

В. М. Тихомиров

Синтетический курс математики

Оглавление

Введение

Лекция 1. Теория линейных уравнений

Лекция 2. Решение нелинейных уравнений

Лекция 3. Правило множителей Лагранжа. Принцип компактности

Лекция 4. Теория линейных уравнений и геометрия

Лекция 5. Приведение квадратичных форм к главным осям. Доказательство теоремы Фредгольма

Лекция 6. Начала дифференциального исчисления и теория дифференциальных уравнений

Лекция 7. Определенный интеграл

Лекция 8. Интегрирование дифференциальных форм

Лекция 9. Ряды и комплексный анализ

Лекция 10. Начала теории вероятностей

Дополнения (приложения к алгебре, геометрии, анализу и к естествознанию)

Введение

С момента образования мех-мата, с 1933 года, преподавание математики было поделено на несколько фактически независимых областей. А далее все закоснело. Программа Государственных Экзаменов на мех-мате свыше пятидесяти лет не меняется, и читая ее особенно отчетливо видно это феодальное разделение нашей науки. Иногда речь идет о согласовании разных курсов, но фактически такого согласования добиться невозможно.

Андрей Николаевич Колмогоров стремился как-то нарушить это положение созданием своего синтетического курса. Синтетическими курсами я называю такие курсы, при которых воедино соединяются несколько математических направлений. Первым синтетическим курсом в университетском образовании был курс, вошедший в историю под названием «Анализа III». Его программа была разработана А. Н. Колмогоровым в сороковые и пятидесятые годы прошлого века. Он объединял в себе начала теории множеств, теорию действительных чисел, начала общей топологии, теорию функций, функциональный анализ, теорию меры, теорию интеграла, теорию интегральных уравнений и вариационное исчисление. Чуть позже, уже по инициативе Юрия Ивановича Манина и Сергея Петровича Новикова появились синтетические курсы по линейной алгебре и геометрии и по дифференциальной геометрии и топологии.

В чем мне видится смысл создания синтетических курсов?

При той системе образования, которая сложилась (причем не только у нас, но фактически во всем мире), наша наука — математика — представлена, как феодальное царство, как собрание фактически независимых провинций, почти не связанных друг с другом.

А на самом деле она едина. И в современное математическое образование, разумно, как мне кажется, включать объединяющие “синтетические” курсы.

Эту необходимость можно мотивировать и теми радикальными переменами, что произошли за последние два десятилетия в нашей стране и в мире.

Что же это за изменения? Я бы указал на три: *изменился государственный строй, изменились интересы молодежи и произошел неслыханной силы информационный взрыв.*

В наше время нет системы государственного трудоустройства, распределения выпускников, планирования профессиональной занятости. Не следует ли из этого, что в математическом образовании нужно выделить главное, наиболее существенное, чтобы дать возможность личности включаться в творческую деятельность в самых разных областях знаний? В представленном читателю курсе я стараюсь *мотивированно* выделить то самое главное, что ныне составляет математическое образование в техническом и экономическом вузе и (на несколько более высоком уровне) в университете.

Каждая лекция предваряется аннотацией, которая рассчитана на читателя, владеющего теми понятиями, которые в ней фигурируют. Читателю, не владеющий этими понятиями, разумно начать с основного текста, в нем все нужные понятия будут введены.

Лекция 1. Теория линейных уравнений

ВСЕ ЕСТЬ ЧИСЛО

Пифагор

В этой лекции речь пойдет о теории линейных уравнений. Будет рассказано о методе Гаусса решения системы n уравнений с n неизвестными. Далее рассказывается об альтернативе Фредгольма для таких систем и доказывается теорема Кронекера – Капелли. В заключительной части лекции приводится формулировка альтернативы Фредгольма в бесконечномерном случае. Эта теорема будет доказана в пятой лекции.

1.1. Конечномерный случай

Рассмотрим следующую систему уравнений:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= y_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= y_2, \\ &\dots \quad \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= y_n. \end{aligned} \tag{1}_n$$

Систему $(1)_n$ называют *системой n линейных алгебраических уравнений с n неизвестными*. Числа a_{ij} называются *коэффициентами системы уравнений*, числа y_i составляют *правую часть системы $(1)_n$* . Если все y_i равны нулю, система $(1)_n$ называется *однородной*. В этой лекции коэффициенты системы и правые части предполагаются вещественными числами.¹

Метод решения систем $(1)_n$ был предложен К. Ф Гауссом.

Метод Гаусса решения системы $(1)_n$ основан на идее *исключения неизвестных*. Применим метод математической индукции. Если $n = 1$, требуется решить уравнение

$$a_{11}x_1 = y_1. \tag{1}_1$$

Возможны три случая:

1) если число a_{11} не равно нулю, тогда уравнение $(1)_1$ имеет единственное решение $x = \frac{y_1}{a_{11}}$; 2) если и число a_{11} , и число y_1 равны нулю, тогда уравнение $(1)_1$ разрешимо, но неоднозначно (более того, каждое число является решением); 3) если же число a_{11} равно нулю, а число y_1 нулю не равно, тогда решений вообще нет.

Эти три тривиальных утверждения мы привели потому, что и в n -мерном и в бесконечномерном случае имеется точно такая же триада. Впрочем, пункт 2) потребует исследования вопроса о том, когда все-таки система (1) разрешима.

Пусть системы $(1)_{n-1}$ мы решать научились. Опишем метод решения систем $(1)_n$.

¹Короткий экскурс в теорию вещественного числа содержится в Дополнении

Если все коэффициенты a_{ij} уравнения $(1)_n$ равны нулю, то решение возможно лишь если $y_1 = y_2 = \dots = y_n = 0$ и им является любые n чисел x_1, x_2, \dots, x_n ; если не все y_i равны нулю система несовместна (т. е. не имеет решения).

Если же, в системе $(1)_n$, скажем, $a_{nn} \neq 0$, то выразим x_n из последнего уравнения через x_1, x_2, \dots, x_{n-1} и подставим полученное выражение в первые $n - 1$ уравнение. В итоге мы приходим к системе $(1)_{n-1}$, которую умеем решать по предположению индукции. Решив полученную систему (если система имеет решение), найдем $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_{n-1}$, а затем и \bar{x}_n . Если получившаяся система $(1)_{n-1}$ несовместна, то и изначальная система $(1)_n$ несовместна. \square

До сих пор мы использовали из математических средств лишь арифметику. В следующем разделе полезно пользоваться простейшими понятиями и обозначениями линейной алгебры.

Совокупность вещественных чисел или, как еще говорят, вещественную прямую, обозначают символом \mathbb{R} . Столбец из n чисел $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ называют *n -мерным вектором-столбцом*.

Совокупность всех n -мерных векторов-столбцов называют n -мерным (векторным) пространством \mathbb{R}^n . Далее для экономии места вектор-столбец с координатами x_1, \dots, x_n мы иногда записываем в строку $(x_1, \dots, x_n)^T$, где символ T означает *транспонирование* — замену столбцов строками (и наоборот).

Наряду с пространством \mathbb{R}^n векторов-столбцов иногда будем рассматривать совокупность \mathbb{R}^{n*} векторов-строк. Для вектора $y = (y_1, \dots, y_n)$ из \mathbb{R}^{n*} и вектора-столбца $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ из \mathbb{R}^n

определяется *внутреннее произведение* $y \cdot x = y_1 x_1 + \dots + y_n x_n$. Выражение $|x| = \sqrt{x^T \cdot x} = (x_1^2 + \dots + x_n^2)^{1/2}$ ($|y| = \sqrt{y \cdot y^T} = (y_1^2 + \dots + y_n^2)^{1/2}$) называется *модулем* (или *длиной*) вектора $x \in \mathbb{R}^n$ ($y \in \mathbb{R}^{n*}$).

Таблицу из коэффициентов $(a_{ij})_{1 \leq i, j \leq n}$ (такие таблицы называют *матрицами*) обозначим буквой A . Матрицу

$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ можно представить, как столбец из n векторов-строк из \mathbb{R}^{n*} ($a^1 = (a_{11}, \dots, a_{1n}), \dots, a^n = (a_{n1}, \dots, a_{nn})$).

Произведение матрицы A на вектор-столбец x (обозначаемое Ax) определяется, как вектор-столбец $(a^1 \cdot x, \dots, a^n \cdot x)^T$. Совокупность всех матриц с n столбцами и n строками обозначим M^{nn} . Таким образом, любая матрица A из M^{nn} задает *отображение* из \mathbb{R}^n в \mathbb{R}^n или, как еще говорят, матрица A задает *оператор* из \mathbb{R}^n в \mathbb{R}^n . Этот оператор обладает *линейным свойством*, заключающемся в том, что если x^1 и x^2 векторы из \mathbb{R}^n , α_1 и α_2 два числа и A матрица из M^{nn} , то имеет место равенство: $A(\alpha_1 x^1 + \alpha_2 x^2) = \alpha_1 A x^1 + \alpha_2 A x^2$. Потому A называют *линейным оператором*.

Теорема 1 (альтернатива² Фредгольма в конечномерном случае). *Имеет место одно из двух: или система уравнений $(1)_n$ разрешима для любой правой части, или однородное уравнение имеет ненулевое решение.*

Доказательство.

Лемма 1.1. *Система*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n+1}x_{n+1} &= 0, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n+1}x_{n+1} &= 0, \\ &\dots \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn+1}x_{n+1} &= 0 \end{aligned} \tag{1}_n$$

из n однородных уравнений с $n + 1$ -м неизвестным имеет ненулевое решение.

Доказательство леммы. Снова применяем метод математической индукции. Если в уравнении $a_{11}x_1 + a_{12}x_2 = 0$ (уравнении $(1_0)_1$) все коэффициенты равны нулю, то любые не равные одновременно нулю x_1 и x_2 будут искомым решением. Если же, скажем, $a_{12} \neq 0$, то одним из решений будет $x_2 = 1$, $x_1 = -\frac{a_{11}}{a_{12}}$.

Допустим теперь, что мы доказали лемму для системы $(1_0)_{n-1}$. Докажем лемму для системы $(1_0)_n$. Если все коэффициенты a_{ij} , $1 \leq i \leq n + 1$, $1 \leq j \leq n$ равны нулю, то любые не равные одновременно нулю числа $x_i \in \mathbb{R}$, $1 \leq i \leq n + 1$ будут искомым решением. А если одно из чисел, скажем, $a_{nn+1} \neq 0$ не равно нулю, тогда, выразив x_{n+1} из n -того уравнения через остальные переменные и подставив то, что получится вместо x_{n+1} , в первые $n - 1$ уравнение, приходим к системе $(1_0)_{n-1}$. Оно по предположению индукции имеет ненулевое решение $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_{n-1}$, а тогда через них выразим \hat{x}_n и тем самым придем к цели. \square

Переходим к доказательству альтернативы Фредгольма в конечномерном случае. В нашем случае утверждение можно разбить на две части: 1) если уравнение $(1)_n$ разрешимо для любой правой части, то однородное уравнение имеет только нулевое решение; 2) если однородное уравнение имеет только нулевое решение, то система разрешима $(1)_n$ при любой правой части.

Докажем первое утверждение. Допустим, что оно не справедливо и система $(1)_n$ разрешима для любых правых частей, но при этом существует вектор $b^1 \neq 0$ такой, что $Ab^1 = 0$. Найдем тогда вектор b^2 такой, что $Ab^2 = b^1$, и так далее, и, наконец, вектор b^{n+1} такой, что $Ab^{n+1} = b^n$. По лемме 1.1 найдутся числа $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_{n+1}$ не равные одновременно нулю и такие, что $\hat{x}_1b^1 + \hat{x}_2b^2 + \dots + \hat{x}_{n+1}b^{n+1} = 0$. Подействовав на это равенство матрицей A , используя линейное свойство матрицы A и определения столбцов b^1, b^2, \dots, b^{n+1} , получим, что $\hat{x}_2b^1 + \hat{x}_3b^2 + \dots + \hat{x}_{n+1}b^n = 0$. Подействовав теперь на это новое равенство оператором A , и поступая далее аналогично, получим, что $\hat{x}_{n+1}b^1 = 0$. Но мы предположили, что $b^1 \neq 0$, значит $\hat{x}_{n+1} = 0$. Аналогично докажем, что $\hat{x}_n = \hat{x}_{n-1} = \dots = \hat{x}_1 = 0$. Пришли к противоречию: мы ведь предполагали, что не все \hat{x}_i равны нулю. Значит, из разрешимости системы следует, что однородное уравнение имеет только нулевое решение.

Докажем второе утверждение. Пусть однородное уравнение имеет только нулевое реше-

²Слово «альтернатива» означает, что возможен лишь один из двух вариантов, взаимно исключающих друг друга; И. Фредгольм (1866 – 1927) — шведский математик.

ние. Применим лемму 1.1 к системе

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + y_1x_{n+1} &= 0, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + y_2x_{n+1} &= 0, \\ &\dots \quad \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n + y_nx_{n+1} &= 0 \end{aligned} \tag{1'_0}_n$$

n уравнений с $n + 1$ неизвестным x_1, x_2, \dots, x_{n+1} . Из леммы 1.1 следует, что найдутся числа $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_{n+1}$, не равные нулю одновременно и удовлетворяющие системе. **Но по условию число \hat{x}_{n+1} не может равняться нулю, ибо иначе однородное уравнение имело бы ненулевое решение.** Значит, положив $x_1 = -\frac{\hat{x}_1}{\hat{x}_{n+1}}, x_2 = -\frac{\hat{x}_2}{\hat{x}_{n+1}}, \dots, x_n = -\frac{\hat{x}_n}{\hat{x}_{n+1}}$, мы разрешаем систему $(1)_n$. Альтернатива Фредгольма доказана. \square

Переходим к условиям разрешимости систем уравнений.

Нам понадобится для этого определение ранга. Приведем его.

Пусть даны векторы a^1, \dots, a^k из \mathbb{R}^n . Скажем, что эти векторы *линейно зависимы*, если найдутся числа $\alpha_1, \dots, \alpha_k$ не равные нулю одновременно такие, что $\alpha_1 a^1 + \dots + \alpha_k a^k = 0$. Если таких чисел подобрать нельзя, векторы называются *линейно независимыми*. Если для вектора a^0 можно найти числа β_1, \dots, β_k такие, что $\beta_1 a^1 + \dots + \beta_k a^k = a^0$, то говорят, что a^0 *линейно выражается через $\{a^i\}_{1 \leq i \leq k}$* . Матрицу $A = (a_{ij})_{1 \leq i, j \leq n}$ назовем *разрешимой*, если уравнение $Ax = y$ разрешимо для любого $y \in \mathbb{R}^n$.

Определение 1. Пусть $C = (c_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \in M^{mn}$ (матрица с n столбцами и m строками). Максимальный размер разрешимой квадратной подматрицы матрицы C называют *рангом* этой матрицы (ранг матрицы C обозначают $\text{rang } C$). На лекции 4 это понятие будет дополнено.

Предложение 1. Для системы $(1)_n$ имеются три возможности:

1) если столбцы $a^1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \dots, a^n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix}$ линейно независимы, тогда имеется единственное решение системы $(1)_n$; 2) если столбцы a^1, \dots, a^n линейно зависимы и более того, столбец $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ линейно выражается через них, тогда система разрешима, но неоднозначно; 3) если столбцы a^1, \dots, a^n линейно зависимы, а столбец y через них не выражается, тогда решений вообще нет.

Доказательство. Утверждение а1) сразу следует из альтернативы Фредгольма: сказать, что векторы-столбцы a^1, \dots, a^n линейно независимы, равносильно тому, что однородная система имеет лишь нулевое решение. А тогда в силу альтернативы Фредгольма система $(1)_n$ разрешима. Если бы существовали два разных решения, то их разность была бы нетривиальным решением однородного уравнения. Отсюда следует единственность. Докажем утверждение 2). Из условий следует, что существуют отличный от нуля вектор $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_n)^T$ — решение однородной системы и вектор $\hat{x} = (\hat{x}_1, \dots, \hat{x}_n)^T$ — решение неоднородной системы,

т. е. $\tilde{x}_1 a^1 + \dots + \tilde{x}_n a^n = 0$ и $\hat{x}_1 a^1 + \dots + \hat{x}_n a^n = y$. Тогда вектор $\hat{x} + t\tilde{x}$ будет решением неоднородной системы при любом $t \in \mathbb{R}$, т. е. решение системы неоднозначно. Утверждение а3) — не что иное, как тавтология. \square

Теорема 2 (Кронекера – Капелли). *Для того, чтобы система $\sum_{j=1}^n c_{ij}x_j = b_i$, $1 \leq i \leq m$, m уравнений с n неизвестными была разрешима, необходимо и достаточно, чтобы ранг матрицы $C = (c_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ этой системы совпадал с рангом “расширенной” матрицы \tilde{C} , получаемой присоединением к матрице C столбца y правых частей.*

Находить решения возможно методом Гаусса.

Доказательство. Нам понадобятся две леммы.

Лемма 1.2. *Если матрица $A \in M^{k,k}$ разрешима, то столбцы этой матрицы линейно независимы, а любой вектор из \mathbb{R}^k единственным образом выражается через эти столбцы.*

Эта лемма сразу следует из определения разрешимости матрицы и альтернативы Фредгольма.

Лемма 1.3. *Ранг матрицы равен максимальному числу линейно независимых столбцов этой матрицы.*

Действительно, пусть матрица $C = (c_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ имеет ранг r . Не ограничив себя в общности, считаем, что разрешимая матрица размера $r \times r$ расположена в верхнем левом углу. Обозначим столбцы матрицы C через c^i , $1 \leq i \leq n$. То, что первые r столбцов матрицы C линейно независимы, следует из леммы 1.2. Допустим, что столбец c^k , $m+1 \leq k \leq n$ не выражается линейно через первые r столбцов. Из леммы 1.2 следует, что существует единственный набор чисел $\alpha_1(k), \dots, \alpha_r(k)$ такой, что $\sum_{1 \leq j \leq r} \alpha_j(k)c_{ij} = c_{ik}$, $1 \leq i \leq r$ (т. е. первые r координат вектора $\alpha_1(k)c^1 + \dots + \alpha_r(k)c^r$ совпадают с первыми r координатами вектора c^k). А предположение о том, что столбец c^k не выражается через первые r столбцов, означает, что для некоторого $r < l \leq n$ числа c_{lk} и $\alpha_1(j)c_{l1} + \dots + \alpha_r(k)c_{lr}$ не совпадают друг с другом. Это означает, что столбцы у матрицы размеров $(r+1) \times (r+1)$, получающейся присоединением столбца $(c_{1k}, \dots, c_{rk}, c_{lk})^T$ и строки, составленной из элементов c_{l1}, \dots, c_{lr} и c_{lk} , линейно независимы, и значит, в силу утверждения а1) теоремы 1.1 она разрешима, что противоречит определению ранга. Лемма доказана.

Возвратимся к доказательству теоремы Кронекера – Капелли. Если система разрешима, это означает, что вектор свободных членов выражается через столбцы матрицы, а значит, в силу леммы 1.1, он выражается через первые r столбцов, т. е. ранг расширенной матрицы равен r . А если система не разрешима, то это значит, что вектор y не выражается через столбцы матрицы C , т. е., в частности, столбцы c^1, \dots, c^r, y линейно независимы, т. е. ранг \tilde{C} больше ранга C . \square

Исторический комментарий. Решения линейных уравнений с одним неизвестным содержатся в первом тексте, где обсуждаются математические сюжеты — в древнеегипетском папирусе, за которым закрепилось название папируса Райнда, в честь английского египтолога, приобретшего его в 1858 году на базаре. Написание папируса Райнда относят к 18 веку до нашей эры, к эпохе Среднего царства. Теория линейных систем n уравнений с n неизвестными была создана К. Ф. Гауссом (1777 – 1855), Г. Грассманом (1809 – 1877), Г. Крамером (1704 – 1752), А. Кэли (1821 – 1895), К. Якоби (1804 – 1851) и другими математиками 18 и 19

веков. Тот фрагмент теории, который был изложен выше, состоит из двух частей — метода Гаусса решения системы и теоремы Кронекера – Капелли — условия разрешимости системы. Всегда с большим удовольствием сообщаю, что первым опубликовал теорему Кронекера – Капелли, не кто иной, как Чарлз Людвиг Доджсон (1832 –1898), псевдоним которого — Льюис Кэррол — известен каждому из нас. Это произошло в 1867 году. Как написано в одной исторической справке, в лекциях Л. Кронекера (1823 – 1791), читанных в 1864 г. “теория систем линейных уравнений получила, по существу, свое завершение”, (включая теорему, о которой мы ведем речь). А. Капелли (1855 – 1910) опубликовал эту теорему в 1892 г.

Эта тема будет продолжена в пятой лекции, где будет рассказано о других критериях разрешимости — правиле Крамера и “второй теореме” Фредгольма, и тогда завершится путь от задач из папируса Райнда до начала двадцатого века.

1.2. Бесконечномерный случай (формулировка альтернативы Фредгольма)

Сформулируем здесь (а докажем потом в пятой лекции) ослабленный вариант альтернативы Фредгольма. Альтернатива Фредгольма принадлежит к одной из вершин университетского образования. В этом курсе лекций она будет доказана не в самом общем виде, не в банаховом пространстве, а лишь в специальном случае — в пространстве l_2 . Но именно этот результат доказывается в учебнике по функциональному анализу Колмогорова и Фомина и читается на третьем курсе мех-мата МГУ. Доказательство основного утверждения — если система уравнений разрешима, то однородное уравнение будет иметь лишь нулевое решение — будет лишь незначительной надстройкой над теми рассуждениями, которые были проведены в предыдущем пункте. А обратное утверждение, по сути дела геометрическое, разумно проводить в геометрическом разделе. Потому и сама теорема доказывается в этом разделе.

Пусть l_2 – множество векторов-столбцов $x = (x_1, x_2, \dots)^T$ таких, что $\sum_{i \in \mathbb{N}} x_i^2 < \infty$, $A = (a_{ij})_{1 \leq i, j \leq \infty}$ – бесконечная матрица. Требуется решить такую систему бесконечного числа уравнений с бесконечным числом неизвестных:

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots = b_1 \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots = b_n \\ \dots \end{array} \Leftrightarrow \sum_{j \geq 1} a_{ij}x_j = b_i, \quad i \geq 1. \quad (1)_\infty$$

Альтернатива Фредгольма в бесконечномерном случае. Пусть в $(1)_\infty$, матрица $A = (a_{ij})_{i, j \in \mathbb{N}}$ представляет собой сумму $I + B$ единичной матрицы и матрицы $B = (b_{ij})_{i, j \in \mathbb{N}}$, удовлетворяющей условию $\|B\| = (\sum_{i, j \in \mathbb{N}} b_{ij}^2)^{1/2} < \infty$. Тогда имеет место альтернатива: или система уравнений $(1)_\infty$ разрешима для любого $b \in l_2$, или однородная система имеет ненулевое решение.

Доказательство этой теоремы будет приведено в пятой лекции.

Лекция 2. Решение нелинейных уравнений

МНЕ ПРЯМО СТЫДНО СКАЗАТЬ, ДО КАКОГО ЧИСЛА ЗНАКОВ Я ДОВЁЛ НА ДОСУГЕ ЭТИ ВЫЧИСЛЕНИЯ.

И. НЬЮТОН

В этой главе описывается модифицированный метод Ньютона решения нелинейных уравнений.

2.1. Конечномерный случай

Нам надлежит научиться решать уравнения вида $F(x) = y$, где аргументом является вектор $x \in \mathbb{R}^n$, y это вектор из \mathbb{R}^m , а $F = (f_1, \dots, f_n)^T$ — отображение из \mathbb{R}^n в \mathbb{R}^m (или, как пишут, $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$). В развернутом виде это означает, что надо решить систему уравнений

$$\begin{aligned} f_1(x_1, x_2, \dots, x_n) &= y_1, \\ f_2(x_1, x_2, \dots, x_n) &= y_2, \\ &\dots \quad \dots \\ f_m(x_1, x_2, \dots, x_n) &= y_m \end{aligned} \tag{2}_{nm},$$

где f_i , $1 \leq i \leq m$ — функции n переменных.

Среди отображений из \mathbb{R}^n в \mathbb{R}^m важную роль играют *линейные отображения*, фрагмент теории которых был построен в лекции 1. Напомним, что при линейном отображении $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ вектору $x = (x_1, \dots, x_n)^T$ ставится в соответствие вектор $Ax = (a_{11}x_1 + \dots + a_{1n}x_n, \dots, a_{m1}x_1 + \dots + a_{mn}x_n)^T$.

Система m линейных уравнений с n неизвестными $Ax = y$ определяется матрицей $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ и вектором $y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$.

При $m < n$, если система уравнений $Ax = y$ разрешима при любом y , у уравнения $(2)_{nm}$ при конкретном y существует, вообще говоря, множество решений, т. е. прообраз y не однозначен. Но в этом случае можно построить *правое обратное* отображение.

Лемма 2.1 (о правом обратном отображении к линейному отображению в конечномерном случае). Пусть $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ — сюръективное линейное отображение (т. е. такое линейное отображение, что уравнение $Ax = y$ разрешимо для любого $y \in \mathbb{R}^m$). Тогда существуют линейное отображение $R : \mathbb{R}^m \rightarrow \mathbb{R}^n$ (правое обратное к A) и константа $\gamma > 0$ такие, что $AR(y) = y$ и $|R(y)| \leq \gamma|y|$ для любого $y \in \mathbb{R}^m$.

Доказательство. Рассмотрим систему векторов $e_1 = (1, 0, \dots, 0)^T, \dots, e_m = (0, \dots, 0, 1)^T$ в \mathbb{R}^m .³ По условию существуют векторы $\{f_k\}_{k=1}^m$, $f_k \in \mathbb{R}^n$ $1 \leq k \leq m$ такие, что $Af_k = e_k$. Для

³В лекции 1 индексы векторов ставились наверху, чтобы не путать векторы с координатами. Начиная с этого места индексы векторов ставятся, как это принято, внизу (чтобы не путалось обозначение вектора a^k с k -той степенью числа a), а координаты вектора a_k имеют два индекса.

каждого вектора $y = \sum_{k=1}^m y_k f_k$ (здесь y_k — числа) положим $R(y) = \sum_{k=1}^m y_k f_k$. Тогда равенство $AR(y) = y$ следует из линейности оператора A , а второе свойство следует из очевидных оценок (и того, что длина вектора не меньше любой его координаты): $|Ry| \leq \sum_{i=1}^m |y_i| |f_i| \leq \max_{1 \leq k \leq m} |y_k| \sum_{k=1}^m |f_k| \leq \gamma |y|$, где $\gamma = \sum_{k=1}^m |f_k|$.

Метод решения системы $(2)_{nm}$ m уравнений с n неизвестными, о котором сейчас будет рассказано, основан на линеаризации, — замене нелинейного уравнения линейным.

Пусть задан вектор y . Возьмем вектор $x_0 = (x_{10}, \dots, x_{n0})^T$, для которого $F(x_0)$ близко по модулю к y и пусть A — сюръективный линейный оператор из \mathbb{R}^n в \mathbb{R}^m . В качестве приближенного решения уравнения $F(x) = y$ возьмем решение x_1 системы линейных уравнений $A(x - x_0) + F(x_0) = y$, получающееся с помощью леммы 2.1 о правом обратном: $x_1 = x_0 + R(y - F(x_0))$. А далее будем поступать аналогично, полагая:

$$x_k = x_{k-1} + R(y - F(x_{k-1})), \quad k = 1, 2, \dots \quad (2.1)$$

Определение 2. Итеративную процедуру (2.1) называют *модифицированным методом Ньютона*.

Теорема 3 (о правом обратном отображении в конечномерном случае). Пусть заданы вектор $x_0 \in \mathbb{R}^n$, число $\delta > 0$, сюръективный линейный оператор A из \mathbb{R}^n в \mathbb{R}^m и отображение $F : U_{\mathbb{R}^n}(x_0, \delta) \rightarrow \mathbb{R}^m$ (где $U_{\mathbb{R}^n}(x_0, \delta) = \{x \in \mathbb{R}^n \mid |x - x_0| < \delta\}$ — открытый шар в \mathbb{R}^n с центром в x_0 радиуса $\delta > 0$). Если существуют такое число θ , $0 < \theta < 1$, что для любых $\xi, x \in U_{\mathbb{R}^n}(x_0, \delta)$ выполняется неравенство:

$$|F(\xi) - F(x) - A(\xi - x)| \leq \frac{\theta}{\gamma} |\xi - x|, \quad \forall \xi, x \in U_{\mathbb{R}^n}(x_0, \delta), \quad (2.2)$$

где число γ взято из леммы 2.1 о правом обратном, то для каждого $y \in U_{\mathbb{R}^m}(F(x_0), \delta_0)$, где $\delta_0 = \delta(1 - \theta)/\gamma$, последовательность (2.1) модифицированного метода Ньютона сходится к такому вектору $x(y) \in U_{\mathbb{R}^n}(x_0, \delta)$, что $F(x(y)) = y$ и при этом $|x(y) - x_0| \leq K|y - F(x_0)|$, где $K = \gamma/(1 - \theta)$.

Отображение $y \rightarrow x(y)$ из $U_{\mathbb{R}^m}(F(x_0), \delta_0)$ в $U_{\mathbb{R}^n}(x_0, \delta)$, построенное в этой теореме, называется *правым обратным отображением к F* .

Доказательство опирается на понятия непрерывности функции n переменных, сходимости векторов и фундаментальной последовательности векторов из \mathbb{R}^n . Напомним соответствующие определения. Пусть F отображает шар $U_{\mathbb{R}^n}(\xi, r)$ с центром в ξ радиуса r в \mathbb{R}^m . Говорят, что отображение F непрерывно в точке ξ , если для любого числа $\varepsilon > 0$ найдется такое число $\delta > 0$, что $|F(\xi + x) - F(\xi)| < \varepsilon$, если только $|x| < \delta$.

Определение 3. Говорят, что последовательность $\{\xi_1, \xi_2, \dots, \xi_k, \dots\}$ векторов из \mathbb{R}^n сходится к вектору $\xi \in \mathbb{R}^n$ (и пишут $\xi_k \rightarrow \xi$ при $k \rightarrow \infty$ или $\lim_{k \rightarrow \infty} \xi_k = \xi$), если для любого числа $\varepsilon > 0$ найдется такое число N , что $|\xi_k - \xi| < \varepsilon$, как только $k > N$. Последовательность $\{\xi_1, \xi_2, \dots, \xi_k, \dots\}$ векторов из \mathbb{R}^n называется *фундаментальной*, если для любого числа $\varepsilon > 0$ найдется такое число N , что $|\xi_k - \xi_m| < \varepsilon$, как только $k, m > N$.

Числовой ряд $a_1 + a_2 + \dots + a_k + \dots$ называется *сходящимся*, если последовательность $S_1 = a_1, S_2 = a_1 + a_2, \dots, S_k = a_1 + \dots + a_k, \dots$ является сходящейся.

Определение 3 в дальнейшем будет уточняться.

В доказательстве будут использованы следующие факты: 1) если отображение F отображает шар $U_{\mathbb{R}^n}(\xi, r)$ в \mathbb{R}^m и является непрерывным в точке ξ , а последовательность $\{\xi_1, \xi_2, \dots, \xi_k, \dots\}$ векторов из \mathbb{R}^n сходится к вектору $\xi \in \mathbb{R}^n$, то последовательность $\{F(\xi_1), F(\xi_2), \dots, F(\xi_k), \dots\}$ векторов из \mathbb{R}^m сходится к вектору $F(\xi) \in \mathbb{R}^m$; 2) предел суммы последовательностей равен сумме пределов последовательностей и предел произведения последовательности на число равен произведению числа на предел последовательности (если $\xi_k \rightarrow \xi$, а $\xi'_k \rightarrow \xi'$, а a — число, то $\xi_k + \xi'_k \rightarrow \xi + \xi'$, а $a\xi_k \rightarrow a\xi$); 3) если $0 < \theta < 1$, то $1 + \theta + \dots + \theta^k = \frac{1-\theta^{k+1}}{1-\theta}$ (формула для суммы геометрической прогрессии); 4) всякая фундаментальная последовательность сходится.

Если первые два утверждения легко следуют из определений, а третье вытекает из непосредственно проверяемого тождества $(1 + \theta + \dots + \theta^k)(1 - \theta) = 1 - \theta^{k+1}$, то четвертое является следствием теории вещественных чисел, с изложения которой обычно начинается курс математического анализа. Об этом речь идет в Дополнении в разделе **Числа**.

Доказательство теоремы 3. Докажем, что а) векторы x_k построенные модифицированным методом Ньютона, для всех $k \geq 0$ лежат в $U_{\mathbb{R}^n}(x_0, \delta)$ и б) что последовательность $\{x_k\}_{k \geq 0}$ фундаментальна. Утверждение а) докажем по индукции. Начальный элемент x_0 принадлежит $U_{\mathbb{R}^n}(x_0, \delta)$ по определению. Пусть $x_s \in U_{\mathbb{R}^n}(x_0, \delta)$, $1 \leq s \leq k$. Применение к равенству (2.1) оператор A , с использованием линейного свойства этого оператора и равенства $AR(y) = y$ для любого y из \mathbb{R}^m , приходим к уравнению:

$$A(x_k - x_{k-1}) = y - F(x_{k-1}). \quad (2.3)$$

Учитывая равенства (2.3), неравенство $|AR(y)| \leq \gamma(y)$ и (2.2), получим:

$$|x_{k+1} - x_k| = |R(y - F(x_k))| \leq \gamma|y - F(x_k) - y + F(x_{k-1}) + A(x_k - x_{k-1})| \leq \theta|x_k - x_{k-1}|.$$

Повторив еще $k - 1$ раз это рассуждение, приходим к неравенству

$$|x_{k+1} - x_k| \leq \theta|x_k - x_{k-1}| \leq \theta^2|x_{k-1} - x_{k-2}| \leq \dots \leq \theta^n|x_1 - x_0|. \quad (2.4)$$

Используя в следующей выкладке неравенство для модулей, неравенство (2.4), формулу для суммы геометрической прогрессии и соотношение (2.3) для $k = 1$, а также выбор y , получаем:

$$|x_{k+1} - x_0| \leq |x_{k+1} - x_k| + \dots + |x_1 - x_0| \leq (\theta^k + \theta^{k-1} + \dots + 1)|x_1 - x_0| < \frac{\gamma}{1-\theta}|y - F(x_0)| < \delta \quad (2.5)$$

т. е. $x_{r+1} \in U_{\mathbb{R}^n}(x_0, \delta)$ и значит, все x_k принадлежат $U_{\mathbb{R}^n}(x_0, \delta)$. Утверждение а) доказано.

Докажем б). Для любых $k, s \in \mathbb{N}$ имеем (в силу неравенств для модулей): $|x_{k+s} - x_s| \leq |x_{k+s} - x_{k+s-1}| + \dots + |x_{k+1} - x_k| \leq (\theta^{k+s-1} + \dots + \theta^k)|x_1 - x_0| \leq \frac{\gamma\theta^k}{1-\theta}|y - F(x_0)| < \delta\theta^k$, откуда вытекает, что $\{x_n\}_{n \in \mathbb{N}}$ фундаментальная последовательность. Значит она сходится. Обозначим ее предел $x(y)$. Переход к пределу в (2.3) (который существует из-за непрерывности F в $U_{\mathbb{R}^n}(x_0, \delta)$) (совсем простое доказательство непрерывности предоставим читателю) приводит к равенству $F(x(y)) = y$, а переход к пределу в (2.5) обеспечивает неравенство $|x(y) - x_0| \leq K|y - F(x_0)|$ с $K = \frac{\gamma}{1-\theta}$. \square

2.2. Бесконечномерный случай.

В бесконечномерном случае вместо \mathbb{R}^n и \mathbb{R}^m фигурируют *банаховы пространства*.

Напомним определение банахова пространства. Множество X называется (вещественным) *векторным пространством*, если в нем определены две операции – суммы векторов и произведения вектора на вещественное число. Эти операции удовлетворяют естественным свойствам чисел и конечномерных векторов, таким, как законы коммутативности и ассоциативности сложения, существование нулевого элемента, противоположного элемента и т. п.

Векторное пространство X называется *нормированным*, если определена функция $\|\cdot\|_X : X \rightarrow \mathbb{R}_+$, сопоставляющая вектору x неотрицательное число $\|x\|_X$, называемое *нормой* вектора x , удовлетворяющее аксиомам: 1) $\|x\|_X \geq 0$ для любого $x \in X$ и $\|x\|_X = 0$ тогда и только тогда, когда $x = 0$, 2) $\|\alpha x\|_X = |\alpha| \|x\|_X \forall x \in X, \alpha \in \mathbb{R}$, 3) $\|x+x'\|_X \leq \|x\|_X + \|x'\|_X$. Такие пространства мы обозначаем $X = (X, \|\cdot\|_X)$. Через $U_X(\hat{x}, r) = \{x \in X \mid \|x - \hat{x}\|_X < r\}$ обозначим открытый шар в X , а через $B_X(\hat{x}, r) = \{x \in X \mid \|x - \hat{x}\|_X \leq r\}$ – замкнутый шар в X с центром в \hat{x} радиуса r . Подмножество U пространства $(X, \|\cdot\|_X)$ называется *открытым*, если для любой точки $x \in U$ найдется открытый шар с центром в этой точке целиком расположенном в U . Множество дополнительное к открытому множеству называется *замкнутым*. Наличие нормы позволяет определить понятия сходимости последовательности и непрерывности отображения.

Определение 3'. Говорят, что последовательность векторов $\{x_n\}_{n \in \mathbb{N}}, x_n \in X$ сходится к вектору \hat{x} и пишут $\lim_{n \rightarrow \infty} x_n = \hat{x}$, если для любого положительного числа ε найдется число N такое, что неравенство $\|x_n - \hat{x}\|_X < \varepsilon$ выполняется для любого $n > N$.

Пусть $Y = (Y, \|\cdot\|_Y)$ – другое нормированное пространство и $F : U_X(\hat{x}, r) \rightarrow Y$. Говорят, что это отображение F непрерывно в точке \hat{x} , если для любого числа $\varepsilon > 0$ найдется такое число $\delta > 0$, что $\|F(\hat{x} + x) - F(\hat{x})\|_Y < \varepsilon$, если $\|x\|_X < \delta$.

Из этого определения непрерывности функции сразу следует, что если функция F непрерывна в точке \hat{x} , то для любой последовательности $\{x_n\}_{n \in \mathbb{N}}$, сходящейся к \hat{x} предел последовательности $\{F(x_n)\}_{n \in \mathbb{N}}$ равен $F(\hat{x})$.

Последовательность $\{x_n\}_{n \in \mathbb{N}}$ векторов из X называется *фундаментальной* или *последовательностью Коши*, если для любого числа $\varepsilon > 0$ найдется такое число N , что неравенство $\|x_{n+k} - x_n\| < \varepsilon$ выполнено для любого натурального $n \geq N$ и любого натурального $k \in \mathbb{N}$.

Говорят, что нормированное пространство *обладает свойством полноты*, если в нем любая фундаментальная последовательность имеет предел. Полное нормированное пространство называется *банаховым*.

Пространство $X = \mathbb{R}^n, n \geq 1$ с нормой $\|x\|_X = |x| = (\sum_{k=1}^n x_k^2)^{\frac{1}{2}}$, а также пространство $C([a, b])$ непрерывных функций на $[a, b]$ с нормой $\|x(\cdot)\|_{C([a, b])} = \max_{x \in [a, b]} |x(t)|$ являются банаховыми пространствами.

Вернемся к вопросу о решении уравнений, но уже в банаховых пространствах.

Лемма 2.2 (о правом обратном отображении в бесконечномерном случае). Пусть X и Y – банаховы пространства $A : X \rightarrow Y$ – линейное непрерывное сюръективное отображение.⁴ Тогда существуют отображение $R : Y \rightarrow X$ (правое обратное) и константа $\gamma > 0$ такие, что $AR(y) = y$ и $\|R(y)\|_X \leq \gamma \|y\|_Y \quad \forall y \in Y$.

Доказательство. Эта лемма равносильна одному из основных принципов линейного анализа – принципу Банаха об открытости.

Согласно принципу Банаха об открытости (см. учебник Колмогорова и Фомина гл. IV, п.5), образ единичного открытого шара в X с центром в нуле, при отображении A содержит открытый шар в Y с центром в нуле радиуса δ . Таким образом, для любого элемента $y \in Y$, по норме меньшего δ , найдется элемент $x(y) \in X$, по норме меньший единицы, такой, что $Ax(y) = y$. Положим $R(y) = \frac{2\|y\|_Y}{\delta} x(\frac{\delta}{2\|y\|_Y} y), R(0) = 0$. Проверка свойств R тривиальна. \square

⁴Это означает, что уравнение $Ax = y$ разрешимо для любого $y \in Y$.

Теорема 3' о (правом) обратном отображении в бесконечномерном случае. Пусть X и Y — банаховы пространства, $F : U_X(x_0, \delta) \rightarrow Y$ (где $U_X(x_0, \delta) = \{x \in X \mid \|x - x_0\|_X < \delta\}$ — открытый шар в X с центром в x_0 радиуса $\delta > 0$) и при этом существуют такое линейное, непрерывное и сюръективное отображение $A : X \rightarrow Y$, а также число θ , $0 < \theta < 1$, что выполняется неравенство:

$$\|F(\xi) - F(x) - A(\xi - x)\|_Y \leq \frac{\theta}{\gamma} \|\xi - x\|_X, \quad \forall \xi, x \in U_X(x_0, \delta), \quad (2.6)$$

где число γ взято из леммы 2.2. Тогда для каждого $y \in U_Y(F(x_0), \delta_0)$, где $\delta_0 = \delta(1 - \theta)/\gamma$, последовательность модифицированного метода Ньютона

$$x_n = x_{n-1} + R(y - F(x_{n-1})), \quad n \in \mathbb{N} \quad (2.7)$$

сходится к такому элементу $x(y) \in B_X(x_0, \delta)$, что $F(x(y)) = y$ и при этом $\|x(y) - x_0\|_X \leq K\|y - F(x_0)\|_Y$, где $K = \gamma/(1 - \theta)$.

Доказательство этого утверждения является дословным повторением доказательства теоремы 3 и с заменой $|x|$ при $x \in X$ на $\|x\|_X$ и $|y|$ при $y \in Y$ на $\|y\|_Y$. \square

Исторический комментарий. Методы приближенного нахождения корней отдельных уравнений применялись еще в Древней Греции. Считается, что первый алгоритм нахождения корня уравнения $x^2 = 2$ был известен в первом веке нашей эры. Этот алгоритм приписывают Герону. Алгоритм Герона описывается так: надо выбрать любую дробь x_0 и затем использовать такую итеративную последовательность: $x_k = \frac{1}{2}(x_{k-1} + \frac{2}{x_{k-1}})$, $r = 1, 2, \dots$. История методов решения общих нелинейных уравнений восходит к 1676 году, к ответу Ньютона на письмо к нему секретаря Королевского общества Ольденбурга, в котором тот обратился к Ньютону за некоторыми разъяснениями. Ньютон на примере решения уравнения $F(x) = 0$, где $F(x) = x^3 - 2x - 5$, изложил метод, который в современных обозначениях представляет собой итеративную процедуру $x_k = x_{k-1} - (F'(x_{k-1}))^{-1}F(x_{k-1})$, $k = 1, 2, \dots$, где F' — это производная F , а начальная точка выбирается вычислителем, исходя из каких-то соображений целесообразности. Примерно в те же времена Ньютон научился разлагать функции в ряды. Комбинируя эти два метода Ньютон составил таблицы многих элементарных функций. О точности этих таблиц свидетельствует эпиграф в начале лекции.

Доказанная нами заключительная теорема принадлежит американскому математику Л.Грейвсу (1896 – 1973). Она была опубликована в 1950 г. Таким образом, в этой лекции мы совершили экскурс в теорию обратных отображений от Ньютона до середины прошлого столетия.

Доказанные нами теоремы об обратных отображениях играют в математике огромную роль. Расскажем здесь о приложении этих теорем к доказательству правила множителей Лагранжа (1736 – 1813).

Лекция 3. Правило множителей Лагранжа. Принцип компактности

Можно высказать следующий общий принцип. Если ищется максимум или минимум некоторой функции многих переменных при условии, что между этими переменными имеется связь, задаваемая одним или несколькими уравнениями, нужно прибавить к функции, о которой говорилось, функции, задающие уравнения связи, умноженные на неопределенные множители, и искать затем максимум или минимум построенной суммы, как если бы переменные были независимы. Полученные уравнения, присоединенные к уравнениям связи, послужат для определения всех неизвестных.

Ж.-Л. Лагранж

Эта лекция посвящена методу Лагранжа решения задач на максимум и минимум и основному принципу существования решений задач таких задач — принципу компактности.

3.1. Правило множителей Лагранжа. Термины “максимум” и “минимум” объединяются общим термином “экстремум”, и задачи на максимум и минимум называют *экстремальными задачами*.

Экстремальные задачи стали исследовать на заре развития математики. Они встречаются в трудах многих математиков древности, например, в сочинениях Зенона, Евклида, Архимеда, Аполлония, Герона. И в новое время (в 16 и 17 веках) было решено множество экстремальных задач.

Вплоть до 17 века каждая экстремальная задача решалась “индивидуально”, приемом специально придуманным именно для нее. Первый общий метод нахождения экстремумов функций был предложен Ферма (1600 – 1665).

Ферма рассматривал функции одного переменного (даже, собственно говоря, полиномы одного переменного) и дал набросок доказательства утверждения, которое сейчас формулируется так: *если в точке локального экстремума функция дифференцируема, то ее производная равна нулю*. В начале двадцатого века его теорема была распространена на дифференцируемые функции, определенные на нормированном пространстве. Этот результат также называют *теоремой Ферма*.

Пусть X — нормированное пространство, V — окрестность точки \hat{x} и f — функция, определенная на V . Говорят, что точка \hat{x} является *точкой локального максимума (минимума) функции* f , если найдется такое открытое множество $U \subset V$ в котором выполнено неравенство $f(x) \leq f(\hat{x})$ ($f(x) \geq f(\hat{x})$) для всякого $x \in U$. Точка \hat{x} является *абсолютным максимумом (минимумом) функции* f на некотором множестве, если для любой точки x из этого множества выполнено неравенство $f(x) \leq f(\hat{x})$ ($f(x) \geq f(\hat{x})$).

Напомним понятие дифференцируемости отображения, определенного в нормированном пространстве.

Определение 4. Пусть X и Y — нормированные пространства, V — окрестность точки $\hat{x} \in X$ и $F : V \rightarrow Y$. Говорят, что F *дифференцируемо в точке* \hat{x} (и пишут $f \in D^1(\hat{x})$), если существует линейный непрерывный оператор Λ из X в Y такой, что для всех $x \in X$, для которых $\hat{x}+x \in V$ справедливо представление $F(\hat{x}+x) = F(\hat{x}) + \Lambda x + r(x)$, где $\|r(x)\|_Y / \|x\|_X \rightarrow 0$ при $\|x\|_X \rightarrow 0$. Оператор Λ (определяемый этим представлением однозначно) называется *производной отображения* F в точке \hat{x} и обозначается $F'(\hat{x})$. Из определения производной в \hat{x} следует, что отображение должно быть непрерывным в этой точке. Отображение $x \rightarrow F'(x)x$

называется *дифференциалом* F в \hat{x} ; его называют *главной линейной частью* отображения F в окрестности точки \hat{x} .

Отдельно обсудим вопрос о дифференцируемости функций. Пусть X и нормированное пространство, V — окрестность точки $\hat{x} \in X$ в X и $f : V \rightarrow \mathbb{R}$ — функция, определенная на V . Согласно общему определению, производной $f'(\hat{x})$ функции f в точке \hat{x} является линейный непрерывный функционал, определенный на X . Совокупность всех таких функционалов называют пространством, сопряженным с X и обозначают X^* . В одномерном случае производная $f'(\hat{x})$ это число, для которого имеется равносильное определение: $f'(\hat{x}) = \lim_{x \rightarrow \hat{x}} \frac{f(x) - f(\hat{x})}{x - \hat{x}}$. В пространстве \mathbb{R}^n производная — это вектор-строка $f'(\hat{x}) = (y_1, \dots, y_n)$, где y_i — это производная в нуле функции $t \rightarrow f(\hat{x} + te_i)$, $1 \leq i \leq n$ (e_i — вектор, у которого i -тая координата равна единице, остальные нули). Число $y_i = \lim_{t \rightarrow 0} \frac{f(\hat{x} + te_i) - f(\hat{x})}{t}$ называют *частной производной функции f в точке \hat{x} по x_i* . Ее обозначают $\frac{\partial f(\hat{x})}{\partial x_i}$. Наличие частных производных в точке не гарантирует дифференцируемости в этой точке.

Это показывает пример функции двух переменных (x_1, x_2) , равной нулю всюду, кроме множества $\{(x_1, x_2) \mid x_1 = \sqrt{x_2}, x_2 > 0\}$. Эта функция разрывна в нуле, т. е. не дифференцируема, в то время, как ее частные производные по x_1 и x_2 существуют и равны нулю.

Имеет место

Теорема Ферма (необходимое условие минимума дифференцируемой функции). *Если в точке локального экстремума функция дифференцируема, то производная в этой точке равна нулю.*

Этот результат легко выводится из определения дифференцируемости.

Равенство нулю производной в точке, где предполагается локальный экстремум, называют *условием стационарности*.

Поставим проблему о задаче на экстремум при ограничениях типа равенств. Пусть на открытом подмножестве V нормированного пространства X определены функция $f_0 : V \rightarrow \mathbb{R}$, отображение $F : V \rightarrow Y$ и *требуется найти (абсолютный) максимум (минимум) $f_0(x)$ при условии, что выполнено равенство $F(x) = 0$* . Будем записывать эту проблему так:

$$f_0(x) \rightarrow \max(\min), \quad F(x) = 0. \quad (P)$$

Говорят, что \hat{x} — *точка локального максимума (минимума) для задачи (P)*, если найдется такой открытый шар $U_X(\hat{x}, \delta)$, что для любой точки x из этого шара, для которой $F(x) = 0$, выполнено неравенство $f_0(x) \leq f_0(\hat{x})$ ($f_0(x) \geq f_0(\hat{x})$). Точка \hat{x} является *абсолютным максимумом (минимумом) функции f в задаче (P)*, если для любой точки $x \in V$, для которой $F(x) = 0$, выполнено неравенство $f_0(x) \leq f_0(\hat{x})$ ($f_0(x) \geq f_0(\hat{x})$).

Метод исследования задач (P) был описан Лагранжем. Слова Лагранжа из его книги “Теория аналитических функций” (1797) были приведены в эпиграфе к этой лекции. Мы позволим себе лишь незначительно изменить рецепт Лагранжа. Во-первых, мы будем рассматривать не только функции многих переменных, но даже функции, определенные на нормированном пространстве. Во-вторых, оказывается удобным и функцию, экстремум которой ищется, также умножить на неопределенный множитель и тем самым составлять выражение

$\mathcal{L}(x, \bar{\lambda}) = \lambda_0 f_0(x) + \langle \lambda, F(x) \rangle$, где $\bar{\lambda} = (\lambda_0, \lambda) \in \mathbb{R}_+ \times X^*$, которое стали называть *функцией Лагранжа* (а число λ_0 и элемент λ из сопряженного пространства X^* — множителями Лагранжа). Наконец, мы не будем “искать минимум или максимум” функции Лагранжа, а просто будем применять к задаче на экстремум функции Лагранжа необходимое условие экстремума «как если бы переменные были независимы», т. е. теорему Ферма.

Для того, чтобы сформулировать соответствующую теорему, необходимо дать определение некоторого усиления этого понятия дифференцируемости.

Определение 4'. Пусть X и Y — нормированные пространства, V — окрестность точки $\hat{x} \in X$, $F : V \rightarrow Y$.

Говорят, что отображение F *строго дифференцируемо* в \hat{x} и пишут $F \in SD(\hat{x}, Y)$, если $F \in D^1(\hat{x})$ и для любого числа $\varepsilon > 0$ найдется число $\delta > 0$ такое, что если $\|x_i - \hat{x}\|_X < \delta$, $i = 1, 2$, то $\|F(x_1) - F(x_2) - F'(\hat{x})(x_1 - x_2)\|_Y \leq \varepsilon \|x_1 - x_2\|_X$.

Понятие строгой дифференцируемости позволяет сформулировать теорему об обратном отображении для гладких отображений.

Теорема 3'' об обратном отображении для строго дифференцируемых отображений: *строго дифференцируемое и регулярное в точке отображение из одного банахова пространства в другое локально обратимо. Или: Пусть X и Y — банаховы пространства, V — окрестность точки $\hat{x} \in X$, F — отображение из V в Y , строго дифференцируемое и регулярное в точке \hat{x} (т. е. отображение $F'(\hat{x}) : X \rightarrow Y$ сюръективно). Тогда найдутся окрестность W точки $F(\hat{x})$, отображение $\varphi : W \rightarrow V$ и константа $K > 0$, такие, что $F(\varphi(y)) = y$ и $\|\varphi(y) - \hat{x}\|_X \leq K \|y - F(\hat{x})\|_Y$ для любого y из W .*

Этот результат вытекает из теоремы 2.2 очевидным образом, если вспомнить определение строгой дифференцируемости.

Сформулируем теперь основной результат:

Теорема 4 (правило множителей Лагранжа). *Пусть в задаче (P) функция f_0 и отображение F определены в некоторой окрестности точки \hat{x} и строго дифференцируемы в этой точке. Тогда, если в точке \hat{x} достигается локальный экстремум в задаче (P), то существует набор множителей Лагранжа $(\lambda_0, \lambda) \in \mathbb{R}_+ \times X^*$, не равных одновременно нулю, такой, что*

$$\mathcal{L}_x(\hat{x}, \bar{\lambda}) = \lambda_0 f'_0(\hat{x}) + \langle \lambda, F'(\hat{x}) \rangle = 0. \quad (2.8)$$

Докажем эту теорему здесь в конечномерном случае, когда $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, $F = (f_1, \dots, f_m)^T$, f_i определены в окрестности точки $\hat{x} \in \mathbb{R}^n$ и строго дифференцируемы в этой точке. Тогда соотношение (2.8) приобретает вид:

$$\lambda_0 f'_0(\hat{x}) + \lambda_1 f'_1(\hat{x}) + \dots + \lambda_m f'_m(\hat{x}) = 0.$$

Доказательство. Теорема утверждает, что если \hat{x} — локальный экстремум, то векторы $f'_i(\hat{x})$, $0 \leq i \leq m$, линейно зависимы. Докажем теорему от противного: предположим, что эти векторы линейно независимы и придем к противоречию с тем, что \hat{x} — локальный экстремум. Рассмотрим отображение $G(x) = (f_0(x), f_1(x), \dots, f_m(x))^T$. По условию функции f_i , $0 \leq i \leq m$ строго дифференцируемы в \hat{x} . Отсюда легко следует, что и отображение G строго дифференцируемо в \hat{x} . Векторы $f'_i(\hat{x})$, $0 \leq i \leq m$ образуют строки матрицы $G'(\hat{x})$. Следовательно (см. определение ранга) ранг этой матрицы равен $m + 1$, т. е. совпадает с числом строк. Значит, по теореме Кронекера – Капелли $G'(\hat{x})$ — сюръективный оператор, и мы

можем применить теорему 2.2'. Положим $y_\nu = (f_0(\hat{x}) + \nu, 0, \dots, 0)^T$ и пусть $x_\nu = \varphi(y_\nu)$, где φ — правое обратное отображение к G . Тогда $G(x_\nu) = (f_0(\hat{x}) + \nu, 0, \dots, 0)$ и $|x_\nu - \hat{x}| \leq K|\nu|$. Это означает, что \hat{x} не доставляет задаче (P) ни локального максимума, ни локального минимума. Пришли к противоречию. Теорема доказана. \square

Исторический комментарий. Теорию экстремума принято отсчитывать от письма Ферма Декарту (1638), где он объяснял свой прием решения гладких задач на экстремум, обобщения которого называют теоремой Ферма. Условия стационарности в терминах производных появились в первых работах по анализу Ньютона и Лейбница, в конечномерном случае у Эйлера, в бесконечномерном случае у Фреше. Принцип Лагранжа для конечномерной гладкой задачи с равенствами был сформулирован Лагранжем в 1797 г., но строгое доказательство этой теоремы стало делом математиков последней четверти девятнадцатого века. Теорема 3.1 была доказана Люстерником в 1934г..

При решении задач на экстремум возникает вопрос о том, существует ли решение задачи. Один из основных принципов существования является принцип компактности, к изложению которого мы переходим.

Я УБЕЖДЕН, ЧТО БУДЕТ ВОЗМОЖНО ДОКАЗЫВАТЬ ТЕОРЕМЫ СУЩЕСТВОВАНИЯ С ПОМОЩЬЮ ОБЩЕГО ПРИНЦИПА, ЧЬЯ СУЩНОСТЬ НАВЕЯНА ПРИНЦИПОМ ДИРИХЛЕ. ЭТОТ ОБЩИЙ ПРИНЦИП, ВОЗМОЖНО ПРИБЛИЗИТ НАС К ОТВЕТУ НА СЛЕДУЮЩИЙ ВОПРОС: ИМЕЕТ ЛИ РЕШЕНИЕ КАЖДАЯ РЕГУЛЯРНАЯ ВАРИАЦИОННАЯ ПРОБЛЕМА, ЕСЛИ САМОМУ ПОНЯТИЮ "РЕШЕНИЕ" ПРИ СЛУЧАЕ ПРИДАВАТЬ РАСШИРЕННОЕ ТОЛКОВАНИЕ.

Д. Гильберт.

3.2. Принцип компактности. Сначала напомним некоторые понятия общей топологии. Общая топология — ветвь математики, изучающая понятия предела и непрерывности. Опишем эти понятия в рамках метрических пространств. Пара (X, d) , где X — множество, а $d = d(x, y)$ — вещественная неотрицательная функция, удовлетворяющая аксиомам: а) $d(x_1, x_2) = 0 \Leftrightarrow x_1 = x_2$, б) $d(x_1, x_2) = d(x_2, x_1) \forall x_1, x_2 \in X$, в) $d(x_1, x_3) \leq d(x_1, x_2) + d(x_2, x_3) \forall x_1, x_2, x_3 \in X$, называется *метрическим пространством*. Последовательность $\{x_n\}_{n \in \mathbb{N}}$ называется *сходящейся*, если существует такой элемент $\xi \in X$, что $\forall \varepsilon > 0 \exists N > 0 : d(x_n, \xi) < \varepsilon \forall n \geq N$; последовательность $\{x_n\}_{n \in \mathbb{N}}$ называется *фундаментальной*, если $\forall \varepsilon > 0 \exists N > 0 : d(x_n, x_m) < \varepsilon \forall n, m \geq N$; вещественная функция f на (X, d) называется непрерывной в точке $x_0 \in X$, если $\forall \varepsilon > 0 \exists \delta > 0 : d(x, x_0) < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$; функцию f называют непрерывной на (X, d) , если она непрерывна в любой точке X .

Определение 5. Метрическое пространство (X, d) называют *компактом*, если из каждой последовательности элементов из X можно выбрать сходящуюся подпоследовательность.

Лемма (о конечномерных компактах). Для того, чтобы подмножество \mathbb{R}^n было компактом, необходимо и достаточно, чтобы оно было ограниченным и замкнутым.

Доказательство. Необходимость. Если множество не является ограниченным, то строится последовательность точек, расстояние между которыми больше некоторого числа. Из нее нельзя выбрать сходящейся подпоследовательности. Если же последовательность ограниче-

на, но не замкнута, то существует последовательность, сходящуюся к точке, принадлежащей замыканию множества, но не самому множеству. Из нее также нельзя выбрать сходящуюся последовательность.

Достаточность. Докажем в случае \mathbb{R}^2 , в остальных случаях доказательство аналогично. Множество ограничено, и значит, его можно поместить в квадрат. Пусть дана какая-то последовательность в нашем ограниченном и замкнутом множестве. Выберем точку x_1 в этом множестве. Разделим квадрат на четыре равных квадрата. В одном из них окажется бесконечное множество точек. Выберем из него точку $x_2 \neq x_1$ и далее будем поступать аналогично. В итоге приходим к фундаментальной последовательности, из которой выберем сходящуюся. В силу замкнутости она будет сходить к точке множества. \square

Лемма (Гейне) Для того, чтобы функция f , определенная на метрическом пространстве (X, d) , была непрерывной в точке x_0 необходимо и достаточно, чтобы она была непрерывной по Гейне, т. е. для любой последовательности $\{x_n\}_{n \in \mathbb{N}}$, сходящейся к x_0 , последовательность $\{f(x_n)_{n \in \mathbb{N}}\}$ сходилась бы к $f(x_0)$.

Доказательство. Из определения непрерывности следует непрерывность по Гейне. Пусть непрерывность по Гейне имеется, но непрерывности нет. Отсутствие непрерывности означает, что имеется такое число $\varepsilon > 0$, что $\forall n \in \mathbb{N} \exists x_n : d(x_n, x_0) < \frac{1}{n}$, но $|f(x_n) - f(x_0)| > \varepsilon$. Это противоречит непрерывности по Гейне. \square

Вещественная функция f на (X, d) называется равномерно непрерывной на X , если $\forall \varepsilon > 0 \exists \delta > 0 : d(x_1, x_2) < \delta \Rightarrow |f(x_1) - f(x_2)| < \varepsilon$.

Теорема 5. а) Принцип компактности Вейерштрасса. Функция, непрерывная на компакте, достигает на нем своего минимального и максимального значения.

б) Теорема Кантора о равномерной непрерывности. Функция, непрерывная на компакте, равномерно непрерывна на нем.

Доказательство. а) Пусть (X, d) — метрический компакт и $f \in C(X)$. 1) Если допустить, что f неограничена, то $\forall n \in \mathbb{N} \exists x_n : |f(x_n)| > n$. По определению компактности $\exists \{n_k\}_{k \in \mathbb{N}} : x_{n_k} \rightarrow x_0$, что противоречит непрерывности по Гейне. 2) Пусть f ограничена сверху и M — верхняя грань значений f . По определению верхней грани $\forall n \in \mathbb{N} \exists x_n : f(x_n) > M - \frac{1}{n}$. По определению компактности $\exists \{n_k\}_{k \in \mathbb{N}} : x_{n_k} \rightarrow x_0$. По лемме Гейне $f(x_0) = M$. \square

б) Если допустить обратное, то найдется такое число $\varepsilon > 0$, что $\forall n \in \mathbb{N} \exists x_n, x'_n : d(x_n, x'_n) < \frac{1}{n}$, но $|f(x_n) - f(x'_n)| > \varepsilon$. По определению компактности $\exists \{n_k\}_{k \in \mathbb{N}} : x_{n_k} \rightarrow x_0$, при этом $f(x'_n) \neq f(x_0)$, что противоречит лемме Гейне. \square

Введем метрику в совокупность непрерывных функций на компакте $X : d_C(f_1, f_2) = \max_{x \in X} |f_1(x) - f_2(x)|$. Пространство непрерывных функций на компакте X с метрикой d_C обозначают $C(X)$.

Теорема о полноте $C(X)$. Пространство $(C(X), d_C)$ полно.

Доказательство. Пусть $\{f_n\}_{n \in \mathbb{N}}$ — фундаментальная (в смысле метрики d_C) последовательность, т. е. $\forall \varepsilon > 0 \exists N \in \mathbb{N} : d_C(f_n, f_m) < \varepsilon$ (i) $\forall n, m \geq N$. Отсюда следует, что для каждого $x \in X \lim_{n \rightarrow \infty} f_n(x) = f(x)$. Переходя к пределу в (i), получим, что при $n \geq N$ $|f(x) - f_n(x)| \leq \varepsilon$ (ii). Пусть $x_0 \in X$. Для некоторого $n_0 \geq N$ найдем $\delta > 0 : d_C(x, x_0) < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon/3$ (iii). Из (ii) и (iii) следует, что $|f_{n_0}(x) - f_{n_0}(x_0)| < \varepsilon$, т. е. f непрерывна. \square

Исторический комментарий. Основные понятия теории множеств и общей топологии были введены Г. Кантором. Принцип компактности претерпел значительную эволюцию, в развитии которой надо назвать имена Вейерштрасса, Бореля Лебега и Бэра, которому принадлежит весьма существенное развитие теоремы Вейерштрасса. Теорема Бэра гласит: *полу-непрерывная функция на компакте достигает своего минимума.*

Лекция 4. Теория линейных уравнений и геометрия.

В этой лекции проводится “визуализация” теории линейных уравнений, построенной в лекции 1. Здесь снова, как и в лекции 1 индексы векторов будем помещать наверху.

4.1. Геометрический смысл определителя матрицы и правило Крамера

Понятие определителя, которое мы собираемся сейчас вводить, играет выдающуюся роль и в алгебре, и в геометрии, и в анализе. Через определители выражаются решения систем уравнений (в которых число уравнений и неизвестных совпадают). Эти выражения называются правилом Крамера. Но на практике системы линейных уравнений в больших размерностях с помощью определителей не решают (такие решения требуют слишком большого числа операций).

Геометрический смысл определителя, порожденного n векторами в n -мерном пространстве — это *ориентированный объем параллелепипеда, порожденного этими векторами*. Отметим свойства такой характеристики, которые геометрически доказываются на примерах $n = 2, 3$. Введем функцию $V : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$, $V = V(a^1, \dots, a^n)$ системы векторов $\{a^j\}_{j=1}^n$ и потребуем от этой функции выполнения таких условий: а) *условие нормировки*, согласно которому объем единичного куба должен равняться единице: $V(e^1, \dots, e^n) = 1$, где $e^1 = (1, 0, \dots, 0)^T, \dots, e^n = (0, \dots, 1)^T$ — набор единичных векторов в \mathbb{R}^n ,

б) *условие “антисимметричности”*:

$$V(a^1, \dots, a^{i-1}, a^i, a^{i+1}, a^{i+2}, \dots, a^n) = -V(a^1, \dots, a^{i-1}, a^{i+1}, a^i, a^{i+2}, \dots, a^n) \text{ и}$$

с) *условие линейности по каждому векторному аргументу* (для этого достаточно потребовать, чтобы $V(\alpha a^1 + \alpha' a^{1'}, a^2, \dots, a^n) = \alpha V(a^1, a^2, \dots, a^n) + \alpha' V(a^{1'}, a^2, \dots, a^n)$).

Условия а) и б) представляются естественными, и они справедливы для определителей матриц второго порядка. Условие линейности для двумерного случая демонстрирует рис. 1

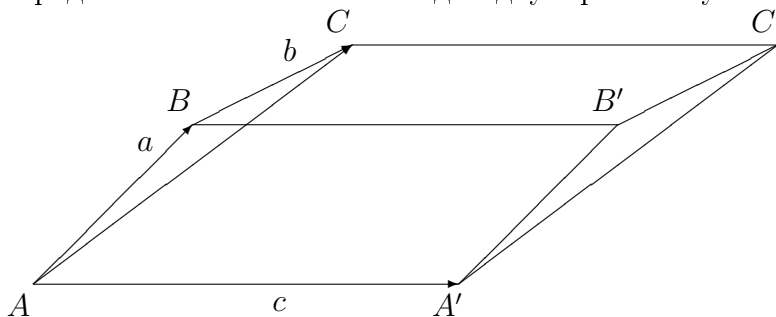


Рис. 1:

Из этого рисунка видно, что $V(a + b, c) = V(a, c) + V(b, c)$. Действительно, $V(a + b, c) =$

$S_{ACC'A'}$, а $V(a.c) + V(b.c) = S_{ACC'A'} - S_{A'B'C'} + S_{ABC}$. Но поскольку $S_{A'B'C'} = S_{ABC}$, мы и получаем искомое равенство.

Если $a^i = a^{i+1}$, то из условия б) следует, что $V(a^1, \dots, a^{i-1}, a^i, a^{i+1}, a^{i+1}, \dots, a^n) = 0$. Нетрудно показать, что при совпадении любых двух векторов (необязательно соседних) это равенство остается в силе.

Выведем из условий а) – с) выражение для определителя в двумерном случае. Здесь $e^1 = (1, 0)^T$, $e^2 = (0, 1)^T$. Имеем: $V = V(a^1, a^2) = V(a_{11}e^1 + a_{21}e^2, a_{12}e^1 + a_{22}e^2) = (a_{11}a_{22} - a_{12}a_{21})V(e^1, e^2) = a_{11}a_{22} - a_{12}a_{21}$.

Аналогично из приведенных аксиом однозначно выводится выражение для объема параллелепипеда, порожденного векторами a^1, a^2 и a^3 в трехмерном случае: $V(a^1, a^2, a^3) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{32}a_{23}$.

Продолжая этот процесс, приходим к такому результату: *существует единственная функция $V = V(a^1, \dots, a^n)$, которая определяется аксиомами а) - с)*. Для этой функции имеет место равенство

$$\sum_P (-1)^{|P|} \prod_{i=1}^n a_{iP(i)},$$

где сумма берется по всем перестановкам P первых n чисел $\{1, 2, \dots, n\}$, где число $|P|$ равно числу транспозиций соседних элементов, требующихся для перехода от $(P(1), P(2), \dots, P(n))$ к $\{1, 2, \dots, n\}$. Например, член $a_{13}a_{21}a_{32}$ имеет знак $+$, ибо переход от (312) к (123) требует двух транспозиций: $(312) \rightarrow (132) \rightarrow (323)$.

Определителем (детерминантом) матрицы A , составленной из столбцов a^1, a^2, \dots, a^n , называется число $\det A = V(a^1, a^2, \dots, a^n)$. Сформулируем сказанное еще раз.

Определение 6. Определителем системы векторов $\{a^1, \dots, a^n\}$ из \mathbb{R}^n называется значение на этой системе векторов единственным образом определяемой функции $\det : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$, удовлетворяющей условиям а) - с) нормировки, антисимметрии и линейности по каждому аргументу. Если матрица A , составлена из столбцов a^1, a^2, \dots, a^n , то $\det(a^1, a^2, \dots, a^n)$ называется *определителем матрицы A* .

Теорема 6 (правило Крамера). *Если определитель системы n уравнений с n неизвестными отличен от нуля, тогда система $(1)_n$ лекции 1 однозначно разрешима для любой правой части, и решение определяется формулами Крамера:*

$$x_i = \frac{\det A_i}{\det A}, \quad \det A_i = V(a^1, \dots, a^{i-1}, y, a^{i+1}, \dots, a^n), \quad 1 \leq i \leq n. \quad (i)$$

Доказательство. Пусть $\det A \neq 0$. Тогда по его столбцы линейно независимы (ибо в противном случае из линейности определителя по столбцам следовало бы, что определитель равен нулю). По теореме 1.1 система $(1)_n$ разрешима. В силу разрешимости системы $Ax = b$, найдутся x_i , $1 \leq i \leq n$, такие, что $x_1 a^1 + \dots + x_n a^n = b$. Имеем

$$\begin{aligned} \det A_i &= V(a^1, \dots, a^{i-1}, b, a^{i+1}, \dots, a^n) = \\ &= V(a^1, \dots, a^{i-1}, x_1 a^1 + \dots + x_n a^n, a^{i+1}, \dots, a^n) = x_i \det A \end{aligned}$$

Отсюда вытекают доказываемые равенства. □

4.2. Евклидово пространство \mathbb{E}^n . Подпространства в \mathbb{E}^n .

Евклидовым пространством \mathbb{E}^n называется пространство векторов из \mathbb{R}^n , снабженное скалярным произведением

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

Число $|x| = \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + \dots + x_n^2}$ называется *модулем вектора* $x \in \mathbb{E}^n$.

Замечание. Скалярное произведение позволяет идентифицировать пространство \mathbb{R}^n и пространство линейных функционалов на нем, обозначенное нами в лекции 1 через \mathbb{R}^{n*} . Это дает возможность изображать производную функции, определенной в \mathbb{E}^n в том же пространстве \mathbb{E}^n . Вектор $f'(\hat{x}) = (\frac{\partial f(\hat{x})}{\partial x_1}, \dots, \frac{\partial f(\hat{x})}{\partial x_n})$ называется *градиентом* f в \hat{x} .

Векторы $x, y \in \mathbb{E}^n$, для которых $\langle x, y \rangle = 0$, называются *ортогональными*. Поясним это определение в плоском случае (рис. 2).

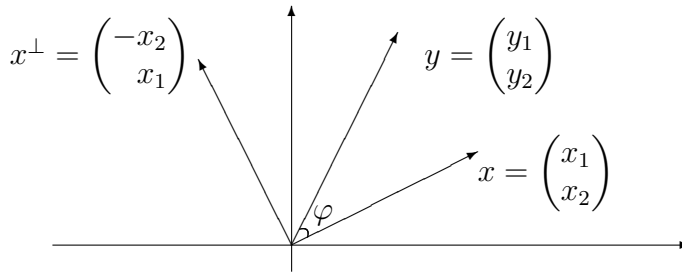


Рис. 2:

Найдем выражение для скалярного произведения векторов $x = (x_1, x_2)$ и $y = (y_1, y_2)$ через координаты. Для этого надо найти выражение для косинуса угла между векторами x и y или, что то же $\sin(y, x^\perp)$ — синуса угла между векторами y и x^\perp . Воспользуемся теперь тем, что площадь $V(y, x^\perp)$ параллелограмма, натянутого на векторы y и x^\perp равна с одной стороны $|x||y| \sin(y, x^\perp) = |x||y| \cos(x, y)$, а с другой — это определитель матрицы $\begin{pmatrix} -x_2 & x_1 \\ y_1 & y_2 \end{pmatrix} = x_1 y_1 + x_2 y_2 = \langle x, y \rangle$, откуда $\langle x, y \rangle = |x||y| \cos(x, y)$. Тем самым, ортогональность векторов равносильна тому, что косинус угла между ними равен нулю.

Множество $L \subset \mathbb{E}^n$ называется *подпространством*, если при всех $x, y \in L$ $\alpha x + \beta y \in L$ для любых $\alpha, \beta \in \mathbb{R}$.

Множество векторов $\{e^1, \dots, e^n\}$ называется *ортогональной системой*, если $\langle e^i, e^j \rangle = 0$ при всех $i \neq j$. Ортогональная система называется *ортонормированной*, если $|e^i| = 1$ при всех $1 \leq i \leq k$. Ортогональная система называется *полной*, если из того, что $\langle x, e^i \rangle = 0$ для всех $1 \leq i \leq k$ следует, что $x = 0$.

Построим в подпространстве $L \subset \mathbb{E}^n$ полную ортонормированную систему. Если L не состоит из одного нуля, в нем имеется ненулевой вектор f^1 . Положим $e^1 = \frac{f^1}{|f^1|}$. Если нет векторов из L , не пропорциональных e^1 , то построение закончено. Пусть существует вектор f^2 , не пропорциональный вектору f^1 . Тогда вектор $g^2 = f^2 - \langle f^2, e^1 \rangle e^1$ будет ортогонален e^1 . Обозначим вектор $\frac{g^2}{|g^2|}$ через e^2 . И далее будем поступать аналогично. Через некоторое число шагов, меньшее или равное n , все закончится построением системы $\{e^i\}_{i=1}^m$. Можно доказать, что m — инвариант подпространства: с какого вектора мы бы ни начинали и как

бы ни продолжали, процесс закончится за m шагов. Число m называется *размерностью* подпространства L .

Для любого вектора ξ , не принадлежащего L , вектор $\xi - \sum_{i=1}^m \langle \xi, e^i \rangle e^i$ ортогонален L , а расстояние от ξ до L равно модулю этого вектора, т. е. числу $\sqrt{|\xi|^2 - \sum_{i=1}^m \langle \xi, e^i \rangle^2}$ (проверьте!). Отметим этот важный факт: *для любого подпространства, размерности меньшей n , существует прямая перпендикулярная этому подпространству.* Это дает возможность сформулировать отличное от теоремы Кронекера – Капелли условие разрешимости уравнения $Ax = y$:

Предложение 4.1. *Для того, чтобы уравнение $Ax = y$ возможно было бы решить, необходимо и достаточно, чтобы вектор y был ортогонален любому решению однородного уравнения с транспонированной матрицей A^T .*

Действительно, если ξ — решение уравнения $Ax = y$, а η — решение уравнения $A^T x = 0$, то $\langle y, \eta \rangle \stackrel{\text{def}}{=} \langle A\xi, \eta \rangle \stackrel{\text{Id}}{=} \langle \xi, A^T \eta \rangle = 0$.

Допустим теперь, что вектор y ортогонален любому решению уравнения $A^T x = 0$, но решения $Ax = y$ не существует. Образ \mathbb{E}^n при отображении A — подпространство и $y \notin A\mathbb{E}^n$. Пусть y' основание перпендикуляра из y на $A\mathbb{E}^n$. Тогда $\langle y - y', Ax \rangle = 0$ для всех x и значит, $\langle A^T(y - y'), A^T(y - y') \rangle = 0$, т. е. $A^T(y - y') = 0$. Тогда по условию $\langle y, y - y' \rangle = 0$, откуда (в силу того, что $y - y'$ ортогонально y') получаем, что $\langle y - y', y - y' \rangle = 0$, т. е. $y \in L$ — противоречие. \square

Исторический комментарий. *Аксиоматическое определение детерминанта было предложено во второй половине 19 века К. Вейерштрассом (1815 – 1897) и Л. Кронекером; это определение следовало из концепции Грассмана, изложенной в его труде «Учение о линейном протяжении» (1847), который долго оставался невостребованным.*

Лекция 5. Приведение квадратичных форм к главным осям. Доказательство теоремы Фредгольма.

Если ничто не сойдёт нас с пути, то мы уподобимся Зигфриду, перед которым огненный вал раступается сам собою

Д. Гильберт.

Здесь доказываются теоремы Фредгольма и Гильберта — центральные в бесконечномерной теории линейных уравнений и квадратичных форм.

5.1. Теория конечномерных квадратик

Квадрика — это множество уровня квадратичной функции, а квадратичная функция в \mathbb{E}^n — это функция, равная сумме квадратичной формы, линейной функции и константы, т. е. $f(x) = \langle Ax, x \rangle + 2\langle b, x \rangle + c = \sum_{1 \leq i, j \leq n} a_{ij} x_i x_j + 2 \sum_{i=1}^n b_i x_i + c$. Здесь A — симметричная матрица, т. е. $A^T = A$ или $a_{ij} = a_{ji}$ для любых i и j . Квадратичную форму назовем *невырожденной*, если матрица A является разрешимой.

У невырожденных квадратичных форм можно избавиться от линейных членов. Для этого

сдвинем систему координат параллельно на вектор a , перейдя к координатам $y = x + a$. Подставив в f вместо x вектор $y - a$ и обозначив $\varphi(y) = f(y - a)$, получим $\varphi(y) = \langle A(y - a), (y - a) \rangle + 2\langle b, (y - a) \rangle + c$. После несложных преобразований будем иметь: $\varphi(y) = \langle Ay, y \rangle + 2\langle b - Aa, y \rangle + c'$, где $c' = c + \langle Aa - 2b, a \rangle$. Если A разрешимая матрица, то можно решить уравнение $Aa = b$, что приведет квадратичную форму к более простому виду $\langle Ay, y \rangle + c'$.

Приведем квадратичную форму $Q(x) = \langle Ax, x \rangle$ к *главным осям* или к *каноническому виду*.

Теорема 7 (теорема о приведении квадратичной формы к главным осям.) Для любой симметричной матрицы A найдутся $m \leq n$ единичных взаимно ортогональных векторов $\{f^i\}_{1 \leq i \leq m}$ и m отличных от нуля чисел λ_i , при которых квадратичная форма Q приобретает вид: $Q(x) = \sum_{i=1}^m \lambda_i \langle f^i, x \rangle^2$. Прямые, порожденные векторами f^i , называются *главными осями*.

Доказательство. Метод доказательства теоремы 5.1 и ее обобщения на бесконечномерный случай связан с постановкой и решением некоторых задач на максимум и минимум с помощью правила множителей Лагранжа.

Рассмотрим задачу на максимум при наличии ограничения типа равенства:

$$f_0(x) = \sum_{1 \leq i, j \leq n} a_{ij} x_i x_j \rightarrow \max, \quad f_1(x) = \sum_{i=1}^n x_i^2 = 1, \text{ или, в сокращенной форме:}$$

$$\langle Ax, x \rangle \rightarrow \max, \quad \langle x, x \rangle = 1. \quad (P_1)$$

Единичная сфера $S^{n-1} = \{x \mid \sum_{i=1}^n x_i^2 = 1\}$ — компакт в \mathbb{E}^n , функция f_0 непрерывна всюду в \mathbb{E}^n , значит, по теореме Вейерштрасса о непрерывной функции на компакте решение поставленной задачи существует. Обозначим его e^1 . Из правила множителей Лагранжа следует, что найдется число λ_1 для которого $Ae^1 = \lambda_1 e^1$ (вектор e^1 называется *собственным вектором матрицы A* , а число λ^1 называется *собственным значением этой матрицы*). Мы доказали существование первого собственного вектора матрицы A . Допустим, мы доказали существование k ортогональных векторов $\{e^j\}_{1 \leq j \leq k}$, $k < n$. Докажем существование $k + 1$ -го вектора. Для этого рассмотрим задачу

$$\langle Ax, x \rangle \rightarrow \max, \quad \langle x, x \rangle = 1, \quad \langle x, e^j \rangle = 0, \quad 1 \leq j \leq k. \quad (P_k)$$

Снова из соображений компактности решение существует. Обозначим его e^{k+1} . Из правила множителей Лагранжа следует, что найдутся константы $\lambda_{k+1}, \mu_1, \dots, \mu_k$ такие, что $Ae^{k+1} = \lambda_{k+1}e^{k+1} + \mu_1 e^1 + \dots + \mu_k e^k$. Умножая это равенство последовательно на e^1, \dots, e^k , получаем, что $\mu_1 = \dots = \mu_k = 0$. Мы построили $k + 1$ -й собственный вектор. Так будут построена система из n ортонормированных векторов $\{e^k\}_{1 \leq k \leq n}$. В силу ортонормированности эта система линейно независима (почему?). А тогда в силу альтернативы для решения системы, присоединение любого вектора x его можно выразить через $\{e^k\}_{1 \leq k \leq n}$: $x = \sum_{k=1}^n x_k e^k$. Умножая это равенство на e^k и пользуясь ортонормированностью, приходим к равенству: $x = \sum_{k=1}^n \langle x, e^k \rangle e^k$, а значит (в силу линейного свойства A) будем иметь $Ax = \sum_{k=1}^n \lambda_k \langle x, e^k \rangle e^k$ и, наконец, в силу ортонормированности e^k , получаем окончательно: $f_0(x) = \sum_{1 \leq k \leq n} \lambda_k \langle x, e^k \rangle^2$. \square

В связи с доказанной теоремой разумно ввести два важных понятия.

Определение 7. Пусть A — матрица размеров $n \times n$. Вектор $x \in \mathbb{R}^n$ называется *собственным вектором* матрицы A , если существует число λ такое, что $Ax = \lambda x$. Если строки матрицы образуют ортонормированную систему в \mathbb{R}^n , то такая матрица называется *ортogonalной*.

Доказанную теорему можно переформулировать так: *квадратичная форма ортогональной заменой переменных приводится к каноническому виду или симметрический оператор ортогональным преобразованием приводится к главным осям.*

Исторический комментарий. Применение методов теории экстремума для сведения квадратичной формы к сумме квадратов было осуществлено Лагранжем в 1759 г.

5.2. Бесконечномерное евклидово пространство — пространство l_2 . Доказательство теоремы Фредгольма

Пространство l_2 образовано векторами $x = (x_1, x_2, \dots, x_n, \dots)$, обладающими свойством $\sum_{n \in \mathbb{N}} x_i^2 < \infty$. Выражение $(\sum_{n \in \mathbb{N}} x_i^2)^{1/2}$ называют *модулем* x и обозначают $|x|$. Пространство l_2 можно снабдить *скалярным произведением* $\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n + \dots$, которое оказывается конечным для любых двух векторов из l_2 .

Это вытекает из следующего неравенства (Коши – Буняковского) для двух векторов x и y из l_2 : $|\langle x, y \rangle| \leq |x||y|$. (В конечномерном случае оно вытекает из соотношений $0 \leq \langle x - ty, x - ty \rangle \stackrel{\text{Id}}{=} |x|^2 - 2t\langle x, y \rangle + |y|^2$, а в бесконечномерном случае надо перейти к пределу).

Из неравенства Коши – Буняковского следует, что линейная комбинация векторов из l_2 не выводит из этого пространства. Определение модуля дает возможность превратить l_2 в метрическое пространство (определение см. в Дополнении), положив расстояние $d(x, y)$ между векторами x и y равным $|x - y|$. Доказывается, что пространство l_2 является полным.

Замечание об определителях в l_2 . Некоторое время в 19 веке делалась попытка построить теорию бесконечных систем линейных уравнений с бесконечным числом неизвестных с помощью определителей, но она оказалась не вполне удачной и была заменена геометрической теорией к изложению которой мы переходим.

Теорема Фредгольма в бесконечномерном случае была нами сформулирована в главе 1 в пространстве l_2^T . Переформулируем ее в пространстве l_2 векторов x , в которых введено скалярное произведение. Переформулировка этой теоремы такова:

Теорема (Фредгольма). Пусть матрица A представляет собой сумму $I + B$ единичной матрицы и матрицы $B = (b_{ij})_{1 \leq i, j < \infty}$, удовлетворяющей условию $\|B\| = (\sum_{i, j \in \mathbb{N}} b_{ij}^2)^{1/2} < \infty$. Тогда:

а) Имеет место альтернатива: или уравнение $Ax = y$ разрешимо для любого $y \in l_2$, или однородное уравнение $Ax = 0$ имеет ненулевое решение.

б) Для того, чтобы уравнение $Ax = y$ было разрешимо необходимо и достаточно, чтобы скалярное произведение $\langle y, x \rangle$ вектора y на любое решение однородного уравнения $A^T x = 0$, было равно нулю (иначе говоря, вектор y должен быть ортогонален подпространству $\text{Ker} A^T$).

Доказательство. А) Допустим, что нашу систему $Ax = y$ можно решить для любой правой части, а однородное уравнение имеет ненулевое решение, и придем к противоречию. Доказа-

тельство в этой части проходит в значительной мере по схеме конечномерных теорем главы 1 и основывается на таком утверждении.

Лемма 5.1. *Не существует бесконечной последовательности единичных векторов $\{f^i\}_{i \in \mathbb{N}}$, такой, что расстояния между Bf^i и Bf^j при $i \neq j$ превосходят некоторое положительное число.*

Доказательство леммы. Пусть $|Bf^i - Bf^j| \geq \alpha$, $i \neq j$. Выберем N настолько большим, чтобы $\|B - B_N\| \leq \frac{\alpha}{4}$ (i), где $B_N = (b_{ij})_{1 \leq i, j \leq N}$. Тогда будем иметь ((TI) — Triangle Inequality — неравенство треугольника): $|B_N f^i - B_N f^j| \stackrel{\text{Id}}{=} |Bf^i - Bf^j + B_N f^i - Bf^i - B_N f^j + Bf^j| \stackrel{\text{TI}}{\geq} |Bf^i - Bf^j| - |B_N f^i - Bf^i| - |B_N f^j - Bf^j|$ (ii). Но из неравенства Коши — Буняковского вытекает, что $|B_N f^i - Bf^i| \leq \|B_N - B\| |f^i| \stackrel{(i)}{\leq} \alpha/4$, аналогично $|B_N f^j - Bf^j| \leq \alpha/4$, откуда и из (ii) следует, что $|B_N f^i - B_N f^j| \geq \alpha/2$. Мы свели дело к конечномерному случаю и должны доказать, что в N -мерном шаре $\text{Ball}^N(o, R)$ с центром в нуле радиуса $R = \|B_N\|$ нельзя расположить слишком много точек x^i , находящихся друг от друга на расстоянии, большем $\alpha/2$. Пусть нам удалось расположить в шаре $\text{Ball}^N(o, R)$ k точек $\{x^1, \dots, x^k\}$. Это значит, что внутренности шаров $\text{Ball}^N(x^i, \alpha/4)$ друг с другом не пересекаются. Но все эти шары лежат внутри шара $\text{Ball}^N(o, R')$, где $R' = R + \alpha/4$. Откуда следует оценка $k \text{VolBall}^N(o, \alpha/4) \leq \text{VolBall}^N(o, R')$, где $\text{VolBall}^N(o, \rho)$ — объем шара с центром в нуле радиуса ρ . Мы пришли к оценке $k \leq (\frac{4R'}{\alpha})^N$. Лемма доказана.

Вернемся к доказательству теоремы. По предположению существует ненулевое решение e^1 уравнения $Ax = 0$, и при этом можно считать, что $|e^1| = 1$. Обозначим через $L_k = \text{Ker} A^k$. Доказано, что L_1 нетривиально. Решение уравнения $Ae^2 = e^1$ и далее $Ae^{k+1} = e^k$ показывает, что все L_k нетривиальны. При этом из леммы I вытекает, что все L_k — конечномерны (иначе мы построили бы в L_k ортонормированную систему из любого числа векторов g^k таких, что $|Bg^i - Bg^j| = |g^i - g^j| = \sqrt{2}$. Рассмотрим последовательность Be^n . Пусть $m > n$. Легко понять, что $z_n = e^n - Ae^n + Ae^m \in \text{Ker} A^{m-1}$. Значит, $\|Be^m - Be^n\| = \|z_n - e_m\| \geq 1/2$. Пришли к противоречию с компактностью оператора B . Следовательно, $\text{Ker} A = 0$.

В) Пусть $AX \neq X$. Для доказательства, утверждения, что тогда $\text{Ker} A \neq 0$, придется использовать два факта, которые будут доказаны в Дополнении): 1) AX — замкнутое подпространство в l_2 , и 2) к замкнутому подпространству l_2 можно провести перпендикуляр. Итак, AX — замкнутое подпространство. Пусть y — перпендикуляр к нему, т. е. $\langle y, Ax \rangle = \langle A^T y, x \rangle = 0 \forall x \in X$. Это означает, что $A^T y = 0$, т. е. $\text{Ker} A^T \neq 0$. По доказанному в А) отсюда следует, что $A^T X \neq X$. А по только что доказанному, отсюда следует, что $\text{Ker} A^{TT} \neq 0$. Но $A^{TT} = A$ (двукратная смена строк и столбцов все возвращает на место). Вот мы и доказали, что хотели: $AX \neq X \Rightarrow \text{Ker} A \neq 0$. □

5.3. Теория квадратик в l_2 . Теорема Гильберта о приведении квадратичной формы к главным осям.

Спектральная теорема Гильберта Пусть матрица $A = (a_{ij})_{1 \leq i, j < \infty}$ такова, что $\sum_{i, j} a_{ij}^2 < \infty$, тогда найдется последовательность единичных взаимно ортогональных векторов $\{f^i\}_{1 \leq i \leq n}$ и последовательность чисел λ_i , при которых квадратичная форма $f(x) =$

$\langle Ax, x \rangle$ приобретает вид: $f(x) = \sum_{i=1}^i \lambda_i \langle f^i, x \rangle^2$.

Доказательство проходит в точности по схеме доказательства конечномерной теоремы.

Рассмотрим задачу на максимум при наличии ограничения типа равенства: $f_0(x) = \sum_{1 \leq i, j < \infty} a_{ij} x_i x_j \rightarrow \max$, $f^1(x) = \sum_{i=1}^n x_i^2 \leq 1$, или, в сокращенной форме:

$$\langle Ax, x \rangle \rightarrow \max, \quad \langle x, x \rangle \leq 1. \quad (P_2)$$

Сразу отметим отличие: раньше у нас было равенство $\langle x, x \rangle = 1$, теперь оно сменилось неравенством $\langle x, x \rangle \leq 1$. Именно шар $B_{l_2}(0, 1) = \{x \in l_2 \mid \langle x, x \rangle \leq 1\}$ обладает свойством (слабой) компактности, позволяющем утверждать, что решение задачи (P_2) существует. И правило множителей Лагранжа в пространстве l_2 применимо. Эти факта оставим без доказательства. А далее доказательство в точности повторяет то, что было проделано в конечномерном случае.

Остались не доказанными три утверждения, относящихся, собственно, к общей топологии и одно к математическому анализу. Это утверждения о замкнутости и перпендикуляре в теореме Фредгольма и существовании максимума в теореме Гильберта.

Лекция 6. Начала дифференциального исчисления и теории дифференциальных уравнений

ПОВОРОТ В МАТЕМАТИКЕ [ВЫЗВАННЫЙ ЗАРОЖДЕНИЕМ МАТЕМАТИЧЕСКОГО АНАЛИЗА] ПРОИЗОШЕЛ В 17 ВЕКЕ ОДНОВРЕМЕННО С СОЗДАНИЕМ ОСНОВ МАТЕМАТИЧЕСКОГО ЕСТЕСТВОЗНАНИЯ. ЗНАЧЕНИЕ ЭТОГО ПОВОРОТА НАСТОЛЬКО ВЕЛИКО, ЧТО ОБРАЗОВАВШИЕСЯ В РЕЗУЛЬТАТЕ НЕГО РАЗДЕЛЫ ОБЪЕДИНЯЮТ ПОД НАЗВАНИЕМ *высшей* МАТЕМАТИКИ В ОТЛИЧИЕ ОТ СЛОЖИВШЕЙСЯ РАНЕЕ *элементарной* МАТЕМАТИКИ”.

А. Н. Колмогоров

В этой лекции мы вначале на короткое время возвращаемся к истокам, к началам дифференциального исчисления. Основная же часть этой лекций посвящена теоремам существования решений дифференциальных уравнений.

6.1. Об истоках математического анализа.

Основоположниками математического анализа были И. Ньютон и Г. Лейбниц (1646 – 1716). Ньютон создавал анализ, как аппарат естествознания, Лейбниц был воодушевлен широчайшей программой развития человеческой мысли, существенной компонентой которой, по его мнению, было создание *исчислений*. Он заложил начала дифференциального исчисления.

Приведем слова Ньютона, где он объясняет суть математического анализа: *Для пояснения искусства анализа надо привести некоторые примеры задач [...]. 1) Пусть длина пути известна. Нужно узнать скорость в данный момент времени. 2) Пусть известна скорость движения. Надо узнать длину пройденного пути.*”

Продумывание первой задачи Ньютона с неизбежностью ведет к понятию производной, как скорости движения. Скорость $v(\tau)$ при малом Δt примерно равна средней скорости

$\frac{s(\tau+\Delta t)-s(\tau)}{\Delta t}$ (где $s(t)$ – длина пути, пройденного объектом к моменту времени t), и величина $v(\tau)$ равна *пределу* таких отношений при Δt стремящемся к нулю, т. е. $v(\tau) = \lim_{\Delta t \rightarrow 0} \frac{s(\tau+\Delta t)-s(\tau)}{\Delta t}$.

Этот предел стали называть *производной пути по времени* и обозначать $\dot{s}(\tau)$, $s'(\tau)$ или $\frac{ds(\tau)}{d\tau}$. Первое обозначение принадлежит Ньютону, второе Лагранжу, третье Лейбницу.

Вернемся к эйлеровым обозначениям, когда независимое переменное обозначается x , а функция f . Приведем три известных определения производной $f'(\hat{x})$ функции f , определенной в некоторой окрестности точки \hat{x} :

- 1) $f'(\hat{x}) = \lim_{x \rightarrow \hat{x}} \frac{f(x)-f(\hat{x})}{x-\hat{x}}$.

- 2) Производная $f'(\hat{x})$ равна тангенсу угла наклона касательной к графику f , проведенной в точке \hat{x} .

- 3) Производная $f'(\hat{x})$ равна главной линейной части отображения $x \rightarrow f(x)$ в точке \hat{x} ; это такое число A , что $f(\hat{x} + x) = f(\hat{x}) + Ax + o(x)$.

Первое определение принадлежит Коши, второе Лейбницу, а третье, собственно говоря, Фреше (который дал определение в бесконечномерном случае, но полагал, что оно неизвестно в конечномерном: описывая производную функции многих переменных он написал “...differentiele à mon sens” (“дифференциал в моем смысле”).

И как тут не вспомнить известный рассказ.

... Однажды, войдя в аудиторию, лорд Кельвин спросил студентов: “Что такое $\frac{dx}{dt}$?” В ответ он получил все возможные логические определения. “Оставьте это, — воскликнул он, — $\frac{dx}{dt}$ — это скорость!”

Ньютон, кстати, очень хорошо осознавал смысл предельного перехода. Он писал: “Количества [...], которые в продолжение любого конечного времени постоянно стремятся к равенству и ранее конца этого времени приближаются друг к другу ближе, нежели на любую заданную величину, будут в пределе равны”. Полную ясность в этот вопрос внес О. Коши (1789 – 1857). Число a называется *пределом* последовательности $\{x^n\}_{n \in \mathbb{N}}$ при n стремящемся к бесконечности (или функции f , определенной на интервале $(x_0 - \alpha, x_0 + \alpha)$ при x , стремящемся к x_0), если для любого числа $\varepsilon > 0$ найдутся число N (или положительное число δ), такие что $|x^n - a| < \varepsilon$, $\forall n \geq N$ ($|f(x) - f(x_0)| < \varepsilon \forall x : |x - x_0| < \delta$).

Из определений легко следуют формулы *исчисления пределов*: предел линейной комбинации, произведения, частного двух последовательностей или функций равен линейной комбинации, произведению, частному пределов этих последовательностей или функций (в последнем случае, если предел знаменателя не равен нулю).

Совершенно аналогично осмысление второй задачи — об определении пути по скорости. Оно неизбежно ведет к понятию римановой суммы. Пусть нам известна скорость $v(t)$ в любой момент t на временном отрезке $[t_0, t_1]$, и мы хотели бы найти длину пути, преодоленного за этот период времени. Естественно тогда разбить отрезок $[t_0, t_1]$ на малые промежутки точками $t_0 = \tau_0 \leq \tau_1 \leq \dots \leq \tau_{N+1} = t_1$ и считать, что на участке $\Delta_k = [\tau_k, \tau_{k+1}]$ скорость *примерно постоянна* и равна $v(\theta_k)$, где θ_k какая-то точка на отрезке Δ_k . Таким образом, преодоленный путь, равный $s(t_1) - s(t_0)$ *примерно равен* $\sum_{k=0}^N v(\theta_k) |\Delta_k|$ (where $|\Delta|$ — это длина отрезка Δ . Эта сумма и называется *римановой суммой* и обозначается $R(v(\cdot), D, S)$, где D (от слова “division” — разбиение) — разбиение отрезка $[t_0, t_1]$ на малые отрезки Δ_k , а S (от слова “selection” — выбор) — выбор точек θ_k на Δ_k . Так мы приходим к понятию интеграла, как *предела римановых сумм* (точное определение будет дано впоследствии). Этот предел обозначают $\int_{t_0}^{t_1} v(t) dt$ и

называют *определенным интегралом* или *интегралом Римана* от функции $v(\cdot)$ на отрезке $[t_0, t_1]$.

Подведем предварительный итог. Мы пришли к формулам: $s'(t) = v(t)$; $\int_{t_0}^{t_1} v(t)dt = s(t_1) - s(t_0)$ и в первом приближении обосновали фундаментальный результат о связи производной и интеграла, которым является

Формула Ньютона – Лейбница или основная формула интегрального исчисления:

$$\int_{t_0}^{t_1} s'(t)dt = s(t_1) - s(t_0). \quad (1)$$

Или еще $s(t) = s(t_0) + \int_{t_0}^t v(s)ds$, т. е. интегрирование в некотором отношении – обратная операция по отношению к дифференцированию. Далее будут даны все необходимые уточнения. Ньютон обозначал независимое переменное буквой t (от time – время). Сейчас чаще независимое переменное обозначают буквой x , а функцию – f . В этих обозначениях формула Ньютона – Лейбница записывается так: $f(x_1) - f(x_0) = \int_{x_0}^{x_1} f'(s)ds$.

6.2. Теоремы существования решения задачи Коши в теории дифференциальных уравнений.

Поворот в математике, о котором говорит А. Н. Колмогоров в приведенном выше эпиграфе, на языке математического анализа может быть выражен следующим образом: **законы, выражающие эволюцию процессов, происходящих в природе, описываются дифференциальными уравнениями.** В частности, законы механики описываются обыкновенными дифференциальными уравнениями. Решения таких уравнений предопределяются начальными данными. Это создало у человечества иллюзию о том, что все в мире предопределено.

Наиболее остро выразил это Пьер Лаплас. Он сказал: “Ум, которому были бы известны для какого-либо данного момента все силы, одушевляющие природу и относительное положение всех ее составляющих частей, [...] он обнял бы в одной формуле движение величайших тел Вселенной наравне с движениями легчайших атомов, и ничто не осталось бы для него недостоверным. Будущее, равно, как и прошедшее, предстало бы перед его взором.”

Это воззрение оказалось иллюзией, но математическая основа справедлива и состоит в теореме существования и единственности решения задачи Коши обыкновенного дифференциального уравнения. Переходим к доказательству этой теоремы.

Пусть $D = [t_0 - a, t_0 + a] \times B_{\mathbb{R}^n}(x_0, b) \rightarrow \mathbb{R}^n$, $f : D \rightarrow \mathbb{R}^n$, $x_0 \in \mathbb{R}^n$. Рассмотрим задачу:

$$\dot{x} = f(t, x), \quad x(t_0) = x_0, \quad (2.1)$$

которую называют *задачей Коши для дифференциального уравнения* $\dot{x} = f(t, x)$.

Теорема 8 (локальная теорема существования решения задачи Коши). *Если f непрерывна в D и липшицева с константой L по x , то на $[t_0 - \alpha, t_0 + \alpha]$ ($\alpha = \min(a, 1/2L, b/2M)$), где $M = \max_{t \in [t_0 - a, t_0 + a]} |f(t, x_0)|$ существует и единственно решение (2.1).*

Доказательство. Здесь следует применить теорему о (правом) обратном отображении в бесконечномерном случае (из второй лекции) в ситуации, когда $X = Y = C([t_0 - \alpha, t_0 + \alpha], \mathbb{R}^n)$, $x_0(\cdot) \equiv x_0$, $F(x(\cdot))(t) = x(t) - x_0 - \int_{t_0}^t f(s, x(s))ds$, $\Lambda = \text{Id}$, $\theta = 1/2$, $\gamma = 1$, $\delta = b$, $\alpha = \frac{b}{2}$.

Проверим выполнимость условий этой теоремы. Имеем: $\|F(x_0(\cdot))(\cdot)\|_{C([t_0-\alpha, t_0+\alpha], \mathbb{R}^n)} \leq \delta M \leq \frac{b}{2}$ (т. е. функция тождественно равная нулю принадлежит $B_{C([t_0-\alpha, t_0+\alpha])}(F(x_0(\cdot))(\cdot), \alpha)$), и поскольку для любых $\xi(\cdot)$, $x(\cdot)$ выполняется неравенство

$$\|F(\xi(\cdot))(\cdot) - F(x(\cdot))(\cdot) - (\xi(\cdot) - x(\cdot))\|_{C([t_0-\alpha, t_0+\alpha], \mathbb{R}^n)} \leq \max_{t \in [t_0-\delta, t_0+\delta]} \left| \int_{t_0}^t (f(s, \xi(s)) - f(s, x(s)))ds \right| \leq |t - t_0|L\|\xi(\cdot) - x(\cdot)\|_{C([t_0-\alpha, t_0+\alpha], \mathbb{R}^n)} \leq \frac{1}{2}\|\xi(\cdot) - x(\cdot)\|_{C([t_0-\alpha, t_0+\alpha], \mathbb{R}^n)},$$

условия теоремы о правом обратном отображении выполнены. Применение этой теоремы немедленно приводит к локальной теореме существования. \square

Теорема 8' (глобальная теорема существования решения задачи Коши для линейной системы). Пусть $\Delta = [t_0, t_1]$, функции $A : \Delta \mapsto \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ и $b : \Delta \mapsto \mathbb{R}^n$ непрерывны на отрезке Δ , $\tau \in \Delta$ и $\xi \in \mathbb{R}^n$. Тогда на Δ существует единственное решение задачи Коши

$$\dot{x} = A(t)x + b(t), \quad x(\tau) = \xi. \quad (2.2)$$

Доказательство. Здесь следует применить теорему о правом обратном к случаю, когда $X = Y = C(\Delta, \mathbb{R}^n)$, тождественному оператору Λ и $F(x(\cdot))(\cdot) = x(\cdot) - G^m(x(\cdot))(\cdot)$, где G

$G(x(\cdot))(t) = \xi + \int_{t_0}^t (A(s)x(s) + b(s))ds$ и $G^m - m$ -тая степень оператора G . По индукции доказывается, что $|F(\xi(\cdot))(t) - F(x(\cdot))(t) - (\xi(t) - x(t))| \leq \frac{c^m}{m!}\|\xi(\cdot) - x(\cdot)\|_{C(\Delta, \mathbb{R}^n)}$, где $c = \int_{\Delta} \|A(t)\|dt$. Подбрав m таким, чтобы $\frac{c^m}{m!} < 1$ и применив теорему о правом обратном, приходим к утверждению теоремы 8'. \square

Лекция 7. Определенный интеграл

Здесь будет рассказано об определенных интегралах функций одного переменного, измерении площадей плоских фигур и объемов многомерных тел и об определенном интеграле функций многих. Параллельно речь пойдет о дифференциальных формах в одномерном, двумерном и трехмерном случаях.

7.1. Определенный интеграл в одномерном случае. Фактически это понятие было введено в п. 6.1. Дадим точное определение.

Пусть f — функция, определенная на конечном отрезке $[a, b]$, $-\infty < a < b, \infty$. Раздробим отрезок $[a, b]$ на примыкающие друг к другу отрезки $\Delta_k = [\alpha_{k-1}, \alpha_k]$, $1 \leq k \leq N$ (иногда вырождающиеся в точку), порожденную системой точек $a = \alpha_0 \leq \alpha_1 \leq \dots \leq \alpha_{n+1} = b$. Обозначим это разбиение буквой D . Рассмотрим сумму $\sum_{k=1}^N f(x_k)|\Delta_k|$ (где $|\Delta|$ — это длина отрезка

Δ), а $S = \{x_i\}_{i=1}^N$ — некоторая выборка $x_k \in \Delta_k$. Эта сумма называется *римановой суммой* и обозначается $R(f(\cdot), D, S)$. Величину $\rho = \rho(D) = \max_{1 \leq k \leq N} |\Delta_k|$ назовем *размером* разбиения.

Определение 8. Число I называют определенным интегралом функции $f : [a, b] \rightarrow \mathbb{R}$ на отрезке $[a, b]$, если для любого $\varepsilon > 0$ найдется число $\delta > 0$, что для любого разбиения D отрезка $[a, b]$ размера меньшего δ , для любой выборки S выполнено неравенство $|R(f(\cdot), D, S) - I| < \varepsilon$. Эту величину обозначают $\int_a^b f(x)dx$, а саму функцию f называют *интегрируемой (по Риману)*.

Теорема 9. *Функция, непрерывная на конечном отрезке, интегрируема на нем.*

Доказательство. Пусть f — непрерывная функция на $[a, b]$. Рассмотрим некоторое разбиение $D = \cup_{i=1}^N \Delta_k$. Нижней выборкой S_- по этому разбиению назовем такую выборку, при которой в каждом отрезке Δ_k выбирается одна из точек этого отрезка, где f принимает наименьшее значение. Аналогично определяется верхняя выборка S_+ (где берется наибольшее значение). Для любой выборки S по разбиению D выполнены очевидные неравенства $R(f, D, S_-) \leq R(f, D, S) \leq R(f, D, S_+)$. Из теоремы Кантора о равномерной непрерывности по любому $\varepsilon > 0$ найдется число $\delta > 0$, для которого из неравенства $\rho(D) < \delta$ следует, что $R(f, D, S_+) - R(f, D, S_-) < \varepsilon$. Отсюда следует, что число I , равное верхней грани по D всех сумм $R(f, D, S_-)$ совпадает с нижней гранью по D всех сумм $R(f, D, S_+)$, и что это число удовлетворяет определению 8. \square

7.2. Плоские области и измерение их площадей. Для того, чтобы найти площадь квадрата, надо измерить длину его стороны, взяв в качестве масштаба, скажем, 1 см, и если длина стороны квадрата равна a см, его площадь по определению равна a^2 см². Пусть теперь мы хотим определить, что такое площадь какой-то фигуры Φ , расположенной на плоскости (скажем, полукруга, радиуса R см) и подсчитать площадь $S(\Phi)$ этой фигуры. Для определенности пусть радиус полукруга 1 см. Разместим основание полукруга на отрезке $[-1, 1]$ на оси Ox_1 в плоскости, так что полукруг оказывается расположенным между графиком функции $x_2 = \sqrt{1 - x_1^2}$ и основанием.

Вырежем (мысленно) некоторое количество квадратиков размера 1 мм² и будем выкладывать их друг к дружке, чтобы они целиком поместились в нашей фигуре (в нашем примере — в полукруге). Подсчитаем их общую площадь (которую будем всегда измерять в см²), пусть она равна a_{12} см² (a_{12} — первая оценка площади двумерной фигуры). При этом, разумеется, останутся непокрытые места. Далее будем поступать уже “умственно”: вырежем (умственно) некоторое количество квадратиков размера 0.1 мм = 10⁻² см и будем их прикладывать к объединению начальных, миллиметровых квадратиков, сколько возможно, и к тому же заполним маленькими квадратиками и миллиметровые квадратики. Пересчитаем полученную площадь снова в см², получим число a_{22} см². Ясно, что $a_{12} < a_{22}$. Далее будем поступать аналогично, вкладывая квадратiki со сторонами 10⁻ⁿ, $n \geq 3$ см, и получая площади a_{n2} см², удовлетворяющие неравенствам $a_{12} < a_{22} < \dots < a_{n2} < \dots$. Согласно известному свойству чисел, монотонно возрастающая ограниченная сверху (в нашем случае двумя квадратными сантиметрами) последовательность чисел имеет предел (этот факт называют иногда аксиомой Вейерштрасса, иногда его теоремой). Этот предел (пусть это будет число $a_{\infty 2}$) назовем *нижней площадью фигуры*. Аналогично поступим с покрытием фигуры. Сначала покроем ее

целиком наименьшим числом сантиметровых, потом миллиметровых и т. д. квадратиков. И будем получать числа $b_{12} > b_{22} > \dots > b_{n2} > \dots$, очевидно, ограниченные снизу. Предел этой последовательности (снова по аксиоме или теореме Вейерштрасса) обозначим $b_{\infty 2}$ и назовем *верхней площадью фигуры*. Ясно, что $a_{\infty 2} \leq b_{\infty 2}$.

Если числа $a_{\infty 2}$ и $b_{\infty 2}$, определенные выше, совпадают, т. е. $a_{\infty 2} = b_{\infty 2} = S$, плоскую фигуру называют *измеримой по Жордану*, а число S см² называется *площадью фигуры* Φ : $S = S(\Phi)$.

Можно доказать, что это определение площади фигуры не зависит от того, как расположена фигура на плоскости относительно декартовой системы координат и какие размеры квадратиков мы последовательно выбираем.

Если освоиться с понятием площади фигуры, то становится понятным равносильность определения 8 и следующего определения:

Определение 8₁. Определенный интеграл $\int_a^b f(x)dx$ от неотрицательной функции f , определенной на отрезке $[a, b]$, по определению — это площадь фигуры, ограниченной прямыми $x_1 = a$, $x_2 = 0$, $x_1 = b$ и графиком функции $x_2 = f(x_1)$. Если функция меняет знак, то $\int_a^b f(x)dx = \int_a^b f_+(x)dx + \int_a^b f_-(x)dx$, где $f_+(x) = \max\{f(x), 0\}$, а $f_-(x) = \min\{f(x), 0\}$.

Исторический комментарий. Измерять площади, как считается, люди научились очень давно. Полагают, что само название науки *геометрия* (*geōmetria* по-гречески “землемерие”) происходит от того, что когда-то в древнем Египте люди научились мерить площади земельных участков. Долгое время ограничивались измерением многоугольных участков (их разбивали на треугольники, площади которых равняются половине произведения основания на высоту). Вопрос о том, что такое площадь круга, потребовал больших умственных усилий. В Древней Греции стали определять площадь круга, как число, равное пределу площадей вписанных правильных многоугольников, и равного ему пределу описанных правильных многоугольников. Исходя из этого определения Архимед первый получил оценки для площади единичного круга $B^2(0, 1) = \{\bar{x} = (x_1, x_2) \mid x_1^2 + x_2^2 \leq 1\}$: $3\frac{10}{71} < S(B^2(0, 1)) < 3\frac{1}{7}$. Приведенное определение измеримости принадлежит французскому ученому К. Жордану (1838 – 1922).

Одномерные дифференциальные формы и их исчисления. Формула Ньютона – Лейбница

В одномерном случае существуют *дифференциальные формы* двух типов — функции $\omega_0 = f_0$ и выражения вида $\omega_1 = f_1(x)dx$. Формы ω_0 можно дифференцировать, и мы приходим к специальному виду формы ω_1 : $df_0(x) = f_0'(x)dx$, вторые можно интегрировать: $\int_a^b f_1(x)dx$. Из формулы Ньютона – Лейбница получаем: $\int_a^b df_0(x) = f_0(b) - f_0(a)$. За этим простым фактом можно уже усмотреть замечательную формулу Пуанкаре $\int_{\Omega} d\omega = \int_{\partial\Omega} \omega$.

7.2. Интегральное исчисление функций двух переменных

В этом разделе мы еще остаемся в школе, точнее, в тех школах, где преподают *стереометрию*. В стереометрии учат вычислять объемы. Начнем этот раздел с уточнения понятия объема пространственной фигуры. Будем двигаться по аналогии с тем, что рассказывалось в п.7.2.

Пространственные фигуры и измерение их объемов. Для того, чтобы найти объем куба, надо измерить длину его стороны, взяв в качестве масштаба, скажем, 1 см, и если длина стороны куба равна a см, его объем по определению равен a^3 см³. Пусть теперь мы хотим определить, что такое объем какой-то фигуры Ψ , расположенной в пространстве (скажем, полушара $B_+^3(O, R) = \{x \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 \leq R^2, x_3 \geq 0\}$, с центром в нуле радиуса R см) и подсчитать объем $V(B_+^3(O, R))$ этой фигуры. Для определенности пусть опять $R = 1$ см. Расположим этот полушар так, как он нарисован на рис. 5.2, так что полушар $B_+^3 = B_+^3(O, 1)$ оказывается расположенным между графиком функции $x_3 = \sqrt{1 - x_1^2 - x_2^2}$ и основанием, являющимся кругом $x_1^2 + x_2^2 = 1$, расположенным в плоскости $x_3 = 0$. Все это можно как бы “материализовать”, разрезав пополам, скажем, сантиметровой резиновый мячик и положив половину мячика на стол.

Возьмем набор одинаковых кубиков с миллиметровой длиной стороны и будем (мысленно) укладывать эти кубики в полушар “до упора”. В итоге получим первую оценку объема полушара: $V(B_+^3) > a_{13}$ см³. Затем, заполним полушар “до упора” маленькими кубиками с длиной стороны, равной 0.1 мм, и получится вторая оценка a_{23} объема $V(B_+^3)$ полушара. Продолжая умственно заполнять полушар кубиками все меньших и меньших размеров, будем получать последовательность оценок $0 < a_{23} < \dots < a_{n3} < \dots$. Последовательность эта, разумеется, ограничена сверху. Значит, она имеет предел. Обозначим его $a_{\infty 3}$ и назовем *нижним объемом фигуры*.

Аналогично поступим с покрытием фигуры. Сначала покроем полушар наименьшим числом кубиков размера кусочков сахара, потом кубиков размера кубиков сахарного песка и т. д. И будем получать числа $b_{13} > b_{23} > \dots > b_{n3} > \dots$. Предел этой последовательности (снова по аксиоме или теореме Вейерштрасса) обозначим $b_{\infty 3}$ и назовем *верхним объемом фигуры*. Ясно, что $a_{\infty 3} \leq b_{\infty 3}$.

Если числа $a_{\infty 3}$ и $b_{\infty 3}$, определенные выше, для измеряемой фигуры Ψ совпадают, т. е. $a_{\infty 3} = b_{\infty 3} = V$, пространственную фигуру называют *измеримой по Жордану*, а число V см³ называют *объемом фигуры* Ψ : $V = V(\Psi)$.

Можно доказать, что это определение не зависит от того, как расположена фигура в пространстве относительно декартовой системы координат и от размеров тех кубиков, которые используются при построении.

Если освоиться с понятием объема фигуры, то становится понятным естественность следующего определения:

Определение 8₂. Определенный интеграл $\int_{\Omega} f(x)dx$ от неотрицательной функции f , определенной на плоском множестве, измеримом по Жордану, по определению — это объем фигуры, ограниченной цилиндром над Ω и графиком функции $x_3 = f(x_1, x_2)$, с естественным дополнением, касающимся функций, меняющих знак.

Аналогичным образом индуктивно вводится определенный интеграл функций в \mathbb{R}^n .

Исторический комментарий. Измерять объемы многогранников, как считается, люди научились в Древней Греции. По свидетельству Архимеда Евдоксу (~ 408 – ~ 355 до н. э.) принадлежит прием вычисления объема пирамиды известный, как *метод исчерпывания*. Выдающуюся роль в понимании того, что такое площадь плоской фигуры и объем фигуры пространственной, сыграл сам Архимед (~ 287 – 212 до н. э.). Шедевром античной матема-

тики является вычисление им объема шара.

Лекция 8. Интегрирование дифференциальных форм

Здесь речь пойдет о формуле Пуанкаре⁵ об интегрировании дифференциальных форм. На самом деле мы ограничимся одномерным, двумерным и трехмерным случаями, о общий случай лишь слегка наметим.

8.1. Дифференциальные формы в одномерном случае и формула Ньютона – Лейбница.

Дифференциальные формы в одномерном случае определяются на отрезке $\mathcal{M}^1 = [a, b]$, который является диффеоморфным образом отрезка $B^1 = B^1(0, 1) = [-1, 1]$. Это означает, что существует непрерывно-дифференцируемое в обе стороны взаимно однозначное отображение $\varphi : B^1 \rightarrow \mathcal{M}^1$, переводящее B^1 в $[a, b]$.

Отрезок являет собой прототип того, что далее будет называться “многообразием с краем”. “Краем” отрезка $\mathcal{M}^1 = [a, b]$ является пара чисел (b, a) , снабженных знаками. Обозначим это так: $\mathcal{M}^0 = \partial\mathcal{M}^1 = \{+b, -a\}$. А все остальные точки x отрезка \mathcal{M}^1 таковы, что содержат некоторый интервал $(x - \delta, x + \delta)$, лежащий в \mathcal{M}^1 .

В одномерном случае существуют два типа дифференциальных форм $\Omega^0 = \Omega^{01}$ и $\Omega^1 = \Omega^{11}$. Ω^0 — это функции на \mathcal{M}^1 ($\omega_0 = f_0(x)$), а Ω^1 — это выражения вида $\omega_1 = f_1(x)dx$. Дифференциальные формы можно дифференцировать и интегрировать. Правила дифференцирования дифференциальных форм ω_0 и ω_1 в одномерном случае таковы: $d\omega_0 = df_0(x) = f_0'(x)dx$, $d\omega_1 = 0$. А вот каковы правила интегрирования этих дифференциальных форм: $\int_{\partial\mathcal{M}^1} \omega_0 = f_0(b) - f_0(a)$, $\int_{\mathcal{M}^1} \omega_1 = \int_a^b f_1(x)dx$.

Вспоминая то, о чем говорилось на лекции 6, приходим к следующей форме теоремы Ньютона – Лейбница:

$$\int_a^b f_0'(x)dx = f_0(b) - f_0(a) \Leftrightarrow \int_{\mathcal{M}^1} d\omega_0 = \int_{\partial\mathcal{M}^1} \omega_0.$$

Две фигуры в пространстве \mathbb{R}^n называют *диффеоморфными*, если существует взаимно-однозначное и непрерывно-дифференцируемое вместе со своим обратным отображение, переводящее одну фигуру в другую.

Дифференциальные формы, будут рассматриваться нами в двумерном случае, в основном, на областях \mathcal{M}^2 , являющихся диффеоморфизмами круга $B^2 = \{t = (t_1, t_2) \in \mathbb{R}^2 \mid t_1^2 + t_2^2 \leq 1\}$ и на кривых \mathcal{M}^1 — диффеоморфных образах отрезка $B^1 = [-1, 1]$ или окружности $\partial B^2 = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$.

Область \mathcal{M}^2 тоже является “многообразием с краем”. Ее “краем” является кривая, являющаяся образом окружности $\mathbb{S}^1 = \{t = (t_1, t_2) \in \mathbb{R}^2 \mid t_1^2 + t_2^2 = 1\}$. Обозначим эту кривую $\partial\mathcal{M}^2$. А все остальные точки x области \mathcal{M}^2 таковы, что некоторый открытый круг с центром в x

⁵точнее формуле Ньютона – Лейбница – Грина – Остроградского – Стокса – Пуанкаре

радиуса δ , лежит в \mathcal{M}^2 . Помимо всего этого существуют еще пары точек $\mathcal{M}^0 = \partial\mathcal{M}^1$, где \mathcal{M}^1 — это диффеоморфизм отрезка.

Имеется три типа дифференциальных форм в двумерном случае: $\Omega^0 = \Omega^{01}$, $\Omega^1 = \Omega^{11}$, $\Omega^2 = \Omega^{21}$. Это функции ($\omega_0 = f_0(x)$), выражения вида $\omega^2 = f_2(x)dx_1 \wedge dx_2$ на \mathcal{M}^2 , а также выражения вида $\omega^1 = f_{11}(x)dx_1 + f_{12}(x)dx_2$ на \mathcal{M}^1 . Дифференциальные формы можно дифференцировать и интегрировать. Дифференцирование превращает дифференциальную форму ω_i из Ω^i в дифференциальную форму ω_{i+1} из Ω^{i+1} , причем $d\omega_2 = 0$. Правила дифференцирования дифференциальных форм $\omega_0 - \omega_2$ таковы: $d\omega_0 = df_0(x) = \frac{\partial f_0(x)}{\partial x_1}dx_1 + \frac{\partial f_0(x)}{\partial x_2}dx_2$, $d\omega_1 = \left(\frac{\partial f_{21}(x)}{\partial x_1} - \frac{\partial f_{11}(x)}{\partial x_2}\right)dx_1 \wedge dx_2$, $d\omega_2 = 0$.

А вот каковы правила интегрирования этих дифференциальных форм: $\int_{\partial\mathcal{M}^1} \omega_0 = f_0(b) - f_0(a)$, $\int_{\mathcal{M}^1} \omega_1 = \int_{B^1} f_1(\varphi(t))\varphi'(t)dt$ (здесь $t \in [-1, 1]$), $\int_{\mathcal{M}^2} \omega_2 = \int_{B^2} f_2(\varphi(t))\varphi'(t)dt$ (здесь $t = (t_1, t_2)^T$, $t_1^2 + t_2^2 \leq 1$, $\varphi: B^2 \rightarrow \Omega^2$, $\varphi(t) = (\varphi_1(t), \varphi_2(t))^T$, $\varphi'(t)dt = \left(\frac{\partial\varphi_1(t)}{\partial t_1}dt_1 + \frac{\partial\varphi_1(t)}{\partial t_2}dt_2\right) \wedge \left(\frac{\partial\varphi_2(t)}{\partial t_1}dt_1 + \frac{\partial\varphi_2(t)}{\partial t_2}dt_2\right) = \left(\frac{\partial\varphi_1(t)}{\partial t_1}\frac{\partial\varphi_2(t)}{\partial t_2} - \frac{\partial\varphi_1(t)}{\partial t_2}\frac{\partial\varphi_2(t)}{\partial t_1}\right)dt_1 \wedge dt_2$

В итоге при интегрировании форм ω_i , $i = 0, 1, 2$ и их дифференциалов, приходим к следующему результату:

$$\int_{\partial\mathcal{M}^{i+1}} \omega_i = \int_{\mathcal{M}} d\omega_i, \quad i = 0, 1. \quad (2)$$

При $i = 0$ получаем “двумерную” формулу Ньютона – Лейбница:

$$\int_a^b \left(\frac{\partial f'_0(x)}{\partial x_1} dx_1 + \frac{\partial f'_0(x)}{\partial x_2} dx_2 \right) = f(b) - f(a). \quad (21)$$

При $i = 1$ приходим к так называемой *формуле Грина*.⁶

$$\int_{\Omega^2} d\omega_1 = \int_{\Omega^2} \left(\frac{\partial f_{12}(x)}{\partial x_1} - \frac{\partial f_{11}(x)}{\partial x_2} \right) dx_1 \vee dx_2 = \int_{\partial\Omega^2} \omega_1 = \int_{\partial\Omega^2} f_{11}(x)dx_1 + f_{12}(x)dx_2. \quad (22)$$

Доказательство этой формулы проведем сразу в многомерном случае.

8.2. Определенный интеграл и дифференциальные формы в многомерном случае. Формула Пуанкаре

Теорема 10. (Пуанкаре). Пусть $X = \mathbb{R}^n$, $B^k = \{x \in \mathbb{R}^k \mid \sum_{j=1}^k x_j^2 \leq 1\}$, $\omega_{kn} = \sum_{i_1 < i_2 < \dots < i_k} f_{i_1 i_2 \dots i_k}(x) dx_{i_1} \wedge dx_{i_2} \dots \wedge dx_{i_k}$ — дифференциальная форма k -того порядка в \mathbb{R}^n , Ω_k^n и $\partial\Omega_k^n$ — диффеоморфы $B^k = \{x \in \mathbb{R}^k \mid \sum_{j=1}^k x_j^2 \leq 1\}$ и $\partial B^k = \mathbb{S}^{k-1} = \{x \in \mathbb{R}^k \mid \sum_{j=1}^k x_j^2 = 1\}$ соответственно. Тогда $\int_{\partial\Omega_k^n} \omega_{kn} = \int_{\Omega_k^n} d\omega_{kn}$.

Доказательства этой теоремы мы не приводим.

Следствия из общей формулы Пуанкаре интегрирования дифференциальных форм в одномерном, двумерном и трехмерном случаях

⁶Дж. Грин (1793 – 1841) — английский математик, опубликовавший формулу (21) в 1828 году; она была известна еще Эйлеру.

$n = 1$ — формула Ньютона – Лейбница, $n = 2, k = 1$ — формула Грина, $n = 3, k = 2$ — формула Остроградского, $n = 3, k = 1$ — формула Стокса.

Формула Остроградского и ее физическая интерпретация. Следующие два пункта опираются на следующую формулу Пуанкаре: $\int_{\Omega} d\omega^{k-1} = \int_{\partial\Omega} \omega^{k-1}$, где Ω область в \mathbb{R}^k с гладкой границей $\partial\Omega$.

Пусть Ω — некоторая область в \mathbb{R}^3 , ограниченная гладкой поверхностью $\partial\Omega$ и $F : \Omega \rightarrow \mathbb{R}^3$ — гладкое отображение. F сопоставляет точке $x \in \mathbb{R}^3$ вектор $(F_1(x), F_2(x), F_3(x))$. Тогда говорят, что на Ω задано *векторное поле*. Применяя формулу Пуанкаре в этом частном случае (для $\omega^2 = F_1 dx_2 \wedge dx_3 + F_2 dx_3 \wedge dx_1 + F_3 dx_1 \wedge dx_2$), получаем $\int_{\partial\Omega} \omega^2 = \int_{\partial\Omega} F_1 dx_2 \wedge dx_3 + F_2 dx_3 \wedge dx_1 + F_3 dx_1 \wedge dx_2 = \int_{\Omega} d(F_1 dx_2 \wedge dx_3 + F_2 dx_3 \wedge dx_1 + F_3 dx_1 \wedge dx_2) = \int_{\Omega} \left(\frac{\partial F_1(x)}{\partial x_1} + \frac{\partial F_2(x)}{\partial x_2} + \frac{\partial F_3(x)}{\partial x_3} \right) dx_1 \wedge dx_2 \wedge dx_3$.

Это соотношение называется *формулой Остроградского*.

Число $\frac{\partial F_1(x)}{\partial x_1} + \frac{\partial F_2(x)}{\partial x_2} + \frac{\partial F_3(x)}{\partial x_3}$ называется *дивергенцией* или *расходимостью* поля F в точке x и обозначается $\operatorname{div} F(x)$. Гладкая поверхность $\partial\Omega$ в каждой точке x имеет *нормаль* $n(x)$ — единичный вектор перпендикулярный касательному пространству к поверхности в точке x . Элемент площади поверхности $\partial\Omega$ вблизи x обозначим $d\sigma(x)$. Формуле Остроградского можно придать такую форму: $\int_{\partial\Omega} \langle n(x), F(x) \rangle d\sigma(x) = \int_{\Omega} \operatorname{div}(x) dx$ (i). Выражение $\int_{\partial\Omega} \langle n(x), F(x) \rangle d\sigma(x)$ называется *поток* векторного поля F через поверхность $\partial\Omega$. Таким образом, формула Остроградского утверждает, что *интеграл от дивергенции векторного поля по объему равен потоку этого векторного поля по поверхности, ограничивающей этот объем*. В этом состоит физическая интерпретация формулы Остроградского.

Докажем эту формулу для частного случая, когда область Ω выпукла. Рассмотрим интеграл $J_1 = \int_{\Omega} \frac{\partial F_1(x)}{\partial x} dx_1 \wedge dx_2 \wedge dx_3 = \int \int \int_V \frac{\partial F_1(x)}{\partial x} dx$. Произведя при фиксированных x_2 и x_3 интегрирование по x_1 с применением формулы Ньютона – Лейбница, будем иметь: $J_1 = \int_{\Omega_{23}} (F_1(B) - F_1(A)) dx_2 \wedge dx_3$, где Ω_{23} — проекция тела Ω на плоскость $x_2 O x_3$, а $[A, B]$ — отрезок интегрирования. Пусть $d\sigma(A)$ и $d\sigma(B)$ — элементы поверхности $\partial\Omega$ в точках A и B и $n(A)$ и $n(B)$ — единичные нормали в этих точках. Тогда, как легко понять, $dx_2 \wedge dx_3(B) = d\sigma(B) \langle n(B), e^1 \rangle$, где $e^1 = (1, 0, 0)^T$ — первый орт \mathbb{R}^3 . Аналогично $dx_2 \wedge dx_3(A) = -d\sigma(A) \langle n(A), e^1 \rangle$. Прделав аналогичную операцию для двух других интегралов, приходим к формуле (i).

Формула Стокса

Пусть Ω — некоторая поверхность в \mathbb{R}^3 , границей которой является гладкая кривая $\partial\Omega$ и $F : \Omega \rightarrow \mathbb{R}^3$ — гладкое отображение из Ω в \mathbb{R}^3 . F сопоставляет точке $x \in \mathbb{R}^3$ вектор $(F_1(x), F_2(x), F_3(x))$. Тогда говорят, что на Ω задано *векторное поле*. Вектор $\det \begin{pmatrix} e^1 e^2 e^3 \\ \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} \frac{\partial}{\partial x_3} \\ F_1 F_2 F_3 \end{pmatrix} = \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \right) e^1 + \left(\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \right) e^2 + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) e^3$ называется *ротором* (или *вихрем*) векторного поля F в точке x и обозначается $\operatorname{rot} F(x)$. Гладкая поверхность Ω в каждой точке x имеет *нормаль* $n(x)$ — единичный вектор перпендикулярный касательному пространству к поверхности в точке x ; элемент площади поверхности $\partial\Omega$ вблизи x обозначим $d\sigma(x)$. Гладкая кривая $\partial\Omega$ в каждой точке x имеет единичный *касательный вектор* $\tau(x)$; элемент длины кривой $\partial\Omega$

вблизи x обозначим $ds(x)$.

Имеет место следующий частный случай формулы Пуанкаре : $\int_{\partial\Omega} \langle n(x), F(x) \rangle ds(x) = \int_{\Omega} F_1 dx_1 + F_2 dx_2 + F_3 dx_3 = \int_{\Omega} d(F_1 dx_1 + F_2 dx_2 + F_3 dx_3) = \int_{\Omega} (\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3}) dx_2 \wedge dx_3 + (\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1}) dx_3 \wedge dx_1 + (\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}) dx_1 \wedge dx_2 = \int_{\Omega} \text{rot} F(x) d\sigma(x)$ (формула Стокса).

Лекция 9. Ряды и комплексный анализ

Напоминание простейших сведений. У комплексного числа $z = x + iy$, $x, y \in \mathbb{R}$ число x называется вещественной частью z , y — мнимой его частью. Они обозначаются, соответственно, $\text{Re}z$ и $\text{Im}z$. Совокупность комплексных чисел обозначают \mathbb{C} . Функция $f : \mathbb{C} \rightarrow \mathbb{C}$ сопоставляет комплексному числу $z = x + iy$ комплексное число $f(z) = u(z) + iv(z) = u(x, y) + iv(x, y)$. Например, если $f(z) = z^2$, то $u(x, y) = x^2 - y^2$, $v(x, y) = 2xy$.

Основными понятиями вещественного анализа являются понятия производной и интеграла. Введем эти понятия для функций комплексного переменного.

9.1. Производная комплексной функции. Если f — вещественная функция, определенная в некоторой окрестности точки \hat{x} , то ее производной в точке \hat{x} называется такое число $a \in \mathbb{R}$ (в предположении, что оно существует), что $f(\hat{x} + x) = f(\hat{x}) + ax + o(|x|)$.

В комплексном случае определение аналогично: если f — комплексная функция, определенная в некоторой окрестности точки \hat{z} , то ее *производной в точке \hat{z}* называется такое число $a \in \mathbb{C}$ (в предположении, что оно существует), что $f(\hat{z} + z) = f(\hat{z}) + az + o(|z|)$. При этом говорят, что f *дифференцируема в \hat{z}* , а число a обозначается $f'(\hat{z})$. Если f дифференцируема в окрестности точки \hat{z} и производная непрерывна, как функция z в окрестности \hat{x} , говорят, что f там *аналитична*.

9.2. Условия Коши-Римана.

Предложение 1. Пусть f аналитична в окрестности точки $\hat{z} = (\hat{x}, \hat{y})$. Тогда для $u(\hat{x}, \hat{y}) = \text{Re}f(\hat{z})$ и $v(\hat{x}, \hat{y}) = \text{Im}f(\hat{z})$ выполнены соотношения: $\frac{\partial u(\hat{x}, \hat{y})}{\partial x} = \frac{\partial v(\hat{x}, \hat{y})}{\partial y}$, $\frac{\partial v(\hat{x}, \hat{y})}{\partial x} = -\frac{\partial u(\hat{x}, \hat{y})}{\partial y}$.

Определение 10. Соотношения $\frac{\partial u(\hat{x}, \hat{y})}{\partial x} = \frac{\partial v(\hat{x}, \hat{y})}{\partial y}$, $\frac{\partial v(\hat{x}, \hat{y})}{\partial x} = -\frac{\partial u(\hat{x}, \hat{y})}{\partial y}$ называются *условиями Коши - Римана*.

Доказательство предложения 1.

Имеем: $f(\hat{z} + z) = f(\hat{z}) + f'(\hat{z})z + o(|z|) = u(\hat{x} + x, \hat{y} + y) + iv(\hat{x} + x, \hat{y} + y) + o(|z|) \stackrel{\text{Id}}{=} f(\hat{z}) + \frac{\partial u}{\partial x}x + \frac{\partial v}{\partial y}y + i(\frac{\partial v}{\partial x}x + \frac{\partial u}{\partial y}y) + o(|z|)$.

Положив $f'(\hat{z}) = \xi + i\eta$, приравняв вещественную и мнимую компоненты левой и правой частей равенства, приходим к соотношению: $\xi x - \eta y + i(\xi y + \eta x) = \frac{\partial u}{\partial x}x + \frac{\partial v}{\partial y}y + i(\frac{\partial v}{\partial x}x + \frac{\partial u}{\partial y}y)$. Откуда получаем: $\frac{\partial u}{\partial x} = \xi$, $\frac{\partial u}{\partial y} = -\eta$, $\frac{\partial v}{\partial x} = \eta$, $\frac{\partial v}{\partial y} = \xi$, и следовательно, $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$, $\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}$.

Нетрудно доказать, что если функции u и v непрерывно дифференцируемы, то функция $f = u + iv$ аналитична, но мы на этом не остановимся.

9.3. Теорема Коши. Знаменитая теорема Коши, лежащая в основании комплексного анализа, является следствием формулы Грина. Формула Грина — это частный случай общей

формулы Пуанкаре, касающейся интегрирования дифференциальных форм: $\int_{\partial\Omega} (adx + bdy) = \int_{\Omega} d(adx + bdy) = \int_{\Omega} (-\frac{\partial a}{\partial y} + \frac{\partial b}{\partial x})dxdy$. Применим эту формулу к случаю, когда \mathcal{O} — область на комплексной плоскости, в которой задана аналитическая функция f и Ω — односвязная область с гладкой границей $\partial\Omega = \gamma$, расположенная в \mathcal{O} .

Теорема 11 (Коши об интеграле по замкнутому контуру). *Интеграл по гладкому замкнутому контуру, охватывающему односвязную область от аналитической функции, равен нулю.*

Доказательство. Действительно, $\int_{\gamma} f(z)dz \stackrel{\text{def}}{=} \int_{\partial\Omega} (u + iv)(dx + idy) \stackrel{\text{Id}}{=} \int_{\partial\Omega} (udx - vdy) + i(vdx + udy) \stackrel{\text{Green}}{=} \int_{\Omega} (-\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x})dxdy + i \int_{\Omega} (\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y})dxdy \stackrel{\text{Cauchy-Riemann}}{=} 0$.

□

9.4. Формула Коши.

Предложение 2. *Пусть f аналитична в некоторой области, содержащей точку z и γ — замкнутая кривая, охватывающая односвязную область, включающую точку z . Имеет место формула Коши: $\int_{\gamma} \frac{f(\zeta)d\zeta}{\zeta-z} = 2\pi i f(z)$.*

Доказательство. Проведем окружность $\gamma_{\varepsilon} = \{\zeta \in \mathbb{C} \mid |\zeta - z| = \varepsilon\}$. Если из области, ограниченной кривой γ исключить круг, ограниченный окружностью γ_{ε} , останется двусвязная область, ограниченная кривыми γ и γ_{ε} . По теореме Коши сумма интегралов, проходимых по γ в положительном направлении (против часовой стрелки) и по γ_{ε} в отрицательном направлении, то получится нуль. Значит, интеграл по γ равен $J_{\varepsilon} = \int_{|\gamma-z|=\varepsilon} \frac{f(\zeta)d\zeta}{\zeta-z}$.

Положим $\zeta = z + \varepsilon e^{it}$, $0 \leq t \leq 2\pi$ и получим, что $J_{\varepsilon} = \int_0^{2\pi} \frac{f(\varepsilon e^{it})\varepsilon i e^{it} dt}{\varepsilon e^{it}} = i \int_0^{2\pi} f(\varepsilon e^{it}) dt = 2\pi i (f(\varepsilon e^{it}) + o(\varepsilon))$. Переходя к пределу при $\varepsilon \rightarrow 0$, приходим к формуле Коши. □

Дифференцируя под знаком интеграла последовательно формулу Коши, получаем, что $f^{(k)}(z) = \frac{k!}{2\pi i} \int_{\gamma} \frac{f(\zeta)d\zeta}{(\zeta-z)^{k+1}}$. Получили, что аналитическая в некоторой области функция бесконечно дифференцируема в ней.

9.5. Ряд Маклорена. Пусть f аналитична в окрестности нуля и круг с центром в нуле радиуса ρ лежит в этой области. Тогда: $f(z) \stackrel{1.4}{=} \frac{1}{2\pi i} \int_{|z|=\rho} \frac{f(\zeta)d\zeta}{\zeta-z} \stackrel{\text{Id}}{=} \frac{1}{2\pi i} \int_{|z|=\rho} \frac{f(\zeta)d\zeta}{\zeta} \left(\sum_{k \geq 0} \frac{z^k d\zeta}{\zeta^k} \right) \stackrel{1.4}{=} \sum_{k \geq 0} \frac{f^{(k)}(0)}{k!} z^k$.

Это соотношение называется *рядом Маклорена*. Аналогично доказывается, что функция f аналитическая в некоторой области, содержащей точку z разлагается в *ряд Тейлора*: $f(z) = \sum_{k \geq 0} \frac{f^{(k)}(\widehat{z})}{k!} (z - \widehat{z})^k$.

Получили, что аналитическая в некоторой области функция не только бесконечно дифференцируема в ней, но разлагается в ней в ряд.

Лекция 10. Начала теории вероятностей

После работ Ньютона в течение длительного времени казалось, что мир детерминирован. Лишь в прошлом веке возникли сомнения в этом, и стало формироваться убеждение, что мир скорее хаотичен. Законы хаоса изучает теория вероятностей, и потому, любой курс математики должен содержать начала этой теории.

Краткий исторический экскурс. Первая глубокая теорема в теории вероятностей — закон больших чисел Якоба Бернулли (1654 – 1705) — была опубликована в его труде по теории вероятностей «Искусство предположений», который увидел свет в 1713 г. В применении к бросанию монеты закон больших чисел утверждает, что при большом числе бросаний число гербов будет примерно равно числу решек. Следующий шаг сделал Пьер Лаплас (1749 — 1827). В своем труде «Аналитическая теория вероятностей» (1812) он определил вероятность отклонения доли числа, скажем, гербов, от половины при большом числе испытаний. Этот результат стал называться *центральной предельной теоремой*. Третье имя, которое надо назвать — Андрей Николаевич Колмогоров (1903 – 1987). Он в своем классическом труде «Основные понятия теории вероятностей» (1933) разработал аксиоматику теории вероятностей (и тем включил ее в ряд точных наук), доказал многие важнейшие теоремы теории вероятностей, начал разработку теории случайных процессов и многое другое.

Теория

10.1. Вероятностное пространство. Основным объектом, изучаемым в теории вероятностей, является *вероятностное пространство*.

Определение 10. *Вероятностное пространство* (согласно Колмогорову), это тройка (Ω, \mathcal{F}, P) , где Ω — это некоторое множество точек $\{\omega\}$, называемых *элементарными событиями*, $\mathcal{F} = \{A\}$ — это некоторая совокупность подмножеств $A \subset \Omega$, называемых *событиями*, а P — это неотрицательная функция на \mathcal{F} , удовлетворяющая следующим аксиомам:

- \mathcal{F} является *алгеброй множеств*, т. е. если A_1 и A_2 события (иначе говоря, если $A_i \in \mathcal{F}$, $i = 1, 2$), то их объединение $A_1 \cup A_2$, пересечение $A_1 \cap A_2$, разность $A_1 \setminus A_2$ и дополнения \bar{A}_i , $i = 1, 2$ тоже являются событиями (т. е. принадлежат \mathcal{F}).
- Каждому множеству $A \in \mathcal{F}$ поставлено в соответствие неотрицательное действительное число $P(A)$. Это число называется *вероятностью события A* .
- $P(\Omega) = 1$.
- Если два события несовместны, (т.е. $A \cap B = \emptyset$) то вероятность того, что произойдет одно из них равна сумме вероятностей каждого: $P(A \cup B) = P(A) + P(B)$.

Два события A и B называются *независимыми*, если $P(A \cap B) = P(A)P(B)$, а если это равенство не выполняется, события называются *зависимыми*.

Из аксиом вероятностного пространства сразу следует

Теорема сложения вероятностей. *Вероятность суммы событий равна сумме вероятностей минус вероятность их совместного исхода* $(P(A \cup B) = P(A) + P(B) - P(A \cap B))$.

Доказательство. Действительно, объединение $A \cup B$ представимо, как объединение двух *непересекающихся* множеств $A \cup (B \setminus A)$ (если некий элемент принадлежит объединению множеств A и B , то он либо принадлежит A , либо принадлежит к тем элементам из B , которые A не принадлежат). По последней аксиоме $P(A \cup B) = P(A) + P(B \setminus A)$. С другой стороны B разлагается в сумму двух непересекающихся множеств $(B \setminus A) \cup (A \cap B)$, откуда $P(B) = P(A \cap B) + P(B \setminus A)$. Откуда $P(A \cap B) = P(B) - P(B \setminus A)$. Подставляя это в первое равенство, получим доказательство теоремы. \square

Вероятность события B при условии, что A произошло, называется *условной вероятностью* и обозначается $P(B|A)$. Имеет место формула: $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$. Это равенство называется *теоремой умножения вероятностей*.

Имеются три наиболее важные для нас модели вероятностного пространства: *конечное вероятностное пространство*, *счетное вероятностное пространство* и *прямая \mathbb{R}* .

10.2. Конечное и счетное вероятностные пространства. *Конечное вероятностное пространство* — это пространство Ω , состоящее из конечного числа элементарных событий $\{\omega_1, \dots, \omega_n\}$, \mathcal{F} — это совокупность всех подмножеств Ω , а вероятности элементарных событий задаются числами $\{p_i\}_{i=1}^n$: $P(\omega_i) = p_i$ ($p_i \geq 0$, $\sum_{i=1}^n p_i = 1$), и если $A = \cup_j^k \omega_{i_j}$, то $P(A) = \sum_j^k p_{i_j}$. Конечное вероятностное пространство будем представлять таблицей

$$(\Omega, \mathcal{F}, P) \leftrightarrow \begin{pmatrix} \omega_1, & \dots, & \omega_n \\ p_1, & \dots, & p_n \end{pmatrix}.$$

В дальнейшем будем иллюстрировать теорию двумя такими простейшими примерами.

Пример 1 (испытание Бернулли). Здесь $\Omega = \{\omega_0, \omega_1\}$. Исход опыта ω_1 назовем “успехом”. Пусть вероятность успеха $p_1 = p$. Тогда вероятность неуспеха равна $p_0 = q = 1 - p$. Всех событий четыре: $\mathcal{F} = \{A_0, A_1, A_2, A_3\}$, где $A_0 = \emptyset$, $A_1 = \omega_0$, $A_2 = \omega_1$, $A_3 = \Omega = A_1 \cup A_2$, и при этом $P(A_0) = 0$, $P(A_1) = p$, $P(A_2) = q$, $P(A_3) = 1$.

Скажем, испытанием Бернулли является бросание монеты, когда событие ω_1 состоит в том, что выпал герб, ω_0 в том, что выпала решка и вероятности обоих событий равны $\frac{1}{2}$ (или бросание кости, при котором в качестве ω_1 рассматривается выпадение единицы, а в качестве ω_0 — выпадение всего остального. Здесь $P(\omega_1) = \frac{1}{6}$, а $P(\omega_0) = \frac{5}{6}$).

Пример 2 (бросание кости). Здесь $\Omega = \{\omega_0, \omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, $p_i = \frac{1}{6}$, $1 \leq i \leq 6$.

Счетное вероятностное пространство — это пространство Ω , состоящее из счетного числа элементарных событий $\{\omega_1, \dots, \omega_n, \dots\}$, \mathcal{F} — это совокупность всех подмножеств Ω , а вероятности элементарных событий задаются числами $\{p_i\}_{i \in \mathbb{N}}$: $P(\omega_i) = p_i$ ($p_i \geq 0$, $\sum_{i=1}^{\infty} p_i = 1$), и если $A = \cup_j^{\mathcal{N}} \omega_{i_j}$, где \mathcal{N} — конечное число или бесконечность, то $P(A) = \sum_j^{\mathcal{N}} p_{i_j}$. Счетное вероятностное пространство будем представлять таблицей

$$(\Omega, \mathcal{F}, P) \leftrightarrow \begin{pmatrix} \omega_1, & \dots, & \omega_n, & \dots \\ p_1, & \dots, & p_n, & \dots \end{pmatrix}.$$

Пример 3 счетного вероятностного пространства представляет таблица

$$(\Omega, \mathcal{F}, P) \leftrightarrow \left(\begin{array}{ccc} \omega_1, & \dots & \omega_n, & \dots \\ e^{-\lambda}, & \dots & e^{-\lambda} \frac{\lambda^{n-1}}{(n-1)!}, & \dots \end{array} \right).$$

О прямой, как вероятностном пространстве будет сказано позже.

10.3. Случайные величины, их математические ожидания и дисперсии. Случайной величиной называется функция $f : \Omega \rightarrow \mathbb{R}$ на пространстве элементарных событий.

Пусть вероятностное пространство представляет таблица $\begin{pmatrix} \omega_1, & \dots & \omega_n \\ p_1, & \dots & p_n \end{pmatrix}$ и функция f , задающая случайную величину, такова: $f(\omega_i) = \xi_i$. Обозначим эту случайную величину буквой ξ . Ее представляет таблица: $\xi = \begin{pmatrix} \xi_1, & \dots & \xi_n \\ p_1, & \dots & p_n \end{pmatrix}$ (1). Если ξ — случайная величина, принимающая значения $\{\xi_1, \dots, \xi_n\}$, она приводит к набору вероятностей $P_\xi(x)$ — это вероятность события: случайная величина принимает значение x . Этот набор чисел называется *распределением вероятностей случайной величины ξ* .

Математическим ожиданием случайной величины ξ , задаваемой таблицей (1) называется число $M\xi = \xi_1 p_1 + \dots + \xi_n p_n$. Дисперсией $D\xi$ является число $M(\xi - M\xi)^2 = M\xi^2 - (M\xi)^2$.

Пример 4. Пусть произведено испытание Бернулли. Рассмотрим случайную величину, которая успеху сопоставляет единицу, а неудаче нуль. Вот ее таблица: $\widehat{\xi} = \begin{pmatrix} 0, & 1 \\ 1-p, & p \end{pmatrix}$ (2). Ее математическое ожидание $M\widehat{\xi} = p$, а дисперсия $D\xi = p - p^2 = p(1-p) = pq$. Легко доказывается, что математическое ожидание суммы случайных величин равно сумме математических ожиданий слагаемых.

Пример 5. Это пример функции на счетном вероятностном пространстве, описанном в примере 3. Здесь случайная величина, принимает значения $\{0, 1, \dots, m, \dots\}$, с вероятностью появления числа m , равной $P_\lambda(m) = e^{-\lambda} \frac{\lambda^m}{m!}$. Это распределение называется *распределением Пуассона*. Математическое ожидание этого распределения равно (как нетрудно подсчитать) λ . При больших n и малых m распределение Пуассона хорошо приближает биномиальное распределение: $P_n(m) \cong P_\lambda(m)$, где $\lambda = \frac{m}{n}$.

10.4. Независимые испытания (на примере испытания Бернулли). Пусть опыт, приведший к вероятностному пространству элементарных событий $\begin{pmatrix} \omega_0, & \omega_1 \\ 1-p, & p \end{pmatrix}$ повторяется еще раз (скажем, кость бросается дважды). Получим тогда четыре элементарных события $\{\omega_0\omega_0, \omega_1\omega_0, \omega_0\omega_1, \omega_1\omega_1\}$ с вероятностями, соответственно, $(1-p)^2, p(1-p), (1-p)p, p^2$, откуда получаем случайную величину для числа успехов: $\begin{pmatrix} 0 & 1 & 2 \\ (1-p)^2, & 2(1-p)p & p^2 \end{pmatrix}$. В случае n испытаний вероятность m успехов равна $C_n^m p^m (1-p)^{n-m}$, где число $C_n^m = \frac{n(n-1)\dots(n-m+1)}{m!} = \frac{n!}{m!(n-m)!}$ называется *биномиальным коэффициентом* (ибо это коэффициент при $p^m q^{n-m}$ в разложении бинома Ньютона $(p+q)^n$). Вероятностное распределение случайной величины, принимающей значения $\{0, 1, \dots, n\}$, с вероятностями $\{P_n(0), P_n(1), \dots, P_n(n)\}$, где $P_n(m) = C_n^m p^m (1-p)^{n-m}$, называется *биномиальным распределением*, равенство $P_n(m) = C_n^m p^m (1-p)^{n-m}$ называется *формулой Бернулли*. Свойства биномиального распределения: математическое ожидание равно np , дисперсия равна npq . Эти формулы вытекают из свойств ма-

тематических ожиданий и дисперсий (в случае независимости испытаний дисперсия суммы случайных величин равна сумме дисперсий).

10.5. Еще одна модель вероятностного пространства — прямая

Здесь вероятностное пространство $\Omega = \mathbb{R}$, элементарные события — точки прямой, события — объединение отрезков, интервалов и полуинтервалов, а вероятность задается монотонной функцией F , предел которой при $x \rightarrow -\infty$ равен нулю, а при $x \rightarrow \infty$ равен единице. Эта функция может иметь скачки, но является непрерывной справа. Такие функции называются *функциями распределения*. Значение функции F в точке x задает вероятность события, равного открытому лучу $(-\infty, x)$. Такое вероятностное пространство можно интерпретировать, как случайную величину ξ_F . В важнейших случаях функция распределения задается плотностью $x \mapsto p(x)$, которая является неотрицательной функцией интеграл от которой по всей прямой равен единице. Тогда $F(x) = \int_{-\infty}^x p(t)dt$.

Математическим ожиданием случайной величины ξ_F , задаваемой функцией распределения F , порожденной плотностью p , называется число $M\xi_F = \int_{\mathbb{R}} xp(x)dx$. *Дисперсией* $D\xi_F$ является число $M(\xi_F - M\xi_F)^2 = M\xi_F^2 - (M\xi_F)^2$.

Приведем три непрерывных распределения.

1. **Равномерное распределение** на отрезке $[a, b]$ имеет плотность, равную нулю всюду, кроме этого отрезка, а на отрезке оно равно константе, равной $\frac{1}{b-a}$.
2. **Показательное распределение** задается функцией распределения $F_{\text{exp}}(x)$, равной нулю, если $x < 0$ и $1 - e^{-\lambda x}$, если $x \geq 0$.
3. **Нормальное распределение**. “Каноническое” нормальное распределение задается плотностью $\varphi(x) = p_{\text{norm}}(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$. Его называют *распределением Гаусса* (или иногда *распределением Лапласа*).

Основные результаты

Неравенство Чебышева. Пусть ξ — некоторая случайная величина. Имеет место неравенство: $P\{|\xi - M\xi| > \alpha\} \leq \frac{D\xi}{\alpha^2}$.

Теорема 10 (закон больших чисел Я. Бернулли). В схеме независимых испытаний Бернулли вероятность того, что доля успехов по модулю отличается от вероятности успеха на любую положительную величину стремится к нулю при числе испытаний, стремящемся к бесконечности. В виде формулы это выглядит так: $P\{|\frac{m}{n} - p| > \varepsilon\} \rightarrow 0$ при $n \rightarrow \infty$.

Центральная предельная теорема П. Лапласа.

а) Локальная: Если число n велико, а m достаточно близко к n , то $\sqrt{npq}P_n(m) : \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \rightarrow 1$ для всех m для которых $x = \frac{m-np}{\sqrt{npq}}$ находится в конечном интервале.

б) Интегральная: Пусть $x_i = \frac{m_i-np}{\sqrt{npq}}, i = 1, 2$. Тогда вероятность того, что число успехов при n испытаниях Бернулли удовлетворяет неравенствам $m_1 \leq m \leq m_2$ примерно равна

$$\frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx.$$

Доказательства

1) *Закон больших чисел.* Вводим случайную величину ξ_k , равную числу наступлений события A при k -том испытании, получаем, что $\xi = \xi_1 + \dots + \xi_n$. При этом $M\xi_k = p$, $D\xi_k = p(1-p)$, откуда $P\{|\frac{m}{n} - p| > \varepsilon\} \stackrel{\text{Cheb.in eq}}{\leq} \frac{np(1-p)}{n^2} \rightarrow 0$, когда $n \rightarrow \infty$. \square

2) *Локальная центральная предельная теорема.* Доказательство см. в книге Б. В. Гнеденко «Курс теории вероятностей».

Дополнения.

Д1. Математика и естествознание

Законы Ньютона.

В 1687 году вышла книга Ньютона “Математические начала натуральной философии”, быть может, величайшая в истории науки. Среди многого другого Ньютон там сформулировал законы динамики. Впоследствии они были истолкованы на языке теории экстремумов.

Уравнение $m\ddot{x} + kx = 0$ как и остальные уравнения динамики пишутся, основываясь на некоем экстремальном принципе, который имеет долгую историю. Назовем некоторых ученых, которые вложили свою лепту в формирование этого принципа. Это Мопертюи, Эйлер, Лагранж и Гамильтон. Некогда он назывался принципом наименьшего действия. При рассмотрении механической системы, имеющей координаты $(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n)$ рассматривается ее лагранжиан $L(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n)$, являющийся разностью кинетической и потенциальной энергий $L = T - U$. Интеграл от лагранжиана называется *действием*. Законы, описывающие движения механических систем, являются условиями стационарности функционала действия (некогда это называли *принципом наименьшего действия*). Эти условия стационарности состоят в системе уравнений: $-\frac{d}{dt}L_{\dot{q}_i} + L_{q_i} = 0$. Например, уравнение $m\ddot{x} + kx = 0$ является условием стационарности лагранжиана $L = m\frac{\dot{x}^2}{2} - k\frac{x^2}{2}$. Сказанное дает возможность дать эскиз решения задачи Кеплера.

Сила F , притягивающая движущееся тело массы m (находящееся в точке плоскости, имеющей декартовы координаты $(x(t), y(t))$ и полярные координаты $(r(t), \varphi(t))$) к Солнцу, расположенному в начале координат, равна по закону всемирного тяготения $-kr^{-2}(t)$. Кинетическая энергия T тела равна $\frac{m}{2}(\dot{x}^2(t) + \dot{y}^2(t)) = \frac{m}{2}(\frac{d}{dt}(r(t) \cos \varphi(t))^2 + (\frac{d}{dt}(r(t) \sin \varphi(t))^2) = \frac{m}{2}(\dot{r}^2(t) + r^2(t)\dot{\varphi}^2(t))$, потенциальная U равна $\frac{mk}{r(t)}$. Лагранжиан $T - U$ не зависит от времени, следовательно, имеет место *закон сохранения энергии* $\frac{m}{2}(\dot{r}^2(t) + r^2(t)\dot{\varphi}^2(t)) + \frac{mk}{r(t)} = H$. Он не зависит также от φ и потому имеет место *закон сохранения импульса* $mr^2(t)\dot{\varphi}(t) = mL$, который перепишем в виде второго кеплеровского закона площадей:

$$\frac{d\varphi}{dt} = \frac{L}{r^2}. \quad (i)$$

Закон сохранения энергии (с учетом закона сохранения импульса) перепишем в виде $\frac{m}{2}(\dot{r}^2(t) + (\frac{L}{mr^2(t)})^2) + \frac{mk}{r(t)} = H$, что равносильно дифференциальному уравнению:

$$\frac{dr}{dt} = \sqrt{\frac{2}{m}(H - \frac{mk}{r}) - \frac{L^2}{m^2r^2}}. \quad (ii)$$

Соотношения (i) и (ii) приводят к интегрируемому дифференциальному уравнению:

$$\frac{d\varphi}{dr} = \frac{L}{r^2 \sqrt{\frac{2}{m}(H - \frac{mk}{r}) - \frac{L^2}{m^2r^2}}}. \quad (iii)$$

(Дифференциальное уравнение (iii) можно проинтегрировать подстановкой Эйлера).

Производя интегрирование, получаем:

$$r = \frac{p}{1 + e \cos \varphi}, \quad p = \frac{L^2}{|k|}, \quad e = \sqrt{\frac{2HL^2}{mk^2} + 1}, \quad L = \sqrt{|k|p}.$$

Получилось, что небесные тела под воздействием Солнца, движутся по кривым второго порядка. При этом, если $e < 1$, то движение проходит по эллипсу (этим доказан первый закон Кеплера). Соотношение (i), как было уже сказано, представляет собой второй закон Кеплера. Докажем третий закон. Интегрируя (i) по периоду, и используя то, что площадь эллипса с полуосями a и b равна πab , приходим к равенству: $\pi ab = LT$, откуда, используя то, что $\frac{b^2}{a} = p$, $L^2 = |k|p$, получаем: $T^2 = \frac{\pi^2 a^2 b^2}{L^2} = \frac{a^3}{|k|}$, что и требовалось.

Кеплер задумал написать книгу об устройстве Вселенной, завершением которой должен был стать третий закон движения планет. Там есть такие слова: “Я пишу эту книгу. Прочтется ли она моими современниками или потомством, мне нет до этого дела — она подождет своего читателя. Разве Господь Бог не ждал шесть тысяч лет созерцателя своего творения?”

Метод Фурье в задачах математической физики.

Колебание струны (Д. Бернулли и Эйлер)

Пусть $u(t, x)$ — положение в момент t колеблющейся, закрепленной в концах отрезка $[0, 1]$ струны, начальное положение которой в точке x было $\varphi(x)$, а скорость в точке x в начальный момент времени была равна $\psi(x)$. Функция $(t, x) \mapsto u(t, x)$ удовлетворяет уравнению струны

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \quad (i)$$

краевым условиям $u(t, 0) = u(t, 1) = 0$ и начальным условиям $u(0, x) = \varphi(x)$, $u_t(0, x) = \psi(x)$. Решим задачу методом разделения переменных. Ищем решение уравнения струны в

виде произведения $T(t)X(x)$. Тогда уравнение $\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}$ приводит к равенству $\frac{T''(t)}{T(t)} = \frac{X''(x)}{X(x)}$.

Значит, и $\frac{T''(t)}{T(t)}$ и $\frac{X''(x)}{X(x)}$ равны константе, которую обозначим через λ . Краевые условия ведут к равенствам $X(0) = X(1) = 0$. Таким образом, функция $X(\cdot)$ является решением задачи Коши $X'' = \lambda X$, $X(0) = X(1) = 0$. Отсюда следует, что $\lambda_k = -k^2$, и решениями задачи Коши являются функции $c_k \sin k\pi x$, $k \in \mathbb{N}$. Остается применить теорему Гильберта, о которой я рассказывал на предыдущей школе. Положим $G(x, \xi) = (1 - \xi)x$ при $0 \leq \xi \leq x$ и $(1 - x)\xi$ при $x \leq \xi \leq 1$. Дважды продифференцировав по x функцию $z(x) = \int_0^1 G(x, \xi)y(\xi)d\xi$ убеждаемся в том, что $z''(x) = y(x)$ и при этом $z(0) = z(1) = 0$. Система функций $\{2^{-1/2} \sin k\pi x\}_{k \in \mathbb{N}}$ является полной ортонормированной системой в $L_2([0, 1])$. Остается решить уравнения $T''(t) + n^2 T(t) = 0$ и тогда общее решение уравнения струны будет таково: $u(t, x) = \sum_{k \in \mathbb{N}} (c_k \sin nt + d_k \cos nt) \sin k\pi x$, где c_k и d_k определяются однозначно из начальных данных.

Уравнение колебаний струны (i) вывел Даламбер в 1747 году. Решение задачи на всей прямой $\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}$, $u(0, x) = \varphi(x)$, $u_t(0, x) = \psi(x)$ задает формула: $u(t, x) = \frac{\varphi(x+t) + \varphi(x-t)}{2} + \frac{1}{2} \int_{x-t}^{x+t} \psi(\tau) d\tau$, которая была получена Даламбером и Эйлером и получила название *формулы Даламбера*.

Подобную формулу можно записать и для смешанной задачи, решение которой было описано выше. Возник известный спор между Эйлером и Д. Бернулли по поводу того, чье решение является более общим. Нельзя сказать, что этот спор поныне окончательно решен.

Стационарное состояние мембраны (Лаплас).

Лаплас был одним из крупнейших ученых 18 века. Он подробно исследовал уравнение $\Delta u = 0$, получившее его имя, которое лежит в основании теории потенциала, теплопроводности, электростатики и гидродинамики.

Проинтегрируем уравнение Лапласа в круге. Пусть на окружности единичного радиуса задана непрерывная функция $\varphi \mapsto f(\varphi)$, $0 \leq \varphi \leq 2\pi$. Требуется решить задачу Дирихле, т. е. найти решение уравнения Лапласа

$$\Delta u(x_1, x_2) = \frac{\partial^2 u(x_1, x_2)}{\partial x_1^2} + \frac{\partial^2 u(x_1, x_2)}{\partial x_2^2} = 0, \quad (i)$$

принимая на границе круга значения, равные f . Уравнение Лапласа в полярных координатах имеет вид: $\frac{\partial^2 u}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial u}{\partial \rho} + \frac{1}{\rho^2} \frac{\partial^2 u}{\partial \varphi^2} = 0$.

Решим задачу снова методом разделения переменных. Как и в случае колебаний струны, ищем решение уравнения Лапласа в виде произведения $R(\rho)\Phi(\varphi)$. В результате уже дважды примененной процедуры, приходим к уравнениям $\Phi'' + \lambda\Phi = 0$ с периодическими условиями и к уравнению $\rho^2 R'' + \rho R' - \lambda R = 0$. Получаем $\lambda_k = k^2$ и полную систему $\{\cos k\cdot, \sin k\cdot\}_{k \geq 0}$. Уравнение $\rho^2 R'' + \rho R' - k^2 R = 0$ имеет решение ρ^k , и тогда общее решение уравнения Лапласа будет таково: $u(t, x) = \frac{a_0}{2} + \sum_{k \in \mathbb{N}} \rho^k (a_k \sin k\varphi + b_k \cos k\varphi)$, где a_k и b_k определяются как коэффициенты Фурье функции f .

Распространение тепла (Фурье)

Фурье — выдающийся математик и математический физик. Он вывел уравнение теплопроводности и проинтегрировал его при различных граничных условиях методом разделения переменных.

Пусть $u(t, x)$ — температура в момент t металлической нити, концы которой, расположенные в точках 0 и 1, содержатся при температуре в ноль градусов, а начальная температура в точке x равна $\varphi(x)$. Функция $(t, x) \mapsto u(t, x)$ удовлетворяет уравнению теплопроводности

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad (i)$$

краевым условиям $u(t, 0) = u(t, 1) = 0$ и начальному условию $u(0, x) = \varphi(x)$. Решим задачу опять-таки методом разделения переменных. Как и в случае колебаний струны, ищем решение уравнения теплопроводности в виде произведения $T(t)X(x)$. Тогда уравнение $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$ приводит к равенству $\frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)}$. Значит, и $\frac{T'(t)}{T(t)}$ и $\frac{X''(x)}{X(x)}$ равны константе, которую обозначим через λ . Краевые условия ведут к равенствам $X(0) = X(1) = 0$. Таким образом, функция $X(\cdot)$ является решением задачи Коши $X'' = \lambda X$, $X(0) = X(1) = 0$. Отсюда следует, что $\lambda_k = -k^2$, и решениями задачи Коши являются функции $c_k \sin k\pi x$, $k \in \mathbb{N}$. Применить теорему Гильберта (как это было проделано в предыдущем пункте) приводит к тому, что система функций $\{2^{-1/2} \sin k\pi x\}_{k \in \mathbb{N}}$ является полной ортонормированной системой в $L_2([0, 1])$. Остается решить уравнения $T'(t) + n^2 T(t) = 0$ и тогда общее решение уравнения струны будет таково: $u(t, x) = \sum_{k \in \mathbb{N}} a_k e^{-n^2 t} \sin k\pi x$, где a_k определяются однозначно из начальных данных.

Как-то возник спор между Фурье и Якоби. Фурье написал, что цель математики в объяснении природы. Якоби же дал отповедь Фурье, сказав, что философам подобного ранга неплохо бы знать, что цель математики в прославлении человеческого разума. Каждый из вас может иметь собственное мнение на этот счет. Мне ближе воззрение Фурье.

Числа. Совокупность натуральных чисел $\{1, 2, 3, \dots\}$ обозначается \mathbb{N} , множество вещественных чисел —

через \mathbb{R} . Вещественные числа можно определять конструктивно, как совокупность всевозможных десятичных дробей, не оканчивающихся на одни девятки, и аксиоматически, как *полное упорядоченное поле*. Десятичная дробь x — это последовательность чисел вида: $m_1m_2 \cdots m_N.n_1n_2 \cdots$, где m_i , $1 \leq i \leq N$ и n_i , $i \in \mathbb{N}$ — целые числа от нуля до девяти. Не допускается, чтобы все n_i , начиная с некоторого номера, были бы все девятки; дробь называется конечной, если, начиная с некоторого номера, все n_i нули; $m_1m_2 \cdots m_N$ называется целой частью дроби, остальная часть называется дробной. Естественно определяется сравнение дробей (когда $x = y$ и $x < y$); с помощью действий над конечными дробями и понятия предела определяются сложение, вычитание, умножение и деление десятичных дробей, что превращает их совокупность (т. е. \mathbb{R}) в поле (т. е. множество, в котором определены операции сложения, умножения и деления с привычными аксиомами). Вещественные числа обладают свойствами а) полноты; согласно ему фундаментальная последовательность $\{x_n\}_{n \in \mathbb{N}}$ (т. е. такая последовательность $\{x_n\}_{n \in \mathbb{N}}$, что $\forall \varepsilon > 0 \exists N \in \mathbb{N} : |x_{n+k} - x_n| < \varepsilon \forall n \geq N, k \in \mathbb{N}$) имеет предел; б) существования верхней (нижней) грани для ограниченного сверху (снизу) числового множества (для верхней грани это означает, что если $a < C$ для любого числа a из числового множества A , то $\exists \alpha \in \mathbb{R} : \alpha \geq a \forall a \in A$ & $\forall \varepsilon > 0 \exists a(\varepsilon) : a(\varepsilon) > \alpha - \varepsilon$ и с) убывающей последовательности отрезков с длинами, стремящимися к нулю иметь единственную общую точку. Таким образом, описанная нами совокупность десятичных дробей — это модель аксиоматически заданного объекта — полного упорядоченного поля.