УДК 517.1
ББК 22.16
Д998

Дятлов Г. В.

Основы математического анализа для студентов-физиков. Лекции. 5. Дифференциальное исчисление функций многих переменных / Г. В. Дятлов. — Новосибирск: Издательство Института математики, 2014. — 68 с.

ISBN 978–5–86134–153–0

Пятый раздел курса лекций, прочитанных автором студентам физического факультета Новосибирского государственного университета в 2014/15 учебном году. Включает в себя элементы дифференциального исчисления функций многих переменных.

Для студентов специальностей с углубленным изучением математического анализа.

УДК 517.1
ББК 22.16
Д998

Д$\frac{1602070000–02}{Я82(03)–14}$ Без объявл.

ISBN 978–5–86134–153–0

§ 1. Нормированное пространство $\mathbb{R}^n$

1.1. Арифметическое конечномерное пространство.

Определение пространства $\mathbb{R}^n$. Пространство $\mathbb{R}^n$ состоит из упорядоченных наборов $x = (x^1, \ldots, x^n)$ вещественных чисел x^i, $i = 1, \ldots, n$, называемых *векторами*. Иначе можно сказать, что $\mathbb{R}^n$ — это прямое произведение n экземпляров множества $\mathbb{R}$:

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ штук}} = \{(x^1, \ldots, x^n) \mid x^1, \ldots, x^n \in \mathbb{R}\}.$$

На $\mathbb{R}^n$ определены сумма элементов и умножение на скаляр, а именно для $x = (x^1, \ldots, x^n)$, $y = (y^1, \ldots, y^n)$, $\lambda \in \mathbb{R}$ полагают

$$x + y = (x^1 + y^1, \ldots, x^n + y^n), \quad \lambda x = (\lambda x^1, \ldots, \lambda x^n).$$

Введенные операции обладают всеми свойствами, предъявляемыми к операциям в векторном пространстве, стало быть, $\mathbb{R}^n$ — векторное пространство. О нем также говорят как об *арифметическом конечномерном пространстве.*

1.2. Расстояние и норма в $\mathbb{R}^n$.

Нам надо будет определить понятие предела последовательностей элементов из $\mathbb{R}^n$, а также предела отображений между такими пространствами. Определение предела в одномерной ситуации опирается на понятие расстояния между точками. Оно определялось как модуль разности двух чисел. Здесь также потребуется расстояние между двумя векторами. Способов определить расстояние в $\mathbb{R}^n$ много, поэтому вводится некоторое общее понятие метрики. Для определения расстояния не требуется никаких дополнительных свойств от того множества, на котором метрика определяется, поэтому дадим определение метрики на произвольном множестве, тем более что в будущем это понадобится.

Определение метрики и метрического пространства. Пусть X — произвольное множество. Функцию $d(x, y)$, определенную на парах $(x, y) \in X^2$, называют *метрикой на* X, если выполнены следующие условия:

(1) $d(x, y) \geqslant 0$ и $d(x, y) = 0$ тогда и только тогда, когда $x = y$;
(2) $d(x, y) = d(y, x)$ (симметричность);

(3) $d(x,z) \leqslant d(x,y) + d(y,z)$ для любых $x, y, z \in X$ (неравенство треугольника).

Множество с заданной на нем метрикой называют *метрическим пространством.*

Пример 1 (расстояние на клетчатой бумаге). Возьмем лист клетчатой бумаги, и пусть множество X состоит из всех точек пересечения линий. В качестве расстояния между точками можно принять наименьшее количество отрезков между вершинами по горизонтали и по вертикали, по которым надо пройти для попадания из одной точки в другую. Нетрудно проверить, что так определенная функция обладает всеми свойствами расстояния. Обратим внимание на то, что здесь никакой векторной структуры нет — мы не может складывать перекрестки или умножать их на число.

Пример 2 (рукопожатия). Определим расстояние между любыми двумя людьми, полагая его равным числу рукопожатий, которые их связывают, т. е. количеству знакомых между ними. Расстояние от данного человека до него самого равно нулю. Расстояние от данного человека до его знакомого равно единице. Если его знакомый знает кого-то, которого не знает данный человек, то расстояние равно двум, и т. д. Нетрудно понять, что так введенная для пар людей функция обладает всеми свойствами метрики.

Вернемся в пространство $\mathbb{R}^n$. Когда есть структура векторного пространства, необязательно определять расстояние между двумя произвольными элементами, как это сделано в метрическом пространстве. Достаточно определить расстояние от каждой точки векторного пространства до нуля и положить расстояние от x до y равным расстоянию от начала координат до точки $y - x$.

Определение нормы и нормированного пространства. Пусть X — векторное пространство. Функцию $\|\cdot\|$, которая произвольному $x \in X$ сопоставляет число $\|x\|$, называют *нормой*, если выполнены условия:

(1) $\|x\| \geqslant 0$, $x \in X$, и $\|x\| = 0$ тогда и только тогда, когда $x = 0$,
(2) $\|\lambda x\| = |\lambda| \|x\|$, $\lambda \in \mathbb{R}$, $x \in X$ (положительная однородность)
(3) $\|x + y\| \leqslant \|x\| + \|y\|$, $x, y \in X$ (неравенство треугольника).

Векторное пространство с заданной на нем нормой называют *нормированным пространством.*

Замечание. Норма всегда порождает метрику

$$d(x,y) = \|x - y\|.$$

Легко проверить, что $d(x, y)$ удовлетворяет аксиомам метрики. Тем самым любое нормированное пространство является метрическим пространством.

Однако не всякая метрика порождается нормой. Во-первых, на том множестве, где задана метрика, может не быть структуры векторного пространства, т. е. нет сложения элементов и умножения их на скаляр. Во-вторых, может оказаться так, что векторная структура есть, однако невозможно при заданной метрике определить норму так, чтобы она порождала данную метрику по указанному выше правилу.

Определение основных норм и метрик в $\mathbb{R}^n$. Для простоты при обозначении основных норм в $\mathbb{R}^n$ будем писать одну черту вместо двух. В определении норм и соответствующих им метрик берутся произвольные элементы $x = (x^1, \ldots, x^n)$, $y = (y^1, \ldots, y^n)$ из $\mathbb{R}^n$.

1. *Евклидова норма* (l_2-*норма*) определяется так:

$$|x|_2 = \sqrt{\sum_{i=1}^{n} (x^i)^2}.$$

Соответствующая метрика равна

$$d_2(x, y) = \sqrt{\sum_{i=1}^{n} (x^i - y^i)^2}.$$

2. l_1-*Норма* определяется так:

$$|x|_1 = \sum_{i=1}^{n} |x^i|.$$

Ей соответствует метрика

$$d_1(x, y) = \sum_{i=1}^{n} |x^i - y^i|.$$

3. l_∞-*Норма* (или *равномерная норма*) определяется равенством

$$|x|_\infty = \max_{i=1,\ldots,n} |x^i|,$$

и ей соответствует метрика

$$d_\infty(x, y) = \max_{i=1,\ldots,n} |x^i - y^i|.$$

4. l_p-*Норма* и соответствующая ей метрика определяются для $p \geqslant 1$ так:

$$|x|_p = \sqrt{\sum_{i=1}^{n} |x^i|^p},$$

$$d(x, y)_p = \sqrt{\sum_{i=1}^{n} |x^i - y^i|^p}.$$

Ясно, что если $p = 1$ или $p = 2$, то l_p-норма совпадает с определенными выше.

Для нас основной будет евклидова норма, потому что она связана со скалярным произведением

$$\langle x, y\rangle = x^1 y^1 + \cdots + x^n y^n$$

следующим образом:

$$|x|_2 = \sqrt{\langle x, x\rangle}.$$

Изредка оказываются удобными l_1- и l_∞-нормы.

Теорема 1 (о сравнении основных норм). *Для любого $x \in \mathbb{R}^n$ верны неравенства*

$$|x|_\infty \leqslant |x|_2 \leqslant |x|_1, \quad \frac{1}{\sqrt{n}}|x|_1 \leqslant |x|_2 \leqslant \sqrt{n}\,|x|_\infty.$$

Доказательство. Шаг 1 ($|x|_\infty \leqslant |x|_2$). Заметим, что для любого фиксированного $i = 1, \ldots, n$

$$(x^i)^2 \leqslant \sum_{j=1}^{n} (x^j)^2 \iff |x^i| \leqslant \sqrt{\sum_{j=1}^{n} (x^j)^2},$$

а значит, и

$$|x|_\infty = \max_{i=1,\ldots,n} |x^i| \leqslant \sqrt{\sum_{j=1}^{n} (x^j)^2}.$$

Шаг 2 ($|x|_2 \leqslant |x|_1$). Имеем

$$|x|_2 \leqslant |x|_1 \iff \sqrt{\sum_{i=1}^{n} (x^i)^2} \leqslant \sum_{i=1}^{n} |x^i| \iff \sum_{i=1}^{n} (x^i)^2 \leqslant (|x^1| + \cdots + |x^n|)^2,$$

но последнее неравенство очевидно верное, ибо в правой части после возведения в квадрат появятся все слагаемые из левой части и еще удвоенные произведения элементов суммы.

Шаг 3 ($\frac{1}{\sqrt{n}}|x|_1 \leqslant |x|_2$). Воспользуемся неравенством Иенсена, согласно которому для выпуклой функции f при любых $x^1, \ldots, x^n \in \mathbb{R}$ и $\alpha_1, \ldots, \alpha_n \geqslant 0$, $\alpha_1 + \cdots + \alpha_n = 1$, верно

$$f(\alpha_1 x^1 + \cdots + \alpha_n x^n) \leqslant \alpha_1 f(x^1) + \cdots + \alpha_n f(x^n), \qquad (*)$$

и выпуклостью функции $f(x) = x^2$. Взяв в $(*)$ $\alpha_1 = \cdots = \alpha_n = \frac{1}{n}$, получаем

$$\left(\frac{1}{n}|x^1| + \cdots + \frac{1}{n}|x^n|\right)^2 \leqslant \frac{1}{n}((x^1)^2 + \cdots + (x^n)^2),$$

или после извлечения корня

$$\frac{1}{\sqrt{n}}|x|_1 \leqslant \sqrt{\sum_{i=1}^{n}(x^i)^2} = |x|_2.$$

Шаг 4 ($|x|_2 \leqslant \sqrt{n}|x|_\infty$). Оценим

$$|x|_2 = \sqrt{\sum_{i=1}^{n}|x^i|^2} \leqslant \sqrt{\sum_{i=1}^{n}(\max_{j=1,\ldots,n}|x^j|)^2} = \sqrt{n}\max_{j=1,\ldots,n}|x^j| = \sqrt{n}\,|x|_\infty.$$

Теорема доказана.

Дадим геометрическую интерпретацию утверждения теоремы 1.

Определение шара нормы. Пусть в $\mathbb{R}^n$ задана какая-то норма $\|\cdot\|$. Множество

$$B(x,r) = \{y \in \mathbb{R}^n \mid \|x-y\| < r\}$$

называют *открытым шаром радиуса $r > 0$ с центром в точке $x \in \mathbb{R}^n$*.

Множество

$$\overline{B}(x,r) = \{y \in \mathbb{R}^n \mid \|x-y\| \leqslant r\}$$

называют *замкнутым шаром радиуса $r > 0$ с центром в точке $x \in \mathbb{R}^n$*.

Множество
$$S(x,r) = \{y \in \mathbb{R}^n \mid \|x-y\| = r\}$$
называют *сферой радиуса* $r > 0$ *с центром в точке* $x \in \mathbb{R}^n$.

В этом определении словосочетания «открытый шар» и «замкнутый шар» надо воспринимать как целостные, т. е. нет просто шара, есть открытый шар и замкнутый шар. Если говорят «шар», то подразумевают открытый шар.

Пример 3 (шары основных норм в $\mathbb{R}^2$).

1. Шар
$$B_2(0,r) = \{(x^1,x^2) \in \mathbb{R}^2 \mid (x^1)^2 + (x^2)^2 < r^2\}$$
изображается как круг на координатной плоскости радиуса r с центром в начале координат, в который не входит ограничивающий его контур (рис. 1.1).

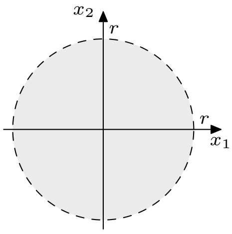

Рис. 1.1.

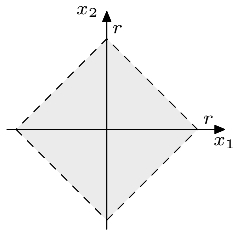

Рис. 1.2.

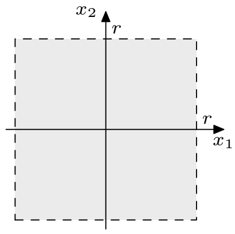

Рис. 1.3.

2. Шар в l_1-норме — это множество
$$B_1(0,r) = \{(x^1,x^2) \in \mathbb{R}^2 \mid |x^1| + |x^2| < r\},$$
представляющее собой квадрат на координатной плоскости, без ограничивающего его контура (рис. 1.2).

3. Шар в норме l_∞ — это множество
$$B_\infty(0,r) = \{(x^1,x^2) \mid \max\{|x^1|,|x^2|\} < r\}.$$
Он выглядит как квадрат на координатной плоскости, в который не входит ограничивающий его контур (рис. 1.3).

Как видно, не все шары круглые, но тем не менее они называются шарами.

Замкнутые шары отличаются от открытых только тем, что они включают ограничивающие их контуры.

ЗАМЕЧАНИЕ (о геометрическом смысле неравенств). Изобразим единичные шары в разных нормах на одной плоскости (рис. 1.4). Видно, что они вкладываются друг в друга: $B_1 \subset B_2 \subset B_\infty$. Такое расположение шаров иллюстрирует неравенство

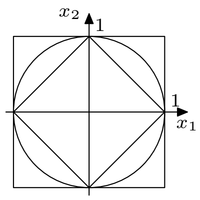

Рис. 1.4.

$$|x|_\infty \leqslant |x|_2 \leqslant |x|_1.$$

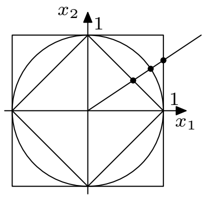

Рис. 1.5.

В этом можно убедиться следующим образом. Пусть $x = (x^1, x^2)$ — некоторая точка плоскости. Как измеряется расстояние от начала координат до этой точки? Проводится луч, идущий из начала через данную точку, на этом луче берется некоторый отрезок в качестве единичного и с использованием этого масштаба проводится измерение. Изобразим на плоскости единичные сферы, т. е. контуры, ограничивающие шары в разных нормах (рис. 1.5), и заметим, что относительно соответствующей нормы единичным окажется отрезок от начала координат до точки пересечения луча со сферой. Ясно, что чем короче отрезок, тем больше соответствующая норма данного элемента x. Самый короткий отрезок будет относительно нормы l_1, затем идет отрезок по норме l_2 и завершает отрезок относительно нормы l_∞. Кстати, из рис. 1.5 видно, что относительно нормы l_2 масштаб, т. е. единичный относительно нормы отрезок, во всех направлениях один и тот же, а для других норм он меняется в зависимости от расположения данной точки на плоскости.

Чему соответствует на плоскости второе неравенство из теоремы 1? Для двух переменных ($n = 2$) оно выглядит так:

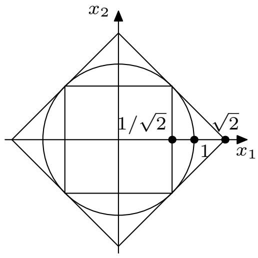

Рис. 1.6.

$$\frac{1}{\sqrt{2}}|x|_1 \leqslant |x|_2 \leqslant \sqrt{2}|x|_\infty.$$

Это неравенство соответствует вложению шаров, изображенному на рис. 1.6.

Отметим важное наблюдение. Если есть две нормы в $\mathbb{R}^n$, то, сжимая или растягивая шары каждой из этих норм, можно добиться того, что один из них, взятый с некоторым коэффициентом, окажется внутри второго и, в свою очередь, второй, также взятый с некоторым коэффициентом, окажется в первом.

Замечание о геометрическом смысле свойств нормы можно перефразировать так. Свойство (1) означает, что по любому направлению отрезок от начала координат до пересечения с единичной сферой не сводится к точке и ограничен. Свойство (2) геометрически означает симметрию шара относительно нуля.

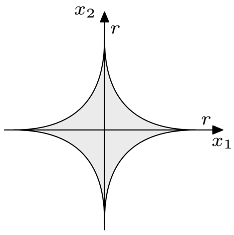

Рис. 1.7.

Свойство (3) означает, что шар должен быть множеством выпуклым, т. е. для любых двух точек шара весь отрезок с концами в этих точках должен содержаться в шаре. Например, если взять на роль нормы выражение

$$\left(\sum_{i=1}^{n} |x^i|^p\right)^{1/p}$$

с $p < 1$, то единичный шар в таком случае выглядит так, как показано на рис. 1.7, и не является выпуклым множеством, стало быть, для таких p не выполнено неравенство треугольника и указанное выражение нормой не будет.

Теорема 2 (об эквивалентности норм). *В пространстве $\mathbb{R}^n$ любые две нормы эквивалентны, т. е. для любых двух норм $\|\cdot\|_a$ и $\|\cdot\|_b$ существуют такие константы $C_1, C_2 > 0$, что для любого $x \in \mathbb{R}^n$ выполнены неравенства*

$$\|x\|_a \leqslant C_1\|x\|_b, \quad \|x\|_b \leqslant C_2\|x\|_a.$$

Мы убедились в справедливости утверждения теоремы на примерах основных норм и в общем случае доказывать его не будем.

Теорема 3 (о классических неравенствах). (1) *Для любых $x, y \in \mathbb{R}^n$ имеет место неравенство Коши — Буняковского*

$$\left|\sum_{i=1}^{n} x^i y^i\right| \leqslant \left(\sum_{i=1}^{n} (x^i)^2\right)^{1/2} \left(\sum_{i=1}^{n} (y^i)^2\right)^{1/2},$$

или, иначе,

$$|\langle x, y\rangle| \leqslant |x|_2 |y|_2.$$

(2) *Для любых $x, y \in \mathbb{R}^n$ верно неравенство Минковского*

$$|x+y|_p \leqslant |x|_p + |y|_p, \quad p \geqslant 1.$$

Доказательство. (1) Если хотя бы один из элементов x, y нулевой, то утверждение очевидно: с обеих сторон находятся нули. Пусть x и y отличны от нуля. Рассмотрим функцию

$$f(t) = |x - ty|_2^2 = \langle x - ty, x - ty \rangle.$$

Ввиду линейности скалярного произведения по каждой из переменных можно раскрыть скобки:

$$f(t) = \langle x, x \rangle - 2\langle x, y \rangle t + \langle y, y \rangle t^2.$$

Функция $f(t)$ является квадратичной функцией, и при этом она как квадрат нормы всегда неотрицательна. Коэффициент при t^2 положителен, стало быть, дискриминант неположителен, т. е.

$$\langle x, y \rangle^2 - \langle x, x \rangle \langle y, y \rangle \leqslant 0,$$

или

$$|\langle x, y \rangle| \leqslant \sqrt{\langle x, x \rangle \langle y, y \rangle}.$$

(2) Докажем неравенство Минковского только для $p = 2$. Пользуясь неравенством Коши — Буняковского, имеем

$$\begin{aligned} |x + y|_2^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\ &\leqslant \langle x, x \rangle + 2|\langle x, y \rangle| + \langle y, y \rangle \leqslant \langle x, x \rangle + 2|x|_2|y|_2 + \langle y, y \rangle \\ &= |x|_2^2 + 2|x|_2|y|_2 + |y|_2^2 = (|x|_2^2 + |y|_2^2)^2, \end{aligned}$$

откуда требуемое неравенство получается в результате извлечения корня.

Теорема доказана.

1.3. Взаимное расположение точки и множества.

Определение внутренней, внешней, граничной и предельной точек. Пусть $X \subset \mathbb{R}^n$ — множество и $a \in \mathbb{R}^n$ — точка.

(1) Точку a называют *внутренней точкой множества* X, если существует такое $r > 0$, что $B_2(a, r) \subset X$. В этом случае также говорят, что *множество X является окрестностью точки a*. Множество всех внутренних точек множества X называют *внутренностью* X и обозначают через $\operatorname{int} X$.

(2) Говорят, что точка a — *внешняя точка множества* X, если существует такое $r > 0$, что $B_2(a, r) \subset \mathbb{R}^n \setminus X$ или, иными словами, если a — внутренняя точка дополнения множества X до $\mathbb{R}^n$. Множество всех внешних точек множества X называют *внешностью* X и обозначают через $\operatorname{ext} X$.

(3) Точку a называют *граничной точкой множества* X, если в любом шаре в центром в этой точке есть как точки множества X, так и точки его дополнения. Множество всех граничных точек называют *границей* X и обозначают через $\mathrm{fr}\, X$.

(4) Точку a называют *предельной точкой множества* X, если в любом шаре с центром в этой точке есть точка множества X, отличная от a. Объединение множества X и множества его предельных точек называют *замыканием множества* X и обозначают либо $\mathrm{cl}\, X$, либо $\overline{X}$.

Замечание. Для любого множества $X \subset \mathbb{R}^n$ всё $\mathbb{R}^n$ распадается на три множества: внутренность $\mathrm{int}\, X$, внешность $\mathrm{ext}\, X$ и границу $\mathrm{fr}\, X$, причем эти множества попарно не пересекаются.

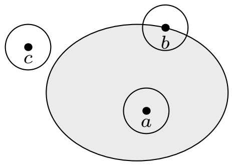

Рис. 1.8.

Сформируем образное представление о внутренних, граничных и внешних точках множества. Изобразим какое-нибудь множество X в $\mathbb{R}^2$ (рис. 1.8), состоящее из всех точек плоскости, охваченных контуром, включая сам контур. Точка a на рис. 1.8 — внутренняя точка множества X, ибо есть шар с центром в этой точке, состоящий только из точек множества X. Ясно, что эта точка не граничная и не внешняя, так как в таком шаре нет точек дополнения множества X. Кстати, эта точка предельная, поскольку в любом шаре с центром в этой точке есть хотя бы одна точка множества X, отличная от a. Далее, точка b граничная ввиду того, что какой бы шар с центром в этой точке мы ни взяли, в нем есть как точки самого множества, так и точки дополнения. Такая точка, как и a, является предельной точкой множества X, но она не может быть ни внутренней, ни внешней, поскольку любой шар с центром в этой точке включает как точки из X, так и точки дополнения. Наконец, точка c внешняя: есть шар с центром в этой точке, полностью расположенный в дополнении множества X. Ясно, что она не может быть ни предельной для X, ни внутренней, ни граничной.

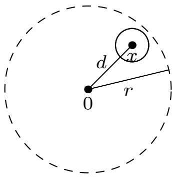

Рис. 1.9.

Пример 4. (1) Найдем внутренность, внешность и границу открытого шара $B_2(0,r)$ (рис. 1.9).

Ясно, что каждая точка этого шара является его внутренней точкой. Действительно, взяв точку $x \in B_2(0,r)$, находящуюся на расстоянии $d < r$ от начала координат, можно найти шар с центром в этой точке, целиком лежащий в $B_2(0,r)$. Например, можно взять любой шар радиуса, меньшего чем $r-d$ (см. рис. 1.9). Тем самым $\operatorname{int} B_2(0,r) = B_2(0,r)$.

Внешность этого шара состоит из всех точек x, находящихся на расстоянии $D > r$ от нуля. Ясно, что шар с центром в x и любым радиуса, меньшего чем $D-r$, свободен от точек исходного шара.

Точки на сфере $S_2(0,r) = \{x \in \mathbb{R}^2 \mid |x|_2 = r\}$ не могут быть ни внутренними, ни внешними. Они образуют границу шара.

Опишем замыкание $\operatorname{cl} B_2(0,r)$. Ясно, что точки самого множества всегда входят в его замыкание. Для получения замыкания к точкам множества надо добавить его предельные точки, не принадлежащие самому множеству. В нашем случае к шару надо добавить сферу как множество всех не принадлежащих открытому шару его предельных точек и в объединении получится, с одной стороны, замыкание открытого шара, а с другой — замкнутый шар. Тем самым замыкание открытого шара есть соответствующий ему замкнутый шар:

$$\operatorname{cl} B_2(0,r) = \overline{B}_2(0,r).$$

(2) Найдем внутренность, внешность и границу замкнутого шара $\overline{B}_2(0,r)$.

Этот шар отличается от открытого шара только сферой $S_2(0,r)$. У него внешность, внутренность и граница такие же, как и у открытого шара. Отличие открытого шара от замкнутого только в том, что открытый шар совпадал со своей внутренностью, а замкнутый совпадает со своим замыканием.

(3) Найдем внутренность, внешность и границу сферы $S_2(0,r)$.

Начнем с точек сферы. Можно ли точку из сферы окружить шаром с центром в этой точке, целиком состоящим из точек сферы? Ясно, что нельзя: в любом таком шаре есть как точки с $|x|_2 < r$, так и с $|x|_2 > r$, стало быть, никакая точка сферы не является ее внутренней точкой и ее внутренность пустая: $\operatorname{int} S_2(0,r) = \varnothing$. Дополнение сферы состоит из внешних точек: $\operatorname{ext} S_2(0,r) = \mathbb{R}^2 \setminus S_2(0,r)$. Наконец, каждая точка сферы является ее граничной точкой, ибо в любом шаре с центром в такой точке есть точки как сферы, так и ее дополнения, стало быть, $\operatorname{fr} S_2(0,r) = S_2(0,r)$. Таким образом, сфера состоит только из граничных точек. Ее замыкание совпадает со сферой: $\operatorname{cl} S_2(0,r) = S_2(0,r)$.

(4) Найдем внутренность, внешность и границу всего $\mathbb{R}^n$.

Очевидно, что $\operatorname{int} \mathbb{R}^n = \mathbb{R}^n$, $\operatorname{ext} \mathbb{R}^n = \varnothing$, $\operatorname{fr} \mathbb{R}^n = \varnothing$, $\operatorname{cl} \mathbb{R}^n = \mathbb{R}^n$.

(5) Найдем внутренность, внешность и границу множества $\mathbb{Q}^n$, где $\mathbb{Q}$ — множество рациональных чисел, т. е. множества упорядоченных наборов из n рациональных чисел. Таких точек довольно много, ибо, как известно, в любом открытом промежутке числовой прямой есть рациональное число. Тем не менее внутренность этого множества пустая: $\operatorname{int}\mathbb{Q}^n = \varnothing$. Это связано с тем, что в любом открытом промежутке числовой прямой есть не только рациональное, но и иррациональное число, так что в любом шаре с центром в точке с рациональными координатами есть точка, хотя бы одна координата которой иррациональна. Ясно, что и внешность его пустая. Стало быть, любая точка из $\mathbb{R}^n$. является его граничной точкой, так что $\operatorname{fr}\mathbb{Q}^n = \mathbb{R}^n$ и $\operatorname{cl}\mathbb{Q}^n = \mathbb{R}^n$.

Определение открытого и замкнутого множеств. Множество $X \subset \mathbb{R}^n$ называют *открытым*, если все его точки внутренние, т. е. если оно совпадает со своей внутренностью: $X = \operatorname{int} X$.

Множество $X \subset \mathbb{R}^n$ называют *замкнутым*, если его дополнение $\mathbb{R}^n \setminus X$ открыто.

Можно показать, что X замкнуто тогда и только тогда, когда оно совпадает со своим замыканием: $X = \operatorname{cl} X$. Этим обстоятельством обычно пользуются при рассмотрении замкнутых множеств.

Пример 5. (1) Открытый шар $B_2(0,r)$ является открытым множеством, ибо он совпадает со своей внутренностью.

(2) Замкнутый шар $B_2(0,r)$ является замкнутым множеством, так как он совпадает со своим замыканием.

(3) Сфера $S_2(0,r)$ является замкнутым множеством, поскольку она совпадает со своим замыканием.

(4) $\mathbb{R}^n$ замкнуто и открыто одновременно.

(5) $\mathbb{Q}^n$ не открыто и не замкнуто.

Определение ограниченного множества. Множество $X \subset \mathbb{R}^n$ называют *ограниченным*, если X содержится в некотором шаре.

Определение компактного множества. Множество $X \subset \mathbb{R}^n$ называют *компактным* (или *компактом*), если оно замкнуто и ограничено.

Понятие компактного множества есть в любом нормированном пространстве, но там оно определяется иначе. Данное здесь определение компакта корректно только для конечномерного векторного пространства $\mathbb{R}^n$.

Определение связного множества и области. Множество $X \subset \mathbb{R}^n$ называют *связным*, если для любых двух его точек есть непрерывная кривая, соединяющая эти точки и полностью лежащая в X. Открытое связное множество в $\mathbb{R}^n$ называют *областью*.

Пример 6. Шар $B(0,r)$, очевидно, множество ограниченное, но не компактное, поскольку не замкнутое. Оно связно, так как любые две его точки можно соединить гладкой кривой, например отрезком. Замкнутый шар $\overline{B}(0,r)$ есть множество ограниченное и замкнутое, стало быть, компактно. Это связное множество. Сфера $S(0,r)$ — множество ограниченное, замкнутое, так что компактное, и связное.

Ясно, что открытый шар — это область, тогда как замкнутый шар и сфера областями не являются.

1.4. Предел и непрерывность отображений арифметических конечномерных пространств.

Определение сходимости в $\mathbb{R}^n$. Говорят, что *последовательность векторов $x_k = (x_k^1, \ldots, x_k^n)$ сходится к вектору $a = (a^1, \ldots, a^n)$ при $k \to \infty$*, если

$$\forall \varepsilon > 0 \ \exists k_0 \in \mathbb{N} \ \forall k \geqslant k_0 \quad |x_k - a|_2 < \varepsilon,$$

при этом пишут $x_k \to a$ или $a = \lim x_k$.

Нетрудно заметить, что это определение практически совпадает с определением предела числовой последовательности с тем только отличием, что здесь вместо модуля разности чисел берется l_2-норма разности векторов. Впрочем, ввиду эквивалентности норм вместо l_2-нормы можно взять любую другую.

Замечание. Сходимость в $\mathbb{R}^n$ — это покомпонентная сходимость, т. е. последовательность $x_k = (x_k^1, \ldots, x_k^n)$ сходится к вектору $a = (a^1, \ldots, a^n)$ тогда и только тогда, когда каждая из компонент последовательности векторов сходится к соответствующей компоненте предельного вектора, т. е. $x_k^i \to a^i$ для каждого $i = 1, \ldots, n$ при $k \to \infty$.

Действительно, ввиду эквивалентности норм имеем

$$\begin{aligned} x_k \to a \iff |x_k - a|_2 \to 0 \iff |x_k - a|_\infty \to 0 \\ \iff \max_{i=1,\ldots,n} \left|x_k^i - a^i\right| \to 0 \iff \forall i = 1, \ldots, n \quad x_k^i \to a^i. \end{aligned}$$

Определение предела отображения. Пусть $f : X \to \mathbb{R}^m$ — отображение, определенное на множестве $X \subset \mathbb{R}^n$, и a — предельная точка множества X. Вектор $b = (b^1, \ldots, b^m)$ называют *пределом* *отображения* f *в точке* a и пишут

$$f(x) \to b \text{ при } x \to a \quad \text{или} \quad b = \lim_{x \to a} f(x),$$

если

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x \in X \quad 0 < |x - a|_2 < \delta \to |f(x) - b|_2 < \varepsilon.$$

Характеризующее предел высказывание по виду ничем не отличается от соответствующего высказывания для функции одной переменной, различие только в том, что здесь взаимная удаленность точек измеряется их нормой.

Замечание о сходимости координатных функций. Ввиду эквивалентности норм в конечномерном арифметическом пространстве сходимость отображения, действующего в $\mathbb{R}^m$, равносильна сходимости всех его координатных функций. Точнее, если

$$f(x) = (f^1(x), \ldots, f^m(x)),$$

то

$$b = \lim_{x \to a} f(x) \Longleftrightarrow \forall j = 1, \ldots, m \quad b^j = \lim_{x \to a} f^j(x).$$

Определение непрерывности отображения. Пусть $f : X \to \mathbb{R}^m$ — отображение, определенное на множестве $X \subset \mathbb{R}^n$, и $a \in X$. Говорят, что *отображение* f *непрерывно в точке* a, если

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x \in X \quad |x - a|_2 < \delta \to |f(x) - f(a)|_2 < \varepsilon$$

Говорят, что отображение f *непрерывно на множестве* X, если оно непрерывно в каждой точке этого множества.

Как и в случае предела, непрерывность отображения в точке равносильна непрерывности в этой точке всех его координатных функций.

Понятия связного и компактного множества обобщают на многомерный случай понятия промежутка и отрезка. Для функций одной переменной были теоремы, посвященные свойствам функций на соответствующих множествах. Они переносятся на многомерный случай с необходимыми изменениями в условиях.

Замечание о теореме о промежуточных значениях. В условиях теоремы о промежуточных значениях предполагалось, что функция задана на промежутке. Эта теорема доказывалась путем деления отрезка пополам и важно, что средняя точка любого отрезка с концами в данном промежутке попадет в исходное множество. Тем самым было важно, что любые две точки данного множества можно соединить лежащей в этом множестве линией, в одномерном случае отрезком. При переносе теоремы на многомерный случай надо предполагать, что область определения связна.

Замечание о теореме Вейерштрасса. В доказательстве этой теоремы мы пользовались тем, что область задания функции являлась замкнутым ограниченным промежутком. Вспомним детали доказательства. Мы строили последовательность точек промежутка и из нее выбирали подпоследовательность, сходящуюся непременно к элементу этого же множества. Гарантией такой возможности были ограниченность промежутка и его замкнутость. Эти свойства вместе в многомерном случае называют компактностью. Стало быть, при обобщении теоремы Вейерштрасса на многомерный случай надо предполагать компактность области определения.

Таким образом, теоремы о свойствах непрерывных функций в одномерном случае переносятся на многомерный, при этом в тех теоремах, где важны промежуточные значения, надо предполагать связность, а там, где важны замкнутость и ограниченность, — предполагать компактность.

§ 2. Линейные отображения

2.1. Определение линейного отображения.

Пусть X и Y — векторные пространства. Отображение $L : X \to Y$ называют *линейным отображением*, если

$$L(\alpha x_1 + \beta x_2) = \alpha L(x_1) + \beta L(x_2), \quad x_1, x_2 \in X,\ \alpha, \beta \in \mathbb{R}.$$

Значение отображения L на элементе x будем обозначать через $L(x)$ или, короче, Lx.

Если $Y = \mathbb{R}$, то линейное отображение называют *линейным функционалом.*

Пример 1 ($m \times n$-матрица). Пусть $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$. Матрица

$$A = \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \dots & a_n^m \end{pmatrix}$$

определяет линейный оператор из $\mathbb{R}^n$ в $\mathbb{R}^m$ по правилу

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \dots & a_n^m \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

или, в координатах,

$$y^j = \sum_{i=1}^{n} a_i^j x^i,$$

или в терминах отображений

$$y = Ax.$$

Легко проверить, что определяемое матрицей A отображение линейно.

Пример 2. В том случае, когда $m = 1$, умножение на матрицу $A = (a_i)$ можно представить как скалярное произведение

$$y = \begin{pmatrix} a_1 & \dots & a_n \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \sum_{i=1}^{n} a_i x^i = \langle a, x \rangle.$$

Тем самым скалярное умножение на фиксированный вектор $a \in \mathbb{R}^n$ — это линейный функционал на $\mathbb{R}^n$.

Пример 3. Пусть

$$X = Y = \mathscr{P}_n = \{p(x) = a_0 + a_1 x + \dots + a_n x^n \mid a_i \in \mathbb{R},\ i = 0, \dots, n\}$$

— множество полиномов степени не выше чем n. Это векторное пространство, ибо в результате сложения двух таких полиномов получается опять полином степени не выше чем n, и в результате умножение такого полинома на вещественное число получается полином из этого же класса. Рассмотрим отображение

$$\frac{d}{dx} : \mathscr{P}_n \to \mathscr{P}_n,$$

действующее по правилу

$$p(x) \mapsto \frac{dp(x)}{dx},$$

т. е.

$$a_0 + a_1 x + \dots + a_n x^n \mapsto a_1 + 2a_2 x + \dots + n a_n x^{n-1}.$$

Это отображение линейно ввиду того, что производная линейной комбинации двух функций есть линейная комбинация их производных.

Пример 4. На векторном пространстве $X = \mathscr{R}[a,b]$ интегрируемых на $[a,b]$ функций зададим отображение I, полагая

$$I(f) = \int\limits_a^b f(x)\,dx, \quad f \in \mathscr{R}[a,b].$$

Это линейный функционал, так как интеграл от линейной комбинации двух функций есть линейная комбинация их интегралов.

Пример 5. В $\mathbb{R}^n$ рассмотрим функцию

$$x \mapsto \langle x, x\rangle = \sum_{i=1}^n (x^i)^2,$$

сопоставляющую каждому вектору его скалярный квадрат. Будет ли это отображение линейным? Нет, оно не линейное, что легко понять хотя бы в случае $n = 1$. Но внимание на него мы обратим.

Немного изменим ситуацию и рассмотрим отображение

$$x \mapsto \sqrt{\langle x, x\rangle} = |x|_2.$$

Это отображение линейно или нет? Посмотрим в простейшей ситуации $n = 1$, тогда отображение выглядит так: $x \mapsto |x|$, и оно нелинейно, хотя на отдельных участках его график совпадает с графиком линейного отображения.

Кстати, если функция, отличная от тождественно нулевой, принимает только неотрицательные значения, то она не может быть линейной. В частности, норма никогда не бывает линейным функционалом.

Пример 6 (закон Ома). Если в пространстве $\mathbb{R}^3$ задано постоянное электрическое поле $\overrightarrow{E} = (E_x, E_y, E_z)$, то возникает ток с плотностью $\vec{j} = (j_x, j_y, j_z)$. Плотность тока пропорциональна вектору напряженности и в общем случае связана с полем следующим образом:

$$\begin{pmatrix} j_x \\ j_y \\ j_z \end{pmatrix} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix},$$

где σ_{ij}, $i, j = x, y, z$, — матрица проводимости. Чаще всего диагональные элементы этой матрицы равны между собой, а внедиагональные равны нулю, хотя бывают ситуации, в которых все элементы матрицы ненулевые.

2.2. Матрица линейного отображения.

Пусть X и Y — конечномерные векторные пространства. Пусть $e_1, \ldots, e_n$ — базис в X, т. е. набор линейно независимых векторов такой, что любой элемент $x \in X$ однозначно представляется в виде

$$x = \sum_{i=1}^{n} x^i e_i. \tag{$*$}$$

Тем самым каждому вектору x сопоставляется набор из n чисел x^i, $i = 1, \ldots, n$, состоящий из коэффициентов разложения этого вектора по базису, называемых *координатами вектора x относительно данного базиса.* Обратно, каждый набор из n чисел при заданном базисе определяет вектор по правилу $(*)$. Обратим внимание на то, что X может не быть арифметическим векторным пространством, т. е. не состоять само из упорядоченных наборов чисел, это могут быть, например, полиномы, и такое векторное пространство нам уже встречалось.

Аналогично пусть $f_1, \ldots, f_m$ — базис в Y и любой $y \in Y$ однозначно представим в виде

$$y = \sum_{j=1}^{m} y^j f_j.$$

Наконец, пусть $L : X \to Y$ — линейное отображение. Тогда действие отображения L можно задать матрицей.

Разберемся в том, как такое задание реализуется. Пусть $x \in X$ и

$$x = \sum_{i=1}^{n} x^i e_i.$$

Подействовав на x линейным отображением L, получим некоторый элемент $y = Lx$. Используя взаимно однозначное соответствие между элементами векторных пространств и их координатами, можно происходящее представить в виде диаграммы:

$$\begin{array}{ccc} x \in X & \longrightarrow & y = Lx \\ \updownarrow & & \updownarrow \\ (x^1, \ldots, x^n) & \longrightarrow & (y^1, \ldots, y^m) \end{array}.$$

Если первоначально было отображение L, т. е. действовала стрелка в верхней строке диаграммы, то появляется линейное отображение между координатами, отраженное в нижней строке. Его линейность обеспечивается тем, что каждое из отображений, определенных уже заданными на диаграмме стрелками, линейно. Оно задается некоторой матрицей, так что

$$y^j = \sum_{i=1}^{n} a_i^j x^i.$$

Матрица $\left(a_i^j\right)$ задается при заданных базисах следующим образом. Возьмем $x \in X$, разложим его по базису, подействуем на него отображением L и результат разложим по базису в Y:

$$\begin{aligned} y = L(x) = L\left(\sum_{i=1}^{n} x^i e_i\right) &= \sum_{i=1}^{n} x^i L(e_i) \\ = \sum_{i=1}^{n} x^i \sum_{j=1}^{m} (L(e_i))^j f_j &= \sum_{j=1}^{m} \sum_{i=1}^{n} (L(e_i))^j x^i f_j. \end{aligned}$$

С другой стороны, y раскладывается по базису:

$$y = \sum_{j=1}^{m} y^j f_j.$$

Ввиду однозначности разложения вектора по базису получаем равенство

$$y^j = \sum_{i=1}^{n} (L(e_i))^j x^i.$$

Получили правило, по которому координаты y^j вектора $y = Lx$ выражаются через координаты x^i вектора x. Следовательно, матрица $A = \left(a_i^j\right)$ линейного отображения L находится так:

$$a_i^j = (L(e_i))^j,$$

т. е. для нахождения элемента матрицы a_i^j, расположенного в i-м столбце и j-й строке, берем i-й базисный вектор, действуем на него отображением L, результат раскладываем по базису в Y и берем j-ю компоненту.

Пример 7 (матрица оператора дифференцирования $\dfrac{d}{dx} : \mathscr{P}^n \to \mathscr{P}_n$). Выберем базис в $\mathscr{P}_n$ следующим образом:

$$e_1 = 1,\ e_2 = x,\ \ldots,\ e_{n+1} = x^n.$$

Размерность пространства $\mathscr{P}_n$ равна $n+1$. Для нахождения матрицы оператора дифференцирования надо подействовать этим оператором на элемент $e_i = x^{i-1}$ и разложить результат по базису:

$$\begin{aligned}\frac{d}{dx}e_i &= \frac{d}{dx}x^{i-1} = (i-1)x^{i-2}\\ &= 0\cdot 1 + 0\cdot x + \dots + (i-1)x^{i-2} + \dots + 0\cdot x^n\\ &= 0\cdot e_1 + 0\cdot e_2 + \dots + (i-1)\cdot e_{i-1} + \dots + 0\cdot e_{n+1}.\end{aligned}$$

В итоге $a_i^{i-1} = i-1$, остальные элементы нулевые, и матрица имеет вид

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0\\ 0 & 0 & 2 & \dots & 0\\ \vdots & \vdots & 0 & \dots & 0\\ \hdotsfor{5}\\ 0 & 0 & 0 & \dots & n\\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

§ 3. Дифференцирование функций многих переменных

3.1. Дифференциал и частные производные.

В общем случае будем рассматривать отображения, определенные на некотором множестве $X \subset \mathbb{R}^n$ и действующие в $\mathbb{R}^m$, т. е. функции

$$f(x) = (f^1(x^1, \dots, x^n), \dots, f^m(x^1, \dots, x^n)).$$

Если функция действует в $\mathbb{R}^m$, мы имеем дело с m функциями, действующими в $\mathbb{R}$.

Примеры функций из $\mathbb{R}^n$ в $\mathbb{R}^m$.

(1) Кривая — это отображение из $[a,b] \subset \mathbb{R}$ в $\mathbb{R}^3$, т. е. отображение, которое каждому $t \in [a,b]$ сопоставляет три числа, и возникают три функции:

$$t \mapsto (x(t), y(t), z(t)).$$

(2) Пусть дана плоскость, определяемая уравнением $ax + by + cz + d = 0$, $a, b, c, d \in \mathbb{R}$. Тогда расстояние от точки (x, y, z) до этой плоскости определяется так:

$$\rho(x,y,z) = \frac{|ax + by + cz + d|}{\sqrt{a^2+b^2+c^2}}.$$

Это отображение из $\mathbb{R}^3$ в $\mathbb{R}$.

(3) Задание поля в пространстве, например электрического, означает, что каждой точке пространства $\mathbb{R}^3$ сопоставляется другая точка пространства $\mathbb{R}^3$. Появляется отображение из $\mathbb{R}^3$ в $\mathbb{R}^3$:

$$\overrightarrow{E}(x, y, z) = (E_x(x, y, z), E_y(x, y, z), E_z(x, y, z)).$$

(4) При параметрическом описании поверхности в $\mathbb{R}^3$ возникает отображение области $U \subset \mathbb{R}^2$ в $\mathbb{R}^3$. Например, на множестве $U = \{(x, y) \mid x^2 + y^2 < r^2\}$ зададим отображение

$$(x, y) \mapsto (x, y, \sqrt{r^2 - x^2 - y^2}).$$

Так описывается верхняя полусфера с центром в нуле радиуса r.

Определение дифференциала. Пусть $f : X \to \mathbb{R}^m$, где $X \subset \mathbb{R}^n$, и x — внутренняя точка множества X, т. е. она содержится в X вместе с некоторым открытым шаром $B(x, r)$, $r > 0$. Говорят, что f *дифференцируемо* или *имеет дифференциал в точке* x, если существует такое линейное отображение $L : \mathbb{R}^n \to \mathbb{R}^m$, что

$$f(x + h) = f(x) + L(h) + o(h) \quad \text{при } h \to 0,$$

где $o(h)$ — некоторая функция $\beta(h) : B(0, r) \to \mathbb{R}^m$ такая, что

$$|\beta(h)|_2 = \alpha(h)|h|_2, \quad \text{где } \alpha : B(0, r) \to \mathbb{R},\ \alpha(h) \to 0 \text{ при } h \to 0.$$

Линейное отображение L при этом называют *дифференциалом* f *в точке* x и обозначают через $df(x)$. Равенство в определении дифференциала понимается покомпонентно, т. е.

$$f^j(x + h) = f^j(x) + L^j(h) + \beta^j(h), \quad j = 1, \ldots, m.$$

Предположение о том, что точка x внутренняя, надо для возможности отступить от нее в произвольном направлении, оставаясь в пределах данного множества, и брать соответствующие приращения. Это требуется для обеспечения единственности дифференциала. Более того, в дальнейшем все связанные с дифференцируемостью рассмотрения будем зачастую проводить для областей, т. е. множеств открытых.

Замечание (дифференцируемая функция непрерывна). Пусть f имеет дифференциал в точке x. Тогда

$$f(x + h) = f(x) + L(h) + o(h) \quad \text{при } h \to 0.$$

Непрерывность функции в точке x означает, что

$$\lim_{h\to 0} f(x+h) = f(x).$$

Для дифференцируемой в точке x функции это свойство выполнено, ибо

$$|f(x+h) - f(x)| \leqslant |L(h)| + |o(h)|$$

и $\lim\limits_{h\to 0} |L(h)| = 0$ как линейное отображение, равное нулю в нуле, а $\lim\limits_{h\to 0} |o(h)| = 0$ по определению o-малого, так что правая часть стремится к нулю и левая также стремится к нулю.

Определение частной производной. Пусть $f : X \to \mathbb{R}^m$, $x \in X$ — внутренняя точка множества X. Если существует предел

$$\begin{aligned} \frac{\partial f^j}{\partial x^i}(x) &= \lim_{t\to 0} \frac{f^j(x^1, \dots, x^i + t, \dots, x^n) - f^j(x^1, \dots, x^i, \dots x^n)}{t} \\ &= \lim_{t\to 0} \frac{f^j(x + te_i) - f^j(x)}{t}, \end{aligned}$$

то его называют *частной производной компоненты* f^j *по* x^i *в точке* x. Здесь $e_i = (0, \dots, 0, 1, 0, \dots, 0)$, где единица находится на i-м месте, — векторы стандартного базиса в $\mathbb{R}^n$. По существу, частная производная — это производная обычной функции, действующей из $\mathbb{R}$, которая получается, если фиксировать все компоненты аргумента, кроме одной, а по этой компоненте рассматривать обычную производную.

Матрицу

$$\frac{\partial f}{\partial x}(x) = \begin{pmatrix} \frac{\partial f^1}{\partial x^1}(x) & \dots & \frac{\partial f^1}{\partial x^n}(x) \\ \vdots & & \vdots \\ \frac{\partial f^m}{\partial x^1}(x) & \dots & \frac{\partial f^m}{\partial x^n}(x) \end{pmatrix},$$

составленную из частных производных, называют *матрицей Якóби отображения* f *в точке* x. Символ $\dfrac{\partial f}{\partial x}(x)$ воспринимается как цельный.

Теорема 4 (о связи дифференциала и частных производных). *Пусть* $f : X \to \mathbb{R}^m$, $x \in X$ *— внутренняя точка множества* X.

(1) *Если* f *имеет дифференциал в точке* x, *то все компоненты* f^j *имеют все частные производные* $\dfrac{\partial f^j}{\partial x^i}(x)$.

(2) *Если f имеет все частные производные $\dfrac{\partial f^j}{\partial x^i}$ в некоторой окрестности U точки x и они непрерывны в точке x, то f имеет дифференциал в точке x.*

При этом матрица дифференциала $df(x)$ как линейного отображения относительно стандартного базиса равна матрице Якоби, т. е.

$$df^j(x)(h) = \sum_{i=1}^{n} \frac{\partial f^j}{\partial x^i}(x)h^i.$$

Доказательство. (1) По определению дифференциала имеет место равенство

$$f^j(x+h) - f^j(x) = df^j(x)(h) + \beta^j(h),$$

где $\beta(h) = (\beta^1(h), \ldots, \beta^m(h)) = o(h)$. Возьмем в качестве h приращение в направлении e_i. Тогда

$$\lim_{t\to 0} \frac{f^j(x+te_i) - f^j(x)}{t} = \lim_{t\to 0}\left(\frac{df^j(x)(te_i)}{t} + \frac{\beta^j(te_i)}{t}\right) = df^j(x)(e_i), \tag{$*$}$$

поскольку ввиду линейности дифференциала множитель t можно вынести, так что

$$\frac{df^j(x)(te_i)}{t} = df^j(x)(e_i),$$

а

$$\lim_{t\to 0}\frac{\beta^j(te_i)}{t} = 0.$$

Действительно,

$$|\beta(h)|_2 = \alpha(h)|h|_2,$$

где $\alpha(h) : \mathbb{R}^n \to \mathbb{R}$ и $\alpha(h) \to 0$ при $h \to 0$. При $h = te_i$ получим

$$|\beta(te_i)|_2 = \alpha(te_i)|te_i|_2 = \alpha(te_i)|t|,$$

стало быть,

$$\left|\frac{\beta^j(te_i)}{t}\right| \leqslant \frac{|\beta(te_i)|_2}{|t|} = \alpha(te_i) \to 0.$$

Равенство $(*)$ означает, что

$$\frac{\partial f^j}{\partial x^i}(x) = df^j(x)(e_i),$$

где справа — элемент матрицы линейного оператора, а в слева — элемент матрицы Якоби. Это означает, что матрица линейного оператора $df(x)$ равна матрице Якоби $\dfrac{\partial f}{\partial x}(x)$.

(2) Докажем теорему в случае $m = 1$, $n = 2$, общий случай отличается только количеством выкладок. Поскольку точка x внутренняя, есть некоторый шар $B(x, r)$, содержащийся в X. Будем брать приращения, не выводящие за пределы этого шара, т. е. такие $h = (h^1, h^2)$, что $h \in B(0, r)$. Тогда $x + h \in B(x, r)$. Рассмотрим приращение функции

$$\begin{aligned}&f(x^1 + h^1, x^2 + h^2) - f(x^1, x^2)\\&\quad = f(x^1 + h^1, x^2 + h^2) - f(x^1 + h^1, x^2) + f(x^1 + h^1, x^2) - f(x^1, x^2).\end{aligned}$$

По теореме Лагранжа существуют числа $\theta^1(h)$ и $\theta^2(h)$ такие, что $0 < \theta^1(h) < 1$, $0 < \theta^2(h) < 1$ и

$$f(x^1 + h^1, x^2 + h^2) - f(x^1 + h^1, x^2) = \frac{\partial f}{\partial x^2}(x^1 + h^1, x^2 + \theta^2(h)h^2)h^2,$$
$$f(x^1 + h^1, x^2) - f(x^1, x^2) = \frac{\partial f}{\partial x^1}(x^1 + \theta^1(h)h^1, x^2)h^1.$$

Ввиду непрерывности частных производных $\dfrac{\partial f}{\partial x^1}$, $\dfrac{\partial f}{\partial x^2}$ в точке x получаем

$$\frac{\partial f}{\partial x^2}(x + h^1, x^2 + \theta^2(h)h^2) = \frac{\partial f}{\partial x^2}(x^1, x^2) + \alpha^2(h)$$
$$\frac{\partial f}{\partial x^1}(x + \theta^1(h)h^1, x^2) = \frac{\partial f}{\partial x^1}(x^1, x^2) + \alpha^1(h),$$

где $\alpha^1(h), \alpha^2(h) \to 0$ при $h \to 0$ ввиду непрерывности производных.

Мы получили, что приращение функции представимо в виде

$$\begin{aligned}&f(x^1 + h^1, x^2 + h^2) - f(x^1, x^2)\\&\quad = \underbrace{\frac{\partial f}{\partial x^1}(x^1, x^2)h^1 + \frac{\partial f}{\partial x^2}(x^1, x^2)h^2}_{L(h)} + \underbrace{\alpha^1(h)h^1 + \alpha^2(h)h^2}_{\beta(h)}.\end{aligned}$$

В итоге получилось линейное отображение $L(h)$ и остаток $\beta(h)$. Докажем, что $\beta(h) = o(h)$. Действительно, по неравенству Коши — Буняковского имеем

$$|\beta(h)| = \left|\sum_{i=1}^{2} \alpha^i(h)h^i\right| \leqslant \underbrace{\left(\sum_{i=1}^{2} |\alpha^i(h)|^2\right)^{1/2}}_{\to 0} \underbrace{\left(\sum_{i=1}^{2} |h^i|^2\right)^{1/2}}_{|h|_2},$$

и первый множитель в правой части стремится к нулю, а второй есть не что иное как l_2-норма вектора приращений. Следовательно,

$$0 \leqslant \frac{|\beta(h)|}{|h|_2} \leqslant \left(\sum_{i=1}^{2} |\alpha^i(h)|^2\right)^{1/2} \to 0,$$

стало быть,

$$|\beta(h)| = \gamma(h)|h|_2, \quad \text{где } \gamma(h) \to 0 \text{ при } h \to 0.$$

Мы доказали, что f дифференцируемо в точке x и заодно выяснили, как действует дифференциал.

Теорема доказана.

Пример 1. Предположение непрерывности частных производных в п. (2) существенно. Приведем пример функции, у которой частные производные есть, а дифференциала нет. Рассмотрим функцию

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

Вдоль координатных осей функция равна нулю, т. е. $f(x,0) = 0$, $f(0,y) = 0$, стало быть, $\dfrac{\partial f}{\partial x}(0,0) = 0$, $\dfrac{\partial f}{\partial y}(0,0) = 0$. Однако эта функция недифференцируема, более того, она разрывна в нуле. Действительно, если входить в точку $(0,0)$ по разным направлениям, то получатся разные пределы в нуле. Если стремиться к началу координат по прямой $y = x$, то значения функции равны $\dfrac{1}{2}$, а если по прямой $y = -x$, то $-\dfrac{1}{2}$. Различие в получаемых пределах говорит о том, что предела функции f в точке $(0,0)$ нет. Тем более нет дифференцируемости этой функции, ибо всякая дифференцируемая в точке функция непрерывна в этой точке.

В одномерном случае была информация о том, как понимают дифференциал физики и математики. Повторим ее содержание для многомерного случая, т. е. ответим на вопрос: что такое dx^i?

Рассмотрим функцию

$$(x^1, \ldots, x^n) \mapsto x^i,$$

сопоставляющую каждому вектору его i-ю координату. Она определена на $\mathbb{R}^n$, действует в $\mathbb{R}$ и линейна. Естественно, хотя и не очень корректно, обозначить ее через x^i, тем самым используя один и тот же символ для обозначения координаты и функции, проецирующей на эту координату. Поскольку эта функция линейна, ее дифференциал действует по тому же правилу, т. е. для любого $x \in \mathbb{R}^n$

$$dx^i(x)(h^1, \ldots, h^n) = h^i.$$

Так как от точки x здесь ничего не зависит, о ней можно не упоминать и писать

$$dx^i(h^1, \ldots, h^n) = h^i.$$

Получился набор $dx^1, \ldots, dx^n$ из n простейших линейных отображений. Согласно теореме 4 дифференциал функции выражается через dx^i:

$$df(x)(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x) h^i = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x) dx^i(h).$$

Последнее равенство понимается в смысле совпадения значений функций. Если перейти к равенству функций, опуская аргумент h, то можно написать

$$df(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x) dx^i.$$

Это стандартная запись дифференциала функции n переменных.

Обсудим, как воспринимают дифференциал физики. Если $\Delta x = (\Delta x^1, \ldots, \Delta x^n)$ — произвольное приращение, то приращение функции, равное

$$\Delta f(x) = f(x + \Delta x) - f(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x) \Delta x^i + o(\Delta x),$$

выражается через приращение аргумента и частные производные с некоторой ошибкой. Если приращения Δx^i настолько малы, что ошибкой $o(\Delta x)$ можно пренебречь, то такие приращение называют *бесконечно малыми* и обозначают через dx^i. Тогда можно сказать, что дифференциал функции — это бесконечно малое приращение функции $df(x)$, вызванное бесконечно малыми приращениями аргумента dx и выражающееся через них и частные производные по формуле

$$df(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x)\, dx^i. \qquad (*)$$

Тем самым у физиков дифференциал — это бесконечно малое приращение функции.

Если необходимо записать дифференциал функции, то независимо от его восприятия надо записать равенство вида $(*)$.

Пример 2. Найдем дифференциал функции

$$f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$$

в точке $(1, 2, 3)$.

Первым делом посчитаем частные производные:

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

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

Запишем формулу для дифференциала:

$$\begin{aligned} df(x, y, z) &= \frac{\partial f}{\partial x}(x, y, z)\, dx + \frac{\partial f}{\partial y}(x, y, z)\, dy + \frac{\partial f}{\partial z}(x, y, z)\, dz \\ &= \frac{x}{\sqrt{x^2 + y^2 + z^2}}\, dx + \frac{y}{\sqrt{x^2 + y^2 + z^2}}\, dy + \frac{z}{\sqrt{x^2 + y^2 + z^2}}\, dz, \end{aligned}$$

и, полагая $(x, y, z) = (1, 2, 3)$, получаем

$$df(1, 2, 3) = \frac{1}{\sqrt{14}}\, dx + \frac{2}{\sqrt{14}}\, dy + \frac{3}{\sqrt{14}}\, dz.$$

Ответим на вопрос: если мы находимся в точке $(1, 2, 3)$, то изменение какого из аргументов вызовет наибольший рост функции? Иначе говоря, встав в эту точку и перемещаясь по направлениям координатных осей, будем следить за тем, насколько быстро изменяется функция в каждом из направлений. В каком направлении надо двигаться, чтобы изменение было самым большим? Простое голосование в аудитории показывает, что двигаться надо вдоль оси z.

Теорема 5 (о дифференцировании и алгебраических операциях). *Пусть $f, g : X \to \mathbb{R}$, $X \subset \mathbb{R}^n$, $x \in X$ — внутренняя точка множества X. Предположим, что функции f и g дифференцируемы в точке x. Тогда их сумма, произведение и отношение (последнее при условии $g(x) \neq 0$) также дифференцируемы в точке x и имеют место формулы:*

для суммы

$$d(f + g)(x) = df(x) + dg(x)$$

или в терминах частных производных

$$\frac{\partial(f+g)}{\partial x^i}(x) = \frac{\partial f}{\partial x^i}(x) + \frac{\partial g}{\partial x^i}(x),$$

для произведения

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

или в терминах частных производных

$$\frac{\partial(fg)}{\partial x^i}(x) = \frac{\partial f}{\partial x^i}(x)g(x) + f(x)\frac{\partial g}{\partial x^i}(x),$$

для частного

$$d\left(\frac{f}{g}\right)(x) = \frac{df(x)g(x) - f(x)dg(x)}{g^2(x)}, \quad g(x) \neq 0,$$

или в терминах частных производных

$$\frac{\partial\left(\frac{f}{g}\right)}{\partial x^i}(x) = \frac{\frac{\partial f}{\partial x^i}(x)g(x) - f(x)\frac{\partial g}{\partial x^i}(x)}{g^2(x)}.$$

Доказательство проводится аналогично тому, как это делается в одномерное случае, и будет опущено.

Теорема 6 (о дифференцировании композиции). *Пусть*

(1) *функция $f : \Omega \to \mathbb{R}^m$, где Ω — область в $\mathbb{R}^n$, дифференцируема в точке $x \in \Omega$,*

(2) *функция $g : U \to \mathbb{R}^k$, где U — область в $\mathbb{R}^m$, содержащая образ $f(\Omega)$ области Ω, дифференцируема в точке $y = f(x)$ (рис. 3.1).*

Тогда композиция $g \circ f$ дифференцируема в точке x и дифференциал композиции равен композиции дифференциалов:

$$d(g \circ f)(x) = dg(f(x)) \circ df(x). \tag{$*$}$$

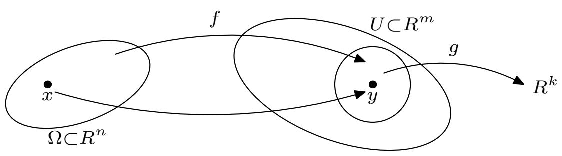

Рис. 3.1.

Матрица Якоби композиции — это произведение матриц Якоби каждого из отображений f, g, так что $(*)$ в матричном виде запишется так:

$$\frac{\partial(g \circ f)}{\partial x}(x) = \frac{\partial g}{\partial y}(f(x))\frac{\partial f}{\partial x}(x),$$

или, подробнее, в терминах частных производных:

$$\frac{\partial(g \circ f)^l}{\partial x^i}(x) = \sum_{j=1}^{m} \frac{\partial g^l}{\partial y^j}(f(x))\frac{\partial f^j}{\partial x^i}(x), \quad i = 1, \dots, n, \ l = 1, \dots, k. \qquad (\dagger)$$

Равенство $(\dagger)$ называют *цепным правилом* и регулярно применяют на практике.

Доказательство. Будем действовать по определению. Тот факт, что функция f имеет дифференциал в точке x, означает, что

$$f^j(x+h) = f^j(x) + df^j(x)(h) + \alpha^j(h), \qquad (1)$$

где $\alpha(h) = (\alpha^1(h), \dots, \alpha^m(h)) = o(h)$ в следующем смысле:

$$|\alpha(h)|_2 = \left(\sum_{j=1}^{m} |\alpha^j(h)|^2\right)^{1/2} = \varphi(h)|h|_2, \quad \text{где } \varphi(h) \to 0 \text{ при } h \to 0. \qquad (2)$$

Аналогично для g в точке y имеем

$$g^l(y+t) = g^l(y) + dg^l(y)(t) + \beta^l(t), \text{ где } \beta(t) = (\beta^1(t), \dots, \beta_k(t)) = o(t). \qquad (3)$$

Здесь x, h, y, t — векторы.

В формуле (3) выберем t так, что $y + t = f(x+h)$. Поскольку $y = f(x)$, надо взять $t = f(x+h) - f(x)$. Имеем

$$\begin{aligned} g^l(f(x+h)) &= g^l(f(x)) + dg^l(f(x))(f(x+h) - f(x)) \\ &\qquad + \beta^l(f(x+h) - f(x)) \\ &= g^l(f(x)) + dg^l(f(x))(df(x)(h) + \alpha(h)) \\ &\qquad + \beta^l(df(x)(h) + \alpha(h)). \end{aligned}$$

Следовательно,

$$\begin{aligned} g^l(f(x+h)) = g^l(f(x)) &+ dg^l(f(x))(df(x)(h)) \\ &+ dg^l(f(x))(\alpha(h)) + \beta^l(df(x)(h) + \alpha(h)). \end{aligned}$$

Получили формулу из определения дифференциала, в которой слагаемое $dg^l(f(x))(df(x)(h))$ представляет собой линейное отображение, являющееся композицией линейных отображений. Тем самым осталось доказать, что сумма

$$dg^l(f(x))(\alpha(h)) + \beta^l(df(x)(h) + \alpha(h))$$

бесконечно малая более высокого порядка, чем h. Действительно, проверим это свойство для первого слагаемого. Для второго можно доказать аналогично, но с бóльшими техническими трудностями.

Докажем, что

$$dg(f(x))(\alpha(h)) = o(h),$$

т. е.

$$|dg(f(x))(\alpha(h))|_2 = \left(\sum_{l=1}^{k}(dg^l(y)(\alpha(h)))^2\right)^{1/2} = b(h)|h|_2,$$

где $b(h) \to 0$ при $h \to 0$ или, иначе говоря, что

$$\frac{|dg(f(x))(\alpha(h))|_2}{|h|_2} \to 0.$$

Оценим l-ю компоненту сверху:

$$\begin{aligned} |dg^l(y)(\alpha(h))| &= \left|\sum_{j=1}^{m}\frac{\partial g^l(y)}{\partial y^j}\alpha^j(h)\right| \\ &\qquad\qquad \text{(по неравенству Коши — Буняковского)} \\ &\leqslant \left(\sum_{j=1}^{m}\left(\frac{\partial g^l(y)}{\partial y^j}\right)^2\right)^{1/2}\left(\sum_{j=1}^{m}(\alpha^j(h))^2\right)^{1/2} \\ &\leqslant C|\alpha(h)|_2 = C\varphi(h)|h|_2, \end{aligned} \tag{4}$$

где C — максимальное из всех чисел $\left|\dfrac{\partial g^l(y)}{\partial y^j}\right|^2$ по всевозможным индексам j, l.

Мы получили оценку отдельной компоненты, а надо было оценить l_2-норму. Воспользовавшись теоремой 1, позволяющей оценивать одну норму через другую, из (4) имеем

$$\begin{aligned} 0 \leqslant \frac{|dg(y)(\alpha(h))|_2}{|h|_2} &\leqslant \frac{\sqrt{k}|dg(y)(\alpha(h))|_\infty}{|h|_2} \\ &\leqslant \frac{\sqrt{k}C\varphi(h)|h|_2}{|h|_2} = \sqrt{k}C\varphi(h) \to 0. \end{aligned}$$

Тем самым доказано, что

$$|dg(y)(\alpha(h))|_2 = o(h).$$

Первая часть теоремы доказана.

Запишем правило (∗) в терминах частных производных. По теореме 4 для функции f имеем

$$df^j(x)(h) = \sum_{i=1}^{n} \frac{\partial f^j}{\partial x^i}(x)h^i$$

и аналогично для функции g

$$dg^l(y)(t) = \sum_{j=1}^{m} \frac{\partial g^l}{\partial y^j} t^j.$$

С одной стороны,

$$\begin{aligned} d(g \circ f)^l(x)(h) &= dg^l(y)(df(x)(h)) = \sum_{j=1}^{m} \frac{\partial g^l}{\partial y^j}(y) \sum_{i=1}^{n} \frac{\partial f^j}{\partial x^i}(x)h^i \\ &= \sum_{i=1}^{n} \left(\sum_{j=1}^{m} \frac{\partial g^l}{\partial y^j}(y) \frac{\partial f^j}{\partial x^i}(x) \right) h^i. \end{aligned} \tag{5}$$

С другой стороны, по теореме 4 должно быть

$$d(g \circ f)^l(x)(h) = \sum_{i=1}^{n} \frac{\partial (g \circ f)^l}{\partial x^i}(x)h^i. \tag{6}$$

Сопоставляя (5) и (6) и замечая, что правые части в них равны и приращение h^i произвольно, получаем

$$\frac{\partial (g \circ f)^l}{\partial x^i}(x) = \sum_{j=1}^{m} \frac{\partial g^l}{\partial y^j}(y) \frac{\partial f^j}{\partial x^i}.(x)$$

Теорема доказана.

3.2. Производная по вектору и по направлению. Градиент.

Определение производной по вектору. Пусть $f : X \to \mathbb{R}$, $X \subset \mathbb{R}^n$, $x \in X$ — внутренняя точка множества X и $v \in \mathbb{R}^n$ — некоторый вектор. *Производной по вектору* v называют предел (разумеется, если он существует)

$$D_v f(x) = \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t}.$$

Если вектор v единичной длины, то производную по нему называют *производной по направлению вектора* v.

Суть определения производной по вектору такова: мы проходим через точку x вдоль вектора v со скоростью, задаваемой этим вектором, и изучаем скорость изменения функции.

Производная по вектору связана с дифференциалом следующим равенством:

$$D_v f(x) = df(x)(v),$$

т. е. производная по вектору равна значению дифференциала на этом векторе. Действительно, для дифференцируемой в точке x функции f ввиду линейности дифференциала имеем

$$\begin{aligned}\frac{f(x+tv)-f(x)}{t} &= \frac{df(x)(tv)+o(tv)}{t} \\ &= df(x)(v) + \frac{o(tv)}{t} \to df(x)(v), \quad t \to 0.\end{aligned}$$

откуда и приходим к требуемому.

Можно сказать, что частная производная $\dfrac{\partial f}{\partial x^i}(x)$ — это производная по вектору $v = e_i$ или, иначе говоря, по направлению соответствующего базисного вектора.

Для дифференцируемой в точке x функции f и вектора v имеет место равенство

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

согласно которому производную по вектору можно найти через частные производные.

Определение градиента. Пусть $f : \Omega \to \mathbb{R}$ — функция, заданная в области $\Omega \subset \mathbb{R}^n$ и имеющая в этой области непрерывные частные производные $\dfrac{\partial f}{\partial x^i}(x)$. *Градиентом функции* f *в точке* $x \in \Omega$ называют вектор, обозначаемый через $\operatorname{grad} f(x)$, который однозначно определяется равенством

$$df(x)(h) = \langle h, \operatorname{grad} f(x) \rangle, \quad h \in \mathbb{R}^n \qquad (*)$$

где угловые скобки означают скалярное произведение.

Напомним факт из курса линейной алгебры. Если есть векторное пространство, на котором задано скалярное произведение, то действие линейного функционала на элемент h выражается в виде скалярного произведения h на некоторый вектор. Конкретнее, пусть L — линейный функционал на $\mathbb{R}^n$. тогда существует единственный вектор $f = (f^1, \dots, f^1) \in \mathbb{R}^n$ такой, что

$$Lh = \langle h, f \rangle = \sum_{i=1}^{n} h^i f^i, \quad h \in \mathbb{R}^n. \tag{$*$}$$

Рассматривая в качестве линейного функционала дифференциал, можно утверждать, что градиент функции f в точке x — это тот единственный вектор $\operatorname{grad} f(x) \in \mathbb{R}^n$, для которого действие дифференциала на приращение h выражается в виде скалярного произведения $(*)$. Этот вектор, естественно, зависит от функции, точки и скалярного произведения. Пока мы рассматриваем только стандартное скалярное произведение

$$\langle a, b \rangle = \sum_{i=1}^{n} a^i b^i.$$

Относительно него градиент записывается в векторном и координатном видах так:

$$\operatorname{grad} f(x) = \frac{\partial f}{\partial x^1}(x) e_1 + \dots + \frac{\partial f}{\partial x^n}(x) e_n = \left(\frac{\partial f}{\partial x^1}(x), \dots, \frac{\partial f}{\partial x^n}(x) \right).$$

В общем случае вид градиента иной.

Геометрический смысл градиента. Градиент указывает направление наибольшего роста функции. Рассмотрим функцию и ее производную по различным направлениям $h \in \mathbb{R}^n$, $|h|_2 = 1$, в фиксированной точке $x \in \Omega$. В случае функции двух переменных прохождение области Ω через точку x в направлении h означает разрезание графика функции f вертикальной плоскостью, проходящей через x в направлении h и нахождение скорости изменения функции одной переменной, получающейся в результате сечения графика плоскостью. Ясно, что в разных направлениях изменение функции разное. Среди всех направлений есть такое, в котором функция имеет наибольшую скорость роста. Это и есть направление градиента. Действительно, запишем производную функции f в направлении h:

$$df(x)(h) = \langle h, \operatorname{grad} f(x) \rangle = |h|_2 |\operatorname{grad} f(x)|_2 \cos\alpha,$$

где α — угол между h и $\operatorname{grad} f(x)$. Это значение будет наибольшим в том случае, если $\cos\alpha = 1$, т. е. $\alpha = 0$ и h смотрит в направлении градиента, так что

$$h = \frac{\operatorname{grad} f(x)}{|\operatorname{grad} f(x)|_2}.$$

Градиент имеет еще одну полезную геометрическую интерпретацию, связанную с множествами уровня. Поясним это на примере функции двух переменных для простоты, хотя все наблюдения годны для произвольной размерности.

Пусть $f : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^2$, — достаточно гладкая функция. Рассмотрим множество M, называемое в общем случае *множеством уровня*, в $\mathbb{R}^2$ — *линией уровня* и в $\mathbb{R}^3$ — *поверхностью уровня*, определяемое так:

$$M = \{x \in \Omega \mid f(x) = C\},$$

где C — некоторая константа. Пусть $x_0 \in M$. Ясно, что если мы пойдем через точку x_0 по линии уровня, то функция не изменяется, если сойдем с этой линии, то она может возрастать или убывать, причем в разных направлениях движения с разными скоростями. Оказывается, что направление градиента указывает направление наибольшей скорости роста функции и перпендикулярно линии уровня, проходящей через данную точку.

§ 4. Старшие производные и формула Тейлора

Определение старших производных. Пусть $f : \Omega \to \mathbb{R}$, где Ω — область в $\mathbb{R}^n$, и $x \in \Omega$ — некоторая точка. Предположим, что f имеет частную производную $\dfrac{\partial f}{\partial x^i}(x)$ в каждой точке $x \in \Omega$. Если $\dfrac{\partial f}{\partial x^i}$ имеет частную производную $\dfrac{\partial}{\partial x^j}\dfrac{\partial f}{\partial x^i}(x)$, то говорят, что f имеет вторую частную производную $\dfrac{\partial^2 f}{\partial x^j \partial x^i}(x)$.

По индукции можно определить производные любого порядка k: если производная $\dfrac{\partial^{k-1} f}{\partial x^{i_{k-1}} \dots \partial x^{i_1}}(x)$ существует в области Ω и имеет частную производную по x^{i_k}, то говорят, что f *имеет частную производную*

$$\frac{\partial^k f}{\partial x^{i_k} \dots \partial x^{i_1}}(x) = \frac{\partial}{\partial x^{i_k}} \left(\frac{\partial^{k-1} f}{\partial x^{i_{k-1}} \dots \partial x^{i_1}}(x) \right)$$

порядка k по переменным $x^{i_1}, \dots, x^{i_k}$ в точке x.

В частных производных выше первого порядка, вообще говоря, существенно, в каком порядке следования переменных берутся производные. Однако в благоприятных условиях, которые практически всегда выполнены, порядок следования переменных неважен.

Теорема 7 (о перестановочности частных производных).

(1) *Пусть функция* $f : \Omega \to \mathbb{R}$, *где* Ω *— область в* $\mathbb{R}^n$, *имеет в каждой точке из* Ω *частные производные* $\dfrac{\partial^2 f}{\partial x^i \partial x^j}$ *и* $\dfrac{\partial^2 f}{\partial x^j \partial x^i}$ *и они непрерывны в точке* x. *Тогда они совпадают*:

$$\frac{\partial^2 f}{\partial x^i \partial x^j}(x) = \frac{\partial^2 f}{\partial x^j \partial x^i}(x).$$

(2) *Если функция* $f : \Omega \to \mathbb{R}$, *где* Ω *— область в* $\mathbb{R}^n$, *имеет в каждой точке из* Ω *частные производные до порядка* k *включительно, то значение производной* $\dfrac{\partial^k f}{\partial x^{i_1} \dots \partial x^{i_k}}(x)$ *не зависит от порядка следования индексов* $i_1, \dots, i_k$.

Эта теорема, доказательство которой опустим, доказывается в стиле доказательства теоремы 4 на основе теоремы Лагранжа.

Определение класса $C^k(\Omega)$. Пусть Ω — область в $\mathbb{R}^n$ и k — целое неотрицательное число. *Класс* $C^k(\Omega)$ состоит из всех k раз непрерывно дифференцируемых функций $f : \Omega \to \mathbb{R}$, т. е. функций, у которых есть все частные производные до порядка k и все они непрерывны на Ω.

В случае $k = 0$ класс $C^0(\Omega)$ — это множество всех непрерывных на Ω функций. Функции из $C^1(\Omega)$ будем называть *гладкими*.

Гладкие функции, имея непрерывные частные производные во всех точках области, имеют дифференциал в каждой точке. Для функции класса $C^k(\Omega)$, $k \geqslant 2$, можно считать старшие частные производные, не задумываясь о порядке следования переменных.

Поскольку ниже мы часто будем иметь дело с частными производными, позаботимся о их более лаконичном обозначении. По мере необходимости вместо $\dfrac{\partial}{\partial x^i}$ будем писать просто ∂_i. Кроме того, нередко будем использовать выражение

$$(h^1 \partial_1 + \dots + h^n \partial_n) f(x),$$

понимая это так: при возведении в степень k, т. е. при раскрытии скобок, переменные h с соответствующими индексами просто перемножаются, а применительно к символам ∂_i рассматривается производная по соответствующим переменным, взятая от функции f в точке x. Например, для функции $f(x^1, x^2)$ и переменных h^1, h^2

$$(h^1 \partial_1 + h^2 \partial_2)^2 f(x) = \frac{\partial^2 f}{\partial (x^1)^2}(h^1)^2 + 2\frac{\partial^2 f}{\partial x^1 x^2} h^1 h^2 + \frac{\partial^2 f}{\partial (x^2)^2}(h^2)^2, \text{ и т. п.}$$

Теорема 8 (формула Тейлора). *Пусть $f : \Omega \to X$ — функция, Ω — область в $\mathbb{R}^n$, $x \in \Omega$ — фиксированная точка и $h \in \mathbb{R}^n$ — вектор приращения такой, что отрезок с концами x, $x + h$ содержится в Ω. Предположим, что $f \in C^{m+1}(\Omega)$. Тогда имеет место формула Тейлора*

$$f(x+h) = \sum_{k=0}^{m} \frac{(h^1\partial_1 + \dots + h^n\partial_n)^k f(x)}{k!} + r_m(x,h), \qquad (*)$$

где $r_n(x,h)$ — остаточный член, который можно записать в интегральном виде:

$$r_m(x,h) = \frac{1}{m!}\int_0^1 (1-t)^m (h^1\partial_1 + \dots + h^n\partial_n)^{m+1} f(x+th)\,dt,$$

или в виде Пеано:

$$r_m(x,h) = o(|h|^m) \quad \text{при } h \to 0,$$

или в форме Лагранжа:

$$r_m(x,h) = \frac{(h^1\partial_1 + \dots + h^n\partial_n)^{m+1} f(x+\theta h)}{(m+1)!},$$

где $\theta \in (0,1)$ — некоторое число, зависящее от x и h.

Доказательство. Фиксируем x и h и рассмотрим функцию одной переменной $\varphi(t) = f(x+th)$, $t \in [0,1]$. Функция $\varphi(t)$ имеет непрерывные производные до порядка $m+1$, что вытекает из предположения о гладкости исходного отображения. Для функции φ запишем формулу Тейлора с интегральным остаточным членом в окрестности нуля:

$$\varphi(\tau) = \sum_{k=0}^{m} \frac{\varphi^{(k)}(0)}{k!}\tau^k + \frac{1}{m!}\int_0^{\tau} (\tau-u)^m \varphi^{(m+1)}(u)\,du. \qquad (\dagger)$$

Формула Тейлора для функции многих переменных получится из формулы ($\dagger$). Для этого надо понять, как производные $\varphi^{(k)}(t)$ функции φ выражаются через частные производные исходной функции. Для первой производной по цепному правилу имеем

$$\varphi'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x+th)h^i = (h^1\partial_1 + \dots + h^n\partial_n) f(x+th).$$

При нахождении второй производной эту процедуру надо применить два раза, и т. д. В итоге производная порядка k запишется так:

$$\varphi^{(k)}(t) = (h^1\partial_1 + \dots + h^n\partial_n)^k f(x+th).$$

Обратим внимание на то, что коэффициенты, на которые домножаются частные производные, т. е. компоненты h^i вектора h, не зависят от x, поэтому, когда будем применять эту операцию много раз, скобки легко раскрываются. Ниже мы распишем первые три члена подробно.

Подставим полученный результат в формулу (†), положив $\tau = 1$ и заметив, что при $t = 0$ мы оказываемся в точке x, а при $t = 1$ — в точке $x + h$:

$$\begin{aligned} f(x+h) = \varphi(1) = \sum_{k=0}^{m} &\frac{(h^1\partial_1 + \dots + h^n\partial_n)^k f(x)}{k!} \\ &+ \frac{1}{m!}\int\limits_0^1 (1-u)^m (h^1\partial_1 + \dots + h^n\partial_n)^{m+1} f(x+uh)\,du. \end{aligned}$$

Другие формы остаточного члена получаются аналогично.

Теорема доказана.

Определение второго дифференциала и матрицы Гессе. Запишем формулу Тейлора до третьего порядка подробно:

$$\begin{aligned} f(x+h) = f(x) + \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x)h^i + \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(x)h^i h^j \\ + \frac{1}{3!}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^3 f}{\partial x^i \partial x^j \partial x^k}(x)h^i h^j h^k + r_3(x,h). \end{aligned} \tag{‡}$$

В представлении (‡) слагаемое $\displaystyle\sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x)h^i$ есть не что иное как значение первого дифференциала на векторе h. Похожее на него следующее слагаемое естественно считать вторым дифференциалом, примененным к вектору h, так что по определению положим

$$d^2 f(x)(h) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(x)h^i h^j. \tag{**}$$

Следующее слагаемое без множителя $\dfrac{1}{3!}$ называют третьим дифференциалом, и т. д.

Если воспользоваться тем, что $dx^i(h) = h^i$, записать формулу $(**)$ в виде

$$d^2 f(x)(h) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(x) dx^i(h) dx^j(h),$$

а затем убрать упоминание h, то получится стандартное выражение для второго дифференциала:

$$d^2 f(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(x) dx^i dx^j,$$

в котором функционалы dx^k можно рассматривать как переменные квадратичной формы.

Второй дифференциал представляет собой квадратичную форму от переменных h^i. Ее матрицу, т. е. матрицу, составленную из вторых производных, называют *матрицей Гессе* и обозначают через $\dfrac{\partial^2 f}{\partial x^2}$, так что

$$\frac{\partial^2 f}{\partial x^2}(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial (x^1)^2}(x) & \frac{\partial^2 f}{\partial x^1 \partial x^2}(x) & \ldots & \frac{\partial^2 f}{\partial x^1 \partial x^n}(x) \\ \multicolumn{4}{c}{\dotfill} \\ \frac{\partial^2 f}{\partial x^n \partial x^1}(x) & \frac{\partial^2 f}{\partial x^n \partial x^2}(x) & \ldots & \frac{\partial^2 f}{\partial (x^n)^2}(x) \end{pmatrix}.$$

Эта матрица симметричная ввиду независимости вторых частных производных от порядка следования переменных дифференцирования.

У нас возникает уже вторая матрица из частных производных. Ранее мы определили матрицу Якоби. Она состоит из первых производных $\dfrac{\partial f^j}{\partial x^i}$. Ее (вертикальный) размер вызван наличием нескольких компонент f^j. Определяя матрицу Гессе, мы рассматриваем скалярную функцию, а матрица получается потому, что берутся вторые производные $\dfrac{\partial^2 f}{\partial x^i \partial x^j}$. Отметим также, что матрица Гессе, в отличие от матрицы Якоби, всегда квадратная.

§ 5. Локальный экстремум

Определение локального экстремума. Пусть $f : \Omega \to \mathbb{R}$ — функция, заданная в области $\Omega \subset \mathbb{R}^n$. Говорят, что f *имеет локальный максимум* (*минимум*) *в точке* $x \in \Omega$, если существует шар

$B(x,\delta)$ такой, что для любого $y \in \Omega \cap B(x,\delta)$ выполняется неравенство $f(y) \leqslant f(x)$ (соответственно $f(y) \geqslant f(x)$). Если функция имеет в данной точке либо локальный минимум, либо локальный максимум, то говорят, что она имеет в этой точке *локальный экстремум.*

Определение критической точки. Пусть $f : \Omega \to \mathbb{R}$ — функция, заданная в области $\Omega \subset \mathbb{R}^n$. Точку $x \in \Omega$, в которой есть все частные производные первого порядка, называют *критической точкой функции* f, если

$$\frac{\partial f}{\partial x^1}(x) = \cdots = \frac{\partial f}{\partial x^n}(x) = 0.$$

Теорема 9 (необходимое условие локального экстремума). *Пусть $f : \Omega \to \mathbb{R}$ — функция класса $C^1(\Omega)$, где $\Omega \subset \mathbb{R}^n$ — область в $\mathbb{R}^n$. Если f имеет локальный экстремум в точке x, то точка x критическая.*

Доказательство. Докажем, что $\dfrac{\partial f}{\partial x^i}(x) = 0$. Для этого рассмотрим функцию $\varphi(t) = f(x + te_i)$ одной переменной t, заданную в окрестности точки $t = 0$. Эта функция имеет локальный экстремум в точке $t = 0$ того же типа, что и функция f в точке x, и дифференцируема в точке $t = 0$ как композиция дифференцируемых функций. Стало быть, по теореме Ферма $\varphi'(0) = 0$. С другой стороны, по определению $\varphi'(0) = \dfrac{\partial f}{\partial x^i}(x)$. Значит, $\dfrac{\partial f}{\partial x^i}(x) = 0$.

Теорема доказана.

Теорема 10 (достаточное условие локального экстремума). *Пусть $f : \Omega \to \mathbb{R}$ — функция класса $C^2(\Omega)$, где $\Omega \subset \mathbb{R}^n$ — область в $\mathbb{R}^n$. Предположим, что в точке $x \in \Omega$ выполнено необходимое условие локального экстремума, т. е. это критическая точка. Рассмотрим квадратичную форму*

$$Q(h) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(x) h^i h^j.$$

(1) *Если $Q(h)$ положительно определенная, т. е. $Q(h) > 0$ для любого $h \in \mathbb{R}^n$, $h \neq 0$, то x — точка локального минимума.*

(2) *Если $Q(h)$ отрицательно определенная, т. е. $Q(h) < 0$ для любого $h \in \mathbb{R}^n$, $h \neq 0$, то x — точка локального максимума.*

(3) *Если $Q(h)$ — знаконеопределенная, т. е. $Q(h)$ принимает значения обоих знаков, то в точке x локального экстремума нет.*

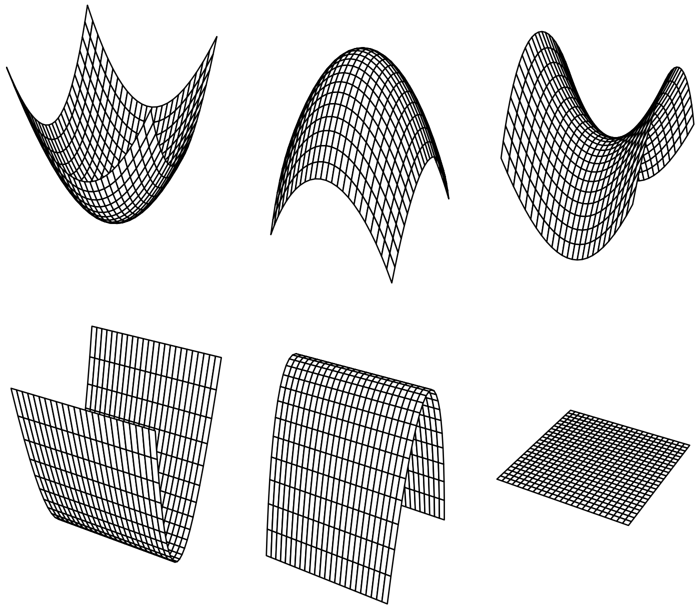

Рис. 5.1.

Замечание о квадратичных формах. Ограничимся формами в $\mathbb{R}^2$, для которых есть наглядная геометрическая интерпретация. Квадратичная форма в $\mathbb{R}^2$ — это функция вида

$$Q(h) = a_{11}(h^1)^2 + 2a_{12}h^1h^2 + a_{22}(h^2)^2$$
$$= (h^1, h^2)\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}\begin{pmatrix} h^1 \\ h^2 \end{pmatrix} = \langle h, Ah \rangle.$$

Можно изобразить графики таких функций, и в зависимости от коэффициентов могут быть поверхности, изображенные на рис. 5.1. Например, если $a_{11} = a_{22} = 1$, $a_{12} = 0$, то получится параболоид, как на рис. 5.1 вверху слева. Если $a_{11} = a_{22} = -1$, $a_{12} = 0$, то увидим параболоид, как на рис. 5.1 вверху в центре. Может получиться седловая поверхность (рис. 5.1, вверху справа), и так будет при коэффициентах $a_{11} = 1$, $a_{22} = -1$, $a_{12} = 0$. Может быть поверхность, изображенная на рис. 5.1 внизу. Может оказаться просто координатная плоскость xOy.

В начале координат любая квадратичная форма равна нулю. Если графиком формы является параболоид, то в любой ненулевой точке значения такой формы положительны. В таком случае говорят, что квадратичная форма *положительно определенная* в том смысле, что она определенно положительна, т. е. принимает только положительные значения на ненулевых векторах. В ситуации, когда значения квадратичной формы на ненулевых векторах отрицательны, такую форму называют *отрицательно определенной.* Если форма принимает как положительные, так и отрицательные значение, то ее называют *знаконеопределенной.* Наконец, если форма принимает только неотрицательные или только неположительные значения и обращается в нуль не только в нуле, то ее называют *полуопределенной.*

Знакоопределенность квадратичной формы можно выяснить несколькими способами. Например, можно записать характеристический многочлен матрицы, найти ее собственные числа, и если они все положительны, то форма положительно определенная, если отрицательны, то — отрицательно. Кроме того, есть критерий Сильвестра, который чаще всего используется на практике. Он состоит в следующем. Запишем матрицу Гессе и найдем главные миноры. Если они все положительны, то квадратичная форма положительно определенная, если их знаки чередуются, начиная с отрицательного, то — отрицательно определенная. Кстати, положительность или отрицательность всех элементов матрицы Гессе не влечет ее знакоопределенности.

Доказательство. Шаг 1 (применение формулы Тейлора). Пусть $x \in \Omega$ — критическая точка и $B(x, r)$ — шар, содержащийся в Ω. Будем брать приращения h так, что $h \in B(0, r)$. По формуле Тейлора

$$f(x+h) = f(x) + \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x) h^i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(x) h^i h^j + \alpha(x,h)|h|_2^2, \tag{$*$}$$

где $\alpha(x, h) \to 0$ при $h \to 0$. По условию точка x критическая, поэтому в ней первые производные равны нулю, стало быть,

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(x) h^i = 0.$$

С учетом этого обстоятельства запишем $(*)$ так:

$$f(x+h) - f(x) = \left[\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(x) \frac{h^i}{|h^i|_2} \frac{h^j}{|h^j|_2} + \alpha(x,h) \right] |h|_2^2. \tag{$**$}$$

В равенстве $(**)$ появилась квадратичная форма второго дифференциала

$$Q(h) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(x) h^i h^j, \qquad (\dagger)$$

примененная к единичному вектору $\dfrac{h}{|h|_2}$. Рассмотрим форму $(\dagger)$ для $h \in S = S(0,1)$. Она непрерывна как функция от h, а сфера является компактным множеством. По теореме Вейерштрасса существуют такие векторы $h_{\min}, h_{\max} \in S$, что для любого $h \in S$

$$Q(h_{\min}) \leqslant Q(h) \leqslant Q(h_{\max}).$$

Вместе с тем функция

$$h \mapsto Q\left(\frac{h}{|h|_2}\right)$$

однородная, т. е. принимает одни и те же значения вдоль лучей. Следовательно, функция $Q\left(\dfrac{h}{|h|_2}\right)$ принимает в точности те значения, которые она принимает на единичной сфере. Тогда для любого $h \neq 0$

$$Q(h_{\min}) \leqslant Q\left(\frac{h}{|h|_2}\right) \leqslant Q(h_{\max}).$$

Шаг 2 (случай $Q > 0$). Поскольку в этом случае форма принимает только положительные значения на ненулевых векторах, ее наименьшее значение положительно, т. е. $Q(h_{\min}) > 0$. По определению предела существует такой шар $B(0,\delta) \subset B(0,r)$, что $|\alpha(x,h)| < \dfrac{1}{4} Q(h_{\min})$ для любого $h \in B(0,\delta)$. Для таких h выражение в квадратных скобках из $(**)$ положительно. Множитель $|h|_2^2$ всегда положителен для $h \neq 0$, стало быть, $f(x+h) - f(x) > 0$, тем самым x — точка локального минимума функции f.

Шаг 3 ($Q < 0$). Можно либо повторить рассуждения, аналогичные проделанным на шаге 2, либо просто применить доказанный на шаге 2 результат к функции $-f$.

Шаг 4 (Q принимает значения разных знаков). В этом случае, естественно, $Q(h_{\min}) < 0$, $Q(h_{\max}) > 0$. По определению предела существует такой шар $B(0,\delta)$, что для любого $h \in B(0,\delta)$

$$|\alpha(x,h)| < \frac{1}{2} \min\{-Q(h_{\min}), Q(h_{\max})\}.$$

Тогда при $0 < t < \delta$ имеем

$$f(x + th_{\min}) - f(x) = \left[\frac{1}{2} Q\left(\frac{th_{\min}}{|th_{\min}|}\right) + \alpha(x, th_{\min})\right] |th_{\min}|_2^2. \qquad (\dagger)$$

Ввиду однородности Q

$$Q\left(\frac{th_{\min}}{|th_{\min}|}\right) = Q\left(\frac{h_{\min}}{|h_{\min}|}\right) = Q(h_{\min})$$

и выражение в скобках в (†) отрицательно, а вместе с этим и

$$f(x + th_{\min}) - f(x) < 0.$$

Рассуждая аналогично, можно показать, что при $0 < t < \delta$

$$f(x + th_{\max}) - f(x) > 0.$$

Следовательно, в любой окрестности x есть точки, в которых значения функции больше чем $f(x)$, и есть точки, в которых значения меньше чем $f(x)$, стало быть, x не является точкой локального экстремума.

Теорема доказана.

Естественно возникает вопрос: а если форма второго дифференциала полуопределенная, т. е. принимает только неотрицательные или только неположительные значения, но может обращаться в нуль не только в нуле, в таком случае можно ли гарантировать локальный экстремум в данной точке, хотя бы нестрогий? Оказывается, нельзя. Приведем две функции, у которых форма второго дифференциала в некоторой точке принимает неотрицательные значения, но у одной из них есть экстремум, а у другой — нет. Пусть $f(x, y) = x^2 + y^4$, $g(x, y) = x^2 - y^4$. Ясно, что в нуле первые частные производные равны нулю, а из вторых производных в нуле производная по x дважды равна 2, а остальные — нулю, так что их матрица Гессе такова: $\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$. Однако функция f имеет в нуле локальный минимум, а g экстремума не имеет.

Тем самым полуопределенность формы означает всего лишь, что по соответствующему направлению квадратичная форма не дает достаточной информации для обнаружения экстремума или его отсутствия.

Пример (метод наименьших квадратов). Предположим, что есть данные приближенных измерений, представляющие собой множество пар (x_i, y_i), $i = 1, \ldots, n$. Отметим их на координатной плоскости и допустим, что результаты наблюдений позволяют предположить, что эти данные лежат неподалеку от некоторой прямой, т. е. предположительно они находятся в аффинной зависимости. Будем предполагать, что измерения по оси Ox точные, а по оси Oy — приближенные. Ставится вопрос: найти зависимость вида $f(x; a, b) = ax + b$, т. е. найти такие a, b, что указанная зависимость наилучшим образом описывает измерения. Естественно, поскольку в задаче есть фраза «наилучшим образом», надо ввести целевую функцию, относительно которой эти слова обретут конкретное содержание. В силу разных обстоятельств логично потребовать, чтобы сумма квадратов отклонений ординат измерений от соответствующих точек графика функции $ax+b$ была наименьшей, т. е. требуется найти такие a, b, чтобы значение функции

$$F(a, b) = \sum_{i=1}^{n} (f(x_i; a, b) - y_i)^2$$

оказалось наименьшим. Тем самым поставлена задача на нахождение минимума функции двух переменных. Будем действовать согласно теории: найдем частные производные, приравняем их к нулю, найдем критическую точку и, скорее всего, это будет ответ, хотя в завершение все-таки проверим, что это действительно точка минимума. Имеем

$$\frac{\partial F}{\partial a} = \sum_{i=1}^{n} 2((ax_i + b) - y_i)x_i = 2\left(a\sum_{i=1}^{n} x_i x_i + b\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i x_i\right) = 0.$$

Заметим, что

$$\sum_{i=1}^{n} x_i x_i = |x|_2^2, \quad \sum_{i=1}^{n} y_i x_i = \langle x, y\rangle,$$

и введем обозначение

$$X = \sum_{i=1}^{n} x_i.$$

Приравняем к нулю производную по b:

$$\frac{\partial F}{\partial b} = 2\sum_{i=1}^{n}((ax_i + b) - y_i) = 2\left(a\sum_{i=1}^{n} x_i + nb - \sum_{i=1}^{n} y_i\right).$$

Обозначим

$$Y = \sum_{i=1}^{n} y_i.$$

Запишем получившуюся систему линейных уравнений в матричном виде:

$$\begin{pmatrix} |x|_2^2 & X \\ X & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \langle x, y\rangle \\ Y \end{pmatrix}.$$

Ее решением будет

$$a = \frac{n\langle x, y\rangle - XY}{n|x|_2^2 - X^2}, \quad b = \frac{|x|_2^2 Y - X\langle x, y\rangle}{n|x|_2^2 - X^2}.$$

Получили координаты критической точки. Проверим, что в этой точке действительно минимум. Для этого выпишем матрицу вторых производных:

$$\frac{\partial F^2}{\partial (a,b)^2} = \begin{pmatrix} |x|_2^2 & X \\ X & n \end{pmatrix}.$$

Изучим ее положительную определенность. Согласно критерию Сильвестра должны быть положительными главные миноры этой матрицы:

$$|x|_2^2 > 0, \quad n|x|_2^2 - \left(\sum_{i=1}^n x_i\right)^2 > 0.$$

Длина вектора строго больше нуля, когда среди его компонент есть хотя бы одна ненулевая. Это требование содержательно означает, что измерения проводятся не только при $x = 0$. Разберемся со вторым неравенством. Преобразуем его:

$$n|x|_2^2 - \left(\sum_{i=1}^n x_i\right)^2 > 0 \Longleftrightarrow \sqrt{n}|x|_2 > \left|\sum_{i=1}^n x_i\right|.$$

Здесь в левой части евклидова норма. Правая часть похожа на l_1-норму, только модулей не хватает. Для норм есть связывающее их неравенство

$$\sum_{i=1}^n |x_i| \leqslant \sqrt{n}\,|x|_2,$$

откуда следует требуемое неравенство:

$$\sqrt{n}\,|x|_2 \geqslant \sum_{i=1}^n |x_i| \geqslant \left|\sum_{i=1}^n x_i\right|.$$

Заметим, что для положительной определенности матрицы Гессе нужно строгое неравенство, а мы имеем нестрогое. Однако равенство в нем достигается, как нетрудно проверить, в единственном случае $x_1 = \dots = x_n$. Значит, если хотя бы две из точек x_i различны, то неравенство строгое и полученная критическая точка является точкой минимума.

§ 6. Теоремы об обратной и неявной функциях

6.1. Теорема об обратной функции.

Вопрос, который будет рассматриваться в этом параграфе, состоит в следующем. Рассмотрим функцию $f : \Omega \to \mathbb{R}^m$, где Ω — область в $\mathbb{R}^n$. Функция f многозначна, так что у нее m координатных функций, пусть

$$(y^1, \dots, y^m) = (f^1(x^1, \dots, x^n), \dots, f^m(x^1, \dots, x^n)).$$

Ставится вопрос: когда у нее есть обратная, т. е. когда для любого $y \in f(\Omega)$ существует единственное $x \in \Omega$, для которого $y = f(x)$? После положительного ответа на этот вопрос естественно возникает вопрос о дифференциальных свойствах обратного отображения, т. е. когда оно дифференцируемо и как его дифференциал и частные производные связаны с дифференциалом и частными производными исходного отображения?

Для ответа на поставленные вопросы обратимся к анализу простейших отображений и появившиеся там идеи попробуем реализовать в общем случае. Любую дифференцируемую в точке функцию можно приблизить линейной, а именно ее дифференциалом, поэтому сначала изучим поставленные вопросы для линейных отображений. Рассмотрим линейное отображение, действующее следующим образом:

$$\begin{pmatrix} y^1 \\ \vdots \\ y^m \end{pmatrix} = \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_1^m & \dots & a_n^m \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix},$$

где $\left(a_i^j\right)$ — матрица этого линейного отображения. Есть три возможности для сопоставления области определения и множества значений: $n < m$, $n = m$ и $n > m$. В первом случае, когда переменных меньше чем значений, даже при взаимной однозначности отображения образ не будет заполнять все пространство. Образно этот случай для малых размерностей можно описать так: это прямая в трехмерном пространстве. Конечно, встав в точку такой прямой, мы можем найти ее прообраз, но дифференцировать такое отображение невозможно — из данной точки мы практически всегда попадем за пределы прямой и дифференциал определить нельзя. Этот случай мы обсудим позже, когда будем говорить о многообразиях.

Здесь подробно обсудим второй и третий случаи. Второй нам даст теорему об обратной функции, третий — теорему о неявной функции.

Начнем со случая $n = m$. Тогда матрица линейного оператора квадратная. Когда можно решить систему линейных уравнений с квадратной матрицей? В том случае, когда ее определитель отличен от нуля. Это и есть то условие, которое появится в общем нелинейном случае.

Теорема 11 (об обратной функции). *Пусть*

(1) f *— отображение области* $\Omega \subset \mathbb{R}^n$ *в* $\mathbb{R}^n$, *все координатные функции которого принадлежат классу* $C^1(\Omega)$.

(2) $x_0 \in \Omega$ *такова, что* $\det \dfrac{\partial f}{\partial x}(x_0) \neq 0$.

Тогда существует окрестность $U(x_0)$ *точки* x_0, *на которой отображение* f *обратимо, т. е. существует обратное к* f *отображение* f^{-1}, *переводящее* $f(U)$ *в* U *и такое, что*

$$f^{-1}(f(x)) = x,\ x \in U, \quad f(f^{-1}(y)) = y,\ y \in f(U).$$

Кроме того, образ $f(U)$ *— область в* $\mathbb{R}^n$ *и окрестность точки* $y_0 = f(x_0)$.

Матрицы Якоби отображений f *и* f^{-1} *связаны равенством*

$$\frac{\partial f^{-1}}{\partial y}(y_0) = \left(\frac{\partial f}{\partial x}(x_0)\right)^{-1}, \quad y_0 = f(x_0).$$

Образно условия и результат теоремы можно представить себе так. Дифференцируемая функция в некоторой окрестности фиксированной точки приближается линейным отображением. Теорема утверждает, что если это линейное отображение обратимо, то и исходная функция тоже обратима по крайней мере в некоторой окрестности рассматриваемой точки.

Несмотря на прозрачность идеи, доказательство этого утверждения технически сложное и мы его опустим. Второе утверждение теоремы касается нахождения матрицы Якоби обратной функции и доказывается просто.

Замечание (о нахождении частных производных). Рассмотрим тождество

$$f^{-1}(f(x)) = x.$$

Это векторное равенство. Его можно продифференцировать:

$$\frac{\partial}{\partial x^i}((f^{-1})^j(f(x))) = \frac{\partial}{\partial x^i}x^j = \delta_i^j = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Выражение δ_i^j называют *символом Кронекера.* Выполним дифференцирование в левой части, пользуясь цепным правилом:

$$\sum_{k=1}^{n} \frac{\partial (f^{-1})^j}{\partial y^k}(y_0)\frac{\partial f^k}{\partial x^i}(x_0) = \delta_i^j,$$

В левой части стоит произведение двух матриц, и последнее равенство означает, что оно равно единичной матрице. В матричном виде это запишется так:

$$\frac{\partial f^{-1}}{\partial y}(y_0)\frac{\partial f}{\partial x}(x_0) = I,$$

где I — единичная матрица, откуда

$$\frac{\partial f^{-1}}{\partial x}(y_0) = \left(\frac{\partial f}{\partial x}(x_0)\right)^{-1}.$$

Замечание (о глобальной обратимости). Теорема об обратной функции дает только локальную обратимость, т. е. обратимость функции в некоторой окрестности той точки, в которой выполнены условия. Ясно, что основным препятствием отсутствия локальной обратимости является вырожденность матрицы Якоби: если $\det \dfrac{\partial f}{\partial x} = 0$, обратной функции может не быть. Это легко видно даже в одномерное случае на примере функции $y = x^2$, $x_0 = 0$. Однако обратная функция иногда существует, несмотря на вырожденность матрицы Якоби, как, например, для функции $y = x^3$ и $x_0 = 0$.

В одномерном случае если $\dfrac{df}{dx}(x) \neq 0$ для любого $x \in (a, b)$, то можно гарантировать глобальную обратимость функции f, т. е. существование обратной функции на всем промежутке (a, b). Как обстоит дело с глобальной обратимостью в многомерном случае? Пусть $n > 1$ и $\det \dfrac{\partial f}{\partial x}(x) \neq 0$ в каждой точке $x \in \Omega$. Можно ли утверждать, что f обратима на всей области Ω? Вообще говоря, необязательно. Например, рассмотрим отображение

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

якобиан которого равен r, так что оно локально обратимо для $r \neq 0$. Будет ли оно обратимо в полуплоскости $r > 0$ переменных r, φ? Ясно, что если к φ добавить величину, кратную 2π, то результат не изменится, и мы будем получать одно и то же значение (x, y) при разных значениях (r, φ). Тем самым глобальной обратимости нет.

Пример 1. Найдем матрицу Якоби обратного к отображению

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

Для начала запишем само обратное отображение, т. е. выразим r, φ через x, y (для первого квадранта):

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

(для других значений x, y надо изменить формулу для φ). Матрицу Якоби можно найти, просто продифференцировав найденное явно обратное отображение, а можно обратить матрицу Якоби исходного отображения. Найдем с помощью теоремы об обратном отображении. Имеем

$$\frac{\partial(r,\varphi)}{\partial(x,y)} = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}^{-1} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\frac{\sin\varphi}{r} & \frac{\cos\varphi}{r} \end{pmatrix}.$$

В полученной обратной матрице остались исходные переменные φ, r, а надо чтобы были x, y. Перейдя к x, y, получим матрицу

$$\begin{pmatrix} \dfrac{x}{\sqrt{x^2+y^2}} & \dfrac{y}{\sqrt{x^2+y^2}} \\ -\dfrac{y}{x^2+y^2} & \dfrac{x}{x^2+y^2} \end{pmatrix}.$$

Другой способ заключается в дифференцировании обратного отображения, и, естественно, дает тот же результат. Однако такой способ годится не всегда, а только тогда, когда можно записать обратное отображение явно.

6.2. Теорема о неявной функции.

Замечание (о разрешимости линейной системы в случае $n > m$). Запишем систему в виде

$$\begin{gathered} a_1^1 x^1 + \cdots + a_n^1 x^n = y^1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_1^m x^1 + \cdots + a_n^m x^n = y^m. \end{gathered}$$

О разрешимости системы в том случае, когда число переменных больше числа уравнений, известно следующее. Сразу можно считать, что ранг матрицы a_i^j максимальный, т. е. $\operatorname{rank} A = m$. В этом случае если решения и есть, то их будет много, появляется многообразие решений, и некоторые из переменных можно брать произвольно, а остальные через них выражать. Поскольку ранг максимальный, можно упорядочить переменные так, что определитель, составленный из первых m столбцов матрицы $\left(a_i^j\right)$, будет отличен от нуля и первые m зависимых переменных $x^1, \dots, x^m$ будут выражаться через остальные $n-m$ (независимых) переменных $x^{m+1}, \dots, x^n$. Перенесем независимые переменные в правую часть:

$$\begin{aligned} a_1^1 x^1 + \dots + a_m^1 x^m &= y^1 - \left(a_{m+1}^1 x^{m+1} + \dots + a_n^1 x^n\right) \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_1^m x^1 + \dots + a_m^m x^m &= y^m - \left(a_{m+1}^m x^{m+1} + \dots + a_n^m x^n\right). \end{aligned}$$

Тогда

$$\begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix} = \begin{pmatrix} a_1^1 & \dots & a_m^1 \\ \vdots & & \vdots \\ a_1^m & \dots & a_m^m \end{pmatrix}^{-1} \times \left[\begin{pmatrix} y^1 \\ \vdots \\ y^m \end{pmatrix} - \begin{pmatrix} a_{m+1}^1 & \dots & a_n^1 \\ \vdots & & \vdots \\ a_{m+1}^m & \dots & a_n^m \end{pmatrix} \begin{pmatrix} x^{m+1} \\ \vdots \\ x^n \end{pmatrix} \right].$$

Применение обратной матрицы $\left(a_i^j\right)^{-1}$ к столбцу с координатами y^j дает частное решение, к нему добавляется общее решение однородной системы, т. е. системы с нулевой правой частью $y = 0$.

Нас будет интересовать случай однородной системы, т. е. у которой все y^j равны нулю, и вопрос будет в возможности выражения зависимых переменных через независимые.

Каковы основные наблюдения в наших рассуждениях? Во-первых, ранг матрицы системы должен быть максимальным. Во-вторых, надо было переставить переменные так, чтобы первые столбцы обеспечивали максимальность ранга, т. е. чтобы соответствующая матрица оказалась невырожденной, остальные переменные независимые, и через них выражаются зависимые. Эти наблюдения отразятся в теореме о неявной функции, связанной с нелинейным отображением.

Теорема 12 (о неявной функции). *Пусть*

(1) $f : \Omega \to \mathbb{R}^m$, *где* Ω — *область в* $\mathbb{R}^{n+m}$, *и все координатные функции* f^j *отображения* f *принадлежат классу* $C^1(\Omega)$. *Переменные в* $\mathbb{R}^{n+m}$ *будем записывать в виде пары* (x, y), *где* $x = (x^1, \dots, x^n)$, $y = (y^1, \dots, y^m)$, *и тем самым* $f(x, y) = f(x^1, \dots, x^n, y^1, \dots, y^m)$.

(2) $(x_0, y_0) \in \Omega$ — *произвольная точка такая, что* $f(x_0, y_0) = 0$, *и*

$$\frac{\partial f}{\partial y}(x_0, y_0) = \begin{pmatrix} \frac{\partial f^1}{\partial y^1} & \cdots & \frac{\partial f^1}{\partial y^m} \\ \vdots & & \vdots \\ \frac{\partial f^m}{\partial y^1} & \cdots & \frac{\partial f^m}{\partial y^m} \end{pmatrix}$$

невырожденная.

Тогда существуют окрестность $U(x_0) \subset \mathbb{R}^n$ *точки* x_0, *окрестность* $W(x_0, y_0) \subset \mathbb{R}^{n+m}$ *точки* (x_0, y_0) *и единственная функция* $g : U \to \mathbb{R}^m$ *такие, что для любого* $x \in U$

$$(x, g(x)) \in W \quad \text{и} \quad f(x, g(x)) = 0.$$

При этом функцию g *называют функцией, неявно заданной посредством отображения* f. *Все компоненты* g^j *функции* g *принадлежат классу* $C^1(U)$.

Производные функций g^j находятся путем применения цепного правила к равенству $f(x, g(x)) = 0$ и выражения искомых производных из получаемой системы уравнений.

Замечание (о локализации). Рассмотрим функцию $f(x, y) = x^2 + y^2 - 1$ и уравнение $f(x, y) = 0$, т. е. $x^2 + y^2 - 1 = 0$. Мы хотим выразить y через x. Это можно сделать явно, только надо помнить, что при извлечении корня получаются два значения, поэтому пишут $y = \pm\sqrt{1 - x^2}$. Это означает, что каждому допустимому значению x можно сопоставить одно из двух значений y, а именно одно со знаком плюс, а другое со знаком минус. Получается раздвоение.

Что получается, если применять теорему о неявном отображении? Во-первых, надо выделить соответствующую матрицу и разобраться в ее невырожденности. В нашем случае это просто требование отличия от нуля производной: $\dfrac{\partial f}{\partial y} = 2y \neq 0$. Из точек окружности этим свойством обладают все точки, кроме $(-1, 0)$, $(1, 0)$. Теперь встанем в некоторую точку (x_0, y_0), $y_0 > 0$, окружности и спроектируем близлежащую часть окружности на ось абсцисс, чем получим окрестность U точки x_0. Кроме этого, возьмем окрестность W в $\mathbb{R}^2$ точки

(x_0, y_0), в которой расположена небольшая часть окружности. Тогда появляется функция, сопоставляющая каждому x из U то единственное $y = \sqrt{1-x^2}$, для которого $(x, y) \in W$ и $x^2 + y^2 - 1 = 0$. Второе значение $-\sqrt{1-x^2}$ в W не попадает.

Таким образом, при задании неявной функции важно указывать не только точку x_0, в окрестности которой будет задана неявная функция, но и точку y_0.

Замечание (о дифференцировании неявно заданной функции). В общем случае запишем подробно определяющее неявную функцию тождество

$$f^j(x^1, \dots, x^n, g^1(x^1, \dots, x^n), \dots, g^m(x^1, \dots, x^n)) = 0,\ j = 1, \dots, m,$$

и продифференцируем его:

$$\frac{\partial f^j}{\partial x^i}(x, g(x)) + \sum_{k=1}^{m} \frac{\partial f^j}{\partial y^k}(x, g(x)) \frac{\partial g^k}{\partial x^i}(x) = 0, \quad j = 1, \dots, m.$$

Получилось m уравнений. Матрица $\dfrac{\partial f^j}{\partial y^k}(x, g(x))$ квадратная, далее расположена искомая производная $\dfrac{\partial g^k}{\partial x^i}(x)$, а первое слагаемое можно перенести в другую часть равенства и записать его так:

$$\sum_{k=1}^{m} \frac{\partial f^j}{\partial y^k}(x, g(x)) \frac{\partial g^k}{\partial x^i}(x) = -\frac{\partial f^j}{\partial x^i}(x, g(x)), \quad j = 1, \dots, m.$$

В матричном виде это выглядит так:

$$\begin{pmatrix} \frac{\partial f^1}{\partial y^1}(x, g(x)) & \dots & \frac{\partial f^1}{\partial y^m}(x, g(x)) \\ \vdots & & \vdots \\ \frac{\partial f^m}{\partial y^1}(x, g(x)) & \dots & \frac{\partial f^m}{\partial y^m}(x, g(x)) \end{pmatrix} \begin{pmatrix} \frac{\partial g^1}{\partial x^i}(x) \\ \vdots \\ \frac{\partial g^m}{\partial x^i}(x) \end{pmatrix} = -\begin{pmatrix} \frac{\partial f^1}{\partial x^i}(x, g(x)) \\ \vdots \\ \frac{\partial f^m}{\partial x^i}(x, g(x)) \end{pmatrix}.$$

Получилась система линейных уравнений для нахождения требуемых производных. Кстати, ее матрица и есть та матрица, невырожденность которой предполагалась в условии.

Рассмотрим несколько частных случаев. Пусть $n = 2$, $m = 1$. Здесь число независимых переменных равно 2, зависимых — 1 и всего — 3 . Это значит, что на три переменные накладывают одно условие вида $f(x, y, z) = 0$. Будем считать переменные x, y независимыми, а z — зависимой. Для гарантии существования неявно заданной функции надо предполагать, что $\dfrac{\partial f}{\partial z} \neq 0$. Тогда в некоторой окрестности такой точки существует функция $z(x, y)$ такая, что

$$f(x, y, z(x, y)) = 0.$$

Дифференцируем это тождество:

$$\frac{\partial f}{\partial x}+\frac{\partial f}{\partial z}\frac{\partial z}{\partial x}=0,\quad \frac{\partial f}{\partial y}+\frac{\partial f}{\partial z}\frac{\partial z}{\partial y}=0,$$

откуда

$$\frac{\partial z}{\partial x}=-\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial z}}.\quad \frac{\partial z}{\partial y}=-\frac{\frac{\partial f}{\partial y}}{\frac{\partial f}{\partial z}}.$$

Пусть теперь $n=1$, $m=2$, т. е. независимая переменная одна, а условий два. Пусть условия заданы в виде уравнений

$$f(x,y,z)=0,\quad g(x,y,z)=0.$$

Они при определенных условиях позволяют однозначно выразить, например, y и z через x, т. е. существуют такие функции $y(x)$ и $z(x)$, что

$$f(x,y(x),z(x))=0,\quad g(x,y(x),z(x))=0. \tag{$*$}$$

При каких условиях такое возможно? Надо предполагать, чтобы соответствующая часть матрицы Якоби была невырожденной. В любом случае это условие появится при необходимости выражения производных функций $y(x)$ и $z(x)$, так что сначала можем попытаться найти эти производные и в процессе указать соответствующее условие. Дифференцируем равенства $(*)$ по x:

$$\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\frac{dy}{dx}+\frac{\partial f}{\partial z}\frac{dz}{dx}=0,\quad \frac{\partial g}{\partial x}+\frac{\partial g}{\partial y}\frac{dy}{dx}+\frac{\partial g}{\partial z}\frac{dz}{dx}=0.$$

В матричном виде эта система выглядит так:

$$\begin{pmatrix} \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \end{pmatrix}\begin{pmatrix} \frac{dy}{dx} \\ \frac{dz}{dx} \end{pmatrix}=-\begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial g}{\partial x} \end{pmatrix}.$$

Решим ее по правилу Крамера. Имеем

$$\begin{pmatrix} \frac{dy}{dx} \\ \frac{dz}{dx} \end{pmatrix}=\frac{-1}{\frac{\partial f}{\partial y}\frac{\partial g}{\partial z}-\frac{\partial f}{\partial z}\frac{\partial g}{\partial y}}\begin{pmatrix} \frac{\partial g}{\partial z} & -\frac{\partial f}{\partial z} \\ -\frac{\partial g}{\partial y} & \frac{\partial f}{\partial y} \end{pmatrix}\begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial g}{\partial x} \end{pmatrix}.$$

Тем самым

$$\frac{dy}{dx}=\frac{\frac{\partial f}{\partial z}\frac{\partial g}{\partial x}-\frac{\partial g}{\partial z}\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}\frac{\partial g}{\partial z}-\frac{\partial f}{\partial z}\frac{\partial g}{\partial y}},\quad \frac{dz}{dx}=\frac{\frac{\partial g}{\partial y}\frac{\partial f}{\partial x}-\frac{\partial f}{\partial y}\frac{\partial g}{\partial x}}{\frac{\partial f}{\partial y}\frac{\partial g}{\partial z}-\frac{\partial f}{\partial z}\frac{\partial g}{\partial y}}.$$

Обратим внимание на то, что деление происходит на определитель, отличие от нуля которого должно предполагаться в условиях.

Пример 2. Состояние газа характеризуется уравнением

$$pV = \nu RT,$$

где ν, R — некоторые постоянные, T — температура, V — объем и p — давление, или, иначе,

$$f(p, V, T) = \frac{pV}{T} - \nu R = 0.$$

Получилась зависимость между тремя переменными p, V, T. Кроме того, при адиабатическом процессе выполнено равенство $g(p, V, T) = pV^\gamma - \mathrm{const} = 0$. Получилось два уравнения, связывающие три переменные. При некоторых условиях две из них могут быть выражены через оставшуюся одну. В качестве условия выступает отличие от нуля соответствующего минора. Выпишем матрицу Якоби, считая p первой переменной, V второй, T третьей:

$$\frac{\partial(f, g)}{\partial(p, V, T)} = \begin{pmatrix} \frac{V}{T} & \frac{p}{T} & -\frac{pV}{T^2} \\ V^\gamma & \gamma p V^{\gamma-1} & 0 \end{pmatrix}.$$

Теперь начинаем считать различные миноры. Например, если считать температуру независимой переменной, а p и V зависимыми, то надо предполагать отличным от нуля минор, составленный из первых двух столбцов:

$$\frac{\gamma p V^\gamma}{T} - \frac{pV^\gamma}{T} = \frac{(\gamma - 1)pV^\gamma}{T} \neq 0,$$

так как в рассматриваемом случае γ никогда не берется равным 1. Это означает, что все можно выразить через температуру.

Аналогично можно посчитать остальные миноры, и легко заметить, что они отличны от нуля. Значит, здесь любую из переменных можно взять в качестве независимой.

Пример 3 (двойной математический маятник). Рассмотрим систему из двух точек (x, y) и (u, v) на плоскости, связанных уравнениями

$$f(x, y) = x^2 + y^2 = R^2, \quad g(x, y, u, v) = (x - u)^2 + (y - v)^2 = r^2.$$

Можно себе представить, что есть потолок, на нем закреплен шарнир, расположенный в начале координат, есть закрепленное на шарнире звено, координаты второго конца которого равны (x, y), и есть еще звено, находящееся на расстоянии r от конца первого, пусть его координаты (u, v) (рис. 6.1). Нас будет интересовать ее кинематика, т. е. как она может двигаться и при помощи каких переменных это можно описать.

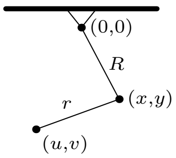

Рис. 6.1.

У нас четыре переменных и два уравнения связи. Следовательно, две из переменных можно выразить через оставшиеся при определенных условиях. В окрестности некоторых точек эти условия выполняются, в других — нет. Чтобы понять, когда они выполняются, надо выписать матрицу Якоби и рассмотреть соответствующие миноры. Имеем

$$\frac{\partial(f,g)}{\partial(x,y,u,v)} = \begin{pmatrix} 2x & 2y & 0 & 0 \\ 2(x-u) & 2(y-v) & -2(x-u) & -2(y-v) \end{pmatrix}$$

Случай 1. Минор, составленный из последних двух столбцов, всегда нулевой. Это означает, что x, y нельзя брать в качестве независимых переменных. Это и понятно — в таком случае вторая пара (u, v) ничем не фиксируется.

Случай 2. Изучим минор, составленный из первых двух столбцов и тем самым ответим на вопрос: можно ли взять в качестве независимых переменные u, v? Он легко вычисляется и условие отличия его от нуля выглядит так: $4(xy - xv - yx + yu) = 4(uy - xv) \neq 0$, или $\begin{vmatrix} u & v \\ x & y \end{vmatrix} \neq 0$. Это означает, что векторы (x, y) и (u, v) линейно независимы, т. е. исключается ситуация, когда звенья лежат на одной прямой (рис. 6.2).

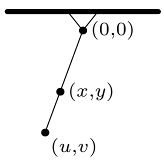

Рис. 6.2.

Почему в таком случае нельзя описать положение всей системы только положением точки (u, v)? Потому, что при смещении точки (u, v) вдоль образованного отрезка точка (x, y) может сместиться как в одну сторону относительно прямой, так и в другую в некоторой окрестности своего начального положения. Но если точку (u, v) хотя бы немного сместить, то локально положение точки (x, y) определяется однозначно.

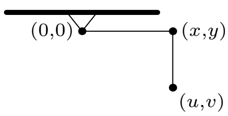

Рис. 6.3.

Случай 3. Из оставшихся случаев рассмотрим, например, тот, который соответствует второму и третьему столбцам матрицы Якоби. Тогда для соответствующего минора должно быть $4y(x - u) \neq 0$. Это значит, что $y \neq 0$ и $x \neq u$. Соотношения $y = 0$, $x = u$ означают, что первый отрезок расположен горизонтально, а второй отходит от него вертикально вниз (рис. 6.3). В этом случае нельзя описать положение при помощи x и v, т. е. только при помощи горизонтального отклонения первого отрезка и разности $y-v$. Действительно, в таком случае при изменении x первый отрезок может качнуться как вверх, так и вниз, следовательно, однозначности описания нет и аналогично при изменении v неясно, куда отклонится нижний отрезок, влево или вправо.

§ 7. Диффеоморфизмы и замена переменных

7.1. Криволинейные координаты.

Определение диффеоморфизма. Отображение $\Phi : U \to \Omega$, где U и Ω — области в $\mathbb{R}^n$, называют *диффеоморфизмом*, если выполнены условия

(1) Φ обратимо, т. е. действует взаимно однозначно между U и Ω.

(2) Отображение Φ и обратное к нему Φ^{-1} принадлежат по крайней мере классу C^1 (каждое в своей области).

Замечание. Матрицы Якоби $\dfrac{\partial \Phi}{\partial x}(x)$ и $\dfrac{\partial \Phi^{-1}}{\partial y}(y)$ невырожденные во всех точках.

Диффеоморфизмы нужны по крайней мере для того чтобы делать замену переменной и переходить к криволинейным координатам. Сначала несколько слов о том, что это такое.

Математическое описание какого-либо процесса нередко сопровождается выделением набора его числовых характеристик и заданием некоторой числовой величины, характеризующей процесс. Иначе говоря, такое описание приводит к определению конечномерного арифметического пространства и заданию (числовой) функции на некоторой его открытой области. Пусть переменные этого пространства обозначены через $x = (x_1, x_2, \dots, x_n)$, а функция — через $f(x)$. Как правило, в описании процесса большую роль играет не только сама функция, но и скорость ее изменения по тем или иным координатным направлениям, т. е. частные производные функции. Иногда бывает полезно посмотреть, как изменяется функция вдоль других линий.

Для формирования подходящего математического аппарата заметим, что положение $x = (x_1, x_2, \dots, x_n)$ точки x можно описать так. Фиксируем какую-то одну из переменных $x_1, x_2, \dots, x_n$, а остальные оставим меняющимися. Множество таких точек образует некоторую координатную (гипер)плоскость в $\mathbb{R}^n$. Тогда наша точка x может быть охарактеризована как точка, находящаяся на пересечении всех таких плоскостей.

Предположим, что мы хотим использовать другой набор чисел $u = (u_1, u_2, \ldots, u_n)$ для описания положения той же самой точки. Естественно, что он должен быть связан с исходным набором $(x_1, x_2, \ldots, x_n)$ некоторой зависимостью. Иначе говоря, предположим, что на открытой области Ω изменения переменной u задано взаимно однозначное отображение Φ (чаще всего предполагается, что это диффеоморфизм) такое, что $x = \Phi(u)$. Расположение точки x на пересечении плоскостей $x_k = \mathrm{const}$ можно в терминах u описать как расположение этой точки на пересечении поверхностей $\{x : x = \Phi(u), u_k = \mathrm{const}\}$, т. е. на этот раз на пересечении множеств, не являющихся плоскими (с точки зрения переменных x). Исходя из такой возможности, набор чисел $u = (u_1, u_2, \ldots, u_n)$ называют *криволинейными координатами точки* $x = (x_1, x_2, \ldots, x_n)$, говоря при этом о самих $x_1, x_2, \ldots, x_n$ как о *исходных координатах*.

Приведем наиболее часто используемые конкретные системы координат.

Полярная система координат. Положение точки $(x, y) \in \mathbb{R}^2$ можно описать при помощи расстояния $r = \sqrt{x^2 + y^2}$ до нее от начала координат и полярного угла φ, т. е. угла между осью абсцисс и лучом, исходящим из начала и проходящим через данную точку. Связь между r, φ и x, y устанавливается следующим диффеоморфизмом Φ:

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

действующим из области $U = (0, +\infty) \times (0, 2\pi)$ в область $\Omega = \mathbb{R}^2 \setminus ([0, +\infty) \times \{0\})$. Бывает удобно включить в U значение $\varphi = 0$. Иногда это можно делать, иногда нет. Так, при замене переменных в дифференциальных выражениях важно обращаться к диффеоморфизмам, заданным на области. Но включение нулевого угла, во-первых, нарушает открытость множества, а во-вторых, приводит к отсутствию однозначности обратного отображения. Поэтому при изучении свойств, связанных с производными, всегда рассматривают указанную выше область или заменяют $(0, 2\pi)$ другим интервалом той же длины, например $-\pi < \varphi < \pi$. Исключение точки $r = 0$ происходит для обеспечения взаимной однозначности отображения Φ.

Якобиан этого отображения равен r.

Как всегда при рассмотрении криволинейных координат полезно изображать координатные линии. Они получаются в плоскости (x, y)

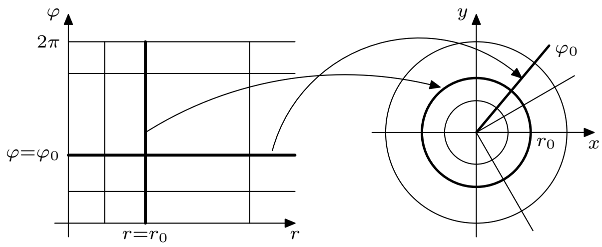

Рис. 7.1.

как образы множеств, на которых изменяется только одна из переменных r, φ. В плоскости (r, φ), изображенной в виде прямоугольной системы координат, множества постоянства r или φ суть либо отрезки $r = C$, $0 < \varphi < 2\pi$, либо лучи $\varphi = C$, $r > 0$. Отрезки, на которых постоянно r, т. е. $r = r_0$, перейдут при отображении Φ в окружности с центром в нуле радиуса r_0, а лучи $\varphi = \varphi_0$ — в исходящие из начала координат лучи (рис. 7.1). Образованные пересечениями координатных линии в плоскости (r, φ) прямоугольники перейдет в соответствующие «криволинейные четырехугольники» на плоскости (x, y). Забегая вперед, можно отметить, что якобиан показывает, во сколько раз изменяется площадь прямоугольника в плоскости (r, φ) при переносе его в плоскость (x, y).

Цилиндрическая система координат. Эта система координат получается из полярной добавлением еще одной координаты. А именно, точка $(x, y, z) \in \mathbb{R}^3$ описывается радиусом $r = \sqrt{x^2 + y^2}$, полярным углом φ, как в полярной системе координат, и координатой z исходной декартовой системы координат. Тем самым цилиндрические координаты задаются диффеоморфизмом

$$x = r\cos\varphi, \quad y = r\sin\varphi, \quad z = z,$$

т. е. диффеоморфизмом

$$\Phi(r, \varphi, z) = (r\cos\varphi, r\sin\varphi, z),$$

определенным на области $(0, +\infty) \times (0, 2\pi) \times (-\infty, +\infty)$.

Изобразим координатные поверхности цилиндрической системы координат. Поверхности $r = \mathrm{const}$ суть цилиндры (рис. 7.2(a)), поверхности $\varphi = \mathrm{const}$ — полуплоскости (рис. 7.2(b)), поверхности $z = \mathrm{const}$ — плоскости, параллельные плоскости xOy.

Нетрудно найти, что якобиан цилиндрической системы координат равен r.

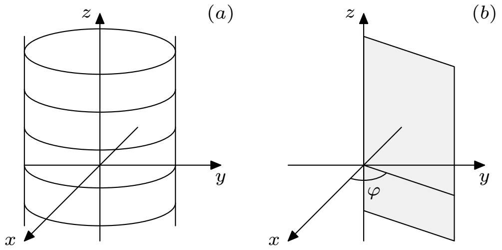

Рис. 7.2.

Порядок следования полярных и цилиндрических координат выбирается так, чтобы якобиан был положителен. Если этот порядок изменить, то может получиться отрицательный якобиан.

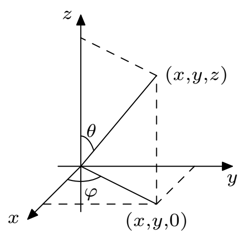

Рис. 7.3.

Сферическая система координат (физическая). Точка $(x, y, z) \in \mathbb{R}^3$ описывается расстоянием $R = \sqrt{x^2 + y^2 + z^2}$ от начала координат до этой точки, полярным углом θ и азимутальным углом φ согласно рис. 7.3. Из общих соображений ясно, что почти все точки пространства можно так охарактеризовать. При этом сферические координаты берутся в таком порядке: (R, θ, φ).

Диффеоморфизм Φ, действующий из (r, θ, φ) в (x, y, z) и задающий сферические координаты, дается равенствами

$$x = R\sin\theta\cos\varphi, \quad y = R\sin\theta\sin\varphi, \quad z = R\cos\theta.$$

Строго математически диффеоморфизм Φ определен на области $(0, +\infty) \times (0, \pi) \times (0, 2\pi)$. Образом этой области при отображении Φ является $\mathbb{R}^3$ за исключением оси Oz и вертикальной полуплоскости, проходящей через положительную часть оси абсцисс. Это связная область в $\mathbb{R}^3$.

Выпишем матрицу Якоби:

$$\frac{\partial\Phi}{\partial(R, \theta, \varphi)} = \begin{pmatrix} \sin\theta\cos\varphi & R\cos\theta\cos\varphi & -R\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & R\cos\theta\sin\varphi & R\sin\theta\cos\varphi \\ \cos\theta & -R\sin\theta & 0 \end{pmatrix}.$$

Легко найти, что ее определитель, т. е. якобиан сферической (физической) системы координат, равен $R^2\sin\theta$. Отметим, что он положителен.

Координатные поверхности в этой системе координат таковы: поверхности постоянства R — это сферы радиуса R. Поверхности $\theta = \text{const}$ суть конусы. Наконец, поверхности $\varphi = \text{const}$ представляют собой полуплоскости.

Сферическая система координат (географическая). Этой системой чаще пользуются математики. Она отличается от физической только тем, что в ней используется не полярный угол, а широта ψ, т. е. угол между плоскостью xOy и лучом, идущим из начала координат в данную точку. При этом немного меняются определяющие замену равенства:

$$x = R\cos\varphi\cos\psi, \quad y = R\sin\varphi\cos\psi, \quad z = R\sin\psi.$$

Здесь переменные берутся в таком порядке: R, φ, ψ, при этом область изменения переменных такова:

$$R > 0, \quad \varphi \in (0, 2\pi), \quad \psi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right),$$

ее якобиан равен $R^2\cos\psi$.

7.2. Замена переменных.

Пусть дана функция $f : \Omega \to \mathbb{R}$, где Ω — область в $\mathbb{R}^n$, и пусть $\Phi : U \to \Omega$, где U — область в $\mathbb{R}^n$, — диффеоморфизм. В таком случае в области Ω вместе с имевшимися координатами $(x^1, \dots, x^n)$ появляются новые координаты $(u^1, \dots, u^n)$, т. е. новый набор чисел, описывающих данную точку. Возникает композиция $\tilde{f}(u) = f(\Phi(u))$. Функция $\tilde{f}(u)$ — функция f, записанная в координатах $(u^1, \dots, u^n)$.

Поскольку часто в соотношениях участвуют производные функции f, важно выяснить, как связаны производные функции $\tilde{f}$ с производными функции f. Для этого надо записать тождество

$$\tilde{f}(u) = f(\Phi(u)),$$

связывающее старую и новую функции, и его дифференцировать по u:

$$\frac{\partial \tilde{f}}{\partial u^i}(u) = \sum_{j=1}^{n} \frac{\partial f}{\partial x^j}(\Phi(u))\frac{\partial \Phi^j}{\partial u^i}(u), \quad i = 1, \dots, n. \tag{$*$}$$

В матричном виде это выглядит так:

$$\begin{pmatrix} \frac{\partial \tilde{f}}{\partial u^1} \\ \vdots \\ \frac{\partial \tilde{f}}{\partial u^n} \end{pmatrix} = \begin{pmatrix} \frac{\partial \Phi^1}{\partial u^1} & \cdots & \frac{\partial \Phi^n}{\partial u^1} \\ \vdots & & \vdots \\ \frac{\partial \Phi^1}{\partial u^n} & \cdots & \frac{\partial \Phi^n}{\partial u^n} \end{pmatrix} \begin{pmatrix} \frac{\partial f}{\partial x^1} \\ \vdots \\ \frac{\partial f}{\partial x^n} \end{pmatrix}, \tag{$**$}$$

где в правой части равенства вектор умножается на транспонированную к матрице Якоби $\left(\frac{\partial\Phi}{\partial u}\right)^{\top}$. Правда, если векторы-столбцы слева и справа рассматривать как матрицы Якоби, то это тоже транспонированные матрицы Якоби. Тем самым $(**)$ можно записать так:

$$\left(\frac{\partial\tilde f}{\partial u}\right)^{\top} = \left(\frac{\partial\Phi}{\partial u}\right)^{\top}\left(\frac{\partial f}{\partial x}\right)^{\top}. \qquad (\dagger)$$

Посмотрим, как будут выражаться вторые производные. Продифференцируем равенство $(*)$. Имеем

$$\frac{\partial^2\tilde f}{\partial u^i\partial u^k} = \sum_{l=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 f(\Phi(u))}{\partial x^l\partial x^j}\frac{\partial\Phi^j(u)}{\partial u^i}\frac{\partial\Phi^l(u)}{\partial u^k} + \sum_{j=1}^{n}\frac{\partial f(\Phi(u))}{\partial x^j}\frac{\partial^2\Phi^j(u)}{\partial u^i\partial u^k}.$$

Пример 1 (запись оператора Лапласа в полярных координатах). Пусть $f:\mathbb{R}^2\to\mathbb{R}$ — функция класса $C^2(\mathbb{R}^2)$. *Оператором Лапласа*, примененным к функции f, называют выражение

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

В полярных координатах $x = r\cos\varphi$, $y = r\sin\varphi$ функция f переходит в функцию

$$\tilde f(r,\varphi) = f(x,y) = f(r\cos\varphi, r\sin\varphi).$$

Для нахождения связи между производными можно продифференцировать любое из тождеств

$$\tilde f(r,\varphi) = f(r\cos\varphi, r\sin\varphi), \quad \tilde f\left(\sqrt{x^2+y^2}, \operatorname{arctg}\frac{y}{x}\right) = f(x,y), \qquad (*)$$

характеризующих взаимосвязь между старой функцией f и новой $\tilde f$. В последнем тождестве выражение для φ через арктангенс, данное в формуле, верно для первого квадранта $x,y>0$, в других квадрантах к арктангенсу нужно прибавить некоторую константу, но ясно, что это не повлияет на производные, так что формулы для связи между производными будут одни и те же при всех допустимых x,y.

Дифференцирование любого из тождеств $(*)$ приведет к результату, разными могут оказаться трудозатраты для его достижения. На первый взгляд, первое равенство в $(*)$ предпочтительнее, так как там проще выражения. Однако при его дифференцировании в одном равенстве окажутся производные функции f по x и по y, стало быть, из одного равенства никакую из них выразить не удастся, и придется составлять и решать систему. Во втором тождестве при дифференцировании его, например, по x производная f'_x может быть выражена из одного равенства, поэтому будем дифференцировать второе. Имеем

$$\frac{\partial f}{\partial x} = \frac{\partial \tilde{f}}{\partial r}\frac{x}{\sqrt{x^2+y^2}} + \frac{\partial \tilde{f}}{\partial \varphi}\frac{-y}{x^2+y^2}.$$

Дифференцируем полученное равенство еще раз:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 \tilde{f}}{\partial r^2}\left(\frac{x}{\sqrt{x^2+y^2}}\right)^2 + 2\frac{\partial^2 \tilde{f}}{\partial r \partial \varphi}\frac{x}{\sqrt{x^2+y^2}}\frac{-y}{x^2+y^2}$$
$$+ \frac{\partial^2 \tilde{f}}{\partial \varphi^2}\left(\frac{-y}{x^2+y^2}\right)^2 + \frac{\partial \tilde{f}}{\partial r}\frac{\partial}{\partial x}\left(\frac{x}{\sqrt{x^2+y^2}}\right) + \frac{\partial \tilde{f}}{\partial \varphi}\frac{\partial}{\partial x}\left(\frac{-y}{x^2+y^2}\right).$$

Найдем вторую производную по y. Имеем

$$\frac{\partial f}{\partial y} = \frac{\partial \tilde{f}}{\partial r}\frac{y}{\sqrt{x^2+y^2}} + \frac{\partial \tilde{f}}{\partial \varphi}\frac{x}{x^2+y^2},$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 \tilde{f}}{\partial r^2}\left(\frac{y}{\sqrt{x^2+y^2}}\right)^2 + 2\frac{\partial^2 f}{\partial r \partial \varphi}\frac{y}{\sqrt{x^2+y^2}}\frac{x}{x^2+y^2}$$
$$+ \frac{\partial^2 \tilde{f}}{\partial \varphi^2}\left(\frac{x}{x^2+y^2}\right)^2 + \frac{\partial \tilde{f}}{\partial r}\frac{\partial}{\partial y}\left(\frac{y}{\sqrt{x^2+y^2}}\right) + \frac{\partial \tilde{f}}{\partial \varphi}\frac{\partial}{\partial y}\left(\frac{x}{x^2+y^2}\right).$$

Сложим полученные производные:

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 \tilde{f}}{\partial r^2} + \frac{\partial^2 \tilde{f}}{\partial \varphi^2}\frac{1}{x^2+y^2}$$
$$+ \frac{\partial \tilde{f}}{\partial r}\underbrace{\left(\frac{\partial}{\partial x}\left(\frac{x}{\sqrt{x^2+y^2}}\right) + \frac{\partial}{\partial y}\left(\frac{y}{\sqrt{x^2+y^2}}\right)\right)}_{=\frac{1}{\sqrt{x^2+y^2}}}$$
$$+ \frac{\partial \tilde{f}}{\partial \varphi}\underbrace{\left(\frac{\partial}{\partial x}\left(\frac{-y}{\sqrt{x^2+y^2}}\right) + \frac{\partial}{\partial y}\left(\frac{x}{\sqrt{x^2+y^2}}\right)\right)}_{=0}.$$

В итоге получаем

$$\frac{\partial^2 f}{\partial x^2}(x,y) + \frac{\partial^2 f}{\partial y^2}(x,y) = \frac{\partial^2 \tilde{f}}{\partial r^2} + \frac{1}{r}\frac{\partial \tilde{f}}{\partial r} + \frac{1}{r^2}\frac{\partial^2 \tilde{f}}{\partial \varphi^2}.$$

Сумму первых двух слагаемых в правой части можно записать так:

$$\frac{\partial^2 \tilde{f}}{\partial r^2} + \frac{1}{r}\frac{\partial \tilde{f}}{\partial r} = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \tilde{f}}{\partial r}\right).$$

Тем самым лапласиан в полярных координатах имеет вид

$$\frac{\partial^2 f}{\partial x^2}(x,y) + \frac{\partial^2 f}{\partial y^2}(x,y) = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \tilde{f}}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 \tilde{f}}{\partial \varphi^2}.$$

Заметим, что выражение оператора Лапласа в полярных координатах не позволяет посчитать его в точке $(0,0)$. Однако бывает потребность рассмотреть уравнение $\Delta f = \delta(x,y)$, более того, интерес может представлять поиск решения, зависящего только от r. Тогда полезно представить его в полярных координатах и рассмотреть уравнение $\Delta f(r,\varphi) = \delta(r,\varphi)$. Если действие происходит вне начала координат, то проблемы нет, все переносится. В точке $(0,0)$ появляется проблема с якобианом.

Пример 2. Найдем решение уравнения Лапласа $\Delta f(x,y) = 0$, зависящее только от $r = \sqrt{x^2+y^2}$. Иначе можно сказать: найти все сферически симметрические решения. Вообще задача ставится и решается в трехмерном пространстве, но пока мы записали лапласиан для двумерного случая, поэтому им и ограничимся. В полярных координатах уравнение примет вид

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \tilde{f}}{\partial r}\right) = 0.$$

Слагаемое $\dfrac{\partial^2 f}{\partial \varphi^2}$ обратилось в нуль по той причине, что ищется решение, не зависящее от φ. Далее, заметим, что $r > 0$, значит, уравнение становится таким:

$$\frac{\partial}{\partial r}\left(r\frac{\partial \tilde{f}}{\partial r}\right) = 0.$$

Отсюда следует, что

$$r\frac{\partial \tilde{f}}{\partial r} = C_1,$$

где C_1 — некоторая константа. Запишем его в виде

$$\frac{\partial \tilde{f}}{\partial r} = \frac{C_1}{r}.$$

Интегрируя последнее равенство, находим, что

$$F(r, \varphi) = C_1 \ln r + C_2.$$

Ясно, что поскольку в исходном уравнении были производные второго порядка, решение должно зависеть от двух произвольных постоянных.

Оглавление

Основы математического анализа
для студентов-физиков. Лекции.
5. Дифференциальное исчисление
функций многих переменных

Дятлов Глеб Владимирович

Подписано в печать 29.12.2014. Формат 60 × 84 1/16. Печать офсетная.
Усл. печ. л. 4.0. Уч.-изд. л. 4.0. Тираж 200 экз. Заказ №   .

Лицензия ЛР № 065614 от 8 января 1998 г.
Издательство Института математики,
пр. Академика Коптюга, 4, 630090 Новосибирск, Россия.

Отпечатано в ООО «Омега Принт»,
пр. Академика М. А. Лаврентьева, 6, 630090 Новосибирск, Россия.