкнутым одновременно.

Можно показать, что только два множества $\emptyset$ и $\mathbb{R}^n$ являются одновременно открытыми и замкнутыми.

ПРЕДЛОЖЕНИЕ 16.15. *Дополнение к открытому множеству в $\mathbb{R}^n$ замкнуто, а к замкнутому множеству — открыто.*

ДОКАЗАТЕЛЬСТВО. Заметим, что если точка x является внутренней для множества M , то она является внешней для дополнения к M : $x \in \text{Int } M \implies x \in \text{Ext}(\mathbb{R}^n \setminus M)$. Если же точка x является внешней для множества M , то она является внутренней для дополнения к M : $x \in \text{Ext } M \implies x \in \text{Int}(\mathbb{R}^n \setminus M)$. Это вытекает из соответствующих определений.

Кроме того, вспомним, что все пространство $\mathbb{R}^n$ по отношению к множеству M распадается на три попарно непересекающиеся части: $\mathbb{R}^n = \text{Int } M + \partial M + \text{Ext } M$.

Отсюда видно, что $\text{Int } M = \text{Ext}(\mathbb{R}^n \setminus M)$, $\text{Ext } M = \text{Int}(\mathbb{R}^n \setminus M)$, $\partial M = \partial(\mathbb{R}^n \setminus M)$

Поэтому если множество M замкнуто, т.е. $M = \text{Int } M + \partial M$, то множество $\mathbb{R}^n \setminus M = \text{Ext } M = \text{Int}(\mathbb{R}^n \setminus M)$, т.е. дополнение состоит только из своих внутренних точек.

Если же множество M открыто, т.е. $M = \text{Int } M$, то множество $\mathbb{R}^n \setminus M = \text{Ext } M + \partial M = \text{Int}(\mathbb{R}^n \setminus M) + \partial M(\mathbb{R}^n \setminus M)$, т.е. включает все свои граничные точки. $\square$

ПРИМЕР 16.16. а) Множество $M = \{x \in \mathbb{R}^n : \rho(x, a) \geq \varepsilon\}$ является замкнутым, поскольку M совпадает с дополнением к открытому множеству $O_\varepsilon(a)$.

б) Множество $M = \{x \in \mathbb{R}^n : \rho(x, a) > r\}$ является открытым, поскольку M совпадает с дополнением к замкнутому множеству $B_r(a)$.

Лекция 17

ПРЕДЛОЖЕНИЕ 17.1. 1) *Любое объединение и конечное пересечение открытых множеств открыто.*

2) *Любое пересечение и конечное объединение замкнутых множеств замкнуто.*

ДОКАЗАТЕЛЬСТВО. 1) Пусть $A = \bigcup A_\alpha$, где все A_α открыты. Пусть $x \in A$. Тогда $x \in A_\alpha$ с некоторым α . Множество A_α открыто, поэтому $O_\varepsilon(x) \subset A_\alpha$, а, значит, тем более, $O_\varepsilon(x) \subset A$. Это означает, что множество A открыто.

Пусть $A = \bigcap_{i=1}^m A_i = A_1 \cap A_2 \cap \dots \cap A_m$. Пусть $x \in A$, т.е. $x \in A_i$ для всех $i = 1, \dots, m$. При этом для каждого $i = 1, \dots, m$ существует $O_{\varepsilon_i} \subset A_i$. Выберем $\varepsilon = \min\{\varepsilon_1, \dots, \varepsilon_m\}$. Тогда $O_\varepsilon(x) \subset A_i$ для любого i ; поэтому $O_\varepsilon(x) \subset \bigcap_{i=1}^m A_i = A$. Это означает, что множество A открыто.

2) Пусть $A = \bigcap A_\alpha$, где A_α — замкнутые множества. Пусть $x \notin A$, тогда $x \notin A_\alpha$ при некотором α . Поскольку A_α замкнуто, $O_\alpha(x) \cap A_\alpha = \emptyset$. Тем более, $O_\alpha(x) \cap A = \emptyset$.

Пусть $A = \bigcup_{i=1}^m A_i$, где A_i — замкнутые множества. Пусть $x \notin A$. Тогда $x \notin A_i$, $i = 1, \dots, m$. При этом для каждого A_i существует $\varepsilon_i > 0$, такое, что $A_i \cap O_{\varepsilon_i}(x) = \emptyset$. Если выбрать $\varepsilon = \min\{\varepsilon_1, \dots, \varepsilon_m\}$, то, конечно, $A_i \cap O_\varepsilon = \emptyset$, $i = 1, \dots, m$, или, что то же самое, $A \cap O_\varepsilon = \emptyset$, т.е. точка x является внешней для A . Это означает, что все граничные точки множества A принадлежат этому множеству, т.е. множество A является замкнутым. $\square$

(КОНТР)ПРИМЕР 17.2. Утверждения типа «любое объединение замкнутых множеств является замкнутым» или «любое пересечение открытых множеств является открытым» ложны. Действительно,

(1) Интервал $(0, 1)$ можно представить в виде объединения точек, из которых он состоит, а точка является замкнутым множеством.

(2) Точка x является пересечением всех открытых множеств, ее содержащих.

ОПРЕДЕЛЕНИЕ 17.3. Точка x называется *предельной* точкой множества M , если существует такая последовательность $\{x^{(k)}\}$ точек из M , что $x^{(k)} \neq x$, но $x^{(k)} \rightarrow x$ при $k \rightarrow \infty$.

Иными словами, точка x называется предельной точкой множества M , если в любой проколотой окрестности $\dot{O}_\varepsilon(x)$ существует хотя бы одна точка множества M .

Покажем, что эти определения эквивалентны. Пусть $x^{(k)} \neq x$, $x^{(k)} \rightarrow x$, $x^{(k)} \in M$. Тогда по определению предела последовательности в любой окрестности $O_\varepsilon(x)$ есть элемент $x^{(k)}$, причем $x^{(k)} \in \dot{O}_\varepsilon(x)$, поскольку $x^{(k)} \neq x$. Наоборот, если для любого k в проколотой окрестности $\dot{O}_{1/k}(x)$ (радиуса $1/k$) есть точка $x^{(k)} \in M$, то $x^{(k)} \rightarrow x$, поскольку $\rho(x^{(k)}, x) \leq 1/k \rightarrow 0$ при $k \rightarrow +\infty$.

ОПРЕДЕЛЕНИЕ 17.4. Точка x множества M называется *изолированной* точкой множества M , если существует такая проколотая окрестность $\dot{O}(x)$, что $M \cap \dot{O}(x) = \emptyset$ или, что то же $M \cap O(x) = \{x\}$.

УТВЕРЖДЕНИЕ 17.5. 1) Любая внутренняя точка является предельной точкой для M .

2) Любая граничная точка множества M является либо предельной точкой, либо изолированной точкой для M .

3) Внешняя точка множества не может быть ни предельной, ни изолированной.

ДОКАЗАТЕЛЬСТВО. 1) Существует целая окрестность $O(x)$, состоящая из точек множества M . В ней можно выбрать нужную последовательность, сходящуюся к x .

2) Если x не является изолированной, то в любой окрестности точки x есть (по крайней мере один) элемент множества M , т.е. точка является предельной.

3) По определению.

ОПРЕДЕЛЕНИЕ 17.6. *Замыканием* множества M называется объединение множества M и предельных точек этого множества.

Замыкание множества M обозначается через $\overline{M}$.

Из утверждения 17.5 вытекает, что замыкание множества M состоит из внутренних и граничных точек множества M :

$$\overline{M} = \text{Int } M + \partial M.$$

Из этой формулы видно, что если M замкнуто, то $\overline{M} = M$.

Множество $\overline{M}$ всегда замкнуто (попробуйте это доказать самостоятельно).

Множество $\overline{M}$ является наименьшим замкнутым множеством, содержащим M .

Предел функции и непрерывность.

ОПРЕДЕЛЕНИЕ 17.7. Рассмотрим отображение $f: \dot{O}(a) \rightarrow \mathbb{R}^m$, где $a \in \mathbb{R}^n$. Это означает, что каждой точке $x \in \dot{O}(a) \subset \mathbb{R}^n$ сопоставляется некоторая точка $f(x) \in \mathbb{R}^m$.

По определению $\lim_{x \rightarrow a} f(x) = b$, если для любого $\varepsilon > 0$ существует $\delta = \delta(\varepsilon) > 0$, такое, что $f(x) \in O_\varepsilon(b)$, если $x \in \dot{O}_\delta(a)$.

Отображение $f: O(a) \rightarrow \mathbb{R}^m$ называется непрерывным в точке a , если $\lim_{x \rightarrow a} f(x) = f(a)$.

Видно, что эти определения дословно повторяют определения из первого семестра. Меняется только смысл понятия окрестность.

ПРИМЕР 17.8. Рассмотрим отображение из интервала $(\alpha, \beta) \in \mathbb{R}^1$ в $\mathbb{R}^2$, т.е. отображение вида $f: (\alpha, \beta) \rightarrow \mathbb{R}^2$, где

$$\begin{cases} x_1 = \varphi(t), \\ x_2 = \psi(t), \end{cases} \quad t \in (\alpha, \beta).$$

Если функции $\varphi(t)$ и $\psi(t)$ непрерывны в точке $t_0 \in (\alpha, \beta)$, то отображение непрерывно в этой точке. Действительно,

$$\lim_{t \rightarrow t_0} (\varphi(t), \psi(t)) = (\varphi(t_0), \psi(t_0)),$$

поскольку

$$\rho((\varphi(t), \psi(t)), (\varphi(t_0), \psi(t_0))) = \sqrt{(\varphi(t) - \varphi(t_0))^2 + (\psi(t) - \psi(t_0))^2} \rightarrow 0$$

при $t \rightarrow t_0$.

ОПРЕДЕЛЕНИЕ 17.9. Отображение $f: M \rightarrow \mathbb{R}$ ($M \subset \mathbb{R}^n$) мы будем называть *функцией многих переменных* (точнее, функцией n переменных) и пользоваться обозначениями

$$f = f(x_1, \dots, x_n) \quad \text{или} \quad f = f(x), \quad \text{где } x \in \mathbb{R}^n, \quad x = (x_1, \dots, x_n).$$

Если $n = 2$ или $n = 3$, то чаще будем писать $f(x, y)$ или $f(x, y, z)$.

ПРИМЕР 17.10. Рассмотрим функцию $f(x_1, x_2) = x_1 x_2$. Докажем, что $f(x_1, x_2) \rightarrow 0$ при $(x_1, x_2) \rightarrow (0, 0)$.

Заметим, что $|x_1 x_2| \leq \frac{1}{2}(x_1^2 + x_2^2)$. Положим $\delta = \sqrt{2\varepsilon}$. Тогда если $\rho((x_1, x_2), (0, 0)) < \delta$, то $|f(x_1, x_2) - 0| < \varepsilon$:

$$|f(x_1, x_2) - 0| = |x_1 x_2| \leq \frac{x_1^2 + x_2^2}{2} < \frac{\delta^2}{2} = \varepsilon.$$

ПРИМЕР 17.11. Рассмотрим функцию $f(x, y) = \frac{2xy}{x^2 + y^2}$. Эта функция определена на всей плоскости $\mathbb{R}^2$, кроме точки $(0, 0)$. При всех x, y имеем

$$|f(x, y)| = \left| \frac{2xy}{x^2 + y^2} \right| \leq 1.$$

Существует ли $\lim_{(x,y) \rightarrow (0,0)} \frac{2xy}{x^2 + y^2}$? Оказывается, не существует.

Действительно, если $x = y$, то $f(x, y) = 1$; если $x = -y$, то $f(x, y) = -1$. А поскольку в любой (сколь угодно малой) окрестности точки $(0, 0)$ есть и точки вида (δ, δ) , и точки вида $(\delta, -\delta)$, это означает, что предела быть не может. (В следующей лекции мы это докажем более аккуратно.)

Лекция 18

ТЕОРЕМА 18.1 (о непрерывности сложного отображения). Пусть определены отображения

$$f: O(a) \rightarrow \mathbb{R}^m \quad (\text{где } a \in \mathbb{R}^n), \quad g: O(b) \rightarrow \mathbb{R}^l \quad (\text{где } b = f(a)).$$

Пусть f непрерывно в точке a и отображение g непрерывно в точке b . Тогда сложное отображение $g \circ f: O(a) \rightarrow \mathbb{R}^l$ непрерывно в точке a .

ДОКАЗАТЕЛЬСТВО. Для любого $\varepsilon > 0$ существует $\sigma > 0$, такое, что если $y \in O_\sigma(b)$, то $g(y) \in O_\varepsilon(g(b))$. Для любого $\sigma > 0$ существует $\delta > 0$, такое, что если $x \in O_\delta(a)$, то $f(x) \in O_\sigma(f(a))$. Таким образом, для любого $\varepsilon > 0$ существует такое $\delta > 0$, что если $x \in O_\delta(a)$, то $g(f(x)) \in O_\varepsilon(g(b))$ (мы воспользовались тем, что $f(a) = b$). $\square$

Обратите внимание, что доказательство дословно повторяет доказательство соответствующей теоремы из прошлого семестра.

ОПРЕДЕЛЕНИЕ 18.2. Будем писать $\lim_{\substack{x \rightarrow a \\ x \in M}} f(x) = b$, если для любого $\varepsilon > 0$ существует $\delta > 0$, такое, что $f(x) \in O_\varepsilon(b)$, если $x \in \dot{O}_\delta(a) \cap M$. Такой предел называется пределом по множеству M .

Смысл этого понятия таков: нас интересует поведение функции только на множестве M .

Например, хорошо знакомый предел $\lim_{x \rightarrow a+0} f(x)$ из первого семестра можно записать так:

$$\lim_{x \rightarrow a+0} f(x) = \lim_{\substack{x \rightarrow a \\ x \in (0, +\infty)}} f(x).$$

ПРИМЕР 18.3. Рассмотрим опять функцию $f(x, y) = \frac{2xy}{x^2 + y^2}$.

Если $M_1 = \{(x, y) : x = y\}$ — биссектриса координатного угла, то

$$\lim_{\substack{(x,y) \rightarrow (0,0) \\ (x,y) \in M_1}} \frac{2xy}{x^2 + y^2} = \lim_{x \rightarrow 0} \frac{2xx}{x^2 + x^2} = \lim_{x \rightarrow 0} 1 = 1.$$

Пусть M_α — луч

$$\begin{cases} x = t \cos \alpha, \\ y = t \sin \alpha, \end{cases} \quad t \in [0, +\infty).$$

Тогда

$$\lim_{\substack{(x,y) \rightarrow (0,0) \\ (x,y) \in M_\alpha}} \frac{2xy}{x^2 + y^2} = \lim_{t \rightarrow 0} \frac{2t \cos \alpha t \sin \alpha}{t^2 \cos^2 \alpha + t^2 \sin^2 \alpha} = \lim_{t \rightarrow 0} \sin 2\alpha = \sin 2\alpha.$$

Этот пример показывает, что предел, вообще говоря, зависит от множества, по которому он берется.

ПРЕДЛОЖЕНИЕ 18.4. Пусть $\lim_{x \rightarrow a} f(x) = b$. Тогда $\lim_{\substack{x \rightarrow a \\ x \in M}} f(x) = b$.

ДОКАЗАТЕЛЬСТВО. Если $f(x) \in O_\varepsilon(b)$ при $x \in \dot{O}_\delta(a)$, то тем более это верно, если $x \in \dot{O}_\delta(a) \cap M$. $\square$

Теперь легко показать, что предел $\lim_{(x,y) \rightarrow (0,0)} \frac{2xy}{x^2 + y^2}$ не существует. В самом деле, если бы этот предел существовал, то все пределы $\lim_{\substack{(x,y) \rightarrow (0,0) \\ (x,y) \in M_\alpha}} \frac{2xy}{x^2 + y^2}$ были бы одинаковы, а они зависят от α (см. выше).

УПРАЖНЕНИЕ. Покажите, что предел функции $\frac{2x^2y}{x^4 + y^2}$ при $(x, y) \rightarrow (0, 0)$ по любому лучу, выходящему из начала координат, равен 0. Тем не менее, предел этой функции при $(x, y) \rightarrow (0, 0)$ отсутствует. Убедитесь в этом, рассмотрев предел $\lim_{\substack{(x,y) \rightarrow (0,0) \\ (x,y) \in M}} \frac{2x^2y}{x^4 + y^2}$, где $M = \{(x, y) : y = x^2\}$ — парабола.

ОПРЕДЕЛЕНИЕ 18.5. Отображение $f: M \rightarrow \mathbb{R}^n$ называется *непрерывным по множеству M* в точке a , если $\lim_{\substack{x \rightarrow a \\ x \in M}} f(x) = f(a)$.

Отображение $f: M \rightarrow \mathbb{R}^n$ называется *непрерывным на M* , если оно непрерывно в любой точке множества M по множеству M .

ОПРЕДЕЛЕНИЕ 18.6. Множество M называется *компактом* в $\mathbb{R}^n$ (или *компактным множеством* в $\mathbb{R}^n$), если оно замкнуто и ограничено в $\mathbb{R}^n$.

Например, отрезок в $\mathbb{R}^1$ является компактом, шар $B_r(a)$ в $\mathbb{R}^n$ является компактом. Однако, окрестность $O_\varepsilon(a)$ в $\mathbb{R}^n$ не является компактом, поскольку это не замкнутое множество.

ТЕОРЕМА 18.7 (об образе компакта). *Образом компакта при непрерывном отображении всегда является компакт. Более точно, пусть отображение $f: M \rightarrow \mathbb{R}^m$ непрерывно на компакте M . Тогда образ $\text{Im } M$ этого компакта*

$$\text{Im } M = \{y \in \mathbb{R}^m : \exists x \in M, f(x) = y\}$$

также является компактом.

ДОКАЗАТЕЛЬСТВО. 1) Допустим, что образ $\text{Im } M$ не является ограниченным. Это означает, что вне любой окрестности $O_k(0)$ радиуса k найдется точка $y^{(k)} \in \text{Im } M$. Рассмотрим последовательность прообразов этих точек $x^{(k)} \in M$ ($f(x^{(k)}) = y^{(k)}$). Поскольку эта последовательность ограничена (M — ограниченное множество), из этой последовательности можно выбрать сходящуюся подпоследовательность $x^{(k_m)} \rightarrow a$, причем $a \in M$ в силу замкнутости множества M . В силу непрерывности функции $f(x)$ в точке a имеем $f(x^{(k_m)}) \rightarrow f(a)$ при $m \rightarrow \infty$. С другой стороны, $f(x^{(k_m)}) \geq k_m \rightarrow +\infty$. Противоречие.

2) Докажем замкнутость $\text{Im } M$. Пусть последовательность точек $y^{(k)}$ такова, что $y^{(k)} \rightarrow b$. Выберем из последовательности прообразов $\{x^{(k)}\}$ сходящуюся подпоследовательность (это возможно, поскольку она ограничена): $x^{(k_m)} \rightarrow a$, причем $a \in M$ в силу замкнутости множества M . В силу непрерывности f в точке a получим, что $f(x^{(k_m)}) \rightarrow f(a)$ при $m \rightarrow \infty$. Таким образом, $b = f(a)$, т.е. $b \in \text{Im } M$. $\square$

СЛЕДСТВИЕ 18.8. *Если $f: M \rightarrow \mathbb{R}^1$ (т.е. f является функцией нескольких переменных) непрерывна на компакте M , то существуют такие точки $x_{\min}$ и $x_{\max}$, что*

$$f(x_{\min}) \leq f(x) \leq f(x_{\max}) \quad \text{для всех } x \in M.$$

Действительно, по теореме 18.7 образ множества M на $\mathbb{R}^1$ является замкнутым ограниченным множеством. Из ограниченности множества $\text{Im } M$ вытекает существование $y_{\min} = \inf(\text{Im } M)$ и $y_{\max} = \sup(\text{Im } M)$. По определению $\inf$ существует последовательность $y^{(k)}$, такая, что $y^{(k)} \rightarrow y_{\min}$. Ее прообраз $x^{(k)}$ на M имеет сходящуюся подпоследовательность $x^{(k_m)} \rightarrow a$. При этом в силу замкнутости M точка a принадлежит M . Значит, в силу непрерывности f в точке a имеем $f(x^{(k_m)}) \rightarrow f(a)$, т.е. $f(a) = y_{\min}$ и мы нашли точку минимума $a = x_{\min}$. Аналогично ищется точка максимума.

По существу, следствие 18.8 заменяет две теоремы Вейерштрасса из первого семестра.

ОПРЕДЕЛЕНИЕ 18.9. Множество $M \subset \mathbb{R}^n$ называется *связным*, если для любых двух точек $x^{(1)}, x^{(2)} \in M$ существует непрерывная кривая, целиком лежащая в M и соединяющая эти точки: т.е. существует непрерывное отображение $\varphi: [\alpha, \beta] \rightarrow M$, такое, что $\varphi(\alpha) = x^{(1)}$, $\varphi(\beta) = x^{(2)}$, $\varphi(t) \in M$ для любого $t \in [\alpha, \beta]$.

Пример связного множества на плоскости — круг. Пример несвязного: $M = \{(x, y) : xy > 0\}$ (в последнем множестве две компоненты связности: $M_1 = \{(x, y) : x > 0, y > 0\}$ и $M_2 = \{(x, y) : x < 0, y < 0\}$).

СЛЕДСТВИЕ 18.10 (из теоремы 18.7). *Если M — связный компакт, то образом непрерывной функции $f: M \rightarrow \mathbb{R}^1$ является отрезок.*

Действительно, возьмем точки $x_{\min}$ и $x_{\max}$, такие, что $f(x_{\min}) \leq f(x) \leq f(x_{\max})$ для всех $x \in M$ (см. следствие 19.8) и соединим их непрерывной кривой $\varphi(t)$, $t \in [\alpha, \beta]$. Тогда функция $f(\varphi(t))$ непрерывна на $[\alpha, \beta]$, $f(\varphi(\alpha)) = f(x_{\min})$, $f(\varphi(\beta)) = f(x_{\max})$. По теореме Коши о промежуточном значении для любого $y \in [f(x_{\min}), f(x_{\max})]$ существует такое t , что $f(\varphi(t)) = y$. Таким образом, функция f принимает значение y в точке $\varphi(t) \in M$. $\square$

Ясно, что следствие 18.10 заменяет теорему Коши о промежуточном значении из первого семестра.

ЗАМЕЧАНИЕ 18.11 (о графике отображения). *Графиком* отображения $f: M \rightarrow \mathbb{R}^m$ ($M \subset \mathbb{R}^n$) мы будем называть множество $\Gamma = \{(x, y) \in \mathbb{R}^{n+m} : x \in M, y = f(x)\}$.

График Γ может лежать в $\mathbb{R}^3$ в двух случаях:

1) $n = 2, m = 1$, т.е. в случае когда f — функция двух переменных. Например, $z = x^2 + y^2$. График этой функции — параболоид вращения.

2) $n = 1, m = 2$, т.е. f задает кривую в пространстве. Например, $\begin{cases} x = R \cos t, \\ y = R \sin t \end{cases} \quad (t \in \mathbb{R}^1)$. Здесь график — винтовая линия.

Дифференциальное исчисление функций нескольких переменных.

Частные производные. Дифференцируемость.

ОПРЕДЕЛЕНИЕ 18.12. Пусть функция $f(x)$ определена в окрестности $O(a)$, $a \in \mathbb{R}^n$: $f(x) = f(x_1, x_2, \dots, x_n)$. «Заморозим» все переменные, кроме первой, т.е. рассмотрим функцию

$$\varphi(t) = f(t, a_2, \dots, a_n).$$

Если функция $\varphi(t)$ дифференцируема в точке a_1 , то $\varphi'(t)$ будем называть *частной производной* функции $f(x)$ в точке $a = (a_1, \dots, a_n)$ по переменной x_1 :

$$\frac{\partial f}{\partial x_1}(a) = \varphi'(a_1).$$

Аналогично определяется частная производная по любой переменной x_k :

$$\frac{\partial f}{\partial x_k}(a) = \varphi'(a_k), \quad \text{где } \varphi(t) = f(a_1, \dots, a_{k-1}, t, a_{k+1}, \dots, a_n).$$

Например, если $f(x, y) = x^y$, $a = (2, 3)$, то

$$\frac{\partial f}{\partial x}(2, 3) = (x^3)'|_{x=2} = 3x^2|_{x=2} = 12, \quad \frac{\partial f}{\partial y}(2, 3) = (2^y)'|_{y=3} = 2^y \ln 2|_{y=3} = 8 \ln 2.$$

Лекция 19

ЗАМЕЧАНИЕ 19.1. В случае $\mathbb{R}^2$ или $\mathbb{R}^3$ вместо $f(x_1, x_2)$ и $f(x_1, x_2, x_3)$ будем писать $f(x, y)$ и $f(x, y, z)$. В случае $\mathbb{R}^n$ фиксированная точка как правило будет обозначаться $a = (a_1, \dots, a_n)$, а в случае $\mathbb{R}^2$ или $\mathbb{R}^3$ — (x_0, y_0) или (x_0, y_0, z_0) .

Таким образом, для функции двух переменных имеем:

$$\begin{aligned} \frac{\partial f}{\partial x}(x_0, y_0) &= \lim_{x \rightarrow x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0} = \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}, \\ \frac{\partial f}{\partial y}(x_0, y_0) &= \lim_{y \rightarrow y_0} \frac{f(x_0, y) - f(x_0, y_0)}{y - y_0} = \lim_{\Delta y \rightarrow 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y}. \end{aligned}$$

Если не фиксировать точки, в которой вычисляется частная производная, то $\frac{\partial f}{\partial x}(x, y)$ и $\frac{\partial f}{\partial y}(x, y)$ становятся функциями от переменных (x, y) .

Например, если $f(x, y) = x^y$, то

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

ОПРЕДЕЛЕНИЕ 19.2. Пусть функция $f(x)$ определена в окрестности $O(x^{(0)})$, $x \in \mathbb{R}^n$. Эта функция называется *дифференцируемой* в точке $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})$, если справедлива формула

$$f(x) = f(x^{(0)}) + \sum_{k=1}^n A_k(x_k - x_k^{(0)}) + o(\rho(x, x^{(0)})), \quad x \rightarrow x^{(0)}. \quad (19.1)$$

Дадим некоторые пояснения. Напомним, что функция вида $L(x) = a_1x_1 + \dots + a_nx_n$ называется *линейной* функцией n переменных.

Перепишем формулу (19.1) в виде

$$\Delta f(x) = \sum_{k=1}^n A_k \Delta x_k + o(\rho(x, x^{(0)})), \quad x \rightarrow x^{(0)}, \quad (19.2)$$

где $\Delta f(x) = f(x) - f(x^{(0)})$ — приращение функции, $\Delta x_k = x_k - x_k^{(0)}$ — приращение независимых переменных.

Из формулы (19.2) виден смысл определения дифференцируемой функции: функция дифференцируема в точке $x^{(0)}$, если приращение функции в этой точке является линейной функцией приращений независимых переменных с точностью до «пренебрежимо малого» добавка.

Наконец, выпишем формулу (19.1) отдельно для случая функции трех переменных:

$$f(x, y, z) = f(x_0, y_0, z_0) + A(x - x_0) + B(y - y_0) + C(z - z_0) + o(\sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}), \quad (x, y, z) \rightarrow (x_0, y_0, z_0).$$

ПРЕДЛОЖЕНИЕ 19.3. Функция $\alpha(x)$ имеет вид $o(\rho(x, x^{(0)}))$ тогда и только тогда, когда

$$\alpha(x) = \sum_{k=1}^n \alpha_k(x)(x_k - x_k^{(0)}), \quad \text{где} \quad \lim_{x \rightarrow x^{(0)}} \alpha_k(x) = 0.$$

ДОКАЗАТЕЛЬСТВО. Проверим это для функции двух переменных.

1) Пусть

$$h(x, y) = \alpha(x, y)(x - x_0) + \beta(x, y)(y - y_0),$$

где α, β — бесконечно малые функции при $(x, y) \rightarrow (x_0, y_0)$. Тогда

$$\begin{aligned} \frac{h(x, y)}{\rho((x, y), (x_0, y_0))} &= \frac{\alpha(x, y)(x - x_0) + \beta(x, y)(y - y_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \\ &= \underbrace{\alpha(x, y)}_{\text{б.м.}} \underbrace{\frac{x - x_0}{\sqrt{(x - x_0)^2 + (y - y_0)^2}}}_{\text{ограниченная функция: } |\dots| \leq 1} + \underbrace{\beta(x, y)}_{\text{б.м.}} \underbrace{\frac{y - y_0}{\sqrt{(x - x_0)^2 + (y - y_0)^2}}}_{\text{ограниченная функция: } |\dots| \leq 1} \rightarrow 0 \end{aligned}$$

при $(x, y) \rightarrow (x_0, y_0)$.

2) Наоборот, если

$$\frac{h(x, y)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \rightarrow 0 \quad \text{при} \quad (x, y) \rightarrow (x_0, y_0),$$

то тем более

$$\gamma(x, y) = \frac{h(x, y)}{|x - x_0| + |y - y_0|} \rightarrow 0 \quad \text{при} \quad (x, y) \rightarrow (x_0, y_0)$$

(ведь $|x - x_0| + |y - y_0| \geq \sqrt{(x - x_0)^2 + (y - y_0)^2}$ согласно неравенству треугольника). Таким образом,

$$h(x, y) = \gamma(x, y)(|x - x_0| + |y - y_0|) = \underbrace{\gamma(x, y)}_{\text{б.м.}} \underbrace{\frac{|x - x_0|}{|x - x_0| + |y - y_0|}}_{\text{огр.}} (x - x_0) + \underbrace{\gamma(x, y)}_{\text{б.м.}} \underbrace{\frac{|y - y_0|}{|x - x_0| + |y - y_0|}}_{\text{огр.}} (y - y_0)$$

В силу доказанного предложения определение 19.2 теперь можно сформулировать так:

ОПРЕДЕЛЕНИЕ 19.2'. Пусть функция $f(x)$ определена в окрестности $O(x^{(0)})$, $x \in \mathbb{R}^n$. Эта функция называется *дифференцируемой* в точке $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})$, если справедлива формула

$$f(x) = f(x^{(0)}) + \sum_{k=1}^n (A_k(x_k - x_k^{(0)}) + \alpha_k(x)(x_k - x_k^{(0)})), \quad (19.1')$$

где $\alpha_k(x)$ — бесконечно малые при $x \rightarrow x^{(0)}$ функции.

Например, функцию $f(x, y) = xy$ в окрестности точки $(1, 2)$ можно записать в виде

$$f(x, y) = \underbrace{2}_{f(1,2)} + 2 \cdot (x - 1) + 1 \cdot (y - 2) + \underbrace{0}_{\text{б.м. при } (x,y) \rightarrow (1,2)} \cdot (x - 1) + \underbrace{(x - 1)}_{\text{б.м. при } (x,y) \rightarrow (1,2)} \cdot (y - 2).$$

ТЕОРЕМА 19.4. Если функция $f(x)$ дифференцируема в точке $x^{(0)}$ ($x \in \mathbb{R}^n$), то эта функция в этой точке имеет частные производные $\partial f / \partial x_k$ по всем переменным x_k , $k = 1, \dots, n$, причем

$$\frac{\partial f}{\partial x_k}(x^{(0)}) = A_k, \quad k = 1, \dots, n,$$

где A_k — коэффициенты из формулы (19.1).

ДОКАЗАТЕЛЬСТВО. Из (19.1') имеем

$$f(x_1, x_2^{(0)}, \dots, x_n^{(0)}) = f(x^{(0)}) + A_1(x_1 - x_1^{(0)}) + \alpha_1(x)(x_1 - x_1^{(0)})$$

(все остальные слагаемые равны 0). Поэтому

$$\begin{aligned}\frac{\partial f}{\partial x_1}(x^{(0)}) &= \lim_{x_1 \rightarrow x_1^{(0)}} \frac{f(x_1, x_2^{(0)}, \dots, x_n^{(0)}) - f(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})}{x_1 - x_1^{(0)}} \\ &= \lim_{x_1 \rightarrow x_1^{(0)}} \frac{A_1(x_1 - x_1^{(0)}) + \alpha_1(x)(x_1 - x_1^{(0)})}{x_1 - x_1^{(0)}} = \lim_{x_1 \rightarrow x_1^{(0)}} (A_1 + \alpha_1(x)) = A_1.\end{aligned}$$

Аналогичные рассуждения нужно провести для остальных $k = 2, \dots, n$. $\square$

Заметим, что из наличия частных производных по всем переменным в точке не вытекает дифференцируемость функции в этой точке. Действительно, функция

$$f(x, y) = \begin{cases} 1, & xy = 0, \\ 0, & xy \neq 0, \end{cases}$$

имеет в точке $(0, 0)$ частные производные

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

поскольку $f(x, 0) \equiv 1$, $f(0, y) \equiv 1$. Тем не менее, эта функция даже не является в точке $(0, 0)$ непрерывной. А дифференцируемая функция должна быть непрерывной:

ТЕОРЕМА 19.5. *Дифференцируемая в точке $x^{(0)}$ функция непрерывна в ней.*

ДОКАЗАТЕЛЬСТВО. Из формулы (19.1) вытекает, что

$$\lim_{x \rightarrow x^{(0)}} f(x) = \lim_{x \rightarrow x^{(0)}} \left(f(x^{(0)}) + \sum_{k=1}^n A_k(x_k - x_k^{(0)}) + o(\rho(x, x^{(0)})) \right) = f(x^{(0)}). \quad \square$$

Геометрический смысл условия дифференцируемости. Рассмотрим функцию двух переменных $z = f(x, y)$. Условие дифференцируемости этой функции в точке (x_0, y_0) имеет вид

$$f(x, y) = f(x_0, y_0) + A(x - x_0) + B(y - y_0) + o(\sqrt{(x - x_0)^2 + (y - y_0)^2}),$$

где $(x, y) \rightarrow (x_0, y_0)$. Здесь

$$z = f(x_0, y_0) + A(x - x_0) + B(y - y_0)$$

— уравнение некоторой плоскости, проходящей через точку с координатами $(x_0, y_0, f(x_0, y_0))$. По определению это *касательная плоскость* к графику дифференцируемой в точке (x_0, y_0) функции $f(x, y)$.

Итак, уравнение касательной плоскости выглядит следующим образом:

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

При этом условие дифференцируемости в точке (x_0, y_0) функции означает, что в окрестности этой точки график функции «похож на плоскость».

Проясним геометрический смысл частных производных. «Замораживая» переменную $y = y_0$, мы тем самым рассматриваем функцию $z = f(x, y_0)$, зависящую от одной переменной. Ее график — кривая, которая является пересечением поверхности $z = f(x, y)$ с плоскостью $y = y_0$. В этой плоскости кривая имеет касательную в точке (x_0, y_0) . Тангенс наклона этой касательной к горизонтали равен значению частной производной $\partial f(x_0, y_0)/\partial x$. Аналогичные рассуждения справедливы для частной производной по y .

ТЕОРЕМА 19.6 (достаточное условие дифференцируемости). *Пусть функция $f(x)$ определена в $O(x^{(0)})$ и все частные производные $\frac{\partial f}{\partial x_k}(x)$, $k = 1, \dots, n$, существуют и непрерывны в точке $x^{(0)}$. Тогда функция $f(x)$ дифференцируема в точке $x^{(0)}$.*

ДОКАЗАТЕЛЬСТВО. Проведем доказательство для $z = f(x, y)$ ($n = 2$). Имеем

$$\begin{aligned}f(x, y) - f(x_0, y_0) &= (f(x, y) - f(x_0, y)) + (f(x_0, y) - f(x_0, y_0)) \\ &= \underbrace{\frac{\partial f}{\partial x}(x_0 + \theta_1(x - x_0), y)(x - x_0)}_{\text{формула Лагранжа для функции } \varphi(x) = f(x, y) - f(x_0, y), \theta_1 = \theta_1(y), 0 < \theta_1 < 1} + \underbrace{\frac{\partial f}{\partial y}(x_0, y_0 + \theta_2(y - y_0))(y - y_0)}_{\text{формула Лагранжа для функции } \psi(y) = f(x_0, y) - f(x_0, y_0), 0 < \theta_2 < 1} \\ &= \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \alpha(x, y)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + \beta(x, y)(y - y_0),\end{aligned}$$

где

$$\alpha(x, y) = \frac{\partial f}{\partial x}(x_0 + \theta_1(x - x_0), y) - \frac{\partial f}{\partial x}(x_0, y_0), \quad \beta(x, y) = \frac{\partial f}{\partial y}(x_0, y_0 + \theta_2(y - y_0)) - \frac{\partial f}{\partial y}(x_0, y_0).$$

Функции $\alpha(x, y)$ и $\beta(x, y)$ являются бесконечно малыми при $(x, y) \rightarrow (x_0, y_0)$ в силу непрерывности функций $\partial f/\partial x$ и $\partial f/\partial y$ в точке (x_0, y_0) (поясним, что $(x_0 + \theta_1(x - x_0), y) \rightarrow (x_0, y_0)$ и $(x_0, y_0 + \theta_2(y - y_0)) \rightarrow (x_0, y_0)$). Остается сослаться на определение 19.2'. $\square$

ПРИМЕР. Функция $f(x, y) = xe^y$ имеет непрерывные частные производные $\frac{\partial f}{\partial x} = e^y$, $\frac{\partial f}{\partial y} = xe^y$. Эти функции непрерывны всюду. По теореме 19.6 функция f дифференцируема в любой точке плоскости.

Лекция 20

Смешанные производные. Теорема Шварца.

ОПРЕДЕЛЕНИЕ 20.1. Пусть $f(x) = f(x_1, \dots, x_n)$ — функция n переменных. Если $\frac{\partial f}{\partial x_i}(x_1, \dots, x_n)$ рассматривать как функцию n переменных, то ее вновь можно попытаться дифференцировать по любой переменной x_j .

Введем обозначения

$$\frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i}(x) \right) = \frac{\partial^2 f}{\partial x_j \partial x_i},$$

а в случае $i = j$

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

Например, если $f(x, y) = \frac{x}{y}$, то

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{1}{y}, \quad \frac{\partial f}{\partial y} = -\frac{x}{y^2}, \\ \frac{\partial^2 f}{\partial x^2} &= 0, \quad \frac{\partial^2 f}{\partial y^2} = \frac{2x}{y^3}, \quad \frac{\partial^2 f}{\partial x \partial y} = -\frac{1}{y^2}, \quad \frac{\partial^2 f}{\partial y \partial x} = -\frac{1}{y^2}. \end{aligned}$$

Заметим, что совпадение частных производных $\frac{\partial^2 f}{\partial x \partial y}$ и $\frac{\partial^2 f}{\partial y \partial x}$ не случайно.

ТЕОРЕМА 20.2 (о смешанных производных или теорема Шварца). Пусть функции $\frac{\partial^2 f}{\partial x \partial y}$ и $\frac{\partial^2 f}{\partial y \partial x}$ непрерывны в окрестности точки (x_0, y_0) . Тогда

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

ДОКАЗАТЕЛЬСТВО. Рассмотрим выражение

$$A = f(x_0 + h, y_0 + h) + f(x_0, y_0) - f(x_0 + h, y_0) - f(x_0, y_0 + h).$$

Если ввести обозначение $\varphi(t) = f(t, y_0 + h) - f(t, y_0)$, то

$$A = \underbrace{[f(x_0 + h, y_0 + h) - f(x_0 + h, y_0)]}_{\varphi(x_0 + h)} - \underbrace{[f(x_0, y_0 + h) - f(x_0, y_0)]}_{\varphi(x_0)}.$$

По формуле Лагранжа

$$A = \varphi(x_0 + h) - \varphi(x_0) = \varphi'(x_0 + \theta_1 h)h = \left(\frac{\partial f}{\partial x}(x_0 + \theta_1 h, y_0 + h) - \frac{\partial f}{\partial x}(x_0 + \theta_1 h, y_0) \right) h.$$

Введя обозначение $\psi(t) = \frac{\partial f}{\partial x}(x_0 + \theta_1 h, t)$, вновь воспользуемся формулой Лагранжа:

$$\begin{aligned} A &= (\psi(y_0 + h) - \psi(y_0))h = \psi'(y_0 + \theta_2 h)h^2 \\ &= \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x}(x_0 + \theta_1 h, y_0 + \theta_2 h) \right) h^2 = \frac{\partial^2 f}{\partial y \partial x}(x_0 + \theta_1 h, y_0 + \theta_2 h)h^2. \end{aligned}$$

Заметим теперь, что выражение A «симметрично» относительно переменных x и y . Поэтому, проведя указанные преобразования в другом порядке, получим формулу

$$A = \frac{\partial^2 f}{\partial x \partial y}(x_0 + \tilde{\theta}_1 h, y_0 + \tilde{\theta}_2 h)h^2.$$

(Поясним, что начинать при этом нужно с формулы

$$A = \underbrace{[f(x_0 + h, y_0 + h) - f(x_0, y_0 + h)]}_{\varphi(y_0 + h)} - \underbrace{[f(x_0 + h, y_0) - f(x_0, y_0)]}_{\varphi(y_0)}.$$

Далее по аналогии.)

Итак сравнивая два различных выражения для A , выписываем равенство

$$\frac{\partial^2 f}{\partial y \partial x}(x_0 + \theta_1 h, y_0 + \theta_2 h) = \frac{\partial^2 f}{\partial x \partial y}(x_0 + \tilde{\theta}_1 h, y_0 + \tilde{\theta}_2 h),$$

где $\theta_1, \theta_2, \tilde{\theta}_1$, и $\tilde{\theta}_2$, вообще говоря, зависят от h , но находятся в пределах от 0 до 1.

Переходя в этом равенстве к пределу при $h \rightarrow 0$, получаем, пользуясь непрерывностью функций $\frac{\partial^2 f}{\partial x \partial y}$ и $\frac{\partial^2 f}{\partial y \partial x}$ в точке (x_0, y_0) , нужное равенство

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

Дифференцирование сложной функции.

ТЕОРЕМА 20.3 (о дифференцировании сложной функции). Пусть функции $x_i(t)$, $i = 1, \dots, n$, $t \in \mathbb{R}$, дифференцируемы в точке t_0 . Пусть функция $f(x)$, $x \in \mathbb{R}^n$, дифференцируема в точке $a = (x_1(t_0), \dots, x_n(t_0))$. Тогда сложная функция $f(x(t))$ дифференцируема в точке t_0 , причем

$$\frac{df}{dt}(x(t_0)) = \sum_{i=1}^n \frac{\partial f(x(t_0))}{\partial x_i} \frac{dx_i(t_0)}{dt}. \quad (20.1)$$

ДОКАЗАТЕЛЬСТВО. Запишем условие дифференцируемости (20.1') функции $f(x)$ в точке a :

$$f(x) - f(a) = \sum_{i=1}^n \left(\frac{\partial f(a)}{\partial x_i} + \alpha_i(x) \right) (x_i - a_i), \quad (20.2)$$

где $\alpha_i(x) \rightarrow 0$ при $x \rightarrow a$. В силу дифференцируемости функций $x_i(t)$ имеем, в свою очередь,

$$x_i(t) - x_i(t_0) = (x'_i(t_0) + \beta_i(t)) \Delta t, \quad (20.3)$$

где $\beta_i(t) \rightarrow 0$ при $t \rightarrow t_0$. Подставляя (20.3) в (20.2), получаем:

$$f(x(t)) - f(a) = \sum_{i=1}^n \left(\frac{\partial f(a)}{\partial x_i} + \alpha_i(x(t)) \right) (x'_i(t_0) + \beta_i(t)) \Delta t.$$

Раскрывая скобки, получаем:

$$\begin{aligned} f(x(t)) &= f(x(t_0)) + \left(\sum_{i=1}^n \frac{\partial f(a)}{\partial x_i} x'_i(t_0) \right) \Delta t + \underbrace{\left(\sum_{i=1}^n \left(\frac{\partial f(a)}{\partial x_i} + \alpha_i(x(t)) \right) \beta_i(t) + \sum_{i=1}^n \alpha_i(x(t)) x'_i(t_0) \right) \Delta t}_{\text{ограниченная функция б.м.}} \\ &= f(x(t_0)) + A \Delta t + o(\Delta t), \quad \Delta t \rightarrow 0, \quad A = \sum_{i=1}^n \frac{\partial f(a)}{\partial x_i} x'_i(t_0). \end{aligned}$$

В силу условия дифференцируемости функции одной переменной (из первого семестра) функция $f(x(t))$ дифференцируема в точке t_0 , а число A — производная $df(x(t))/dt$. $\square$

ПРИМЕР 20.4 (не слишком содержательный). Пусть $f(x, y, z) = x^2 y^3 z^4$, $x = \cos t$, $y = \sin t$, $z = t$. Имеем

$$\begin{aligned} \frac{d}{dt} f(x(t), y(t), z(t)) &= \frac{\partial f}{\partial x} x'_t(t) + \frac{\partial f}{\partial y} y'_t(t) + \frac{\partial f}{\partial z} z'_t(t) = 2xy^3z^4(-\sin t) + 3x^2y^2z^4 \cos t + 4x^2y^3z^3 \\ &= -2t^4 \cos t \sin^4 t + 3t^4 \cos^2 t \sin^2 t + 4t^3 \cos^2 t \sin^3 t. \end{aligned}$$

ЗАМЕЧАНИЕ 20.5. Каждая из координат x_i может быть функцией не одной переменной t , а m переменных $t_1, \dots, t_m$ и можно рассмотреть следующую сложную функцию m переменных:

$$f(x(t)) = f(x_1(t_1, \dots, t_m), \dots, x_t(t_1, \dots, t_m)).$$

Можно показать, что если все функции $x_i(t)$ дифференцируемы в точке $t^{(0)} = (t_1^{(0)}, \dots, t_m^{(0)})$, а функция $f(x)$ дифференцируема в точке $x = x(t^{(0)})$, то сложная функция $f(x(t))$ дифференцируема в точке $t^{(0)}$. При

этом, «заморозив» все переменные t кроме t_j , из формулы 20.1 нетрудно получить следующую формулу для $\partial f / \partial t_j$:

$$\frac{\partial f(x(t))}{\partial t_j} = \sum_{i=1}^n \frac{\partial f(x(t))}{\partial x_i} \frac{\partial x_i(t)}{\partial t_j}. \quad (20.4)$$

Производная по направлению. Градиент.

ОПРЕДЕЛЕНИЕ 20.6. Пусть в окрестности точки $a = (a_1, \dots, a_m)$ задана функция $f(x)$. Рассмотрим луч в пространстве $\mathbb{R}^m$ с началом в точке a , определяемый единичным вектором $\vec{n} = (n_1, \dots, n_m)$:

$$\begin{cases} x_1(t) = a_1 + n_1 t, \\ \dots\dots\dots \\ x_m(t) = a_m + n_m t, \end{cases} \quad t \in [0, +\infty)$$

(это параметрическое запись луча).

Рассмотрим функцию $\varphi(t) = f(x(t)) = f(a_1 + n_1 t, \dots, a_m + n_m t)$. По определению производная $\varphi'_+(0)$ (если она существует) называется *производной функции $f(x)$ по направлению $\vec{n}$ в точке a* и обозначается $\frac{\partial f(a)}{\partial \vec{n}}$:

$$\frac{\partial f(a)}{\partial \vec{n}} = \varphi'_+(0) = \left. \frac{df(a_1 + n_1 t, \dots, a_m + n_m t)}{dt} \right|_{t=0}.$$

Смысл производной по направлению (как и любой производной) — скорость изменения функции, но не вообще, а в направлении, определяемом вектором $\vec{n}$.

Вычислим производную по направлению, предположив, что функция $f(x)$ дифференцируема в точке a . Воспользуемся формулой (20.1):

$$\frac{\partial f}{\partial \vec{n}} = \sum_{i=1}^m \frac{\partial f(a)}{\partial x_i} x'_i(0) = \sum_{i=1}^m \frac{\partial f(a)}{\partial x_i} \cdot n_i.$$

С геометрической точки зрения правая часть этой формулы представляет собой скалярное произведение вектора

$$\vec{g} = \left(\frac{\partial f(a)}{\partial x_1}, \dots, \frac{\partial f(a)}{\partial x_m} \right)$$

на вектор $\vec{n}$:

$$\frac{\partial f}{\partial \vec{n}} = (\vec{g}, \vec{n}).$$

Вектор $\vec{g}$ имеет специальное название; она называется *градиентом* функции $f(x)$ в точке a и обозначается $\text{grad } f(a)$.

Выясним геометрический смысл градиента функции. Запишем формулу для производной по направлению через скалярное произведение

$$\frac{\partial f}{\partial \vec{n}} = (\text{grad } f(a), \vec{n}) = \cos \alpha \cdot |\text{grad } f(a)| \cdot |\vec{n}| = \cos \alpha \cdot |\text{grad } f(a)|,$$

где α — угол между градиентом и вектором $\vec{n}$, и будем менять угол α , пытаясь добиться, чтобы производная по направлению стала максимальной. Нетрудно видеть, что этого можно добиться, положив $\alpha = 0$. Это означает, что *направление, в котором функция возрастает быстрее всего, совпадает с направлением градиента, а модуль градиента равен производной функции по этому направлению*.

Заметим, что если $\text{grad } f(a) = \vec{0} = (0, \dots, 0)$, то производная по любому направлению в точке a равна 0. Такая точка называется *критической*.

Лекция 21

Дифференциал функции многих переменных.

ОПРЕДЕЛЕНИЕ 21.1. Пусть функция $f(x)$ имеет в точке $x \in \mathbb{R}^n$ все частные производные первого порядка. Ее дифференциалом в точке x называется выражение

$$df(x) = \sum_{k=1}^n \frac{\partial f(x)}{\partial x_k} \Delta x_k. \quad (21.1)$$

Здесь Δx_k — приращение независимой переменной x_k . Часто вместо Δx_k пишут dx_k , так как если $f(x_1, \dots, x_n) = x_k$, то $df(x) = dx_k = 0 \cdot \Delta x_1 + \dots + 1 \cdot \Delta x_k + \dots + 0 \cdot \Delta x_n = \Delta x_k$.

Таким образом,

$$df(x) = \sum_{k=1}^n \frac{\partial f(x)}{\partial x_k} dx_k. \quad (21.1')$$

Если сравнить определение дифференциала с условием дифференцируемости функции $f(x)$ в точке a :

$$f(x + \Delta x) = f(x) + \sum_{k=1}^n \frac{\partial f(x)}{\partial x_k} \Delta x_k + o(\rho), \quad \rho \rightarrow 0$$

($\rho = |\Delta x| = \sqrt{(\Delta x_1)^2 + \dots + (\Delta x_n)^2}$), то мы обнаружим, что это условие эквивалентно условию

$$\Delta f(x) = df(x) + o(|\Delta x|), \quad |\Delta x| \rightarrow 0.$$

Здесь $\Delta f(x) = f(x + \Delta x) - f(x)$ — приращение функции $f(x)$, отвечающее приращению аргумента $\Delta x = (\Delta x_1, \dots, \Delta x_n)$.

ПРИМЕР 21.2. Пусть $f = f(x, y)$. Тогда

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

Например, в случае $f(x, y) = x^y + xy$ имеем

$$df(x, y) = (yx^{y-1} + y) \Delta x + (x^y \ln x + x) \Delta y.$$

Если $(x, y) = (2, 1)$, то

$$df(2, 1) = 2 \Delta x + 2(\ln 2 + 1) \Delta y.$$

Геометрический смысл дифференциала. Для графика функции $z = f(x, y)$ уравнение касательной в точке (x_0, y_0) имеет вид

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

Это уравнение можно записать в виде

$$\Delta z = df(x_0, y_0).$$

В этом смысле дифференциал функции в точке определяет уравнение касательной плоскости в этой точке.

ЗАМЕЧАНИЕ 21.3. Дифференциал $df(x)$ является функцией от $2n$ переменных $(x_1, \dots, x_n, \Delta x_1, \dots, \Delta x_n)$, причем $f(x)$ (при фиксированном x) является линейной функцией $\Delta x_1, \dots, \Delta x_n$.

Свойства дифференциала $df(x)$:

1. *Линейность:* $d(\lambda u + \mu v) = \lambda du + \mu dv$.
2. $d(uv) = u dv + v du$.
3. $d\left(\frac{u}{v}\right) = \frac{v du - u dv}{v^2}$.

Докажем, например, второе свойство:

$$\begin{aligned} d(uv) &= d(u(x)v(x)) = \sum_{k=1}^n \frac{\partial(uv)}{\partial x_k} dx_k = \sum_{k=1}^n \left(v \frac{\partial u}{\partial x_k} + u \frac{\partial v}{\partial x_k} \right) dx_k \\ &= v \sum_{k=1}^n \frac{\partial u}{\partial x_k} dx_k + u \sum_{k=1}^n \frac{\partial v}{\partial x_k} dx_k = v du + u dv. \end{aligned}$$

Инвариантность дифференциала относительно замены переменных. Допустим, что $x_k = x_k(t)$, $t \in \mathbb{R}^m$.

Тогда $dx_k = \sum_{l=1}^m \frac{\partial x_k}{\partial t_l} dt_l$. Подставив это в формулу (21.1'), получим следующий результат:

$$df(x) = \sum_{k=1}^n \frac{\partial f}{\partial x_k} \sum_{l=1}^m \frac{\partial x_k}{\partial t_l} dt_l = \sum_{l=1}^m \left(\sum_{k=1}^n \frac{\partial f}{\partial x_k} \frac{\partial x_k}{\partial t_l} \right) dt_l.$$

С другой стороны, если $f(x(t))$ рассматривать как сложную функцию, то

$$df(x(t)) = \sum_{l=1}^m \frac{\partial f}{\partial t_l} dt_l.$$

А поскольку (см. формулу (20.4))

$$\frac{\partial f}{\partial t_l} = \sum_{k=1}^n \frac{\partial f}{\partial x_k} \frac{\partial x_k}{\partial t_l},$$

мы вновь имеем формулу

$$df(x) = \sum_{l=1}^m \left(\sum_{k=1}^n \frac{\partial f}{\partial x_k} \frac{\partial x_k}{\partial t_l} \right) dt_l.$$

Итак, формула для $df(x)$ не зависит от того являются ли $x_1, \dots, x_n$ независимыми переменными или, в свою очередь, зависят от других переменных. Это свойство называют *инвариантностью дифференциала относительно замены переменных*.

Дифференциалы высших порядков

ОПРЕДЕЛЕНИЕ 21.4. По определению

$$d^2 f(x) = d(df(x)).$$

Что это означает? Дело в том, что $df(x)$ является (при фиксированных $\Delta x_1, \dots, \Delta x_n$) дифференцируемой функцией от x , то мы можем выписать дифференциал этой функции:

$$\begin{aligned} d^2 f(x) &= d(df(x)) = d\left(\sum_{k=1}^n \frac{\partial f(x)}{\partial x_k} \Delta x_k\right) = \sum_{k=1}^n d\left(\frac{\partial f(x)}{\partial x_k}\right) \Delta x_k \\ &= \sum_{k=1}^n \left(\sum_{s=1}^n \frac{\partial}{\partial x_s} \left(\frac{\partial f(x)}{\partial x_k}\right) \Delta x_s\right) \Delta x_k = \sum_{k=1}^n \sum_{s=1}^n \frac{\partial^2 f(x)}{\partial x_s \partial x_k} \Delta x_s \Delta x_k. \end{aligned}$$

Отметим, что здесь реально не n^2 , а $n(n+1)/2$ слагаемых (по теореме Шварца). Второй дифференциал функции $f(x)$ является квадратичной формой (определение будет позже) переменных $\Delta x_1, \dots, \Delta x_n$.

Дифференциалы и производные порядка ≥ 2 определяются рекуррентным образом. В частности,

$$d^m f(x) \stackrel{\text{def}}{=} d(d^{m-1} f(x)).$$

Выпишем формулы для $d^2 f(x)$ в частных случаях $f(x, y)$ и $f(x, y, z)$:

$$\begin{aligned} df(x, y) &= \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy, \\ d^2 f(x, y) &= \frac{\partial^2 f}{\partial x^2} dx^2 + 2 \frac{\partial^2 f}{\partial x \partial y} dx dy + \frac{\partial^2 f}{\partial y^2} dy^2, \\ df(x, y, z) &= \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz, \\ d^2 f(x, y, z) &= \frac{\partial^2 f}{\partial x^2} dx^2 + \frac{\partial^2 f}{\partial y^2} dy^2 + \frac{\partial^2 f}{\partial z^2} dz^2 + 2 \frac{\partial^2 f}{\partial x \partial y} dx dy + 2 \frac{\partial^2 f}{\partial x \partial z} dx dz + 2 \frac{\partial^2 f}{\partial y \partial z} dy dz. \end{aligned}$$

Можно показать, что для $d^m f$ справедлива формула

$$d^m f = \left(\sum_{k=1}^n \frac{\partial}{\partial x_k} dx_k \right)^m f.$$

которую нужно воспринимать так: для применения этой формулы нужно, воспринимая символы $\frac{\partial}{\partial x_k}$ и dx_k как буквы (иероглифы), раскрыть скобки и привести подобные слагаемые.

Например,

$$\begin{aligned} df(x, y) &= \left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy \right)^2 f = \left(\left(\frac{\partial}{\partial x} dx \right)^2 + 2 \frac{\partial}{\partial x} dx \frac{\partial}{\partial y} dy + \left(\frac{\partial}{\partial y} dy \right)^2 \right) f \\ &= \left(\frac{\partial^2}{\partial x^2} dx^2 + 2 \frac{\partial^2}{\partial x \partial y} dx dy + \frac{\partial^2}{\partial y^2} dy^2 \right) f = \frac{\partial^2 f}{\partial x^2} dx^2 + 2 \frac{\partial^2 f}{\partial x \partial y} dx dy + \frac{\partial^2 f}{\partial y^2} dy^2. \end{aligned}$$

Формула Тейлора. Будем говорить, что функция $f(x)$ принадлежит классу C^m в области $\Omega \subset \mathbb{R}^n$ (будем писать $f \in C^m(\Omega)$), если все частные производные до порядка m включительно существуют и непрерывны в области Ω .

ТЕОРЕМА 21.5 (формула Тейлора с остаточным членом в форме Лагранжа). Пусть $f(x)$ — функция n переменных, принадлежащая классу C^{m+1} в окрестности точки x . Тогда при достаточно малых $|\Delta x|$ справедлива формула

$$f(x + \Delta x) = f(x) + \sum_{k=1}^m \frac{d^k f(x)}{k!} + \frac{d^{m+1}(x + \theta \Delta x)}{(m+1)!}. \quad (21.2)$$

ДОКАЗАТЕЛЬСТВО. Рассмотрим функцию $\varphi(t) = f(x(t))$, где $x(t) = x + t\Delta x = (x_1 + t\Delta x_1, \dots, x_n + t\Delta x_n)$, $x_k(t) = x_k + t\Delta x_k$. По теореме о дифференцировании сложной функции

$$\varphi'(t) = \sum_{k=1}^n \frac{\partial f(x(t))}{\partial x_k} \frac{dx_k}{dt} = \sum_{k=1}^n \frac{\partial f(x(t))}{\partial x_k} \Delta x_k = df(x(t)).$$

Итак, $\varphi'(t) = df(x(t))$. Воспользуемся этим результатом еще раз:

$$\varphi''(t) = (\varphi'(t))' = d(df(x(t))) = d^2 f(x(t)).$$

Применяя этот «фокус» несколько раз, получаем:

$$\varphi^{(k)}(t) = d^k f(x(t)), \quad k = 1, \dots, m+1.$$

Запишем теперь формулу Тейлора с остаточным членом в форме Лагранжа с центром в точке $t_0 = 0$ для функции $\varphi(t)$ при $t = 1$:

$$\varphi(1) = \varphi(0) + \sum_{k=1}^m \frac{\varphi^{(k)}(0)}{k!} 1^k + \frac{\varphi^{(m+1)}(\theta)}{(m+1)!} 1^{m+1}.$$

Заменив $\varphi^{(k)}(t)$ на $d^k f(x(t))$, получим

$$f(x + \Delta x) = f(x) + \sum_{k=1}^m \frac{d^k f(x)}{k!} + \frac{d^{m+1}(x + \theta \Delta x)}{(m+1)!}. \quad \square$$

ЗАМЕЧАНИЕ 21.6. Остаточный член в формуле (21.2) можно записать в виде $o(|\Delta x|^m)$, $|\Delta x| \rightarrow 0$. Получится формула Тейлора с остаточным членом в форме Пеано:

$$f(x + \Delta x) = f(x) + \sum_{k=1}^m \frac{d^k f(x)}{k!} + o(|\Delta x|^m), \quad |\Delta x| \rightarrow 0. \quad (21.2')$$

При этом требования на функцию $f(x)$ можно снизить: достаточно требовать, чтобы функция принадлежала классу C^m в окрестности точки x .

ЗАМЕЧАНИЕ 21.7. При $m = 1$ формула (21.2') совпадает условием дифференцируемости функции.

Лекция 22

Экстремум функции нескольких переменных.

ОПРЕДЕЛЕНИЕ 22.1. Пусть функция $f(x)$ определена в окрестности $O(a)$.

Точка a называется точкой *минимума* функции $f(x)$, если существует $\delta > 0$, такое, что $f(x) \geq f(a)$ для любого $x \in O_\delta(a)$.

Точка a называется точкой *строгого минимума* функции $f(x)$, если существует $\delta > 0$, такое, что $f(x) > f(a)$ для любого $x \in \dot{O}_\delta(a)$.

Точка a называется точкой *максимума* функции $f(x)$, если существует $\delta > 0$, такое, что $f(x) \leq f(a)$ для любого $x \in O_\delta(a)$.

Точка a называется точкой *строгого максимума* функции $f(x)$, если существует $\delta > 0$, такое, что $f(x) < f(a)$ для любого $x \in \dot{O}_\delta(a)$.

Точки указанных типов называются точками *экстремума* (строгого или нестрогого).

ТЕОРЕМА 22.2 (необходимое условие экстремума). Если в точке экстремума существует частная производная, то она обращается в 0.

ДОКАЗАТЕЛЬСТВО. Рассмотрим вспомогательную функцию $\varphi(t) = f(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n)$. Эта функция определена в некоторой окрестности $O(0)$, так как точка $(a_1, \dots, a_{j-1}, a_j + t, a_{j+1}, \dots, a_n)$ находится на расстоянии $|t|$ от точки $(a_1, \dots, a_{j-1}, a_j, a_{j+1}, \dots, a_n)$. Кроме того, функция $\varphi(t)$ имеет в точке 0 экстремум. Действительно, если a — точка минимума для $f(x)$, то $f(x) \geq f(a)$ для любой точки $x \in O_\delta(a)$. Следовательно,

$$\underbrace{f(a_1, \dots, a_j + t, \dots, a_n)}_{\varphi(t)} \geq \underbrace{f(a)}_{\varphi(0)}$$

для всех t , таких, что $|t| \leq \delta$. Согласно необходимому условию экстремума функции одной переменной $\varphi'(t) = 0$, или, что то же $\frac{\partial f(a)}{\partial x_j} = 0$. $\square$

СЛЕДСТВИЕ 22.3. Если функция $f(x)$ дифференцируема в точке экстремума a , то $df(a) = 0$.

Действительно, $df(a) = \sum_{j=1}^n \frac{\partial f(a)}{\partial x_j} dx_j \equiv 0$.

Еще одна интерпретация необходимого условия экстремума: $\text{grad } f(a) = 0$.

Геометрически, в точке экстремума касательная плоскость (если она существует) горизонтальна $z = f(x_0, y_0) + 0 \cdot (x - x_0) + 0 \cdot (y - y_0)$, т.е. $z = f(x_0, y_0)$.

ПРИМЕР. а) $z = x^2 + y^2$. Эта функция принадлежит C^1 всюду на плоскости. Необходимое условие экстремума $\text{grad } f(x) = (0, 0)$ выполняется в единственной точке $(0, 0)$:

$$\begin{cases} \partial z / \partial x = 0, \\ \partial z / \partial y = 0, \end{cases} \iff \begin{cases} 2x = 0, \\ 2y = 0. \end{cases}$$

И эта точка действительно является точкой минимума: $x^2 + y^2 \geq 0$.

б) $z = xy$. Эта функция принадлежит C^1 всюду на плоскости. Необходимое условие экстремума $\text{grad } f(x) = (0, 0)$ выполняется в единственной точке $(0, 0)$

$$\begin{cases} \partial z / \partial x = 0, \\ \partial z / \partial y = 0, \end{cases} \iff \begin{cases} y = 0, \\ x = 0. \end{cases}$$

Однако, эта точка не является точкой экстремума: в любой окрестности $O_\delta(0, 0)$ есть точка $(\delta/2, \delta/2)$, в которой $f \geq f(0, 0) = 0$, и точка $(-\delta/2, \delta/2)$, в которой $f \leq f(0, 0) = 0$.

Ничего удивительного в этом нет. Мы нашли точку при помощи необходимого условия. Оно позволяет только найти точки, подозрительные на экстремум. Подозрительная точка есть (только одна), но подозрения не оправдались.

ОПРЕДЕЛЕНИЕ 22.4. Точка a называется *стационарной* или *критической точкой* функции $f(x)$, если $\text{grad } f(a) = 0$.

Итак, если функция $f(x)$ дифференцируема в точке экстремума, то эта точка является стационарной. Обратное утверждение неверно (см. пример $z = xy$).

ЛЕММА 22.5 (необходимое условие экстремума функции одной переменной). Пусть точка 0 является точкой минимума функции $\varphi(t)$, $t \in \mathbb{R}$, и существуют $\varphi'(0)$ и $\varphi''(0)$. Тогда $\varphi'(0) = 0$, а $\varphi''(0) \geq 0$.

ДОКАЗАТЕЛЬСТВО. По теореме Ферма $\varphi'(0) = 0$. Рассмотрим разность $\frac{\varphi(t) - \varphi(0)}{t^2}$. по определению минимума эта разность неотрицательна при $t \in \dot{O}_\delta(0)$. С другой стороны, по формуле Тейлора с остаточным членом в форме Пеано

$$\frac{\varphi(t) - \varphi(0)}{t^2} = \frac{\varphi'(0)t + \frac{1}{2}\varphi''(0)t^2 + o(t^2)}{t^2} = \frac{1}{2}\varphi''(0) + o(1), \quad t \rightarrow 0.$$

Таким образом, $\frac{1}{2}\varphi''(0) + o(1) \geq 0$ при $t \in \dot{O}_\delta(0)$. Переходя в этом неравенстве к пределу при $t \rightarrow 0$, получаем $\varphi''(0) \geq 0$. $\square$

При помощи этой леммы доказывается

ТЕОРЕМА 22.6 (необходимое условие экстремума). Пусть функция $f(x)$ имеет в точке a все частные производные до второго порядка включительно.

Тогда если a — точка минимума, то выполняются условия

$$df(a) = 0, \quad d^2f(a) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(a)}{\partial x_j \partial x_i} dx_i dx_j \geq 0.$$

Если a — точка максимума, то

$$df(a) = 0, \quad d^2f(a) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(a)}{\partial x_j \partial x_i} dx_i dx_j \leq 0.$$

Поясним: условие $d^2f(a) \geq 0$ означает, что $d^2f(a) \geq 0$ при всех значениях $dx_1, \dots, dx_n$.

ДОКАЗАТЕЛЬСТВО. Пусть для определенности a — точка минимума. Рассмотрим вспомогательную функцию

$$\varphi(t) = f(a + t\Delta x) = f(a_1 + t\Delta x_1, \dots, a_n + t\Delta x_n).$$

Если $|t|$ достаточно мало ($|t| < \delta/|\Delta x|$), то $a + t\Delta x \in O_\delta(a)$. Поэтому в точке 0 функция $\varphi(t)$ имеет минимум. Кроме того (см. доказательство формулы Тейлора)

$$\varphi'(0) = \sum_{i=1}^n \frac{\partial f(a)}{\partial x_i} \Delta x_i, \quad \varphi''(0) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(a)}{\partial x_i \partial x_j} \Delta x_i \Delta x_j.$$

По лемме 22.5 $\varphi'(0) = 0$ и $\varphi''(0) \geq 0$.

Случай, когда a — точка максимума сводится к разобранному заменой функции $f(x)$ на функцию $-f(x)$. $\square$

ПРИМЕР. Рассмотрим функцию $f(x, y) = \sin x \sin y$.

Чтобы найти стационарные точки, приравняем $\partial f / \partial x = \cos x \sin y$ и $\partial f / \partial y = \sin x \cos y$ к нулю. Получим систему уравнений

$$\begin{cases} \cos x \sin y = 0, \\ \sin x \cos y = 0, \end{cases}$$

решениями которой, в частности, являются точки $(0, 0)$ и $(\pi/2, \pi/2)$ (разумеется, есть и другие решения). Подсчитаем второй дифференциал в этих точках:

$$d^2 f(0, 0) = \frac{\partial^2 f(0, 0)}{\partial x^2} dx^2 + 2 \frac{\partial^2 f(0, 0)}{\partial x \partial y} dx dy + \frac{\partial^2 f(0, 0)}{\partial y^2} dy^2 = 0 \cdot dx^2 + 2 dx dy + 0 \cdot dy^2 = 2 dx dy.$$

Итак, $d^2 f(0, 0) = 2 dx dy$. Ясно, что $d^2 f(0, 0) = 2 dx dy$ может менять знак (например $d^2 f = 2$ при $dx = dy = 1$, но $d^2 f = -1$ при $dx = -1, dy = 1$). Это означает, что точка $(0, 0)$ не является точкой экстремума.

Рассмотрим точку $(\pi/2, \pi/2)$. Второй дифференциал в этой точке таков:

$$\begin{aligned} d^2 f\left(\frac{\pi}{2}, \frac{\pi}{2}\right) &= \frac{\partial^2 f\left(\frac{\pi}{2}, \frac{\pi}{2}\right)}{\partial x^2} dx^2 + 2 \frac{\partial^2 f\left(\frac{\pi}{2}, \frac{\pi}{2}\right)}{\partial x \partial y} dx dy + \frac{\partial^2 f\left(\frac{\pi}{2}, \frac{\pi}{2}\right)}{\partial y^2} dy^2 \\ &= -1 \cdot dx^2 + 2 \cdot 0 \cdot dx dy - 1 \cdot dy^2 = -dx^2 - dy^2. \end{aligned}$$

Таким образом, в точке $(\pi/2, \pi/2)$ выполнено необходимое условие максимума. Но это, к сожалению, не означает, что это точка максимума. Остается только подозрение, что это максимум. Нужны какие-нибудь достаточные условия.

Для того, чтобы сформулировать достаточные условия, нужно вспомнить кое-что из курса линейной алгебры:

ОПРЕДЕЛЕНИЕ. *Квадратичной формой* в $\mathbb{R}^n$ называется функция вида

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

Симметричная матрица

$$(a_{ij}) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

называется матрицей квадратичной формы $\Phi(\xi)$.

Например, $\Phi(\xi_1, \xi_2, \xi_3) = \xi_1^2 - 2\xi_1\xi_2 + \xi_1\xi_3 - \xi_3^2$ — квадратичная форма с матрицей $\begin{pmatrix} 1 & -1 & \frac{1}{2} \\ -1 & 0 & 0 \\ \frac{1}{2} & 0 & -1 \end{pmatrix}$.

ОПРЕДЕЛЕНИЕ. Квадратичная форма $\Phi(\xi)$ называется *знакоположительной*, если $\Phi(\xi) > 0$ при любом $\xi \neq 0$ (т.е. $(\xi_1, \dots, \xi_n) \neq (0, \dots, 0)$).

Квадратичная форма $\Phi(\xi)$ называется *знакоотрицательной*, если $\Phi(\xi) < 0$ при любом $\xi \neq 0$.

В этих двух случаях форма называется *знакоопределенной*.

Квадратичная форма $\Phi(\xi)$ называется *неопределенной*, если существуют две точки $\xi^{(1)}$ и $\xi^{(2)}$, такие, что $\Phi(\xi^{(1)}) < 0$, а $\Phi(\xi^{(2)}) > 0$.

Например, форма $\Phi(\xi_1, \xi_2) = -\xi_1^2 - \xi_2^2$ является знакоотрицательной: $-\xi_1^2 - \xi_2^2 < 0$, если $(\xi_1, \xi_2) \neq (0, 0)$.

Форма $\Phi(\xi_1, \xi_2) = 2\xi_1^2 + \xi_2^2$ знакоположительна.

Форма $\Phi(\xi_1, \xi_2) = \xi_1\xi_2$ является неопределенной.

ТЕОРЕМА 22.7 (критерий Сильвестра). *Квадратичная форма*

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

знакоположительна тогда и только тогда, когда все главные миноры положительны:

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \dots, \quad \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} > 0.$$

Квадратичная форма знакоотрицательна тогда и только тогда, когда главные миноры чередуют знаки, причем $a_{11} < 0$:

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

Например, квадратичная форма $\Phi(\xi_1, \xi_2, \xi_3) = \xi_1^2 - 2\xi_1\xi_2 + \xi_1\xi_3 - \xi_3^2$ с матрицей $\begin{pmatrix} 1 & -1 & \frac{1}{2} \\ -1 & 0 & 0 \\ \frac{1}{2} & 0 & -1 \end{pmatrix}$ (см. выше) является неопределенной:

$$a_{11} = 1 > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = 1 \cdot \dots \cdot 0 - (-1) \cdot (-1) = -1 < 0.$$

ЛЕММА. Пусть $\Phi(\xi)$ — положительно определенная квадратичная форма. Тогда существует такое $\gamma > 0$, что $\Phi(\xi) \geq \gamma|\xi|^2$ при всех $\xi \in \mathbb{R}^n$.

ДОКАЗАТЕЛЬСТВО. Пусть $\Gamma = \{\xi \in \mathbb{R}^n : |\xi| = 1\}$ — сфера радиуса 1 в $\mathbb{R}^n$. Эта сфера является компактом, так как это множество замкнуто (если $\xi_n \rightarrow \xi$ и $|\xi_n| = 1$, то $|\xi| = 1$) и ограничено. По теореме Вейерштрасса (точнее, по следствию 19.8) существует точка $\xi_0 \in \Gamma$, такая, что $\Phi(\xi) \geq \Phi(\xi_0)$ для всех ξ . Обозначим $\Phi(\xi_0)$ через γ . Из положительной определенности Φ следует, что $\gamma > 0$. Далее, при любом $\xi \neq 0$ точка $\xi/|\xi|$ находится на Γ , поскольку

$$\left| \frac{\xi}{|\xi|} \right| = \frac{|\xi|}{|\xi|} = 1.$$

Поэтому при любом $\xi \neq 0$

$$\Phi\left(\frac{\xi}{|\xi|}\right) \geq \gamma.$$

Остается заметить, что $\Phi\left(\frac{\xi}{|\xi|}\right) = \frac{1}{|\xi|^2}\Phi(\xi)$ при $\xi \neq 0$

$$\left(\Phi\left(\frac{\xi}{|\xi|}\right) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} \frac{\xi_i}{|\xi|} \frac{\xi_j}{|\xi|} = \frac{1}{|\xi|^2} \sum_{i=1}^n \sum_{j=1}^n a_{ij} \xi_i \xi_j = \frac{1}{|\xi|^2} \Phi(\xi)\right)$$

Итак, $\frac{1}{|\xi|^2}\Phi(\xi) \geq \gamma$ для всех $\xi \neq 0$. □

ТЕОРЕМА (о достаточном условии экстремума). Пусть $f(x)$ принадлежит C^2 в окрестности точки a . Если $df(a) = 0$ и $d^2f(a) > 0$ (дифференциал $d^2f(a)$ положительно определен), то точка a является точкой строгого минимума.

Если $df(a) = 0$ и $d^2f(a) < 0$ (дифференциал $d^2f(a)$ отрицательно определен), то точка a является точкой строгого максимума.

ДОКАЗАТЕЛЬСТВО. Пусть $df(a) = 0$, $d^2f(a) > 0$. Запишем формулу Тейлора с остаточным членом в форме Пеано

$$f(x) = f(a) + df(a) + \frac{1}{2}d^2f(a) + o(|\Delta x|^2), \quad |\Delta x| \rightarrow 0.$$

Учитывая, что $df(a) = 0$, имеем

$$f(x) - f(a) = \frac{1}{2}d^2f(a) + o(|\Delta x|^2), \quad |\Delta x| \rightarrow 0.$$

По лемме имеем $d^2f(a) \geq \gamma|\Delta x|^2$. Отсюда

$$f(x) - f(a) \geq \frac{1}{2}\gamma|\Delta x|^2 + o(|\Delta x|^2) \geq \frac{1}{2}\gamma|\Delta x|^2(1 + o(1)), \quad |\Delta x| \rightarrow 0.$$

Величина $1 + o(1)$ положительна при достаточно малых $|\Delta x|$, т.е. при $x \in O_\delta(a)$. Это означает, что $f(x) - f(a) > 0$ при $x \in \dot{O}_\delta(a)$.

Второй случай (точка максимума) доказывается аналогично. □

ПРИМЕР. Вернемся к примеру $f(x, y) = \sin x \sin y$. Проверим является ли стационарная точка $(\pi/2, \pi/2)$ точкой экстремума.

Имеем $d^2f(\pi/2, \pi/2) = -dx^2 - dy^2 < 0$. Это отрицательно определенная форма. Поэтому $(\pi/2, \pi/2)$ — точка строгого максимума для $f(x, y)$.

Лекция 23

Функции, заданные неявно.

Попробуем ответить на следующий вопрос: при каких условиях из уравнения $F(x, y) = 0$ можно «выразить» y ? Точнее, при каких условиях это уравнение определяет функцию $y = f(x)$.

Рассмотрим «типичный» пример: $F(x, y) = x^2 + y^2 - 1$. Отметим, что «школьное решение» $y = \pm\sqrt{1-x^2}$ нас не устраивает, поскольку это выражение задает целых две функции. Ясно также, что «в целом» уравнение $x^2 + y^2 - 1 = 0$ не задает функцию $y = f(x)$, поскольку задает на множестве окружность, которая не может быть графиком функции (почему?). Однако, в окрестности почти любой своей точки (т.е. локально) эта кривая является графиком некоторой функции. Например, в окрестности точки $(-\sqrt{2}/2, -\sqrt{2}/2)$ окружность является графиком функции $y = -\sqrt{1-x^2}$. Легко сообразить, что в малой окрестности любой точки (x_0, y_0) окружности с $y_0 > 0$ окружность задается (локально) функцией $y = \sqrt{1-x^2}$. Если же $y_0 < 0$, то локально $y = -\sqrt{1-x^2}$. Есть две «неудобные» точки $(-1, 0)$ и $(1, 0)$, ни в какой окрестности которых невозможно выразить y через x . Однако, в окрестности этих точек можно выразить x через y .

ОПРЕДЕЛЕНИЕ. Пусть $F(x, y)$ — заданная в некоторой окрестности точки (x_0, y_0) , причем $F(x_0, y_0) = 0$. Будем говорить, что уравнение $F(x, y) = 0$ задает в окрестности точки (x_0, y_0) функцию $y = f(x)$ неявно, если существует такая окрестность $O_\delta(x_0, y_0)$, в которой уравнения $F(x, y) = 0$ и $y = f(x)$ определяют одно и то же множество точек.

Но бывает ли так, что уравнение $F(x, y) = 0$ нельзя разрешить (хотя бы локально) в окрестности некоторой точки. Да, например, уравнение $x^2 - y^2 = 0$ неразрешимо в окрестности точки $(0, 0)$.

ТЕОРЕМА 23.1 (1-я теорема о неявной функции). Пусть $F(x, y)$ принадлежит классу C^1 в некоторой окрестности $O(x_0, y_0)$. Пусть $F(x_0, y_0) = 0$. Пусть $\frac{\partial F(x_0, y_0)}{\partial y} \neq 0$. Тогда в некоторой (малой) окрестности точки (x_0, y_0) уравнение $F(x, y) = 0$ задает неявно дифференцируемую функцию $y = f(x)$, причем

$$f'(x_0) = -\frac{\frac{\partial F}{\partial x}(x_0, y_0)}{\frac{\partial F}{\partial y}(x_0, y_0)}.$$

ДОКАЗАТЕЛЬСТВО. Пусть для определенности $\frac{\partial F}{\partial y}(x_0, y_0) > 0$. В силу непрерывности функции $\partial F/\partial y$ это неравенство справедливо в некоторой окрестности $O_\delta(x_0, y_0)$.

Нетрудно видеть, что $F(x_0, y_0 - \delta/2) < 0$, а $F(x_0, y_0 + \delta/2) > 0$. В самом деле, функция $\varphi(y) = F(x_0, y)$ на интервале $(y_0 - \delta, y_0 + \delta)$ обладает следующими свойствами: $\varphi(y_0) = 0$, $\varphi'(y) = \partial F/\partial y > 0$, т.е. монотонно возрастает и обращается в 0 в точке y_0 . Тогда, конечно, $\varphi(y_0 - \delta/2) < 0$ и $\varphi(y_0 + \delta/2) > 0$.

Далее, $F(x, y)$ непрерывна в $O_\delta(x_0, y_0)$. Поэтому существует такая окрестность $O_{\delta_1}(x_0, y_0 + \delta/2)$, в которой $F(x, y) > 0$ и окрестность $O_{\delta_2}(x_0, y_0 - \delta/2)$, в которой $F(x, y) < 0$. Положим $\delta_3 = \min\{\delta_1, \delta_2\}$.

Ограничим все наши рассуждения прямоугольником

$$\Pi = \begin{cases} x_0 - \delta_3 \leq x \leq x_0 + \delta_3, \\ y_0 - \delta/2 \leq y \leq y_0 + \delta/2. \end{cases}$$

Возьмем произвольное число $x^* \in [x_0 - \delta_3, x_0 + \delta_3]$ и рассмотрим функцию $\varphi(y) = F(x^*, y)$ при $y \in [y_0 - \delta/2, y_0 + \delta/2]$. Имеем $\varphi(y_0 - \delta/2) < 0$, $\varphi(y_0 + \delta/2) > 0$, $\varphi'(y) = \partial F/\partial y > 0$. Это означает, что на отрезке существует (по теореме Коши) и единственна (из монотонности) точка y^* , в которой $\varphi(y^*) = 0$.

Итак, для любого $x^* \in [x_0 - \delta_3, x_0 + \delta_3]$ существует и единственна точка $(x^*, y^*) \in \Pi$, такая, что $F(x^*, y^*) = 0$. Эту точку мы будем обозначать $f(x^*)$. Мы построили функцию $f(x)$.

Докажем теперь, что функция обладает нужными свойствами. Запишем формулу Тейлора с остаточным членом в форме Лагранжа с центром в точке (x_0, y_0) и $m = 1$:

$$F(x, y) = F(x_0, y_0) + \frac{\partial F}{\partial x}(x_0 + \theta\Delta x, y_0 + \theta\Delta y)\Delta x + \frac{\partial F}{\partial y}(x_0 + \theta\Delta x, y_0 + \theta\Delta y)\Delta y, \quad 0 < \theta < 1.$$

Здесь $F(x_0, y_0) = 0$ и если $y = f(x)$, то $F(x, y) = 0$; поэтому

$$\frac{\partial F}{\partial x}(x_0 + \theta\Delta x, y_0 + \theta\Delta y)\Delta x + \frac{\partial F}{\partial y}(x_0 + \theta\Delta x, y_0 + \theta\Delta y)\Delta y = 0,$$

откуда

$$\frac{\Delta y}{\Delta x} = -\frac{\frac{\partial F}{\partial x}(x_0 + \theta\Delta x, y_0 + \theta\Delta y)}{\frac{\partial F}{\partial y}(x_0 + \theta\Delta x, y_0 + \theta\Delta y)}. \quad (23.1)$$

Покажем сначала, что $\Delta y \rightarrow 0$ при $\Delta x \rightarrow 0$. Действительно, $\partial F/\partial y$ непрерывна на компакте Π ; поэтому на Π существует такая точка, в которой $\partial F/\partial y$ принимает свое минимальное значение m . При этом $m > 0$, поскольку $\partial F/\partial y > 0$ в прямоугольнике Π . Кроме того, будучи непрерывной в Π , функция $\partial F/\partial x$ ограничена: $|\partial F/\partial x| \leq M$. Таким образом,

$$\left| \frac{\Delta y}{\Delta x} \right| = \left| \frac{\frac{\partial F}{\partial x}(x_0 + \theta\Delta x, y_0 + \theta\Delta y)}{\frac{\partial F}{\partial y}(x_0 + \theta\Delta x, y_0 + \theta\Delta y)} \right| \leq \frac{M}{m}.$$

Отсюда видно, что если $\Delta x \rightarrow 0$, то $\Delta y \rightarrow 0$ тоже, т.е. функция $f(x)$ непрерывна. Теперь, учитывая это обстоятельство, переходим к пределу в формуле (23.1):

$$f'(x_0) = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = - \lim_{\Delta x \rightarrow 0} \frac{\frac{\partial F}{\partial x}(x_0, y_0)}{\frac{\partial F}{\partial y}(x_0, y_0)}.$$

Остается заметить, что функция $f'(x)$ непрерывна в точке x_0 , поскольку задается она формулой

$$f'(x) = - \frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))},$$

а в точке x_0 функция $f(x)$ непрерывна, функции $\partial F/\partial x$, $\partial F/\partial y$ также непрерывны и можно применить теорему о непрерывности сложной функции.

СЛЕДСТВИЕ 23.2. Пусть $F(x, y)$ принадлежит C^1 в некоторой окрестности точки (x_0, y_0) . Пусть $F(x_0, y_0) = 0$. Тогда если $\text{grad } F(x_0, y_0) \neq 0$, то хотя бы одна из переменных x , y локально (в некоторой окрестности точки (x_0, y_0)) выражается через другую.

Действительно, либо $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$, и тогда действует теорема 23.1, либо $\frac{\partial F}{\partial x}(x_0, y_0) \neq 0$, и тогда x и y меняются ролями: x можно выразить через y (по той же теореме).

ОПРЕДЕЛЕНИЕ 23.3. Точка (x_0, y_0) кривой $F(x, y) = 0$ называется *неособой* точкой, если $F(x, y)$ в окрестности этой точки принадлежит классу C^1 и $\text{grad } F(x_0, y_0) \neq 0$.

Например, кривая $x^2 + y^2 - 1 = 0$ не имеет особых точек, поскольку $\text{grad } F(x, y) = (2x, 2y) \neq (0, 0)$ в точках кривой.

С другой стороны, кривая $x^2 - y^2 = 0$ имеет особую точку $(0, 0)$, поскольку $\text{grad } F(x, y) = (2x, -2y) = (0, 0)$ в этой точке.

Запишем теперь уравнение касательной к кривой $F(x, y) = 0$ в неособой точке (x_0, y_0) . Если, скажем, $\frac{\partial F}{\partial y} \neq 0$, то по теореме о неявной функции $f'(x_0) = -\frac{\partial F(x_0, y_0)/\partial x}{\partial F(x_0, y_0)/\partial y}$ и стандартное уравнение касательной плоскости $y = f(x_0) + f'(x_0)(x - x_0)$ приобретает вид

$$y = y_0 - \frac{\partial F(x_0, y_0)/\partial x}{\partial F(x_0, y_0)/\partial y}(x - x_0)$$

или, что то же самое,

$$\boxed{\frac{\partial F(x_0, y_0)}{\partial x}(x - x_0) + \frac{\partial F(x_0, y_0)}{\partial y}(y - y_0) = 0.}$$

Нетрудно проверить, что если выразить x через y (при условии $\partial F(x_0, y_0)/\partial \neq 0$), то получится то же самое уравнение. Таким образом, уравнение в рамке является уравнением касательной к кривой $F(x, y) = 0$ в неособой точке.

Из уравнения касательной видно, что вектор $\text{grad } F(x_0, y_0) = 0$ перпендикулярен к кривой $(x_0, y_0) = 0$ (по определению вектор перпендикулярен к кривой, если он перпендикулярен к касательной к этой кривой).

ОПРЕДЕЛЕНИЕ 23.4. Кривая вида $F(x, y) = C$ называется *линией уровня* функции $F(x, y)$ (отвечающей уровню C).

Рассмотрим теперь уравнение вида $F(x_1, \dots, x_n, x_{n+1}) = 0$, где $n > 1$. Это уравнение определяет некоторое множество в пространстве $\mathbb{R}^{n+1}$. При $n = 2$ это множество называют поверхностью в $n = 2$, а при $n > 2$ — гиперповерхностью.

ТЕОРЕМА 23.5 (2-я теорема о неявной функции). Пусть $F(x, y) = F(x_1, \dots, x_n, y)$ — функция класса C^1 в окрестности точки $(x^{(0)}, y_0) = (x_1^{(0)}, \dots, x_n^{(0)}, y_0)$, причем $F(x^{(0)}, y_0) = 0$. Если $\frac{\partial F(x^{(0)}, y_0)}{\partial y} \neq 0$, то существует окрестность точки $(x^{(0)}, y_0)$ вида $O_\delta(x^{(0)}) \times O_\delta(y_0)$, в которой уравнение $F(x, y) = 0$ эквивалентно уравнению $y = f(x)$, где функция $f(x)$ принадлежит классу C^1 в $O_\delta(x^{(0)})$, причем частные производные вычисляются при помощи формул

$$\frac{\partial f}{\partial x_j}(x^{(0)}, y_0) = - \frac{\frac{\partial F}{\partial x_j}(x^{(0)}, y_0)}{\frac{\partial F}{\partial y}(x^{(0)}, y_0)}, \quad j = 1, \dots, n.$$

Поясним, что множество $O_\delta(x^{(0)}) \times O_\delta(y_0)$ состоит из точек (x, y) , таких, что $x \in O_\delta(x^{(0)})$, а $y \in O_\delta(y_0)$.

НАБРОСОК ДОКАЗАТЕЛЬСТВА. Первая часть доказательства почти дословно повторяет соответствующую часть доказательства первой теоремы о неявной функции.

Вторая часть доказательства также похожа на вторую часть доказательства первой теоремы: если $y = f(x_1, \dots, x_n)$ — неявная функция, которая определяется уравнением $F(x_1, \dots, x_n, y) = 0$, то для любой точки $(x, y) = (x_0 + \Delta x, y_0 + \Delta y)$, удовлетворяющей уравнению $F(x, y) = 0$, имеем согласно формуле Тейлора

$$F(x^{(0)} + \Delta x, y_0 + \Delta y) = F(x^{(0)}, y_0) + \sum_{j=1}^n \frac{\partial F(x^{(0)} + \theta \Delta x, y_0 + \theta \Delta y)}{\partial x_j} \Delta x_j + \frac{\partial F(x^{(0)} + \theta \Delta x, y_0 + \theta \Delta y)}{\partial y} \Delta y,$$

или, с учетом того, что $F(x^{(0)} + \Delta x, y_0 + \Delta y) = F(x^{(0)}, y_0) = 0$,

$$\sum_{j=1}^n \frac{\partial F(x^{(0)} + \theta \Delta x, y_0 + \theta \Delta y)}{\partial x_j} \Delta x_j + \frac{\partial F(x^{(0)} + \theta \Delta x, y_0 + \theta \Delta y)}{\partial y} \Delta y = 0.$$

Из этой формулы имеем

$$\Delta y = - \sum_{j=1}^n \frac{\frac{\partial F(x^{(0)} + \theta \Delta x, y_0 + \theta \Delta y)}{\partial x_j}}{\frac{\partial F(x^{(0)} + \theta \Delta x, y_0 + \theta \Delta y)}{\partial y}} \Delta x_j.$$

Далее доказывается, что функции $\frac{\frac{\partial F(x^{(0)} + \theta \Delta x, y_0 + \theta \Delta y)}{\partial x_j}}{\frac{\partial F(x^{(0)} + \theta \Delta x, y_0 + \theta \Delta y)}{\partial y}}$ ограничены в окрестности точки $(x^{(0)}, y_0)$. Отсюда выводится непрерывность в этой точке: $\Delta y \rightarrow 0$ при $|\Delta x| \rightarrow 0$. Наконец, в силу непрерывности функций $\frac{\partial F(x, y)}{\partial x_j}$, $\frac{\partial F(x, y)}{\partial y}$ в точке $(x^{(0)}, y_0)$ имеем

$$\Delta y = \sum_{j=1}^n \left(- \frac{\frac{\partial F(x^{(0)}, y_0)}{\partial x_j}}{\frac{\partial F(x^{(0)}, y_0)}{\partial y}} + \alpha_j(x, y) \right) \Delta x_j,$$

где $\alpha_j(x, y) \rightarrow 0$ при $(x, y) \rightarrow (x^{(0)}, y_0)$. Это означает, что функция $y = f(x)$ дифференцируема в точке $x^{(0)}$ и справедлива формула для частных производных (см. формулировку теоремы).

СЛЕДСТВИЕ 23.6. Пусть $F(x) = F(x_1, \dots, x_n)$ — функция класса C^1 в окрестности точки $x^{(0)}$. Тогда если $\text{grad}(x^{(0)}) \neq 0$, то локально одну из переменных x_j , $j = 1, \dots, n$, можно выразить через остальные (посредством функции класса C^1).

Например, в $\mathbb{R}^3$ уравнение $F(x, y, z) = 0$ в окрестности точки (x_0, y_0, z_0) , такой, что $\text{grad} F(x_0, y_0, z_0) \neq 0$, можно записать явно в виде одного из уравнений $z = f(x, y)$, $y = f(x, z)$ или $x = f(y, z)$.

ОПРЕДЕЛЕНИЕ 23.7. Точка (x_0, y_0, z_0) поверхности $F(x, y, z) = 0$ называется *неособой* точкой, если в ее окрестности функция F принадлежит классу C^1 и $\text{grad} F(x_0, y_0, z_0) \neq 0$.

Запишем уравнение касательной плоскости к поверхности $F(x, y, z) = 0$ в окрестности неособой точки (x_0, y_0, z_0) . Пусть для определенности $\partial F(x_0, y_0, z_0)/\partial z \neq 0$. Тогда поверхность в окрестности этой точки можно записать явно в виде $z = f(x, y)$. При этом уравнение касательной плоскости имеет вид

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

Применяя формулы для частных производных функции, заданной неявно, получим уравнение

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

или, что то же,

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

Нетрудно проверить, что если выразить y или x , через остальные переменные, то мы все равно придем к такому же уравнению.

Из уравнения видно, что вектор с координатами $(\frac{F(x_0, y_0, z_0)}{\partial x}, \frac{F(x_0, y_0, z_0)}{\partial y}, \frac{F(x_0, y_0, z_0)}{\partial z})$, т.е. градиент функции f в точке (x_0, y_0, z_0) , является вектором нормали к касательной плоскости.

ПРИМЕР. Уравнение касательной плоскости к поверхности сферы $x^2 + y^2 + z^2 = 1$ имеет вид $2x_0(x - x_0) + 2y_0(y - y_0) + 2z_0(z - z_0) = 0$ (здесь $(2x_0, 2y_0, 2z_0)$ — компоненты градиента в точке (x_0, y_0, z_0)).

ОПРЕДЕЛЕНИЕ 23.8. Поверхность вида $F(x, y, z) = C$ называется поверхностью уровня функции $F(x, y, z)$ (отвечающей уровню C).

Из уравнения касательной плоскости видно, что вектор $\text{grad} F(x_0, y_0, z_0)$ перпендикулярен к поверхности уровня функции, проходящей через точку (x_0, y_0, z_0) .

Лекция 24

Приведем пример особой точки на поверхности. Уравнение $x^2 + y^2 - z^2 = 0$, как известно, задает конус в пространстве. Его вершина $(0, 0, 0)$ является особой точкой (нетрудно проверить, что градиент функции $x^2 + y^2 - z^2$ в этой точке равен 0). Ни в какой окрестности этой точки уравнение $x^2 + y^2 - z^2 = 0$ невозможно записать в виде $z = f(x, y)$, $y = f(x, z)$ или $x = f(y, z)$.

Функции, заданные неявно системой уравнений. Рассмотрим систему уравнений

[illegible]

Поскольку количество уравнений m меньше количества неизвестных $n + m$, интуитивно ясно, что «лишние» m неизвестных можно выразить через остальные.

ТЕОРЕМА 24.1 (3-я теорема о неявной функции (для систем уравнений)). Пусть $F_i(x, y)$, $i = 1, \dots, m$, — функции класса C^1 в окрестности точки $(x^{(0)}, y^{(0)})$ ($x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $y = (y_1, \dots, y_m) \in \mathbb{R}^m$).

Пусть $F_i(x^{(0)}, y^{(0)}) = 0$ для всех $i = 1, \dots, m$. Пусть выполнено условие невырожденности

$$\left| \begin{array}{ccc} \frac{\partial F_1(x^{(0)}, y^{(0)})}{\partial y_1} & \cdots & \frac{\partial F_1(x^{(0)}, y^{(0)})}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m(x^{(0)}, y^{(0)})}{\partial y_1} & \cdots & \frac{\partial F_m(x^{(0)}, y^{(0)})}{\partial y_m} \end{array} \right| \neq 0.$$

Тогда в некоторой окрестности точки $(x^{(0)}, y^{(0)})$ система уравнений

[illegible]

эквивалентна системе

[illegible]

где $f_i(x_1, \dots, x_n)$, $i = 1, \dots, t$, — функции класса C^1 в окрестности точки $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})$. (Это означает, что существует такая «прямоугольная» окрестность вида $O_\delta(x^{(0)}) \times O_\varepsilon(y^{(0)})$, в которой выписанные системы задают одно и то же множество точек.)

НАБРОСОК ДОКАЗАТЕЛЬСТВА. Пусть для простоты $m = 2$. Рассмотрим систему

$$\begin{cases} F_1(x, y_1, y_2) = 0, \\ F_2(x, y_1, y_2) = 0, \end{cases}$$

где $x \in \mathbb{R}^n$, точка $(x^{(0)}, y_1^{(0)}, y_2^{(0)})$ удовлетворяет этой системе, и выполнено условие невырожденности

$$\left| \frac{\frac{\partial F_1(x^{(0)}, y_1^{(0)}, y_2^{(0)})}{\partial y_1}}{\frac{\partial F_2(x^{(0)}, y_1^{(0)}, y_2^{(0)})}{\partial v_1}} - \frac{\frac{\partial F_1(x^{(0)}, y_1^{(0)}, y_2^{(0)})}{\partial y_2}}{\frac{\partial F_2(x^{(0)}, y_1^{(0)}, y_2^{(0)})}{\partial v_2}} \right| \neq 0.$$

Из условия невырожденности вытекает, что хотя бы одна из частных производных $\partial F/\partial y_1$, $\partial F/\partial y_2$ не равна нулю в точке $(x^{(0)}, y_1^{(0)}, y_2^{(0)})$ (иначе определитель имеет нулевую первую строку). Пусть для определенности $\partial F/\partial y_1 \neq 0$. Тогда по второй теореме о неявной функции из уравнения $F_1(x, y_1, y_2) = 0$ «можно выразить переменную y_1 »: это уравнение локально (в окрестности точки $(x^{(0)}, y_1^{(0)}, y_2^{(0)})$) эквивалентно уравнению $y_1 = \tilde{f}_1(x, y_2)$, где $\tilde{f}_1(x, y_2)$ — функция класса C^1 в окрестности точки $(x^{(0)}, y_2^{(0)})$, причем

$$\frac{\partial \tilde{f}_1}{\partial y_2} = - \frac{\frac{\partial F_1(x^{(0)}, y_1^{(0)}, y_2^{(0)})}{\partial y_2}}{\frac{\partial F_1(x^{(0)}, y_1^{(0)}, y_2^{(0)})}{\partial y_1}}$$

Подставим $y_1 = \tilde{f}_1(x, y_2)$ во второе уравнение:

$$F_2(x, \tilde{f}_1(x, y_2), y_2) = 0.$$

Здесь функция слева принадлежит C^1 в окрестности точки $(x^{(0)}, y_2^{(0)})$. По второй теореме о неявной функции это уравнение в окрестности этой точки можно разрешить относительно переменной y_2 , если частная

производная по этой переменной в точке $(x^{(0)}, y_2^{(0)})$ не равна 0. Убедимся, что это условие выполнено:

$$\frac{\partial}{\partial y_2} F_2(x, \tilde{f}_1(x, y_2), y_2) = \frac{\partial F_2}{\partial y_1} \frac{\partial \tilde{f}_1}{\partial y_2} + \frac{\partial F_2}{\partial y_2} = -\frac{\partial F_2}{\partial y_1} \frac{\frac{\partial F_1}{\partial y_2}}{\frac{\partial F_1}{\partial y_1}} + \frac{\partial F_2}{\partial y_2} = \frac{1}{\frac{\partial F_1}{\partial y_1}} \left| \frac{\partial F_1}{\partial y_1} \frac{\partial F_1}{\partial y_2} \right|$$

Здесь правая часть не обращается в 0 в точке $(x^{(0)}, y_1^{(0)}, y_2^{(0)})$ в силу условия невырожденности.

Итак, локально $y_2 = f_2(x)$, где $f_2(x)$ — функция класса C^1 в окрестности точки $x^{(0)}$. Наконец, $y_1 = \tilde{f}_1(x, y_2) = \tilde{f}_1(x, f_2(x))$. Это также функция от x класса C^1 в окрестности точки $x^{(0)}$. Обозначим ее через $f_1(x)$. $\square$

ЗАМЕЧАНИЕ 24.2. Пусть точка $x^{(0)} = (x_1^{(0)}, \dots, x_{n+m}^{(0)})$ удовлетворяет системе уравнений

$$\begin{cases} F_1(x_1, \dots, x_{n+m}) = 0, \\ \vdots \\ F_m(x_1, \dots, x_{n+m}) = 0 \end{cases}$$

и, кроме того, функции F_j , $j = 1, \dots, m$, принадлежат C^1 в окрестности этой точки. При каких условиях (локально) какие-нибудь m переменных можно выразить через остальные n ?

Ответ: в случае когда матрица Якоби

$$\frac{DF(x^{(0)})}{Dx} = \begin{pmatrix} \frac{\partial F_1(x^{(0)})}{\partial x_1} & \cdots & \frac{\partial F_1(x^{(0)})}{\partial x_{n+m}} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m(x^{(0)})}{\partial x_1} & \cdots & \frac{\partial F_m(x^{(0)})}{\partial x_{n+m}} \end{pmatrix}$$

невырожденна, т. е. в случае когда ее ранг (максимальное количество линейно независимых столбцов или строк) равен m : $\text{rank} \frac{DF(x^{(0)})}{Dx} = m$. При этом можно выразить те переменные, которым отвечают линейно независимые столбцы.

Это замечание объясняет разумность следующего определения:

ОПРЕДЕЛЕНИЕ 24.3. Точка $x^{(0)}$ называется *неособой* точкой гиперповерхности, заданной системой уравнений

$$\begin{cases} F_1(x_1, \dots, x_{n+m}) = 0, \\ \vdots \\ F_m(x_1, \dots, x_{n+m}) = 0, \end{cases}$$

если матрица Якоби $\frac{DF(x^{(0)})}{Dx}$ невырождена.

ПРИМЕР. Рассмотрим систему уравнений

$$\begin{cases} x^2 + y^2 + z^2 = 1, \\ x + 2y + 3z = \frac{6}{\sqrt{3}} \end{cases}$$

в окрестности точки $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$. Проверим можно ли выразить переменные y и z через x в окрестности этой точки. Определитель

$$\left. \frac{Df}{D(y, z)} \right|_{(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})} = \left| \begin{array}{cc} 2y & 2z \\ 2 & 3 \end{array} \right|_{(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})} = \frac{2}{\sqrt{3}}$$

отличен от 0. Поэтому ответ на вопрос утвердительный.

ЗАМЕЧАНИЕ. В формулировке теоремы 24.1 нет формул для вычисления частных производных.

Для вычисления частных производных $\partial y_i / \partial x_j$, $i = 1, \dots, m$, запишем равенства

[illegible]

и продифференцируем их по x_j :

$$\begin{cases} \frac{\partial f_1}{\partial x_j} + \frac{\partial f_1}{\partial y_1} \frac{\partial y_1}{\partial x_j} + \cdots + \frac{\partial f_1}{\partial y_m} \frac{\partial y_m}{\partial x_j} = 0, \\ \dots\dots\dots \\ \frac{\partial f_m}{\partial x_j} + \frac{\partial f_m}{\partial y_1} \frac{\partial y_1}{\partial x_j} + \cdots + \frac{\partial f_m}{\partial y_m} \frac{\partial y_m}{\partial x_j} = 0. \end{cases}$$

$\begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_m} \end{pmatrix}$ и столбцом правых частей $\begin{pmatrix} -\frac{\partial f_1}{\partial x_j} \\ \vdots \\ -\frac{\partial f_m}{\partial x_j} \end{pmatrix}$. В силу условия невырожденности (см. теорему 24.1) определитель матрицы системы отличен от нуля. Поэтому она имеет единственное решение.

ПРИМЕР. Рассмотрим систему уравнений $\begin{cases} xy + 2yz + 3xz = 6, \\ 4xy - yz + 2xz = 5. \end{cases}$ В окрестности точки $(1, 1, 1)$ она определяет две функции $y(x)$, $z(x)$. Проверим это, вычислив определитель соответствующей матрицы:

Далее, продифференцируем по x соотношения

Получим

Нас интересует точка $(1, 1, 1)$. Поэтому, подставляя $x = y = z = 1$, получаем

Лекция 25

$$\begin{cases} y_1 = f_1(x_1, \dots, x_n), \\ \dots\dots\dots \\ y_m = f_m(x_1, \dots, x_n), \end{cases}$$
$$\Delta y_i = \sum_{j=1}^n \frac{\partial f_i(x^{(0)})}{\partial x_j} \Delta x_j + o(\Delta x), \quad \Delta x \rightarrow 0,$$
$$\begin{pmatrix} \Delta y_1 \\ \vdots \\ \Delta y_m \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1(x^{(0)})}{\partial x_1} & \cdots & \frac{\partial f_1(x^{(0)})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(x^{(0)})}{\partial x_1} & \cdots & \frac{\partial f_m(x^{(0)})}{\partial x_n} \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \vdots \\ \Delta x_n \end{pmatrix} + \begin{pmatrix} o(\Delta x) \\ \vdots \\ o(\Delta x) \end{pmatrix}, \quad \Delta x \rightarrow 0,$$

68

$$A = \begin{pmatrix} \frac{\partial f_1(x^{(0)})}{\partial x_1} & \cdots & \frac{\partial f_1(x^{(0)})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(x^{(0)})}{\partial x_1} & \cdots & \frac{\partial f_m(x^{(0)})}{\partial x_n} \end{pmatrix}$$

называется *матрицей Якоби* отображения f в точке $x^{(0)}$.

Геометрический смысл дифференцируемости отображения таков: приращение Δy с точностью до $o(\Delta x)$ задается линейным оператором с матрицей Якоби, примененным к приращению аргумента Δx .

ТЕОРЕМА 25.1 (о дифференцируемой сложной функции). Пусть отображение $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ дифференцируемо в точке $x^{(0)}$ и имеет в этой точке матрицу Якоби A . Пусть отображение $g: \mathbb{R}^m \rightarrow \mathbb{R}^k$ дифференцируемо в точке $y^{(0)} = f(x^{(0)})$ и имеет в этой точке матрицу Якоби B . Тогда сквозное отображение $g(f(x))$ из $\mathbb{R}^n$ в $\mathbb{R}^k$ дифференцируемо в точке $x^{(0)}$ и имеет в этой точке матрицу Якоби $C = BA$.

В самом деле, если $g_i(y_1, \dots, y_m)$ и $y_s = y_s(x_1, \dots, x_n)$, то по формуле дифференцирования сложной функции

$$\frac{\partial g_i}{\partial x_j} = \sum_{s=1}^m \frac{\partial g_i}{\partial y_s} \frac{\partial y_s}{\partial x_j}.$$

Эта формула совпадает с формулой для элементов произведения соответствующих матриц Якоби B и A . $\square$

ТЕОРЕМА 25.2 (об обратном отображении). Пусть отображение $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ принадлежит классу C^1 в окрестности точки $x^{(0)}$. Тогда если определитель матрицы Якоби (Якобиан) в этой точке не равен 0, то отображение f локально (в окрестности этой точки) обратимо, т. е. существует такое отображение $g = f^{-1}$, что $f^{-1}(f(x)) = x$ для всех x из некоторой окрестности $O(x^{(0)})$, и $f(f^{-1}(y)) = y$ для всех y из некоторой окрестности $O(f(x^{(0)}))$.

НАБРОСОК ДОКАЗАТЕЛЬСТВА. Запишем систему уравнений

[illegible]

и разрешим ее локально относительно переменных $x_1, \dots, x_n$. Это возможно в силу 3-й теоремы о неявной функции. Все ее условия выполнены и, в частности, выполнено условие невырожденности:

$$\det \begin{pmatrix} \frac{\partial f_1(x^{(0)})}{\partial x_1} & \cdots & \frac{\partial f_1(x^{(0)})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(x^{(0)})}{\partial x_1} & \cdots & \frac{\partial f_n(x^{(0)})}{\partial x_n} \end{pmatrix} \neq 0.$$

Из формулы $C = BA$ для матрицы Якоби композиции двух отображений видно, что если $f^{-1}(y)$ — отображение, обратное к отображению $f(x)$ с матрицей Якоби A , то его матрица Якоби совпадает с матрицей A^{-1} . В самом деле, по определению обратного отображения композиция обратного и прямого отображений является тождественным отображением (т.е., отображением, сопоставляющим точке x саму эту точку). Матрица Якоби такого отображения, как нетрудно проверить, является единичной матрицей E . Таким образом, в силу теоремы о композиции матрица Якоби B обратного отображения f^{-1} должна удовлетворять соотношениям $BA = E$ и $AB = E$. Таким образом, $B = A^{-1}$.

ПРИМЕР 25.3. Рассмотрим отображение перехода к полярным координатам. Оно действует из $\mathbb{R}^2$ в $\mathbb{R}^2$ посредством формул

$$\begin{cases} x = \rho \cos \varphi, \\ y = \rho \sin \varphi \end{cases}$$

с областью определения $M = \{(\rho, \varphi) : \rho \geq 0\}$. Вычислим матрицу Якоби $\frac{D(x, y)}{D(\rho, \varphi)}$ и ее определитель:

$$\frac{D(x, y)}{D(\rho, \varphi)} = \begin{pmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\rho \sin \varphi \\ \sin \varphi & \rho \cos \varphi \end{pmatrix},$$

$$\left| \frac{D(x, y)}{D(\rho, \varphi)} \right| = \begin{vmatrix} \cos \varphi & -\rho \sin \varphi \\ \sin \varphi & \rho \cos \varphi \end{vmatrix} = \rho.$$

Таким образом, при $\rho > 0$ есть локальная обратимость (при $\rho = 0$ локальной обратимости нет, поскольку все точки с $\rho = 0$ попадают под действием этого отображения в одну точку $(0, 0)$).

1) либо являются особыми для поверхности S , т. е.

$$\text{rank} \begin{pmatrix} \frac{\partial f_1(x^{(0)})}{\partial x_1} & \dots & \frac{\partial f_1(x^{(0)})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(x^{(0)})}{\partial x_1} & \dots & \frac{\partial f_m(x^{(0)})}{\partial x_n} \end{pmatrix} < m,$$

2) либо координаты точка $(x^{(0)}, \lambda^{(0)})$ при некоторых $\lambda_1^{(0)}, \dots, \lambda_n^{(0)}$ является стационарной точкой функции Лагранжа.

Выпишем условие $dL(x^{(0)}, \lambda^{(0)}) = 0$, приравняв все частные производные этой функции нулю:

[illegible]

Видно, что последние m уравнений просто совпадают с уравнениями связи.

Отметим, что если $x^{(0)}$ не является особой точкой поверхности, то решение системы уравнений относительно $\lambda_1, \dots, \lambda_m$ единственно. Это вытекает из того, что матрица этой системы невырождена.

НАБРОСОК ДОКАЗАТЕЛЬСТВА. Рассмотрим для простоты случай $m = 1$ (одно уравнение связи):

$$\begin{cases} F(x_1, \dots, x_n) \rightarrow \text{extr}, \\ f(x_1, \dots, x_n) = 0 \end{cases}$$

Согласно условиям теоремы функции F и f принадлежат C^1 в окрестности точки $x^{(0)}$ и $\text{grad } f(x^{(0)}) \neq 0$. По теореме о неявной функции одну из переменных в условии связи (будем считать для определенности, что это x_n) можно (локально) выразить через остальные: $x_n = \varphi(x_1, \dots, x_{n-1})$. Поэтому поиск условного экстремума сводится поиску безусловного экстремума функции $F(x_1, \dots, x_{n-1}, \varphi(x_1, \dots, x_{n-1}))$.

Выпишем необходимое условие безусловного экстремума:

$$\begin{cases} \frac{\partial F}{\partial x_1} + \frac{\partial F}{\partial x_n} \frac{\partial \varphi}{\partial x_1} = 0, \\ \dots\dots\dots \\ \frac{\partial F}{\partial x_{n-1}} + \frac{\partial F}{\partial x_n} \frac{\partial \varphi}{\partial x_{n-1}} = 0. \end{cases}$$

Отсюда, пользуясь формулой для частных производных функции φ , заданной неявно, получим

$$\left\{ \begin{array}{l} \frac{\partial F}{\partial x_1} + \frac{\partial F}{\partial x_n} \left(-\frac{\frac{\partial f}{\partial x_1}}{\frac{\partial f}{\partial x_n}} \right) = 0, \\ \dots\dots\dots \\ \frac{\partial F}{\partial x_{n-1}} + \frac{\partial F}{\partial x_n} \left(-\frac{\frac{\partial f}{\partial x_{n-1}}}{\frac{\partial f}{\partial x_n}} \right) = 0. \end{array} \right.$$

Введем обозначение $\lambda_0 = -\frac{\partial F}{\partial x_n} / \frac{\partial f}{\partial x_n}$. Теперь наша система переписывается в виде

$$\begin{cases} \frac{\partial F}{\partial x_1} + \lambda_0 \frac{\partial f}{\partial x_1} = 0, \\ \dots\dots\dots \\ \frac{\partial F}{\partial x_{n-1}} + \lambda_0 \frac{\partial f}{\partial x_{n-1}} = 0. \end{cases}$$

К этой системе можно добавить равенство $\frac{\partial F}{\partial x_n} + \lambda_0 \frac{\partial f}{\partial x_n} = 0$ (вытекающее из самого обозначения). Итак, мы доказали существование λ_0 , такого, что $(x^{(0)}, \lambda_0)$ — стационарная точка функции Лагранжа. $\square$

