

Выборный Евгений Викторович
email: evybornyi@hse.ru

Математический анализ
Тема 6: Функции многих переменных

Москва 2016

Определение. Вещественное n -мерное пространство \mathbb{R}^n

Множество упорядоченных наборов из n действительных чисел называют **вещественным n -мерным пространством \mathbb{R}^n** :

$$x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n.$$

Эти наборы чисел из \mathbb{R}^n называют точками или векторами. В \mathbb{R}^n определена сумма векторов и операция умножения вектора на число:

$$x + y = (x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n),$$

$$\alpha x = \alpha \cdot (x_1, x_2, \dots, x_n) = (\alpha x_1, \alpha x_2, \dots, \alpha x_n).$$

Определено понятие расстояния между точкам:

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

Выполнено неравенство треугольника:

$$\|x + y\| \leq \|x\| + \|y\|.$$

Шар в \mathbb{R}^n

Ключевым понятием для определения сходимости в одномерном случае была ϵ -окрестность точки a . Определим аналогичные понятия в многомерном пространстве \mathbb{R}^n .

Определение. Шар в \mathbb{R}^n

Открытым шаром в \mathbb{R}^n с центром в точке a и радиусом r называют множество точек $x \in \mathbb{R}^n$, удовлетворяющих условию

$$\|x - a\| < r \iff (x_1 - a_1)^2 + \dots + (x_n - a_n)^2 < r^2.$$

Иногда это множество называют r -окрестностью точки a , сохраняя обозначение $O_r(a)$. Тогда **проколотой r -окрестностью** точки a , называют множество точек:

$$\dot{O}_r(a) = O_r(a) \setminus \{a\} = \{x \in \mathbb{R}^n \mid 0 < \|x - a\| < r\}.$$

Определение. Ограниченное множество

Множество $A \subset \mathbb{R}^n$ называется **ограниченным**, если A полностью лежит в некотором шаре. В этом случае существует R такое, что

$$\|x\| < R \quad \forall x \in A.$$

Предел последовательности точек

Определение. Предел последовательности точек

Говорят, что последовательность точек $\{x^{(k)}\}$, $x^{(k)} \in \mathbb{R}^n$ сходится к точке $y \in \mathbb{R}^n$, пишут $x^{(k)} \rightarrow y$, если к нулю стремится расстояние между y и $x^{(k)}$ при $k \rightarrow +\infty$:

$$\lim_{k \rightarrow +\infty} d(y, x^{(k)}) = 0.$$

Эквивалентные записи имеют вид:

$$\forall \varepsilon > 0 \exists N : x^{(k)} \in O_\varepsilon(y) \quad \forall k \geq N.$$

$$\forall \varepsilon > 0 \exists N : \|x^{(k)} - y\| < \varepsilon \quad \forall k \geq N.$$

Таким образом, последовательность точек стремится к y тогда и только тогда, когда в любом открытом шаре с центром в точке y лежит бесконечно много точек последовательности, а вне его — лишь конечное число.

Упражнение

Докажите, что множество точек сходящейся последовательности является ограниченным.

Предел последовательности точек

Предложение

Сходимость последовательности точек $\{x^{(k)}\}$ к точке y эквивалентна сходимости координат точек $x^{(k)} = (x_1^{(k)}, \dots, x_n^{(k)})$ к координатам точки $y = (y_1, \dots, y_n)$:

$$\lim_{k \rightarrow +\infty} x^{(k)} = y \iff \lim_{k \rightarrow +\infty} x_1^{(k)} = y_1, \dots, \lim_{k \rightarrow +\infty} x_n^{(k)} = y_n.$$

Доказательство

Доказательство теоремы непосредственно следует из очевидных неравенств:

$$|x_j^{(k)} - y_j| \leq \sqrt{(x_1^{(k)} - y_1)^2 + \dots + (x_n^{(k)} - y_n)^2} \leq n \max_{1 \leq j \leq n} |x_j^{(k)} - y_j|.$$

Замечание

Иногда в \mathbb{R}^n вводят другое понятие расстояния по формуле:

$$\tilde{d}(x, y) = \max_{1 \leq j \leq n} |x_j - y_j|.$$

Следовательно, сходимость последовательности точек относительно расстояния d и \tilde{d} эквивалентна.

Открытые множества

Определение. Внутренние точки множества

Точка $a \in A \subset \mathbb{R}^n$, которая принадлежит множеству A вместе с некоторым открытым шаром с центром в точке a , называется **внутренней точкой** множества A .

Определение. Открытое множество

Множество точек $A \subset \mathbb{R}^n$ называется **открытым**, если для каждой точки $a \in A$ этого множества существует открытый шар с центром в точке a , который полностью лежит в A :

$$A \text{ — открыто} \iff \forall a \in A \exists r > 0 : O_r(a) \subset A.$$

Пустое множество \emptyset полагается открытым по определению.

Таким образом, открытое множество — это множество, которое полностью состоит из внутренних точек.

Пример

Открытый шар является открытым множеством. Действительно, $\forall x \in A = O_R(a)$ положим $r = R - \|x - a\| > 0$. Тогда

$$y \in O_r(x) \Rightarrow \|y - a\| = \|y - x + x - a\| \leq \|y - x\| + \|x - a\| < r + \|x - a\| = R \Rightarrow y \in A.$$

Свойства открытых множеств

- ❶ Все пространство \mathbb{R}^n является открытым.
- ❷ Любое объединение открытых множеств является открытым.
- ❸ Конечное пересечение открытых множеств является открытым.

Замечание

Пересечение бесконечного числа открытых множеств может не быть открыто. Например,

$$A_k = (-1/k, +1/k) \subset \mathbb{R}, \quad k = 1, 2, \dots$$

Множества A_k открыты, но

$$\bigcap_{k=1}^{+\infty} A_k = \{x \in \mathbb{R} \mid x \in A_k \ \forall k\} = \{0\},$$

а множество, состоящее только из одной точки, не является открытым.

Замкнутые множества

Определение. Предельные и изолированные точки множества

Точка $a \in \mathbb{R}^n$ называется **предельной точкой** множества A или точкой сгущения, если в любой окрестности точки a существуют точки из множества A , отличные от a :

$$\forall r > 0 \quad O_r(a) \cap A \neq \{a\}.$$

Точка $a \in A$ называется **изолированной точкой** множества A , если существует окрестность точки a , в которой нет других точек из множества A .

Предельные точки могут как принадлежать, так и не принадлежать рассматриваемому множеству.

Определение. Замкнутое множество

Множество называется **замкнутым**, если оно содержит все свои предельные точки. Пустое множество считают замкнутым по определению.

Предложение. Замкнутость в терминах последовательностей

Множество замкнуто тогда и только тогда, когда предел любой сходящейся последовательности точек этого множества также принадлежит этому множеству.

Свойства замкнутых множеств

Свойства замкнутых множеств

- ❶ Множество является замкнутым тогда и только тогда, когда его дополнение является открытым:

$$A \text{ — замкнуто} \iff (\mathbb{R}^n \setminus A) \text{ — открыто.}$$

- ❷ Все пространство \mathbb{R}^n является замкнутым.
❸ Конечное объединение замкнутых множеств является замкнутым.
❹ Любое пересечение замкнутых множеств является замкнутым.

Определение. Компакт

Замкнутое ограниченное множество в \mathbb{R}^n называют **компактом**.

Понятие компакта является естественным обобщением понятия отрезка в многомерном пространстве.

Предложение. Компактность в терминах последовательностей

Множество является компактом тогда и только тогда, когда из любой последовательности точек множества можно выбрать подпоследовательность, сходящуюся к точке из заданного множества.

Граница множества

Определение. Граница множества

Точка $x \in \mathbb{R}^n$ называется **граничной точкой** для множества $M \subset \mathbb{R}^n$, если в любой окрестности точки x есть как точки из множества M , так и точки не принадлежащие M . Граничные точки могут принадлежать или не принадлежать множеству M .

Множество всех граничных точек для заданного множества M называют **границей** M и обозначают ∂M .

Несложно доказать, что замкнутое множество всегда содержит свою границу.

Объединение множества и его границы всегда является замкнутым. Это множество называют замыканием множества M и обозначают

$$\bar{M} = M \cup \partial M.$$

Пример

Несложно найти границы следующих множеств:

$$\partial [a, b] = \{a, b\}, \quad \partial (a, b) = \{a, b\};$$

$$\partial O_r(a) = \{x \in \mathbb{R}^n \mid \|x - a\| = r\},$$

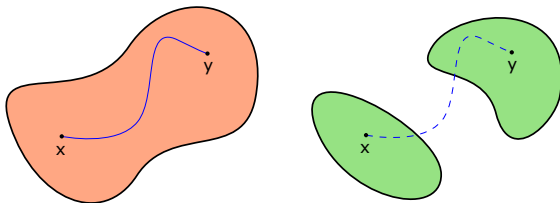
$$\partial \mathbb{R}^n = \emptyset.$$

Определение. Связное множество

Множество $M \subset \mathbb{R}^n$ является **связным** (линейно связным), если для любой пары точек x и y из M существует непрерывный путь (кривая), которая соединяет точки x и y , и при этом полностью лежит в M .

Определение. Область

Областью в $M \subset \mathbb{R}^n$ называют открытое связное множество.

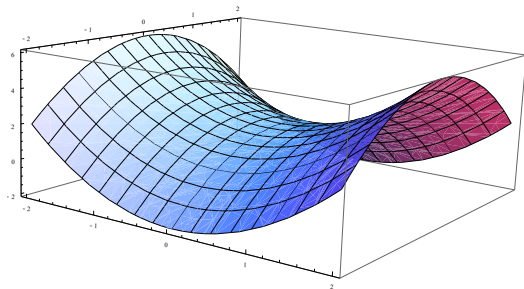


Множество, изображенное на рисунке слева, является связным, а множество, изображенное справа, не является связным (состоит из двух частей).

Функция нескольких переменных

Определение. Функция нескольких переменных

Числовой функцией нескольких переменных называют отображение $f : E \rightarrow \mathbb{R}$, где $E \subset \mathbb{R}^n$ — некоторое множество, называемое **множеством определения** функции. Значение функции f в точке $x \in E$ записывают, как $f(x) = f(x_1, \dots, x_n)$, при этом x_j называют независимыми переменными, а $z = f(x)$ — зависимой переменной, так как ее значение определяется выбором точки x .



При рассмотрении функций двух переменных $z = f(x, y)$ можно рассматривать график функции как поверхность Γ в трехмерном пространстве \mathbb{R}^3 :

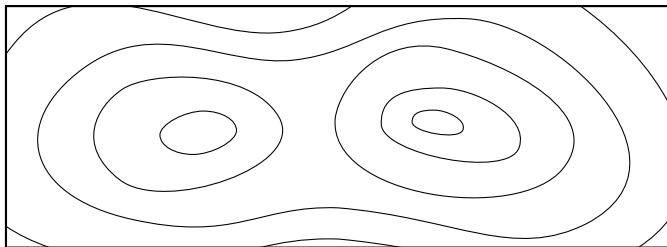
$$\Gamma = \{(x, y, z) \mid z = f(x, y), (x, y) \in E\}.$$

Линии уровни

Другой способ визуально представить функцию двух независимых переменных — это рассмотреть семейство кривых на плоскости, вдоль которых функция является постоянной

$$f(x, y) = \text{const}$$

Данные кривые называют **линиями уровня** для функции f .



Предел функции

Определение. Предел функции

Пусть функция f определена в некоторой проколотой окрестности точки $a \in \mathbb{R}^n$. Говорят, что число f_0 является **пределом** $f(x)$ при $x \rightarrow a$, если

$$\forall \varepsilon > 0 \exists \delta > 0 : |f(x) - f_0| < \varepsilon, \forall x \in \dot{O}_\delta(a).$$

В этом случае пишут $\lim_{x \rightarrow a} f(x) = f_0$.

В случае двух переменных иногда пишут

$$\lim_{x \rightarrow x_0, y \rightarrow y_0} f(x, y) = f_0,$$

а соответствующий предел называют двойным.

Как и в одномерном случае, можно определить сходимость в терминах последовательностей (по Гейне).

Предел $f(x)$ равен f_0 при $x \rightarrow a$ тогда и только тогда, когда для любой сходящейся к a последовательности точек $\{x^{(k)}\}$ из проколотой окрестности точки a последовательность значений функции в этих точках $f(x^{(k)})$ сходится к f_0 :

$$\forall \{x^{(k)}\} : x^{(k)} \rightarrow a, x^{(k)} \neq a \Rightarrow f(x^{(k)}) \rightarrow f_0.$$

По аналогии с одномерным случаем определяются и бесконечные пределы функций.

Предел функции

Пусть функция f определена на множестве M и точка a является предельной точкой множества M . Тогда уместно говорить о стремлении $x \rightarrow a$ при условии $x \in M, x \neq a$.

Определение. Предел функции по множеству

Говорят, что число f_0 является пределом функции $f(x)$ при $x \rightarrow a$ по множеству M ($x \rightarrow a, x \in M$), если

$$\forall \varepsilon > 0 \exists \delta > 0 : |f(x) - f_0| < \varepsilon, \forall x \in \dot{O}_\delta(a) \cap M.$$

В этом случае пишут $\lim_{x \rightarrow a, x \in M} f(x) = f_0$.

Частным случаем предела по множеству служат односторонние пределы функции одной переменной.

Непрерывность функции

Определение непрерывности в точке

Говорят, что функция $f(x)$, определенная в некоторой окрестности точки $A \in \mathbb{R}^n$, **непрерывна в точке** $x = A$, если

$$\exists \lim_{x \rightarrow A} f(x) = f(A).$$

Предположим, что функция определена на множестве M и точка $A \in M$ является предельной точкой этого множества. Если точка A не является внутренней для множества M , а принадлежит его границе, то рассматривают следующие определение непрерывности:

Определение непрерывности по множеству

Говорят, что функция $f(x)$, определенная на множестве $M \subset \mathbb{R}^n$, **непрерывна по множеству** M в точке $x = A$, если

$$\exists \lim_{x \rightarrow A, x \in M} f(x) = f(A).$$

Функция считается по определению непрерывной в изолированных точках множества M .

Непрерывность функции

Определение непрерывности в точке

Говорят, что функция $f(x)$, определенная в некоторой окрестности точки $A \in \mathbb{R}^n$, **непрерывна в точке** $x = A$, если

$$\exists \lim_{x \rightarrow A} f(x) = f(A).$$

Предположим, что функция определена на множестве M и точка $A \in M$ является предельной точкой этого множества. Если точка A не является внутренней для множества M , а принадлежит его границе, то рассматривают следующие определение непрерывности:

Определение непрерывности по множеству

Говорят, что функция $f(x)$, определенная на множестве $M \subset \mathbb{R}^n$, **непрерывна по множеству** M в точке $x = A$, если

$$\exists \lim_{x \rightarrow A, x \in M} f(x) = f(A).$$

Функция считается по определению непрерывной в изолированных точках множества M .

Непрерывность функции

Определение непрерывности в точке

Говорят, что функция $f(x)$, определенная в некоторой окрестности точки $A \in \mathbb{R}^n$, **непрерывна в точке** $x = A$, если

$$\exists \lim_{x \rightarrow A} f(x) = f(A).$$

Предположим, что функция определена на множестве M и точка $A \in M$ является предельной точкой этого множества. Если точка A не является внутренней для множества M , а принадлежит его границе, то рассматривают следующие определение непрерывности:

Определение непрерывности по множеству

Говорят, что функция $f(x)$, определенная на множестве $M \subset \mathbb{R}^n$, **непрерывна по множеству** M в точке $x = A$, если

$$\exists \lim_{x \rightarrow A, x \in M} f(x) = f(A).$$

Функция считается по определению непрерывной в изолированных точках множества M .

Непрерывность функции

Определение непрерывности по множеству

Говорят, что функция $f(x)$, определенная на множестве $M \subset \mathbb{R}^n$, **непрерывна** на множестве M , если она непрерывна в каждой точке множества M .

Свойства непрерывных функций нескольких переменных во многом совпадают со свойствами непрерывных функций одной переменной.

Свойства непрерывных функций

- ❶ Сумма и произведение непрерывных функций непрерывны. Частное непрерывных функций заведомо непрерывно, если делитель (знаменатель) не обращается в ноль.
- ❷ Композиция непрерывных функций непрерывна.
- ❸ Элементарные функции непрерывны на своем множестве определения.

Теорема Вейерштрасса

Пусть функция $f(x)$ непрерывна на компакте $K \subset \mathbb{R}^n$. Тогда она ограничена на множестве K и достигает на нем своей верхней и нижней грани:

$$\exists x_{\min} \in K : f(x_{\min}) = \inf_{x \in K} f(x),$$

$$\exists x_{\max} \in K : f(x_{\max}) = \sup_{x \in K} f(x).$$

Непрерывность функции

Определение непрерывности по множеству

Говорят, что функция $f(x)$, определенная на множестве $M \subset \mathbb{R}^n$, **непрерывна** на множестве M , если она непрерывна в каждой точке множества M .

Свойства непрерывных функций нескольких переменных во многом совпадают со свойствами непрерывных функций одной переменной.

Свойства непрерывных функций

- ❶ Сумма и произведение непрерывных функций непрерывны. Частное непрерывных функций заведомо непрерывно, если делитель (знаменатель) не обращается в ноль.
- ❷ Композиция непрерывных функций непрерывна.
- ❸ Элементарные функции непрерывны на своем множестве определения.

Теорема Вейерштрасса

Пусть функция $f(x)$ непрерывна на компакте $K \subset \mathbb{R}^n$. Тогда она ограничена на множестве K и достигает на нем своей верхней и нижней грани:

$$\exists x_{\min} \in K : f(x_{\min}) = \inf_{x \in K} f(x),$$

$$\exists x_{\max} \in K : f(x_{\max}) = \sup_{x \in K} f(x).$$

Непрерывность функции

Определение непрерывности по множеству

Говорят, что функция $f(x)$, определенная на множестве $M \subset \mathbb{R}^n$, **непрерывна** на множестве M , если она непрерывна в каждой точке множества M .

Свойства непрерывных функций нескольких переменных во многом совпадают со свойствами непрерывных функций одной переменной.

Свойства непрерывных функций

- ❶ Сумма и произведение непрерывных функций непрерывны. Частное непрерывных функций заведомо непрерывно, если делитель (знаменатель) не обращается в ноль.
- ❷ Композиция непрерывных функций непрерывна.
- ❸ Элементарные функции непрерывны на своем множестве определения.

Теорема Вейерштрасса

Пусть функция $f(x)$ непрерывна на компакте $K \subset \mathbb{R}^n$. Тогда она ограничена на множестве K и достигает на нем своей верхней и нижней грани:

$$\exists x_{\min} \in K : f(x_{\min}) = \inf_{x \in K} f(x),$$

$$\exists x_{\max} \in K : f(x_{\max}) = \sup_{x \in K} f(x).$$

Частные производные. Определение

Рассмотрим функцию $f(x) = f(x_1, \dots, x_n)$, заданную в окрестности точки $x^0 = (x_1^0, \dots, x_n^0) \in \mathbb{R}^n$. Тогда можно рассмотреть функцию одной переменной x_k , фиксируя остальные переменные:

$$\phi(x_k) = f(x_1^0, \dots, x_{k-1}^0, x_k, x_{k+1}^0, \dots, x_n^0).$$

Для определения скорости изменения значения функции $f(x)$ при изменении только одной переменной x_k можно рассмотреть производную функции $\phi(x_k)$. Эту производную называют **частной производной** функции f по переменной x_k в точке x^0 и обозначают:

$$\left. \frac{\partial f(x)}{\partial x_k} \right|_{x=x^0} = \left. \frac{\partial}{\partial x_k} \right|_{x=x^0} f(x) = \frac{\partial f(x^0)}{\partial x_k} = \frac{\partial f}{\partial x_k}(x^0) = f'_{x_k}(x^0) = f_{x_k}(x^0) = \partial_k f(x^0).$$

Определение

Говорят, что функция f , определенная в окрестности точки $x^0 \in \mathbb{R}^n$, имеет **частную производную** по переменной x_k в точке x^0 , если существует предел

$$\lim_{h \rightarrow 0} \frac{f(x_1^0, \dots, x_{k-1}^0, x_k^0 + h, x_{k+1}^0, \dots, x_n^0) - f(x^0)}{h} = \left. \frac{\partial f(x)}{\partial x_k} \right|_{x=x^0}.$$

Пример

Вычислим частные производные функции трех переменных

$$u = x^2 + 2xy + \cos(xz).$$

Тогда

$$\frac{\partial u}{\partial x} = 2x + 2y - \sin(xz) z;$$

$$\frac{\partial u}{\partial y} = 2x;$$

$$\frac{\partial u}{\partial z} = -x \sin(xz).$$

В данном случае вычисления производились в произвольной точке $(x, y, z) \in \mathbb{R}^3$. Тогда при дифференцировании по одной из переменных все остальные переменные можно считать постоянными.

Производная по направлению

Частная производная отражает изменение функции при изменении только одной переменной. Иногда удобно рассматривать производные и по другим направлениям не связанным с координатными осями.

Определение

Говорят, что функция $f(x)$, определенная в некоторой окрестности точки $x_0 \in \mathbb{R}^n$, имеет в точке x_0 **производную по направлению** l :

$$l = (l_1, \dots, l_n), \quad l_1^2 + \dots + l_n^2 = 1,$$

если существует предел:

$$\frac{\partial f}{\partial l}(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + hl) - f(x_0)}{h}.$$

Например, в двумерном случае направление l можно представить в виде:

$$l = (\cos \alpha, \sin \alpha).$$

Тогда производная по направлению l функции $f(x, y)$ примет вид:

$$\frac{\partial f}{\partial l} = \lim_{h \rightarrow 0} \frac{f(x + h \cos \alpha, y + h \sin \alpha) - f(x, y)}{h}.$$

Определение. Дифференцируемая функция

Говорят, что функция f , определенная в окрестности точки $x^0 \in \mathbb{R}^n$, **дифференцируема** в точке x^0 , если для любого достаточно маленького приращения Δx переменной x :

$$\Delta x = (\Delta x_1, \dots, \Delta x_n),$$

приращение значения функции представимо в виде:

$$\Delta f(x^0) = f(x^0 + \Delta x) - f(x^0) = \sum_{k=1}^n A_k \Delta x_k + o(\|\Delta x\|),$$

где A_k — постоянные.

Линейную часть приращения функции называют **дифференциалом** и обозначают:

$$df(x^0)(\Delta x) = \sum_{k=1}^n A_k \Delta x_k.$$

Дифференциал является линейной формой порядка n , как функция от Δx , а также неявно зависит от выбора точки x^0 .

Теорема. Необходимое условие дифференцируемости

Пусть функция f дифференцируема в точке $x^0 \in \mathbb{R}^n$ и

$$df(x^0)(\Delta x) = \sum_{k=1}^n A_k \Delta x_k.$$

Тогда у функции f существуют частные производные по всем переменным x_k в точке x^0 и

$$A_k = \frac{\partial f}{\partial x_k}(x^0).$$

Таким образом, для дифференцируемости функции f необходимо наличие у нее всех частных производных в заданной точке.

В одномерном случае существования производной было вполне достаточно для дифференцируемости функции, но в многомерном случае это уже не так.

Упражнения

- Проведите доказательство теоремы о необходимом условии дифференцируемости, аналогично доказательству в одномерном случае.
- Докажите, что дифференцируемая функция непрерывна.
- Покажите, что функция

$$f(x, y) = \frac{xy}{x^2 + y^2}, \quad f(0, 0) = 0,$$

не является дифференцируемой, но имеет частные производные в точке $(0, 0)$.
Покажите, что эта функция не является даже непрерывной.

Дифференцируемость

Вычислим дифференциал функции

$$u(x_1, \dots, x_n) = x_k.$$

Тогда

$$u(x + \Delta x) - u(x) = (x_k + \Delta x_k) - x_k = \Delta x_k.$$

Следовательно, дифференциал имеет вид

$$du(x)(\Delta x) = dx_k(\Delta x) = \Delta x_k.$$

Таким образом, общую формулу для дифференциала функции f можно переписать в виде:

$$df(x^0) = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(x^0) dx_k.$$

Упражнение

Проверьте, что

$$d\left(\sqrt{x^2 + y^2}\right) = \frac{xdx + ydy}{\sqrt{x^2 + y^2}}.$$

Дифференцируемость. Достаточное условие

Теорема. Достаточное условие дифференцируемости

Пусть все частные производные $\frac{\partial f}{\partial x_k}$ функции $f(x)$ определены в окрестности точки x_0 и непрерывны в точке x^0 . Тогда функция f является дифференцируемой в точке x_0 .

Доказательство

Проведем доказательство для $n = 2$. Пусть $f = f(x, y)$. Тогда, применяя формулу конечных приращений, получаем:

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

Из непрерывности частных производных следует, что

$$\frac{\partial f}{\partial y}(x, \eta) = \frac{\partial f}{\partial y}(x_0, y_0) + o(1), \quad \frac{\partial f}{\partial x}(\xi, y_0) = \frac{\partial f}{\partial x}(x_0, y_0) + o(1).$$

Таким образом,

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

Частные производные. Градиент

Определение. Градиент

Если функция $f(x)$, $x \in \mathbb{R}^n$, имеет частные производные по всем переменным (n штук) в фиксированной точке x^0 , то числовой вектор

$$\operatorname{grad} f \Big|_{x=x^0} = \left(\frac{\partial f}{\partial x_1}(x^0), \dots, \frac{\partial f}{\partial x_n}(x^0) \right)$$

называют **градиентом** функции f в точке x^0 . То есть градиент — это вектор, составленный из частных производных функции f в точке x^0 .

Определение. Производные высших порядков

Если частная производная функции $f(x)$ по переменной x_k определена в некоторой окрестности точки $x^0 \in \mathbb{R}^n$, то можно определить **частные производные второго порядка**:

$$\frac{\partial^2 f}{\partial x_k \partial x_m}(x^0) = \frac{\partial}{\partial x_m} \Big|_{x=x^0} \frac{\partial f(x)}{\partial x_k}, \quad k = 1, \dots, n,$$

как частные производные по переменной x_m от функции $\frac{\partial f(x)}{\partial x_k}$. Аналогично определяются производные третьего порядка и далее.

Частные производные. Пример

Пример

Пусть

$$f = x^2 + 2xy.$$

Тогда

$$\frac{\partial f}{\partial x} = 2x + 2y;$$

$$\frac{\partial f}{\partial y} = 2x;$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x} (2x + 2y) = 2; \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial y} (2x + 2y) = 2;$$

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

Таким образом, для функции двух переменных существует четыре частных производных второго порядка. Они образуют матрицу 2×2 . В общем случае размерности n мы получим квадратную матрицу размера $n \times n$.

Предложение

Пусть функция f дифференцируема в точке x . Тогда для производной по направлению l справедлива формула:

$$\frac{\partial f}{\partial l}(x) = \sum_{k=1}^n l_k \cdot \frac{\partial f}{\partial x_k} = \left(l_1 \frac{\partial}{\partial x_1} + \cdots + l_n \frac{\partial}{\partial x_n} \right) f(x) = \langle l, \text{grad } f \rangle.$$

То есть производная по направлению l совпадает с проекцией вектора $\text{grad } f$ на вектор l .

Доказательство

Функция f является дифференцируемой:

$$f(x + \Delta x) - f(x) = \sum_{k=1}^n \frac{\partial f}{\partial x_k} \Delta x_k + o(\|\Delta x\|).$$

Пусть $\Delta x = hl$. Тогда $\|\Delta x\| = |h| \|l\| = |h|$, $\Delta x_k = l_k h$.

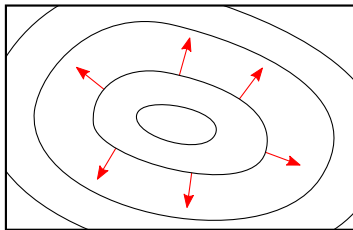
Следовательно,

$$\frac{\partial f}{\partial l} = \lim_{h \rightarrow 0} \frac{f(x + hl) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{1}{h} \left(\sum_{k=1}^n \frac{\partial f}{\partial x_k} l_k h + o(h) \right) = \sum_{k=1}^n \frac{\partial f}{\partial x_k} l_k.$$

Геометрический смысл градиента

Производная по направлению l характеризует скорость роста функции в этом направлении. Мы видим, что производная по направлению совпадает с проекцией направляющего вектора l на вектор градиента $\text{grad } f$. Следовательно, направление, в котором функция растет сильнее всего, совпадает с направлением градиента.

Направление градиента — это направление наискорейшего роста функции, а модуль градиента — это скорость роста функции в этом направлении.



Если рассмотреть линии уровни функции, то есть кривые на которых функция постоянна, то градиент будет ортогонален линиям уровня. **Докажите это** в двумерном случае.

Правила дифференцирования

Поскольку частные производные — это производные от функции многих переменных, как от функции одной переменной при фиксации остальных переменных, то все правила дифференцирования сохраняют силу. Остановимся только на дифференцировании сложной функции.

Правило дифференцирования сложной функции

Пусть функция $f(x) = f(x_1, \dots, x_n)$ дифференцируема в точке $x^0 \in \mathbb{R}^n$, n функций $g_j(t) = g_j(t_1, \dots, t_m)$ дифференцируемы в точке $t^0 \in \mathbb{R}^m$, и $g_j(t^0) = x_j^0$, где $j = 1, \dots, n$. Тогда сложная функция $F(t) = f(g(t))$ дифференцируема в точке t^0 и справедливы формулы

$$\frac{\partial F}{\partial t_k} = \frac{\partial}{\partial t_k} \Big|_{t=t^0} f(g_1(t), \dots, g_n(t)) = \sum_{s=1}^n \frac{\partial f}{\partial x_s}(g_1(t^0), \dots, g_n(t^0)) \frac{\partial g_s}{\partial t_k}(t^0).$$

Пример

Пусть на прямой задана потенциальная энергия $U(x)$. Тогда уравнение Ньютона имеет вид:

$$m \frac{dv}{dt} = - \frac{dU}{dx}.$$

Определим функцию

$$E(x, v) = \frac{mv^2}{2} + U(x).$$

Пусть $x(t)$, $v(t) = \dot{x}(t)$ — решение уравнений Ньютона. Тогда

$$\frac{d}{dt} E(x(t), v(t)) = \frac{\partial E}{\partial x} \frac{dx}{dt} + \frac{\partial E}{\partial v} \frac{dv}{dt} = U'(x(t))v(t) + mv(t) \left(-\frac{1}{m} U'(x(t)) \right) = 0.$$

Таким образом, функция $E(x, v)$ сохраняет свое значение — это закон сохранения энергии.

Теорема Шварца

Теорема о равенстве смешанных производных

Пусть частные производные $\frac{\partial^2 f}{\partial x \partial y}$ и $\frac{\partial^2 f}{\partial y \partial x}$ функция $f(x, y)$ определены в некоторой окрестности точки (x_0, y_0) и **непрерывны** в этой точке. Тогда они совпадают:

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

Теорема естественным образом обобщается на случай функции нескольких переменных и на производные выше второго порядка. Так значение смешанной производной не зависит от порядка дифференцирования, если соответствующие частные производные непрерывны.

Замечание

Условие непрерывности является достаточным, но отнюдь не является необходимым. Смешанные производные могут совпадать и в случае, когда непрерывность не имеет места.

Упражнения

- Докажите теорему.
- Рассмотрите пример
$$f(x, y) = \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2}, & x^2 + y^2 > 0; \\ 0, & x = y = 0. \end{cases}$$

Теорема Шварца

Теорема о равенстве смешанных производных

Пусть частные производные $\frac{\partial^2 f}{\partial x \partial y}$ и $\frac{\partial^2 f}{\partial y \partial x}$ функция $f(x, y)$ определены в некоторой окрестности точки (x_0, y_0) и **непрерывны** в этой точке. Тогда они совпадают:

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

Теорема естественным образом обобщается на случай функции нескольких переменных и на производные выше второго порядка. Так значение смешанной производной не зависит от порядка дифференцирования, если соответствующие частные производные непрерывны.

Замечание

Условие непрерывности является достаточным, но отнюдь не является необходимым. Смешанные производные могут совпадать и в случае, когда непрерывность не имеет места.

Упражнения

- Докажите теорему.
- Рассмотрите пример
$$f(x, y) = \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2}, & x^2 + y^2 > 0; \\ 0, & x = y = 0. \end{cases}$$

Старшие дифференциалы. Формула Тейлора

Для дифференцируемой функции $f(x)$, $x \in \mathbb{R}^n$, справедлива формула линеаризации:

$$f(x + y) = f(x) + \langle \text{grad } f(x), y \rangle + o(\|y\|).$$

Часто возникает необходимость приближения более высокой точности.

Теорема. Формула Тейлора

Пусть все частные производные функция f до порядка m включительно определены в некоторой окрестности точки $x \in \mathbb{R}^n$ и непрерывны в этой точке. Тогда справедлива формула Тейлора:

$$\begin{aligned} f(x + y) = f(x) + \sum_{k=1}^n \frac{\partial f(x)}{\partial x_k} y_k + \frac{1}{2} \sum_{k,r=1}^n \frac{\partial^2 f(x)}{\partial x_k \partial x_r} y_k y_r + \dots \\ + \frac{1}{m!} \sum_{k_1, \dots, k_m=1}^n \frac{\partial^m f(x)}{\partial x_{k_1} \dots \partial x_{k_m}} y_{k_1} \dots y_{k_m} + o(\|y\|^m). \end{aligned}$$

Старшие дифференциалы. Формула Тейлора

Пример

Для функции двух переменных получаем:

$$f(x + \Delta x, y + \Delta y) = f(x, y) + \frac{\partial f}{\partial x} \Delta x + \frac{\partial f}{\partial y} \Delta y + \\ \frac{1}{2} \left(\frac{\partial^2 f}{\partial x^2} \Delta x^2 + 2 \frac{\partial^2 f}{\partial x \partial y} \Delta x \Delta y + \frac{\partial^2 f}{\partial y^2} \Delta y^2 \right) + \dots,$$

где производные взяты в точке (x, y) .

Второй дифференциал функции $f(x, y)$ определяется как:

$$d^2 f(x, y) = d(df(x, y)),$$

Следовательно,

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

Аналогично определяются дифференциалы старших порядков.

Мы видим, что формулу Тейлора можно записать в виде:

$$f(x + \Delta x, y + \Delta y) = f(x, y) + df(x, y)(\Delta x, \Delta y) + \frac{1}{2} d^2 f(x, y)(\Delta x, \Delta y) + \dots$$

Локальные экстремумы

Определение. Локальный экстремум

Точка $x^0 \in \mathbb{R}^n$ называется **точкой локального максимума** (или минимума) функции $f(x)$, определенной в некоторой окрестности этой точки, если существует окрестность U точки x^0 такая, что для всех $x \in U$, $x \neq x^0$ справедливо неравенство:

$$f(x) > f(x^0), \quad \text{или} \quad f(x) < f(x^0).$$

Точки локальных минимумов и максимумов называют **точками локального экстремума** функции.

Пусть функция рассматривается на множестве $M \subset \mathbb{R}^n$, для которого точка $x^0 \in M$ не является внутренней, а например, является точкой границы. Тогда данное выше определение дополняют условием $x \in M$.

Определение. Условный локальный экстремум

Точка $x^0 \in M$ называется **точкой условного локального максимума** (или минимума) функции $f(x)$ при условии $x \in M$, если существует окрестность U точки x^0 такая, что для всех $x \in U \cap M$, $x \neq x^0$ справедливо неравенство:

$$f(x) > f(x^0), \quad \text{или} \quad f(x) < f(x^0).$$

Иногда говорят о внутренних и граничных точках локального экстремума.

Определение. Локальный экстремум

Точка $x^0 \in \mathbb{R}^n$ называется **точкой локального максимума** (или минимума) функции $f(x)$, определенной в некоторой окрестности этой точки, если существует окрестность U точки x^0 такая, что для всех $x \in U$, $x \neq x^0$ справедливо неравенство:

$$f(x) > f(x^0), \quad \text{или} \quad f(x) < f(x^0).$$

Точки локальных минимумов и максимумов называют **точками локального экстремума** функции.

Пусть функция рассматривается на множестве $M \subset \mathbb{R}^n$, для которого точка $x^0 \in M$ не является внутренней, а например, является точкой границы. Тогда данное выше определение дополняют условием $x \in M$.

Определение. Условный локальный экстремум

Точка $x^0 \in M$ называется **точкой условного локального максимума** (или минимума) функции $f(x)$ при условии $x \in M$, если существует окрестность U точки x^0 такая, что для всех $x \in U \cap M$, $x \neq x^0$ справедливо неравенство:

$$f(x) > f(x^0), \quad \text{или} \quad f(x) < f(x^0).$$

Иногда говорят о внутренних и граничных точках локального экстремума.

Достаточное условие экстремума

В многомерном случае справедлив аналог теоремы Фурье:

Теорема. Достаточные условия локального экстремума

Пусть функция $f(x)$ дифференцируема в точке $x^0 \in \mathbb{R}^n$ и имеет в этой точке локальный экстремум. Тогда в этой точке градиент функции равен нулю:

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

Таким образом, во всех точках локального экстремума равны нулю все частные производные. Такие точки принято называть **критическими**.

Доказательство

Докажем от противного. Пусть в точке x^0 имеет место локальный экстремум, а градиент не обращается в ноль. Можно считать, что $\frac{\partial f}{\partial x_1}(x^0) \neq 0$. Тогда функция $g(t) = f(x_1^0 + t, x_2^0, \dots, x_n^0)$ одной переменной t , очевидно, дифференцируема и имеет экстремум в точке $t = 0$, но

$$g'(0) = \frac{d}{dt} f(x_1^0 + t, x_2^0, \dots, x_n^0) = \frac{\partial f}{\partial x_1}(x^0) \neq 0,$$

что противоречит теореме Ферма.

Необходимое условие

Как и в одномерном случае, для исследования критической точки можно рассмотреть второй дифференциал функции.

Теорема. Необходимые условия локального экстремума

Пусть функция $f(x)$ дважды дифференцируема в точке $x^0 \in \mathbb{R}^n$, точка x^0 является критической $\text{grad } f(x^0) = 0$ и второй дифференциал функции f в этой точке является положительно (отрицательно) определенной квадратичной формой. Тогда эта точка является минимумом (максимумом).

Двумерный случай

Для функции двух переменных условие максимума примет вид:

$$d^2f = \frac{\partial^2 f}{\partial x^2} dx^2 + 2 \frac{\partial^2 f}{\partial x \partial y} dx dy + \frac{\partial^2 f}{\partial y^2} dy^2 < 0,$$

для произвольных приращений dx и dy , кроме $dx = dy = 0$.

Вопрос о знаке квадратичной формы детально изучается в линейной алгебре.

Достаточное условие экстремума

Пример

Рассмотрим функцию $f(x) = x^2 + x + 2y^2 + xy$. Задача стоит в поиске локальных и глобальных точек минимума и максимума.

Функция, очевидно, является дифференцируемой. Критические точки определяются из системы уравнений:

$$\begin{cases} \frac{\partial f}{\partial x} = 0; \\ \frac{\partial f}{\partial y} = 0, \end{cases} \iff \begin{cases} 2x + 1 + y = 0; \\ 4y + x = 0, \end{cases} \iff \begin{cases} x = -4/7; \\ y = 1/7. \end{cases}$$

Вычислим d^2f в этой точке:

$$d^2f = 2dx^2 + 2dxdy + 4dy^2 = dx^2 + (dx + dy)^2 + 3dy^2 > 0,$$

при dx и dy неравных нулю одновременно. Следовательно, рассматриваемая точка является точкой минимума.

Критерий Сильвестра

В алгебре доказывается теорема, позволяющая легко проверять знак второго дифференциала.

Теорема. Критерий Сильвестра

Пусть F — квадратичная форма порядка n :

$$F(y) = \langle Ay, y \rangle = \sum_{i,j=1}^n a_{ij} y_i y_j,$$

где A — симметричная матрица $a_{ij} = a_{ji}$. Тогда F положительно определена тогда и только тогда, когда все угловые миноры A_k , $k = 1, \dots, n$ положительны, и F отрицательно определена тогда и только тогда, когда знаки A_k чередуются так, что $A_1 < 0$.

Угловые миноры A_k матрицы A — это определители:

$$A_1 = a_{11}, \quad A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \dots \quad A_n = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$