МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Физический факультет

Кафедра высшей математики физического факультета

А. П. Ульянов

ОСНОВЫ МАТЕМАТИЧЕСКОГО АНАЛИЗА ДЛЯ СТУДЕНТОВ-ФИЗИКОВ

Часть 3. Функции нескольких переменных

Учебное пособие
по курсу основ математического анализа

Новосибирск
~~2013~~ 2019

УДК: 510
ББК: В14я73-1 + В151.54я73-1
У 517

Ульянов А. П. Основы математического анализа для студентов-физиков. Часть 3. Функции нескольких переменных. Учеб. пособие / Новосиб. гос. ун-т. Новосибирск, 2013.

ISBN ??

Пособие содержит конспект лекций, прочитанных автором для студентов 1-го курса физического и геолого-геофизического факультетов НГУ. Ввиду большого объёма курса и разнообразия материала пособие разделено на несколько частей. Часть 3 содержит фундамент многомерного анализа, включающий следующие темы: эвклидово пространство; гладкие функции; гладкие отображения.

Пособие предназначено для студентов первого курса физического и геолого-геофизического факультетов НГУ.

Рецензент
проф. В.А. Александров

Учебное пособие подготовлено в рамках реализации Программы развития НИУ-НГУ на 2009-2018 гг.

Версия от 24 марта 2019 г.

Глава 8. ТОПОЛОГИЯ ЭВКЛИДОВА ПРОСТРАНСТВА

8.1. Эвклидово пространство

Арифметическое пространство. Множество $\mathbb{R}$ всех вещественных чисел с обычными арифметическими операциями и есть одномерное арифметическое пространство $\mathbb{R}^1$. Геометрически оно представляется точками числовой прямой, то есть прямой с выбранной системой координат: началом, положительным направлением и единицей длины. Каждой точке соответствует её координата. лекция 23 23.11.15

Арифметическое пространство $\mathbb{R}^2 = \mathbb{R}\times\mathbb{R}$ состоит из всех (упорядоченных) пар вещественных чисел. Геометрически оно представляется точками плоскости; выбор системы декартовых координат устанавливает взаимно однозначное соответствие между точками плоскости и парами вещественных чисел:

$$\text{точка } x \leftrightarrow \text{пара координат } (x_1, x_2).$$

Вместо стрелочки, указывающей это соответствие, обычно небрежно пишут знак равенства. Нередко удобнее брать координаты (x, y), но тогда саму точку обозначают иначе. Следует быть внимательным к смыслу букв в контексте, поскольку иногда x будет означать точку, а иногда лишь её первую координату.

Арифметическое пространство $\mathbb{R}^3 = \mathbb{R}\times\mathbb{R}\times\mathbb{R}$ состоит из всех (упорядоченных) троек вещественных чисел. Геометрически оно представляется точками привычного «трёхмерного» пространства,

$$\text{точка } x \leftrightarrow \text{тройка координат } (x_1, x_2, x_3),$$

опять же посредством выбора системы координат. Нередко удобнее брать координаты (x, y, z).

Для любого положительного целого числа n арифметическое пространство $\mathbb{R}^n$ состоит из всех (упорядоченных) наборов из n вещественных чисел:

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \mid x_\alpha \in \mathbb{R}\}.$$

Непосредственная геометрическая визуализация этого пространства при $n > 3$ затруднена. Она проходит наиболее успешно лишь в тех вопросах, где ситуация в произвольной размерности n концептуально не усложняется по сравнению со случаем $n = 2$ или $n = 3$.

Векторы. Известное из элементарной геометрии понятие вектора чрезвычайно полезно и в многомерной ситуации. Напомним, что векторы бывают двух типов: приложенные и свободные.

Приложенный вектор в $\mathbb{R}^n$ есть направленный отрезок; он определяется упорядоченной парой точек — началом и концом. Координатами

рисунок

вектора из точки $x = (x_1, \ldots, x_n)$ в точку $y = (y_1, \ldots, y_n)$ называют набор разностей $(y_1 - x_1, \ldots, y_n - x_n)$. Сложение векторов соответствует сложению наборов их координат; также естественно устроено умножение вектора на число.

Свободный вектор — это класс эквивалентности приложенных векторов: эквивалентны приложенные векторы с одинаковыми наборами координат. Свободный вектор можно откладывать от любой понравившейся точки.

Ещё одной повсеместной небрежностью является отождествление точки y и вектора из начала координат в y, называемого радиус-вектором этой точки. Точка и её радиус-вектор имеют одинаковые координаты, что и провоцирует.

Радиус-векторы

$$\mathbf{e}_\alpha = (\overbrace{0, \ldots, 0}^{\alpha-1}, 1, \overbrace{0, \ldots, 0}^{n-\alpha})$$

точек, отмечающих единицы на осях координат, называют стандартными базисными векторами в $\mathbb{R}^n$, или ортами для краткости. Координаты каждого вектора непосредственно видны в его разложении по стандартному базису:

$$\mathbf{u} = (u_1, \ldots, u_n) = \sum_{1 \leqslant \alpha \leqslant n} u_\alpha \mathbf{e}_\alpha.$$

Украшения на знаке суммы вскоре надоедают и начинают постепенно опускаться: сперва границы суммирования, затем индекс суммирования, а в итоге, особенно у физиков, и сам знак суммирования.

Эвклидова геометрия. Многие важные величины элементарной геометрии, и в частности, длины и углы выражаются через стандартное скалярное произведение векторов. Для векторов $\mathbf{u} = (u_1, u_2, u_3)$ и $\mathbf{v} = (v_1, v_2, v_3)$ в пространстве $\mathbb{R}^3$ его вычисляют как

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3.$$

Для векторов $\mathbf{u} = (u_1, \ldots, u_n)$ и $\mathbf{v} = (v_1, \ldots, v_n)$ в пространстве $\mathbb{R}^n$ это

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \cdots + u_n v_n = \sum_{1 \leqslant \alpha \leqslant n} u_\alpha v_\alpha.$$

Скалярное произведение обозначают по-разному. В математике популярно обозначение $(\mathbf{u},\mathbf{v})$, которое следует уметь отличать от списка координат.

Непосредственно из координатной формулы видна симметричность скалярного произведения, $(\mathbf{u},\mathbf{v}) = (\mathbf{v},\mathbf{u})$, и его линейность по каждому сомножителю: по первому

$$(c_1\mathbf{u}_1 + c_2\mathbf{u}_2, \mathbf{v}) = c_1(\mathbf{u}_1, \mathbf{v}) + c_2(\mathbf{u}_2, \mathbf{v})$$

и аналогично по второму.

Определение. **Длиной**, или **эвклидовой нормой** вектора $\mathbf{u}$ в $\mathbb{R}^n$ для любого n называют число $\|\mathbf{u}\| = \sqrt{(\mathbf{u},\mathbf{u})}$.

Расстояние между двумя точками эквлидова пространства определяют как длину вектора из одной в другую:

$$\rho(x,y) = \|\mathbf{x} - \mathbf{y}\| = \Big(\sum_{1\leqslant\alpha\leqslant n} (x_\alpha - y_\alpha)^2\Big)^{1/2}.$$

Здесь в левой части точки, в средней их радиус-векторы, а в правой координаты.

Угол φ между ненулевыми векторами $\mathbf{u}$ и $\mathbf{v}$ вычисляют по формуле

$$\cos\varphi = \frac{(\mathbf{u},\mathbf{v})}{\|\mathbf{u}\|\cdot\|\mathbf{v}\|}.$$

Уверенность в том, что для этого числа всегда найдётся соответствующий угол, возникает благодаря важному неравенству.

Теорема (Неравенство Коши)**.** *Всегда выполнено неравенство*

$$|(\mathbf{u},\mathbf{v})| \leqslant \|\mathbf{u}\|\cdot\|\mathbf{v}\|.$$

Можно написать и более изящное

$$(\mathbf{u},\mathbf{v})(\mathbf{v},\mathbf{u}) \leqslant (\mathbf{u},\mathbf{u})(\mathbf{v},\mathbf{v}).$$

Доказательство. Для всех $t \in \mathbb{R}$ имеем $\|t\mathbf{u}+\mathbf{v}\|^2 = (t\mathbf{u}+\mathbf{v}, t\mathbf{u}+\mathbf{v}) \geqslant 0$. Пользуясь линейностью и симметричностью, получим

$$t^2(\mathbf{u},\mathbf{u}) + 2t(\mathbf{u},\mathbf{v}) + (\mathbf{v},\mathbf{v}) \geqslant 0.$$

Сохранение знака происходит только при отрицательном дискриминанте квадратного трёхчлена; но $D/4 = (\mathbf{u},\mathbf{v})^2 - (\mathbf{u},\mathbf{u})(\mathbf{v},\mathbf{v})$.

Совершенно иное доказательство (матричное) представлено в лекциях по алгебре и геометрии за второй семестр; на мой вкус оно лучше. Здесь дискриминант, там детерминант. $\square$

Упражнение. *Когда возможно достижение равенства?*

Теорема (Неравенство треугольника)**.** *Всегда выполнено неравенство*

$$\|\mathbf{u}+\mathbf{v}\| \leqslant \|\mathbf{u}\| + \|\mathbf{v}\|.$$

Доказательство. Раскроем квадрат левой части и применим неравенство Коши:

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 = (\mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v}) = (\mathbf{u},\mathbf{u}) + 2(\mathbf{u},\mathbf{v}) + (\mathbf{v},\mathbf{v})\\ \leqslant \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 = (\|\mathbf{u}\| + \|\mathbf{v}\|)^2.\end{aligned}$$

После извлечения корня отсюда следует неравенство треугольника, ибо обе его части по определению неотрицательны. □

Нормы. Выпишем абстрактные свойства эвклидовой нормы:

- $\|\mathbf{u}\| \geqslant 0$;
- $\|\mathbf{u}\| = 0 \;\Leftrightarrow\; \mathbf{u} = \mathbf{0}$;
- $\|\lambda\mathbf{u}\| = |\lambda|\cdot\|\mathbf{u}\|$;
- $\|\mathbf{u}+\mathbf{v}\| \leqslant \|\mathbf{u}\| + \|\mathbf{v}\|$.

В терминах расстояния между точками эти свойства, кроме третьего, становятся

- $\rho(x,y) \geqslant 0$;
- $\rho(x,y) = 0 \;\Leftrightarrow\; x = y$;
- $\rho(x,z) \leqslant \rho(x,y) + \rho(y,z)$.

Чего ещё мы ждём от функции, измеряющей расстояния?

Ситуация хорошо иллюстрирует типичный ход мысли математика, делающего маленький шаг в развитии своей науки. Четыре выписанных свойства нормы объявляются своего рода аксиомами, а любая функция, обладающая ими, называется **нормой**. Каждой норме соответствует своя функция расстояния, или **метрика**; наоборот не всегда, хотя бы из-за одного пропущенного свойства.

Пример. Укажем две метрики, соответствующие другим простым и важным нормам (эвклидово расстояние это ρ_2):

$$\rho_1(x,y) = \sum_{1\leqslant\alpha\leqslant n} |x_\alpha - y_\alpha|; \qquad \rho_\infty(x,y) = \max_{1\leqslant\alpha\leqslant n} |x_\alpha - y_\alpha|.$$

Упражнение. *Проверьте неравенство треугольника для ρ_1 и ρ_∞.*

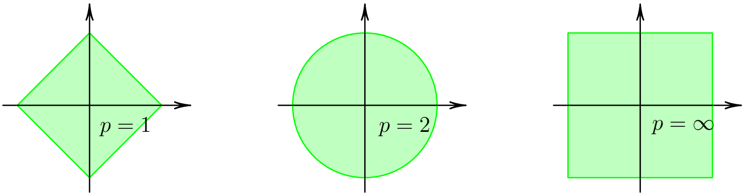
Пример. Для каждого вещественного числа $1 \leqslant p < +\infty$ функция

$$\|\mathbf{u}\| = \Big(\sum_{1\leqslant\alpha\leqslant n} |u_\alpha|^p\Big)^{1/p}$$

является нормой. Для $p \neq 1, 2$ они менее употребительны.

Нестандартные способы измерения расстояния имеют свои области приложения, но они победнее, чем у эвклидовой нормы.

Шарики и кубики. Наглядное представление о норме может дать определяемая ею геометрическая фигура. Множество всех точек, находящихся по норме не дальше от начала координат, чем фиксированное число, называют **шаром** нормы. Для эвклидовой нормы в $\mathbb{R}^3$ получается обычный шар. Для «бесконечной» нормы из примера получается куб, а для первой — октаэдр (упражнение).



Полезно посмотреть и на всё семейство p-норм. Шар p-нормы при $p > 2$ имеет угловатую форму с округлёнными углами, нечто между обычным шаром и кубом. Округлённости заостряются при росте p, и в пределе при $p \to +\infty$ шар p-нормы становится кубом. При $p \to 1$ происходит аналогичное заострение, но в других точках.

В дальнейшем мы будем часто употреблять (открытые) маленькие ε-шарики и ε-кубики, поэтому введём обозначения

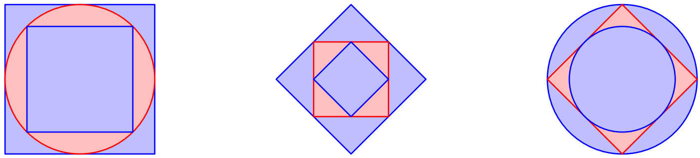
$$B_\varepsilon(z) = \{x \in \mathbb{R}^n \mid \rho(z, x) < \varepsilon\},$$
$$C_\varepsilon(z) = \{x \in \mathbb{R}^n \mid \rho_\infty(z, x) < \varepsilon\}.$$

Их роль в том, что это ε-окрестности точки при разных способах измерения расстояния. Шарики кажутся привычнее, но для некоторых целей кубики удобнее. Возникает вопрос: насколько сильно могут отличаться теории многомерного анализа, построенные на шариках и на кубиках? Ответ замечателен: они не будут отличаться.

Теорема (Эквивалентность норм)**.** *Для каждого положительного числа n найдётся такая постоянная $M = M(n)$, что в $\mathbb{R}^n$ верно*

$$B_\varepsilon(z) \subset C_{M\varepsilon}(z), \quad C_\varepsilon(z) \subset B_{M\varepsilon}(z).$$

Упражнение. *Укажите подходящее значение $M(n)$.*



Значение постоянной $M(n)$ необходимо редко. Причина в том, что в анализе при рассуждениях вблизи некоторой точки характерным шагом является переход к меньшей окрестности, как только это потребуется. Содержание же теоремы — внутри маленького шарика найдётся ещё меньший кубик и наоборот. Вот и всё, что нам нужно, чтобы всегда свободно применять наиболее удобную форму окрестностей.

Эта теорема на самом деле гораздо более общая: на арифметическом пространстве $\mathbb{R}^n$ эквивалентны не только эти две нормы, но вообще все нормы. Однако в этом курсе мы не будем пользоваться прочими.

8.2. Предел и непрерывность

Предел последовательности. Понятие сходимости последовательности точек в $\mathbb{R}^k$ легко получить, напрямую обобщая происходящее на числовой прямой $\mathbb{R}^1$. Видно два естественных пути обобщения.

На одном пути достаточно заменить способ измерения расстояния. Если на прямой этой цели служит модуль разности, то в эвклидовом пространстве — норма разности как векторов. Тогда $x_n \to x$ означает

$$\forall \varepsilon > 0 \ \exists N = N(\varepsilon)\colon n > N \Rightarrow \|x_n - x\| < \varepsilon.$$

С помощью шариков запишем это формально как

$$\forall \varepsilon > 0 \ \exists N = N(\varepsilon)\colon n > N \Rightarrow x_n \in B_\varepsilon(x),$$

а неформально выразим словами, что любой (сколь угодно малый) шарик вокруг точки x включает хвост последовательности $\{x_n\}$.

рисунки

Другой путь покоординатный. Скажем, что $x_n \to x$ тогда и только тогда, когда по каждой координате $x_{n,\alpha} \to x_\alpha$. Формально

$$\forall \varepsilon > 0 \ \exists N_\alpha = N_\alpha(\varepsilon)\colon n > N_\alpha \Rightarrow |x_{n,\alpha} - x_\alpha| < \varepsilon.$$

Выбирая тогда $N(\varepsilon) = \max\{N_\alpha(\varepsilon)\}$, перепишем это как

$$\forall \varepsilon > 0 \; \exists N = N(\varepsilon) \colon n > N \Rightarrow x_n \in C_\varepsilon(x),$$

а неформально выразим словами, что любой (сколь угодно малый) кубик вокруг точки x включает хвост последовательности $\{x_n\}$.

Поскольку маленькие шарики и кубики вкладываются друг в друга, указанные пути равносильны: они приводят к идентичным понятиям сходимости.

Непрерывность функции. Интуитивное представление о непрерывности функции нескольких переменных не отличается от одномерного: если $x \approx p$, то $f(x) \approx f(p)$.

Формализация в подходе Гейне происходит автоматически, потому что опирается на предел последовательности: функция $f \colon X \to \mathbb{R}$ по определению непрерывна в точке $p \in X$, когда $f(x_n) \to f(p)$ для любой последовательности $\{x_n\}$ в X, сходящейся к p.

Формализация в подходе Коши явным образом ссылается на расстояние между точками. Значит, для обобщения на многомерный случай нужно опять же заменить способ измерения расстояния. Функция $f \colon X \to \mathbb{R}$ по определению непрерывна в точке $p \in X$, когда

$$\forall \varepsilon > 0 \; \exists \delta = \delta(\varepsilon) > 0 \colon \|x - p\| < \delta \Rightarrow |f(x) - f(p)| < \varepsilon.$$

Внешнее отличие от одномерного случая сводится к удвоению пары палочек. Переписав определение в терминах окрестностей, мы и вовсе избавимся от внешних изменений, ибо все модификации укладываются в расширение смысла букв:

$$\forall \varepsilon > 0 \; \exists \delta = \delta(\varepsilon) > 0 \colon x \in B_\delta(p) \cap X \Rightarrow f(x) \in B_\varepsilon(f(p)).$$

Доказательство равносильности подходов Коши и Гейне не отличается от одномерного случая. Такова же ситуация с проверкой того, что арифметические операции с непрерывными функциями дают непрерывные функции.

По сравнению с одномерным, в многомерном случае пропадает изучение односторонних пределов, потому что фактически оно лишено особого смысла, хотя иногда пределы по направлению нужны.

Пример. На плоскости рассмотрим функцию

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

Вне начала координат она непрерывна, но при приближении к нему вдоль различных прямых имеет различные постоянные значения от -1 до 1. Поэтому в начале координат эта функция терпит разрыв.

Непрерывность отображения. Следующий шаг — переход от числовой функции на множестве $X \subseteq \mathbb{R}^k$ к отображению $f\colon X \to \mathbb{R}^m$. На уровне определений он не требует ничего, кроме расширения смысла букв: в подходе Гейне изменений нет; в подходе Коши нужно заменить и второе вхождение модуля на норму, хотя и эта модификация пропадает из виду при употреблении шариков.

В то же время, для облегчения восприятия, особенно поначалу, бывает полезно «разбирать» отображение на части. Аргументы его есть точки $x = (x_1, \ldots, x_k)$, а результаты есть точки $f(x) = (f_1, \ldots, f_m)$, где каждая координата является функцией $f_\beta\colon X \to \mathbb{R}$,

$$x = (x_1, \ldots, x_k) \mapsto f_\beta(x) = f_\beta(x_1, \ldots, x_k).$$

Их называют **координатными функциями** отображения f.

Лемма. *В заданной точке, либо на заданном множестве, отображение непрерывно $\Leftrightarrow$ все его координатные функции непрерывны.*

Доказательство. Прямо следует из определений при использовании кубических окрестностей или подхода Гейне. □

Опять же автоматически получается непрерывность композиции непрерывных отображений. Вопрос об обратном отображении гораздо сложнее и потому откладывается.

8.3. Открытые и замкнутые множества

лекция 24
26.11.15

Понятия, вводимые в большом количестве в этом и следующем разделах, составляют основу языка математического анализа. Проникали они и в наше предыдущее знакомство с одномерным анализом, но там многие из них недостаточно богато раскрываются из-за тесноты одномерного пространства. Рассматриваемые здесь типы множеств тут же связываются с многомерными обобщениями основных теорем о непрерывных функциях на отрезке.

правки
01.12.15

Различные типы точек. В пространстве $\mathbb{R}^k$ зафиксируем множество X и дадим названия различным точкам пространства, выражающие расположение точки относительно X. Идеи здесь аналогичны делению точек множества на точки сгущения и изолированные, но есть отличия. Обозначим через $\neg X$ множество $\mathbb{R}^k \setminus X$ всех точек пространства, не лежащих в X; его называют дополнением X.

Определение. **Расстоянием** от точки z до множества $X \subseteq \mathbb{R}^k$ называют

$$\rho(z, X) = \inf\{\rho(z, x) \mid x \in X\}.$$

Отметим, что $X \subset Y$ влечёт $\rho(z, X) \geqslant \rho(z, Y)$ для любой точки z.

Определение. Если всякий шарик $B_\varepsilon(z)$ пересекает множество X, то z называют точкой **прикосновения** для X. Это требование равносильно условию $\rho(z, X) = 0$.

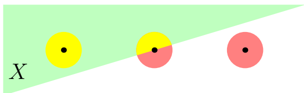
Если тут заменить шарик на выколотый шарик $B_\varepsilon(z)\setminus\{z\}$, то вместо точки прикосновения получится точка сгущения, также называемая **предельной**. Если же точка $x \in X$ не предельная, то она изолированная; это равносильно условию $\rho(x, X \setminus \{x\}) > 0$.

Упражнение. *Точные грани непустого ограниченного числового множества являются его точками прикосновения.*

Упражнение. *Всякая точка произвольного множества $X \subseteq \mathbb{R}^k$ является его точкой прикосновения.*

Упражнение. *Всякая точка прикосновения множества X является пределом сходящейся последовательности точек X, и наоборот.*

Разобьём всё пространство на три множества: внутренние точки X; граничные точки X; точки, отдалённые от X. Последние обычно не играют роли в анализе, поэтому общепринятого термина для них нет. Тип точки определяется ε-шариками вокруг неё.



Определение. Для множества $X \subseteq \mathbb{R}^k$ точка z является:

- **внутренней**, если $B_\varepsilon(z) \subset X$ для некоторого $\varepsilon > 0$;
- **граничной**, если всякий шарик $B_\varepsilon(z)$ пересекает и X, и $\neg X$;
- **отдалённой**, если $B_\varepsilon(z) \subset \neg X$ для некоторого $\varepsilon > 0$.

Внутренние точки задаются условием $\rho(z, \neg X) > 0$, отдалённые — условием $\rho(z, X) > 0$, а граничные — условием $\rho(z, X) = \rho(z, \neg X) = 0$. Внутренние точки множества обязательно ему принадлежат, отдалённые обязательно не принадлежат, а граничные ничего не обязаны в плане принадлежности, но каждая граничная точка для X является точкой прикосновения одновременно для X и для его дополнения.

Внутренность, граница, замыкание. Основываясь на классификации точек, введём новые термины для множеств.

Определение. Множество всех внутренних точек X называют **внутренностью** множества X и обозначают через X° (иногда иначе, например, $\operatorname{Int} X$ от слова *interior*).

Определение. Множество всех граничных точек X называют **границей** множества X и обозначают через ∂X.

Определение. Множество всех точек прикосновения X называют **замыканием** множества X и обозначают через $\overline{X}$ (иногда иначе, например, $\operatorname{Cl} X$ от слова *closure*).

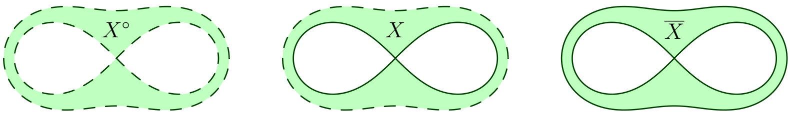
Поскольку всегда $x \in B_\varepsilon(x)$, определения влекут $X^\circ \subseteq X \subseteq \overline{X}$.

Пример. Для полуинтервала $X = [a, b) \subset \mathbb{R}$ получаем

$$X^\circ = (a, b), \quad \overline{X} = [a, b], \quad \partial X = \{a, b\}.$$

Пример. Для множества всех рациональных чисел $\mathbb{Q} \subset \mathbb{R}$ получаем

$$\mathbb{Q}^\circ = \varnothing, \quad \overline{\mathbb{Q}} = \mathbb{R}, \quad \partial\mathbb{Q} = \mathbb{R}.$$



Упражнение. *Проверьте, что*

$$X^\circ = X \setminus \partial X, \quad \overline{X} = X \cup \partial X, \quad \partial X = \overline{X} \setminus X^\circ.$$

Особый интерес представляют множества, для которых в цепочке $X^\circ \subseteq X \subseteq \overline{X}$ одно из включений реализуется как равенство.

Baire 1899

Определение. Множество $X \subseteq \mathbb{R}^k$ называют **открытым**, если $X^\circ = X$.

Открытость можно выразить иными словами: все точки открытого множества внутренние; все граничные точки открытого множества не принадлежат ему.

Cantor 1884

Определение. Множество $X \subseteq \mathbb{R}^k$ называют **замкнутым**, если $X = \overline{X}$.

Замкнутость можно выразить иными словами: замкнутое множество содержит предел всякой сходящейся последовательности своих точек; все граничные точки замкнутого множества принадлежат ему.

Упражнение. *Открыты ли и замкнуты ли: всё пространство $\mathbb{R}^k$; одноточечное множество $\{x\}$; пустое множество $\varnothing$?*

Упражнение. *Проверьте, что $X \subseteq \mathbb{R}^k$ открыто $\Leftrightarrow \mathbb{R}^k \setminus X$ замкнуто, и наоборот.*

Упражнение. *Найдите все такие $X \subseteq \mathbb{R}^k$, что $X^\circ = X = \overline{X}$.*

Множества, заданные неравенствами. Для любого числа c определим **лебеговские множества** функции f на множестве $X \subseteq \mathbb{R}^k$:

$$\{f < c\} = \{x \in X \mid f(x) < c\}$$

и аналогично $\{f > c\}$, $\{f \leqslant c\}$, $\{f \geqslant c\}$ и $\{f = c\}$.

Пример. Замкнутый единичный круг на плоскости описывают как множество $\{(x, y) \mid x^2 + y^2 \leqslant 1\}$. Открытый единичный круг задаётся так же, но со строгим неравенством. Граница круга есть единичная окружность; она задаётся тем же условием с равенством.

Следствие. *Для непрерывной функции $f\colon X \to \mathbb{R}$:*

(1) *если X открыто, то $\{f < c\}$ и $\{f > c\}$ открыты;*
(2) *если X замкнуто, то $\{f \leqslant c\}$, $\{f \geqslant c\}$ и $\{f = c\}$ замкнуты.*

Доказательство. Речь идёт о прообразах открытых и замкнутых бесконечных промежутков, а также одноточечного множества. Ниже мы получим общую лемму о прообразах при непрерывном отображении, из неё это утверждение и будет следствием. $\square$

Конкретные множества в анализе часто возникают в такой форме, причём функций может быть несколько и тогда берётся пересечение. Их ещё называют множествами, заданными неравенствами. Набор из только строгих неравенств приводит к открытости, набор из только нестрогих — к замкнутости.

Пример. Множество $X \subset \mathbb{C}$ на предыдущем рисунке задано условием $1 \leqslant |z^2 - 1| < 1.2$, которое несложно выразить в вещественном виде. Оно не открыто и не замкнуто; часть его границы входит в само множество, а часть нет.

Во втором семестре часто будут полезны множества вида $M \cap B_\varepsilon(p)$, где множество M является линией или поверхностью, заданной уравнением $F = 0$, а точка p лежит на ней. Это окрестности точки на фигуре. Их также называют открытыми в M, хотя они не открыты как подмножества всего эвклидова пространства.

Пересечения, объединения, произведения. Открытость и замкнутость хорошо себя ведут при основных операциях на множествах. Можно провести аналогию с аксиомами нормы: на более высоком и очень абстрактном уровне, куда мы не планируем выходить в этом курсе, эти операционные свойства кладут в основу общей топологии.

Лемма. *Всегда открыто:*

(1) *объединение произвольного семейства открытых множеств;*
(2) *пересечение конечного набора открытых множеств.*

Лемма. *Всегда замкнуто:*

(1) *пересечение произвольного семейства замкнутых множеств;*
(2) *объединение конечного набора замкнутых множеств.*

Пример. Пересечение бесконечного семейства открытых промежутков $(-\frac{1}{n}, 1+\frac{1}{n})$ для всех положительных целых чисел n не открыто; оно равно отрезку $[0,1]$.

Наоборот, объединение бесконечного семейства замкнутых промежутков $[\frac{1}{n}, 1-\frac{1}{n}]$ не замкнуто; оно равно интервалу $(0,1)$.

Доказательство первой леммы. (1) Нужно установить, что вместе с каждой точкой x объединение включает некоторый шарик $B_\varepsilon(x)$. Однако x принадлежит хотя бы одному множеству семейства, которое, ввиду своей открытости, включает требуемый шарик.

(2) Нужно установить, что вместе с каждой точкой x пересечение включает некоторый шарик $B_\varepsilon(x)$. Однако x принадлежит каждому множеству набора, которые, ввиду своей открытости, включают требуемые шарики с центрами в x. Пересечение гарантированно включает наименьший из них.

Заменить конечный набор на бесконечное семейство тут нельзя, ибо наименьшего шарика тогда может не быть. □

Доказательство второй леммы. Следует из первой леммы, одного из упражнений и свойств операций $\cap$, $\cup$, $\setminus$: дополнение замкнутого открыто и наоборот; дополнение пересечения равно объединению дополнений и наоборот. □

Упражнение. *Докажите вторую лемму напрямую.*

Лемма.

(1) *Если множества X и Y открыты, то $X \times Y$ открыто.*
(2) *Если множества X и Y замкнуты, то $X \times Y$ замкнуто.*

Здесь речь может идти об $X \subseteq \mathbb{R}^k$ и $Y \subseteq \mathbb{R}^m$; тогда $X \times Y \subseteq \mathbb{R}^{k+m}$.

Доказательство. (1) Здесь удобно пользоваться кубиками вместо шариков. Нужно установить, что $X \times Y$ вместе с каждой своей точкой (x, y) включает некоторый кубик с центром в ней. Ввиду открытости сомножителей, найдутся кубики $C_\varepsilon(x) \subset X$ и $C_\varepsilon(y) \subset Y$. Их произведение есть требуемый кубик в $X \times Y$. рисунок

(2) Нужно установить, что $X \times Y$ содержит все свои предельные точки. Возьмём в нём последовательность пар $\{(x_n, y_n)\}$, сходящуюся к точке (x, y). Тогда $x_n \to x$ и $y_n \to y$. Ввиду замкнутости сомножителей получаем $x \in X$ и $y \in Y$, поэтому $(x, y) \in X \times Y$. □

Произведение открытого и замкнутого множеств, с некоторыми исключениями, не открыто и не замкнуто, ибо включает только часть своей границы. Это видно на примере отрезка и интервала.

Упражнение. *Проверьте равенство*

$$\partial(X \times Y) = (\partial X \times Y) \cup (X \times \partial Y).$$

Упражнение. *Докажите, что граница всегда замкнута.*

Непрерывность и прообразы. Отображение $f\colon X \to \mathbb{R}^m$ по определению непрерывно в точке $p \in X$, когда для любого $\varepsilon > 0$ существует такое $\delta = \delta(\varepsilon) > 0$, что $x \in B_\delta(p) \cap X$ влечёт $f(x) \in B_\varepsilon(f(p))$. Окончание формулировки равносильно более ёмкому $f(B_\delta(p) \cap X) \subseteq B_\varepsilon(f(p))$. Когда X открыто, дополнительным уменьшением шарика всегда можно добиться включения $B_\delta(p) \subset X$, поэтому определение становится более красивым:

$$\forall \varepsilon > 0\ \exists \delta = \delta(\varepsilon) > 0\colon f(B_\delta(p)) \subseteq B_\varepsilon(f(p)).$$

Лемма. *Даны множества $X \subseteq \mathbb{R}^k$ и $Y \subseteq \mathbb{R}^m$ и непрерывное отображение $f\colon X \to \mathbb{R}^m$.*

(1) *Если X и Y открыты, то прообраз $f^{-1}(Y)$ открыт.*
(2) *Если X и Y замкнуты, то прообраз $f^{-1}(Y)$ замкнут.*

Доказательство. (1) Нужно установить, что $f^{-1}(Y)$ вместе с каждой своей точкой x включает некоторый шарик вокруг неё. Ввиду открытости Y, найдём такое $\varepsilon > 0$, что $B_\varepsilon(f(x)) \subset Y$. Затем по непрерывности найдём такое $\delta = \delta(\varepsilon) > 0$, что $f(B_\delta(x)) \subseteq B_\varepsilon(f(x))$. Значит, $B_\delta(x) \subseteq f^{-1}(Y)$. Открытость X спрятана в предыдущей формуле.

(2) Возьмём в $f^{-1}(Y)$ сходящуюся последовательность $\{x_n\}$. Ввиду замкнутости X, её предел x принадлежит X. По непрерывности в версии Гейне, $f(x_n) \to f(x)$. Однако $f(x_n) \in Y$, поэтому $f(x) \in Y$ ввиду замкнутости Y. Значит, $x \in f^{-1}(Y)$, что и требовалось. □

лекция 25
30.11.15

Следует отметить особую роль первого утверждения этой леммы. Дело в том, что эволюция формального понятия непрерывности привела математиков к тому, чтобы именно это свойство считать определением непрерывности (на множестве): отображение называют непрерывным, если прообраз каждого открытого множества открыт.

Лемма. *Непрерывность в этом смысле равносильна введённой ранее.*

Доказательство. Равносильность означает две взаимно обратных импликации. Одну из них обеспечивает предыдущая лемма, так что осталось установить другую.

Возьмём отображение $f\colon X \to Y$, при котором прообраз каждого открытого множества в Y открыт в X, а также точку $x \in X$. Для любого $\varepsilon > 0$ шарик $B_\varepsilon(f(x))$ открыт, поэтому его прообраз тоже открыт, а значит, включает некоторый шарик $B_\delta(x)$. Тем самым, истинно высказывание, определяющее непрерывность отображения f в точке x в прежнем смысле. □

Упражнение. *Найдите примеры таких непрерывных отображений, при которых образ открытого множества не обязательно открыт, а образ замкнутого множества не обязательно замкнут. Указание: далеко ходить не надо.*

Линейно связные множества. Зафиксируем множество $X \subseteq \mathbb{R}^k$. Образ непрерывного отображения $\gamma\colon [0,1] \to X$ называют **путём** в X. Отрезок $[0,1]$ здесь выбран совершенно произвольно: единственная причина в том, что это «стандартный» отрезок. Точки $\gamma(0)$ и $\gamma(1)$ называются соответственно **началом** и **концом** пути γ. Говорят, что путь γ **соединяет** эти точки. Как правило, важно, что путь не покидает X. Никакие формулы для γ обычно не возникают, поскольку существенен лишь сам факт существования пути.

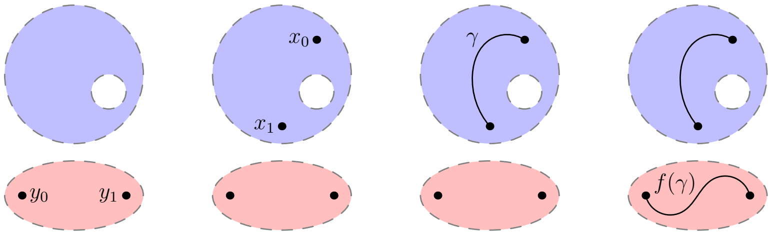
рисунок

Определение. Множество называют **линейно связным**, если любую пару его точек можно соединить непрерывным путём. Иначе множество называют **линейно несвязным**.

Обобщением какого важного в одномерном анализе понятия являются линейно связные множества?

Упражнение. *Опишите все линейно связные множества на числовой прямой* $\mathbb{R}^1$.

Лемма. *Всякий непрерывный образ линейно связного множества линейно связен.*



Доказательство. Возьмём линейно связное множество X и непрерывное отображение $f\colon X \to f(X) = Y$. Чтобы установить связность Y, нужно каждую пару точек $y_i = f(x_i)$ суметь соединить непрерывным путём в Y. Ввиду линейной связности X найдётся путь $\gamma\colon [0,1] \to X$, соединяющий точки x_i. Тогда композиция $f \circ \gamma\colon [0,1] \to Y$ непрерывна и даёт требуемый непрерывный путь в Y, соединяющий точки y_i. $\square$

Следствие. *Для всякого линейно связного множества* X *и всякой непрерывной функции* $f\colon X \to \mathbb{R}$ *верно*

$$y_0, y_1 \in f(X) \;\Rightarrow\; [y_0, y_1] \subseteq f(X).$$

Это утверждение является многомерным аналогом теоремы о промежуточных значениях.

Доказательство. Упражнение. $\square$

Упражнение. *Докажите, что линейно связное множество нельзя разбить на два непересекающихся открытых множества.*

8.4. Компактные множества

Ограниченные множества. Ограниченные числовые множества у нас определялись как имеющие конечные нижние и верхние грани. Этот подход не работает в многомерной ситуации, поскольку нет никакой разумно устроенной операции сравнения точек, а значит, нет никаких «граней». Поэтому определение будет иным.

Определение. Множество $X \subset \mathbb{R}^k$ называется **ограниченным**, если оно целиком заключено в некотором шаре (или кубе, что равносильно): существуют такая точка $c \in \mathbb{R}^k$ и такое число $R > 0$, что $X \subset B_R(c)$.

Аналогично, последовательность точек называется ограниченной, если вся она лежит в некотором шарике.

Лемма. *Всегда ограничены:*

(1) *произвольное подмножество ограниченного множества;*
(2) *пересечение любого семейства ограниченных множеств;*
(3) *объединение конечного набора ограниченных множеств;*
(4) *произведение ограниченных множеств.*

Доказательство. Упражнения. □

Теорема. *Всякая ограниченная последовательность в $\mathbb{R}^k$ имеет сходящуюся подпоследовательность, или, равносильно, точку сгущения.*

Первое доказательство. Предположим сначала, что $k = 2$. Если последовательность пар $\{(x_n, y_n)\}$ ограничена, то ограничены и последовательности координат $\{x_n\}$ и $\{y_n\}$. К ним применим известную нам одномерную версию доказываемой теоремы. Именно: выберем сперва такую подпоследовательность пар $\{(x'_n, y'_n)\}$, что сходится последовательность $\{x'_n\}$; затем уже из $\{(x'_n, y'_n)\}$ выберем такую подпоследовательность $\{(x''_n, y''_n)\}$, что сходится $\{y''_n\}$. Теперь обе последовательности координат сходятся, поэтому $\{(x''_n, y''_n)\}$ сходится.

Обозначения подпоследовательностей вспомогательные, и суть вовсе не в них. К подпоследовательности всегда переходят, прореживая последовательность индексов. Здесь на первом шаге условие прореживания определяют первые координаты, а на втором — вторые. Это же рассуждение работает в любой размерности k, но состоит оно из k таких шагов прореживания исходной последовательности. □

Второе доказательство. Предположим сначала, что $k = 2$. Возьмём замкнутый квадрат K_0, содержащий всю последовательность. Разделим его на четыре части координатными линиями через центр. Хотя бы одна из частей содержит бесконечно много элементов последовательности. Обозначим одну такую (замкнутую) часть через K_1, затем повторим деление и выбор бесконечное количество раз. В итоге получим последовательность $K_0 \supset K_1 \supset K_2 \supset \ldots$ вложенных квадратов. Их общая точка является для исходной последовательности точкой сгущения, существование которой и требовалось доказать.

Аналогичное рассуждение работает в любой размерности k, только на каждом шаге кубик делится на 2^k вдвое меньших кубиков. $\square$

Упражнение. *Второе доказательство опирается на «принцип вложенных кубиков». Сформулируйте его явно и затем докажите.*

Компактные множества. Предыдущая теорема на самом деле говорит, что замкнутый кубик в $\mathbb{R}^k$, как и отрезок в $\mathbb{R}^1$, является компактом. Это особенно сильно просвечивает между строк её второго доказательства. Теперь, наконец, пора разобраться с этим понятием.

Очень общих определений компактного множества несколько, но в эвклидовом пространстве все они равносильны.

Теорема (Критерий компактности)**.** *Следующие условия на множество $X \subseteq \mathbb{R}^k$ равносильны:*

(1) *оно замкнуто и ограничено;*
(2) *из всякой последовательности его точек можно извлечь подпоследовательность, сходящуюся к некоторой его точке;*
(3) *из всякого его покрытия открытыми множествами можно извлечь конечное подпокрытие.*

Полезность первого условия — в его простоте и сравнительной наглядности. Второе и третье условия возникли, когда математики заметили регулярное удобство их применения в различных доказательствах; поэтому они обычно формулируются в такой активной форме со словами навроде «можно извлечь».

Определение. Множество, для которого эти условия выполнены, называют **компактным**, а также кратко **компактом**. Fréchet 1906

Доказательство. (1 ⇒ 2) Поскольку X ограничено, всякая последовательность его точек имеет сходящуюся подпоследовательность. Её предел является точкой прикосновения для замкнутого множества X, а значит, в нём лежит.

(1 ⇒ 3) В одномерной ситуации мы применили метод деления отрезка пополам как для отыскания точки сгущения, так и для установ-

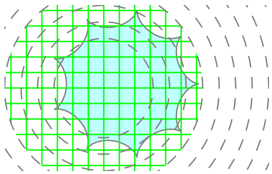
ления возможности извлечения конечного подпокрытия. В двумерной ситуации мы применили в предыдущей теореме аналогичный метод деления квадрата на четыре части; он легко переносится на любую (конечную) размерность. Абсолютно также можно поступить и здесь.

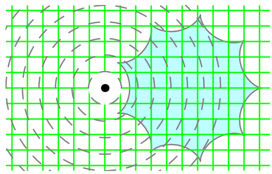
($2 \Rightarrow 1$) Возьмём предельную точку z множества X. Тогда существует последовательность его точек, сходящаяся к z. Иных точек сгущения у сходящейся последовательности не может быть, поэтому $z \in X$. Значит, X замкнуто.

рисунок

Предположим теперь от противного, что X не лежит ни в каком шаре, и выберем $x_n \in X \setminus B_n(0)$. Поскольку $\|x_n\| \to +\infty$, последовательность $\{x_n\}$ не имеет подпоследовательности, сходящейся к точке с конечной нормой. Значит, предположение неверно и X ограничено.

($3 \Rightarrow 1$) Открытые шары $B_r(0)$ для всех $r > 0$ составляют покрытие X, ибо их объединение равно всему пространству. Извлечём конечное подпокрытие. В нём самый большой шар покрывает остальные, а потому и всё X. Значит, X ограничено.





Покажем, что $z \notin X$ не может быть точкой прикосновения для X. Открытые дополнения $U_r = \mathbb{R}^k \setminus \overline{B_r(z)}$ замкнутых шаров для всех $r > 0$ составляют покрытие X, ибо их объединение равно $\mathbb{R}^k \setminus \{z\}$. Извлечём конечное подпокрытие. В нём множество U_r с наименьшим значением $r(z)$ покрывает остальные, а потому и всё X. Следовательно,

$$\rho(z, X) \geqslant \rho(z, U_r) = r > 0.$$

Значит, множество X содержит все свои точки прикосновения, то есть оно замкнуто. $\square$

Упражнение. *Чем можно объяснить схожесть двух рассуждений в доказательстве последней импликации?*

Лемма. *Всегда компактны:*

(1) *замкнутое подмножество компакта;*
(2) *пересечение произвольного семейства компактов;*
(3) *объединение конечного набора компактов;*
(4) *произведение компактов.*

Доказательство. Упражнения. □

Непрерывные функции на компакте. Не только формулировка, но и доказательство теоремы Вейерштрасса об экстремумах ничем не отличаются от случая отрезка.

Теорема. *Всякая непрерывная функция на непустом компакте ограничена и достигает своего минимума и своего максимума.*

Доказательство. Укажем лишь получение максимума, ибо для минимума всё делается аналогично.

Для непрерывной функции f на компакте X положим

$$\beta = \sup\{f(x) \mid x \in X\}.$$

Возьмём такую последовательность $\{x_n\}$ точек X, что $f(x_n) \to \beta$, и извлечём из неё подпоследовательность $\{x'_n\}$, сходящуюся к некоторой точке $c \in X$. По непрерывности $f(x'_n) \to f(c)$. Поэтому $\beta = f(c)$ конечно и является максимальным значением функции f на X. □

Теорема. *Все нормы на $\mathbb{R}^k$ эквивалентны.*

Идеи доказательства. Достаточно понять, что каждая норма эквивалентна эвклидовой. Непрерывность любой нормы как функции координат выводят из неравенства треугольника. Эвклидова сфера единичного радиуса с центром в начале координат компактна. Минимум и максимум значений выбранной нормы на этой сфере и определяют подходящие постоянные для эквивалентности норм. □

Теорема. *Всякий непрерывный образ компакта компактен.*

Доказательство. Возьмём непрерывное отображение $f\colon X \to f(X)$ компакта. Покроем $f(X)$ семейством открытых множеств (годятся шары с центрами в точках образа). Их прообразы в X открыты и покрывают X. Извлечём конечное открытое покрытие X и вернёмся в $f(X)$, где образы дадут конечное подпокрытие исходного покрытия. □ рисунок

Упражнение. *Выведите первую теорему подраздела из третьей.*

Упражнение. *Сформулируйте определение равномерной непрерывности функции на множестве в $\mathbb{R}^k$ и докажите, что всякая непрерывная функция на компакте равномерно непрерывна на нём.*

8.5. Принцип сжимающих отображений

Отображения множества в себя встречаются в самых разнообразных и неожиданных приложениях. При этом мощным средством в доказательствах существования и единственности решения задачи являются теоремы о неподвижной точке, коих в математике нашли много. Одна из них вполне достижима уже доступными нам средствами. лекция 26 03.12.15

Определение. Точку $c \in X$ называют **неподвижной точкой** отображения $f\colon X \to X$, если $f(c) = c$.

Определение. Отображение $f\colon X \to X$ называют **сжимающим**, если существует такое число $0 < \lambda < 1$, что

$$\|f(x_1) - f(x_2)\| \leqslant \lambda \|x_1 - x_2\|$$

для всех точек $x_i \in X$.

гиперссылка

Сжимающее отображение липшицево, а потому всегда непрерывно.

Picard 1880?

Теорема. *Всякое сжимающее отображение замкнутого множества в себя имеет единственную неподвижную точку.*

Доказательство. Условие $\lambda < 1$ исключает наличие двух различных неподвижных точек сжимающего отображения $f\colon X \to X$. Чтобы найти одну, введём вспомогательную непрерывную функцию

$$g\colon X \to \mathbb{R}, \quad x \mapsto \|f(x) - x\|.$$

Определение сжимающего отображения влечёт $g(f(x)) \leqslant \lambda g(x)$ для всех $x \in X$. Если X компактно, то g достигает на X своего минимума в некоторой точке $c \in X$, однако при этом последнее неравенство не нарушается только при $g(c) = 0$ и, следовательно, $f(c) = c$.

Рассуждение более низкого уровня использует только замкнутость множества X, добывая замену недостающей ограниченности из специфики сжимания. Происходящее здесь далее представляет собой форму метода последовательных приближений, важность которого выходит далеко за рамки доказательства этой теоремы. □

Возьмём любую точку $x_0 \in X$ и построим последовательность $\{x_n\}$ по правилу $x_n = f(x_{n-1}) = f^n(x_0)$ для всех целых $n > 0$. Здесь f^n означает (называемую **итерацией**) кратную композицию отображения с собой. Сохраним эти обозначения, как и $g(x)$, до завершения раздела.

Лемма. *При итерациях сжимающего отображения последовательность образов любой точки сходится.*

Доказательство. По неравенству «многоугольника»

$$\begin{aligned}\|x_n - x_0\| &\leqslant \|x_n - x_{n-1}\| + \cdots + \|x_2 - x_1\| + \|x_1 - x_0\| \\ &= g(x_{n-1}) + \cdots + g(x_1) + g(x_0).\end{aligned}$$

Свойство $g(f(x)) \leqslant \lambda g(x)$ даёт $g(x_1) \leqslant \lambda g(x_0)$ и далее $g(x_i) \leqslant \lambda^i g(x_0)$. Следовательно,

$$\|x_n - x_0\| \leqslant (\lambda^{n-1} + \cdots + \lambda + 1)\, g(x_0) < \tfrac{1}{1-\lambda}\, g(x_0).$$

Обозначим число в правой части через r. Это радиус шара $B_r(x_0)$, включающего всю последовательность $\{x_n\}$. Аналогично получаем, рисунок
что её хвост из элементов с $n > N$ лежит в шаре радиуса $\lambda^N r$ с центром в точке x_N. Поскольку $\lambda < 1$, эти радиусы стремятся к нулю и последовательность фундаментальна. □

Следствие. *При итерациях сжимающего отображения замкнутого множества в себя последовательность образов любой точки сходится к неподвижной точке.*

рисунок!

Доказательство. Замкнутое множество X обязательно содержит предел c сходящейся последовательности $\{f^n(x_0)\}$ своих точек. Этим установлена основная теорема раздела и для неограниченных множеств.

Используя непрерывность отображения f, перейдём к пределу в равенстве $x_n = f(x_{n-1})$ и получим $c = f(c)$. Этим установлено уточняющее теорему следствие. □

Глава 9. ГЛАДКИЕ ФУНКЦИИ

В этой главе у нас две основные цели. Во-первых, нужно перенести основные понятия дифференциального исчисления с одномерного случая на многомерный. При этом мы увидим, как сильно изложение загромождается избытком различных букв и индексов. Вместе с геометрической интуицией, это подталкивает ко второй цели. Во-вторых, нужно ввести новые, укрупнённые понятия, ёмкость которых позволит заменить громоздкие формулы на лаконичные, часто внешне очень похожие на формулы одномерного анализа, но обогащённые содержанием. Эти понятия будут основаны на векторах и матрицах; поэтому, вся глава очень существенно опирается на линейную алгебру.

9.1. Дифференциал и градиент

Частные производные. Основные общие понятия и идеи в этой главе адекватно представляются в частном случае функции двух переменных. Такую функцию мы обычно обозначаем через $f(x,y)$. Область определения удобно считать открытой, а ещё удобнее вспоминать о ней пореже, только при крайней необходимости.

Определение. Когда существует конечный предел

$$\lim_{\delta\to 0}\frac{f(x+\delta,y)-f(x,y)}{\delta},$$

его называют **частной производной** функции f по переменной x в точке (x,y) и обозначают через $\frac{\partial f}{\partial x}(x,y)$. Круглая форма буквы ∂ важна!

Таким образом, при вычислении частной производной функции по одной переменной все другие переменные просто считаются постоянными («замораживаются»).

Пример. Найдём частные производные функции $f(t,\varphi)=\cos(\omega t+\varphi)$:

$$\frac{\partial f}{\partial t}=-\omega\sin(\omega t+\varphi),\qquad \frac{\partial f}{\partial\varphi}=-\sin(\omega t+\varphi).$$

Частные производные функции в точке улавливают её поведение далеко не полностью.

Пример. Рассмотрим функцию

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

Заклеим неопределённость в начале координат, полагая $f(0,0)=0$. Тогда обе частные производные $\frac{\partial f}{\partial x}(0,0)$ и $\frac{\partial f}{\partial y}(0,0)$ существуют и равны нулю, хотя сама функция в этой точке разрывна. Например, $f(x,x) = 1$ и $f(x,-x) = -1$ при всех $x \neq 0$. Приближаясь к началу координат вдоль иных прямых, получим иные предельные значения.

В то же время, поворот системы координат на 45 градусов преобразует эту функцию в функцию

$$g(x,y) = \frac{x^2 - y^2}{x^2 + y^2},$$

частные производные которой в начале координат разрывны.

Производные по вектору. Графиком функции $z = f(x,y)$ является поверхность в $\mathbb{R}^3$. При геометрическом взгляде сразу видна стеснённость координатного подхода, ведущего к понятию частной производной: изменяя одну переменную, мы движемся по поверхности в одном направлении. Движение в произвольном направлении приводит к понятию производной по вектору, которое предпочтительно ввести сразу для функции нескольких переменных.

Определение. Когда существует конечный предел

$$\lim_{\delta\to 0} \frac{f(x+\delta v) - f(x)}{\delta},$$

его называют **производной** функции f **по вектору** v в точке x и обозначают через $\frac{\partial f}{\partial v}(x)$. Также встречаются иные обозначения: $D_v f(x)$, $\partial_v f(x)$, $f'_v(x)$ и даже совсем краткое $f_v(x)$.

Уже на уровне определений производная по стандартному базисному вектору $\mathbf{e}_\alpha$ равна частной производной:

$$\frac{\partial f}{\partial \mathbf{e}_\alpha}(x) = \frac{\partial f}{\partial x_\alpha}(x).$$

Арифметические правила вычисления производных по вектору и, в том числе, частных производных идентичны соответствующим правилам для обычной производной функции одной переменной.

Иногда вместо производной по вектору говорят о производной **по направлению**. Отличие обусловлено тем, что направлением называют единичный вектор, а тогда направление вектора $v \neq 0$ есть единичный вектор $v/\|v\|$. Деление на $\|v\|$ перекочёвывает и в формулу производной по направлению (которую я считаю вредной).

Дифференциал и дифференцируемость. Наша задача теперь — перенести понятие дифференциала на случай функции нескольких переменных. Удобнее всего это делать для формулы

$$\Delta f = \mathcal{A}(\Delta x) + o(|\Delta x|),$$

выделяющей из приращения функции его главную линейную часть.

Для функции нескольких переменных условие дифференцируемости формально остаётся таким же, но Δx теперь является вектором из точки p в (близкую) точку x, так что вместо модуля нужно писать норму. Линейная функция $\mathcal{A}$ есть дифференциал df по определению, а функция f дифференцируема в точке, если вблизи неё имеется такое представление; равносильно, но без символа Δ, выразим это как

$$f(x) = f(p) + \mathcal{A}(x - p) + o(\|x - p\|).$$

Общий вид линейной функции векторного аргумента известен из линейной алгебры. Каждой такой функции $\mathcal{A}$ взаимно однозначно соответствует такой вектор A, что

$$\mathcal{A}(v) = A \cdot v$$

для всех векторов v. В координатной записи $A = (A_1, \ldots, A_n)$ и тогда

$$\mathcal{A}(v) = A_1 v_1 + \ldots + A_n v_n = \sum_{1 \leqslant \alpha \leqslant n} A_\alpha v_\alpha.$$

Для дальнейшего развития предмета полезно считать вектор v столбцом, а вектор A строкой; тогда их скалярное произведение становится матричным произведением Av. Ниже мы назовём вектор A по имени.

Пример. Линейную функцию, которая выделяет из точки (вектора) в $\mathbb{R}^n$ одну определённую координату, называют **координатной функцией** и обозначают через x_α. Значит, для $1 \leqslant \alpha \leqslant n$ имеем

$$x_\alpha(v_1, \ldots, v_n) = v_\alpha, \qquad x_\alpha(\Delta x_1, \ldots, \Delta x_n) = \Delta x_\alpha.$$

Она совпадает со своим дифференциалом dx_α, поскольку

$$dx_\alpha(\Delta x_1, \ldots, \Delta x_n) = \Delta x_\alpha$$

обобщает одномерную формулу $dx(\Delta x) = \Delta x$.

Упражнение. *Всякая линейная функция на $\mathbb{R}^n$ дифференцируема и совпадает со своим дифференциалом.*

Лемма. *Функция, дифференцируемая в точке, непрерывна в ней.*

Доказательство. При $x \to p$ имеем $\mathcal{A}(x - p) \to 0$ и $o(\|x - p\|) \to 0$. Поэтому $f(x) \to f(p)$, что и является непрерывностью.

Результат лёгок, поскольку мы лишь отмечаем, что более сильное свойство влечёт более слабое. □

При необходимости дифференциал функции помечают точкой приложения в нижнем индексе: $df_p(v)$. Следует отличать это обозначение от прочих индексов. Оно подчёркивает, что точка p зафиксирована, а дифференциал является функцией приложенного вектора v. Бывает и наоборот, когда производную по зафиксированному свободному вектору v изучают как функцию точки приложения p. В таком случае предпочитают обозначения типа $\partial_v f(p)$.

Теорема. *Если функция f дифференцируема в точке $x = p$, то:*

(1) *все её частные производные существуют и равны соответствующим коэффициентам её дифференциала:*

$$df_p = \sum_{1 \leqslant \alpha \leqslant n} \frac{\partial f}{\partial x_\alpha}(p)\, dx_\alpha;$$

(2) *её производная по любому вектору существует и равна значению дифференциала на нём:*

$$\partial_v f(p) = df_p(v).$$

Эта теорема даёт координатную формулу

$$\partial_v f(p) = df_p(v) = \sum_{1 \leqslant \alpha \leqslant n} \frac{\partial f}{\partial x_\alpha}(p)\, v_\alpha$$

для вычисления производной по любому вектору.

Пример. Выпишем дифференциал функции $f(x, y, z) = xy^2z^3$. Для этого сперва найдём частные производные:

$$\tfrac{\partial f}{\partial x} = y^2z^3, \quad \tfrac{\partial f}{\partial y} = 2xyz^3, \quad \tfrac{\partial f}{\partial z} = 3xy^2z^2.$$

Значит,

$$df = y^2z^3\, dx + 2xyz^3\, dy + 3xy^2z^2\, dz.$$

В точке $(x, y, z) = (1, 1, 1)$ получим $df = dx + 2dy + 3dz$.

Доказательство. (1) Запишем общий вид дифференциала

$$df_p(\Delta x) = \sum A_\alpha \Delta x_\alpha$$

и отыщем его коэффициенты A_α. Выберем для этого приращение Δx, у которого лишь компонента Δx_α отлична от нуля. Тогда

$$\Delta f(\Delta x) = A_\alpha \Delta x_\alpha + o(\Delta x_\alpha).$$

Поделим на Δx_α и в пределе при $\Delta x_\alpha \to 0$ получим $A_\alpha = \frac{\partial f}{\partial x_\alpha}(p)$.

(2) По определению $\partial_v f(p)$ равно пределу при $\delta \to 0$ выражения

$$\frac{1}{\delta}\big(f(p+\delta v) - f(p)\big) = \frac{1}{\delta}\big(df_p(\delta v) + o(\delta)\big) = df_p(v) + o(1).$$

Избавление от δ в главном слагаемом происходит благодаря линейности дифференциала: $df_p(\delta v) = \delta\, df_p(v)$. □

Упражнение. *Выразите $d(f \pm g)$, $d(f \cdot g)$ и $d(f/g)$ через df, dg и сами функции.*

Свобода выбора вектора приводит к новому свойству производных.

Следствие. *Производные дифференцируемой функции по векторам ведут себя линейно по отношению к ним:*

$$\partial_{\alpha u + \beta v} f(x) = \alpha \cdot \partial_u f(x) + \beta \cdot \partial_v f(x)$$

для любых числовых постоянных α и β.

Градиент. Выше упомянуто взаимно однозначное соответствие между линейными функциями на $\mathbb{R}^n$ и векторами в $\mathbb{R}^n$. Дифференциал df_p это линейная функция; соответствующий вектор называют **градиентом** дифференцируемой функции f в точке p и обозначают через $\nabla f(p)$. Сам треугольный значок называют **набла**. В координатной записи

$$\nabla f = \big(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\big).$$

Производную по произвольному вектору теперь можно представить как скалярное произведение: $\partial_v f(p) = \nabla f(p) \cdot v$. Значит, $\nabla f(p) = 0$ влечёт $\partial_v f(p) = 0$ для всех векторов v.

Градиент функции имеет важный геометрический смысл.

Теорема. *Если $\nabla f(p) \neq 0$, то его направление совпадает с направлением скорейшего роста функции f вблизи точки p, а его (эвклидова) норма есть скорость этого роста.*

Доказательство. Направлением иногда называют единичный вектор (его эвклидова норма равна 1), поэтому направлением скорейшего роста является единичный вектор, производная по которому в точке p максимальна. Поскольку

$$\partial_v f(p) = \nabla f(p) \cdot v = \|\nabla f(p)\| \cdot \|v\| \cdot \cos\angle(\nabla f(p), v),$$

максимум достигается при нулевом угле, то есть совпадении направлений, и равен норме градиента. □

Гладкие функции. В то время как для функций одной переменной дифференцируемость в точке и существование в ней производной равносильны практически по определению, для функций нескольких переменных это разные понятия. Дифференцируемость функции в точке влечёт существование частных производных, но обратное неверно. Мы уже видели пример функции, имеющей нулевые частные производные в точке, но при этом разрывной в ней. Изучим теперь важнейшее условие на частные производные, гарантирующее дифференцируемость.

Определение. Функцию на открытом множестве называют **гладкой**, если все её частные производные существуют и непрерывны на нём.

Область определения предполагают открытой во избежание глупых ситуаций и частых длинных оговорок, когда, например, при переходе из одной точки множества в близкую ей теряются хорошие свойства функции, кои желательно сохранять в некоторой окрестности, что и обеспечивает открытость. Предположение естественно ещё и потому, что качественные потери часто происходят в отдельных точках или на иных замкнутых подмножествах. Исключая их, получают открытые.

Теорема. *Всякая гладкая функция всюду дифференцируема.*

Доказательство. Разберём случай функции двух переменных. В открытой области определения возьмём точку (x, y) и близкую точку $(x + \Delta x, y + \Delta y)$. Приращение функции разобьём на сегменты вдоль осей координат:

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

Преобразуем посегментно, пользуясь формулой конечных приращений для функций одной переменной, пока другая сохраняет значение: рисунок

$$\Delta f = \frac{\partial f}{\partial x}(\xi, y + \Delta y)\Delta x + \frac{\partial f}{\partial y}(x, \eta)\Delta y.$$

Число ξ лежит между x и Δx, а число η — между y и $y + \Delta y$. Ввиду непрерывности частных производных $\frac{\partial f}{\partial x}$ и $\frac{\partial f}{\partial y}$ получаем

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

Итак, приращение f вблизи точки (x, y) представлено как сумма линейной и бесконечно малой функций приращения аргумента, что и означает дифференцируемость f в этой точке.

Рассуждение для произвольной конечной размерности отличается лишь количеством сегментов: по одному на каждую координату. □

средние?

9.2. Дифференцирование композиции

Правило цепочки для функций. Возьмём функцию $y(x_1, \ldots, x_n)$, аргументы которой зависят от других переменных: $x_\alpha = x_\alpha(t_1, \ldots, t_m)$. Указание областей определения лежит в стороне от существа дела. Можно было бы начать со случая двух переменных, $n = 2$ и $m = 2$, но он практически ничем не отличается от общего.

Теорема. *Композиция функций сохраняет дифференцируемость.*

Следствие. *Производные композиции вычисляются по формуле*

$$\frac{\partial y}{\partial t_\beta} = \sum_\alpha \frac{\partial y}{\partial x_\alpha} \cdot \frac{\partial x_\alpha}{\partial t_\beta}.$$

Следствие. *Композиция функций сохраняет гладкость.*

Упражнение. *Перепишите формулу для частных производных композиции в матричном виде.*

Для разминки перед доказательством общего правила рассмотрим несколько его типичных применений.

Полная производная. Функция $f(x(t), y(t), z(t))$ в конечном итоге зависит только от t. По общему правилу цепочки получаем

$$\frac{df}{dt} = \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial f}{\partial z} \cdot \frac{dz}{dt}.$$

Производные по времени полные. Запись можно сократить, вводя для $\frac{\partial f}{\partial x}$ обозначения f_x или $\partial_x f$, далее свободно — причём без штрихов! — используемые, когда сами переменные без индексов. Тогда

$$f' = f_x \cdot x' + f_y \cdot y' + f_z \cdot z'.$$

Функция $f(t, y(t), z(t))$ явно зависит от t. Её полная производная по t находится подстановкой $x = t$ в предыдущую формулу:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial f}{\partial z} \cdot \frac{dz}{dt}.$$

Здесь особенно хорошо видно, что нужно различать полную и частную производные $\frac{df}{dt}$ и $\frac{\partial f}{\partial t}$.

Пример. Возьмём $f(t,x) = t + x^2$, где $x = \sin t$. Тогда $f_t = 1$, $f_x = 2x$ и $x' = \cos t$, поэтому

$$f' = 1 + 2x\cos t = 1 + 2\sin t \cos t.$$

Замена независимых переменных. При повороте декартовой системы координат на плоскости на угол θ вокруг начала старые координаты (x,y) и новые координаты (u,v) связаны формулами подраздел 17.02.17

$$x = u\cos\theta - v\sin\theta, \quad y = u\sin\theta + v\cos\theta,$$

или, в обратную сторону,

$$u = x\cos\theta + y\sin\theta, \quad v = -x\sin\theta + y\cos\theta.$$

Тогда любую функцию можно выразить через новые координаты, $w = f(x,y) = g(u,v)$. Правило цепочки сразу даёт

$$w_u = w_x\cos\theta + w_y\sin\theta, \quad w_v = -w_x\sin\theta + w_y\cos\theta.$$

Видим, что частные производные преобразуются по тому же закону, что и независимые переменные. Это верно и для поворота в пространстве, хотя сами формулы с углами будут гораздо сложнее.

Другая важная замена осуществляет переход от декартовых координат к полярным:

$$x = \rho\cos\varphi, \quad y = \rho\sin\varphi.$$

При обратном переходе традиционно пишут

$$\rho = \sqrt{x^2 + y^2}, \quad \varphi = \operatorname{arctg}\frac{y}{x},$$

хотя у арктангенса есть подвох со сдвигом на половину оборота. Пожалуй, оптимальным решением тут является функция `atan2`, завоевавшая популярность в эру компьютеров.

При таком переходе часто требуются значения производных правка 03.01.18

$$\rho_x = \frac{x}{\sqrt{x^2+y^2}} = \frac{x}{\rho} = \cos\varphi, \quad \varphi_x = -\frac{y}{x^2+y^2} = -\frac{y}{\rho^2} = -\frac{\sin\varphi}{\rho},$$
$$\rho_y = \frac{y}{\sqrt{x^2+y^2}} = \frac{y}{\rho} = \sin\varphi, \quad \varphi_y = \frac{x}{x^2+y^2} = \frac{x}{\rho^2} = \frac{\cos\varphi}{\rho}.$$

С их помощью, для $w = f(x,y) = g(\rho,\varphi)$ по правилу цепочки находят

$$w_x = w_\rho\cdot\rho_x + w_\varphi\cdot\varphi_x, \quad w_y = w_\rho\cdot\rho_y + w_\varphi\cdot\varphi_y.$$

Неявные функции. Говорят, что функция $y(x)$ задана **неявно** уравнением $F(x,y)=0$, если $F(x,y(x))=0$ для всех значений x, обычно вблизи некоторой точки. Это определение мы уточним в дальнейшем.

Примеры. Уравнение $x^2+y^2-1=0$ задаёт две функции; их можно явно выразить как $y=\pm\sqrt{1-x^2}$. Однако даже небольшое усложнение уравнения даёт функцию, явное выражение которой затруднительно: $x^3+y^3-3xy=0$. Явное выражение обычно неэлементарно при более высокой степени полинома и, тем более, когда уравнение включает трансцендентные функции.

Отсутствие явного выражения не мешает дифференцировать неявно. Как и выше, выпишем полную производную по x:

$$F(x,y(x))'=F_x+F_y\cdot y'.$$

Условие $F(x,y(x))=0$ влечёт $F(x,y(x))'=0$, поэтому $y'(x)=-F_x/F_y$.

Пример. Для $F(x,y)=x^2+y^2-1$ легко найдём $y'(x)=-x/y$.

Упражнение. *Сравните с вычислением в явном и параметрическом видах.*

Пример. Для $F(x,y)=x^3+y^3-3xy$ опять же легко найдём

$$y'(x)=\frac{x^2-y}{x-y^2}.$$

Неявно заданы бывают и функции нескольких переменных. Простейший такой случай — когда уравнение $F(x,y,z)=0$ определяет z как функцию $z(x,y)$. Геометрически он даёт поверхность в трёхмерном пространстве. Тут по правилу цепочки получаем

$$0=\partial_x F(x,y,z(x,y))=F_x+F_y\cdot y_x+F_z\cdot z_x \quad\Rightarrow\quad z_x=-F_x/F_z;$$
$$0=\partial_y F(x,y,z(x,y))=F_x\cdot x_y+F_y+F_z\cdot z_y \quad\Rightarrow\quad z_y=-F_y/F_z.$$

Частные производные x_y и y_x равны нулю в предположении независимости переменных x и y друг от друга.

Пример. Уравнение $x^2-y^2+z^2-1=0$ задаёт однополостный гиперболоид. Дифференцируя неявно, получим $z_x=-x/z$ и $z_y=y/z$.

Основы работы с неявными функциями в более сложных ситуациях мы изучим подробнее в следующей главе, а пока для повышения аппетита полезно сделать следующее упражнение.

рисунок

Упражнение. *Линия в пространстве бывает задана как пересечение двух поверхностей: $F(x,y,z)=0$ и $G(x,y,z)=0$. Найдите формулы для производных соответствующих неявных функций $y(x)$ и $z(x)$.*

Доказательство правила цепочки. Представим приращения функций суммами дифференциалов и бесконечно малых:

$$\Delta y = dy(\Delta x) + o(\|\Delta x\|), \quad \Delta x_\alpha = dx_\alpha(\Delta t) + o(\|\Delta t\|).$$

Для частных производных введём сокращения

$$A_\alpha = \frac{\partial y}{\partial x_\alpha}, \quad B_{\alpha\beta} = \frac{\partial x_\alpha}{\partial t_\beta}$$

и через них распишем дифференциалы подробнее:

$$\Delta y = \sum_\alpha A_\alpha \Delta x_\alpha + o(\|\Delta x\|), \quad \Delta x_\alpha = \sum_\beta B_{\alpha\beta}\Delta t_\beta + o(\|\Delta t\|),$$

Подставим в первое выражение остальные и поменяем порядок суммирования:

$$\begin{aligned}\Delta y &= \sum_\alpha A_\alpha \Big(\sum_\beta B_{\alpha\beta}\Delta t_\beta + o(\|\Delta t\|)\Big) + o(\|\Delta x\|)\\ &= \sum_\beta \Big(\sum_\alpha A_\alpha B_{\alpha\beta}\Big)\Delta t_\beta + \sum_\alpha A_\alpha o(\|\Delta t\|) + o(\|\Delta x\|).\end{aligned}$$

Ввиду линейной связи Δx_α и Δt, можно заменить $o(\|\Delta x\|)$ на $o(\|\Delta t\|)$, и тогда все наличные бесконечно малые собираются в одну $o(\|\Delta t\|)$. Вот мы и представили приращение композиции $y(x(t))$ как сумму линейной функции от Δt и бесконечно малой, доказав теорему.

Искомая формула цепочки содержится в последней большой скобке. Нужно лишь вспомнить, что под A_α и $B_{\alpha\beta}$ скрываются соответствующие частные производные. Далее, когда все они непрерывны, левая часть формулы тоже непрерывна, и тем самым установлено следствие про гладкость.

Однородные функции. Продолжим теперь рассмотрение применений правила цепочки более специфическими вещами.

Функцию $f(x_1, \dots, x_n)$ называют **однородной** степени λ, если для всех $t > 0$ выполнено

$$f(tx_1, \dots, tx_n) = t^\lambda f(x_1, \dots, x_n).$$

Допускаются любые действительные показатели λ.

Пример. Функция $f(x, y) = \sqrt{x^2 + y^2}$ однородна степени 1. Функция $g(x, y) = \frac{2xy}{x^2+y^2}$ однородна степени 0. Вычислим для интереса

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = f, \qquad x\frac{\partial g}{\partial x} + y\frac{\partial g}{\partial y} = 0.$$

Конструкции наподобие $x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$ называют **дифференциальными операторами**. Такой оператор представляет собой правило, содержащее дифференцирования наряду с арифметическими операциями, по которому он преобразует данную ему функцию в другую функцию.

Euler 1736

Теорема. *Всякая гладкая однородная степени λ функция f удовлетворяет тождеству*

$$\sum_\alpha x_\alpha \frac{\partial f}{\partial x_\alpha} = \lambda f.$$

Это тождество имеет приложения в физике (теорема о вириале).

Доказательство. Найдём полную производную по t от $f(tx_1, tx_2) = t^\lambda f(x_1, x_2)$, считая x_α не зависящими от t. С одной стороны,

$$\frac{d}{dt} f(tx_1, tx_2) = \sum_\alpha \frac{\partial f}{\partial x_\alpha}(tx_1, tx_2) \cdot \frac{d}{dt}(tx_\alpha) = \sum_\alpha \frac{\partial f}{\partial x_\alpha}(tx_1, tx_2) \cdot x_\alpha.$$

С другой стороны,

$$\frac{d}{dt}\big(t^\lambda f(x_1, x_2)\big) = \lambda t^{\lambda-1} f(x_1, x_2).$$

Полагая теперь $t = 1$, получим искомое тождество. □

Необозначенные аргументы. Иногда необходимо выразить производную композиции, обладая неполной информацией о функции. Специфическая трудность возникает, когда аргументам не присвоены никакие обозначения. Один способ решения — ввести эти обозначения, однако тем самым загромождается изложение. Укажем другой встречающийся в задачниках способ, легко могущий поставить читателя в тупик: на аргументы функции ссылаются по их номерам. Именно, f_1 означает частную производную по первому аргументу, и так далее.

Пример. Выразим частные производные функции $f(x+y, x-y)$ через частные производные самой функции f. По общему правилу

$$\begin{aligned}
\tfrac{\partial}{\partial x} f(x+y, x-y) &= f_1 \cdot \tfrac{\partial}{\partial x}(x+y) + f_2 \cdot \tfrac{\partial}{\partial x}(x-y) = f_1 + f_2;\\
\tfrac{\partial}{\partial y} f(x+y, x-y) &= f_1 \cdot \tfrac{\partial}{\partial y}(x+y) + f_2 \cdot \tfrac{\partial}{\partial y}(x-y) = f_1 - f_2.
\end{aligned}$$

При этом в аргументы f_i всюду нужно подставить $x + y$ и $x - y$.

9.3. Высшие частные производные

Здесь мы рассматриваем функции на открытом множестве $U \subseteq \mathbb{R}^n$. Как уже говорилось, открытость удобна тем, что любая точка U имеет окрестность, на которой функция определена; тем самым пропадает необходимость в некоторых оговорках.

Перестановочность частных производных. Возьмём функцию f на U. Частная производная $\frac{\partial f}{\partial x_\alpha}$ — тоже функция на U, каждая частная производная которой, если существует, является второй частной производной функции f. Тогда пишут

$$\frac{\partial^2 f}{\partial x_\beta \partial x_\alpha} = \frac{\partial}{\partial x_\beta}\left(\frac{\partial f}{\partial x_\alpha}\right), \qquad \frac{\partial^2 f}{\partial x_\alpha^2} = \frac{\partial}{\partial x_\alpha}\left(\frac{\partial f}{\partial x_\alpha}\right).$$

Аналогично определяют частные производные более высоких порядков. Всё заметнее выгода кратких обозначений типа $f_{xy} = (f_x)_y$.

Высшие, особенно вторые, частные производные, включающие дифференцирование по разным переменным, называют **смешанными**.

Пример. Найдём смешанные вторые частные производные функции $f(x,y) = x^y$. Первые производные равны $f_x = yx^{y-1}$ и $f_y = x^y \ln x$. Искомые смешанные производные совпадают:

$$(f_x)_y = (f_y)_x = x^{y-1}(y \ln x + 1).$$

Теорема. *Равенство*

$$\frac{\partial}{\partial x_\alpha}\left(\frac{\partial f}{\partial x_\beta}\right) = \frac{\partial}{\partial x_\beta}\left(\frac{\partial f}{\partial x_\alpha}\right)$$

верно, как только обе его части непрерывны.

Доказательство. Поскольку утверждение касается пары переменных, наличие у функции дополнительных переменных не существенно и только загромождает рассуждение. Итак, будем работать с функцией $f(x,y)$, имеющей непрерывные смешанные производные $(f_x)_y$ и $(f_y)_x$.

Зафиксируем точку (x,y) и выберем малое приращение $(\Delta x, \Delta y)$. Начнём с выражения

рисунок

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

симметричного относительно букв x и y. Обозначим его загадочно через $\Delta^2 f$ и перепишем через вспомогательную функцию

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

одной переменной. Применим формулу конечных приращений:

$$\Delta^2 f = g(x+\Delta x) - g(x) = g'(\xi)\,\Delta x$$

для некоторого ξ между x и $x+\Delta x$. Подставим определение $g(x)$ и ещё раз применим ту же формулу, но теперь по другой переменной:

$$\Delta^2 f = \big(f_x(\xi, y+\Delta y) - f_x(\xi, y)\big)\Delta x = (f_x)_y(\xi,\eta)\,\Delta y\,\Delta x$$

для некоторого η между y и $y+\Delta y$. Значит,

$$\frac{\Delta^2 f}{\Delta x\,\Delta y} = (f_x)_y(\xi,\eta) \to (f_x)_y(x,y)$$

в силу предполагаемой непрерывности вторых частных производных. Вследствие равноправия букв, симметричным рассуждением получим

$$\frac{\Delta^2 f}{\Delta x\,\Delta y} \to (f_y)_x(x,y),$$

что и даёт требуемое равенство $(f_x)_y = (f_y)_x$. □

переставил 18.02.17

Второй дифференциал и гессиан. Если все вторые частные производные функции непрерывны на некотором множестве U, то её называют **дважды гладкой** на U. Для дважды гладкой функции $f(x,y)$ запишем первый дифференциал

$$df = f_x\,dx + f_y\,dy.$$

Дифференциал этого выражения есть искомый второй дифференциал $d^2 f$. Упростим его вычисление, считая переменные независимыми, ибо в таком случае $d(dx)=0$ и $d(dy)=0$. Получаем

$$\begin{aligned} d^2 f = d(df) &= d(f_x)\,dx + f_x\,d(dx) + d(f_y)\,dy + f_y\,d(dy) \\ &= (f_{xx}\,dx + f_{xy}\,dy)\,dx + (f_{yx}\,dx + f_{yy}\,dy)\,dy \\ &= f_{xx}\,dx^2 + 2f_{xy}\,dx\,dy + f_{yy}\,dy^2. \end{aligned}$$

Пример. Для функции $f(x,y) = x^2 - y^2$ найдём

$$\begin{aligned} df &= 2x\,dx - 2y\,dy, \\ d^2 f &= 2dx^2 - 2dy^2. \end{aligned}$$

Пример. Для функции $f(x,y) = e^{-x^2-y^2}$ найдём

$$\begin{aligned} df &= -2f\cdot(x\,dx + y\,dy), \\ d^2 f &= \dots = -2f\cdot\big((1-2x^2)\,dx^2 - 4xy\,dx\,dy + (1-2y^2)\,dy^2\big). \end{aligned}$$

В начале координат получаем $d^2 f_{(0,0)} = -2(dx^2+dy^2)$.

В выбранной точке каждый дифференциал есть функция приращения $(\Delta x, \Delta y)$. При этом первый дифференциал — линейная функция, значит, однородная степени 1; второй дифференциал — квадратичная функция, или же однородная степени 2. И так далее.

Однородные функции на $\mathbb{R}^n$ в некоторых контекстах издавна называют **формами**. Соответственно степени, формы бывают линейные, квадратичные, кубические... Квадратичные формы имеют много приложений и довольно подробно изучаются в курсе линейной алгебры. Кубические и дальнейшие сложны и встречаются редко.

Коэффициенты квадратичной формы удобно записать в матрице. Для формы второго дифференциала эту матрицу называют **гессианом**. В случае функции $f(x, y)$ гессиан имеет вид

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix},$$

а в общем случае функции $f(x_1, \ldots, x_n)$ это квадратная $n \times n$ матрица H с элементами $h_{\alpha\beta} = \frac{\partial}{\partial x_\alpha}\left(\frac{\partial f}{\partial x_\beta}\right)$. Для дважды гладких функций эта матрица симметрична: $h_{\beta\alpha} = h_{\alpha\beta}$, или $H^\top = H$.

Классы гладкости. Функцию называют m **раз гладкой** на U, или m-**кратно непрерывно-дифференцируемой**, или функцией **класса** $C^m(U)$, если существуют и непрерывны на U все её частные производные порядка m. Производная порядка 0 есть сама функция.

Упражнение. *Встречаются формулировки со словами «вплоть до порядка m». Почему здесь это «вплоть до» излишне?*

Этому определению можно придать интересный индуктивный вид.

Определение. Функцию называют m **раз гладкой** на U, или ...,

- при $m = 0$, если она непрерывная;
- при $m = 1$, если она гладкая;
- при $m > 1$, если все её (первые) частные производные $m - 1$ раз гладкие.

Все условия здесь — на U.

Классы гладкости вложены друг в друга:

$$C^0(U) \supset C^1(U) \supset C^2(U) \supset \ldots$$

Выделяют также класс $C^\infty(U)$ **бесконечно гладких функций**, у которых все производные любого порядка непрерывны на U.

Ранее в курсе говорилось, что все элементарные функции непрерывны, затем — что они непрерывно дифференцируемы (с оговорками о концевых точках). Поскольку производная элементарной функции элементарна, все элементарные функции на самом деле бесконечно гладкие вне отдельных точек.

Следствие. *Смешанные производные порядка m достаточно гладкой функции не зависят от порядка выполнения дифференцирований.*

Доказательство. Достаточная гладкость здесь означает класс C^m.

По теореме о смешанных вторых частных производных, можно обменивать соседние дифференцирования. Остаётся сослаться на чисто комбинаторное утверждение: каждую перестановку упорядоченного конечного списка можно получить некоторой цепочкой преобразований, обменивающих соседние элементы. □

Например, для третьего порядка имеется $3! = 6$ различных перестановок. Расположим их в цепочку с обменами соседних:

6-угольник!

$$f_{xyz} = f_{xzy} = f_{zxy} = f_{zyx} = f_{yzx} = f_{yxz}.$$

Выпишем теперь все частные производные второго порядка для дважды гладкой функции $f(x, y, z)$. Их получается шесть:

$$f_{xx}, f_{xy}, f_{xz}, f_{yy}, f_{yz}, f_{zz}.$$

А теперь — все частные производные третьего порядка для трижды гладкой функции $f(x, y)$. Их получается четыре:

$$f_{xxx}, f_{xxy}, f_{xyy}, f_{yyy}.$$

Упражнение. *Сколько различных частных производных порядка m может иметь бесконечно гладкая функция n переменных?*

Многомерная формула Тейлора. Формула Тейлора для достаточно гладких функций одной переменной записывается с любым количеством высших производных. При нескольких переменных следует ожидать огромной формулы со всеми частными производными вплоть до нужного порядка, свести которую к обозримому виду можно, только применяя составные понятия.

Чтобы задача не выглядела столь абстрактно, для начала выпишем явно формулу Тейлора второго порядка для функции $f(x, y)$ вблизи начала координат. Её основная составляющая — полином Тейлора второй степени от двух букв x и y. Коэффициент при каждом мономе состоит из частной производной, причём довольно быстро становится ясно, какой именно, и числового множителя комбинаторной природы.

Простейший способ получить эту формулу состоит в двухшаговом разложении: сперва по одной переменной, затем по другой. На каждом шаге применяем одномерную формулу Тейлора. Первый шаг — разложим по x вблизи $x = 0$, совсем не отслеживая o-малые:

$$f(x,y) = f(0,y) + f_x(0,y)\,x + \tfrac{1}{2!} f_{xx}(0,y)\,x^2 + \dots.$$

На втором шаге разложим функции $f(0,y)$, $f_x(0,y)$ и $f_{xx}(0,y)$ по y вблизи $y = 0$:

$$\begin{aligned}
f(0,y) &= f(0,0) + f_y(0,0)\,y + \tfrac{1}{2!} f_{yy}(0,0)\,y^2 + \dots,\\
f_x(0,y) &= f_x(0,0) + f_{xy}(0,0)\,y + \tfrac{1}{2!} f_{xyy}(0,0)\,y^2 + \dots,\\
f_{xx}(0,y) &= f_{xx}(0,0) + f_{xxy}(0,0)\,y + \tfrac{1}{2!} f_{xxyy}(0,0)\,y^2 + \dots.
\end{aligned}$$

Эти разложения подставим в записанное на первом шаге, отбрасывая выделенные слагаемые, поскольку после умножения на нужную степень x общая степень каждого из них выше второй. Получим

$$\begin{aligned}
f(x,y) = f(0,0)&\\
&+ f_x(0,0)\,x + f_y(0,0)\,y\\
&+ \tfrac{1}{2} f_{xx}(0,0)\,x^2 + f_{xy}(0,0)\,xy + \tfrac{1}{2} f_{yy}(0,0)\,y^2 + \dots.
\end{aligned}$$

Более аккуратно, вместо многоточия здесь следует написать $o(x^2+y^2)$, а при разложении порядка k даже $o\big(\|(x,y)\|^k\big)$.

Разобравшись с высшими дифференциалами, мы вернёмся к этому простому подходу в следующем разделе, чтобы связать его с более формальным. Выглядеть это докапывание до причин будет несравнимо сложнее, а потому рекомендуется пропускать его при первом, втором, а то и третьем чтении.

Упражнение. *Выпишите подробно следующие по сложности формулы Тейлора: второго порядка для функции $f(x,y,z)$ и третьего порядка для функций $f(x,y)$ и $f(x,y,z)$.*

Упражнение. *Найдите общую комбинаторную формулу для появляющихся числовых множителей.*

9.4. Высшие дифференциалы и формула Тейлора

Индексные обозначения. Теперь напишем высшие дифференциалы достаточно гладкой функции $f(x_1,\dots,x_n)$.

Первый дифференциал функции в точке — это пример линейной формы. Её коэффициенты есть первые частные производные:

$$df_p(\Delta x) = \sum_\alpha \frac{\partial f}{\partial x_\alpha}(p)\,\Delta x_\alpha.$$

Координатные формы dx_α действуют по правилу $dx_\alpha(\Delta x) = \Delta x_\alpha$, а первый дифференциал функции равен их линейной комбинации

$$df_p = \sum_\alpha \frac{\partial f}{\partial x_\alpha}(p)\,dx_\alpha.$$

Коэффициенты, а потому и вся линейная форма, меняются от точки к точке. Как и в случае функций одной или двух переменных, возникает второй дифференциал $d^2 f = d(df)$. Это уже квадратичная форма:

$$d^2 f = \sum_\alpha d\Big(\frac{\partial f}{\partial x_\alpha}(x)\Big)\,dx_\alpha = \sum_{\alpha,\beta} \frac{\partial^2 f}{\partial x_\alpha \partial x_\beta}(x)\,dx_\alpha\,dx_\beta.$$

Последняя сумма двойная; индексы α и β независимо пробегают значения от 1 до n. Каждое слагаемое есть произведение коэффициента, зависящего от точки приложения, и двух координатных дифференциалов. Значит, в каждой точке $x = p$ получается квадратичная функция приращения Δx.

Построив второй дифференциал из первого, продолжим этот процесс. Получим третий дифференциал

$$d^3 f = d(d^2 f) = \sum_{\alpha,\beta,\gamma} \frac{\partial^3 f}{\partial x_\alpha \partial x_\beta \partial x_\gamma}(x)\,dx_\alpha\,dx_\beta\,dx_\gamma.$$

Это кубическая форма. Коэффициенты её образуют трёхмерный массив, работать с которым вручную весьма неудобно. К счастью, это требуется значительно реже, чем для второго дифференциала.

При необходимости записать развёрнутое выражение для дифференциалов более высокого порядка m мы сталкиваемся с новым уровнем сложности в обозначениях. Простые индексы α, β, γ приходится заменять на $\alpha_1, \ldots, \alpha_m$. Заметной лаконичности достигают, применяя **мульти-индексы** (наборы натуральных чисел), но нам нет нужды вдаваться в детали этого формализма.

Лемма об одномерном сечении. Исследуя гладкую функцию f нескольких переменных вблизи точки $x = p$, мы иногда будем рассматривать вспомогательную функцию одной переменной

$$\varphi(t) = f(p + tv),$$

где вектор v фиксирован и обычно мал: по смыслу это Δx, но писать так меньше.

Лемма. *При достаточной гладкости функции f имеем*

$$\varphi^{(k)}(t) = d^k f_{p+tv}(v).$$

Доказательство. Давнишнее соглашение, что производная нулевого порядка означает самую функцию, распространим и на дифференциал. Тогда $d^0 f$ зависит лишь от точки приложения и $d^0 f_{p+tv} = f(p+tv)$.

Перейдём к покоординатным обозначениям:

$$\varphi(t) = f(p_1 + tv_1, \ldots, p_n + tv_n).$$

Дифференцируя по правилу цепочки, получим

$$\varphi'(t) = \sum_\alpha \frac{\partial f}{\partial x_\alpha}(p + tv) \cdot v_\alpha,$$

что совпадает со значением дифференциала $df_{p+tv}(v)$.

На следующем шаге по линейности и опять по правилу цепочки

$$\varphi''(t) = \sum_\beta \Big(\sum_\alpha \frac{\partial^2 f}{\partial x_\alpha \partial x_\beta}(p + tv) \cdot v_\alpha \Big) \cdot v_\beta,$$

что совпадает со значением $d^2 f_{p+tv}(v)$ второго дифференциала.

Уяснив происходящее, дальше применим индукцию. □

Формула Тейлора в дифференциалах. Она имеет совершенно тот же вид, что и для функций одной переменной, хотя там этот вид указан лишь в упражнении, а теперь его избежать трудно:

$$\begin{aligned} f(p + \Delta x) &= f(p) + df_p(\Delta x) + \tfrac{1}{2!} d^2 f_p(\Delta x) + \tfrac{1}{3!} d^3 f_p(\Delta x) + \ldots \\ &= \sum_{0 \leqslant k \leqslant m} \tfrac{1}{k!} d^k f_p(\Delta x) + o(\|\Delta x\|^m). \end{aligned}$$

Теорема. *Формула Тейлора в дифференциалах верна для любой функции класса $C^m(U)$.*

Доказательство. Выберем точку $p \in U$ и близкую точку $p + \Delta x$. Избавимся от многомерности, поняв, что на самом деле нас интересует лишь отрезок, соединяющий эти точки. Поэтому введём функцию одной переменной $\varphi(t) = f(p + t\Delta x)$. По последней лемме

$$\varphi(0) = f(p), \quad \varphi'(0) = df_p(\Delta x), \quad \ldots, \quad \varphi^{(m)}(0) = d^m f_p(\Delta x).$$

Обычная одномерная формула Тейлора для $\varphi(1) = f(p + \Delta x)$ даёт требуемое разложение. □

Первый и второй дифференциалы должны стать хорошо знакомы, поскольку они нужны во многих других вопросах. При необходимости явным образом приблизить функцию с привлечением и третьего дифференциала (мультипольное разложение в электродинамике), нужно использовать его развёрнутое выражение.

Операторные обозначения. Дифференцирование по x есть операция на функциях: $f(x) \mapsto f'(x)$ и $f(x,y) \mapsto \frac{\partial f}{\partial x}(x,y)$. Бывает полезно иметь обозначение для самой этой операции, отвлечённое от конкретных функций. Естественнее всего взять $\frac{\partial}{\partial x}$, либо краткое ∂_x.

Ранее в этой главе уже промелькнул пример дифференциального оператора — выражения, комбинирующего дифференцирование с арифметическими операциями. Теперь возьмём другую комбинацию: оператор $\frac{\partial}{\partial x}\,\Delta x$ превращает функцию $f(x,y)$ в функцию $\frac{\partial f}{\partial x}(x,y)\,\Delta x$. Однако мы хотим построить формальный агрегат, также зависящий (линейно) от приращения независимых переменных, как и дифференциал df, чтобы его действие на функциях отвечало правилу $f \mapsto df$. Поэтому для $f(x,y)$ запишем

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

Затем можно выразить второй дифференциал как

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

Раскрывая оператор в скобках, получим оператор

$$\frac{\partial^2}{\partial x^2}\,dx^2 + 2\frac{\partial^2}{\partial x \partial y}\,dx\,dy + \frac{\partial^2}{\partial y^2}\,dy^2,$$

действие которого отвечает правилу $f \mapsto d^2 f$. Функция здесь должна быть дважды гладкой.

При работе с высшими дифференциалами выгодно не раскрывать операторную скобку. Таким образом, для $f(x,y)$ имеем

$$d^m f = \Big(\frac{\partial}{\partial x}\,dx + \frac{\partial}{\partial y}\,dy\Big)^m f,$$

а для $f(x_1,\dots,x_n)$ имеем

$$d^m f = \Big(\sum_\alpha \frac{\partial}{\partial x_\alpha}\,dx_\alpha\Big)^m f.$$

Если же раскрыть сумму, возведённую в степень, то и придём к развёрнутой формуле, обычно записываемой с помощью мульти-индексов.

Гораздо чаще встречается операторное обозначение градиента: это просто символ ∇, действующий по правилу $f \mapsto \nabla f$. Умножая его скалярно на «вектор» $dx = (dx_1, \ldots, dx_n)$, можно ещё ужать скобку:

$$d^m f = \bigl(\nabla \cdot dx\bigr)^m f, \qquad d^m f(\Delta x) = \bigl(\nabla \cdot \Delta x\bigr)^m f.$$

Ряд Тейлора и экспонента. Пропустите остаток раздела при наличии тени сомнения в том, что вы действительно хотите прочесть и понять его. Без малейшего ущерба для понимания остального курса.

Займёмся здесь обещанными ещё менее понятными формальными преобразованиями. А именно, вернёмся к формуле Тейлора в дифференциалах и перейдём от неё к ряду

$$f(\mathbf{p} + \Delta\mathbf{x}) = \sum_{k \geqslant 0} \frac{1}{k!}\, d^k f_{\mathbf{p}}(\Delta\mathbf{x}) = \Bigl(\sum_{k \geqslant 0} \frac{1}{k!} \bigl(\nabla \cdot \Delta\mathbf{x}\bigr)^k f\Bigr)(\mathbf{p}).$$

При раскрытии правой части сперва «вычисляются» все производные как функции, только после этого подставляется точка $\mathbf{p}$; поэтому она и вынесена наружу большой формальной скобки.

Сумма в правой части поразительно похожа на разложение экспоненциальной функции, однако тут вместо аргумента стоит оператор! Выходит, операторные обозначения полезно развить ещё дальше, позволяя формальную запись $\exp(\nabla \cdot \Delta\mathbf{x})$, с помощью которой ряд Тейлора выражается формулой

$$f(\mathbf{p} + \Delta\mathbf{x}) = \bigl(\exp(\nabla \cdot \Delta\mathbf{x}) f\bigr)(\mathbf{p}).$$

Жирный шрифт здесь призван подчеркнуть, что в формулу входят векторы в $\mathbb{R}^n$, в отличие от следующих ниже.

А теперь применим эти операторы в случае функции $f(x, y)$ вблизи начала координат, с которого начат текущий раздел. Приращения независимых переменных составят вектор $(\Delta x, \Delta y)$, а формальный оператор, действующий на функцию в точке $(0, 0)$, превратится в

$$\exp\bigl(\partial_x \Delta x + \partial_y \Delta y\bigr) = \exp\bigl(\partial_y \Delta y\bigr) \cdot \exp\bigl(\partial_x \Delta x\bigr).$$

Двухшаговое разложение, выполненное в начале раздела, можно выразить в итоге как

$$f(\Delta x, \Delta y) = \exp\bigl(\partial_y \Delta y\bigr) f(\Delta x, 0),$$
$$f(\Delta x, 0) = \exp\bigl(\partial_x \Delta x\bigr) f(0, 0),$$

что и даёт

$$f(\Delta x, \Delta y) = \exp\bigl(\partial_x \Delta x + \partial_y \Delta y\bigr) f(0, 0).$$

Аналогично можно поступить для функции нескольких переменных. Хотя эти фокусы нельзя считать доказательством многомерной формулы Тейлора на основе одномерной, они приносят весьма интересное описание её структуры.

9.5. Локальные экстремумы

Продолжаем изучать функции на открытом множестве $U \subseteq \mathbb{R}^n$.

Определение. Точку p называют точкой **локального максимума** функции f, когда $f(x) \leqslant f(p)$ вблизи неё. Формально: когда существует такое $\varepsilon > 0$, что неравенство верно для всех $x \in B_\varepsilon(p)$.

Обратное неравенство определяет точку **локального минимума**. Максимумы и минимумы вместе называют **экстремумами**. Строгие неравенства определяют строгие экстремумы.

Необходимые условия. Определение дословно совпадает с одномерным случаем, условия же локального экстремума в многомерном анализе лаконично выражаются через дифференциалы. Однако прежде всего посмотрим на простейшие случаи с двумя переменными.

Постоянная функция не имеет строгих экстремумов, но нестрогие максимумы и одновременно минимумы — везде; график является горизонтальной плоскостью. рисунки Ненулевая линейная функция (форма) не имеет экстремумов; график является наклонной плоскостью. Эти наблюдения приводят к первому необходимому условию экстремума.

Квадратичные формы дают несколько типов картинок. Графики называют параболоидами. Какие из них имеют локальные экстремумы в начале координат? рисунки Вот почти полный список форм, к которым остальные сводятся заменами, изучаемыми в линейной алгебре:

- форма $x^2 + y^2$ имеет строгий минимум (чашечка);
- форма $-x^2 - y^2$ имеет строгий максимум (шапочка);
- форма $x^2 - y^2$ не имеет экстремума (седловинка);
- форма x^2 имеет только нестрогий минимум (долинка).

Эти наблюдения приводят ко второму необходимому условию экстремума.

Теорема. *В каждой точке локального экстремума гладкой функции равны нулю:*

(1) *все её частные производные;*
(2) *её производная по любому вектору;*
(3) *её дифференциал.*

Доказательство. (1) Первое утверждение следует из второго.

(2) Вблизи точки p локального экстремума гладкой функции f одномерное сечение $\varphi(t) = f(p+tv)$ с любым вектором v гладкое и имеет экстремум в точке $t = 0$. По необходимому условию экстремума для функций одной переменной $\varphi'(0) = 0$, а по лемме об одномерном сечении $\varphi'(0) = df_p(v)$, что совпадает с производной $\partial_v f(p)$ по вектору v.

(3) Только что вектор v был любым, поэтому дифференциал df_p равен нулю тождественно — как линейная функция. $\square$

Определение. Если $df_p = 0$, то точку p называют **стационарной** точкой функции f.

Примеры. Функции $f(x, y) = x^2 + y^2$ и $g(x, y) = x^2 - y^2$ имеют единственную стационарную точку $(0, 0)$.

В стационарной точке поверхности может находиться неподвижный шарик, отсюда и название. Говорят также о **критических** точках, где дифференциал или равен нулю, или вообще не существует. Иными словами, в критической точке p нарушается условие $df_p \neq 0$, или $\nabla f(p) \neq 0$, что равносильно, но чаще сопутствует этому термину. Однако для дифференцируемых функций стационарные и критические точки это одно и то же.

Теорема. *Для всякой дважды гладкой функции f:*

(1) *в точке p локального минимума всегда $d^2f_p \geqslant 0$;*
(2) *в точке p локального максимума всегда $d^2f_p \leqslant 0$.*

Неравенства означают соответственно неотрицательность и неположительность квадратичной формы d^2f_p. Эти понятия вводятся в курсе линейной алгебры и геометрии: $d^2f_p \geqslant 0$ пишут, когда $d^2f_p(v) \geqslant 0$ для всех векторов v. Если $d^2f_p(v) > 0$ для всех $v \neq 0$, то пишут $d^2f_p > 0$ и называют форму положительно определённой.

Доказательство. Как и в предыдущей теореме, для одномерного сечения $\varphi(t) = f(p + tv)$ с любым вектором v аналогичное условие из одномерного анализа даёт $\varphi''(0) \geqslant 0$ либо обратное неравенство, соответственно случаю. По лемме о сечении $\varphi''(0) = d^2f_p(v)$. $\square$

Следствие. *Если квадратичная форма d^2f_p знакопеременна, то f не имеет экстремума в точке p.*

Достаточные условия. Естественно ожидать, что основное достаточное условие локального экстремума упоминает второй дифференциал. Доказательство же довольно интересно.

Лемма. *Для всякой положительно определённой квадратичной формы q найдётся такое $\mu > 0$, что $q(v) \geqslant \mu\|v\|^2$ для всех векторов v.*

Доказательство. Сфера радиуса 1 вокруг начала координат компактна. По теореме Вейерштрасса, непрерывная функция q на ней достигает минимального значения, которое положительно, поскольку $q(v) > 0$ для всех ненулевых векторов v.

Возьмём любое $\mu > 0$ ниже минимума. Требуемое неравенство выполнено для всех единичных векторов (векторов из начала координат в точку на единичной сфере). Поскольку обе его стороны однородны и квадратичны, оно верно и вообще для всех векторов. □

Теорема. *В каждой стационарной точке p гладкой функции f:*

(1) *если $d^2f_p > 0$, то p есть строгий локальный минимум;*
(2) *если $d^2f_p < 0$, то p есть строгий локальный максимум.*

Доказательство. (1) Вблизи стационарной точки

$$\Delta f = f(p + \Delta x) - f(p) = \tfrac{1}{2}d^2f_p(\Delta x) + o(\|\Delta x\|^2).$$

Если $d^2f_p > 0$, то оценим его значение снизу по лемме и получим

$$\Delta f \geqslant (\tfrac{\mu}{2} + o(1))\|\Delta x\|^2,$$

что положительно при малых $\Delta x \neq 0$, и локальный минимум налицо.

(2) Заменив функцию на $-f$, обнаружим первый случай. □

Совершенно другое и очень короткое доказательство можно дать, сославшись на достижения линейной алгебры и геометрии. Поворотом системы координат квадратичную форму q приводят к виду

$$q(v) = \sum_\alpha \lambda_\alpha v_\alpha^2.$$

Коэффициенты λ_α называются собственными числами квадратичной формы. Вы уже встречали их в механике в обличии главных моментов инерции, а вскоре подробно займётесь ими в курсе линейной алгебры. Если знаки всех λ_α для формы d^2f_p в стационарной точке p одинаковые, то функция имеет там строгий локальный экстремум.

Таким образом, в простых случаях алгоритм отыскания локальных экстремумов состоит из двух шагов. На первом шаге находим стационарные точки из уравнения $df = 0$. На втором шаге проверяем знакоопределённость d^2f алгебраическими методами, обычно с помощью критерия Сильвестра.

Примеры. Рассмотрим простенькие примеры, где сами функции

$$f_1(x,y) = x^2 + y^2, \qquad f_3(x,y) = 2xy,$$
$$f_2(x,y) = x^2 - y^2, \qquad f_4(x,y) = x^2 + 2xy + y^2$$

есть квадратичные формы. Поэтому их матрицы

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

равны матрицам соответствующих квадратичных форм $\frac{1}{2}d^2f_i(0,0)$.

Лишь первая из них положительно определена и соответствует минимуму функции. Вторая и третья функции не имеют экстремума. Четвёртая форма неотрицательно полуопределена и вырождена; выбранная здесь функция имеет нестрогий экстремум.

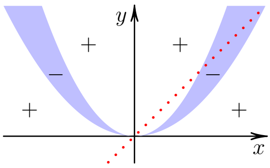
Вырожденность второго дифференциала не даёт оснований что-либо заключить о локальном экстремуме, поскольку тогда существенны дифференциалы более высоких порядков, очень сложные для исследования. За счёт вклада высших дифференциалов в разложение горизонтальная долинка на графике второго дифференциала может изогнуться в вытянутую чашечку или седловинку на графике самой исследуемой функции. Тогда требуется дополнительное рассмотрение значений функции в точках, близких к стационарной, универсальных методов для которого нет. рисунки

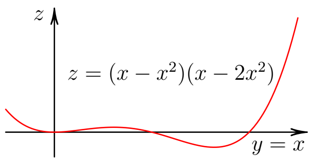
Пример. Рассмотрим функцию Peano 1884

$$z(x,y) = 2x^4 - 3x^2y + y^2 = (y - x^2)(y - 2x^2).$$

Второй дифференциал $d^2z(0,0) = 2dy^2$ вырожден.

Локального минимума в начале координат нет: в точках между двух парабол функция отрицательна, а в других точках вне этих парабол она положительна.





Это популярный пример, потому что здесь каждое одномерное сечение функции через начало координат имеет там локальный минимум.

Глава 10. ГЛАДКИЕ ОТОБРАЖЕНИЯ

10.1. Гладкие отображения и матрица Якоби

Этот раздел является переходным между двумя главами курса. Исходные понятия непосредственно развивают некоторые идеи предыдущей главы и в то же время играют главные роли в дальнейшем. Здесь мы обращаемся к более сложным объектам, хотя сначала делаем с ними немногое. Всё по-прежнему происходит на открытом множестве $U \subseteq \mathbb{R}^n$, но теперь мы изучаем отображение $f\colon U \to \mathbb{R}^m$.

01.04.17 **Ассортимент отображений.** Разнообразные математические объекты, с которыми мы будем работать во втором семестре, представляют собой частные случаи отображений из (подмножества) пространства $\mathbb{R}^n$ в пространство $\mathbb{R}^m$, в основном для малых размерностей.

	в $\mathbb{R}^1$	в $\mathbb{R}^2$	в $\mathbb{R}^3$
из $\mathbb{R}^1$	функция	плоская линия	линия
из $\mathbb{R}^2$	плоское скалярное поле	плоское вект. поле, система координат	поверхность
из $\mathbb{R}^3$	скалярное поле	—	векторное поле, система координат

Мы уже встречались с параметризованными плоскими линиями:

$$x = x(t), \quad y = y(t).$$

Это примеры отображений из подмножества $\mathbb{R}^1$ в $\mathbb{R}^2$. Будем иногда выражать их краткой записью $t \mapsto (x, y)$. Совершенно аналогично абстрактное представление о параметризованной линии в пространстве как отображении $t \mapsto (x, y, z)$.

Поверхность в пространстве параметризуют двумя параметрами:

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v),$$

так что это отображение $(u, v) \mapsto (x, y, z)$.

Случай отображений с $m = 1$ мы изучали в этой главе: это функции нескольких переменных. В физических приложениях, а также в тесно связанном с ними векторном анализе, функции на пространстве зачастую называют полями, уточняя тип поля прилагательным, указывающим тип его значения в каждой точке; прежде всего выделяют скалярные и векторные поля. Температура является примером скалярного поля, а скорость в жидкости — векторного. Поэтому скалярные поля стоят в первом столбце таблицы, а векторные по диагонали.

Полевые термины нам пока не нужны, но другие отображения с $m = n$ исключительно важны: это системы координат и различные преобразования. Последний термин подразумевает обратимость. Мы остановимся на этом подробнее в следующем разделе, ограничившись в основном плоским случаем.

Матрица Якоби. Представим отображение f набором его координатных функций $f_i \colon U \to \mathbb{R}$. Так,

$$f(x) = [f_1(x), \dots, f_m(x)]^\top$$

для всех $x = [x_1, \dots, x_n]^\top$. Значки транспонирования указывают на то, что все векторы мы теперь считаем столбцами. Однако зачастую вместо $[x_1, \dots, x_n]^\top$ проще писать $(x_1, \dots, x_n)$.

Определение. Отображение называют **гладким**, если все его координатные функции гладкие.

Также по координатам можно говорить о дифференцируемом отображении, хотя это определение мы слегка обработаем с целью удаления лишнего.

Если вблизи точки $x \in U$ приращение каждой координатной функции $f_i(x)$ представимо суммой линейной и малой функций,

$$f_i(x + \Delta x) - f_i(x) = \mathcal{A}_i(\Delta x) + o(\|\Delta x\|),$$

то мы вправе утверждать это же непосредственно об отображении f,

$$f(x + \Delta x) - f(x) = \mathcal{A}(\Delta x) + o(\|\Delta x\|),$$

с той лишь разницей, что $\mathcal{A}$ здесь — линейное отображение $\mathbb{R}^n \to \mathbb{R}^m$, а не линейная функция; также и o-малое стало вектором. Естественно, тогда $\mathcal{A}$ называют **дифференциалом** отображения f в точке x. Упирать на это понятие в нашем курсе мы не будем, сосредоточив внимание на явном выражении для $\mathcal{A}$.

Из линейной алгебры вам должен быть известен общий вид линейного отображения пространств столбцов $\mathbb{R}^n \to \mathbb{R}^m$. Оно всегда представимо как умножение на $m \times n$ матрицу, поэтому

$$\Delta f = A \cdot \Delta x + o(\|\Delta x\|).$$

Элементы её находятся легко, ибо мы уже знаем их в случае $m = 1$. Там это частные производные; значит, теперь в i-ой строке матрицы размещаются частные производные i-ой координатной функции.

Определение. Функциональную $m \times n$ матрицу с элементами $\partial f_i / \partial x_j$ называют **матрицей Якоби** отображения $f(x)$ и обозначают через

$$\frac{Df}{Dx} = \frac{D(f_1, \ldots, f_m)}{D(x_1, \ldots, x_n)}.$$

Следствие. *Отображение является гладким $\Leftrightarrow$ его матрица Якоби непрерывна.*

Главные свойства матриц Якоби — их поведение при композиции и обращении отображений.

Дифференцирование композиции. Правило цепочки для функций нескольких переменных, переписанное в матричном виде, подскажет нам общий вид этой формулы. Больше того, в том частном случае уже присутствует композиция отображения $x(t) = [x_1(t), \ldots, x_n(t)]^\top$ и функции $y(x)$, а матричный вид правила связывает градиенты посредством матрицы Якоби: развёрнуто

$$\left[\frac{\partial y}{\partial t_1}, \ldots, \frac{\partial y}{\partial t_k}\right] = \left[\frac{\partial y}{\partial x_1}, \ldots, \frac{\partial y}{\partial x_m}\right] \cdot \begin{bmatrix} \dfrac{\partial x_1}{\partial t_1} & \cdots & \dfrac{\partial x_1}{\partial t_m} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial x_n}{\partial t_1} & \cdots & \dfrac{\partial x_n}{\partial t_m} \end{bmatrix},$$

или кратко $\nabla_t y = \nabla_x y \cdot \frac{Dx}{Dt}$. Действительно, выполнив матричное умножение, увидим знакомое поэлементное равенство

$$\frac{\partial y}{\partial t_j} = \sum_i \frac{\partial y}{\partial x_i} \cdot \frac{\partial x_i}{\partial t_j}.$$

Теперь остаётся только заметить, что $\nabla_t y = \frac{Dy}{Dt}$ и аналогично по x, и для композиции отображений $y(x(t))$ обнаруживается общая формула

$$\frac{Dy}{Dt} = \frac{Dy}{Dx} \cdot \frac{Dx}{Dt}.$$

Она матричная, поэтому порядок сомножителей нужно соблюдать.

Теорема. *Композиция дифференцируемых отображений является дифференцируемым отображением.*

Доказательство. Рассуждение практически дословно повторяет случай композиции функций, но теперь более сжато. Для двух дифференцируемых отображений $x \mapsto y(x)$ и $t \mapsto x(t)$ запишем приращения

$$\Delta y = A(x)\Delta x + o(\|\Delta x\|),$$
$$\Delta x = B(t)\Delta t + o(\|\Delta t\|),$$

где $A(x)$ и $B(t)$ это матрицы Якоби. Подставим второе представление в первое и упростим o-малые:

$$\Delta y = A(x(t))B(t)\Delta t + o(\|\Delta t\|).$$

Произведение матриц AB, конечно, тоже матрица, поэтому формула представляет $\Delta t \mapsto \Delta y$ суммой линейного отображения и малого. □

Следствие. *Для гладких отображений $y = y(x)$ и $x = x(t)$ равенство*

$$\frac{Dy}{Dt}(t_0) = \frac{Dy}{Dx}(x(t_0)) \cdot \frac{Dx}{Dt}(t_0)$$

верно в каждой точке t_0, где эти матрицы определены.

Доказательство. Формула обоснована в доказательстве теоремы. □

Следствие. *Композиция отображений сохраняет гладкость.*

Доказательство. Перемножение непрерывных матриц Якоби даёт непрерывную матрицу Якоби. □

Дифференцирование обратного отображения. Мы изучали правило дифференцирования обратной функции одной переменной. Вопрос об обращении функции нескольких переменных лишён смысла и обретает его лишь при переходе к отображениям.

Пример (примитивная замена). Замены переменных применяются в самых разных задачах. Один из случаев замены — переход от переменных (x, v) к переменным (u, v) подстановкой $x = x(u, v)$. Отметим, что, понимая это как отображение, следует писать $(u, v) \mapsto (x, v)$, то есть направление стрелки обратно смыслу слова «переход». Ввиду независимости v от u, матрица Якоби этого отображения равна 02.04.17

$$\frac{D(x, v)}{D(u, v)} = \begin{bmatrix} x_u & x_v \\ 0 & 1 \end{bmatrix}.$$

При условии $x_u \neq 0$ легко найти обратную матрицу

$$\left(\frac{D(x, v)}{D(u, v)}\right)^{-1} = \begin{bmatrix} 1/x_u & -x_v/x_u \\ 0 & 1 \end{bmatrix}.$$

Это же условие гарантирует, что замена обратима, то есть определена функция $u = u(x, v)$ и тем самым отображение $(x, v) \mapsto (u, v)$. Его матрица Якоби равна

$$\frac{D(u, v)}{D(x, v)} = \begin{bmatrix} u_x & u_v \\ 0 & 1 \end{bmatrix}.$$

При этом $u_x = 1/x_u$ из теоремы об обратной функции, а $u_v = -x_v/x_u$ по формуле для производной неявной функции; для её применения нужно учесть, что по определению частной производной $u_v = \partial u/\partial v$ вычисляется при постоянном x. Таким образом,

$$\frac{D(u,v)}{D(x,v)} = \left(\frac{D(x,v)}{D(u,v)}\right)^{-1}.$$

Получим теперь такую формулу для любых обратимых гладких отображений. Обозначим через E_n единичную матрицу размера $n \times n$.

Лемма. *Если матрицы A и B удовлетворяют $AB = E_m$ и $BA = E_n$, то они квадратны и обратны друг другу.*

Доказательство. Это линейная алгебра: ранг произведения матриц не превосходит рангов сомножителей; ранг единичной матрицы равен её размеру ($m = n$); квадратная матрица полного ранга обратима. □

Теорема. *Если два гладких отображения взаимно обратны, то взаимно обратны и их матрицы Якоби.*

Доказательство. Взаимная обратность пары отображений $x \mapsto y(x)$ и $y \mapsto x(y)$ означает, что $y = y(x(y))$ и $x = x(y(x))$. В доказательстве предыдущей теоремы возьмём $t = y$ и придём к условию $AB = E_m$ для композиции $y(x(y))$. Для композиции в обратном порядке таким же образом получим $BA = E_n$. Алгебраическая лемма завершает рассуждение. □

Заключение теоремы выражается формулой

$$\frac{Dx}{Dy}(y(x_0)) = \left(\frac{Dy}{Dx}(x_0)\right)^{-1}.$$

Следствие. *Если два гладких взаимно обратных отображения определены на открытых подмножествах $U \subseteq \mathbb{R}^n$ и $V \subseteq \mathbb{R}^m$, то $m = n$.*

10.2. Криволинейные системы координат

Сетка координат на плоскости. Рассмотрим для начала два примера отображений: переход от декартовых координат к полярным, осу-
01.04.17 ществляемый отображением $(\rho, \varphi) \mapsto (x, y)$ по правилу

$$x = \rho\cos\varphi, \quad y = \rho\sin\varphi,$$

и какое-нибудь линейное отображение $(u, v) \mapsto (x, y)$, скажем,

$$x = 2u + v, \quad y = -u + v.$$

Второе отображение задаёт косоугольные прямолинейные координаты. Изобразим на плоскости Oxy **координатные линии** этих двух систем: окружности $\rho = \text{const}$ и лучи $\varphi = \text{const}$ для полярной системы; прямые $u = \text{const}$ и $v = \text{const}$ для взятой наугад косоугольной системы. Совокупность всех координатных линий системы называют **координатной сеткой**. Нужно оговаривать область изменения новых координат; тем самым, сетка может покрывать лишь часть плоскости. В каждой системе криволинейных координат через каждую точку покрытой сеткой части как правило проходит ровно две координатных линии (ниже будет ясно, почему в полярной системе это нарушается при $\rho = 0$). рисунок

Криволинейные системы координат целесообразно применять, когда они упрощают какое-то из условий задачи. Нередко удаётся распознать в одной из заданных линий координатную линию подходящей системы, но в других случаях выбор бывает не очевиден. Иногда приходится изобретать систему под геометрию задачи.

Цилиндрические и сферические координаты. Трёхмерные задачи, лишённые всякой симметрии, зачастую оказываются недоступны анализу. Однако наиболее важны задачи с вращательной или сферической симметрией. При наличии первой и отсутствии второй применяют цилиндрическую систему координат (ρ, φ, z). Она связана с декартовой (x, y, z) формулами

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

Таким образом, можно считать, что цилиндрическая система получена поступательным движением полярной системы в плоскости Oxy вдоль оси Oz. Как и в других системах в трёхмерном пространстве, координатная сетка состоит из трёх семейств поверхностей. Здесь это прямые круговые цилиндры $\rho = \text{const}$, полуплоскости $\varphi = \text{const}$ и плоскости $z = \text{const}$.

Сферическую систему координат вводят немного по-разному, поэтому сто́ит привыкнуть всегда обращать внимание на смысл букв и формулы связи с декартовыми координатами.

В физике стандартизирована сферическая система (r, θ, φ), пришедшая из астрономии: расстояние, зенитный угол, азимут. Тут зенитный угол θ отсчитывается от зенита (аналога северного полюса). Тогда

$$\begin{cases} x = r\sin\theta\cos\varphi, \\ y = r\sin\theta\sin\varphi, \\ z = r\cos\theta. \end{cases}$$

Координатные поверхности $r = \mathrm{const}$ это конечно же сферы; $\varphi = \mathrm{const}$ это полуплоскости, проходящие через ось Oz; условие $\theta = \mathrm{const}$ задаёт половину кругового конуса, а также плоскость Oxy в случае $\theta = {}^{\pi}/_{2}$.

В иных сферических системах, помимо отличия в отсчёте одного из углов, могут быть переставлены местами обозначения углов. Порядок записи углов важен, ибо от него зависит, будет система правой или левой; физики же всегда стараются работать в правой системе.

Мне больше нравится система (r, φ, θ) принятая в географии, с отсчётом широты θ от экватора. Тогда

рисунок

$$\begin{cases} x = r\cos\theta\cos\varphi, \\ y = r\cos\theta\sin\varphi, \\ z = r\sin\theta. \end{cases}$$

Формулы перехода можно разбить на два перехода к цилиндрическим координатам, причём первый стандартен:

$$\begin{cases} x = \rho\cos\varphi, \\ y = \rho\sin\varphi, \\ z = z; \end{cases} \qquad \begin{cases} \rho = r\cos\theta, \\ \varphi = \varphi, \\ z = r\sin\theta. \end{cases}$$

Значит, вторая тройка равенств связывает стандартную цилиндрическую и географическую сферическую системы.

Упражнение. *Найдите матрицу Якоби цилиндрической и сферической систем. Проследите перемножение матриц, соответствующее разбиению перехода от декартовых координат к сферическим на два цилиндрических перехода.*

Дальнейшие примеры плоских координат. Познакомимся теперь кратко с некоторыми другими полезными системами криволинейных координат на плоскости.

Пример. Параболические координаты задают отображением

$$x = \sigma\tau, \quad y = (\tau^2 - \sigma^2)/2.$$

Координатная сетка состоит из двух семейств парабол с общим фокусом $(x, y) = (0, 0)$ и общей осью, совпадающей с Oy. Параболы разных семейств всегда пересекаются под прямым углом. Такие криволинейные координаты называют **ортогональными**. Координаты (σ, τ) и $(-\sigma, -\tau)$ соответствуют одной точке, поэтому для взаимной однозначности нужно ограничиваться половиной плоскости $O\sigma\tau$.

рисунки

Для обратного отображения $(x, y) \mapsto (\sigma, \tau)$ можно отметить, что координатные линии $x = \text{const}$ и $y = \text{const}$ на плоскости $O\sigma\tau$ являются гиперболами и также ортогональны. Можно найти и формулы, задающие это отображение, решая два похожих биквадратных уравнения. Несмотря на обилие гипербол, это отображение не связано напрямую с гиперболическими системами координат.

Пример. Симметрией относительно единичной окружности, или инверсией, называют преобразование плоскости, которое проще всего записать в полярных координатах: точка с координатами (ρ, φ) переходит в точку с координатами $(1/\rho, \varphi)$. Начало координат обменивается с «бесконечно удалённой точкой», а при неверии в неё просто исключается. Это преобразование обратно само себе, иначе бы его не назвали инверсией.

Вводя для прообразов декартовы координаты (x, y), а для образов декартовы координаты (u, v), находим $u = x/\rho^2$ и $v = y/\rho^2$, так что

$$u = \frac{x}{x^2 + y^2}, \quad v = \frac{y}{x^2 + y^2}.$$

Координатная сетка состоит из двух ортогональных семейств окружностей, проходящих через точку $(x, y) = (0, 0)$.

рисунок

Обратное отображение $(u, v) \mapsto (x, y)$, конечно, задаётся такими же формулами, нужно только поменять местами буквы $u \leftrightarrow x$, $v \leftrightarrow y$.

Пример. Гиперболические координаты встречаются в двух вариантах. Первый вариант, довольно старинный, задают отображением

$$x = ve^{u}, \quad y = ve^{-u}.$$

Оно попадает только в два квадранта, ибо получается $xy \geqslant 0$, но удобнее потребовать $v > 0$ и вследствие того покрывать только первый квадрант $x > 0$, $y > 0$. Соответственно, формулы

$$u = \frac{1}{2} \ln \frac{x}{y}, \quad v = \sqrt{xy}$$

задают обратное отображение. Координатная сетка состоит из открытых лучей и ветвей гипербол, которые пересекаются под различными углами; значит, в отличие от всех предыдущих примеров, эта система координат не ортогональна.

рисунки

Второй вариант используется в специальной теории относительности и поэтому вам уже встречался. Он отличается от первого иными обозначениями переменных, $(\xi, \theta) \mapsto (x, t)$, и поворотом декартовых осей на 45°, собирающим две экспоненты в гиперболические функции:

$$x = \xi \operatorname{ch} \theta, \quad y = \xi \operatorname{sh} \theta.$$

Иногда бывают полезны системы координат, связанные с двумя точками, которые называют **фокусами**. Декартову систему обычно выбирают так, что фокусы лежат в точках $(x, y) = (\pm 1, 0)$, или $(\pm a, 0)$ при необходимости.

Пример. Эллиптические координаты задают отображением

$$x = a \operatorname{ch} \mu \cos \nu, \quad y = a \operatorname{sh} \mu \sin \nu$$

рисунок

с ограничениями $\mu \geqslant 0$ и $0 \leqslant \nu < 2\pi$. Координатная сетка ортогональна и состоит из эллипсов и гипербол с фокусами в точках $(\pm a, 0)$, причём на каждой гиперболе лежит четыре координатных линии: по одной в квадранте.

Пример. Биполярные координаты задают отображением

$$x = h \operatorname{sh} \tau, \quad y = h \sin \sigma, \quad \text{где } h = \frac{a}{\operatorname{ch} \tau - \cos \sigma}.$$

При этом σ равно углу, под которым из точки виден отрезок между фокусами, а τ равно логарифму отношения расстояний от точки до фокусов. Координатная сетка ортогональна и состоит из двух семейств окружностей. Каждая окружность одного семейства проходит через оба фокуса и составлена из двух дуг со значениями σ, отличающимися прибавлением π. Окружности второго семейства известны с античности; их называют окружностями Аполлония.

Упражнение. *Проверьте для биполярных координат соотношения*

$$\operatorname{tg} \sigma = \frac{2ay}{x^2 + y^2 - a^2}, \quad \operatorname{th} \tau = \frac{2ax}{x^2 + y^2 + a^2}.$$

Упражнение. *Проверьте два «предельных случая» биполярных координат: при удалении от их плоскости они выглядят как система, заданная инверсией; вблизи фокуса они выглядят как система, получаемая из полярной заменой в радиальной переменной.*

Дальнейшие примеры трёхмерных координат. Способ получения цилиндрических координат из полярных применим к любым плоским системам. Так из приведённых выше примеров получают параболические цилиндрические координаты, эллиптические цилиндрические координаты и бицилиндрические (то есть биполярные цилиндрические) координаты.

Другой способ заключается во вращении плоскости в пространстве. Обычную цилиндрическую систему получают так из декартовой системы в плоскости Oxz. А взяв там криволинейные координаты, можно получить другую полезную трёхмерную систему. Два семейства координатных поверхностей всегда будут состоять из поверхностей вращения, а третье — из полуплоскостей через ось.

Пример. Вращение параболической системы вокруг общей оси всех координатных линий (парабол) даёт параболоидальную систему координат. Координатных поверхности двух семейств — параболоиды вращения.

Примеры. Биполярную систему можно вращать либо вокруг прямой, содержащей фокусы, либо вокруг прямой, равноотстоящей от фокусов. В первом случае результат называют бисферической системой; в ней одно семейство координатных поверхностей состоит из сфер, а другое сложнее (поверхности четвёртого порядка). Во втором случае помимо сфер образуются торы, а результат называют тороидальной системой.

Примеры. Эллиптическую систему тоже можно вращать двумя способами, получая эллипсоиды и гиперболоиды вращения в качестве координатных поверхностей. Это даёт системы координат вытянутого сфероида и сплющенного сфероида.

С помощью аналогий придём теперь к двум системам трёхмерных координат, не получаемым вращением плоских систем.

Пример. По аналогии с инверсией относительно окружности, устроим инверсию в пространстве относительно единичной сферы. В сферических координатах точка с координатами (r, θ, φ) переходит в точку с координатами $(1/r, \theta, \varphi)$. Вводя для прообразов декартовы координаты (x, y, z), а для образов декартовы координаты (u, v, w), находим

$$u = \frac{x}{x^2 + y^2 + z^2}, \quad v = \frac{y}{x^2 + y^2 + z^2}, \quad w = \frac{z}{x^2 + y^2 + z^2}.$$

Координатные поверхности являются сферами, касающимися декартовых координатных плоскостей в начале. Эту систему иногда называют 6-сферной.

Пример. Этот пример посложнее предшествующих. Чтобы выстроить нужную аналогию, сперва зафиксируем числа $a > b > 0$ и рассмотрим в декартовой системе на плоскости уравнение

$$\frac{x^2}{a^2 - \delta} + \frac{y^2}{b^2 - \delta} = 1$$

при различных значениях δ. Оно задаёт эллипсы при $\delta < b^2$, гиперболы при $b^2 < \delta < a^2$ и пустышки (мнимые эллипсы) при $\delta > a^2$. Фокусы всегда находятся в точках $(\pm c, 0)$, где

$$c^2 = (a^2 - \delta) - (b^2 - \delta) = a^2 - b^2.$$

Поскольку фокусы общие, все эти кривые являются координатными для эллиптической системы координат; более того, все её координатные линии описываются так, за исключением прямолинейных, появляющихся в пределах при $\delta \to a^2$ и $\delta \to b^2$.

Наоборот, для заданной точки (x, y), лучше вне координатных осей, получаем квадратное уравнение на δ. Два его решения окажутся вещественными в «правильных» диапазонах и будут эллиптическими координатами (μ, ν) этой точки: $\mu < b^2 < \nu < a^2$.

Теперь перейдём к трёхмерному пространству. Зафиксируем числа $a > b > c > 0$ и рассмотрим уравнение

$$\frac{x^2}{a^2 - \delta} + \frac{y^2}{b^2 - \delta} + \frac{z^2}{c^2 - \delta} = 1.$$

Аналогично предыдущему, при значениях δ в трёх диапазонах получаем эллипсоиды, однополостные гиперболоиды и двуполостные гиперболоиды. Никаких фокусов у них нет, но можно показать, что две таких поверхности разных типов всегда пересекаются ортогонально. Это координатные поверхности эллипсоидальной системы координат.

Наоборот, для заданной точки (x, y, z), лучше вне координатных плоскостей, получаем кубическое уравнение на δ. Все три его решения окажутся вещественными в «правильных» диапазонах и будут эллипсоидальными координатами (λ, μ, ν) этой точки:

$$\lambda < c^2 < \mu < b^2 < \nu < a^2.$$

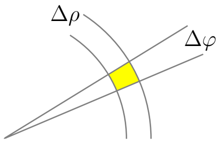
Пытаться выразить (λ, μ, ν) через (x, y, z) смысла нет, а вот наоборот сделать несложно, только громоздко. Три копии исходного уравнения, в которые по очереди подставлены λ, μ и ν вместо δ, собираются в систему линейных уравнений на x^2, y^2 и z^2; решение можно записать по формуле Крамера, а затем упростить определители. Это даёт ответ:

$$x^2 = \frac{(a^2-\lambda)(a^2-\mu)(a^2-\nu)}{(a^2-b^2)(a^2-c^2)},$$

а выражения для y и z находятся отсюда циклической перестановкой букв $x \mapsto y \mapsto z \mapsto x$, $a \mapsto b \mapsto c \mapsto a$. Взаимной однозначности достигают, ограничиваясь одним октантом: $x>0$, $y>0$, $z>0$.

10.3. Обратимость отображения и якобиан

Наводящие соображения об искажении объёма. Координатные линии полярной системы координат разбивают плоскость Oxy на кольцевые секторы. Найдём площадь ΔS такого сектора с радиусами от ρ до $\rho+\Delta\rho$ и углами от φ до $\varphi+\Delta\varphi$.



Площадь кольца равна

$$\pi(\rho+\Delta\rho)^2 - \pi\rho^2 = 2\pi\rho\Delta\rho + \pi(\Delta\rho)^2 = 2\pi\big(\rho+o(1)\big)\Delta\rho.$$

Здесь 2π как раз равно углу полного оборота, поэтому

$$\Delta S = \big(\rho+o(1)\big)\Delta\rho\Delta\varphi.$$

С другой стороны, формулы перехода от полярной системы к декартовой можно понимать как отображение $(\rho,\varphi)\mapsto(x,y)$ из области на плоскости (ρ,φ) на плоскость (x,y). При этом прообраз сектора — прямоугольник площади $\Delta\rho\Delta\varphi$. Таким образом, множитель ρ в площади сектора выступает как локальный коэффициент искажения площади при этом отображении. Возникает вопрос: как найти этот коэффициент для любого (обратимого) отображения?

Хорошей точкой старта для поиска ответа является простейшая формула для площади параллелограмма на векторах (x_1, y_1) и (x_2, y_2):

$$S = \det \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix}.$$

Точнее, так находится ориентированная площадь, чувствительная к ориентации параллелограмма (это поясняется в лекциях по алгебре и геометрии). Аналогичным образом, ориентированный объём параллелепипеда на трёх векторах вычисляют как определитель третьего порядка, составленный из координат направляющих векторов.

При линейном отображении

$$\begin{cases} x = x(u, v) = a_{11}u + a_{12}v, \\ y = y(u, v) = a_{21}u + a_{22}v \end{cases}$$

прямоугольник плоскости (u, v) переходит в параллелограмм плоскости (x, y). Отношение площади второго к площади первого, равное определителю матрицы коэффициентов $[a_{ij}]$, и является коэффициентом искажения площади при этом линейном отображении. Для бо́льших размерностей, где говорят уже об объёме и n-мерном объёме, аналогично получаем определитель.

Искажение объёма и якобианы. Перейдём от линейных отображений $(u, v) \mapsto (x, y)$ к гладким. Образ малого прямоугольника со сторонами Δu и Δv будет, вероятно, криволинейным параллелограммом. Его площадь мало отличается от площади параллелограмма, полученного заменой приращений Δx и Δy на их главные линейные части:

$$\Delta x = \frac{\partial x}{\partial u}\Delta u + \frac{\partial x}{\partial v}\Delta v + o(|\Delta u| + |\Delta v|),$$
$$\Delta y = \frac{\partial y}{\partial u}\Delta u + \frac{\partial y}{\partial v}\Delta v + o(|\Delta u| + |\Delta v|).$$

Применяя к ним предыдущий вывод, видим, что локальный коэффициент искажения площади равен определителю матрицы Якоби

$$\frac{D(x, y)}{D(u, v)} = \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix}.$$

Определение. Определитель (квадратной) матрицы Якоби дифференцируемого отображения называют **якобианом** этого отображения.

Пример. При переходе от декартовых координат к полярным

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

Пример. При переходе от декартовых координат к параболическим

$$\frac{\partial(x,y)}{\partial(\sigma,\tau)} = \det\begin{bmatrix} \tau & \sigma \\ -\sigma & \tau \end{bmatrix} = \sigma^2+\tau^2.$$

Упражнение. *Найдите якобиан отображения, задающего симметрию относительно единичной окружности. Упростите ответ.*

Отметим ещё раз, что изменение ориентации (например, от перестановки переменных) приводит к смене знака якобиана. Когда чувствительность якобиана к ориентации мешает, его берут по модулю. Это требуется главном образом при вычислении двойных интегралов посредством замены переменных.

Упражнение. *Найдите якобианы цилиндрической и сферической систем координат двумя способами: геометрическим рассуждением, как для полярных; вычислением определителя матрицы Якоби.*

Якобиан отображения $(u_1,\ldots,u_n)\mapsto(x_1,\ldots,x_n)$ обозначают через

$$\frac{\partial(x_1,\ldots,x_n)}{\partial(u_1,\ldots,u_n)} = \det\frac{D(x_1,\ldots,x_n)}{D(u_1,\ldots,u_n)}.$$

В соответствующей $u=(u_1,\ldots,u_n)$ и $x=(x_1,\ldots,x_n)$ краткой записи

$$\frac{\partial x}{\partial u} = \det\frac{Dx}{Du}$$

можно ошибочно принять якобиан за простую частную производную.

Теорема. *При композиции отображений якобианы перемножаются.*

Доказательство. Представим матрицу Якоби композиции $y(x(t))$ как произведение

$$\frac{Dy}{Dt} = \frac{Dy}{Dx}\cdot\frac{Dx}{Dt}$$

и возьмём определитель. Утверждение следует из того, что определитель произведения матриц равен произведению их определителей. $\square$

Следствие. *Якобианы взаимно обратных отображений взаимно обратны.*

Следствие. *Для обратимости отображения вблизи некоторой точки необходимо, чтобы его якобиан в этой точке отличался от нуля.*

Важнейшим фактом является достаточность этого условия. Мы её установим в одной из ближайших теорем.

17.02.17
02.04.17

Функции комплексной переменной. Простейшие функции комплексной переменной z строятся при помощи арифметических операций. Так мы получаем комплексные полиномы и дроби. Вспомним также, что уже познакомились с комплексной экспонентой.

Все эти функции являются примерами гладких отображений из $\mathbb{R}^2$ в $\mathbb{R}^2$. Попробуем выяснить неформально, какое свойство хороших комплексных функций выделяет их среди всех гладких отображений. Тут достаточно применить правило цепочки для функций.

Соберём из пары (x, y) независимых вещественных переменных комплексную переменную $z = x + iy$, а из пары вещественных функций $u(x, y)$ и $v(x, y)$ попытаемся собрать функцию $w = u + iv$ от z. Это не всегда возможно, но мы как раз и найдём правильные условия.

Сперва приходится заметить, что из исходной пары переменных можно составить, помимо z, равноправную комбинацию $\bar{z} = x - iy$. Наоборот, пара $(z, \bar{z})$ однозначно определяет пару (x, y). Мы попадаем в ситуацию замены переменных и получаем функцию $w(z, \bar{z})$ вместо желаемой $w(z)$. Правило цепочки даёт производные

$$w_{x \pm iy} = \tfrac{1}{2}\big((u_x \pm v_y) + i(v_x \mp u_y)\big).$$

Упражнение. *Проведите это вычисление.*

Чтобы функция w не зависела от $\bar{z}$, естественно потребовать $w_{\bar{z}} = 0$. Все четыре производные в формуле вещественны и комплексное равенство распадается на два вещественных. Это даёт искомые условия

$$u_x - v_y = 0, \quad v_x + u_y = 0,$$

традиционно называемые **условиями Коши — Римана**. Если они выполнены, то формула для w_z упрощается, причём в четырёх вариантах:

$$w_z = u_x + iv_x = u_x - iu_y = v_y - iu_y = v_y + iv_x.$$

Приведённое рассуждение нестрогое, а обычный способ вывода условий Коши — Римана исходит из определения производной как предела и требования, что w_z не зависит от направления приращения z.

Упражнение. *Сравните такие пределы вдоль осей координат.*

Итак, комплексно-дифференцируемая функция из $\mathbb{C}$ в $\mathbb{C}$ по правилу $z = x + iy \mapsto w = u + iv$ является частным случаем отображения из $\mathbb{R}^2$ в $\mathbb{R}^2$, а именно, подчинённым условиям Коши — Римана. Применим эти соотношения к якобиану и раскроем его:

$$\det \begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix} = \det \begin{bmatrix} u_x & u_y \\ -u_y & u_x \end{bmatrix} = u_x^2 + u_y^2 = |w_z|^2.$$

При детальном изучении на втором курсе выяснится, что при таком отображении сохраняются углы. Значит, вблизи каждой точки происходят подобие и поворот; модуль производной равен искажению длины подобием, а аргумент производной равен углу поворота.

Пример. Распишем функцию $z \mapsto w = z^2$ как вещественное отображение. Поскольку $(x+iy)^2 = (x^2-y^2) + 2ixy$, получим

$$(x, y) \mapsto (x^2 - y^2, 2xy).$$

Формулы незначительно отличаются от отображения, задающего параболические координаты. Сетка декартовых координат на плоскости переменной z отображается в сетку парабол на плоскости w.

Пример. Распишем функцию $z \mapsto w = 1/z$ как вещественное отображение. Получим

$$(x, y) \mapsto \frac{(x, -y)}{x^2 + y^2}.$$

Формулы незначительно отличаются от отображения, задающего симметрию относительно окружности. Сетка декартовых координат на плоскости переменной z отображается в сетку окружностей на плоскости w.

Пример. Эллиптические координаты связаны с отображением

$$x + iy = a \operatorname{ch}(\mu + i\nu).$$

Пример. Биполярные координаты связаны с отображением

$$x + iy = ai \operatorname{ctg} \frac{\sigma + i\tau}{2}.$$

Гомеоморфизмы и диффеоморфизмы. Особое место в современной математике занимает общее понятие отображения с определённым свойством, обратное к которому также им обладает. Такие отображения называют **изоморфизмами**, а в конкретных ситуациях — сложными двухкоренными словами с морфизмом во второй части. Примеры нам уже встречались в этом курсе, и особенно интенсивно ранее в этом разделе. Соответствующие термины эпизодически будут полезны во втором семестре.

Определение. Отображение $f\colon X \to Y$ называют **гомеоморфизмом**, реже **непрерывным изоморфизмом** из X в Y, если

- f непрерывное;
- f обратимое;
- обратное отображение $f^{-1}\colon Y \to X$ тоже непрерывное.

Определение. Отображение $f\colon X \to Y$ называют **диффеоморфизмом**, или **гладким изоморфизмом** из X в Y, если

- f гладкое;
- f обратимое;
- обратное отображение $f^{-1}\colon Y \to X$ тоже гладкое.

Непрерывное отображение не имеет «разрывов»: образы достаточно близких точек тоже близки. Непрерывность обратного отображения характеризуется отсутствием «склеек»: если образы точек близки, то и сами точки близки. Поэтому гомеоморфизмы характеризуют обратимые деформации без разрывов и склеек.

Всякий диффеоморфизм является гомеоморфизмом, но диффеоморфизм требует большей гладкости деформации. Скажем, исключено появление угла (излома).

Пример. Отображение отрезка $[-1, 1]$ на себя по правилу $x \mapsto x^3$ является гомеоморфизмом, но не диффеоморфизмом: обратное отображение $x \mapsto \sqrt[3]{x}$ не гладко в нуле.

Обратимые линейные отображения векторных пространств появляются в линейной алгебре, так что тоже на первом курсе, и называются изоморфизмами векторных пространств.

Главный смысл общего понятия изоморфизма (от греческого «сохранение формы», а на деле скорее структуры) в том, что различия между изоморфными объектами, сколь велики бы они ни казались на первый взгляд, оказываются математически несущественными.

10.4. Теорема о неявной функции

лекция 1
09.02.17

Наводящие соображения. Простейшими уравнениями и системами являются линейные. Соответствующие им геометрические фигуры (размерности 1 и 2) — это прямые и плоскости. Напомним их общие уравнения:

$$\text{прямая на плоскости:}\quad Ax + By + C = 0;$$

$$\text{плоскость в пространстве:}\quad Ax + By + Cz + D = 0;$$

$$\text{прямая в пространстве:}\begin{cases} A_1x + B_1y + C_1z + D_1 = 0,\\ A_2x + B_2y + C_2z + D_2 = 0.\end{cases}$$

Нелинейные уравнения задают линии и поверхности:

$$\begin{aligned}\text{линия на плоскости:}\quad & F(x,y)=0;\\ \text{поверхность в пространстве:}\quad & F(x,y,z)=0;\\ \text{линия в пространстве:}\quad & \begin{cases}F_1(x,y,z)=0,\\ F_2(x,y,z)=0.\end{cases}\end{aligned}$$

Все функции считаем гладкими.

Примеры.

$$\begin{aligned}\text{окружность}: F(x,y)&=x^2+y^2-1;\\ \text{гипербола}: F(x,y)&=x^2-y^2-1;\\ \text{крест}: F(x,y)&=x^2-y^2;\\ \text{сфера}: F(x,y,z)&=x^2+y^2+z^2-1.\end{aligned}$$

Основной вопрос сейчас: когда можно выразить какую-то переменную как явную функцию других?

В линейном случае легко усмотреть условие на коэффициенты уравнения. Так, z выражается из уравнения плоскости, если $C \neq 0$. Для прямой в пространстве условие, что y и z выражаются через x, даётся в аналитической геометрии:

$$\det\begin{bmatrix}B_1 & C_1\\ B_2 & C_2\end{bmatrix}\neq 0.$$

Нелинейные примеры показывают, что нужно работать **локально**, то есть в окрестности какой-либо точки фигуры. Гладкость фигуры по определению означает, что вблизи каждой своей точки фигура мало отличается от касательной линии или плоскости:

$$F(x,y)-F(x_0,y_0)\approx A(x-x_0)+B(y-y_0),$$

где $A=\frac{\partial F}{\partial x}(x_0,y_0)$ и $B=\frac{\partial F}{\partial y}(x_0,y_0)$. Поэтому искомое условие получается таким же, как для касательной линии или плоскости.

Одна функция двух переменных. Возьмём открытую плоскую область (множество) $D\subset\mathbb{R}^2$ и гладкую функцию $F\colon D\to\mathbb{R}$. Предположим, что значение F в некоторой точке $p\in D$ равно 0.

Определение. Будем говорить, что в окрестности точки p уравнение $F(x,y)=0$ **определяет** y как функцию $y(x)$, если существует такой прямоугольник $P\times Q\subset D$ с точкой p внутри, что для каждого $x\in P$ найдётся единственное $y\in Q$ с условием $F(x,y)=0$. **рисунок**

То есть, в некоторой окрестности имеем единственным образом $y = y(x)$, при котором $F(x, y(x)) = 0$ выполнено тождественно. Глядя на примеры, можно догадаться, какое условие нужно наложить.

Теорема. *Если* $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$*, то в окрестности точки* $p = (x_0, y_0)$ *уравнение* $F(x, y) = 0$ *определяет* $y = y(x)$ *как гладкую функцию.*

Доказательство. Искомый прямоугольник $P \times Q$ получается серией шагов, на каждом из которых желаемое свойство функции требует его уменьшения. В качестве первого приближения возьмём за $P \times Q$ такую прямоугольную окрестность точки p, на которой сохраняется основное условие на $\frac{\partial F}{\partial y}$, и в частности, там $F(x, y)$ монотонна по y при каждом x. При этом, сначала выберем подходящее $\varepsilon > 0$ и положим $Q = (y_0 - \varepsilon, y_0 + \varepsilon)$, а затем подберём P.

рисунок

Затем уменьшим P до $(x_0 - \delta, x_0 + \delta)$, выбирая $\delta = \delta(\varepsilon)$ так, что $F(x, y)$ сохраняет знак на верхней и нижней гранях $P \times Q$. При том, эти знаки противоположны. Поскольку для любого $x \in P$ значения F в точках $(x, y_0 \pm \varepsilon)$ имеют разные знаки, по теореме Больцано — Коши о промежуточных значениях, на отрезке между этими точками найдётся точка $(x, y(x))$, где значение F нулевое. Благодаря монотонности по y, она единственная.

гиперссылка

Условие непрерывности

$$\forall \varepsilon > 0 \ \exists \delta > 0 : |x - x_0| < \delta \Rightarrow |y(x) - y_0| < \varepsilon$$

функции $y(x)$ в точке x_0 выполнено по построению. Однако вместо $p = (x_0, y_0)$ можно начинать это построение в любой точке $(x, y(x))$ внутри $P \times Q$, поэтому $y(x)$ непрерывна внутри P.

Пользуясь гладкостью функции F, распишем её приращение между двумя близкими точками линии $F(x, y) = 0$ внутри $P \times Q$:

$$0 = F(x + \Delta x, y + \Delta y) - F(x, y) = (A + \alpha)\Delta x + (B + \beta)\Delta y,$$

где α, $\beta = o(1)$. Отсюда при $\Delta x \to 0$ получаем

$$\frac{\Delta y}{\Delta x} = -\frac{A + \alpha}{B + \beta} \to -\frac{A}{B} = -\frac{F_x(x, y)}{F_y(x, y)} = \frac{dy}{dx}(x).$$

Получилось, что функция $y'(x) = -F_x/F_y$ непрерывна на окрестности точки x_0, а значит, $y(x)$ на ней гладкая. □

Одна функция трёх переменных. Возьмём теперь открытую область $D \subset \mathbb{R}^3$ и гладкую функцию $F\colon D \to \mathbb{R}$. Предположим, что значение F в некоторой точке $p \in D$ равно 0.

Определение. Будем говорить, что в окрестности точки p уравнение $F(x,y,z)=0$ **определяет** z как функцию $z(x,y)$, если существует такой параллелепипед $P\times Q\subset D$ с точкой p внутри, что для каждой точки $(x,y)\in P$ найдётся единственное $z\in Q$ с условием $F(x,y,z)=0$. рисунок

То есть, в некоторой окрестности имеем единственным образом $z=z(x,y)$, при котором $F(x,y,z(x,y))=0$ выполнено тождественно.

Теорема. *Если* $\frac{\partial F}{\partial z}(p)\neq 0$, *то в окрестности точки* p *уравнение* $F(x,y,z)=0$ *определяет* $z=z(x,y)$ *как гладкую функцию.*

Доказательство. Явная функция строится точно так же, как и выше, меняется лишь P: вместо интервала на оси Ox теперь это прямоугольник в плоскости Oxy. Непрерывность снова не требует особых усилий. рисунок

Гладкость $z(x,y)$ вблизи точки (x_0,y_0) тоже получается аналогично:

$$0=\Delta F=(A+\alpha)\Delta x+(B+\beta)\Delta y+(C+\gamma)\Delta z$$

влечёт при $\Delta y=0$, $\Delta x\to 0$

$$\frac{\Delta z}{\Delta x}\to -\frac{A}{C}=-\frac{F_x(x,y,z)}{F_z(x,y,z)}=\frac{\partial z}{\partial x}(x,y),$$

и соответственно при $\Delta x=0$, $\Delta y\to 0$

$$\frac{\Delta z}{\Delta y}\to -\frac{B}{C}=-\frac{F_y(x,y,z)}{F_z(x,y,z)}=\frac{\partial z}{\partial y}(x,y).$$

Получена непрерывность частных производных $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ на окрестности точки (x_0,y_0), а это обеспечивает на ней гладкость функции $z(x,y)$. □

Две функции трёх переменных. Геометрическая картина тут — пересечение двух поверхностей. Это совсем новое явление, а потому рассуждения в этом случае включают новые элементы. Полезно начать с примера.

Пример. Система уравнений

$$\begin{cases} F_1=xz-y=0,\\ F_2=x-yz=0\end{cases}$$

вблизи точки $p=(1,1,1)$ определяет гладкие функции $y(x)$ и $z(x)$. Точный смысл этих слов определён ниже, а сейчас мы просто найдём эти функции в два шага. Для краткости обозначим

$$F_{i,y}=\frac{\partial F_i}{\partial y},\quad F_{i,z}=\frac{\partial F_i}{\partial z}$$

и аналогично для функций G_i, которые появятся ниже.

На первом шаге выразим $z = z(x,y)$ из уравнения $F_2 = 0$, отмечая при этом, что $F_{2,z}(p) = -1 \neq 0$. Значит, $z = x/y$. Для подготовки второго шага подставим это выражение в функцию $F_1(x,y,z)$ и получим новую функцию

$$G_1(x,y) = F_1(x,y,z(x,y)) = F_1(x,y,x/y) = x^2/y - y.$$

Также обратим внимание на то, что аналогично

$$G_2(x,y) = F_2(x,y,z(x,y)) = F_2(x,y,x/y) = 0.$$

На втором шаге выразим $y = y(x)$ из уравнения $G_1 = 0$. Отметим при этом, что $G_{1,y}(p) = -2 \neq 0$. Имеем $y = \pm x$, причём вблизи точки p нужно выбрать $y = x$. Наконец, это выражение нужно подставить в $z(x,y)$. В итоге видим, что линию пересечения заданных поверхностей вблизи точки p можно параметризовать как $(x,y,z) = (x,x,1)$.

Обе поверхности в этом примере являются гиперболическими параболоидами, а их пересечение состоит из трёх прямых линий.

Возьмём теперь открытую область $D \subset \mathbb{R}^3$ и две гладкие функции $F_1,\ F_2\colon D \to \mathbb{R}$. Предположим, что значения и F_1, и F_2 в некоторой точке $p \in D$ равны 0.

Определение. Будем говорить, что в окрестности точки p уравнения

$$F_1(x,y,z) = 0, \quad F_2(x,y,z) = 0$$

определяют y и z как функции от x, если существует такой параллелепипед $P \times Q \subset D$ с точкой p внутри, что для каждого $x \in P$ найдётся единственная пара $(y,z) \in Q$ с условием $F(x,y,z) = 0$.

рисунок

То есть, в некоторой окрестности имеем единственным образом $y = y(x)$ и $z = z(x)$, при которых $F_1(x,y(x),z(x)) = 0$ и $F_2(x,y(x),z(x)) = 0$ выполнены тождественно.

Линейная алгебра и аналитическая геометрия опять подсказывают: в этой ситуации условие на один коэффициент уравнения заменяется на условие на соответствующий определитель из коэффициентов. В нашем же нелинейном случае условие на одну частную производную превращается в условие на якобиан.

лекция 2 13.02.17

Теорема. *Если*

$$\frac{\partial(F_1,F_2)}{\partial(y,z)} \neq 0$$

в точке p, то в её окрестности два уравнения $F_i(x,y,z) = 0$ определяют $y = y(x)$ и $z = z(x)$ как гладкие функции.

Доказательство. Если определитель матрицы Якоби в точке p ненулевой, то хотя бы один её элемент там ненулевой, и можно считать, что таков правый нижний.

Построение проводится в два шага: сперва строится функция $z(x, y)$, затем функция $y(x)$. На каждом шагу обеспечивается гладкость полученной явной функции на некоторой окрестности.

Первый шаг осуществим по предыдущей теореме, ибо $F_{2,z}(p) \neq 0$, а законность второго надо проверять. Для этого обозначим

$$G_i(x, y) = F_i(x, y, z(x, y))$$

и убедимся, что для G_1 выполнены все условия теоремы о неявной функции двух переменных. Обе функции G_i гладкие как композиции гладких F_i и z. Остаётся проверить лишь условие на $\frac{\partial G_1}{\partial y}(p)$.

Преобразуем исходную матрицу Якоби, сохраняя значение определителя. Умножим второй столбец на $z_y = \frac{\partial z}{\partial y}$ и прибавим к первому:

$$\begin{vmatrix} F_{1,y} & F_{1,z} \\ F_{2,y} & F_{2,z} \end{vmatrix} = \begin{vmatrix} F_{1,y} + F_{1,z} z_y & F_{1,z} \\ F_{2,y} + F_{2,z} z_y & F_{2,z} \end{vmatrix} = \begin{vmatrix} G_{1,y} & F_{1,z} \\ G_{2,y} & F_{2,z} \end{vmatrix}.$$

В первом столбце получились в точности частные производные $G_{i,y}$, согласно правилу дифференцирования композиции. Однако по построению $z(x, y)$ в некоторой окрестности $G_2(x, y) = 0$ тождественно, поэтому якобиан упрощается до $G_{1,y} F_{2,z}$. Значит, $G_{1,y} \neq 0$. □ пояснить?

Общая теорема о неявной функции. Возьмём теперь открытую область $D \subset \mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^m$ и m гладких функций $F_1, \ldots, F_m \colon D \to \mathbb{R}$. Предположим, что значения всех F_i в некоторой точке $p \in D$ равны 0.

Теорема (Дини)**.** *Если* Dini 1876

$$\frac{\partial(F_1, \ldots, F_m)}{\partial(y_1, \ldots, y_m)} \neq 0$$

в точке p, то в её окрестности m уравнений

$$F_i(x_1, \ldots, x_k, y_1, \ldots, y_m) = 0$$

определяют m гладких функций $y_i = y_i(x_1, \ldots, x_k)$.

Доказательство. Рассуждение аналогично проведённому в предыдущем случае $k = 1$, $m = 2$. Логически оно устроено как нисходящая индукция: выразив одну переменную из одного уравнения, мы проверяем, что для оставшихся соблюдены все условия этой же теоремы, но с меньшим на одну количеством переменных y_i и уравнений. Окрестность точки p, на которой всё происходит, будет уменьшаться на каждом таком шаге.

Покажем ещё раз, как работает матричная выкладка для $m = 3$. Будущие независимые переменные x_j не участвуют в ней.

Считаем ненулевым правый нижний элемент $F_{33} = \frac{\partial F_3}{\partial y_3}$ матрицы Якоби исходных неявных функций F_i по будущим зависимым переменным y_j. Выразим y_3 из связи $F_3 = 0$. Полагая

$$G_i(x_1, \ldots, x_k, y_1, y_2) = F_i(x_1, \ldots, x_k, y_1, y_2, y_3(x_1, \ldots, x_k, y_1, y_2))$$

для всех $1 \leqslant i \leqslant 3 = m$, получим $G_3 = 0$ тождественно. При тех же i и при $1 \leqslant j \leqslant 2 = m - 1$ получим

$$\frac{\partial G_i}{\partial y_j} = \frac{\partial F_i}{\partial y_j} + \frac{\partial F_i}{\partial y_m} \frac{\partial y_m}{\partial y_j}$$

по правилу дифференцирования композиции. Обозначим частные производные так, чтобы это равенство стало

$$G_{ij} = F_{ij} + F_{im} Y_{mj}.$$

Умножим последний столбец матрицы Якоби на Y_{mj} и прибавим к остальным её столбцам. Увидим частные производные G_{ij}:

$$\begin{vmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{vmatrix} = \begin{vmatrix} F_{11} + F_{13}Y_{31} & F_{12} + F_{13}Y_{32} & F_{13} \\ F_{21} + F_{23}Y_{31} & F_{22} + F_{23}Y_{32} & F_{23} \\ F_{31} + F_{33}Y_{31} & F_{32} + F_{33}Y_{32} & F_{33} \end{vmatrix} = \left|\begin{array}{cc|c} G_{11} & G_{12} & F_{13} \\ G_{21} & G_{22} & F_{23} \\ \hline G_{31} & G_{32} & F_{33} \end{array}\right|.$$

Поскольку $G_3 = 0$ тождественно, видим тут $G_{31} = G_{32} = 0$, так что получена блочная верхнетреугольная матрица. Её определитель вычисляется перемножением определителей диагональных блоков:

$$\begin{vmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{vmatrix} = \left|\begin{array}{cc|c} G_{11} & G_{12} & * \\ G_{21} & G_{22} & * \\ \hline 0 & 0 & F_{33} \end{array}\right| = \begin{vmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{vmatrix} \cdot F_{33}.$$

Итак, если исходный якобиан ненулевой в точке p, то то же самое можно сказать про меньшего порядка якобиан для G_i. □

10.5. Теорема об обратном отображении

Постановка вопроса. Для примера, на прямоугольнике

$$U = \{(\rho, \varphi) \in \mathbb{R}^2 \mid 0 \leqslant \rho \leqslant 1,\ 0 \leqslant \varphi \leqslant \pi\}$$

определим отображение $f\colon U \to \mathbb{R}^2$, $(\rho, \varphi) \mapsto (x, y)$, по правилу

$$x = \rho \cos\varphi, \quad y = \rho \sin\varphi.$$

рисунок Образом $V = f(U)$ является полукруг.

Обратимо ли это гладкое отображение? Нет: в одну точку переходит вся сторона прямоугольника, где $\rho = 0$. Заметим, что именно там равен нулю якобиан

$$\frac{\partial(x, y)}{\partial(\rho, \varphi)} = \rho.$$

Ограничивая f на внутренность U, получим диффеоморфизм на внутренность V.

Перейдём к общей постановке вопроса. Когда из системы гладких функций

$$\begin{cases} y_1 = f_1(x_1, x_2), \\ y_2 = f_2(x_1, x_2) \end{cases}$$

возможно выразить все x_j как гладкие функции от y_i? Далее, что, если и тех, и других переменных не по две, а по n штук? На каком множестве ответ годится? Окончательно приходим к такому вопросу: для двух открытых областей U, $V \subset \mathbb{R}^n$ и гладкого отображения $f\colon U \to V$, когда и где существует гладкое обратное отображение?

Локальная обратимость. Теорема об умножении якобианов при композиции отображений указывает, какое условие необходимо: композиция $f^{-1} \circ f$ есть тождественное отображение, и его якобиан равен 1 всюду на U; значит, якобиан f не должен быть нулевым. Локально (вблизи выбранной точки) это условие оказывается достаточным для ответа на поставленный вопрос.

гиперссылка

Теорема. *Гладкое отображение является локальным диффеоморфизмом вблизи каждой точки, где его якобиан отличен от нуля.*

Эта формулировка подразумевает открытость областей U и V и равенство их размерностей: матрица Якоби должна быть квадратной. Заключение теоремы для точки $\tilde{x} \in U$ означает, что найдутся такие открытые множества $\tilde{U}$ и $\tilde{V}$, удовлетворяющие

рисунок

$$\tilde{x} \in \tilde{U} \subset U, \quad f(\tilde{x}) \in \tilde{V} \subset V,$$

что гладкое отображение $f\colon \tilde{U} \to \tilde{V}$ взаимно однозначно, а обратное ему отображение $f^{-1}\colon \tilde{V} \to \tilde{U}$ также гладко.

Доказательство. Постараемся применить теорему о неявной функции, ограничиваясь разбором случая $n = 2$. Общий случай принципиально не отличается. Перепишем систему в виде

$$\begin{cases} F_1(x_1, x_2, y_1, y_2) = f_1(x_1, x_2) - y_1 = 0, \\ F_2(x_1, x_2, y_1, y_2) = f_2(x_1, x_2) - y_2 = 0. \end{cases}$$

Если точка $p \in U \times V$ является её решением и

$$\frac{\partial(F_1, F_2)}{\partial(x_1, x_2)}(p) = \frac{\partial(f_1, f_2)}{\partial(x_1, x_2)}(p) \neq 0,$$

то найдётся такой открытый параллелепипед $P \times Q \subset U \times V$, содержащий p, в котором x_1 и x_2 находятся единственным образом как гладкие функции от (y_1, y_2). Это и есть координатные функции гладкого обратного отображения $g = f^{-1}\colon Q \to P$.

Однако образ $g(Q)$ не обязан покрывать весь прямоугольник P и рисунок равняться полному прообразу $f^{-1}(Q)$. Взаимной однозначности добиваемся, ограничивая f на пересечение $\tilde{U} = P \cap f^{-1}(Q)$, которое также открыто. При этом, $\tilde{V} = Q$. Теперь

$$x \in \tilde{U} \Rightarrow f(x) \in \tilde{V}, \quad y \in \tilde{V} \Rightarrow f^{-1}(y) \in \tilde{U},$$

так что $\tilde{U}$ и $\tilde{V}$ удовлетворяют заключению теоремы. □

Неудача глобализации. Можно ли распространить эту локальную теорему до глобальной? Это значит: можно ли утверждать, что если якобиан отображения f отличен от нуля на всей области определения, то f является диффеоморфизмом всей области? Нет, нельзя, и причина в том, что невозможно гарантировать взаимную однозначность.

Пример. Комплексная функция $f(z) = z^2$ в вещественном представлении даёт отображение

$$f(x, y) = (x^2 - y^2, 2xy),$$

якобиан которого равен $4(x^2 + y^2)$. Однако окрестность половины единичной окружности отображается не взаимно однозначно на окрест- рисунок ность окружности, ибо $f(1, 0) = f(-1, 0) = (1, 0)$.

10.6. Дифференцирование неявных функций

Первые производные и дифференциалы. В случае одной неявной функции, $F(x, y) = 0$, формула

$$dy/dx = -F_x/F_y$$

получена прямо в доказательстве простейшего варианта теоремы о неявной функции, а в следующем варианте для $F(x_1, x_2, y) = 0$ с другими буквами мы нашли

$$\partial y/\partial x_j = -F_{x_j}/F_y.$$

Эта же формула аналогично получается для одной неявной функции нескольких переменных, $F(x_1, \ldots, x_k, y) = 0$. Важно в частности то, что неявную функцию можно дифференцировать, не зная её явного вида, который часто и вовсе недоступен.

Обобщим теперь эту формулу двумя способами на ситуацию общей теоремы о неявной функции: систему $F(x, y) = 0$, где

$$F = [F_1, \dots, F_m]^\top, \quad x = [x_1, \dots, x_k]^\top, \quad y = [y_1, \dots, y_m]^\top.$$

Все векторы удобно считать столбцами. Обобщение устанавливается аккуратным формальным переходом к матрицам.

Теорема. *Матрица Якоби неявно заданных функций y_i по независимым переменным x_j равна*

$$\frac{Dy}{Dx} = -\left[\frac{DF}{Dy}\right]^{-1} \cdot \frac{DF}{Dx}.$$

Доказательство. Как и в простейшем случае, только теперь векторно, запишем дифференциалы связей и затем выразим дифференциалы зависимых переменных через дифференциалы независимых. Искомая формула следует из того, что

$$dF = \frac{DF}{Dx}\,dx + \frac{DF}{Dy}\,dy = \frac{DF}{Dx}\,dx + \frac{DF}{Dy} \cdot \frac{Dy}{Dx}\,dx$$

должно равняться нулю при любом значении dx. □

Конкретные частные производные $\partial y_i/\partial x_j$ можно вычислить чуть иначе, избегая обращения матриц и пользуясь другим обобщением. Разберём случай двух функций от двух независимых переменных: из уравнений связи лекция 3 16.02.17

$$\begin{cases} F_1(x_1, x_2, y_1, y_2) = 0, \\ F_2(x_1, x_2, y_1, y_2) = 0 \end{cases}$$

выражаются гладкие функции $y_1(x_1, x_2)$ и $y_2(x_1, x_2)$. Матричная основа всех действий позволяет затем легко получить обобщение на произвольные количества переменных и функций.

Теорема. *Частные производные $\partial y_i/\partial x_j$ находятся как отношения якобианов:*

$$\frac{\partial y_1}{\partial x_j} = -\frac{\partial(F_1, F_2)}{\partial(x_j, y_2)} \Big/ \frac{\partial(F_1, F_2)}{\partial(y_1, y_2)}, \qquad \frac{\partial y_2}{\partial x_j} = -\frac{\partial(F_1, F_2)}{\partial(y_1, x_j)} \Big/ \frac{\partial(F_1, F_2)}{\partial(y_1, y_2)}.$$

Не забываем минус перед дробью! Более существенное замечание приведём после вывода формулы.

Доказательство. Обозначим известные частные производные через

$$A_{ij} = \frac{\partial F_i}{\partial x_j}, \quad B_{ij} = \frac{\partial F_i}{\partial y_j}.$$

Дифференцируя равенства $F_i = 0$, получим $dF_i = 0$. Распишем дифференциалы через частные производные:

$$\begin{cases} dF_1 = A_{11}\,dx_1 + A_{12}\,dx_2 + B_{11}\,dy_1 + B_{12}\,dy_2 = 0, \\ dF_2 = A_{21}\,dx_1 + A_{22}\,dx_2 + B_{21}\,dy_1 + B_{22}\,dy_2 = 0. \end{cases}$$

Из этой системы линейных уравнений можно выразить неизвестные dy_1 и dy_2, ибо, по предположению теоремы о неявной функции,

$$J = \begin{vmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{vmatrix} \neq 0$$

в рассматриваемой точке. Сперва перенесём в правую часть «лишние» слагаемые и обозначим их через C_i:

$$\begin{cases} B_{11}\,dy_1 + B_{12}\,dy_2 = C_1 = -A_{11}\,dx_1 - A_{12}\,dx_2, \\ B_{21}\,dy_1 + B_{22}\,dy_2 = C_2 = -A_{21}\,dx_1 - A_{22}\,dx_2. \end{cases}$$

Формула Крамера даёт единственное решение

$$dy_1 = \frac{1}{J}\begin{vmatrix} C_1 & B_{12} \\ C_2 & B_{22} \end{vmatrix}, \quad dy_2 = \frac{1}{J}\begin{vmatrix} B_{11} & C_1 \\ B_{21} & C_2 \end{vmatrix}.$$

Подставляя выражения для C_i и пользуясь линейностью определителя по первому столбцу, получаем

$$dy_1 = -\frac{1}{J}\begin{vmatrix} A_{11} & B_{12} \\ A_{21} & B_{22} \end{vmatrix} dx_1 - \frac{1}{J}\begin{vmatrix} A_{12} & B_{12} \\ A_{22} & B_{22} \end{vmatrix} dx_2,$$

и аналогично для dy_2. Сравнивая это с

$$dy_i = \frac{\partial y_i}{\partial x_1}\,dx_1 + \frac{\partial y_i}{\partial x_2}\,dx_2,$$

получим искомые частные производные. □

Новая тонкость в частных производных. Выведенная формула подразумевает, что независимыми переменными считаются x_j, а зависимыми y_i. Напоминанием об этом служит якобиан в знаменателе.
12.02.16 В приложениях обозначения переменных не обязаны отражать такое разделение, а кроме того, их роли могут меняться; такая ситуация ти-

пична в термодинамике. Так, в случае двух связей $F_i(x,y,z,t)=0$ есть два разных значения $\partial z/\partial t$:

$$\left(\frac{\partial z}{\partial t}\right)_{x=\mathrm{const}} = -\frac{\partial(F_1,F_2)}{\partial(y,t)}\bigg/\frac{\partial(F_1,F_2)}{\partial(y,z)},$$
$$\left(\frac{\partial z}{\partial t}\right)_{y=\mathrm{const}} = -\frac{\partial(F_1,F_2)}{\partial(x,t)}\bigg/\frac{\partial(F_1,F_2)}{\partial(x,z)}.$$

Здесь выделенный внешний индекс левой части указывает зафиксированную независимую переменную.

Пример. Из системы двух связей 23.03.19

$$\begin{cases} F = xy + zt, \\ G = x + y + z + t \end{cases}$$

найдём производные z_x, фиксируя y или t. Для первой вычислим

$$\frac{\partial(F,G)}{\partial(t,x)} = \begin{vmatrix} F_t & F_x \\ G_t & G_x \end{vmatrix} = \begin{vmatrix} z & y \\ 1 & 1 \end{vmatrix} = z - y,$$
$$\frac{\partial(F,G)}{\partial(t,z)} = \begin{vmatrix} F_t & F_z \\ G_t & G_z \end{vmatrix} = \begin{vmatrix} z & t \\ 1 & 1 \end{vmatrix} = z - t,$$

и аналогично для второй, с обменом $y \leftrightarrow t$ и $x \leftrightarrow z$. Получим

$$\left(\frac{\partial z}{\partial x}\right)_{y=\mathrm{const}} = \frac{z-y}{t-z}, \qquad \left(\frac{\partial z}{\partial x}\right)_{t=\mathrm{const}} = \frac{x-y}{t-x}.$$

Можно убедиться, что ответы различны, подставив подходящую точку, например, $(x,y,z,t)=(1,-6,3,2)$. Выражение производных опирается на равенства $dF=0$ и $dG=0$, то есть на постоянство значений связей F и G, но сами эти значения никуда не входят.

Отметим ещё, что простота взятых связей тут позволяет исключить одну из переменных, $y = G - x - z - t$, и обходиться одной связью

$$E = x(G - x - z - t) + zt.$$

Отсюда находим «обычную» частную производную

$$\frac{\partial z}{\partial x} = -\frac{E_x}{E_z} = -\frac{G-2x-z-t}{x-t} = \frac{x-y}{t-x}$$

и обнаруживаем совпадение с найденной выше при $t=\mathrm{const}$.

Вторые производные и дифференциалы. Второй дифференциал функции $F(x,y)$, в предположении независимости её аргументов, равен

$$d^2F = d(dF) = d(F_x\,dx + F_y\,dy) = F_{xx}\,dx^2 + 2F_{xy}\,dx\,dy + F_{yy}\,dy^2.$$

Считая, что связь $F(x,y)=0$ влечёт зависимость y от x, мы получаем добавочное слагаемое, потому что теперь $d^2y \neq 0$:

$$F_{xx}\,dx^2 + 2F_{xy}\,dx\,dy + F_{yy}\,dy^2 + F_y\,d^2y = 0.$$

Чтобы выразить отсюда d^2y, достаточно иметь $F_y \neq 0$, а также подставить уже известное выражение для dy.

Когда из двух связей $F_1(x,y,z)=0$ и $F_2(x,y,z)=0$ мы находим две функции $y(x)$ и $z(x)$, аналогично получаются два уравнения

$$H_i + F_{i,y}\,d^2y + F_{i,z}\,d^2z = 0,$$

где H_i это полный второй дифференциал (гессиан) функции F_i в предположении независимости её аргументов. Из полученной системы по формуле Крамера находим

$$d^2y = -\frac{1}{J}\begin{vmatrix} H_1 & F_{1,z} \\ H_2 & F_{2,z} \end{vmatrix}, \quad d^2z = -\frac{1}{J}\begin{vmatrix} F_{1,y} & H_1 \\ F_{2,y} & H_2 \end{vmatrix}; \quad J = \frac{\partial(F_1,F_2)}{\partial(y,z)}.$$

Подставляя в H_i выражения dy и dz через dx, сводим гессианы к ранее найденным функциям, умноженным на dx^2.

Аналогично находятся вторые дифференциалы и при большем количестве переменных и связей, но растёт порядок определителей.

Высшие производные и дифференциалы. Задачи, требующие оных, возникают намного реже. Метод тот же, главное — не потерять слагаемые. Полные дифференциалы связей включают дополнительные слагаемые уже на втором шаге. Их нужно аккуратно дифференцировать дальше. Получится система уравнений того же вида

$$(\text{известные дифференциалы})_i + F_{i,y}\,d^ny + F_{i,z}\,d^nz = 0,$$

также и решаемая.

Глава 11. МНОГООБРАЗИЯ

11.1. Особенности линий и поверхностей

Три способа задания. Основных способа задания линий и поверхностей три: явный, неявный, параметрический. Явное задание является простым частным случаем как неявного, так и параметрического.

	Явно	Неявно	Параметрически
Линия в $\mathbb{R}^2$	$y = y(x)$	$F(x, y) = 0$	$x = x(t)$, $y = y(t)$
Поверхность в $\mathbb{R}^3$	$z = z(x, y)$	$F(x, y, z) = 0$	$x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$
Линия в $\mathbb{R}^3$	$y = y(x)$, $z = z(x)$	$F_1(x, y, z) = 0$, $F_2(x, y, z) = 0$	$x = x(t)$, $y = y(t)$, $z = z(t)$

Что называть гладкой линией или поверхностью? Считаем, конечно, все функции гладкими; однако уже простые примеры линий ниже показывают, что как при неявном, так и при параметрическом задании мало предполагать лишь это. При явном задании приходится разбивать фигуру на куски с разными явными заданиями, как мы уже не раз отмечали на примере окружности.

Примеры особенностей плоских линий. Чтобы получить из неявного задания $F(x, y) = 0$ плоской линии явное задание $y = y(x)$ вблизи точки p на линии, достаточно иметь $F_y(p) \neq 0$. При условии $F_x(p) \neq 0$ аналогично выразимо $x = x(y)$. Следовательно, выполнение условия $\nabla F(p) \neq \mathbf{0}$ гарантирует возможность явного задания.

16.02.16

лекция 4
20.02.17

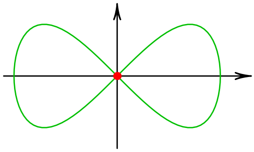
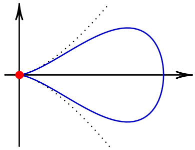
Примеры. Параметризация

$$x = \cos t, \quad y = \cos t \sin t$$

задаёт «восьмёрку», знакомую многим по работе с осциллографом. Здесь естественно считать t углом. Вся фигура проходится однократно на участке $0 \leqslant t < 2\pi$, замыкаясь при достижении $t = 2\pi$. Нехитрая тригонометрия позволяет составить также неявное уравнение

$$x^4 - x^2 + y^2 = 0.$$

Линия имеет особенность в начале координат: она не является графиком функции ни в какой окрестности этой точки.

Нарисуем также линию, заданную немного изменённым уравнением

$$x^4 - x^3 + y^2 = 0.$$

Точек с отрицательными абсциссами нет, ибо в этом случае $x^4 - x^3 > 0$. Вблизи начала координат слагаемое x^4 мало, и линия имеет особенность (точку возврата) подобно полукубической параболе $y^2 = x^3$, на рисунке изображённой пунктиром.

В обоих примерах условие $\nabla F \neq \mathbf{0}$ нарушается в начале координат.

Итак, при неявном задании могут быть точки, вблизи которых явного задания нет. Подобное происходит и при параметрическом задании, но тут ещё возможны нюансы с кратными точками и даже дугами.

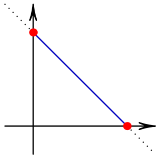
Удобно считать параметр t временем и обозначать производные по нему точками сверху. Чтобы получить из параметризации $x = x(t)$, $y = y(t)$ явное задание $y = y(x)$ вблизи точки p на линии, нужно выразить $t = t(x)$. Для этого по теореме об обратной функции одной переменной достаточно иметь $\dot{x}(p) \neq 0$. При условии $\dot{y}(p) \neq 0$ аналогично $t = t(y)$ даёт $x = x(t(y))$, то есть в итоге $x = x(y)$. В отличие от ситуации с неявным заданием, выполнение условия $(\dot{x}, \dot{y})(p) \neq \mathbf{0}$ не гарантирует возможность явного задания: в первом примере вектор скорости оказывается отличен от нуля в каждой точке линии, однако начало координат проходится дважды и это даёт особенность. Во втором примере скорость в начале координат оказывается нулевой и это приводит к особенности.

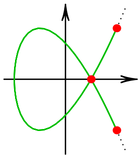
Когда линия параметризована гладкими функциями и речь идёт фактически о траектории, может оказаться полезной (нестандартная) точка зрения, выставляющая все особенности фигуры как иллюзорные. Для этого достаточно вместо траектории $(x(t), y(t))$ в $\mathbb{R}^2$ думать о явно заданной линии $(t, x(t), y(t))$ в $\mathbb{R}^3$. Это исключает как особенности ввиду повторного прохождения, так и особенности ввиду остановок. Точки возврата и излома превращаются в обыкновенные точки, в которых скорость направлена вдоль оси t. Особые точки исходной фигуры возникают при проектировании, забывающем о параметре.

Примеры. Параметризация

$$x = \cos^2 t, \quad y = \sin^2 t$$

задаёт отрезок. Когда t имеет смысл угла, при полном обороте этот отрезок проходится четыре раза, а по краям точки возврата. Неявное же уравнение $x + y = 1$ отвечает не отрезку, а всей прямой.





Вот чуть более сложный вариант того же явления. Параметризация

$$x = \cos 2t, \quad y = \cos 3t$$

задаёт «рыбку», другую простую фигуру с экрана осциллографа. При полном обороте параметра линия проходится дважды. Из тригонометрии можно получить неявное уравнение $2y^2 = (x+1)(2x-1)^2$, но оно, мало того, что некрасивое, отвечает рыбке с бесконечным хвостом.

Примеры. Уравнение $x^2 + y^2 = 0$ задаёт единственную точку.

Уравнение $(x+1)(x^2+y^2) = 0$ задаёт фигуру, состоящую из прямой и точки. Точка является особой.

Эффекты последних двух пар примеров возникают в вещественной геометрии, но не в комплексной.

Нарушение отмеченных условий на градиент связи при неявном задании и на вектор скорости при параметрическом не всегда приводит к особенностям.

Пример. Уравнение $x^2 = 0$ на плоскости Oxy задаёт ту же прямую линию, что и уравнение $x = 0$, но градиент связи $F(x, y) = x^2$ на ней тождественно равен нулю.

Пример. Параметризация $(x, y) = (t^3, t^3)$ задаёт ту же прямую линию, что и уравнение $x = y$, но скорость зануляется при $t = 0$.

Особенности пересечения. Линии и поверхности без особых точек мы далее назовём гладкими, а сейчас отметим, что пересечение двух гладких поверхностей не обязано быть гладкой линией.

Примеры. Рассмотрим сперва сечения явно заданной поверхности горизонтальными плоскостями $z = c$. Они называются **линиями уровня** функции $z(x, y)$ и широко используются для её изображения; например, их часто рисуют на географических картах.

рисунок

Уравнение $z = x^2 + y^2$ задаёт параболоид вращения. При $c > 0$ в сечении окружность, но при $c = 0$ там всего одна точка, а при $c < 0$ пересечение пустое.

Уравнение $z = x^2 - y^2$ задаёт гиперболический параболоид. При $c \neq 0$ его типичные линии уровня являются гиперболами, но при $c = 0$ это пара прямых, то есть линия с особой точкой — самопересечением.

Итак, некоторым **критическим** значениям параметра соответствуют пересечения с особенностями. Более того, при переходе параметра через критическое значение вид пересечения качественно меняется.

Пример. В начале текущей главы упомянуто пересечение двух неявно заданных гиперболических параболоидов:

$$\begin{cases} F_1 = xz - y = 0, \\ F_2 = x - yz = 0. \end{cases}$$

Эта линия состоит из трёх прямых: $\{(0, 0, z)\}$, $\{(x, x, 1)\}$ и $\{(x, -x, -1)\}$. Точки $(0, 0, \pm 1)$ лежат на двух прямых и поэтому являются особыми.

Найдём матрицу Якоби функций связи:

$$\frac{\partial(F_1, F_2)}{\partial(x, y, z)} = \begin{bmatrix} z & -1 & x \\ 1 & -z & -y \end{bmatrix}.$$

Подставляя точки на линии, видим, что ранг этой матрицы равен 2 везде на линии, кроме двух особых точек, где ранг падает до 1. Теорема о неявной функции работает лишь в обыкновенных точках. Для дальнейших приложений полезно осознать падение ранга матрицы Якоби в особой точке как появление линейной зависимости между градиентами функций связи.

Примеры особенностей поверхностей. Они весьма разнообразны, и мы отметим тут лишь две простейшие и одну любопытную.

Пример (конус). Уравнение

$$x^2 + y^2 - z^2 = 0$$

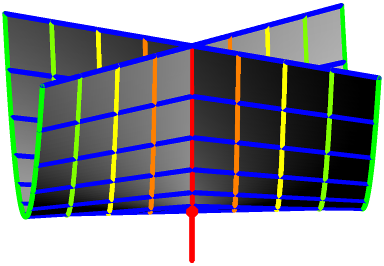
задаёт прямой круговой конус, имеющий особую точку в начале координат: ни в какой её окрестности конус не может быть графиком функции. При этом в вершине ожидаемо зануляется градиент связи.

Пример. Уравнение $xy = 0$ задаёт поверхность, состоящую из двух координатных плоскостей. Ось Oz, по которой они пересекаются, это линия особых точек поверхности — точек самопересечения.

Пример (зонтик Уитни). Фигура с уравнением

$$x^2 - y^2 z = 0$$

выглядит довольно странно. По положительной части оси Oz поверхность имеет самопересечение, а по отрицательной части идёт линия! Особую точку в начале координат могли бы называть точкой пинча (pinch point). Грубую модель можно свернуть из листа бумаги.



Пример (коноид Плюккера). Для функций двух и более переменных наличие частных производных и наличие дифференциала — разные свойства. Мы увидели это на классическом примере, а теперь обратимся к нему же с другой целью. От явной функции

$$z(x, y) = \frac{2xy}{x^2 + y^2}$$

с серьёзным разрывом в начале координат перейдём к уравнению

$$(x^2 + y^2)z - 2xy = 0.$$

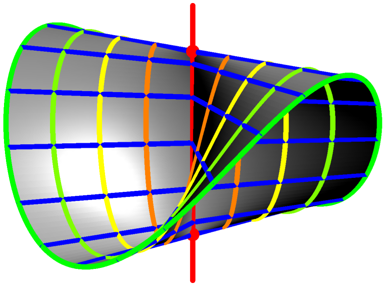
Градиент этой связи равен нулю вдоль оси Oz. Все точки оси являются особыми точками фигуры, причём их три типа, как и в предыдущем примере. При $|z| < 1$ это точки самопересечения поверхности; вблизи них коноид похож на пару плоскостей. Линия самопересечения коноида (отрезок оси) оканчивается в точках пинча с $|z| = 1$. При $|z| > 1$ получаются два луча — «ручки», не входящие в коноид.

Коноид Плюккера без ручек можно параметризовать, полагая

$$x = v \cos u, \quad y = v \sin u, \quad z = \sin 2u.$$

Фиксируя произвольно u, получим прямую, ортогональную оси коноида. При изменении u прямая движется и заметает коноид. Движение состоит из равномерного вращения вокруг оси и гармонического колебания вдоль оси. Благодаря такому сочетанию, коноид Плюккера применяется в машиностроении.

ещё?



Следует отметить, что в точках пинча происходит «остановка»: в матрице Якоби отображения $(u, v) \mapsto (x, y, z)$ зануляется весь столбец производных по u, так что падает ранг.

Ребро возврата. Другим типом особенности обладает целый класс поверхностей. Возьмём в $\mathbb{R}^3$ гладкую параметризованную линию $\mathbf{r}(u)$ ненулевой кривизны, где $\mathbf{r} = (x, y, z)$ для краткости. Касательная к ней в точке с параметром u задаётся уравнением

$$\mathbf{s}(t, u) = \mathbf{r}(u) + t\,\mathbf{r}'(u).$$

При изменении u касательная заметает поверхность S с таким же параметрическим заданием. Её называют тангенциальной развёртывающейся поверхностью исходной линии. Случай плоской линии проще понять по рисунку, чем случай скрученной, но интересен последний. Исходная скрученная линия L состоит из особых точек поверхности, называемых ребром возврата, потому что сечение S нормальной плоскостью линии имеет точку возврата.

рисунок

Найдём матрицу Якоби отображения $\mathbf{s}(t, u)$. В ней два столбца

$$\frac{\partial \mathbf{s}}{\partial t} = \mathbf{r}'(u), \qquad \frac{\partial \mathbf{s}}{\partial u} = \mathbf{r}'(u) + t\,\mathbf{r}''(u),$$

линейно независимые при $t \neq 0$, ибо по условию ненулевой кривизны $\mathbf{r}' \times \mathbf{r}'' \neq \mathbf{0}$. Значит, вблизи каждой точки поверхности вне L ранг матрицы Якоби равен 2 и хотя бы один из максимальных миноров невырожден. По теореме о неявной функции можно выразить (t, u) как функции соответствующей пары декартовых координат и подставить в третью. Так мы получим явное задание кусочка поверхности.

Наоборот, в каждой точке на L ранг матрицы Якоби равен 1, и его падение указывает на возможное появление особенности.

Ласточкин хвост. Познакомимся теперь с одной замечательной поверхностью, заслужившей редкую почесть увлечь великого художника и стать его последней моделью.

22.03.16

Источником служит вопрос, имеющий приложения в разных областях: расположение корней полинома на комплексной плоскости. Коэффициенты полинома вещественны, а степень мы выберем так, чтобы задача была максимально интересной — и не слишком тривиальной, и в то же время имеющей достаточно наглядное решение. Это случай степени $n = 4$, но предварительно следует разобрать степени 2 и 3.

Сразу отметим простой трюк, позволяющий понизить сложность задачи и тем самым суметь разобраться со степенью 4. Будем изучать только взаимное расположение корней, а потому наложим условие, что их сумма равна 0. Иные суммы достигаются сдвигом всей картинки. Тогда

$$(x - x_1)(x - x_2)(x - x_3)(x - x_4) = x^4 + ax^2 + bx + c$$

и коэффициент при x^3, равный сумме корней, уничтожен. Каждому такому полиному соответствует точка (a, b, c) в трёхмерном пространстве. Наша задача — описать точки, соответствующие полиномам с кратными корнями. Эту поверхность, дискриминант приведённой квартики, и называют ласточкиным хвостом.

В случае степени 2 тот же трюк сводит полином к

$$(x - x_1)(x - x_2) = x^2 + a,$$

и кратный корень есть только при $a = 0$. Одномерное пространство параметров разделено этой точкой на две части: при $a < 0$ корни вещественны, а при $a > 0$ образуют комплексно-сопряжённую пару.

В случае степени 3 имеем

$$(x - x_1)(x - x_2)(x - x_3) = x^3 + ax + b.$$

Первый шаг метода решения кубического уравнения состоит как раз в переходе к такой форме. Пространство параметров двумерное, а дискриминант приведённой кубики является линией. Если уравнение имеет двойной корень u, то его третий корень равен $-2u$, так что

$$(x - u)(x - u)(x + 2u) = x^3 - 3u^2x + 2u^3.$$

Мы нашли параметризацию дискриминанта: $(a, b) = (-3u^2, 2u^3)$. Его явное уравнение $4a^3 + 27b^2 = 0$, и левая часть тут знакома тем, кто всё-таки решал кубические уравнения. Эта линия — полукубическая парабола. Начало координат является особой точкой (точкой возврата) и соответствует полиному x^3 с тройным корнем. Дискриминант делит плоскость параметров на две части: по одну сторону лежат полиномы с тремя вещественными корнями, а по другую — с одним вещественным и комплексно-сопряжённой парой.

рисунок

Наконец, займёмся дискриминантом приведённой квартики. Самое интересное тут тоже происходит вблизи начала координат, соответствующего совпадению всех четырёх корней, но поверхность имеет и другие особые точки. Найдём полиномы с тройным корнем:

$$(x-u)(x-u)(x-u)(x+3u) = x^4 - 6u^2x^2 + 8u^3x - 3u^4.$$

Получившаяся линия

$$(a,b,c) = (-6u^2, 8u^3, -3u^4)$$

может иметь особенность при $u = 0$, что в дальнейшем подтвердится. Найдём полиномы с двумя двойными корнями:

$$(x^2+w)^2 = x^4 + 2wx^2 + w^2.$$

Получилась линия $(a,b,c) = (2w, 0, w^2)$, то есть парабола. Все её точки также окажутся особыми точками поверхности, но разных типов.

Теперь найдём полиномы с хотя бы двойным вещественным корнем u. Два других корня в сумме дают $-2u$, а в качестве второго параметра v можно взять их произведение. Тогда

$$(x-u)^2(x^2+2ux+v) = x^4 + (v-3u^2)x^2 - 2u(v+u^2)x + u^2v,$$

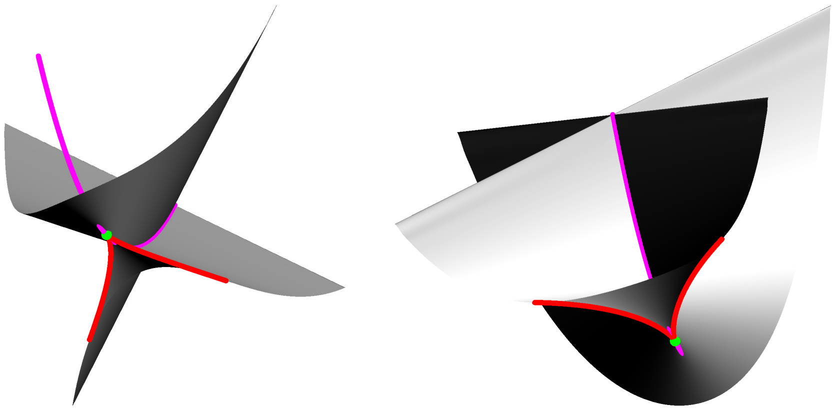
и мы получили параметризацию дискриминанта квартики:

$$(a,b,c) = (v-3u^2, 2u^3-2uv, u^2v).$$

Однако нужно учесть и полиномы, двойные корни которых образуют комплексно-сопряжённую пару. Это половина полученной выше параболы, образующая ручку дискриминанта навроде ручки зонтика Уитни и ручек коноида Плюккера; при описании ласточкиного хвоста про ручку зачастую забывают. На второй половине параболы $w < 0$, так что можно взять $w = -u^2$. На поверхности значения параметров $(\pm u, u^2)$ дают именно эти точки; выходит, это точки самопересечения, что и видно на рисунке.

Чтобы понять остальные особые точки, выпишем матрицу Якоби отображения $(u,v) \mapsto (a,b,c)$:

$$\frac{D(a,b,c)}{D(u,v)} = \begin{bmatrix} -6u & 1 \\ 6u^2-2v & -2u \\ 2uv & u^2 \end{bmatrix}.$$



Ранг её падает, когда все три минора второго порядка зануляются. Посчитав миноры, видим, что один из них равен $6u^2 + 2v$, а другие делятся на него. Подставив $v = -3u^2$ в параметризацию поверхности, получим в точности найденную выше параметризацию линии, на которой полином имеет тройной корень. Она является ребром возврата ласточкиного хвоста и притом сама имеет точку возврата в начале координат.

11.2. Гладкие линии, поверхности и многообразия

Гладкие через элементарные. Разобрав несколько примеров, мы видим, что все перечисленные способы задания фигур обобщаются, но простейший путь к сложному лежит через явное задание.

Определение. Назовём **элементарной** гладкой линией на плоскости $\mathbb{R}^2$ график гладкой функции на интервале.

рисунок

Определение. Назовём **элементарной** гладкой линией в пространстве $\mathbb{R}^3$ «совместный» график двух гладких функций на интервале.

рисунок

Любая из координат может выступать в качестве независимой переменной, а остальные будут функциями от неё. Аналогично можно определить элементарную гладкую линию в $\mathbb{R}^n$ для $n > 3$ как совместный график $n-1$ гладкой функции (или отображения в $\mathbb{R}^{n-1}$).

Определение. Назовём **элементарной** гладкой поверхностью в $\mathbb{R}^3$ график гладкой функции на открытом подмножестве плоскости.

рисунок

Любая пара координат может выступать в качестве независимой переменной, а третья будет функцией от них. Аналогично можно определить элементарную гладкую поверхность в $\mathbb{R}^n$ для $n > 3$ как совместный график $n-2$ гладких функций двух переменных.

Определение. Множество $L \subset \mathbb{R}^n$ называется **гладкой линией**, если всякая точка $p \in L$ имеет такую (открытую) окрестность U, что $L \cap U$ есть элементарная гладкая линия.

рисунок

Определение. Множество $S \subset \mathbb{R}^n$ называется **гладкой поверхностью**, если всякая точка $p \in S$ имеет такую окрестность U, что $S \cap U$ есть элементарная гладкая поверхность.

Примеры. Сфера, тор, крендель.

рисунки крендели?

Также мы отметили в нескольких примерах, как теорема о неявной функции выявляет условие, исключающее особенности неявного задания и почти исключающее особенности параметризации. Оба условия равносильны требованию, что ранг матрицы Якоби функций связи максимален. При параметризации приходится дополнительно оговорить, что отображение является вложением; обычно этого достигают, оговаривая область определения параметров, но формально можно потребовать непрерывной обратимости.

Теорема. *Уравнение $F(x, y, z) = 0$ с гладкой функцией $F\colon U \to \mathbb{R}$, удовлетворяющей на открытой области $U \subset \mathbb{R}^3$ условию $\nabla F \neq 0$, задаёт гладкую поверхность.*

Доказательство. Упражнение. □

Теорема. *Система уравнений*

$$F_1(x, y, z) = 0, \quad F_2(x, y, z) = 0$$

с гладкими функциями F_i, удовлетворяющими условию

$$\operatorname{rk} \frac{D(F_1, F_2)}{D(x, y, z)} = \operatorname{rk} \begin{bmatrix} F_{1,x} & F_{1,y} & F_{1,z} \\ F_{2,x} & F_{2,y} & F_{2,z} \end{bmatrix} = 2$$

всюду на непустом множестве её решений, задаёт гладкую линию.

Доказательство. Вытащите рассуждение из примера с двумя гиперболическими параболоидами (упражнение). □

Теорема. *Образ в $\mathbb{R}^3$ любого гладкого гомеоморфизма*

$$(u, v) \mapsto (x, y, z)$$

открытой плоской области, удовлетворяющего на ней условию

$$\operatorname{rk}\frac{D(x,y,z)}{D(u,v)} = \operatorname{rk}\begin{bmatrix} x_u & x_v \\ y_u & y_v \\ z_u & z_v \end{bmatrix} = 2,$$

есть гладкая поверхность.

Доказательство. Вытащите рассуждение из примера с ребром возврата (упражнение). □

Пример. Ь — не гомео (тем более, не диффео).

куда воткнут

Гладкие многообразия. Понятие многообразия обобщает на любую размерность понятия гладкой линии и поверхности. Линии являются 1-мерными многообразиями, а поверхности — 2-мерными.

Немедленную и вечную трудность для восприятия доставляет факт, что уже 3-мерные многообразия «живут» в пространствах размерности выше 3, где во многом отказывают наглядные образы.

Определение. Назовём **элементарным** гладким k-мерным многообразием в $\mathbb{R}^{k+m} = \mathbb{R}^k \times \mathbb{R}^m$ график гладкого отображения $f\colon D \to \mathbb{R}^m$ открытой области $D \subset \mathbb{R}^k$.

Определение. Множество $M \subset \mathbb{R}^{k+m}$ называется **гладким k-мерным многообразием**, если всякая точка $p \in M$ имеет такую окрестность U, что $M \cap U$ есть элементарное гладкое k-мерное многообразие. Poincaré 189

В пространстве $\mathbb{R}^3$ для поверхности имеем $k = 2$ и $m = 1$, а для линии $k = 1$ и $m = 2$. Это нужно помнить, думая о многообразиях.

Упражнение. *Обобщите три теоремы предыдущего подраздела на k-мерные многообразия.*

Примеры. Кроме обыденных, первыми примерами гладких многообразий, с которыми сталкивается образованный физик, вероятно являются конфигурационные пространства механических систем. Так называется множество всевозможных положений системы.

Сюда можно отнести и окружность S^1 (плоский маятник), и сферу S^2 (сферический маятник), и тор $S^1 \times S^1$ (двойной маятник).

Более сложный пример: все положения в пространстве твёрдого тела с неподвижным центром масс. Связав с телом репер системы координат, видим, что конфигурационным пространством является множество всех одинаково ориентированных реперов. Каждый из них можно однозначно представить ортогональной матрицей с определителем 1. Получается 3-мерное многообразие под названием $\mathbf{SO}_3$.

Касательное пространство. Касательной к графику функции мы давно назвали предел секущей. Отсюда легко получить определение касательной к гладкой линии, ибо вблизи любой своей точки такая линия является графиком.

Определение. **Касательным вектором** к гладкой линии с параметрическим заданием $\mathbf{r}(t)$ в точке $\mathbf{r}(t_0)$ называется $\dot{\mathbf{r}}(t_0) = \frac{d\mathbf{r}}{dt}(t_0)$.

При смене параметризации длина касательного вектора меняется, но направление может смениться только на противоположное.

Определение. **Касательным вектором** к гладкой поверхности S в какой-то её точке p называется касательный вектор в p к любой гладкой линии $L \subset S$, проходящей через p.

рисунок

Определение. **Касательным пространством** T_pS в точке p поверхности S называется линейная оболочка всех касательных векторов к S в p.

Абсолютно так же, как для поверхности, определяется касательное пространство гладкого k-мерного многообразия. Следует различать касательное пространство T_pM и касательное линейное многообразие $p+T_pM$. В случае линий и поверхностей последние обыкновенно называют касательной прямой и плоскостью; их уравнения пишут в системе координат самой фигуры. Координатами в касательном пространстве T_pM обычно служат дифференциалы независимых переменных, а начало лежит в точке p.

рисунок

Теорема. *Размерность касательного пространства всякого гладкого k-мерного многообразия в каждой его точке равна k.*

Доказательство. Разберём случай $k = 2$. Выбрав на поверхности S любую точку $p = (x_0, y_0, z_0)$, ограничимся элементарной гладкой поверхностью $S \cap U$, являющейся графиком функции $z(x, y)$.

рисунок

Линии $z(x, y_0)$ и $z(x_0, y)$ на $S \cap U$ имеют касательные векторы

$$\mathbf{r}_x = (1, 0, z_x), \quad \mathbf{r}_y = (0, 1, z_y),$$

где частные производные взяты в точке (x_0, y_0). Каждая гладкая линия на $S \cap U$ задаётся параметризацией $(x(t), y(t), z(t))$, где $(x(t), y(t))$ задаёт гладкую плоскую линию, а $z(t) = z(x(t), y(t))$. Касательный вектор этой линии равен

$$(\dot{x}, \dot{y}, \dot{z}) = (\dot{x}, \dot{y}, z_x\dot{x} + z_y\dot{y}) = \dot{x}\mathbf{r}_x + \dot{y}\mathbf{r}_y,$$

где производные взяты при значении $t = t_0$, соответствующем точке $p = (x_0, y_0, z_0)$. Таким образом, все касательные векторы к S в p являются линейными комбинациями двух независимых. $\square$

Теорема. *Касательное пространство T_pM совпадает со множеством касательных векторов к M в p.*

Доказательство. Упражнение. Не забудьте про нулевой вектор. ☐

Уравнения касательной плоскости. Для параметризованной поверхности они фактически найдены выше в доказательстве теоремы: хотя там использовано явное задание на элементарном куске, рассуждение дословно переносится на параметрический случай. В точке $\mathbf{r}_0 = \mathbf{r}(u_0, v_0)$ на поверхности с параметризацией $\mathbf{r} = \mathbf{r}(u, v)$ векторы $\mathbf{r}_u(u_0, v_0)$ и $\mathbf{r}_v(u_0, v_0)$ частных производных радиус-вектора по параметрам составляют естественный базис касательной плоскости. Поэтому она задаётся параметрическим уравнением 20.02.17

$$\mathbf{r} = \mathbf{r}_0 + \alpha\mathbf{r}_u + \beta\mathbf{r}_v, \quad \alpha, \beta \in \mathbb{R}.$$

Напомним также уравнение касательной плоскости для поверхности, заданной неявно уравнением связи $F(x, y, z) = 0$, появлявшееся в задачах ещё в первом семестре. Возьмём на поверхности параметризованную линию через точку $\mathbf{r}_0 = (x_0, y_0, z_0)$. Функция F постоянна на поверхности и на этой линии в частности, поэтому её производная по вектору скорости этой линии равна нулю. Найдём её по правилу цепочки и затем заменим дифференциалы конечными приращениями:

$$F_x(\mathbf{r}_0) \cdot (x - x_0) + F_y(\mathbf{r}_0) \cdot (y - y_0) + F_z(\mathbf{r}_0) \cdot (z - z_0) = 0,$$

или в векторной форме, со скалярным произведением:

$$\nabla F(\mathbf{r}_0) \cdot (\mathbf{r} - \mathbf{r}_0) = 0.$$

Это и есть искомое уравнение касательной плоскости. Здесь виден второй геометрический смысл градиента: это вектор нормали к касательной плоскости. Логично будет называть его нормальным вектором к самой поверхности, но об этом чуть ниже в более общей ситуации.

Градиенты связей. Теорема о неявной функции даёт условие, при выполнении которого набор уравнений связи $\{F_i = 0\}$ определяет гладкое многообразие. Для применимости теоремы матрица Якоби $\frac{DF}{Dx}$ в каждой удовлетворяющей этим уравнениям точке должна иметь хотя бы один ненулевой минор максимального размера. Это равносильно тому, что её ранг максимален. Также это равносильно тому, что её строки линейно независимы в каждой точке, а строки эти являются градиентами ∇F_i.

Определение. Ортогональное дополнение к касательному пространству T_pM называют **нормальным пространством** N_pM.

Лемма. *Когда в точке гладкого многообразия градиенты связей линейно независимы, их линейная оболочка совпадает с нормальным пространством.*

Доказательство. Подсчёт размерностей для $M \subset \mathbb{R}^{k+m}$ сразу даёт

$$\dim T_pM = \dim M = k, \qquad \dim N_pM = m,$$

так что остаётся установить ортогональность касательных векторов и градиентов связей. Для этого нужно повторить сделанное при выводе уравнения касательной плоскости. При движении из точки p по любой траектории на многообразии $M = \{F_i = 0\}$ значение каждой функции F_i остаётся неизменным. Их производные по касательному вектору в p, определённому этой траекторией, равны нулю. Значит, этот вектор ортогонален всем градиентам связей. □

Следующий пример покажет, что для произвольного неявного задания утверждение леммы ложно. Это приходится учитывать в следующем разделе. Любое многообразие имеет (бесконечно) много различных неявных заданий; гладкость самого многообразия не влечёт линейную независимость градиентов связей в каждой точке.

Пример. Как уже отмечено в предыдущем разделе, уравнение $x^2 = 0$ на плоскости Oxy задаёт, по нашим относительно наивным представлениям, ту же прямую линию, что и уравнение $x = 0$. Градиент связи $F(x, y) = x^2$ на ней тождественно равен нулю, а потому в одиночестве образует линейно зависимую систему в каждой точке.

20.02.17 Можно понять, в чём тут дело, представив эту прямую как пересечение поверхностей $z = x^2$ и $z = 0$. Поскольку координата y никуда не входит, забудем о ней вообще и увидим на плоскости Oxz касание параболы и прямой. Пошевелим эту прямую, взяв вместо неё $z = c$, где $c \to 0$ справа. Пересечение состоит из двух точек, в пределе сливающихся в одну. Поэтому более точно говорить, что уравнение $x^2 = 0$ задаёт «двойную точку» (и двойную прямую в исходной постановке). При $c < 0$ пересечение также состоит из двух точек, но они мнимые.

Дальше мы постараемся не затрагивать эти идеи — там всё хитро.

Набор гладких уравнений связи с всюду зависимыми градиентами бывает избыточен: упрощение его может сократить количество неявных уравнений для того же самого многообразия. Однако набор, у которого градиенты зависимы в отдельных точках, задаёт многообразие с вероятными **особенностями**. Простейший пример: конус $x^2+y^2-z^2=0$ и его плоская версия, крест $x^2-y^2=0$.

В особой точке тоже можно ввести касательное пространство как линейную оболочку касательных векторов; размерность его может превосходить размерность самого многообразия. Изучение особенностей многообразий входит в сложный раздел математики с невинным названием **алгебраическая геометрия**.

11.3. Условные экстремумы

Терминология. Экстремум функции при условии, что переменные подчинены каким-то связям, называют условным. Задача непосредственно связана с нуждами аналитической механики, а также, например, теории оптимального управления. лекция 5 27.02.17

Будем рассматривать гладкие **целевые** функции $f\colon D\to\mathbb{R}$ на открытой области $D\subset\mathbb{R}^n$ с гладким многообразием $M\subset D$.

Определение. Точка p многообразия M называется точкой **условного минимума** функции f на M, если она имеет такую окрестность U, что $f(p)\leqslant f(q)$ для всех $q\in M\cap U$.

Аналогично, с обратным неравенством, определяется точка **условного максимума**. Точки условного минимума и максимума вместе называются точками **условного экстремума**. Аналогично, со строгими неравенствами, определяются точки **строгого** условного экстремума. рисунки?

Пример. Функция $f(x,y)=x+y$ на окружности $x^2+y^2=1$. Локальных экстремумов на $D=\mathbb{R}^2$ нет, условные есть.

При параметрическом задании многообразия, а также явном как его частном случае, условный экстремум сводится к локальному экстремуму простой подстановкой в целевую функцию выражений для зависимых переменных через параметры. Например, из целевой функции $f(x,y,z)$ на элементарной поверхности $z=z(x,y)$ получаем целевую функцию $\hat{f}(x,y)=f(x,y,z(x,y))$ на плоской области.

Необходимое условие условного экстремума. Ввиду сказанного в предыдущем абзаце, задача поиска условного экстремума нова для нас только при неявном задании многообразия.

Теорема (необходимое условие). *В каждой точке p условного экстремума функции f на многообразии M имеем:*

(1) $\nabla f(p)\perp T_pM$;
(2) $\partial_{\mathbf{v}}f(p)=0$ *для всех* $\mathbf{v}\in T_pM$.

Ввиду тесной связи $\partial_{\mathbf{v}} f = \mathbf{v} \cdot \nabla f$ производной по вектору и градиента, эти два утверждения равносильны.

Доказательство. Возьмём в M гладкую линию $\mathbf{r}(t)$ с касательным вектором $\mathbf{v}$ в точке $p = \mathbf{r}(0)$ и положим $\varphi(t) = f(\mathbf{r}(t))$. Если в p есть условный экстремум функции f на M, то функция φ имеет обычный локальный экстремум при $t = 0$, а тогда $\partial_{\mathbf{v}} f(p) = \varphi'(0) = 0$. □

Упражнение. *В тех же условиях докажите, что при задании гладкого многообразия M системой связей $\{F_i = 0\}$ градиент $\nabla f(p)$ является линейной комбинацией градиентов $\nabla F_i(p)$.*

не гладкий!

Пример. $f(x, y) = x + y$ на листе Декарта $x^3 + y^3 - 3xy = 0$. Условный экстремум находится одним взглядом: в единственной точке касания.

Упражнение. *В тех же условиях докажите, что «обычно» M касается* **гиперповерхности уровня** *целевой функции, то есть многообразия*

$$N = \{q \in D \mid f(q) = f(p)\}.$$

Здесь это фактически означает, что $T_pM \subset T_pN$. Правда, могут появиться проблемы из-за строения N. Например, для точки p строгого локального экстремума будет $N = \{p\}$.

Метод исключения дифференциалов. Этот метод поиска точек условного экстремума функции f основан на том, из равенства $df = 0$, расписанного через дифференциалы dx_j и dy_i независимых и зависимых переменных, можно исключить все dy_i, выражая их через dx_j. Этим мы занимались неделю назад; для нахождения m штук неизвестных dy_i используются m штук (независимых между собой) дифференциалов связей: $dF_i = 0$.

гиперссылка

Теорема. *Если в окрестности точки p условного экстремума функции f есть представление*

$$df = \sum_{1 \leqslant j \leqslant k} A_j(x_1, \ldots, y_m)\, dx_j,$$

то значения всех функций A_j в точке p равны нулю.

Таким образом, для отыскания $k + m$ неизвестных координат подозрительных на экстремум f точек имеется столько же уравнений: m изначальных связей $F_i = 0$ и k зависящих от f условий $A_j = 0$.

Применение этого метода на практике осложняет то обстоятельство, что набор зависимых переменных обычно не удаётся выбрать глобально (на всём M целиком). При каждом выборе рассмотрение ограничивается элементарным куском многообразия. Чтобы не потерять точки возможного экстремума, приходится перебирать все $\binom{k+m}{m}$ якобианов и для каждого ненулевого из них повторять процедуру исключения дифференциалов. Более того, одна и та же точка экстремума может при этом попадать в несколько разных, но перекрывающихся элементарных кусков, так что в списках подозрительных точек ещё надо выискивать дубликаты.

Метод множителей Лагранжа. В этом красивом методе не приходится среди всех $n = k + m$ переменных выискивать зависимые.

Вводим m новых вспомогательных переменных λ_i в расчёте, что они станут коэффициентами линейной комбинации

$$\nabla f(p) = \sum_{1 \leqslant i \leqslant m} \lambda_i \nabla F_i(p).$$

Далее, для сокращения записи вводим **функцию Лагранжа**

$$L = f - \sum \lambda_i F_i.$$

Она придаёт необходимому условию условного экстремума следующий лаконичный вид.

Следствие. *Каждая точка p условного экстремума функции f на многообразии удовлетворяет уравнению*

$$\nabla L(p) = 0.$$

Это векторное уравнение в $\mathbb{R}^n$. Вместе со связями $F_i = 0$ набирается $n + m$ уравнений для отыскания всех m штук **множителей Лагранжа** λ_i и n координат точек многообразия, подозрительных на экстремум.

Пример. Можно тот же лист Декарта!

пример?

Упражнение. *Где в методе множителей Лагранжа действительно используется гладкость многообразия M?*

Пример (экстремумы квадратичной формы на сфере). Здесь замечательно переплелись геометрия, алгебра и анализ. Для простоты, разберём плоский случай: условные экстремумы квадратичной формы

$$f(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2$$

на единичной окружности $F(x, y) = x^2 + y^2 - 1 = 0$.

Пользуясь следствием, найдём подозрительные точки. Дифференцируем функцию Лагранжа и видим, что условие в данном случае выражается системой линейных уравнений, которую мы запишем в матричном виде с симметричной матрицей:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \lambda x \\ \lambda y \end{bmatrix},$$

или $(A - \lambda E)\mathbf{x} = \mathbf{0}$. Ненулевые решения есть лишь при $\det(A - \lambda E) = 0$.

Следовательно, тут множители Лагранжа оказываются собственными числами матрицы коэффициентов формы, а точки экстремума — собственными векторами. Можно убедиться, что максимум отвечает (наи)большему собственному числу, а минимум (наи)меньшему. Эти выводы сохранятся и в многомерной задаче; собственные же векторы, соответствующие промежуточным собственным числам, окажутся обобщёнными седловыми точками, а не экстремумами.

окрасить сферу?

Упражнение. *Убедитесь.*

Достаточные условия условного экстремума. Зачастую вместо громоздкого исследования характера точек, подозрительных на экстремум функции f на многообразии, достаточно сравнить её значения в этих точках. Точнее, это работает на компактных многообразиях.

Исследование же опирается на второй дифференциал. Нужно выразить $d^2 f$ в подозрительной точке p как квадратичную форму от дифференциалов независимых переменных. В случае её строгой знакоопределённости или знакопеременности делается заключение о наличии или отсутствии строгого условного экстремума в p. Фактически это сведение задачи на элементарном куске многообразия к поиску локального экстремума.

Метод множителей Лагранжа упрощает задачу: поскольку f и L совпадают на многообразии, можно исследовать не $d^2 f$, а $d^2 L$. Последний находят, не взирая на различие между независимыми и зависимыми переменными x_i. Хитрость функции Лагранжа в том, что $d^2 L$ априори не содержит слагаемых с $d^2 x_i$, потому что коэффициенты $\partial L/\partial x_i$ при них нулевые ввиду необходимого условия!

Увидим это для простейшего случая неявного задания: $F(x, y) = 0$ с одномерными переменными. Тогда

$$d^2 L = d(L_x\,dx + L_y\,dy) = L_{xx}\,dx^2 + 2L_{xy}\,dx\,dy + L_{yy}\,dy^2 + L_y\,d^2 y.$$

На практике здесь всегда $L_y = 0$, ибо $\nabla L = 0$ в точках, подозрительных на экстремум. При бо́льшем количестве переменных появляются лишь дополнительные индексы и суммирования.

Экстремальные значения функции на области. Точки области, где некоторая функция f достигает своего максимума или минимума, могут находиться внутри области или на её границе. Соответственно, поиск подозрительных точек разбивается на две задачи: внутри по обычному методу для локальных экстремумов; на границе, которая часто является гладким многообразием или составлена из нескольких гладких кусков, по одному из методов для условных экстремумов. Затем сравниваются значения f во всех подозрительных точках.

Простые примеры показывают, что, в отличие от одномерного случая (промежутка), забывать о границе нельзя, даже если локальный экстремум один. Причина кроется в седловых точках.

рисунок

Содержание