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:

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

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 $(1)_n$.

[1] Короткий экскурс в теорию вещественного числа содержится в Дополнении

Если все коэффициенты a_{ij} уравнения $(1)_n$ равны нулю, то решение возможно лишь если $y_1 = y_2 = \ldots = y_n = 0$ и им является любые n чисел $x_1, x_2, \ldots, x_n$; 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 $x_1, x_2, \ldots, x_{n-1}$ 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 $\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_{n-1}$, а затем и $\overline{x}_n$. 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 $x_1, \ldots, x_n$ мы иногда записываем в строку $(x_1, \ldots, x_n)^T$, 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 $y = (y_1, \ldots, y_n)$ из $\mathbb{R}^{n*}$ и вектора-столбца $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ из $\mathbb{R}^n$ определяется *внутреннее произведение* $y \cdot x = y_1x_1 + \ldots + y_nx_n$. Выражение $|x| = \sqrt{x^T \cdot x} = (x_1^2 + \ldots + x_n^2)^{1/2}$ $(|y| = \sqrt{y \cdot y^T} = (y_1^2 + \ldots + y_n^2)^{1/2})$ называется *модулем* (или *длиной)* вектора $x \in \mathbb{R}^n$ $(y \in \mathbb{R}^{n*})$.

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

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

Произведение матрицы A на вектор-столбец x (обозначаемое Ax) определяется, как вектор-столбец $(a^1 \cdot x, \ldots, a^n \cdot x)^T$. 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: $A(\alpha_1 x^1 + \alpha_2 x^2) = \alpha_1 Ax_1 + \alpha_2 Ax_2$. Потому A называют *линейным оператором.*

Теорема 1 (альтернатива[2] 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*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n+1}x_{n+1} &= 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n+1}x_{n+1} &= 0, \\ \ldots \quad \ldots & \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn+1}x_{n+1} &= 0 \end{aligned} \tag{$(1_0)_n$}$$

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 $a_{ij},\ 1 \le i \le n+1,\ 1 \le j \le n$ равны нулю, то любые не равные одновременно нулю числа $x_i \in \mathbb{R},\ 1 \le i \le n+1$ 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 $\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_{n-1}$, а тогда через них выразим $\widehat{x}_n$ и тем самым придем к цели. □

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 $\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_{n+1}$ не равные одновременно нулю и такие, что $\widehat{x}_1 b^1 + \widehat{x}_2 b^2 + \ldots + \widehat{x}_{n+1} b^{n+1} = 0$. Подействовав на это равенство матрицей A, используя линейное свойство матрицы A и определения столбцов $b^1, b^2, \ldots, b^{n+1}$, получим, что $\widehat{x}_2 b^1 + \widehat{x}_3 b^2 + \ldots + \widehat{x}_{n+1} b^n = 0$. Подействовав теперь на это новое равенство оператором A, и поступая далее аналогично, получим, что $\widehat{x}_{n+1} b^1 = 0$. Но мы предположили, что $b^1 \neq 0$, значит $\widehat{x}_{n+1} = 0$. Аналогично докажем, что $\widehat{x}_n = \widehat{x}_{n-1} = \ldots = \widehat{x}_1 = 0$. Пришли к противоречию: мы ведь предполагали, что не все $\widehat{x}_i$ равны нулю. Значит, из разрешимости системы следует, что однородное уравнение имеет только нулевое решение.

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

[2]Слово «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

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

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

Переходим к *условиям разрешимости систем уравнений.*

Нам понадобится для этого определение ранга. Приведем его.

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

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

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

1) если столбцы $a^1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \ldots, a^n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix}$ *линейно независимы, тогда имеется единственное решение системы $(1)_n$; 2) если столбцы $a^1, \ldots, a^n$ 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 $a^1, \ldots, a^n$ 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 $a^1, \ldots, a^n$ 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 $\tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_n)^T$ — решение однородной системы и вектор $\widehat{x} = (\widehat{x}_1, \ldots, \widehat{x}_n)^T$ — решение неоднородной системы, т. е. $\tilde{x}_1 a^1 + \ldots + \tilde{x}_n a^n = 0$ и $\widehat{x}_1 a^1 + \ldots + \widehat{x}_n a^n = y$. Тогда вектор $\widehat{x} + t\tilde{x}$ 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 $\widetilde{C}$, 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 $\alpha_1(k), \ldots, \alpha_r(k)$ 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 $\alpha_1(k) c^1 + \ldots + \alpha_r(k) c^r$ 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 $\alpha_1(j) c_{l1} + \ldots + \alpha_r(k)_{lr}$ не совпадают друг с другом. Это означает, что столбцы у матрицы размеров $(r+1) \times (r+1)$, получающейся присоединением столбца $(c_{1k}, \ldots, c_{rk}, c_{lk})^T$ и строки, составленной из элементов $c_{l1}, \ldots, c_{lr}$ 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 $c^1, \ldots, c^r, y$ линейно независимы, т.е. ранг $\widetilde{C}$ 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 Людвидж 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 $x = (x_1, x_2, \ldots)^T$ таких, что $\sum_{i\in\mathbb{N}} x_i^2 < \infty$, $A = (a_{ij})_{1\le i,j\le\infty}$ – бесконечная матрица. Требуется решить такую систему бесконечного числа уравнений с бесконечным числом неизвестных:

$$\begin{matrix} a_{11}x_1 + a_{12}x_2 + \ldots = b_1 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots = b_n \\ \cdots \end{matrix} \quad \Leftrightarrow \sum_{j\ge1} a_{ij}x_j = b_i, \quad i \ge 1. \tag{$1)_\infty$}$$

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.