Лекции[*] по математическому анализу
ФН, I курс, 2 семестр
2 модуль „Функции нескольких переменных"

Пугачёв О.В.

весна 2017

1 Множества и функции в $\mathbb{R}^n$

Точками пространства $\mathbb{R}^n$ являются наборы n действительных чисел, называемых координатами точки. Расстояние между точками $A = (a_1, \ldots, a_n)$ и $B = (b_1, \ldots, b_n)$ определяется формулой

$$\varrho(A, B) = |\vec{AB}| = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2}$$

Для $\mathbb{R}^n$ и ϱ выполнены все аксиомы метрического пространства:

1. Положительность: $\varrho(A, B) \geq 0$; $\varrho(A, B) = 0 \Longleftrightarrow A = B$;

2. Симметричность: $\varrho(A, B) = \varrho(B, A)$;

3. Неравенство треугольника: $\varrho(A, C) \leq \varrho(A, B) + \varrho(B, C)$.
Расстоянием от точки A до множества S называется величина $\varrho(A, S) = \inf\limits_{B \in S} \varrho(A, B)$.

Определение 1.1. *Последовательность точек* $\{A_k\}_{k=1}^{\infty}$ *в пространстве* $\mathbb{R}^n$ **сходится** *к точке* B, *если* $\varrho(A_k, B) \underset{k\to\infty}{\longrightarrow} 0$. *Обозначается* $A_k \underset{k\to\infty}{\longrightarrow} B$ *или* $\lim\limits_{k\to\infty} A_k = B$.

Лемма 1.1. *Сходимость последовательности точек* $A_k = (a_{1k}, \ldots, a_{nk})$ *к точке* $B = (b_1, \ldots, b_n)$ *равносильна покоординатной сходимости:*

$$A_k \underset{k\to\infty}{\longrightarrow} B \quad \Longleftrightarrow \quad a_{ik} \underset{k\to\infty}{\longrightarrow} b_i \quad \forall\ i = 1, \ldots, n.$$

Доказательство. ($\Rightarrow$): Если $\forall\varepsilon > 0\ \exists M$: $\forall k > M\ \ \varrho(A_k, B) < \varepsilon$, то для каждой координаты получим $|a_{ik} - b_i| < \varepsilon$.
($\Leftarrow$): Найдём такие M_1,...,M_n, что $\forall k > M_i\ |a_{ik} - b_i| < \varepsilon/\sqrt{n}$. Тогда при всех $k > \max\{M_1, \ldots, M_n\}$ получим $\varrho(A_k, B) < \sqrt{n(\varepsilon/\sqrt{n})^2} = \varepsilon$. $\square$

Следствие 1.1. *Если* $\exists \lim\limits_{k\to\infty} A_k$, *то он единственный.*

Определение 1.2. ε**-окрестностью** *точки* $A \in \mathbb{R}^n$ ($\varepsilon > 0$) *называется множество* $U_\varepsilon(A) = \{P \in \mathbb{R}^n :\ \varrho(A, P) < \varepsilon\}$ (*шар радиуса* ε *с центром* A). *Проколотой окрестностью называется* $\overset{o}{U}_\varepsilon(A) = U_\varepsilon(A) \backslash \{A\}$.

Открытые и замкнутые множества

Определение 1.3. *Пусть* $S \subset \mathbb{R}^n$. *Точка* A *называется* **внутренней** *точкой множества* S, *если* $\exists\varepsilon > 0$: $U_\varepsilon(A) \subset S$.
Точка B *называется* **граничной** *точкой множества* S, *если* $\forall\varepsilon > 0$ *одновременно* $U_\varepsilon(B) \cap S \neq \emptyset$ *и* $U_\varepsilon(B) \cap (\mathbb{R}^n \backslash S) \neq \emptyset$.

Внутренние точки принадлежат множеству S, а граничная точка может либо принадлежать S, либо нет. Множество граничных точек S называется его *границей* и обозначается ∂S.

Определение 1.4. *Множество $G \subset \mathbb{R}^n$ называется* **открытым***, если все его точки внутренние, т. е. граничных точек оно не содержит: $\partial G \subset \mathbb{R}^n \backslash G$.*

Множество $F \subset \mathbb{R}^n$ называется **замкнутым***, если оно содержит все свои граничные точки: $\partial F \subset F$.*

Замыканием *множества $S \subset \mathbb{R}^n$ называется множество $[S] = S \cup \partial S$.*

Замечание 1.1. 1) Поскольку граница у множеств S и $\mathbb{R}^n \backslash S$ одна и та же, из определения следует, что дополнение открытого множества замкнуто и наоборот. 2) Множество S будет одновременно открытым и замкнутым, лишь если у него нет граничных точек, т.е. если $S = \emptyset$ или $S = \mathbb{R}^n$.

Лемма 1.2. *Если последовательность точек $A_k \in S \subset \mathbb{R}^n$ сходится к точке B, то $B \in [S]$.*

Доказательство. Очевидно, $\varrho(B, S) = 0$. Если при этом $\varrho(B, \mathbb{R}^n \backslash S) = 0$, то точка B — граничная для S, если же $\varrho(B, \mathbb{R}^n \backslash S) > 0$, то внутренняя. $\square$

Пример 1.1. Рассмотрим несколько множеств на плоскости $\mathbb{R}^2$.
1) $U = \{x^2 + y^2 < 1\}$ — открытое. Его граница $\partial U = \{x^2 + y^2 = 1\}$; его замыкание $[U] = \{x^2 + y^2 \leq 1\}$.
2) $\Pi = \{x \geq 0\}$ — замкнутое; его граница $\partial \Pi = \{x = 0\}$.
3) $\ell = \{y = x^2\}$ — замкнутое; его граница $\partial \ell = \ell$ (внутренних точек нет).
4) $\mathbb{Q}^2 = \{x, y \in \mathbb{Q}\}$ — не открытое и не замкнутое. Граница $\partial \mathbb{Q}^2 = \mathbb{R}^2$.

Лемма 1.3. 1) *Объединение любого набора открытых множеств и пересечение конечного набора открытых множеств являются открытыми.*
2) *Объединение конечного набора замкнутых множеств и пересечение любого набора замкнутых множеств являются замкнутыми.*

Доказательство. 1) Если $P \in \bigcup G_i$, то $\exists G_k \ni P \Longrightarrow \exists \varepsilon > 0$: $U_\varepsilon(P) \subset G_k \subset \bigcup G_i$. Если же $P \in G_1 \cap \ldots \cap G_m$, то при каждом $i = 1, \ldots, m$ найдём ε_i: $U_{\varepsilon_i}(P) \subset G_i$; взяв $\varepsilon = \min\{\varepsilon_1, \ldots, \varepsilon_m\}$, получаем $U_\varepsilon(P) \subset G_1 \cap \ldots \cap G_m$.
2) Следует из 1) в силу замечания 1.1 и известных соотношений из теории множеств (множества G_i берутся в любом количестве):

$$\bigcup (X \backslash G_i) = X \backslash \bigcap G_i; \quad \bigcap (X \backslash G_i) = X \backslash \bigcup G_i.$$

$\square$

Определение 1.5. *Множество $K \subset \mathbb{R}^n$ называется* **компактом***, если из любой последовательности точек $A_j \in K$ можно выбрать подпоследовательность A_{j_k} $(j_1 < j_2 < \ldots)$, сходящуюся к точке $B \in K$.*

Теорема 1.1. *Множество $K \subset \mathbb{R}^n$ компактно $\Longleftrightarrow$ оно замкнуто и ограничено.*

Доказательство. ($\Rightarrow$): Если бы K не было ограниченным, то нашлась бы последовательность точек, уходящих в ∞, а из неё нельзя выбрать сходящуюся подпоследовательность. Если бы K не было замкнутым, то нашлась бы граничная точка $C \notin K$, но $\forall j \in \mathbb{N}\ \exists A_j \in K$: $\varrho(A_j, C) < 1/j \Longrightarrow A_j \to C \Longrightarrow$ любая подпоследовательность этой последовательности $\{A_j\}$ сходится только к $C \notin K$ (другого предела нет по следствию 1.1).
($\Leftarrow$): Пусть K замкнуто и ограничено; $\{A_j = (a_{1j}, \ldots, a_{nj})\}_{j=1}^{\infty}$ – последовательность точек K. Последовательности координат $\{a_{ij}\}_{j=1}^{\infty}$, $i = 1, \ldots, n$ ограничены $\Longrightarrow$ можно выбрать подпоследовательность номеров $j_1 < j_2 < \ldots$ так, чтобы a_{1j_k} сходились к некоторому $b_1 \in \mathbb{R}$; затем из этой подпоследовательности выберем новую подпоследовательность так, чтобы a_{2j_k} сходились к некоторому $b_2 \in \mathbb{R}$, и т. д. Проделав эту процедуру n раз (на каждом шаге будет получаться бесконечная подпоследовательность), мы получим подпоследовательность точек, покоординатно сходящихся к $B = (b_1, b_2, \ldots) \in \mathbb{R}^n$. Поскольку K замкнуто, то по лемме 1.2 получаем $B \in K$. $\square$

Скалярные и векторные ФНП

Скалярной функцией n переменных называется отображение $f\colon D(f) \mapsto \mathbb{R}$, где множество $D(f) \subset \mathbb{R}^n$ называется *областью определения* функции f. *Поверхностями уровня* (в случае $n = 2$ *линиями уровня*) называются множества $\{P \in D(f) :\ f(P) = C\}$, где $C \in \mathbb{R}$ — константы.

Пример 1.2. Пусть $n = 2$; $f(x, y) = y/x$. Область определения $D(f) = \{(x; y) : x \neq 0\}$. Область значений $= \mathbb{R}$. Линия уровня C — прямая $y = Cx$, за исключением точки $(0; 0)$.

Пример 1.3. Пусть $n = 3$; $f(x, y, z) = x^2 + y^2 - z^2$. Область определения $D(f) = \mathbb{R}^3$. Область значений $= \mathbb{R}$. Поверхности уровня:
при $C < 0$ — двуполостные гиперболоиды,
при $C = 0$ — конус,
при $C > 0$ — однополостные гиперболоиды.

Векторной (k-мерной) функцией n переменных называется отображение $\vec{V}\colon D(\vec{V}) \mapsto \mathbb{R}^k$, где множество $D(\vec{V}) \subset \mathbb{R}^n$ называется *областью определения* функции $\vec{V}$. В случае $k = n$ отображение $\vec{V}$ иногда называют *векторным полем*, так как его можно представить в виде совокупности векторов $\vec{V}(P)$, торчащих из соотвествующих точек P.

Если $\vec{V}(P) = (v_1(P), \ldots, v_k(P))$, то скалярные функции $v_i(P)$, $i = 1,\ldots,k$, называются *координатными функциями* векторной функции $\vec{V}$. *Множествами уровня* называются множества вида $\{P \in D(\vec{V}) :\ \vec{V}(P) = (C_1, \ldots, C_k)\}$, где $C_i \in \mathbb{R}$ — константы. Очевидно, такое множество является пересечением поверхностей уровня координатных функций $\{P \in D(\vec{V}) :\ v_i(P) = C_i\}$, $i = 1, \ldots, k$.

2 Непрерывность ФНП

Известны два разных определения предела ФНП в точке.

Определение 2.1. (Коши) *Пусть* $f(x_1, \dots, x_n)$ *— скалярная или векторная ФНП; точка* $A \in [D(f)]$. *Говорят, что предел функции* f *в точке* A *равен* y, *если*

$$\forall \varepsilon > 0\ \exists \delta > 0: \quad \forall P \in \overset{o}{U}_\delta(A) \cap D(f) \qquad |f(P) - y| < \varepsilon.$$

Определение 2.2. (Гейне) *... предел функции* f *в точке* A *равен* y, *если для любой последовательности точек* $A_k \in D(f)$, $A_k \neq A$, *сходящейся к* A, *имеем* $f(A_k) \to y$.

Обозначение: $\lim\limits_{P \to A} f(P) = y$ или $f(P) \underset{P \to A}{\longrightarrow} y$. Докажем равносильность двух определений.

Доказательство. (Коши$\Rightarrow$Гейне): Если $D(f) \ni A_k \to A$, $A_k \neq A$, то $\forall \varepsilon > 0\ \exists N$: при всех $k > N$ $A_k \in U_\delta(A) \Longrightarrow A_k \in \overset{o}{U}_\delta(A) \cap D(f) \Longrightarrow |f(A_k) - y| < \varepsilon$.
(Гейне$\Rightarrow$Коши): Предположим, что нет сходимости по Коши. Тогда $\exists \varepsilon > 0$: $\forall \delta > 0$, в частности, для $\delta = 1/k$, $\exists A_k \in \overset{o}{U}_\delta(A) \cap D(f)$: $|f(P) - y| \geq \varepsilon$. Мы видим, что $A_k \to A$, но $f(A_k) \nrightarrow y$. □

Свойства пределов ФНП

1. Если существует предел $\lim\limits_{P \to A} f(P)$, то он единственный.

2. Сходимость векторной ФНП $\lim\limits_{P \to A} \vec{V}(P) = Y = (y_1, \dots, y_k)$ равносильна сходимости её координатных функций $\lim\limits_{P \to A} v_i(P) = y_i$, $\forall\ i = 1, \dots, k$ (вытекает из леммы 1.1, применённой к последовательности $f(A_k)$, см. определение 2.2).

3. Для скалярных ФНП: если $\lim\limits_{P \to A} f(P) = y$, $\lim\limits_{P \to A} g(P) = z$, то существуют пределы $\lim\limits_{P \to A} (f(P) + g(P)) = y + z$, $\lim\limits_{P \to A} f(P)g(P) = yz$, а при $z \neq 0$ — также $\lim\limits_{P \to A} \dfrac{f(P)}{g(P)} = \dfrac{y}{z}$.

4. Для k-мерных векторных ФНП: если $\lim\limits_{P \to A} \vec{V}(P) = Y$, $\lim\limits_{P \to A} \vec{W}(P) = Z$, то существуют пределы линейной комбинации $\lim\limits_{P \to A} (c\vec{V}(P) + k\vec{W}(P)) = cY + kZ$; скалярного произведения $\lim\limits_{P \to A} (\vec{V}(P), \vec{W}(P)) = (Y, Z)$; при $k = 3$ — также векторного произведения.

Определение 2.3. *ФНП* **непрерывна** *в точке* $A \in D(f)$, *если* $\lim\limits_{P \to A} f(P) = f(A)$.

Из свойств пределов ФНП вытекает:

I. Непрерывность векторной ФНП равносильна непрерывности её координатных функций.

II. Линейные комбинации, произведения и частные (если знаменатель не обнуляется в A) ФНП, непрерывных в точке A, также непрерывны в A.

Ещё одно свойство: **Предел сложной функции**.

Лемма 2.1. *Если $\exists$ предел $y = \lim\limits_{P \to A} f(P)$; y — внутренняя точка $D(\Phi) \subset \mathbb{R}^m$, и если функция Φ непрерывна в точке y, то существует предел $\lim\limits_{P \to A} \Phi(f(P)) = \Phi(y)$.*

Доказательство. Если $A_k \in D(f)$, $A_k \neq A$, $A_k \to A$, то $f(A_k) \to y \Longrightarrow \Phi(f(A_k)) \to \Phi(y)$. $\square$

Следствие 2.1. *Если функция f непрерывна в точке P, $y = f(P)$ — внутренняя точка $D(\Phi) \subset \mathbb{R}^m$, и функция Φ непрерывна в точке y, то функция $\Phi \circ f$ непрерывна в точке P.*

Определение 2.4. *Будем называть скалярную ФНП $f(x_1, \ldots, x_n)$* **элементарной***, если её значения получаются из $x_1, \ldots, x_n$ и констант за конечное число шагов через арифметические действия и функции $\sqrt[k]{x}$, e^x, $\ln x$, $\sin x$ и $\operatorname{arctg} x$. Векторная ФНП элементарная, если элементарны все её координатные функции.*

Элементарная функция f непрерывна во всех точках $D(f)$ (доказывается индукцией по числу шагов её построения).

Если ФНП непрерывна в точке A, то она в ней непрерывна по любой координате при фиксированных остальных координатах. Однако, обратное неверно, как показывает следующий

Пример 2.1. *В каждой точке $(x; y) \in \mathbb{R}^2$ функция*

$$f(x; y) = \begin{cases} \dfrac{2xy}{x^2 + y^2} & \text{при} \quad (x; y) \neq (0; 0) \\ 0 & \text{при} \quad x = y = 0 \end{cases}$$

непрерына по x при фиксированном y и непрерывна по y при фиксированном x. Но она разрывна в точке $O = (0; 0)$.

Доказательство. Рассмотрим элементарную функцию $f_o(x; y) = \dfrac{2xy}{x^2 + y^2}$, её область определения $D(f_o) = \mathbb{R}^2 \backslash \{O\}$, на которой она совпадает с f. Она непрерывна в каждой точке $(x; y) \neq O \Longrightarrow f$ также непрерывна в $(x; y) \neq O$, поскольку в определении предела по Коши всегда можно выбирать $\delta < \varrho((x; y), O)$.

В точке O функция f непрерывна по x при фиксированном $y = 0$ и по y при фиксированном $x = 0$. Но как ФНП она разрывна в O, и разрыв неустраним: можно построить две последовательности точек, сходящиеся к O, но имеющие разные пределы:

$$A_k = (1/k; 1/k) \to O, \quad f(A_k) = 1; \qquad B_k = (1/k; 0) \to O, \quad f(A_k) = 0.$$

$\square$

Отображения множеств

Определение 2.5. *Пусть $D(f) \subset \mathbb{R}^n$, функция f принимает значения в $\mathbb{R}^m$.*
1) **Образом** *множества $S \subset D(f)$ под действием f называется множество $f(S) = \{f(A): \ A \in S\} \subset \mathbb{R}^m$.*
2) **Прообразом** *множества $T \subset \mathbb{R}^m$ называется множество $f^{-1}(T) = \{A \in D(f): \ f(A) \in T\}$.*

Определение 2.6. *Множество $S \subset \mathbb{R}^n$ называется* **свя́зным***, если для любых двух точек A, $B \in S$ найдётся кривая $\ell \subset S : A, B \in \ell$.*

В одномерном случае S связно $\iff S$ — промежуток.

Теорема 2.1. *Если множество $S \subset \mathbb{R}^n$ связно и функция $f : S \mapsto \mathbb{R}^m$ непрерывна в каждой точке S, то образ $f(S)$ — связное множество в $\mathbb{R}^m$. В частности, для скалярной $f \quad f(S)$ — промежуток.*

Доказательство. Пусть X, $Y \in f(S)$. Тогда $X = f(A)$, $Y = f(B)$, A, $B \in S$. Соединим A и B кривой $\ell \subset S$, пусть $\ell = \{P(t): \ 0 \leq t \leq 1\}$, $P(t)$ непрерывна на $[0;1]$; $P(0) = A$, $P(1) = B$. Тогда функция $Q(t) = f \circ P(t)$ тоже непрерывна, кривая $f(\ell) = \{Q(t): \ 0 \leq t \leq 1\}$ лежит в $f(S)$ и содержит точки X и Y. $\square$

Следствие 2.2. *Если скалярная функция непрерывна на связном множестве S и не обращается на нём в 0, то она сохраняет знак на S.*

Теорема 2.2. *Если множество $K \subset \mathbb{R}^n$ компактно и функция $f : K \mapsto \mathbb{R}^m$ непрерывна в каждой точке K, то $f(K)$ — компакт в $\mathbb{R}^m$ (замкнуто и ограничено). В частности, скалярная непрерывная f ограничена на K и достигает на нём минимум и максимум.*

Доказательство. Пусть $\{B_k = f(A_k)\}$ — последовательность точек в $f(K)$. Из последовательности $A_k \in K$ выберем подпоследовательность A_{k_j}, сходящуюся к точке $A_o \in K$. Тогда $B_{k_j} \to f(A_o) \in f(K)$ при $j \to \infty$. $\square$

Теорема 2.3. *Если $D(f) = \mathbb{R}^n$, $f : \ \mathbb{R}^n \mapsto \mathbb{R}^m$ непрерывна в каждой точке, то прообразы открытых множеств — открытые, а прообразы замкнутых – замкнутые.*

Доказательство. Из непрерывности по Коши следует, что если Y – внутренняя точка множества $G \subset \mathbb{R}^m$ ($\exists \varepsilon > 0$: $U_\varepsilon(Y) \subset G$), и если $A \in f^{-1}(Y)$, то $\exists \delta > 0$: $U_\delta(A) \subset f^{-1}(G)$. Значит, если G открыто, то $f^{-1}(G) \subset \mathbb{R}^n$ также открыто. Для замкнутых множеств вспомним замечание 1.1 и то, что прообраз дополнения = дополнение прообраза. $\square$

Примеры: $\{(x;y): \ x^2 - y^2 > 1\}$ — открытое (прообраз $(1;+\infty)$);
$\{(x;y): \ x^2 - y^2 = 1\}$ — замкнутое (прообраз одноточечного множества).

3 Дифференцирование скалярных ФНП

Пусть $D(f) \subset \mathbb{R}^n$; $A = (x_1, \ldots, x_n)$ — внутренняя точка $D(f)$. В отличие от функции одного переменного, теперь приращение аргумента — вектор, поэтому возможны разные определения производных.

Определение 3.1. **Частной производной** *функции f в точке A по i-ой переменной называется предел*

$$\frac{\partial f}{\partial x_i}(A) = \lim_{t \to 0} \frac{f(A_t) - f(A)}{t} \qquad (\text{где} \quad \vec{AA_t} = t\vec{e}_i),$$

если он существует. Другими словами, $\dfrac{\partial f}{\partial x_i}$ – производная по переменной x_i при фиксированных остальных $n-1$ переменных.

Это определение является частным случаем следующего:

Определение 3.2. *Пусть $\vec{v}$ — вектор с нормой 1. Функция f дифференцируема по направлению $\vec{v}$ в точке A, если существует предел*

$$\frac{\partial f}{\partial \vec{v}}(A) = \lim_{t \to 0} \frac{f(A_t) - f(A)}{t} \qquad (\text{где} \quad \vec{AA_t} = t\vec{v}),$$

называемый производной функции f по направлению $\vec{v}$.

Но главное определение таково:

Определение 3.3. *Пусть A — внутренняя точка $D(f)$. Функция f* **дифференцируема** *в точке A, если существует такой вектор-строка g, что при $P \to A$*

$$f(P) - f(A) = g \cdot \vec{AP} + o(|\vec{AP}|).$$

Вектор g называется **градиентом** *f в точке A и обозначается $g = \nabla f(A)$.*

Изучим связь этих определений между собой и с непрерывностью функции.

Теорема 3.1. *Если функция f дифференцируема в точке $A = (x_1, \ldots, x_n)$, то она непрерывна в A и дифференцируема по любому направлению в A, при этом градиент равен*

$$\nabla f(A) = \left(\frac{\partial f}{\partial x_1}(A), \ldots, \frac{\partial f}{\partial x_n}(A) \right),$$

а производная по направлению $\vec{v}$ равна $\dfrac{\partial f}{\partial \vec{v}}(A) = \nabla f(A) \cdot \vec{v}$.

Доказательство. Пусть $U_r(A) \subset D(f)$. Обозначим $g = \nabla f(A)$. При $P \to A$ имеем $|f(P) - f(A)| = |g \cdot \vec{AP}| + o(|\vec{AP}|) \to 0$, поэтому f непрерывна в точке A. Пусть теперь P_t — такая точка, что $\vec{AP_t} = t\vec{v}$, $|t| < r$. Вычислим производную f по направлению $\vec{v}$:

$$\frac{\partial f}{\partial \vec{v}}(A) = \lim_{t\to 0} \frac{f(P_t) - f(A)}{t} = \lim_{t\to 0} \frac{g \cdot t\vec{v} + o(t)}{t} = g \cdot \vec{v} = \sum_{i=1}^{n} g_i v_i.$$

Подставив в качестве $\vec{v}$ j-й базисный вектор $\vec{e}_j$, получаем $\dfrac{\partial f}{\partial x_j}(A) = g \cdot \vec{e}_j = g_j$. □

Таким образом, градиент и производные по всем направлениям вычисляются через частные производные. Но существование частных производных (и даже производных по всем направлениям) не гарантирует дифференцируемость функции (и даже непрерывность):

Пример 3.1. Рассмотрим функцию 2 переменных

$$f(x,y) = \begin{cases} \psi(y/x^2) & \text{при} \quad x \neq 0 \\ 0 & \text{при} \quad x = 0 \end{cases} \quad \text{где} \quad \psi(u) = \begin{cases} \sin^2 u & \text{при} \quad 0 < u < \pi \\ 0 & \text{иначе.} \end{cases}$$

В точке $O = (0;0)$ функция f имеет производные по всем направлениям, равные 0. Действительно, $f(x,y)$ отлична от 0 лишь на открытом множестве U, лежащем между осью Ox и параболой $y = \pi x^2$. Оси Ox и Oy не пересекают U; другие прямые, проходящие через точку O, пересекают U на некотором расстоянии от O.

Однако, функция $f(x,y)$ разрывна в O: если взять последовательность точек $A_k = (1/k; \pi/2k^2)$, то $A_k \to O$, но $f(A_k) = 1 \to 1 \neq f(O)$.

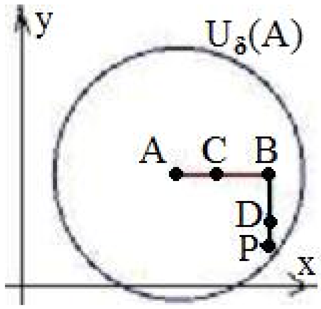
Теорема 3.2. (достаточное условие дифференцируемости) *Если существует такое $r > 0$, что в каждой точке окрестности $U_r(A)$ функция f имеет все частные производные, и они* **непрерывны** *в точке A, то функция f дифференцируема в A.*

Доказательство. Для простоты обозначений рассмотрим случай $n = 2$. Пусть $A = (x_o, y_o)$. Поскольку частные производные f непрерывны в точке A, для любого $\varepsilon > 0$ можно так подобрать $\delta \in (0; r]$, что

$$\left|\frac{\partial f}{\partial x}(P) - \frac{\partial f}{\partial x}(A)\right|, \ \left|\frac{\partial f}{\partial y}(P) - \frac{\partial f}{\partial y}(A)\right| < \varepsilon \quad \forall\ P \in U_\delta(A).$$

Пусть $P = (x_1, y_1) \in U_\delta(A)$. Возьмём вспомогательную точку $B = (x_1, y_o)$, она также лежит в $U_\delta(A)$. Выразим $f(P) - f(A)$ по теореме Лагранжа, фиксируя переменные поочерёдно:

$$\big(f(B) - f(A)\big) + \big(f(P) - f(B)\big) = \frac{\partial f}{\partial x}(C) \cdot (x_1 - x_o) + \frac{\partial f}{\partial y}(D) \cdot (y_1 - y_o),$$



что отличается от выражения

$$\frac{\partial f}{\partial x}(A)\cdot(x_1-x_o)+\frac{\partial f}{\partial y}(A)\cdot(y_1-y_o)=g\cdot\vec{AP}$$

на величину, по модулю не превышающую

$$\left(\left|\frac{\partial f}{\partial x}(C)-\frac{\partial f}{\partial x}(A)\right|+\left|\frac{\partial f}{\partial y}(D)-\frac{\partial f}{\partial y}(A)\right|\right)|\vec{AP}|<2\varepsilon|\vec{AP}|.$$

Поскольку $\varepsilon\to 0$ при $\delta\to 0$, мы получаем

$$f(P)-f(A)=g\cdot\vec{AP}+o(|\vec{AP}|)\quad\text{при}\quad P\to A,$$

где компоненты вектора g — частные производные функции f в точке A. Таким образом, f дифференцируема в A, и её градиент в A равен вектору g. □

Пример 3.2. Функция $f(x,y)=\sqrt{x^2+2y^2}$ определена на всей плоскости $\mathbb{R}^2$. В точке $O=(0;0)$ она не имеет частных производных. В любой другой точке $(x;y)$ имеем $\dfrac{\partial}{\partial x}\sqrt{x^2+2y^2}=\dfrac{x}{\sqrt{x^2+2y^2}}$, $\dfrac{\partial}{\partial y}\sqrt{x^2+y^2}=\dfrac{2y}{\sqrt{x^2+2y^2}}$. Обе частные производные — элементарные функции $\Longrightarrow$ они непрерывны на всей области определения (на $\mathbb{R}^2\backslash\{O\}$). Следовательно, $f(x,y)=\sqrt{x^2+2y^2}$ дифференцируема во всех точках этой области.

Геометрический смысл градиента. Если $\vec{v}$ — единичный вектор, то производная по его направлению равна

$$\frac{\partial f}{\partial\vec{v}}(A)=\nabla f(A)\cdot\vec{v}=|\nabla f|\cdot\cos(\widehat{\vec{v},\nabla f}).$$

Из этой формулы видно, что максимальное значение производной по направлению (равное норме градиента) достигается в том случае, когда $\vec{v}$ сонаправлен градиенту. Таким образом, *вектор градиента направлен в сторону максимальной скорости роста функции и по норме равен этой скорости.*

Определение 3.4. *Если скалярная функция $f(x_1,\ldots,x_n)$ дифференцируема в точке A, то* **полным дифференциалом** *функции f в точке A называется выражение*

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

Пример 3.2 (продолжение).

$$d\sqrt{x^2+2y^2} = \frac{x}{\sqrt{x^2+2y^2}}dx + \frac{2y}{\sqrt{x^2+2y^2}}dy = \frac{xdx+2ydy}{\sqrt{x^2+2y^2}}.$$

В точке $A=(1;2)$ получаем $d\sqrt{x^2+2y^2}\Big|_A = \dfrac{1}{3}dx + \dfrac{4}{3}dy$; производная по направлению вектора $\vec{v} = \dfrac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$ равна

$$\frac{\partial}{\partial\vec{v}}\sqrt{x^2+2y^2}\Big|_A = \nabla f(A)\cdot\vec{v} = \begin{pmatrix}\frac{1}{3} & \frac{4}{3}\end{pmatrix}\cdot\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} = \frac{5}{3\sqrt{2}}.$$

Максимальная производная по направлению получится, если взять $\vec{w} \upuparrows \nabla f(A)$, т.е. $\vec{w} = \nabla f(A)/|\nabla f(A)| = \dfrac{1}{\sqrt{17}}\begin{bmatrix}1\\4\end{bmatrix}$:

$$\frac{\partial}{\partial\vec{w}}\sqrt{x^2+2y^2}\Big|_A = \nabla f(A)\cdot\vec{w} = \begin{pmatrix}\frac{1}{3} & \frac{4}{3}\end{pmatrix}\cdot\frac{1}{\sqrt{17}}\begin{bmatrix}1\\4\end{bmatrix} = \frac{\sqrt{17}}{3}.$$

---

4 Дифференцирование векторных ФНП

Определение 4.1. *Пусть A — внутренняя точка $D(F)\subset\mathbb{R}^n$; $F : D(F)\mapsto\mathbb{R}^k$. Функция F* **дифференцируема** *в точке A, если существует линейный оператор $DF(A):\mathbb{R}^n\mapsto\mathbb{R}^k$, такой, что*

$$F(P)-F(A) = DF(A)\cdot\vec{AP} + o(|\vec{AP}|) \quad \text{при} \quad P\to A,$$

т. е. $\forall\varepsilon>0\ \exists\delta>0 : |F(P)-F(A)-DF(A)\cdot\vec{AP}| < \varepsilon|\vec{AP}|$ при всех $P\in U_\delta(A)$.

Матрица оператора $DF(A)$ называется матрицей Якоби функции F в точке A и обозначается $J_F(A)$.

Лемма 4.1. *Векторная функция F дифференцируема в точке $A=(a_1,\ldots,a_n) \iff$ все её координатные функции $f_1,\ldots,f_k$ дифференцируемы в A. При этом i-я строка матрицы $J_F(A)$ равна $\nabla f_i(A)$.*

Доказательство. Обозначим через J_i i-ю строку матрицы $J_F(A)$. Тогда дифференцируемость F в точке A с заданной матрицей Якоби равносильна существованию нулевого предела векторной функции

$$\vec{V}(P) = \frac{1}{|\vec{AP}|}\left(F(P)-F(A)-\begin{bmatrix}J_1\cdot\vec{AP}\\ \vdots \\ J_k\cdot\vec{AP}\end{bmatrix}\right)\to\vec{0}$$

при $P \to A$. Это, в свою очередь, равносильно стремлению к 0 координатных функций $f_i(P) - f_i(A) - J_i \cdot \vec{AP} \quad \forall\, i = 1, \ldots, k$, что означает дифференцируемость всех f_i в точке A, причём $\nabla f_i(A) = J_i$. $\square$

Следствие 4.1. *Если F дифференцируема в точке A, то она непрерывна в A.*

Следствие 4.2. *Если координатные функции $f_1, \ldots, f_k$ имеют частные производные в окрестности точки A, непрерывные в A, то F дифференцируема в A. При этом её матрица Якоби имеет вид*

$$J_F(A) = \begin{bmatrix} \partial f_1/\partial x_1 & \ldots & \partial f_1/\partial x_n \\ \ldots & \ldots & \ldots \\ \partial f_k/\partial x_1 & \ldots & \partial f_k/\partial x_n \end{bmatrix}.$$

Полным дифференциалом векторной функции F в точке A называется выражение

$$dF(A) = \begin{bmatrix} df_1(A) \\ \vdots \\ df_k(A) \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(A)dx_1 + \ldots + \dfrac{\partial f_1}{\partial x_n}(A)dx_n \\ \ldots \\ \dfrac{\partial f_k}{\partial x_1}(A)dx_1 + \ldots + \dfrac{\partial f_k}{\partial x_n}(A)dx_n \end{bmatrix} = J_F(A) \cdot \begin{bmatrix} dx_1 \\ \vdots \\ dx_n \end{bmatrix}$$

Пример 4.1. Рассмотрим двумернозначную функцию трёх переменных

$$F(x, y, z) = \begin{bmatrix} x/y \\ xyz \end{bmatrix} = \begin{bmatrix} f_1(x, y, z) \\ f_2(x, y, z) \end{bmatrix}.$$

Первая координатная функция f_1 дифференцируема на всей своей области определения $\{y \neq 0\}$; вторая f_2 дифференцируема на всём $\mathbb{R}^3$. Следовательно, F дифференцируема на $D(F) = \{y \neq 0\}$. Её матрица Якоби равна

$$J_F(x, y, z) = \begin{bmatrix} \nabla f_1(x, y, z) \\ \nabla f_2(x, y, z) \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} & \dfrac{\partial f_1}{\partial z} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} & \dfrac{\partial f_2}{\partial z} \end{bmatrix} = \begin{bmatrix} 1/y & -x/y^2 & 0 \\ yz & xz & xy \end{bmatrix};$$

полный дифференциал функции F выглядит так:

$$dF(x, y, z) = \begin{bmatrix} \dfrac{1}{y}dx - \dfrac{x}{y^2}dy \\ yzdx + xzdy + xydz \end{bmatrix}.$$

Теорема 4.1. (о дифференцировании сложной ФНП). *Пусть A — внутренняя точка $D(F) \subset \mathbb{R}^n$, $F : D(F) \mapsto \mathbb{R}^k$; $B = F(A)$ — внутренняя точка $D(G) \subset \mathbb{R}^k$, $G : D(G) \mapsto \mathbb{R}^m$. Если F дифференцируема в точке A, G дифференцируема в точке B, то функция $G \circ F$ дифференцируема в A, и при этом $J_{G \circ F}(A) = J_G(B) \cdot J_F(A)$.*

В этой теореме функции F и G могут быть как векторными, так и скалярными ($k \geq 1$, $m \geq 1$).

Доказательство. Сначала докажем, что A – внутренняя точка области определения $G \circ F$. Действительно, $\exists \varepsilon > 0$: $U_\varepsilon(B) \subset D(G)$. Поскольку F непрерывна в A, то $\exists \delta > 0$: для всех $P \in U_\delta(A)$ будет $F(P) \in U_\varepsilon(B) \subset D(G)$, следовательно, $P \in D(G \circ F)$.

Пусть $P \in D(G \circ F)$, $P \to A$. Тогда приращение функции $G \circ F$ имеет вид

$$G \circ F(P) - G \circ F(A) = G(F(P)) - G(B) = J_G(B) \cdot \overrightarrow{BF(P)} + o(|\overrightarrow{BF(P)}|) =$$

$$= J_G(B)(J_F(A)\vec{AP} + o(|\vec{AP}|)) + o(|\vec{AP}|) = J_G(B)J_F(A)\vec{AP} + o(|\vec{AP}|). \qquad \square$$

Следствие 4.3. *Инвариантность полного дифференциала:*

$$d(G \circ F)(x_1, \ldots, x_n) = dG(y_1, \ldots, y_k)\big|_{y_i = f_i(x_1, \ldots, x_n)}$$

Доказательство.

$$d(G \circ F)(X) = J_{G \circ F}(X) \cdot \begin{bmatrix} dx_1 \\ \vdots \\ dx_n \end{bmatrix} = J_G(Y) \cdot J_F(X) \cdot \begin{bmatrix} dx_1 \\ \vdots \\ dx_n \end{bmatrix} = J_G(Y) \cdot \begin{bmatrix} df_1(X) \\ \ldots \\ df_k(X) \end{bmatrix}$$

$\square$

Пример 4.2. Получим полный дифференциал скалярной функции

$$\Phi(x, y) = \sqrt{xy + e^{x-2y}}.$$

Представим её в виде $\Phi = G \circ F$, где

$$F(x, y) = \begin{bmatrix} xy \\ e^{x-2y} \end{bmatrix}, \quad G(u, v) = \sqrt{u + v}.$$

Тогда

$$d\Phi(x, y) = dG(u, v)\Bigg|_{\substack{u = xy \\ v = e^{x-2y}}} =$$

$$= \frac{du + dv}{2\sqrt{u + v}}\Bigg|_{\substack{u = xy \\ v = e^{x-2y}}} = \frac{d(xy) + de^{x-2y}}{2\sqrt{xy + e^{x-2y}}} = \frac{ydx + xdy + e^{x-2y}dx - 2e^{x-2y}dy}{2\sqrt{xy + e^{x-2y}}}$$

Если требуется найти полный дифференциал сложной функции, такой способ бывает удобнее, чем вычислять частные производные по отдельности.

5 Дифференцирование высших порядков

Пусть $D(f) \subset \mathbb{R}^n$; функция $f\colon D(f) \mapsto \mathbb{R}^k$. Если функция f дифференцируема в каждой точке открытого множества $U \subset D(f)$, то на U определена её первая производная Df — для скалярной f ($k = 1$) это градиент $\nabla f\colon U \mapsto \mathbb{R}^k$, а для векторной f это матрица Якоби из k строк и n столбцов $J_f\colon U \mapsto \mathbb{R}^{kn}$. Если функция Df, в свою очередь, дифференцируема в некоторой точке $A \in U$, то говорят, что f **дважды дифференцируема** в A.

Если f скалярная, то её вторая производная $D^2 f$ принимает значения в $\mathbb{R}^{n^2}$: её координатные функции — вторые частные производные $\dfrac{\partial^2 f}{\partial x_j \partial x_i} = \dfrac{\partial}{\partial x_j}\Big(\dfrac{\partial f}{\partial x_i}\Big)$. Их удобно записать в виде **матрицы Гессе** $n \times n$:

$$\mathbf{H}_f = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dots & \dots & \dots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix}.$$

Пример 5.1. Функция $f(x, y, z) = xy^2/z$ дифференцируема в каждой точке своей области определения $U = \mathbb{R}^3 \backslash Oxy$. Её частные производные равны

$$\frac{\partial f}{\partial x} = \frac{y^2}{z}; \quad \frac{\partial f}{\partial y} = \frac{2xy}{z}; \quad \frac{\partial f}{\partial z} = -\frac{xy^2}{z^2}.$$

Каждая из этих функций дифференцируема на всей $U \Longrightarrow$ векторная функция ∇f дифференцируема на $U \Longrightarrow f$ дифференцируема дважды в любой внутренней точке U, т.е. в $\forall\ (x; y; z) \in U$ (U открытое). Её матрица Гессе выглядит так:

$$\mathbf{H}_f(x; y; z) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial x \partial z} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} & \dfrac{\partial^2 f}{\partial y \partial z} \\ \dfrac{\partial^2 f}{\partial z \partial x} & \dfrac{\partial^2 f}{\partial z \partial y} & \dfrac{\partial^2 f}{\partial z^2} \end{bmatrix} = \begin{bmatrix} 0 & 2y/z & -y^2/z^2 \\ 2y/z & 2x/z & -2xy/z^2 \\ -y^2/z^2 & -2xy/z^2 & 2xy^2/z^3 \end{bmatrix}.$$

Если $D^2 f$ также определена на открытом множестве и дифференцируема в некоторой его точке A, то f трижды дифференцируема в A и определены третьи частные производные

$$\frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i} = \frac{\partial}{\partial x_k}\left(\frac{\partial^2 f}{\partial x_j \partial x_i}\right)$$

И так далее. Из m-кратной дифференцируемости функции следует существование всех m-ых частных производных, но обратное неверно (как мы видели, даже при $m = 1$).

Теорема 5.1. *Пусть $D(f) \subset \mathbb{R}^n$; $f : D(f) \mapsto \mathbb{R}$. Пусть A – внутренняя точка $D(f)$. Если все частные производные функции f имеют свои частные производные в $U_r(A)$, непрерывные в A, то f дважды дифференцируема в A, причём*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(A) = \frac{\partial^2 f}{\partial x_j \partial x_i}(A). \tag{1}$$

Доказательство. Из теоремы 3.2 следует, что каждая частная производная функции f дифференцируема в A. Это по лемме 4.1 равносильно дифференцируемости векторной функции ∇f, т. е. двукратной дифференцируемости f, в точке A.

Теперь докажем равенство (1). Пусть $A = (a_1, \ldots, a_n)$; все частные производные f дифференцируемы в $U_r(A)$. При вычислении вторых частных производных в (1) мы фиксируем все координаты, кроме x_i и x_j, т. е. рассматриваем плоскость $\Pi \ni A$, параллельную i-ой и j-ой координатным осям. Пусть $P_{y,z} \in \Pi \cap U_r(A)$ — такая точка, что $\vec{AP}_{y,z} = y\vec{e}_i + z\vec{e}_j$. Рассмотрим функцию двух переменных

$$Q(y,z) = \frac{f(P_{y,z}) - f(P_{y,0}) - f(P_{0,z}) + f(A)}{yz},$$

определённую на множестве $\{(y;z) : \ y^2 + z^2 < r^2, \ yz \neq 0\}$. Применим теорему Лагранжа к функции $\Phi(y) = f(P_{y,z}) - f(P_{y,0})$:

$$Q(y,z) = \frac{\Phi(y) - \Phi(0)}{yz} = \frac{\Phi'(\eta)}{z} = \frac{1}{z}\left(\frac{\partial f}{\partial x_i}(P_{\eta,z}) - \frac{\partial f}{\partial x_i}(P_{\eta,0})\right) = \frac{\partial^2 f}{\partial x_j \partial x_i}(P_{\eta,\theta})$$

– вновь по теореме Лагранжа. Поскольку $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ непрерывна в точке A и поскольку $\varrho(A, P_{\eta,\theta}) < \varrho(A, P_{y,z})$, мы получаем $Q(y,z) \to \dfrac{\partial^2 f}{\partial x_j \partial x_i}(A)$ при $(y;z) \to (0;0)$.

Теперь применим теорему Лагранжа к функции $\Psi(z) = f(P_{y,z}) - f(P_{0,z})$:

$$Q(y,z) = \frac{\Psi(z) - \Psi(0)}{yz} = \frac{\Psi'(\xi)}{y} = \frac{1}{y}\left(\frac{\partial f}{\partial x_j}(P_{y,\xi}) - \frac{\partial f}{\partial x_j}(P_{0,\xi})\right) = \frac{\partial^2 f}{\partial x_i \partial x_j}(P_{\zeta,\xi})$$

– вновь по теореме Лагранжа. Поскольку $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ также непрерывна в точке A и $\varrho(A, P_{\zeta,\xi}) < \varrho(A, P_{y,z})$, мы получаем $Q(y,z) \to \dfrac{\partial^2 f}{\partial x_i \partial x_j}(A)$ при $(y;z) \to (0;0)$. Но функция может иметь в точке только один предел $\Longrightarrow$ верно равенство (1). □

Следствие 5.1. *В условиях теоремы 5.1 матрица Гессе в точке A симметрична:* $\mathbf{H}_f^{\mathrm{T}}(A) = \mathbf{H}_f(A)$.

Следствие 5.2. *Если функция f дифференцируема $k-1$ раз в открытом множестве U, и если все её $(k-1)$-ые частные производные имеют свои частные производные, непрерывные в U, то f дифференцируема k раз в каждой точке U, и значения частных производных* **не меняются при перестановке последовательности дифференцирований.**

Доказательство. Применяя теорему 5.1 к самой функции f и её частным производным от первого до $(k-2)$-го порядка, мы докажем инвариантность частной производной при перестановке двух соседних дифференцирований. Любую перестановку можно получить конечной цепочкой таких элементарных перестановок. Например:

$$\frac{\partial^4 f}{\partial x \partial y \partial z \partial x} = \frac{\partial^4 f}{\partial x \partial z \partial y \partial x} = \frac{\partial^4 f}{\partial z \partial x \partial y \partial x} = \frac{\partial^4 f}{\partial z \partial^2 x \partial y} \qquad \square$$

Определение 5.1. *Если скалярная функция $f(x_1, \dots, x_n)$ дифференцируема k раз в точке A, то полным* **дифференциалом k-го порядка** *функции f в точке A называется выражение*

$$d^k f = \sum_{i_1=1}^{n} \dots \sum_{i_k=1}^{n} \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(A) dx_{i_1} \dots dx_{i_k}$$

В этой сумме n^k слагаемых. Однако, если все k-ые частные производные непрерывны, то среди слагаемых будут встречаться одинаковые. Второй и третий дифференциалы будут выглядеть так:

$$d^2 f = (dx_1 \dots, dx_n) \cdot \mathbf{H}_f \cdot \begin{bmatrix} dx_1 \\ \vdots \\ dx_n \end{bmatrix} = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2} dx_i^2 + 2 \sum_{1 \le i < j \le n} \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i dx_j$$

$$d^3 f = \sum_{i=1}^{n} \frac{\partial^3 f}{\partial x_i^3} dx_i^3 + 3 \sum_{i=1}^{n} \sum_{j \ne i} \frac{\partial^3 f}{\partial x_i \partial x_j^2} dx_i dx_j^2 + 6 \sum_{i<j<k} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k} dx_i dx_j dx_k$$

Частная производная $\dfrac{\partial^k f}{\partial x_{i_1}^{k_1} \dots \partial x_{i_m}^{k_m}}$ будет повторяться $\dfrac{k!}{k_1! \cdot \ldots \cdot k_m!}$ раз.

Пример 5.2. Найдём 4-й дифференциал функции $f(x, y) = x^2/y$:

$$\begin{aligned} \partial f/\partial x &= 2x/y \\ \partial f/\partial y &= -x^2/y^2 \end{aligned} \qquad \begin{aligned} \partial^2 f/\partial x^2 &= 2/y \\ \partial^2 f/\partial x \partial y &= -2x/y^2 \\ \partial^2 f/\partial y^2 &= 2x^2/y^3 \end{aligned} \qquad \begin{aligned} \partial^3 f/\partial x^3 &= 0 \\ \partial^3 f/\partial x^2 \partial y &= -2/y^2 \\ \partial^3 f/\partial x \partial y^2 &= 4x/y^3 \\ \partial^3 f/\partial y^3 &= -6x^2/y^4 \end{aligned} \qquad \begin{aligned} \partial^4 f/\partial x^4 &= 0 \\ \partial^4 f/\partial x^3 \partial y &= 0 \\ \partial^4 f/\partial x^2 \partial y^2 &= 4/y^3 \\ \partial^4 f/\partial x \partial y^3 &= -12x/y^4 \\ \partial^4 f/\partial y^4 &= 24x^2/y^5 \end{aligned}$$

С учётом повторяемости частных производных получается

$$d^4 f = \frac{24}{y^3} dx^2 dy^2 - \frac{48x}{y^4} dx dy^3 + \frac{24x^2}{y^5} dy^4$$

6 Формула Тейлора для ФНП

Вспомним формулу Тейлора для функции скалярного переменного: Пусть функция $f(x)$ $m+1$ раз дифференцируема в $U_r(a) \subset \mathbb{R}$. Тогда для любого $x \in U_r(a)$ верна *формула Тейлора с остаточным членом в форме Лагранжа*:

$$f(x) = f(a) + \sum_{k=1}^{m} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(m+1)}(\xi)}{(m+1)!}(x-a)^{m+1}, \tag{2}$$

где ξ — некоторая точка, лежащая между a и x.

Выведем формулы Тейлора для функции n переменных.
Пусть $A_0 \in U_r(A_0) \subset D(f) \subset \mathbb{R}^n$; $f : D(f) \mapsto \mathbb{R}$. Зафиксируем точку $A_1 \in U_r(A_0)$ и обозначим через A_t такие точки, что $A_0\vec{A}_t = t \cdot A_0\vec{A}_1$, $0 \le t \le 1$. Пусть $A = (a_1, \dots, a_n)$, $A_0\vec{A}_1 = (v_1, \dots, v_n)$, тогда точка A_t имеет координаты $(a_1 + tv_1, \dots, a_n + tv_n)$.

Рассмотрим функцию $g(t) = f(A_t)$, определённую на $[0; 1]$. Если f дифференцируема в $U_r(A_0)$, то g дифференцируема, и её производная по теореме о дифференцировании сложной ФНП равна

$$\frac{dg}{dt} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(A_t) \frac{dx_j(t)}{dt} = \sum_{j=1}^{n} v_j \frac{\partial f}{\partial x_j}(A_t)$$

Если f дважды дифференцируема в $U_r(A_0)$, то dg/dt дифференцируема, и

$$\frac{d^2 g}{dt^2} = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\sum_{j=1}^{n} v_j \frac{\partial f}{\partial x_j}(A_t)\right)\frac{dx_i(t)}{dt} = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i v_j \frac{\partial^2 f}{\partial x_i \partial x_j}(A_t)$$

И т. д. Если f дифференцируема в $U_r(A_0)$ k раз, то

$$\frac{d^k g}{dt^k} = \sum_{i_1=1}^{n} \dots \sum_{i_k=1}^{n} v_{i_1} \dots v_{i_k} \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(A_t).$$

Эта формула легко запоминается: нужно в k-ый дифференциал $d^k f(A_0)$ подставить v_i вместо dx_i.

Теорема 6.1. *Пусть $A = (a_1, \dots, a_n)$ — внутренняя точка $D(f) \subset \mathbb{R}^n$. Если скалярная ФНП f $m+1$ раз дифференцируема в $U_r(A) \subset D(f)$, то для любой точки $X = (x_1, \dots, x_n) \in U_r(A)$ верна* **формула Тейлора с остаточным членом в форме Лагранжа:**

$$f(X) = f(A) + \sum_{k=1}^{m}\left(\sum_{i_1=1}^{n} \dots \sum_{i_k=1}^{n} \frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(A) \frac{(x_{i_1} - a_{i_1}) \dots (x_{i_k} - a_{i_k})}{k!}\right) +$$

$$+ \sum_{i_1=1}^{n} \dots \sum_{i_{m+1}=1}^{n} \frac{\partial^{m+1} f}{\partial x_{i_1} \dots \partial x_{i_{m+1}}}(Y) \frac{(x_{i_1} - a_{i_1}) \dots (x_{i_{m+1}} - a_{i_{m+1}})}{(m+1)!},$$

где Y — некоторая точка, лежащая на отрезке, соединяющем A и X.

Доказательство. Применим формулу (2) к функции $g(t) = f(X_t)$, где $X_t = (a_1 + t(x_1 - a_1), \dots a_n + t(x_n - a_n))$. $\square$

Теорема 6.2. *Пусть $A = (a_1, \dots, a_n)$ — внутренняя точка $D(f) \subset \mathbb{R}^n$. Если скалярная ФНП f m раз дифференцируема в $U_r(A) \subset D(f)$ и её m-ые частные производные непрерывны в точке A, то верна* **формула Тейлора с остаточным членом в форме Пеано:** $f(X) =$

$$f(A) + \sum_{k=1}^{m}\left(\sum_{i_1=1}^{n}\dots\sum_{i_k=1}^{n}\frac{\partial^k f}{\partial x_{i_1}\dots\partial x_{i_k}}(A)\frac{(x_{i_1}-a_{i_1})\dots(x_{i_k}-a_{i_k})}{k!}\right) + o(|\vec{AX}|^m).$$

Доказательство. Применим теорему 6.1 к функции f, взяв $m-1$ вместо m:

$$f(X) = f(A) + \sum_{k=1}^{m}\left(\sum_{i_1=1}^{n}\dots\sum_{i_k=1}^{n}\frac{\partial^k f}{\partial x_{i_1}\dots\partial x_{i_k}}(A)\frac{(x_{i_1}-a_{i_1})\dots(x_{i_k}-a_{i_k})}{k!}\right) +$$

$$+\sum_{i_1=1}^{n}\dots\sum_{i_m=1}^{n}\underbrace{\left(\frac{\partial^m f}{\partial x_{i_1}\dots\partial x_{i_m}}(Y) - \frac{\partial^m f}{\partial x_{i_1}\dots\partial x_{i_m}}(A)\right)}_{\to 0 \text{ при } X\to A}\underbrace{\frac{(x_{i_1}-a_{i_1})\dots(x_{i_m}-a_{i_m})}{m!}}_{\leq|\vec{AX}|^m/m!}$$

Последняя строка и есть остаточный член в форме Пеано. $\square$

Формулу Тейлора легко запомнить в таком виде:

$$f(A+\vec{dx}) = f(A) + \sum_{k=1}^{m}\frac{1}{k!}d^k f(A) + \text{остаточный член},$$

где на место dx_i надо будет подставить $x_i - a_i$.

Пример 6.1. Выпишем формулу Тейлора с остаточным членом в форме Пеано при $m = 1$ и $m = 2$ для функции $f(x,y) = \sqrt{x^2+y^2}$ в окрестности точки $A(3;4)$.

$$df = \frac{xdx + ydy}{\sqrt{x^2+y^2}} \implies df(A) = \frac{3}{5}dx + \frac{4}{5}dy;$$

$$d^2 f = \frac{y^2dx^2 + x^2dy^2 - 2xydxdy}{(x^2+y^2)^{3/2}} \implies d^2 f(A) = \frac{16}{125}dx^2 + \frac{9}{125}dy^2 - \frac{24}{125}dxdy.$$

$$\text{Для } m=1: \quad f(x,y) = 5 + \frac{3}{5}(x-3) + \frac{4}{5}(y-4) + o\big(\sqrt{(x-3)^2+(y-4)^2}\big);$$

$$\text{Для } m=2: \quad f(x,y) = 5 + \frac{3}{5}(x-3) + \frac{4}{5}(y-4) +$$

$$+\frac{16}{250}(x-3)^2 + \frac{9}{250}(y-4)^2 - \frac{12}{125}(x-3)(y-4) + o\big((x-3)^2+(y-4)^2\big)$$

при $(x;y) \to (3;4)$.

7 Восстановление ФНП по градиенту

Теорема 7.1. *Пусть* $\Omega \subset \mathbb{R}^n$ *открытое связное;* $\vec{V} = (v_1, \ldots, v_n) : \Omega \mapsto \mathbb{R}^n$ *— векторное поле, дифференцируемое в* Ω*, причём его частные производные непрерывны. Тогда для того, чтобы* $\vec{V}$ *было градиентом некоторой скалярной функции* f *на* Ω*, необходимы равенства*

$$\frac{\partial v_i}{\partial x_j} = \frac{\partial v_j}{\partial x_i} \quad \text{на } \Omega, \qquad 1 \le i < j \le n. \tag{3}$$

(Если $n = 2$ или 3, и Ω не имеет сквозных дыр, то условие (3) является достаточным, что будет доказано в 3 семестре).

Доказательство. Если $\vec{V} = \nabla f$, то $\dfrac{\partial v_i}{\partial x_j} = \dfrac{\partial^2 f}{\partial x_j \partial x_i}$, $\dfrac{\partial v_j}{\partial x_i} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}$, и эти частные производные по условию непрерывны. Следовательно, по теореме 5.1 они равны. □

Алгоритм восстановления ФНП по её градиенту объясним на примере.

Пример 7.1. *Построить на области* $\mathbb{R}^3 \supset \Omega = \{x > |z|\}$ *функцию* $f(x, y, z)$*, для которой*

$$df = \frac{2x^2 - z^2 + xy}{\sqrt{x^2 - z^2}}dx + (\sqrt{x^2 - z^2} + yz^2)dy + \Big(y^2 z - \frac{(x+y)z}{\sqrt{x^2 - z^2}}\Big)dz =$$

$= Pdx + Qdy + Rdz$. Сначала проверим необходимое условие:

$$\frac{\partial P}{\partial y} = \frac{x}{\sqrt{x^2 - z^2}} = \frac{\partial Q}{\partial x}; \quad \frac{\partial P}{\partial z} = \frac{z(z^2 + xy)}{(x^2 - z^2)^{3/2}} = \frac{\partial R}{\partial x};$$

$$\frac{\partial Q}{\partial z} = \frac{-z}{\sqrt{x^2 - z^2}} + 2yz = \frac{\partial R}{\partial y}$$

выполнены на всей Ω, и поскольку Ω не имеет сквозных дыр, искомая f должна существовать.

Зафиксируем y и z и проинтегрируем $P = \dfrac{\partial f}{\partial x}$ по x при $x > |z|$:

$$f(x, y, z) = \int \frac{2x^2 - z^2 + xy}{\sqrt{x^2 - z^2}}dx = (x + y)\sqrt{x^2 - z^2} + C_1(y, z).$$

Продифференцировав полученное выражение по y и по z, мы получим, соответственно,

$$\sqrt{x^2 - z^2} + \frac{\partial C_1}{\partial y} = \frac{\partial f}{\partial y} = \sqrt{x^2 - z^2} + yz^2;$$

$$\frac{(x+y)z}{\sqrt{x^2 - z^2}} + \frac{\partial C_1}{\partial z} = \frac{\partial f}{\partial z} = \frac{(x+y)z}{\sqrt{x^2 - z^2}} + y^2 z.$$

Отсюда $\dfrac{\partial C_1}{\partial y} = yz^2; \quad \dfrac{\partial C_1}{\partial z} = y^2 z.$

Зафиксировав z, проинтегрируем по y: $C_1(y,z) = \displaystyle\int yz^2 dy = \frac{y^2z^2}{2} + C_2(z).$

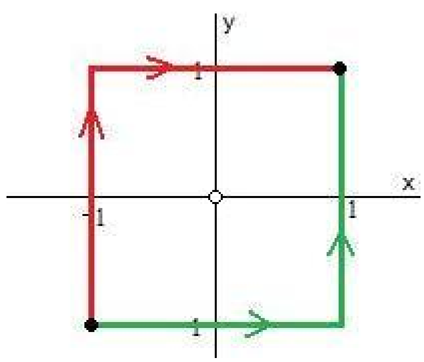
Продифференцировав такое выражение по z, получаем

$$y^2z + C_2'(z) = \frac{\partial C_1}{\partial z} = y^2 z \Longrightarrow C_2(z) = const.$$

Ответ: $f(x,y,z) = (x+y)\sqrt{x^2-z^2} + \dfrac{y^2z^2}{2} + const.$

Теперь приведём пример, показывающий, что на области с дырой условие (3) не является достаточным.

Пример 7.2. Рассмотрим векторное поле $\vec{V} = \left(\dfrac{-y}{x^2+y^2}; \dfrac{x}{x^2+y^2}\right)$ на области $\Omega = \mathbb{R}^2 \backslash \{O\}$. Условие (3) выполнено. Предположим, что существует $f\colon \Omega \mapsto \mathbb{R}$, $\nabla f = \vec{V}$. Пусть $f(-1,-1) = C$. Тогда



$$f(1,1) = C + \int_{-1}^{1} \frac{-(-1)}{x^2+(-1)^2} dx + \int_{-1}^{1} \frac{1}{1^2+y^2} dy = C + \frac{\pi}{2} + \frac{\pi}{2};$$

с другой стороны,

$$f(1,1) = C + \int_{-1}^{1} \frac{-1}{(-1)^2+y^2} dy + \int_{-1}^{1} \frac{-1}{x^2+1^2} dx = C - \frac{\pi}{2} - \frac{\pi}{2}.$$

Противоречие.

8 Безусловные экстремумы

Безусловные экстремумы

Определение 8.1. *Пусть A — внутренняя точка $D(f) \subset \mathbb{R}^n$; f скалярная. Говорят, что A –*

точка **максимума**, *если* $\exists \varepsilon > 0 : \ \forall P \in \overset{o}{U}_\varepsilon(A) \ \ f(P) < f(A)$;

точка **минимума**, *если* $\exists \varepsilon > 0 : \ \forall P \in \overset{o}{U}_\varepsilon(A) \ \ f(P) > f(A)$;

точка **нестрогого максимума**, *если* $\exists \varepsilon > 0 : \ \forall P \in \overset{o}{U}_\varepsilon(A) \ \ f(P) \le f(A)$;

точка **нестрогого минимума**, *если* $\exists \varepsilon > 0 : \ \forall P \in \overset{o}{U}_\varepsilon(A) \ \ f(P) \ge f(A)$.

В первых двух случаях A называют точкой **экстремума**, *в последних двух — точкой* **нестрогого экстремума**.

Следующая теорема обобщает теорему Ферма для ФНП и даёт

Необходимые условия экстремума

Теорема 8.1. *Пусть A — внутренняя точка $D(f) \subset \mathbb{R}^n$; скалярная функция f имеет в точке A все первые частные производные. Если A – точка (нестрогого) экстремума, то*

$$\frac{\partial f}{\partial x_1}(A) = \ldots = \frac{\partial f}{\partial x_n}(A) = 0. \tag{4}$$

Точки, удовлетворяющие условию (4), называются **критическими**.

Доказательство. Пусть точка $A = (a_1, \ldots, a_n)$ не критическая. Предположим, что $\dfrac{\partial f}{\partial x_1}(A) = C > 0$. Тогда

$$f(x_1 + t, x_2, \ldots, x_n) = f(A) + Ct + o(t) \begin{cases} > f(A) \text{ при } 0 < t < \varepsilon \\ < f(A) \text{ при } -\varepsilon < t < 0 \end{cases}$$

при некотором $\varepsilon > 0$. Таким образом, A не является точкой (нестрогого) экстремума. Аналогично рассматриваются случаи $\dfrac{\partial f}{\partial x_k}(A) > 0$ и $\dfrac{\partial f}{\partial x_k}(A) < 0$, $k = 1, \ldots, n$. $\square$

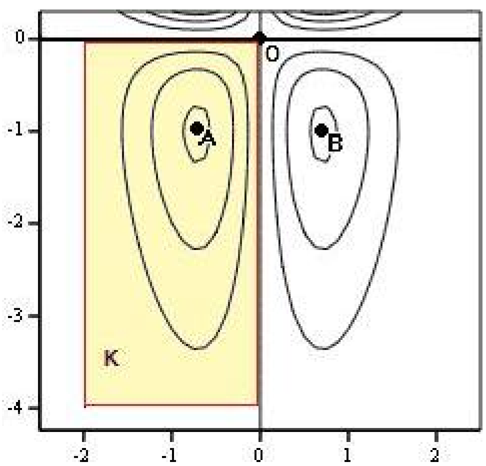
Необходимое условие экстремума (4) не является достаточным, даже если потребовать дифференцируемости f в точке A (тогда $\nabla f(A) = \vec{0}$). Если $\nabla f(A) = \vec{0}$, но при этом $\forall \varepsilon > 0 \ \ \exists P, Q \in \overset{o}{U}_\varepsilon(A)$: $f(P) > f(A) > f(Q)$, то A называется **седловой точкой**.

Пример 8.1. Рассмотрим функцию, определённую на $\mathbb{R}^2$: $f(x, y) = xye^{y-x^2}$.

Её частные производные: $\dfrac{\partial f}{\partial x} = (y - 2x^2 y)e^{y-x^2}$; $\quad \dfrac{\partial f}{\partial y} = (x + xy)e^{y-x^2}$.

Проверяем необходимое условие экстремума:

$$\begin{cases} y(1 - 2x^2) = 0 \\ x(1 + y) = 0 \end{cases} \Longrightarrow \begin{cases} y = 0 \\ x = 0 \end{cases} \text{ или } \begin{cases} x = \pm 1/\sqrt{2} \\ y = -1 \end{cases}.$$



Найдены три критические точки: $O(0;0)$ — седловая, поскольку $f(O) = 0$, в I и III четвертях $f > 0$, а во II и IV четвертях $f < 0$; $A(-1/\sqrt{2};-1)$ — максимум, поскольку только в ней может достигаться максимальное значение f на компакте $K = [-2;0] \times [-4;0]$: на его правой и верхней границах $f = 0$, на левой и нижней $f \leq 8e^{-4} < e^{-3/2} = f(A)$, так что $\max\limits_K f$ достигается во внутренней точке K, но все остальные внутренние точки K, кроме A, не экстремумы. Точка $B(1/\sqrt{2};-1)$ — минимум (аналогично точке A). Как видим, классифицировать критические точки на (нестрогие) максимумы/минимумы и седловые весьма непросто. Нам нужны удобные

Достаточные условия экстремума

Теорема 8.2. *Пусть A — внутренняя точка $D(f) \subset \mathbb{R}^n$; скалярная функция f дважды дифференцируема в окрестности точки A, $\nabla f(A) = \vec{0}$, и вторые частные производные непрерывны в A. Пусть $\lambda_1,\ldots,\lambda_n$ — собственные значения матрицы $\mathbf{H} = \mathbf{H}_f(A)$. Тогда:*

1. *Если $\forall\lambda_i > 0$, то A — минимум;*

2. *Если $\forall\lambda_i < 0$, то A — максимум;*

3. *Если $\forall\lambda_i \geq 0$, $\exists\lambda_j > 0$, то A — (нестрогий) минимум или седловая;*

4. *Если $\forall\lambda_i \leq 0$, $\exists\lambda_j < 0$, то A — (нестрогий) максимум или седловая;*

5. *Если $\exists\lambda_i > 0$, $\exists\lambda_j < 0$, то A — седловая;*

6. *Если $\forall\lambda_i = 0$, то возможны все случаи.*

Доказательство. В окрестности точки A разложим функцию f по формуле Тейлора до второй степени, учитывая $\nabla f(A) = \vec{0}$:

$$f(P) = f(A) + \frac{1}{2}\vec{AP}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{AP} + r(P), \qquad r(P) = o(|\vec{AP}|^2).$$

Пусть $\{\vec{v_i}\}_{i=1}^n$ — ортонормированный базис $\mathbb{R}^n$ из собственных векторов матрицы $\mathbf{H}$; $\mathbf{H}\vec{v_i} = \lambda_i\vec{v_i}$. Разложим $\vec{AP} = \sum\limits_{i=1}^n c_i\vec{v_i}$, тогда $f(P) - f(A) = \dfrac{1}{2}\sum\limits_{i=1}^n \lambda_i\, c_i^2 + r(P)$.

1. Пусть $0 < \lambda_1 \leq \ldots \leq \lambda_n$. Найдём такое $\varepsilon > 0$, что $\forall P \in \overset{o}{U}_\varepsilon(A)$ $|r(P)| < \dfrac{\lambda_1}{4}|\vec{AP}|^2$. Тогда $\forall P \in \overset{o}{U}_\varepsilon(A)$

$$f(P) > f(A) + \frac{1}{2}\sum_{i=1}^{n}\lambda_i\ c_i^2 - \frac{\lambda_1}{4}|\vec{AP}|^2 \geq f(A) + \frac{\lambda_1}{2}\underbrace{\sum_{i=1}^{n}c_i^2}_{|\vec{AP}|^2} - \frac{\lambda_1}{4}|\vec{AP}|^2 > f(A).$$

2. Аналогично.

3. Пусть $\forall\lambda_i \geq 0$; $\lambda_1 > 0$. Надо доказать, что невозможен (нестрогий) максимум. Найдём такое $\varepsilon > 0$, что $\forall P \in \overset{o}{U}_\varepsilon(A)$ $|r(P)| < \dfrac{\lambda_1}{4}|\vec{AP}|^2$. Рассмотрим такие точки P_t, что $\vec{AP_t} = t\vec{v_1}$, $0 < t < \varepsilon$. Тогда

$$f(P_t) > f(A) + \frac{1}{2}\lambda_1\ t^2 - \frac{\lambda_1}{4}\underbrace{|\vec{AP_t}|^2}_{t^2} > f(A).$$

Но точки P_t встречаются в любой окрестности A.

4. Аналогично доказывается, что невозможен (нестрогий) минимум.

5. Поскольку $\exists\lambda_i > 0$, невозможен (нестрогий) максимум. С другой стороны, $\exists\lambda_j < 0$ $\Longrightarrow$ невозможен (нестрогий) минимум. Значит, точка A седловая.

6. Ничего не утверждается — доказывать нечего. $\square$

Пример 8.2. *Исследуем на экстремум функцию* $f(x,y,z) = (x^2+y^2-z^2)e^{x-2z}$.

$$\frac{\partial f}{\partial x} = (x^2+2x+y^2-z^2)e^{x-2z};\quad \frac{\partial f}{\partial y} = 2ye^{x-2z};\quad \frac{\partial f}{\partial z} = 2(z^2-z-x^2-y^2)e^{x-2z}.$$

Ищем критические точки:

$$\begin{cases} 2x &= z^2-x^2-y^2 \\ y &= 0 \\ z &= z^2-x^2-y^2 \end{cases} \Longrightarrow \begin{cases} 2x &= 3x^2 \\ y &= 0 \\ z &= 2x \end{cases} \Longrightarrow \begin{array}{l} O(0;0;0) \\ A\left(\dfrac{2}{3};0;\dfrac{4}{3}\right) \end{array}$$

Вычислим матрицу Гессе:

$$\mathbf{H}_f(x,y,z) = e^{x-2z}\begin{bmatrix} 2+4x+Q & 2y & -2z-4x-2Q \\ 2y & 2 & -4y \\ -2z-4x-2Q & -4y & -2+8z+4Q \end{bmatrix},$$

где $Q = x^2+y^2-z^2$. В найденных критических точках имеем

$$\mathbf{H}_f(O) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix};\qquad \mathbf{H}_f(A) = \frac{e^{-2}}{3}\begin{bmatrix} 10 & 0 & -8 \\ 0 & 6 & 0 \\ -8 & 0 & 10 \end{bmatrix}$$

Собственные значения $\mathbf{H}_f(O)$: $\lambda_{1,2} = 2$, $\lambda_3 = -2 \Longrightarrow O$ — седловая точка; квадратичная форма с матрицей $\mathbf{H}_f(A)$ положительно определена по критерию Сильвестра:

$$10 > 0, \qquad \begin{vmatrix} 10 & 0 \\ 0 & 6 \end{vmatrix} > 0, \qquad \begin{vmatrix} 10 & 0 & -8 \\ 0 & 6 & 0 \\ -8 & 0 & 10 \end{vmatrix} > 0,$$

следовательно, A — точка минимума.

9 Дифференцирование неявных ФНП

Теорема 9.1. (об обратной функции) *Пусть $A \in \mathbb{R}^n$ — внутренняя точка области определения функции F, принимающей значения в $\mathbb{R}^n$. Пусть F дифференцируема в окрестности $U_R(A)$, и её частные производные непрерывны в A. Если матрица Якоби в точке A невырождена* ($\det J_F(A) \neq 0$), *то существует непрерывная функция Φ, определённая в некоторой окрестности U точки $B = F(A)$, такая, что*

$$\forall Y \in U, \quad \forall X \in \Phi(U): \qquad \Phi(Y) = X \Longleftrightarrow F(X) = Y,$$

т.е. функция Φ — обратная к сужению функции F на область $\Phi(U)$. Кроме того, Φ дифференцируема в B, и

$$J_\Phi(B) = \big(J_F(A)\big)^{-1}.$$

Пример 9.1. Пусть $n = 2$; $F(x, y) = (x + y; xy)$. Матрица Якоби

$$J_F(x,y) = \begin{bmatrix} \dfrac{\partial F_1}{\partial x} & \dfrac{\partial F_1}{\partial y} \\ \dfrac{\partial F_2}{\partial x} & \dfrac{\partial F_2}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ y & x \end{bmatrix}$$

невырождена при $y \neq x$. Частные производные непрерывны во всех точках, поэтому теорема 9.1 применима в любой точке $(x; y)$, $y \neq x$. Возьмём точку $A(1; 2)$. Её образ — точка $B = F(A) = (3; 2)$. В некоторой окрестности точки B определена обратная к F функция Φ, для которой

$$J_\Phi(x+y, xy) = \big(J_F(x,y)\big)^{-1} = \frac{1}{x-y}\begin{bmatrix} x & -1 \\ -y & 1 \end{bmatrix}; \quad J_\Phi(3,2) = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}.$$

(В данном примере для Φ удаётся найти явную формулу:

$$\Phi(u, v) = \big(\frac{1}{2}(u - \sqrt{u^2 - 4v}); \frac{1}{2}(u + \sqrt{u^2 - 4v})\big).$$

Так что вычислить матрицу Якоби для Φ можно и непосредственно. Но явно выразить обратную функцию удаётся далеко не всегда! Даже если удаётся, то её непосредственное дифференцирование может оказаться более трудоёмким, чем нахождение обратной матрицы к J_F.)

Теорема 9.2. (о неявно заданной скалярной функции). *Пусть $A \in \mathbb{R}^n$ — внутренняя точка области определения скалярной функции F; $F(A) = 0$. Пусть F дифференцируема в окрестности $U_R(A)$, и её частные производные непрерывны в A. Если при некотором $k \in \{1, \dots, n\}$ $\dfrac{\partial f}{\partial x_k}(A) \neq 0$, то существует $\varepsilon \in (0; R]$ и непрерывная функция f, определённая в некотором открытом множестве $U \subset \mathbb{R}^{n-1}$,* *содержащем точку A' (проекцию A на $(n-1)$-мерное пространство, ортогональное k-ой оси), такая, что*

$$\{F=0\}\cap U_\varepsilon(A)=\{x_k=f\underbrace{(x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)}_{\in U}\}.$$

Кроме того, функция f дифференцируема в A', и

$$\frac{\partial f}{\partial x_i}(A')=-\ \frac{\partial F/\partial x_i(A)}{\partial F/\partial x_k(A)}\quad \forall i\neq k.$$

Пример 9.2. Дано уравнение

$$F(x,y,z)=x^2+y^2+z^2+x-2y+3z-5=0.$$

Градиент F равен $\nabla F(x,y,z)=(2x+1;2y-2;2z+3)$. В точке $A(1;1;1)$ имеем $\nabla F(A)=(3;0;5)\Longrightarrow$ мы можем в окрестности A представить $x=X(y,z)$ или $z=Z(x,y)$. Выберем вторую возможность. Тогда $Z(1,1)=1$; по теореме 9.2 в точке A и в точках её окрестности, где выполнено условие $\partial F/\partial z>0$, существуют частные производные

$$\frac{\partial Z}{\partial x}=-\frac{\partial F/\partial x}{\partial F/\partial z}=-\ \frac{2x+1}{2z+3};\qquad \frac{\partial Z}{\partial y}=-\frac{\partial F/\partial y}{\partial F/\partial z}=-\ \frac{2y-2}{2z+3}.$$

Теорема 9.3. (о неявно заданной векторной функции). *Пусть $A\in\mathbb{R}^n$ — внутренняя точка области определения k-мерно-значной функции $F=(f_1,\ldots,f_k)$, $1<k<n$, причём $F(A)=(0,\ldots,0)$. Пусть F дифференцируема в окрестности $U_R(A)$, и её частные производные непрерывны в A. Предположим, что ранг матрицы $J_F(A)$ равен k, т. е. $\nabla f_1(A),\ldots,\nabla f_k(A)$ линейно независимы. Пусть столбцы $J_F(A)$ с номерами $j_1,\ldots,j_k$ образуют базисный минор; номера остальных столбцов обозначим $i_1,\ldots,i_{n-k}$. Тогда $\exists\ \varepsilon\le R$ и непрерывная k-мерно-значная функция G, определённая в некотором открытом множестве $U\subset\mathbb{R}_{n-k}$, содержащем точку A' (проекцию A на $(n-k)$-мерное пространство, ортогональное j_1-ой,…,j_k-ой осям), такая, что*

$$\{F=\vec{0}\}\cap U_\varepsilon(A)=\left\{\begin{bmatrix}x_{j_1}\\ \vdots\\ x_{j_k}\end{bmatrix}=G\underbrace{(x_{i_1},\ldots,x_{i_{n-k}})}_{\in U}\right\}.$$

Кроме того, функция $G=(g_1,\ldots,g_k)$ дифференцируема в A', и её частные производные находятся из системы линейных уравнений

$$\frac{\partial f_m}{\partial x_{i_q}}(A)+\sum_{p=1}^{k}\frac{\partial f_m}{\partial x_{j_p}}(A)\frac{\partial g_p}{\partial x_{i_q}}(A')=0,\qquad \begin{matrix}m=1,\ldots,k\\ q=1,\ldots,n-k\end{matrix}\tag{5}$$

(всего $k(n-k)$ уравнений; искомых частных производных столько же).

Правило для запоминания формулы (5): продифференцируем

$$df_m = \sum_{i=1}^{n} \frac{\partial f_m}{\partial x_i} dx_i = 0, \qquad m = 1, \ldots, k,$$

и распишем $dx_{j_1}, \ldots, dx_{j_k}$ как дифференциалы функций от остальных $n-k$ переменных. Получившиеся дифференциалы k сложных функций от $n-k$ переменных приравняем к 0 покомпонентно.

Пример 9.3. Дана система уравнений

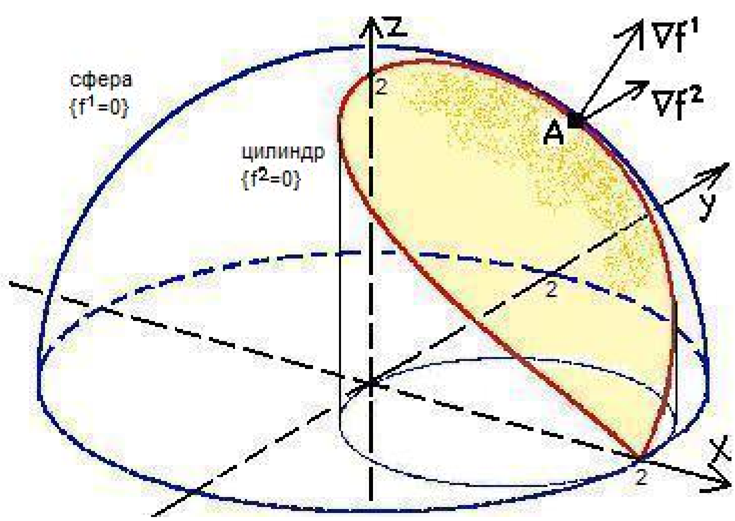
$$\begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} x^2+y^2+z^2-4 \\ x^2-2x+y^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \qquad J_F = \begin{bmatrix} 2x & 2y & 2z \\ 2x-2 & 2y & 0 \end{bmatrix}.$$

В точке $A(1;1;\sqrt{2})$ и в её окрестности первые 2 столбца J_F могут служить её базисным минором $\Longrightarrow$ x и y локально выражаются через z: $x = X(z)$, $y = Y(z)$. Распишем дифференциалы F как сложной функции:

$$\begin{cases} df_1 = 2x\dfrac{dX}{dz}dz + 2y\dfrac{dY}{dz}dz + 2zdz & = 0 \\ df_2 = (2x-2)\dfrac{dX}{dz}dz + 2y\dfrac{dY}{dz}dz & = 0 \end{cases} \Longrightarrow \begin{cases} x\dfrac{dX}{dz} + y\dfrac{dY}{dz} & = -z \\ (x-1)\dfrac{dX}{dz} + y\dfrac{dY}{dz} & = 0 \end{cases}$$

Отсюда $\dfrac{dX}{dz} = -z$, $\dfrac{dY}{dz} = \dfrac{(x-1)z}{y}$ в окрестности точки A;

в самой точке A: $\dfrac{dX}{dz}(\sqrt{2}) = -\sqrt{2}$, $\dfrac{dY}{dz}(\sqrt{2}) = 0$.



10 Гладкие k-мерные поверхности (ГkМП)

Определение 10.1. *Множество $\Sigma \subset \mathbb{R}_n$ называется простой ГkМП, если оно является образом $\Sigma = \vec{r}(U)$ связного открытого множества $U \subset \mathbb{R}^k$ под действием инъективного отображения $\vec{r} : U \mapsto \mathbb{R}^n$, причём $\vec{r}$ имеет непрерывные частные производные, и его матрица Якоби имеет всюду ранг k (т.е. $\partial\vec{r}/\partial u_1, \ldots, \partial\vec{r}/\partial u_k$ линейно независимы в каждой точке $(u_1, \ldots, u_k) \in U$).*
Связное множество $\Sigma \subset \mathbb{R}^n$ называется **ГkМП**, *если $\forall A \in \Sigma \;\; \exists\varepsilon > 0 :$ множество $\Sigma \cap U_\varepsilon(A)$ является простой ГkМП.*

Пример 10.1. (Г2МП в $\mathbb{R}^3$). 1) Половина конуса $x^2+y^2 = z^2$, $z > 0$ — простая Г2МП, поскольку она имеет вид

$$x = u,\; y = v,\; z = \sqrt{u^2 + v^2};\quad (u; v) \in U = \mathbb{R}^2 \backslash \{O\}.$$

2) Сфера $S = \{x^2 + y^2 + z^2 = 1\}$ не является простой Г2МП. Но полусфера — простая Г2МП: например,

$$S \cap \{x > 0\} = \left\{ (\sqrt{1 - u^2 - v^2}; u; v) :\; (u; v) \in U = \{u^2 + v^2 < 1\} \right\}.$$

Любая точка S лежит в некоторой полусфере $\Longrightarrow$ сфера — Г2МП.

Важнейшим частным случаем простой ГkМП является *график* m-мерно-значной функции от k переменных, $m = n - k$, имеющей непрерывные частные производные, т. е. множество вида

$$\Gamma_G = \left\{ \begin{bmatrix} x_{j_1} \\ \vdots \\ x_{j_m} \end{bmatrix} = G(x_{i_1}, \ldots, x_{i_k}) : \quad (x_{i_1}, \ldots, x_{i_k}) \in U \right\}. \tag{6}$$

Теорема 10.1. *Если $\Sigma \subset \mathbb{R}^n$ — ГkМП, то $\forall A \in \Sigma \;\; \exists\varepsilon > 0 :$ множество $\Sigma \cap U_\varepsilon(A)$ имеет вид (6) при некотором разбиении n переменных на два набора:* $\{x_{j_\alpha}\}_{\alpha=1}^m$ *и* $\{x_{i_\beta}\}_{\beta=1}^k$.

Другой способ задания поверхности — неявный (через уравнения). Следующая теорема (следствие теорем 9.2 и 9.3) даёт достаточные условия того, чтобы поверхность получилась гладкой.

Теорема 10.2. *Пусть $\Omega \subset \mathbb{R}^n$ — открытое связное множество; на нём определена функция $F : \Omega \mapsto \mathbb{R}^m$, $1 \le m < n$. Пусть во всех точках Ω F дифференцируема, её частные производные непрерывны, и ранг J_F равен m. Тогда связные множества уровня функции F являются ГkМП при $k = n - m$.*

Пример 10.2. $F(x, y, z) = x^2 + y^2 - z^2$ удовлетворяет условиям теоремы на $\Omega = \mathbb{R}^3 \backslash \{O\}$. Получается семейство Г2МП: однополостные гиперболоиды, половины конуса и половины двуполостных гиперболоидов.

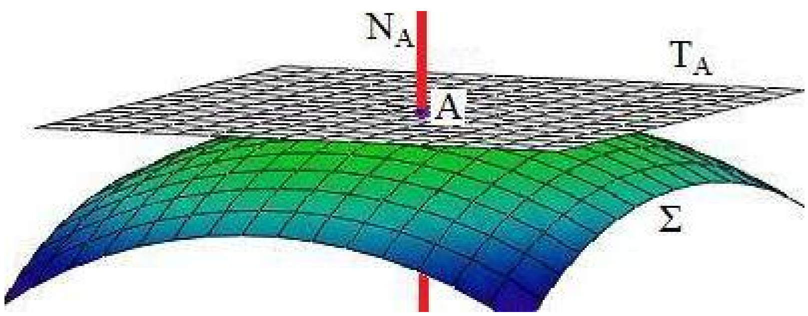
Пример 10.3. Пусть $F(x, y, z) = \begin{bmatrix} x^2 + y^2 + z^2 \\ z \end{bmatrix}$. Её матрица Якоби

$$J_F(x, y, z) = \begin{pmatrix} 2x & 2y & 2z \\ 0 & 0 & 1 \end{pmatrix}$$

непрерывна и имеет ранг 2 на $\Omega = \mathbb{R}^3 \backslash Oz$. Множества уровня F – горизонтальные окружности в Ω — являются Г1МП.

Касательные и нормали

Определение 10.2. *Пусть $\Sigma \subset \mathbb{R}^n$ — ГkМП; $A \in \Sigma$.*
Вектор $\vec{V} \in \mathbb{R}^n$ называется **касательным** *в точке A, если $\exists$ гладкая кривая $\ell \subset \Sigma$, $\ell \ni A$, для которой $\vec{V}$ – касательный вектор в точке A.*
Касательная плоскость T_A *в точке A — плоскость, проходящая через A и содержащая все касательные вектора.*
Нормальная плоскость N_A *в точке A — плоскость, проходящая через A и содержащая все вектора, ортогональные всем касательным.*



Если поверхность задана явно, для неё легко строится касательная плоскость:

Теорема 10.3. *Пусть ГkМП Σ в окрестности точки A имеет вид $\vec{r}(U)$; $A = \vec{r}(B)$. Тогда* $\dim T_A = k$ *(соответственно,* $\dim N_A = n - k$*); базисом T_A является набор векторов $\{\partial\vec{r}/\partial u_i\}_{i=1}^k$.*

Доказательство. Пусть через $A = \vec{r}(B)$ проходит гладкая $\ell = \{(x_1(t), \ldots, x_n(t)) : |t| \le 1\}$; $x_i(0) = a_i$. Тогда прообраз ℓ в U — тоже гладкая кривая в окрестности B, в частности, существуют $du_i/dt(0)$. Следовательно, касательный вектор к ℓ в точке A имеет вид $\vec{V} = \sum\limits_{i=1}^{k} \frac{du_i}{dt}(0)\frac{\partial\vec{r}}{\partial u_i}$. □

Если же поверхность задана неявно, то легче построить сначала нормальную плоскость:

Теорема 10.4. *Пусть ГkМП Σ задана неявно*: $\Sigma = \{P \in \Omega : \ F(P) = \vec{0}\}$; $A \in \Sigma$. *Тогда базисом N_A является $\{\nabla F_1(A), \ldots, \nabla F_m(A)\}$.*

Доказательство. Пусть гладкая кривая $\ell \subset \Sigma$, $\ell \ni A$. Пусть $\vec{\tau}$ — касательный вектор к ℓ в точке A; $|\vec{\tau}| = 1$. Поскольку $F_i = const$ на Σ, получаем $\dfrac{\partial F_i}{\partial \vec{\tau}}(A) = 0$, т. е. $\nabla F_i(A) \perp \vec{\tau}$. Таким образом, линейная оболочка $\{\nabla F_1(A), \ldots, \nabla F_m(A)\}$ лежит в N_A. Но $\dim N_A = m \Longrightarrow$ других векторов там нет. $\square$

Теперь посмотрим, как выглядят касательные и нормальные плоскости для поверхностей в $\mathbb{R}^3$.

Г2МП. Касательные плоскости двумерные, нормали — одномерные.

1. Заданная явно. $\Sigma = \{(x(u,v), y(u,v), z(u,v)) : \ (u,v) \in U\}$. Тогда

$$T_A = \operatorname{span}\left\{ \begin{bmatrix} \partial x/\partial u \\ \partial y/\partial u \\ \partial z/\partial u \end{bmatrix}; \begin{bmatrix} \partial x/\partial v \\ \partial y/\partial v \\ \partial z/\partial v \end{bmatrix} \right\}; \quad N_A = \operatorname{span}\left\{ \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \partial x/\partial u & \partial y/\partial u & \partial z/\partial u \\ \partial x/\partial v & \partial y/\partial v & \partial z/\partial v \end{vmatrix} \right\}.$$

2. Заданная неявно. $\Sigma = \{F(x,y,z) = C\}$. Вектор $\nabla F(A)$ — направляющий для прямой N_A и нормальный для плоскости T_A.

Г1МП. Касательные одномерные, нормальные плоскости двумерные.

1. Заданная явно. $\ell = \{(x(u), y(u), z(u)) : \ u \in U\}$. Пусть для точки A $u = a$. Вектор $\begin{bmatrix} dx/du \\ dy/du \\ dz/du \end{bmatrix}(a)$ является направляющим для прямой T_A и нормальным для плоскости N_A.

2. Заданная неявно. $\ell = \left\{ \begin{array}{c} F(x,y,z) = C_1 \\ G(x,y,z) = C_2 \end{array} \right\}$.

Базис N_A: $\{\nabla F(A); \nabla G(A)\}$. Направляющий для T_A: $\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \partial F/\partial x & \partial F/\partial y & \partial F/\partial z \\ \partial G/\partial x & \partial G/\partial y & \partial G/\partial z \end{vmatrix}$

Пример 10.4. Рассмотрим кривую из примера 9.3. Дана система уравнений

$$\begin{bmatrix} F(x,y,z) \\ G(x,y,z) \end{bmatrix} = \begin{bmatrix} x^2 + y^2 + z^2 - 4 \\ x^2 - 2x + y^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

В точке $A(1;1;\sqrt{2})$ базисом нормальной плоскости будет пара векторов $\nabla F(A) = (2;2;2\sqrt{2})$ и $\nabla G(A) = (0;2;0)$. Их векторное произведение равно $4(-\sqrt{2};0;1)$, отсюда получаем уравнения нормальной плоскости и касательной:

$$N_A = \{-\sqrt{2}(x-1) + (z-\sqrt{2}) = 0\}; \qquad T_A = \left\{ \frac{x-1}{-\sqrt{2}} = \frac{y-1}{0} = \frac{z-\sqrt{2}}{1} \right\}.$$

11 Условные экстремумы

Определение 11.1. *Пусть* $\Omega \subset \mathbb{R}^n$ *— открытое множество,* $f\colon \Omega \mapsto \mathbb{R}$; $u\colon \Omega \mapsto \mathbb{R}^k$, $1 \le k < n$. *Пусть* $A \in \Omega$, $u(A) = \vec{0}$. *Говорят, что* A *–*
точка **условного максимума** (*при условии* $u = \vec{0}$), *если*

$$\exists \varepsilon > 0 : \ \forall P \in \overset{o}{U}_\varepsilon(A) \cap \{u = \vec{0}\} \qquad f(P) < f(A);$$

точка **условного минимума**, *если*

$$\exists \varepsilon > 0 : \ \forall P \in \overset{o}{U}_\varepsilon(A) \cap \{u = \vec{0}\} \qquad f(P) > f(A);$$

точка **нестрогого условного максимума**, *если*

$$\exists \varepsilon > 0 : \ \forall P \in \overset{o}{U}_\varepsilon(A) \cap \{u = \vec{0}\} \qquad f(P) \le f(A);$$

точка **нестрогого условного минимума**, *если*

$$\exists \varepsilon > 0 : \ \forall P \in \overset{o}{U}_\varepsilon(A) \cap \{u = \vec{0}\} \qquad f(P) \ge f(A).$$

Далее будем предполагать, что функция u дифференцируема и имеет непрерывные частные производные в Ω, причём ранг матрицы J_u всюду равен k (в случае $k = 1$ это означает $\nabla u \ne \vec{0}$). Тогда множество $\Sigma = \Omega \cap \{u = \vec{0}\}$ состоит из $\Gamma(n-k)$МП по теореме 10.2.

Необходимые условия условного экстремума. Функции Лагранжа

Теорема 11.1. *Пусть* $A \in \Sigma$. *Пусть функция* f *дифференцируема в точке* A. *Если* A *— точка (нестрогого) условного экстремума, то существуют* $\lambda_1, \dots, \lambda_k$:

$$\nabla f(A) = \sum_{i=1}^{k} \lambda_i \nabla u_i(A).$$

Доказательство. Поскольку вектора $\nabla u_1(A), \dots, \nabla u_k(A)$ образуют базис плоскости N_A, нам нужно доказать, что $\nabla f(A) \in N_A$. Предположим, что это не так. Тогда найдётся касательный вектор $\vec{\tau}$: $\nabla f(A) \cdot \vec{\tau} > 0$. Пусть гладкая кривая $\ell = \{\vec{r}(t) : |t| < 1\} \subset \Sigma$; $\vec{r}(0) = A$; $\dfrac{d}{dt}\vec{r}(0) = \vec{\tau}$. Тогда для сложной функции $f \circ \vec{r}$ получаем:

$$\frac{d}{dt} f(\vec{r}(t))\bigg|_{t=0} = \nabla f(\vec{r}(0)) \cdot \frac{d\vec{r}}{dt}(0) = \nabla f(A) \cdot \vec{\tau} > 0.$$

Следовательно, $\exists \varepsilon \in (0; 1]$: $f(\vec{r}(t)) > 0$ при $0 < t < \varepsilon$, $f(\vec{r}(t)) < 0$ при $-\varepsilon < t < 0$. Но поскольку $\vec{r}(t) \in \Sigma$, точка A не может быть (нестрогим) условным экстремумом. □

Следствие 11.1. *Пусть* $A \in \Sigma$; *функция* f *дифференцируема в точке* A. *Если* A *— точка (нестрогого) условного экстремума, то существует* **функция Лагранжа**

$$\mathcal{L}(x_1, \dots, x_n) = f(x_1, \dots, x_n) - \sum_{i=1}^{k} \lambda_i \ u_i(x_1, \dots, x_n),$$

такая, что $\nabla \mathcal{L}(A) = \vec{0}$. *Очевидно,* $\mathcal{L} = f$ *на* Σ.

Пример 11.1. *Найдём условные экстремумы функции* $f(x, y) = y/x$ *при условии* $u(x, y) = x^2 + y^2 - 6x - 2y + 5 = 0$. Здесь $\Omega = \{x > 0\}$; Σ — окружность радиуса $\sqrt{5}$ с центром $(3; 1)$. Функция Лагранжа имеет вид

$$\mathcal{L}(x, y) = \frac{y}{x} - \lambda(x^2 + y^2 - 6x - 2y + 5).$$

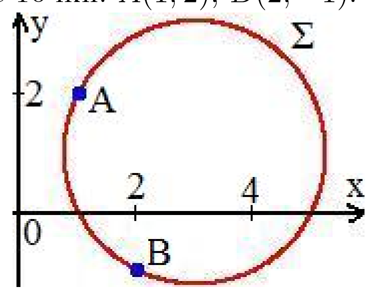
Решаем систему уравнений:

$$\begin{cases} \nabla\mathcal{L}(x, y) = \vec{0} \\ u(x, y) = 0 \end{cases} \iff \begin{cases} -y/x^2 - \lambda(2x - 6) = 0 & (1) \\ 1/x - \lambda(2y - 2) = 0 & (2) \\ x^2 + y^2 - 6x - 2y + 5 = 0 & (3) \end{cases}$$

Из (2) $\Longrightarrow \lambda = \big(2x(y-1)\big)^{-1}$. Домножим (1) на x, подставив λ:

$$-\frac{y}{x} = \frac{x-3}{y-1} \Longrightarrow -y(y-1) = x(x-3) \iff x^2 + y^2 - 3x - y = 0.$$

Вычтя отсюда (3), получим $y = 5 - 3x$. Подставив такой y в (3), получим $10x^2 - 30x + 20 = 0 \Longrightarrow x_1 = 1$, $x_2 = 2$. Итак, найдены две точки: $A(1; 2)$, $B(2; -1)$.



Поскольку Σ — компакт, $\min_\Sigma f$ и $\max_\Sigma f$ должны достигаться. Сравнив значения $f(A)$ и $f(B)$, получаем, что A — точка условного максимума, B — точка условного минимума.

Если мы вычислим $\lambda = \dfrac{1}{2x(y-1)}$, то получим $\lambda = 1/2$ для A и $\lambda = -1/8$ для B. Так что для одних и тех же f и u могут быть разные функции Лагранжа, соответствующие разным критическим точкам.

Пример 11.2. *В эллипсоид* $\{x^2+2y^2+5z^2 = 1\}$ *вписать прямоугольный параллелепипед с рёбрами, параллельными координатным осям, имеющий максимальную площадь поверхности.*

Рекомендую рассмотреть на семинаре.

Ответ: $S_{\max} = 1$.

Достаточные условия условного экстремума

Теорема 11.2. Пусть f, u дважды дифференцируемы *в окрестности точки* A, *и вторые частные производные непрерывны в* A. *Пусть существует функция Лагранжа* $\mathcal{L} = f - \sum_{i=1}^k \lambda_i u_i$, *такая, что* $\nabla\mathcal{L}(A) = \vec{0}$. *Обозначим* $\mathbf{H} = \mathbf{H}_{\mathcal{L}}(A)$. *Тогда:*

1. *Если для* $\forall$ *касательного вектора* $\vec{\tau} \neq \vec{0}$ $\quad \vec{\tau}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{\tau} > 0$, *то* A *— условный минимум;*

2. *Если* $\forall \vec{\tau} \neq \vec{0}$ $\quad \vec{\tau}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{\tau} < 0$, *то* A *— условный максимум;*

3. *Если* $\forall \vec{\tau}$ $\quad \vec{\tau}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{\tau} \geq 0$, $\quad \exists \vec{\tau}$ $\quad \vec{\tau}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{\tau} > 0$, *то* A *— (нестрогий) условный минимум или условная седловая точка;*

4. *Если* $\forall \vec{\tau}$ $\quad \vec{\tau}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{\tau} \leq 0$, $\quad \exists \vec{\tau}:$ $\quad \vec{\tau}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{\tau} < 0$, *то* A *— (нестрогий) условный максимум или условная седловая;*

5. *Если* $\exists \vec{\tau}_1, \vec{\tau}_2$*:* $\quad \vec{\tau}_1^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{\tau}_1 < 0 < \vec{\tau}_2^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{\tau}_2$, *то* A *— условная седловая;*

6. *Если* $\forall \vec{\tau}$ $\quad \vec{\tau}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{\tau} = 0$, *то возможны все случаи.*

Доказательство. Поскольку $\mathcal{L} = f$ на поверхности Σ, мы будем рассматривать функцию $\mathcal{L}$ вместо f. По теореме 10.1 в окрестности точки A можно представить k переменных $x_{j_1},\ldots,x_{j_k}$ (для простоты обозначений предположим, что это $x_1,\ldots,x_k$) как дифференцируемые функции от остальных $n-k$ переменных. Тогда уравнение касательной плоскости T_A будет выглядеть так:

$$x_j - a_j = \sum_{i=k+1}^{n} \frac{\partial x_j}{\partial x_i}(A')(x_i - a_i), \qquad 1 \leq j \leq k < i \leq n.$$

Следовательно, $\varrho(X, T_A) = o(|\vec{AX}|)$, т.е. $\forall \varepsilon > 0$ $\exists \delta > 0$: если $X \in \Sigma \cap \overset{o}{U}_\delta(A)$, то $\varrho(X, T_A) < \varepsilon |\vec{AX}|$.
Обозначим $\mathbf{H} = \mathbf{H}_{\mathcal{L}}(A)$. Выпишем формулу Тейлора для функции $\mathcal{L}$, учитывая, что $\nabla \mathcal{L}(A) = \vec{0}$:

$$\mathcal{L}(X) - \mathcal{L}(A) = \frac{1}{2} \vec{AX}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{AX} + o(|\vec{AX}|^2).$$

Пусть Y — ортогональная проекция точки X на плоскость T_A. Тогда при $X \to A$ имеем $|\vec{AY}| \sim |\vec{AX}|$, $|\vec{YX}| = o(|\vec{AX}|)$, и получаем

$$\begin{aligned} \mathcal{L}(X) - \mathcal{L}(A) &= \frac{1}{2}(\vec{AY} + \vec{YX})^{\mathrm{T}} \cdot \mathbf{H} \cdot (\vec{AY} + \vec{YX}) + o(|\vec{AX}|^2) = \\ &= \frac{1}{2} \vec{AY}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{AY} + \vec{AY}^{\mathrm{T}} \cdot \underbrace{\mathbf{H} \cdot \vec{YX}}_{o(|\vec{AX}|)} + \frac{1}{2} \underbrace{\vec{YX}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{YX}}_{o(|\vec{AX}|^2)} + o(|\vec{AX}|^2) = \\ &= \frac{1}{2} \vec{AY}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{AY} + o(|\vec{AY}|^2). \end{aligned}$$

Рассмотрим квадратичную форму, порождённую $\mathbf{H}$ на $(n-k)$-мерной плоскости T_A; её собственные значения обозначим в порядке возрастания $\mu_1, \ldots, \mu_{n-k}$ (эти значения не обязаны быть подмножеством собственных значений $\mathbf{H}$ на $\mathbb{R}^n$).

1. Если $\mu_1 > 0$, то $\mathcal{L}(X) - \mathcal{L}(A) \geq \dfrac{1}{2}\ \mu_1 |\vec{AY}|^2 + o(|\vec{AY}|^2) > 0$ при $\forall X \in \overset{o}{U}_\delta(A) \cap \Sigma$ для некоторого $\delta > 0$, т. е. A — точка условного минимума.

2. Случай $\mu_{n-k} < 0$ рассматривается аналогично.

3. Если $\mu_1 = 0$, но $\mu_{n-k} = L > 0$, то выберем единичный касательный вектор $\vec{\tau}$, для которого $\vec{\tau}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{\tau} = L$, и рассмотрим гладкую кривую $\ell = \{X_t : \ |t| < 1\} \subset \Sigma$; $X_0 = A$; $\dfrac{d}{dt}X_t(0) = \vec{\tau}$. Пусть $Y_t \in T_A$ – такая точка, что $\vec{AY_t} = t\vec{\tau}$, тогда $|\vec{Y_tX_t}| = o(t)$. Получаем

$$\mathcal{L}(X_t) - \mathcal{L}(A) = \frac{t^2}{2}\ \vec{\tau}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{\tau} + t\ \vec{\tau}^{\mathrm{T}} \cdot \underbrace{\mathbf{H} \cdot \vec{Y_tX_t}}_{o(t)} + \frac{1}{2}\ \underbrace{\vec{Y_tX_t}^{\mathrm{T}} \cdot \mathbf{H} \cdot \vec{Y_tX_t}}_{o(t^2)} + o(t^2)$$

$= t^2(L/2 + o(1)) > 0$ при $|t| < \delta$. Следовательно, в точке A невозможен (нестрогий) условный максимум.

4. Если $\mu_{n-k} = 0$, но $\mu_1 < 0$, то аналогичные рассуждения показывают, что невозможен (нестрогий) условный минимум.

5. Если $\mu_1 < 0 < \mu_{n-k}$, то невозможен никакой (нестрогий) условный экстремум $\Longrightarrow$ точка условная седловая. $\square$

Пример 11.3. *Исследуем функцию* $f(x,y,z) = (x-3)^2 + y^2 + z^2$ *при условии* $u(x,y,z) = x^2 + y^2 + 4z^2 - 25 = 0$.

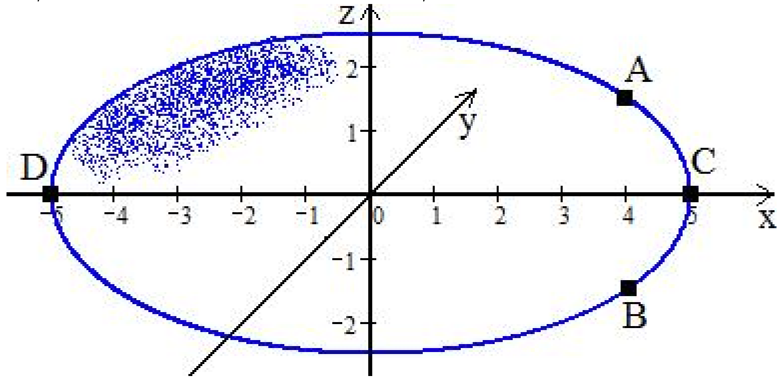
Функция Лагранжа имеет вид

$$\mathcal{L}(x,y,z) = (x-3)^2 + y^2 + z^2 - \lambda(x^2 + y^2 + 4z^2 - 25).$$

Решаем систему уравнений:

$$\left\{ \begin{array}{r} \nabla\mathcal{L}(x,y,z) = \vec{0} \\ u(x,y,z) = 0 \end{array} \right. \iff \left\{ \begin{array}{cll} x - 3 - \lambda x & = 0 & (1) \\ y - \lambda y & = 0 & (2) \\ z - 4\lambda z & = 0 & (3) \\ x^2 + y^2 + 4z^2 & = 25 & (4) \end{array} \right.$$

Если $\lambda = 1$, то (1) невыполнимо.
Если $\lambda = 1/4$, то $x = 4$, $y = 0$, и из (4) находим $z = \pm 3/2$.
Если $1 \neq \lambda \neq 1/4$, то $x = 3/(1-\lambda)$, $y = z = 0$, тогда (4)$\Longrightarrow$ $x = \pm 5$; если $x = 5$, то $\lambda = 2/5$, а если $x = -5$, то $\lambda = 8/5$.



Найдено 4 критические точки. Матрица Гессе функции Лагранжа равна

$$\mathbf{H} = \begin{bmatrix} 2-2\lambda & 0 & 0 \\ 0 & 2-2\lambda & 0 \\ 0 & 0 & 2-8\lambda \end{bmatrix}.$$

Точка $A(4;0;3/2)$: здесь $\nabla u(A) = (8;0;12) \Longrightarrow T_A = \{2x+3z = \dfrac{25}{2}\}$, касательные вектора имеют вид $\vec{\tau} = (3s;t;-2s)$. Перемножим:

$$\vec{\tau}^{\mathrm{T}} \cdot \mathbf{H}_A \cdot \vec{\tau} = (3s;t;-2s) \begin{bmatrix} 3/2 & 0 & 0 \\ 0 & 3/2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3s \\ t \\ -2s \end{bmatrix} = \frac{27}{2}s^2 + \frac{3}{2}t^2 > 0.$$

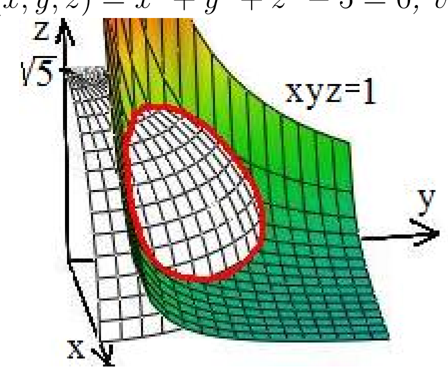
Следовательно, A — точка условного минимума.
Точка $B(4;0;-3/2)$ — то же самое.
Точка $C(5;0;0)$: здесь $\nabla u(C) = (10;0;0) \Longrightarrow T_C = \{x = 5\}$, касательные вектора имеют вид $\vec{\tau} = (0;t;s)$. Перемножим:

$$\vec{\tau}^{\mathrm{T}} \cdot \mathbf{H}_C \cdot \vec{\tau} = (0;t;s) \begin{bmatrix} 6/5 & 0 & 0 \\ 0 & 6/5 & 0 \\ 0 & 0 & -6/5 \end{bmatrix} \begin{bmatrix} 0 \\ t \\ s \end{bmatrix} = \frac{6}{5}(t^2 - s^2) - \text{ переменного знака.}$$

Следовательно, C — условная седловая.
Точка $D(-5;0;0)$: поскольку $\lambda = 8/5$, матрица $\mathbf{H}_D$ отрицательно определена на всём $\mathbb{R}^3 \Longrightarrow$ и на $T_D \Longrightarrow D$ — точка условного максимума.

Теперь рассмотрим задачу на условный экстремум с двумерным условием.

Пример 11.4. *Исследовать функцию* $f(x,y,z) = x+y+z$ *на экстремум при условиях* $u(x,y,z) = x^2+y^2+z^2-5 = 0$, $v(x,y,z) = xyz-1 = 0$ *в области* $x,y,z > 0$.



Функция Лагранжа

$$\mathcal{L}(x,y,z) = x+y+z-\lambda(x^2+y^2+z^2-5)-\mu(xyz-1).$$

Решаем систему уравнений:

$$\left\{\begin{array}{ll} 1-2\lambda x-\mu yz & =0 \\ 1-2\lambda y-\mu xz & =0 \\ 1-2\lambda z-\mu xy & =0 \\ x^2+y^2+z^2 & =5 \\ xyz & =1. \end{array}\right. \qquad \begin{array}{c} \text{Домножив первые три уравнения} \\ \text{соотв. на } x,y,z, \text{ получим} \\ x-2\lambda x^2=y-2\lambda y^2=z-2\lambda z^2=\mu. \end{array}$$

Если $\lambda=0$, то получим $x=y=z$, что невозможно при условии $u=v=0$. Если $\lambda\neq 0$, то имеем $\phi(x)=\phi(y)=\phi(z)$, где $\phi(t)=t-2\lambda t^2$, график ϕ парабола с осью симметрии $t=\dfrac{1}{4\lambda}\implies\left|x-\dfrac{1}{4\lambda}\right|=\left|y-\dfrac{1}{4\lambda}\right|=\left|z-\dfrac{1}{4\lambda}\right|\implies$ из координат x,y,z две принимают значение a и одна b, причём $a+b=\dfrac{1}{2\lambda}$. Ищем a и b.

$$\left\{\begin{array}{r} 2a^2+b^2=5 \\ a^2b=1 \end{array}\right.\implies\left\{\begin{array}{r} 2/b+b^2=5 \\ a=1/\sqrt{b} \end{array}\right.\implies b^3-5b+2=0\implies\left[\begin{array}{l} b=2 \\ b=\sqrt{2}-1 \end{array}\right.$$

(третий корень $b=-\sqrt{2}-1<0$). Если $b=2$, то $a=1/\sqrt{2}$; если $b=\sqrt{2}-1$, то $a=\sqrt{\sqrt{2}+1}$. Параметры λ и μ выражаются через них так:
$a+b=\dfrac{1}{2\lambda}\implies\lambda=\dfrac{1}{2(a+b)};\qquad \mu=b-2\lambda b^2=b-\dfrac{b^2}{a+b}=\dfrac{ab}{a+b}.$
Матрица Гессе функции Лагранжа в точке $(a;a;b)$ равна

$$\mathbf{H}=-\left[\begin{array}{ccc} 2\lambda & \mu b & \mu a \\ \mu b & 2\lambda & \mu a \\ \mu a & \mu a & 2\lambda \end{array}\right]=\frac{-1}{a+b}\left[\begin{array}{ccc} 1 & ab^2 & a^2b \\ ab^2 & 1 & a^2b \\ a^2b & a^2b & 1 \end{array}\right].$$

Найдём касательный вектор в точке $(a;a;b)$.

$$\nabla u=2(a;a;b),\ \nabla v=(ab;ab;a^2)\implies\vec{\tau}\parallel\left|\begin{array}{ccc} \vec{i} & \vec{j} & \vec{k} \\ a & a & b \\ b & b & a \end{array}\right|=(a^2-b^2)\left[\begin{array}{c} 1 \\ -1 \\ 0 \end{array}\right].$$

Получаем $\vec{\tau}^{\mathrm{T}}\cdot\mathbf{H}\cdot\vec{\tau}=$

$$=(1;-1;0)\frac{-1}{a+b}\left[\begin{array}{ccc} 1 & ab^2 & a^2b \\ ab^2 & 1 & a^2b \\ a^2b & a^2b & 1 \end{array}\right]\left[\begin{array}{c} 1 \\ -1 \\ 0 \end{array}\right]=\frac{-2}{a+b}(1-ab^2)\left\{\begin{array}{l} <0 \text{ при } a>b \\ >0 \text{ при } a<b. \end{array}\right.$$

Ответ: точки условного максимума $(\sqrt{\sqrt{2}+1};\sqrt{\sqrt{2}+1};\sqrt{2}-1)$, $(\sqrt{\sqrt{2}+1};\sqrt{2}-1;\sqrt{\sqrt{2}+1})$, $(\sqrt{2}-1;\sqrt{\sqrt{2}+1};\sqrt{\sqrt{2}+1})$; точки условного минимума $(1/\sqrt{2};1/\sqrt{2};2)$, $(1/\sqrt{2};2;1/\sqrt{2})$, $(2;1/\sqrt{2};1/\sqrt{2})$.

12 Интеграл, зависящий от параметра

Теорема 12.1. *Пусть функция $f(x,y)$ непрерывна на множестве $D = \{(x;y):\ a \le x \le b,\quad g(x) \le y \le h(x)\}$, где g и h непрерывны на $[a;b]$, $h(x) > g(x)$ на $(a;b)$. Тогда функция*

$$F(x) = \int\limits_{g(x)}^{h(x)} f(x,y)dy \qquad \text{непрерывна на } [a;b].$$

При доказательстве нам понадобится следующее свойство функций, непрерывных на компакте:

Теорема 12.2. *Если скалярная функция f непрерывна на компакте $K \subset \mathbb{R}^n$, то она* **равномерно непрерывна** *на нём, т.е.*

$$\forall \varepsilon > 0\ \exists \delta > 0 : \text{если } X,\ Y \in K,\ |\vec{XY}| < \delta, \text{ то } |f(X) - f(Y)| < \varepsilon.$$

Доказательство. Предположим, что равномерной непрерывности нет, т.е. $\exists \varepsilon > 0 : \forall \delta > 0\ \exists X,\ Y \in K : |\vec{XY}| < \delta$, но $|f(X) - f(Y)| \ge \varepsilon$. Выберем такие X_k, Y_k для $\delta = 1/k$, $k = 1, 2, \dots$, и в силу компактности K выберем из последовательности $\{X_k\}$ подпоследовательность $\{X_{k_i}\}$, сходящуюся к некоторой точке $X_o \in K$. Тогда по неравенству треугольника $|X_o\vec{Y}_{k_i}| \le |X_o\vec{X}_{k_i}| + 1/k_i \to 0$, значит, $Y_{k_i} \to X_o$ при $i \to \infty$. Отсюда по непрерывности f будет следовать

$$\lim_{i\to\infty} f(X_{k_i}) = \lim_{i\to\infty} f(Y_{k_i}) = f(X_o) \Longrightarrow \lim_{i\to\infty} |f(X_{k_i}) - f(Y_{k_i})| = 0,$$

но это противоречит предположению $|f(X_{k_i}) - f(Y_{k_i})| \ge \varepsilon$. □

Доказательство теор. 12.1. Пусть $A = \inf\limits_{[a;b]} f$, $B = \sup\limits_{[a;b]} g$ (оба конечны). Функция f непрерывна на компакте $D \Longrightarrow$ ограничена и равномерно непрерывна на нём. Пусть $M = \max\limits_D |f|$. Зададим $\varepsilon > 0$. Тогда $\exists \delta_1 > 0$:
если $(x_1, y_1), (x_2, y_2) \in D$, $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} < \delta_1$, то $|f(x_1, y_1) - f(x_2, y_2)| < \varepsilon$. Далее, $\exists \delta_2 > 0$: если $a \le x_1, x_2 \le b$, $|x_1 - x_2| < \delta_2$, то $|g(x_1) - g(x_2)| < \varepsilon$ и $|h(x_1) - h(x_2)| < \varepsilon$.
Пусть $x, u \in [a;b]$, $\quad |u - x| < \delta = \min\{\delta_1, \delta_2\}$. Тогда

$$|F(u) - F(x)| \le M \cdot |g(u) - g(x)| + M \cdot |h(u) - h(x)| + \\ + \int\limits_{\max\{g(u),g(x)\}}^{\min\{h(u),h(x)\}} |f(u,y) - f(x,y)|dy < 2M\varepsilon + (B - A)\varepsilon.$$

Следовательно, при $u \to x$ имеем $F(u) \to F(x)$. □

Теорема 12.3. *Пусть функции $f(x,y)$ и $\dfrac{\partial f}{\partial x}$ непрерывны на множестве $D = \{(x;y) : a \le x \le b, \quad g(x) \le y \le h(x)\}$, где g и h дифференцируемы на $[a;b]$, $h(x) > g(x)$ на $(a;b)$. Тогда на $[a;b]$ дифференцируема функция*

$$F(x) = \int\limits_{g(x)}^{h(x)} f(x,y)dy; \quad F'(x) = \int\limits_{g(x)}^{h(x)} \frac{\partial f}{\partial x}(x,y)dy + f(x,h(x))h'(x) - f(x,g(x))g'(x).$$

Несобственный интеграл, зависящий от параметра

Для удобства обозначений будем рассматривать несобственные интегралы I рода по полупрямой $[0;+\infty)$.

Определение 12.1. *Пусть функция $f(x,y)$ определена на $[a;b]\times[0;+\infty)$. Несобственный интеграл I рода $\displaystyle\int_0^{+\infty} f(x,y)dy$* **сходится равномерно** *на $[a;b]$, если*

$$\forall\varepsilon > 0 \ \ \exists M > 0 : \quad \forall N > M \ \ \forall x \in [a;b] \quad \left|\int_N^{+\infty} f(x,y)dy\right| < \varepsilon.$$

Теорема 12.4. *Пусть функция $f(x,y)$ непрерывна на множестве $[a;b]\times[0;+\infty)$. Тогда функция*

$$F(x) = \int_0^{+\infty} f(x,y)dy, \quad x \in [a;b] \tag{7}$$

непрерывна при условии, что интеграл (7) сходится равномерно.

Доказательство. Пусть $x \in [a;b]$. Зададим $\varepsilon > 0$. В силу равномерной сходимости $\exists\ M > 0$:

$$\forall t \in [a;b] \quad \left|\int_0^M f(t,y)dy - F(t)\right| < \frac{\varepsilon}{3}. \tag{8}$$

Далее, по теореме 12.1, $\exists\ \delta > 0$:

$$\forall u \in U_\delta(x)\cap[a;b] \quad \left|\int_0^M f(u,y)dy - \int_0^M f(x,y)dy\right| < \frac{\varepsilon}{3}. \tag{9}$$

Сложив (9) с оценками (8) для $t = x$ и $t = u$, получаем $|F(u) - F(x)| < \varepsilon$. □

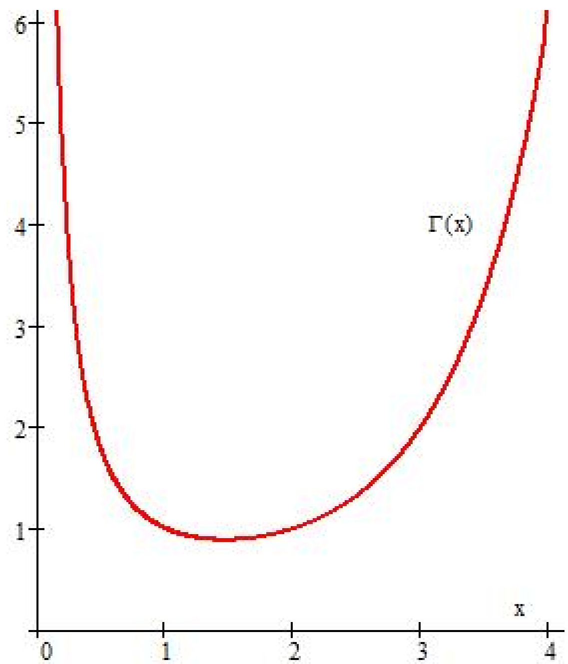
Следующий пример показывает необходимость равномерной сходимости.

Пример 12.1. Рассмотрим интеграл $F(x) = \displaystyle\int_0^{+\infty} x^2 y e^{-xy} dy$, $x \in [0;1]$. Подинтегральная функция непрерывна и ограничена на $[0;1]\times[0;+\infty)$, однако в 0 функция F разрывна: $F(0) = 0$, $F(x) = 1$ при $x > 0$.

Пример 12.2. *Гамма-функция Эйлера*

$$\Gamma(x) = \int_0^{+\infty} y^{x-1} e^{-y} dy, \quad x > 0.$$

Интеграл сходится равномерно на каждом отрезке $x \in [a; b] \subset (0; +\infty)$. Поэтому Γ непрерывна на $(0; +\infty)$. Интегрированием по частям получается рекуррентное соотношение $\Gamma(x+1) = x \cdot \Gamma(x)$. Поскольку $\Gamma(1) = 1$, получаем $\Gamma(n+1) = n!$



Теорема 12.5. *Пусть функции* $f(x, y)$ *и* $\dfrac{\partial f}{\partial x}$ *непрерывны на множестве* $[0; 1] \times [0; +\infty)$. *Тогда функция*

$$F(x) = \int_0^{+\infty} f(x, y) dy \qquad \text{дифференцируема на} \ \ [a; b], \tag{10}$$

$$F'(x) = \int_0^{+\infty} \frac{\partial f}{\partial x}(x, y) dy, \tag{11}$$

при условии, что оба интеграла (10) *и* (11) *сходятся равномерно.*

Пример 12.2 (продолжение). Гамма-функция дифференцируема, и её производная равна $\Gamma'(x) = \displaystyle\int_0^{+\infty} \ln y \cdot y^{x-1} e^{-y} dy \qquad \forall\ x > 0.$