Функции многих переменных
Лекции для ИУ9

Далее ссылки на учебники:

[**З**]: Зорич В.А. Математический анализ. Ч. I. - М.: Наука, 1981. - 544с.

[**АСЧ**]: Архипов Г.И., Садовничий В.А., Чубариков В.Н. Лекции по математическому анализу / Под ред. В.А. Садовничего. - М.: Высш. шк., 1999. - 695с.

[**ККЧ**]: Канатников А.Н., Крищенко А.П., Четвериков В.Н. Дифференциальное исчисление функций многих переменных. - М.: Изд-во МГТУ, 2000. - 456с.

Многие величины, встречающиеся в природе, зависят не от одного, а от нескольких факторов, и если сама величина и каждый из определяющих ее факторов могут быть охарактеризованы некоторым числом, то указанная зависимость сводится к тому, что упорядоченному набору $(x_1, \ldots, x_n)$ чисел, каждый из которых описывает состояние соответствующего фактора, ставится в соответствие значение $y = f(x_1, \ldots, x_n)$ исследуемой величины, которое она приобретает при этом состоянии факторов. Такая зависимость называется функцией многих переменных.

Теорию функций многих переменных мы будем строить, обобщая соответствующие понятия теории функций одного переменного. Первым и основным среди них является понятие открытого множества. Семейство всех рассматриваемых открытых множеств называют топологией. Отметим, что топологией также называют раздел математики, изучающий в самом общем виде явление непрерывности. Мы используем далее только первое упомянутое здесь значении этого термина.

1 Топология в $\mathbb{R}^n$

См. [ККЧ, §1.1, стр. 20–30], [АСЧ, стр. 297, 303–305, 309–310], [З, гл.7, §1, стр. 403–408].

1.1 Открытые множества в $\mathbb{R}^n$

Множество упорядоченных наборов $(x_1, \ldots, x_n)$ из n действительных чисел x_1, ..., x_n есть n–я декартова степень $\mathbb{R}^n$ множества $\mathbb{R}$ действительных чисел. Элементы $\mathbb{R}^n$ будем называть не векторами, а точками. Для элемента $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ числа x_1, ..., x_n будем называть **координатами точки** x **в** $\mathbb{R}^n$.

Расстоянием $\rho(x,y)$ между точками $x=(x_1,\dots,x_n)$ и $y=(y_1,\dots,y_n)$ назовем число

$$\rho(x,y)=\sqrt{(x_1-y_1)^2+\dots+(x_n-y_n)^2}. \tag{1}$$

Вспомним, что $\mathbb{R}^n$ — это евклидово пространство со скалярным произведением $(x,y)=x_1y_1+\dots+x_ny_n$. В евклидовом пространстве можно ввести евклидову норму $|x|=\sqrt{(x,x)}$ и в соответствии с этой нормой расстояние $\rho(x,y)=|x-y|$, которое совпадает с расстоянием, введенным согласно формуле (1).

Пусть a — произвольная точка из $\mathbb{R}^n$, ε — положительное число. Множество

$$U_\varepsilon(a)=\{x\in\mathbb{R}^n\colon\ |x-a|<\varepsilon\},$$

называют ε**–окрестностью точки** a, а множество

$$\dot U_\varepsilon(a)=U_\varepsilon(a)\setminus\{a\}=\{x\in\mathbb{R}^n\colon\ 0<|x-a|<\varepsilon\}\quad —$$

выколотой ε**–окрестностью** точки a.

Точку a множества $E\subseteq\mathbb{R}^n$ называют **внутренней точкой** этого **множества**, если существует ε-окрестность $U_\varepsilon(a)$ точки a, целиком содержащаяся в E: $U_\varepsilon(a)\subseteq E$. Множество всех внутренних точек E называют **внутренностью множества** E и обозначают $\operatorname{Int}E$. Если каждая точка множества E является его внутренней точкой: $E\subseteq\operatorname{Int}E$, то само множество E называют **открытым множеством**. Пустое множество по определению считают открытым.

В прошлом семестре открытыми множествами в $\mathbb{R}$ мы назвали любое объединение интервалов, но потом доказали, что у открытых множеств и только у них любая точка является внутренней. Таким образом, в случае $n=1$ определения открытого множества, введеное здесь и в прошлом семестре, эквивалентны.

Пример 1. Открытыми множествами в $\mathbb{R}^n$ являются ε-окрестности точек. Действительно, рассмотрим произвольную точку $a\in\mathbb{R}^n$ и ее ε-окрестность $U_\varepsilon(a)$ (для $n=2$ см. рис. 1). Если $x\in U_\varepsilon(a)$, то по определению имеем $\rho(x,a)<\varepsilon$. Выберем положительное число $\varepsilon_1=\varepsilon-\rho(x,a)$. Если точка y принадлежит ε_1-окрестности $U_{\varepsilon_1}(x)$ точки x, то $\rho(y,x)<\varepsilon_1$. Согласно неравенству треугольника,

$$\rho(y,a)\leqslant\rho(y,x)+\rho(x,a)<\varepsilon_1+\rho(x,a)=\varepsilon.$$

Значит, точка y принадлежит ε-окрестности $U_\varepsilon(a)$ точки a. Поскольку точка $y\in$ $\in U_{\varepsilon_1}(x)$ может быть выбрана произвольно, заключаем, что $U_{\varepsilon_1}(x)\subseteq U_\varepsilon(a)$. Итак, любая точка $x\in U_\varepsilon(a)$ имеет ε_1-окрестность $U_{\varepsilon_1}(x)$, целиком попадающую в $U_\varepsilon(a)$. Это означает, что точка x внутренняя для множества $U_\varepsilon(a)$, которое, следовательно, является открытым.

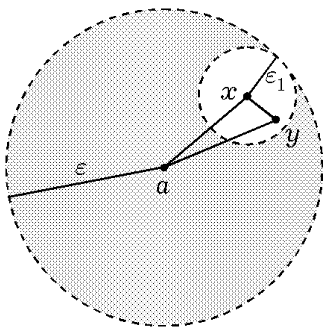


Рис. 1

Поэтому ε-окрестности точек в $\mathbb{R}^n$ называют также **открытыми n-мерными шарами**. Множество точек $x \in \mathbb{R}^n$, для которых $\rho(x, a) = \varepsilon$, называют $(n-1)$–**мерной сферой** радиуса ε с центром в точке a и обозначают $S_\varepsilon^{n-1}(a)$. Сфера не яляется открытым множеством.

1.2 Понятие топологического пространства

Для построения теории пределов открытые множества должны обладать рядом свойств. Сформулируем базовые из них, т.е. те, из которых следуют остальные, и покажем, что этими свойствами обладают введенные выше открытые множества.

Пусть дано множество X. Система τ его подмножеств называется **топологией** на X, если выполнены следующие условия (**аксиомы топологического пространства**):

1. Объединение произвольного семейства множеств, принадлежащих τ, принадлежит τ, то есть если $U_\alpha \in \tau$ при $\alpha \in I$, то $\bigcup\limits_{\alpha \in I} U_\alpha \in \tau$.

2. Пересечение конечного семейства множеств, принадлежащих τ, принадлежит τ, то есть если $U_i \in \tau$ при $i = 1, \ldots, n$, то $\bigcap\limits_{i=1}^{n} U_i \in \tau$.

3. $X, \varnothing \in \tau$.

Пара (X, τ) называется **топологическим пространством**. Множества, принадлежащие τ, называются **открытыми множествами в** X.

Отметим, что на одном множестве X существует много топологий. Например, мы можем рассмотреть все подмножества X и получить самую большую топологию, или положить $\tau = \{X, \varnothing\}$ и получить самую маленькую топологию на X. Топологическим пространством поэтому называют пару — множество X вместе с топологией τ.

Окрестностью точки $a \in X$ называют любое открытое множество U в X, включающее в себя эту точку. Это фактически означает, что открытое множество является окрестностью каждой своей точки. При этом множество $\dot{U} = U \setminus \{a\}$ называют **выколотой окрестностью** точки a. Как следует из примера 1, ε-окрестность точки $a \in \mathbb{R}^n$ является ее окрестностью. Будем обозначать окрестности точки a через $U(a)$, $V(a)$ и т.п., выколотые окрестности точки a — через $\dot{U}(a)$, $\dot{V}(a)$ и т.п.

Топологическое пространство X называется **хаусдорфовым**, если любые две различных точки x, y из X обладают непересекающимися окрестностями $U(x), V(y)$: $U(x) \cap V(y) = \varnothing$ (**условие Хаусдорфа**).

Теорема 1. $\mathbb{R}^n$ — хаусдорфово топологическое пространство.

Док–во. Необходимо доказать аксиомы 1–3 топологического пространства и условие Хаусдорфа.

Аксиома 3 следует из определений.

Для доказательства аксиомы 1 рассмотрим множество

$$V = \bigcup_{\alpha \in I} V_\alpha,$$

где множества $V_\alpha \subseteq \mathbb{R}^n$, $\alpha \in I$, открытые, а I — некоторое множество индексов. В случае пустого множества V утверждение очевидно, и мы будем считать, что V не пусто. Если точка a принадлежит множеству V, то по определению операции объединения множеств точка a принадлежит множеству V_α хотя бы для одного значения индекса $\alpha = k$. Так как V_k — открытое множество, то существует ε-окрестность $U_\varepsilon(a)$ точки a, содержащаяся в V_k. Следовательно, эта окрестность содержится и в V. Но это значит, что a — внутренняя точка V, а так как она может быть выбрана в V произвольно, множество V открытое.

Докажем аксиому 2. Пусть множества U_i, $i = 1, \ldots, n$, открыты и

$$U = \bigcap_{i=1}^{n} U_i.$$

Если множество U пустое, то оно открыто по определению. Для непустого множества U рассмотрим произвольную точку $a \in U$. Согласно определению пересечения множеств, она принадлежит каждому из множеств U_i, $i = 1, \ldots, n$. Так как эти множества открыты, то по определению для каждого множества U_i существует такое число $\varepsilon_i > 0$, что ε_i-окрестность точки a содержится в U_i. Положим $\varepsilon = \min\{\varepsilon_1, \ldots, \varepsilon_n\}$. Тогда при всех $i = 1, \ldots, n$ выполнены неравенства $\varepsilon \leqslant \varepsilon_i$. Имеем $U_\varepsilon(a) \subseteq U_{\varepsilon_i}(a) \subseteq U_i$, $i = 1, \ldots, n$. Поэтому ε-окрестность $U_\varepsilon(a)$ содержится и в пересечении всех множеств U_i, т.е. в множестве U, а это по определению означает, что a — внутренняя точка для множества U. Поскольку в качестве точки a может быть выбрана любая точка множества U, это множество открытое.

Для доказательства условия Хаусдорфа рассмотрим две произвольные точки a и b из $\mathbb{R}^n$ и их окрестности $U_\varepsilon(a)$ и $U_\varepsilon(b)$, где $\varepsilon = \rho(a, b)/2$. Тогда $U_\varepsilon(a) \cap U_\varepsilon(b) = \varnothing$. $\triangleright$

1.3 Другие классы множеств в $\mathbb{R}^n$

Как и для пространства $\mathbb{R}$, для $\mathbb{R}^n$ вводят следующие понятия. Дополнение открытого множества называют **замкнутым множеством**. Точку $x \in \mathbb{R}^n$ называют **предельной точкой множества** $E \subseteq \mathbb{R}^n$, если любая окрестность точки x содержит хотя бы одну точку множества E, отличную от x. **Замыканием множества** E называют объединение E с множеством предельных точек E. Обозначение: $[E]$. **Открытым покрытием множества** $E \subseteq \mathbb{R}^n$ называют семейство $\{G_\alpha\}$ открытых подмножеств $\mathbb{R}^n$, такое, что $E \subseteq \cup_\alpha G_\alpha$. Множество $E \subseteq \mathbb{R}^n$ называют **компактом**, если из любого его открытого покрытия можно выделить конечное подпокрытие. Точку $a \in \mathbb{R}^n$ называют **граничной точкой** множества $E \subset \mathbb{R}^n$, если любая ε-окрестность точки a содержит как точки, принадлежащие множеству E, так и точки, не принадлежащие этому множеству. Множество всех граничных точек множества E называют его **границей** и обозначают ∂E.

Пример 2. На плоскости границей замкнутого круга

$$\{(x_1, x_2) : \ (x_1 - a_1)^2 + (x_2 - a_2)^2 \leqslant \varepsilon^2\}$$

радиуса ε с центром в точке $a = (a_1, a_2)$ является окружность. Для открытого круга аналогично.

Задача 1. Докажите:
множество замкнуто $\quad\Leftrightarrow\quad$ оно содержит все свои граничные точки.

Точку $b \in \mathbb{R}^n$ называют **внешней точкой множества** $E \subset \mathbb{R}^n$, если существует такая ε-окрестность этой точки, которая не пересекается с множеством E. Множество всех внешних точек множества E называют **внешностью множества** E.

Если точка $b \in \mathbb{R}^n$ не принадлежит множеству $E \subset \mathbb{R}^n$, то существуют две возможности: а) любая ε-окрестность точки b содержит точки множества E и, следовательно, точка b является граничной точкой множества E; б) некоторая ε-окрестность точки b не пересекается с E и, следовательно, точка b является внешней точкой множества E.

Точку $a \in E \subset \mathbb{R}^n$ называют **изолированной** точкой множества E, если в некоторой ее окрестности нет других точек множества E, кроме a.

Задача 2. Могут ли изолированные или предельные точки множества быть его внутренними, граничными или внешними точками? Если могут, приведите пример множества с такими точками. Если нет, докажите, что не могут.

Множество $E \subseteq \mathbb{R}^n$, любые две точки которого можно соединить кривой, целиком лежащей в этом множестве, называют **линейно связным**. Открытое линейно связное множество называют **областью**.

Пример 3. Следующие множества являются областями:
– любая ε-окрестность $U_\varepsilon(a)$ точки $a \in \mathbb{R}^n$;
– выколотая ε-окрестность $\dot{U}_\varepsilon(a)$ точки $a \in \mathbb{R}^n$;
– открытое кольцо в $\mathbb{R}^2$ с центром в точке (a_1, a_2) и радиусами r и R:

$$r^2 < (x_1 - a_1)^2 + (x_2 - a_2)^2 < R^2, \quad (x_1, x_2) \in \mathbb{R}^2.$$

Пример 4. $\mathbb{R}^n \setminus S_r^{n-1}(a)$ — не линейно связное множество (докажем это позже).

В заключениие отметим, что все понятия этого пункта и многие понятия теории пределов и непрерывности, вводимые далее, обобщаются на случай произвольного хаусдорфова пространства. Эта особенность данной теории будет использоваться в следующих математических курсах.

2 Компакты в $\mathbb{R}^n$

См. [3, гл.7, §1, стр. 408–409].

Множество

$$I^n = \{x \in \mathbb{R}^n \colon\ a_i \leqslant x_i \leqslant b_i,\ i = 1, \ldots, n\}$$

называют n–**мерным брусом**.

Теорема 1. n–Мерный брус есть компакт.

Док–во. Предположим, что из некоторого открытого покрытия n–мерного бруса $I = I^n$ нельзя выделить конечное подпокрытие. Разделив каждый из координатных отрезков $[a_i, b_i], i = 1, \ldots, n$, пополам, мы разобьем брус I на 2^n брусов, из которых по крайней мере один не допускает конечного покрытия множествами нашей системы. С ним поступим так же, как и с исходным брусом.

Продолжая этот процесс деления, получим последовательность вложенных брусов $I = I_1 \supset I_2 \supset \ldots \supset I_k \supset \ldots$, ни один из которых не допускает конечного подпокрытия. Если

$$I_k = \{x \in \mathbb{R}^n \colon\ a_i^k \leqslant x_i \leqslant b_i^k,\ i = 1, \ldots, n\}, \quad k = 1, 2, \ldots,$$

то при каждом $i = 1, \ldots, n$ координатные отрезки $[a_i^k, b_i^k]$, $k = 1, 2, \ldots$, образуют, по построению, систему вложенных отрезков, длины которых стремятся к нулю. Найдя при каждом $i = 1, \ldots, n$ точку $\xi_i \in [a_i^k, b_i^k]$, общую для всех этих отрезков, получим точку $\xi = (\xi_1, \ldots, \xi_n)$, принадлежащую всем брусам $I = I_1, I_2, \ldots, I_k, \ldots$. Поскольку $\xi \in I$, то найдется такое открытое множество G_α нашей системы покрывающих множеств, что $\xi \in G_\alpha$. Тогда при некотором $\varepsilon > 0$ также $U_\varepsilon(\xi) \subseteq G_\alpha$. Так как длины отрезков $[a_i^k, b_i^k], i = 1, \ldots, n$, стремятся к нулю, то найдется номер N такой, что $|b_i^k - a_i^k| < \dfrac{\varepsilon}{\sqrt{n}}$ при $k > N$, $i = 1, \ldots, n$. Тогда для любой точки $x = (x_1, \ldots, x_n)$ из I_k имеем

$$|x - \xi| = \sqrt{(x_1 - \xi_1)^2 + \cdots + (x_n - \xi_n)^2} \leqslant \sqrt{(b_1^k - a_1^k)^2 + \cdots + (b_n^k - a_n^k)^2} < \varepsilon.$$

Поэтому $I_k \subseteq U_\varepsilon(\xi) \subseteq G_\alpha$ при $k > N$, и мы вступаем в противоречие с тем, что брус I_k не допускает конечного покрытия множествами данной системы. $\triangleright$

Множество $E \subset \mathbb{R}^n$ называют **ограниченным множеством**, если существует такое положительное число r, что r-окрестность точки $O = (0, \ldots, 0)$ содержит множество E.

Теорема 2. Множество $K \subset \mathbb{R}^n$ является компактом $\Leftrightarrow$ оно замкнутое и ограниченное.

Док–во аналогично случаю $n = 1$ и состоит из трех частей.

1) Докажем замкнутость компакта. Используя условие Хаусдорфа покрываем компакт K окрестностями $U_x(x), x \in K$, непересекающимися с окрестностями $V_x(z), x \in K$, фиксированной точки $z \in \overline{K}$. Выбираем конечное подпокрытие $\{U_{x_i}\}_{i=1,\ldots,l}, x_i \in K$. Тогда $\bigcap_{i=1}^l V_{x_i}$ — окрестность точки $z \in \overline{K}$, непересекающаяся с $\bigcup_{i=1}^l U_{x_i} \supseteq K$. Т.е. z — внутренняя точка $\overline{K}$, $\overline{K}$ открыто, а значит, K замкнуто.

2) Для доказательства ограниченности компакта K рассмотрим последовательность открытых шаров $U_k(O)$, $k = 1, 2, \ldots$. Они образуют открытое покрытие $\mathbb{R}^n$, а следовательно, и K. Если бы K не было ограниченным множеством, то из этого покрытия нельзя было бы извлечь конечное подпокрытие K.

3) Пусть K — замкнутое и ограниченное множество. Тогда найдется n–мерный брус I, содержащий K. Компактность K следует из теоремы 1 и следующей леммы.

Лемма. Любое замкнутое подмножество компакта есть компакт .

Док–во леммы. Пусть I — компакт в $\mathbb{R}^n$, F — замкнутое в $\mathbb{R}^n$ множество и $F \subseteq I$. Пусть $\{G_\alpha\}_{\alpha \in I}$ — покрытие F множествами, открытыми в $\mathbb{R}^n$. Присоединив к нему еще одно открытое множество $G = \mathbb{R}^n \setminus F$, получим открытое покрытие $\mathbb{R}^n$ и, в частности, I, из которого извлекаем конечное подпокрытие I. Это конечное покрытие I будет покрывать также множество F. Замечая, что $G \cap F = \varnothing$, можно сказать, что если G входит в это конечное покрытие, то, даже удалив G, мы получим конечное покрытие F множествами исходной системы $\{G_\alpha\}_{\alpha \in I}$. $\triangleright$

3 Предел функции многих переменных

См. [ККЧ, §1.2, 1.3, стр. 31–52], [З, гл.7, §2, стр. 410–415].

3.1 Предел последовательности в $\mathbb{R}^n$

Как и в случае числовых последовательностей, **последовательностью в** $\mathbb{R}^n$ называют любое отображение $\mathbb{N} \longrightarrow \mathbb{R}^n$. Последовательность будем обозначать $\{a_k\}$, если $k \longmapsto a_k \in \mathbb{R}^n$ при $k \in \mathbb{N}$.

Пусть для последовательности $\{a_k\}$ существует такая точка $a \in \mathbb{R}^n$, что для любой ее ε-окрестности $U_\varepsilon(a)$ можно указать такой номер $N \in \mathbb{N}$, что для любого $k > N$ верно соотношение $a_k \in U_\varepsilon(a)$, т.е.

$$\forall \varepsilon > 0 \quad \exists N = N(\varepsilon) \in \mathbb{N}: \quad \forall k > N \quad |a_k - a| < \varepsilon.$$

Тогда $\{a_k\}$ называют **сходящейся последовательностью в** $\mathbb{R}^n$, а точку a — **пределом последовательности** $\{a_k\}$ **в** $\mathbb{R}^n$. Если указанной точки a не существует, то последовательность $\{a_k\}$ называют **расходящейся последовательностью в** $\mathbb{R}^n$.

Для предела последовательности в $\mathbb{R}^n$ сохраняются основные свойства числовых последовательностей, которые можно рассматривать как частный случай последовательностей в $\mathbb{R}^n$ при $n = 1$. Последовательность $\{a_k\}$ в $\mathbb{R}^n$ называют **фундаментальной**, если для любого числа $\varepsilon > 0$ можно указать такой номер $N \in \mathbb{N}$, что для любых $k > N$ и $m > N$ выполняется неравенство $|a_k - a_m| < \varepsilon$.

Критерий Коши. Последовательность сходится т. и т. т., к. она является фундаментальной.

Без док–ва.

3.2 Определение предела функции

Отображение вида $f\colon\ A \to \mathbb{R}^m$, где $A \subseteq \mathbb{R}^n$, $n > 1$, называют **функцией многих переменных** или **функцией нескольких переменных**. В случае $m = 1$ значением такого отображения является действительное число и отображение называют **скалярной функцией многих переменных**. При $m > 1$ будем называть такое отображение **векторной функцией многих переменных**.

Графиком функции многих переменных $f\colon\ A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ называют подмножество

$$\Gamma(f) = \{(x; y) \in \mathbb{R}^{n+m}\colon\ x \in A,\ y = f(x)\} \ \subset\ \mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m.$$

Поскольку элемент линейного пространства $\mathbb{R}^m$ при $m > 1$ является совокупностью m действительных чисел, то векторную функцию многих переменных $f\colon\ A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ можно рассматривать как совокупность m скалярных функций f_i, полагая, что

$$f(x) = \big(f_1(x), \ldots, f_m(x)\big), \quad x \in A. \tag{1}$$

Функции многих переменных f_i, $i = 1, \ldots, m$, называют **координатными функциями** векторной функции f. Для представления векторной функции наряду с матричной формой записи (1) используют координатную форму записи

$$y_1 = f_1(x), \quad \dots, \quad y_m = f_m(x), \quad x \in A.$$

Пусть $A \subseteq \mathbb{R}^n$ — область определения функции, $\mathcal{B}$ — база в множестве A. Точку $d \in \mathbb{R}^m$ называют **пределом функции** $f\colon A \to \mathbb{R}^m$ **по базе** $\mathcal{B}$, если

$$\forall \varepsilon > 0 \quad \exists b = b(\varepsilon) \in \mathcal{B}: \quad \forall x \in b \quad |f(x) - d| < \varepsilon.$$

Обозначение: $\lim\limits_{\mathcal{B}} f(x) = d$.

Теорема 1 (**об эквивалентности определений предела функции**).

$$\begin{aligned} \lim_{\mathcal{B}} f(x) = d \quad &\Leftrightarrow \quad a)\ \lim_{\mathcal{B}} \rho\big(f(x), d\big) = 0 \\ &\Leftrightarrow \quad b)\ \forall U(d)\ (\text{окрест.}) \quad \exists b \in \mathcal{B}: \quad f(b) \subseteq U(d) \\ &\Leftrightarrow \quad c)\ \lim_{\mathcal{B}} f_i(x) = d_i, \quad i = 1, \dots, m, \quad d = (d_1, \dots, d_m). \end{aligned}$$

Док–во следует из определений.

Для функции $f\colon A \to \mathbb{R}^m$ будем обозначать:

$x \to a$ — база $\{\dot{U}_\delta(a) \cap A \ : \ \delta > 0\}$ выколотых δ–окрестностей точки a в множестве A, если a — предельная точка множества A,

$x \to \infty$ — база $\{A \setminus [U_r(O)] \ : \ r > 0\}$ окрестностей бесконечности в множестве A, если A — неограниченное множество ($[U_r(O)]$ — замыкание множества $U_r(O)$).

Задача 1. Докажите, что база $\{\dot{U}_\delta(a) \cap A \ : \ \delta > 0\}$ эквивалентна базе $\{\dot{U}(a) \cap A\}$, а база $\{A \setminus [U_r(O)] \ : \ r > 0\}$ эквивалентна базе $\{A \setminus [U_r(a)] \ : \ r > 0,\ a \in \mathbb{R}^n\}$. Т.е. предел функции по какой–либо базе существует и равен d т. и т.т., к. существует предел этой функции по эквивалентной базе и он также равен d.

Имеем

$$\begin{aligned} \lim_{x \to a} f(x) = d \quad &\Leftrightarrow \quad \forall \varepsilon > 0\ \exists \delta = \delta(\varepsilon) > 0: \ \forall x \in A\ 0 < |x - a| < \delta \Rightarrow |f(x) - d| < \varepsilon \\ \lim_{x \to \infty} f(x) = d \quad &\Leftrightarrow \quad \forall \varepsilon > 0 \quad \exists r > 0 \ : \quad \forall x \in A \quad |x| > r \quad \Rightarrow \quad |f(x) - d| < \varepsilon \\ \lim_{x \to a} f(x) = \infty \quad &\Leftrightarrow \quad \forall r > 0\ \exists \delta = \delta(r) > 0: \ \forall x \in A\ 0 < |x - a| < \delta \Rightarrow |f(x)| > r \end{aligned}$$

Функцию $\alpha(x)$ называют **б.м. по базе** $\mathcal{B}$ $\quad \Leftrightarrow \quad \lim\limits_{\mathcal{B}} \alpha(x) = O$.

Функцию $f(x)$ называют **б.б. по базе** $\mathcal{B}$ $\quad \Leftrightarrow \quad \lim\limits_{\mathcal{B}} f(x) = \infty$.

3.3 Свойства предела функции

Функция многих переменных $f\colon A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ **ограничена на множестве** $B \subseteq A$, если множество $f(B) = \{y \in \mathbb{R}^m\colon\ y = f(x),\ x \in B\}$ ограничено.

Теорема 2 (**свойства предела функции**).

а) Фун. $f\colon A \to \mathbb{R}^m$ может иметь не более одного предела по данной базе $\mathcal{B}$ в A.

б) $\lim\limits_{\mathcal{B}} f(x) = d \quad \Rightarrow \quad \exists b \in \mathcal{B}: \ f$ огранич. на b.

Док–во аналогично случаю функции одного переменного.

Теорема 3 (**предел сложной функции или замена переменной в пределе**).

$$\left.\begin{array}{ll} 1) & g\colon Y \to \mathbb{R}^m, \quad \mathcal{B}_y \text{ — база в } Y \subseteq \mathbb{R}^k, \quad \exists \lim\limits_{\mathcal{B}_y} g(y), \\ 2) & f\colon X \to Y, \quad \mathcal{B}_x \text{ — база в } X \subseteq \mathbb{R}^n, \\ 3) & \forall b_y \in \mathcal{B}_y \quad \exists b_x \in \mathcal{B}_x : \quad f(b_x) \subseteq b_y \end{array}\right. \quad \Rightarrow \quad \exists \lim_{\mathcal{B}_x} g[f(x)] = \lim_{\mathcal{B}_y} g(y).$$

Док–во аналогично случаю функции одного переменного.

Теорема 4. Если $B \subset A$, a — предельная точка множества B, и $d = \lim\limits_{x\to a} f(x)$ существует, то существует $\lim\limits_{x\to a} f\big|_B(x)$, равный d.

Док–во следует из определений.

Пример 1. Рассмотрим функцию двух переменных

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

и исследуем ее на существование предела в точке $a = (0;0)$. Пусть

$$B_k = \{(x,y) \in \mathbb{R}^2 \ : \ y = kx\}, \qquad k \in \mathbb{R}.$$

Воспользуемся тем, что в точках B_k функцию f можно рассматривать как функцию одного действительного переменного $f\big|_{B_k}(x) \equiv f(x,kx)$, которая при $x \neq 0$ принимает постоянное значение:

$$f\big|_{B_k}(x) = f(x,kx) = \frac{kx^2}{x^2+k^2x^2} = \frac{k}{1+k^2}.$$

Поэтому при $(x;y) \to (0;0)$ существует предел функции $f\big|_{B_k}$, равный этому постоянному значению. Так как это значение зависит от k, то по теореме 3 функция f не имеет предела в точке $(0;0)$.

4 Непрерывность функции многих переменных

См. [ККЧ, §1.4, 1.5, 1.7], [З, гл.7, §2, стр. 415–420].

4.1 Непрерывность функций в точке

Функцию многих переменных $f\colon A \to \mathbb{R}^m$ называют **непрерывной в точке** $a \in A \subseteq \mathbb{R}^n$ в двух случаях: 1) a — изолированная точка множества A, или 2) a — предельная точка множества A и $\lim_{x\to a} f(x) = f(a)$.

Теорема 1 (**об эквивалентности определений непрерывности**).
Функция $f\colon A \to \mathbb{R}^m$ непрерывна в точке $a \in A \subseteq \mathbb{R}^n \quad \Leftrightarrow$
$\forall \varepsilon > 0 \quad \exists \delta = \delta(\varepsilon) > 0 : \quad \forall x \in A \cap U_\delta(a) \quad f(x) \in U_\varepsilon\big(f(a)\big)$.

Док–во следует из определений.

Теорема 2. Векторная фун. $f(x) = \big(f_1(x), \ldots, f_m(x)\big)$ непрерывна в т. $a \;\; \Leftrightarrow$ $\forall i = 1, \ldots, m$ функция $f_i(x)$ непрерывна в т. a.

Док–во следует из определений и теоремы 1 из §3 (см. условие с).

Теорема 3 (**локальные свойства непрерывных функций**).

а) Если функции $f_i \colon\ A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $i = 1, \ldots, k$, непрерывны в некоторой точке $a \in A$, то любая их линейная комбинация непрерывна в этой точке.

б) Если скалярные функции f, $g \colon\ A \subseteq \mathbb{R}^n \to \mathbb{R}$ непрерывны в точке $a \in A$, то их произведение fg, а при $g(a) \neq 0$ и частное f/g непрерывны в точке a.

в) Если функция $f \colon\ A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ непрерывна в точке $a \in A$, то она ограничена в пересечении множества A с некоторой окрестностью точки a.

г) Если скалярная функция $f \colon\ A \subseteq \mathbb{R}^n \to \mathbb{R}$ непрерывна в точке a и $f(a) > 0$ (или $f(a) < 0$), то существует такая окрестность $U(a)$ точки a, что функция f положительна (соответственно отрицательна) в точках множества $U(a) \cap A$.

д) Если функция $f \colon\ A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ непрерывна в точке $a \in A$, $f(A) \subseteq B$ и функция $g \colon\ B \subseteq \mathbb{R}^m \to \mathbb{R}^p$ непрерывна в точке $d = f(a)$, то сложная функция $(g \circ f)(x) = g\big(f(x)\big)$, $x \in A$, непрерывна в точке a.

Док–во следует из свойств предела функции многих переменных.

Пример 1. При $i = 1, \ldots, n$ функцию $\pi_i \colon\ (x_1, \ldots, x_n) \mapsto x_i$ из $\mathbb{R}^n$ в $\mathbb{R}$ называют **проекцией на** i**–ую координату**. Эта функция непрерывна в любой точке $a = (a_1, \ldots, a_m)$, так как $\lim\limits_{x \to a} \pi_i(x) = a_i = \pi_i(a)$.

Пример 2. Любую функцию $x \mapsto f(x)$ одного переменного можно рассматривать и как функцию $F \colon\ (x_1, \ldots, x_n) \mapsto f(x_i)$ многих переменных. При этом если f непрерывна как функция одного переменного в точке $a_i \in \mathbb{R}$, то F непрерывна как функция многих переменных в любой точке $(x_1, \ldots, a_i, \ldots, x_m) \in \mathbb{R}^n$, поскольку F есть композиция непрерывных функций: $F = f \circ \pi_i$. В частности, отсюда следует, что непрерывна функция $f(x, y) = \sin x + e^{xy}$.

Точками разрыва функции $f \colon\ A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ называют точки двух типов:
1) точки, в которых фукция определена, но не является непрерывной,
2) предельные точки области A определения функции, не принадлежащие A.

Точки разрыва могут образовывать подмножества в $\mathbb{R}^n$, которые в зависимости от их вида называют **линиями** или **поверхностями разрыва функции**.

4.2 Непрерывность функций на множествах

Функцию $f \colon\ A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, непрерывную во всех точках множества A, называют **непрерывной на** этом **множестве**. Через $C(A, \mathbb{R}^m)$ ($C(A)$ при $m = 1$) обозначают класс непрерывных на множестве A функций.

Доказанные в прошлом семестре свойства функций, непрерывных на множествах, обощается на случай функций многих переменных следующим образом.

Пусть A — подмножество $\mathbb{R}^n$. Подмножество $V \subseteq A$ называют **открытым в** A множеством, если существует такое открытое в $\mathbb{R}^n$ множество U, что $V = U \cap A$.

Теорема 4. Функция $f \colon\ A \to \mathbb{R}^m$ непрерывна на $A \subseteq \mathbb{R}^n \;\Leftrightarrow\;$ прообраз любого открытого в $\mathbb{R}^m$ множества открыт в A.

Док–во. $\Rightarrow$: Рассмотрим открытое множество $V \subseteq \mathbb{R}^m$. Если $f^{-1}(V)$ — пустое множество, то оно открыто. Пусть $a \in f^{-1}(V)$. Рассмотрим точку $f(a) \in V$. Так как V — открытое множество, то $\exists \varepsilon > 0 :\ U_\varepsilon\big(f(a)\big) \subseteq V$. Так как функция f непрерывна в точке a, то

$$\exists \delta = \delta(\varepsilon) > 0 : \quad \forall x \in A \cap U_\delta(a) \quad f(x) \in U_\varepsilon\big(f(a)\big).$$

Поэтому

$$a \in A \cap U_\delta(a) \subseteq f^{-1}\Big(U_\varepsilon\big(f(a)\big)\Big) \subseteq f^{-1}(V).$$

Имеем: $U = \bigcup_{a \in f^{-1}(V)} U_\delta(a)$ — открытое в $\mathbb{R}^n$ множество, и $f^{-1}(V) = U \cap A$, а значит, $f^{-1}(V)$ — открытое в A множество.

$\Leftarrow$: Фиксируем $a \in A$ и $\varepsilon > 0$. Рассмотрим открытое множество $V = U_\varepsilon\big(f(a)\big)$. Тогда $f^{-1}(V)$ открыто в $A \quad \Rightarrow \quad f^{-1}(V) = U \cap A \ni a \quad \Rightarrow$

$$\exists \delta > 0 :\ U_\delta(a) \subseteq U \quad \Rightarrow \quad \forall x \in A \cap U_\delta(a) \subseteq f^{-1}(V) \quad f(x) \in V = U_\varepsilon\big(f(a)\big)$$

$\Rightarrow \quad f$ непрерывна в т. a. $\triangleright$

Теорема 5. Пусть функция $f\colon\ A \to \mathbb{R}^m$ непрерывна на A.

а) Если A — компакт в $\mathbb{R}^n$, то и $f(A)$ — компакт в $\mathbb{R}^m$.

б) Если A — линейно связное множ., то и $f(A)$ — линейно связное множ.

Док–во следует из определений и теоремы 4.

а): пусть $\{G_\alpha\}$ — открытое покрытие $f(A)$. Из теоремы 4 следует, что $\forall \alpha\ \ f^{-1}(G_\alpha) = U_\alpha \cap A$, где U_α — открытое в $\mathbb{R}^n$ множество. Кроме того, $A \subseteq \bigcup_\alpha f^{-1}(G_\alpha) \subseteq \bigcup_\alpha U_\alpha$. Т.е. $\{U_\alpha\}$ — открытое покрытие компакта $A \ \Rightarrow$ существует конечное подпокрытие $\{U_{\alpha_i}\ :\ i = 1, \ldots, k\}$ множества $A \quad \Rightarrow$ $\{G_{\alpha_i} :\ i = 1, \ldots, k\}$ — конечное подпокрытие $f(A) \quad \Rightarrow \quad f(A)$ — компакт.

б): $x, y \in f(A)$ — произвольные $\quad \Rightarrow \quad$ существуют $a, b \in A : f(a) = x, f(b) = y \quad \Rightarrow$ существует кривая $\gamma : \vec{r} = \vec{r}(t), t \in [\alpha, \beta],\ \vec{r}(\alpha) = a, \vec{r}(\beta) = b,\ \forall t\ \vec{r}(t) \in A \quad \Rightarrow \quad f(\gamma)\ :\ \vec{r} = f\big(\vec{r}(t)\big), t \in [\alpha, \beta]$, — кривая, соединяющая $f\big(\vec{r}(\alpha)\big) = x$ и $f\big(\vec{r}(\beta)\big) = y,\ \forall t\ f\big(\vec{r}(t)\big) \in f(A) \quad \Rightarrow \quad f(A)$ — линейно связное множество. $\triangleright$

Пример 3. Докажем несвязность $\mathbb{R}^n \backslash S_r^{n-1}(a)$. Рассмотрим непрерывную функцию $f(x) = |x - a|$. Так как множество $f\big(\mathbb{R}^n \backslash S_r^{n-1}(a)\big) = [0, r) \cap (r, +\infty)$ не является линейно связным, то по теореме 5 п.б множество $\mathbb{R}^n \setminus S_r^{n-1}(a)$ также не линейно связно. $\triangleright$

Теорема 5 справедлива и в случае отображений произвольных хаусдорфовых пространств. При этом отображение из хаусдорфова пространства X в хаусдорфово пространство Y называют **непрерывным**, если прообраз любого открытого в Y множества открыт в X (ср. с теоремой 4).

Функция $f\colon\ A \to \mathbb{R}^m$ **равномерно непрерывна на** A, если

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

Теорема 6 (**свойства функций, непрерывных на множествах**).

а) Если функция $f\colon\ K \to \mathbb{R}^m$ непрерывна на компакте $K \subset \mathbb{R}^n$, то она равномерно непрерывна на K.

б) Если функция $f\colon K \to \mathbb{R}^m$ непрерывна на компакте $K \subset \mathbb{R}^n$, то она ограничена на K.

в) Если функция $f\colon K \to \mathbb{R}$ непрерывна на компакте $K \subset \mathbb{R}^n$, то она достигает на K своих наибольшего и наименьшего значений, т.е. существуют такие точки x_*, $x^* \in K$, что $f(x_*) \leqslant f(x) \leqslant f(x^*)$, $x \in K$.

г) Если функция $f\colon A \to \mathbb{R}$ непрерывна на линейно связном множестве $A \subseteq \mathbb{R}^n$, принимает в точках $a, b \in A$ значения $f(a) = c, f(b) = d$, то для любого числа μ, лежащего между c и d существует точка $x_\mu \in A$, для которой $f(x_\mu) = \mu$.

Док–во а) аналогично случаю функции одного переменного.

Док–во б), в) и г) следует из теоремы 5. Действительно, компакт на числовой оси — это замкнутое ограниченное множество, а линейно связное множество на числовой оси — промежуток. Если K — компакт, то и $f(K)$ является компактом, т.е. ограниченным замкнутым множеством. Ограниченность множества $f(K)$ равносильна ограниченности функции на множестве K (утверждению б)). Замкнутость множества $f(K)$ на числовой оси означает, что это множество содержит все свои предельные точки, в том числе (при $n = 1$) точную верхнюю и точную нижнюю грани. Другими словами, точная верхняя и точная нижняя грани множества $f(K)$ являются значениями функции, а это равносильно утверждению в). Наконец, если A — линейно связное множество, то и $f(A)$ — линейно связное множество, которое на числовой оси может быть только промежутком. Но если $f(A)$ — промежуток, то получаем утверждение г).

5 Дифференцируемые функции многих переменных

См. [ККЧ, §2.1–2.3, §2.4 стр. 77–79, §2.7 стр. 91–92], [З, гл.8, §2, стр. 426–430].

5.1 Дифференцируемость и дифференциал функции многих переменных

Пусть функция $f\colon A \to \mathbb{R}^m$ определена на множестве $A \subseteq \mathbb{R}^n$, $x \in A$, и $\Delta x = (\Delta x_1, \dots, \Delta x_n)^{\mathrm{T}}$ — такой вектор приращений независимых переменных, что точка $x + \Delta x$ тоже принадлежит A. В этом случае определено **полное приращение функции** f

$$\Delta f(x) = f(x + \Delta x) - f(x),$$

соответствующее приращению Δx переменных в точке x. Полное приращение векторной функции $f = (f_1, \dots, f_n)^{\mathrm{T}}$ в точке x можно выразить через полные приращения координатных функций f_1, …, f_n:

$$\Delta f(x) = (\Delta f_1(x), \dots, \Delta f_m(x))^{\mathrm{T}}.$$

Кроме того, напомним, что $|\Delta x| = \sqrt{(\Delta x_1)^2 + \dots + (\Delta x_n)^2}$.

Пусть x — предельная точка множества $A \subseteq \mathbb{R}^n$. Функцию $f\colon A \to \mathbb{R}^m$ называют **дифференцируемой в точке** x, если ее приращение в этой точке можно представить в виде суммы двух слагаемых, первое — линейное относительно Δx, второе — o–малое при $\Delta x \to 0$:

$$\Delta f(x) = L\Delta x + \alpha(\Delta x)|\Delta x|, \tag{1}$$

где L — матрица типа $m\times n$, элементы которой не зависят от Δx, а функция $\alpha(\Delta x)$ является бесконечно малой при $\Delta x \to 0$. Линейную относительно Δx часть полного приращения функции f, дифференцируемой в точке x, называют (**полным**) **дифференциалом функции** f и обозначают через df.

Заметим, что любое o–малое при $\Delta x \to 0$ представляется в виде $\alpha(\Delta x)|\Delta x|$, где $\alpha(\Delta x)$ — б. м. при $\Delta x \to 0$.

Функцию f называют **дифференцируемой в области** $X \subseteq \mathbb{R}^n$, если она дифференцируема в каждой точке этой области.

При $m = 1$ равенство (1) упрощается. В этом случае функция f скалярная, и в равенстве (1) матрица L является строкой длины n, т.е. $L = (l_1, \dots, l_n)$, а функция $\alpha(\Delta x)$ — это бесконечно малая при $\Delta x \to 0$ скалярная функция. Значит, при $m = 1$ соотношение (1) можно представить в виде

$$\Delta f(x) = l_1\Delta x_1 + \dots + l_n\Delta x_n + \alpha(\Delta x)|\Delta x|. \tag{2}$$

Теорема 1. Векторная функция $f\colon\ A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ дифференцируема в точке x тогда и только тогда, когда в этой точке дифференцируемы все ее координатные функции.

Док–во состоит в переходе от матричной формы записи условия (1) дифференцируемости векторной функции к его записи в координатной форме. $\triangleright$

5.2 Частные производные

Пусть скалярная функция многих переменных $f\colon\ A \to \mathbb{R}$ определена в некоторой окрестности точки $a = (a_1, \dots, a_n) \in A \subseteq \mathbb{R}^n$. Тогда в некоторой окрестности точки $a_1 \in \mathbb{R}$ определена функция $g(x_1) = f(x_1, a_2, \dots, a_n)$, которая получается из функции $f(x)$ при фиксированных значениях всех аргументов, кроме первого. Производную $g'(a_1)$ функции $g(x_1)$ в точке $a_1 \in \mathbb{R}$ называют **частной производной функции многих переменных** f **в точке** a по переменной x_1. Аналогично можно определить частные производные функции f и по другим переменным. Частную производную функции f в точке a по переменной x_i обозначают

$$\frac{\partial f(a)}{\partial x_i} \qquad \text{или} \qquad f'_{x_i}(a).$$

Вычисление частных производных скалярной функции сводится к дифференцированию функции одного действительного переменного, когда все переменные функции, кроме одного, "замораживаются".

Теорема 2 (**1-ое необходимое условие дифференцируемости**). Если скалярная функция $f\colon\ A \subseteq \mathbb{R}^n \to \mathbb{R}$ дифференцируема во внутренней точке $x \in A$ своей области определения, то у этой функции в точке x существуют все частные производные и дифференциал функции равен

$$df(x) = f'_{x_1}(x)\,\Delta x_1 + \dots + f'_{x_n}(x)\,\Delta x_n.$$

Док–во. В условиях теоремы представление (2) верно для любого достаточно малого приращения Δx. В частности, это представление верно, если приращение Δx имеет вид

$$\Delta x = (0, \ldots, 0, \Delta x_i, 0, \ldots, 0)^{\text{т}}, \quad \Delta x_i \neq 0,$$

где номер i выбран произвольным образом и зафиксирован. В этом случае $|\Delta x| = |\Delta x_i|$, соответствующее полное приращение есть

$$\Delta f(x) = f(x_1, \ldots, x_{i-1}, x_i + \Delta x_i, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_n),$$

а равенство (2) принимает вид

$$\Delta f(x) = l_i \Delta x_i + \alpha(\Delta x)|\Delta x_i|.$$

Разделив последнее равенство на Δx_i и перейдя к пределу при $\Delta x_i \to 0$, получим

$$\lim_{\Delta x_i \to 0} \frac{\Delta f(x)}{\Delta x_i} = l_i + \lim_{\Delta x_i \to 0} \left(\alpha(\Delta x) \frac{|\Delta x_i|}{\Delta x_i} \right) = l_i,$$

так как $\alpha(\Delta x)$ — бесконечно малая, а $\dfrac{|\Delta x_i|}{\Delta x_i}$ — ограниченная функция. Т.е. производная $f'_{x_i}(x)$ существует и равна l_i. $\triangleright$

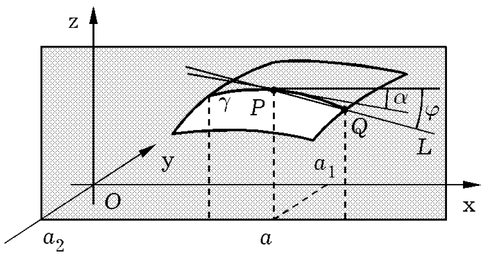


Рис. 2

Дадим **геометрическую интерпретацию частным производным**. Пусть функция двух переменных $f(x, y)$ определена в некоторой окрестности точки $a = (a_1, a_2) \in \mathbb{R}^2$. Графиком этой функции в пространстве является поверхность, которая в прямоугольной системе координат $Oxyz$ описывается уравнением $z = f(x, y)$ (см. рис. 2). Рассмотрим плоскость $y = a_2$ и систему координат $O'xz$ на ней. Линия пересечения γ этой плоскости с графиком $z = f(x, y)$ представляет собой в этой системе координат график функции $z = g(x) = f(x, a_2)$. Из курса теории функций одного действительного переменного известно, что существует касательная к линии γ в точке $(a_1, g(a_1))$ т. и т.т., к. существует производная функции $g(x)$ в точке a_1, причем значение производной $g'(a_1)$ равно тангенсу угла α, который эта касательная образует с положительным направлением оси Ox. Но $g'(a_1) = f'_x(a_1, a_2)$. Поэтому

если существует частная производная $f'_x(a_1, a_2)$, то в точке $P(a_1, a_2, f(a_1, a_2))$ существует касательная к линии пересечения поверхности $z = f(x, y)$ с плоскостью $y = a_2$, причем значение частной производной $f'_x(a_1, a_2)$ равно тангенсу угла, который эта касательная образует с положительным направлением оси Ox.

Аналогично, если существует частная производная $f'_y(a_1, a_2)$, то в точке P существует касательная к линии пересечения поверхности $z = f(x, y)$ с плоскостью $x = a_1$, причем значение частной производной $f'_y(a_1, a_2)$ равно тангенсу угла, который эта касательная образует с положительным направлением оси Oy.

5.3 Матрица Якоби

Пусть a — внутренняя точка области $A \subseteq \mathbb{R}^n$ определения функции $f\colon A \to \mathbb{R}$. **Матрицей Якоби** векторной функции $f(x)$ в точке a называют матрицу

$$f'(a) = \frac{\partial f(a)}{\partial x} = \begin{pmatrix} \frac{\partial f_1(a)}{\partial x_1} & \dots & \frac{\partial f_1(a)}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_m(a)}{\partial x_n} & \dots & \frac{\partial f_m(a)}{\partial x_n} \end{pmatrix}.$$

Часто используют запись матрицы Якоби в виде блочной матрицы–строки

$$f'(a) = \left(\frac{\partial f(a)}{\partial x_1} \ \dots \ \frac{\partial f(a)}{\partial x_n} \right)$$

или блочной матрицы–столбца

$$f'(a) = \begin{pmatrix} \frac{\partial f_1(a)}{\partial x} \\ \vdots \\ \frac{\partial f_m(a)}{\partial x} \end{pmatrix}.$$

В последнем случае каждый блок представляет собой матрицу Якоби соответствующей координатной функции.

Пример 1. Для векторной функции

$$f(x,y) = \left(xe^{-y}, x^2y^3, y\right)^{\mathrm{T}}$$

двух переменных x и y найдем все частные производные и запишем матрицу Якоби:

$$f'(x,y) = \begin{pmatrix} e^{-y} & -xe^{-y} \\ 2xy^3 & 3x^2y^2 \\ 0 & 1 \end{pmatrix}.$$

Теорема 3. Если функция $f\colon\ A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ дифференцируема во внутренней точке $x \in A$, то в этой точке существуют частные производные всех ее координатных функций по всем переменным, определена ее матрица Якоби $f'(x)$ и полное приращение $\Delta f(x)$ можно представить в виде

$$\Delta f(x) = f'(x)\Delta x + \alpha(\Delta x)|\Delta x|,$$

где $\alpha(\Delta x) \to 0$ при $\Delta x \to 0$.

Док–во. Согласно теореме 1, из дифференцируемости функции $f = (f_1, \dots, f_m)^T$ в точке x следует дифференцируемость в этой точке всех ее координатных функций f_i. По теореме 2:

$$\Delta f_i(x) = \frac{\partial f_i(x)}{\partial x_1}\Delta x_1 + \dots + \frac{\partial f_i(x)}{\partial x_n}\Delta x_n + \alpha_i(\Delta x)|\Delta x|.$$

Значит, в представлении для $\Delta f(x)$ линейное отображение имеет матрицу, составленную из значений частных производных координатных функций в точке x, т.е.

матрицу Якоби $f'(x)$, а векторная функция $\alpha(\Delta x)$ имеет своими координатными функциями функции $\alpha_i(\Delta x)$. Так как все функции $\alpha_i(\Delta x)$ являются бесконечно малыми при $\Delta x \to 0$, то и векторная функция $\alpha(\Delta x)$ является бесконечно малой при $\Delta x \to 0$. $\triangleright$

Дифференциалы независимых переменных x_i, $i = 1, \dots, n$, как и для функции одного переменного, по определению равны приращениям этих переменных: $dx_i = \Delta x_i$. С учетом этого дифференциал функции f можно записать в виде

$$df(x) = f'(x)\, dx, \qquad dx = (dx_1, \dots, dx_n)^T.$$

Если векторная функция дифференцируема в некоторой области, то во всех точках этой области существуют частные производные ее координатных функций и, следовательно, в области существует ее матрица Якоби.

6 Условия дифференцируемости функций многих переменных

См. [ККЧ, §2.4 стр. 80–82, §2.5], [З, гл.8, §2 стр. 431–432, §4 стр. 449–450].

6.1 Непрерывность и дифференцируемость

Теорема 1 (**2-ое необходимое условие дифференцируемости**). Если функция многих переменных дифференцируема в некоторой точке, то она непрерывна в этой точке.

Док–во. Пусть функция $f = (f_1, \dots, f_m)^{\mathrm{т}}$ дифференцируема в точке a. Тогда по теореме 1 из §5 все ее координатные функции f_i дифференцируемы в точке a, а их полные приращения в точке a можно записать в виде

$$\Delta f_i(a) = \sum_{k=1}^{n} l_{ik} \Delta x_k + \alpha_i(\Delta x)|\Delta x|.$$

Следовательно, существует предел $\lim_{\Delta x \to 0} \Delta f_i(a) = 0$, означающий, что функции f_i, $i = 1, \dots, m$, непрерывны в точке a. Так как все координатные функции f_i непрерывны в точке a, то и векторная функция f непрерывна в точке a. $\triangleright$

Следующие два примера показывают, что необходимые условия дифференцируемости, о которых говорится в теоремах 1 здесь и 2 из §5, не являются достаточными условиями дифференцируемости, т.е. обращения соответствующих теорем неверны.

Пример 1. Функция двух переменных

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

в начале координат имеет частные производные $f'_x(0, 0) = 0$, $f'_y(0, 0) = 0$, так как $f(0, y) = 0 = f(x, 0)$. Если бы эта функция была дифференцируемой в точке $(0, 0)$, то по теореме 1 она была бы непрерывной в этой точке $(0,0)$. Однако это не так (см. пример из §4). Следовательно, функция $f(x,y)$ не дифференцируема в точке $(0,0)$, хотя и имеет частные производные в этой точке. $\triangleright$

Пример 2. Функция двух переменных $f(x,y) = |x| + |y|$ непрерывна в точке $(0,0)$, но в этой точке не существуют ее частные производные $f'_x(0,0)$ и $f'_y(0,0)$. Поэтому данная функция не может быть дифференцируемой в точке $(0,0)$. $\triangleright$

Два необходимых условия (непрерывность в точке и существование частных производных) также не гарантируют дифференцируемости функции в точке.

Пример 3. Функция двух переменных

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

непрерывна при $x^2 + y^2 \neq 0$ как отношение двух непрерывных функций. Эта функция непрерывна и в точке $(0,0)$, поскольку из двойного неравенства

$$0 \leqslant \left| \frac{x^2 y}{x^2+y^2} \right| = \frac{|x|\,|y|}{|x|^2 + |y|^2}|x| \leqslant \frac{1}{2}|x| \leqslant \frac{1}{2}|(x,y)|$$

и свойства пределов следует существование предела

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

Следовательно, функция $f(x,y)$ непрерывна в $\mathbb{R}^2$.

В точке $(0,0)$ частные производные существуют: $f'_x(0,0) = 0$, $f'_y(0,0) = 0$, так как $f(0,y) = 0 = f(x,0)$. Если бы эта функция была дифференцируемой в этой точке, то, учитывая значение частных производных, мы имели бы равенство

$$\Delta f(0,0) = \frac{(\Delta x)^2 \Delta y}{\Delta x^2 + \Delta y^2} = \alpha(\Delta x, \Delta y)\sqrt{\Delta x^2 + \Delta y^2},$$

где $\alpha(\Delta x, \Delta y) \to 0$ при $(\Delta x, \Delta y) \to (0,0)$. Но последнее равенство при $\Delta y = \Delta x$ принимает вид

$$\frac{1}{2}\Delta x = \sqrt{2}|\Delta x|\alpha(\Delta x, \Delta x),$$

откуда $|\alpha(\Delta x, \Delta x)| = \sqrt{2}/4$, а это противоречит тому, что $\alpha(\Delta x, \Delta y)$ является бесконечно малой при $(\Delta x, \Delta y) \to (0,0)$.

6.2 Достаточное условие дифференцируемости

В случае действительных функций одного действительного переменного дифференцируемость функции в точке эквивалентна существованию в этой точке конечной производной функции. Однако уже для функций двух переменных дифференцируемость в точке не следует из существования частных производных (см. пример 1), так как существование частных производных является необходимым, но не достаточным условием дифференцируемости. Чтобы при наличии частных производных гарантировать дифференцируемость функции, нужны дополнительные условия.

Теорема 2 (**достаточное условие дифференцируемости**). Если скалярная функция $f\colon\ A \subseteq \mathbb{R}^n \to \mathbb{R}$ в некоторой окрестности точки $a \in A$ определена и имеет частные производные по всем переменным, причем все производные непрерывны в самой точке a, то функция f дифференцируема в точке a.

Док–во. Упрощая выкладки, докажем утверждение теоремы для частного случая функции двух независимых переменных, т.е. при $n = 2$. В общем случае доказательство аналогично.

Пусть заданная точка $a \in A \subseteq \mathbb{R}^2$ в системе координат Oxy имеет координаты x и y. Согласно условию теоремы, можно выбрать такое число $\delta > 0$, что функция $f(x, y)$ будет иметь частные производные в любой точке $(x + \Delta x, y + \Delta y)$ при $|\Delta x| < \delta$ и $|\Delta y| < \delta$. Приращение функции $f(x, y)$ удобно представить как сумму двух слагаемых:

$$\begin{aligned}\Delta f(x,y) &= f(x+\Delta x, y+\Delta y) - f(x,y) = \big(f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y)\big) + \\ &+ \big(f(x, y+\Delta y) - f(x,y)\big) = \Delta\varphi(x) + \Delta\psi(y),\end{aligned}$$

где

$$\varphi(t) = f(t, y+\Delta y), \qquad \psi(v) = f(x, v)$$

— дифференцируемые и, следовательно, непрерывные функции на отрезках $t \in [x,\ x+\Delta x]$ и $v \in [y,\ y+\Delta y]$ соответственно. При этом

$$\varphi'(t) = f'_x(t, y+\Delta y), \qquad \psi'(v) = f'_y(x, v).$$

Из теоремы Лагранжа для данных функций на указанных отрезках получаем:

$$\Delta f(x,y) = f'_x(x+\theta_1\Delta x, y+\Delta y)\Delta x + f'_y(x, y+\theta_2\Delta y)\Delta y, \qquad \theta_1, \theta_2 \in (0,\ 1).$$

Учитывая непрерывность частных производных в точке (x, y), получаем, что

$$f'_x(x+\theta_1\Delta x, y+\Delta y) = f'_x(x,y) + \beta, \quad f'_y(x, y+\theta_2\Delta y) = f'_y(x,y) + \gamma,$$

где β и γ — бесконечно малые при $(\Delta x, \Delta y) \to (0, 0)$. Следовательно,

$$\begin{aligned}\Delta f(x,y) &= \big(f'_x(x,y) + \beta\big)\Delta x + \big(f'_y(x,y) + \gamma\big)\Delta y = \\ &= f'_x(x,y)\Delta x + f'_y(x,y)\Delta y + \beta\Delta x + \gamma\Delta y.\end{aligned}$$

Обозначим $|\,(\Delta x, \Delta y)\,| = \sqrt{(\Delta x)^2 + (\Delta y)^2}$ через ρ. Тогда

$$\frac{|\beta\Delta x + \gamma\Delta y|}{\rho} \leqslant |\beta|\frac{|\Delta x|}{\rho} + |\gamma|\frac{|\Delta y|}{\rho} \leqslant |\beta| + |\gamma| \to 0$$

при $\rho \to 0$, так как $|\Delta x|/\rho \leqslant 1$, $|\Delta y|/\rho \leqslant 1$, а β, $\gamma \to 0$ при $\rho \to 0$. Следовательно, $\beta\Delta x + \gamma\Delta y = o(\rho)$ при $\rho \to 0$. Итак,

$$\Delta f(x,y) = f'_x(x,y)\Delta x + f'_y(x,y)\Delta y + o(\rho) \qquad \text{при} \ \ \rho \to 0.$$

Другими словами, функция f дифференцируема в точке (x, y). $\triangleright$

Если векторная функция f имеет матрицу Якоби всюду в некоторой окрестности точки x, причем все элементы матрицы Якоби непрерывны в самой точке x, то эта функция дифференцируема в точке x.

Функцию, имеющую в каждой точке области $A \subseteq \mathbb{R}^n$ непрерывные частные производные по всем переменным, называют **непрерывно дифференцируемой в области** A. Множество всех скалярных функций $f\colon A \to \mathbb{R}$, непрерывно дифференцируемых в области A, обозначают через $C^1(A)$. Аналогичное множество векторных функций $f\colon A \to \mathbb{R}^m$, имеющих в A непрерывные частными производные по всем переменным, обозначают $C^1(A, \mathbb{R}^m)$.

7 Правила дифференцирования

См. [ККЧ, §2.6, 2.7 стр. 93–94], [З, гл.8, §3, стр. 432–436].

7.1 Свойства производных и дифференциалов ФМП

На функции многих переменных можно распространить правила дифференцирования, установленные для функций одного действительного переменного.

Теорема 1 (**линейность операции дифференцирования**). Для функций $f, g\colon A \to \mathbb{R}^m$, дифференцируемых во внутренней точке $x \in A \subseteq \mathbb{R}^n$, и произвольного действительного числа c верны равенства:

а) $(cf)'(x) = cf'(x)$;

б) $(f \pm g)'(x) = f'(x) \pm g'(x)$;

в) $d(cf)(x) = cdf(x)$;

г) $d(f \pm g)(x) = df(x) \pm dg(x)$.

Теорема 2 (**правило Лейбница**). Для скалярных функций $f, g\colon A \to \mathbb{R}$, дифференцируемых во внутренней точке $x \in A \subseteq \mathbb{R}^n$, верны равенства:

а) $(fg)'(x) = f(x)g'(x) + g(x)f'(x)$;

б) $d(fg)(x) = f(x)dg(x) + g(x)df(x)$;

в) $(\dfrac{f}{g})'(x) = \dfrac{g(x)f'(x) - f(x)g'(x)}{[g(x)]^2}$ (в точках, где $g(x) \neq 0$);

г) $d(\dfrac{f}{g})(x) = \dfrac{g(x)df(x) - f(x)dg(x)}{[g(x)]^2}$ $(g(x) \neq 0)$.

Док–ва следуют из определений.

Теорема 3 (**дифференцируемость сложной функции**). Если $X \subseteq \mathbb{R}^n, Y \subseteq \mathbb{R}^m$, функция $f\colon X \to Y$ дифференцируема во внутренней точке $a \in X$, а функция многих переменных $g\colon Y \to \mathbb{R}^k$ дифференцируема во внутренней точке $b = f(a) \in Y \subseteq \mathbb{R}^m$, то сложная функция $g \circ f\colon X \to \mathbb{R}^k$ дифференцируема в точке a и выполнено равенство

$$(g \circ f)'(a) = g'(b)f'(a). \tag{1}$$

Док–во. Пусть функция g определена в окрестности $U_\sigma(b)$ точки b. Так как функция f дифференцируема в точке a, она определена в некоторой окрестности этой точки и является непрерывной функцией в точке a. Значит, согласно определению непрерывности, существует такая окрестность $U_\delta(a)$ из области определения функции f, для которой $f\big(U_\delta(a)\big) \subseteq U_\sigma(b)$. В окрестности $U_\delta(a)$ определена сложная функция $g \circ f$.

Пусть x — произвольная точка в $U_\delta(a)$ и $y = f(x)$, $z = g(y)$. Обозначим $\Delta x = x - a$, $\Delta y = y - b$, $\Delta z = z - c$, где $c = g(b)$. В силу дифференцируемости функции f в точке a имеем представление

$$\Delta y = f'(a)\Delta x + \alpha(\Delta x)|\Delta x|, \qquad \alpha(\Delta x) \to 0 \quad \text{при} \quad \Delta x \to 0. \tag{2}$$

В силу дифференцируемости g в точке b имеем аналогичное представление

$$\Delta z = g'(b)\Delta y + \beta(\Delta y)|\Delta y|, \qquad \beta(\Delta y) \to 0 \quad \text{при} \quad \Delta y \to 0. \tag{3}$$

Подставив (2) в (3), получим

$$\Delta(g \circ f)(a) = \Delta z = g'(b)f'(a)\Delta x + \gamma(\Delta x)|\Delta x|, \tag{4}$$

где

$$\gamma(\Delta x) = g'(b)\alpha(\Delta x) + \beta(\Delta y)\frac{|\Delta y|}{|\Delta x|}.$$

Имеем

$$\left.\begin{array}{llll} \Delta x \to 0 & \Rightarrow & \alpha(\Delta x) \to 0 & \\ & \Rightarrow & \Delta y \to 0 \quad \Rightarrow & \beta(\Delta y) \to 0 \end{array}\right| \;\Rightarrow\; \gamma(\Delta x) \to 0,$$

так как

$$\frac{|\Delta y|}{|\Delta x|} = \left| f'(a)\frac{\Delta x}{|\Delta x|} + \alpha(\Delta x) \right|$$

— ограниченная функция. Представление (4) означает, что функция $g \circ f$ дифференцируема в точке a. При этом произведение $g'(b)\,f'(a)$ двух матриц Якоби является, согласно (4), матрицей Якоби сложной функции $g \circ f$, т.е. имеет место равенство (1). $\triangleright$

Композицию $(g \circ f)(x) = g\big(f(x)\big)$ двух функций f и g часто задают в виде $z = g(y)$, $y = f(x)$, вводя дополнительный набор переменных $y \in R^m$. Переменные в наборе $y \in R^m$ называют **промежуточными переменными**, подчеркивая роль, которую они играют при задании сложной функции. Указанная роль проявляется и при вычислении частных производных сложной функции. Используя координатные функции $z_i(y)$, $y_s(x)$ и матрицы Якоби

$$g' = \left(\frac{\partial z_i}{\partial y_s}\right), \qquad f' = \left(\frac{\partial y_s}{\partial x_j}\right)$$

векторных функций g и f, равенство (1) матриц Якоби можно записать в координатной форме

$$\frac{\partial z_i}{\partial x_j} = \sum_{s=1}^{m} \frac{\partial z_i}{\partial y_s}\,\frac{\partial y_s}{\partial x_j} = \frac{\partial z_i}{\partial y_1}\,\frac{\partial y_1}{\partial x_j} + \cdots + \frac{\partial z_i}{\partial y_m}\,\frac{\partial y_m}{\partial x_j}, \quad i = 1, \ldots, k,\ j = 1, \ldots, n. \tag{5}$$

Получены формулы для вычисления частных производных координатных функций сложной функции. Отметим, что частные производные в (5) вычисляются в соответствующих точках, а именно: $\dfrac{\partial z_i}{\partial x_j}$ и $\dfrac{\partial y_s}{\partial x_j}$ — в точке a, а $\dfrac{\partial z_i}{\partial y_s}$ — в точке $b = f(a)$.

В равенствах (5) следует обратить внимание на то, как в них входят промежуточные и остальные переменные. Запись частных производных сложной функции в виде (5) называют **правилом дифференцирования сложной функции** или **цепным правилом**.

Рассмотрим некоторые частные случаи дифференцирования сложных функций при различных значениях n, m и k. Будем предполагать, не оговаривая этого специально, что условия теоремы 3 выполнены для этих функций в соответствующих точках.

Если $n = 1$ (или $m = 1$), то у функции f (соответственно g) будет всего лишь один аргумент и частная производная будет фактически обыкновенной производной. Это должно быть отражено в обозначениях производных. Например, если $z = g(y)$, $y = f(t) = \big(f_1(t), \ldots, f_m(t)\big)^{\mathrm{T}}$, где $z \in \mathbb{R}$, $t \in \mathbb{R}$, $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$, то сложная функция $z = g\big(f(t)\big)$ зависит только от одного переменного t и

$$\frac{dz}{dt} = \sum_{s=1}^{m} \frac{\partial z}{\partial y_s}\,\frac{dy_s}{dt} = \frac{\partial z}{\partial y_1}\,\frac{dy_1}{dt} + \cdots + \frac{\partial z}{\partial y_m}\,\frac{dy_m}{dt}, \tag{6}$$

где частные и обыкновенные производные вычисляются в соответствующих точках.

Производную сложной функции $z = g\big(f(t)\big)$ в случае $n = 1$, $k = 1$ (т.е. действительной функции действительного переменного), вычисляемую по формуле (6), называют **полной производной функции** $g\big(f(t)\big)$.

Пример 1. Найдем полную производную сложной функции $z = f(t, x, y)$, $x = \sin t$, $y = \cos t$, предполагая, что функция $f\colon\ \mathbb{R}^3 \to \mathbb{R}$ дифференцируема в $\mathbb{R}^3$.

В данном случае промежуточных переменных два, но удобно ввести третье промежуточное переменное $w = t$ и рассмотреть сложную функцию $z = f(w, x, y)$, $w = t$, $x = \sin t$, $y = \cos t$. Используя правило дифференцирования сложной функции, находим

$$\frac{dz}{dt} = \frac{\partial f}{\partial w}\,\frac{dw}{dt} + \frac{\partial f}{\partial x}\,\frac{dx}{dt} + \frac{\partial f}{\partial y}\,\frac{dy}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}\cos t - \frac{\partial f}{\partial y}\sin t,$$

где частные производные функции f вычисляются в точке $(t, \sin t, \cos t)$.

Следствие 1 (инвариантность формы первого дифференциала). Во внутренних точках области определения функции формула для дифференциала $dz = g'(y)dy$ одинакова для случая, когда y — вектор аргументов функции g, и для случая, когда y — векторная функция каких–либо других аргументов.

Док–во. Пусть функция $f\colon\ X \to Y$ дифференцируема во внутренней точке $a \in X$, а функция $g\colon\ Y \to \mathbb{R}^k$ дифференцируема во внутренней точке $b = f(a) \in Y$. Согласно теореме 3, композиция $g \circ f$ двух функций дифференцируема в точке

a, а ее дифференциал в точке a в соответствии с правилом дифференцирования сложной функции имеет вид

$$d(g \circ f) = (g \circ f)'(a)\,dx = g'(b)f'(a)\,dx.$$

Вводя набор промежуточных переменных $y = f(x)$, запишем композицию $g \circ f$ в виде $z = g(y)$, $y = f(x)$. Тогда $dy = f'(a)\,dx$ — дифференциал функции f, а $dz = d(g \circ f)$ — дифференциал композиции $g \circ f$. Т.е. получается указанная формула при $y = b$. Точно такой же вид имеет дифференциал функции $z = g(y)$. $\triangleright$

Геометрическая интерпретация этого следствия. Формула для дифференциала не меняется при переходе от координат y к координатам x. Т.е. дифференциал не зависит от выбора системы координат и характеризует собственно геометрические свойства функции, а не ее представление в координатах.

8 Частные производные второго порядка

См. [ККЧ, §3.1].

8.1 Матрица Гессе

Предположим, что скалярная функция $f\colon\ \mathbb{R}^n \to \mathbb{R}$ во всех точках некоторой окрестности $U_\delta(a)$ точки a имеет частную производную $f'_{x_i}(x)$. Эта частная производная сама является функцией многих переменных, определенной в окрестности $U_\delta(a)$, и может оказаться, что она имеет частную производную в точке a, например по переменному x_j. Частную производную

$$\left(f'_{x_i}\right)'_{x_j}(a) = \frac{\partial}{\partial x_j}\left(\frac{\partial f(x)}{\partial x_i}\right)\bigg|_{x=a}$$

функции $f'_{x_i}(x)$ называют **частной производной второго порядка** функции $f(x)$ в точке a по переменным x_i и x_j и обозначают

$$\frac{\partial^2 f(a)}{\partial x_j \partial x_i} \qquad \text{или} \qquad f''_{x_i x_j}(a).$$

Производные $f'_{x_i}(x)$ в связи с этим называют **частными производными первого порядка**. При $i \neq j$ частные производные второго порядка называют **смешанными**. При $j = i$ используют обозначения

$$\frac{\partial^2 f(a)}{\partial x_i^2}, \qquad f''_{x_i x_i}(a) \qquad \text{или} \qquad f''_{x_i^2}(a).$$

Если для скалярной функции $f = f(x_1, \ldots, x_n)$ в точке x существуют все частные производные второго порядка, то из них можно составить квадратную матрицу порядка n

$$f''(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \ldots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \ldots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \ldots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix},$$

которую называют **матрицей Гессе** функции f в точке x.

Пример 1. Найдем все частные производные второго порядка для функции двух переменных $f(x,y)=3x^2y+y^2+5$ и запишем для нее матрицу Гессе:

$$f''(x)=\begin{pmatrix} 6y & 6x \\ 6x & 2 \end{pmatrix}.$$

Обратим внимание на то, что в ответе получилась симметрическая матрица, т.е. в данном случае значение смешанных производных второго порядка не зависит от порядка дифференцирования (последовательности, в которой вычисляются частные производные первого порядка). $\triangleright$

8.2 Совпадение смешанных производных

Как показывает рассмотренный пример, в некоторых случаях смешанные производные, которые отличаются лишь порядком дифференцирования, совпадают. Следующая теорема дает достаточные условия для такого совпадения, т.е. условия, при выполнении которых значение смешанной производной не зависит от порядка дифференцирования.

Функцию $f(x)$ n переменных называют **непрерывной в точке** $a=(a_1,\ldots,a_n)$ **по части переменных** x_i, x_j, если функция двух переменных

$$g(x_i,x_j)=f(a_1,\ldots,x_i,\ldots,x_j,\ldots,a_n)$$

непрерывна в точке $x_i=a_i$, $x_j=a_j$.

Задача 1. Выведите из теоремы 4 из §3, что если функция непрерывна в точке, то она непрерывна в этой точке по любой части переменных.

Теорема 1. Пусть скалярная функция $f(x_1,\ldots,x_n)$ $(n>1)$ в некоторой окрестности точки $a\in\mathbb{R}^n$ имеет частные производные f'_{x_i}, f'_{x_j} $(i\neq j)$, $f''_{x_ix_j}$ и $f''_{x_jx_i}$, смешанные производные $f''_{x_ix_j}$ и $f''_{x_jx_i}$ непрерывны в точке a по части переменных x_i и x_j. Тогда $f''_{x_ix_j}(a)=f''_{x_jx_i}(a)$.

Док–во. Значения всех переменных, кроме x_i и x_j, можно считать фиксированными. Поэтому можно вести речь о функции, имеющей только два аргумента x_i и x_j, которые удобно переобозначить: $x_i=x$, $x_j=y$, $a=(p;q)$.

Выберем такое число $\delta>0$, что при $|\Delta x|<\delta$, $|\Delta y|<\delta$ точка $(p+\Delta x;q+\Delta y)$ попадает в окрестность U точки $(p;q)$, где функция $f=f(x,y)$ имеет частные производные первого порядка и смешанные производные f''_{xy} и f''_{yx}. В квадрате $|\Delta x|<\delta$, $|\Delta y|<\delta$ определена функция

$$g(\Delta x,\Delta y)=f(p{+}\Delta x,q{+}\Delta y)-f(p{+}\Delta x,q)-f(p,q{+}\Delta y)+f(p,q).$$

Для функции $\varphi(x)=f(x,q+\Delta y)-f(x,q)$ имеем $g(\Delta x,\Delta y)=\varphi(p+\Delta x)-\varphi(p)$. Функция $\varphi(x)$ на отрезке $[p,\,p+\Delta x]$ имеет производную

$$\varphi'(x)=f'_x(x,q+\Delta y)-f'_x(x,q)$$

и потому непрерывна на этом отрезке. Следовательно, к функции $\varphi(x)$ на указанном отрезке можно применить теорему Лагранжа:

$$g(\Delta x, \Delta y) = \varphi'(p+\theta\Delta x)\Delta x = \left[f'_x(p+\theta\Delta x, q+\Delta y) - f'_x(p+\theta\Delta x, q)\right]\Delta x, \quad \theta \in (0,\, 1).$$

Выражение в квадратных скобках представляет собой приращение функции $\lambda(y) = f'_x(p+\theta\Delta x, y)$ одного переменного на отрезке $[q,\, q+\Delta y]$. На отрезке $[q,\, q+\Delta y]$ функция $\lambda(y)$ имеет производную

$$\lambda'(y) = f''_{xy}(p+\theta\Delta x, y)$$

и является поэтому непрерывной на этом отрезке. Значит, и к этой функции на указанном отрезке можно применить теорему Лагранжа:

$$\lambda(q+\Delta y) - \lambda(q) = \lambda'(q+\theta_1\Delta y)\Delta y = f''_{xy}(p+\theta\Delta x, q+\theta_1\Delta y)\Delta y, \quad \theta_1 \in (0,\, 1).$$

В результате находим, что

$$g(\Delta x, \Delta y) = f''_{xy}(p+\theta\Delta x, q+\theta_1\Delta y)\Delta x\Delta y, \quad \theta, \theta_1 \in (0,\, 1). \tag{1}$$

Равенство (1) было получено в результате двукратного применения теоремы Лагранжа, причем сперва она применялась по переменному x, затем — по переменному y. Но те же рассуждения можно повторить, поменяв лишь порядок переменных. Тогда получим равенство, аналогичное (1), но включающее другую смешанную производную:

$$g(\Delta x, \Delta y) = f''_{yx}(p+\theta_3\Delta x, q+\theta_2\Delta y)\Delta y\Delta x, \quad \theta_2, \theta_3 \in (0,\, 1). \tag{2}$$

Соединяя равенства (1) и (2), получаем

$$f''_{xy}(p+\theta\Delta x, q+\theta_1\Delta y) = f''_{yx}(p+\theta_3\Delta x, q+\theta_2\Delta y).$$

Переходя в этом равенстве к пределу при $(\Delta x; \Delta y) \to (0; 0)$ и используя непрерывность в точке $(p; q)$ смешанных производных, заключаем, что $f''_{xy}(p, q) = f''_{yx}(p, q)$. $\triangleright$

Условие непрерывности смешанных производных в доказанной теореме нельзя опустить: при нарушении этого условия смешанные частные производные могут отличаться.

Пример 2. Покажем, что смешанные частные производные второго порядка функции двух переменных

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

различны в точке $(0; 0)$. Найдем частные производные первого порядка:

$$\begin{aligned} f'_x(x,y) &= \begin{cases} y\left(\dfrac{x^2-y^2}{x^2+y^2} + \dfrac{4x^2y^2}{(x^2+y^2)^2}\right), & x^2+y^2 \neq 0; \\ \qquad 0, & x=y=0, \end{cases} \\ f'_y(x,y) &= \begin{cases} x\left(\dfrac{x^2-y^2}{x^2+y^2} - \dfrac{4x^2y^2}{(x^2+y^2)^2}\right), & x^2+y^2 \neq 0; \\ \qquad 0, & x=y=0. \end{cases} \end{aligned}$$

Имеем $f'_x(0, y) = -y$, откуда $f''_{xy}(0, 0) = -1$. Аналогично $f'_y(x, 0) = x$, и поэтому $f''_{yx}(0, 0) = 1$. Следовательно, $f''_{xy}(0, 0) \neq f''_{xy}(0, 0)$, что связано с нарушением условия непрерывности смешанных производных в точке $(0; 0)$ (см. теорему 1).

9 Производные и дифференциалы высших порядков

См. [ККЧ, §3.2–3.5].

9.1 Частные производные высших порядков

Частные производные высшего порядка вводятся так же, как и частные производные второго порядка. Частную производную k-го порядка ($k > 1$) функции многих переменных определяют как частную производную первого порядка от некоторой частной производной ($k-1$)-го порядка этой функции.

Например, для функции $f = f(x, y)$ двух переменных могут существовать восемь частных производных третьего порядка: $f'''_{xxx} \equiv f'''_{x^3}$, $f'''_{xxy} \equiv f'''_{x^2y}$, f'''_{xyx}, $f'''_{yxx} \equiv f'''_{yx^2}$ и т.д. Возможные другие обозначения: $f''''_{xxyy} \equiv f^{(4)}_{x^2y^2} \equiv f^{\mathrm{IV}}_{x^2y^2}$ и т.п.

Пусть U — область в $\mathbb{R}^n$. Функцию $f\colon\ U \to \mathbb{R}^m$, у которой все частные производные до порядка k включительно у всех ее координатных функций непрерывны в U, называют k **раз непрерывно дифференцируемой** в области U. Говорят также, что она имеет k–й **порядок гладкости** в области U или что она принадлежит **классу** C^k в области U. Через $C^k(U, \mathbb{R}^m)$ обозначают множество k раз непрерывно дифференцируемых в области U векторных функций из U в $\mathbb{R}^m$, а через $C^k(U)$ — множество k раз непрерывно дифференцируемых в области U скалярных функций. Множества $C^k(U, \mathbb{R}^m)$, $C^k(U)$ называют классами. В этих обозначениях допусти́м случай $k = \infty$, означающий, что соответствующая функция имеет непрерывные частные производные любого порядка. Такую функцию обычно называют **гладкой** или **бесконечно дифференцируемой функцией**.

Для функций k–го порядка гладкости значение частной производной порядка не выше k не зависит от последовательности, в которой выполняется дифференцирование. Например, при $k = 4$ равенство $f^{\mathrm{IV}}_{xyxy} = f^{\mathrm{IV}}_{xxyy}$ можно рассматривать как равенство частных производных по переменному y функций f'''_{xyx} и f'''_{xxy}, или g''_{xy} и g''_{yx}, где $g = f'_x$. Но $g''_{xy} = g''_{yx}$, так как эти частные производные, являющиеся частными производными функции f'_x, непрерывны. Значит, $f'''_{xyx} = f'''_{xxy}$ и $f^{\mathrm{IV}}_{xyxy} = f^{\mathrm{IV}}_{xxyy}$.

Свойство независимости частных производных от порядка дифференцирования, которым обладают функции классов C^k, позволяет в обозначениях частных производных группировать одни и те же переменные и тем самым упрощать запись. Например, для скалярной функции $f(x_1, \ldots, x_n)$ частные производные r-го порядка обозначают следующим образом:

$$D^\sigma f \equiv \frac{\partial^r f(x)}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}} \equiv \frac{\partial^{|\sigma|} f(x)}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}},$$

где $\sigma = (i_1, \ldots, i_n)$ и $r = |\sigma|$ обозначает сумму $i_1 + \cdots + i_n$ всех индексов.

9.2 Дифференциалы высших порядков

Пусть скалярная функция f дифференцируема в окрестности точки $x \in \mathbb{R}^n$. Тогда ее дифференциал

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

как функция от x с постоянными $dx_1, \ldots, dx_n$ может оказаться дифференцируемой функцией в точке x. В этом случае выражение

$$d\big(df(x)\big) = d\left(\sum_{j=1}^{n} \frac{\partial f(x)}{\partial x_j}\, dx_j\right) = \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f(x)}{\partial x_j \partial x_i}\, dx_i\, dx_j, \tag{1}$$

представляющее собой дифференциал от дифференциала функции $f(x)$, называют **дифференциалом второго порядка** функции $f(x)$ в точке x и обозначают $d^2 f(x)$. В этой связи дифференциал $df(x)$ называют **дифференциалом первого порядка** функции f.

Из 1-ого необходимого условия дифференцируемости следует, что для существования в точке x дифференциала второго порядка функции f необходимо существование всех частных производных второго порядка этой функции в точке x. Достаточным же условием существования дифференциала является условие, что указанные производные являются непрерывными функциями в точке x. В частности, если $f \in C^2(U)$, где U — некоторая окрестность точки $x \in \mathbb{R}^n$, то в этой окрестности существуют непрерывные частные производные первого и второго порядка, а значит, в U существуют как дифференциал первого порядка df, так и дифференциал второго порядка $d^2 f$.

Отметим, что дифференциал первого порядка является линейной функцией переменных $dx = (dx_1, \ldots, dx_n)^{\mathrm{T}}$, а дифференциал второго порядка, согласно представлению (1), является квадратичной формой относительно этих переменных. В матричной записи дифференциал второго порядка имеет вид

$$d^2 f(x) = dx^{\mathrm{T}} f''(x)\, dx,$$

где $f''(x)$ — матрица Гессе функции f.

Дифференциал второго порядка, зависящий от набора независимых переменных x и вектора их приращений dx (дифференциалов независимых переменных), может оказаться дифференцируемой функцией по совокупности переменных x. Повторяя последовательно процесс вычисления дифференциалов, приходим к **дифференциалу** функции **k-го порядка**, который является дифференциалом первого порядка от дифференциала $(k-1)$-го порядка функции f:

$$d^k f(x) = d\big(d^{k-1} f(x)\big).$$

Достаточным условием существования дифференциала k-го порядка в области U является k-й порядок гладкости функции в этой области, т.е. условие $f \in C^k(U)$.

С помощью оператора

$$\frac{\partial}{\partial x_1}\, dx_1 + \cdots + \frac{\partial}{\partial x_n}\, dx_n$$

дифференциал k-го порядка функции $f \in C^k$ удобно записывать в виде

$$d^k f(x) = \Big(\frac{\partial}{\partial x_1}\, dx_1 + \dots + \frac{\partial}{\partial x_n}\, dx_n\Big)^k f(x).$$

Здесь выражение в скобках возводится в степень k по обычным алгебраическим правилам, причем полагают, что

$$\Big(\frac{\partial}{\partial x_i}\, dx_i\Big)^m = \frac{\partial^m}{\partial x_i^m}\, dx_i^m, \qquad \text{где} \;\; dx_i^m = (dx_i)^m.$$

В случае функции двух независимых переменных $z = f(x, y) \in C^2$ имеем

$$d^2 f(x, y) = \Big(\frac{\partial}{\partial x}\, dx + \frac{\partial}{\partial y}\, dy\Big)^2 f(x, y) = f''_{xx}\, dx^2 + 2f''_{xy}\, dxdy + f''_{yy}\, dy^2.$$

Пример 1. Для функции $z = x^2 y^3$ ее второй дифференциал имеет вид

$$d^2 z = 2y^3 dx^2 + 12xy^2 dxdy + 6x^2 y\, dy^2.$$

А ее матрица Гессе есть

$$\begin{pmatrix} 2y^3 & 6xy^2 \\ 6xy^2 & 6x^2 y \end{pmatrix}. \qquad \triangleright$$

Отметим, что дифференциал второго порядка не обладает свойством *инвариантности формы записи*. Например, если x, y — функции каких–либо других аргументов, то вместо формулы (2) получаем:

$$d^2 f(x, y) = d\left(f'_x\, dx + f'_y\, dy\right) = f''_{xx}\, dx^2 + 2f''_{xy}\, dxdy + f''_{yy}\, dy^2 + f'_x\, d^2x + f'_y\, d^2y.$$

9.3 Формула Тейлора

Напомним, что если у функции $g(t)$ одного переменного t на отрезке $[0,\, T]$ существует конечная производная $(m+1)$–го порядка, то при любом $t \in [0,\, T]$ имеет место формула Маклорена с остаточным членом в форме Лагранжа:

$$g(t) = g(0) + \frac{g'(0)}{1!} t + \frac{g''(0)}{2!} t^2 + \dots + \frac{g^{(m)}(0)}{m!} t^m + \frac{g^{(m+1)}(\theta t)}{(m+1)!} t^{m+1}, \qquad (2)$$

где $\theta \in (0,\, 1)$ — некоторое число.

Следующая теорема обобщает формулу Тейлора на случай скалярной функции многих переменных.

Теорема 1 (**теорема Тейлора**). Пусть U — окрестность точки $a \in \mathbb{R}^n$, $f \in C^{m+1}(U)$, а отрезок, соединяющий точки $a = (a_1, \dots, a_n)$ и $a + \Delta x = (a_1 + \Delta x_1, \dots, a_n + \Delta x_n)$, содержится в U. Тогда существует такое число $\theta \in (0,\, 1)$, что для функции $f(x)$ имеет место **формула Тейлора**:

$$f(a + \Delta x) = f(a) + \sum_{k=1}^{m} \frac{d^k f(a)}{k!} + \frac{d^{m+1} f(a + \theta \Delta x)}{(m+1)!} \qquad \text{при} \;\; dx = \Delta x. \qquad (3)$$

Док–во. Рассмотрим функцию одного переменного $g(t) = f(a+t\Delta x)$. По условию $a+t\Delta x \in U$ при $t \in [0,\, 1]$. Поэтому функция $g(t)$ имеет конечную производная $(m+1)$–го порядка на отрезке $[0,\, 1]$, а значит для нее справедливо равенство (2) при $t = 1$. Покажем, что это равенство можно преобразовать в равенство (3). Действительно,

$$g(1) = f(a + 1 \cdot \Delta x) = f(a + \Delta x), \qquad g(0) = f(a).$$

Для вычисления производных функции $g(t)$ рассмотрим ее как сложную функцию $g(t) = f(x)$, $x = a + t\Delta x$. Согласно свойству инвариантности формы записи дифференциала первого порядка,

$$dg(t) = df(x) = \Big(\frac{\partial}{\partial x_1}dx_1 + \cdots + \frac{\partial}{\partial x_n}dx_n\Big)f(x) = \Big(\frac{\partial}{\partial x_1}\Delta x_1 + \cdots + \frac{\partial}{\partial x_n}\Delta x_n\Big)f(x)dt,$$

так как при фиксированных a и Δx имеем $dx_i = d(a_i + t\Delta x_i) = \Delta x_i dt$, $i = 1, \ldots, n$. Повторяя процесс дифференцирования, находим

$$d^k g(t) = d\big(d^{k-1} f(u)\big) = \Big(\frac{\partial}{\partial x_1}\Delta x_1 + \cdots + \frac{\partial}{\partial x_n}\Delta x_n\Big)^k f(x)dt^k.$$

Из найденных дифференциалов функции $g(t)$ при $t = 0$ (что равносильно $x = a$) получаем

$$g^{(k)}(0) = d^k f(a)\Big|_{dx_i = \Delta x_i}, \qquad k = 1, \ldots, m, \quad i = 1, \ldots, n,$$

и при $k = m + 1$ и $t = \theta \in (0,\, 1)$

$$g^{(m+1)}(\theta) = d^{m+1} f(a + \theta\Delta x)\Big|_{dx_i = \Delta x_i}, \qquad i = 1, \ldots, n.$$

Заменяя в (2) t на 1 и производные функции g согласно полученным формулам, приходим к равенству (3). $\triangleright$

Как и в случае функций одного переменного, при $a = 0$ формулу Тейлора (3) часто называют **формулой Маклорена**. Число m, определяющее количество слагаемых в формуле Тейлора, называют **порядком формулы Тейлора**. Последнее слагаемое в формуле Тейлора (3) называют **остаточным членом в форме Лагранжа**. Остаточный член можно также записать в виде $o(|\Delta x|^m)$, и тогда его называют **остаточным членом в форме Пеано**. Таким образом, формула Тейлора с остаточным членом в форме Пеано имеет вид

$$f(x + \Delta x) = f(x) + \sum_{k=1}^{m} \frac{d^k f(x)}{k!} + o(|\Delta x|^m), \qquad dx = \Delta x. \tag{4}$$

Пример 2. Запишем формулу (3) для функции двух переменных $f(x, y)$ в случае $m = 2$:

$$f(x + \Delta x, y + \Delta y) = f(x, y) + f'_x(x, y)\Delta x + f'_y(x, y)\Delta y +$$
$$+ \frac{1}{2}\Big(f''_{x^2}(x, y)(\Delta x)^2 + 2f''_{xy}(x, y)\Delta x\Delta y + f''_{y^2}(x, y)(\Delta y)^2\Big) + \frac{1}{3!}d^3 f(x + \theta\Delta x, y + \theta\Delta y).$$

Формула Тейлора (4) с остаточным членом в форме Пеано справедлива при более слабых предположениях о функции f, чем формула Тейлора (3) с остаточным членом в форме Лагранжа: она справедлива, если функция f имеет непрерывные частные производные до порядка m включительно в окрестности точки x и частные производные порядка $m+1$ в окрестности точки x, непрерывные в самой точке x.

Следующий пример демонстрирует, что формула Тейлора (3) с остаточным членом в форме Лагранжа не распространяется на случай векторной функции.

Пример 3. Для функции $\varphi(t) = (\cos t; \sin t; t)^{\mathrm{T}}$, имеющей непрерывные производные любого порядка, формула Тейлора нулевого порядка ($m = 0$ в формуле (3)) на отрезке $[0,\, 2\pi]$ ($a = 0$, $\Delta x = 2\pi$) должна иметь вид

$$\varphi(2\pi) = \varphi(0) + d\varphi(\theta \cdot 2\pi).$$

Но $\varphi(2\pi) - \varphi(0) = (0; 0; 2\pi)^{\mathrm{T}}$, в то время как $d\varphi(t) = (-\sin t; \cos t; 1)^{\mathrm{T}}\, dt$. Легко понять, что в любой точке $t \in [0,\, 2\pi]$ хотя бы одна из первых двух координат дифференциала $d\varphi(t)$ не обращается в нуль. Поэтому ни в одной точке $t \in [0,\, 2\pi]$ значение $d\varphi(t)$ не может быть равно $\varphi(2\pi) - \varphi(0)$. $\triangleright$

При $m = 1$ формула Тейлора (4) с остаточным членом в форме Пеано в окрестности точки $a \in \mathbb{R}^n$ имеет вид

$$f(a + \Delta x) = f(a) + df(a) + o(|\Delta x|) = f(a) + f'(a)\, dx + o(|\Delta x|).$$

Отбрасывая в этой формуле остаточный член, получаем приближенное представление l_f функции f в окрестности точки a. Его обычно записывают в виде

$$l_f(x) = f(a) + f'(a)(x - a)$$

и называют **линейным** (или **первым**) **приближением функции** f в точке a.

Линейные приближения функций широко используют при изучении локальных свойств (т.е. в окрестности заданной точки) математических моделей объектов, которые описываются сложными функциональными зависимостями.

Формула Тейлора с остаточным членом в форме Лагранжа порядка $m = 0$ имеет вид

$$f(a + \Delta x) = f(a) + f'(a + \theta\Delta x)\Delta x, \quad 0 < \theta < 1,$$

и известна как **формула конечных приращений**.

Дифференциалы функций многих переменных и формула Тейлора для функций многих переменных могут использоваться в **приближенных вычислениях** примерно так же, как и в случае действительных функций одного действительного переменного. Применение аппарата функций многих переменных предпочтительнее, когда в вычисляемом выражении есть несколько величин, которые могут меняться независимо друг от друга. Покажем это на примере.

Пример 4. Вычислим приближенное значение $\sqrt{12{,}01^2 + 4{,}98^2}$. Это выражение можно рассматривать как значение функции $f(x, y) = \sqrt{x^2 + y^2}$ в точке с координатами $x = 12{,}01$, $y = 4{,}98$. Полагаем $a = 12$, $b = 5$, $\Delta x = 0{,}01$, $\Delta y = -0{,}02$. Тогда

$$f(x, y) = f(a + \Delta x, b + \Delta y) \approx f(a, b) + f'_x(a, b)\Delta x + f'_y(a, b)\Delta y.$$

В точке $(12;5)$ значение функции равно $\sqrt{12^2+5^2}=13$. Вычисляем частные производные функции:

$$f'_x(x,y)=\frac{x}{\sqrt{x^2+y^2}},\qquad f'_y(x,y)=\frac{y}{\sqrt{x^2+y^2}}.$$

В результате получаем $\sqrt{12{,}01^2+4{,}98^2}\approx 13{,}0015$.

Полученное значение отличается от точного лишь в пятом знаке после запятой. Если необходимо оценить точность R полученного приближения, можно использовать формулу Тейлора с остаточным членом в форме Лагранжа. В рассматриваемом случае

$$R=\frac{1}{2}d^2f(a+\theta\Delta x,b+\theta\Delta y)=\frac{y^2(\Delta x)^2-2xy\Delta x\Delta y+x^2(\Delta y)^2}{2(x^2+y^2)^{3/2}},$$

где $x=a+\theta\Delta x$, $y=b+\theta\Delta y$, $0<\theta<1$. Используя неравенства $x\leqslant 13$, $y\leqslant 5$ в числителе дроби и оценивая снизу знаменатель дроби с помощью неравенства $\sqrt{x^2+y^2}>12$, получаем

$$R<\frac{25(\Delta x)^2+130|\Delta x\Delta y|+169(\Delta y)^2}{2\cdot 12^3}=\frac{961}{2\cdot 12^3}\cdot 10^{-4}<0{,}3\cdot 10^{-4}.$$

Отметим, что значение R, вычисленное при $x=12$, $y=5$, которое, очевидно, можно рассматривать как приближение остаточного члена в форме Лагранжа, равно $0{,}19\cdot 10^{-4}$. Это близко к полученной нами оценке сверху. $\triangleright$

10 Неявные функции

См. [ККЧ, §4.1, 4.2].

10.1 Постановка задачи

Рассмотрим систему уравнений

$$\begin{cases} f_1(x_1,\ldots,x_{n+m})=0,\\ \qquad\ldots\\ f_m(x_1,\ldots,x_{n+m})=0,\end{cases}\tag{1}$$

где функции $f_1(x)$, …, $f_m(x)$ определены в некоторой области $X\subseteq\mathbb{R}^{n+m}$. Предположим, что эта система разрешима относительно части переменных, например x_{n+1}, …, x_{n+m}. Разрешимость системы в данном случае следует понимать как существование такой векторной функции $y=h(x)$, где $x=(x_1,\ldots,x_n)\in\mathbb{R}^n$, $y=(x_{n+1},\ldots,x_{n+m})\in\mathbb{R}^m$, что

$$y=h(x)\ \Leftrightarrow\ f(x,y)=0,\qquad f(x,y)=\bigl(f_1(x,y),\ldots,f_m(x,y)\bigr).$$

В этом случае о функции $h(x)$ говорят как о **неявной функции**, или **неявно заданной функции**. Отметим, что термин "неявная функция" относится не к виду или структуре функции, а лишь к способу ее задания. Переменные x_{n+1}, …, x_{n+m}

называют **зависимыми**, переменные $x_1, \ldots, x_n$ — **свободными** (или **независимыми**).

В математическом анализе важную роль играют условия, при выполнении которых система уравнений вида (1) разрешима относительно части переменных. Отметим, что весьма не просто определить, разрешима ли система в заданной области. Условия же локальной разрешимости системы (1), т.е. ее разрешимости в некоторой окрестности заданной точки, достаточно просты и связаны в первую очередь с дифференциальными свойствами функций $f_1, \ldots, f_m$. Рассмотрим несколько примеров.

Пример 1. Из уравнения $x^2 + y^2 - 9 = 0$ можно выразить y через x: $y = \pm\sqrt{9 - x^2}$. При этом каждому значению $x \in (-3, 3)$ соответствуют два значения y. Мы получаем из уравнения две функции, определенные на отрезке $[-3, 3]$: $h_1(x) = \sqrt{9 - x^2}$ и $h_2(x) = -\sqrt{9 - x^2}$.

Пример 2. Уравнение $x^2 + y^2 + 1 = 0$ не имеет решений и потому не задает ни одно из переменных как функцию от другого.

Пример 3. Задает ли уравнение $e^y + y - x + \ln x = 0$ переменное y как функцию переменного x или переменное x как функцию переменного y? Ответить на этот вопрос сложно, так как не ясно, каким образом одно из переменных можно выразить через другое, преобразуя это уравнение. $\triangleright$

Задача 1. Дайте геометрическую интерпретацию условий существования неявной функции.

10.2 Случай уравнения с двумя неизвестными

Изучение вопроса начнем с частного случая, когда $n = m = 1$, т.е. переменные x и y являются скалярными. Пусть $z = f(x, y)$ — функция двух переменных. Уравнение $f(x, y) = 0$ будем называть **уравнением с двумя неизвестными**. Остановимся на вопросе о том, при каких условиях это уравнение определяет переменное y как неявную функцию переменного x.

Теорема 1 (**о неявной функции**). Пусть уравнение $f(x, y) = 0$, x, $y \in \mathbb{R}$, в точке (a, b) и в ее окрестности U удовлетворяет следующим трем условиям:

а) $f(a, b) = 0$;
б) $f \in C^1(U)$;
в) $f'_y(a, b) \neq 0$.

Тогда точка (a, b) имеет окрестность вида

$$P = \{|x - a| < \delta_x,\ |y - b| < \delta_y\},$$

в которой уравнение $f(x, y) = 0$ разрешимо относительно y, т.е.

$$\exists h \in C^1(T), \quad T = (a - \delta_x,\ a + \delta_x): \qquad \forall (x, y) \in P \quad \Big(f(x, y) = 0 \ \Leftrightarrow\ y = h(x)\Big).$$

При этом

$$h'(x) = -\frac{f'_x(x, y)}{f'_y(x, y)}\bigg|_{y = h(x)}. \tag{2}$$

Док–во. Будем для определенности считать, что $f'_y(a,b) > 0$, так как в противном случае вместо функции $f(x,y)$ можно рассмотреть функцию $-f(x,y)$. В силу непрерывности частных производных существует такое число $\Delta > 0$, что при $|x-a| \leqslant \Delta$ и $|y-b| \leqslant \Delta$ точка (x,y) попадет в U и будет верно неравенство $f'_y(x,y) > 0$. Тогда при фиксированном $x \in [a-\Delta,\ a+\Delta]$ функция $g_x(y) = f(x,y)$ одного переменного y определена на отрезке $[b-\Delta,\ b+\Delta]$ и на этом отрезке возрастает, так как имеет положительную производную. При этом для $x = a$ функция $g_x(y)$ обращается в нуль в точке b отрезка $[b-\Delta,\ b+\Delta]$. Следовательно, $f(a, b-\Delta) = g_a(b-\Delta) < 0$ и $f(a, b+\Delta) = g_a(b+\Delta) > 0$. Отметим, что функция двух переменных $f(x,y)$ имеет непрерывные частные производные в U. Значит, она дифференцируема и непрерывна в U. Поэтому можно указать такое число δ, $0 < \delta \leqslant \Delta$, что при $|x-a| \leqslant \delta$ будем иметь $f(x,\ b-\Delta) < 0$ и $f(x,\ b+\Delta) > 0$ (рис. 3).

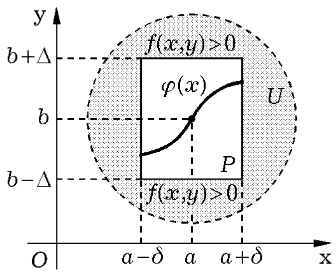


Рис. 3

Выберем произвольное значение $x \in [a-\delta,\ a+\delta]$ и зафиксируем. Функция $g_x(y) = f(x,y)$ на отрезке $[b-\Delta,\ b+\Delta]$ непрерывна, возрастает, причем на концах отрезка имеет значения разных знаков. Следовательно, на этом отрезке существует единственное значение y, при котором функция $g_x(y)$ принимает нулевое значение. Иначе говоря, при $x \in [a-\delta,\ a+\delta]$ уравнение $f(x,y) = 0$ на отрезке $[b-\Delta,\ b+\Delta]$ имеет единственное решение относительно y, т.е. на отрезке $x \in [a-\delta,\ a+\delta\,]$ определена однозначным образом функция $y = h(x)$, удовлетворяющая соотношению $f\big(x, h(x)\big) \equiv 0$. Тем самым доказана разрешимость уравнения $f(x,y) = 0$ относительно y в некотором прямоугольнике P. В качестве δ_x и δ_y следует взять соответственно δ и Δ.

Отметим, что функция $y = h(x)$ непрерывна в точке a, в которой значение функции равно b, т.е. $h(a) = b$. Действительно, приведенные выше рассуждения верны при любом сколь угодно малом значении $\Delta > 0$. Это на самом деле означает, что для любого числа $\Delta > 0$ существует такое число $\delta > 0$, что при $|x-a| \leqslant \delta$ уравнение $f(x,y) = 0$ разрешимо относительно y, причем значение y попадает в отрезок $[-\Delta,\ \Delta]$. Но уменьшение Δ приведет не к изменению функции $h(x)$ (в виду ее единственности), а лишь к сужению ее области определения. Значит, при $|x-a| \leqslant \delta$ верно неравенство $|h(x) - b| = |y-b| \leqslant \Delta$. А это и есть условие непрерывности функции $y = h(x)$ в точке a.

В качестве точки (a,b) можно взять любую точку в прямоугольнике P на графике функции $y = h(x)$ (условие теоремы выполняется в любой такой точке) и повторить все рассуждения. Мы получим, что функция $y = h(x)$ непрерывна в любой точке x интервала $(a-\delta_x, a+\delta_x)$.

Выберем в прямоугольнике P произвольную точку (x,y), для которой $y = h(x)$, или, по-другому, $f(x,y)=0$. Обозначим $\Delta x = x-a$, $\Delta y = y-b$. Тогда

$$\Delta y = y-b = h(x)-h(a) = h(a+\Delta x)-h(a).$$

Функция $f(x,y)$ имеет непрерывные частные производные в P. Поэтому к ней можно применить теорему Тейлора при $m=0$. В результате получаем

$$0 = f(x,y)-f(a,b) = f'_x(a+\theta\Delta x, b+\theta\Delta y)\Delta x + f'_y(a+\theta\Delta x, b+\theta\Delta y)\Delta y, \quad \theta\in(0,\,1),$$

откуда, учитывая, что $f'_y\neq 0$ в прямоугольнике P, находим

$$\Delta y = -\frac{f'_x(\widetilde{x},\widetilde{y})}{f'_y(\widetilde{x},\widetilde{y})}\,\Delta x, \tag{3}$$

где $\widetilde{x}=a+\theta\Delta x$, $\widetilde{y}=b+\theta\Delta y$.

Если x устремить к значению a ($x\to a$), то в силу непрерывности функции $y=h(x)$ будем иметь $y\to b$, откуда заключаем, что точка (x,y) на плоскости будет стремиться к точке (a,b). Но тогда $\Delta x\to 0$, $\Delta y\to 0$ и $(\widetilde{x},\widetilde{y})\to(a,b)$. В силу непрерывности частных производных функции $f(x,y)$ отношение $f'_x(\widetilde{x},\widetilde{y})/f'_y(\widetilde{x},\widetilde{y})$ можно представить в виде

$$-\frac{f'_x(\widetilde{x},\widetilde{y})}{f'_y(\widetilde{x},\widetilde{y})} = -\frac{f'_x(a,b)}{f'_y(a,b)} + \alpha(\Delta x),$$

где $\alpha(\Delta x)$ — бесконечно малая функция при $\Delta x\to 0$. Учтем это представление в равенстве (3):

$$h(a+\Delta x)-h(a) = \Delta y = -\frac{f'_x(a,b)}{f'_y(a,b)}\,\Delta x + \alpha(\Delta x)\,\Delta x.$$

Полученное соотношение по определению означает, что функция одного переменного $y=h(x)$ дифференцируема в точке a. При этом

$$h'(a) = -\frac{f'_x(a,b)}{f'_y(a,b)}.$$

Ясно, что вместо точки (a,b) можно взять любую точку в прямоугольнике P, лежащую на графике функции $y=h(x)$, и повторить рассуждения. Мы приходим к выводу, что функция $y=h(x)$ дифференцируема в интервале $(a-\delta_x,\ a+\delta_x)$, а ее производная может быть вычислена по формуле (2). Наконец, отметим, что правая часть формулы (2) представляет собой композицию непрерывных функций: отношение частных производных есть непрерывная функция двух переменных как отношение непрерывных функций (при этом $f'_y(x,y)\neq 0$ в прямоугольнике P), а переменное y есть непрерывная функция $y=h(x)$ переменного x. Следовательно, функция $h'(x)$ непрерывна в интервале $(a-\delta_x,\ a+\delta_x)$. $\triangleright$

Отметим, что прямоугольник P, построенный в доказательстве теоремы 1, выбран так, что график функции $y=h(x)$ расположен внутри этого прямоугольника и соединяет две точки на противоположных вертикальных сторонах прямоугольника.

Пример 4. Уравнение $f(x,y) = x^2 + y^2 - 9 = 0$ из примера 1 задает окружность радиуса 3. Условия теоремы 1 выполнены во всех точках окружности, кроме точек $(3,0)$ и $(-3,0)$, в которых $f'_y = 0$. Для любой точки (a,b) окружности в верхней полуплоскости существует прямоугольник P, в котором уравнение разрешимо относительно y: $y = \sqrt{9-x^2}$ (рис. 4).

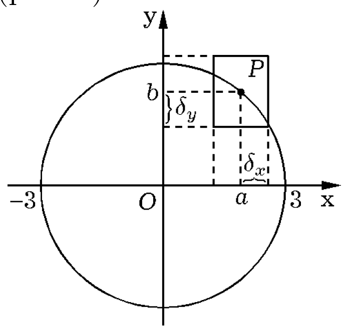


Рис. 4

В качестве такого прямоугольника подходит любой прямоугольник в верхней полуплоскости со сторонами, параллельными осям координат, центром в точке $(a,\ b)$, такой, что окружность пересекается с его боковыми сторонами. Для точек окружности в нижней полуплоскости в соответствующем прямоугольнике $y = -\sqrt{9-x^2}$. В окрестности точек $(3,0)$ и $(-3,0)$ уравнение неразрешимо относительно y. Например, возьмем точку $(3,0)$. Тогда при любом $x > 3$ уравнение не имеет решений, а при $x < 3$ оно имеет два решения, т.е. в окрестности этой точки уравнение не определяет y как функцию x. Отметим, что в этих особых точках касательные к окружности вертикальны.

Пример 5. Функция $f(x,y) = e^y + y - x + \ln x$ из примера 3 удовлетворяет условию $f'_y(x,y) = e^y + 1 > 0$ всюду в своей области определения: $x > 0$. Значит, если это уравнение имеет хотя бы одно решение (a,b), то в окрестности этого решения уравнение разрешимо относительно переменного y и задает y как функцию переменного x.

10.3 Общий случай

Теорему 1 несложно обобщить на случай одного уравнения с $n+1$ неизвестными.

Теорема 2 (**о неявной функции**). Пусть в окрестности V точки (a,b), $a \in \mathbb{R}^n$, $b \in \mathbb{R}$, задана скалярная функция $f(x,y)$ от $n+1$ переменного ($x \in \mathbb{R}^n$, $y \in \mathbb{R}$), удовлетворяющая условиям:

а) $f(a,b) = 0$;

б) $f \in C^1(V)$;

в) $f'_y(a,b) \neq 0$.

Тогда точка (a,b) имеет окрестность вида

$$P = \{(x,y) \in \mathbb{R}^{n+1}\colon\ x \in U_{\delta_x}(a),\ |y-b| < \delta_y\},$$

в которой уравнение $f(x,y) = 0$ разрешимо относительно y, т.е.

$$\exists h \in C^1\big(U_{\delta_x}(a)\big): \qquad \forall (x,y) \in P \quad \Big(f(x,y) = 0 \ \Leftrightarrow\ y = h(x)\Big).$$

При этом

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

Док–во повторяет док–во теоремы 1, но вместо частной производной по переменному x используется матрица Якоби функции f по части переменных x. $\triangleright$

Опираясь на более общий вариант теоремы 1, можно доказать теорему о неявной функции для системы уравнений, используя метод математической индукции.

Пусть $f\colon \mathbb{R}^{m+n} \to \mathbb{R}^m$ — векторная функция $m+n$ переменных. Запишем эту функцию в виде $f(x, y)$, где $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ обозначает группу из n свободных переменных, а $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ — группу из m зависимых переменных. Тогда для функции $f(x,y) = \big(f_1(x,y), \ldots, f_m(x,y)\big)^{\mathrm{T}}$ векторное уравнение $f(x,y) = 0$ равносильно системе (1). Поставим вопрос: при каких условиях эта система разрешима относительно переменных y_1, $\ldots$, y_m в окрестности данной точки $(a,b) \in \mathbb{R}^{n+m}$? В формулировке и доказательстве через $f'_x(x,y) = \dfrac{\partial f(x,y)}{\partial x}$ и $f'_y(x,y) = \dfrac{\partial f(x,y)}{\partial y}$ будем обозначать соответственно матрицы Якоби функции f по части переменных x и по части переменных y, т.е.

$$f'_x(x,y) = \begin{pmatrix} \frac{\partial f_1(x,y)}{\partial x_1} & \frac{\partial f_1(x,y)}{\partial x_2} & \ldots & \frac{\partial f_1(x,y)}{\partial x_n} \\ \frac{\partial f_2(x,y)}{\partial x_1} & \frac{\partial f_2(x,y)}{\partial x_2} & \ldots & \frac{\partial f_2(x,y)}{\partial x_n} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial f_m(x,y)}{\partial x_1} & \frac{\partial f_m(x,y)}{\partial x_2} & \ldots & \frac{\partial f_m(x,y)}{\partial x_n} \end{pmatrix}$$

и

$$f'_y(x,y) = \begin{pmatrix} \frac{\partial f_1(x,y)}{\partial y_1} & \frac{\partial f_1(x,y)}{\partial y_2} & \ldots & \frac{\partial f_1(x,y)}{\partial y_m} \\ \frac{\partial f_2(x,y)}{\partial y_1} & \frac{\partial f_2(x,y)}{\partial y_2} & \ldots & \frac{\partial f_2(x,y)}{\partial y_m} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial f_m(x,y)}{\partial y_1} & \frac{\partial f_m(x,y)}{\partial y_2} & \ldots & \frac{\partial f_m(x,y)}{\partial y_m} \end{pmatrix}.$$

Отметим, что матрица $f'_y(x,y)$ является квадратной порядка m, а матрица Якоби $f'(x,y)$ по всей совокупности переменных может быть записана как блочная матрица $\big(f'_x(x,y), f'_y(x,y)\big)$. Отметим также, что определитель квадратной матрицы Якоби (по части переменных или по всем переменным — неважно) называют **якобианом**.

Теорема 3 (**о неявной функции**). Пусть в окрестности $V \subseteq \mathbb{R}^{n+m}$ точки (a,b), $a \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, задана функция $f\colon \mathbb{R}^{n+m} \to \mathbb{R}^m$, удовлетворяющая условиям:

а) $f(a,b) = 0$;

б) $f \in C^1(V, \mathbb{R}^m)$;

в) $\det f'_y(a,b) \neq 0$.

Тогда точка (a,b) имеет окрестность вида

$$P = \{(x,y) \in \mathbb{R}^{n+m}\colon\ x \in U_{\delta_x}(a),\ y \in U_{\delta_y}(b)\},$$

в которой уравнение $f(x,y)=0$ разрешимо относительно y, т.е.

$$\exists h \in C^1\big(U_{\delta_x}(a), \mathbb{R}^m\big): \qquad \forall (x,y) \in P \quad \Big(f(x,y)=0 \;\Leftrightarrow\; y=h(x)\Big).$$

При этом

$$h'(x) = -\Big(f'_y\big(x,h(x)\big)\Big)^{-1} f'_x\big(x,h(x)\big). \tag{4}$$

Доказательство теоремы проведем индукцией по количеству уравнений системы, т.е. по параметру m. При $m=1$ мы имеем одно уравнение с $n+1$ неизвестными и утверждение теоремы сводится к утверждению теоремы 2.

Предположим, что утверждение теоремы верно для случая $m-1$ уравнений, и рассмотрим систему $f(x,y)=0$ с m уравнениями, которая удовлетворяет условиям теоремы. В точке (a,b) матрица Якоби $f'_y(x,y)$ является невырожденной. Следовательно, последняя строка этой матрицы, состоящая из частных производных $\dfrac{\partial f_m}{\partial y_j}(a,b)$, $j=1,\ldots,m$, имеет хотя бы один ненулевой элемент. Перенумеруем переменные $y_1,\ldots,y_m$ так, чтобы это был последний элемент строки, т.е. $\dfrac{\partial f_m}{\partial y_m}(a,b) \neq 0$. Тогда уравнение $f_m(x,\tilde{y},y_m)=0$, где $\tilde{y}=(y_1,\ldots,y_{m-1})$, в некоторой окрестности $U_*(a,b)$ точки (a,b) вида

$$U_*(a,b) = \big\{(x,\tilde{y},y_m) \in \mathbb{R}^{n+m}\colon |x-a|<\delta_1,\ |\tilde{y}-\tilde{b}|<\delta_2,\ |y_m-b_m|<\delta_3\big\},$$

где $b=(\tilde{b},b_m)=(b_1,\ldots,b_{m-1},b_m)$, согласно теореме 2, разрешимо относительно переменного y_m, т.е. в окрестности точки $(a,\tilde{b})$ существует такая непрерывно дифференцируемая функция $\psi(x,\tilde{y})$, что

$$f(x,y)=0 \;\Leftrightarrow\; \left\{ \begin{array}{c} \tilde{f}(x,\tilde{y},y_m)=0 \\ y_m=\psi(x,\tilde{y}) \end{array} \right. \;\Leftrightarrow\; \left\{ \begin{array}{c} \tilde{f}\big(x,\tilde{y},\psi(x,\tilde{y})\big)=0 \\ y_m=\psi(x,\tilde{y}) \end{array} \right. .$$

Для системы уравнений $g(x,\tilde{y})=\tilde{f}\big(x,\tilde{y},\psi(x,\tilde{y})\big)=0$ выполняются условия теоремы. Действительно, используя правило дифференцирования сложной функции, из определения g и тождества $f_m\big(x,\tilde{y},\psi(x,\tilde{y})\big)=0$ выводим равенства

$$\frac{\partial g}{\partial \tilde{y}} = \frac{\partial \tilde{f}}{\partial \tilde{y}} + \frac{\partial \tilde{f}}{\partial y_m}\frac{\partial \psi}{\partial \tilde{y}}, \qquad 0 = \frac{\partial f_m}{\partial \tilde{y}} + \frac{\partial f_m}{\partial y_m}\frac{\partial \psi}{\partial \tilde{y}},$$

которые можно записать в виде матричного равенства

$$\begin{pmatrix} g'_{\tilde{y}} \\ 0 \end{pmatrix} = f'_y \begin{pmatrix} E \\ \psi'_{\tilde{y}} \end{pmatrix}.$$

Так как единичная матрица E невырождена, а матрица $f'_y(a,b)$ невырождена по условию теоремы, то ранг матрицы, стоящей справа, максимален и равен количеству строк. Поэтому ранг матрицы, стоящей слева, также равен количеству строк, т.е. $\det g'_{\tilde{y}}(a,\tilde{b}) \neq 0$.

Согласно предположению математической индукции, система $g(x,\tilde{y})=0$ разрешима относительно переменных $\tilde{y}$ в окрестности точки $(a,\tilde{b})$ и может быть представлена в виде $\tilde{y}=\tilde{\psi}(x)$. Но тогда система

$$\tilde{y}=\tilde{\psi}(x), \qquad y_m=\psi\big(x,\tilde{\psi}(x)\big)$$

эквивалентна системе $f(x,y) = 0$. Функция $h(x) = \Big(\tilde\psi(x), \psi\big(x, \tilde\psi(x)\big)\Big)^{\mathrm{т}}$ непрерывно дифференцируема в окрестности точки $a \in \mathbb{R}^n$ как композиция непрерывно дифференцируемых функций ψ и $\tilde\psi$. Поэтому сложная функция $G(x) = f\big(x, h(x)\big)$ дифференцируема в некоторой окрестности точки a. Согласно правилу дифференцирования сложной функции,

$$G'(x) = \big(f'_x(x,y) + f'_y(x,y)h'(x)\big)\Big|_{y=h(x)}.$$

Но в то же время в силу выбора функции $h(x)$ имеем $G(x) \equiv 0$ в некоторой окрестности точки a. Значит,

$$f'_x\big(x, h(x)\big) + f'_y\big(x, h(x)\big)\, h'(x) = 0$$

в этой окрестности точки a. Отметим, что матрица Якоби $f'_y(a,b)$ невырожденная. Поэтому матрица $f'_y(x,y)$ является невырожденной в некоторой окрестности

$$U(a,b) = \{(x,y) : \ |x-a| < \varepsilon_1, \ |y-b| < \varepsilon_2\}$$

точки (a,b). Число ε_1 можно взять настолько малым, что при $|x-a| < \varepsilon_1$ будет выполняться соотношение $|h(x) - b| < \varepsilon_2$. В этом случае матрица $f'_y\big(x, h(x)\big)$ при $|x-a| < \varepsilon_1$ является невырожденной, и мы приходим к формуле (4). $\triangleright$

Теоремы 1–3 содержат три условия, которые являются достаточными для локального существования неявной функции и ее дифференцируемости. В случае нарушения хотя бы одного из этих условий применение указанных теорем невозможно и следует искать другие подходы к выявлению разрешимости системы нелинейных уравнений. Покажем на примерах, что при нарушении условий теоремы о неявной функции ее утверждение в одних случаях верно, а в других нет.

Пример 6. Уравнение $y^2 - x^3 = 0$ (рис. 5) в окрестности точки $(0,0)$ не разрешимо относительно переменного y, так как любому значению $x > 0$ соответствуют два противоположных по знаку значения y, в то время как при $x < 0$ уравнение вообще не имеет решений. Теорема 1 не применима, так как нарушено третье условие теоремы. Отметим, что в других точках в $\mathbb{R}^2$, удовлетворяющих уравнению $y^{2/3} - x = 0$ (или $y^2 - x^3 = 0$), условия теоремы 1 выполнены, а уравнение в области $x > 0$ задает неявную функцию $y = x^{3/2}$ для точек выше оси абсцисс и $y = -x^{3/2}$ для точек ниже оси абсцисс.

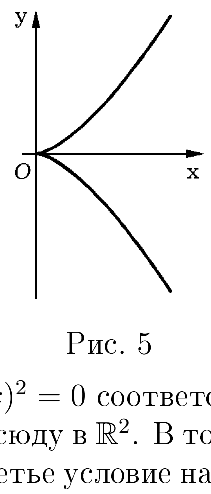


Рис. 5

Пример 7. Уравнению $(y-x)^2 = 0$ соответствует функция $f(x,y) = (y-x)^2$, непрерывно дифференцируемая всюду в $\mathbb{R}^2$. В точке $(0,0)$ выполнены первое и второе условия теоремы 1, однако третье условие нарушено, так как $f'_y(x,y) = 2(y-x)$

и $f'_y(0,0) = 0$. Тем не менее уравнение разрешимо относительно переменного y и задает функцию $y = x$, определенную и непрерывно дифференцируемую всюду в $\mathbb{R}$. В данном случае утверждение теоремы (в части существования неявной функции) верно, хотя применение этой теоремы невозможно из-за нарушения ее условий.

11 Геометрические приложения

См. [ККЧ, §5.1–5.3].

11.1 Производная по направлению

Пусть скалярная функция многих переменных f определена в некоторой окрестности точки $a \in \mathbb{R}^n$ и задан вектор $\vec{\nu} \neq \vec{0}$. Обозначим через $\vec{\nu}_0$ единичный вектор, имеющий то же направление, что и вектор $\vec{\nu}$: $\vec{\nu}_0 = \dfrac{\vec{\nu}}{|\vec{\nu}|}$. **Производной** функции $f\colon A \to \mathbb{R}$ в точке $a \in A \subseteq \mathbb{R}^n$ **по направлению вектора** $\vec{\nu}$ называют число

$$\frac{\partial f(a)}{\partial \vec{\nu}} = \lim_{s \to +0} \frac{f(a + s\vec{\nu}_0) - f(a)}{s},$$

если этот предел существует.

Непосредственно из определения следует, что

$$\exists\, \frac{\partial f(a)}{\partial x_i} \quad \Rightarrow \quad \frac{\partial f(a)}{\partial x_i} = \frac{\partial f(a)}{\partial \vec{\nu}}, \quad \text{где} \quad \vec{\nu} = e_i = (0, \ldots, 1, \ldots, 0)\,.$$

Таким образом, производная по направлению базисного вектора совпадает с соответствующей частной производной. Однако производная по направлению определяется односторонним пределом, а частная производная — двусторонним. Поэтому возможна ситуация, когда производная по базисному направлению существует, а соответствующая частная производная — нет. Производная по направлению вектора обобщает понятие частной производной первого порядка, распространяя это понятие на случай произвольного направления в заданной точке.

Из определения следует **механическая интерпретация производной по направлению вектора**: $\dfrac{\partial f(a)}{\partial \vec{\nu}}$ представляет собой скорость изменения значений функции f в окрестности точки a в направлении вектора $\vec{\nu}$, а также **геометри-**

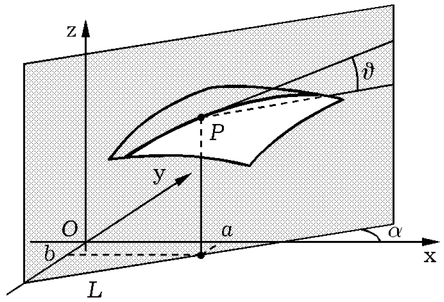


Рис. 6

ческая интерпретация: производная функции двух переменных $f(x,y)$ в точке

(a,b) по направлению вектора $\vec{\nu}$ равна тангенсу угла наклона односторонней касательной в точке (a,b) к сечению графика функции плоскостью, параллельной вектору $\vec{\nu}$ и оси Oz (рис. 6).

Теорема 1. Если скалярная функция f определена в окрестности точки $a \in \mathbb{R}^n$ и дифференцируема в точке a, то в этой точке она имеет производную по направлению любого ненулевого вектора $\vec{\nu}$, причем

$$\frac{\partial f(a)}{\partial \vec{\nu}} = \sum_{i=1}^{n} \frac{\partial f(a)}{\partial x_i} \nu_i, \qquad \text{где} \quad (\nu_1, \ldots, \nu_n) = \frac{\vec{\nu}}{|\vec{\nu}|} = \vec{\nu}_0.$$

Док–во. Рассмотрим функцию $g(s) = f(a + s\vec{\nu}_0)$ одного действительного переменного s. Поскольку функция многих переменных $f(x)$ дифференцируема в точке $a = (a_1, \ldots, a_n)$, то сложная функция $g(s) = f\bigl(x(s)\bigr)$, где $x(s) = a + s\vec{\nu}_0$, дифференцируема в точке $s = 0$ и

$$\frac{dg(0)}{ds} = \left.\frac{df(a_1 + s\nu_1, \ldots, a_n + s\nu_n)}{ds}\right|_{s=0} = \sum_{i=1}^{n} \frac{\partial f(a)}{\partial x_i}\, \nu_i.$$

В то же время, согласно определению производной функции, имеем

$$\frac{dg(0)}{ds} = \lim_{s \to 0} \frac{g(s) - g(0)}{s} = \lim_{s \to 0} \frac{f(a + s\vec{\nu}_0) - f(a)}{s}.$$

Из существования последнего предела вытекает и существование равного ему одностороннего предела при $s \to +0$. Поэтому

$$\frac{dg(0)}{ds} = \lim_{s \to +0} \frac{f(a + s\vec{\nu}_0) - f(a)}{s} = \frac{\partial f(a)}{\partial \vec{\nu}}.$$

Приравнивая правые части полученных равенств, получаем утверждение теоремы. $\triangleright$

Задача 1. Приведите пример функции f, которая недифференцируема в точке a, но $\forall \vec{\nu} \neq \vec{0} \quad \exists \dfrac{\partial f(a)}{\partial \vec{\nu}}$.

11.2 Градиент

Пусть скалярная функция многих переменных f в точке $a \in \mathbb{R}^n$ имеет все частные производные первого порядка. Вектор

$$\operatorname{grad} f(a) = \left(f'_{x_1}(a), \ldots, f'_{x_n}(a)\right),$$

составленный из частных производных первого порядка функции f в точке a, называют **градиентом функции** f в точке a.

Теорема 2 (**свойства градиента**). Если скалярная функция f определена в окрестности точки $a \in \mathbb{R}^n$ и дифференцируема в точке a, то

а) $\dfrac{\partial f(a)}{\partial \vec{\nu}} = (\operatorname{grad} f(a), \vec{\nu}_0)$;

б) $\dfrac{\partial f(a)}{\partial \vec{\nu}} = \text{пр}_{\vec{\nu}} \operatorname{grad} f(a)$ — проекция вект. $\operatorname{grad} f(a)$ на направ. вектора $\vec{\nu}$;

в) вектор $\operatorname{grad} f(a)$ — направление наибольшего роста ф. f в т. a;

г) $\dfrac{\partial f(a)}{\partial \big(\operatorname{grad} f(a)\big)} = |\operatorname{grad} f(a)|$ — наибольшая скорость роста функции в т. a.

Док–во. Утв. а) следует из теоремы 1, б) — из определения ортогональной проекции вектора $\vec{y}$ на направление вектора $\vec{x}$: $(\vec{x}, \vec{y}) = |\vec{x}| \, \text{пр}_{\vec{x}} \vec{y}$.

в, г): в силу *неравенства Коши—Буняковского* для любого вектора $\vec{\nu}$ имеем

$$\frac{\partial f(a)}{\partial \vec{\nu}} = (\operatorname{grad} f(a), \vec{\nu}_0) \leqslant |\operatorname{grad} f(a)| \, |\vec{\nu}_0| = |\operatorname{grad} f(a)|.$$

где равенство только, когда векторы $\vec{\nu}_0$ и $\operatorname{grad} f(a)$ однонаправлены. $\triangleright$

11.3 Касательная плоскость и нормаль

Рассмотрим некоторую поверхность S в пространстве и точку $M \in S$. Пусть существует такая плоскость π, проходящая через точку M, которая содержит все касательные в точке M ко всем возможным гладким кривым, лежащим на поверхности S и проходящим через точку M. Тогда плоскость π называют **касательной плоскостью** к поверхности S в точке M (рис. 7).

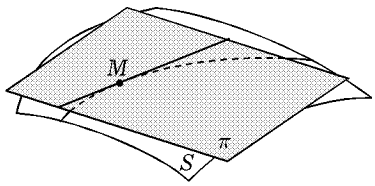


Рис. 7

Прямую L, проходящую через точку M и перпендикулярную плоскости π, называют **нормалью к поверхности** S в точке M.

Теорема 3 (**достаточное условие существования касательной плоскости**). Пусть в прямоугольной системе координат $Oxyz$ поверхность S задана уравнением $F(x, y, z) = 0$, функция $F(x, y, z)$ дифференцируема в точке M, (a, b, c) — координаты точки M, градиент функции $F(x, y, z)$ в точке M отличен от нуля: $\operatorname{grad} F(a, b, c) \neq \vec{0}$. Тогда касательная плоскость к поверхности S в точке M существует и имеет уравнение

$$F'_x(a, b, c)(x - a) + F'_y(a, b, c)(y - b) + F'_z(a, b, c)(z - c) = 0. \tag{1}$$

Нормаль к поверхности S в точке M также существует и имеет уравнение

$$\frac{x - a}{F'_x(a, b, c)} = \frac{y - b}{F'_y(a, b, c)} = \frac{z - c}{F'_z(a, b, c)}. \tag{2}$$

Док–во. Рассмотрим кривую γ, лежащую на поверхности S, проходящую через точку M и имеющую касательную в точке M. Тогда эту кривую можно задать такими параметрическими уравнениями

$$x = \varphi(t), \quad y = \psi(t), \quad z = \chi(t),$$

что значение параметра t_0 соответствует точке M:

$$\varphi(t_0) = a, \quad \psi(t_0) = b, \quad \chi(t_0) = c,$$

а вектор $\vec{\tau} = (\varphi'(t_0), \psi'(t_0), \chi'(t_0))$ является направляющим вектором касательной к γ в точке M. Так как γ принадлежит поверхности S, то

$$F\big(\varphi(t), \psi(t), \chi(t)\big) \equiv 0, \tag{3}$$

причем сложная функция в левой части тождества дифференцируема в точке t_0. Поэтому, дифференцируя (3) в точке t_0 по правилу дифференцирования сложной функции, получаем

$$\frac{\partial F(a,b,c)}{\partial x}\varphi'(t_0) + \frac{\partial F(a,b,c)}{\partial y}\psi'(t_0) + \frac{\partial F(a,b,c)}{\partial z}\chi'(t_0) = 0.$$

Записанное равенство означает, что вектор $\vec{\tau}$ ортогонален вектору $\operatorname{grad} F(a,b,c)$, не зависящему от выбора кривой γ.

Итак, все касательные в точке $M \in S$ к всевозможным кривым, лежащим на поверхности S и проходящим через точку M, ортогональны градиенту $\operatorname{grad} F(a,b,c)$ функции $F(x,y,z)$. Построим плоскость π, проходящую через точку M и имеющую нормальный вектор $\operatorname{grad} F(a,b,c)$. Так как касательные к любой кривой, лежащей на поверхности S, в точке M принадлежат плоскости π, плоскость π является касательной к поверхности S в точке M.

Зная координаты a, b, c точки M, через которую проходит плоскость π, и координаты нормального вектора $\operatorname{grad} F(a,b,c)$ этой плоскости, можем записать общее уравнение (1) плоскости π.

Нормаль в точке M поверхности S определяется той же точкой M и тем же вектором $\operatorname{grad} F(a,b,c)$, который является направляющим вектором этой прямой. Ее каноническое уравнение есть (2). $\triangleright$

Из доказанной теоремы вытекает еще одно свойство градиента функции.

Следствие (свойство градиента функции). В точке M дифференцируемости функции $F(x,y,z)$ ее градиент $\operatorname{grad} F(M)$ ортогонален поверхности уровня $F(x,y,z) = C$, где $C = F(M)$, при этом ортогональность к поверхности понимается как ортогональность к ее касательной плоскости.

Док–во. Так как градиенты функций $F(x,y,z) - C$ и $F(x,y,z)$ совпадают, то вектор $\operatorname{grad} F(M)$ является нормальным вектором касательной плоскости к поверхности $F(x,y,z) - C = 0$ в точке M. $\triangleright$

Пусть поверхность S задана уравнением $z = f(x,y)$, где функция $f(x,y)$ дифференцируема в окрестности точки $M(a,b)$. Найдем уравнения касательной плоскости и нормали в точке $P(a,b,c)$, где $c = f(a,b)$. Положим $F(x,y,z) = f(x,y) - z$.

Тогда условия теоремы 3 будут выполнены, и, следовательно, в точке P существуют касательная плоскость и нормаль к поверхности S. Уравнения касательной плоскости найдем по формуле (1). Так как $F'_x(a,b,c) = f'_x(a,b)$, $F'_y(a,b,c) = f'_y(a,b)$, $F'_z(a,b,c) = -1$, то уравнение касательной плоскости имеет вид

$$f'_x(a,b)(x-a) + f'_y(a,b)(y-b) - (z-c) = 0. \tag{4}$$

Аналогично по формуле (2) находим канонические уравнения нормали

$$\frac{x-a}{f'_x(a,b)} = \frac{y-b}{f'_y(a,b)} = \frac{z-c}{-1}.$$

Понятие касательной плоскости позволяет дать **геометрическую интерпретацию дифференциалу функции многих переменных**. Пусть функция $z = f(x,y)$ двух переменных дифференцируема в точке $M\,(a,b)$. Тогда ее дифференциал dz в этой точке равен

$$dz = f'_x(a,b)\,dx + f'_y(a,b)\,dy. \tag{5}$$

В то же время уравнение $z = f(x,y)$, рассматриваемое в прямоугольной системе координат $Oxyz$, задает поверхность в пространстве, и эта поверхность в точке $\big(a,b,f(a,b)\big)$ имеет касательную плоскость, уравнение (4) которой можно записать в виде

$$z - c = f'_x(a,b)(x-a) + f'_y(a,b)(y-b).$$

Обозначив $x-a = \Delta x$, $y-b = \Delta y$, $z-c = \Delta z$, перепишем это уравнение в виде

$$\Delta z = f'_x(a,b)\,\Delta x + f'_y(a,b)\,\Delta y. \tag{6}$$

Сравнивая (5) и (6), заключаем, что дифференциал dz совпадает с Δz, так как приращения Δx и Δy независимых переменных x и y в точке $M(a,b)$ в то же самое время являются дифференциалами этих переменных. Таким образом, получаем такую **геометрическую интерпретацию**: дифференциал функции двух переменных есть приращение в точке M аппликаты точки на касательной плоскости, соответствующее приращениям dx, dy независимых переменных (рис. 8).

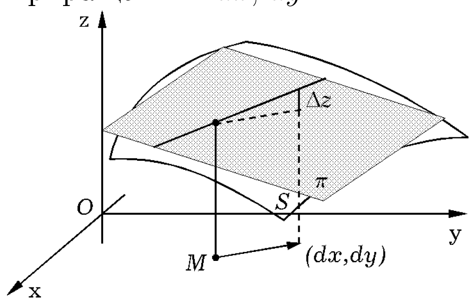


Рис. 8

12 Экстремум функции многих переменных

См. [ККЧ, §6.1–6.4].

12.1 Необходимое условие экстремума

Пусть скалярная функция $f\colon \mathbb{R}^n \to \mathbb{R}$ определена в некоторой окрестности точки $a \in \mathbb{R}^n$. Говорят, что функция f имеет в точке a **локальный максимум** (**минимум**), если существует такая проколотая окрестность $\dot{U}_\varepsilon(a)$ точки a, что для любой точки $x \in \dot{U}_\varepsilon(a)$ выполнено неравенство $f(x) \leqslant f(a)$, $(f(x) \geqslant f(a))$. Понятия локального минимума и локального максимума функции объединяют под общим названием **экстремум функции**.

Если неравенства в этом определении являются строгими, то говорят о **строгом экстремуме функции**.

Теорема 1 (**необходимое условие экстремума функции**). Пусть скалярная функция $f\colon \mathbb{R}^n \to \mathbb{R}$ имеет в точке $a \in \mathbb{R}^n$ экстремум и частную производную $f'_{x_i}(a) = 0, (1 \leqslant i \leqslant n)$. Тогда эта частная производная равна нулю: $f'_{x_i}(a) = 0$.

Док–во. Пусть $a = (a_1, \ldots, a_n)$. Рассмотрим действительную функцию

$$g(t) = f(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_n)$$

одного переменного t, которая в точке $t = a_i$ имеет локальный экстремум. В самом деле, пусть, например, $f(x)$ имеет в точке a локальный максимум. Тогда существует такая проколотая окрестность $\dot{U}_\varepsilon(a)$, что $f(x) \leqslant f(a)$ при $x \in \dot{U}_\varepsilon(a)$. Но в таком случае $g(t) \leqslant g(a_i)$ при $0 < |t - a_i| < \varepsilon$, что соответствует определению локального максимума функции одного переменного.

Функция $g(t)$ дифференцируема в точке $t = a_i$, так как $g'(a_i) = f'_{x_i}(a)$. Согласно необходимому условию локального экстремума для функции одного переменного, выполнено равенство $g'(a_i) = 0$. Следовательно, $f'_{x_i}(a) = 0$. $\triangleright$

Следствие. Пусть скалярная функция $f\colon \mathbb{R}^n \to \mathbb{R}$ имеет в точке $a \in \mathbb{R}^n$ экстремум. Тогда:
(1) если в точке a определен градиент функции $f(x)$, то $\operatorname{grad} f(a) = 0$;
(2) если функция дифференцируема в точке a, то $df(a) = 0$.

Таким образом, точки экстремума скалярной функции $f(x)$ надо искать либо среди точек, в которых $\operatorname{grad} f(x) = 0$, либо среди точек, в которых градиент не определен (не существует одна или несколько частных производных). Точки, в которых градиент функции равен нулю или не определен, называют **критическими точками функции**. **Стационарными точками функции** называют точки, в которых $\operatorname{grad} f(x) = 0$. Укажем эквивалентные условия:

$$\operatorname{grad} f(x) = 0 \quad \Longleftrightarrow \quad df(x) = 0 \quad \Longleftrightarrow \quad f'_{x_i}(x) = 0 \quad i = 1, \ldots, n.$$

Пример 1. Покажем, что у функции $g(x, y) = x^2 - y^2$ нет экстремумов (в этом, кстати, можно убедиться, изобразив в прямоугольной системе координат в пространстве график этой функции). Функция $g(x, y)$ дифференцируема на всей плоскости. Поэтому ее точки экстремума могут быть лишь среди стационарных точек. Вычислим частные производные функции и запишем систему уравнений, приравняв частные производные нулю:

$$g'_x = 2x = 0, \quad g'_y = -2y = 0.$$

Эта система имеет единственное решение $x = 0$, $y = 0$. Значит, функция $g(x, y)$ может иметь экстремум лишь в точке $(0, 0)$. Однако при $y = 0$ функция одного переменного $g(x, 0) = x^2$ в точке $x = 0$ имеет строгий локальный минимум, так как $g(x, 0) = x^2 > 0 = g(0, 0)$, $x \neq 0$, а при $x = 0$ функция одного переменного $g(0, y)$ при $y = 0$ имеет строгий локальный максимум, так как $g(0, y) = -y^2 < 0 = g(0, 0)$, $y \neq 0$. Поэтому точка $(0, 0)$ не может быть точкой экстремума функции $g(x, y)$. $\triangleright$

Пример 2. Функция двух переменных $h(x, y) = |x| + y^2$ дифференцируема во всех точках плоскости xOy, кроме точек оси Oy. При этом $h'_x(x, y) = 1$ при $x > 0$ и $h'_x(x, y) = -1$ при $x < 0$. Значит, точки экстремума могут располагаться только на оси Oy, в точках которой не существует частная производная h'_x. Обратим внимание, что частная производная функции $h(x, y)$ по переменному y существует во всех критических точках, но обращается в нуль только при $y = 0$, т.е. в начале координат. Поэтому единственная точка, в которой может быть экстремум функции, — это точка $(0, 0)$. Так как слагаемое $|x|$ имеет строгий локальный минимум при $x = 0$, а слагаемое y^2 — при $y = 0$, то в точке $(0, 0)$ функция $h(x, y)$ имеет строгий локальный минимум. $\triangleright$

12.2 Достаточные условия экстремума

Исследование стационарных точек функции многих переменных на экстремум, как и в случае функций одного переменного, можно проводить, анализируя дифференциал второго порядка. Напомним, что дифференциал второго порядка функции многих переменных представляет собой квадратичную форму относительно приращений (дифференциалов) независимых переменных.

Теорема 2 (**достаточные условия экстремума**). Пусть U — окрестность точки $a \in \mathbb{R}^n$, $f \in C^2(U)$ и $df(a) = 0$. Тогда:

1) КФ $d^2 f(a)$ положительно опр. $\Rightarrow$ a — т. строгого лок. минимума ф. f;
2) КФ $d^2 f(a)$ отрицательно опр. $\Rightarrow$ a — т. строгого лок. максимума ф. f;
3) КФ $d^2 f(a)$ знакопеременная $\Rightarrow$ в точке a функция f не имеет экстремума.

Док–во. Через $\Delta x = (\Delta x_1, \ldots, \Delta x_n)^{\mathrm{T}}$ обозначим ненулевой вектор приращений независимых переменных, а через $\nu = \Delta x / |\Delta x|$ — единичный вектор с тем же направлением, что и вектор Δx. Второй дифференциал $d^2 f(a)$ функции $f(x)$ в точке a является квадратичной формой $k(\Delta x)$ от вектора приращений Δx. Имеем $k(\Delta x) = k(|\Delta x|\nu) = k(\nu)|\Delta x|^2$. Отсюда и из формулы Тейлора с остаточным членом в форме Пеано для функции $f(x)$ в точке a получаем

$$f(a+\Delta x)-f(a) = df(a)+\frac{1}{2}d^2 f(a)+o(|\Delta x|^2) = \frac{1}{2}k(\Delta x)+o(|\Delta x|^2) = \frac{1}{2}\Big(k(\nu)+\alpha\Big)|\Delta x|^2, \tag{1}$$

где $\alpha \to 0$ при $\Delta x \to 0$.

Поскольку $k(y)$ — квадратичная форма, то она непрерывна при всех y, в том числе и на множестве $S^{n-1} = \{y \in \mathbb{R}^n \colon\ |y| = 1\}$. Множество $S^{n-1} \subset \mathbb{R}^n$ ограничено и замкнуто (это $(n-1)$-мерная сфера), т.е. является компактом. Поэтому функция $k(y)$, будучи непрерывной на компакте S^{n-1} достигает на этом множестве своего наименьшего m_* и наибольшего m^* значений. Обозначим через y_* и y^* точки минимума и масимума функции $k(y)$ на S^{n-1}. Возможны следующие четыре случая

1) $0 < m_* \leqslant m^*$;
2) $m_* \leqslant m^* < 0$;
3) $m_* < 0 < m^*$;
4) $m_* = 0$ или $m^* = 0$.

Покажем, что теорема верна в каждом из этих случаев.

Случай 1. Если $m_* > 0$, то $k(y) \geqslant m_* > 0$ при $|y| = 1$. Значит, $k(\Delta x) = k(\nu)|\Delta x|^2 > 0$ при $\Delta x \neq 0$, т.е. квадратичная форма $d^2 f(a)$ положительно определенная. В силу того, что $\alpha \to 0$ при $\Delta x \to 0$, для $\varepsilon = m_*/2$, можно выбрать такое число $\delta > 0$, что $|\alpha| < \varepsilon = m_*/2$ при $0 < |\Delta x| < \delta$. Но тогда $\alpha > -m_*/2$ и в силу представления (1) находим, что

$$f(a + \Delta x) - f(a) > \frac{1}{2}\Big(m_* - \frac{1}{2}m_*\Big)|\Delta x|^2 = \frac{1}{4}m_*|\Delta x|^2 > 0,$$

если $0 < |\Delta x| < \delta$.

Таким образом, при $m_* > 0$ в точке a функция $f(x)$ имеет строгий локальный минимум, а ее второй дифференциал является положительно определенной квадратичной формой.

Случай 2. Если $m^* < 0$, то $k(y) \leqslant m^* < 0$ при $|y| = 1$. Значит, $k(\Delta x) = k(\nu)|\Delta x|^2 < 0$ при $\Delta x \neq 0$ и квадратичная форма $d^2 f(a)$ отрицательно определена. Выбирая $\varepsilon = |m^*|/2$, заключаем, как и выше, что для достаточно малого числа $\delta > 0$ при $0 < |\Delta x| < \delta$ выполнено неравенство $|\alpha| < |m^*|/2$. В силу представления (1) находим, что

$$f(a + \Delta x) - f(a) < \frac{1}{2}\Big(m^* + \frac{1}{2}|m^*|\Big)|\Delta x|^2 = \frac{1}{4}m^*|\Delta x|^2 < 0,$$

если $0 < |\Delta x| < \delta$.

Итак, при $m^* < 0$ в точке a функция $f(x)$ имеет в точке a строгий локальный максимум, а ее второй дифференциал в этой точке является отрицательно определенной квадратичной формой.

Случай 3. Пусть $m_* < 0 < m^*$. Тогда $k(y^*) = m^* > 0$, $k(y_*) = m_* < 0$. Значит, квадратичная форма $k(y) = y^{\mathrm{T}} f''(a) y$ (она же $d^2 f(a)$) знакопеременная. Выберем $\varepsilon = \min\Big\{\frac{m^*}{2}, \frac{|m_*|}{2}\Big\}$. Тогда существует такое число $\delta > 0$, что $|\alpha| < \varepsilon$ при $0 < |\Delta x| < \delta$. Поэтому при $\Delta x = t y^* \delta$, $0 < t < 1$, имеем $\nu = y^*$ и

$$f(a + \Delta x) - f(a) = \frac{1}{2}\Big(k(y^*) + \alpha\Big)|\Delta x|^2 > \frac{1}{2}\Big(m^* - \frac{1}{2}m^*\Big)|\Delta x|^2 = \frac{1}{4}m^*|\Delta x|^2 > 0,$$

так как $\alpha > -\varepsilon \geqslant -m^*/2$. Аналогично при $\Delta x = t y_* \delta$, $0 < t < 1$, имеем

$$f(a + \Delta x) - f(a) = \frac{1}{2}\Big(k(y_*) + \alpha\Big)|\Delta x|^2 < \frac{1}{2}\Big(m_* + \frac{1}{2}|m_*|\Big)|\Delta x|^2 = \frac{1}{4}m_*|\Delta x|^2 < 0,$$

так как $\alpha < \varepsilon \leqslant |m_*|/2$. Следовательно, в любой окрестности точки a есть значения функции $f(x)$, бо́льшие $f(a)$, и есть значения, меньшие $f(a)$.

Итак, в данном случае точка a не является точкой локального экстремума функции $f(x)$, а ее второй дифференциал в этой точке является знакопеременной квадратичной формой.

С л у ч а й 4. Если $m_* = 0$ или $m^* = 0$, то квадратичная форма $k(y)$ является неотрицательно определенной, неположительно определенной или вообще тождественно равна нулю (если $m_* = m^* = 0$). В этом случае вид квадратичной формы не позволяет сделать какое-либо заключение о характере точки a: в этой точке может быть экстремум, а может и не быть. Но при этом квадратичная форма $d^2 f(a)$ не является ни знакоположительной, ни знакоотрицательной, ни знакопеременной, т.е. он не подпадает ни под одно из трех утверждений теоремы.

Итак, каждый из первых трех случаев соответствует одному из утверждений теоремы, а четвертый случай не соответствует ни одному из ее утверждений. Это завершает доказательство теоремы. $\triangleright$

Напомним, что тип квадратичной формы $d^2 f(a)$ можно определить с помощью критерия Сильвестра или приведением ее к каноническому виду.

Пример 3. У функции $f(x, y) = x^2 + y^4$ только одна критическая точка $(0, 0)$. $d^2 f(0, 0) = 2\,dx^2$ — вырожденная квадратичная форма. Значит, теорема 3 в данном случае ничего не дает, хотя в точке $(0, 0)$ функция $f(x, y)$ имеет локальный минимум (> 0).

Пример 4. Функция $g(x, y) = x^2 - y^4$ также имеет единственную критическую точку $(0, 0)$, причем $d^2 g(0, 0) = 2dx^2$. Но при этом функция $g(x, y)$ не имеет в точке $(0, 0)$ экстремума, так как она в этой точке достигает максимума при фиксированном $x = 0$ и минимума при фиксированном $y = 0$.

12.3 Достаточные условия экстремума функции двух переменных

В случае функции двух переменных достаточные условия экстремума функции в сочетании с критерием Сильвестра приводит к простым правилам проверки.

Предположим, что функция $f(x, y)$ дважды непрерывно дифференцируема в окрестности точки (a, b) и в этой точке выполнено необходимое условие экстремума функции, т.е.

$$f'_x(a, b) = f'_y(a, b) = 0.$$

Матрицей квадратичной формы $d^2 f(a, b)$ является матрица Гессе $f''(a, b)$, которая имеет вид

$$f''(a, b) = \begin{pmatrix} A & B \\ B & C \end{pmatrix}, \qquad A = f''_{xx}(a, b), \quad B = f''_{xy}(a, b), \quad C = f''_{yy}(a, b).$$

Применим к исследованию КФ $d^2 f(a, b)$ критерий Сильвестра и переформулируем утверждения теоремы 2 в случае двух переменных:

1) $A > 0,\ AC - B^2 > 0 \ \Rightarrow\ (a, b)$ — т. строгого лок. минимума ф. f;
2) $A < 0,\ AC - B^2 > 0 \ \Rightarrow\ (a, b)$ — т. строгого лок. максимума ф. f;
3) $AC - B^2 < 0 \ \Rightarrow$ ф. f не имеет в точке (a, b) экстремума.

Приведенные утверждения не охватывают случай $AC = B^2$. В этом случае функция может иметь в точке (a, b) локальный экстремум, а может и не иметь его (см. примеры 3 и 4).

Задачу исследования скалярной функции многих переменных $f\colon A \to \mathbb{R}$ на экстремум часто записывают в виде

$$f(x) \to \mathrm{extr}\,.$$

Такие задачи решают в два этапа. На первом этапе с помощью необходимых условий экстремума отбирают точки, подозрительные на экстремум (критические точки). На втором этапе каждую отобранную точку исследуют на наличие в ней экстремума функции. Это исследование может выполняться либо с помощью различных достаточных условий экстремума, либо с помощью непосредственного анализа поведения функции в окрестности исследуемой точки.

13 Условный экстремум

См. [ККЧ, §7.1–7.4].

13.1 Общая постановка задачи

В приложениях часто встречаются задачи поиска экстремумов функций многих переменных при дополнительных ограничениях на возможные изменения переменных. Такие ограничения могут иметь различный характер. Например, значения переменных должны удовлетворять одному или нескольким уравнениям или неравенствам.

Пример 1. Рассмотрим задачу определения прямоугольника с заданным периметром наибольшей площади. Обозначив через x и y длины сторон прямоугольника, через $2p$ — его периметр, мы придем к задаче поиска максимума площади прямоугольника $S(x,y) = xy$ при дополнительном условии (ограничении) $2(x+y) = 2p$, что кратко можно записать следующим образом:

$$S(x,y) = xy \to \max, \qquad x + y = p. \tag{1}$$

Нас интересует решение задачи в области $x > 0$, $y > 0$.

В данном случае решение задачи легко можно найти, выразив из уравнения $x+y=p$ одно из переменных и подставив найденное выражение в функцию $S(x,y)$. В результате мы придем к задаче поиска минимума действительной функции одного действительного переменного. Например, из уравнения связи находим $y = p - x$. Тогда площадь прямоугольника при заданном ограничении можно представить как функцию только переменного x: $S(x) = x(p-x)$. Исходя из естественных ограничений $x > 0$, $y > 0$, находим область изменения переменного x: $0 < x < p$. Функция $S(x)$ достигает максимума в интервале $(0, p)$ при $x = p/2$, что дает решение рассматриваемой задачи: $x = y = p/2$. Итак, среди всех прямоугольников с заданным периметром наибольшую площадь имеет квадрат.

Отметим, что функция двух переменных $S(x,y) = xy$ не имеет экстремумов, а у рассмотренной задачи решение существует. Это связано с тем, что для задачи (1) не играют роли значения функции $S(x,y)$ в тех точках, которые не удовлетворяют ограничениям. В задачах такого типа все зависит от поведения функции лишь на части ее области определения, а именно на множестве тех точек в области определения, которые подчиняются установленным ограничениям. $\triangleright$

Пусть функции $f\colon\ \mathbb{R}^n \to \mathbb{R}$ и $\varphi\colon\ \mathbb{R}^n \to \mathbb{R}^m$ определены в некоторой окрестности точки $a \in \mathbb{R}^n$. Говорят, что функция f в точке $a \in \mathbb{R}^n$ достигает **условного локального максимума** (**минимума**) при условии $\varphi(x) = 0$, если $\varphi(a) = 0$ и

$$\exists\delta > 0: \quad \forall x \in \dot{U}_\delta(a) \quad \Big(\varphi(x) = 0 \ \ \Rightarrow \ \ f(x) \leqslant f(a) \quad \big(f(x) \geqslant f(a)\big)\Big). \tag{2}$$

Понятия условного локального максимума и минимума объединяют под общим названием **условный экстремум функции**. Если в определении неравенства строгие, то говорят о **строгом условном экстремуме функции**.

Задачу исследования функции $f\colon\ \mathbb{R}^n \to \mathbb{R}$ на условный экстремум при ограничениях $\varphi(x) = 0$ часто записывают в виде

$$f(x) \to \text{extr}, \tag{3}$$

$$\varphi(x) = 0 \tag{4}$$

и называют **задачей на условный экстремум**. При этом функцию $f(x)$ называют **целевой функцией**. Условие (4) в общем случае представляет собой систему нелинейных уравнений, называемых **уравнениями связи**.

Метод решения, использованный в примере 1, может применяться лишь в простейших ситуациях. Распространение этого метода на общий случай наталкивается на трудности, связанные с исключением части переменных из аргументов целевой функции при помощи уравнений связи. Такой подход приводит к необходимости решать систему нелинейных уравнений, а это, как известно, — сложная задача. Отметим, что исключение неизвестных с помощью уравнений связи приводит затем к задаче поиска локального экстремума функции многих переменных, т.е. к решению еще одной системы нелинейных уравнений, которые получаются приравниванием нулю частных производных. Исключение неизвестных нужно лишь затем, чтобы вычислить эти частные производные, но частные производные можно также вычислить и с помощью теоремы о неявной функции. В этом случае исключение неизвестных фактически уже не нужно, и решение задачи упрощается. Развитию этого подхода на основе теоремы о неявной функции мы и уделим внимание.

13.2 Необходимое условие условного экстремума

Остановимся сначала на простейшем случае функции двух переменных.

Теорема 1 (**необходимое условие условного экстр. при** $n = 2, m = 1$). Пусть $f\colon\ \mathbb{R}^2 \to \mathbb{R}$ и $\varphi\colon\ \mathbb{R}^2 \to \mathbb{R}$ — функции двух переменных, определенные и непрерывно дифференцируемые в окрестности точки $P(a, b)$, функция $f(x, y)$ имеет в точке P условный экстремум при условии $\varphi(x, y) = 0$, причем $\operatorname{grad}\varphi(a, b) \neq 0$. Тогда существует такое число λ, которое вместе с координатами a и b точки P удовлетворяет системе уравнений

$$f'_x(a, b) + \lambda\varphi'_x(a, b) = 0, \qquad f'_y(a, b) + \lambda\varphi'_y(a, b) = 0, \qquad \varphi(a, b) = 0. \tag{5}$$

Док–во. Поскольку $\operatorname{grad}\varphi(a, b) \neq 0$, то одна из частных производных первого порядка функции $\varphi(x, y)$ в точке P отлична от нуля. Пусть, например, $\varphi'_y(a, b) \neq 0$. По теореме о неявной функции в некотором прямоугольнике

$$U = \{(x, y): \ |x - a| < \delta_1, \ |y - b| < \delta_2\}$$

уравнение $\varphi(x,y)=0$ разрешимо относительно переменного y, т.е. задает неявную функцию $y=h(x)$, непрерывно дифференцируемую в окрестности точки a, причем

$$b=h(a), \qquad h'(a)=-\frac{\varphi'_x(a,h(a))}{\varphi'_y(a,h(a))}=-\frac{\varphi'_x(P)}{\varphi'_y(P)}. \tag{6}$$

В прямоугольнике U точки, удовлетворяющие условию $\varphi(x,y)=0$, имеют вид $\bigl(x,h(x)\bigr)$, где $x\in(a-\delta_1,\,a+\delta_1)$. Значит, если функция $f(x,y)$ имеет в точке P условный экстремум при условии $\varphi(x,y)=0$, то функция $g(x)=f\bigl(x,h(x)\bigr)$ одного переменного имеет в точке a локальный экстремум. Эта функция, как композиция дифференцируемых функций, является дифференцируемой в точке a. Следовательно, в силу необходимого условия локального экстремума верно соотношение $g'(a)=0$. Согласно правилу дифференцирования сложной функции и равенствам (6), находим

$$g'(a)=f'_x(P)+f'_y(P)h'(a)=f'_x(P)-f'_y(P)\frac{\varphi'_x(P)}{\varphi'_y(P)}=f'_x(P)-\frac{f'_y(P)}{\varphi'_y(P)}\,\varphi'_x(P)=0.$$

Введем обозначение $\lambda=-f'_y(P)/\varphi'_y(P)$. Тогда

$$f'_x(P)+\lambda\varphi'_x(P)=0, \qquad f'_y(P)+\lambda\varphi'_y(P)=0,$$

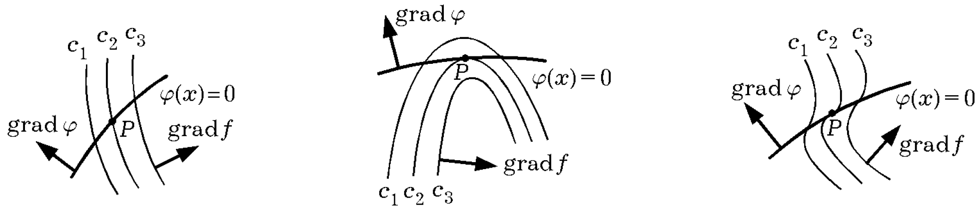
где первое из этих уравнений вытекает из условия $g'(a)=0$, а второе эквивалентно равенству, определяющему число λ. Добавив к этим уравнениям равенство $\varphi(P)=0$, которое должно выполняться в точке условного локального экстремума, получим систему уравнений (5).

Доказательство теоремы в случае, когда $\varphi'_x(P)\neq 0$, проводится аналогично. $\triangleright$

Систему уравнений (5) можно записать в виде

$$\operatorname{grad} f(P)=-\lambda\operatorname{grad}\varphi(P), \qquad \varphi(P)=0.$$

Так как градиент ортогонален линии уровня, получаем следующую **геометрическую интерпретацию необходимых условий условного экстремума**: линия уровня целевой функции касается кривой, заданной уравнением связи. На рис. 9, *а* показано, почему в случае нарушения необходимого условия в точке P нет условного экстремума. Представлены линии уровня $f(x)=c_1$, $f(x)=c_2$ и



а *б* *в*

Рис. 9

$f(x)=c_3$. В изображенной ситуации $c_1<c_2<c_3$ (это определяется направлением

градиента функции $f(x, y)$, являющимся направлением ее роста) и функция $f(x, y)$ на кривой $\varphi(x, y) = 0$ не может иметь экстремума. На рис. 9, *б* показано поведение функции в окрестности условного максимума P. В соответствии с указанным направлением градиента функции $f(x, y)$ имеем $c_1 < c_2 < c_3$, что и обеспечивает локальный максимум $f(x, y)$ в точке P на кривой $\varphi(x, y) = 0$. На рис. 9, *в* изображена ситуация, при которой необходимое условие условного экстремума выполнено, но экстремума тем не менее нет (в соответствии с направлением $\operatorname{grad} f$ в точке P имеем $c_1 < c_2 < c_3$).

Введем функцию

$$L(x, y, \lambda) = f(x, y) + \lambda\varphi(x, y),$$

которую называют **функцией Лагранжа**, где λ — **множитель Лагранжа**. Тогда система (5) будет иметь вид

$$L'_x(x, y, \lambda) = 0, \quad L'_y(x, y, \lambda) = 0, \quad L'_\lambda(x, y, \lambda) = 0. \tag{7}$$

Таким образом, задача на условный экстремум

$$f(x, y) \to \text{extr}, \quad \varphi(x, y) = 0$$

при выполнении условий теоремы 1 сводится к поиску стационарных точек функции Лагранжа и их анализу.

Пример 2. Найдем точки, подозрительные на условный экстремум, в задаче

$$f(x, y) = xy \to \text{extr}, \quad 2(x + y) = 2p, \quad x > 0,\ y > 0,$$

сформулированной в примере 1. Функции $f(x, y) = xy$ и $\varphi(x, y) = 2(x + y) - 2p$ удовлетворяют условиям теоремы 1, поэтому решать задачу можно при помощи функции Лагранжа. Составим функцию Лагранжа:

$$L(x, y, \lambda) = xy + \lambda(x + y - p).$$

Необходимые условия (7) условного экстремума приводят к системе уравнений

$$L'_x(x, y, \lambda) = y + \lambda = 0, \qquad L'_y(x, y, \lambda) = x + \lambda = 0, \qquad L'_\lambda(x, y, \lambda) = x + y - p = 0.$$

Выражая x и y из первых двух уравнений и подставляя эти выражения в третье уравнение, находим $-2\lambda - p = 0$, откуда $\lambda = -p/2$ и $x = y = p/2$. Следовательно, условный экстремум в рассматриваемой задаче может быть только в точке $P(p/2, p/2)$ (рис. 10). ▹

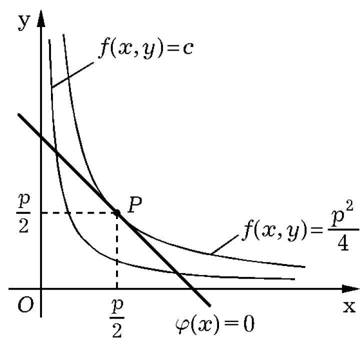


Рис. 10

Необходимое условие для задачи общего вида (3), (4) может быть получено по той же схеме, что и в частном случае двух переменных. Для задачи (3), (4) **функция Лагранжа** по определению имеет вид

$$L(x,\lambda) = f(x) + \lambda\varphi(x) = f(x) + \sum_{i=1}^{m} \lambda_i \varphi_i(x),$$

где $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$, $\varphi(x) = \left(\varphi_1(x), \ldots, \varphi_m(x)\right)^{\mathrm{T}}$. При этом числа λ_i, $i = 1, \ldots, m$, называют **множителями Лагранжа**.

Теорема 2 (**необходимое условие условного экстремума**). Пусть функции $f\colon \mathbb{R}^n \to \mathbb{R}$ и $\varphi\colon \mathbb{R}^n \to \mathbb{R}^m$ определены и непрерывно дифференцируемы в окрестности точки $a \in \mathbb{R}^n$, причем $\operatorname{rang}\varphi'(a) = m$, в точке a функция $f(x)$ имеет условный локальный экстремум при условии $\varphi(x) = 0$. Тогда

$$\exists \lambda_a \in \mathbb{R}^m : \qquad \begin{cases} \frac{\partial L(a,\lambda_a)}{\partial x_1} = 0, \\ \ldots\ldots\ldots\ldots, \\ \frac{\partial L(a,\lambda_a)}{\partial x_n} = 0, \\ \frac{\partial L(a,\lambda_a)}{\partial \lambda_1} = 0, \\ \ldots\ldots\ldots\ldots, \\ \frac{\partial L(a,\lambda_a)}{\partial \lambda_m} = 0. \end{cases} \tag{8}$$

Доказательство этой теоремы опирается на теорему о неявной функции и в основном повторяет доказательство теоремы 1 (случай $n = 2$, $m = 1$).

13.3 Достаточные условия условного экстремума

Достаточные условия условного экстремума в задаче (3), (4) можно сформулировать с помощью функции Лагранжа. Пусть в задаче на условный экстремум функции $f\colon \mathbb{R}^n \to \mathbb{R}$ при условии $\varphi(x) = 0$, заданном функцией $\varphi\colon \mathbb{R}^n \to \mathbb{R}^m$, в точке $a \in \mathbb{R}^n$ выполнено необходимое условие условного экстремума. В этом случае в точке a определен вектор λ_a множителей Лагранжа. Зафиксируем в функции Лагранжа $L(x,\lambda)$ значения множителей Лагранжа, представив ее как функцию только переменных x: $L(x) = L(x,\lambda_a)$. Чтобы выяснить, является ли точка a точкой условного экстремума рассматриваемой функции, нужно проанализировать дифференциал второго порядка $d^2L(a)$ функции $L(x)$ в точке a. Рассмотрим квадратичную форму $d^2L(a)_H$, равную ограничению дифференциала $d^2L(a)$ на линейное подпространство H в $\mathbb{R}^n$, заданное системой линейных уравнений $d\varphi(a) = 0$ (или, по-другому, $\varphi'(a)\,\Delta x = 0$). Т.е.

$$H = \{\Delta x : d\varphi(a) = 0\}, \qquad d^2L(a)_H = d^2L(a)\big|_{d\varphi(a)=0}.$$

Теорема 3 (**достаточные условия условного экстремума**). Пусть функции $f\colon \mathbb{R}^n \to \mathbb{R}$ и $\varphi\colon \mathbb{R}^n \to \mathbb{R}^m$ дважды непрерывно дифференцируемы в окрестности точки $a \in \mathbb{R}^n$, $\varphi(a) = 0$, $\operatorname{rang}\varphi'(a) = m$ и координаты точки a вместе с некоторым вектором λ_a удовлетворяют системе уравнений (8). Тогда:

1) КФ $d^2L(a)_H$ положит. опр. $\Rightarrow$ a — т. строгого услов. лок. минимума ф. f;

2) КФ $d^2L(a)_H$ отрицат. опр. $\Rightarrow$ a — т. строгого услов. лок. максимума ф. f;
3) КФ $d^2L(a)_H$ знакопеременная $\Rightarrow$ в т. a ф. f не имеет услов. экстремума.

Док–во основано на сведение задачи на условный экстремум к задаче на безусловный экстремум. А именно, выразив из уравнений связи $\varphi(x) = 0$ часть переменных (y) через остальные (z) и подставив найденное выражение $(y = h(z))$ в функцию $f(x)$ $(x = (z, y))$, мы придем к задаче поиска экстремума функции $f\big(z, h(z)\big) = L\big(z, h(z), \lambda_a\big)$. Второй дифференциал функции $L\big(z, h(z), \lambda_a\big)$ в точке a отождествляется с квадратичной формой $d^2L(a)_H$, поэтому из достаточных условий экстремума (см. теорему 2 из §12) получаем утверждение данной теоремы. Точные рассуждения следующие.

Ранг матрицы Якоби $\varphi'(a)$ функции $\varphi(x)$ равен m, и в этой матрице можно выделить базисный минор порядка m, равного количеству строк матрицы. Обозначим через $y = (y_1, \ldots, y_m)$ ту часть переменных $x = (x_1, \ldots, x_n)$, которые соответствуют базисным строкам этого минора, а через $z = (z_1, \ldots, z_{n-m})$ — остальные переменные. Тогда функции f и φ можно представить в виде $f(z, y)$ и $\varphi(z, y)$. Определитель матрицы Якоби $\varphi'_y(b, c)$ функции $\varphi(x, y)$ по части переменных y в точке $a = (b, c)$ совпадает с базисным минором и, поэтому, не равен нулю. Следовательно, в точке a выполняются условия теоремы о неявной функции. По этой теореме в некоторой окрестности точки $a = (b, c)$ уравнение $\varphi(z, y) = 0$ неявно задает y как функцию z, т.е. уравнение $\varphi(z, y) = 0$ эквивалентно уравнению $y = h(z)$, где $h(z)$ — непрерывно дифференцируемая функция в окрестности точки $b \in \mathbb{R}^{n-m}$, а $h(b) = c$. А значит, $a = (b, c)$ — точка условного локального экстремум функции $f(x)$ при условии $\varphi(x) = 0$ т. и т.т., когда b — точка локального экстремум функции $g(z) = f\big(z, h(z)\big)$.

По построению функция Лагранжа $L(z, y, \lambda_a)$ на множестве точек, удовлетворяющих уравнениям связи $\varphi(z, y) = 0$, совпадает с функцией $g(z) = f\big(z, h(z)\big)$. Убедимся в том, что в условиях теоремы для функции $g(z) = L(z, h(z), \lambda_a)$ в точке b выполнены достаточные условия экстремума. Для этого вычислим дифференциалы первого и второго порядка этой функции. Согласно свойству инвариантности формы записи дифференциала первого порядка,

$$dg = L'_z\,dz + L'_y\,dy,$$

где $dy = h'(z)\,dz$. Запишем это равенство в точке $(z = b, y = h(b) = c)$ в координатной форме:

$$dg(b) = \sum_{i=1}^{n} L'_{x_i}(b, c)\,dx_i.$$

Согласно равенствам (8), $dg(b) = 0$, т.е. в точке b функция $g(z)$ удовлетворяет необходимым условиям экстремума.

Теперь вычислим дифференциал второго порядка и покажем, что $d^2g(b) = d^2L(b, c)_H$. Имея в виду, что в данном случае переменные, входящие в набор y, являются промежуточными переменными, получаем:

$$d^2g = \sum_{i,j=1}^{n} L''_{x_i x_j} dx_i\,dx_j + \sum_{i=1}^{m} L'_{y_i} d^2y_i.$$

Так как в силу условий теоремы в точке (b, c) выполнены равенства (8), то

$$d^2g(b) = \sum_{i=1}^{n} L''_{x_i x_j}(b, c) dx_i\, dx_j. \tag{9}$$

Чтобы получить окончательный вид дифференциала второго порядка функции $g(z)$ в точке b, нужно исключить дифференциалы промежуточных переменных y, т.е. в правой части равенства (9) выполнить замену $dy = h'(b)\, dz$.

Из теоремы о производной неявной функции следует, что матрица Якоби $h'(b)$ может быть вычислена по формуле

$$h'(b) = -\big(\varphi'_y(b, c)\big)^{-1} \varphi'_z(b, c).$$

Поэтому, умножая систему $h'(b)dz - dy = 0$ на невырожденную матрицу $-\varphi'_y(b, c)$, получаем систему $d\varphi(b, c) = \varphi'_z(b, c)dz + \varphi'_y(b, c)dy = 0$. Таким образом, дифференциал второго порядка функции $g(z) = L\big(z, h(z)\big)$ в точке $z = b$ совпадает с ограничением $d^2L(b, c)_H$ на линейное подпространство H дифференциала второго порядка функции Лагранжа $L(z, y)$.

Применение достаточных условий экстремума для функций многих переменных завершает доказательство теоремы. Например, если квадратичная форма $d^2L(b, c)_H$ положительно определенная, то и $d^2g(b)$ является положительно определенной квадратичной формой. Согласно достаточным условиям экстремума, функция $g(z) = L\big(z, h(z)\big)$ имеет в точке b локальный минимум. Это равносильно тому, что функция $L(z, y)$ (а значит, и функция $f(z, y)$) имеет в точке $a = (b, c)$ условный локальный минимум при условии $y = h(z)$ (или $\varphi(z, y) = 0$). $\triangleright$

Теорема 3 утверждает, что для проверки точек, подозрительных на условный экстремум, необходимо проанализировать квадратичную форму $d^2L(a)$, т.е. дифференциал второго порядка функции Лагранжа при $\lambda = \lambda_a$ и значениях приращений Δx, которые удовлетворяют системе линейных уравнений

$$d\varphi_i = \sum_{j=1}^{m} \frac{\partial \varphi_i(a)}{\partial x_j} \Delta x_j = 0, \quad i = 1, \ldots, m. \tag{10}$$

Матрица этой системы линейных алгебраических уравнений совпадает с матрицей Якоби $\varphi'(a)$ функции $\varphi(x)$ в точке a, ранг которой по условию теоремы 3 равен m. Следовательно, система (10) позволяет выразить m приращений через оставшиеся $n - m$ приращений. Зафиксируем известные значения множителей Лагранжа (координат вектора λ). Рассматривая функцию Лагранжа $L(x)$ как функцию только переменных x_1, ..., x_n, вычислим ее дифференциал второго порядка d^2L в точке a. Исключим из квадратичной формы d^2L указанные m дифференциалов. Получим квадратичную форму относительно $n - m$ дифференциалов. Если эта квадратичная форма является положительно определенной (отрицательно определенной, знакопеременной), то в точке a функция $f(x)$ имеет условный локальный минимум (условный локальный максимум, не имеет условного локального экстремума). Если указанная квадратичная форма от $n - m$ переменных вырождена, но сохраняет знак (неположительно или неотрицательно определена), то в точке a функция $f(x)$ может иметь условный локальный экстремум, а может и не иметь. В этом случае по виду второго дифференциала в точке выявить поведение функции $f(x)$ нельзя и нужны другие методы исследования.

Пример 3. В примере 2 уравнение $d\varphi = 0$ дает $dx + dy = 0$, откуда можно, например, выразить dx через dy: $dx = -dy$. Дифференциал второго порядка функции Лагранжа $L(x, y, \lambda)$ при фиксированном значении $\lambda = -p/2$ в точке $(p/2, p/2)$ имеет вид $d^2L = 2dxdy$. Исключая из второго дифференциала dx, получаем квадратичную форму $2(-dy)dy = -2(dy)^2$, которая отрицательно определена. Следовательно, в точке $(p/2, p/2)$ мы имеем условный локальный максимум. $\triangleright$

Рассмотрим **задачу вычисления наименьшего и наибольшего значений** непрерывной функции n переменных на компактном множестве K, которое задано, например, m неравенствами:

$$\begin{aligned}&f(x) \to \mathrm{extr},\\ &g_i(x) \leqslant 0, \quad i = 1, \ldots, m.\end{aligned}$$

Для решения этой задачи отбираются точки, в которых может достигаться наибольшее или наименьшее значение:

а) среди внутренних точек множества K;
б) на всех k–мерных частях границы ($k = 1, \ldots, n-1$);
в) все нульмерные элементы границы.

Например, при $n = 2$ имеем

а) внутренние точки K — это все точки, удовлетворяющие неравенствам $g_i(x) < 0, \quad i = 1, \ldots, m$;
б) 1-мерные части границы — точки, удовлетворяющие одному равенству $g_j(x) = 0$ и $m - 1$ неравенствам $g_i(x) < 0, \quad i \neq j$;
в) нульмерные элементы границы — точки, удовлетворяющие двум равенствам $g_j(x) = 0, g_l(x) = 0$ и $m - 2$ неравенствам $g_i(x) < 0, \quad i \neq j, l$.

Отбор точек внутри K приводит к задаче на локальный экстремум, отбор точек на k–мерных частях границы приводит к задаче на условный локальный экстремум. При этом проверяются только необходимые условия данных задач. Затем во всех точках, удовлетворяющих необходимым условиям, и нульмерных элементах границы вычисляются значения функции и выбираются точки с наименьшим и наибольшим значением. Проверять достаточные условия экстремума и условного экстремума нет необходимости, так как из свойств функций, непрерывных на компакте, следует, что на K функция f достигает своих наименьшего и наибольшего значений, а из необходимых условий следует, что эти значения могут достигаться только в рассматриваемых точках.

14 Обратная функция

См. [ККЧ, §4.3].

Рассмотрим вопрос, при каких условиях функция многих переменных $G\colon \mathbb{R}^n \to \mathbb{R}^n$ имеет обратную функцию G^{-1}, а также вопрос о том, дифференцируема ли обратная функция. Соответствующие условия в окрестности фиксированной точки можно получить с помощью теоремы о неявной функции.

Теорема 1 (**об обратной функции**). Пусть функция $G\colon\ \mathbb{R}^n \to \mathbb{R}^n$ в окрестности V точки $a \in \mathbb{R}^n$ удовлетворяет условиям:

1) $G \in C^1(V, \mathbb{R}^n)$;

2) $\det G'(a) \neq 0$.

Тогда $\exists U$ (окр. т. $b = G(a)$), $\exists G^{-1}\colon\ U \to V$:

а) $\forall y \in U \quad G\big(G^{-1}(y)\big) = y \quad$ и $\quad \forall x \in G^{-1}(U) \quad G^{-1}\big(G(x)\big) = x$;

б) $G^{-1} \in C^1(U, \mathbb{R}^n)$ и

$$\big(G^{-1}\big)'(y) = \big(G'(x)\big)^{-1}\Big|_{x=G^{-1}(y)}. \tag{1}$$

Док–во. Рассмотрим функцию $f\colon\ \mathbb{R}^{2n} \to \mathbb{R}^n$, определяемую равенством $f(x,y) = G(x) - y$. Эта функция непрерывно дифференцируема в окрестности точки $(x = a; y = b) \in \mathbb{R}^{2n}$, а множество решений системы n уравнений $f(x,y) = 0$ представляет собой график функции $G(x)$, т.е. множество точек $(x; y)$, удовлетворяющих условию $y = G(x)$. В частности, $f(a,b) = 0$. Так как $\det G'(a) \neq 0$, то матрица Якоби $f'_x(a,b) = G'(a)$ невырождена. Таким образом, для функции $f(x,y)$ в окрестности точки $(a;b)$ выполнены условия теоремы 3 о неявной функции. Это значит, что система уравнений $f(x,y) = 0$ в некоторой окрестности W вида $W = \{(x;y) \in \mathbb{R}^{2n}\colon\ |x-a| < \delta_x,\ |y-b| < \delta_y\}$ разрешима относительно переменных x, т.е. существует такая функция $h(y)$, определенная в окрестности $U_{\delta_y}(b)$ точки b, что

$$f\big(h(y), y\big) \equiv 0, \tag{2}$$

причем функция $h(y)$ непрерывно дифференцируема, а ее матрица Якоби равна

$$h'(y) = -\Big(f'_x\big(h(y), y\big)\Big)^{-1} f'_y\big(h(y), y\big). \tag{3}$$

Так как $f(x,y) = G(x) - y$, тождество (2) означает, что $G\big(h(y)\big) \equiv y$, т.е. функция $h(y)$ является обратной к функции $G(x)$. Кроме того, матрица $f'_y(x,y)$ совпадает с матрицей $-E$, противоположной единичной матрице E. Поэтому равенство (3) сводится к равенству $h'(y) = \big(G'(h(y))\big)^{-1}$, равносильному (1) $\triangleright$

Пример 1. Рассмотрим отображение $G\colon\ \mathbb{R}^2 \to \mathbb{R}^2$, заданное уравнениями $z_1 = x_1 + e^{x_2}$, $z_2 = e^{x_1} - x_2$. Имеем $G \in C^1(\mathbb{R}^2, \mathbb{R}^2)$ и

$$\forall\,(x_1; x_2) \in \mathbb{R}^2, \qquad G'(x_1, x_2) = \begin{pmatrix} 1 & e^{x_2} \\ e^{x_1} & -1 \end{pmatrix}, \qquad \det G'(x_1, x_2) = -1 - e^{x_1 + x_2} \neq 0.$$

По теореме об обратной функции, в любой точке $b \in \mathbb{R}^2$, $b = G(a)$, существует окрестность, в которой определено обратное отображение G^{-1}, причем $G^{-1}(b) = a$. $\triangleright$

Доказательство теоремы 1 показывает, что теорема об обратной функции сводится к теореме о неявной функции. Любопытно, что можно делать и наоборот: теорему о неявной функции выводить из теоремы об обратной функции. В самом деле, пусть для функции f выполнены условия теоремы 3 из §10. Систему уравнений $f(x,y) = 0$, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, можно трактовать как частный случай системы уравнений $f(x,y) = z$ при фиксированном $z \in \mathbb{R}^n$. Рассмотрим отображение

$$H(x,y) = \begin{pmatrix} x \\ f(x,y) \end{pmatrix},$$

которое получается добавлением новых n координатных функций. Это отображение дифференцируемо в окрестности точки $(a;b)$, а его матрица Якоби может быть записана как блочная матрица

$$H'(x,y) = \begin{pmatrix} E & 0 \\ f'_x(x,y) & f'_y(x,y) \end{pmatrix},$$

в которой символ 0 обозначает нулевой блок соответствующего типа, а E — единичную матрицу соответствующего порядка. Видно, что $\det H'(a,b) = \det E \cdot \det f'_y(a,b) \neq 0$ и для функции H выполнены все условия теоремы об обратной функции в окрестности точки $(a;b)$, причем $H(a,b) = (a;0)$. Значит, по теореме об обратной функции в окрестности точки $(a;0)$ существует обратное отображение $H^{-1}(x,z)$, которое можно представить в виде

$$H^{-1}(x,z) = \begin{pmatrix} \varphi(x,z) \\ \psi(x,z) \end{pmatrix},$$

где $\varphi(x,z)$ — совокупность первых n координатных функций, а $\psi(x,z)$ — совокупность оставшихся m координатных функций. Тождество $H\bigl(H^{-1}(x,z)\bigr) \equiv (x;z)$ в блочной форме имеет вид

$$\begin{pmatrix} \varphi(x,z) \\ f\bigl(\varphi(x,z),\, \psi(x,z)\bigr) \end{pmatrix} \equiv \begin{pmatrix} x \\ z \end{pmatrix}.$$

Отсюда находим, что $\varphi(x,z) \equiv x$ и

$$f\bigl(\varphi(x,z),\, \psi(x,z)\bigr) \equiv f\bigl(x,\, \psi(x,z)\bigr) \equiv z.$$

Но последнее тождество и означает, что уравнение $f(x,y) = z$ разрешимо относительно переменных y и может быть представлено в виде $y = \psi(x,z)$. В частном случае $z = 0$ получаем разрешимость уравнения $f(x,y) = 0$. Используя правила операций с блочными матрицами, можно из формулы (1) для матрицы Якоби обратной функции получить формулу для матрицы Якоби неявной функции. Действительно, для невырожденной квадратной матрицы B порядка m и матрицы A типа $m \times n$

$$\begin{pmatrix} E & 0 \\ A & B \end{pmatrix}^{-1} = \begin{pmatrix} E & 0 \\ -B^{-1}A & B^{-1} \end{pmatrix}.$$

Исходя из этого, заключаем, что, согласно теореме об обратной функции,

$$(H^{-1})'(x,z) = \begin{pmatrix} E & 0 \\ -\bigl(f'_y(x,y)\bigr)^{-1} f'_x(x,y) & \bigl(f'_y(x,y)\bigr)^{-1} \end{pmatrix}.$$

Следовательно,

$$\psi'(x,z) = \Bigl(-\bigl(f'_y(x,y)\bigr)^{-1} f'_x(x,y) \qquad \bigl(f'_y(x,y)\bigr)^{-1}\Bigr).$$

В частном случае $z = 0$ для функции $\psi(x,0)$, неявно заданной уравнением $f(x,y) = 0$, получаем $\psi'_x(x,0) = -\bigl(f'_y(x,y)\bigr)^{-1} f'_x(x,y)$, что соответствует формуле производной неявно заданной функции.