ЧАСТЬ I.
ОСНОВЫ
ЛИНЕЙНОЙ АЛГЕБРЫ

Введение. Основные определения

Алгебра — это часть математики, занимающаяся решением различных алгебраических уравнений, в которые неизвестные могут входить в любой степени. **Степенью уравнения** называется наибольший из показателей степеней при неизвестном. Решение уравнений первой и второй степени известно еще с древности (ученые Вавилона и Древнего Китая еще за 2000 лет до н. э. умели решать эти уравнения), в XVI в. итальянские математики нашли решение уравнений третьей и четвертой степени. Решение уравнений выше четвертой степени, как доказали ***Н. Абель*** и ***Э. Галуа*** в XIX веке (1830 г.), нельзя выразить в общем случае через коэффициенты уравнения посредством алгебраических операций, хотя и, как установлено ***К. Гауссом***, всякое алгебраическое уравнение n-й степени имеет n действительных или мнимых решений.

Теория решения алгебраических уравнений первой степени выделилась в самостоятельный раздел алгебры — линейную алгебру, которая в настоящее время получила большое распространение в различных приложениях.

Алгебраическим уравнением называется уравнение, обе части которого являются многочленом или одночленом по отношению к неизвестным величинам. **Одночлен** — это произведение двух или нескольких сомножителей, **многочлен** — это сумма нескольких одночленов.

Линейная алгебра — это часть алгебры, содержащая теорию линейных уравнений. Она получила свое дальнейшее развитие после возникновения учений об определителях и матрицах.

Линейное уравнение — это алгебраическое уравнение, в котором есть неизвестные в первой степени и отсутствуют члены, содержащие произведения неизвестных. Так, например, линейное уравнение с одним неизвестным имеет вид $ax = b$ (его решение $x = \frac{b}{a}$). Линейное уравнение с двумя неизвестными имеет вид $ax + by = c$, оно имеет бесчисленное множество решений. Если неизвестных несколько, то для их нахождения используются системы линейных уравнений (линейные системы). Например, система линейных уравнений с двумя неизвестными имеет следующий общий вид:

$$\begin{cases} ax + by = c, \\ dx + ey = f, \end{cases}$$

но обычно коэффициенты системы обозначаются следующим образом:

$$\begin{cases} a_{11}x + a_{12}y = b_1, \\ a_{21}x + a_{22}y = b_2, \end{cases}$$

В общем виде линейная система m уравнений с n неизвестными записывается следующим образом:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases} \tag{1}$$

Числа a_{ij} — вещественные (действительные) коэффициенты, индекс i обозначает номер уравнения, индекс j — номер неизвестного x_j, b_i ($i = 1, 2, \ldots, m$) — свободные члены уравнений системы. Если среди свободных членов имеются не равные нулю, то линейная система называется **неоднородной**. Если все числа $b_i = 0$ ($i = 1, 2, \ldots, m$), то линейная система называется **однородной** и имеет вид:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0. \end{cases}$$

Данная однородная система иначе называется **ассоциированной** с системой (1) (или **приведенной** системой для системы (1)).

Решением линейной системы с n неизвестными называется упорядоченная совокупность n чисел, подстановка которых вместо неизвестных обращает в тождества все уравнения этой системы.

Если у системы линейных уравнений имеется несколько решений, то эта система называется **совместной неопределенной**, если решение единственное — **совместной определенной**, если решений нет — **несовместной**.

Две линейные системы называются **эквивалентными (равносильными)**, если они обе несовместны (не имеют решений) или если их решения одинаковы.

Определителем (или детерминантом) называется составленное по определенному правилу математическое выражение из n^2 чисел, которое применяется при решении и исследовании систем линейных уравнений. Число n является порядком определителя.

Матрицей называется составленная по определенному правилу прямоугольная таблица с определенным количеством строк и столбцов, применяемая в различных областях математики и физики. Матрица, составленная из m строк и n столбцов, применяется при решении и исследовании систем m линейных уравнений с n неизвестными.

Понятие линейности применяется не только по отношению к алгебраическим уравнениям, оно используется и в других областях математики, например при решении линейных дифференциальных уравнений.

ЛЕКЦИЯ №1. Простейшие случаи. Решение систем

1. *Простейшие случаи. Решение системы двух линейных уравнений с двумя неизвестными (способ подстановки и способ сложения или вычитания)*

Рассмотрим простейшие способы решения систем двух (трех) линейных уравнений с двумя (тремя) неизвестными, известные из школьного курса алгебры. Однако принципы работы этих способов применяются и при работе (и помогают в понимании) с системами n линейных уравнений в общем случае.

Решение системы двух линейных уравнений с двумя неизвестными можно свести к решению одного линейного уравнения с одним неизвестным способом подстановки или способом сложения (вычитания).

Суть **способа подстановки** состоит в том, что из одного линейного уравнения системы одно из неизвестных выражается через другое неизвестное, а затем полученное выражение подставляется в другое уравнение системы. Таким образом, получается линейное уравнение с одним неизвестным. Например, имеется система $\begin{cases} 2x + y = 6 \\ 3x - 2y = 2 \end{cases}$. Выражаем y из первого уравнения через x: $y = 6 - 2x$. Подставляя это выражение во второе уравнение системы, получаем: $3x - 2(6 - 2x) = 2$, отсюда $7x = 14$, $x = 2$, следовательно, $y = 2$.

Суть **способа сложения (вычитания)** состоит в том, что каждое из уравнений системы умножается на свой множитель. Эти множители подбираются так, чтобы после умножения на них коэффициенты при одном из неизвестных в обоих уравнениях по абсолютной величине были бы равны. Затем сложением (или вычитанием) уравнений системы исключается неизвестное, получившее одинаковые по абсолютной величине коэффициенты. Получается одно уравнение с одним неизвестным, после нахождения которого при подстановке его в любое уравнение исходной системы определяется другое неизвестное.

Например, имеется система $\begin{cases} 2x + y = 6, \\ 3x - 2y = 5{,}5. \end{cases}$

Первое уравнение умножается на 3, второе — на 2. Получается: $\begin{cases} 6x + 3y = 18 \\ 6x - 4y = 11 \end{cases}$. После вычитания из первого уравнения системы второго получаем: $7y = 7$ или $y = 1$.

Если умножить первое уравнение исходной системы на 2, а второе — на 1 (т. е. оставить без изменения), то получаем

$\begin{cases} 4x + 2y = 12 \\ 3x - 2y = 5{,}5 \end{cases}$. После сложения уравнений системы получаем: $7x = 17{,}5$ или $x = 2{,}5$.

После подстановки найденного неизвестного в любое уравнение исходной системы находим второе неизвестное.

В общем случае система двух линейных уравнений имеет вид:

$$\begin{cases} a_{11}x + a_{12}y = b_1, \\ a_{21}x + a_{22} = b_2. \end{cases}$$

Возможны три случая.

1. Коэффициенты уравнений системы не пропорциональны: $a_{11}/a_{21} \neq a_{12}/a_{22}$. В этом случае при любых свободных членах система уравнений имеет единственное решение.

2. Коэффициенты уравнений и свободные члены пропорциональны: $a_{11}/a_{21} = a_{12}/a_{22} = b_1/b_2$. В этом случае система имеет бесчисленное множество решений (одно уравнение является следствием другого).

3. Коэффициенты уравнения пропорциональны, а свободные члены не находятся в том же соотношении: $a_{11}/a_{21} = a_{12}/a_{22} \neq b_1/b_2$. В этом случае система уравнений не имеет решений (уравнения системы противоречат друг другу).

Выражение вида $a_{11}a_{22} - a_{12}a_{2}$ (это число) называется определителем второго порядка и обозначается:

$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{2}$ (в дальнейшем будет введено более строгое определение определителя в общем случае), т. е. из коэффициентов данной системы можно составить определитель второго порядка. Тогда, используя введенное таким образом понятие определителя, можно сказать, что система двух линейных уравнений имеет единственное решение при $|A| \neq 0$. Если определитель $|A| = 0$, то в зависимости от величин свободных членов система имеет либо бесчисленное множество решений, либо не имеет решений.

Решение исходной системы с помощью определителя записывается следующим образом: $x = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} / |A|, \quad y = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} / |A|$.

Как видим, в числителях формул стоят определители, полученные из определителя $|A|$ заменой столбца из коэффициентов при соответствующем неизвестном на свободные члены. Если оба определителя, стоящие в числителях не равны нулю при равном нулю определителе $|A| = 0$, то система не имеет решений. Если определители в числителях равны нулю (если один определитель в числителе равен нулю, то и другой тоже равен нулю) при $|A| = 0$, то исходная система имеет бесчисленное множество решений.

2. Решение системы трех линейных уравнений с тремя неизвестными

Для решения системы трех и более линейных уравнений с тремя и более неизвестными используются аналогичные способы и правила. Например, пусть имеется линейная система трех уравнений с тремя неизвестными: $\begin{cases} x + y - z = 5 \\ 2x - 3y + 4z = 10 \\ 5x + 2y - 3z = 9 \end{cases}$. Сначала рассмотрим любые два уравнения этой системы, например первое и второе. Будем считать, что z известно, тогда, используя приемы решения линейной системы двух уравнений, находим неизвестные x и y: $x = 5 - z / 5$, $y = 6z / 5$. Подставляя найденные значения x и y в третье уравнение системы, находим неизвестное z: $5(5 - z / 5) + 2(6z / 5) — 3z = 9$, отсюда $z = 10$, и тогда можно окончательно определить x и y: $x = 3$, $y = 12$.

Данную систему можно решать и способом подстановки: из одного уравнения системы одно неизвестное, например x, выражается через y и z, подставляется в два других уравнения системы, получаются два уравнения относительно двух неизвестных y и z, которые решаются как система двух линейных уравнений. Затем, найдя *y* и *z*, определяем неизвестное *x*.

В общем виде линейная система трех уравнений с тремя неизвестными записывается следующим образом: $\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3. \end{cases}$

К данной системе также можно применить понятие определителя, но в этом случае определителем будет называться выражение вида (число) $a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$, обозначается определитель третьего порядка следующим образом:

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} -$$

$$- a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Данное выражение легко запомнить, если использовать следующий графический способ: составим новую таблицу, прибавив к столбцам определителя еще раз первый и второй столбцы:

$$\begin{matrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{matrix}.$$

Произведения элементов, стоящих на полных диагоналях, параллельных линии $a_{11}a_{22}a_{33}$, берутся со знаком плюс, а произведения элементов, стоящих на полных диагоналях, параллельных линии $a_{13}a_{22}a_{31}$, берутся со знаком минус.

Решение линейной системы трех уравнений с тремя неизвестными может быть записано и с помощью определителей по тем же правилам, что и в случае двух линейных уравнений с двумя неизвестными:

$$x_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix} / |A|, x_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix} / |A|, x_3 = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix} / |A|.$$

Как видим, в числителях стоят определители, полученные из определителя $|A|$ заменой соответствующего столбца коэффициентов на столбец свободных членов.

3. Эквивалентность линейных систем, элементарные преобразования. Приведение к ступенчатому виду (метод Гаусса)

Две линейные системы с одинаковым количеством уравнений и неизвестных являются эквивалентными, если они несовместны или совместны и имеют одни и те же решения. Пусть имеется исходная линейная система

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases} \quad (1)$$

К **элементарным преобразованиям** линейных систем относятся следующие преобразования:

1) первого типа: перемена мест двух или более уравнений исходной системы (1) (например, i и k уравнения системы (1) поменялись местами);

2) второго типа: замена одного уравнения исходной системы (1), например i-го уравнения $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = b_i$ на уравнение вида (c — некоторое число)

$$(a_{i1} + ca_{k1})x_1 + (a_{i2} + ca_{k2})x_2 + \ldots + (a_{in} + ca_{km})x_n = b_i + cb_k,$$

a_{kj}, b_k — коэффициенты и свободный член k-го уравнения). Или иначе: к i-му уравнению исходной системы (1) прибавляется k-е уравнение, умноженное на число c. К этому же типу относится и умножение какого-либо уравнения системы на не равное нулю число.

Теорема (признак эквивалентности систем). Две линейные системы эквивалентны, если одна получается из другой путем применения конечной последовательности элементарных преобразований.

Доказательство данной теоремы основано на том, что элементарные преобразования обратимы. Для 1-го типа: если одна система получилась из другой переменой местами i-го и k-го уравнений, то, еще раз поменяв эти уравнения местами, вернемся к исходной системе; для второго типа: если к i-му уравнению исходной системы прибавили k-е уравнение, умноженное на число с, то, еще раз прибавив k-е уравнение, умноженное на ($-c$), вернемся к исходной системе.

Из данной теоремы следует, что несовместность исходной системы влечет несовместность системы, полученной из исходной посредством элементарных преобразований. К линейной системе можно несколько раз применять элементарные преобразования, в результате получится система, эквивалентная исходной. В результате элементарных преобразований линейной системы можно прийти к ситуации, когда все коэффициенты какого-либо уравнения равны нулю. Тогда если свободный член такого уравнения равен нулю, то это уравнение можно не рассматривать (исключить из рассматриваемой системы), так как ему будут удовлетворять любые значения неизвестных. Если свободный член не равен нулю, то полученная система и эквивалентная ей исходная система несовместны.

Линейную систему посредством элементарных преобразований можно привести к **ступенчатому (трапециевидному или квазитреугольному) виду** (это более простой вид):

$$\begin{aligned}
\tilde{a}_{11}x_1 + \tilde{a}_{12}x_2 + \ldots\ldots\ldots\ldots\ldots + \tilde{a}_{1n}x_n &= \tilde{b}_1;\\
\tilde{a}_{2k}x_k + \ldots\ldots\ldots\ldots\ldots + \tilde{a}_{2n}x_n &= \tilde{b}_2;\\
\tilde{a}_{3l}x_l + \ldots\ldots\ldots\ldots + \tilde{a}_{3n}x_n &= \tilde{b}_3;\\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\tilde{a}_{rs}x_s + \ldots\ldots + \tilde{a}_{rn}x_n &= \tilde{b}_r;\\
0 &= \tilde{b}_{r+1};\\
\ldots\ldots\\
0 &= \tilde{b}_m.
\end{aligned}$$

Здесь $1 < k < l < \ldots < s$, $\tilde{a}_{11}, \tilde{a}_{2k}, \tilde{a}_{3l}, \ldots, \tilde{a}_{rs} \neq 0$. Возможно, что $r = m$, тогда уравнений типа $0 = \tilde{b}_j$ нет. Для того чтобы привести линейную систему (1) к ступенчатому виду, надо на первое место исходной системы (1) поставить уравнение, у которого коэффициент при неизвестном x_1 был бы не равен нулю, если первое уравнение системы (1) не удовлетворяет такому требованию. Из всех уравнений полученной системы, начиная со второго (i = 2, 3, ..., m), затем надо вычесть первое уравнение, умноженное на $c = a_{i1}^* / a_{11}^*$ (a_{i1}^*, a_{11}^* — коэффициенты системы, полученной после перестановки уравнений в случае необходимости местами). Таким образом, были выполнены элементарные преобразования первого и второго типов (действия аналогичны способу сложения или вычитания для линейной системы двух уравнений). Затем, не обращая внимания на первое уравнение, с остальными уравнениями вновь полученной системы повторяем те же действия. И так далее до тех пор, пока очередное преобразование не приведет к равенству типа $0=\tilde{b}_{r+1}$.

Метод приведения линейной системы к ступенчатому виду и дальнейшего решения ее называется **методом Гаусса** или **методом последовательного исключения неизвестных**. Данный метод удобен при небольших n и при реализации на ЭВМ.

Теорема. Всякая система линейных уравнений эквивалентна системе, имеющий ступенчатый вид (следует из способа приведения к ступенчатому виду).

Пример. Привести систему

$$\begin{cases} 2x_2+x_3-2x_4=3, \\ 4x_1+x_2-2x_3+x_4=1, \\ 2x_1-x_3+2x_4=0, \\ 5x_1+2x_2+3x_3-x_4=2 \end{cases}$$

к ступенчатому виду. Это система четырех линейных уравнений с четырьмя неизвестными, она приводится к частному случаю ступенчатого вида — треугольному. Поскольку в первом уравнении нет слагаемого с x_1, или, иначе, $a_{11}=0$, то поменяем местами первое и второе уравнения и получим:

$$\begin{cases} 4x_1+x_2-2x_3+x_4=1, \\ 2x_2+x_3-2x_4=3, \\ 2x_1-x_3+2x_4=0, \\ 5x_1+2x_2+3x_3-x_4=2. \end{cases}$$

Как видно, первые два уравнения соответствуют ступенчатому виду. Далее к третьему уравнению полученной системы прибавим первое, умноженное на число $-{}^2/_4=-{}^1/_2$, а к четвертому прибавим первое, умноженное на число $-{}^5/_4$, второе уравнение оставляем без изменений, так как в него не входит слагаемое с x_1.

В результате получаем: $$\begin{cases} 4x_1+x_2-2x_3+x_4=1, \\ 2x_2+x_3-2x_4=3, \\ -\frac{1}{2}x_2+\frac{3}{2}x_4=-\frac{1}{2}, \\ \frac{3}{4}x_2+\frac{11}{2}x_3-\frac{9}{4}x_4=\frac{3}{4}. \end{cases}$$

Теперь к третьему уравнению полученной системы прибавляем второе, умноженное на число $(-{}^1/_2):2={}^1/_4$, а к четвертому прибавляем второе, умноженное на число $-(^3/_4) : 2 = -^3/_8$. В результате получается:

$$\begin{cases} 4x_1 + x_2 - 2x_3 + x_4 = 1, \\ 2x_2 + x_3 - 2x_4 = 3, \\ \dfrac{1}{4}x_3 + x_4 = \dfrac{1}{4}, \\ \dfrac{41}{8}x_3 - \dfrac{3}{2}x_4 = -\dfrac{3}{8}. \end{cases}$$

И наконец, к четвертому уравнению полученной системы прибавляем третье, умноженное на число $-(^{41}/_8)$: $(^1/_4) = -^{41}/_2$. Получается требуемое:

$$\begin{cases} 4x_1 + x_2 - 2x_3 + x_4 = 1, \\ 2x_2 + x_3 - 2x_4 = 3, \\ \dfrac{1}{4}x_3 + x_4 = \dfrac{1}{4}, \\ -22x_4 = -\dfrac{11}{2}. \end{cases}$$

Таким образом, исходная система была приведена к ступенчатому (точнее, треугольному) виду путем последовательного исключения неизвестных. Решая полученную систему снизу вверх последовательно, находим: $x_4 = {}^1/_4$, $x_3 = 0$, $x_2 = {}^7/_4$, $x_1 = -{}^1/_4$.

4. Совместность и определенность линейных систем

Напомним, что если линейная система имеет решения, то она совместна, если решение единственное — определена. Теоремы о совместности и определенности линейных систем имеют дело с системами ступенчатого вида, поскольку любую линейную систему можно привести к ступенчатому виду. Пусть имеется некоторая линейная система, приведенная к следующему ступенчатому виду:

$$\begin{gathered} a_{11}x_1 + a_{12}x_2 + \ldots\ldots\ldots\ldots\ldots\ldots + a_{1n}x_n = b_1 \\ a_{2k}x_k + \ldots\ldots\ldots\ldots\ldots\ldots\ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{rs}x_s + \ldots\ldots + a_{rn}x_n = b_r \\ 0 = b_{r+1} \\ \ldots\ldots \\ 0 = b_m \end{gathered}$$

$(1 < k < l < \ldots < s)$. Неизвестные x_1, x_2, ..., x_s называются **главными**, r первых уравнений системы ступенчатого вида начинаются со слагаемых, в состав которых входят эти неизвестные.

Если имеются уравнения типа $0 = b_j$, то неизвестные, начиная с $xs+1$, называются **свободными**. Главных неизвестных *r* штук.

Главные неизвестные определяются однозначно для любых заданных свободных неизвестных, иначе говоря, задав некоторым образом свободные неизвестные, по уравнениям ступенчатой системы, начиная с конца системы (с *r*-уравнения), можно определить все главные неизвестные.

Теорема. Для совместности системы линейных уравнений необходимо и достаточно, чтобы после приведения к ступенчатому виду в ней не оказалось уравнений вида $0 = b_j$ при $b_j \neq 0$.

Если это условие выполнено, то свободным неизвестным можно придать произвольные значения.

Главные значения при заданных свободных однозначно определяются из системы (доказательство теоремы следует из самого способа определения неизвестных, указанного выше).

Теорема. Совместная линейная система является определенной тогда и только тогда, когда в полученной из нее ступенчатой системе выполняется равенство *r* = *n*.

Доказательство следует из самого вида получающейся в этом случае ступенчатой системы. При *r* = *n* = *m* линейная система приводится к **треугольному** виду (количество уравнений равно количеству неизвестных):

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots & \\ a_{nn}x_n &= b_n \end{aligned}$$

(это частный случай систем ступенчатого вида). Если $a_{ij} = 0$ при $i > j$, то система называется **системой верхнего треугольного вида** (возможен и нижний треугольный вид системы).

Следствия.

1. Линейная система в случае *m* = *n* является совместной и определенной тогда и только тогда, когда после приведения к треугольному виду получается система с коэффициентами $a_{11}, a_{22}, \ldots, a_{nn} \neq 0$.

Данное условие означает, что ассоциированная с исходной линейной системой однородная система имеет только нулевое решение.

2. Совместная система при $n > m$ является неопределенной. Однородная система при $n > m$ всегда имеет ненулевое решение (неопределенность однородной системы означает наличие нескольких решений, а значит, и наличие ненулевого решения).

ЛЕКЦИЯ № 2. Матрицы

1. Матрицы. Основные определения

Матрицей размерности *m* × *n* называется прямоугольная таблица вида

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

или кратко $A = (a_{ij})$, где индексы элементов $i = 1, 2, ..., m$; $j = 1, 2, ...,$ т. е. матрица — это таблица, состоящая из m строк и n столбцов. Числа a_{ij} называются **элементами матрицы**. Матрица может быть обозначена следующим образом:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \ddots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

или

$$A = \begin{Vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \ddots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{Vmatrix}$$

(кратко $A = \left[a_{ij}\right]$ или $A = \left\|a_{ij}\right\|$). Если $m = 1$, а $n > 1$, то матрица состоит из одной строки, т. е. является **матрицей-строкой**, или **строчной матрицей**: $A = \left[a_{11}, a_{12}, a_{1n}\right]$. Если $m > 1$, а $n = 1$, то матрица состоит из одного столбца, т. е. является **матрицей-столбцом**, или **столбцевой матрицей**:

$$A = \begin{bmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{n1} \end{bmatrix}.$$

Строки и столбцы матрицы называются **рядами матрицы**. Матрица A размерности $m \times n$ может быть обозначена кратко: A_{mn} или $A_{m \times n}$.

В случае, когда $m = n$ (число строк равно числу столбцов), матрица называется **квадратной**, а число ее строк (или столбцов) называется **порядком матрицы**. Квадратная матрица первого порядка состоит из одного элемента a_{11}. Квадратная матрица имеет две диагонали: **главная диагональ** образована элементами $a_{11}, a_{22}, \ldots, a_{nn}$; вторая диагональ образована элементами $a_{1n}, a_{2n-1}, \ldots, a_{n1}$.

Симметрической называется квадратная матрица, элементы которой, симметричные относительно главной диагонали, равны, например,

$$\begin{pmatrix} 2 & 4 \\ 4 & 5 \end{pmatrix} \text{ или } \begin{pmatrix} a_{11} & b & c \\ b & a_{22} & d \\ c & d & a_{33} \end{pmatrix}.$$

Если все элементы квадратной матрицы, кроме элементов главной диагонали, равны нулю, то такая матрица называется **диагональной**, она имеет следующий вид:

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & 0 & a_{nn} \end{pmatrix}.$$

Единичной матрицей называется квадратная диагональная матрица, у которой элементы главной диагонали, равны единице, а остальные элементы равны нулю, она обозначается:

$$E = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Единичную матрицу можно записать с помощью **символа Кронекера** $\delta_{ij} = \begin{cases} 1, \text{ при } i \neq j \\ 0, \text{ при } i = j \end{cases}$ следующим образом: $E = (\delta_{ij})$.

Если все элементы матрицы равны нулю, то такая матрица называется **нулевой**, она обозначается следующим образом:

$$0 = \begin{pmatrix} 0 & \cdots & 0 \\ \cdots & \ddots & \cdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

Если все элементы квадратной матрицы, расположенные выше (или ниже) главной диагонали, то такая квадратная матрица называется **нижней (или верхней) треугольной матрицей**, соответственно она имеет вид:

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \text{ или } \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Если все элементы произвольной матрицы размерности $m \times n$, расположенные ниже линии из элементов $a_{11}, a_{22}, ..., a_{rr}$, равны нулю, то такая матрица называется **ступенчатой**, или **трапециевидной**, или **квазитреугольной**, она имеет вид:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1r} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2r} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{rr} & \cdots & a_{rn} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}.$$

Две матрицы A и B называются **равными**, если их размер одинаков и их соответствующие элементы равны $a_{ij} = b_{ij}$.

Матрицей, **транспонированной** к матрице A размерности $m \times n$, называется матрица A^{T} размерности $n \times m$, полученная из матрицы A, если строки поменять местами со столбцами. Например, если имеется матрица

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix},$$

то транспонированная матрица имеет вид:

$$A^T = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}.$$

Операция нахождения транспонированной матрицы называется **транспонированием**.

2. Линейные действия над матрицами

Над матрицами одинакового размера можно производить следующие линейные действия: сложение, вычитание, умножение на число. Для матриц различного размера сложение и вычитание не определено.

Суммой двух однотипных матриц A и B называется матрица C, элементы которой равны сумме соответствующих элементов матриц A и B, записывается следующим образом: $C = A + B$ (где $c_{ij} = a_{ij} + b_{ij}$). Если складывается более двух матриц, то их сумма получается в результате последовательного сложения этих матриц.

При сложении с нулевой матрицей исходная матрица не меняется: $A + O = A$.

Например, пусть матрица

$$A = \begin{pmatrix} 1 & 2 & 4 & 5 & 7 \\ 3 & 4 & 5 & 8 & 9 \end{pmatrix},$$

матрица

$$B = \begin{pmatrix} 7 & 6 & 4 & 3 & 2 \\ 8 & 2 & 4 & 3 & 1 \end{pmatrix}.$$

Тогда суммой матриц A и B будет матрица, равная

$$C = A + B = \begin{pmatrix} 8 & 8 & 8 & 8 & 9 \\ 11 & 6 & 9 & 11 & 10 \end{pmatrix}.$$

Разностью двух однотипных матриц A и B называется матрица C, элементы которой равны разности соответствующих элементов матриц A и B: $C = A - B$ (где $c_{ij} = a_{ij} - b_{ij}$).

Например, разностью матриц A и B, приведенных выше, будет матрица, равная

$$C = A - B = \begin{pmatrix} -6 & -4 & 0 & 2 & 5 \\ -5 & 2 & 1 & 5 & 8 \end{pmatrix}.$$

Произведением матрицы A **на число** c называется матрица B, элементы которой равны произведению этого числа с на соответствующие элементы матрицы A, обозначается $B = cA$ (где $b_{ij} = ca_{ij}$).

Например, если матрица

$$A = \begin{pmatrix} 4 & 6 & 1 & 0 \\ 2 & 5 & 7 & 9 \end{pmatrix},$$

то при умножении на $c = 5$ получается матрица

$$B = \begin{pmatrix} 20 & 30 & 5 & 0 \\ 10 & 25 & 35 & 45 \end{pmatrix}.$$

При умножении матрицы на нулевую матрицу получается нулевая матрица. При умножении матрицы A на (-1) получается матрица $-A$, которая называется **противоположной** матрице A.

Для линейных операций справедливы следующие свойства:

1) $A + B = B + A$ (сочетательное свойство);
2) $(A + B) + C = A + (B + C)$ (переместительное свойство);
3) $A + 0 = A$;
4) $A + (-A) = 0$;
5) $1A = A$;
6) $\alpha(\beta A) = (\alpha\beta)A$ (ассоциативность относительно умножения на число);
7) $(A + B)\,\alpha = \alpha A + \alpha B$ (дистрибутивность относительно сложения);
8) $(\alpha + \beta)A = \alpha A + \beta A$ (дистрибутивность относительно сложения чисел).

3. Умножение матриц

Кроме линейных операций над матрицами, возможно еще и умножение матриц, но оно определено только для согласованных матриц.

Матрица A называется **согласованной с матрицей** B, если количество столбцов матрицы A равно количеству строк матрицы B. Однако если матрица A согласована с матрицей B, то матрица B может быть несогласованной с матрицей A.

Например, если матрица

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{pmatrix},$$

а матрица

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \\ b_{41} & b_{42} & b_{43} \end{pmatrix},$$

то матрица A будет согласована с матрицей B, так как у матрицы A четыре столбца, а у матрицы B четыре строки (одинаковое количество). Но матрица B не согласована с матрицей A, так как матрица B имеет три столбца, а матрица A — две строки (неравное количество).

Две квадратные матрицы одного порядка всегда согласованы, так как имеют одинаковое количество строк и столбцов.

Произведением двух согласованных матриц A и B называется матрица C, элемент которой, находящийся на пересечении i-й строки и k-го столбца, является суммой парных произведений элементов i-той строки первой матрицы A на соответствующий элемент k-го столбца второй матрицы B:

$$C = AB(c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \ldots + a_{in}b_{nk} = \sum_{j=1} a_{ij}b_{jk}).$$

При умножении двух ненулевых матриц может получиться нулевая матрица. Произведение матриц AB — матрица, у которой число строк равно числу строк матрицы A, а число столбцов — числу столбцов матрицы B.

Пример. Даны две матрицы: $A = \begin{pmatrix} 2 & 4 & 0 \\ 3 & 6 & 7 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 & 6 \\ 6 & 3 & 5 \\ 7 & 8 & 4 \end{pmatrix}$.

Матрица A имеет три столбца, а матрица B — три строки, следовательно, матрица A согласована с матрицей B и их произведение равно:

$$AB = \begin{pmatrix} 2\times1+4\times6+0\times7 & 2\times2+4\times3+0\times8 & 2\times6+4\times5+0\times4 \\ 3\times1+6\times6+7\times7 & 3\times2+6\times3+7\times8 & 3\times6+6\times5+7\times4 \end{pmatrix} =$$

$$= \begin{pmatrix} 26 & 16 & 32 \\ 88 & 80 & 76 \end{pmatrix}.$$

Но матрицу В нельзя умножить на матрицу A, так как B не согласована с A (B имеет три столбца, а матрица A имеет две строки), т. е. произведение BA не определено.

В общем случае произведение двух матриц не подчиняется переместительному закону, т. е. $AB \neq BA$; матрицы, для которых $AB = BA$, называются **коммутирующими (перестановочными, коммутативными)**.

Для операции умножения матриц справедливы следующие свойства (если соответствующие действия определены):

1) $EA = AE = A$;
2) $A(BC) = (AB)C$ (сочетательное или ассоциативное свойство);

3) $\alpha(AB) = (\alpha A)B = A(\alpha B)$;

4) $(A_1 + A_2)B = A_1B + A_2B$, $A(B_1 + B_2) = AB_1 + AB_2$ (свойство дистрибутивности);

5) $AO = OA = A$.

Для матриц возможна операция, аналогичная операции возведения в степень. Такая операция определена для квадратных матриц. **Целой положительной степенью** A^k ($k > 1$) квадратной матрицы A называется произведение $A \times A \times ... \times A$ (k раз). Очевидно, что порядок матрицы A^k равен порядку матрицы A. Первая степень матрицы равна самой матрице $A^1 = A$, нулевая степень матрицы равна единичной матрице равного порядка $A^0 = E$.

4. Многочлены от матриц, линейные комбинации матриц, комплексные матрицы

Из матриц одинаковой размерности (однотипных, одного порядка) можно составлять линейные комбинации. **Линейной комбинацией** матриц одинаковой размерности $A_1, A_2, ..., A_k$ с коэффициентами $\lambda_1, \lambda_2, ..., \lambda_k$ называется новая матрица той же размерности вида:

$$A = A_1\lambda_1 + A_2\lambda_2 + ... + A_k\lambda_k.$$

Матрицы $A_1, A_2, ..., A_k$ называются **линейнозависимыми**, если найдутся такие одновременно не равные нулю числа $\lambda_1, \lambda_2, ..., \lambda_k$, что линейная комбинация этих матриц с этими числами равна нулю: $A_1\lambda_1 + A_2\lambda_2 + ... + A_k\lambda_k = 0$.

Если данное равенство имеет место, когда все числа $\lambda_1, \lambda_2, ..., \lambda_k$ одновременно равны нулю, то матрицы $A_1, A_2, ..., A_k$ называются **линейно независимыми**.

Теорема. Матрицы $A_1, A_2, ..., A_k$ линейнозависимы тогда и только тогда, когда, по крайней мере, одна из них является линейной комбинацией остальных.

Следствие. Любая подсистема линейно независимой системы матриц также линейно независима.

Для квадратичных матриц возможно составление многочленов. **Многочленом** степени n от квадратичной матрицы A называется следующее выражение:

$$a_nA^n + a_{n-1}A^{n-1} + ... + a_1A^1 + a_0A^0 \ (a_n \neq 0)$$

или $$P(A) = a_nA^n + a_{n-1}A^{n-1} + ... + a_1A^1 + a_0\text{E}.$$

Если $P(A) = 0$, то матрица А называется **корнем многочлена** P, а сам многочлен называется **аннулирующим** для матрицы A.

Кроме действительных чисел, часто приходится иметь дело с комплексными числами, которые также могут быть элементами матриц. Комплексным числом называется упорядоченная пара чисел (x, y). Оно может быть записано в виде $z = x + iy$, где x, y — действительные числа, называемые соответственно **действительной** $x = \mathrm{Re}\, z$ и **мнимой частью** $y = \mathrm{Im}\, z$ комплексного числа z, ai — **мнимая единица**, определенная следующим образом $i^2 = -1$.

Введение комплексных чисел позволяет решить уравнения вида $z^2 + a^2 = 0$ (а $\neq 0$), решение имеет вид: $z = \pm\, ia$.

Сопряженными комплексными числами называются два числа вида $x + iy$ и $x - iy$, обозначается $\overline{x+iy} = x - iy$ или $\overline{x-iy} = x+iy$.

Комплексной матрицей называется матрица Z, все элементы которой — комплексные числа: $z_{kj} = x_{kj} + iy_{kj}$. Если обозначить символом X матрицу с элементами x_{kj}, а символом Y — матрицу с элементами y_{kj}, то комплексную матрицу можно записать в виде $Z = X + iY$. Для этой комплексной матрицы определена комплексно-сопряженная матрица $\overline{Z}$, элементы которой равны $\overline{z}_{kj} = x_{kj} - iy_{kj}$. Для комплексно-сопряженных матриц справедливы следующие свойства (если соответствующие действия определены): $\overline{Z_1 + Z_2} = \overline{Z}_1 + \overline{Z}_2$, $\overline{Z_1 Z_2} = \overline{Z}_1 \overline{Z}_2$.

5. Свойства транспонированных матриц

Пусть дана матрица A размером $m \times n$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Если каждую строку матрицы A заменить столбцом с таким же номером, то получим матрицу A^T размером $n \times m$, транспонированную относительно данной матрицы A:

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

Для транспонированных матриц справедливы следующие свойства:

1) $(A^T)^T = A$;

2) $(cA)^T = cA^T$ (c — постоянная);
3) $(A + B)^T = A^T + B^T$;
4) $(AB)^T = B^T A^T$ (в случае определенности данных операций).

Для доказательства этих свойств требуется убедиться, что матрицы, стоящие в разных частях равенств, имеют одинаковый размер и равные соответствующие (стоящие на одинаковых местах) элементы. Доказательство первых двух свойств следует из определения операции транспонирования, третьего — из определения сложения матриц.

Докажем последнее свойство. Пусть матрица A имеет размер $m \times n$, матрица B — размер $n \times r$. Тогда произведение (считается, что оно определено) AB имеет размер $m \times r$, а матрица, транспонированная относительно этого произведения $(AB)^T$, имеет размер $r \times m$. Матрица A^T имеет размер $\mathrm{n} \times \mathrm{m}$, матрица B^T — размер $r \times n$; тогда их произведение $B^T A^T$ имеет размер $r \times m$, т. е. размеры матриц $(AB)^T$ и $B^T A^T$ совпадают. Докажем равенство соответствующих элементов. Определим, какие элементы стоят на месте ik (i — строка, k — столбец) в матрицах $(AB)^T$ и $B^T A^T$. На этом месте в матрице A стоит элемент a_{ik}, в матрице B — элемент b_{ik}.

Тогда в матрице AB на этом месте стоит элемент, равный $\sum_{j=1}^{n} a_{ij} b_{jk}$,

в матрице $(AB)^{\mathrm{т}}$ — элемент, равный $\sum_{j=1}^{n} a_{kj} b_{ji}$. В матрице B^T на этом

месте стоит элемент b_{ki}, в матрице A^T — элемент a_{ki}, тогда в мат-

рице $B^T A^T$ — элемент, равный $\sum_{j=1}^{n} b_{ji} a_{kj}$ (следует из правила умно-

жения матриц).Но $\sum_{j=1}^{n} a_{kj} b_{ji} = \sum_{j=1}^{n} b_{ji} a_{kj}$. Свойство доказано.

Очень легко эти свойства проверяются в цифрах.

6. Определители второго и третьего порядка. Их свойства

Определитель (детерминант) первого порядка, соответствующий матрице $A = a_{11}$, равен числу $|A| = a_{11}$. **Определителем (детерми-**

нантом) второго порядка, соответствующим матрице $A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, называется число, равное $|A|=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}$.

Определитель (детерминант) третьего порядка, соответствующий матрице

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

равен числу

$$|A|=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}=$$

$$=a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}$$

(графический способ вычисления определителя третьего порядка приведен в лекции 1).

Элементы a_{11}, a_{22} (или a_{11}, a_{22}, a_{33}) образуют главную диагональ, а элементы a_{12}, a_{21} (или a_{13}, a_{22}, a_{31}) — побочную диагональ. Определитель $|A|$ может быть обозначен **det *A*** или знаком $\boldsymbol{\Delta}$.

Теорема (свойства определителей второго и третьего порядка).

1. Величина определителя не меняется, если заменить его строки соответствующими столбцами, т. е. при транспонировании.

2. Определитель поменяет знак при перестановке двух строк или двух столбцов.

3. Определитель будет равен нулю, если элементы какого-либо столбца (или строки) равны нулю или элементы двух строк (или столбцов) соответственно равны.

4. Множитель, общий для элементов некоторого столбца (или строки), можно выносить за знак определителя.

5. Величина определителя не меняется, если к элементам какой-либо его строки или столбца прибавить соответствующие элементы другой строки или столбца, умноженные на одно и то же число.

Доказательство приводится для более общего случая — определителя третьего порядка (для определителей второго порядка проверяется непосредственно).

Пусть имеется определитель

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

1. Поменяем строки и столбцы местами и вычислим определитель:

$$\begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} =$$

$$= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23}$$

Сравнивая данное выражение с исходным определителем, видим, что они равны.

2. Поменяем местами в исходном определителе, например, первую и вторую строку и вычислим полученный определитель:

$$\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$= a_{21}a_{12}a_{33} + a_{22}a_{13}a_{31} + a_{23}a_{11}a_{32} - a_{23}a_{12}a_{31} - a_{22}a_{11}a_{33} - a_{21}a_{13}a_{32} =$$
$$= -(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}).$$

3. Если элементы, например, первой строки исходного определителя равны нулю, то и сам определитель равен нулю, так как в составе каждого слагаемого суммы

$$(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33})$$

в качестве множителя есть элемент первой строки определителя (аналогично для других строк или столбцов).

Пусть равны элементы первой и второй строки исходного определителя, вычислим его:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$= a_{11}a_{12}a_{33} + a_{12}a_{13}a_{31} + a_{13}a_{11}a_{32} - a_{13}a_{12}a_{31} - a_{11}a_{13}a_{32} - a_{12}a_{11}a_{33} = 0.$$

4. Пусть имеется определитель

$$\begin{vmatrix} ca_{11} & ca_{12} & ca_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

Он равен сумме:

$$ca_{11}a_{22}a_{33} + ca_{12}a_{23}a_{31} + ca_{13}a_{21}a_{32} - ca_{13}a_{22}a_{31} - ca_{11}a_{23}a_{32} - ca_{12}a_{21}a_{33} =$$
$$= c(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}) =$$
$$= c\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

5. Пусть к первой строке исходного определителя прибавлена вторая строка, умноженная на число c, тогда полученный определитель можно представить в виде суммы двух определителей:

$$\begin{vmatrix} a_{11} + ca_{21} & a_{12} + ca_{22} & a_{13} + ca_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$
$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} ca_{21} & ca_{22} & ca_{23} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

так как второе слагаемое-определитель равен нулю по третьему свойству. Теорема доказана.

Следствие. Определитель с двумя пропорциональными столбцами (или строками) равен нулю.

7. Минор. Алгебраическое дополнение

Минором какого-либо элемента исходного определителя $|A|$ называется определитель, полученный из $|A|$ вычеркиванием той строки и того столбца, которым принадлежит этот элемент. Минор элемента a_{ik} обозначается M_{ik}. Для определителя третьего порядка минором является определитель второго порядка. Например, минор элемента a_{11} определителя

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \text{ равен } M_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}.$$

Алгебраическим дополнением A_{ik} элемента a_{ik} определителя $|A|$ называется его минор, взятый со знаком $(-1)^{i+k}$, обозначается $A_{ik} = (-1)^{i+k}M_{ik}$.

Теорема о разложении определителя по элементам строки или столбца. Определитель равен сумме произведений от умножения элементов любой строки (столбца) на их алгебраические дополнения.

Доказательство. Проверим данное утверждение для элементов, например, первого столбца определителя третьего порядка $|A|$:

$$|A| = (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}) =$$
$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{21}(a_{13}a_{32} - a_{12}a_{33}) + a_{31}(a_{12}a_{23} - a_{13}a_{22}) =$$
$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{21}\begin{vmatrix} a_{13} & a_{12} \\ a_{33} & a_{32} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = a_{11}A_{11} + a_{21}A_{21} + a_{31}A_{31},$$

здесь

$$A_{21} = -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{13} & a_{12} \\ a_{33} & a_{32} \end{vmatrix}.$$

Аналогичным образом можно разложить определитель по элементам любой строки или столбца. Теорема доказана.

Теорема (теорема замещения). Пусть $|A|$ — некоторый определитель третьего порядка. Сумма произведений от умножения алгебраических дополнений элементов какой-либо строки (или столбца) на любые числа c_1, c_2, c_3 равна определителю, который получается из исходного определителя $|A|$ заменой этой строки (столбца) строкой (или столбцом) из чисел c_1, c_2, c_3.

Доказательство. Заменим в исходном определителе |A| элементы первой строки соответственно на числа c_1, c_2, c_3. Вычислим, чему равен полученный определитель. Его (по предыдущей теореме) можно разложить по элементам первой строки, являющимся числами c_1, c_2, c_3, следующим образом:

$$\begin{vmatrix} c_1 & c_2 & c_3 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = c_1A_1 + c_2A_2 + c_3A_3.$$

Таким образом, теорема доказана.

Теорема аннулирования. Сумма произведений от умножения элементов какой-нибудь строки (или столбца) определителя на алгебраические дополнения соответствующих элементов другой строки (или столбца) этого определителя равна нулю.

Доказательство. Для некоторого определителя $|A|$ сумма произведений в результате умножения элементов, например, первой строки на алгебраические дополнения соответствующих элементов, например, второй строки равна $a_{11}A_{21}+a_{12}A_{22}+a_{13}A_{23}$.

По предыдущей теореме данная сумма равна определителю, полученному из определителя $|A|$ при замене элементов второй строки соответственно на

$$a_{11}, a_{12}, a_{13}: \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix},$$

а данный определитель равен нулю, так как имеет две одинаковые строки. Теорема доказана.

Применяя теорему о разложении определителя по элементам строки (столбца) и используя свойства определителя, можно вычислить определитель третьего порядка не только непосредственно по формуле, но и следующими способами. Пусть имеется определитель

$$|A| = \begin{vmatrix} 1 & 2 & 4 \\ 4 & 7 & 5 \\ 3 & 9 & 6 \end{vmatrix}.$$

1. Разложить определитель по элементам, например, третьей строки. При разложении получается:

$$|A| = 3\begin{vmatrix} 2 & 4 \\ 7 & 5 \end{vmatrix} - 9\begin{vmatrix} 1 & 4 \\ 4 & 5 \end{vmatrix} + 6\begin{vmatrix} 1 & 2 \\ 4 & 7 \end{vmatrix} = 3(-18) - 9(-11) + 6(-1) = 39.$$

2. Используя свойства определителя, надо привести его к такому виду, чтобы в некоторой строке (или столбце) два элемента были бы равны нулю, и затем разложить его по элементам этой строки (или столбца). Умножим первую строку определителя $|A|$ на (-4) и прибавим ее ко второй строке, а к третьей строке прибавим первую, умноженную на (-3).

Тогда исходный определитель принимает вид:

$$\begin{vmatrix} 1 & 2 & 4 \\ 0 & -1 & -11 \\ 0 & 3 & -6 \end{vmatrix},$$

а после разложения по элементам первого столбца он равен:

$$1\times\begin{vmatrix} -1 & -11 \\ 3 & -6 \end{vmatrix} = 6+33=39.$$

8. Определитель n-го порядка

Пусть имеется квадратная матрица n-го порядка

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Понятие минора и алгебраического дополнения для нее вводится аналогично случаю матрицы третьего порядка.

Минор элемента a_{ik} квадратной матрицы n-го порядка — это определитель порядка (n – 1), соответствующий матрице, полученной из исходной вычеркиванием i-той строки и k-го столбца, обозначается M_{ik}.

Алгебраическим дополнением A_{ik} элемента a_{ik} называется минор этого элемента, взятый со знаком $(-1)^{i+k}$, $A_{ik} = (-1)^{i+k}M_{ik}$.

С учетом этих понятий определение определителя n-го порядка выводится следующим образом.

Определителем n-го порядка (обозначается $|A|$, или Δ, или $\det A$), соответствующим квадратной матрице A n-го порядка, называется число, равное сумме парных произведений в результате умножения элементов какой-либо строки (столбца) на их алгебраические дополнения. Например, для элементов первой строки имеем:

$$\Delta = \det A = |A| = \sum_{k=1}^{n}(-1)^{1+k}a_{1k}M_{1k} = \sum_{k=1}^{n}a_{1k}A_{1k}.$$

Аналогично определителям второго и третьего порядка определители n-го порядка могут быть разложены по любой строке или столбцу, это утверждение основано на определении определителя n-го порядка. При разложении по какой-либо i-той строке справедливо выражение: $\Delta = \sum_{k=1}^{n} a_{ik} A_{ik}$, при разложении по элементам k-го столбца имеем: $\Delta = \sum_{i=1}^{n} a_{ik} A_{ik}$.

Таким образом, учитывая данное определение, вычисление определителя n-го порядка сводится к вычислению определителей $(n - 1)$-го порядка.

Для определителей n-го порядка справедливы те же свойства, что и для определителей второго и третьего порядка. Теоремы замещения и аннулирования для определителя n-го порядка имеют следующую формулировку (доказательства аналогичны случаю определителя третьего порядка).

Теорема (замещения). Сумма произведений от умножения произвольных чисел $c_1, c_2, ..., c_n$ соответственно на алгебраические дополнения элементов некоторой строки (или столбца) матрицы n-го порядка равна определителю матрицы, которая получается из исходной заменой элементов этой строки (или столбца) числами $c_1, c_2, ..., c_n$. Например, для k-го столбца имеем $\Delta = c_1A_{1k} + c_2A_{2k} + ... + c_nA_{nk}$.

Теорема (аннулирования). Сумма произведений элементов одной из строк (или столбцов) квадратной матрицы n-го порядка на соответствующие алгебраические дополнения элементов другой строки (или столбца) равна нулю. Например, если взять k-тую и j-тую строки определителя Δ, причем k ≠ j, то имеем:

$$a_{k1}A_{j1} + a_{k2}A_{j2} + \ldots + a_{kn}A_{jn} = 0\ (k \neq j, 1 \leq k, j \leq n).$$

Теорема об определителе произведения. Если A и B — квадратные матрицы одного порядка с определителями $|A|$ и $|B|$, то определитель матрицы $C = AB$ равен произведению в результате умножения определителей перемножаемых матриц: $|C| = |A||B|$ (или $\det(AB) = \det A \det B$) (доказывается путем непосредственного вычисления определителей).

Вычисление определителей больших порядков — довольно громоздкая операция. Но чем больше нулей среди элементов определителя, тем легче его вычислять. Хорошо известны следующие частные случаи:

1) определитель треугольной матрицы, у которой все элементы выше (или ниже) главной диагонали равны нулю, равен произведению в результате умножения элементов, стоящих на главной диагонали.

2) для определителя порядка $n + m$, имеющего на пересечении первых n столбцов и последних m строк нули (имеющего угол нулей), справедлива формула:

$$\begin{vmatrix} a_{11}\ldots\ldots a_{1n} & a_{1,n+1} & \cdots & a_{1,n+m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}\ldots\ldots a_{nn} & a_{n,n+1} & \cdots & a_{n,n+m} \\ 0\ldots\ldots\ldots 0 & b_{11} & \cdots & b_{1m} \\ \cdots & \cdots & \cdots & \cdots \\ 0\ldots\ldots\ldots 0 & b_{m1} & \cdots & b_{mm} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \begin{vmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mm} \end{vmatrix}.$$

9. Обратная матрица. Присоединенная матрица. Вырожденная матрица

Обратной матрицей для квадратной матрицы A называется такая матрица $A - 1$, которая удовлетворяет условиям $AA - 1 = A - - 1A = E$ (здесь E — единичная матрица). Если матрица $A - 1$ обратная для A, то A будет обратной для A–1. Матрица, обратная для квадратной матрицы A, тоже квадратная и имеет тот же порядок, что и матрица A.

Присоединенной матрицей для квадратной матрицы A называется такая матрица C (тоже квадратная), элементами которой являются алгебраические дополнения элементов матрицы A, причем алгебраические дополнения j-той строки матрицы A расположены в j-том столбце матрицы C, т. е. для матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

присоединенная матрица имеет вид:

$$C = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}.$$

Теорема. Если A — квадратная матрица порядка n, а матрица C присоединена к ней, то их произведение равно $AC = CA = E\,\det A$, где E — единичная матрица того же порядка.

Доказательство. По определению операции умножения матриц произведением матриц AC является матрица P, каждый элемент p_{ik} которой равен сумме произведений от умножения элементов i-той строки матрицы A на элементы k-того столбца матрицы C. Все элементы p_{ii}, находящиеся на главной диагонали матрицы P (по теореме о разложении определителя), равны $\det A$, а остальные элементы матрицы P (по теореме аннулирования) равны нулю. Таким образом, произведение AC может быть представлено в виде:

$$P = AC = \begin{pmatrix} \det A & & 0 \\ & \ddots & \\ 0 & & \det A \end{pmatrix} = \det A \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = E \det A.$$

Для произведения CA доказательство аналогично. Теорема доказана.

Матрица A называется **вырожденной (или особенной)**, если ее определитель $|A|$ равен нулю, в противном случае матрица A называется **невырожденной (неособенной)**. Вырожденная матрица не имеет обратной матрицы.

Теорема. Матрица

$$\begin{pmatrix} A_{11}/|A| & A_{21}/|A| & \dots & A_{n1}/|A| \\ A_{12}/|A| & A_{22}/|A| & \dots & A_{n2}/|A| \\ \dots & \dots & \dots & \dots \\ A_{1n}/|A| & A_{2n}/|A| & \dots & A_{nn}/|A| \end{pmatrix},$$

где A_{ik} — алгебраическое дополнение элемента a_{ik} невырожденной матрицы A, является обратной для A.

Другая формулировка этой теоремы. Для невырожденной матрицы A существует единственная обратная матрица A^{-1}, определяемая формулой $A^{-1} = C/\det A$, где матрица C присоединенная к матрице A.

Доказательство. Найдем произведение матриц A и $C / \det A$ ($\det A \neq 0$ по определению невырожденности). Оно равно (по предыдущей теореме) $A(C / \det A) = E$. Следовательно, матрица $C / \det A$ является обратной матрице A, $C / \det A = A^{-1}$. Если бы такая матрица не была единственной, т. е. существовала бы еще одна матрица $A_*^{-1} = C / \det A$, то из равенств $A_*^{-1}AA^{-1} = (A_*^{-1}A)A^{-1} = EA^{-1} = A^{-1}$ и $A_*^{-1}AA^{-1} = A_*^{-1}(AA^{-1}) = A_*^{-1}E = A_*^{-1}$ следует, что $A^{-1} = A_*^{-1}$. Теорема доказана.

Данная теорема указывает способ нахождения обратных матриц, т. е. чтобы найти матрицу, обратную матрице A, надо вычислить определитель det$A$, затем для каждого элемента матрицы А вычислить алгебраическое дополнение и из них составить присоединенную для матрицы A матрицу C. Тогда элементы обратной матрицы будут равны соответствующим элементам присоединенной матрицы C, деленным на детерминант матрицы A (определитель матрицы A).

Справедливо следующее утверждение (как следствие из последней теоремы): определитель равен нулю тогда и только тогда, когда его строки (или столбцы) линейнозависимы.

Невырожденные матрицы обладают следующими свойствами:

1) $\det A^{-1} = 1 / \det A$;
2) $(A^{-1})^{-1} = A$;
3) $(A^m)^{-1} = (A^{-1})^m$;
4) $(AB)-1 = \mathrm{B}^{-1}A^{-1}$.

10. Ранг матрицы

Пусть имеется некоторая матрица A размером $m \times n$ с элементами a^{ik}. Пусть число k удовлетворяет неравенству $1 < k \le \min(m, n)$. Тогда **минором порядка *k* матрицы *A*** называется определитель матрицы, состоящей из элементов матрицы A, стоящих на пересечении произвольно выбранных k строк и k столбцов.

Одна и та же матрица может иметь несколько миноров одного порядка, миноры могут быть равны нулю.

Например, для матрицы

$$\begin{pmatrix} 2 & -3 & 4 & 6 \\ 3 & 5 & -8 & 8 \\ 1 & 4 & 6 & 9 \end{pmatrix}$$

минором второго порядка будут определители

$$\begin{vmatrix} 2 & -3 \\ 3 & 5 \end{vmatrix}$$

(взяты первая и вторая строки и первый и второй столбцы),

$$\begin{vmatrix} -3 & 6 \\ 5 & 8 \end{vmatrix}$$

(взяты первая и вторая строки и второй и четвертый столбец). Один из миноров третьего порядка для данной матрицы имеет вид:

$$\begin{vmatrix} 2 & 4 & 6 \\ 3 & -8 & 8 \\ 1 & 6 & 9 \end{vmatrix}$$

(взяты все строки и первый, третий, четвертый столбцы). Возможны и другие варианты миноров второго и третьего порядка данной матрицы.

Базисный минор матрицы A — это любой не равный нулю минор максимального порядка матрицы A. Столбцы и строки матрицы А, пересекающие базисный минор, называются базисными.

Ранг матрицы *r* — это максимальный порядок ее миноров, отличных от нуля. В случае, когда все миноры матрицы равны нулю, ее ранг считается равным нулю. Из данного определения вытекает, что:

1) ранг матрицы размером $m \times n$ — целое число, причем $0 \leq r \leq \min(m, n)$;
2) если матрицы квадратная, то ее ранг равен ее порядку только в случае, если эта матрица невырожденная;
3) ранг матрицы равен нулю только в случае нулевой матрицы;
4) ранг транспонированной матрицы равен рангу исходной матрицы;
5) ранг матрицы не меняется при вычеркивании или добавлении к ней ряда, в котором все элементы равны нулю.

Ранг матрицы может быть найден при последовательном рассмотрении так называемых окаймляющих миноров. Данный способ основан на свойствах миноров.

Из теоремы о разложении определителя по элементам строки (или столбца) следует, что если все миноры j-го порядка рассматриваемой матрицы равны нулю, то и миноры более высоких порядков будут равны нулю.

Следовательно, чтобы найти ранг матрицы надо определить порядок миноров, выше которого все миноры равны нулю, т. е. если не все миноры i-го порядка равны нулю, а миноры порядка $(i + 1)$ (окаймляющие) все равны нулю, то ранг матрицы равен i $(r = i)$.

Как видим, ранг определяется последовательной проверкой миноров, начиная с минора первого порядка, в сторону увеличения порядка. При таком способе определения ранга матрицы определяются одновременно столбцы и строки этой матрицы, составляющие максимальную линейно независимую систему.

Ранг матрицы может быть определен при применении элементарных преобразований матрицы, но при этом элементы столбцов и строк этой матрицы, составляющих максимальную линейно независимую систему исходной матрицы, не определяются.

К элементарным преобразованиям матрицы относятся следующие действия (аналогично элементарным преобразованиям линейных систем):

1) перестановка местами двух параллельных рядов (строк или столбцов) матрицы;
2) умножение строки или столбца матрицы на не равное нулю число;
3) прибавление к какой-нибудь строке (или столбцу) другой строки (или столбца), умноженной на любое число.

Применяя эти элементарные преобразования, всякую матрицу можно привести к ступенчатому (квазитреугольному) виду.

Если квазитреугольная матрица имеет вид

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1r} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2r} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{rr} & \cdots & a_{rn} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix},$$

то ее ранг равен r, так как миноры более высокого порядка содержат строки нулей, а значит, равны нулю, т. е. для определения ранга матрицы строки или столбцы исходной матрицы подвергают элементарным преобразованиям до тех пор, пока не получится строка нулей (нулевая строка).

Применение элементарных преобразований не влияет на ранг матрицы, т. е. ранг исходной матрицы равен рангу матрицы, полученной в результате элементарных преобразований.

Например, найдем ранг матрицы путем элементарных преобразований:

$$\begin{matrix} 1 & -1 & 0 & 3 \\ 4 & -2 & 5 & 0 \\ 2 & -3 & 0 & 6 \\ 7 & -6 & 5 & 9 \end{matrix} \rightarrow \begin{matrix} 1 & -1 & 0 & 3 \\ 0 & 2 & 5 & -12 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 5 & -12 \end{matrix} \rightarrow \begin{matrix} 1 & -1 & 0 & 3 \\ 0 & 1 & 5 & -12 \\ 0 & -1 & 0 & 0 \\ 0 & 2 & 5 & -12 \end{matrix} \rightarrow$$

$$\rightarrow \begin{pmatrix} 1 & -1 & 0 & 3 \\ 0 & 1 & 5 & -12 \\ 0 & 0 & 5 & -12 \\ 0 & 0 & -5 & 12 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & 0 & 3 \\ 0 & 1 & 5 & -12 \\ 0 & 0 & 5 & -12 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Здесь вторая матрица была получена при прибавлении ко второй строке первой, умноженной на (–4), к третьей строке — прибавлением первой, умноженной на (–2), к четвертой — прибавлением первой, умноженной на (–7). Третья матрица получилась при перемене местами второй и четвертой строки, четвертая матрица получилась из предыдущей путем прибавления к третьей строке второй, а к четвертой строке — прибавлением второй, умноженной на (–2). И наконец, в четвертой матрице, прибавив к четвертой строке третью, получили пятую матрицу, соответствующую ступенчатому виду. Отсюда видно, что ранг исходной матрицы равен 3.

Теория определителей и матриц широко применяется при решении линейных систем m уравнений с *n* неизвестными. Линейные системы записывают в матричном виде и при решении используют свойства матриц и определителей.

ЛЕКЦИЯ № 3. Линейные системы. Решение линейных систем

1. Системы линейных уравнений (матричная запись)

Пусть дана система m линейных уравнений с n неизвестными:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases} \quad (1)$$

Здесь a_{ik} — коэффициенты системы, $b_1, b_2, \ldots, b_m$ — свободные члены. Если все свободные члены равны нулю, то система называется **однородной**.

Матрица A, составленная из коэффициентов системы (1), т. е.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

называется **матрицей системы** (1) (или **основной матрицей)**, а ее определитель $|A|$ является **определителем исходной системы** уравнений (1).

Расширенной матрицей линейной системы (1) называется матрица, которая получается из основной прибавлением столбца свободных членов. Для системы (1) расширенная матрица имеет вид:

$$\tilde{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}.$$

Решением системы (1) называется совокупность чисел $x_i = \lambda_i$ ($i = 1, 2, \ldots, n$), обращающих все уравнения системы в тождества. Если система линейных уравнений не имеет решений, то она называется **несовместной**. Если она имеет хотя бы одно решение, то она называется **совместной**.

Если обозначить буквой X матрицу-столбец из неизвестных и буквой B — матрицу-столбец их свободных членов:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

то тогда исходную систему уравнений можно записать следующим образом: $AX = B$ (умножение матриц A и X возможно, так как матрица A согласована с матрицей X, количество столбцов матрицы A и количество строк матрицы X равно числу n). Эта запись называется **матричным уравнением** (или **матричной записью линейной системы)**. Действительно, по правилам умножения матриц произведение AX равно

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix}.$$

Справедливо утверждение, что определенной линейной системе соответствует своя единственная пара матриц A и B, обратное тоже верно.

Имеют место следующие равенства: $A^{-1}(AX) = A^{-1}B$, $X = A^{-1}B$.

Используя матрицы-столбцы коэффициентов линейной системы (1), эту систему можно записать в следующем виде:

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} x_2 + \ldots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Если некоторая совокупность чисел p_1, p_2, ..., p_n является решением линейной системы (1), то вектор-решение имеет вид матрицы-столбца

$$P = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix}.$$

2. *Решение линейных систем n уравнений с n неизвестными с помощью определителей. Теорема Крамера*

Пусть Δ (или $\det A$) — определитель системы n линейных уравнений с n неизвестными

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \qquad\qquad \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n, \end{cases} \tag{1}$$

т. е.

$$\Delta = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

Δ^i — определитель, полученный из определителя Δ заменой его i-го столбца из коэффициентов при неизвестном x_i столбцом свободных членов.

Теорема Крамера (формулы Крамера). Если определитель Δ системы линейных уравнений (1) не равен нулю, то эта система имеет единственное решение: $x_i = \Delta^i/\Delta$, где $i = 1, 2, \ldots, n$.

Доказательство. Так как по условию теоремы определитель матрицы А линейной системы не равен нулю, то существует матрица A^{-1}, обратная для матрицы A. Умножим на A^{-1} матричную форму исходной линейной системы: : $A^{-1}AX = A^{-1}B$. Но $A^{-1}AX = EX = X$, следовательно, $X = A^{-1}B$.

Распишем полученное матричное равенство:

$$\begin{matrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{matrix} = \frac{1}{\Delta} \begin{matrix} A_{11} & A_{21} & \cdots & A_{n1} \\ \vdots & \vdots & \cdots & \vdots \\ A_{1i} & A_{2i} & \cdots & A_{ni} \\ \vdots & \vdots & \vdots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{matrix} \quad \begin{matrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{matrix} =$$

$$= \frac{1}{\Delta} \begin{pmatrix} A_{11}b_1 + A_{21}b_2 + \ldots + A_{n1}b_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ A_{1i}b_1 + A_{2i}b_2 + \ldots + A_{ni}b_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ A_{1n}x_1 + A_{2n}x_2 + \ldots + A_{nn}x_n \end{pmatrix}.$$

Но так как $\Delta^i = A_{1i}b_1 + A_{2i}b_2 + \ldots + A_{ni}x_x$ (по теореме замещения), то из расписанного матричного равенства видно, что $x_i = \Delta^i/\Delta$, где $i = 1, 2, \ldots, n$. Теорема доказана.

Следствие. Если однородная система имеет ненулевое решение, то ее определитель Δ равен нулю.

По формулам Крамера удобно решать линейные системы при небольших значениях n.

Решать линейную систему можно и путем элементарных преобразований ее матрицы, так как при этом получаются матрицы, соответствующие эквивалентным системам.

Пример. Решить линейную систему

$$\begin{cases} 2x_1 - x_2 + 3x_3 = 1, \\ x_1 - 2x_2 + 4x_3 = 10, \\ 5x_1 + x_2 - x_3 = 5, \end{cases}$$

используя формулы Крамера. Определитель этой системы

$$\Delta = \begin{vmatrix} 2 & -1 & 3 \\ 1 & -2 & 4 \\ 5 & 1 & -1 \end{vmatrix} = 8$$

(т. е. не равен нулю), определители со столбцом свободных членов равны:

$$\Delta^1 = \begin{vmatrix} 1 & -1 & 3 \\ 10 & -2 & 4 \\ 5 & 1 & -1 \end{vmatrix} = 28,$$

$$\Delta^2 = \begin{vmatrix} 2 & 1 & 3 \\ 1 & 10 & 4 \\ 5 & 5 & -1 \end{vmatrix} = -174,$$

$$\Delta^3 = \begin{vmatrix} 2 & -1 & 1 \\ 1 & -2 & 10 \\ 5 & 1 & 5 \end{vmatrix} = -74.$$

Тогда по формулам Крамера $x_1 = {}^{\Delta}/_{\Delta} = {}^{28}/_8 = {}^7/_2$, $x_2 = {}^{\Delta}/_{\Delta} = -{}^{174}/_8 = -{}^{87}/_4$, $x_3 = {}^{\Delta}/_{\Delta} = -{}^{74}/_8 = -{}^{37}/_4$.

Пример. Решить линейную систему

$$\begin{cases} x_1+x_2+x_3-x_4=1, \\ 3x_1+x_2+2x_3-2x_4=2, \\ 2x_1-4x_2-3x_3+6x_4=7, \\ 7x_1+5x_2+6x_3-6x_4=6, \end{cases}$$

используя элементарные преобразования. Составим матрицу следующим образом: элементы первых четырех столбцов — соответствующие коэффициенты при неизвестных, последний столбец состоит из свободных членов. К этой матрице применим элементарные преобразования:

$$\begin{pmatrix} 1 & 1 & 1 & -1 & 1 \\ 3 & 1 & 2 & -2 & 2 \\ 2 & -4 & -3 & 6 & 7 \\ 7 & 5 & 6 & -6 & 6 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & -1 & 1 \\ 0 & -2 & -1 & 1 & -1 \\ 0 & -6 & -5 & 8 & 5 \\ 0 & -2 & -1 & 1 & -1 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 1 & 1 & -1 & 1 \\ 0 & -2 & -1 & 1 & -1 \\ 0 & 0 & -2 & 5 & 8 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Здесь на первом шаге ко второму уравнению прибавили первое уравнение, умноженное на (–3), к третьему прибавили первое, умноженное на (–2), к четвертому — первое, умноженное на (–7). На втором шаге к третьему уравнению прибавили второе, умноженное на (–3), к четвертому прибавили второе, умноженное на (–1). В результате получилась матрица ступенчатого вида, которая соответствует системе трех линейных уравнений: $\begin{cases} x_1+x_2+x_3-x_4=1 \\ -2x_2-x_3+x_4=-1. \\ -2x_3+5x_4=8 \end{cases}$

В данном случае одно из неизвестных, например x_4, может принимать любые значения, а остальные выражаются через него: $x_1 = {}^5/_2 - {}^{3x}/_4$, $x_2 = {}^5/_2 - {}^{3x}/_4$, $x_3 = {}^{5x}/_2 - 4$.

Линейная однородная система в любом случае имеет нулевое решение, которое называется **тривиальным**, но она может иметь и ненулевое, т. е. **нетривиальное**, решение. Для однородных линейных систем справедливы следующие утверждения:

1) если определитель однородной системы не равен нулю, то эта система имеет только тривиальное решение;

2) если однородная система уравнений имеет нетривиальное решение, то ее определитель равен нулю;

3) если определитель однородной линейной системы уравнений равен нулю, то эта система имеет нетривиальное решение.

3. Исследование систем m линейных уравнений с n неизвестными

Пусть дана система m линейных уравнений с n неизвестными:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases},$$

A — ее основная матрица, составленная из коэффициентов при неизвестных, $\tilde{A}$ — расширенная матрица этой системы, состоящая из коэффициентов при неизвестных и столбца свободных членов.

По рангу основной матрицы можно судить о наличии решения (совместности) рассматриваемой системы. Для данной системы справедливы следующие утверждения:

1) для совместности линейной системы необходимо и достаточно, чтобы ранг ее основной матрицы А был бы равен рангу расширенной матрицы $\tilde{A}$ (теорема Кронекера—Капелли) ;

2) если ранг матрицы совместной системы равен числу неизвестных, то система имеет единственное решение;

3) если ранг матрицы совместной системы меньше числа неизвестных, то множество ее решений является бесконечным.

Доказательство **теоремы Кронекера—Капелли**. Исходную систему можно записать в следующем векторном виде:

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} x_2 + \ldots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

т. е. столбец свободных членов является линейной комбинацией столбцов основной матрицы A. Если исходная линейная система совместна (имеет решение), то ранг основной матрицы равен рангу расширенной. И наоборот, если ранги расширенной и основной систем совпадают, то столбец свободных членов может быть представлен как линейная комбинация столбцов основной матрицы, которые можно считать базисными. Значит, исходная система совместна.

Таким образом, при исследовании системы сначала определяют ранги основной и расширенной матриц исходной системы. Если эти ранги не равны, то система несовместна. Если ранги равны, то в случае, когда порядок базисного минора (максимального не равного нулю минора основной матрицы) равен числу неизвестных, система имеет единственное решение; если меньше числа неизвестных, система имеет бесконечно много решений, которые могут быть определены, например, с помощью элементарных преобразований матрицы рассматриваемой системы.

ЛЕКЦИЯ № 4. Линейные пространства

1. Линейные (арифметические) пространства и подпространства

Понятие арифметического пространства широко применимо в теории матриц. **Линейным пространством (арифметическим, действительным) размерности *n* над *R* (или действительным векторным пространством)** называется множество R^n, элементами которого являются векторы или векторы-строки (векторы-столбцы), рассматриваемые вместе с операциями сложения векторов и умножения их на скалярные величины, удовлетворяющими своим аксиомам. Здесь n — определенное натуральное число, R — множество действительных чисел. Вектор-строка рассматривается как матрица размером $1 \times n$ (вектор-столбец — матрица размером $n \times 1$). Понятие линейного пространства можно ввести следующим образом: если некоторой упорядоченной паре элементов x, y, принадлежащих множеству V (или R^n), соответствует элемент $z \in V$, являющийся суммой ($z = x + y$), и определено умножение на скаляр $\alpha \in R$, таким образом: $z = ax, z \in V$, причем эти операции сложения и умножения удовлетворяют своим аксиомам, то это множество V называется **линейным пространством** (или **действительным векторным пространством)**. Если вместо множества действительных чисел R взять множество комплексных чисел, то пространство будет называться **комплексным линейным пространством**.

Аксиомы действительного линейного пространства:

1) сложение коммутативно: $x + y = y + x$ для любых $x, y \in V$;
2) сложение ассоциативно: $(x + y) + z = x + (y + z)$ для любых $x, y, z \in V$;
3) существует нулевой элемент 0 (нулевой вектор), для которого справедливо $0 + x = x$ для любых $x \in V$;
4) для любого $x \in V$ существует противоположный элемент $(-x)$, для которого $(-x) + x = 0$;
5) для любого $x \in V$ произведение $1x = x$;
6) для любых $x \in V$ и $\alpha, \delta \in R$ справедливо $\alpha(\delta x) = (\alpha\delta)x$;

7) умножение дистрибутивно относительно скалярной величины: $(\alpha + \delta)x = \alpha x + \delta x$ для любых $x \in V$, $\alpha, \delta \in R$;

8) умножение дистрибутивно относительно векторов (элементов пространства V): $\alpha(x + y) = \alpha x + \alpha y$ для любых x, $y \in V$, $\alpha \in R$.

Разность двух элементов линейного пространства x, y определяется как сумма $x + (-y)$ (сумма x и элемента, противоположного по знаку элементу y).

К линейным пространствам относится, например, множество свободных векторов (нулевой элемент — нулевой вектор); множество матриц размером $m \times n$ (нулевой элемент — нулевая матрица); множество алгебраических многочленов степени меньше n (нулевой элемент — многочлен, у которого все коэффициенты равны нулю); множество непрерывных функций $f(x)$, определенных на отрезке $[a, b]$ (нулевой элемент — тождественно равная нулю функция).

Для линейных пространств справедливы следующие утверждения:

1) нулевой элемент — единственный (действительно, если бы было два нулевых элемента 0_1 и 0_2, то из равенств, следующих из аксиомы 3, $0_1 + 0_2 = 0_1$ и $0_1 + 0_2 = 0_2$, вытекает, что $0_1 = 0_2$);

2) для любого элемента x линейного пространства противоположный по знаку элемент единственный (действительно, если бы было два противоположных элемента x_1 и x_2, т. е. по аксиоме 4$x + x_1 = 0$ и $x + x_2 = 0$, то преобразованиями получаем: $x_1 = x_1 + 0 = x_1 + (x + x_2) = (x_1 + x) + x_2 = 0 + x_2 = x_2$);

3) элемент $(-x)$ является противоположным для элемента x (действительно, основываясь на аксиомах 1 и 3, имеем: $-x + x = x + (-x) = 0$);

4) для любого элемента x линейного пространства произведение в результате умножения его на число 0 является нулевым элементом (действительно, из следующего преобразования суммы $0x + 0 = 0x + (x + (-x)) = (0x + x) + (-x) = x(0 + 1) + (-x) = x + (-x) = 0$ получаем, что $0x = 0$);

5) для любого элемента линейного пространства произведение в результате умножения его на число (-1) есть элемент, противоположный элементу x, т. е. $(-1)x = -x$ (действительно, рассматривая сумму и используя предыдущее утверждение $(-1)x + x = (-1 + 1)x = 0x = 0$, получаем, что $(-1)x = -x$);

6) произведение в результате умножения нулевого элемента линейного пространства на любое действительное число есть

нулевой элемент (действительно, при вычислении получаем: $\alpha 0 = \alpha(x + (-x)) = \alpha(x + (-1)x) = \alpha x + \alpha(-1)x = \alpha x + (-\alpha x) = 0$);

7) если произведение $\alpha x = 0$ (равно нулевому элементу) при α, отличном от нуля, то x — нулевой элемент (действительно, умножив обе части равенства $\alpha x = 0$ на число $^1/_\alpha$, с учетом предыдущего утверждения получаем, что $x = 0$ (нулевой элемент));

8) если произведение $\alpha x = 0$ (нулевому элементу) в случае, когда x не является нулевым элементом, то $\alpha = 0$ (действительно, предположив, что α не равно нулю, сможем произвести следующие действия: $(^1/_\alpha)(\alpha x) = (^1/_\alpha)0 = 0$, откуда получим, что $x = 0$ (является нулевым элементом), что противоречит условию).

Линейным подпространством *W* называется непустое подмножество линейного пространства V (или Rn), если выполняются следующие условия: на множестве W определены те же операции, что и на множестве V, и если $x, y \in W$, то и $x + y \in W$, $\alpha x \in W$, где α — число. Таким образом, в линейном подпространстве выполняются все аксиомы линейного пространства, т. е. линейное подпространство само является линейным пространством. Нулевой элемент линейного пространства сам является линейным подпространством (нулевым подпространством), так как для него удовлетворяются все условия линейного подпространства. Например, линейным подпространством будет являться множество свободных векторов, удовлетворяющих какому-либо условию, например параллельных некоторой прямой.

Евклидовым пространством называется линейное действительное пространство, в котором определена операция скалярного умножения векторов следующим образом: упорядоченной паре векторов x, y ставится в соответствие число (x, y), принадлежащее этому же пространству, и выполняются следующие аксиомы:

$$(x, y) = (y, x),$$
$$(x + y, z) = (x, z) + (y, z),$$
$$(\alpha x, y) = \alpha(x, y),\ \alpha \in R;\ (x, x) > 0,$$

для всех $x \neq 0$, $(x, x) = 0$ при $x = 0$.

Норма вектора евклидова пространства равна $\|x\| = \sqrt{(x,x)} = \sqrt{x^2}$. Она обладает следующими свойствами, вытекающими из аксиом скалярного умножения:

1) $\|x\| = 0$ тогда и только тогда, когда $x = 0$;

2) $\|\alpha x\| = |\alpha|\,\|x\|$, где α — действительное число;

3) $|(x, y)| \le \|x\|\,\|y\|$; $\|x + y\| \le \|x\| + \|y\|$.

Например, в евклидовом пространстве с обычным определением скалярного произведения, элементами которого являются трехмерные свободные векторы, норма вектора совпадает с его длиной; в пространстве, элементами которого являются определенные и непрерывные на отрезке $[a, b]$ функции вида $x = (t)$, а скалярное произведение двух функций $x(t)$, $y(t)$ определено формулой

$$(x, y) = \int_a^b x(t)y(t)dt,$$

норма определяется формулой

$$\|x(t)\| = \sqrt{\int_a^b x^2(t)dt};$$

в n-мерном линейном пространстве упорядоченных совокупностей n действительных чисел вида $x = (x_1, x_2, ..., x_n)$, для двух элементов которого скалярное произведение определено следующим соотношением $(x, y) = x_1y_1 + x_2y_2 + ... + x_ny_n$ (или иначе можно сказать, что скалярное произведение векторов $x = (x_1, x_2, ..., x_n)$ и $y = (y_1, y_2, ..., y_n)$ определено как сумма произведений их соответствующих координат, норма определяется формулой:

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}.$$

Углом между двумя векторами *x* и *y* евклидова пространства называется угол φ, для которого

$$\cos\varphi = \frac{(x, y)}{\|x\| \cdot \|y\|}, \quad (0 \le \varphi \le 2\pi).$$

Ортогональными называются два вектора евклидова пространства, если их скалярное произведение равно нулю. Система попарно ортогональных векторов называется **ортогональной** системой.

Вектор называется **нормированным** (единичным), если его норма равна единице. Система векторов называется **ортонормированной**, если она ортогональна и все ее векторы нормированы.

2. Линейные комбинации. Линейная зависимость

Линейной комбинацией векторов (элементов линейного пространства) $x_1, x_2, ..., x_n$ называется вектор $y = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n$, где α_i $(i = 1, 2, ..., n)$ — числа, которые называются **коэффициентами** линейной комбинации. Линейная комбинация называется **тривиальной**, если все ее коэффициенты равны нулю; если не все коэффициенты равны нулю, то такая линейная комбинация называется **нетривиальной**. С одним и тем же набором векторов можно составлять различные линейные комбинации с разными коэффициентами. Множество линейных комбинаций векторов $x_1, x_2, ..., x_n$ является линейным подпространством множества векторов, это подпространство называют линейной оболочкой векторов $x_1, x_2, ..., x_n$ или говорят, что оно порождено векторами $x_1, x_2, ..., x_n$.

Линейнозависимой называется такая система векторов $x_1, x_2, ..., x_n$, для которой существуют коэффициенты α_i $(i = 1, 2, ..., n)$, не все одновременно равные нулю, такие, что линейная комбинация $\alpha_1 x_1 + \alpha_2 x^2 + ... + \alpha_n x_n = 0$. Если данная линейная комбинация равна нулю только тогда, когда все коэффициенты αi $(i = 1, 2, ..., n)$ равны нулю, эта линейная комбинация называется **линейно независимой**. Система единичных векторов линейно независима. Линейная зависимость (независимость) системы векторов не зависит от порядка векторов в этой системе (от перемены мест слагаемых сумма не меняется). Два (три) линейнозависимых вектора линейного пространства иначе называются **коллинеарными (компланарными)**, линейно независимые векторы называются **неколлинеарными (некомпланарными)**.

Справедливы следующие утверждения:

1) система векторов, содержащая нулевой вектор, линейнозависима (действительно, если, например, $x_1 = 0$ (нулевой вектор), то для такой системы существует набор не всех одновременно равных нулю коэффициентов, а именно $a_1 \neq 0, 0, ..., 0$, при котором линейная комбинация равна нулю);

2) если m $(m < n)$ векторов системы $x_1, x_2, ..., x_n$ линейнозависимы, то и вся система линейно зависима (действительно, если коэффициенты α_i $(i = 1, 2, ..., m)$ не все одновременно равны нулю, но $\alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_m x_m = 0$, то, добавив к этому набору коэффициентов нули в качестве коэффициентов для векторов $x_{m+1}, ..., x_n$, получим новый набор коэффициентов,

в котором есть не равные нулю, при котором линейная комбинация $\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = 0$);

3) любая часть линейно независимой системы векторов линейно независима, т. е. при отбрасывании части векторов линейно независимая система остается линейно независимой (действительно, данное утверждение следует из предыдущего при рассуждениях от противного);

4) если какие-либо два вектора, например x_i и x_j, системы x_1, x_2, ..., x_n связаны соотношением $xi = \beta x_j$ (β — число), то вся система векторов линейнозависима (действительно, векторы x_i и x_j являются линейно зависимой частью системы x_1, x_2, ..., x_n, так как $1x_i - \beta x_j = 0$, следовательно, и вся система векторов линейнозависима);

5) система векторов x_1, x_2, ..., x_n линейнозависима тогда и только тогда, когда хотя бы один из этих векторов является линейной комбинацией остальных (действительно, необходимость видно, если условие линейной зависимости $\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = 0$ при $a_1 \neq 0$ записать в виде: $x_1 = -\alpha_2 x_2/a_1 - \ldots - \alpha_n x_n/a_1$; достаточность следует (например, x_1 является линейной комбинацией остальных векторов системы), если равенство $x_1 = \alpha_2 x_2 + \ldots + \alpha_n x_n$ записать в виде $(-1)x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = 0$, т. е. в данной комбинации хотя бы первый коэффициент не равен нулю).

Если линейная система состоит только из нулевого вектора, то она линейнозависима.

Если линейная система состоит из одного ненулевого вектора, то она линейно независима.

3. Базис, разложение вектора по базису. Размерность. Изоморфизм

Размерностью *n* (dim *V*) линейного пространства *V* называется число (максимальное) линейно независимых векторов (если они существуют) этого пространства, т. е. для данного пространства существует *n* линейно независимых векторов, а любая система ($n + 1$) вектора этого пространства линейнозависима. Размерность линейного пространства, состоящего из одного нулевого вектора, равна нулю.

Базис n-мерного линейного пространства — это любая упорядоченная система n линейно независимых векторов этого пространства.

Если базис некоторого линейного пространства состоит из конечного числа векторов, то это пространство называется **конечномерным**.

Если для любого натурального числа m в линейном пространстве найдется m линейно независимых векторов, то такое пространство называется **бесконечномерным**.

Стандартным базисом n-мерного линейного пространства является линейно независимая система вида $e_1(1, 0, .., 0)$, $e_2(0, 1, 0, ..., 0)$, ..., $e_n(0, 0, ..., 0, 1)$; любой элемент данного линейного пространства можно представить в виде линейной комбинации векторов стандартного базиса: $x = \alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n$, где числа $\alpha_1, \alpha_2, ..., \alpha_n \in$ $\in R$ и являются **координатами вектора** x относительно базиса $e_1, e_2,$ $..., e_n$.

Одно и то же линейное пространство может иметь разные базисы, состоящие из разных векторов, но количество их одинаково. Базис двумерного (трехмерного) пространства состоит из двух неколлинеарных (трех некомпланарных) векторов. Каждое подпространство линейного n-мерного пространства обладает конечным базисом.

Лемма. Если система векторов $x_1, x_2, ..., x_r$ — базис линейного подпространства, а система векторов этого же подпространства $y_1,$ $y_2, ..., y_s$ линейно независима, то $s \leq r$ (доказывается от противного).

Во всяком евклидовом n-мерном пространстве ($n \geq 2$) существует ортонормированный базис (базис называется ортонормированным, если базисные векторы образуют ортонормированную систему).

Теорема. Если $e_1, e_2, ..., e_n$ — базис n-мерного линейного пространства V, то любой вектор x этого пространства линейно выражается через базисные векторы $x = \alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n$, и коэффициенты этого разложения определяются однозначно.

Доказательство. Система векторов по условию является базисом, т. е. она линейно независима, но, добавив к ней еще один вектор x, получаем уже линейнозависимую систему, т. е. линейная комбинация $\beta_1 e_1 + \beta_2 e_2 + ... + \beta_n e_n + \beta_{n+1} x = 0$, и не все коэффициенты равны нулю. Причем именно $\beta_{n+1} \neq 0$, в противном случае (при $\beta_{n+1} = 0$) для выполнения последнего равенства и сохранения условия линейной зависимости среди коэффициентов β_1, ..., β_n должен быть хотя бы один не равный нулю, что противоречит определению базиса. Но если $\beta_{n+1} \neq 0$, то x можно представить в виде $x = \alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n$, где $\alpha_i = -\beta_i/\beta_{n+1}$. Покажем, что это разложение единственное.

Если для вектора x возможны два разложения по базису с разными коэффициентами α_i и β_i, т. е. $x = \alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n = \beta_1 e_1 + \beta_2 e_2 + ... + \beta_n e_n$, то отсюда $(\alpha_1 - -\beta_1)e_1 + ... + (\alpha_n - \beta_n)e_n = 0$, что возможно в силу линейной независимости базиса лишь при $\alpha_i - \beta_i = 0$ или $\alpha_i = \beta_i$ ($_i = 1, 2, ..., _n$). Теорема доказана.

Справедливы следующие утверждения (они легко доказываются на основании аксиом линейного пространства):

1) вектор x является нулевым тогда и только тогда, когда его координаты при любом базисе равны нулю (действительно, при равных нулю координатах из свойств линейного пространства следует, что x — нулевой вектор; и наоборот, если вектор x — нулевой, то его разложение по базису равно 0, что возможно при условии линейной независимости базиса, лишь когда все коэффициенты равны нулю);
2) два вектора равны тогда и только тогда, когда равны их соответствующие координаты в одном и том же базисе;
3) вектор x является линейной комбинацией некоторой системы векторов тогда и только тогда, когда каждая координата вектора x является линейной комбинацией с этими же коэффициентами соответствующих координат векторов системы в одном и том же базисе;
4) в одном и том же базисе координаты суммы двух векторов равны сумме соответствующие координат этих векторов, координаты произведения вектора на число равны произведению координат этого вектора на это число.

Изоморфными называются такие линейные пространства V, U, между элементами которых $x_i \in V$ и $y_i \in U$ установлено взаимно-однозначное соответствие, такое, что сумме $x_1 + x_2$ соответствует сумма $y_1 + y_2$, произведению αx_i соответствует произведение αy_i, где $\alpha \in \mathrm{R}$.

Необходимым и достаточным условием изоморфности двух линейных пространств является равенство их размерностей.

4. Матрица системы векторов линейного пространства. Преобразование координат при изменении базиса

Пусть имеется некоторое n-мерное линейное пространство с определенным базисом, и пусть задана система m векторов этого пространства, разложенных по выбранному базису:

$x_1(a_{11}, a_{21}, ..., a_{n1}), x_2(a_{12}, a_{22}, ..., a_{n2}), ..., x_{\mathrm{m}}(a_{1m}, a_{2m}, ..., a_{nm})$. Матрица A данной системы векторов составляется по следующему правилу: элементами любого i-го столбца матрицы являются координаты i-го вектора. Такая матрица имеет размерность $n \times m$ (m столбцов и n строк), ее ранг является рангом рассматриваемой системы векторов. Если векторы исходной системы линейнозависимы, то и столбцы матрицы A будут также линейнозависимыми. Матрица A имеет следующий вид:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}.$$

Основываясь на свойствах матриц, имеем следующую теорему со следствиями.

Теорема. Для того чтобы m векторов n-мерного линейного пространства были линейно зависимыми, необходимо и достаточно, чтобы ранг матрицы этой системы был бы равен m.

Следствия. 1. Система n векторов n-мерного линейного пространства линейнонезависима тогда и только тогда, когда матрица этой системы векторов является невырожденной.

2. Если ранг матрицы системы m векторов линейного пространства равен r, то максимальное число линейно независимых векторов этой системы равно r.

Таким образом, определение максимального числа линейно независимых векторов некоторой системы векторов сводится к определению ранга матрицы этой системы векторов.

Пусть имеются два базиса в одном n-мерном линейном пространстве: $e_1, e_2, ..., e_n$ и $e^*_1, e^*_2, ..., e^*_n$. Векторы одного базиса можно разложить по другому базису, например векторы второго базиса разложены по первому: $e^*_i = t_{i1}e_1 + t_{i2}e_2 + ... + t_{in}e_n$ ($i = 1, 2, ..., n$). Матрица составленная из коэффициентов данного разложения, называется **матрицей перехода от базиса** $e_1, e_2, ..., e_n$ к базису $e^*_1, e^*_2, ..., e^*_n$.

Она имеет следующий вид:

$$T = \begin{pmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{n1} & \cdots & t_{nn} \end{pmatrix}.$$

Так как векторы базиса линейно независимы, то данная матрица невырожденная, т. е. ее определитель не равен нулю. Матрица T^{-1} является матрицей перехода от базиса $e^*_1, e^*_2, ..., e^*_n$ к базису $e_1, e_2, ..., e_n$.

Теорема. Если $x_1, x_2, ..., x_n$ — координаты вектора х в базисе $e_1, e_2, ..., e_n$, $x^*_1, x^*_2, ..., x^*_n$ — координаты этого же вектора в другом базисе $e^*_1, e^*_2, ..., e^*_n$, то справедливо $X = TX^*$, где T — матрица перехода от базиса $e_1, e_2, ..., e_n$ к базису $e^*_1, e^*_2, ..., e^*_\mathrm{n}$, а X и X^* — матрицы-столбцы, соответственно равные

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad X^* = \begin{pmatrix} x^*_1 \\ \vdots \\ x^*_n \end{pmatrix}.$$

Данная теорема доказывается при непосредственном сравнении разложения вектора х по первому базису с матрицей, равной произведению матриц TX^*. Эта теорема выражает координаты вектора в первом (старом) базисе через координаты этого вектора во втором (новом) базисе. Можно новые координаты выразить через старые по формуле: $X^* = T^{-1}X$ (данная формула получается из формулы $X = TX^*$ при умножении ее слева на обратную матрицу T^{-1}).

ЛЕКЦИЯ № 5. Линейное преобразование

1. Линейное преобразование (отображение)

Пусть имеется некоторое линейное пространство V. **Задать преобразование** f пространства V самого в себя — значит поставить каждому вектору x (который при этом будет называться **прообразом**) пространства V единственный вектор у (**образ**) этого же пространства V. При отображении у каждого вектора пространства V имеется образ, но прообраза может и не быть, или прообраз может быть неединственным. **Взаимно-однозначным (биективным)** называется такое преобразование, при котором каждый вектор имеет единственный прообраз.

Линейным (или линейным оператором) называется преобразование f линейного пространства, при котором справедливо равенство

$$f(\alpha x_1 + \beta x_2) = \alpha f(x_1) + \beta f(x_2) \text{ для любых чисел } \alpha, \beta.$$

При линейном преобразовании нулевой вектор переходит в нулевой. Действительно, $f(0) = f(0x) = 0\,f(x) = 0$.

Тождественное преобразование $f(x) = x$ является примером простейшего линейного преобразования.

Если имеются два базиса в одном n-мерном линейном пространстве $e_1, e_2, ..., e_n$ и $e^*_1, e^*_2, ..., e^*_n$ и линейное преобразование f переводит базис $e_1, e_2, ..., e_n$ в базис $e^*_1, e^*_2, ..., e^*_n$: $e^*_i = f(e_i)$, то образ любого вектора $x = \alpha_1 e_1 + ... + \alpha_n e_n$ этого пространства выражается через образы базисных векторов следующим образом: $f(x) = \alpha_1 e^*_1 + ... + \alpha_n e^*_n$.

Если f переводит базис $e_1, e_2, ..., e_n$ в базис $e^*_1, e^*_2, ..., e^*_n$, а векторы базиса $e^*_1, e^*_2, ..., e^*$ могут быть разложены по базису $e_1, e_2, ..., e_n$: $e^*_i = f(e_i)$, то матрицей данного линейного преобразования f будет матрица

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

в которой элементами любого i-го столбца являются координаты вектора e^*_i при разложении по базису $e_1, e_2, ..., e_n$: $e^*_i = a_{1i}e_1 + a_{2i}e_2 + ... + a_{ni}e_n$ ($i = 1, 2, ..., n$). Рангом данного преобразования r является ранг матрицы A, а разность $n - r$ называется **дефектом** этого преобразования. Для тождественного преобразования матрица единичная. Данное преобразование называется невырожденным, если определитель $|A| \neq 0$.

Пусть вектор $y = f(x) = y_1e_1 + ... + y_ne_n$, т. е. является образом вектора $x = x_1e_1 + ...+ x_ne_n$ (они разложены по одному и тому же базису $e_1, e_2, ..., e_n$). Тогда их координаты связаны следующим соотношением: $Y = AX$ (в матричной форме) или $y_i = a_{i1}x_1 + a_{i2}x_2 + +... +a_{in}x_n$ ($i = 1, 2, ..., n$), где X, Y — матрицы-столбцы соответственно с элементами (x_i) и (y_i).

Теорема. Если $e_1, e_2, ..., e_n$ (старый базис) и $e^*_1, e^*_2, ..., e^*_n$ (новый базис) — два базиса в одном n-мерном линейном пространстве, A — матрица линейного преобразования в старом базисе, то матрица B этого преобразования в новом базисе имеет вид: $B = T^{-1}AT$, где T — матрица перехода из старого базиса к новому.

Действительно, пусть имеются два вектора x и y, разложенные по обоим базисам: $x = x_1e_1 + ... + x_ne_n = x^*_1e^*_1 + ... + x^*_ne^*_n$ и $y = y_1e_1 + ... + y_ne_n = y^*_1e^*_1 + ... + y^*_ne^*_n$. Но известно, что при переходе от одного базиса к другому координаты вектора связаны соотношениями $X = TX^*$, $Y = TY^*$, а в соответствии с матричной записью линейного преобразования имеем: $Y = AX$ и $Y^* = BX^*$.

Если равенство $X = TX^*$ умножить слева на матрицу A, то получаем: $AX = ATX^*$, но $AX = Y = TY^*$, тогда $TY^* = ATX^*$ или $Y^* = T^{-1}ATX^*$.

При сравнении полученного выражения с равенством $Y = BX^*$ получаем, что $B = \mathbf{T}^{-1}AT$.

Следствие. Если линейное преобразование имеет невырожденную матрицу в некотором базисе, то матрица этого преобразования будет невырожденной в любом другом базисе.

Подобными называются такие две матрицы A и B, для которых существует невырожденная матрица C, такая, что выполняется равенство $B = C^{-1}AC$ (в этом случае говорят, что матрица B подобна матрице A).

2. Характеристическое уравнение и собственный вектор линейного преобразования

Характеристическим уравнением линейного преобразования f называется уравнение вида $\det(A - \lambda E) = 0$, где λ — любое действительное число, A — матрица этого линейного преобразования f, E — единичная матрица того же порядка n. Многочлен $\det(A-\lambda E)$ называется **характеристическим многочленом $P_n(\lambda)$** матрицы A (линейного преобразования f). В матричном виде характеристическое уравнение имеет следующий вид:

$$\det\left(\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n} \end{pmatrix} - \lambda \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}\right) = 0$$

или

$$P_n(\lambda) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0.$$

Следовательно, приравняв характеристический многочлен к нулю, получаем уравнение степени n, где в качестве неизвестного выступает λ. Решив это уравнение относительно λ, получаем значения его корней — характеристических чисел данной матрицы.

Пример. Найти характеристические числа матрицы

$$A = \begin{pmatrix} 4-\lambda & 2 \\ 3 & 1-\lambda \end{pmatrix}.$$

Приравнивая характеристический многочлен к нулю, получаем (в данном случае матрица второго порядка) квадратное уравнение:

$$\begin{vmatrix} 4-\lambda & 2 \\ 3 & 1-\lambda \end{vmatrix} = (4-\lambda)(1-\lambda) - 2\times 3 = \lambda^2 - 5\lambda - 2 = 0$$

Тогда корни этого уравнения равны $\lambda_{1,2} = (5 \pm \sqrt{17})/2$.

Теорема. Если линейное преобразование f в базисе $e_1,\dots,e_n$ (первый базис) имеет матрицу A и в базисе $e^*_1,\dots,e^*_n$ (второй базис) — матрицу B, то имеет место равенство: $\det(A-\lambda \mathrm{E}) = \det(\mathrm{B}-\lambda \mathrm{E})$. Следовательно, при переходе к новому базису характеристический многочлен линейного преобразования не меняется, хотя матрица линейного преобразования меняется. Действительно, если T — матрица перехода от первого базиса ко второму, то $B = T^{-1}AT$. Тогда преобразуем правую часть равенства: $\det(B - \lambda E) = \det(T^{-1}AT - \lambda E) = \det(T^{-1}AT - \lambda T^{-1}ET) = \det(T^{-1}(A - \lambda E)T) = \det T^{-1} \det(A-\lambda E) \det T = \det(A-\lambda E)$, т. е. получили требуемое.

Собственный вектор x линейного преобразования f — это такой ненулевой вектор линейного пространства, для которого существует такое число k, что выполняется равенство $f(x) = kx$ (число k называется **собственным значением** вектора x относительно преобразования f), или в матричном виде $AX = kX$, где A — матрица линейного преобразования f, X — матрица-столбец, элементами которой являются координаты вектора x.

Для собственных векторов и значений справедливы следующие свойства:

1) каждому собственному вектору соответствует единственное собственное значение k. Действительно, если вектору x соответствуют два числа, т. е. $f(x) = kx$ и $f(x) = k^*x$, то тогда выполняется равенство $kx = k^*x$ или $(k - k^*)x = 0$, последнее равенство выполняется при ненулевом векторе x только в случае $k = k^*$;

2) если собственному вектору x некоторого линейного преобразования соответствует собственное число k, то вектор αx также является собственным вектором этого же линейного преобразования с тем же собственным значением k, действительно, так как k и α — числа, то из свойств линейного пространства следует: $f(\alpha x) = \alpha f(x) = \alpha(kx) = k(\alpha x)$;

3) сумма двух линейно независимых собственных векторов x и y линейного преобразования f, имеющих одно и то же собственные значение k (т. е. $f(x) = kx$ и $f(y) = ky$), является собственным вектором того же линейного преобразования с тем же собственным значением $f(x + y) = k(x + y)$; действительно, сумма x + y не является нулевым вектором, поскольку x и y линейно независимы, и по свойствам линейного преобразования $f(x + y) = f(x) + f(y) = kx + ky = k(x + y)$;

4) если собственные значения k и m двух собственных векторов x и y одного линейного преобразования не равны, то векторы x и y линейно независимы; действительно, если бы векторы x и y были линейнозависимы, то можно было бы записать $y = \alpha x$ $(\alpha \neq 0)$. А исходя из первых двух свойств и этого равенства следует, что $k = m$, т. е. получили противоречие условию, следовательно, векторы x и y линейно независимы.

Обобщая данные свойства, можно сказать, что любая нетривиальная линейная комбинация системы собственных линейно независимых векторов, имеющих одинаковое собственное значение k, является собственным вектором того же линейного преобразования с тем же собственным значением k.

Теорема. В комплексном линейном пространстве все корни характеристического уравнения (и только они) являются собственными значениями линейного преобразования (следует из матричной записи определения собственного вектора, представленной в виде системы уравнений).

Для линейного преобразования действительного пространства собственные значения являются действительными числами. Если какое-либо собственное значение является корнем кратности m характеристического уравнения, то это собственное значение называется m-кратным (кратности m).

Если матрица A линейного преобразования является симметрической (симметричные относительно главной диагонали элементы равны), то для нее справедливо:

1) корни характеристического уравнения симметрической матрицы являются действительными числами (действительная симметрическая матрица имеет только действительные собственные векторы);
2) собственные векторы действительной симметрической матрицы, соответствующие разным собственным значениям, ортогональны (два вектора называются ортогональными, если их скалярное произведение равно нулю, выражение скалярного произведения векторов через их координаты имеет вид:

$$(x,y) = x_1y_1 + x_2y_2 + ... + x_ny_n).$$

3. Диагональный вид матрицы линейного преобразования

Пусть имеется некоторое линейное преобразование с матрицей A.

Теорема (условие диагональности матрицы). Матрица линейного преобразования имеет диагональный вид тогда и только тогда, когда каждый базисный вектор является собственным вектором этого преобразования.

Действительно, если матрица линейного преобразования имеет диагональный вид

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

то векторы соответствующего базиса $e_1, e_2, ..., e_n$ могут быть представлены в виде $f(e_i) = \lambda$ei ($I = 1, 2, ..., n$), что означает, что они являются собственными векторами. И обратно, если векторы базиса являются собственными, т. е. имеет место равенство $f(e_i) = \lambda e_i$ ($I = 1, 2,..., n$), то матрица A имеет диагональный вид.

Приводимой к диагональному виду называется такая матрица A, для которой существует невырожденная матрица T, для которой матрица $T^{-1}AT$ является диагональной. Чтобы построить матрицу T, надо определить собственные числа λi из характеристического уравнения для матрицы A.

Теорема (о приводимости к диагональному виду). Матрица A линейного преобразования f n-мерного линейного пространства приводима к диагональному виду тогда и только тогда, когда существует базис этого пространства, состоящий из собственных векторов данного преобразования (доказательство основано на предыдущей теореме и определении приводимости к диагональному виду).

Матрица будет приводиться к диагональному виду, если все ее собственные числа λi попарно различны.

4. Действия над линейными преобразованиями

Произведением (композицией) линейного преобразования f на линейное преобразование g называется преобразование, являющееся последовательным применением преобразований f и g, обозначается $g \times f$, т. е. для вектора x имеем $g \times f(x) = g(f(x))$. Произведение линейных преобразований само является линейным преобразованием; действительно, для любых векторов x и y исходя из определения линейного пространства имеем: $g \times f(\lambda x + \beta y) = g(f(\lambda x + \beta y)) = g(\lambda f(x) + \beta f(y)) = \lambda g(f(x)) + \beta g(f(y)) = \lambda g \times f(x) + \beta g \times f(y)$.

Теорема: Если в некотором базисе линейные преобразования f и g имеют соответственно матрицы A и B, то их произведение $g \times f$ имеет матрицу BA.

Действительно, если для векторов x, y, z имеем $y = f(x)$ (в векторной форме $Y = AX$), $z = g(y)$ ($Z = BY$), произведение $z = g \times f(x)$ (матрица преобразования произведения $Z = DX$), то отсюда $Z = BY = BAX$, а сравнивая с векторной формой для произведения, получаем $B = BA$.

Сумма линейных преобразований f и g некоторого линейного пространства — это такое преобразование h, что для любого вектора x выполняется равенство $h(x) = f(x) + g(x)$; сумма преобразований обозначается $f + g$. Справедливо: $f + g = g + f$. Сумма линейных преобразований сама является линейным преобразованием (доказывается аналогично случаю произведения).

Теорема. Если линейные преобразования f и g в некотором базисе имеют соответственно матрицы A и B, то преобразование $f + g$ имеет в том же базисе матрицу $A + B$ (доказывается аналогично предыдущей теореме).

Линейные преобразования могут быть вырожденными (имеют вырожденную матрицу) и невырожденными (имеют невырожденную матрицу).

Теорема. Линейное преобразование является невырожденным тогда и только тогда, когда оно взаимно-однозначно.

Действительно, если матрица A линейного преобразования невырожденная (т. е. ее определитель не равен нулю), то для векторов x и y имеем $y = f(x)$ или в матричной форме $Y = AX$, где X и Y — матрицы-столбцы из коэффициентов соответственно вектора x и y. Для конкретного векторе y^* имеем $AX = Y^*$, но поскольку определитель матрицы А не равен нулю, то последнее уравнение имеет единственное решение x^*. И наоборот, если имеется взаимно-однозначное соответствие, т. е. каждому вектору y^* соответствует единственный вектор x^*, что означает, что система $AX^* = Y^*$ имеет единственное решение, то определитель матрицы A не равен нулю, преобразование невырожденное.

Теорема. Произведение двух линейных невырожденных преобразований есть невырожденное линейное преобразование. Действительно, если A и B матрицы преобразований f и g, то матрица произведения преобразований равна BA, ее определитель равен $\det BA = \det B \det A \neq 0$, поскольку $\det A \neq 0$ и $\det B \neq 0$.

Преобразование g называется **обратным** преобразованию f, если произведение этих преобразований — тождественное: $f \times g(x) = g \times f(x) = x$. Преобразования f и g в этом случае называются взаимнообратными. Если A и B — соответственно матрицы взаимно обратных преобразований f и g, то $AB = BA = E$, следовательно, $B = A^{-1}$. Следовательно, для любого невырожденного преобразования с матрицей А существует единственное обратное с матрицей A^{-1}.

5. Ортогональные матрицы

Пусть имеется евклидово n-мерное пространство E_n (линейное пространство, в котором определено скалярное произведение). Базис $e_1, e_2, \ldots, e_n$ такого пространства называется **ортонормированным**, если базисные векторы образуют ортонормированную систему. Система векторов является ортонормированной, если она ортогональна (векторы системы попарно ортогональны, т. е. их скалярные произведения равны нулю) и норма каждого вектора равна единице (норма вектора — это арифметическое значение корня из скалярного квадрата этого вектора), т. е. имеем для данного базиса:

$$(e_i, e_k) = \begin{cases} 0, & \text{при } i \neq k, \\ 1, & \text{при } i = k. \end{cases}$$

Матрица ортонормированной системы векторов $a_1(a_{11}, a_{21}, ..., an_1)$, $a_2(a_{12}, a_{22}, .., an_2), ..., an(a_{1n}, a_{2\mathbf{n}}, ..., a_{nn})$ называется **ортогональной**. Для таких ортонормированных векторов имеем:

$$\sum_{k=1}^{n} a_{ki} a_{kj} = \begin{cases} 0, & \text{при } i \neq j, \\ 1, & \text{при } i = j. \end{cases}$$

Единичные матрицы ортогональны.

Например, ортогональными являются следующие единичные матрицы:

$$\begin{matrix} 0{,}8 & -0{,}6 \\ -0{,}6 & 0{,}8 \end{matrix}, \quad \begin{matrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{matrix}.$$

Теорема (необходимое и достаточное условие ортогональности). Для того чтобы матрица A была ортогональной, необходимо и достаточно, чтобы выполнялось равенство $A^T A = E$ (A^T — транспонированная матрица). Действительно, если обозначим $B = A^T A$, то элементы этой матрицы В будут равны:

$$b_{ij} = \sum_{k=1}^{n} a^{\bullet}_{ik} a_{kj} = \sum_{k=1}^{n} a_{ki} a_{kj} = \begin{cases} 0, \, i \neq j, \\ 1, \, i = j, \end{cases}$$

(a_{ik}— элементы транспонированной матрицы A^T). Но это означает, что $B = E$ или $A^{\mathrm{т}} A = E$. И обратно, если $A^T A = E$, имеем равенство

$$\sum_{k=1}^{n} a^{\bullet}_{ik} a_{kj} = \sum_{k=1}^{n} a_{ki} a_{kj} = \begin{cases} 0, \, i \neq j, \\ 1, \, i = j, \end{cases}$$

что означает ортогональность матрицы A.

Следствия.

1. Модуль определителя ортогональной матрицы равен единице.
2. Ортогональная матрица — невырожденная.
3. Произведение двух ортогональных матриц — ортогональная матрица.
4. Необходимым и достаточным условием ортогональности матрицы A является равенство $A^T = A^{-1}$.

5. При транспонировании ортогональной матрицы получается ортогональная матрица.

6. Матрица, обратная ортогональной, тоже ортогональна.

Но сумма ортогональных матриц не является ортогональной матрицей.

Теорема. Матрица перехода T от одного ортонормированного базиса к другому является ортогональной.

Действительно, если $e_1, e_2, ..., e_n$ и $e^*_1, e^*_2, ..., e^*_n$ — ортонормированные базисы, то $(e^*_i, e^*_k) = \begin{cases} 0, & \text{при } i \neq k \\ 1, & \text{при } i = k \end{cases}$ и $\left(e^*_i, e^*_j\right) = \sum_{k=1}^{n} t_{ki} t_{kj}$ (t_{ij} — элементы матрицы T перехода от первого базиса ко второму), а это означает ортогональность матрицы T.

Ортогональное преобразование — это линейное преобразование евклидова пространства, матрица которого ортогональна в некотором ортонормированном базисе.

Теорема. Линейное преобразование евклидова пространства является ортогональным тогда и только тогда, когда оно переводит ортонормированный базис в ортонормированный.

Действительно, по предыдущей теореме матрица перехода от одного ортонормированного базиса к другому является ортогональной, следовательно, линейное преобразование, соответствующее данной матрице, ортогонально. И обратно, если имеется линейное преобразование в некотором ортонормированном базисе с ортогональной матрицей, то из ортогональности следует, что $(e^*_i, e^*_k) = \begin{cases} 0, & при\ i \neq k \\ 1, & при\ i = k \end{cases}$, где каждый из векторов второго базиса $e^*_1, e^*_2, ..., e^*_n$ равен $e^*_k = f(e_k) = a_{1k}e_1 + a_{2k}e_2 + ... + a_{nk}e_k$ ($k = 1, 2, ..., n$), коэффициенты этого разложения составляют k-ый столбец ортогональной матрицы перехода. Отсюда следует ортонормированность базиса $e^*_1, e^*_2, ..., e^*_n$.

Известно, что ортогональное преобразование не меняет скалярного произведения векторов (следует из выражения скалярного произведения через координаты векторов в ортонормированном базисе), а следовательно, не меняются норма вектора и угол между двумя векторами.

Ортогональные преобразования евклидова пространства с матрицей A невырожденные, имеют обратное ортогональное преобразование с матрицей, равной $A^{\text{т}}$, и их произведение тоже является ортогональным преобразованием.

ЧАСТЬ II
ПРИЛОЖЕНИЯ
ЛИНЕЙНОЙ АЛГЕБРЫ

ЛЕКЦИЯ № 6. Квадратичные формы

1. Квадратичные формы

Квадратичная форма n переменных $x_1, x_2, ..., x_n$ — это сумма произведений $x_i x_j$ с соответствующими коэффициентами, обозначается $f(x_1, x_2, ..., x_n)$ и равна $f(x_1, x_2, ..., x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j$, числа a_{ij} называются коэффициентами квадратичной формы (могут быть и действительными, и комплексными). Матрица A, составленная из коэффициентов a_{ij}, называется **матрицей квадратичной формы**, а ее ранг — **рангом квадратичной формы**:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Матрица A квадратичной формы является симметрической, так как ее элементы, симметричные относительно главной диагонали, равны между собой потому, что произведение чисел $x_i x_j = x_j x_i$, а равенство $a_{ij} = a_{ji}$ следует из следующих рассуждений: $a_{ij} x_i x_j + a_{ji} x_j x_i = (a_{ij} + a_{ji}) x_i x_j = S\,(a_{ij} + a_{ji}) x_i x_j + S\,(a_{ij} + a_{ji}) x_j x_i = a_{ij}' x_i x_j + a_{ji}' x_j x_i$, и для нее справедливо $A = A^{\mathrm{т}}$. Всякой симметрической матрице соответствует единственная квадратичная форма. В матричном виде квадратичная форма равна $f(x_1, x_2, ..., x_n) = X^T A T$, где X — матрица-столбец переменных $x_1, x_2, ..., x_n$, $X^{\mathrm{т}}$ — матрица-строка переменных $x_1, x_2, ..., x_n$.

Например, квадратичной форме $f(x_1, x_2, x_3) = x_1^2 + 2x_1x_2 + 4x_1x_3 - 2x_2^2 + 5x_2x_3 = 0$ соответствует квадратичная матрица

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & -2 & 5 \\ 4 & 5 & 0 \end{pmatrix}.$$

Канонической называется квадратичная форма, не содержащая произведений различных переменных, т. е.

$$f(x_1, x_2, ..., x_n) = \sum_{i=1}^{r} a_{ii} x_i^2 \quad (r \le n).$$

Например, каноническая квадратичная форма $f(x_1, x_2, ..., x_n) = 3x_1^2 + 2x_2^2 - 4x_3^2$. Если коэффициенты канонической квадратичной формы равны ± 1, то такая форма называется **нормальной**, например нормальная квадратичная форма $f(x_1, x_2,...,x_n) = x_1^2 - x_2^2 + x_3^2$.

Пусть переменные $x_1, x_2, ..., x_n$ являются линейной комбинацией новых переменных $y_1, y_2, ..., y_n$, т. е. имеет место равенства: $x_i = b_{i1}y_1 + b_{i2}y_2 + ... + b_{in}y_n$ $(i = 1, 2,..., n)$ или в матричной форме: $X = BY$ (X, Y — матрицы-столбцы переменных). Данное линейное преобразование называется невырожденным, если определитель матрицы B не равен нулю. Подставив в выражение $f(x_1, x_2, ..., x_n) = X^T AT$ значения $x_i = b_{i1}y_1 + b_{i2}y_2 + ... + b_{in}y_n$, получим новую квадратичную форму $g(y_1, y_2, ..., y_n)$ со своей матрицей C. Следовательно, квадратичная форма $f(x_1, x_2, ..., x_n)$ перешла в квадратичную форму $g(y_1,y_2, ..., y_n)$.

Конгруэнтными называются такие две квадратичные формы, для которых существует невырожденное линейное однородное преобразование ($X = BY$), переводящее одну форму в другую; обозначается $f(x_1, x_2, ..., x_n) \sim g(y_1, y_2, ..., y_n)$.

Справедливо:

1) если $f(x_1, x_2, ..., x_n) \sim g(y_1,y_2, ..., y_n)$, а $g(y_1, y_2, ..., y_n) \sim q(z_1, z_2, ..., z_n)$, то $f(x_1, x_2, ..., x_n) \sim q(z_1, z_2, ..., z_n)$ (так как произведение невырожденных линейных преобразований — невырожденное);

2) $f(x_1, x_2, ...,x_n) \sim f(x_1, x_2, ..., x_n)$ (так как тождественное преобразование невырожденное).

Теорема: Квадратичная форма $f(x_1, x_2, ..., x_n)$ с матрицей A линейным однородным преобразованием $X = BY$ переводится в квадратичную форму $g(y_1,y_2,...,y_n)$, имеющую матрицу, равную $C = B^T AB$.

Данное утверждение доказывается следующими преобразованиями: $X^T AX = (BY)TA(BY) = (YTBT)ABY = YT(BTAB)Y$, т. е. $C = BTAB$, невырожденность следует из: $CT = (BTAB)T = (BT(AB))T = (AB)T(BT)T = (BTAT)B = BTAB = C$.

Следствия

1. Определители матриц конгруэнтных невырожденных действительных квадратичных форм имеют одинаковые знаки.

2. Конгруэнтные квадратичные формы имеют одинаковые ранги.

Теорема. Любая квадратичная форма некоторым невырожденным преобразованием может быть приведена к каноническому виду (доказывается методом математической индукции по числу переменных).

Теорема. Любую действительную квадратичную форму линейным невырожденным преобразованием можно привести к нормальному виду (доказывается на основе предыдущей теоремы).

Приведение действительной квадратичной формы к каноническому виду подчиняется следующим теоремам.

Теорема. Если существует ортогональное преобразование с матрицей C, приводящее действительную квадратичную форму $f = f(x_1, x_2, ..., x_n)$ к каноническому виду $g(y_1, y_2, ..., y_n) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + ... + \lambda_n y_n^2$, то λ_i $(i = 1, ..., n)$ — то характеристические числа матрицы A квадратичной формы f, причем столбцами матрицы C являются собственные векторы-столбцы матрицы А с собственными числами λ_i.

Теорема. Для любой действительной квадратичной формы существует ортогональное преобразование, приводящее ее к каноническому виду (доказывается методом математической индукции).

Теорема. Для любой действительной симметрической матрицы A существует такая ортогональная матрица T, что $T^{-1}AT$ — диагональная матрица.

Следствие. Любая действительная симметрическая матрица может быть приведена к диагональному виду.

Теорема. Если линейное преобразование действительного линейного пространства имеет действительную симметрическую матрицу в некотором ортонормированном базисе, то существует ортонормированный базис, состоящий из собственных векторов этого преобразования.

2. Закон инерции квадратичных форм

Закон инерции квадратичных форм говорит о том, что квадратичная форма с точностью до обозначения переменных и нумерации приводится только к одному нормальному виду.

Положительным (отрицательным) индексом инерции квадратичной формы называется число положительных (отрицательных) квадратов в нормальной форме. **Сигнатурой** квадратичной формы называется разность между положительным и отрицательным индексами инерции этой формы.

Теорема (закон инерции квадратичных форм). Число положительных и число отрицательных квадратов в нормальном виде, к которому приводится данная квадратичная форма невырожденным действительным линейным преобразованием, не зависит от выбора преобразования.

Доказательство от противного: пусть имеется квадратичная форма $f(x_1, x_2, ..., x_n)$, которая двумя способами приведена к нормальному виду. Пусть при первом способе приведения число положительных квадратов в нормальном виде равно k, при втором способе приведения — равно m и пусть (для определенности) $k < m$ (не равны). Следовательно, для данной квадратичной формы $f(x_1, x_2, ..., x_n)$ имеем два нормальных вида и выполняется равенство: $y_1^2 + ... + y_k^2 - y_{k+1}^2 - ... - y_r^2 = z_1^2 + ... + z_m^2 - z_{m+1}^2 - ... - z_r^2$ (где r — ранг квадратичной формы). Для действительных линейных невырожденных преобразований в данном случае справедливо:

$$y_i = \sum_{s=1}^{n} b_{is} x_s, \quad z_j = \sum_{i=1}^{n} c_{ji} x_i \ (i, j = 1, 2, ..., n).$$

Предположим, что в этих равенствах переменные $y_1, y_2, ..., y_k$, $z_{m+1}, z_{m+2}, ..., z_r, ..., z_n = 0$ равны нулю, тогда приходим к однородной системе $(n - m + k)$ уравнений с n неизвестными (уравнения с не равными нулю переменными в эту систему не включаются). Причем количество данных уравнений будет меньше n (поскольку $k < m$), а, следовательно, данная система имеет ненулевое решение $\alpha(\alpha_1, \alpha_2, ..., \alpha_n)$. Подставив это решение в выражение $y_1^2 + ... + y_k^2 - y_{k+1}^2 - ... - y_r^2 = z_1^2 + ... + z_m^2 - z_{m+1}^2 - ... - z_r^2$ и учитывая равенство части переменных нулю, получаем: $-y_{k+1}^2(\alpha) - ... - y_r^2(\alpha) = z_1^2(\alpha) + ... + z_m^2(\alpha)$. Для действительных чисел такое равенство возможно лишь при $z_1^2(\alpha) = ... = z_m^2(\alpha) = 0$. Добавив к этим равенствам равенства (то, что предположили) $z_{m+1}, z_{m+2}, ..., z_r, ..., z_n = 0$, получаем систему n линейных однородных уравнений, имеющую ненулевое решение. По теореме Крамера определитель такой системы равен нулю, что противоречит невырожденности преобразования:

$$z_j = \sum_{i=1}^{n} c_{ji} x_i \ (i, j = 1, 2, ..., n).$$

Значит, предположение $k < m$ (аналогично $m < k$) неверно, т. е. $k = m$. Теорема доказана.

Теорема. Две действительные квадратичные формы от n переменных тогда и только тогда конгруэнтны, когда они имеют одинаковые ранги и одинаковые сигнатуры. Действительно, перевод одной квадратичной формы в другую невырожденным действительным линейным преобразованием не меняет ранга, а следовательно и сигнатуры, иначе формы приводились бы к различным нормальным видам, что противоречит предыдущей теореме о законе инерции квадратичных форм.

3. Знакоопределенные квадратичные формы

Положительно (отрицательно) определенной квадратичной формой называется действительная квадратичная форма, если она невырожденная и приводится к нормальному виду, содержащему только положительные (отрицательные) квадраты всех переменных. Положительно и отрицательно определенные квадратичные формы называются **знакоопределенными**. **Полуопределенными** называются вырожденные квадратичные формы, приводящиеся к нормальному виду из квадратов одного знака. Если квадратичная форма приводится к нормальному виду, содержащему и положительные, и отрицательные квадраты переменных, то такая форма называется **неопределенной**.

Теорема. Действительная квадратичная форма $f(x_1, x_2,...,x_n)$ является положительно определенной тогда и только тогда, когда она принимает положительные значения при любой ненулевой системе значений переменных $x_1, x_2, ..., x_n$. Действительно, если квадратичная форма $f(x_1, x_2, ..., x_n)$ является положительно определенной, то она приводится к нормальному виду преобразованием

$$y_i = \sum_{j}^{n} b_{ij} x_j \ \ (j = 1, ..., n)$$

с ненулевым определителем. Тогда для не равной нулю системы $x_1, x_2, ..., x_n$ получаем не равные нулю $y_1, y_2, ..., y_n$ (иначе была бы однородная система с не равным нулю определителем и ненулевым решением, что невозможно). Следовательно, $y_1^2 + y_2^2 + ... + y_n^2 > 0$. И обратно, если квадратичная форма не является положительно определенной, то ее нормальный вид может быть таким: $y_1^2 + y_2^2 + {} + ... y_{n-1}^2 - y_n^2$ или $y_1^2 + y_2^2 +...+ y_{n-1}^2$. Следовательно, не равной нулю системе $x_1, x_2, .., x_n$ найдется ненулевая система $y_1, y_2, ..., y_n$, для которой нормальный вид будет меньше или равен нулю (≤ 0).

Следующая теорема позволяет по элементам матрицы квадратичной формы узнать, положительно или отрицательно определенная эта форма.

Теорема.

1. Квадратичная форма $f(x_1, x_2, ..., x_n)$ с действительной матрицей является положительно определенной тогда и только тогда, когда все ее главные миноры положительны (главными минорами матрицы называются миноры, элементы которых расположены в левом верхнем углу этой матрицы). Доказывается методом математической индукции.

2. Квадратичная форма является отрицательно определенной тогда и только тогда, когда ее главные миноры четного порядка положительны, а нечетного — отрицательны (необходимость следует из вида матрицы).

4. Упрощение уравнений второго порядка с помощью квадратичных форм

Уравнение второго порядка в общем случае имеет вид $a_{11}x_2 + 2a_{12}xy + a_{22}y_2 + a_{13}x + a_{23}y + a_{33} = 0$. Данное уравнение соответствует фигурам второго порядка на плоскости. Выражение $f(x, y) = a_{11}x_2 + 2a_{12}xy + a_{22}y_2$ можно рассматривать как квадратичную форму с матрицей $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Матрица А — симметрическая. С помощью ортогонального преобразования квадратичная форма $f(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2$ приводится к каноническому виду $(\lambda_1 x^{*2} + \lambda_2 y^{*2})$, где λ_1, λ_2 являются корнями (характеристическими числами матрицы A) характеристического уравнения $(a_{11} - \lambda)(a^{22} - \lambda) - a_{12}^2 = 0$. Следовательно, исходное уравнение второго порядка ортогональным преобразованием приводится к следующему виду $\lambda_1 x^{*2} + \lambda_2 y^{*2} + a_{13}{}^*x^* + a_{23}{}^*y^* + a_{33}{}^* = 0$.

Далее при выделении полных квадратов получается выражение типа $\lambda_1(x^* - d_1)^2 + \lambda_2(x^* - d_2)^2 = q$.

Если ввести новые переменные $X = x^* - d_1$, $Y = x^* - d_2$ (что соответствует параллельному переносу осей координат) и обозначить $a^2 = 1 / \lambda_1$, $b^2 = 1 / \lambda_2$, то исходное уравнение принимает вид $X^2 / a^2 \pm Y^2 / b^2 = q$. Знак в этом выражении совпадает со знаком произведения $\lambda_1\lambda_2$ (если произведение положительно, то берется знак плюс; если отрицательно — знак минус), q принимает значения 0, 1, –1. Получаемые при этом выражения в зависимости от знака и величины q соответствуют различным фигурам (эллипсу, мнимому эллипсу, точке, гиперболе, паре пересекающихся прямых).

Если произведение $\lambda_1\lambda_2 = 0$ (одно из характеристических чисел матрицы A равно нулю, оба одновременно не могут быть равны нулю, поскольку квадратичная форма невырожденная)), то в случае $a_{23}^* \neq 0$ и $\lambda_1 \neq 0$ исходное уравнение приводится к виду $X^2 = 2px$ (что соответствует параболе), в случае $a_{23}^* = 0$, $\lambda_1 \neq 0$ исходное уравнение приводится к виду $X^2 = \pm\, a^2$ или $X^2 = 0$ (что соответствует паре действительных, или мнимых, или совпадающих прямых).

Переход уравнения фигуры второго порядка $a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + a_{13}x + a_{23}y + a_{33} = 0$ при ортогональном преобразовании к одному из перечисленных видов (к каноническому виду) называется **отнесением фигуры к главным осям, главными направлениями фигуры** называются направления ортогональных собственных векторов матрицы A квадратичной формы. Фигура второго порядка называется **центральной**, если определитель матрицы A не равен нулю, если этот определитель равен нулю, то фигура называется **нецентральной**.

Фигурам второго порядка в пространстве в декартовых координатах соответствует уравнение вида $a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + a_{33}z^2 + 2a_{13}xz + 2a_{23}yz + a_{14}x + a_{24}y + a_{34}z + a_{44} = 0$. Здесь квадратичная форма от трех переменных и имеет вид: $a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + a_{33}z^2 + 2a_{13}xz + 2a_{23}yz = 0$, причем сумма квадратов ее коэффициентов не равна нулю $a_{11}{}^2 + a_{12}{}^2 + a_{22}{}^2 + a_{33}{}^2 + a_{13}{}^2 + a_{23}{}^2 \neq 0$. Матрица квадратичной формы также симметрическая и имеет размерность 3×3. Аналогичными ортогональными преобразованиями пространственное уравнение приводится к каноническому виду, который соответствует в зависимости от характеристических чисел матрицы A различным центральным (в случае, когда определитель матрицы A не равен нулю) или нецентральным (определитель матрицы A равен нулю) пространственным фигурам.

ЛЕКЦИЯ № 7. Группы и подгруппы

1. Группы и подгруппы (определения)

Существуют различные множества, над элементами которых определены операции сложения и умножения, обладающие свойством ассоциативности, и в которых существуют нейтральные для данных операций элементы (нуль — для сложения и единица — для умножения), для каждого элемента существует обратный (противоположный) элемент (например, множество действительных чисел, множество целых чисел, множество векторов, множество невырожденных квадратных матриц, множество линейных преобразований). Такие множества называются **группой**. Следовательно, если некоторое множество G является группой, то в нем определены операции сложения или умножения так, что при этом каждой паре элементов соответствует элемент из этого же множества: $a, b, c \in G, c = a \times b$; эти операции ассоциативны $a \times (b \times c) = (a \times b) \times c$; в G имеется нейтральный элемент е: $a \times e = e \times a = a$; для каждого элемента a имеется обратный элемент a^{-1}: $a \times a^{-1} = e$.

Группа G, в которой для любых $a, b \in G$ выполняется условие $a \times b = b \times a$, называется **абелевой (коммутативной)** группой. Если число элементов группы конечное, то и группа называется **конечной**, а количество элементов называется **порядком группы**. Бесконечная группа имеет бесконечный порядок. Множество всех элементов группы называется **основным множеством группы**. **Мощностью группы** G называется число элементов группы, обозначается: $|G|$, $|\mathrm{Card}\ G|$.

Если операция, определенная в группе, — сложение, то группа называется аддитивной, нейтральный элемент — нуль. Если в группе определена операция умножения, то группа называется **мультипликативной**, нейтральный элемент — единица, в этой группе однозначно разрешимы уравнения вида $ax = b$. Например, аддитивной группой является множество целых чисел с операцией сложения, но это множество не является мультипликативной группой, поскольку обратные числа не являются целыми (например, $3^{-1} = {}^{1}/_{3}$ — дробь). К аддитивным группам относится также множество из одного элемента 0, множество векторов трехмерного пространства, множество матриц $m \times n$. Множество действительных чисел без нуля является мультипликативной группой, а с нулем — нет, так как для нуля нет обратного числа. К мультипликативным группам относится также множество из одного элемента 1, множество из двух элементов 1 и -1, множество невырожденных квадратных матриц.

Пусть имеется группа G, если ее подмножество H является группой относительно определенной в группе G операции, то говорят, что H является **подгруппой** группы G. Следовательно, и в группе, и в ее подгруппе определена одна и та же операция.

Например, к аддитивным подгруппам аддитивной группы действительных чисел относится множество целых чисел, множество рациональных чисел. К мультипликативным подгруппам мультипликативной группы действительных, не равных нулю чисел относится множество положительных действительных чисел, множество не равных нулю рациональных чисел. К мультипликативным подгруппам мультипликативной группы невырожденных квадратных матриц относится множество ортогональных матриц, множество диагональных матриц, множество матриц с положительным определителем.

Пересечение двух подгрупп некоторой группы тоже является подгруппой этой группы. Сама группа является своей подгруппой. Объединение двух подгрупп является подгруппой только в случае, когда одна из этих подгрупп содержит другую. Объединение возрастающей цепи подгрупп (предыдущая — подгруппа последующей) является подгруппой. **Тривиальной (несобственной, порядка 1)** подгруппой группы G является нейтральный элемент этой группы G, остальные подгруппы группы G — истинные (собственные). Подгруппа подгруппы группы G является и подгруппой группы G. Например, аддитивная группа целых чисел (подгруппа подгруппы) — подгруппа группы рациональных чисел (подгруппа) — подгруппа группы действительных чисел (группа).

Определение группы может быть дано в других терминах (с помощью понятий моноида, бинарной алгебраической операции). Пусть имеется некоторое множество X. **Бинарной алгебраической операцией** называется произвольное отображение $X \times X \to X$ (любой упорядоченной паре элементов множества X ставится в соответствие однозначно определенный элемент этого же множества). **Моноидом** называется множество X с заданной на нем бинарной ассоциативной операцией и с единичным (нейтральным) элементом. Тогда группа определяется как моноид G, все элементы которого обратимы (имеют обратные элементы в этой же группе), и справедливо, что на множестве G определена ассоциативная бинарная операция, существует нейтральный элемент и для каждого элемента существует обратный.

Полной линейной группой является мультипликативная группа невырожденных матриц порядка n (множество, рассматриваемое вместе с бинарной операцией, степени n).

2. Циклические группы. Симметрические группы

Циклическая группа — это группа, образованная степенями одного из своих элементов (порождена одним из своих элементов). Пусть G — некоторая группа, a — элемент этой группы. Если при $n \neq m$ не равны степени $a^m \neq a^n$, то говорят, что элемент a группы G имеет бесконечный порядок. Если $am = an$ при $m \neq n$ (например, $m > n$), т. е. имеются такие m и n, что $am - n = e$ (e — единичный элемент), то число q, равное минимальному числу, для которого выполняется равенство $a^q = e$, называется конечным порядком элемента a (или a — элемент конечного порядка q). Если группа G — конечная, то все ее элементы имеют конечный порядок.

Простейшей аддитивной циклической группой является множество целых чисел, порожденное единицей (порождено одним своим элементом — 1, каждое число представимо в виде суммы единиц, например $2 = 1 + 1$, $3 = (1 + 1) + 1$ и т. д.). Данная группа — бесконечная.

Примером циклической группы порядка n является множество поворотов правильного n-угольника на плоскости вокруг его центра O (каждый поворот совмещает многоугольник сам с собой). Каждый поворот, при котором многоугольник совмещается сам с собой, происходит на угол, равный $2\pi/n$. Следовательно, $\varphi_{n-1} = 2\pi(n - 1) / n$, здесь $\varphi 0 = 0$. Данная группа будет являться мультипликативной, если под умножением понимать последовательное выполнение поворотов одного за другим. Определенное так умножение подчиняется ассоциативному закону, коммутативно, единичным элементом является поворот на 0°, обратный элемент — поворот в обратную сторону. Для правильного треугольника порядок такой группы поворотов равен 3, для квадрата он равен 4.

Справедливо утверждение: перестановочные элементы a, b произвольной группы G, имеющие взаимно-простые порядки s и t, порождают в группе G циклическую подгруппу порядка st.

Симметрической группой называется множество взаимнооднозначных отображений множества $X \to X$ (преобразований множества X). Если X — конечное множества из n элементов, то данная симметрическая группа будет иметь порядок, равный n, а взаимнооднозначные отображения называются **подстановками** (иногда называют перестановками). Например, элементами множества X являются числа 1, 2, 3, 4, 5, оно отображается в множество 3, 5, 1, 2, 4. Обозначается данная подстановка (перестановка) следующим образом:

$$\begin{matrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 2 & 4 \end{matrix}$$

или в общем виде:

$$\begin{matrix} 1 & 2 & 3 & 4 & 5 \\ a_1 & a_2 & a_3 & a_4 & a_5 \end{matrix},$$

где a_1, a_2, a_3, a_4, a_5 — записанные в произвольном порядке числа 1, 2, 3, 4, 5.

Порядок записи столбцов не играет роли, т. е. две подстановки, отличающиеся порядком записи столбцов, считаются равными:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 4 & 1 & 5 \\ 1 & 5 & 2 & 4 & 3 \end{pmatrix}.$$

Подстановка, в которой каждый элемент соответствует сам себе, называется тождественной, например:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}.$$

Умножение подстановок определяется как последовательное их выполнения, начиная с правого сомножителя, например:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 2 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 4 & 5 \end{pmatrix}.$$

Определенное так умножение некоммутативно, т. е. если A и B — две подстановки, то $AB \neq BA$. Роль единицы играет тождественная подстановка. Обратная подстановка получается при перемене строк местами:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ a_1 & a_2 & a_3 & a_4 & a_5 \end{pmatrix}^{-1} = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}.$$

Число подстановок для множества X из n элементов равно $n!$, т. е. порядок группы подстановок равен $n!$.

3. Морфизмы групп

Группы G^1 и G^2 **называются изоморфными**, если между их элементами можно установить взаимно однозначное соответствие, сохраняющее групповую операцию, т. е. если

$$x_1, y_1 \in G_1, x_1, y_1 \in G_2 \text{ и } x_1 \Leftrightarrow x_2, y_1 \Leftrightarrow y_2, \text{ то } x_1 \times y_1 \Leftrightarrow x_2 \times y_2.$$

Для мультипликативных групп данное соответствие рассматривается как взаимно-однозначное отображение одной первой группы на вторую. **Абстрактной группой** называется множество групп, изоморфных данной.

Пусть имеется правильный треугольник. Сам в себя он переводится не только поворотом вокруг центра, но и поворотами вокруг осей симметрии, которыми являются медианы (биссектрисы, высоты) этого правильного треугольника. Следовательно, имеется шесть преобразований (перестановок вершин треугольника). Симметрическая группа подстановок (перестановок) из трех элементов и группа симметрий (преобразований) правильного треугольника изоморфны, с точки зрения теории групп они не различаются между собой. Операции в изоморфных группах могут называться по-разному. Циклическая группа порядка n изоморфна группе вращений правильного n-угольника, бесконечная циклическая группа изоморфна аддитивной группе целых чисел.

Простейшие свойства изоморфизма: если f — отображение группы G_1 в группу G_2, то при изоморфизме:

1) единица переходит в единицу;
2) $f(a^{-1}) = f(a)^{-1}$;
3) обратное отображение f^{-1} группы G_2 в G_1 тоже является изоморфным;
4) все циклические группы одного и того же порядка (в том числе и бесконечного) изоморфны;
5) любая конечная группа порядка n изоморфна некоторой подгруппе симметрической группы порядка n (теорема Кэли).

Автоморфизмом группы G называется изоморфное отображение этой группы самой на себя: $G \to G$.

Гомоморфизмом (эндоморфизмом) называется изоморфное отображение, не обладающее свойством биективности (взаимной однозначности). Например, аддитивная группа целых чисел гомоморфно отображается на конечную циклическую группу, образованную элементом g, порядка q при отображении $n \to g^n$; полная линейная группа вещественных матриц A с отличным от нуля определителем гомоморфно отображается на мультипликативную группу не равных нулю вещественных чисел при равенстве $f = \det A$. Иначе определение гомоморфизма можно сформулировать следующим образом: если G_1 и G_2 — группы и $f: G_1 \to G_2$ — такое отображение, при котором для любых элементов x, y группы G $f(xy) = f(x)f(y)$, то f называется гомоморфным отображением (или гомоморфизмом) группы G_1 в группу G_2. Изоморфизм является частным случаем гомоморфизма. Каждая группа гомоморфна самой себе, гомоморфна единичной группе. Гомоморфизм сохраняет групповую операцию, нейтральный и обратный элементы. **Ядром гомоморфизма** f называется множество $\mathrm{Ker} f = \{g G_1 \mid f(g) = e^*$ — единица группы $G_2\}$ (иначе, ядро гомоморфизма — это полный прообраз единичной подгруппы при гомоморфизме).

4. Разложение группы по подгруппе

Пусть даны некоторая группа G (элемент x $\in G$) и ее подгруппа H (элемент $h \in H$). Множество элементов $x \times h$ при фиксированном x называется **левым смежным классом** группы G по подгруппе H, порожденной элементом x, обозначается $x \times$ H; множество $H \times x$ элементов $h \times x$ называется **правым смежным классом** группы G по подгруппе H, порожденным элементом x. Сам элемент x принадлежит смежному классу, поскольку $x \times e = x$, где e — нейтральный элемент подгруппы H. Всякий смежный класс порождается любым своим элементом, т. е. если y содержится в левом смежном классе, то справедливо: $y \times H = x \times H$. Два любых смежных класса одной группы по одной и той же подгруппе либо совпадают, либо не пересекаются (не имеют ни одного общего элемента). Следовательно, группа распадается на непересекающиеся левые (или правые) смежные группы по одной и той же подгруппе H. Иначе говоря, имеется левостороннее (или правостороннее) разложение группы G по подгруппе H, причем сама подгруппа H является одним из смежных классов разложения. Для абелевой группы левостороннее разложение совпадает с правосторонним разложением.

Теорема Лагранжа. Порядок подгруппы конечной группы является делителем порядка группы.

Действительно, если n — порядок группы G, k — порядок ее подгруппы H, m — количество левых смежных классов разложения группы по подгруппе H, то каждый смежный класс состоит из k элементов, тогда количество элементов группы равно $n = km$, следовательно, n делится на k. Число m называется **индексом подгруппы** H в группе G.

Следствия.

1. Порядок любого элемента конечной группы будет делителем порядка группы.
2. Всякая конечная группа, порядок которой есть простое число, является циклической.

Если правостороннее разложение группы G по подгруппе H совпадает с левосторонним, то подгруппа H называется **нормальным делителем** (инвариантной подгруппой, нормальной подгруппой) группы G. Нормальными делителями являются единичная подгруппа группы G и сама группа G, любая подгруппа коммутативной группы — нормальный делитель. В группе невырожденных квадратных матриц нормальным делителем является подгруппа матриц, определитель которых равен единице. Подгруппа Н является нормальным делителем группы G, если для всех $x \in G$ справедливо равенство: $x \times H = H \times x$.

Два элемента a, $b \in G$ называются **сопряженными**, если в группе G найдется хотя бы один такой элемент x, что выполняется равенство: $b = x^{-1}ax$ (элемент b получатся из элемента а транспонированием с помощью элемента x).

Теорема. Подгруппа H группы G тогда и только тогда будет нормальным делителем группы G, если вместе с любым своим элементом h она содержит все элементы, сопряженные с ним в группе G.

Теорема. Пересечение двух нормальных делителей группы G является нормальным делителем этой группы.

Фактор-группой группы G по нормальному делителю H называется группа всех смежных классов группы G по подгруппе H.

Теорема о гомоморфизмах. Каждая группа гомоморфна любой своей фактор-группе. Обратно, если группа G гомоморфна группе G^*, то G^* изоморфна фактор-группе G по некоторому нормальному делителю H.

ЛЕКЦИЯ № 8. Кольца и поля

1. Кольца

Кольцом называется абелева группа, в которой определено умножение, связанное со сложением дистрибутивными законами умножение дистрибутивно по сложению): $(a + b)c = ac + bc$, $a(b + c) = ab + ac$. Если K — кольцо, т. е. непустое множество с заданными на нем бинарными операциями сложения и умножения, то оно обозначается $(K+)$. Например, кольцами является множество целых чисел, множество действительных квадратных матриц, множество трехмерных векторов.

Структура $(K+)$ называется **аддитивной группой**, (K) — **мультипликативной полугруппой** (мультипликативным группоидом). Единичный элемент обычно обозначается единицей. Одноэлементное кольцо называется нулевым.

Подмножество S кольца K называется подкольцом, если $0 \in S$ и $a + b$, $-a$, $ab \in S$ всякий раз, когда $a, b \in S$. Например, множество всех четных чисел является подкольцом кольца целых чисел. Тривиальными подкольцами являются одноэлементное множество и само кольцо K. Каждое подкольцо является кольцом относительно заданных в кольце K операций. Пересечение любого семейства подколец в K также является подкольцом. Наименьшее подкольцо, содержащее множество X, являющееся подмножеством K, совпадает с множеством всех сумм всевозможных произведений элементов из X. В этом случае говорят, что оно порождается множеством X.

Кольцо K называется **ассоциативным**, если для любых $a, b, c \in K$ справедливо $(a \times b)c = a(b \times c)$, если имеет место равенство $ab = ba$, то кольцо называется **коммутативным**. Если выполняются оба равенства, то кольцо называется **ассоциативно-коммутативным**. **Левой (правой) единицей** кольца K называется такой элемент e, для которого $e \times a = a(a \times e = a)$. Если элемент е является одновременно и левой, и правой единицей, т. е. $e \times a = a \times e = a$, то он называется просто единицей. Единица в кольце только одна. В нулевом кольце единица совпадает с нулем. Для элементов коммутативного кольца справедлива биноминальная формула (бином Ньютона).

Линейной алгеброй (алгеброй) называется кольцо K над ассоциативно-коммутативным кольцом Φ с единицей, если для любых a, $b \in K$ и $\lambda \in \Phi$ выполняется равенство $(\lambda ab) = (\lambda a)b = a(\lambda b)$. В случае, когда Φ является полем, K будет линейным пространством над этим полем. Любое кольцо является алгеброй над кольцом целых чисел. Алгебраическим называется такой элемент а линейной алгебры над полем, для которого выполняется равенство $f(a) = 0$, где $f(x)$ — ненулевой многочлен над тем же полем.

Определения морфизмов для колец аналогично определению для групп. Отображение φ кольца K в кольцо S называется **гомоморфизмом**, если выполняются равенства $\varphi(a + b) = \varphi(a) + \varphi(b)$ и $\varphi(ab) = \varphi(a)\varphi(b)$ для любых a, b, $c \in K$. **Изоморфизмом** называется взаимно-однозначный гомоморфизм. Два кольца называются изоморфными, если отображения одного в другое и наоборот изоморфные. **Эндоморфизмом** называется гомоморфизм кольца в себя, изоморфизм кольца на себя называется **автоморфизмом**. **Ядром гомоморфизма φ** называется множество $\mathrm{Rer}\,\varphi = \{a \in \in K \,|\, \varphi(a) = = 0^*\}$. Ядро является подкольцом в K.

Подмножество P кольца K называется двусторонним идеалом (идеалом), если $0 \in P$, а принадлежность элементов a, $b \in P$ влечет принадлежность $(a - b) \in P$ и принадлежность ar, $ra \in P$ для любых $r \in K$. Идеал кольца является его подкольцом и является идеалом для любого подкольца данного кольца.

Характеристикой кольца K называется натуральное число n, для которого $na = 0$ при любых $a \in K$ и $kb \neq 0$ $0 \neq b \in K$, $0 \neq k \in Z$ и $k < n$. В случае, когда $na \neq 0$ при любых $0 \neq a \in K$ и $0 \neq n \in Z$, характеристика кольца равна нулю (Z — множество целых чисел).

Пусть имеется множество целых чисел Z. Два целых числа n_1, n_2 называются сравнимыми по модулю m ($m \in N$), если при делении на m (**модуль сравнения**) получаются одинаковые остатки. По этому принципу множество Z разбивается на смежные классы — **классы вычетов по модулю m**. Множество классов вычетов будет являться коммутативным кольцом с единицей.

Пусть имеется некоторое кольцо K, a, $b \in K$. Если при отличных от нуля a и b их произведение равно нулю $ab = 0$, то а называется **левым**, b — **правым делителем нуля**, нуль в кольце — тривиальным делителем нуля. Если в кольце имеется только тривиальный делитель нуля, то кольцо называется **кольцом без делителей нуля**.

Целостным кольцом называется коммутативное кольцо с единицей без делителей нуля. Элемент a кольца K называется **обратимым** (делителем единицы), если в кольце существует элемент a^{-1}, для которого $aa^{-1} = a^{-1}a = 1$. Обратимый элемент не является делителем нуля.

Справедливы следующие утверждения.

1. Нетривиальное коммутативное кольцо K с единицей является целостным тогда и только тогда, когда при a ≠ 0 из равенства $ab = ac$ следует равенство $b = c$ (a, b, $c \in K$).

2. Все обратимые элементы кольца K с единицей составляют группу по умножению (мультипликативную).

2. Поле. Модуль

Полем называется коммутативное кольцо с единицей $1 \neq 0$, в котором каждый элемент обратим. Это своеобразный гибрид аддитивной и мультипликативной групп, подчиненный дистрибутивному закону. Определение поля можно дать и в других терминах: поле — это коммутативное тело (тело — это коммутативное кольцо с единицей, в котором каждый ненулевой элемент обратим). Например, полем являются множества рациональных чисел, вещественных чисел. Если a, b ∈ П (П — поле), то произведение ab^{-1} записывается в виде дроби a/b (b не равно нулю), и для этого произведения (дроби) справедливы все школьные правила действий с дробями, подчиняющиеся аксиомам поля.

Подполе поля П — это подкольцо в П, само являющееся полем. Например, рациональные числа — подполе вещественных чисел. Нуль и единица поля являются нулем и единицей в подполе.

Поля являются изоморфными, если они изоморфны как кольца. Если поле не обладает никаким собственным подполем, то оно называется **простым полем**.

Теорема. Кольцо классов вычетов по модулю m является полем тогда и только тогда, когда модуль сравнения m — простое число.

Следствие (малая теорема Ферма). Для любого целого числа m, не делящегося на простое число p, имеет место сравнение $mp^{-1} \equiv$ $\equiv 1 \pmod{p}$.

Теорема. В каждом поле П содержится одно и только одно простое подполе, которое изоморфно либо множеству рациональных чисел, либо множеству всех целых чисел.

Поле П имеет характеристику нуль, если его простое подполе изоморфно множеству рациональных чисел; поле П имеет простую (конечную) характеристику p, если его простое подполе изоморфно множеству целых чисел, состоящему из p элементов.

Пусть K (x, y — его элементы) — ассоциативное кольцо с единицей, V (v, u — ее элементы) — абелева группа. Если задано отображение из $K \times V$ в V $(x,v) \to xv$, удовлетворяющее условиям: $x(u + v) = xu + xv$, $(x + y)v = xv + yv$, $(xy)v = x(yv)$, $1v = v$ (условие унитарности), то V называется **левым *K*-модулем** (правый K-модуль определяется аналогично). **Подмодулем *U*** K-модуля V называется подгруппа U группы V, если $xu \in U$ для всех $x \in K$, $u \in U$. Гомоморфизмом K-модулей U и V называется такое отображение φ: $U \to V$, что $\varphi(u_1 + u_2) = \varphi(u_1) + \varphi(u_2)$, $\varphi(xu_1) = x\varphi(u_1)$ $(u_1, u_2 \in U$, $x \in K)$. Ядро данного гомоморфизма является подмодулем.

ЛЕКЦИЯ № 9. Функции от матриц. Жордановы цепочки

1. Аннулирующие многочлены

Пусть имеется множество многочленов от одной переменной. Оно образует линейное пространство (комплексное или действительное в зависимости от коэффициентов многочленов) по отношению к операциям сложения и умножения на число с тождественно равным нулю (все коэффициенты равны нулю) многочленом в качестве нулевого элемента, в нем определена операция умножения многочленов. Пусть p и q — два многочлена этого множества. Многочлен h будет называться **частным** от деления многочлена q на p, а многочлен r степени меньшей степени многочлена p — **остатком** от деления, если выполняется равенство $q = ph + r$.

Справедливо утверждение: каждый многочлен можно разделить с остатком на любой ненулевой многочлен, причем частное и остаток однозначно определены (данное утверждение доказывается алгоритмом деления с остатком).

Теорема Безу. Многочлен p делится на двучлен $x - \mu$ тогда и только тогда, когда $p(\mu) = 0$. Действительно, поскольку двучлен $x - \mu$ имеет степень, равную единице, то остаток r от деления будет иметь нулевую степень, т. е. будет постоянной. Тогда имеет место равенство $p = h(x - \mu) + r$, подставляя в него вместо x μ, получаем $r = p(\mu)$, и тогда $p(\mu) = 0$.

Понятия наибольшего общего делителя (НОД) и наименьшего общего кратного (НОК) для многочленов аналогичны этим понятиям для чисел.

Теорема. Для того чтобы комплексные многочлены $p_1, ..., p_n$ не были взаимно просты, необходимо и достаточно, чтобы они имели общий корень (два многочлена являются взаимно простыми, если их наибольший общий делитель равен нулю). Действительно, достаточность следует из теоремы Безу. А если многочлены имеют общий делитель d, степень которого не равна нулю, то каждый из многочленов может быть представлен в виде $p_i = h_i d$ (h_i — многочлен); делитель d как многочлен ненулевой степени имеет хотя бы один корень μ. Тогда: $p_i(\mu) = h_i(\mu)d(\mu) = 0$, значит, μ — общий корень.

Теорема. Пусть $p_1, ..., p_n$ — многочлены, среди которых имеется не равный нулю многочлен, d — НОД. Тогда найдутся такие многочлены $u_1, ..., u_n$, что $d = u_1p_1 + ... + u_np_n$. Действительно, пусть $d = u_1p_1 + ... + u_np_n$ — ненулевой многочлен минимальной степени в множестве многочленов вида $f_1p_1 + ... + f_np_n$ (в нем определены операции сложения и умножения многочленов). Тогда имеет место равенство $q = hd + r$, где q и h принадлежат рассматриваемому множеству многочленов, а r тоже принадлежит этому множеству и имеет степень, меньшую степени многочлена d, что противоречит условию теоремы.

Следствие. Многочлены $p_1, ..., p_n$ взаимно-просты тогда и только тогда, когда найдутся такие многочлены $u_1, ..., u_n$, что $u_1p_1 + ... + u_np_n = 1$.

Теорема. Общее кратное q многочленов $p_1, ..., p_n$ является их наименьшим общим кратным (НОК) тогда и только тогда, когда частные от деления q на эти многочлены взаимно-просты.

В качестве переменой в многочлене можно брать линейные преобразования. В множестве линейных преобразований (n-мерное линейное пространство) определены операции сложения, умножения, а нулевой степенью является тождественное преобразование E пространства. Тогда многочлен p от линейного преобразования А сам является линейным преобразованием и имеет вид: $p(A) = a_0E + a_1A + ... + a_mA^m$. Если А является матрицей некоторого линейного преобразования φ, то преобразование $p(\varphi)$ имеет матрицу вида $p(A) = a_0E + a_1A + ... + a_mA^m$ (т. е. в данном случае A — матрица преобразования, E — единичная матрица). Это матричный многочлен.

Аннулирующим преобразование *A* многочленом (аннулирующим многочленом) называется ненулевой многочлен p в случае, если $p(A) = O$ (здесь O — нулевое преобразование линейного пространства, которое любому вектору сопоставляет нулевой вектор; оно имеет нулевую матрицу). Для аннулирующего многочлена матричный многочлен равен нулю. В множестве P аннулирующих некоторое преобразование A многочленов справедливы свойства: если $g, h \in P$, то $g + h \in P$; если $g \in P$, то $gh \in P$ для любого многочлена h.

Теорема. Для каждого линейного преобразования существует аннулирующие его многочлен.

Теорема. Среди многочленов, аннулирующих некоторое преобразование, существует многочлен минимальной степени, который является делителем всех аннулирующих многочленов и определен с точностью до числового множителя (если старший коэффициент такого многочлена равен единице, то он называется **минимальным многочленом** этого преобразования).

Если минимальный многочлен некоторого линейного преобразования имеет вид t^m, то это преобразование называется **нильпотентным**, а число m — **показателем нильпотентности**. Все собственные значения нильпотентного преобразования равны нулю; все собственные векторы вместе с нулевым составляют ядро преобразования (ядро преобразования можно определить как множество векторов, переводимых данным преобразованием в нулевой вектор). Показатель нильпотентности не превосходит размерности пространства.

2. Корневые подпространства

Пусть имеется комплексное линейное пространство, а линейное преобразование имеет минимальный многочлен вида:

$$p(t) = (t - \lambda_1)^{k_1} \cdots (t - \lambda_n)^{k_n},$$

где $\lambda_1, \ldots, \lambda_n$ — попарно различные корни этого многочлена кратности $k_1, \ldots, k_n$ соответственно, причем сумма $(k_1 + \ldots + k_n)$ равна степени минимального многочлена линейного преобразования.

Корневыми подпространствами K_i преобразования A называются подпространства, являющиеся ядрами $Ker(FA - \lambda_i E)^{k_i}, i = 1, \ldots, n$, где λ_i, k_i соответствуют виду минимального многочлена.

Справедливо утверждение: корневые подпространства некоторого преобразование инвариантны относительно этого преобразования.

Сумма подпространств $L_1, \ldots, L_n$ называется **прямой суммой**, если $L_k \cap \sum_{i \neq k} L_i = 0,\ k = 1, \ldots, n.$ Размерность прямой суммы равна сумме размерностей слагаемых.

Блочно-диагональным (клеточно-диагональным) видом называется следующий вид матрицы A:

$$A = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_n \end{pmatrix},$$

где A_1, ..., A_n — квадратные матрицы (матрицы ограничения преобразования), имеющих порядок соответственно m_1, ..., m_n. Блочно-диагональная матрица обозначается $\mathrm{diag}(A_1, ..., A_n)$. Детерминант (определитель) такой матрицы равен произведению определителей квадратных матриц.

Характеристический многочлен преобразования с матрицей A равен произведению характеристических многочленов его ограничений $q(\lambda) = \det(A - \lambda E) = \prod_{i=1}^{n} \det(A_i - \lambda E_{m_i})$, где E_{mi} — единичная матрицы порядка mi.

Справедливы утверждения.

1. Множество корней минимального многочлена некоторого преобразования совпадает с множеством корней его характеристического многочлена. Кратности корней в характеристическом многочлене равны размерностям корневых подпространств.

2. Показатель нильпотентности преобразования в точности равен кратности соответствующего корня в минимальном многочлене.

Теорема Гамильтона — Кэли: Характеристический многочлен преобразования является его аннулирующим многочленом. Или в матричной формулировке. Каждая матрица удовлетворяет своему характеристическому уравнению (данное утверждение не зависит от выбора базиса).

3. Жордановы цепочки. Теорема Жордана. Жорданова матрица

Пусть имеется m-мерное пространство K, B — нильпотентное преобразование этого пространства с показателем нильпотентности k. Пусть для некоторого вектора x $B^h(x) \neq 0$, $B^{h+1}(x) = 0$.

Если векторы e_1, ..., e_h вместе с собственным вектором e^0 удовлетворяют следующим условиям: $B(e^1) = e^0$, $B(e^2) = e^1$, ..., $B(e^h) = e^{h-1}$, то они называются соответственно первым, вторым и так далее **присоединенными** к e^0 векторами, и векторы e^0, e^1, ..., e^h образуют **жорданову цепочку с началом в e^0**.

Теорема. Вектор e^0 имеет ровно h присоединенных векторов (начало цепочки) тогда и только тогда, когда $e^0 \in \mathrm{Im}\, B^h \cap KerB$, $e^0 \notin \mathrm{Im}\, B^{h+1}$ ($\mathrm{Im}\ B$ — образ отображения (преобразования)). Действительно, если вектор e^0 принадлежит ядру $\mathrm{Ker}\ B$, то он собственный; а если он принадлежит множеству $\operatorname{Im} B^h$, то найдется вектор x, для которого $e^0 = B^h(x)$. А поскольку $e^0 \neq B^{h+1}(x)$, то цепочки большей длины нет.

Теорема. Система из собственных и присоединенных ним векторов линейно независима, если входящие в нее собственные векторы линейно независимы (доказывается методом математической индукции по числу векторов). Следовательно, при добавлении к каждому вектору базиса присоединенных к нему векторов получается система линейно независимых векторов в случае линейной независимости собственных векторов.

Теорема. Каждый вектор пространства $\operatorname{Im} B^h \cap KerB$ раскладывается в линейную комбинацию векторов системы из собственных и присоединенных к ним векторов.

Жордановым базисом пространства для некоторого преобразования называется объединение жордановых базисов корневых подпространств этого пространства, построенных для преобразований этих подпространств.

Жордановой клеткой с собственным значением λ_i называется матрица вида:

$$J(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix},$$

она имеет единственное характеристическое число λ_i. Минимальный многочлен жордановой клетки имеет вид $(t - \lambda_i)$m, где m — порядок клетки. Диагональная матрица, элементами диагонали которой являются жордановы клетки, называется **жордановой матрицей**. Определение жордановой матрицы называется приведением матрицы к жордановой нормальной форме.

Справедливо утверждение. Максимальный порядок клетки с собственным значением λ_i равен кратности корня λ_i в минимальном многочлене.

Преобразованием простой структуры (преобразование с простым спектром) называется линейное преобразование, минимальный многочлен которого не имеет кратных корней.

Линейное преобразование имеет в некотором базисе диагональную матрицу тогда и только тогда, когда оно является преобразованием простой структуры.

Матрица преобразования A в его жордановом базисе (жорданова матрица) есть клеточно-диагональная матрица, диагональными летками которой являются жордановы клетки порядков, равных размерностям циклических подпространств, на которые распадаются все корневые подпространства, с собственными значениями, равными соответствующим корням минимального многочлена этого преобразования.

Теорема Жордана. Для каждого преобразования комплексного линейного пространства существует базис, в котором его матрица имеет жорданову нормальную форму. При этом жорданова нормальная форма матрицы однозначно определена по преобразованию с точностью до порядка, в котором расположены диагональные клетки.

Следствие. Каждая матрица подобна некоторой жордановой матрице. Для подобия двух матриц необходимо и достаточно, чтобы соответствующие жордановы матрицы совпадали с точностью до порядка следования клеток (матрицы A_1 и A_2 называются подобными, если существует такая невырожденная матрица B, что справедливо равенство $A_1 = B^{-1}A_2B$).

Каждое линейное преобразование комплексного линейного пространства представимо как сумма преобразования простой структуры и нильпотентного преобразования, которые коммутируют между собой. Для нильпотентной матрицы жорданова нормальная форма существует.

Пусть линейное преобразование задано матрицей A. Чтобы построить ее жорданов базис, надо найти корни характеристического многочлена матрицы A и их кратности (число жордановых клеток определенного порядка, соответствующих определенному собственному значению). Пусть λ^* — корень кратности m. Если ранг матрицы $(A - \lambda \times \mathrm{E})$ больше разности $n - m$, то найдем k — кратность корня λ^* в минимальном многочлене. При этом k будет выполняться равенство $(A - \lambda \times E) = n - m$. Далее рассматриваются столбцы базисного минора матрицы $(A - \lambda \times E)^{\kappa-1}$ и строятся собственные векторы, пока общая длина всех цепочек не станет равной m. Присоединенные векторы определяются при решении линейных систем вида $(A - \lambda \times E)e^i = e^{i-1}$. Как только будут определены длины всех цепочек, соответствующих каждому корню характеристического многочлена, определяется жорданова форма матрицы.

Например, для матрицы

$$A = \begin{pmatrix} 3 & -1 & 0 & -1 \\ 1 & 1 & 0 & -1 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

одним их корней характеристического многочлена $(\lambda - 1)^2(\lambda - 2)^2$ является 1 ($\lambda = 1$). Составляем для него матрицу вида $A - \lambda E$:

$$\begin{pmatrix} 2 & -1 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix},$$

ее ранг равен 3. При возведении в квадрат получается матрица

$$\begin{pmatrix} 3 & -2 & -1 & 0 \\ 2 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

с рангом, равным 2, значит, кратность рассматриваемого корня в минимальном многочлене равна 2. Собственные векторы определяются при решении системы

$$\begin{pmatrix} 2 & -1 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Решение — столбец

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Присоединенный вектор определяется при решении системы

$$\begin{pmatrix} 2 & -1 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Одно из решений — столбец

$$\begin{matrix} 1 \\ 1 \\ 1 \\ 0 \end{matrix}.$$

4. Функции от матриц

Понятие функции от матрицы вводится в соответствии с общим определением функции, то есть функция от матрицы определяется как отображение, ставящее в соответствие каждой матрице определенного множества матриц единственный элемент некоторого другого множества, элементами которого также могут являться матрицы.

Пусть A — квадратная матрица порядка n с элементами a_{ij}, $\{a_k\}$ — последовательность комплексных чисел. Тогда **степенным рядом относительно матрицы** ***A*** называется сумма

$$a_0 E + \alpha_1 A + \alpha_2 A^2 + \ldots = \sum^{\infty} \alpha_k A^k.$$

Частичными суммами данного степенного ряда называются конечные суммы вида $\sum_{k=0}^{N} \alpha_k A^k$. **Матричная норма** — это число

$$\|A\| = n \max_{i,j} |a_{ij}|.$$

Степенной ряд относительно матрицы A называется **сходящимся** к матрице F, если к ней сходится по норме последовательность частичных сумм, т. е если для любого положительного числа $\varepsilon > 0$ найдется такой номер N_0, что для всех $N > N_0$ выполняется неравенство:

$$\left\| F - \sum_{k=0}^{N} \alpha_k A^k \right\| < \varepsilon.$$

В случае сходимости матрица F называется **суммой ряда,** и

$$F = \sum_{k=0}^{\infty} \alpha_k A^k.$$

Пусть K_1 и K_2 — множества комплексных квадратных матриц порядка n. Функция f на множестве K_1 со значениями в множестве K_2 называется **регулярной**, если существует степенной ряд, сходящийся на K_1, такой, что для каждой матрицы A из множества K_1

справедливо равенство $f(A)=\sum_{k=0}^{\infty}\alpha_k(A-A_0)^k$, где A_0 — фиксированная матрица. Регулярная функция f $(f(A)=\sum_{k=0}^{\infty}\alpha_k A^k)$ обладает следующим свойством: для любой невырожденной матрицы S выполняется равенство:

$$S^{-1}f(A)S=\sum_{k=0}^{\infty}\alpha_k(S^{-1}AS)^k.$$

Данное свойство позволяет упростить исследование матричных степенных рядов благодаря замене функции от матрицы на матрицу вида $S^{-1}AS$.

Справедливы следующие утверждения.

1. Пусть A — клеточно-диагональная матрица $A = diag(A_1, ..., A_s)$. Тогда для любой регулярной функции f, определенной на матрице A, выполняется равенство: $f(A) = diag(f(A_1), ..., f(A_s))$. Действительно, ведь при возведении в степень клеточно-диагональной матрицы получается опять клеточно-диагональная матрица, по диагонали которой стоят степени диагональных клеток исходной матрицы.

2. Регулярная функция f определена на матрице A линейного преобразования тогда и только тогда, когда она определена на матрицах A_i ограничений линейного преобразования на его корневых подпространствах. Причем $f(A) = S\,diag(f(A_1), ..., f(A_s))S^{-1}$, где S — матрица перехода к базису, являющемуся объединением базисов корневых подпространств этого линейного преобразования.

Теорема. Для того чтобы регулярная функция f, задаваемая скалярным степенным рядом от комплексной переменной

$$f(\xi)=\sum_{i=0}^{\infty}\alpha_i\xi^i,$$

была определена на матрице A, достаточно, чтобы корни характеристического многочлена матрицы A лежали внутри круга сходимости данного ряда. Необходимым условием является принадлежность этих корней замыканию круга сходимости. При этом если S — матрица перехода к базису, который расположен в корневых подпространствах преобразования, соответствующего матрице A, и $A = S\,diag(A_1, ..., A_s)\mathrm{S}^{-1}$, то

$$f(A)=Sdiag\left\{\sum_{j=0}^{k_1-1}\frac{1}{j!}f^{(j)}(\lambda_1)(A_1-\lambda_1E)^j,..,\sum_{j=0}^{k_s-1}\frac{1}{j!}f^{(j)}(\lambda_s)(A_s-\lambda_sE)^j\right\}S^{-1}$$

(понятие круга сходимости и его замыкания аналогично этим понятиям в математическом анализе).

Значениями функции f на спектре матрицы A называется совокупность $f(\lambda_i), f'(\lambda_i), ..., f^{(ki-1)}(\lambda_i)$ (i = 1, 2, .., s), где λi — это корни минимального многочлена матрицы A кратности k_i.

Если корни минимального многочлена матрицы A лежат внутри круга сходимости скалярного ряда, определяющего функцию f, то значение этой функции на матрице A однозначно определено значениями функции на спектре матрицы A (следует из предыдущей теоремы).

Для регулярной функции f корни характеристического многочлена матрицы $f(A)$ равны $f(\lambda_1), ..., f(\lambda_n)$, где $\lambda_1, ..., \lambda n$ — корни характеристического многочлена матрицы A.

5. Матричные функции скалярного аргумента

Пусть R — множество вещественных чисел t ($t \in R$). На нем задана матричная функция F, которая каждому числу множества ставит в соответствие матрицу $F(t)$ размером $m \times n$, т. е. с каждым скалярным аргументом t сопоставляются отдельные элементы матрицы $F(t)$ (задается mn числовых функций на множестве вещественных чисел, называемых элементами функции F). Такие понятия, как непрерывность, дифференцируемость матричной функции, вводятся посредством определений для их элементов.

Матричная функция считается непрерывной (дифференцируемой), если все ее элементы непрерывны (дифференцируемы). Производная матричной функции равна:

$$\frac{dF(t)}{dt} = \begin{pmatrix} f'_{11}(t) & \cdots & f'_{1n}(t) \\ & \cdots & \\ f'_{m1}(t) & \cdots & f'_{mn}(t) \end{pmatrix}.$$

Для дифференциала матричной функции справедливы следующие свойства: постоянный множитель выносится за знак дифференциала:

$$\frac{d}{dt}(F^k) = \frac{dF}{dt}F^{k-1} + F\frac{dF}{dt}F^{k-2} + ... + F^{k-1}\frac{dF}{dt};$$

$$\frac{d}{dt}(F^{-1}) = -F^{-1}\frac{dF}{dt}F^{-1};$$

$$\frac{d(AB)}{dt} = \frac{dA}{dt}B + A\frac{dB}{dt}.$$

Считается, что матричная функция скалярного аргумента разложена в степенной ряд, если в ряды разложены все ее элементы. Частичные суммы равны:

$$F_N(t) = \sum_{k=0}^{N} A_k (t - t_0)^k.$$

Степенной ряд матричной функции скалярного аргумента:

$$F(t) = \sum_{k=0}^{\infty} A_k (t - t_0)^k.$$

Матричный ряд сходится, если сходятся все ряды для элементов функции. Функции-столбцы высоты n от скалярного аргумента образуют бесконечномерное вещественное линейное пространство R_t^n.

ЛЕКЦИЯ № 10. Норма

1. Нормированные пространства. Норма

Сравнение элементов какого-нибудь линейного пространства осуществляется с помощью нормы. Линейное пространство, в котором задана норма, называется **нормированным** линейным пространством. Понятие нормы вводится следующим образом: функция f, сопоставляющая с каждым вектором вещественного или комплексного линейного пространства некоторое вещественное число, называется **нормой**, а ее значения на векторе x называются **нормой этого вектора** и обозначаются $\|x\|$, если для любых векторов x и y и произвольного числа α выполнены следующие условия:

1) $f(x) > 0$ для всех не равных нулю $x \neq 0$;
2) свойство положительной однородности: $f(\alpha x) = |\alpha| f(x)$, и отсюда $f(0) = 0$;
3) свойство выпуклости: $f(x + y) \leq f(x) + f(y)$;
4) $f(x - y) \geq |f(x) - f(y)|$.

Расстояние между двумя векторами в нормированном пространстве — это норма разности этих векторов. Для евклидова пространства определенное так расстояние совпадает с обычным (известным со школы) понятием расстояния.

***Е*-окрестностью** некоторого вектора a называется множество векторов нормированного пространства, расстояние от которых до вектора a не превышает заданного числа ε.

В одном и том же нормированном линейном пространстве могут быть определены разные нормы. Например, в n-мерном арифметическом пространстве для вектора или столбца $X(x_1, ..., x_n)$ могут применяться следующие нормы: октаэдрическая норма $\|x\| = \sum\limits_i |x_i|$, евклидова (для комплексного пространства унитарная) норма $\|x\| = \left(\sum\limits_i x_i^2\right)^{1/2}$ **норма Гельдера** $\|x\| = \left(\sum\limits_i |x_i|^p\right)^{1/p}$ $p > 1$, кубическая норма $\|x\| = \max\limits_i |x_i|$. Для произвольных линейных пространств используются такие же нормы при фиксированном базисе.

Существуют различные линейные преобразования, которые сохраняют норму, т. е. $\|A(x)\| = \|x\|$ (например, ортогональные преобразования в евклидовом пространстве).

Единичная сфера нормированного пространства — это множество векторов, имеющих единичную норму (например, для кубической нормы единичная сфера — куб, для октаэдрической — октаэдр). Множество векторов, норма которых не превышает единицы, называется **единичным шаром**. При известной единичной сфере всегда можно вычислить норму любого вектора этого пространства.

Отрезком с концами x_1 и x_2 в нормированном пространстве называется множество векторов вида $y = ax_1 + (1 - a)x_2$, где a принадлежит отрезку $[0, 1]$, x_1 и x_2 — векторы этого нормированного пространства. Если концы отрезка принадлежат единичному шару, то и весь отрезок принадлежит этому шару.

Две нормы линейного пространства могут быть связаны друг с другом (например, отношением эквивалентности). Нормы f_1 и f_2 линейного нормированного пространства называются **эквивалентными**, если существуют такие положительные числа a_1 и a_2, что выполняется неравенство $a_1 f_1 \le f_2 \le a_2 f_1$. Каждая норма эквивалентна самой себе (эквивалентность рефлексивна). В арифметическом пространстве все нормы эквивалентны. Поэтому выбор той или иной нормы зависит только от удобства вычислений.

Справедливо утверждение: последовательность векторов $\{x_i\}$ сходится по норме f_1 тогда и только тогда, когда она сходится по любой эквивалентной ей норме f_2 (следует из определений сходимости и эквивалентности).

Пусть даны две нормы некоторого линейного пространства f_1 и f_2. Считается, что норма f_1 **мажорирует** норму f_2, если для любого вектора x из этого линейного пространства выполнено неравенство $f_1(x) \ge f_2(x)$. Обозначается $f_1 \ge f_2$. Каждая норма в арифметическом пространстве мажорируется произведением октаэдрической нормы на некоторое число a, произведением кубической нормы на некоторое число c и произведением евклидовой нормы на число c.

2. Нормы матриц

Пусть имеется линейное пространство M матриц размеров $m \times n$. Как и в любом другом линейном пространстве, в пространстве матриц также могут быть определены различные нормы.

Норма линейного пространства M матриц называется **согласованной** с нормами в арифметических пространствах, если для любой матрицы A и любого столбца X (высотой n) выполнено неравенство $\|AX\| \le \|A\|\,\|X\|$ (высота столбца $AX - m$). Нормы столбцов определены в соответствующих арифметических пространствах. Каковы бы ни были нормы арифметических пространств, согласованной нормой в пространстве матриц будет функция вида:

$$f(A) = \sup \frac{\|AX\|}{\|X\|}$$

или

$$f(A) = \sup_{\|X\|=1} \|AX\|.$$

Такая норма называется **индуцированной нормой** (индуцированная в арифметических пространствах норма). Легко проверить, что свойства нормы для этих функций выполняются. Известно, что каждая согласованная норма мажорирует индуцированную норму.

Считается, что норма **сохраняет единицу**, если в пространстве квадратных матриц $\|E\| = 1$. Если для любых квадратных матриц выполняется неравенство $\|AB\| \le \|A\|\,\|B\|$, то норма называется **матричной (кольцевой)**, а само неравенство называется **кольцевым свойством** нормы. Для кольцевых норм справедливы следующие неравенства: $\|A\| \le \|E\|\,\|A\|$, $\|E\| \ge 1$, $\|A^k\| \le \|A\|^k$, $\|A^{-1}\| \ge \|A\|^{-1}$ (k — натуральное). Очевидно, что любая индуцированная норма в пространстве квадратных матриц сохраняет единицу и обладает кольцевым свойством.

Кольцевым свойством могут обладать и прямоугольные матрицы, если их размеры таковы, что возможно произведение матриц. Для них кольцевое свойство $\|AB\|_3 \le \|A\|_1\,\|B\|_2$ выполняется при следующих условиях: подходящие для умножения размеры матриц A и B, норма 1 — согласованная, нормы 2 и 3 — индуцированные.

Для матриц наиболее употребительными являются следующие нормы.

1. **Спектральная норма** матрицы A равна $\sup\limits_{X \ne 0} \dfrac{\|AX\|}{\|X\|}$, при этом нормы арифметических пространств — евклидовы или унитарные. Спектральная норма матрицы не меняется при умножении этой матрицы справа или слева на ортогональную матрицу. Дейст-

вительно, евклидова норма столбца не меняется при умножении его на ортогональную матрицу $\|UX\| = \|X\|$, следовательно, спектральная норма ортогональной матрицы равна единице. Тогда имеем $\|UA\| \leq \|U\|\,\|A\| = \|A\|$ и $\|A\| = \|U^{-1}UA\| \leq \|U^{-1}\|\,\|UA\| = \|UA\|$. Для умножения слева доказательство аналогичное.

Спектральная норма матрицы А равна ее максимальному сингулярному числу $\|A\| = a_1$ (для произвольного линейного преобразования F евклидова пространства максимальное сингулярное число равно $a_1 = \max\limits_{x\neq 0}\frac{|F(x)|}{|x|}$, минимальное сингулярное число равно $a_n = \min\limits_{x\neq 0}\frac{|F(x)|}{|x|}$). Для невырожденной квадратной матрицы $\|A^{-1}\| = a_n^{-1}$.

2. **Норма, индуцированная октаэдрической нормой** арифметического пространства, равна $\|A\| = \max\limits_{k}\sum\limits_{j}|a_{jk}|$ (наибольшая из сумм модулей элементов матрицы A по столбцам).

3. **Норма, индуцированная кубической нормой** арифметического пространства, равна $\|A\| = \max\limits_{j}\sum\limits_{k}|a_{jk}|$ (наибольшая из сумм модулей элементов матрицы A по строкам).

4. **Евклидова норма** матрицы имеет вид: $\|A\| = \left(\sum\limits_{i,j}|a_{ij}|^2\right)^{1/2}$ Евклидова норма матрицы A не меняется при умножении A справа или слева на ортогональную матрицу и обладает кольцевым свойством.

5. Реже используются следующие нормы матриц: $\|A\| = \max\limits_{i,j}|a_{ij}|$ (сохраняет единицу, кольцевым свойством не обладает), $\|A\| = n\max\limits_{i,j}|a_{ij}|$ (для квадратных матриц, обладает кольцевым свойством, согласована с тремя основными нормами арифметического пространства), $\left\|A = \sum\limits_{i,j}|a_{ij}|\right\|$ (используется редко), $\|A\| = \frac{1}{n}\sum\limits_{i,j}|a_{ij}|$ (сохраняет единицу).

ЛЕКЦИЯ № 11. Численные методы

1. Численные методы. Введение

В настоящее время для решения многих задач линейной алгебры используется вычислительная техника. Однако работа с ЭВМ подразумевает получение результатов вычислений с какой-либо точностью. Точность расчетов до необходимого можно увеличивать при разработке новых алгоритмов и программ, но при этом программы усложняются, и время выполнения расчетов растет.

В ЭВМ используется приближенная запись чисел. При этом возникают ошибки округления, которые увеличиваются при увеличении числа выполняемых операций. Следовательно, в результате вычисления на ЭВМ получается приближенное решение той или иной задачи или системы линейных уравнений. Довольно часто для оценки получившихся погрешностей применяется обратный анализ ошибок округления. Согласно этому принципу определяется система линейных уравнений (если вычислялось ее решение), для которой полученное приближенной решение является точным, а затем оценивается отличие коэффициентов полученной системы от коэффициентов исходной системы.

Кроме этого, очень часто бывает, что коэффициенты и свободные члены исходной системы заданы с погрешностями, что тоже влияет на точность вычислений.

Исходной системой (невозмущенной) называется система линейных уравнений $Ax = b$ с коэффициентами a_{ij} и свободными членами b_i. **Возмущенной системой** для системы $Ax = b$ называется система с коэффициентами $a_{ij} \pm \Delta a_{ij}$ и свободными членами $b_i \pm \Delta b_i$. **Погрешность** решения — это разность между решениями исходной и возмущенной систем. Одно и то же возмущение коэффициентов и свободных членов может вызывать разную погрешность решения. Например, каждое из уравнений с двумя неизвестными системы двух линейных уравнений соответствует прямой. Чем больше угол между этими прямыми, тем лучше **обусловлена заданная** система, т. е. вызываемая погрешность решения при одинаковом возмущении меньше.

Решение исходной системы не всегда можно заменить решением возмущенной системы, поскольку возмущение коэффициентов и свободных членов может привести к тому, что вместо единственного решения получится либо бесчисленное множество решений, либо решений не будет совсем. При наличии возмущений приходится иметь дело с **почти вырожденными матрицами** (их элементы изменяются в пределах заданной точности). Множество таких матриц зависит от выбранной точности. В ε-окрестности почти вырожденной матрицы найдется вырожденная матрица, т. е. при малом изменении элементов почти вырожденная матрица превращается в вырожденную.

С целью экономии места в оперативной памяти применяются вычислимые матрицы (элементы не задаются, а вычисляются по некоторым данным), разреженные матрицы (большинство элементов — нули, расположенные в каком-либо порядке), матрицы специальных видов (промежуточные, часть их элементов вычисляется или равна нулю, например симметричные или диагональные матрицы, ленточные матрицы).

Ленточной матрицей называется матрица, если при некотором фиксированном числе k и при $|i - j| < k$ элементы $a_{ij} = 0$. В этом случае не равные нулю элементы попадают внутрь полосы — ленты шириной $2k + 1$, которая расположена вдоль главной диагонали.

Преимущества разреженных матриц заметно в случае, когда количество не равных нулю элементов имеет порядок числа строк данной матрицы.

При работе с разными матрицами применяются различные методы. Например, для матриц небольших размеров лучше использовать **метод Гаусса**, а для матриц больших размеров предпочтение отдается итерационным методам.

Для решения линейных систем используются прямые методы и итерационные методы. Прямые методы позволяют теоретически точно решить исходную систему. К ним относятся метод Гаусса, дающий LU-разложение, и методы разложений и вращений для получения QR-разложения. К итерационным методам относятся **метод Зейделя**, метод верхней релаксации.

2. Метод Гаусса

Метод Гаусса относится к прямым методам решения линейных систем и часто используется при решении линейных систем вида $AX = B$ или

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases}$$

относительно n неизвестных x_1, x_2, ..., x_n (A — невырожденная квадратная матрица порядка n). лекции 1 о приведении системы линейных уравнений к ступенчатому виду был изложен в упрощенном варианте метод Гаусса — алгоритм решения линейных систем. В общем случае под названием «метод Гаусса» объединена группа способов, которые позволяют посредством элементарных преобразований привести матрицу исходной системы к единичному виду, причем если к исходной матрице добавить столбец свободных членов, то в результате преобразований он превращается в решение исходной системы. В зависимости от выбранной последовательности элементарных преобразований различаются схема единственного деления, обратная подстановка, обратный ход, метод оптимального исключения.

Схема единственного деления (метод прямого исключения, метод Гаусса в узком смысле) применяется при отличных от нуля главных минорах. При этом не происходит перестановок строк и столбцов. На первом шаге метода первый столбец матрицы преобразуется к виду $\begin{matrix} 1 \\ 0 \\ \vdots \\ 0 \end{matrix}$ для чего из каждой i-строки вычитается первая умноженная, число $c = a_{i1}/a_{11}$ ($i = 2, 3, \ldots, n$) (следовательно, из всех уравнений, начиная со второго, исключаются слагаемые с x_1). Затем аналогичная операция применяется к части полученной матрицы, включающей строки и столбцы с номерами 2, 3, ... n, и т. д. В результате получатся матрица с единицами на главной диагонали и нулями ниже главной диагонали; выше главной диагонали будут находиться не равные нулю элементы (верхняя треугольная матрица).

Решение системы с треугольной матрицей определяется при обратной подстановке: последнее уравнение системы имеет вид $x_n = b_n{}^*$, затем, подставляя найденное x_n в предыдущее уравнение, находим x_{n-1}. Подставляя найденные x_n и x_{n-1} в следующее вышестоящее уравнение полученной системы, определяем x_{n-2} и т. д.

Треугольную матрицу, полученную при прямом исключении, можно элементарными преобразованиями столбцов привести к единичному виду (обратный ход метода Гаусса), т. е. решить систему.

Метод оптимального исключения состоит в переменном выполнении операций прямого и обратного хода, т. е., как и при прямом ходе, приводим первый элемент второй строки к нулю, а затем с помощью второй строки приводим к нулю первый элемент второго столбца (операция обратного хода). И далее аналогичные операции применяются к части полученной матрицы, включающей строки и столбцы с номерами 2, 3, ..., n. Метод оптимального исключения позволяет использовать гораздо меньший объем оперативной памяти.

При данных способах не происходит перестановок строк или столбцов, поскольку главные миноры (или элементы, у которых оба индекса равны) получающихся матриц не равны нулю. Используются обычные арифметические операции с числами, их количество определяет продолжительность вычислений.

3. LU-разложение

Данное разложение основано на умножении матрицы исходной линейной системы на преобразующую матрицу. ***LU*-разложение** матрицы A — это разложение ее в произведение LU невырожденной нижней треугольной матрицы L и верхней треугольной матрицы U с единицами на главной диагонали.

Пусть A — невырожденная квадратная матрица порядка n. Совокупность элементарных преобразований матрицы A равносильна умножению этой матрицы слева на некоторую матрицу B.

Справедливы следующие утверждения.

1. Для любой матрицы A с главными минорами, отличными от нуля, найдется такая невырожденная нижняя треугольная матрица B, что произведение BA есть верхняя треугольная матрица U с единицами на главной диагонали.

2. Матрица L, обратная к нижней треугольной матрице B, сама является нижней треугольной матрицей и имеет вид:

$$L = \begin{pmatrix} a_{11} & 0 & \cdots & 0 & \cdots & 0 \\ a_{12} & a_{22}^{(1)} & \cdots & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & a_{kk}^{(k-1)} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2}^{(1)} & \cdots & a_{nk}^{(n-1)} & \cdots & a_{nn}^{(n-1)} \end{pmatrix},$$

где $a_{ik}^{(k-1)}$ — элементы матрицы $A^{(k-1)}$, получаемой по схеме единственного деления на $(k-1)$-м шаге (доказывается при непосредственной проверке произведения $B^{(n)}, ..., B^{(2)}B^{(1)}L$, которое должно быть равно E, где $B^{(k)}$ — матрица, для которой выполняется равенство $A^{(k)} = B^{(k)}A^{(k-1)}$.

Из определения матрицы L видно, что элементы ее k-го столбца являются элементами матрицы $A^{(k+1)}$.

Следовательно, при LU-разложении получается матрица, в которой выше главной диагонали расположены элементы матрицы U (не равные нулю или единице), а остальные элементы совпадают с соответствующими элементами матрицы L.

Теорема. LU-разложение матрицы A существует, если все ее главные миноры (и определитель) не равны нулю. Отличие от нуля главных миноров является необходимым и достаточным условием существования такого разложения. Причем если разложение существует, то оно единственное (точнее, единственным является такое разложение, когда на главной диагонали именно второго множителя стоят единицы).

Действительно, пусть нижняя треугольная матрица $B = L^{-1}$, $U = BA$. Если U_k, B_k, A_k — матрицы, состоящие соответственно из элементов матриц U, B, A, стоящих в строках и столбцах с номерами 1, 2, ..., k, то для них будет справедливо $U_k = B_kA_k$. Поскольку определитель $\det U_k \neq 0$, то определитель $\det A_k \neq 0$ (главные миноры не равны нулю), значит, LU-разложение существует. Каждая k-тая строка матрицы B_k определена однозначно, так как последняя строка матрицы U_k — последняя строка единичной матрицы, а строка матрицы A_k линейно независима. Значит, и k-тая строка матрицы B определена однозначно, ведь она получается из k-той строки матрицы B_k при добавлении $(n-k)$ нулей. Следовательно, определены однозначно матрицы $L = B^{-1}$ и $U = BA$.

В свою очередь, матрица L представима в виде произведения матрицы L_1 с единицами на главной диагонали и диагональной матрицы D. Тогда исходная матрица A будет разложена в произведение L_1DU, где L_1 и U являются соответственно нижней и верхней треугольной матрицей с единицами на главных диагоналях. Такое разложение также является единственным (следует из единственности LU-разложения).

Преимущество данного метода состоит в том, что решение исходной системы заменяется решением двух систем с треугольными матрицами, что значительно проще (т. е. вместо системы $AX = B$

решаются две системы: $LY = B$ и $UX = Y$). Причем место, занимаемое в оперативной памяти LU-разложением, не превосходит места, занимаемого исходной системой.

UL-разложение может быть получено посредством алгоритма, называемого **компактной схемой метода Гаусса**. Элементы матрицы произведения LU равны $\sum_{k=1}^{n} l_{ik}u_{kj}$, т. е. элементы матрицы A равны соответственно $a_{ij} = \sum_{k=1}^{n} l_{ik}u_{kj}$, что равносильно n^2 числовым равенствам. Причем при $k > i$ выполняется равенство $l_{ik} = 0$, а при $k < 1$ — равенство $u_{kj} = 0$, поскольку матрицы L и U — треугольные. Тогда $a_{ij} = \sum_{k=1}^{m} l_{ik}u_{kj}$, где $m = \min(i, j)$ Используя эти уравнения, последовательно определяют все элементы матриц L и U. При вычислениях уравнения разбиваются на группы в зависимости от величины m и соотношения между i и j. Например, при $m = 1$ и при $j > i = 1$ или при $i \geq j = 1$ имеем $a_{1j} = l_{11}u_{1j}$ $(j > 1)$ или $a_{i1} = l_{i1}u_{11} = l_{i1}$, поскольку $u_{11} = 1$ $(i \geq 1)$. Затем вычисляется $u_{kj} = a_{1j} / l_{11}$ $(j > 1)$. По найденным значениям определяются последовательно остальные. Компактная схема оптимальна по точности и занимаемому объему памяти.

4. Выбор главного элемента. Масштабирование

Довольно часто приходится иметь дело с матрицами, главные миноры которых могут быть равны нулю (например, $a_{11} = 0$). В этом случае приходится прибегать к перестановке строк и столбцов исходной матрицы, в результате чего главные миноры становятся не равными нулю. Любую невырожденную матрицу A порядка n перестановкой только строк (или только столбцов) можно перевести в матрицу, главные миноры которой не равны нулю. Не равный нулю минор точно имеется в первых $(n - 1)$ столбцах матрицы А, поскольку определитель матрицы A не равен нулю (матрица невырожденная) и столбцы линейно независимы.

Перестановка строк (столбцов) равносильна умножению исходной матрицы A слева (справа) на **матрицу перестановок**, которая получается из единичной матрицы при аналогичной перестановке строк (столбцов). Строки (столбцы) могут быть возвращены в исходное состояние при обратной перестановке, матрица которой является обратной к матрице перестановки. Матрица перестановки ортогональна.

Теорема. Для произвольной невырожденной матрицы A существуют следующие разложения: $A = PLU$, $A = PMDV$, $A = L \times U \times Q$, $A = M \times D \times VQ$, где P и Q — матрицы перестановок соответственно строк и столбцов, U, U^*, V,V^* — верхние треугольные матрицы с единицами на главной диагонали, L, L^*, M, M^* — нижние треугольные матрицы, причем матрицы M и M^* имеют единицы на диагонали, а D и D^* — диагональные матрицы.

При перестановке строк (или столбцов) стремятся к тому, чтобы в левый верхний угол матрицы попадал бы элемент, не равный нулю. Такой элемент называется главным (ведущим). Для каждого шага (при необходимости перестановок) будет свой главный элемент.

Для увеличения точности вычисляемого решения в качестве главного элемента желательно брать максимальный по модулю элемент преобразуемой части матрицы или очередной строки (столбца). Существуют и другие критерии выбора главного элемента, зависимые от вида исходной матрицы и возможностей ЭВМ.

Масштабирование — это умножение строк или столбцов на числовые множители. Масштабирование строк (столбцов) равносильно умножению исходной матрицы слева (справа) на диагональную матрицу. Переход к другим единицам измерения равносилен масштабированию столбцов. Применение масштабирования не приводит к дополнительным ошибкам округления, поскольку обычно в качестве множителей для масштабирования применяются степени основания системы счисления. Масштабирование часто применяется при выборе главного элемента, ведь оно позволяет сделать любой ненулевой элемент максимальным по модулю.

Справедливо следующее утверждение: пусть две линейные системы решаются по схеме единственного деления, причем для обеих систем на каждом шаге берутся в качестве главных элементы, занимающие в матрицах одинаковые позиции. Тогда вычисленные решения обеих систем связаны равенством $x_i^* = 10^{-q_i} x_i$, $i = 1, 2, ..., n$, а мантиссы x_i и x_i^* совпадают, если при решении одной из систем не появится машинный нуль или не возникнет переполнения.

5. Разложение на ортогональный и треугольный множители (QR-разложение)

Кроме LU-разложения, применяется QR-разложение матрицы A линейной системы, где Q — ортогональная матрица, R — верхняя треугольная матрица. В случаях, когда найдено QR-разложение матрицы A, исходная система $AX = B$ сводится к системе вида $RX = Q^T B$, имеющей треугольную матрицу и легко решаемой методом Гаусса.

QR-разложение матрицы A можно получить **методом отражений**. **Отражение** в евклидовом пространстве — это самосопряженное преобразование, имеющее собственное значение 1 кратности $n - 1$ и собственное значение -1 кратности 1. Следовательно, в ортонормированном базисе из собственных векторов матрица отражения имеет вид:

$$P_o = \begin{pmatrix} -1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix} = E - 2E_1,$$

где E_1— матрица с единицей в левом верхнем углу и остальными элементами — нулями. P_0 — ортогональная матрица. В произвольном ортонормированном базисе матрица отражения равна $\mathrm{P} = \mathrm{E} - 2\sigma\sigma^{\mathrm{т}}$, где σ — столбец координат определяющего отражение вектора.

Справедливы следующие утверждения.

1. Каковы бы ни были ненулевой вектор x и вектор f, по длине равный единице, найдется такое отражение, которое переводит вектор x в вектор αf, причем $|\alpha| = |x|$.

2. Квадратная матрица A может быть разложена в произведение QR, где Q — ортогональная, а R — верхняя треугольная матрицы.

3. Если матрица A невырожденная, то ее QR-разложение, в котором диагональные элементы матрицы R положительны, единственно.

QR-разложение матрицы A может применяться для вычисления определителя (детерминанта) и для обращения матрицы: $\det A = \det Q \det R = (-1)^h \det R$, где h — число произведенных отражений, det R равен произведению диагональных элементов матрицы R, $A^{-1} = \mathrm{R}^{-1}Q^T$.

Кроме метода отражений, QR-разложение матрицы А можно получить **методом вращений**. Вращением в плоскости i-ого и j-ого векторов ортонормированного базиса ($i > j$) называется преобразование с матрицей, у которой элементы с индексами ii, jj равны $\cos\varphi$, с индексами ij, ji равны $\pm\sin\varphi$ (φ — угол поворота в плоскости i-ого и j-ого векторов), а остальные элементы равны соответствующим элементам единичной матрицы:

$$P = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & \cos\varphi & & & -\sin\varphi & & \\ & & & \ddots & & & & \\ & & & & 1 & & & \\ & & & & & \ddots & & \\ & & \sin\varphi & & & \cos\varphi & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}.$$

Матрицу R можно привести к верхнему треугольному виду последовательностью матриц вращения, содержащей для каждого столбца матрицы A столько матриц, сколько не равных нулю элементов содержится в этом столбце под элементом главной диагонали.

При использовании метода вращений начинают работу с поворота в плоскости первого и второго векторов ортонормированного базиса. Вращение выбирается так, чтобы элемент a_{12} обратился в нуль. Затем следует вращение в плоскости первого и третьего векторов базиса, которое должно обратить в нуль элемент с индексом 1, 3. В конце концов, в результате $(n - 1)$ вращений получается матрица со всеми нулями в первом столбце, кроме верхнего элемента (индекс 1, 1). Затем аналогичные вращения производятся с элементами второго, третьего и так далее столбцов, при этом нули предыдущих столбцов не меняются, и обращаются в нуль элементы рассматриваемого столбца, лежащие ниже соответствующего диагонального элемента.

Методы вращений и отражений дают почти одинаковую точность, но метод вращений требует большего объема памяти и использует большее число арифметических операций. Точность этих методов не зависит от выбора главного элемента.

6. Итерационные методы решения линейных систем. Метод простой итерации

Итерационные методы позволяют построить некоторую последовательность, сходящуюся к искомому решению. За приближенное решение исходной системы принимается максимально близкий к пределу член полученной последовательности. Несмотря на это, итерационные методы позволяют определить решение даже с большей точностью, чем прямые методы. Их часто применяют для проверки решений, полученных прямыми методами. Не для всех линейных систем возможно применение итерационных методов, к некоторым эти методы могут быть применены только после необходимых преобразований исходной системы.

Пусть имеется линейная система вида $AX = B$. Преобразуем ее следующим образом:

$$0 = -AX + B, X = X - AX + B, X = (E - A)\,X + B.$$

Или в общем виде при умножении на невырожденную матрицу H имеем:

$$X = X + H(B - AX).$$

Отсюда получаем вид рекуррентной формулы:

$$X_{k+1} = X_k + H(B - AX_k).$$

Данное выражение соответствует стационарному методу простой итерации. Если его представить в виде

$$S\frac{X_{k+1} - X_k}{\tau} + AX_k = B,$$

где $H = \tau S^{-1}$, то получается выражение, соответствующее общему неявному методу простой итерации. Можно обобщить эти два выражения следующей формулой:

$$X_{k+1} = PX_k + F,$$

где $P = E - HA = E - \tau S^{-1}A$, а $F = HB = \tau S^{-1}B$.

Если X^* — решение исходной линейной системы, то равенство $X_{k+1} = PX_k + F$ можно преобразовать следующим образом:

$$X_{k+1} - X^* = PX_k - X^* + F.$$

Поскольку

$$F = HB = HAX^*,$$

имеем

$$X_{k+1} - X^* = = P(X_k - X^*).$$

Если обозначить разность $D_k = X_k — X^*$, то имеем: $D_k = P_k D_0$. Отсюда видно, что последовательность $\{Dk\}$ сходится при любом D_0 только в случае сходимости последовательности степеней матрицы P. Действительно, $D_k \to QD_0$, где $Q = \lim_{k\to\infty} P^k$. Или $X - X^* = QD_0$, где X является пределом последовательности $\{X_k\}$. Но поскольку X является решением исходной системы, то $AQD_0 = 0$ (потому что QD_0 удовлетворяет приведенной однородной системе). Поскольку D_0 — любое, то $AQ = 0$.

Справедливы следующие утверждения.

1. Если матрица A линейной системы $AX = B$ не вырождена, то последовательность столбцов, задаваемая рекуррентной формулой

$$X_{k+1} = PX_k + F,$$

сходится при любом начальном векторе X_0 тогда и только тогда, когда спектральный радиус матрицы P меньше единицы (спектральным радиусом матрицы называется максимальное по модулю из характеристических чисел этой матрицы). Чем меньше спектральный радиус матрицы P, тем итерационный процесс сходится быстрее.

2. Если норма матрицы $\|P\| < 1$, то рекуррентная последовательность, задаваемая формулой

$$X_{k+1} = PX_k + F,$$

при любом начальном векторе сходится не медленнее, чем сумма геометрической прогрессии со знаменателем $\|P\|$.

Для ускорения итерационного процесса матрица H (или параметр τ и матрица S) выбирается из условия минимальности нормы $\|E - HA\|$ или максимальной приближенности к матрице A^{-1}. В случае равенства $H = A^{-1}$ потребовалась бы одна итерация.

Если каждый диагональный элемент исходной матрицы A по модулю больше суммы модулей остальных элементов той же строки (случай доминирующей главной диагонали), в качестве матрицы H берется матрица G^{-1}, где диагональная матрица $G = diag(a_{11}, ..., a_{nn})$. Это частный случай метода простой итерации — **метод Якоби**.

Для положительно определенной симметрической матрицы исходной системы матрица H берется в виде $H = 2E/\alpha$. В этом случае матрица $P = E - 2A/\alpha$, а ее характеристические числа принадлежат интервалу от -1 до 1.

7. Метод Зейделя. Метод верхней релаксации

Метод последовательных смещений (метод Зейделя) — частный случай общего неявного метода простой итерации, а именно случай, когда в выражении

$$S\frac{X_{k+1} - X_k}{\tau} + AX_k = B$$

считается, что $\tau = 1$, а матрица S является нижней треугольной матрицей, получаемой из матрицы A исходной системы при замене на нули элементов, лежащих выше главной диагонали:

$$S = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \ddots & 0 \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Матрица S будет невырожденной при отличных от нуля элементах главной диагонали. В этом случае формула простой итерации принимает вид: $S(X_{k+1} - X_k) + AX_k = B$ или $SX_{k+1} + UX_k = B$, где $U = A - S$ — матрица, в которой диагональные элементы и элементы, находящиеся под главной диагональю, равны нулю. Если характеристические числа матрицы $E - S^{-1}A$ (или матрицы $-S^{-1}U$) по абсолютной величине меньше единицы, то метод Зейделя сходится, а сами характеристические числа являются корнями уравнения $\det(U + \lambda S) = 0$.

Справедливо утверждение: для сходимости метода Зейделя достаточно, чтобы матрица A была симметричной и положительно определенной.

По сравнению с методом Якоби метод Зейделя сходится быстрее, а введением специального параметра ω по методу верхней релаксации его сходимость можно еще более ускорить. Метод верхней релаксации определяется выражением $S\frac{X_{k+1}-X_k}{\tau}+AX_k=B$ при $\lambda = \omega$ и $S = D + \omega L$, здесь матрица A представлена в виде суммы матриц $A = L + D + U$, где L и U — матрицы, в которых элементы, расположенные соответственно выше или ниже главной диагонали, равны нулю, а остальные — равны соответствующим элементам матрицы A, а D — диагональная матрица diag $(a_{11},.., a_{nn})$. Следовательно, имеем выражение $(D + \omega L)X_{k+1} = [(1 - \omega)D - \omega U]X_k + \omega B$. Для метода Зейделя параметр ω равен единице. Для сходимости метода Зейделя спектральный радиус матрицы $P = E - (D + + \omega L)^{-1}\omega A$ должен быть меньше единицы, причем характеристические числа матрицы P являются корнями уравнения $\det[(1-\omega)D - \omega U - \lambda(D + \omega L)] = 0$. Если ввести новую переменную η следующим образом: $\lambda = (\eta + 1)/(\eta - 1)$, то из предыдущего получаем следующее уравнение: $\det[-\omega\eta \mathrm{A} - (2 - \omega)\mathrm{D} + \omega(\mathrm{U} - \mathrm{L})] = 0$.

Справедливы следующие утверждения.

1. Метод верхней релаксации сходится тогда и только тогда, когда вещественные части всех корней уравнения $\det[-\omega\eta \mathrm{A} - (2 - \omega)D + \omega(U - L)] = 0$ отрицательны ($\mathrm{Re}\,\eta < 0$).

2. Для сходимости метода верхней релаксации достаточно, чтобы матрица A была симметричной и положительно определенной и чтобы $\omega \in (0, 2)$.

ЛЕКЦИЯ № 12. Псевдорешения и псевдообратные матрицы

1. Псевдорешения

На практике довольно часто бывают случаи, когда необходимо решить какую-либо задачу, описываемую системой линейных уравнений, но при этом должны еще удовлетворяться различные условия. Наличие таких условий зачастую приводит к тому, что линейная система оказывается несовместной. Тогда, чтобы все-таки получить наиболее оптимальное решение, приходится какие-то условия принимать целиком, какие-то частично, от каких-то, может, вообще придется отказаться.

Рассмотрим систему линейных уравнений вида $AX = B$. Пусть матрица A имеет размерность $m \times n$, r — ее ранг. Считаем, что норма X (столбца X) равна $\|X\| = \sqrt{X^T X} = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$.

Столбец $U = B - AX$ называется невязкой. Для решения невязка равна нулю. Если исходная линейная система $AX = B$ несовместна, то ее обобщенным решением будет такой столбец X, при котором норма невязки минимальна (для совместной системы обобщенным решением является ее решение). Норму невязки можно записать следующим образом: $\|U\|^2 = (B - AX)^T(B - AX)$. Известно, что, чтобы найти минимум какой-нибудь функции, надо приравнять к нулю дифференциал (производную) этой функции. Дифференциал нормы невязки равен: $d\|U\|^2 = -dX^TA^T(\mathrm{B} - AX) - (B - AX)^TAdX = -2dX^TA^T(B - AX)$ (поскольку второе слагаемое не меняется при транспонировании). Данный дифференциал может быть равен нулю только в случае, когда выполняется равенство $A^TAX = A^TB$. Полученное уравнение (линейная система) называется **нормальной системой** для системы $AX = B$.

Справедливы следующие утверждения.

1. Нормальная система уравнений обязательно совместна. Действительно, для системы $A^TAX = A^TB$ однородная система имеет вид $A^TAY = 0$ (ведь A^TA — симметричная матрица). Тогда верно: $Y^TA^TAY = (AY)^T(AY) = 0$, $AY = 0$, $YT(A^TB) = 0$.

2. Точная нижняя грань квадрата нормы невязки достигается для всех решений нормальной системы и только для них (доказывается при вычислении квадрата нормы U для столба $X + \Delta X$).

3. Нормальная система имеет единственное решение тогда и только тогда, когда система $AZ = 0$ имеет только тривиальное решение, т. е. столбцы матрицы А линейно независимы, в частности если матрица А не вырождена (здесь Z — произвольное решение однородной системы $A^T AZ = 0$).

Нормальное псевдорешение (псевдорешение) линейной системы — это минимальный по норме столбец среди всех столбцов, которые дают минимальную по норме невязку при подстановке в эту систему.

Теорема. Каждая система линейных уравнений имеет одно и только одно нормальное псевдорешение (т. е. требуется доказать, что существует только один столбец с минимальной нормой).

Единственное решение нормальной системы — это нормальное псевдорешение вида $X_0 = A^T Z$. Если X_B и X_C являются нормальными псевдорешениями соответственно линейных систем $AX = B$ и $AX = C$, то $\alpha X_B + \beta X_C$ будет нормальным псевдорешением системы $AX = \alpha B + \beta C$. Аналогичное утверждение справедливо для линейных комбинаций произвольного числа столбцов.

Например, для уравнения $ax = c$ (его можно рассматривать как систему одного линейного уравнения с одним неизвестным) при а $\neq 0$ его решение $x = c/a$ будет являться и псевдорешением, а при $a = 0$ псевдорешение равно нулю; для уравнения $x_1 + x_2 = 2$ нормальная система имеет вид: $\begin{pmatrix}1\\1\end{pmatrix}(1,\ \ 1)\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix}2$, ее общее решение равно $\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}2\\0\end{pmatrix} + \alpha\begin{pmatrix}-1\\1\end{pmatrix}$, псевдорешение равно $\begin{matrix}x_1\\x_2\end{matrix} = \begin{pmatrix}1\\1\end{pmatrix}$.

2. Псевдообратные матрицы

Псевдообратная матрица для матрицы A размеров $m \times n$ называется матрица A^+, столбцы которой являются псевдорешениями систем линейных уравнений вида $AX = E_i$, где E_i — столбцы единичной матрицы порядка m. Псевдообратная матрица для A имеет размерность $n \times m$ (совпадает с размерностью матрицы A^T). Для каждой матрицы A имеется только одна псевдообратная матрица. Если матрица A — невырожденная квадратная, то ее обратная матрица является и псевдообратной, псевдорешение совпадает с решением. Для нулевой матрицы размером $m \times n$ псевдоматрица — нулевая матрица размером $n \times m$.

Псевдорешение линейной системы $AX = B$ может быть записано в виде $X_0 = A^+B$, поскольку столбец свободных членов является линейной комбинацией столбцов единичной матрицы порядка m, а псевдорешение X_0 — линейная комбинация столбцов псевдообратной матрицы, имеющая такие же коэффициенты.

Интересно следующее свойство псевдообратной матрицы: для любой матрицы X размеров $n \times m$ выполнено соотношение между нормами: $\|AA^+ - E\| \le \|AX - E\|$, причем если для какой-нибудь матрицы X, отличной от A^+, имеет место равенство, то $\|A^+\| < \|X\|$ (норма матрицы евклидова пространства есть квадратный корень из суммы квадратов ее элементов).

Справедливо следующее утверждение: матрица X является псевдообратной для матрицы A тогда и только тогда, когда $A^TAX = A^T$ и существует такая квадратная матрица Z, что $X = A^TZ$. Из данного утверждения следует, что для любого не равного нулю числа c справедлива формула $(cA)^+ = c^{-1}A^+$; и для ортогональной матрицы U справедливы выражения: $(UA)^+ = A^+ U^T$ (если матрица U имеет порядок m) или $(AU)^+ = U^TA^+$ (если матрица U имеет порядок n). Если матрица U будет не ортогональной, а невырожденной квадратной, то приведенные равенства уже не будут справедливы. Кроме того, на псевдообратные матрицы не распространяется формула обращения произведения, т. е. $(AB)^+ \neq B^+A^+$.

Если строки или столбцы матрицы A линейно независимы, то для такой матрицы легко найти псевдообратную матрицу. Имеют место следующие выражения: $A^+ = (A^TA)^{-1}A^T$ (столбцы матрицы A линейно независимы) или $A^+ = A^T(AA^T)^{-1}$ (строки матрицы A линейно независимы). Действительно, пусть, например, линейно независимы столбцы матрицы A. Тогда нормальная система $A^TAX = A^TB$ имеет единственное решение вида $(A^TA)^{-1}A^TB$. Если вместо B брать столбцы единичной матрицы соответствующего порядка, то получаем первое выражение. Пусть столбцы матрицы A являются линейно независимыми, тогда минимальное по норме решение системы $AX = B$ (эта система совместна при любой B) является одновременно и ее псевдорешением и имеет вид $X = A^TZ$, где Z можно взять в виде $Z = (AA^T)^{-1}B$. Тогда нормальное псевдорешение имеет вид: $X = A^T(AAT)^{-1}B$, и если сюда подставить вместо B столбцы единичной матрицы, то получается второе выражение для псевдообратной матрицы. Если матрица A произвольная размерами $m \times n$, то применяется **скелетное разложение** матрицы A вида $A = BC$, где матрицы B и C имеют соответственно размеры $m \times r$ и $r \times n$, r — ранг матрицы A. Скелетное разложение (построение матриц B и C) можно произвести, например, следующим образом: матрица B составляется из столбцов базисного минора матрицы A (ее размер равен $m \times r$). Матрица C составляется из записанных в столбцы коэффициентов линейной комбинации столбцов матрицы B, линейная комбинация составляется для столбцов матрицы A (каждый столбец матрицы A — линейная комбинация столбцов матрицы B), получается матрица С размерами $r \times n$. Существуют и другие способы получения скелетного разложения.

Для матрицы, имеющей скелетное разложение, псевдообратная матрица определяется по следующей формуле: $A^+ = C^+B^+ = C^T(CC^T)^{-1}(B^TB)^{-1}B^T$. Из этого выражения следует, что для псевдообратных матриц справедливо $(A^+)^T = (A^T)^+$.

Общее решение нормальной системы $A^TAX = A^TB$ для исходной системы $AX = B$ определяется следующей формулой: $X = A^+B + (E - A^+A)c$, где c — произвольный столбец высоты n. Для совместной исходной системы приведенная формула дает ее общее решение. Действительно, столбец A^+B — нормальное псевдорешение, т. е. частное решение нормальной системы. Столбец $Z = (E - A^+A)\ c$ является общим решением нормальной однородной системы, поскольку: $A^TA\ ((E - A^+A)\ c) = A^TAc - A^TAA^+Ac = A^TAc - A^TAc = 0$ и для любого решения $Z\ (A^TAZ = 0)$ найдется c, что $Z = (E - A^+A)\ c$. Тогда, взяв $c = Z$, получаем $(E - A^+A)\ Z = Z - A^+AZ = Z$.

3. Метод наименьших квадратов

Метод наименьших квадратов является некоторым приложением теории псевдорешений и псевдообратных матриц. Он часто применяется в случаях, когда требуется решить задачу с числом переменных, значительно меньшим числа уравнений. Например, имеются две переменные величины η и ξ, значения которых определяются экспериментально. Требуется определить функцию f, максимально соответствующую реальности, такую, что $\eta = f(\xi)$. Данная задача решается с помощью заранее заданных базисных функций $\varphi_1(\xi)$, ..., $\varphi_n(\xi)$, вид которых выбирается с учетом реальных условий решаемой задачи. Функция f ищется в виде линейной комбинации этих базисных функций. Обычно базисные функции выбираются так, чтобы они были линейно независимыми.

Суть метода наименьших квадратов состоит в том, что функция f должна быть такова, чтобы сумма квадратов

$$(\eta_1 - f(\xi_1))^2 + \ldots + (\eta_m - f(\xi_m))^2$$

была минимальна; считается, что пары η_i, ξ_i лишь приближенно отражают реальную зависимость. Условие минимальности суммы квадратов можно рассматривать как минимальность квадрата нормы невязок $\eta_i - f(\xi_i)$ ($i = 1, 2, \ldots, m$). Определяемая функция f будет лишь аппроксимировать реальность.

Поскольку искомая функция по предположению является линейной комбинацией базисных функций, можно записать

$$f(\xi) = \beta_1 \varphi_1(\xi) + \ldots + \beta_n \varphi_n(\xi),$$

что равносильно следующей системе уравнений:

$$\begin{cases} a_{11}\beta_1 + \ldots + a_{1n}\beta_n = \eta_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}\beta_1 + \ldots + a_{mn}\beta_n = \eta_m \end{cases}.$$

Здесь величины $a_{ij} = \varphi_j(\xi_\mathrm{i})$, их можно считать соответствующими элементами матрицы A. В случае, когда столбцы матрицы A линейно независимы, нормальная система будет иметь единственное решение, являющееся псевдорешением полученной системы. Однако при линейной независимости базисных функций столбцы матрицы A не обязательно будут линейно независимыми, они могут оказаться близкими к линейно-зависимым.

Справедливо следующее утверждение, позволяющее оценить увеличение суммы квадратов: пусть δ_i — столбцы, удовлетворяющие условию $A^T A\delta\mathrm{i} = \alpha_i^2\delta_i$ ($\alpha < \varepsilon$, $i = k, \ldots, r$). Тогда если к решению нормальной системы прибавить произвольную линейную комбинацию $\mathrm{H} = h_k\delta_k + \ldots + h_r\delta_r$ столбцов δ_i, то сумма квадратов $(\eta_1 - f(\xi_1))^2 + \ldots + (\eta_m - f(\xi_m))^2$, соответствующих решению β, увеличится не более чем на $\varepsilon^2\|H\|^2$ (при доказательстве оценивается квадрат норма невязки, даваемой столбцом $\beta + H$).

Следовательно, функция f оценивается по экспериментальным данным (или приближенным измерениям) своих значений. В более общей постановке эта задача относится к области математической статистики, когда математическое ожидание случайной величины рассматривается как линейная комбинация от n переменных, значения которых являются элементами матрицы A (матрица А называется матрицей регрессии, а переменные — регрессорами).

ЛЕКЦИЯ № 13. Системы линейных неравенств

1. Однородные системы линейных неравенств

Пусть дана система неравенств вида

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n \geq 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \geq 0 \end{cases}$$

с вещественными коэффициентами. Если в системе будут неравенства разного типа, то умножением на -1 (где необходимо) все неравенства можно привести к одному типу (≥ 0). В матричном виде данная система имеет вид $AX \geq 0$. **Жестким** называется неравенство этой системы в случае, если оно выполняется как равенство для всех решений системы.

Замкнутым полупространством называется множество векторов, координаты которых удовлетворяют однородному линейному неравенству $a_1x_1 + \ldots + a_nx_n \geq 0$. В случае строгого неравенства полупространство называется **открытым**. В случае равенства имеем **граничное подпространство**.

Выпуклый многогранный конус (конус) K — это пересечение конечного числа полупространств. Например, в трехмерном пространстве конусами являются двугранный угол ($x_1 \geq 0, x_2 \geq 0$), неотрицательный квадрант ($x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$), плоскость ($x_1 \geq 0, -x_1 \geq 0$), четырехгранный конус ($x_1 \geq 0$, $x_2 \geq 0$, $x_3 \geq 0$, $x_1 + x - x_3 \geq 0$), плоский угол ($x_1 \geq 0$, $x_2 \geq 0$, $x_3 \geq 0$, $x_1 + x - x_3 \geq 0$, $-x_1 - x + x_3 \geq 0$); конусами также являются одномерное подпространство, луч, нулевой вектор.

Если векторы x_1 и x_2 принадлежат выпуклому многогранному конусу K, то векторы $x_1 + x_2$ и cx_1 для любого $c \geq 0$ также принадлежат конусу K (следует из непосредственной подстановки координат векторов в задающую конус систему неравенств).

Справедливы следующие утверждения.

1. Если r — ранг матрицы, составленной из коэффициентов жестких неравенств, то конус K, определяемый всей системой неравенств, лежит в $(n - r)$-мерном подпространстве и определяется в нем системой, которая не содержит жестких неравенств.

2. Если система однородных линейных неравенств не содержит жестких неравенств, то соответствующая система строгих неравенств

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n > 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n > 0 \end{cases}$$

имеет решение. Действительно, поскольку все неравенства (нестрогие) исходной системы однородных линейных неравенств нежесткие, то для любого i-го неравенства системы найдется такое решение X_i, которому бы соответствовало соответствующее строгое неравенство. При подстановке в i-ое неравенство столбца

$$X_1 + \ldots + X_m$$

получаем обязательно строгое неравенство, поскольку слагаемое с X_i строго больше нуля, а остальные слагаемые неотрицательны. Значит, $X_1 + \ldots + X_m$ — это решение системы строгих неравенств.

Если в исходной системе однородных (нестрогих) неравенств нет жестких неравенств, то векторы, удовлетворяющие системе строгих неравенств, называются **внутренними векторами конуса**. **Внутренность конуса** — это множество всех внутренних векторов. **Относительной внутренностью конуса** называется множество решений системы строгих неравенств. Каждый внутренний вектор конуса принадлежит ему вместе с некоторой своей окрестностью относительно произвольной нормы.

Размерность конуса K равна размерности вещественного линейного пространства L_n, если система линейных неравенств, соответствующая конусу, не содержит жестких неравенств. Действительно, если вектор $X_0(x_{01}, \ldots, x_{0n})$ — внутренний вектор конуса K, то для любого положительного ε возьмем η такое, что $|\eta| < \varepsilon$ (причем η не является характеристическим числом матрицы, все строки которой совпадают со строкой $x_{01}, ..., x_{0n}$), и тогда векторы с координатами

$$\begin{array}{cccc} x_{01}-\eta, & x_{02}, & \cdots, & x_{0n} \\ x_{01}, & x_{02}-\eta, & \cdots, & x_{0n} \\ \cdots, & \cdots, & \cdots, & \cdots \\ x_{01}, & x_{02}, & \cdots, & x_{0n}-\eta \end{array}$$

будут линейно независимыми и будут принадлежать конусу K.

Если исходная система однородных нестрогих неравенств содержит жесткие неравенства, то размерность конуса K, соответствующего этой исходной системе, равна (n — r), где r — ранг матрицы, составленной из коэффициентов жестких неравенств.

Если в исходной системе некоторые жесткие неравенства заменить на равенства, то для новой системы будет свой конус, размерностью меньше. Такие конусы являются **гранями** конуса K. **Минимальная грань** будет соответствовать системе линейных уравнений

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = 0. \end{cases}$$

Если минимальная грань является нулевым подпространством, то соответствующий ей конус называется **заостренным**. Если в исходной системе нестрогих линейных неравенств имеется только одно нежесткое независимое неравенство, а остальные неравенства обращены в равенства, то имеем одномерную грань — **луч**, которую называют еще **ребром** многогранного конуса. Заостренный конус является суммой своих ребер.

Теорема. Каждый замкнутый выпуклый многогранный конус является суммой конечного числа лучей, или, что то же, совокупностью линейных комбинаций конечного числа векторов с неотрицательными коэффициентами. Иная формулировка теоремы: Общее решение системы однородных линейных неравенств может быть записано в виде $X = \alpha_1 X_1 + \ldots + \alpha_N X_N$, где $X_1, ..., X_N$ — некоторый набор решений, а коэффициенты $\alpha_1, ..., \alpha_N$ — неотрицательны. Данный набор решений называется **полной системой**

решений, **фундаментальной системой** решений называется минимальная по количеству полная система решений.

2. Неоднородная система линейных неравенств

Пусть дана система неравенств вида

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n \geq b_1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \geq b_m \end{cases}$$

с вещественными коэффициентами. Это есть неоднородная система линейных неравенств. Считаем, что все неравенства данной системы одного типа. Если в системе будут неравенства разного типа, то умножением на -1 (где необходимо) все неравенства можно привести к одному типу (≥ 0). В матричном виде данная система имеет вид $AX \geq B$. Геометрически решение такой системы — точка в аффинном (точечном) пространстве, множество решений — плоскость, решение приведенной однородной системы — векторы.

Множество решений неоднородной системы неравенств выпукло (т. е. в этом множестве для каждых двух точек содержится и соединяющий их отрезок). **Выпуклыми комбинациями** X называются такие линейные комбинации решений исходной системы неравенств

$$X = \alpha_1 X_1 + \ldots + \alpha_n X_n,$$

для коэффициентов которых выполняются следующие условия:

$$\alpha_1 \geq 0, \ldots, \alpha_n \geq 0, \alpha_1 + \ldots + \alpha_n = 1.$$

Выпуклой оболочкой некоторого множества называется множество всевозможных выпуклых комбинаций точек исходного множества. Например, отрезок являюется выпуклой оболочкой своих концов. Выпуклая оболочка любого множества является выпуклым множеством.

Выпуклое многогранное множество — это пересечение полупространств аффинного пространства; если это множество ограниченно, то оно называется **выпуклым многогранником**. Именно выпуклым многогранным множеством и является множество решений неоднородной системы линейных неравенств.

Для того чтобы определить решение линейной неоднородной системы неравенств, приведем эту систему к однородному виду. Для этого добавим еще одну переменную x_{n+1} и исходную систему неравенств запишем в следующем виде:

$$AX - x_{n+1}B \geq 0, x_{n+1} \geq 0$$

(добавилось еще одно неравенство, т. е. как бы перешли от n-мерного пространства к $(n+1)$-мерному пространству). Если $x_{n+1} = 1$, то первые n элементов получившейся однородной системы неравенств будут удовлетворять исходной системе неравенств $AX \geq B$.

Пусть P_1, ..., P_s — фундаментальная система решений однородной системы неравенств

$$AX - x_{n+1}B \geq 0, x_{n+1} \geq 0.$$

Считаем, что в столбцах первых t решений последние компоненты равны единице (если не равны, то к этому можно прийти элементарными преобразованиями), а в столбцах оставшихся решений последние компоненты равны нулю. Если исходная система неоднородных неравенств $AX \geq B$ совместна, то $t \geq 1$.

Решения однородной системы

$$AX - x_{n+1}B \geq 0, x_{n+1} \geq 0$$

имеют вид:

$$Q = \alpha_1 P_1 + ... + \alpha_t P_t + \beta_1 P_{t+1} + \beta_s P_s,$$

где коэффициенты α_i, $\beta_j \geq 0$ (неотрицательные), причем последняя компонента Q равна единице только при выполнении условия: $\alpha_1 + ... + \alpha_s = 1$. Тогда общее решение системы неоднородных линейных неравенств $AX \geq B$ получается при отбрасывании последних элементов всех столбцов в выражении

$$Q = \alpha_1 P_1 + ... + \alpha_t P_t + \beta_1 P_{t+1} + \beta_s P_s.$$

Понятия «жесткое неравенство», «размерность выпуклого многогранного множества», «грань», «ребро» при рассмотрении неоднородной системы линейных неравенств определяются аналогично случаю системы однородных неравенств.

Неоднородные системы линейных неравенств не обязательно являются совместными.

Теорема. Система линейных неравенств $AX \geq$ В совместна тогда и только тогда, когда из $UA = 0$ и $U \geq 0$ следует, что $UB \leq 0$ (U — неотрицательная строка длиной m).

Доказательство. Если исходная система неоднородных неравенств совместна и X — ее решение, то для любой строки U, определенной в условии теоремы, справедливо $UAX \geq UB$. Тогда при $UA = 0$ получаем требуемое $UB \leq 0$. Для доказательства обратного надо показать, что для несовместной системы существует строка U такая, что $UA = 0$ и $UB > 0$.

Для неоднородных систем справедливы следующие утверждения.

1. Система линейных неравенств $AX \geq B$ совместна при любой правой части B тогда и только тогда, когда система уравнений $UA = 0$ имеет только тривиальное неотрицательное решение.

2. Система линейных неравенств $AX \geq B$ имеет неотрицательное решение $X \geq 0$ тогда и только тогда, когда из неравенств $UA \leq 0$ и $U \geq 0$ следует $UB \leq 0$.

Литература

1. Кострикин А. И. Введение в алгебру. М.: Наука, 1977.
2. Гусак А. А. Высшая математика. Минск: ТетраСистес, 2001.
3. Общая алгебра / Под ред. Л. А. Скорнякова. М.: Наука, 1990.
4. Беклемишев Д. Б. Дополнительные главы линейной алгебры. М.: Наука, 1983.
5. Выгодский М. Я. Справочник по высшей математике. М.: Наука, 1976.

Содержание