Механико-математический факультет МГУ имени М. В. Ломоносова

Примерный конспект курса математического анализа (второй семестр)

А. В. Бегунц, С. В. Шапошников

2020 год

Курс математического анализа, в первую очередь, призван познакомить с самыми базовыми и необходимыми математическими идеями, фактами и методами, составляющими важную часть образования профессионала практически любой математической и естественнонаучной специализации. Другой важной целью курса анализа, особенно на физико-математических факультетах и направлениях, является знакомство на простейших примерах практически со всеми разделами современной математики от алгебры и геометрии до функционального анализа и теории оптимального управления. Поэтому, по нашему мнению, в курсе математического анализа должны обязательно присутствовать, пусть и в небольшом объеме, элементы других математических дисциплин, иллюстрируя дальнейшее развитие изучаемых идей и методов. Кроме того, в отличие от многих специальных курсов высокого уровня, математический анализ является в определённом смысле синтетическим курсом, сочетающим в себе элементы большого числа математических дисциплин, и тем самым даёт возможность взглянуть на математику в целом как на единую науку.

Настоящий курс фактически состоит из двух достаточно независимых разделов, один из которых посвящён функциям нескольких переменных, а другой — интегралу Римана. Отметим некоторые особенности данного курса. При обсуждении метрических и нормированных пространств на примере конечномерного евклидова пространства доказывается теорема об эквивалентности норм. С одной стороны, это удивительное и впечатляющее утверждение для студентов первых курсов, а с другой стороны, это фундаментальное утверждение и отличная иллюстрация теоремы Больцано и свойств предела последовательности в многомерном пространстве. Кроме евклидова конечномерного пространства обсуждается пространство ограниченных функций и формируется взгляд на равномерную сходимость последовательностей функций как на сходимость последовательностей элементов метрического пространства. Доказывается, что поточечная сходимость в общем случае не задаётся метрикой. Это простое, но примечательное утверждение позволяет осознать, что поточечная сходимость сложнее равномерной. При обсуждении компактных множеств в метрическом и нормированном пространстве доказывается теорема о взаимосвязи компактности шара и конечномерности пространства. Конечно, и это утверждение, и утверждение об эквивалентности норм входят в стандартный курс функционального анализа, но там эти факты несколько теряются на фоне других более трудных теорем, в то время как в курсе математического анализа они, не усложняя его, являются отличной иллюстрацией изучаемых понятий. Отметим, что при обсуждении компактности мы не обсуждаем эпсилон-сети и вполне ограниченные множества, поскольку этот материал можно осознать только решая задачи на построение эпсилон-сетей, что лежит совсем в стороне от основных тем как лекций, так и семинарских занятий. Значимым отличием этого курса является изложение дифференциального исчисления для отображений нормированных пространств. Это связано не с целью максимального обобщения, а для упрощения и наглядности, так как безкоординатная запись определения дифференцируемого отображения практически не отличается от аналогичного определения в одномерном случае и гораздо легче воспринимается. Конечно, требуется некоторая работа по переписыванию общего определения на случай функций нескольких переменных, но и это также кажется весьма полезным. Теорема о неявной функции обсуждается после теоремы об обратном отображении, что позволяет говорить о заменах координат и использовать при изложении наглядные геометрические образы. При обсуждении интеграла Римана мы сразу после определения приводим формулу Ньютона — Лейбница и всячески подчёркиваем роль интеграла в задаче о восстановлении первообразной. Это кажется более естественным и простым, чем апеллировать к трудному понятию площади под графиком. Фактически изложение теории интеграла делится на две части, первая из которых посвящена основным свойствам интеграла и заканчивается на теореме об интегрируемости функции, являющейся равномерным пределом интегрируемых, и доказательством существования первообразной у непрерывной функции, а вторая часть посвящена критерию Дарбу, критерию Лебега и свойствам интеграла с переменным верхним пределом от интегрируемой функции. Отметим, что мы очень кратко (только на уровне определения) обсуждаем несобственный интеграл, поскольку подробное его изучение на семинарах во втором семестре не представляется возможным. Это же относится к интегралу Стилтьеса, который естественным образом возникает и обсуждается в курсе теории меры и интеграла Лебега.

Авторы искренне благодарят своих учителей, раскрывших красоту математического анализа: профессоров В. И. Богачева, В. А. Зорича, Т. П. Лукашенко, В. Н. Чубарикова, Т. А. Шапошникову, а также коллег, в разное время оказавших поддержку при работе над курсом и настоящим конспектом: В. В. Галатенко, Д. В. Горяшина, Е. Д. Косова, С. М. Лыткина, Ю. В. Межевову, А. П. Солодова, А. А. Флёрова, О. Ю. Черкасова, А. Г. Якушева. Наконец, благодарим всех студентов, предоставивших записи лекций, а также тех, кто на протяжении ряда лет своим неподдельным интересом и вдумчивыми вопросами стимулировал совершенствование курса и побудил авторов к написанию настоящего пособия.

Евклидово пространство $\mathbb{R}^n$

Множество $\mathbb{R}^n$ является декартовым произведением n множеств $\mathbb{R}$, т. е. состоит из упорядоченных наборов $(x_1, x_2, \ldots, x_n)$ вещественных чисел x_i. Эти наборы мы будем называть векторами и обозначать латинскими буквами без стрелок, так как отличие векторных величин от скалярных будет ясно из контекста. Величины, обозначаемые греческими буквами α и β (возможно, с индексами), будут всегда скалярными. Кроме того, несмотря на используемую запись набора в строчку, мы будем иметь в виду, что вектор есть вектор-столбец, и исходить из этого, в частности, при выполнении операции умножения на матрицу.

На $\mathbb{R}^n$ можно ввести операцию сложения элементов и операцию умножения элемента на число:

- $(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$,
- $\alpha(x_1, x_2, \ldots, x_n) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n), \quad \alpha \in \mathbb{R}$.

Можно показать, что $\mathbb{R}^n$ является линейным пространством над полем $\mathbb{R}$. Векторы

$$e_1 = (1, 0, \ldots, 0), \quad e_2 = (0, 1, 0, \ldots, 0), \quad \ldots, \quad e_n = (0, 0, \ldots, 1)$$

образуют базис, который мы будем называть *стандартным*. Всякий вектор $x \in \mathbb{R}^n$ единственным образом представляется в виде

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n.$$

Рассмотрим функцию $\langle \cdot\,, \cdot \rangle \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, заданную формулой

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n, \quad \text{где} \quad x = (x_1, x_2, \ldots, x_n),\ y = (y_1, y_2, \ldots, y_n).$$

Эта функция удовлетворяет свойствам:

(i) $\langle x, x \rangle \geqslant 0$ и $\langle x, x \rangle = 0 \Leftrightarrow x = 0$ (положительная определённость);

(ii) $\langle x, y \rangle = \langle y, x \rangle$ (симметричность);

(iii) $\langle \alpha_1 x_1 + \alpha_2 x_2, y \rangle = \alpha_1 \langle x_1, y \rangle + \alpha_2 \langle x_2, y \rangle$ (линейность).

Всякая функция $\langle \cdot\,, \cdot \rangle$ на линейном пространстве над полем $\mathbb{R}$, удовлетворяющая этим трём свойствам, называется *скалярным произведением*. Линейное пространство со скалярным произведением называется *евклидовым* пространством. Таким образом, $\mathbb{R}^n$ — евклидово пространство.

Теорема 1 (неравенство Коши — Буняковского — Шварца)**.** *Для всех $x, y \in \mathbb{R}^n$ выполняется неравенство*
$$|\langle x, y\rangle| \leqslant \sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle},$$
причём равенство достигается тогда и только тогда, когда векторы x и y линейно зависимы.

$\triangleright$ Заметим, что при $y = 0$ утверждение выполнено. Пусть теперь $y \neq 0$. Рассмотрим квадратичную функцию
$$f(t) = \langle x - ty, x - ty\rangle = \langle x, x\rangle - 2t\langle x, y\rangle + t^2\langle y, y\rangle.$$
Поскольку $f(t) \geqslant 0$ для всех t, получаем
$$D = 4|\langle x, y\rangle|^2 - 4\langle x, x\rangle\langle y, y\rangle \leqslant 0.$$
Итак, требуемое неравенство доказано. Остаётся заметить, что $D = 0$ тогда и только тогда, когда $x - ty = 0$ для некоторого t. $\triangleleft$

Величина $\|x\| = \sqrt{\langle x, x\rangle}$ называется *евклидовой нормой* или *длиной* вектора x. Отметим, что функция $\|\cdot\|\colon \mathbb{R}^n \to [0; +\infty)$ удовлетворяет следующим свойствам:

(i) $\|x\| = 0 \Leftrightarrow x = 0$;

(ii) $\|\alpha x\| = |\alpha|\|x\|$;

(iii) $\|x + y\| \leqslant \|x\| + \|y\|$.

Неравенство из последнего свойства называют *неравенством треугольника.* Покажем, что оно действительно выполняется:
$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \leqslant \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$
Пусть x и y отличны от нуля. Число $\varphi \in [0; \pi]$, для которого верно равенство
$$\cos\varphi = \frac{\langle x, y\rangle}{\|x\|\|y\|}$$
называется *углом* между векторами x и y. В силу доказанного выше неравенства Коши — Буняковского — Шварца правая часть по модулю не превосходит единицы, поэтому такое число φ всегда существует. Если $\langle x, y\rangle = 0$, то векторы x и y называются *ортогональными.*

Метрические и нормированные пространства

Пусть X — произвольное непустое множество. Функция $\rho(\,\cdot\,, \cdot\,)\colon X \times X \to [0; +\infty)$, которая обладает следующими свойствами:

(i) $\rho(x, y) = 0 \Leftrightarrow x = y$;

(ii) $\rho(x, y) = \rho(y, x)$ (симметричность);

(iii) $\rho(x, y) \leqslant \rho(x, z) + \rho(z, y)$ (неравенство треугольника),

называется *метрикой*, а пара (X, ρ) называется *метрическим пространством.*

Рассмотрим несколько примеров.

1) X — произвольное непустое множество и $\rho(x, y) = 0$ при $x = y$ и $\rho(x, y) = 1$ при $x \neq y$. Такую метрику называют *дискретной.*

2) $X = \mathbb{R}$ и $\rho(x, y) = |x - y|$. Это обычное, знакомое из школьного курса расстояние между точками на числовой прямой.

3) $X = \mathbb{R}$ и $\rho(x, y) = |f(x) - f(y)|$, где f — произвольное инъективное отображение из $\mathbb{R}$ в $\mathbb{R}$. На первый взгляд, эта метрика не очень сильно отличается от $|x - y|$, но метрическое пространство может получиться весьма экзотическим. Если $f(x) = \mathrm{arctg} x$, то все точки лежат на расстоянии не больше π друг от друга. Если $f(x) = 1/x$ при $x \neq 0$ и $f(0) = 0$, то расстояние от нуля до $1/n$ равно n и стремится к бесконечности.

4) $X = \mathbb{R}^n$ и $\rho(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}$. Легко видеть, что $\rho(x, y) = \|x - y\|$, где $\|\cdot\|$ — евклидова норма. Такую метрику называют *евклидовой.*

5) $X = \mathbb{R}^n$ и $\rho(x, y) = |x_1 - y_1| + \ldots + |x_n - y_n|$. Это так называемая *манхэттенская метрика.* Пусть перед нами лист бумаги в клетку и точке разрешается двигаться только по линиям сетки — границам клеток. Предположим, что сторона клетки равна единице и начало координат находится в вершине одной из клеток. Тогда длина кратчайшего пути от точки $(0, 0)$ до точки (a, b) равна

$|a|+|b|$. Если рассмотреть эту метрику на булевом кубе $\{0,1\}^n$, то получим расстояние Хемминга, которое играет важную роль в теории кодирования.

Следующее утверждение следует из неравенства треугольника.

Предложение 1. *Пусть (X,ρ) — метрическое пространство. Тогда для всяких $x,y,z \in X$ верно неравенство*

$$|\rho(x,y)-\rho(x,z)| \leqslant \rho(y,z).$$

Множество $B_r(a)=\{x\in X\colon \rho(a,x)<r\}$ называется *открытым шаром* радиусом r, а множество $\overline{B}_r(a)=\{x\in X\colon \rho(a,x)\leqslant r\}$ называется *замкнутым шаром* радиусом r.

Отметим, что в метрическом пространстве шар большего радиуса может содержаться строго внутри шара меньшего радиуса. Действительно, пусть $X=\{1,3,6\}$ и $\rho(x,y)=|x-y|$. Открытый шар с центром в 3 и радиусом 4 содержит все X, а открытый шар с центром в 1 и радиусом 5 содержит только точки 1 и 3.

Говорят, что последовательность $\{x_m\}$ точек метрического пространства (X,ρ) сходится к точке x, если $\rho(x_m,x)\to 0$ при $m\to\infty$. В этом случае будем писать $\lim\limits_{m\to\infty} x_m = x$ или $x_m\to x$ при $m\to\infty$ (далее при использовании записи $x_m\to x$ будем подразумевать, что $m\to\infty$).

Предложение 2. *Если $x_m\to x$ и $x_m\to y$, то $x=y$.*

$\rhd$ По неравенству треугольника

$$\rho(x,y)\leqslant \rho(x,x_m)+\rho(x_m,y)\to 0$$

и, следовательно, $\rho(x,y)=0$ и $x=y$. $\lhd$

Предложение 3. *Если последовательность сходится, то она ограниченна, т. е. лежит целиком в некотором шаре.*

$\rhd$ Пусть $x_m\to x$. Тогда найдётся такое N, что для всех $m>N$ выполняется неравенство $\rho(x,x_m)<1$. Пусть $R>\max\{\rho(x_1,x),\dots,\rho(x_N,x),1\}$. Тогда $x_m\in B_R(x)$ для всякого $m\in\mathbb{N}$. $\lhd$

Лемма 1. *Для всякого $x\in\mathbb{R}^n$ имеет место неравенство*

$$\max_{1\leqslant i\leqslant n}|x_i|\leqslant \sqrt{\sum_{i=1}^n x_i^2}\leqslant \sqrt{n}\max_{1\leqslant i\leqslant n}|x_i|.$$

Из леммы следует эквивалентность сходимости по евклидовой метрике и покоординатной сходимости, а именно, имеет место следующее утверждение.

Теорема 2. *Пусть $x^m=(x_1^m,\dots,x_n^m)\in\mathbb{R}^n$. Последовательность x^m сходится к $x=(x_1,\dots,x_n)$ в пространстве $\mathbb{R}^n$ с евклидовой метрикой тогда и только тогда, когда $x_i^m\to x_i$ для каждого $i=1,2,\dots,n$.*

Предложение 4 (Больцано). *Если последовательность $x^m\in\mathbb{R}^n$ ограниченна в евклидовой метрике, то из неё можно выбрать сходящуюся подпоследовательность.*

$\rhd$ Рассмотрим только случай $n=2$, так как общий случай совершенно аналогичен, но записывается более громоздко. Числовые последовательности x_1^m и x_2^m являются ограниченными, и к ним применима теорема Больцано для числовых последовательностей. Сначала выберем сходящуюся подпоследовательность $x_1^{m_k}$ из последовательности первых координат. Далее из последовательности $x_2^{m_k}$ вторых координат векторов с номерами m_k выберем сходящуюся подпоследовательность $x_2^{m_{k_j}}$. Подпоследовательность векторов $x^{m_{k_j}}$ сходится, так как сходятся последовательности первых и вторых координат. $\lhd$

Последовательность $\{x_m\}$ элементов метрического пространства (X,ρ) называется *фундаментальной*, если для всякого $\varepsilon>0$ найдётся такое N, что для всех $n,m>N$ выполняется неравенство $\rho(x_n,x_m)<\varepsilon$.

Метрическое пространство (X,ρ) называется *полным*, если всякая фундаментальная последовательность его элементов сходится к элементу этого пространства. Если из числовой прямой с обычным расстоянием выбросить точку, то получится неполное пространство. Другой пример неполного пространства можно получить, если на числовой прямой вместо обычной метрики задать расстояние формулой $\rho(x,y)=|\mathrm{arctg} x-\mathrm{arctg} y|$. Действительно, последовательность $x_n=n$ фундаментальна по этой метрике, но не имеет предела.

Теорема 3. *Пространство $\mathbb{R}^n$ с евклидовой метрикой является полным.*

$\triangleright$ Если последовательность x^m фундаментальна, то по лемме 1 каждая из числовых последовательностей x_i^m фундаментальна. По критерию Коши для числовых последовательностей последовательность x_i^m сходится для каждого i. Следовательно, по доказанному выше, последовательность x^m сходится. $\triangleleft$

Далее, если не сказано противного, будем по умолчанию считать, что пространство $\mathbb{R}^n$ наделено евклидовой метрикой. Аналогично тому, как на числовой прямой удобным выражением полноты является принцип вложенных отрезков, так и в полном пространстве имеет место *принцип вложенных шаров*.

Теорема 4. *В полном метрическом пространстве всякая последовательность вложенных замкнутых шаров, радиусы которых стремятся к нулю, имеет общую точку.*

$\triangleright$ Достаточно заметить, что центры этих шаров образуют фундаментальную последовательность, которая в силу полноты сходится. В силу замкнутости и вложенности шаров предел этой последовательности является общей точкой. $\triangleleft$

Полезно иметь в виду, что в неполном метрическом пространстве всегда есть последовательность вложенных замкнутых шаров, радиусы которых стремятся к нулю, но пересечение которых пусто. Таким образом, принцип вложенных шаров равносилен полноте метрического пространства.

Нормированные пространства

Пусть X — линейное пространство над $\mathbb{R}$. Функция $\|\cdot\|\colon X \to [0;+\infty)$, которая обладает свойствами:

(i) $\|x\| = 0 \Leftrightarrow x = 0$;

(ii) $\|\alpha x\| = |\alpha|\|x\|$;

(iii) $\|x+y\| \leqslant \|x\| + \|y\|$ (неравенство треугольника),

называется *нормой*, а пара $(X, \|\cdot\|)$ называется *нормированным пространством.*

Важным примером нормы является евклидова норма на $\mathbb{R}^n$: $\|x\| = \sqrt{\sum\limits_{k=1}^{n} x_k^2}$. Всякое нормированное пространство является метрическим с метрикой $\rho(x,y) = \|x-y\|$. Надо иметь в виду, что не всякая метрика задаётся какой-то нормой. В отличие от метрических пространств, в нормированных шар большого радиуса не может содержаться в шаре маленького радиуса.

Предложение 5. *Если $B_r(a) \subset B_R(b)$, то $r + \|a-b\| \leqslant R$.*

$\triangleright$ Случай $a = b$ очевиден. Пусть теперь $a \neq b$. Если $0 \leqslant t < r\|a-b\|^{-1}$, то $a + t(a-b) \in B_r(a)$, а это, в свою очередь, влечёт $a + t(a-b) \in B_R(b)$. Таким образом, для всех $0 \leqslant t < r\|a-b\|^{-1}$ верно неравенство $(1+t)\|a-b\| \leqslant R$. Устремляя t к $r\|a-b\|^{-1}$, получаем требуемое неравенство. $\triangleleft$

Предложение 6. *Пусть $(X, \|\cdot\|)$ — нормированное пространство. Пусть, далее, $x_n, y_n \in X$ сходятся соответственно к x и y, числовая последовательность α_n сходится к α. Тогда*

$$x_n + y_n \to x + y, \quad \alpha_n x_n \to \alpha x.$$

$\triangleright$ Достаточно заметить, что по неравенству треугольника

$$\|(x_n + y_n) - (x+y)\| \leqslant \|x_n - x\| + \|y_n - y\| \to 0,$$
$$\|\alpha_n x_n - \alpha x\| \leqslant |\alpha_n|\|x_n - x\| + |\alpha_n - \alpha|\|x\| \to 0. \qquad \triangleleft$$

Предложение 7. *Пусть $(X, \|\cdot\|)$ — нормированное пространство. Тогда верно неравенство $|\|x\| - \|y\|| \leqslant \|x-y\|$. В частности, если $x_n \to x$, то $\|x_n\| \to \|x\|$.*

$\triangleright$ Требуемое неравенство является следствием неравенства треугольника. $\triangleleft$

На $\mathbb{R}^n$ кроме евклидовой нормы можно ввести и другие нормы: например, $\|x\| = \max_{1\leqslant k\leqslant n} |x_k|$ или $\|x\| = \sum\limits_{1\leqslant k\leqslant n} |x_k|$. Однако, оказывается, что все нормы на $\mathbb{R}^n$ эквивалентны между собой, а именно, справедливо следующее утверждение.

Теорема 5. *На пространстве $\mathbb{R}^n$ все нормы эквивалентны, т. е. для любых двух норм $\|\cdot\|_a$ и $\|\cdot\|_b$ существуют такие положительные числа m и M, что для всех x верно неравенство*

$$m\|x\|_a \leqslant \|x\|_b \leqslant M\|x\|_a.$$

$\triangleright$ Достаточно рассмотреть случай, когда $\|\cdot\|_b$ — евклидова норма. Пусть задана произвольная норма $\|\cdot\|_a$ на $\mathbb{R}^n$. Заметим, что

$$\|x\|_a = \|x_1e_1 + \ldots + x_ne_n\|_a \leqslant (|x_1| + \ldots + |x_n|)\max_k \|e_k\|_a \leqslant C_a\|x\|_b, \text{ где } C_a = n\max_k \|e_k\|_a.$$

Таким образом, остаётся лишь установить неравенство $\|x\|_b \leqslant C_b\|x\|_a$ для всех x и некоторого числа $C_b > 0$. Предположим, что это неравенство не выполняется. Тогда для всякого $N \in \mathbb{N}$ найдётся такой отличный от нуля вектор x^N, что $\|x^N\|_b \geqslant N\|x^N\|_a$. Разделив это неравенство на $\|x^N\|_b$ и на N, получаем

$$1/N > \|y^N\|_a, \quad y^N = x^N/\|x^N\|_b, \quad \|y^N\|_b = 1.$$

Итак, по норме $\|\cdot\|_a$ последовательность y^N сходится к нулю, а по евклидовой норме все векторы y^N имеют единичную длину. По теореме Больцано можно выбрать сходящуюся по евклидовой норме подпоследовательность y^{N_j}. Пусть $y^{N_j} \to y$. Тогда $\|y\|_b = 1$ и, следовательно, $y \neq 0$. В силу оценки $\|x\|_a \leqslant C_a\|x\|_b$ сходимость по евклидовой норме влечёт сходимость по норме $\|\cdot\|_a$, по которой данная последовательность сходится к нулю. Противоречие. Таким образом, справедлива требуемая оценка, где $m = 1/C_a$, $M = C_b$. $\triangleleft$

Пусть $(X, \|\cdot\|)$ — нормированное пространство. Формальная бесконечная сумма $\sum\limits_{n=1}^{\infty} x_n$ элементов $x_n \in X$ называется *рядом.* По определению ряд *сходится*, если сходится последовательность его частичных сумм $S_N = \sum\limits_{n=1}^{N} x_n$.

Предложение 8. *Пусть $(X, \|\cdot\|)$ — полное нормированное пространство. Если сходится числовой ряд $\sum\limits_{n=1}^{\infty} \|x_n\|$, то сходится ряд $\sum\limits_{n=1}^{\infty} x_n$.*

$\triangleright$ Достаточно показать, что последовательность частичных сумм фундаментальна. По неравенству треугольника

$$\|S_N - S_M\| \leqslant \sum_{n=M+1}^{N} \|x_n\|.$$

Ряд из $\|x_n\|$ сходится, поэтому для всякого $\varepsilon > 0$ найдётся такое N_0, что для всех $M, N > N_0$ выполняется неравенство $\sum\limits_{n=M+1}^{N} \|x_n\| < \varepsilon$, что и означает фундаментальность последовательности S_n. $\triangleleft$

Полное нормированное пространство называют *банаховым.*

В заключение обратим внимание на двойственную природу пространства $\mathbb{R}^n$: его можно рассматривать и как векторное, и как аффинное. Выше в определении $\mathbb{R}^n$ как множества упорядоченных наборов $(x_1, x_2, \ldots, x_n)$ вещественных чисел x_i мы условились называть эти наборы векторами (так удобно полагать, например, тогда, когда мы используем возможность сложения элементов и умножения на константу). С другой стороны, в ряде случаев более естественно считать набор $(x_1, x_2, \ldots, x_n)$ не вектором, а точкой, которую можно представлять себе как конец вектора $(x_1, x_2, \ldots, x_n)$, приложенного к началу координат (например, так поступают при изучении топологических свойств пространства). В дальнейшем мы будем пользоваться обоими взглядами на пространство $\mathbb{R}^n$, не оговаривая это особо.

Пространство ограниченных функций и равномерная сходимость

В предыдущих разделах основным примером метрического и нормированного пространства было евклидово пространство $\mathbb{R}^n$. В этом разделе мы познакомимся ещё с одним важным и нетривиальным примером — пространством ограниченных функций.

Пусть X — непустое множество. Через $B(X)$ обозначим линейное пространство ограниченных функций $f\colon X \to \mathbb{R}$. Выше мы видели, что метрика позволяет говорить о сходимости последовательностей. На множестве функций есть два очень важных и часто встречающихся в приложениях вида сходимости: 1) поточечная сходимость, когда $f_n(x) \to f(x)$ для всякого $x \in X$, и 2) равномерная сходимость, когда $\sup\limits_X |f_n(x) - f(x)| \to 0$. Поточечная сходимость с точки зрения определения кажется более простой и понятной, но это обманчивое впечатление. Это гораздо более сложный вид сходимости, чем сходимость равномерная, что иллюстрируется, в частности, следующей теоремой.

Теорема 6. *Не существует такой метрики ρ на $B[0;1]$, что $f_n \to f$ поточечно тогда и только тогда, когда $\rho(f_n, f) \to 0$.*

$\triangleright$ Предположим противное, т. е. пусть такая метрика ρ существует. Выберем произвольное $r > 0$. Во всяком интервале $(\alpha;\beta) \subset [0;1]$ существует такой отрезок Δ, что индикатор I_Δ принадлежит шару с центром в нуле и радиусом r по метрике ρ. Действительно, возьмём последовательность попарно непересекающихся отрезков Δ_n в интервале $(\alpha;\beta)$. Последовательность индикаторов I_{Δ_n} поточечно сходится к нулю. Следовательно, найдётся n, при котором $\rho(0, I_{\Delta_n}) < r$. Используя это наблюдение, можно построить такую систему вложенных отрезков Δ_n, что $\rho(0, I_{\Delta_n}) < 1/n$. С одной стороны, I_{Δ_n} стремятся к нулю поточечно, а с другой стороны, в общей точке c построенных отрезков имеем $I_{\Delta_n}(c) = 1$ для всех n. Противоречие. $\triangleleft$

Равномерная сходимость, напротив, задаётся не просто метрикой, а даже нормой. Пространство $B(X)$ является нормированным с нормой

$$\|f\| = \sup_{x\in X} |f(x)|.$$

Последовательность f_n сходится к f в $B(X)$ тогда и только тогда, когда $\|f_n - f\| \to 0$, т. е.

$$\sup_{x\in X} |f_n(x) - f(x)| \to 0.$$

Теорема 7. *Пространство $B(X)$ полное.*

$\triangleright$ Пусть f_n — фундаментальная последовательность. Тогда при каждом $x \in X$ числовая последовательность $f_n(x)$ является фундаментальной и, следовательно, сходится к некоторому числу $f(x)$. Покажем, что $f \in B(X)$ и $\|f_n - f\| \to 0$. Опять воспользуемся тем, что последовательность f_n фундаментальна. Для всякого $\varepsilon > 0$ найдётся такое N, что для всех $n, m > N$ и всех $x \in X$ выполняется неравенство $|f_n(x) - f_m(x)| < \varepsilon$. Устремим $m \to \infty$. Получаем $|f_n(x) - f(x)| \leqslant \varepsilon$ для всех $x \in X$, что влечёт ограниченность f и сходимость f_n к f в $B(X)$. $\triangleleft$

Таким образом, при исследовании на равномерную сходимость можно пользоваться критерием Коши.

Следствие 1 (признак Вейерштрасса). *Если $f_n \in B(X)$, $\sup\limits_{x\in X} |f_n(x)| \leqslant a_n$ и ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится, то ряд $\sum_{n=1}^{\infty} f_n(x)$ сходится в $B(X)$, т. е. сходится равномерно на множестве X.*

$\triangleright$ Поскольку $B(X)$ — полное пространство, из сходимости ряда $\sum\limits_n \|f_n\|$ следует сходимость ряда $\sum\limits_n f_n$. $\triangleleft$

Итак, признак Вейерштрасса — в точности аналог утверждения про числовые ряды: если ряд сходится абсолютно, то он сходится.

В пространстве ограниченных функций на отрезке выделяют важное подпространство непрерывных функций.

Следствие 2. *Множество $C[a;b]$ непрерывных на отрезке $[a;b]$ функций является полным нормированным пространством с нормой $\|f\| = \max\limits_{x\in[a;b]} |f(x)|$.*

Другими словами, равномерный предел последовательности непрерывных на отрезке функций — также непрерывная на этом отрезке функция.

$\triangleright$ Поскольку $C[a;b] \subset B[a;b]$, фундаментальная последовательность f_n сходится по данной норме к некоторой функции $f \in B[a;b]$ и надо лишь показать, что эта функция непрерывна. Пусть $c \in [a;b]$. Имеет место неравенство

$$|f(x) - f(c)| \leqslant |f(x) - f_n(x)| + |f_n(x) - f_n(c)| + |f_n(c) - f(c)|.$$

Для всякого $\varepsilon > 0$ выберем n столь большим, чтобы первое и третье слагаемое в правой части стали меньше ε. Для такого n найдём $\delta > 0$, при котором из неравенства $|x - c| < \delta$ следует неравенство $|f_n(x) - f_n(c)| < \varepsilon$. Получаем, что если $|x - c| < \delta$, то $|f(x) - f(c)| < 3\varepsilon$. $\triangleleft$

Рассмотрим пространство $C^1[a;b]$ непрерывных на $[a;b]$ функций, имеющих непрерывную производную на $[a;b]$. Это пространство будет нормированным относительно нормы $\|f\| = \max\limits_{x\in[a;b]}|f(x)|$, но не будет полным. Действительно, возьмём последовательность функций $\sqrt{x^2+n^{-2}}$ в пространстве $C^1[-1;1]$. В $C[-1;1]$ эта последовательность сходится к $|x|$, значит, она является фундаментальной. Однако данная последовательность не имеет предела в $C^1[-1;1]$. Этот пример показывает, что равномерная сходимость не сохраняет дифференцируемость функций.

Теорема 8. *Пусть $f_n \in C^1[a;b]$. Если $f_n \rightrightarrows f$ и $f_n' \rightrightarrows g$, то $f' = g$.*

$\triangleright$ Зафиксируем $x \in [a;b]$. Рассмотрим функции

$$g_n(y) = \frac{f_n(y)-f(x)}{y-x} \quad \text{при} \quad y \neq x, \quad g_n(x) = f_n'(x).$$

Заметим, что $g_n(y)$ — непрерывная функция. Кроме того, последовательность g_n фундаментальна в $C[a;b]$. Действительно, по теореме Лагранжа

$$g_n(y) - g_m(y) = \frac{(f_n(y)-f_m(y)) - (f_m(x)-f_m(x))}{y-x} = f_n'(c) - f_m'(c)$$

и

$$\max_{y\in[a;b]}|g_n(y)-g_m(y)| \leqslant \max_{y\in[a;b]}|f_n'(y)-f_m'(y)|.$$

Значит, последовательность $g_n(y)$ сходится к непрерывной функции $h(y)$, причём

$$h(y) = \frac{f(y)-f(x)}{y-x} \quad \text{при} \quad y \neq x, \quad h(x) = g(x).$$

Из непрерывности h получаем $\lim\limits_{y\to x} h(y) = h(x)$, т. е. $f'(x) = g(x)$. $\triangleleft$

Заметим, что условия этой теоремы можно ослабить: предполагать вместо равномерной сходимости f_n лишь сходимость в некоторой точке $x_0 \in [a;b]$. Действительно, по теореме Лагранжа

$$|f_n(x) - f_m(x)| \leqslant |f_n(x_0) - f_m(x_0)| + |f_n'(c) - f_m'(c)|,$$

и из равномерной сходимости производных и сходимости $f_n(x_0)$ следует фундаментальность последовательности f_n в $C[a;b]$.

Следствие 3. *Пространство $C^k[a;b]$ с нормой*

$$\|f\| = \max_{x\in[a;b]}|f(x)| + \max_{x\in[a;b]}|f'(x)| + \ldots + \max_{x\in[a;b]}|f^{(k)}(x)|$$

является банаховым пространством.

Отметим, что в отличие от конечномерного пространства $\mathbb{R}^n$ в бесконечномерном пространстве всегда можно построить две неэквивалентные нормы. Например, на $C^1[a;b]$ таковыми будут $\|f\|_1 = \max\limits_{[a;b]}|f(x)|$ и $\|f\|_2 = \max\limits_{[a;b]}|f(x)| + \max\limits_{[a;b]}|f'(x)|$.

Открытые и замкнутые множества. Компакты

Множество U в метрическом пространстве называется *открытым*, если для всякого элемента $a \in U$ найдётся такой шар $B_r(a)$, что $B_r(a) \subset U$. Множество называется *замкнутым*, если его дополнение является открытым множеством.

Можно показать, что открытый шар является открытым множеством, а замкнутый шар является замкнутым множеством.

Предложение 9. *Объединение любого набора открытых множеств и пересечение конечного набора открытых множеств является открытым множеством. Пересечение любого набора замкнутых множеств и объединение конечного набора замкнутых множеств является замкнутым множеством.*

Отметим, что всякое открытое множество (просто в силу определения) является объединением шаров.

Точка a называется *граничной точкой* множества A в метрическом пространстве, если во всяком шаре $B_r(a)$ есть точки множества A и точки дополнения множества A. Множество граничных точек множества A называется *границей* A и обозначается через ∂A.

Точка a называется *предельной точкой* множества A, если во всяком шаре $B_r(a)$ лежит бесконечно много точек множества A.

Предложение 10. *Следующие утверждения равносильны.*

(i) *A — замкнутое множество.*

(ii) *A содержит все свои граничные точки.*

(iii) *A содержит все свои предельные точки.*

(iv) *Если $a_n \in A$ и $a_n \to a$, то $a \in A$.*

Покажем, что всякое конечномерное подпространство нормированного пространства является замкнутым множеством.

Предложение 11. *Пусть X — нормированное пространство и $F \subset X$ — конечномерное подпространство. Тогда F — замкнуто.*

$\triangleright$ Достаточно заметить, что на конечномерном пространстве все нормы эквивалентны и, следовательно, с любой из норм пространство полно. Из полноты следует замкнутость, так как сходящаяся последовательность элементов подпространства F является фундаментальной в F и в силу полноты сходится к элементу F, что и означает замкнутость F. $\triangleleft$

Множество $\overline{A} = A \cup \partial A$ называется *замыканием* множества A.

Предложение 12. *Множество $\overline{A}$ является замкнутым. Более того, множество A замкнуто тогда и только тогда, когда $\overline{A} = A$.*

$\triangleright$ Предположим, что граничная точка a множества $\overline{A}$ не является граничной точкой множества A. Тогда существует шар $B_r(a)$, который не пересекается с A. Поскольку $B_r(a)$ — открытое множество, никакая его точка не является граничной точкой множества A. Противоречие. $\triangleleft$

Отметим, что в метрическом пространстве замкнутый шар не всегда совпадает с замыканием открытого шара. В нормированном пространстве замкнутый шар является замыканием открытого шара.

Теорема 9 (Бэр). *Если полное метрическое пространство является объединением не более чем счётного набора замкнутых множеств, то хотя бы одно из этих замкнутых множеств содержит открытый шар.*

$\triangleright$ Пусть X — полное метрическое пространство и $X = \cup_n F_n$, где F_n — замкнутые множества. Предположим, что ни одно из них не содержит открытый шар. Множество $X \setminus F_1$ открыто и непусто, поэтому найдётся такой замкнутый шар $\overline{B}_1$ радиуса меньше 1, что $\overline{B}_1 \subset X \setminus F_1$. Открытый шар B_1 не содержится целиком в F_2, поэтому его пересечение с $X \setminus F_2$ является открытым и непустым множеством. Следовательно, найдётся такой замкнутый шар $\overline{B}_2$ радиусом меньше $1/2$, что $\overline{B}_2 \subset B_1$ и $\overline{B}_2 \subset X \setminus F_2$. Продолжая построение, получаем последовательность замкнутых вложенных шаров B_n, радиусы которых стремятся к нулю, и каждый шар B_n не имеет общих точек с $F_1, \ldots, F_n$. По теореме о вложенных шарах существует точка $c \in \bigcap\limits_n B_n$. По построению c не принадлежит ни одному F_n, что противоречит условию $X = \bigcup\limits_n F_n$. $\triangleleft$

Компакты

Множество K в метрическом пространстве называется *компактом*, если из любого его покрытия открытыми множествами можно выбрать конечное подпокрытие.

Всякое открытое множество является объединением шаров, поэтому в этом определении можно было бы рассматривать лишь покрытия шарами.

Пусть (X, ρ) — метрическое пространство. Тогда всякое его непустое подмножество $A \subset X$ также является метрическим пространством с метрикой ρ. Заметим, что шар $B_r^A(x)$ в A является пересечением шара $B_r(x)$ в X с множеством A.

Предложение 13. *Подмножество K метрического пространства (X, ρ) является компактом тогда и только тогда, когда K является компактом в метрическом пространстве (K, ρ).*

$\triangleright$ Как отмечалось выше, при проверке компактности достаточно рассматривать покрытия шарами. Пусть K является компактом в X. Предположим, что набор шаров $\{B_\alpha^K\}$ покрывает K в метрическом пространстве (K, ρ). Тогда такие шары B_α, что $B_\alpha^K = B_\alpha \bigcap K$, покрывают K в X. Выберем конечное подпокрытие $B_{\alpha_1}, \ldots, B_{\alpha_N}$. Тогда шары $B_{\alpha_1}^K, \ldots, B_{\alpha_N}^K$ покрывают K в (K, ρ).

Пусть теперь K является компактом в (K, ρ). Предположим, что набор шаров $\{B^\alpha\}$ покрывает K в X. Тогда шары B_α^K покрывают K в (K, ρ) и можно выбрать конечное подпокрытие $B_{\alpha_1}^K, \ldots, B_{\alpha_N}^K$. Тогда шары $B_{\alpha_1}, \ldots, B_{\alpha_N}$ покрывают K в X. $\lhd$

Таким образом, при доказательстве компактности множества K можно считать, что рассматривается метрическое пространство (K, ρ).

Предложение 14. *Параллелепипед $[a_1; b_1] \times [a_2; b_2] \times \ldots \times [a_n; b_n]$ является компактом в пространстве $\mathbb{R}^n$.*

$\rhd$ Предположим противное: существует бесконечное покрытие открытыми множествами, из которого нельзя выбрать конечное подпокрытие. Разделим каждое ребро $[a_k; b_k]$ пополам. Тогда весь параллелепипед разбивается на 2^n частей. Какой-то из параллелепипедов разбиения не имеет конечного подпокрытия. Продолжая построение, получаем последовательность вложенных параллелепипедов, длины рёбер которых стремятся к нулю, а сами рёбра образуют систему вложенных отрезков. Следовательно, существует общая точка, которая покрывается некоторым открытым множеством, а значит, этим множеством покрывается какой-то из построенных параллелепипедов. Противоречие. $\lhd$

Теорема 10. *Пусть K — компакт в метрическом пространстве X. Тогда*

(i) *K — ограниченное множество;*

(ii) *K — замкнутое множество;*

(iii) *всякое замкнутое подмножество K является компактом;*

(iv) *всякое бесконечное подмножество K имеет предельную точку в K;*

(v) *всякая последовательность элементов компакта K содержит подпоследовательность, сходящуюся к элементу из K.*

$\rhd$ (i) Пусть $a \in X$. Поскольку $K \subset X = \bigcup_n B_n(a)$, существует натуральное число N, для которого $K \subset B_N(a)$.

(ii) Пусть $a \in X \setminus K$. Поскольку $K \subset X \setminus \{a\} = \bigcup (X \setminus \overline{B}_{1/n}(a))$, найдётся N, для которого $K \subset X \setminus B_{1/N}(a)$ и $B_{1/N}(a) \subset X \setminus K$.

(iii) Пусть $F \subset K$ является замкнутым множеством. Если открытые множества $\{U_\alpha\}$ покрывают F, то эти же множества вместе с $X \setminus F$ покрывают K. Следовательно, существует конечный набор $U_{\alpha_1}, \ldots, U_{\alpha_n}$ и, возможно, $X \setminus F$, покрывающие K. Тогда $U_{\alpha_1}, \ldots, U_{\alpha_n}$ покрывают F.

(iv) Пусть E — бесконечное подмножество K, причём всякая точка из K не является предельной для E. Тогда для всякого $a \in K$ существует шар $B_{r_a}(a)$, в котором лежит лишь конечное множество точек из E. Эти шары покрывают K, поэтому существует конечный их набор, покрывающий K. Но это означает, что множество E конечно. Противоречие.

(v) Этот пункт следует из предыдущего. $\lhd$

Следствие 4 (критерий компактности в $\mathbb{R}^n$). *Множество $K \subset \mathbb{R}^n$ компактно тогда и только тогда, когда множество K является ограниченным и замкнутым.*

$\rhd$ Утверждение вытекает из п. iii теоремы, так как ограниченное и замкнутое множество можно поместить внутрь параллелепипеда, который сам уже является компактом. $\lhd$

В бесконечномерном случае замкнутости и ограниченности не хватает для компактности. Например, единичный шар в $C[a; b]$ не является компактом. Пусть f_n — непрерывная функция, которая равна нулю вне отрезка $[a + \frac{b-a}{n+2}; a + \frac{b-a}{n+1}]$ и единице в середине этого отрезка, причём $0 \leqslant f_n \leqslant 1$. Имеем $\|f_n - f_m\| = 2$ при $n \neq m$. Следовательно, из этой последовательности нельзя выбрать сходящуюся подпоследовательность. Более того, имеет место следующее фундаментальное утверждение.

Теорема 11. *Если в нормированном пространстве замкнутый шар положительного радиуса является компактом, то пространство конечномерно.*

$\triangleright$ Заметим, что если некоторый замкнутый шар положительного радиуса является компактом, то и всякий шар положительного радиуса является компактом. Пусть $\overline{B}_2(0)$ является компактом. Тогда существует конечный набор шаров $\{B_1(a_k)\}_{k=1}^N$, покрывающих шар $\overline{B}_2(0)$. Пусть L — линейная оболочка векторов $a_1, \ldots, a_N$. Тогда $\overline{B}_2(0) \subset \bigcup_{a\in L} B_1(a)$ и, следовательно, $\overline{B}_2(c) \subset \bigcup_{a\in L} B_1(a)$ для всякого $c \in L$, так как $a + c \in L$ для всех $a \in L$. Предположим, что $\overline{B}_r(0) \subset \bigcup_{a\in L} B_1(a)$ для некоторого $r > 0$. Тогда

$$\overline{B}_{2r}(0) \subset \bigcup_{a\in L} B_2(2a) = \bigcup_{a\in L} B_2(a) = \bigcup_{a\in L} B_1(a).$$

Таким образом, всякий шар $\overline{B}_{2^n}(0)$ содержится в $\bigcup_{a\in L} B_1(a)$. Значит, $X \subset \bigcup_{a\in L} B_1(a)$. Покажем теперь, что $X = L$. Предположим, что существует $z \in X \setminus L$. Оболочка L замкнута, поэтому в $X \setminus L$ найдётся шар $B_\delta(z)$. Это означает, что $\|a - z\| > \delta$ для всех $a \in L$. Тогда $\|a - z\delta^{-1}\| > 1$ для всех $a \in L$, но это противоречит тому, что $X \subset \bigcup_{a\in L} B_1(a)$. $\triangleleft$

Предел функции

Пусть (X, ρ_X) и (Y, ρ_Y) — метрические пространства и a — предельная точка множества X. Пусть также задана функция $f\colon X \setminus \{a\} \to Y$.

Определение Гейне. Элемент $b \in Y$ называется *пределом* функции f при $x \to a$, если для всякой такой последовательности $x_n \in X$, что $x_n \to a$ и $x_n \neq a$, имеем $f(x_n) \to b$. Далее используем обозначения: $\lim\limits_{x\to a} f(x) = b$ или $f(x) \to b$ при $x \to a$.

Предложение 15 (единственность предела)**.** *Если $\lim\limits_{x\to a} f(x) = b$ и $\lim\limits_{x\to a} f(x) = c$, то $b = c$.*

$\triangleright$ Пусть $x_n \to a$ и $x_n \neq a$. Тогда $f(x_n) \to b$ и $f(x_n) \to c$. Последовательность в метрическом пространстве имеет не более одного предела, поэтому $b = c$. $\triangleleft$

Предложение 16 (арифметика предела)**.** *Пусть X — метрическое пространство, a — предельная точка X. Пусть Y — нормированное пространство. Если $f, g\colon X\setminus\{a\} \to Y$ и $\alpha\colon X\setminus\{a\} \to \mathbb{R}$ и $\lim\limits_{x\to a} f(x) = b$, $\lim\limits_{x\to a} g(x) = c$, $\lim\limits_{x\to a} \alpha(x) = \gamma$, то $\lim\limits_{x\to a}(f(x) + g(x)) = b + c$ и $\lim\limits_{x\to a} \alpha(x)f(x) = \gamma b$.*

$\triangleright$ Утверждение следует из определения предела по Гейне и соответствующих свойств предела последовательности. $\triangleleft$

Предложение 17 (предел сложной функции)**.** *Пусть X, Y, Z — метрические пространства, a и b — предельные точки X и Y соответственно. Пусть $f\colon X \setminus \{a\} \to Y \setminus \{b\}$ и $g\colon Y \setminus \{b\} \to Z$. Если $\lim\limits_{x\to a} f(x) = b$ и $\lim\limits_{y\to b} g(y) = c$, то $\lim\limits_{x\to a} g(f(x)) = c$.*

$\triangleright$ Если $x_n \to a$ и $x_n \neq a$, то $f(x_n) \to b$ и $f(x_n) \neq b$, а это, в свою очередь, влечёт $g(f(x_n)) \to c$. $\triangleleft$

Пусть, как и выше, (X, ρ_X) и (Y, ρ_Y) — метрические пространства, a — предельная точка множества X и задана функция $f\colon X \setminus \{a\} \to Y$.

Определение Коши. Элемент $b \in Y$ называется *пределом* функции f при $x \to a$, если для всякого $\varepsilon > 0$ найдётся такое $\delta > 0$, что из неравенств $0 < \rho_X(x, a) < \delta$ следует неравенство $\rho_Y(f(x), b) < \varepsilon$.

Теорема 12. *Предел $\lim\limits_{x\to a} f(x)$ в смысле определения Гейне существует тогда и только тогда, когда существует предел $\lim\limits_{x\to a} f(x)$ в смысле определения Коши, причём если эти пределы существуют, то они равны.*

$\triangleright$ Предположим, что $f(x) \to b$ при $x \to a$ по определению Коши. Пусть $x_n \to a$ и $x_n \neq a$. Для всякого $\varepsilon > 0$ найдётся такое $\delta > 0$, что неравенство $0 < \rho_X(x, a) < \delta$ влечёт $\rho_Y(f(x), b) < \varepsilon$. По определению предела последовательности существует такой номер N, что $\rho_X(x_n, a) < \delta$ для всех $n > N$. Поскольку $x_n \neq a$, для всех $n > N$ верно неравенство $\rho_Y(f(x_n), b) < \varepsilon$. Получаем, что $f(x_n) \to b$.

Пусть теперь $f(x) \to b$ при $x \to a$ по определению Гейне, но определение Коши не выполняется. Существует такое $\varepsilon > 0$, что для всех $\delta > 0$ найдётся x, для которого $0 < \rho_X(x, a) < \delta$ и $\rho_Y(f(x), b) \geqslant \varepsilon$. Пусть x_n соответствует $\delta = 1/n$. Тогда $x_n \to a$, $x_n \neq a$ и $\rho(f(x_n), b) \geqslant \varepsilon$, а это противоречит определению Гейне. $\triangleleft$

Следующие утверждения наиболее просто доказываются именно с помощью определения Коши.

Предложение 18 (ограниченность и отделимость)**.** (i) *Если* $\lim\limits_{x\to a} f(x) = b$, *то существует такой шар* $B_\delta(a)$, *что* $f(x) \in B_\varepsilon(b)$ *для всех* $x \in B_\delta(a) \setminus \{a\}$.

(ii) *Если* $\lim\limits_{x\to a} f(x) = b$ *и* $c \neq b$, *то существует такой шар* $B_\delta(a)$, *что* $f(x) \notin B_{r/2}(c)$ *для всех* $x \in B_\delta(a) \setminus \{a\}$, *где* $r = \rho_Y(b,c)$.

$\triangleright$ Для доказательства пункта (i) достаточно положить в определении Коши $\varepsilon = 1$, а для доказательства пункта (ii) — положить $\varepsilon = r/2$. $\triangleleft$

Теорема 13 (критерий Коши)**.** *Пусть* X *и* Y — *метрические пространства, причём* Y — *полное пространство. Пусть, далее,* a — *предельная точка пространства* X *и* $f\colon X \setminus \{a\} \to Y$. *Предел* $\lim\limits_{x\to a} f(x)$ *существует тогда и только тогда, когда выполняется условие Коши: для всякого* $\varepsilon > 0$ *существует такое* $\delta > 0$, *что из неравенств* $0 < \rho_X(x_1, a) < \delta$ *и* $0 < \rho_X(x_2, a) < \delta$ *следует неравенство* $\rho_Y(f(x_1), f(x_2)) < \varepsilon$.

$\triangleright$ Докажем лишь достаточность условия Коши. Пусть $x_n \to a$ и $x_n \neq a$. Для всякого $\varepsilon > 0$ найдётся такое $\delta > 0$, что из неравенств $0 < \rho_X(x_1, a) < \delta$ и $0 < \rho_X(x_2, a) < \delta$ следует неравенство $\rho_Y(f(x_1), f(x_2)) < \varepsilon$. Существует такой номер N, что для всех $n > N$ верно неравенство $0 < \rho_X(x_n, a) < \delta$. Следовательно, для всех $n, m > N$ выполняется $\rho_Y(f(x_n), f(x_m)) < \varepsilon$ и последовательность $f(x_n)$ фундаментальна. Поскольку Y является полным пространством, существует предел последовательности $f(x_n)$. Этот предел не зависит от выбора последовательности x_n, иначе можно было бы построить такую последовательность $x_n \to a$, $x_n \neq a$, для которой $f(x_n)$ не имеет предела. $\triangleleft$

Непрерывные функции. Связные пространства.

Пусть (X, ρ_X) и (Y, ρ_Y) — метрические пространства. Функция $f\colon X \to Y$ *непрерывна* в точке $a \in X$, если для всякого $\varepsilon > 0$ найдётся такое $\delta > 0$, что из неравенства $\rho_X(x, a) < \delta$ следует неравенство $\rho_Y(f(x), f(a)) < \varepsilon$.

Предложение 19. *Пусть* $f\colon X \to Y$. *Следующие утверждения равносильны:*

(i) *функция* f *непрерывна в точке* a;

(ii) *если* $x_n \to a$, *то* $f(x_n) \to f(a)$;

(iii) a — *предельная точка пространства* X *и* $\lim\limits_{x\to a} f(x) = f(a)$, *или* a — *изолированная точка.*

Предложение 20. *Пусть* X, Y, Z — *метрические пространства. Если функция* $f\colon X \to Y$ *непрерывна в точке* $a \in X$ *и функция* $g\colon Y \to Z$ *непрерывна в точке* $f(a)$, *то функция* $g(f(x))$ *непрерывна в точке* a.

$\triangleright$ Пусть $x_n \to a$. Тогда $f(x_n) \to f(a)$, а это влечёт $g(f(x_n)) \to g(f(a))$. $\triangleleft$

С помощью теоремы о композиции можно, например, доказать, что если числовые функции f и g непрерывны в точке a метрического пространства X, то функции $f+g$ и fg непрерывны в точке a. По определению проверяется, что отображение $\Psi\colon x \to (f(x), g(x))$ из X в $\mathbb{R}^2$ непрерывно в точке a и отображения $\sum\colon (x_1, x_2) \to x_1 + x_2$ и $\prod\colon (x_1, x_2) \to x_1 x_2$ непрерывны всюду. Поскольку $f+g = \sum \circ \Psi$ и $fg = \prod \circ \Psi$, по теореме о композиции непрерывных функций $f+g$ и fg непрерывны в точке a.

Будем говорить, что $f\colon X \to Y$ непрерывное отображение из X в Y или непрерывная функция на X, если f непрерывна в каждой точке множества X.

Теорема 14. *Отображение* $f\colon X \to Y$ *непрерывно тогда и только тогда, когда для всякого открытого множества* $V \subset Y$ *множество* $f^{-1}(V)$ *открыто.*

$\triangleright$ Докажем необходимость. Пусть $a \in f^{-1}(V)$. Найдётся такое $\varepsilon > 0$, что $B_\varepsilon(f(a)) \subset V$. Существует такое число $\delta > 0$, что $B_\delta(a) \subset f^{-1}(B_\varepsilon(f(a)))$. Следовательно, $B_\delta(a) \subset f^{-1}(V)$. Для доказательства достаточности заметим, что для всякого шара $B_\varepsilon(f(a))$ множество $f^{-1}(B_\varepsilon(f(a)))$ открыто, а значит, точка a содержится в этом множестве с некоторой своей окрестностью. $\triangleleft$

Следствие 5. *Образ компакта при непрерывном отображении является компактом.*

$\triangleright$ Пусть K — компакт и f — непрерывное отображение K в некоторое метрическое пространство Y. Для всякого покрытия $\{U_\alpha\}$ открытыми множествами $f(K)$ открытые множества $f^{-1}(U_\alpha)$ образуют покрытие самого компакта K. Пусть $f^{-1}(U_{\alpha_1}), \dots, f^{-1}(U_{\alpha_N})$ — конечное подпокрытие. Если $y \in f(K)$, то найдётся такой $x \in K$, что $y = f(x)$. Следовательно, найдётся такое α_k, что $x \in f^{-1}(U_{\alpha_k})$, а это в свою очередь влечёт $y = f(x) \in U_{\alpha_k}$. Следовательно, $U_{\alpha_1}, \dots, U_{\alpha_N}$ покрывают $f(K)$. $\triangleleft$

Следствие 6 (теорема Вейерштрасса). *Пусть $f\colon K \to \mathbb{R}$ — непрерывная функция и K — компакт. Тогда f ограниченна и достигает наибольшего и наименьшего значений.*

$\triangleright$ Образ $f(K)$ является компактом в $\mathbb{R}$. Следовательно, $f(K)$ — ограниченное и замкнутое множество, которое содержит граничные точки $\inf f(K)$ и $\sup f(K)$. $\triangleleft$

Используя теорему Вейерштрасса, можно дать несколько иное и более короткое доказательство эквивалентности норм в конечномерном пространстве. Напомним, что в начале доказательства теоремы 5 было установлено, что если $\|x\|_a$ — произвольная норма, а $\|x\|_b$ — стандартная евклидова норма на конечномерном пространстве $\mathbb{R}^n$, то $\|x\|_a \leqslant C_a\|x\|_b$, где $C_a = n\max_k \|e_k\|_a$. Из этой оценки следует, что функция $x \mapsto \|x\|_a$ является непрерывной относительно нормы $\|\cdot\|_b$. Единичная сфера $\{x : \|x\|_b = 1\}$ является компактом в $\mathbb{R}^n$, поэтому функция $\|x\|_a$ достигает на сфере своего минимума m и максимума M, причём $m > 0$. Пусть $x \neq 0$. Тогда $y = \frac{x}{\|x\|_b}$ принадлежит единичной сфере и, следовательно, верна оценка $m \leqslant \|y\|_a \leqslant M$. Используя свойства нормы и умножая последнее неравенство на $\|x\|_b$, получаем $m\|x\|_b \leqslant \|x\|_a \leqslant M\|x\|_b$.

Следствие 7 (теорема Кантора). *Пусть $f\colon K \to Y$ непрерывная функция и K — компакт. Тогда f равномерно непрерывна на K, т. е. для всякого $\varepsilon > 0$ существует такое $\delta > 0$, что из неравенства $\rho_K(x_1, x_2) < \delta$ следует неравенство $\rho_Y(f(x_1), f(x_2)) < \varepsilon$.*

$\triangleright$ Функция f непрерывна на K, поэтому для всякого $a \in K$ существует такой шар $B_{\delta_a}(a)$, что $\rho_Y(f(x_1), f(x_2)) < \varepsilon$ для всех $x_1, x_2 \in B_{\delta_a/3}(a)$. Шары $B_{\delta_a/3}(a)$ покрывают компакт K. Пусть $B_{\delta_1/3}(a_1), \dots, B_{\delta_N/3}(a_N)$ образуют конечное подпокрытие. Покажем, что число $\delta = \min_k \delta_k/3$ искомое. Пусть $\rho_K(x_1, x_2) < \delta$. Точка x_1 принадлежит некоторому $B_{\delta_k/3}(a_k)$. Тогда $x_2 \in B_{\delta_k}(a_k)$ и $\rho_Y(f(x_1), f(x_2)) < \varepsilon$. $\triangleleft$

Пусть $f\colon X \to Y$ — непрерывная биекция. Обратное отображение может и не быть непрерывным. Действительно, пусть $X = [0;1) \cup [2;3]$ и $f(x) = x$ при $x \in [0;1)$, $f(x) = x - 1$ при $x \in [2;3]$. Это биективное отображение X на $[0;2]$, но обратное отображение разрывно в точке $y = 1$.

Если биекция f такова, что f и обратное к нему отображение непрерывно, то отображение f называют *гомеоморфизмом.*

Теорема 15. *Если $f\colon X \to Y$ — непрерывная биекция и X — компакт, то f — гомеоморфизм.*

$\triangleright$ Пусть $y_n \to y$. Из последовательности $x_n = f^{-1}(y_n)$ можно выбрать сходящуюся подпоследовательность $x_{n_k} \to x$. В силу непрерывности f имеем $f(x) = y$, и значит, всякая сходящаяся подпоследовательность сходится к одному и тому же пределу. Следовательно, $f^{-1}(y_n) \to f^{-1}(y)$. $\triangleleft$

Для непрерывных на отрезке функций кроме теорем Вейерштрасса и Кантора важнейшим утверждением была теорема Коши о промежуточном значении. В общем случае свойство промежуточного значения связано не только с непрерывностью функции, но и с метрическим пространством, на котором определена функция.

Связные пространства

Метрическое пространство X называется *несвязным*, если его можно представить в виде объединения двух непустых непересекающихся открытых множеств (поскольку эти множества дополняют друг друга до X, они одновременно являются и замкнутыми). Метрическое пространство, которое не является несвязным, называют *связным.*

Множество $M \subset X$ называется *несвязным*, если существуют такие открытые множества U и V, что $U \cap M \neq \varnothing$, $V \cap M \neq \varnothing$, $U \cap V = \varnothing$ и $M \subset U \cup V$. Далее будем говорить, что эти множества U и V *разделяют* множество M. Множество $M \subset X$, которое не является несвязным, называют *связным.*

Предложение 21. *Пусть (X, ρ) — метрическое пространство и $M \subset X$. Множество M несвязно тогда и только тогда, когда метрическое пространство (M, ρ) несвязно.*

$\triangleright$ Пусть M — несвязное подмножество метрического пространства (X, ρ). Тогда существуют открытые множества U и V, разделяющие M. Получаем, что $U \cap M$ и $V \cap M$ разделяют метрическое пространство M. Пусть теперь метрическое пространство (M, ρ) несвязно, т. е. существуют открытые в M (но, вообще говоря, не в X) множества U_M и V_M, разделяющие M. Для всякого $a \in U_M$ существует такой шар $B_{\delta_a}(a)$, что $B_{\delta_a}(a) \cap M \subset U_M$, т. е. $B_{\delta_a}(a)$ не пересекается с V_M. Для всякого $b \in V_M$ существует такой шар $B_{\delta_b}(b)$, что $B_{\delta_b}(b) \cap M \subset V_M$, т. е. $B_{\delta_b}(b)$ не пересекается с U_M. Положим $U = \cup_{a \in U_M} B_{\delta_a/2}(a)$ и $V = \cup_{b \in V_M} B_{\delta_b/2}(b)$. Покажем, что открытые множества U и V разделяют M. Достаточно только проверить, что $U \cap V = \varnothing$. Если существует $c \in U \cap V$, то $c \in B_{\delta_a/2}(a)$ или $c \in B_{\delta_b/2}(b)$ для некоторых $a \in U_M$ и $b \in V_M$. Тогда

$$\rho(a, b) \leqslant \rho(a, c) + \rho(c, b) < (\delta_a + \delta_b)/2 \leqslant \max\{\delta_a, \delta_b\}.$$

Это означает, что $a \in B_{\delta_b}(b)$ или $b \in B_{\delta_a}(a)$, однако обе ситуации невозможны. $\triangleleft$

Теорема 16. *Метрическое пространство X несвязно тогда и только тогда, когда существует непрерывная функция $f\colon X \to \mathbb{R}$, принимающая ровно два значения: 0 и 1.*

$\triangleright$ Если X несвязно, то существуют открытые множества U и V, разделяющие X. Функция $f(x) = 0$ при $x \in U$ и $f(x) = 1$ при $x \in V$ является непрерывной и принимает ровно два значения 0 и 1. Обратно, если на X определена непрерывная функция f, принимающая ровно два значения 0 и 1, то открытые множества $f^{-1}((-\infty; 1/2))$ и $f^{-1}((1/2; +\infty))$ разделяют X. $\triangleleft$

Опишем все связные множества на $\mathbb{R}$.

Предложение 22. *Множество $M \subset \mathbb{R}$ связно тогда и только тогда, когда M — промежуток, т. е. вместе со всякими двумя своими точками содержит соединяющий их отрезок.*

$\triangleright$ Промежуток является связным множеством, так как по теореме о промежуточном значении не существует непрерывной функции, принимающей на промежутке ровно два различных значения. Если множество M не является промежутком, то существуют такие точки $a < c < b$, что $a, b \in M$ и $c \notin M$. Открытые лучи $(-\infty; c)$ и $(c; +\infty)$ разделяют M, значит, множество M не является связным. $\triangleleft$

Теорема 17. *Если $f\colon X \to Y$ — непрерывное отображение и X — связно, то $f(X)$ — связно. В частности, если $Y = \mathbb{R}$, то $f(X)$ — промежуток.*

$\triangleright$ Если $f(X)$ не является связным, то на $f(X)$ существует непрерывная функция g, принимающая ровно два значения 0 и 1. Тогда на X определена непрерывная функция $g(f(x))$, принимающая ровно два значения, что противоречит связности X. $\triangleleft$

Метрическое пространство X называется *линейно связным*, если для любых двух точек $x_0, x_1 \in X$ существует такое непрерывное отображение $\gamma\colon [0; 1] \to X$ (непрерывный путь), что $\gamma(0) = x_0$ и $\gamma(1) = x_1$.

Теорема 18. *Если X линейно связно, то X связно.*

$\triangleright$ Предположим, что X несвязно. Тогда существует непрерывная функция $f\colon X \to \mathbb{R}$, принимающая ровно два значения: 0 и 1. Пусть $f(x_0) = 0$ и $f(x_1) = 1$. По условию существует путь $\gamma\colon [0; 1] \to X$, соединяющий точки x_0 и x_1. Тогда на отрезке $[0; 1]$ определена непрерывная функция $f(\gamma(t))$, принимающая ровно два значения, что невозможно. Противоречие. $\triangleleft$

Утверждение, обратное к теореме, вообще говоря, неверно. Примером является подмножество плоскости, состоящее из графика функции $y = \sin\frac{1}{x}$ и отрезка $[-1; 1]$ оси Oy. Оказывается, что для открытых подмножеств нормированного пространства из связности следует линейная связность.

Предложение 23. *Пусть X — нормированное пространство. Открытое множество $M \subset X$ связно тогда и только тогда, когда M линейно связно.*

$\triangleright$ Покажем, что из связности следует линейная связность. Введём отношение эквивалентности: $x \sim y$ тогда и только тогда, когда в M существует путь, соединяющий x и y. Классы эквивалентности являются линейно связными множествами. Кроме того, класс эквивалентности, содержащий некоторый элемент x, содержит открытый шар с центром в точке x. Действительно, x можно соединить с каждой точкой шара отрезком. Значит, классы эквивалентности являются открытыми множествами. В случае, когда классов эквивалентности не менее двух, множество M не является связным. Следовательно, класс эквивалентности один, и он совпадает с M. $\triangleleft$

Рассмотренные в доказательстве теоремы классы эквивалентности называют *компонентами линейной связности.*

В заключение обсудим важный пример непрерывных отображений — линейные.

Линейные отображения

Пусть X, Y — нормированные пространства над $\mathbb{R}$. Отображение $L\colon X \to Y$ называется *линейным отображением*, если для всех $\alpha_1, \alpha_2 \in \mathbb{R}$, $x_1, x_2 \in X$ верно равенство

$$L(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 L(x_1) + \alpha_2 L(x_2).$$

Если $Y = \mathbb{R}$, то линейные отображения часто называют *линейными функционалами.*

Предложение 24. *Всякое линейное отображение $L\colon \mathbb{R}^n \to \mathbb{R}^m$ имеет вид $L(x) = Ax$, где A — матрица $n \times m$, в частности, в случае $m = 1$ имеем $L(x) = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n$.*

$\triangleright$ Пусть $e_1, \ldots, e_n$ — стандартный базис в $\mathbb{R}^n$. Тогда

$$L(x) = L(x_1 e_1 + \ldots + x_n e_n) = x_1 L(e_1) + \ldots + x_n L(e_n).$$

Пусть теперь $(a^{1j}, \ldots, a^{nj})$ — координаты вектора $L(e_j)$. Матрица $A = (a^{ij})$ искомая. $\triangleleft$

Из этого утверждения следует, что всякое линейное отображение конечномерных пространств является непрерывным. В случае бесконечномерного пространства X всегда можно построить разрывный линейный функционал. Например, линейный функционал $L(f) = f'(0)$ на пространстве $C^1[-1;1]$ с нормой $\|f\| = \max\limits_{[0;1]} |f|$ не является непрерывной функцией.

Предложение 25. *Пусть X, Y — нормированные пространства над $\mathbb{R}$. Следующие утверждения равносильны*:

(i) *линейное отображение $L\colon X \to Y$ непрерывно*;

(ii) *линейное отображение L непрерывно в нуле*;

(iii) *существует такое число $C > 0$, что $\|L(x)\| \leqslant C\|x\|$ для всех $x \in X$.*

Наименьшее из чисел C, для которых выполняется неравенство пункта (iii) называют *нормой* $\|L\|$ линейного отображения L, т. е.

$$\|L\| = \sup_{x \neq 0} \frac{\|L(x)\|}{\|x\|}.$$

Можно показать, что величина $\|L\|$ действительно является нормой на линейном пространстве непрерывных линейных отображений. Кроме того, если $H\colon Z \to X$ и $L\colon X \to Y$ — линейные непрерывные отображения, то композиция $L \circ H$ является линейным непрерывным отображением, причём выполняется неравенство $\|L \circ H\| \leqslant \|L\|\|H\|$.

В заключение отметим, что если линейное отображение $L\colon X \to Y$ является биекцией, то обратное отображение L^{-1} также является линейным. Однако из непрерывности L, вообще говоря, не следует непрерывность L^{-1}.

Дифференцируемые функции

Пусть X, Y — нормированные пространства и отображение $f\colon X \to Y$ определено в некоторой окрестности точки $a \in X$ (формально запись $f\colon X \to Y$ означает, что отображение f определено на всём пространстве X, но мы будем для краткости так писать и в тех случаях, когда f определено на подмножестве X, причём из контекста всегда будет ясно, на каком именно подмножестве). Будем говорить, что f *дифференцируемо* в точке a, если существует такое линейное непрерывное отображение $L\colon X \to Y$, что для всех h из некоторой окрестности нуля верно равенство

$$f(a + h) - f(a) = L(h) + \alpha(h)\|h\|,$$

где $\alpha\colon X \to Y$ удовлетворяет условию $\lim\limits_{h \to 0} \alpha(h) = 0$. Линейное отображение L называют *дифференциалом* функции f в точке a и обозначают через $df(a, h)$ или $df(h)$.

Простейшими (отличными от констант) примерами дифференцируемых отображений являются линейные непрерывные отображения. Действительно, для непрерывного линейного отображения L верно равенство $L(a + h) - L(a) = L(h)$, т. е. $dL(h) = L(h)$ и $\alpha(h) = 0$.

Предложение 26. *Если отображение $f\colon X \to Y$ дифференцируемо в точке a, то f непрерывно в точке a.*

$\triangleright$ Достаточно заметить, что

$$\lim_{h\to 0} f(a+h) = f(a) + \lim_{h\to 0}\Big(df(a,h) + \alpha(h)\|h\|\Big) = f(a). \qquad \triangleleft$$

В частности, разрывное линейное отображение (см. замечание перед предложением 25) не является дифференцируемым ни в одной точке.

Теорема 19. *Пусть $f\colon X \to Y$ и $a \in X$.*

(i) *Если f дифференцируемо в точке a и df — его дифференциал, то для всякого вектора v в некоторой окрестности нуля на числовой прямой определена функция $t \mapsto f(a+tv)$ и имеет место равенство*

$$\lim_{t\to 0}\frac{f(a+tv)-f(a)}{t} = df(v).$$

В частности, дифференциал определён единственным образом.

(ii) *Отображение f дифференцируемо в точке a и df — его дифференциал тогда и только тогда, когда в некоторой окрестности нуля определено отображение $h \mapsto f(a+h)$ и имеет место равенство*

$$\lim_{h\to 0}\frac{\|f(a+h)-f(a)-df(h)\|}{\|h\|} = 0.$$

$\triangleright$ Докажем первое утверждение. В определении дифференцируемости положим $h = tv$:

$$f(a+tv)-f(a) = df(tv) + \alpha(tv)\cdot\|tv\| = tdf(v) + \alpha(tv)\cdot|t|\cdot\|v\|.$$

Тогда

$$\lim_{t\to 0}\frac{f(a+tv)-f(a)}{t} = \lim_{t\to 0}\Big(df(v) + \alpha(tv)\ \mathrm{sgn}\, t\|v\|\Big) = df(v).$$

Второе утверждение является переформулировкой определения дифференцируемости, которое состоит в том, что функция

$$\alpha(h) = \frac{f(a+h)-f(a)-df(h)}{\|h\|}$$

стремится к нулю при $h \to 0$. $\triangleleft$

Предел

$$\lim_{t\to 0}\frac{f(a+tv)-f(a)}{t}$$

называется *производной функции f по вектору v* и обозначается через $\frac{\partial f}{\partial v}(a)$. Таким образом, если функция f дифференцируема в точке a, то

$$df(a,v) = \frac{\partial f}{\partial v}(a).$$

Производная по вектору может существовать и в случае, когда функция не является дифференцируемой и даже непрерывной. Рассмотрим функцию $f\colon \mathbb{R}^2 \to \mathbb{R}$, равную нулю всюду, за исключением множества $\{(x,y)\colon y = x^2, x > 0\}$, в точках которого f равна единице. Для всякого вектора v функция $t \to f(0+tv)$ равна нулю в некоторой окрестности точки $t = 0$, и производная f по вектору v в нуле равна нулю. Однако эта функция разрывна в нуле.

Рассмотрим важный частный случай функции $f\colon \mathbb{R}^n \to \mathbb{R}$. Если f дифференцируема в точке a, то дифференциал df в точке a является линейным отображением $\mathbb{R}^n \to \mathbb{R}$ и имеет вид

$$df(h) = c_1h_1 + \ldots + c_nh_n.$$

Более того,

$$c_k = df(e_k) = \frac{\partial f}{\partial e_k}(a) = \lim_{t\to 0}\frac{f(a_1,\ldots,a_k+t,\ldots,a_n)-f(a_1,\ldots,a_n)}{t}.$$

Производную по вектору e_k называют *частной производной* по переменной x_k и обозначают через $\frac{\partial f}{\partial x_k}$ или $\partial_{x_k} f$. С учётом новых обозначений дифференциал f имеет следующий вид:

$$df(h) = \frac{\partial f}{\partial x_1}(a)h_1 + \ldots + \frac{\partial f}{\partial x_n}(a)h_n.$$

Дифференциал функции $f(x) = x_k$ сопоставляет вектору h его k-ю координату h_k, т. е. $dx_k(h) = h_k$. Таким образом, дифференциал можно записать так:

$$df = \frac{\partial f}{\partial x_1} dx_1 + \ldots + \frac{\partial f}{\partial x_n} dx_n.$$

Далее будет показано, что данная форма дифференциала является наиболее удобной при заменах переменных.

Вектор из частных производных $(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n})$ называется *градиентом* функции f и обозначается через $\mathrm{grad} f$ или ∇f. Имеет место равенство

$$df(v) = \frac{\partial f}{\partial v} = \langle \nabla f, v \rangle.$$

Предложение 27. *Пусть функция f дифференцируема в точке a и $\nabla f(a) \neq 0$. Тогда производная по единичному вектору v функции f в точке a принимает наибольшее значение при $v = \nabla f(a)/\|\nabla f(a)\|$ и принимает наименьшее значение при $v = -\nabla f(a)/\|\nabla f(a)\|$.*

$\rhd$ Утверждение следует из формулы $\frac{\partial f}{\partial v} = \langle \nabla f, v \rangle$ и неравенства Коши — Буняковского — Шварца. $\lhd$

На этом простом наблюдении про градиент основан известный *метод градиентного спуска*, когда минимум функции f ищут как предел рекуррентной последовательности

$$x_{n+1} = x_n - \gamma \nabla f(x_n),$$

где положительное число γ выбирается, исходя из свойств функции f. Например, для функции $f(x, y) = x^2 + y^2$ градиент равен $(2x, 2y)$. Пусть $0 < \gamma < 1/2$. Получаем последовательность

$$(x_{n+1}, y_{n+1}) = (1 - 2\gamma)(x_n, y_n).$$

Имеем $0 < 1 - 2\gamma < 1$ и $(x_n, y_n) \to (0, 0)$. Точка $(0, 0)$ является точкой минимума f.

Выше уже отмечалось, что наличие производных по всякому вектору ещё не влечёт дифференцируемость функции. Однако справедливо следующее достаточное условие дифференцируемости функции.

Теорема 20. *Если все частные производные функции $f\colon \mathbb{R}^n \to \mathbb{R}$ существуют в окрестности точки a и непрерывны в точке a, то функция f дифференцируема в точке a.*

$\rhd$ Рассмотрим случай функции двух переменных. Пусть $a = (a_1, a_2)$ и шар $B_r(a)$ содержится в окрестности из условия теоремы. Будем далее рассматривать такие приращения аргумента h_1 и h_2, что $(a_1 + h_1, a_2 + h_2) \in B_r(a)$. Представим приращение функции f в виде

$$f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) = f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) + f(a_1, a_2 + h_2) - f(a_1, a_2).$$

По условию функция $f(x, a_2 + h_2)$ одной переменной x дифференцируема на интервале, содержащем отрезок с концами a_1 и $a_1 + h_1$, а значит, она непрерывна на этом отрезке и дифференцируема внутри него. Таким образом, выполняются условия теоремы Лагранжа, поэтому справедливо равенство $f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) = \frac{\partial f}{\partial x_1}(c_1, a_2 + h_2) h_1$, где c_1 лежит между a_1 и $a_1 + h_1$. Аналогично применяя теорему Лагранжа ко вторым двум слагаемым, получаем

$$f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) = \frac{\partial f}{\partial x_1}(c_1, a_2 + h_2) h_1 + \frac{\partial f}{\partial x_2}(a_1, c_2) h_1,$$

где c_2 лежит между a_2 и $a_2 + h_2$. Представим правую часть в следующем виде:

$$\frac{\partial f}{\partial x_1}(a_1, a_2) h_1 + \frac{\partial f}{\partial x_2}(a_1, a_2) h_1 + \alpha(h) \|h\|,$$

где

$$\alpha(h) = \Big(\frac{\partial f}{\partial x_1}(c_1, a_2 + h_2) - \frac{\partial f}{\partial x_1}(a_1, a_2) \Big) \frac{h_1}{\|h\|} + \Big(\frac{\partial f}{\partial x_2}(a_1, c_2) - \frac{\partial f}{\partial x_2}(a_1, a_2) \Big) \frac{h_2}{\|h\|}.$$

Частные производные непрерывны в точке a, поэтому $\lim\limits_{h \to 0} \alpha(h) = 0$. $\lhd$

Рассмотрим теперь отображение f из $\mathbb{R}^n$ в $\mathbb{R}^m$. Пусть f дифференцируемо в точке a. Тогда дифференциал является линейным отображением из $\mathbb{R}^n$ в $\mathbb{R}^m$ и имеет вид $df(h) = Ch$, где C — матрица $n \times m$. Матрица C называется *матрицей Якоби* отображения f в точке a и обозначается через $J_f(a)$ или $Df(a)$.

Всякое отображение $f\colon \mathbb{R}^n \to \mathbb{R}^m$ имеет вид

$$f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \dots \\ f_m(x) \end{pmatrix},$$

где $f_i\colon \mathbb{R}^n \to \mathbb{R}$.

Предложение 28. *Отображение f дифференцируемо в точке a тогда и только тогда, когда в точке a дифференцируема каждая функция f_i. Если отображение f дифференцируемо в точке a, то элементы матрицы Якоби имеют вид*

$$\frac{\partial f_i}{\partial x_j}(a).$$

$\triangleright$ Равенство

$$\lim_{h\to 0} \frac{\|f(a+h) - f(a) - df(h)\|}{\|h\|} = 0$$

равносильно системе равенств

$$\lim_{h\to 0} \frac{|f_i(a+h) - f_i(a) - L_i(h)|}{\|h\|} = 0, \quad i = 1, 2, \dots, m.$$

Здесь $L_i(h)$ — i-я координата вектора $df(h)$, т. е. $c_{i1}h_1 + c_{i2}h_2 + \ldots + c_{in}h_n$, где c_{ij} — элементы матрицы Якоби отображения f. Выше было показано, что

$$L_i(h) = \frac{\partial f_i}{\partial x_1} h_1 + \ldots + \frac{\partial f_i}{\partial x_n} h_n,$$

а это значит, что

$$c_{ij} = \frac{\partial f_i}{\partial x_j}, \quad j = 1, 2, \dots, n. \qquad \triangleleft$$

Таким образом, матрица Якоби отображения f — матрица, строки которой являются градиентами функций f_i.

Важно понимать, что дифференцируемое отображение f локально мало отличается от аффинного отображения:

$$f(x) \approx f(a) + J_f(a)(x-a) \quad \text{при} \quad x \approx a.$$

Это примерное равенство можно трактовать ещё так: приращению аргумента $h = x - a$ соответствует приращение функции $f(a+h) - f(a)$, которое примерно равно $J_f(a)h$.

В некоторых задачах это позволяет без вычислений частных производных выписывать матрицу Якоби. Например, рассмотрим отображение $x = r\cos\varphi$, $y = r\sin\varphi$, переводящее полярные координаты (r, φ) в декартовы (x, y). Приращению $(\Delta r, \Delta\varphi)$ переменных r и φ в декартовых координатах (x, y) соответствует приращение Δr от начала координат по лучу l, расположенному под углом φ к оси Ox, и поворот на угол $\Delta\varphi$ по окружности радиуса r с центром в нуле, т. е. смещение по дуге длины $r\Delta\varphi$. При малом $\Delta\varphi$ можно заменить движение по дуге окружности смещением перпендикулярно радиусу по прямой на $r\Delta\varphi$. Итак, в декартовой системе координат, в которой одна ось направлена вдоль луча l, а другая ось направлена перпендикулярно l, получили смещение на вектор $(\Delta r, r\Delta\varphi)$. Для переходя в исходную систему координат надо сделать поворот на угол φ. Таким образом, смещение в декартовой системе координат $(\Delta x, \Delta y)$ выражается через смещение $(\Delta r, \Delta\varphi)$ в полярной системе координат следующим образом:

$$\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} \Delta r \\ r\Delta\varphi \end{pmatrix} = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix} \begin{pmatrix} \Delta r \\ \Delta\varphi \end{pmatrix}.$$

Можно прямым вычислением показать, что

$$\begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$

действительно является матрицей Якоби отображения, переводящего полярные координаты в декартовы.

Правила дифференцирования. Дифференцирование сложной функции

Теорема 21. *Пусть $f, g\colon X \to Y$ и $\lambda\colon X \to \mathbb{R}$ — дифференцируемые отображения в точке a. Тогда*

(i) *$f+g$ дифференцируемо в точке a и $d(f+g) = df + dg$;*

(ii) *λf дифференцируемо в точке a и $d(\lambda f) = \lambda(df) + (d\lambda)f$;*

(iii) *если $\lambda(a) \neq 0$, то отображение $\lambda^{-1}f$ дифференцируемо в точке a и*

$$d\Big(\lambda^{-1}f\Big) = \lambda^{-2}\Big(\lambda(df) - (d\lambda)f\Big).$$

$\rhd$ Докажем только второе утверждение. По условию справедливы равенства

$$\lambda(a+h) - \lambda(a) = d\lambda(h) + \gamma_1(h)\|h\|, \quad f(a+h) - f(a) = df(h) + \gamma_2(h)\|h\|,$$

где $\lim\limits_{h\to 0}\gamma_1(h) = 0$ и $\lim\limits_{h\to 0}\gamma_2(h) = 0$. Получаем

$$\lambda(a+h)f(a+h) - \lambda(a)f(a) = \lambda(a+h)(f(a+h) - f(a)) + (\lambda(a+h) - \lambda(a))f(a) = \\ = \lambda(a)df(h) + d\lambda(h)f(a) + r(h)\|h\|,$$

где

$$r(h) = (\lambda(a+h) - \lambda(a))df(h)\|h\|^{-1} + \lambda(a+h)\gamma_2(h) + \gamma_1(h)f(a).$$

Заметим, что в силу непрерывности линейного отображения $df(h)$ существует такое число $C > 0$, что $\|df(h)\|\|h\|^{-1} \leqslant C$. Следовательно, $\lim\limits_{h\to 0} r(h) = 0$. $\lhd$

Теорема 22. *Пусть отображение $g\colon Z \to X$ дифференцируемо в точке a и отображение $f\colon X \to Y$ дифференцируемо в точке $g(a)$. Тогда композиция $f \circ g$ дифференцируема в точке a и $d(f\circ g) = df \circ dg$.*

$\rhd$ По условию справедливы равенства

$$g(a+h) - g(a) = dg(h) + \gamma_1(h)\|h\|, \quad f(g(a)+p) - f(g(a)) = df(p) + \gamma_2(p)\|p\|,$$

где $\lim\limits_{h\to 0}\gamma_1(h) = 0$ и $\lim\limits_{h\to 0}\gamma_2(h) = 0$. Определим отображение γ_2 в нуле нулём; теперь γ_2 непрерывно в нуле. Имеем

$$f(g(a+h)) - f(g(a)) = df(g(a+h) - g(a)) + \gamma_2(g(a+h) - g(a))\|g(a+h) - g(a)\| = df(dg(h)) + r(h)\|h\|,$$

где

$$r(h) = df(\gamma_1(h)) + \gamma_2(g(a+h) - g(a))\|g(a+h) - g(a)\|\|h\|^{-1}.$$

Существуют такие числа C_1 и C_2, что $\|dg(h)\| \leqslant C_1\|h\|$ и $\|df(p)\| \leqslant C_2\|p\|$. Тогда

$$\|df(\gamma_1(h))\| \leqslant C_2\|\gamma_1(h)\| \to 0,$$

$$\|\gamma_2(g(a+h) - g(a))\|\|g(a+h) - g(a)\|\|h\|^{-1} \leqslant (C_1 + \|\gamma_1(h)\|)\|\gamma_2(g(a+h) - g(a))\| \to 0$$

при $h \to 0$. Получаем равенство $\lim\limits_{h\to 0} r(h) = 0$, тем самым, для $f \circ g$ выполнено определение дифференцируемости в точке a. $\lhd$

Пусть $g\colon \mathbb{R}^n \to \mathbb{R}^m$ и $f\colon \mathbb{R}^m \to \mathbb{R}^k$ — дифференцируемые отображения. Тогда их композиция дифференцируема и дифференциал композиции является композицией дифференциалов. Следовательно, матрица Якоби композиции является произведением матриц Якоби.

Пусть $n = 1$ и $k = 1$, т. е. композиция является функцией вида $f(g_1(t), g_2(t), \ldots, g_m(t))$.

Матрица Якоби отображения g имеет вид

$$\begin{pmatrix} \frac{dg_1}{dt} \\ \frac{dg_2}{dt} \\ \ldots \\ \frac{dg_m}{dt} \end{pmatrix},$$

матрица Якоби отображения f имеет вид

$$\Big(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_m}\Big).$$

Следовательно,

$$\frac{d}{dt}f\big(g_1(t),g_2(t),\dots,g_m(t)\big)=$$

$$=\Big(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\dots,\frac{\partial f}{\partial x_m}\Big)\left(\begin{array}{c}\frac{dg_1}{dt}\\ \frac{dg_2}{dt}\\ \dots\\ \frac{dg_m}{dt}\end{array}\right)=$$

$$=\frac{\partial f}{\partial x_1}\cdot\frac{dg_1}{dt}+\frac{\partial f}{\partial x_2}\cdot\frac{dg_2}{dt}+\dots+\frac{\partial f}{\partial x_m}\cdot\frac{dg_m}{dt}.$$

Полезно провести отдельное доказательство в этом частном случае. Итак, пусть f — дифференцируемая функция m переменных и заданы m дифференцируемых функций g_k одной переменной t. Найдём по определению производную сложной функции $f(g(t))=f\big(g_1(t),g_2(t),\dots,g_m(t)\big)$. Имеем

$$\frac{d}{dt}f\big(g_1(t),g_2(t),\dots,g_m(t)\big)=\lim_{\tau\to 0}\frac{f(g_1(t+\tau),\dots,g_n(t+\tau))-f(g_1(t),\dots,g_n(t))}{\tau}.$$

Функция f дифференцируема, поэтому

$$f(g_1(t+\tau),\dots,g_n(t+\tau))-f(g_1(t),\dots,g_n(t))=$$

$$=\frac{\partial f}{\partial x_1}(g(t))(g_1(t+\tau)-g_1(t))+\dots+\frac{\partial f}{\partial x_n}(g(t))(g_m(t+\tau)-g_m(t))+\alpha(g(t+\tau)-g(t))\|g(t+\tau)-g(t)\|.$$

Здесь мы, как и выше при доказательстве теоремы, положили $\alpha(0)=0$. Заметим, что $\|g(t+\tau)-g(t)\|/\tau$ — ограниченная величина при $\tau\to 0$. Получаем

$$\frac{d}{dt}f\big(g_1(t),g_2(t),\dots,g_m(t)\big)=$$

$$=\frac{\partial f}{\partial x_1}(g(t))\lim_{\tau\to 0}\frac{g_1(t+\tau)-g_1(t)}{\tau}+\dots+\frac{\partial f}{\partial x_n}(g(t))\lim_{\tau\to 0}\frac{g_m(t+\tau)-g_m(t)}{\tau}=$$

$$=\frac{\partial f}{\partial x_1}\cdot\frac{dg_1}{dt}+\frac{\partial f}{\partial x_2}\cdot\frac{dg_2}{dt}+\dots+\frac{\partial f}{\partial x_m}\cdot\frac{dg_m}{dt}.$$

Теперь рассмотрим более общую ситуацию, когда только $k=1$. Матрица Якоби композиции получается умножением строки

$$\Big(\frac{\partial f}{\partial y_1},\frac{\partial f}{\partial y_2},\dots,\frac{\partial f}{\partial y_m}\Big)$$

на матрицу

$$\left(\begin{array}{cccc}\frac{\partial g_1}{\partial x_1}&\frac{\partial g_1}{\partial x_2}&\dots&\frac{\partial g_1}{\partial x_n}\\ \frac{\partial g_2}{\partial x_1}&\frac{\partial g_2}{\partial x_2}&\dots&\frac{\partial g_2}{\partial x_n}\\ \dots&\dots&\dots&\dots\\ \frac{\partial g_m}{\partial x_1}&\frac{\partial g_m}{\partial x_2}&\dots&\frac{\partial g_m}{\partial x_n}\end{array}\right).$$

Выпишем элемент матрицы Якоби композиции (в данном случае — элементы вектора градиента):

$$\frac{\partial f(g(x))}{\partial x_i}=\frac{\partial f}{\partial y_1}\cdot\frac{\partial g_1}{\partial x_i}+\frac{\partial f}{\partial y_2}\cdot\frac{\partial g_2}{\partial x_i}+\dots+\frac{\partial f}{\partial y_m}\cdot\frac{\partial g_m}{\partial x_i}.$$

Заметим, что этот результат можно формально получить из выражения

$$df=\frac{\partial f}{\partial y_1}dy_1+\frac{\partial f}{\partial y_2}dy_2+\dots+\frac{\partial f}{\partial y_m}dy_m,$$

подставляя вместо y_k функцию g_k и вычислив дифференциалы dg_k. Это наблюдение называют *инвариантностью первого дифференциала.*

Теорема 23. *Пусть $f\colon X\to Y$, U — окрестность точки a в X, V — окрестность точки $f(a)$ в Y. Предположим, что $f\colon U\to V$ гомеоморфизм, f дифференцируемо в точке a и df имеет непрерывное обратное отображение. Тогда $f^{-1}\colon V\to U$ дифференцируемо в точке $f(a)$ и $d(f^{-1})=(df)^{-1}$.*

$\triangleright$ Найдём предел

$$\lim_{p\to 0}\frac{\|f^{-1}(f(a)+p)-f^{-1}(f(a))-(df)^{-1}(p)\|}{\|p\|}.$$

Сделаем замену переменных $h = f^{-1}(f(a)+p) - a$. Тогда $h \to 0$ при $p \to 0$. Имеем
$$p = f(a+h) - f(a) = df(h) + \gamma(h)\|h\|, \quad \lim_{h\to 0}\gamma(h) = 0.$$
Получаем
$$\frac{\|f^{-1}(f(a)+p) - f^{-1}(f(a)) - (df)^{-1}(p)\|}{\|p\|} = \frac{\|h\|\|(df)^{-1}(\gamma(h))\|}{\|df(h)+\gamma(h)\|h\|\|} \leqslant \frac{\|h\|\|(df)^{-1}(\gamma(h))\|}{\|df(h)\| - \|\gamma(h)\|\|h\|}.$$
Существует такое число $C > 0$, что $\|df(h)\| \geqslant C\|h\|$. Следовательно,
$$\frac{\|f^{-1}(f(a)+p) - f^{-1}(f(a)) - (df)^{-1}(p)\|}{\|p\|} \leqslant \frac{\|(df)^{-1}(\gamma(h))\|}{C - \|\gamma(h)\|} \to 0, \quad h \to 0. \qquad \lhd$$

Отметим, что матрица Якоби обратного отображения является обратной матрицей к матрице Якоби прямого отображения.

В заключение приведём ещё одно полезное утверждение, обобщающее теорему Лагранжа на многомерный случай.

Предложение 29. *Пусть отображение $f\colon \mathbb{R}^n \to \mathbb{R}^m$ дифференцируемо в шаре $B_r(a)$ и все частные производные $\frac{\partial f_i}{\partial x_j}$ ограниченны на этом шаре. Положим*
$$M = \sup_{x\in B_r(a)} \|df(x)\|.$$
Тогда для всяких $x, y \in B_r(a)$ выполняется неравенство
$$\|f(x) - f(y)\| \leqslant M\|x-y\|.$$

$\rhd$ Рассмотрим функцию
$$\varphi(t) = \langle f(x+t(y-x)), f(y)-f(x)\rangle.$$
Тогда
$$\varphi'(t) = \langle df(x+t(y-x), y-x), f(y)-f(x)\rangle$$
и по определению нормы линейного отображения и неравенству Коши — Буняковского — Шварца имеет место оценка:
$$|\varphi'(t)| \leqslant \|df(x+t(y-x))\|\|y-x\|\|f(y)-f(x)\|.$$
По теореме Лагранжа $\varphi(1) - \varphi(0) = \varphi'(c)$. Получаем
$$\|f(y)-f(x)\|^2 = \varphi(1)-\varphi(0) = \varphi'(c) \leqslant \|df(x+c(y-x))\|\|y-x\|\|f(y)-f(x)\|.$$
Если $\|f(y)-f(x)\| > 0$, то остаётся разделить обе части неравенства на $\|f(y)-f(x)\|$ и оценить $\|df(x+c(y-x))\|$ константой M. Если же $\|f(y)-f(x)\| = 0$, то доказываемое неравенство выполняется немедленно. $\lhd$

Теорема об обратной функции и теорема о неявной функции

Пусть задана функция $f\colon \mathbb{R} \to \mathbb{R}$. Если на некотором промежутке I_x эта функция непрерывна и строго монотонна, то $I_y = f(I_x)$ является промежутком, $f\colon I_x \to I_y$ является биекцией и обратная функция f^{-1} непрерывна. В частности, если в окрестности некоторой точки x_0 функция f непрерывно дифференцируема и $f'(x_0) \neq 0$, то существуют такие окрестности $U(x_0)$ и $V(f(x_0))$, что $f\colon U(x_0) \to V(x_0)$ — гомеоморфизм и, более того, обратная функция f^{-1} непрерывно дифференцируема. В настоящем разделе мы обобщим последнее утверждение на случай отображения из $\mathbb{R}^n$ в $\mathbb{R}^n$. Однако соображения монотонности хорошо работают для функций одной вещественной переменной, но плохо переносятся на многомерную ситуацию. Поэтому, оставаясь пока на числовой прямой, рассмотрим другой метод построения обратной функции.

Итак, пусть $f\colon \mathbb{R} \to \mathbb{R}$ непрерывно дифференцируема в окрестности точки x_0 и $f'(x_0) \neq 0$. Далее мы всегда будем рассматривать только такую окрестность точки x_0, в которой производная f' отлична от нуля. Положим $y_0 = f(x_0)$. Для построения обратной функции фактически надо научиться решать уравнение $y = f(x)$ относительно переменной x. Перепишем это уравнение в новом виде:
$$x = x + q(y - f(x)),$$

где q некоторое отличное от нуля число, которое мы укажем позднее. Равенство $x = x+q(y-f(x))$ выполнено тогда и только тогда, когда $y = f(x)$. Положим

$$G_y(x) = x + q(y - f(x)).$$

При каждом фиксированном y это функция от x и мы ищем неподвижную точку x, для которой выполнено равенство $x = G_y(x)$. Отметим, что очень важно не только доказать существование неподвижной точки, но и её единственность.

Следующее утверждение даёт простое достаточное условие существования и единственности неподвижной точки.

Теорема 24 (Банах). *Пусть (X, ρ) — полное метрическое пространство и $F\colon X \to X$ является сжимающим отображением, т. е. выполняется неравенство*

$$\rho(F(x), F(y)) \leqslant \lambda\rho(x, y), \quad 0 < \lambda < 1.$$

Тогда существует единственная такая точка x, что $F(x) = x$.

$\triangleright$ Пусть x_0 — произвольная точка пространства (X, ρ). Положим $x_{n+1} = F(x_n)$. Поскольку

$$\rho(x_{n+1}, x_n) \leqslant \lambda\rho(x_n, x_{n-1}) \leqslant \ldots \leqslant \lambda^n\rho(x_1, x_0),$$

для $n > m$ справедливо неравенство

$$\rho(x_n, x_m) \leqslant \frac{\lambda^m}{1-\lambda}\rho(x_1, x_0).$$

Следовательно, x_n является фундаментальной последовательностью и имеет предел. Пусть $x_n \to x$. В силу непрерывности F заключаем, что $F(x) = x$. Предположим, что кроме x есть другая неподвижная точка z. Тогда

$$\rho(x, z) = \rho(F(x), F(z)) \leqslant \lambda\rho(x, z)$$

и, следовательно, $x = z$. $\triangleleft$

Попробуем так выбрать отрезок $I_x = [x_0-\alpha; x_0+\alpha]$ и интервал $I_y = (y_0-\beta; y_0+\beta)$, чтобы для всякого $y \in I_y$ функция G_y являлась сжимающим отображением I_x в I_x. По теореме Лагранжа

$$|G_y(x) - G_y(z)| = \left|\frac{dG_y}{dx}(c)\right| |x - z|,$$

где

$$\frac{dG_y}{dx}(c) = 1 - qf'(c).$$

Положим $q = 1/f'(x_0)$. Тогда $\frac{dG_y}{dx}(x_0) = 0$. По условию f' — непрерывная функция в окрестности x_0. Выбирая α достаточно малым, можно считать, что на I_x выполняется неравенство $|\frac{dG_y}{dx}(x)| = |1-qf'(x)| \leqslant 1/2$, поэтому $|G_y(x)-G_y(z)| \leqslant \frac{1}{2}|x-z|$. Теперь осталось сделать так, чтобы значения G_y лежали в I_x. Заметим, что $G_{y_0}(x_0) = x_0$. Имеем

$$|G_y(x) - x_0| \leqslant |G_y(x) - G_y(x_0)| + |G_y(x_0) - G_{y_0}(x_0)| \leqslant \frac{1}{2}|x - x_0| + |q||y - y_0| \leqslant \frac{1}{2}\alpha + |q|\beta.$$

Выберем β столь малым, что $|q|\beta < \alpha/4$. Итак, мы выбрали I_x и I_y так, что $G_y\colon I_x \to I_x$ является сжимающим отображением для всякого $y \in I_y$. Более того, $G_y(I_x) \subset (x_0 - 3\alpha/4; x_0 + 3\alpha/4) \subset (x_0 - \alpha; x_0 + \alpha)$. По теореме Банаха существует единственное такое $x \in I_x$, что $G_y(x) = x$. Положим $V(y_0) = I_y$ и $U(x_0) = f^{-1}(V(y_0)) \cap (x_0 - 3\alpha/4; x_0 + 3\alpha/4)$. Тогда $U(x_0)$ и $V(y_0)$ — открытые множества, содержащие x_0 и y_0 соответственно. Для всякого $y \in V(y_0)$ существует единственный такой $x \in U(x_0)$, что $G_y(x) = x$ и, значит, $y = f(x)$. Следовательно, $f\colon U(x_0) \to V(y_0)$ — биекция, и существует обратная функция $f^{-1}\colon V(y_0) \to U(x_0)$. Однако ещё надо обосновать непрерывность и дифференцируемость обратной функции f^{-1}.

Предложение 30. *Пусть (X, ρ) — полное метрическое пространство и $F_1, F_2\colon X \to X$ — сжимающие отображения с коэффициентом сжатия $\lambda \in (0; 1)$. По теореме Банаха существуют единственные неподвижные точки $x_1 = F_1(x_1)$ и $x_2 = F_2(x_2)$. Тогда имеет место оценка*

$$\rho(x_1, x_2) \leqslant (1 - \lambda)^{-1}\rho(F_1(x_1), F_2(x_1)).$$

$\triangleright$ Достаточно заметить, что справедливо неравенство

$$\rho(x_1, x_2) = \rho(F_1(x_1), F_2(x_2)) \leqslant \rho(F_1(x_1), F_2(x_1)) + \lambda\rho(x_1, x_2). \quad \triangleleft$$

Пусть $y_1, y_2 \in V(y_0)$. Поскольку

$$|G_{y_1}(x) - G_{y_2}(x)| = |q||y_1 - y_2|,$$

для соответствующих неподвижных точек $x_1 = f^{-1}(y_1)$ и $x_2 = f^{-1}(y_2)$ выполняется оценка

$$|x_1 - x_2| \leqslant (1-\lambda)^{-1}|q||y_1 - y_2|, \text{ где } \lambda = \frac{1}{2}.$$

Таким образом, обратная функция является не просто непрерывной, а даже липшицевой. Осталось показать, что обратная функция дифференцируема, но это следует из доказанной выше теоремы о дифференцируемости обратной функции.

Данный метод построения обратной функции практически без изменений переносится на многомерный случай.

Теорема 25 (об обратной функции)**.** *Пусть отображение $f\colon \mathbb{R}^n \to \mathbb{R}^n$ непрерывно дифференцируемо в окрестности точки x_0 и $y_0 = f(x_0)$. Если матрица Якоби $J_f(x_0)$ обратима, то существуют такие окрестности $U(x_0)$ и $V(y_0)$ точек x_0 и y_0 соответственно, что отображение $f\colon V(y_0) \to U(x_0)$ обратимо и обратное отображение f^{-1} непрерывно дифференцируемо в каждой точке окрестности $U(x_0)$.*

$\triangleright$ Будем действовать так же, как и выше. Рассмотрим отображение

$$G_y(x) = x + Q(y - f(x)), \quad \text{где} \quad Q = J_f(x_0)^{-1}.$$

Построим такие шары $\overline{B}_\alpha(x_0) = \{x\colon \|x - x_0\| \leqslant \alpha\}$ и $B_\beta(y_0) = \{y\colon \|y - y_0\| < \beta\}$, что для всякого $y \in B_\beta(y_0)$ отображение $G_y\colon \overline{B}_\alpha(x_0) \to \overline{B}_\alpha(x_0)$ является сжимающим, причём

$$G_y(\overline{B}_\alpha(x_0)) \subset \{x\colon \|x - x_0\| < 3\alpha/4\}.$$

Выберем такое $\alpha > 0$, что на шаре $B_{2\alpha}(x_0)$ выполняется неравенство $\|\frac{dG_y}{dx}(x)\| \leqslant 1/2$. Следовательно, для всех $x, z \in \overline{B}_\alpha(x_0)$ получаем

$$\|G_y(x) - G_y(z)\| \leqslant \frac{1}{2}\|x - z\|.$$

Теперь сделаем так, чтобы G_y принимало значения в шаре $\{x\colon \|x - x_0\| < 3\alpha/4\}$. Тогда

$$\|G_y(x) - x_0\| \leqslant \frac{1}{2}\|x - x_0\| + \|Q\|\|y - y_0\| \leqslant \frac{1}{2}\alpha + \|Q\|\beta.$$

Выберем $\beta > 0$ так, что $\|Q\|\beta < \alpha/4$. Применим теорему Банаха: для всякого $y \in B_\beta(y_0)$ существует единственный такой $x \in \overline{B}_\alpha(x_0)$, что $G_y(x) = x$ и, значит, $y = f(x)$. Более того, неподвижная точка x принадлежит шару $\{x\colon \|x - x_0\| < 3\alpha/4\}$. Положим

$$V(y_0) = B_\beta(y_0) \quad \text{и} \quad U(x_0) = f^{-1}(V(y_0)) \cap \{x\colon \|x - x_0\| < 3\alpha/4\}.$$

Как и выше, можно показать, что $f\colon U(x_0) \to V(y_0)$ является биекцией и что обратная функция непрерывно дифференцируема. $\triangleleft$

Взаимно однозначное непрерывно дифференцируемое отображение, которое имеет непрерывно дифференцируемое обратное, называется *диффеоморфизмом.* Если задан диффеоморфизм $f\colon U \to V$, то задана новая система координат. Действительно, всякой точке $x \in U$ однозначно сопоставляется $y \in V$, и наоборот. Кроме того, переход к новым координатам сохраняет свойство непрерывной дифференцируемости: если на множестве U определена непрерывно дифференцируемая функция g, то на V ей соответствует непрерывно дифференцируемая функция $g(f^{-1}(y))$. Выбирая подходящую систему координат часто можно значительно упростить решаемую задачу.

Следствие 8 (теорема о неявной функции)**.** *Пусть отображение $F\colon \mathbb{R}^n_x \times \mathbb{R}^m_y \to \mathbb{R}^m$ непрерывно дифференцируемо в окрестности точки (x_0, y_0), $F(x_0, y_0) = 0$ и матрица Якоби отображения $y \to F(x_0, y)$ обратима в точке y_0. Тогда существуют такие окрестности $U(x_0)$ и $V(y_0)$ и непрерывно дифференцируемое отображение $f\colon U(x_0) \to V(y_0)$, что на $U(x_0) \times V(y_0)$*

$$F(x, y) = 0 \Leftrightarrow y = f(x).$$

$\triangleright$ Рассмотрим отображение $H\colon \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$, переводящее переменные (x, y) в переменные (u, v) следующим образом: $u_i = x_i$ при $i = 1, 2, \ldots, n$ и $v_j = F^{j-n}(x, y)$ при $j = n+1, \ldots, n+m$. Заметим, что точка (x_0, y_0) переходит в точку $(x_0, 0)$, и матрица Якоби отображения H имеет вид

$$\begin{pmatrix} E & 0 \\ * & D_yF \end{pmatrix}.$$

Матрица $D_yF(x_0,y_0)$ обратима, поэтому и матрица Якоби отображения H в точке (x_0,y_0) обратима. По теореме об обратной функции существуют такие окрестности $W((x_0,y_0))$ и $Q((x_0,0))$, что $H\colon W\to Q$ является диффеоморфизмом. Пусть открытые шары $B(x_0)$ и $B(y_0)$ выбраны так, что $B(x_0)\times B(y_0)\subset W$. Положим $U(x_0)=\{x\colon (x,0)\in H(B(x_0)\times B(y_0))\}$ и $V(y_0)=B(y_0)$. Отображение H отображает множество $\{(x,y)\in U(x_0)\times V(y_0)\colon F(x,y)=0\}$ на множество $\{(u,0)\colon u\in U(x_0)\}$. Обратное отображение имеет вид $x_i=u_i$ при $i=1,2,\ldots,n$ и $y_j=g^j(u,v)$ при $j=n+1,\ldots m$. Положим $f(x)=g(x,0)$ при $x\in U(x_0)$. Тогда $f\colon U(x_0)\to V(y_0)$ и на $U(x_0)\times V(y_0)$ равенство $F(x,y)=0$ равносильно равенству $y=f(x)$. $\lhd$

Используемый в теореме о неявной функции диффеоморфизм H значительно упрощает отображение F. В новых координатах (u,v) функция F имеет вид $F(x(u,v),y(u,v))=v$. Сформулируем в качестве ещё одного следствия теоремы об обратной функции частный случай этого наблюдения.

Следствие 9. *Пусть функция $f\colon\mathbb{R}^n\to\mathbb{R}$ непрерывно дифференцируема в окрестности нуля. Если $f(0)=0$ и $\nabla f(0)\neq 0$, то существуют такие окрестности $U(0)$ и $V(0)$ и диффеоморфизм $x=x(v)$, отображающий $V(0)\to U(0)$, что $f(x(v))=v_n$.*

$\rhd$ Пусть $\frac{\partial f}{\partial x_n}(0)\neq 0$ (если иначе, то сделаем линейное преобразование, переставляющее координаты). Искомый диффеоморфизм задаётся равенствами

$$v_1=x_1,\quad \ldots,\quad v_{n-1}=x_{n-1},\quad v_n=f(x). \qquad \lhd$$

Таким образом, для всякой непрерывно дифференцируемой функции в некоторой окрестности точки, в которой градиент функции отличен от нуля, можно ввести такую систему координат, в которой функция станет простейшей линейной или аффинной функцией. При изучении функции нескольких переменных существенную роль играют *множества уровня*

$$\{x\colon f(x)=\mathrm{const}\},$$

изображение которых является аналогом построения графика функции в одномерном случае. Если градиент функции отличен от нуля, то локально в подходящей системе координат функция имеет вид $f(v)=v_n+\mathrm{const}$ и её множества уровня — гиперплоскости. Теорема о неявной функции утверждает, что в исходных координатах множество уровня такой функции является графиком функции. Для осознания глубины этих утверждений полезно иметь в виду, что множеством уровня непрерывно дифференцируемой функции может быть произвольное замкнутое множество.

Предложение 31. *Для всякого замкнутого множества $A\subset\mathbb{R}^n$ существует такая непрерывно дифференцируемая функция $f\colon\mathbb{R}^n\to\mathbb{R}$, что $A=\{x\colon f(x)=0\}$.*

$\rhd$ Дополнение к A является открытым множеством и представляется в виде объединения не более чем счётного набора открытых шаров $B_{r_j}(a_j)$. Пусть функция ψ непрерывно дифференцируема, строго положительна на открытом шаре $B_1(0)$ и равна нулю вне $B_1(0)$. В качестве такой функции можно взять $\psi(x)=e^{1/(|x|^2-1)}$ при $|x|<1$ и $\psi(x)=0$ при $|x|\geqslant 0$. Положим $\psi_j(x)=c_j\psi((x-a_j)/r_j)$, где c_j выбраны так, что $|\psi_j|+\|\nabla\psi_j\|\leqslant 1/2^j$. Функция $f(x)=\sum\limits_j\psi_j(x)$ искомая. Действительно, данный ряд и ряды из производных $\sum\limits_j\frac{\partial\psi_j}{\partial x_k}$ сходятся равномерно, поэтому функция f непрерывно дифференцируема. Остаётся заметить, что для всякого x из дополнения к A хотя бы одна из функций $\psi_j(x)$ и $f(x)$ положительна, а для всякого x из A все $\psi_j(x)=0$ и $f(x)=0$. $\lhd$

С теоремой о неявной функции тесно связан важный геометрический объект — гладкая поверхность.

Гладкая поверхность и касательное пространство.

Обсудим сначала двумерные поверхности в $\mathbb{R}^3$ и начнём с нескольких простых примеров.

1. *График функции.* Пусть $f\colon\mathbb{R}^2\to\mathbb{R}$. Тогда график функции $z=f(x,y)$ естественно назвать двумерной поверхностью. С одной стороны, график является множеством точек $(x,y,f(x,y))$, заданных параметрически с помощью двух параметров x и y, а с другой стороны, график — множество решений одного уравнения $z-f(x,y)=0$.

2. *Двумерная плоскость.* Пусть заданы два линейно независимых вектора $a=(a_1,a_2,a_3)$ и $b=(b_1,b_2,b_3)$. Тогда плоскость, натянутая на эти векторы и проходящая через точку (x_0,y_0,z_0),

задаётся параметрически так:

$$x(s,t) = x_0 + a_1 s + b_1 t, \quad y(s,t) = y_0 + a_2 s + b_2 t, \quad z(s,t) = z_0 + a_3 s + b_3 t,$$

где $(s,t) \in \mathbb{R}^2$. Можно эту же самую плоскость задать уравнением

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0,$$

где вектор (A, B, C) является векторным произведением векторов a и b.

3. *Двумерная сфера единичного радиуса.* Это множество задаётся уравнением $x^2 + y^2 + z^2 = 1$. Сферу можно задать параметрически:

$$x(\varphi, \psi) = \cos\varphi \cos\psi, \quad y(\varphi, \psi) = \sin\varphi \cos\psi, \quad z(\varphi, \psi) = \sin\psi,$$

где точки (φ, ψ) принадлежат прямоугольнику $[-\pi; \pi] \times [-\pi/2; \pi/2]$ (отметим, что эта параметризация не является взаимно-однозначным соответствием между точками прямоугольника и сферы).

4. *Двумерный тор.* Пусть $0 < r < R$ и $(\varphi, \psi) \in [0; 2\pi] \times [0; 2\pi]$. Тогда тор — поверхность, задаваемая параметрически так:

$$x(\varphi, \psi) = (R + r\cos\psi)\cos\varphi, \quad y(\varphi, \psi) = (R + r\cos\psi)\sin\varphi, \quad z(\varphi, \psi) = r\sin\psi.$$

Как и сферу, тор можно задать уравнением.

Естественно определить двумерную поверхность как множество, которое в окрестности всякой своей точки задаётся параметрически с помощью двух параметров. Эти параметры являются координатами на поверхности, т. е. по значению параметров однозначно определяется точка поверхности и наоборот.

Скажем, что множество M_2 является *двумерной поверхностью* в $\mathbb{R}^3$, если для всякой точки $p \in M_2$ существуют такие окрестность $U(p)$, открытое множество $V \subset \mathbb{R}^2$ и инъективное непрерывно дифференцируемое отображение $f\colon V \to U(p)$, что $M_2 \bigcap U(p) = f(V)$ и ранг матрицы Якоби J_f равен 2 в каждой точке множества V. Если f задано функциями $x = x(v_1, v_2), y = y(v_1, v_2), z = z(v_1, v_2)$, то предполагается, что эти функции непрерывно дифференцируемы, причём

$$\operatorname{rk} \begin{pmatrix} \frac{\partial x}{\partial v_1} & \frac{\partial y}{\partial v_1} & \frac{\partial z}{\partial v_1} \\ \frac{\partial x}{\partial v_2} & \frac{\partial y}{\partial v_2} & \frac{\partial z}{\partial v_2} \end{pmatrix} = 2.$$

Обсудим последнее условие подробнее. Предположим, что в точке p

$$\begin{vmatrix} \frac{\partial x}{\partial v_1} & \frac{\partial y}{\partial v_1} \\ \frac{\partial x}{\partial v_2} & \frac{\partial y}{\partial v_2} \end{vmatrix} \neq 0.$$

По теореме о неявной функции (или даже по теореме об обратной функции) в некоторой окрестности точки p можно выразить v_1 и v_2 через x и y, т. е. $v_1 = v_1(x, y)$ и $v_2 = v_2(x, y)$. Тогда $z = z(v_1, v_2) = z(v_1(x, y), v_2(x, y)) = g(x, y)$ и поверхность является графиком функции $z = g(x, y)$.

Обобщая понятие двумерной поверхности, можно дать следующее определение.

Множество M_k в $\mathbb{R}^n$ называется *гладкой k-мерной поверхностью*, если для всякой точки $p \in M_k$ существуют такие окрестность $U(p)$ точки p в $\mathbb{R}^n$, открытое непустое множество $V \subset \mathbb{R}^k$ и непрерывно дифференцируемое *инъективное* отображение $f\colon V \to \mathbb{R}^n$, что $f(V) = M_k \bigcap U(p)$ и ранг матрицы Якоби $J_f(u)$ в каждой точке $u \in V$ равен k.

Предложение 32. *Следующие утверждения равносильны*:

(i) *M_k — гладкая k-мерная поверхность в $\mathbb{R}^n$;*

(ii) *для всякой точки $p \in M_k$ существуют окрестность $U(p)$ и такое непрерывно дифференцируемое отображение $F\colon \mathbb{R}^n \to \mathbb{R}^{n-k}$, что ранг матрицы Якоби $J_F(p)$ равен $(n-k)$ и M_k совпадает с множеством $F(x) = 0$ в $U(p)$.*

$\triangleright$ Утверждение следует из теоремы о неявной функции. $\triangleleft$

Отметим, что (ii) можно переформулировать следующим образом: в некоторой окрестности $U(p)$ каждой точки $p \in M_k$ можно так ввести координаты, что в новых координатах множество $M_k \cap U(p)$ задаётся уравнениями $x_{k+1} = 0, \ldots, x_n = 0$, т. е. поверхность в этих координатах является частью гиперплоскости, соответствующей координатам $x_1, \ldots, x_k$. Эти координаты и можно взять в качестве параметров, параметризующих поверхность M_k в окрестности точки p.

Отображение $\gamma\colon (-1;1) \to \mathbb{R}^n$ называется *гладкой кривой*, если $\gamma(t) = (x_1(t), \ldots, x_n(t))$, где $x_i(t)$ — непрерывно дифференцируемые функции на $(-1;1)$. Пусть M_k — гладкая k-мерная поверхность в $\mathbb{R}^n$ и γ — такая гладкая кривая, что $\gamma(t) \in M_k$ для всякого t и $\gamma(0) = p$. Отметим, что если в окрестности точки p поверхность M_k задана отображением $f : V \to \mathbb{R}^n$, где ранг матрицы Якоби отображения f равен k, то можно считать, что $\gamma(t) = f(u(t))$, где $u(t)$ — гладкая кривая в V. Вектор скорости $\dot\gamma(0)$ называется *касательным вектором* к поверхности M_k в точке p.

Предложение 33. *Множество касательных векторов к гладкой k-мерной поверхности M_k в точке p является k-мерным линейным пространством.*

$\triangleright$ Пусть поверхность M_k в окрестности точки p задана отображениями $x_i = f_i(u_1, \ldots, u_k)$. Тогда всякая кривая $\gamma(t)$ на M_k задаётся так: $x_i(t) = f_i(u_1(t), \ldots, u_k(t))$. Найдём $\dot\gamma(0)$:

$$\dot\gamma(0) = \dot{u_1}(0)\frac{\partial f}{\partial u_1}(0) + \ldots + \dot{u_k}(0)\frac{\partial f}{\partial u_k}(0).$$

Таким образом, всякий касательный вектор является линейной комбинацией линейно независимых касательных векторов

$$\frac{\partial f}{\partial u_1}, \quad \frac{\partial f}{\partial u_2}, \quad \ldots, \quad \frac{\partial f}{\partial u_k}.$$

Можно показать, что всякая линейная комбинация этих векторов является касательным вектором. $\triangleleft$

Линейное пространство касательных векторов в точке p называется *касательным пространством* к M_k в точке p и обозначается через T_pM_k.

Предложение 34. *Пусть для точки $p \in M_k$ существуют окрестность $U(p)$ и такое непрерывно дифференцируемое отображение $F\colon \mathbb{R}^n \to \mathbb{R}^{n-k}$, что ранг матрицы Якоби $J_F(p)$ равен $(n-k)$ и M_k совпадает с множеством $\{x\colon F(x) = 0\}$ в $U(p)$. Тогда*

$$v \in T_pM_k \Leftrightarrow \begin{cases} \langle \nabla F_1, v\rangle = 0, \\ \langle \nabla F_2, v\rangle = 0, \\ \ldots \\ \langle \nabla F_{n-k}, v\rangle = 0. \end{cases}$$

$\triangleright$ Необходимость этих условий вытекает из дифференцирования равенств $F_i(\gamma(t)) = 0$. Следовательно, касательное пространство является подпространством решений этой однородной системы. Размерность касательного пространства равна k и размерность пространства решений этой системы равна k. Значит, эти пространства совпадают. $\triangleleft$

В частности, если поверхность является множеством уровня $\{x\colon f(x) = \mathrm{const}\}$ некоторой функции $f\colon \mathbb{R}^n \to \mathbb{R}$, то её касательное пространство ортогонально вектору градиента ∇f. Таким образом, в точке x_0 поверхности уровня функции f касательная плоскость задаётся уравнением

$$\langle \nabla f(x_0), x - x_0\rangle = 0.$$

Например, если поверхность является графиком непрерывно дифференцируемой функции $z = f(x,y)$ (т. е. линией уровня $F = 0$, где $F(x,y,z) = z - f(x,y)$), то уравнение касательной плоскости в точке $(x_0, y_0, f(x_0,y_0))$ имеет вид

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

Производные высокого порядка. Формула Тейлора. Экстремум

Пусть $f\colon \mathbb{R}^n \to \mathbb{R}$. Предположим, что у f в каждой точке некоторой окрестности $U(a)$ есть частная производная $\frac{\partial f}{\partial x_i}$ и функция $x \mapsto \frac{\partial f}{\partial x_i}(x)$ в точке a имеет частную производную по x_j. Тогда выражение

$$\frac{\partial}{\partial x_j}\Big(\frac{\partial f}{\partial x_i}(x)\Big)\Big|_{x=a}$$

называется *частной производной второго порядка* по переменным x_i и x_j и обозначается через

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

Аналогично определяются *частные производные высокого порядка*:
$$\frac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\dots\partial x_{i_m}}(a) = \frac{\partial}{\partial x_{i_1}}\Big(\frac{\partial}{\partial x_{i_2}}\Big(\dots\Big(\frac{\partial f}{\partial x_{i_m}}(x)\Big)\Big)\Big)\Big|_{x=a}.$$
Возникает естественный вопрос о зависимости частных производных от последовательности дифференцирования. В общем случае последовательность важна и, вообще говоря, $\frac{\partial^2 f}{\partial x_1\partial x_2} \neq \frac{\partial^2 f}{\partial x_2\partial x_1}$.

Теорема 26 (Шварц)**.** *Если у функции $f\colon \mathbb{R}^2 \to \mathbb{R}$ в некоторой окрестности $U(a)$ существуют смешанные частные производные второго порядка по переменным x и y, которые непрерывны в точке a, то*
$$\frac{\partial^2 f}{\partial x\partial y}(a) = \frac{\partial^2 f}{\partial y\partial x}(a).$$

Теорема 27 (Юнг)**.** *Если $f\colon \mathbb{R}^2 \to \mathbb{R}$ дифференцируема в окрестности точки a и её частные производные $\frac{\partial f}{\partial x}$ и $\frac{\partial f}{\partial y}$ дифференцируемы в точке a, то*
$$\frac{\partial^2 f}{\partial x\partial y}(a) = \frac{\partial^2 f}{\partial y\partial x}(a).$$

Доказательства теорем Шварца и Юнга весьма схожи, поэтому мы ограничимся только доказательством теоремы Юнга.

$\triangleright$ Будем считать, что $a = (0,0)$. Рассмотрим выражение, в определённом смысле приближающее смешанную частную производную функции f по x и y:
$$F(t,t) = f(t,t) - f(0,t) - f(t,0) + f(0,0).$$
Заметим, что по теореме Лагранжа, применённой к функции $g(u) = f(t,u) - f(0,u)$, имеем
$$F(t,t) = g(t) - g(0) = g'(c)t = \Big(\frac{\partial f}{\partial y}(t,c) - \frac{\partial f}{\partial y}(0,c)\Big)t,$$
где c лежит между 0 и t. Функция $\partial f \partial y$ дифференцируема в точке $(0,0)$, а значит,
$$\frac{\partial f(x,y)}{\partial y} = \frac{\partial f}{\partial y}(0,0) + \frac{\partial^2 f}{\partial x\partial y}(0,0)\,x + \frac{\partial^2 f}{\partial y^2}(0,0)\,y + o(\sqrt{x^2+y^2}).$$
Следовательно,
$$\frac{\partial f}{\partial y}(t,c) - \frac{\partial f}{\partial y}(0,c) = \frac{\partial^2 f}{\partial x\partial y}(0,0)\,t + o(t).$$
Таким образом, имеем
$$F(t,t) = \frac{\partial^2 f}{\partial x\partial y}(0,0)\,t^2 + o(t^2).$$
Аналогично можно получить равенство
$$F(t,t) = \frac{\partial^2 f}{\partial y\partial x}(0,0)\,t^2 + o(t^2).$$
Приравняем правые части двух последних равенств и сократим на t^2. Устремляя $t \to 0$, получим равенство
$$\frac{\partial^2 f}{\partial x\partial y}(0,0) = \frac{\partial^2 f}{\partial y\partial x}(0,0). \qquad \triangleleft$$

Предположим, что f дифференцируема в некоторой окрестности $U(a)$ точки a. Тогда в этой окрестности определены новые функции (частные производные функции f)
$$x \mapsto \frac{\partial f}{\partial x_i}(x), \quad i = 1, 2, \dots n.$$
Если все частные производные функции f дифференцируемы в точке a, то говорят, что функция f *дважды дифференцируема* в точке a.

Пусть уже определено, что такое m раз дифференцируемая функция. Предположим, что функция f дифференцируема m раз в каждой точке некоторой окрестности $U(a)$. Если все её частные производные m-го порядка
$$\frac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\dots\partial x_{i_m}}$$

являются дифференцируемыми функциями в точке a, то функция f называется *дифференцируемой $(m+1)$ раз* в точке a.

Из теоремы Юнга получаем следующее утверждение.

Следствие 10. *Если функция f дифференцируема m раз в точке a, то значение частных производных*

$$\frac{\partial^k f}{\partial x_{i_1}\dots\partial x_{i_k}}(a), \quad k \leqslant m,$$

не зависит от порядка $i_1, \dots, i_k$.

Пусть теперь функция f дважды дифференцируема в точке a. По определению это означает, что в некоторой окрестности $U(a)$ существуют все частные производные первого порядка, которые ещё и дифференцируемы в точке a. В частности, для всякого h функция

$$x \mapsto df(x,h) = \frac{\partial f}{\partial x_1}(x)h_1 + \ldots + \frac{\partial f}{\partial x_n}(x)h_n$$

дифференцируема в точке a. Найдём дифференциал этой функции:

$$d(df(x,h))(a,v) = \sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}(a)h_i v_j = B(h,v).$$

Отображение $(h,v) \to B(h,v)$ является билинейной формой. Часто именно эту билинейную форму называют вторым дифференциалом функции f в точке a. Однако в силу теоремы Юнга $B(h,v)$ симметричная билинейная форма, т. е. $B(h,v) = B(v,h)$, а симметричная билинейная форма однозначно определяется соответствующей квадратичной формой $h \to B(h,h)$. Именно эту квадратичную форму мы будем называть вторым дифференциалом. Итак, *вторым дифференциалом* дважды дифференцируемой функции f в точке a называется квадратичная форма:

$$d^2 f(a,h) = \sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}(a)h_i h_j.$$

Если функция f дифференцируема m раз в точке a, то её *дифференциалом m-го порядка* называется функция $h \to d^m f(a,h)$, заданная формулой:

$$d^m f(a,h) = \sum_{1\leqslant i_1, i_2, \dots i_m \leqslant n}\frac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\dots\partial x_{i_m}}(a)h_{i_1}h_{i_2}\dots h_{i_m}.$$

Поскольку $dx_k(h) = h_k$, дифференциал $d^m f$ часто записывают в следующем виде:

$$d^m f = \sum_{1\leqslant i_1, i_2, \dots i_m \leqslant n}\frac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\dots\partial x_{i_m}}dx_{i_1}dx_{i_2}\dots dx_{i_m}.$$

Далее мы проясним роль дифференциалов высокого порядка.

Формула Тейлора

Предложение 35. *Пусть функция f дифференцируема m раз в окрестности точки a. Тогда для всякого вектора h функция $\varphi(t) = f(a+th)$ дифференцируема m раз в некоторой окрестности нуля и*

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

$\triangleright$ Пусть $m=1$. Тогда по правилу дифференцирования сложной функции получаем

$$\varphi'(t) = \frac{\partial f}{\partial x_1}(a+th)h_1 + \ldots + \frac{\partial f}{\partial x_n}(a+th)h_n = df(a+th,h).$$

Предположим, что для m утверждение доказано, и докажем его для $m+1$. Имеем

$$\varphi^{(m+1)}(t) = \frac{d}{dt}\Big(d^m f(a+th,h)\Big) = \frac{d}{dt}\Big(\sum_{i_1,\dots,i_m}\frac{\partial^m f}{\partial x_{i_1}\dots\partial x_{i_m}}(a+th)h_{i_1}\cdots h_{i_m}\Big).$$

По правилу дифференцирования сложной функции и по определению частных производных получаем

$$\varphi^{(m+1)}(t) = \sum_{i_1,\dots,i_m}\sum_j \frac{\partial^{m+1} f}{\partial x_j \partial x_{i_1}\dots\partial x_{i_m}}(a+th)h_{i_1}\cdots h_{i_m}\cdot h_j = d^{m+1}f(a+th,h). \qquad \triangleleft$$

Следующее утверждение является обобщением и прямым следствием одномерной формулы Тейлора с остаточным членом в форме Лагранжа.

Теорема 28. *Пусть функция f дифференцируема $(m+1)$ раз в открытом шаре $B_r(a)$. Тогда для всякого вектора h, $\|h\| < r$, существует такое число $c \in (0;1)$, что верно равенство*

$$f(a+h) = f(a) + df(a,h) + \frac{d^2 f(a,h)}{2!} + \ldots + \frac{d^m f(a,h)}{m!} + \frac{d^{m+1} f(a+ch,h)}{(m+1)!}.$$

$\rhd$ Достаточно записать формулу Тейлора с остаточным членом в форме Лагранжа для функции $\varphi(t) = f(a+th)$ в точке $t=0$ и подставить $t=1$. $\lhd$

Следствие 11. *Пусть частные производные функции f порядка $(m+1)$ непрерывны в точке a. Тогда*

$$f(a+h) = f(a) + df(a,h) + \frac{d^2 f(a,h)}{2!} + \ldots + \frac{d^m f(a,h)}{m!} + \frac{d^{m+1} f(a,h)}{(m+1)!} + o(\|h\|^{m+1}).$$

$\rhd$ Заметим, что $\frac{h_i}{\|h\|} \leqslant 1$ и

$$\frac{\partial^{m+1} f}{\partial x_{i_1} \ldots \partial x_{i_{m+1}}}(a) - \frac{\partial^{m+1} f}{\partial x_{i_1} \ldots \partial x_{i_{m+1}}}(a+ch) \to 0 \quad \text{при} \quad h \to 0.$$

Следовательно,

$$d^{m+1} f(a+ch,h) - d^{m+1} f(a,h) = \Big(d^{m+1} f(a+ch, h/\|h\|) - d^{m+1} f(a, h/\|h\|)\Big)\|h\|^{m+1} \to 0. \qquad \lhd$$

В качестве применения формулы Тейлора рассмотрим задачу поиска локальных экстремумов.

Локальный экстремум

Пусть $f\colon \mathbb{R}^n \to \mathbb{R}$. Если существует такая окрестность $U(a)$ точки a, что f определена в этой окрестности и $f(x) \leqslant f(a)$ $(f(x) < f(a))$ для всякого $x \in U(a) \setminus \{a\}$, то точка a называется *точкой локального максимума* (*строгого локального максимума*) функции f. Аналогично определяется точка локального минимума (строгого локального минимума).

Теорема 29. *Если функция f дифференцируема в точке своего локального экстремума, то в этой точке $df = 0$.*

$\rhd$ Пусть a — точка локального экстремума функции f. Тогда для каждого k функция одной переменной $x_k \to f(a_1, \ldots, a_{k-1}, x_k, a_{k+1}, \ldots, a_n)$ в точке a_k имеет локальный экстремум. По теореме Ферма производная этой функции в этой точке равна нулю, т. е. верно равенство $\frac{\partial f}{\partial x_k}(a) = 0$. $\lhd$

Если в случае дифференцируемой функции одной переменной для доказательства того, что некоторая точка является точкой локального экстремума, достаточно исследовать знак первой производной, то в случае функции нескольких переменных уже существенную роль играют производные второго порядка.

Теорема 30. *Пусть f дважды дифференцируема в окрестности точки a и частные производные второго порядка непрерывны в точке a. Предположим, что $df(a,h) = 0$ для всякого h. Тогда*

(i) *если $d^2 f(a,h) > 0$ для всякого $h \neq 0$, то a — точка строгого локального минимума;*

(ii) *если $d^2 f(a,h) < 0$ для всякого $h \neq 0$, то a — точка строгого локального максимума;*

(iii) *если существуют такие u, v, что $d^2 f(a,u) > 0$, $d^2 f(a,v) < 0$, то точка a не является точкой локального экстремума функции f.*

$\rhd$ По формуле Тейлора

$$f(a+h) - f(a) = \big(2^{-1} d^2 f(a, h/\|h\|) - \alpha(h)\big)\,\|h\|^2, \quad \lim_{h\to 0} \alpha(h) = 0.$$

Функция $h \mapsto d^2 f(a,h)$ непрерывна, а единичная сфера в $\mathbb{R}^n$ компактна, поэтому по теореме Вейерштрасса эта функция достигает на сфере своего наибольшего и наименьшего значения. Пусть $m = \min_{\|h\|=1} d^2 f(a,h) > 0$. Найдём такую окрестность $U(0)$, что $|\alpha(h)| < m/4$ для всех $h \in U(0) \setminus \{0\}$. Тогда для всех $h \in U(0) \setminus \{0\}$ верно неравенство $f(a+h) - f(a) > 0$. Аналогично разбирается случай, когда $M = \max\limits_{\|h\|=1} d^2 f(a,h) < 0$. Пусть теперь существуют такие u и v, что $d^2 f(a,u) > 0$ и $d^2 f(a,v) < 0$. Функция $\varphi(t) = f(a+tu)$ удовлетворяет следующим условиям: $\varphi'(0) = 0$ и $\varphi''(0) = d^2 f(a,u) > 0$. Следовательно, у этой функции точка 0 является точкой строгого локального минимума. В то же самое время функция $\psi(t) = f(a + tv)$ имеет в точке 0 строгий локальный максимум. Значит, в любой окрестности точки a встречаются точки (на прямой $a + tu$), в которых значение строго больше $f(a)$, и встречаются точки (на прямой $a + tv$), в которых значение строго меньше $f(a)$. Таким образом, точка a не является точкой локального экстремума функции f. $\triangleleft$

Условный экстремум

Пусть M_k — гладкая k-мерная поверхность в $\mathbb{R}^n$ и $f\colon \mathbb{R}^n \to \mathbb{R}$ – непрерывно дифференцируемая функция. Точка $p \in M_k$ называется *точкой локального условного максимума* (*строгого локального условного максимума*) функции f, если существует такая окрестность $U(p)$ точки p, что для всех $x \in U(p) \bigcap M_k$ верно неравенство $f(x) \leqslant f(p)$ ($f(x) < f(p)$ при $x \neq p$). Аналогично определяется точка локального условного минимума (строгого локального условного минимума).

Теорема 31 (необходимое условие локального условного экстремума)**.** *Если p является точкой локального условного экстремума, то $\nabla f(p) \perp T_pM_k$. Более того, если $\nabla f \neq 0$, то в точке p локального условного экстремума f касательное пространство T_pM_k является подпространством касательного пространства к поверхности $\{x\colon\ f(x) = f(p)\}$ в точке p.*

$\triangleright$ Пусть $v \in T_pM_k$. Тогда существует такая кривая γ на M_k, что $\gamma(0) = p$ и $\dot\gamma(0) = v$. Функция $f(\gamma(t))$ имеет в точке $t = 0$ локальный экстремум. Следовательно,

$$0 = \frac{d}{dt} f(\gamma(t))\Big|_{t=0} = \langle \nabla f(p), \dot\gamma(0)\rangle = \langle \nabla f(p), v\rangle. \qquad \triangleleft$$

Пусть M_k задано уравнениями $F_1 = 0, \ldots, F_{n-k} = 0$. Векторы $\nabla F_1(p)$, $\nabla F_2(p), \ldots, \nabla F_{n-k}(p)$ составляют базис ортогонального дополнения к пространству T_pM_k, поэтому

$$\nabla f(p) = \lambda_1 \nabla F_1(p) + \lambda_2 \nabla F_2(p) + \ldots + \lambda_{n-k} \nabla F_{n-k}(p),$$

причём λ_i определены единственным образом.

Это наблюдение называют *правилом множителей Лагранжа*: для поиска точек экстремума функции f при условии $F_1 = 0, \ldots, F_{n-k} = 0$ составляют функцию Лагранжа

$$L(x, \lambda) = f(x) - \lambda_1 F_1(x) - \ldots - \lambda_{n-k} F_{n-k}(x)$$

и приравнивают к нулю её частные производные по x_i и λ_j.

Теперь обсудим достаточное условие.

Теорема 32. *Пусть в точке p выполняется необходимое условие локального условного экстремума и $L(x) = L(x, \lambda)$. Тогда*

(i) *если $d_x^2 L(p, h) > 0$ для всякого $h \neq 0$ и $h \in T_pM_k$, то p — точка строгого локального условного минимума*;

(ii) *если $d_x^2 L(p, h) < 0$ для всякого $h \neq 0$ и $h \in T_pM_k$, то p — точка строгого локального условного максимума*;

(iii) *если существуют такие $h, v \in T_pM_k$, что $d_x^2 L(p, h) > 0$ и $d_x^2 L(p, v) < 0$, то точка p не является точкой локального условного экстремума.*

$\triangleright$ Заметим, что $f(x) = L(x)$ при $x \in M_k$. Пусть $x_i = g_i(t_1, \ldots, t_k)$ задают M_k в окрестности точки p. Вычислим второй дифференциал функции $H(t) = L(g(t))$:

$$d^2 H(0, u) = \sum_{i,j,l,m} \frac{\partial^2 L}{\partial x_i \partial x_j} \cdot \frac{\partial g_j}{\partial t_l} \cdot \frac{\partial g_i}{\partial t_m} u_l u_m.$$

Заметим, что слагаемых, содержащих частные производные L первого порядка, нет, так как мы рассматриваем точку, в которой эти производные в силу необходимого условия равны нулю. Итак, $d^2 H(0, u) = d_x^2 L(p, h)$, где вектор

$$h = \frac{\partial g}{\partial t_1} u_1 + \ldots + \frac{\partial g}{\partial t_k} u_k$$

является касательным к M_k в точке p. Теперь достаточно заметить, что условный экстремум для f является обычным локальным экстремумом для $H(t)$ и можно применить достаточные условия локального экстремума. $\triangleleft$

Во многих случаях приходится рассматривать более общую задачу на условный экстремум. Пусть на $\mathbb{R}^n$ заданы непрерывно дифференцируемые функции f, $F_1, \ldots, F_m$ и $G_1, \ldots, G_k$. Требуется исследовать на экстремум функцию f на множестве, заданном следующими равенствами и неравенствами:

$$F_1(x) = 0,\ \ldots,\ F_m(x) = 0, \quad G_1(x) \leqslant 0,\ \ldots,\ G_m(x) \leqslant 0.$$

Оказывается, можно действовать так же, как и выше. Выпишем функцию Лагранжа:

$$L(x, \lambda, \mu) = \lambda_0 f(x) - \lambda_1 F_1(x) - \ldots - \lambda_m F_m(x) - \mu_1 G_1(x) - \ldots - G_k(x).$$

Предложение 36. *Если точка p является локальным экстремумом функции f при выполнении условий, заданных функциями F_i и G_j, то найдутся такие λ_i и μ_j, не все равные нулю, что выполнены условия Куна — Таккера:*

$$\nabla_x L(p, \lambda, \mu) = 0, \quad \mu_1 G_1(p) = 0,\ \ldots,\ \mu_k G_k(p) = 0.$$

$\triangleright$ Если $G_j(p) < 0$, то μ_j полагаем равным нулю и далее такие G_j не рассматриваем. Пусть теперь $F_i(p) = 0$ и $G_j(p) = 0$ для всех i и j. Если градиенты функций ∇F_i, ∇G_j в точке p линейно зависимы, то утверждение выполнено с $\lambda_0 = 0$ и λ_i, μ_j — коэффициентами соответствующей тривиальной линейной комбинации. Если градиенты функций ∇F_i, ∇G_j в точке p линейно независимы, то равенства $F_i = 0$, $G_j = 0$ в окрестности точки p задают гладкую поверхность, на которой точка p является точкой условного локального экстремума функции f. В этом случае утверждение следует из уже доказанного выше. $\triangleleft$

В случае выпуклых функций f, F_i, G_j последнее утверждение можно значительно обобщить.

Неопределённый интеграл

Пусть $f\colon (a;b) \to \mathbb{R}$. Функция $F\colon (a;b) \to \mathbb{R}$ называется *первообразной* функции f, если F дифференцируема на $(a;b)$ и $F' = f$.

Предложение 37. *Любые две первообразные F_1 и F_2 функции f на интервале $(a;b)$ отличаются на константу.*

$\triangleright$ Положим $F = F_1 - F_2$. Поскольку

$$F'(x) = (F_1 - F_2)'(x) = f(x) - f(x) = 0 \quad \text{при всех } x \in (a;b),$$

по теореме Лагранжа получаем $F(x_1) - F(x_2) = F'(c)(x_1 - x_2) = 0$, а значит, функция F на интервале $(a;b)$ является константой. $\triangleleft$

Совокупность всех первообразных функций функции $f(x)$ на данном интервале называется *неопределённым интегралом* от функции $f(x)$. Эта совокупность обозначается через

$$\int f(x)\,dx.$$

В силу доказанного, если F — некоторая первообразная функции f на интервале, совокупность всех первообразных является множеством функций вида $F + C$, где C — произвольная константа, что кратко выражается в виде равенства

$$\int f(x)\,dx = F + C.$$

Отметим, что далее при выполнении операций с неопределёнными интегралами (дифференцирование, сложение и т. п.) операция выполняется с каждой первообразной по отдельности, а результат записывается с помощью неопределённых интегралов и понимается как совпадение множеств.

Следующие равенства иллюстрируют данное выше определение и подчёркивают, что операции дифференцирования и нахождения первообразной являются «взаимно обратными»:

$$d\Big(\int f(x)\,dx\Big) = f(x)\,dx, \qquad \int F'(x)\,dx = F + C.$$

Обсудим эти равенства. Первое из них означает, что дифференциал любой первообразной функции f равен $f(x)\,dx$. Второе равенство означает, что если функция F дифференцируема, то все первообразные функции F' имеют вид $F + C$. В этом случае подынтегральное выражение имеет вид $F'(x)\,dx$ и совпадает с дифференциалом функции F, поэтому второе равенство можно символически записать в виде $\int dF = F + C$.

Найти первообразные некоторых элементарных функций можно, просто прочитав справа налево таблицу производных. Следующие интегралы называют *табличными*:

$$1)\ \int x^a\,dx = \frac{x^{a+1}}{a+1} + C,\ a \neq -1; \qquad 6)\ \int \cos x\,dx = \sin x + C;$$

$$2)\ \int \frac{dx}{x} = \ln|x| + C; \qquad 7)\ \int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C;$$

$$3)\ \int a^x\,dx = \frac{a^x}{\ln a} + C,\ a > 0,\ a \neq 1, \qquad 8)\ \int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + C;$$

$$4)\ \int e^x\,dx = e^x + C; \qquad 9)\ \int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C;$$

$$5)\ \int \sin x\,dx = -\cos x + C; \qquad 10)\ \int \frac{dx}{1+x^2} = \operatorname{arctg} x + C.$$

Все перечисленные равенства имеют место на любом промежутке из области дифференцируемости правой части.

Свойства неопределённого интеграла получаются путём переформулирования свойств операции дифференцирования на языке интегралов.

1. (*Линейность*) Пусть на интервале $(a;b)$ заданы дифференцируемые функции F и G, причём $F' = f$ и $G' = g$. Свойство линейности для производных имеет вид

$$(\alpha F + \beta G)' = \alpha F' + \beta G' = \alpha f + \beta g,$$

т. е. функция $\alpha F + \beta G$ является первообразной функции $\alpha f + \beta g$, что в терминах неопределённого интеграла выражается равенством

$$\int (\alpha f(x) + \beta g(x))\,dx = \alpha \int f(x)\,dx + \beta \int g(x)\,dx$$

(равенство верно, если хотя бы одно из чисел α, β отлично от нуля; если оба числа равны нулю, то в правой части равенства исчезает произвольная постоянная, а выражение становится тождественно равным нулю, поэтому при $\alpha = \beta = 0$ в правой части должна быть дописана константа C).

2. (*Формула интегрирования по частям*) Пусть на интервале $(a;b)$ заданы дифференцируемые функции f и g, причём функция $f'g$ имеет первообразную. Правило Лейбница имеет вид

$$(fg)' = f'g + fg',$$

следовательно, совокупность первообразных функции $f'g + fg'$ состоит из функций вида $fg + C$. В силу линейности неопределённого интеграла из существования первообразной функции $f'g$ следует существование первообразной функции fg' и выполняется равенство

$$f(x)g(x) + C = \int f'(x)g(x)\,dx + \int f(x)g'(x)\,dx.$$

Последнее равенство обычно записывают в виде соотношения

$$\int f(x)g'(x)\,dx = f(x)g(x) - \int f'(x)g(x)\,dx,$$

которое называют *формулой интегрирования по частям.* Поскольку $g'(x)\,dx = dg$ и $f'(x)\,dx = df$, эту формулу часто символически записывают в виде

$$\int f\,dg = fg - \int g\,df.$$

3. (*Формула замены переменной*) Пусть функция F дифференцируема на интервале $(a;b)$ и $F' = f$. Пусть также $\varphi\colon (\alpha;\beta) \to (a;b)$ — дифференцируемая функция. По теореме о дифференцировании сложной функции композиция $F(\varphi(t))$ дифференцируема и

$$F(\varphi(t))' = F'(\varphi(t))\varphi'(t) = f(\varphi(t))\varphi'(t).$$

Записывая это равенство с помощью интегралов, получаем *формулу замены переменной*:

$$\int f(\varphi(t))\varphi'(t)\,dt = \int f(x)\,dx\Big|_{x=\varphi(t)}.$$

На этой формуле основан один из ключевых приёмов нахождения первообразной произведения функций — *подведение под знак дифференциала*: если один из множителей является производной некоторой функции φ, а другой множитель имеет вид $f \circ \varphi$, то функцию φ *подводят* под знак дифференциала и пишут

$$\int f(\varphi(t))\varphi'(t)\,dt = \int f(\varphi(t))\,d\varphi(t),$$

где можно рассматривать φ как новую переменную. Например,

$$\int \sin^9 t \cos t\,dt = \int \sin^9 t\,d\sin t = \frac{1}{10}\sin^{10} t + C.$$

Вообще, основная идея вычисления первообразных состоит в применении указанных правил интегрирования и сведении исходного интеграла к табличному. В некоторых случаях отыскание первообразных алгоритмизировано. Например, для отыскания первообразной рациональной функции раскладывают эту функцию в сумму простейших дробей и затем вычисляют интеграл отдельно от каждого слагаемого, а эти интегралы мало отличаются от табличных. В некоторых других ситуациях подходящей заменой переменной подынтегральное выражение можно преобразовать к рациональной функции. Однако в общем случае вычисление первообразных является сложной задачей. Более того, довольно часто встречаются первообразные, которые не выражаются в элементарных функциях: такова, например, первообразная функции e^{-x^2}.

Интеграл Римана

Выше было предложено находить первообразные, преобразуя подынтегральное выражение с помощью алгебраических преобразований, замены переменной и интегрирования по частям к функциям, чьи первообразные известны. Однако такой способ даёт ответ лишь для весьма узкого класса относительно простых функций. Рассмотрим другой способ построения первообразной. Предположим, что на интервале $(\alpha;\beta)$ задана функция f и мы хотим найти такую функцию F, что $F' = f$ на этом интервале. Пусть $x_0 \in (\alpha;\beta)$. Поскольку мы ищем первообразную с точностью до константы, можно считать, что $F(x_0) = 0$. Пусть $x > x_0$. В силу определения производной

$$F(x) = f(x_0)(x - x_0) + o(x - x_0),$$

т. е., нестрого говоря, $F(x) \approx f(x)(x - x_0)$, если величина $x - x_0$ мала. Пусть теперь $x - x_0$ нельзя считать малой величиной. Тогда разобьем отрезок $[x_0; x]$ на маленькие отрезки точками $x_0 < x_1 < \ldots < x_N = x$ так, что на каждом из них

$$F(x_{k+1}) - F(x_k) \approx f(x_{k-1})(x_k - x_{k-1}).$$

Тогда

$$F(x) = \sum_{k=1}^{N}(F(x_k) - F(x_{k-1}) \approx \sum_{k=1}^{N} f(x_{k-1})(x_k - x_{k-1}).$$

Устремляя N к бесконечности и длину отрезков $[x_{k-1}; x_k]$ к нулю, эвристически получаем, что

$$F(x) = \lim \sum_{k=1}^{N} f(x_{k-1})(x_k - x_{k-1}).$$

Конечно, такие неформальные рассуждения оставляют много вопросов. Почему погрешность, возникающая при замене $F(x_{k+1}) - F(x_k)$ на $f(x_{k-1})(x_k - x_{k-1})$, не накапливается при суммировании? В каком смысле следует понимать предел в последнем равенстве? Для каких функций f такой предел существует? Каков геометрический смысл этого предела? Ответу на эти вопросы посвящены данный и следующий разделы.

Пусть функция f определена на отрезке $[a;b]$.

- *Разбиением* $\mathbb{T}$ отрезка $[a;b]$ называется набор точек

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b.$$

- Отрезки $\Delta_i = [x_{i-1}; x_i]$, где $i = 1, 2, \dots, n$, называются *отрезками разбиения.*
- Число $\lambda(\mathbb{T}) = \max\limits_{1 \leqslant i \leqslant n} |x_i - x_{i-1}|$ называется *параметром разбиения* или *масштабом.*
- *Отмеченным разбиением* $(\mathbb{T}, \xi)$ отрезка $[a; b]$ называется разбиение $\mathbb{T}$ отрезка $[a; b]$ с отмеченными точками $\xi = (\xi_1, \xi_2, \dots, \xi_n)$, где $\xi_k \in [x_{k-1}; x_k]$.
- *Римановой интегральной суммой* функции f, соответствующей отмеченному разбиению $(\mathbb{T}, \xi)$ отрезка $[a; b]$, называется выражение

$$\sigma(f, \mathbb{T}, \xi) = \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}).$$

Функция f *интегрируема по Риману* на отрезке $[a; b]$ и число I является её интегралом, если для всякого $\varepsilon > 0$ найдётся такое $\delta > 0$, что для всякого отмеченного разбиения $(\mathbb{T}, \xi)$ с параметром разбиения $\lambda(\mathbb{T}) < \delta$ выполняется неравенство

$$|\sigma(f, \mathbb{T}, \xi) - I| < \varepsilon.$$

Число I обозначают через

$$\int_a^b f(x)\, dx$$

и называют *интегралом Римана.*

Интеграл Римана можно понимать как предел по базе. Напомним, что если X — непустое множество, то *базой* называется непустой набор $\mathcal{B}$ его подмножеств, удовлетворяющий условиям: 1) $\varnothing \notin \mathcal{B}$ и 2) для всяких $U, V \in \mathcal{B}$ существует такое $W \in \mathcal{B}$, что $W \subset U \cap V$. Пусть $f\colon X \to \mathbb{R}$. Число A называется *пределом* функции f по базе $\mathcal{B}$, если для всякого $\varepsilon > 0$ найдётся такой элемент базы $U \in \mathcal{B}$, что $|f(x) - A| < \varepsilon$ при $x \in U$. Обозначают число A через $\lim\limits_{\mathcal{B}} f$.

Пусть теперь X — все возможные отмеченные разбиения отрезка $[a; b]$, а база $\mathcal{B}$ — набор множеств B_δ, состоящих из отмеченных разбиений $(\mathbb{T}, \xi)$ с параметром разбиения $\lambda(\mathbb{T}) < \delta$, где $\delta > 0$. Тогда приведённое выше определение в точности означает, что

$$\int_a^b f(x)\, dx = \lim_{\mathcal{B}} \sigma(f, \mathbb{T}, \xi).$$

Конечно, многие свойства интеграла можно вывести из свойств предела по базе, но мы проведём прямые рассуждения, не опирающиеся на понятие предела по базе.

Рассмотрим несколько примеров.

1. Пусть $f \equiv 1$ на $[a; b]$. Тогда для всякого отмеченного разбиения риманова интегральная сумма функции f равна числу $(b - a)$. Следовательно, по определению получаем

$$\int_a^b 1\, dx = b - a.$$

2. Пусть Δ — промежуток с концами $c \leqslant d$ в отрезке $[a; b]$. Индикатор $I_\Delta(x) = 1$ при $x \in \Delta$ и $I_\Delta(x) = 0$ при $x \notin \Delta$, интегрируем по Риману и

$$\int_a^b I_\Delta(x)\, dx = d - c.$$

Распишем риманову сумму:

$$\sigma(I_\Delta, \mathbb{T}, \xi) = \sum_{k\colon c \text{ или } d \in \Delta_k} I_\Delta(\xi_k)|\Delta_k| + \sum_{k\colon \Delta_k \subset (c,d)} |\Delta_k|.$$

Пусть $\lambda(\mathbb{T}) < \delta$. Тогда

$$(d - c) - 2\delta \leqslant \sigma(I_\Delta, \mathbb{T}, \xi) \leqslant 4\delta + (d - c).$$

Для всякого $\varepsilon > 0$ возьмём $\delta < \varepsilon/4$. Тогда для всякого отмеченного разбиения $(\mathbb{T}, \xi)$ с масштабом разбиения $\lambda(\mathbb{T}) < \delta$ верно неравенство

$$|\sigma(I_\Delta, \mathbb{T}, \xi) - (d - c)| < \varepsilon.$$

3. Функция Дирихле $D(x) = 1$ при $x \in \mathbb{Q}$ и $D(x) = 0$ при $x \notin \mathbb{Q}$ не интегрируема по Риману ни на каком отрезке $[a; b]$. Действительно, выбором отмеченных точек риманову сумму можно сделать равной нулю или равной $b - a$ при любом масштабе разбиения.

Прежде чем обсуждать свойства интеграла Римана и выяснять, какие же функции интегрируемы, формализуем рассуждения из начала этого раздела и покажем, как интеграл Римана связан с первообразной.

Теорема 33 (формула Ньютона — Лейбница). *Пусть функция F дифференцируема на отрезке $[a;b]$ и её производная F' интегрируема по Риману на $[a;b]$. Тогда*

$$\int_a^b F'(x)\,dx = F(b) - F(a).$$

$\triangleright$ Рассмотрим произвольное разбиение $\mathbb{T} = \{a = x_0 < x_1 < \ldots < x_n = b\}$ отрезка $[a;b]$. Тогда по теореме Лагранжа $F(x_k) - F(x_{k-1}) = F'(\xi_k)|\Delta_k|$. Следовательно, верны равенства

$$F(b) - F(a) = \sum_{k=1}^{N}(F(x_k) - F(x_{k-1})) = \sum_{k=1}^{N} F'(\xi_k)|\Delta_k|.$$

Остаётся заметить, что при $\lambda(\mathbb{T}) \to 0$ правая часть стремится к интегралу от F' по $[a;b]$. $\triangleleft$

Эта теорема позволяет по производной восстановить первообразную, однако только при выполнении двух условий: 1) эта первообразная действительно существует, и 2) производная интегрируема по Риману. Вполне может случиться, что интегрируемая по Риману функция не является ничьей производной, или что производная не является интегрируемой функцией. Поэтому далее основные усилия будут приложены к описанию интегрируемых функций и выделению тех из них, у которых существует первообразная. Начнём со следующего простого наблюдения.

Предложение 38. *Если функция интегрируема по Риману на отрезке, то эта функция ограниченна на этом отрезке.*

$\triangleright$ Если функция не является ограниченной, то для всякого разбиения найдётся отрезок разбиения, на котором эта функция принимает сколь угодно большие по модулю значения. Следовательно, для всякого разбиения выбором отмеченных точек интегральную сумму можно сделать сколь угодно большой, что противоречит определению интеграла Римана. $\triangleleft$

Прямо из определения выводятся три важнейших свойства интеграла: линейность, монотонность и аддитивность.

Предложение 39 (линейность интеграла). *Если f и g интегрируемы на $[a;b]$, то функция $\alpha f + \beta g$ интегрируема на $[a;b]$ и верно равенство*

$$\int_a^b (\alpha f + \beta g)\,dx = \alpha\int_a^b f\,dx + \beta\int_a^b g\,dx.$$

$\triangleright$ Справедливо равенство

$$\sigma(\alpha f + \beta g, \mathbb{T}, \xi) = \alpha\sigma(f, \mathbb{T}, \xi) + \beta\sigma(g, \mathbb{T}, \xi).$$

Для всякого $\varepsilon > 0$ существует такое $\delta > 0$, что

$$\left|\sigma(f, \mathbb{T}, \xi) - \int_a^b f\,dx\right| < \varepsilon, \quad \left|\sigma(g, \mathbb{T}, \xi) - \int_a^b g\,dx\right| < \varepsilon$$

для всех $(\mathbb{T}, \xi)$ с масштабом разбиения $\lambda(\mathbb{T}) < \delta$. Тогда для таких отмеченных разбиений верно неравенство

$$\left|\sigma(\alpha f + \beta g, \mathbb{T}, \xi) - \alpha\int_a^b f\,dx - \beta\int_a^b g\,dx\right| \leqslant (|\alpha| + |\beta|)\varepsilon.$$ $\triangleleft$

Следствие 12. *Пусть $\Delta_1, \ldots, \Delta_N$ — произвольные промежутки в отрезке $[a;b]$. Тогда для всяких чисел $c_k \in \mathbb{R}$ функция $f(x) = \sum\limits_k c_k I_{\Delta_k}(x)$ интегрируема на $[a;b]$ и*

$$\int_a^b f\,dx = \sum_k c_k|\Delta_k|.$$

Функции такого вида будем называть ступенчатыми.

$\triangleright$ Выше было доказано, что функция I_{Δ_k} интегрируема и её интеграл равен $|\Delta_k|$. Поэтому требуемое утверждение является следствием линейности интеграла. $\triangleleft$

Предложение 40 (монотонность интеграла)**.** *Если f и g интегрируемы на $[a;b]$ и $f \leqslant g$, то*

$$\int_a^b f\,dx \leqslant \int_a^b g\,dx.$$

$\triangleright$ В силу линейности интеграла это свойство достаточно доказать в случае $f = 0$, т. е. из неотрицательности g вывести неотрицательность интеграла от g. Имеем $\sigma(g, \mathbb{T}, \xi) \geqslant 0$. Для всякого $\varepsilon > 0$ существует $\delta > 0$, при котором

$$\int_a^b g\,dx \geqslant \sigma(g, \mathbb{T}, \xi) - \varepsilon \geqslant -\varepsilon.$$

Устремляя ε к нулю, заключаем, что интеграл от g неотрицателен. $\triangleleft$

Следствие 13 (теорема о среднем)**.** *Пусть f интегрируема на $[a;b]$, $m = \inf\limits_{[a;b]} f$, $M = \sup\limits_{[a;b]} f$. Тогда найдётся такое число $\mu \in [m;M]$, что*

$$\int_a^b f\,dx = \mu(b-a).$$

Более того, если f непрерывна, то $\mu = f(c)$ для некоторого $c \in [a;b]$.

$\triangleright$ Достаточно заметить, что в силу монотонности интеграла верно следствие:

$$m \leqslant f(x) \leqslant M \Rightarrow m \leqslant \frac{1}{b-a}\int_a^b f\,dx \leqslant M. \qquad \triangleleft$$

Следующая теорема имеет важное значение в теории интеграла Римана и позволяет доказывать интегрируемость и вычислять интеграл, приближая подынтегральную функцию последовательностью более простых функций.

Теорема 34. *Если последовательность интегрируемых на $[a;b]$ функций f_n равномерно на $[a;b]$ сходится к f (т. е. $\sup\limits_{[a;b]} |f_n(x) - f(x)| \to 0$), то функция f интегрируема на $[a;b]$ и верно равенство*

$$\lim_{n\to\infty} \int_a^b f_n\,dx = \int_a^b f\,dx.$$

$\triangleright$ По теореме о среднем

$$\left|\int_a^b f_n\,dx - \int_a^b f_m\,dx\right| = |\mu_{n,m}||b-a|,$$

где

$$\inf_{[a;b]}(f_n(x) - f_m(x)) \leqslant \mu_{n,m} \leqslant \sup_{[a;b]}(f_n(x) - f_m(x)).$$

Следовательно, $|\mu_{n,m}| \leqslant \sup\limits_{[a;b]} |f_n - f_m|$ и $\mu_{m,n} \to 0$ при $n, m \to \infty$. Таким образом, числовая последовательность

$$\int_a^b f_n\,dx$$

сходится к некоторому числу I. Покажем, что функция f интегрируема и её интеграл равен I. Пусть $(\mathbb{T}, \xi)$ — отмеченное разбиение отрезка $[a;b]$. Имеет место неравенство:

$$\Big|\sigma(f, \mathbb{T}, \xi) - I\Big| \leqslant \sigma(|f - f_n|, \mathbb{T}, \xi) + \Big|\sigma(f_n, \mathbb{T}, \xi) - \int_a^b f_n\,dx\Big| + \Big|\int_a^b f_n\,dx - I\Big|.$$

Пусть $\varepsilon > 0$. Выберем номер n столь большим, чтобы первое и третье слагаемые стали меньше ε. Зафиксируем такое n. Теперь выберем масштаб разбиения так, чтобы для всех отмеченных разбиений с таким или меньшим масштабом второе слагаемое было меньше ε. Получаем оценку

$$\Big|\sigma(f, \mathbb{T}, \xi) - I\Big| < 3\varepsilon,$$

которая завершает доказательство. $\triangleleft$

Следствие 14. *Если функция f непрерывна на отрезке $[a;b]$, то функция f интегрируема на $[a;b]$.*

$\triangleright$ Рассмотрим разбиение отрезка $[a;b]$, заданное точками $x_k = a + \frac{k(b-a)}{n}$. Определим функцию f_n следующим образом: $f_n(x) = \sum\limits_k f(x_{k-1}) I_{\Delta_k}$, где $\Delta_k = [x_{k-1}; x_k)$ при $k \leqslant n-1$ и $\Delta_n = [x_{n-1}; x_n]$. Функция f_n интегрируема на $[a;b]$. Поскольку f равномерно непрерывна на $[a;b]$, для всякого $\varepsilon > 0$ найдётся такое N, что для всех $n > N$ верно неравенство $|f(x) - f(x_k)| < \varepsilon$ при $x \in [x_{k-1}, x_k]$. Следовательно, $\sup\limits_{[a;b]} |f(x) - f_n(x)| \leqslant \varepsilon$ для всех $n > N$. Таким образом, последовательность f_n равномерно сходится к f, а это влечёт интегрируемость f. $\triangleleft$

Функция f называется *кусочно непрерывной* на отрезке $[a;b]$, если есть такое разбиение $a = x_0 < x_1 < \ldots < x_n = b$, что функция f непрерывна на каждом интервале $(x_{i-1}; x_i)$ и существуют конечные пределы $\lim\limits_{x \to x_i - 0} f(x)$ и $\lim\limits_{x \to x_i + 0} f(x)$ в каждой точке x_i. Незначительно модифицируя рассуждения следствия 14, можно доказать, что всякая кусочно непрерывная функция интегрируема.

Следствие 15. *Если функция f монотонна на $[a;b]$, то f интегрируема на $[a;b]$.*

$\triangleright$ Пусть для определённости функция f неубывающая и $f(a) < f(b)$ (случай $f(a) = f(b)$ тривиален, так как тогда $f = \mathrm{const}$). Разобьем отрезок $[f(a), f(b)]$ точками $y_k = f(a) + \frac{(f(b)-f(a))k}{n}$, где $k = 1, 2, \ldots, n$. Положим $\Delta_k = \{x \colon y_{k-1} \leqslant f(x) < y_k\}$ при $k \leqslant n-1$ и $\Delta_n = \{x \colon y_{n-1} \leqslant f(x) \leqslant y_n\}$. Заметим, что $\Delta_k \cap \Delta_l = \varnothing$ при $k \neq l$ и для каждого k множество Δ_k или пусто, или является промежутком. Пусть $f_n(x) = \sum\limits_{k=1}^{n} y_{k-1} I_{\Delta_k}(x)$. Поскольку для $x \in \Delta_k$ справедлива оценка

$$|f_n(x) - f(x)| = |y_{k-1} - f(x)| \leqslant y_k - y_{k-1} = \frac{f(b) - f(a)}{n},$$

получаем $\sup\limits_{[a;b]} |f_n(x) - f(x)| \to 0$, поэтому f интегрируема на $[a;b]$. $\triangleleft$

Отметим, что, развивая способ приближения функции f из последнего следствия, можно построить знаменитый интеграл Лебега. Однако для этого надо уметь измерять длину прообразов промежутков, что само по себе является нетривиальной задачей. В случае монотонной функции такие прообразы оказываются промежутками и их длина — просто разность координат концов, а в общем случае прообраз может быть устроен очень сложно.

Итак, построен естественный класс интегрируемых по Риману функций — множество ограниченных функций, которые равномерно приближаются ступенчатыми функциями. Такие функции образуют линейное пространство, содержащее кусочно непрерывные и монотонные функции. Однако этот класс функций не содержит все интегрируемые по Риману функции, и для получения критерия интегрируемости потребуются более тонкие рассуждения, которые мы вскоре проведём.

Наконец, обсудим третье свойство интеграла — аддитивность.

Предложение 41 (аддитивность интеграла)**.** *Предположим, что функция f интегрируема по Риману на отрезке $[a;b]$ и отрезках $[a;c]$ и $[c;b]$, где $a < c < b$. Тогда верно равенство*

$$\int_a^b f\,dx = \int_a^c f\,dx + \int_c^b f\,dx.$$

$\triangleright$ По условию f интегрируема на $[a;b]$, $[a;c]$ и $[c;b]$, поэтому римановы суммы сходятся к интегралам от f по этим отрезкам для любых разбиений, масштаб которых стремится к нулю. Пусть $(\mathbb{T}_{[a;c]}, \xi_{[a;c]})$ состоит из точек $x_k = a + (c-a)k/n$, $\xi_k = x_{k-1}$, и $(\mathbb{T}_{[c;b]}, \xi_{[c;b]})$ состоит из точек $x_k = c + (b-c)k/n$, $\xi_k = x_{k-1}$. Объединение этих разбиений является отмеченным разбиением $\mathbb{T}_{[a;b]}$ отрезка $[a;b]$ и

$$\sigma(f, \mathbb{T}_{[a;b]}, \xi) = \sigma(f, \mathbb{T}_{[a,c]}, \xi) + \sigma(f, \mathbb{T}_{[c,b]}, \xi).$$

Остаётся заметить, что при $n \to \infty$ левая часть стремится к интегралу по $[a;b]$, а правая часть стремится к сумме интегралов по $[a;c]$ и $[c;b]$. $\triangleleft$

В дальнейшем мы покажем, что из интегрируемости по $[a;c]$ и $[c;b]$ следует интегрируемость по $[a;b]$ и, наоборот, из интегрируемости по $[a;b]$ следует интегрируемость по любому подотрезку.

Далее удобно использовать следующие хорошо согласующиеся с аддитивностью соглашения:

$$\int_a^a f(x)\,dx = 0, \quad \int_b^a f(x)\,dx = -\int_a^b f(x)\,dx.$$

В заключение этого раздела приведём утверждения, которые хорошо иллюстрируют доказанные выше свойства интеграла и дают ответ на вопрос о существовании первообразных у непрерывных функций.

Теорема 35. *Пусть f непрерывна на $[a;b]$. Тогда функция*

$$F(x) = \int_a^x f(t)\,dt$$

является первообразной функции f на $[a;b]$, т. е. для всех $x \in [a;b]$ верно равенство

$$F'(x) = f(x).$$

Более того, для всякой первообразной $\mathcal{F}$ функции f выполнена формула Ньютона — Лейбница:

$$\int_a^b f(x)\,dx = \mathcal{F}(b) - \mathcal{F}(a).$$

▷ Запишем определение производной и воспользуемся аддитивностью интеграла:

$$F'(x) = \lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h\to 0} \frac{1}{h} \int_x^{x+h} f(t)\,dt.$$

По теореме о среднем существует такое c из отрезка с концами x и $x+h$, что

$$\frac{1}{h} \int_x^{x+h} f(t)\,dt = f(c).$$

Поскольку f — непрерывная функция, получаем, что $f(c)$ стремится к $f(x)$ при $h \to 0$. Последнее утверждение следует из того, что любые две первообразные отличаются на константу. ◁

Следствие 16 (формула интегрирования по частям)**.** *Пусть f, g — непрерывно дифференцируемые на отрезке $[a;b]$ функции. Тогда*

$$\int_a^b f(x)g'(x)\,dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x)\,dx.$$

▷ Достаточно заметить, что по формуле Ньютона — Лейбница

$$\int_a^b (fg)'\,dx = f(b)g(b) - f(a)g(a)$$

и применить правило Лейбница $(fg)' = f'g + fg'$. ◁

Следствие 17 (формула замены переменных)**.** *Пусть f — непрерывная на отрезке $[a;b]$ функция и $\varphi : [\alpha;\beta] \to [a;b]$ — непрерывно дифференцируемая функция. Тогда*

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(x)\,dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)\,dt.$$

▷ Пусть F — первообразная функции f. Тогда $F(\varphi(t))$ — первообразная $f(\varphi(t))\varphi'(t)$. Применяя формулу Ньютона — Лейбница, получаем

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f\,dx = F(\varphi(\beta)) - F(\varphi(\alpha)) = \int_\alpha^\beta f(\varphi(t))\varphi'(t)\,dt. \qquad ◁$$

Критерий Дарбу. Множества меры нуль. Критерий Лебега

Обсудим геометрический подход к построению интеграла Римана. По определению интеграл Римана от функции f по отрезку $[a;b]$ является пределом римановых сумм

$$\sigma(f, \mathbb{T}, \xi) = \sum_k f(\xi_k)|\Delta_k|.$$

Пусть функция f неотрицательна. Тогда риманова сумма в точности равна сумме площадей прямоугольников, каждый из которых имеет основание Δ_k и высоту $f(\xi_k)$. Интуитивно ясно, что при стремлении масштаба к нулю эта сумма площадей прямоугольников приближается к площади под графиком функции $y = f(x)$ над отрезком $[a;b]$ оси абсцисс. Таким образом, интеграл Римана выражает площадь под графиком функции, но понятие площади (как и длины, и объёма) очень нетривиально и выходит за рамки настоящего курса. Однако ещё со школы известен следующий подход к определению площади фигуры. У некоторых простых фигур (например, составленных из прямоугольников), площади известны. Поэтому для определения площади данной фигуры вокруг неё описывают и в неё вписывают простые фигуры, стремясь сделать это как можно точнее. Тогда площадью данной фигуры естественно назвать предел площадей описанных и вписанных фигур, если эти пределы существуют и совпадают. Именно этот подход выражают так называемые *суммы Дарбу*.

Пусть f — ограниченная на отрезке $[a;b]$ функция и $\mathbb{T}$ — разбиение отрезка $[a;b]$. Выражения

$$s(f,\mathbb{T})=\sum_k \inf_{\Delta_k} f(x)|\Delta_k|, \quad S(f,\mathbb{T})=\sum_k \sup_{\Delta_k} f(x)|\Delta_k|$$

называются *нижней и верхней суммой Дарбу* соответственно. Здесь $\Delta_k=[x_{k-1};x_k]$.

Лемма 2. *Имеют место следующие утверждения.*

(i) *Для всякого разбиения $\mathbb{T}$ верны неравенства*

$$s(f,\mathbb{T})\leqslant \inf_\xi \sigma(f,\mathbb{T},\xi)\leqslant \sigma(f,\mathbb{T},\xi)\leqslant \sup_\xi \sigma(f,\mathbb{T},\xi)\leqslant S(f,\mathbb{T}).$$

(i) *Если разбиение $\mathbb{T}\subset\mathbb{T}'$ (говорят, что $\mathbb{T}'$ является измельчением $\mathbb{T}$), то*

$$s(f,\mathbb{T})\leqslant s(f,\mathbb{T}'), \quad S(f,\mathbb{T}')\leqslant S(f,\mathbb{T}).$$

(iii) *Если $\mathbb{T}_1$ и $\mathbb{T}_2$ — разбиения отрезка $[a;b]$, то $s(f,\mathbb{T}_1)\leqslant S(f,\mathbb{T}_2)$.*

Нижним интегралом Дарбу функции f называется величина

$$\underline{I}=\sup_{\mathbb{T}} s(f,\mathbb{T}),$$

а *верхним интегралом Дарбу* функции f называется величина

$$\overline{I}=\inf_{\mathbb{T}} S(f,\mathbb{T}).$$

Лемма 3. *Для всякого $\varepsilon>0$ найдётся такое $\delta>0$, что для всякого разбиения $\mathbb{T}$, $\lambda(\mathbb{T})<\delta$, верно неравенство*

$$\underline{I}-s(f,\mathbb{T})<\varepsilon.$$

Аналогичное утверждение выполнено для $S(f,\mathbb{T})$.

$\triangleright$ Пусть $\varepsilon>0$. Найдётся такое $\mathbb{T}_\varepsilon$, что

$$\underline{I}-s(f,\mathbb{T}_\varepsilon)<\varepsilon.$$

Пусть $\mathbb{T}$ — произвольное разбиение. Тогда

$$\underline{I}-s(f,\mathbb{T})=\underline{I}-s(f,\mathbb{T}_\varepsilon)+s(f,\mathbb{T}_\varepsilon)-s(f,\mathbb{T})<\varepsilon+s(f,\mathbb{T}_\varepsilon\cup\mathbb{T})-s(f,\mathbb{T}).$$

Рассмотрим разность $s(f,\mathbb{T}_\varepsilon\cup\mathbb{T})-s(f,\mathbb{T})$. Пусть $a=x_0<x_1<\ldots<x_N$ — точки разбиения $\mathbb{T}_\varepsilon$, Δ'_k — отрезки разбиения $\mathbb{T}_\varepsilon\cup\mathbb{T}$, Δ_k — отрезки разбиения $\mathbb{T}$. Если $x_j\notin\Delta'_k$, то Δ'_k совпадает с одним из отрезков разбиения $\mathbb{T}$. Следовательно,

$$s(f,\mathbb{T}_\varepsilon\cup\mathbb{T})-s(f,\mathbb{T})=\sum_{k:\ x_j\in\Delta'_k}\inf_{\Delta'_k} f|\Delta'_k|-\sum_{k:\ x_j\in\Delta_k}\inf_{\Delta_k} f|\Delta_k|\leqslant 4N\sup_{[a;b]}|f|\lambda(\mathbb{T}).$$

Пусть $\delta>0$ столь мало, что $4N\sup\limits_{[a;b]}|f|<\varepsilon$. Тогда для всякого $\mathbb{T}$ с условием $\lambda(\mathbb{T})<\delta$ выполняется неравенство $\underline{I}-s(f,\mathbb{T})<\varepsilon$. $\triangleleft$

Из последних двух лемм следует критерий интегрируемости.

Теорема 36 (критерий Дарбу). *Ограниченная функция f интегрируема по Риману на отрезке $[a;b]$ тогда и только тогда, когда на этом отрезке её нижний и верхний интегралы Дарбу совпадают. Более того, если f интегрируема, то её верхний и нижний интегралы Дарбу совпадают с интегралом Римана.*

$\triangleright$ Если функция f интегрируема по Риману, то для всякого $\varepsilon>0$ найдётся такое $\delta>0$, что для всех отмеченных разбиений $(\mathbb{T},\xi)$ верно двойное неравенство

$$\int\limits_a^b f(x)\,dx-\varepsilon\leqslant\sigma(f,\mathbb{T},\xi)\leqslant\int\limits_a^b f(x)\,dx+\varepsilon.$$

Поскольку $s(f,\mathbb{T}) = \inf_\xi \sigma(f,\mathbb{T},\xi)$ и $S(f,\mathbb{T}) = \sup\limits_\xi \sigma(f,\mathbb{T},\xi)$, получаем

$$\int\limits_a^b f(x)\,dx - \varepsilon \leqslant s(f,\mathbb{T}) \leqslant \underline{I} \leqslant \overline{I} \leqslant S(f,\mathbb{T}) \leqslant \int\limits_a^b f(x)\,dx + \varepsilon.$$

Следовательно, интегралы Дарбу отличаются от $\int\limits_a^b f(x)\,dx$ меньше, чем на ε. В силу произвольности ε получаем, что интегралы Дарбу совпадают с интегралом Римана.

Обратно, пусть теперь известно, что интегралы Дарбу совпадают. Докажем, что функция f интегрируема по Риману. Пусть $\underline{I} = \overline{I} = I$. Для всякого $\varepsilon > 0$ существует такое $\delta > 0$, что для всех разбиений $\mathbb{T}$ с масштабом меньшим δ верны неравенства

$$I - \varepsilon < s(f,\mathbb{T}) \leqslant S(f,\mathbb{T}) < I + \varepsilon.$$

Остаётся заметить, что для всякого отмеченного разбиения $(\mathbb{T},\xi)$ с масштабом, меньшим δ, выполнены неравенства $s(f,\mathbb{T}) \leqslant \sigma(f,\mathbb{T},\xi) \leqslant S(f,\mathbb{T})$ и $|\sigma(f,\mathbb{T},\xi) - I| < \varepsilon$. $\lhd$

Следствие 18. *Ограниченная функция f интегрируема по Риману на отрезке $[a;b]$ тогда и только тогда, когда для всякого $\varepsilon > 0$ найдётся такое разбиение $\mathbb{T}$ отрезка $[a;b]$, что*

$$S(f,\mathbb{T}) - s(f,\mathbb{T}) < \varepsilon.$$

Заметим, что

$$S(f,\mathbb{T}) - s(f,\mathbb{T}) = \sum_k (\sup_{\Delta_k} f - \inf_{\Delta_k} f)|\Delta_k|.$$

Разность $\sup\limits_{\Delta_k} f - \inf\limits_{\Delta_k} f$ обозначают через $\omega(f,\Delta_k)$ и называют *колебанием* функции f на отрезке Δ_k. Таким образом, интегрируемость ограниченной функции f равносильна тому, что для всякого $\varepsilon > 0$ найдётся такое разбиение $\mathbb{T}$ отрезка $[a;b]$, что

$$\sum_k \omega(f,\Delta_k)|\Delta_k| < \varepsilon.$$

При внимательном взгляде на выражение $\sum\limits_k \omega(f,\Delta_k)|\Delta_k|$ становится ясно, что малость этой суммы зависит от малости $\omega(f,\Delta_k)$ и от малости $|\Delta_k|$. Малость колебания функции на малых отрезках связана с непрерывностью функции. Справиться с точками разрыва функции можно с помощью малости $|\Delta_k|$, но таких точек не должно быть слишком много, иначе малые длины отрезков разбиения в сумме могут дать уже не очень малую величину. Точное описание непрерывности интегрируемой функции даёт критерий Лебега.

Множество $E \subset \mathbb{R}$ называется *множеством меры нуль по Лебегу*, если для всякого $\varepsilon > 0$ существует такой не более чем счётный набор интервалов I_k, что $E \subset \bigcup\limits_k I_k$ и $\sum\limits_k |I_k| < \varepsilon$.

Рассмотрим несколько примеров.

1. Точка является множеством меры нуль по Лебегу. Действительно, для всякого $\varepsilon > 0$ эта точка накрывается интервалом с центром в этой точке и радиусом $\varepsilon/4$.

2. Всякое не более чем счётное множество точек является множеством меры нуль по Лебегу. Пусть x_k — точки этого множества. Положим $I_k = (x_k - \varepsilon 2^{-k-2}; x_k + \varepsilon 2^{-k-2})$. Тогда интервалы I_k накрывают множество и сумма их длин равна $\varepsilon/2$.

3. Отрезок $[a;b]$, где $a < b$, не является множеством меры нуль по Лебегу. Это следует из следующего утверждения.

Лемма 4. *Если $[a;b] \subset \bigcup_k I_k$, где $\{I_k\}$ — не более чем счётная система интервалов, то $b - a \leqslant \sum\limits_k |I_k \bigcap [a;b]| \leqslant \sum\limits_k |I_k|$.*

$\rhd$ Отрезок $[a;b]$ является компактом, поэтому можно считать, что он покрыт конечным набором интервалов $I_1, \ldots, I_N$. Докажем требуемое неравенство индукцией по N. База $N = 1$ очевидна: включение $[a;b] \subset (\gamma;\delta)$ влечёт $[a;b] \bigcap (\gamma;\delta) = [a;b]$. Предположим, что для N утверждение доказано, и докажем его для $N+1$. Без ограничения общности можно считать, что $b \in I_{N+1} = (\gamma;\delta)$. Если $\gamma < a$, то утверждение очевидно выполнено. Пусть теперь $a \leqslant \gamma$. Тогда отрезок (или точка)

$[a;\gamma]$ покрыт интервалами $I_1, \ldots, I_N$, поскольку он заведомо не пересекается с интервалом I_{N+1}. По предположению индукции $\gamma - a \leqslant \sum\limits_{k=1}^{N} |I_k \bigcap [a;\gamma]|$. Промежуток $I_k \bigcap [a;\gamma]$ является подмножеством промежутка $I_k \bigcap [a;b]$, поэтому

$$\gamma - a \leqslant \sum_{k=1}^{N} |I_k \bigcap [a;b]|.$$

Прибавив к обеим частям неравенства $b-\gamma$ и замечая, что $|I_N \bigcap [a;b]| = b-\gamma$, получаем требуемое неравенство для $N+1$ интервала. $\lhd$

Следующие свойства множеств меры нуль используются наиболее часто.

Предложение 42. (i) *В определении множества меры нуль интервалы можно заменить отрезками, т. е. при такой замене получится эквивалентное определение.*

(ii) *Если D — множество меры нуль и $E \subset D$, то E — множество меры нуль.*

(iii) *Если $\{E_n\}$ — не более чем счётный набор множеств меры нуль, то $\bigcup E_n$ является множеством меры нуль.*

$\rhd$ Обоснуем пункт (i). Если множество покрыто интервалами, то оно тем более покрыто отрезками, которые получились добавлением к этим интервалам их концов. Следовательно, если множество является множеством меры нуль по определению с интервалами, то оно является множеством меры нуль по определению с отрезками. Пусть теперь некоторое множество покрыто не более чем счётной системой отрезков $[a_k;b_k]$, сумма длин которых меньше ε. Тогда это множество покрыто интервалами $(a_k - \varepsilon 2^{-k}, b_k + \varepsilon 2^{-k})$, сумма длин которых меньше 3ε. Таким образом, если множество является множеством меры нуль по определению с отрезками, то оно является множеством меры нуль по определению с интервалами.

Пункт (ii) следует из определения, а доказательство пункта (iii) дословно повторяет рассуждения из примера со счётным множеством точек. $\lhd$

Отметим, что множество меры нуль может быть континуальным. Например, множество Кантора континуально и имеет меру нуль.

Если некоторым свойством обладают точки множества, дополнение к которому является множеством меры нуль по Лебегу, то говорят, что свойство имеет место *почти всюду.*

Теорема 37 (критерий Лебега)**.** *Функция f интегрируема по Риману на $[a;b]$ тогда и только тогда, когда f ограниченна на $[a;b]$ и почти всюду непрерывна на $[a;b]$.*

$\rhd$ Поскольку для интегрируемости функции f необходима её ограниченность, далее будем считать, что функция f ограниченна. Покажем, как из интегрируемости следует, что множество точек разрыва имеет меру нуль по Лебегу. Напомним, что колебание $\omega(f,x)$ функции f в точке x определяется как предел $\lim\limits_{\delta\to 0} \omega(f, B_\delta(x))$, и непрерывность функции f в точке x равносильна равенству $\omega(f,x) = 0$. Таким образом, множество точек разрыва является объединением множеств $E_n = \{x\colon \omega(f,x) \geqslant 1/n\}$, и достаточно доказать, что каждое из этих множеств является множеством меры нуль. Для всякого $\varepsilon > 0$ существует такое разбиение $\mathbb{T}$, что

$$\sum_k \omega(f,\Delta_k)|\Delta_k| < \varepsilon.$$

Пусть в точке x колебание функции f не меньше $1/n$. Если x — внутренняя точка отрезка Δ_k, то $\omega(f,\Delta_k) \geqslant 1/n$. Если x является общей точкой отрезков Δ_{k-1} и Δ_k, то $\omega(f,\Delta_{k-1}) \geqslant 1/(3n)$ или $\omega(f,\Delta_k) \geqslant 1/(3n)$, иначе $\omega(f,x) \leqslant 2/(3n)$. Пусть J состоит из всех таких номеров k, что $\omega(f,\Delta_k) \geqslant 1/(3n)$. Тогда $E_n \subset \bigcup_{k\in J} \Delta_k$. Поскольку $1 \leqslant 3n\omega(f,\Delta_k)$ для $k \in J$, имеют место неравенства

$$\sum_{k\in J} |\Delta_k| \leqslant 3n \sum_{k\in J} \omega(f,\Delta_k)|\Delta_k| < 3n\varepsilon.$$

В силу произвольности $\varepsilon > 0$ заключаем, что E_n является множеством меры нуль.

Пусть теперь множество E точек разрыва ограниченной функции f имеет меру нуль. Докажем, что f интегрируема. Для всякого $\varepsilon > 0$ существует такая не более чем счётная система интервалов I_k, что $E \subset \bigcup I_k$ и $\sum\limits_k |I_k| < \varepsilon$. Если $x \notin E$, то f непрерывна в точке x и существует такая окрестность $B_{\delta_x}(x)$, что $\omega(f, B_{\delta_x}(x)) < \varepsilon$. Окрестности $B_{\delta_x/3}(x)$ для всех $x \notin E$ и интервалы I_k вместе образуют покрытие отрезка $[a;b]$. Существует конечное подпокрытие $B_i = B_{\delta_{x_i}/3}(x_i)$ и I_j, где $i = 1, 2, \ldots, l$ и $j = 1, 2, \ldots, m$. Пусть $\delta = \min_j \delta_{x_j}$. Рассмотрим произвольное разбиение $\mathbb{T}$ с масштабом, меньшим $\delta/3$. Имеем

$$\sum_k \omega(f, \Delta_k)|\Delta_k| = \sum_{k\colon \Delta_k \cap \cup_i B_i \neq \varnothing} \omega(f, \Delta_k)|\Delta_k| + \sum_{k\colon \Delta_k \cap \cup_i B_i = \varnothing} \omega(f, \Delta_k)|\Delta_k|.$$

Если Δ_k пересекается с некоторым $B_i = B_{\delta_{x_i}/3}(x_i)$, то $\Delta_k \subset B_{\delta_{x_i}}(x_i)$ и $\omega(f, \Delta_k) < \varepsilon$. Следовательно, имеет место оценка

$$\sum_{k\colon \Delta_k \cap \cup_i B_i \neq \varnothing} \omega(f, \Delta_k)|\Delta_k| \leqslant \varepsilon \sum_k |\Delta_k| \leqslant \varepsilon(b-a).$$

Теперь оценим сумму по тем k, для которых Δ_k не пересекается ни с каким B_j. В этом случае $\Delta_k \subset \cup I_j$ и $|\Delta_k| \leqslant \sum\limits_j |\Delta_k \cap I_j|$. Пусть $M = \sup\limits_{[a;b]} |f|$. Тогда

$$\sum_{k\colon \Delta_k \cap \cup_i B_i = \varnothing} \omega(f, \Delta_k)|\Delta_k| \leqslant 2M \sum_k \Big(\sum_j |\Delta_k \cap I_j|\Big) = 2M \sum_j \Big(\sum_k |\Delta_k \cap I_j|\Big) \leqslant 2M \sum_j |I_j| < 2M\varepsilon.$$

Итак, $\sum\limits_k \omega(f, \Delta_k)|\Delta_k| \leqslant (b - a + 2M)\varepsilon$. Согласно критерию Дарбу функция f интегрируема. $\lhd$

Следствие 19. (i) *Если функции f и g интегрируемы на $[a;b]$, то их произведение fg также интегрируемо на $[a;b]$.*

(ii) *Если функция f интегрируема на $[a;b]$ и ψ непрерывна на $[-\inf f; \sup f]$, то $\psi(f)$ интегрируема на $[a;b]$.*

Например, из интегрируемости функции f на $[a;b]$ следует интегрируемость $|f|$ на $[a;b]$. Более того, из-за монотонности интеграла верна оценка

$$\Big|\int\limits_a^b f(x)\,dx\Big| \leqslant \int\limits_a^b |f(x)|\,dx.$$

Важно иметь в виду, что из интегрируемости $|f|$ не следует интегрируемость f

Следствие 20. (i) *Если функция f интегрируема на отрезке $[a;b]$, то f интегрируема на всяком подотрезке $[c;d] \subset [a;b]$.*

(ii) *Если $a < c < b$ и функция f интегрируема на $[a;c]$ и $[c;b]$, то f интегрируема на $[a;b]$.*

Следствие 21. *Пусть f интегрируема на $[a;b]$ и $f \geqslant 0$. Тогда из равенства нулю интеграла от f по $[a;b]$ следует, что функция f равна нулю почти всюду на $[a;b]$.*

$\rhd$ Согласно критерию Лебега достаточно доказать, что f равна нулю в каждой точке непрерывности, причём достаточно рассмотреть только точки непрерывности из интервала $(a;b)$. Если f непрерывна в точке $c \in (a;b)$ и $f(c) > 0$, то существует отрезок $\Delta \subset (a;b)$ с центром в точке c, на котором $f(x) \geqslant f(c)/2$. Тогда $f(x) \geqslant \frac{f(c)}{2} I_\Delta(x)$ и в силу монотонности интеграла

$$\int\limits_a^b f(x)\,dx \geqslant \int\limits_a^b \frac{f(c)}{2} I_\Delta(x)\,dx = \frac{f(c)}{2}|\Delta| > 0.$$

Следовательно, из равенства нулю интеграла следует равенство нулю функции во всех точках непрерывности из интервала $(a;b)$, т. е. равенство нулю почти всюду. $\lhd$

Интеграл с переменным верхним пределом

Пусть f интегрируема на отрезке $[a;b]$. Функция F, заданная равенством

$$F(x) = \int\limits_a^x f(t)\,dt,$$

называется *интегралом с переменным верхним пределом.* Выше мы видели, что для непрерывной функции f интеграл с переменным верхним пределом является первообразной. Более того, если интегрируемая функция f имеет какую-то первообразную $\mathcal{F}$, то в силу формулы Ньютона — Лейбница (см. теорему 33) эта первообразная отличается лишь на константу от интеграла с переменным верхним пределом. Поэтому представляет интерес описание свойств F в случае произвольной интегрируемой функции f.

Предложение 43. *Имеет место оценка*

$$|F(x) - F(y)| \leqslant M|x - y|, \quad M = \sup|f|,$$

в частности, функция F непрерывна.

$\triangleright$ Пусть $x > y$. Тогда в силу аддитивности интеграла получаем

$$F(x) - F(y) = \int\limits_y^x f(t)\,dt$$

и, следовательно, приходим к оценке

$$|F(x) - F(y)| \leqslant \int\limits_y^x |f(t)|\,dt \leqslant M|x - y|. \qquad \triangleleft$$

Теорема 38. *Если функция f непрерывна в точке $c \in [a;b]$, то F дифференцируема в точке c и $F'(c) = f(c)$.*

$\triangleright$ Используя определение F, получаем

$$\frac{F(c+h) - F(c)}{h} - f(c) = h^{-1} \int\limits_c^{c+h} (f(t) - f(c))\,dt.$$

Таким образом,

$$\Bigl|\frac{F(c+h) - F(c)}{h} - f(c)\Bigr| \leqslant \sup_{[c;c+h]} |f(t) - f(c)|,$$

где правая часть в силу непрерывности f в точке c стремится к нулю. $\triangleleft$

По критерию Лебега интегрируемая функция почти всюду непрерывна, поэтому интеграл с переменным верхним пределом почти всюду дифференцируем. Итак, F непрерывна на $[a;b]$, почти всюду дифференцируема и почти всюду её производная совпадает с f. Однако такие функции уже не называют первообразными функции f. Это связано, в частности, с тем, что если допустить такие функции в качестве первообразных, то у нуля первообразной, кроме констант, будет являться, например, лестница Кантора, которая на константу совсем не похожа.

Формула замены переменной

Выше была доказана формула замены переменной для интеграла от непрерывной функции. Рассмотрим теперь случай интегрируемой функции. Нам потребуется следующее утверждение.

Лемма 5. *Если $\varphi\colon [a;b] \to \mathbb{R}$ — липшицева функция, то для всякого множества $E \subset [a;b]$ меры нуль по Лебегу его образ $\varphi(E)$ является множеством меры нуль по Лебегу.*

$\triangleright$ Пусть $|\varphi(x_1) - \varphi(x_2)| \leqslant M|x_1 - x_2|$ для всех $x_1, x_2 \in [a;b]$. Рассмотрим интервал $I = (c - \delta; c + \delta)$. Если $y \in \varphi(I)$, то существует такой x, что $|x - c| < \delta$ и $y = \varphi(x)$. Тогда $|\varphi(x) - \varphi(c)| \leqslant M|x - c| < 2M\delta$. Следовательно, $\varphi(I)$ содержится в интервале $(\varphi(c) - 2M\delta; \varphi(c) + 2M\delta)$. Если некоторое множество E покрыто не более чем счётной системой интервалов $I_k = (c_k - \delta_k; c_k + \delta_k)$, причём сумма их длин меньше положительного числа ε, то образ $\varphi(E)$ покрыт интервалами $J_k = (\varphi(c_k) - 2M\delta_k; \varphi(c_k) + 2M\delta_k)$, сумма длин которых меньше $2M\varepsilon$. Следовательно, если E — множество меры нуль, то и $\varphi(E)$ — множество меры нуль. $\triangleleft$

Теорема 39. *Пусть $\varphi\colon [\alpha;\beta] \to [a;b]$ — диффеоморфизм и $f\colon [a;b] \to \mathbb{R}$ — интегрируемая функция. Тогда функция $f(\varphi(t))|\varphi'(t)|$ интегрируема на $[\alpha;\beta]$ и*

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f(\varphi(t))|\varphi'(t)|\,dt.$$

$\triangleright$ Интегрируемость функции $f(\varphi(t))|\varphi'(t)|$ следует из леммы и критерия Лебега. Поскольку φ является диффеоморфизмом, на всём отрезке $[\alpha;\beta]$ или выполнено неравенство $\varphi' > 0$, или выполнено неравенство $\varphi' < 0$. Проведём доказательство для случая $\varphi' > 0$. Тогда $\varphi(\alpha) = a$ и $\varphi(\beta) = b$. Для всякого разбиения отрезка $[\alpha;\beta]$ точками $\alpha = t_0 < t_1 < \ldots < t_N = \beta$ точки $\varphi(t_k)$

образуют разбиение отрезка $[a;b]$. Поскольку $|\varphi(t_k)-\varphi(t_{k-1})| \leqslant \max|\varphi'|\cdot|t_k-t_{k-1}|$, стремление масштаба разбиения отрезка $[\alpha;\beta]$ к нулю влечёт стремление к нулю масштаба соответствующего разбиения отрезка $[a;b]$. По теореме Лагранжа для каждого k найдётся такая точка $c_k \in (t_{k-1};t_k)$, что

$$\varphi(t_k)-\varphi(t_{k-1}) = \varphi'(c_k)(t_k-t_{k-1}).$$

Для отмеченного разбиения отрезка $[\alpha;\beta]$ на отрезки $[t_{k-1};t_k]$ с отмеченными точками c_k верно равенство

$$\sum_k f(\varphi(c_k))\varphi'(c_k)(t_k-t_{k-1}) = \sum_k f(\varphi(c_k))(\varphi(t_k)-\varphi(t_{k-1})).$$

Остаётся заметить, что при стремлении масштаба разбиения к нулю левая часть равенства стремится к интегралу от $f(\varphi(t))\varphi'(t)$ (что совпадает с $f(\varphi(t))|\varphi'(t)|$) по $[\alpha;\beta]$, а правая часть стремится к интегралу от $f(x)$ по $[a;b]$.

Если же выполнено неравенство $\varphi' < 0$, то $\varphi(\alpha) = b$ и $\varphi(\beta) = a$ и, дословно повторяя проведённые рассуждения, получим то же самое итоговое равенство, но при стремлении масштаба разбиения к нулю левая часть будет по-прежнему стремиться к интегралу от $f(\varphi(t))\varphi'(t)$ по $[\alpha;\beta]$, а правая часть будет стремиться к интегралу от $f(x)$ по $[a;b]$, взятому со знаком «минус». Таким образом, доказываемая формула снова будет иметь место. $\lhd$

Приложения и обобщения интеграла Римана

Формула Тейлора. Лемма Морса.

С помощью интегрирования по частям можно вывести формулу Тейлора с остаточным членом в интегральной форме.

Предложение 44. *Если функция f непрерывно дифференцируема $(m+1)$ раз на отрезке $[a;x]$, то верно равенство*

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \ldots + \frac{1}{m!}f^{(m)}(a)(x-a)^m + \frac{1}{m!}\int_a^x (x-t)^m f^{(m+1)}(t)\,dt.$$

$\rhd$ Имеем

$$\begin{aligned} f(x) &= f(a) + \int_a^x f'(t)\,dt = f(a) + \int_a^x (t-x)'f'(t)\,dt = \\ &= f(a) + f'(a)(x-a) - \int_a^x (t-x)f''(t)\,dt = f(a) + f'(a)(x-a) - \int_a^x \frac{1}{2}\Big((t-x)^2\Big)'f''(t)\,dt = \ldots \\ &= f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \ldots + \frac{1}{m!}f^{(m)}(a)(x-a)^m + \frac{1}{m!}\int_a^x (x-t)^m f^{(m+1)}(t)\,dt. \end{aligned}$$

$\lhd$

Из одномерной формулы Тейлора получаем многомерную: пусть функция $f\colon \mathbb{R}^n \to \mathbb{R}$ непрерывно дифференцируема $(m+1)$ раз в $B_r(a)$. Тогда для всякого вектора h с условием $\|h\| < r$ верно равенство

$$f(a+h) = f(a) + df(a,h) + \ldots + \frac{1}{m!}d^m f(a,h) + \frac{1}{m!}\int_0^1 (1-t)^m d^{m+1}f(a+th,h)\,dt.$$

Действительно, достаточно записать формулу Тейлора в нуле для функции $\varphi(t) = f(a+th)$ при $t = 1$ (см. теорему 28 на с. 29).

Из многомерной формулы Тейлора с остаточным членом в интегральной форме следуют замечательные утверждения о том, как устроены функции нескольких переменных. Так, следующее утверждение является аналогом известной теоремы Безу для многочленов: если многочлен $P(x)$ имеет корень x_0, то он представим в виде $P(x) = (x-x_0)Q(x)$, где $Q(x)$ — некоторый многочлен, в частности, если $x_0 = 0$, то $P(x) = xQ(x)$.

Следствие 22. *Предположим, что $f\colon \mathbb{R}^n \to \mathbb{R}$ непрерывно дифференцируема в $B_r(0)$ и $f(0) = 0$. Тогда существуют такие непрерывные в $B_r(0)$ функции $g_1, \ldots, g_n$, что*

$$f(x) = x_1 g_1(x) + \ldots + x_n g_n(x), \quad x \in B_r(0).$$

$\triangleright$ Запишем формулу Тейлора для $m=0$, $a=0$ и $h=x$:
$$f(x)=\int_0^1\Big(\frac{\partial f}{\partial x_1}(tx)x_1+\ldots+\frac{\partial f}{\partial x_n}(tx)x_n\Big)\,dt=x_1\int_0^1\frac{\partial f}{\partial x_1}(tx)\,dt+\ldots+x_n\int_0^1\frac{\partial f}{\partial x_n}(tx)\,dt.$$
Непрерывную зависимость интегралов от x мы оставляем без доказательства. Заметим только, что это следует из перестановочности интеграла и равномерного предела. $\triangleleft$

Следствие 23. *Предположим, что $f\colon\mathbb{R}^n\to\mathbb{R}$ дважды непрерывно дифференцируема в $B_r(0)$, причём $f(0)=0$ и $\nabla f(0)=0$. Тогда существуют такие непрерывные в $B_r(0)$ функции g_{ij}, что*
$$f(x)=\sum_{i,j}g_{ij}(x)x_ix_j.$$

$\triangleright$ Достаточно записать формулу Тейлора для $m=1$, $a=0$ и $h=x$:
$$f(x)=\sum_{i,j}\Big(\frac{1}{2}\int_0^1(1-t)^2\frac{\partial^2 f}{\partial x_i\partial x_j}(tx)\,dt\Big)x_ix_j. \qquad \triangleleft$$

Из последнего утверждения выводится знаменитая лемма Морса, которая объясняет, как устроены функции нескольких переменных в окрестности точки, в которой градиент функции равен нулю, но матрица вторых производных невырожденна. Напомним, что если градиент функции отличен от нуля, то можно ввести новые координаты, в которых функция будет иметь вид простейшей линейной функции $f(x)=x_n$. Оказывается, что в окрестности точки, в которой градиент функции равен нулю и матрица вторых производных невырожденна, можно ввести новые координаты, в которых функция имеет вид квадратичной формы.

Теорема 40 (лемма Морса)**.** *Предположим, что $f\colon\mathbb{R}^n\to\mathbb{R}$ три раза непрерывно дифференцируема в окрестности нуля, причём $f(0)=0$, $\nabla f(0)=0$ и матрица вторых производных в нуле невырожденна. Тогда существуют такие окрестности нуля U, V и диффеоморфизм $x=x(u)$, отображающий U на V, что*
$$f(x(u))=u_1^2+\ldots+u_r^2-u_{r+1}^2-\ldots-u_n^2,\quad u\in U.$$

$\triangleright$ Как доказано выше, справедливо равенство
$$f(x)=\sum_{i,j}g_{ij}(x)x_ix_j.$$

Имеем $g_{ij}(0)=\frac{1}{2}\frac{\partial^2 f}{\partial x_i\partial x_j}(0)$. С помощью линейной замены координат с некоторой матрицей C можно привести квадратичную форму $Q(h)=\sum\limits_{i,j}g_{ij}(0)h_ih_j$ к виду
$$Q(h)=h_1^2+\ldots+h_r^2-h_{r+1}^2-\ldots-h_n^2.$$
Переходя от x к Cx, будем далее считать, что $Q(h)$ имеет именно такой вид. Выделим полные квадраты:
$$\sum_{i,j}g_{ij}(x)x_ix_j=\Big(g_{11}(x)^{1/2}x_1+g_{12}(x)g_{11}(x)^{-1/2}x_2+\ldots+g_{1n}(x)g_{11}(x)^{-1/2}x_n\Big)^2+\sum_{i,j\geqslant 2}\widetilde{g}_{ij}(x)x_ix_j,$$
где
$$\widetilde{g}_{ii}(x)=g_{ii}(x)-g_{1i}(x)^2g_{11}(x)^{-1},\quad \widetilde{g}_{ij}(x)=g_{ij}(x)-g_{1i}(x)g_{1j}(x)g_{11}(x)^{-1}.$$
Заметим, что $\widetilde{g}_{ij}(0)=\pm1$. Сделаем замену координат:
$$u_1=g_{11}(x)^{1/2}x_1+g_{12}(x)g_{11}(x)^{-1/2}x_2+\ldots+g_{1n}(x)g_{11}(x)^{-1/2}x_n,\ u_2=x_2,\ \ldots,\ u_n=x_n.$$
Якобиан замены в нуле равен 1, и по теореме об обратной функции это действительно является диффеоморфизмом в некоторой окрестности нуля. В новых координатах функция имеет вид
$$f(x(u))=u_1^2+\sum_{i,j\geqslant 2}\widetilde{g}_{ij}(x(u))u_iu_j.$$
Повторяя аналогичные рассуждения к $\sum\limits_{i,j\geqslant 2}\widetilde{g}_{ij}(x(u))u_iu_j$, получаем требуемый вид функции. $\triangleleft$

Интеграл Стилтьеса. Теорема Рисса.

На конечномерном пространстве $\mathbb{R}^n$ линейные функционалы устроены очень просто:
$$l(h)=a_1h_1+\ldots+a_nh_n=\langle a,h\rangle.$$

Важной задачей является описание линейных непрерывных функционалов на пространстве $C[a;b]$. Пусть функция g интегрируема на $[a;b]$. Можно показать, что отображение

$$l(f) = \int_a^b f(x)g(x)\,dx$$

является линейным непрерывным функционалом на $C[a;b]$. Однако не все линейные непрерывные функционалы могут быть так заданы. Например, функционал $l(f) = f(a)$ не может быть задан в виде интеграла от fg ни для какой интегрируемой функции g. Действительно, предположим, что

$$f(a) = \int_a^b f(x)g(x)\,dx$$

для всех $f \in C[a;b]$. Рассмотрим последовательность функций $f_n(x) = 0$ при $x > a + 1/n$ и $f_n(x) = \frac{a+1/n-x}{a+1/n}$ при $x \leqslant a + 1/n$. Имеем

$$1 = f_n(0) = \int_a^b f_n(x)g(x)\,dx \leqslant \frac{1}{n}\sup|g| \to 0.$$

Противоречие.

Для описания всех непрерывных функционалов на $C[a;b]$ используется обобщение интеграла Римана.

Пусть на отрезке $[a;b]$ определены две функции f и g. Как ранее, *разбиением* $\mathbb{T}$ отрезка $[a;b]$ будем называть набор точек $a = x_0 < x_1 < x_2 < \ldots < x_n = b$, а *отмеченным разбиением* $(\mathbb{T},\xi)$ отрезка $[a;b]$ — разбиение $\mathbb{T}$ отрезка $[a;b]$ с отмеченными точками $\xi = (\xi_1,\xi_2,\ldots,\xi_n)$, где $\xi_k \in [x_{k-1};x_k]$. Для всякого отмеченного разбиения $(\mathbb{T},\xi)$ *интегральной суммой Римана — Стилтьеса* называется выражение

$$\sigma(f dg, \mathbb{T}, \xi) = \sum_k f(\xi_k)(g(x_k) - g(x_{k-1})).$$

Если $g(x) = x$, то получаем обычную риманову сумму.

Число I называется *интегралом Римана — Стилтьеса* функции f по функции g на отрезке $[a;b]$, если для всякого $\varepsilon > 0$ существует такое $\delta > 0$, что для всех отмеченных разбиений $(\mathbb{T},\xi)$ с масштабом меньшим δ верно неравенство

$$|\sigma(f dg, \mathbb{T}, \xi) - I| < \varepsilon.$$

Интеграл Римана — Стилтьеса I обозначают через

$$\int_a^b f\,dg.$$

Аналогично тому, как это было сделано для интеграла Римана, можно показать, что интеграл Римана — Стилтьеса линеен по функции f.

Вариацией функции $g : [a;b] \to \mathbb{R}$ называется величина

$$V_a^b(g) = \sup_{\mathbb{T}} \sum_{k=1}^m |g(x_k) - g(x_{k-1})|$$

(точная верхняя грань берётся по всем разбиениям $\mathbb{T}$ отрезка $[a;b]$ точками $a = x_0 < x_1 < x_2 < \ldots < x_m = b$). Если $V_a^b(g) < +\infty$, то говорят, что функция g имеет *ограниченную вариацию* на $[a;b]$. Можно показать, что всякая функция, монотонная на отрезке, имеет на этом отрезке ограниченную вариацию.

Справедливо следующее достаточное условие существования интеграла Римана — Стилтьеса.

Предложение 45. *Предположим, что функция f непрерывна на $[a;b]$ и функция g имеет ограниченную вариацию на $[a;b]$. Тогда существует интеграл Римана — Стилтьеса от f по g, причём*

$$\left|\int_a^b f\,dg\right| \leqslant V_a^b(g)\max_{[a;b]}|f|.$$

В частности, интеграл Римана — Стилтьеса по функции ограниченной вариации определяет линейный непрерывный функционал на $C[a;b]$.

$\triangleright$ Рассмотрим два отмеченных разбиения $(\mathbb{T},\xi)$ и $(\mathbb{T}',\xi')$. Для всякого отрезка $\Delta_k = [x_{k-1};x_k]$ через $\Delta_k g$ обозначим $g(x_k) - g(x_{k-1})$. Кроме того, через Δ_{ij} обозначим пересечение Δ_i и Δ_j. Можно показать, что $\Delta_i g = \sum\limits_j \Delta_{ij} g$ и $\Delta_j' = \sum\limits_i \Delta_{ij} g$. Здесь предполагается, что для пустого Δ_{ij} величина $\Delta_{ij} g$ равна нулю. Оценим разность интегральных сумм $\sigma(fdg,\mathbb{T},\xi)$ и $\sigma(fdg,\mathbb{T}',\xi')$:

$$\Big|\sum_i f(\xi_i)\Delta_i g - \sum_j f(\xi_j')\Delta_j' g\Big| \leqslant \sum_{ij} |f(\xi_i) - f(\xi_j')||\Delta_{ij} g|.$$

Пусть Δ_{ij} непусто. Выберем точку $\xi_{ij} \in \Delta_{ij}$. Тогда

$$|f(\xi_i) - f(\xi_j')| \leqslant |f(\xi_i) - f(\xi_{ij})| + |f(\xi_{ij}) - f(\xi_j')| \leqslant \omega(f,\Delta_i) + \omega(f,\Delta_j').$$

Пусть теперь $\varepsilon > 0$. Найдём такое $\delta > 0$, что колебание функции f меньше ε на всяком отрезке, длина которого меньше δ. Для отмеченных разбиений $(\mathbb{T},\xi)$ и $(\mathbb{T}',\xi')$ с масштабом, меньшим δ, получаем оценку

$$|\sigma(fdg,\mathbb{T},\xi) - \sigma(fdg,\mathbb{T}',\xi')| \leqslant 2\varepsilon V_a^b(g).$$

Возьмём произвольную последовательность отмеченных разбиений $(\mathbb{T}^n,\xi^n)$, масштабы которых стремятся к нулю. В силу полученной выше оценки последовательность интегральных сумм $\sigma(fdg,\mathbb{T}^n,\xi^n)$ фундаментальна и, следовательно, сходится к некоторому числу I. Опять используя полученную выше оценку, заключаем, что I — интеграл Римана — Стилтьеса функции f по g.

Требуемое неравенство для интеграла получаем из неравенства для интегральной суммы:

$$|\sigma(fdg,\mathbb{T},\xi)| \leqslant \sum_k |f(\xi_k)||\Delta_k g| \leqslant \Big(\sum_k |\Delta_k g|\Big) \max_{[a;b]} |f| \leqslant V_a^b(g) \max_{[a;b]} |f|. \qquad \triangleleft$$

Следующая теорема даёт полное описание непрерывных линейных функционалов на пространстве непрерывных функций.

Теорема 41 (Рисс). *Для всякого линейного непрерывного функционала l на $C[a;b]$ существует такая функция ограниченной вариации g, что*

$$l(f) = \int_a^b f\,dg.$$

Более того, $\|l\| = V_a^b(g)$.

Доказательство этой теоремы выходит за рамки нашего курса. Отметим только, что функционал $l(f) = f(a)$ задаётся с помощью интеграла Римана — Стилтьеса по функции $g(x) = 1$ при $x > a$ и $g(a) = 0$. Действительно, распишем интегральную сумму Римана — Стилтьеса:

$$\sigma(fdg,\mathbb{T},\xi) = f(\xi_1), \quad \xi_1 \in [a;x_1].$$

Для непрерывной функции f значение $f(\xi_1)$ стремится к $f(a)$, если масштаб разбиения стремится к нулю.

Геометрические приложения интеграла Римана.

К геометрическим приложениям интеграла Римана обычно относят задачи вычисления площадей плоских фигур, объёмов тел вращения, длин кривых и др. Каждая из перечисленных задач может быть формализована (см., например, [2, гл. VI] и [7, гл. 7]). Развивая идеи, обсуждавшиеся при определении сумм Дарбу, можно дать определение и указать метод вычисления площади криволинейной трапеции. Пусть функция f определена и неотрицательна на отрезке $[a;b]$. *Криволинейной трапецией* назовём множество

$$\{(x,y) : a \leqslant x \leqslant b,\ 0 \leqslant y \leqslant f(x)\}.$$

Тогда всякая нижняя сумма Дарбу функции f на отрезке $[a;b]$ представляет собой площадь фигуры, содержащейся в этой криволинейной трапеции («вписанная» фигура), а верхняя сумма Дарбу равна площади фигуры, содержащей эту криволинейную трапецию («описанная» фигура). Фигуры, соответствующие суммам Дарбу, являются объединениями прямоугольников, а понятие площади прямоугольника мы считаем известным. Если функция f интегрируема по Риману на отрезке $[a;b]$, то согласно критерию Дарбу её интеграл по этому отрезку является одновременно точной верхней гранью площадей «вписанных» фигур и точной нижней гранью площадей «описанных» фигур её криволинейной трапеции. Поэтому естественно положить по определению, что *площадь криволинейной трапеции* равна интегралу от соответствующей функции по данному отрезку.

Несобственный интеграл Римана.

Пусть функция f определена на полуинтервале $[a;b)$, где допускается $b=+\infty$. Предположим, что для всякого $c\in(a;b)$ функция f интегрируема по Риману на $[a;c]$. Тогда определена функция

$$F(c)=\int_a^c f(x)\,dx.$$

Предел $\lim\limits_{c\to b-}F(c)$ обозначают через $\int\limits_a^b f(x)\,dx$ и называют *несобственным интегралом Римана.* Если этот предел существует и конечен, то говорят, что несобственный интеграл сходится, а функция f интегрируема в несобственном смысле. Из непрерывности интеграла с переменным верхним пределом следует, что интегрируемая по Риману на $[a;b]$ функция интегрируема на $[a;b)$ в несобственном смысле и несобственный интеграл совпадает с обычным. Используя свойства предела функции, можно показать, что несобственный интеграл является линейным и монотонным по f. Если F — первообразная интегрируемой функции f на $[a;b)$, то имеет место обобщение формулы Ньютона — Лейбница:

$$\int\limits_a^b f(x)\,dx=\lim_{c\to b-}F(c)-F(a).$$

Аналогичным образом определяется несобственный интеграл на $(a;b]$ от функции, интегрируемой по Риману на всяком отрезке $[c;b]\subset(a;b]$.

Рассмотрим следующие примеры.

1) Несобственный интеграл от $1/x^p$ по промежутку $[1;+\infty)$ сходится тогда и только тогда, когда $p>1$.

2) Несобственный интеграл от $1/x^p$ по промежутку $(0;1]$ сходится тогда и только тогда, когда $p<1$.

Список литературы

1. Фихтенгольц Г. М. Курс дифференциального и интегрального исчисления. В 3-х тт. — М.: Физматлит, 2001. — Т. 1. — 616 с.

Учебник знаменитого отечественного математика Г. М. Фихтенгольца является одним из самых популярных пособий по математическому анализу и, несмотря на несколько архаичный язык изложения, содержит большое число интереснейших результатов, которые в настоящее время не включаются в курс анализа. Кроме того, в книге разобрано много интересных примеров, знакомство с которыми будет полезно современному студенту.

2. Зорич В. А. Математический анализ. Части I и II. — Изд. 8-е, испр. — М.: МЦНМО, 2017.

Книга представляет собой одно из самых полных и удачных изложений современного курса математического анализа и, несомненно, будет полезна студентам. Обратим особое внимание на задачи и упражнения, которые приводятся после каждого параграфа. Решение этих задач поможет читателю глубоко понять и прочувствовать основные определения и методы математического анализа.

3. Архипов Г. И., Садовничий В. А., Чубариков В. Н. Лекции по математическому анализу. — Изд. 4-е, испр. — М.: Дрофа, 2004. — 640 с.

Учебник представляет собой запись реальных лекций, которые авторы многие годы читали на механико-математическом факультете МГУ имени М. В. Ломоносова, и отличается лаконичностью и отточенностью изложения. Книга содержит интересные дополнительные темы, которые не всегда включаются в курс анализа, например, теорему А. А. Карацубы о быстром умножении двух чисел и критерий Г. Вейля равномерного распределения элементов числовой последовательности.

4. Натансон И. П. Теория функций вещественной переменной. — Изд. 2-е, перераб. — М.: ГИТТЛ, 1957. — 552 с.

Книга И. П. Натансона является введением в теорию функций и, безусловно, будет полезным дополнением к базовому курсу анализа. Материал излагается ясно и очень наглядно, проводится большое число красивых фактов, которые в другой литературе не затрагиваются вовсе. Читателю также рекомендуется обратить внимание на задачи и упражнения после каждой главы.

5. Шварц Л. Анализ. I-й том. — М.: Мир, 1972. — 824 с.

Книга известного французского математика Л. Шварца может быть рекомендована как дополнение к базовому курсу математического анализа и потребует от читателя изначально глубокого понимания его основ. Однако знакомство с этим учебником позволит шире взглянуть на изучаемый материал и лучше осознать суть идей и методов, которые встречаются в базовом курсе и часто кажутся набором весьма изобретательных трюков.

6. Демидович Б. П. Сборник задач и упражнений по математическому анализу: Учеб. пособие. — 13-е изд., испр. — М.: Изд-во Московского университета; ЧеРо, 1997. — 624 с.

Это классический сборник задач по всем темам курса математического анализа, который успешно используется для проведения практических занятий на математических факультетах ведущих вузов. Читатель найдёт здесь как простые упражнения алгоритмического характера, так и сложные глубокие утверждения, доказательства которых требуют уверенного владения методами и идеями анализа.

7. Виноградова И. А., Олехник С. Н., Садовничий В. А. Математический анализ в задачах и упражнениях: В 3-х т. Том 1: Дифференциальное и интегральное исчисление. — Новое изд. — М.: Изд-во Московского университета; МЦНМО, 2017. — 412 с.

Пособие содержит отличную подборку задач по всем темам курса математического анализа и основано на обобщении авторами богатого опыта многолетнего ведения практических занятий на механико-математическом факультете МГУ имени М. В. Ломоносова. Отличительной особенностью задачника является наличие в начале каждой главы подробного введения с обзором методов и разбором примеров. Данное пособие может быть рекомендовано для самостоятельного освоения методов решения задач по математическому анализу.

8. Макаров Б. М., Голузина М. Г., Лодкин А. А., Подкорытов А. Н. Избранные задачи по вещественному анализу: Учеб. пособие для вузов. — М.: Наука. Гл. ред. физ.-мат. лит., 1992. — 432 с.

Пособие содержит разнообразные задачи, в основном — высокого уровня сложности, почти по всем темам математического анализа. Среди задач встречается ряд важных теорем, примеров и контрпримеров, знание которых существенным образом дополняет стандартный курс анализа. Кроме того, при решении задач из данного сборника можно познакомиться со многими методами и приемами, практически не обсуждаемыми на обычных занятиях по анализу.

9. Львовский С. М. Лекции по математическому анализу. — Изд. 2-е, испр. — М.: МЦНМО, 2013.— 296 с.

В книге представлены записи специального курса математического анализа, прочитанного автором в Независимом московском университете. Данный курс представляет собой оригинальное изложение как стандартных для анализа тем, так и весьма интересных и трудных дополнительных разделов, лежащих на стыке геометрии, дифференциальных уравнений и анализа. Подробно описана процедура построения вещественных чисел как классов эквивалентности фундаментальных последовательностей рациональных чисел, разбираются целые p-адические числа, вместо интеграла Римана рассматривается его более слабый аналог — интеграл Коши, обсуждается теория меры и интеграл Лебега, доказывается теорема Сарда и разбирается сложный пример Уитни, обсуждаются элементы анализа на многообразиях. Книга, несомненно, является полезным и увлекательным дополнением классического курса анализа.

10. Прасолов В. В. Элементы комбинаторной и дифференциальной топологии. — Изд. 2-е, испр. и доп. — М.: МЦНМО, 2014. — 360 с.

Книга является введением в современные методы комбинаторной топологии. Изучающему курс многомерного анализа будет очень полезно обратить внимание на вторую главу книги, посвящённую топологии конечномерного евклидова пространства. Здесь обсуждаются весьма полезные утверждения о функциях нескольких переменных, в частности, лемма Урысона и теорема Титце о продолжении непрерывной функции, число Лебега и теорема Лебега о размерности, теоремы Брауэра и лемма Жордана.

Отметим, что наряду с указанными выше подойдут и другие издания перечисленных книг. Конечно, охватить всю литературу, которая может быть полезна изучающим курс математического анализа, не представляется возможным.