КМБО -19 1-й семестр

1-я лекция. Матрицы. 3.09.2019

Матрица A - это прямоугольная таблица, состоящая из чисел. У нее есть **строки**, **столбцы**, **элементы** a_{ij}; в такой записи i – номер строчки, j – номер столбца. Мы пишем, что матрица имет размер $[m \times n]$, если у нее m строчек, n столбцов. Если $m = n$, то матрица называется **квадратной**. У квадратной матрицы есть две **диагонали** - **главная** (из левого верхнего угла в правый нижний)и **побочная**.

Бывают матрицы-столбцы и матрицы-строки. Иногда удобно матрицу рассматривать как совокупность строк или столбцов:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ \dots \\ A_m \end{pmatrix} = \begin{pmatrix} A^1 & A^2 & \dots & A^n \end{pmatrix}.$$

Здесь $A_i = (a_{i1}\ a_{i2}\ \dots\ a_{in})$ – матрицы-строки, $A^j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{mj} \end{pmatrix}$ – матрицы-столбцы.

Будем называть две матрицы **равными**, если их размеры совпадают, а также совпадают все элементы, стоящих на одинаковых местах. Иными словами, $A = B$ – если $a_{ij} = b_{ij}$ для всех i, j.

Операции над матрицами.

Смысл этих операций станет ясен после того, как будут введены понятия линейного оператора в линейном пространстве, матрицы линейного оператора в некотором базисе, а также операций над операторами.

Если A и B – матрицы одинакового порядка $[m \times n]$, $A = (a_{ij})$; $B = (b_{ij})$, то **суммой** этих матриц $A + B$ называется матрица $C = (c_{ij})$ того же порядка, чьи элементы вычисляются по формуле

$$c_{ij} = a_{ij} + b_{ij}.$$

Заметим, что эта операция аналогична операции сложения векторов – там при сложении также происходит сложение координат.

Если A - матрица, а λ – число (пока мы ограничиваемся случаем действительных чисел), то **произведением** λA называется матрица того же порядка, чьи элементы получаются из элементов матрицы A умножением на λ:

$$\lambda(a_{ij}) = (\lambda a_{ij})$$

Аналогия с умножением вектора на число очевидна.

Введенные операции делают множество $\mathbb{R}^{m\times n}$ матриц порядка $[m \times n]$ так называемым **линейным пространством** – так называется множество $\mathcal{L}$, на котором заданы две операции – сложение элементов этого множества, и умножение элементов этого множества на числа – в нашем случае на действительные (вместо действительных чисел мы могли бы рассматривать элементы любого так называемого поля - о поле речь позже). Эти операции должны удовлетворять следующим 8 аксиомам:
1) $A + B = B + A \quad \forall A,\ B \in \mathcal{L}$ (коммутативность сложения)
2) $(A + B) + C = A + (B + C) \quad \forall A,\ B, C \in \mathcal{L}$ (ассоциативность сложения)
3) $\exists\ 0 \in \mathcal{L}$: $A + 0 = 0 + A = A \quad \forall A \in \mathcal{L}$(так называемый нулевой элемент линейного пространства; в множестве матриц роль нулевого элемента исполняет **нулевая матрица**, то есть матрица, состоящая только из нулей)
4) $\forall A \in \mathcal{L}\ \ \exists\ (-\mathcal{A})$: $A + (-A) = (-A) + A = 0$ (так называемый противоположный элемент; в множестве матриц роль матрицы, противоположной к A, играет матрица, полученная из матрицы A умножением на -1)
5) $1 \cdot A = A \quad \forall A \in \mathcal{L}$
6) $\lambda \cdot (\mu A) = (\lambda \cdot \mu)A \quad \forall A \in \mathcal{L};\ \lambda, \mu \in \mathbb{R}$
7) $(\lambda + \mu)A = \lambda A + \mu A \quad \forall A \in \mathcal{L};\ \lambda, \mu \in \mathbb{R}$
8) $\lambda(A + B) = \lambda A + \lambda B \quad \forall A, B \in \mathcal{L};\ \lambda \in \mathbb{R}$.

Теорема Множество матриц размера $[m \times n]$ является линейным пространством относительно введенных операций.

Переходим к **произведению** матриц. Оно задается не так очевидно, как две предыдущие операции. Если $A[m{\times}n]$; $B[n{\times}k]$, то есть длина строки первой матрицы равна длине столбца второй матрицы (только в этом случае произведение матриц будет определено), то $C = AB$ – это матрица размером $[m \times k]$, причем

$$c_{ij} = \sum_{l=1}^{n} a_{il}b_{lj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots a_{in}b_{nj}$$

Запомнить эту операцию может помочь такая аналогия: скалярное произведение двух векторов вычисляется как сумма произведений соответствующих координат (как окажется в дальнейшем, не всегда,но по крайней мере в так называемом ортонормированном базисе). Так вот, произведение матриц состоит из подобных образований: на (i, j)-ом месте матрицы C стоит "скалярное произведение"i-й строчки матрицы A – мы договорились обозначать её A_i, и j-го столбца матрицы B – мы договорились обозначать её B^j. Эти строчка и столбец рассматривются как элементы некоторого n-мерного пространства – ведь у них n координат.

Свойства произведения матриц.

1. $(AB)C = B(AC)$ (точнее, если одна из частей этого равенства существует, то существует и другая, и они равны друг другу.

Это элементарно проверяется выписыванием общего элемента левой части и общего элемента правой части. Лучше всего сначала проделать это на частном случае, так как техника работы со знаком суммирования у нас еще не развита. А потом можно записать это и в общем виде. Впрочем, это свойство станет совершенно очевидным, когда можно будет апеллировать к свойствам произведения операторов – для операторов левая и правая части сводятся к последовательному применению этих операторов. На самой лекции, если останется время, можно выкладку привести, а сейчас делать это лень.

2. $(\lambda A)B = \lambda(AB) = A(\lambda B)$.

Предлагается доказательство провести самостоятельно.

3. $A(B + C) = AB + AC; \quad (A + B)C = AC + BC$

Снова предлагается доказать эти равенства самостоятельно (дистрибутивность).

4. Матрица, состоящая только из нулей, называется **нулевой** независимо от ее размера. Ясно, что умножение произвольной матрицы на нулевую (с любой стороны) дает нулевую матрицу (при условии, что размеры позволяют произвести это умножение.

5. Квадратная матрица, у которой на главной диагонали стоят единицы, а остальные элементы равны нулю, называется **единичной** и обозначается буквой E. Пятое свойство говорит о том, что единичная матрица играет роль единицы в множестве матриц:

$$AE = A; \quad EA = A.$$

Конечно, если A – неквадратная матрица, то единичные матрицы в этих двух равенствах имеют разные размеры.

Продолжим выписывать свойства произведения матриц.

Наряду с приведенными "хорошими"свойствами, произведение матриц обладает и "плохими"свойствами (антисвойствами).

6. Произведение матриц, вообще говоря, некоммутативно, то есть не для всех матриц выполнено равенство $AB = BA$.

Можно выделить три уровня этого антисвойства.

Во-первых бывает так, что AB существует, а BA не существует. Такое происходит, когда $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times k}$, где $m \neq k$.

Во-вторых, бывает так, что AB и BA существуют, но имеют разные размеры. Такое происходит, когда $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times m}$, где $m \neq n$.

В-третьих, даже если оба произведения существуют и имеют одинаковые размеры, может оказаться, что $AB \neq BA$.

Пример. $\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} \cdot \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$, но $\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} \cdot \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} = \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$

7. В множестве матриц есть так называемые **делители нуля**, это такие ненулевые матрицы, чье произведение равно нулевой матрице.

Соответствующий пример можно найти тремя строчками выше.

7. Не для каждой ненулевой матрицы можно подобрать другую так, чтобы их произведение хотя бы в одном порядке равнялось единичной матрице.

Например, если взять матрицу $\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$, то умножая ее справа на любую матрицу получим матрицу с нулевой второй строчкой, а умножая слева на любую матрицу, получим матрицу с нулевым вторым столбцом.

Введем еще одну операцию с матрицами - **транспонирование**, при котором строки матрицы становятся столбцами (а столбцы, соответственно, – строчками). Этот процесс можно еще описать таким образом: при транспонировании элемент, стоящий на (i, j)-м месте, перемещается на (j, i)-е место. Матрицу, транспонированную матрице A, будем обозначать $A^\top$.

Свойства транспонирования.

1. $(A^\top)^\top = A$;
2. $(\lambda A)^\top = \lambda A^\top$;
3. $(A + B)^\top = A^\top + B^\top$.
4. Менее очевидно равенство

$$(AB)^\top = B^\top A^\top$$

Докажем его. Пусть $AB = C$; $C = (c_{ij})$; $C^\top = (c^\top_{ij})$.
$c^\top_{ij} = c_{ji} = \sum\limits_l a_{jl}b_{li} = \sum\limits_l b^\top_{il}a^\top_{lj}$, а это и есть элемент произведения матриц $B^\top$ и $A^\top$.
Первые три равенства предлагается доказать самостоятельно.

Рассмотрим несколько частных случаев квадратной матрицы. **Диагональная** матрица - это матрица, все элементы которой вне главной диагонали равны нулю. **Скалярная** матрица - это диагональная матрица, у которой на главной диагонали все элементы равны. **Единичная** матрица - это скалярная матрица, у которой на диагонали стоят единицы. **Верхнетреугольная** матрица - у нее все элементы ниже главной диагонали равны нулю. **Нижнетреугольная** матрица - у нее все элементы выше главной диагонали равны нулю. Верхнетреугольные матрицы и нижнетреугольные матрицы называются **треугольными** матрицами.

И, наконец, полезно ввести такое обозначение. Матрица, у которой на месте (i, j) стоит 1, а остальные элементы равны 0, обозначается E_{ij}.

Задача. Доказать, что $E_{ij} \cdot E_{pq} = \left[\begin{array}{l} 0, \text{ если } j \neq p \\ E_{iq}, \text{ если } j = p \end{array}\right.$

2-я и 3-я лекции. Определители. 10 и 17.09.2019

Определители возникают во многих ситуациях – например, при решении систем линейных уравнений, при нахождении площадей. Напишем здесь сразу общее определение, случаи определителей матриц второго и третьего порядков выведем из общей формулы.

Введем сначала понятие *перестановки* и *инверсии* в перестановке. Переставлять можно что угодно, но мы будем переставлять числа $1,\ 2,\ \ldots, n$. Располагая эти числа в том или ином порядке, будем получать различные перестановки из n элементов. Число перестановок обозначается как P_n.

Утверждение. $P_n = n!$
$\triangleleft\ \triangleright$

Инверсия в перестановке – это когда большее число стоит левее меньшего. **Четная** перестановка – если в ней четное число инверсий. **Нечетная** перестановка – если в ней нечетное число инверсий. Полезно ввести понятие **знака** перестановки sgn σ (σ – так мы обозначили перестановку). Пусть $\sigma = (i_1, i_2, \ldots, i_n)$; обозначим через $[i_1, i_2, \ldots, i_n]$ число инверсий в σ. Тогда по определению знак σ – это

$$\text{sgn }\sigma = (-1)^{[i_1, i_2, \ldots, i_n]}$$

Поэтому знак четной перестановки равен 1, знак нечетной перестановки равен -1. Докажем, что количество четных перестановок совпадает с количеством нечетных перестановок. Для этого нам понадобится еще одно понятие - **транспозиция** двух элементов в перестановке. Это когда мы меняем местами два числа (любые)в перестановке. Например, поменяв местами числа 3 и 5 в перестановке $[4, 3, 1, 2, 5]$, получаем перестановку $[4, 5, 1, 2, 3]$.

Утверждение. Применив транспозицию к перестановке, мы меняем ее четность.

$\triangleleft$ Если меняем местами соседние числа в перестановке, то изменение ее четности очевидна. Если же между этими числами есть другие числа, мы для доказательства утверждения не будем сразу применять нужную транспозицию. Можно добиться того же результата, несколько раз меняя соседние числа. Легко заметить, что при этом придется транспозицию производить нечетное число раз, причем каждый раз четность перестановки меняется. $\triangleright$

Утверждение. Если $n > 1$ – любое натуральное число, то четных перестановок из n элементов столько же, сколько и нечетных.

$\triangleleft$ Применив какую-то транспозицию ко всем четным перестановкам (одну и ту же для всех), получим такое же количество нечетных перестановок, причем очевидно, что все они будут разные.Значит, четных перестановок не больше, чем нечетных. Проделав то же самое с нечетными перестановками, доказываем, что нечетных перестановок не больше, чем четных. Тем самым утверждение доказано. $\triangleright$

Определение. Определителем матрицы $A = (a_{ij})$ $n-$го порядка (то есть квадратной матрицы $[n \times n]$) называется число $\det A$, задаваемое формулой

$$\det A = \sum_{(j_1, j_2, \ldots, j_n)} (-1)^{[j_1, j_2, \ldots, j_n]} a_{1j_1} a_{2j_2} \cdot a_{nj_n}$$

Иными словами, $\det A = \sum\limits_{\sigma = (j_1, j_2, \ldots, j_n)} \text{sgn }(\sigma) a_{1j_1} a_{2j_2} \cdot a_{nj_n}$

Расписав эту формулу для $n = 2$ и $n = 3$, получаем стандартные формулы для определителей 2-го и 3-го порядков.

Для матрицы второго порядка имеем:

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = (-1)^{[12]} a_{11} a_{22} + (-1)^{[21]} a_{12} a_{21} = a_{11} a_{22} - a_{12} a_{21}$$

Для матрицы третьего порядка имеем:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = (-1)^{[123]}a_{11}a_{22}a_{33} + (-1)^{[132]}a_{11}a_{23}a_{32} + (-1)^{[213]}a_{12}a_{21}a_{33} + (-1)^{[231]}a_{12}a_{23}a_{31} +$$

$$(-1)^{[312]}a_{13}a_{21}a_{32} + (-1)^{[321]}a_{13}a_{22}a_{31} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

Проанализировав получившуюся формулу, можно результат оформить в виде так называемой **формулы Саррюса**, которую легче изобразить мелом на доске или ручкой в тетради, чем в печатном тексте. Поэтому мы выпишем ее на лекции.

Для дальнейшего, в частности для доказательства первого свойства определителей, к которым мы сейчас переходим, полезно доказать следующее простое утверждение.

Утверждение. Если в одном из слагаемых, входящих в определение определителя, произвольным образом поменять местами множители, получив при этом произведение $a_{i_1,k_1}a_{i_2,k_2}\dots a_{i_n,k_n}$, то четность суммы числа инверсий в первых индексах и числа инверсий во вторых индексах

$$[i_1, i_2, \dots, i_n] + [k_1, k_2, \dots, k_n]$$

не зависит от того, в каком порядке мы расставили эти сомножители. Поэтому знак перед произведением можно искать как

$$\mathrm{sgn}\ (i_1, i_2, \dots, i_n) \cdot \mathrm{sgn}\ (k_1, k_2 \dots, k_n).$$

Это утверждение следует из доказанного утверждения о том, что четность перестановки меняется при транспозиции.

Свойства определителей.

1 $\det(A^\top) = \det(A)$ (при транспонировании определитель не меняется).

При транспонировании меняются местами первый и второй индексы, так что суммы инверсий в первых индексах и вторых индексах не изменится.

Следствие. Свойства определителя, справедливые для строчек, справедливы и для столбцов.

2. Если в матрице поменять местами 2 строки (или 2 столбца), то определитель изменит знак (не изменившись по абсолютной величине).

Доказательства утверждений такого типа проще разбирать самостоятельно. Но попробуем привести короткое обоснование здесь. Пусть поменялись местами i-я и j-я строчки, $i < j$. Была матрица A, получилась матрица B, причем элементы этих матриц вне рассматриваемых строчек совпадают: $b_{pq} = a_{pq}$, если $p \neq i$, $p \neq j$, а для рассматриваемых строчек $b_{ik} = a_{jk}$; $b_{jk} = a_{ik}$.

Рассмотрим одно из произведений, из которых складывается $\det B$ (а также складывается $\det A$, но, как мы сейчас докажем, в $\det B$ и в $\det A$ оно входит с разными знаками):

$$b_{1t_1}b_{2t_2}\dots b_{it_i}\dots b_{jt_j}\dots b_{nt_n} = a_{1t_1}a_{2t_2}\dots a_{jt_i}\dots a_{it_j}\dots a_{nt_n}$$

В левой части первые индексы расположены в естественном порядке, поэтому инверсии могут быть только во вторых индексах. В правой части вторые индексы такие же, как и в левой части, поэтому у них инверсий столько же, как и у вторых индексов левой части, в первых же индексах произошла одна транспозиция - поменялись местами индексы i и j. Поэтому суммарно по первым и вторым индексам сумма инверсий изменилась на 1. Это и доказывает, что указанное произведение входит в $\det B$ и $\det A$ с разными знаками.

3. Определитель матрицы с двумя равными строчками (или столбцами) равен нулю.

Это свойство является простым следствием второго свойства – если поменять местами равные строки, то определитель не изменится, так как ничего в нем не изменится, но он должен изменить знак в силу второго свойства.

4. Если все элементы i-ой строки(или стобца) умножить на λ, то $\det(A)$ умножится на λ.

$\lhd$ В каждое слагаемое (из которых состоит определитель) из каждой строчки входит ровно один элемент,в частности это верно для i-й строчки, той, которая умножена на число. $\rhd$

5 Определитель матрицы с нулевой строкой(или столбцом) равен нулю.

$\lhd$ $\rhd$

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

$\lhd$ $\rhd$

7. Если элементы i-ой строчки (или столбца) определителя разбиты в сумму двух чисел, то такой определитель равен сумме двух определителей - в одном из них в i-ой строчке стоят первые числа этих сумм, а во втором - вторые числа. Числа в остальных строчках при этом не меняются.

$\lhd$ $\rhd$

8. Определитель не изменится, если ко всем элементам одной строчки добавить элементы другой строчки, умноженные на одно и то же число.

$\lhd$ $\rhd$

9. Определитель треугольной матрицы равен произведению диагональных элементов (а значит и для скалярной матрицы верно это же утверждение).

$\lhd$ $\rhd$

10. Если матрица A имеет вид

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix},$$

где B и D – квадратные матрицы, возможно, разного размера (и если они разного размера, C – прямоугольная матрица, 0 - нулевая матрица подходящего размера), то

$$\det(A) = \det(B) \cdot \det(D)$$

Естественно, аналогичное свойство справедливо для матрицы, у которой нулевой угол не в левом нижнем углу, а в правом верхнем.Или оба они нулевые.

Разбираться в общем виде с этим свойством лучше в "тиши кабинет", но чтобы понять суть, достаточно рассмотреть, скажем, случай, когда матрицы B и D второго порядка (а тогда и матрицы C и 0 будут такими же).

Разложение определителя по строке или столбцу.

В принципе этот пункт можно считать просто очередным свойством определителя, но в силу его важности выделяем в отдельный раздел.

Вводим необходимые определения. **Минором** M_{ij} элемента a_{ij} определителя $\det(A)$ называется определитель матрицы, полученной из матрицы A вычеркиванием i-й строчки и j-го столбца (то есть строчки и столбца, в которых находится этот элемент). **Алгебраическим дополнением** A_{ij} элемента a_{ij} называется число

$$A_{ij} = (-1)^{i+j} M_{ij}.$$

Тем самым алгебраирческое дополнение или совпадает с соответствующим минором, или отличается от него знаком. Чтобы лучше запомнить, когда минус нужен, а когда нет, можете представить себе матрицу в виде шахматной доски. Для элементов, стоящих на клетках того же цвета, что и a_{11}, минус писать не надо, для остальных – надо.

Теорема.

$$\det A = \sum_{j=1}^{n} a_{ij} A_{ij} = \sum_{i=1}^{n} a_{ij} A_{ij}$$

Например, раскладывая по первой строчке, получаем

$$\det A = a_{11}A_{11} + a_{12}A_{12} + \ldots + a_{1n}A_{1n};$$

раскладывая по второму столбцу, получаем

$$\det A = a_{12}A_{12} + a_{22}A_{22} + \ldots a_{n2}A_{n2}.$$

Времени на аккуратное доказательство этого факта в общем виде у нас нет. Поэтому проиллюстрируем суть утверждения на примере определителя третьего порядка. Он состоит из шести слагаемых, три с плюсом, три с минусом. Чтобы получить разложение, скажем по первой строчке, объединим эти слагаемые по два по признаку совпадения множителя из первой строчки. Получим $\det A = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}$, что и требовалось.

Теорема о "ложном" определителе. Если мы (по ошибке или с намерением) заменим в теореме о разложении определителя по строке или столбцу алгебраические дополнения на алгебраические дополнения из параллельного ряда, то вместо определителя мы получим ноль.

Например, $a_{11}A_{21} + a_{12}A_{22} + \ldots + a_{1n}A_{2n} = 0$

Доказательство сводится к замене (в этом примере) второй строки на первую. Определитель при этом станет равен нулю (так как у него будут две одинаковые строчки), при этом выписанная сумма станет его разложением по второй строчке.

Теорема об определителе произведения квадратных матриц.

$$\det(AB) = \det(A) \cdot \det(B)$$

Времени на доказательство этой формулы у нас нет. Поэтому предлагается кому-нибудь взять доклад на эту тему, возможно с добавлением доказательства формулы Бине-Коши, показывающей, как ищется определитель произведения матриц порядка $m \times n$ и $n \times m$.

Определитель Вандермонда. Так называется определитель

$$W(x_1, x_2, \ldots, x_n) = \begin{vmatrix} 1 & 1 & 1 & \ldots & 1 \\ x_1 & x_2 & x_3 & \ldots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \ldots & x_n^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \ldots & x_n^{n-1} \end{vmatrix}$$

На семинарских занятиях мы научимся вычислять его.

Метод рекуррентных соотношений вычисления определителей.

Этот метод безупречно работает при вычислении "трехдиагональных"определителей, у которых все элементы равны нулю, кроме чисел на диагонали (все они равны a), над диагональю (все они равны b) и под диагональю (все они равны c). Впрочем, чаще всего встречаются определители, у которых $b = c$.

Заметим, что для понимания этого метода неплохо бы знать начала теории линейных пространств, а заодно и комплексные числа.

Выведем рекуррентное соотношение для такого определителя (здесь n – количество строк (и столбцов)):

$$D_n = \begin{vmatrix} a & b & 0 & 0 & \dots & 0 & 0 & 0 \\ c & a & b & 0 & \dots & 0 & 0 & 0 \\ 0 & c & a & b & \dots & 0 & 0 & 0 \\ 0 & 0 & c & a & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & a & b & 0 \\ 0 & 0 & 0 & 0 & \dots & c & a & b \\ 0 & 0 & 0 & 0 & \dots & 0 & c & a \end{vmatrix}.$$

Раскладывая его по первой строчке, а второй из получившихся определителей по первому столбцу, получаем соотношение

$$D_n = aD_{n-1} - bcD_{n-2},$$

которое справедливо при всех $n > 2$. Поскольку вычислить D_1 и D_2 не представляет труда ($D_1 = a$; $D_2 = a^2 - bc$), полученное соотношение позволяет найти последовательно D_3, D_4 и так далее. Зачастую знание значений D_n при небольших n позволяет выдвинуть гипотезу о значении D_n для произвольного n и доказать эту гипотезу (в случае ее истинности).

Приведем два примера "угадывания".

1) $a = 2$; $b = c = 1$; $D_n = 2D_{n-1} - D_{n-2}$; $D_1 = 2$; $D_2 = 3$; $D_3 = 2D_2 - D_1 = 4$. Гипотеза: $D_n = n + 1$, которая элементарно проверяется с помощью метода математической индукции.

2) $a = b = c = 1$; $D_n = D_{n-1} - D_{n-2}$; $D_1 = 1$; $D_2 = 0$; $D_3 = D_2 - D_1 = -1$; $D_4 = D_3 - D_2 = -1$; $D_5 = D_4 - D_3 = 0$; $D_6 = D_5 - D_4 = 1$; $D_7 = D_6 - D_5 = 1$; $D_8 = D_7 - D_8 = 0$. Получили последовательность значений 1; 0; -1; -1; 0; 1; 1; 0. Два последних совпадают с двумя первыми, а поскольку принцип поиска следующего одинаков всюду, дальше числа будут периодически (с периодом 6) повторяться. Для нахождения D_n надо поделить n на 6 с остатком. Если остаток равен 0 или 1, $D_n = 1$; если остаток равен 2 или 5, $D_n = 0$; если остаток равен 3 или 4, $D_n = -1$. В принципе можно написать формулу, по которой можно искать D_n ($D_n = \frac{2}{\sqrt{3}} \cos \frac{(2n-1)\pi}{6}$), но вряд ли использование этой формулы кому-нибудь принесет удовольствие.

К сожалению, метод угадывания срабатывает крайне редко. В этом случае приходится применять "тяжёлую артиллерию". Приведем здесь нужную теорию.

Рассмотрим множество $\mathcal{L}$ всевозможных посдедовательностей $\{u_n\}$, удовлетворяющих при всех $n > 2$ условию

$$u_n = au_{n-1} - bcu_{n-2}.$$

Одной из этих последовательностей является искомая последовательность определителей D_n. Замечаем, что если $\{u_n\} \in \mathcal{L}$ и $\{v_n\} \in \mathcal{L}$, то и $\lambda\{u_n\} + \mu\{v_n\} = \{\lambda u_n + \mu v_n\} \in \mathcal{L}$. Поэтому $\mathcal{L}$ является линейным подпространством линейного пространства всех последовательностей, и поэтому само является линейным пространством. В качестве базиса этого пространства можно взять последовательности $\{e_n\}$ и $\{f_n\}$, начинающиеся с $e_1 = 1$; $e_2 = 0$; $f_1 = 0$; $f_2 = 1$. Произвольная последовательность $\{u_n\} \in \mathcal{L}$ раскладывается по этому базису: $\{u_n\} = u_1\{e_n\} + u_2\{f_n\}$. Таким образом, $\dim \mathcal{L} = 2$. Дальнейшая наша деятельность сводится к угадыванию двух решений рекуррентного соотношения.Постараемся найти хотя бы одно решение в виде $\{u_n = x^n\}$ при некотором $x \neq 0$ (при $x = 0$ получаем нулевую последовательность, которая нас мало интересует). Получаем $x^n = ax^{n-1} - bcx^{n-2}$, а после сокращения на x^{n-2} и перенесения налево возникает квадратное уравнение

$$x^2 - ax + bc = 0.$$

Возможны три случая – когда дискриминант этого уравнения положителен, отрицателен или равен нулю. Рассмотрим их отдельно.

1) Самый простой случай, если дискриминант этого уравнения положителен: $a^2 - 4bc > 0$. В этом случае мы находим два различных значения x_1 и x_2, что дает два линейно независимых элемента из $\mathcal{L}$ – $\{x_1^n\}$ и $\{x_2^n\}$. Нужная нам последовательность значений определителя будет находиться в виде линейной комбинации этих двух последовательностей. При этом нам понадобятся начальные условия – значения D_1 и D_2. А именно,

$$D_n = C_1 x_1^n + C_2 x_2^n;$$

C_1 и C_2 ищутся из системы

$$\begin{cases} C_1 x_1 + C_2 x_2 = D_1 \\ C_1 x_1^2 + C_2 x_2^2 = D_2 \end{cases}$$

Полезное замечание. Вместо последовательностей $\{x_1^n\}$ и $\{x_2^n\}$ можно рассмотреть последовательности $\{x_1^{n-1}\}$ и $\{x_2^{n-1}\}$, полученные из ранее выбранных делением на x_1 и x_2 соответственно. Конечно, они также образуют базис нашего линейного пространства. В этом случае D_n ищется как

$$D_n = C_1 x_1^{n-1} + C_2 x_2^{n-1},$$

причем

$$\begin{cases} C_1 + C_2 = D_1 \\ C_1 x_1 + C_2 x_2 = D_2 \end{cases}$$

Пример. $a = 3$; $b = 2$; $c = 1$; $x^2 - 3x + 2 = 0$; $x_1 = 2$; $x_2 = 1$ (можно обратить внимание, что $x_1 = b$, $x_2 = c$; это произошло из-за того, что $a = b+c$; подумайте сами, почему так происходит); $D_n = C_1 \cdot 2^{n-1} + C_2$; $\begin{cases} C_1 + C_2 = 3 \\ 2C_1 + C_2 = 7 \end{cases}$; $C_1 = 4$; $C_2 = -1$; $D_n = 4 \cdot 2^{n-1} - 1$; $D_n = 2^{n+1} - 1$.

2) В качестве второго случая разберем случай с отрицательным дискриминантом. Формально ничего не меняется (на всякий случай: мы считаем, что все элементы матрицы – действительные числа). Надо только подчеркнуть, что ответ выражается с помощью комплексных чисел, которые после преобразований должны исчезнуть.

Пример. $a = b = c = 1$; $x^2 - x - 1 = 0$; $x_{1,2} = \frac{1 \pm i\sqrt{3}}{2} = e^{\pm i\frac{\pi}{3}}$; $D_n = C_1 e^{i\frac{(n-1)\pi}{3}} + C_2 e^{-i\frac{(n-1)\pi}{3}}$;
$\begin{cases} C_1 + C_2 = 1 \\ C_1 e^{i\frac{\pi}{3}} + C_2 e^{-i\frac{\pi}{3}} = 0 \end{cases}$; $C_1 = -C_2 e^{-i\frac{2\pi}{3}} = C_2 e^{i\frac{\pi}{3}}$; $C_2(1 + e^{i\frac{\pi}{3}}) = 1$; $C_2(1 + \cos\frac{\pi}{3} + i\sin\frac{\pi}{3}) = 1$;
$C_2(2\cos^2\frac{\pi}{6} + 2i\sin\frac{\pi}{6}\cos\frac{\pi}{6}) = 1$; $2C_2\cos\frac{\pi}{6} \cdot e^{i\frac{\pi}{6}} = 1$; $C_2 = \frac{1}{\sqrt{3}} \cdot e^{-i\frac{\pi}{p}}$; $C_1 = \frac{1}{\sqrt{3}} e^{i\frac{\pi}{6}}$;
$D_n = \frac{1}{\sqrt{3}} e^{i\frac{\pi}{6}} \cdot e^{i\frac{(n-1)\pi}{3}} + \frac{1}{\sqrt{3}} e^{-i\frac{\pi}{6}} \cdot e^{-i\frac{(n-1)\pi}{3}} = \frac{2}{\sqrt{3}} \cdot \frac{e^{i\frac{(2n-1)\pi}{6}} + e^{-i\frac{(2n-1)\pi}{6}}}{2} = \frac{2}{\sqrt{3}} \cdot \cos\frac{(2n-1)\pi}{6}$

3) Переходим к последнему случаю – когда дискриминант равен нулю, то есть $a^2 = 4bc$; $x_{1,2} = \frac{a}{2}$. Поэтому вместо двух линейно независимых элементов линейного пространства $\mathcal{L}$ мы имеем только один элемент $\{(\frac{a}{2})^n\}$ (как мы помним, возможно выгоднее взять $\{(\frac{a}{2})^{n-1}\}$

Рассмотрим сначала частный случай $a = 2$; $b = c = 1$ (естественно, можно было взять любые другие b и c, лишь бы их произведение равнялось 1). Этот пример был разобран уже раньше. Мы тогда просто угадали ответ $D_n = n + 1$ и отметили, что его истинность можно доказать методом математической индукции.

В общем же случае, разделив матрицу на $\frac{a}{2}$ (при этом определитель разделится на $(\frac{a}{2})^n$), мы получим матрицу, рассмотренную как частный случай, чей определитель мы вычислять умеем. Поэтому определитель исходной матрицы равен $D_n = (\frac{a}{2})^n(n + 1)$

Педантам, которые обратили внимание на то, что мы делили матрицу на число, не проверив случай, когда это число равно нулю, сообщаем, что мы думали на эту тему и решили, что Вы разберетесь в ней самостоятельно. Отметим только, что формула $D_n = (\frac{a}{2})^n(n+1)$ остается справедливой и в этой ситуации.

4 лекция. Обратная матрица. 24.09.2019

Сначала введем дополнительный термин, который позволит красиво формулировать последующие утверждения.

Квадратная матрица A называется *невырожденной*, если $\det A \neq 0$.

Матрица (которую будем обозначать A^{-1}) называется *обратной* к матрице A, если

$$A \cdot A^{-1} = A^{-1} \cdot A = E.$$

Если обратная матрица к матрице A существует, то A называется *обратимой.*

Первое очевидное замечание. Обратимая матрица обязана быть квадратной.

Второе (почти) очевидное замечание. Обратимая матрица невырождена (совсем недавний термин. Вспоминаем: невырожденная - означает $\det A \neq 0$). В самом деле, $A \cdot A^{-1} = E \Rightarrow \det(A \cdot A^{-1}) = \det E; \det A \cdot \det A^{-1} = \det E = 1$.

Теорема Если A-обратимая матрица, то обратная к ней матрица - единственна.

В самом деле, пусть B и C обратны к A. Тогда в силу ассоциативности умножения матриц $B(AC) = (BA)C$; $BE = EC$; $B = C$.

Мы пока выяснили, что если матрица обратима, то она квадратная, и ее определитель отличен от нуля. Оказывается, верно и обратное, то есть мы имеем критерий обратимости матрицы.

Теорема A - обратимая матрица тогда и только тогда, когда она квадратная и невырожденная. Короче: $A^{-1}\exists \iff \det(A) \neq 0$.

При этом

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \ddots & \ddots & \ddots & \ddots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix}^{\top}$$

Теорему можно переформулировать таким образом: A обратима тогда и только тогда, когда она невырождена.

Доказательство теоремы в одну сторону мы уже знаем, в обратную - достаточно перемножить матрицы A и "претендента"на роль обратной матрицы. При этом придется вспомнить формулы разложения определителя по строке (столбцу) и теорему о ложном определителе.

Бывает такая ситуация, когда известно только, что $AB = E$. В этом случае гарантировать, что матрица A обратима, и B – матрица, обратная к A, невозможно. Простейший пример:

$A = (1\ 0)$; $B = (1\ 0)^{\top} \Rightarrow AB = (1) = E$ (единичная матрица первого порядка).

Кстати, в ситуации $AB = E$ матрица B называется правой обратной к A, A называется левой обратной к B.

Но: если известно, что A – квадратная матрица, ситуация упрощается.

Теорема Если A – квадратная матрица, причем $AB = E$, то A обратима, и $B = A^{-1}$.

В самом деле, раз A – квадратная и E – квадратная, то и B квадратная того же размера, что и A, а тогда $\det(AB) = \det A \cdot \det B = \det E = 1 \Rightarrow \det A \neq 0 \Rightarrow \exists A^{-1}$. Умножим обе части равенство $AB = E$ слева на A^{-1}; получим $A^{-1}AB = A^{-1}E$; $B = A^{-1}$.

Естественно, можно сформулировать аналогичную теорему с $BA = E$.

Свойства обратной матрицы.

1. Если A и B – обратимые матрицы одинакового порядка, то AB – обратимая матрица, и

$$(AB)^{-1} = B^{-1}A^{-1}$$

Для доказательства достаточно перемножить AB и $B^{-1}A^{-1}$:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E$$

Так же просто доказываются и другие свойства.

2. Если A – обратимая матрица и $\lambda \neq 0$, то λA – обратимая матрица, и

$$(\lambda A)^{-1} = \frac{1}{\lambda}A^{-1}$$

3. Если A – обратимая матрица, то A^{-1} – обратимая матрица, и

$$(A^{-1})^{-1} = A$$

4. Если A – обратимая матрица, то

$$\det(A^{-1}) = \frac{1}{\det A}$$

В самом деле, $AA^{-1} = E \Rightarrow \det(AA^{-1} = \det E; \det A \cdot \det(A^{-1}) = 1$.

Решение матричных уравнений $AX = B$ и $XA = B$ с невырожденной матрицей A.

$$AX = B \Rightarrow X = A^{-1}B$$

$$XA = B \Rightarrow X = BA^{-1}$$

$$AXC = B \Rightarrow X = A^{-1}BC^{-1}$$

СЛАУ (Системы линейных алгебраических уравнений)

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \ldots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_n \end{cases} \Leftrightarrow \boxed{AX = B}\text{ , где}$$

$$A = \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{mn} \end{bmatrix}, X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Метод Крамера

Пусть число уравнений равно числу неизвестных, то есть $m = n$, причем $\det(A) \neq 0$ В этом случае $\exists A^{-1}$, и $X = A^{-1}B$. Распишем это произведение:

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} A_{11} & \ldots & A_{n1} \\ A_{12} & \ldots & A_{n2} \\ \ddots & \ddots & \ddots \\ A_{1n} & \ldots & A_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \Rightarrow x_i = \frac{1}{\det(A)}(b_1A_{1i} + b_2A_{2i} + \cdots + b_nA_{ni}) = \frac{\triangle_i}{\det(A)}\text{ , где}$$

$\triangle_i$ -определитель, полученный из матрицы A заменой i-го столбца на столбец свободных членов.

5–6 лекция. Комплексные числа

Как известно, действительных чисел не хватает для "полного счастья". В каком-то смысле полного счастья мы добиваемся, расширяя множество действительных чисел до большего множества - множества комплексных чисел. Причем множество действительных чисел – это не просто множество, это так называемое *поле.* Переходя к большему множеству, нам бы хотелось, чтобы это большее множество оставалось полем. Конкретнее, мы сначала сформулируем определение поля, потом убедимся, что множество действительных чисел удовлетворяет всем аксиомам поля, затем расширим поле действительных чисел до множества комплексных чисел и убедимся, что и это множество является полем.

Определение. Поле – это такое множество F (F - первая буква слова field – поле)(состоящее хотя бы из двух элементов), в котором заданы две операции – сложение и умножение, причем эти операции должны удовлетворять следующим аксиомам:

1) Коммутативность сложения: $a + b = b + a$.

2) Ассоциативность сложения: $a + (b + c) = (a + b) + c$.

3) Наличие нулевого элемента (нуля; нейтрального элемента относительно сложения) – это такой элемент 0, что $a + 0 = 0 + a = a$ (ноль плюс a равно a можно и не писать благодаря первой аксиоме).

4) Наличие для каждого элемента a противоположного элемента (который мы будем обозначать как $-a$), то есть такого элемента, что $a + (-a) = (-a) + a = 0$ (конечно, $(-a) + a = 0$ можно не требовать, сославшись снова на свойство 1).

5) Коммутативность умножения: $ab = ba$.

6) Ассоциативность умножения: $a(bc) = (ab)c$.

7) Наличие единичного элемента (единицы; нейтрального элемента относительно умножения) – это такой элемент 1, что $a \cdot 1 = 1 \cdot a = 1$ (когда мы с самого начала говорили, что в поле должно быть хотя бы два элемента, мы и имели в виду, что там есть 0 и 1, и они не равны друг другу).

8) Дистрибутивность умножения относительно сложения: $a(b + c) = ab + ac$; $(a + b)c = ac + bc$.

9) Для каждого $a \neq 0$ должен существовать *обратный* элемент a^{-1}, то есть такой элемент, что $aa^{-1} = a^{-1}a = 1$.

То, что множество действительных чисел образует поле, думаю, сомнений не вызывает. Кстати, это поле обозначается $\mathbf{R}$.

Действительные числа обычно изображаются точками горизонтальной прямой с выбранным началом отсчета, масштабом и выбранным направлением (то есть осью OX). Комплексные числа будут изображаться точками плоскости с выбранной прямоугольной декартовой системой координат OXY, при этом действительные числа будут образовывать подмножество множества комплексных чисел, располагаясь на оси OX. Итак, с нашей точки зрения комплексные числа - это точки плоскости, иными словами - пары действительных чисел $(x; y)$ (координат этих точек). Пока мы не ввели операции, это просто плоскость, и больше ничего, или просто множество всевозможных наборов вида $(x; y)$, где x и y пробегают все действительные значения.

Операция сложения вводится самым естественным образом – покоординатно:

$$(x_1; y_1) + (x_2; y_2) = (x_1 + x_2; y_1 + y_2)$$

Кстати, никто не может помешать нам идентифицировать комплексные числа не с точками плоскости, а с векторами – радиус-векторами этих точек. Тем самым сложение комплексных чисел определяется с помощью сложения векторов.

Очень важно, что введенная операция не противоречит операции сложения действительных чисел, которые, повторим, лежат на горизонтальной прямой. Складывая числа x_1 и x_2 как элементы нового множества, мы должны складывать пары $(x_1;0)$ и $(x_2;0)$; получаем пару $(x_1+x_2;0)$, которой соответствует действительное число x_1+x_2.

Операцию умножения на первый взгляд мы вводим по очень странной формуле. Нас оправдывает глобальная цель – ввести операцию таким образом, чтобы среди комплексных чисел существовали такие, которые в квадрате равнялись бы минус единице. Итак, по определению

$$(x_1;y_1)(x_2;y_2)=(x_1x_2-y_1y_2;x_1y_2+x_2y_1)$$

Рассмотрим некоторые частные случаи этой формулы.

1) $(x_1;0)(x_2;0)=(x_1x_2-0\cdot 0;x_1\cdot 0+x_2\cdot 0)=(x_1x_2;0)$.

Таким образом эта странная формула по крайней мере "не портит"старое умножение действительных чисел.

2) $(x_1;0)(x_2;y_2)=(x_1x_2;x_1y_2)$.

Таким образом, при умножении комплексного числа на то, что раньше было действительным числом, обе координаты соответствующего вектора умножаются на это действительное число.

3) $(0;1)(0;1)=(0-1;0)=(-1;0)$

Таким образом, пара $(0;1)$ в квадрате равна паре $(-1;0)$, которая раньше была действительным числом -1.

Осталось совсем немного, чтобы перейти к более известному обозначению комплексных чисел. По поводу того, что нам хотелось бы пары $(x;0)$ записывать в виде x мы фактически уже высказывались. В частности пара $(1;0)$ в упрощенном виде будет записываться просто как 1. Остается ввести обозначение для пары $(0;1)$:

$$(0;1)=i.$$

Отсюда $i^2=i\cdot i=(0;1)(0;1)=(-1;0)=-1$, то есть

$$i^2=-1$$

Дальше всё просто: произвольную пару $(x;y)$ можно представить в виде

$$(x;y)=(x;0)+(0;y)=x+y(0;1)=x+yi=x+iy$$

В этом случае "странное"умножение комплексных чисел принимает более комфортную форму:

$(x_1+iy_1)(x_2+iy_2)=(x_1x_2-y_1y_2;x_1y_2+x_2y_1)=x_1x_2-y_1y_2+i(x_1y_2+x_2y_1)$

Иными словами, при перемножении комплексных чисел, записанных с помощью обозначения i для пары $(0;1)$, мы раскрываем скобки по дистрибутивному закону, а i^2 заменяем на -1.

Проверку того, что множество комплексных чисел удовлетворяет всем аксиомам поля, оставляем слушателям. Естественно, действительный ноль становится и нулем комплексным, действительная единица становится единицей комплексной. Единственная аксиома, которая может вызвать затруднение - последняя, о существовании обратного элемента для всех чисел, кроме нуля. Чтобы доказать это, нам понадобится определение комплексно-сопряженного числа.

Да, чтобы не писать всегда $x+iy$, мы будем комплексные числа обозначать одной буквой, чаще всего z (иногда w). При этом x назывется действительной частью z и обозначается $x=\operatorname{Re} z$, y называется мнимой частью z и обозначается $y=\operatorname{Im} z$. Если $z=x+iy$, *сопряженным* к нему (или *комплексно-сопряженным*) будем называть число $\bar{z}=x-iy$. На плоскости эти числа расположены в точках, симметричных друг другу относительно горизонтальной оси. Заметим, что

$$z\bar{z}=(x+iy)(x-iy)=x^2-i^2y^2=x^2+y^2$$

Теперь предъявить число, обратное к ненулевому числу $z=x+iy$, очень просто:

$$z^{-1}=\frac{1}{z}=\frac{\bar{z}}{z\bar{z}}=\frac{x-iy}{x^2+y^2}=\frac{x}{x^2+y^2}-i\frac{y}{x^2+y^2}$$

Без труда мы теперь делим одно комплексое число на другое, естественно, ненулевое:

$$\frac{z_1}{z_2}=z_1z_2^{-1}=\frac{z_1\bar{z}_2}{z_2\bar{z}_2}$$

Модуль и аргумент комплексного числа.

Как любой уважающий себя модуль, модуль $|z|$ комплексного числа равен расстоянию от 0 до z.
Каждый знающий теорему Пифагора сообразит, что

$$|z| = |x + iy| = \sqrt{x^2 + y^2} = \sqrt{z\bar{z}}$$

При этом искать модуль комплексного числа по этой формуле в простых случаях глупо. Так, без вычислений легко сообразить, что $|1| = |-1| = |i| = |-i| = 1$; $|0| = 0$; $|\pm 1 \pm i| = \sqrt{2}$; $|\pm 3 \pm 4i| = |\pm 4 \pm 3i| = 5$; $|\pm\cos\alpha \pm i\sin\alpha| = |\pm\sin\alpha \pm i\cos\alpha| = 1$ и т.д.

Аргумент комплексного числа z – это угол, на который нужно повернуть положительную полуось OX, чтобы она прошла через z. При этом поворот против часовой стрелки считается положительным, по часовой стрелке – отрицательным. Ясно, что аргумент определен неоднозначно – с точностью до $2\pi n$, где n – любое целое число. Если все эти значения свалить в одну кучу, получится $\operatorname{Arg} z$, если взять один конкретный – $\arg z$. Обычно под $\arg z$ понимается или угол в пределах $[0; 2\pi)$, или угол в пределах $(-\pi; \pi]$. Для $z = 0$ аргумент не определен. Для нахождения аргумента (если ответ сразу не очевиден), можно воспользоваться (в случае, когда $x \neq 0$) тем, что его тангенс равен $\frac{y}{x}$. По значению тангенса угол однозначно восстановить невозможно, но если вспомнить, что арктангенс "обслуживает"углы в пределах $(-\pi/2; \pi/2)$, можно сказать, что в случае $x > 0$ в качестве аргумента можно взять $\operatorname{arctg}\frac{y}{x}$, а в случае $x < 0$ к $\operatorname{arctg}\frac{y}{x}$ нужно добавить (или вычесть) π. В качестве случаев, когда формула для нахождения аргумента не нужна, можно привести числа, лежащие на положительной полуоси OX – у них аргумент равен 0 (в частности аргумент 1 равен 0); числа, лежащие на отрицательной полуоси OX – у них аргумент равен π; числа, лежащие на положителдьной полуоси OY – у них аргумент равен $\pi/2$; числа, лежащие на отрицательной полуоси OY – у них аргумент равен $-\pi/2$ (при желании можно взять и $3\pi/2$). Если $x = \pm y$, $x = \pm y\sqrt{3}$; $y = \pm x\sqrt{3}$, любой, кто хотя бы чуть-чуть знает тригонометрию, с легкостью определит аргумент без всяких формул.

Тригонометрическая и показательная формы комплексного числа.

Перед введением новых форм комплексного числа назовем старую форму комплексного числа *алгебраической*: $z = x + iy$. Пусть $z \neq 0$; тогда $|z| \neq 0$, и мы можем написать

$$z = x + iy = |z|\left(\frac{x}{|z|} + i\frac{y}{|z|}\right) = |z|(\cos\varphi + i\sin\varphi),$$

где φ – аргумент z. Получили тригонометрическую форму комплексного числа:

$$z = |z|(\cos\varphi + i\sin\varphi)$$

Для получения показательной формы комплексного числа не обойтись без нескольких пассов и ссылок на высшие силы – имеется в виду ссылка на так называемую *формулу Эйлера*, которую пока приходится принять на веру:

Формула Эйлера

$$e^{i\varphi} = \cos\varphi + i\sin\varphi$$

С учетом формулы Эйлера имеем:

$$z = |z|e^{i\varphi}$$

Это – *показательная* форма комплексного числа.
Заметим, что

$$|e^{i\varphi}| = 1;\ \arg(e^{i\varphi}) = \varphi.$$

Очевидно также, что

$$e^{i\pi/2} = i; \quad e^{i\pi} = -1; \quad e^{-\pi/2} = -i; \quad e^{0i} = e^{2\pi i} = 1.$$

Заметим также, что при переходе к комплексно сопряженному числу модуль не меняется, а аргумент меняет свой знак: если $z = re^{i\varphi}$, то $\bar{z} = re^{-i\varphi}$.

Умножение, деление и возведение в целую степень чисел, записанных в тригонометрической или показательной форме.

Для понимания последующих преобразований требуется знание основных тригонометрических формул.

$z_1 = |z_1|(\cos\varphi_1 + i\sin\varphi_1) = |z_1|e^{i\varphi_1}$, $z_2 = |z_2|(\cos\varphi_2 + i\sin\varphi_2) = |z_2|e^{i\varphi_2} \Rightarrow z_1z_2 = |z_1||z_2|(\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2 + i(\sin\varphi_1\cos\varphi_2 + \cos\varphi_1\sin\varphi_2) = |z_1||z_2|(\cos(\varphi_1+\varphi_2) + i\sin(\varphi_1+\varphi_2)) = |z_1||z_2|e^{i(\varphi_1+\varphi_2)}$.

Запишем ответ в более короткой показательной форме:

$$z_1z_2 = |z_1||z_2|e^{i(\varphi_1+\varphi_2)}$$

Посмотрев на эту формулу, можно результат красиво сформулировать так: при перемножении комплексных чисел модули перемножаются, а аргументы складываются:

$$|z_1z_2| = |z_1||z_2|; \quad \arg(z_1z_2) = \arg z_1 + \arg z_2.$$

Отдельно выводить формулу, по которой комплексные числа делятся одно на другое, не требуется, так как деление мы можем свести к умножению: $z_1/z_2 = z_3$ означает, что $z_1 = z_2z_3$, откуда $|z_1| = |z_2z_3| = |z_2||z_3| \Rightarrow |z_3| = |z_1|/|z_2|$; $\arg z_1 = \arg(z_2z_3) = \arg z_2 + \arg z_3 \Rightarrow \arg z_3 = \arg z_1 - \arg z_2$. Таким образом,

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}; \quad \arg\frac{z_1}{z_2} = \arg z_1 - \arg z_2$$

То есть при делении одного комплексного числа на другое модули делятся, а аргументы вычитаются,

$$\frac{z_1}{z_2} = \frac{|z_1|}{|z_2|}e^{i(\varphi_1-\varphi_2)}$$

Частный случай этой формулы – когда числитель равен 1, то есть когда мы ищем число, обратное к z:

$$\frac{1}{z} = \frac{1}{|z|}e^{-i\varphi}$$

Заметим, что в случае, когда $|z| = 1$, то $\frac{1}{z} = e^{-i\varphi} = \bar{z}$.

Из всего полученного следует формула возведения в целую степень, которая не требует отдельного вывода:

$$z^n = |z|^n e^{in\varphi}; \; n \in \mathbf{Z}$$

Извлечение корня из комплексного числа.

Определим $\sqrt[n]{z}$ как любое число w, которое в n-й степени равно z:

$$w = \sqrt[n]{z} \Leftrightarrow w^n = z.$$

Если $z = |z|e^{i\varphi}$; $w = |w|e^{i\psi}$, то $|w|^n e^{in\psi} = |z|e^{i\varphi}$, откуда $|w|^n = |z| \Rightarrow |w| = \sqrt[n]{|z|}$; $n\psi = \varphi + 2\pi k \Rightarrow \psi = \frac{\varphi+2\pi k}{n}$. Можно в эту формулу подставлять любое целое k, но на практике достаточно подставить n соседних целых чисел; после этого пойдут повторения. Чаще всего берут k от нуля до $n-1$. Окончательно:

$$\sqrt[n]{z} = \sqrt[n]{|z|}e^{i\frac{\varphi+2\pi k}{n}}; \; k = 0, 1, 2, \ldots, n-1$$

7–8 лекция. Многочлены

Числа, которые мы будем рассматривать в качестве коэффициентов, будут принадлежать тому или иному полю F – $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$. В следующем году будем рассматривать многочлены и над конечными полями.

Многочлен (или *полином*) – это выражение вида

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n,\ a_i \in F.$$

Понятно, что можно рассмотреть и функцию, задаваемую этой формулой: $f(x) = \sum\limits_{i=0}^{n} a_ix^i$. Далее мы докажем, что по значениям этой функции все коэффициенты определены однозначно, но только в случае бесконечного поля.

Многочлены можно естественным образом складывать и перемножать (а также умножать на числа; впрочем, умножение на число можно считать умножением на многочлен, сводящийся к числу).

Степенью многочлена $f(x)$ будем называть

$$\deg f(x) = \max\{n :\ a_n \neq 0\}$$

Как видим, степень многочлена равна нулю в случае, когда этот многочлен сводится к ненулевой константе. Степень нулевой константы для простоты будем считать равной $-\infty$.

Утверждение. Степень суммы двух многочленов не превосходит большей из степеней слагаемых, при этом, если степени разные, степень суммы равна большей степени. Степень произведения двух многочленов равна сумме степеней сомножителей. Таким образхом,

$$\deg(f(x) + g(x)) \leq \max\{\deg f(x); \deg g(x)\};$$
$$\deg(f(x) \cdot g(x)) = \deg f(x) + \deg g(x).$$

Доказательство здесь приводить лень, по большому счету оно очевидно и опирается на то, что в поле нет делителей нуля.

Деление с остатком.

Все знают про деление с остатком в кольце целых чисел. Аналогичная операция существует и в кольце многочленов с коэффициентами в поле (мы пока не акцентируем свое внимание на термине «кольцо», используя его правильно, но по сути заменяя им термин множество. На втором курсе термин кольцо, наряду с терминами группа и поле будет всречаться постоянно и уже осознанно).

Теорема. Для любых многочленов $f(x)$ и $g(x)$ (второй, конечно, должен быть отличен от нуля – ведь мы на него делим) существуют и притом единственные многочлены $q(x)$ и $r(x)$ такие, что выполнено равенство

$$f(x) = g(x)q(x) + r(x),$$

причем $\deg r(x) < \deg g(x)$ (как частный случай $r(x)$ может равняться нулю, и тогда $f(x) = g(x)q(x)$).

$\lhd$ Проще всего поиск $q(x)$ и $r(x)$ показать на примере, делением столбиком. А вот то, что они определены однозначно, несложно доказать. Пусть

$$f(x) = g(x)q_1(x) + r_1(x) = g(x)q_2(x) + r_2(x);$$

$$g(x)(q_1(x) - q_2(x)) = r_2(x) - r_1(x);$$

степень правой части меньше степени $g(x)$, и если бы $q_1(x)$ не равнялся $q_2(x)$, последнее равенство было бы невозможно. Итак, $q_1(x) = q_2(x)$, а тогда и $r_1(x) = r_2(x)$. $\rhd$

Важный частный случай – когда $g(x)$ первой степени. Будем считать, что старший коэффициент равен 1, а тогда можно считать, что $g(x) = x - c$ (мы пишем $-c$, а не c, чтобы корень многочлена равнялся c, а не $-c$). В этом случае $r(x)$ становится константой:

$$f(x) = (x - c)q(x) + r.$$

Подставив в это равенство вместо x число c, получаем

$$f(c) = r.$$

Таким образом доказан факт, известный во всем мире как

Теорема Безу. Остаток от деления многочлена $f(x)$ на двучлен $x - c$ равен значению этого многочлена в точке c.

Следствие из теоремы Безу. Число c является корнем многочлена $f(x)$ (то есть $f(c) = 0$) тогда и только тогда, когда $f(x)$ делится на $(x - c)$, то есть $f(x) = (x - c)q(x)$.

Следствие из следствия из теоремы Безу. Число корней ненулевого многочлена $f(x)$ не превосходит степени этого многочлена.

Применяя несколько раз следствие из теоремы Безу, можем получить цепочку

$$f(x) = (x - c_1)q(x) = (x - c_1)(x - c_2)h(x) = \ldots = (x - c_1)(x - c_2)\ldots(x - c_k)p(x),$$

где $p(x)$ корней уже не имеет, откуда c_1, c_2, $\ldots$, c_k – полный набор корней $f(x)$.

Учитывая, что среди корней многочлена могут быть одинаковые, можно разложение $f(x)$ написать в виде

$$f(x) = (x - c_1)^{m_1}(x - c_2)^{m_2}\ldots(x - c_t)^{m_t}p(x),$$

где $p(x)$ корней в поле F не имеет, а среди c_i нет одинаковых. В этом случае говорят, что корень c_i имеет кратность m_i. Если кратность больше 1, корень называется кратным, иначе – простым.

Для поиска кратных корней можно использовать следующее утверждение.

Теорема. Пусть F – это поле $\mathbf{C}$ или его подполе (скажем, $\mathbf{R}$). Число c является кратным корнем многочлена $f(x)$ тогда и только тогда, когда c является решением системы

$$\begin{cases} f(x) = 0 \\ f'(x) = 0 \end{cases}$$

При этом кратность c как корня $f'(x)$ на единицу меньше кратности c как корня $f(x)$.

В самом деле, если $f(x) = (x - c)^k q(x)$, причем $q(c) \neq 0$, то $f'(x) = k(x - c)^{k-1}q(x) + (x - c)^k q'(x) = (x - c)^{k-1}(kq(x) + (x - c)q'(x))$, причем, как легко увидеть, последняя скобка не обращается в ноль при подстановке c.

Нам пришлось оговорить, в каком поле справедлива эта теорема, так как при ее доказательстве используется, что поле имеет нулевую характеристику. Что это такое, узнаем в четвертом семестре (хотя Артамкин, думаю, расскажет об этом намного раньше).

Для осуществления деления $f(x)$ на $(x-c)$ (кроме деления столбиком) существует еще так называемая схема Горнера. Лично мне ближе деление столбиком, когда каждое действие производится осознанно, но многие предпочитают Горнера. Я приводить эту схему не собираюсь, но если у кого-то есть желание, он может сделать доклад.

Из следствия из следствия из теоремы Безу следует

Теорема (тем самым она является следствием из следствия из следствия из теоремы Безу). Если поле F бесконечно, то два многочлена тождественно совпадают только в случае, когда все их коэффициенты совпадают.

Действительно, если $f(x) \equiv g(x)$, то $f(x) - g(x) \equiv 0$, то есть разность многочленов имеет бесконечно много корней, а ненулевой многочлен не может иметь корней больше, чем его степень.

Основная теорема алгебры. Если $f(z)$ – многочлен степени $n > 1$ с коэффициентами из поля $\mathbf{C}$, то он имеет хотя бы один корень (естественно, вообще говоря, комплексный).

Предполагаемое доказательство – в докладе.

Следствие из основной теоремы алгебры и следствия из теоремы Безу. Многочлен n-й степени имеет ровно n корней (учитывая их кратность) и поэтому может быть разложен на линейные множители.

Доказательство фактически очевидно; студентам предлагается провести его сам остоятельно.

Чтобы построить мостик, соединяющий этот материал с материалом 4-го семестра, далим определение *алгебраически замкнутого поля* – это такое поле, для которого "справедлива основная теорема алгебры". Таким образом, поле комплексных чисел алгебраически замкнуто, а поля действительных чисел и рациональных чисел – нет.

Теорема Виета для многочлена n-ой степени. Пусть $f(z) = \sum\limits_{i=0}^{n} a_i z^i = a_n(z-z_1)(z-z_2)\dots(z-z_n)$ (среди скобок могут быть и одинаковые, так как корни могут быть кратные). Тогда

$$\sum_{i=1}^{n} z_i = -\frac{a_{n-1}}{a_n};$$
$$\sum_{i<j} z_i z_j = \frac{a_{n-2}}{a_n};$$
$$\sum_{i<j<k} z_i z_j z_k = -\frac{a_{n-3}}{a_n};$$
$$\dots\dots\dots\dots\dots$$
$$z_1 z_2 z_3 \dots z_n = (-1)^n \frac{a_0}{a_n}$$

Доказательство очевидно.

Формулы Виета имеют разнообразные применения, например, из них следует, что сумма корней n-й из любого числа равна нулю, что произведение всех корней нечетной степени из a равна a, произведение всех корней четной степени из a равна $-a$.

Также использование формул Виета позволяет вычислить значения тригонометрических функций, когда аргумент кратен 18 градусам, то есть кратен $\frac{\pi}{10}$. Имеем уравнение $z^5 - 1 = 0$; $z_k = e^{i\frac{2\pi k}{5}} = \cos\frac{2\pi k}{5} + i\sin\frac{2\pi k}{5}$, $i = 0,\ 1,\ 2,\ 3,\ 4$. Поскольку сумма корней равна нулю, равна нулю и сумма действительных частей (сумма мнимых частей ничего нам не дает):

$$1 + \cos\frac{2\pi}{5} + \cos\frac{4\pi}{5} + \cos\frac{6\pi}{5} + \cos\frac{8\pi}{5} = 0.$$

Применив простые преобразования и обозначив $\cos\frac{2\pi}{5} = t > 0$, получаем уравнение

$$4t^2 + 2t - 1 = 0;\ t = \frac{-1 \pm \sqrt{5}}{4};\ \cos\frac{2\pi}{5} = \cos 72^\circ = \sin 18^\circ = \frac{\sqrt{5}-1}{4}$$

Далее будем разбираться с многочленами с действительными коэффициентами.

Теорема. Если $c = \alpha + i\beta$, $\beta \neq 0$ – корень многочлена $P(z)$ с *действительными* коэффициентами, то $\bar{c} = \alpha - i\beta$ также является корнем этого многочлена.

Для доказательства этой теоремы нам потребуется простое утверждение (приведем в нем большее количество свойств, чем нам здесь потребуется).

Лемма.
(i) $\overline{z_1 \pm z_2} = \bar{z}_1 \pm \bar{z}_2$; (ii) $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$; (iii) $\overline{z^n} = (\bar{z})^n$; (iv) $\overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2}$. (v) $\bar{c} = c \Rightarrow c \in R$.

Первое и последнее свойства очевидны, остальные проще всего доказывать используя показательную форму комплексного числа.

После этой леммы доказать теорему очень просто: $P(c) = a_n c^n + \ldots + a_1 c + a_0 = 0 \Rightarrow \overline{P(c)} = \overline{a_n c^n} + \ldots + \overline{a_1 c} + \bar{a}_0 = \bar{a}_n \overline{c^n} + \ldots + \bar{a}_1 \bar{c} + \bar{a}_0 = a_n \bar{c}^n + \ldots + a_1 \bar{c} + a_0 = P(\bar{c}) = 0$.

Теорема. Любой многочлен положительной степени с действительными коэффициентами может быть разложен на линейные и квадратичные множители с отрицательным дискриминантом.

Доказательство почти очевидно: По основной теореме алгебры такой многочлен имеет хотя бы один корень c, вообще говоря, комплексный. Если он комплексный, $c = \alpha + i\beta$, $\beta \neq 0$, то $\bar{c} = \alpha - i\beta$ также является корнем, а тогда $P(z)$ имеет в своем разложении скобки $(z-c)$ и $(z-\bar{c})$. Перемножив их: $(z-c)(z-\bar{c}) = ((z-\alpha)-i\beta)((z-\alpha)+i\beta) = (z-\alpha)^2+\beta^2 = z^2-2z\alpha+\alpha^2+\beta^2$, получаем, что $P(z)$ раскладывается в произведение этого квадратного трехчлена с действительными коэффициентами и отрицательным дискриминантом, и другого многочлена, с которым операцию можно повторить.

Кстати, из этой теоремы следует, что кратности корней c и $\bar{c}$ многочлена с действительными коэффициентами совпадают.

Способы разложения многочлена на множители.

1) *Угадывание целых корней многочлена с целыми действительными коэффициентами.*

Пусть $f(x) = \sum\limits_{i=0}^{n} a_i x^i$ – многочлен с целыми коэффициентами, $x_0 = p$ – целое число. Если x_0 – корень многочлена, то x_0 – делитель свободного члена.

В самом деле, имеем:

$$a_n p^n + a_{n-1} p^{n-1} + \ldots + a_1 p + a_0 = p(a_n p^{n-1} + a_{n-1} p^{n-2} + \ldots + a_1) + a_0 = 0,$$

откуда a_0 оказывается произведением двух множителей, один из которых равен p.

2) *Угадывание рациональных корней многочлена с целыми действительными коэффициентами.*

Пусть $f(x) = \sum\limits_{i=0}^{n} a_i x^i$ – многочлен с целыми коэффициентами, $x_0 = \frac{p}{q}$ – несократимая дробь. Если x_0 – корень многочлена, то p – делитель свободного члена, а q – делитель старшего коэффициента. Если старший коэффициент равен 1, то отсюда следует, что все рациональные корни такого многочлена на самом деле являются целыми.

Доказательство этого факта лишь немного сложнее предыдущего. Подставляем вместо x дробь $\frac{p}{q}$, избавляемся от знаменателя, сначала выносим за скобку p в первых n слагаемых, откуда делаем вывод, что a_0 делится на p, затем возвращаемся к равенству без скобок и выносим за скобку q в последних n слагаемых. Делаем вывод, что a_n делится на q.

3) *Поиск кратных корней* описан выше.

4) *Поиск парных корней.* Если $x_0 \neq 0$ и $-x_0$ являются корнями многочлена, будем называть их *парными.* Если $x_0 = 0$ является корнем многочлена, будем называть его парным, если он кратный.

Теорема. Парные корни многочлена совпадают с корнями системы

$$\begin{cases} \dfrac{f(x)+f(-x)}{2} = 0 \\ \dfrac{f(x)-f(-x)}{2} = 0 \end{cases}$$

Система может испугать своим видом, но суть ее проста: мы по отдельности приравниваем к нулю четные и нечетные степени многочлена.

Доказательство.Если $f(x_0) = f(-x_0) = 0$, то их сумма и разность равны нулю, а тогда x_0 (как и $-x_0$) являются решениями системы. Если x_0 является решением системы, то, взяв сумму уравнений системы, получаем, что x_0 является корнем $f(x)$, а взяв разность – что $-x_0$ является корнем $f(x)$.

5) *Метод неопределенных коэффициентов.*
6) *Метод Кардано для решения уравнений третьей степени.*
7) *Метод Феррари решения уравнений четвертой степени.*
8) *Решение однородных уравнений.*
9) *Решение возвратных уравнений.*

9 - 12 лекция. Линейные пространства

Определение *линейного* (в другой терминологии – *векторного*) *пространства* над полем F (в этом семестре и в большей части следующего этим полем будет являться поле действительных чисел). Это – множество, в котором заданы две операции – сложение элементов (*векторов*) этого множества, и умножение элементов этого множества на элементы поля (*числа*). При этом должны быть выполнены 8 аксиом. Мы их выписывали на первой лекции, когда говорили о линейном пространстве матриц $[m \times n]$.

Важные примеры линейных пространств – V^3; V^2; V^1 – множество векторов в трехмерном пространстве, на плоскости и на прямой соответственно (во втором семестре, когда будем изучать евклидовы пространства, эти же пространства будем обозначать, заменив букву V на E) с обычными операциями сложения и умножения на число; $\mathbf{R}^n$; $n \in \mathbf{N}$;$\mathbf{R}^{m\times n}$ – множество матриц данного размера; P_n – пространство многочленов степени не выше n; $F(D)$ – множество числовых функций, заданных на множестве D; $\mathbf{R}[x] = P$ – множество всех многочленов и т. д.

Следствия из аксиом.

1. Единственность нуля.
2. Единственность противоположного элемента.
3. $\lambda\bar{0} = \bar{0}$
4. $0x = \bar{0}$
5. $\lambda(-x) = -(\lambda x) = (-\lambda)x$
6. $(-1)x = -x$
7. Уравнение $a + x = b$ имеет единственное решение $b + (-a)$ (будем его обозначать $b - a$ и называть разностью b и a).
8. $\lambda(x - y) = \lambda x - \lambda y$
9. $(\lambda - \eta)x = \lambda x - \eta x$
10. Если $\lambda x = \bar{0}$, то или $\lambda = 0$, или $x = \bar{0}$.

Линейное подпространство. Непустое подмножество линейного пространства называется *линейным подпространством* линейного пространства L, если оно замкнуто относительно операций сложения векторов и умножения на число.

Теорема. Линейное подпространство само является линейным пространством относительно тех же операций. Доказательство почти очевидно.

Примеры подмножеств линейных пространств, являющихся или не являющихся линейными подпространствами – на лекции.

Линейная оболочка системы векторов $e_1,\ e_2,\ \ldots, e_n$ линейного пространства L – это множество $\mathcal{L}[e_1,\ e_2, \ldots, e_n]$ всевозможных линейных комбинаций этих векторов $\{\lambda_1 e_1 + \ldots + \lambda_n e_n\}$.

Теорема. Линейная оболочка системы векторов является линейным подпространством пространства L; более того, оно является наименьшим по включению линейным подпространством, содержащим векторы этой системы. Доказательство почти очевидно.

Кстати, нелишним сформулировать любимые выражения, которые постоянно будут встречаться в этой науке. Линейная комбинация системы векторов; вектор является линейной комбинацией векторов некоторой системы; вектор линейно выражается через векторы некоторой системы векторов.

Ясно, что следующие утверждения равносильны: вектор принадлежит линейной оболочке системы векторов: вектор является линейной комбинацией векторов системы; вектор линейно выражается через векторы этой системы.

Линейная зависимость и линейная независимость системы векторов.

Линейная комбинация $C_1 e_1 + \ldots + C_n e_n$ называется **тривиальной**, если $C_1 = \ldots = C_n = 0$. Она же называется **нетривиальной**, если она не является тривиальной (то есть если хотя бы один коэффициент отличен от нуля). Иногда тривиальность или нетривиальность комбинации записывают в виде условия $|C_1| + \ldots |C_n| = 0$ или $\neq 0$ соответственно.

Существует два эквивалентных определения линейной зависимости системы векторов. Первое интуитивно понятнее, но работает только для системы, в которой хотя бы 2 вектора. Кроме того, на практике работать с ним сложнее. Мы сформулируем оба определения и докажем их равносильность.

Первое определение линейной зависимости $n > 1$ векторов. Векторы $e_1,\ e_2,\ \ldots,\ e_n$ называются **линейно зависимыми** (или так: система векторов $[e_1,\ e_2,\ \ldots,\ e_n]$ называется линейно зависимой), если хотя бы один из этих векторов может быть линейно выражен через остальные (иными словами, хотя бы один из них является линейной комбинацией остальных).

Второе определение линейной зависимости.Векторы $e_1,\ e_2,\ \ldots,\ e_n$ называются **линейно зависимыми** (или так: система векторов $[e_1,\ e_2,\ \ldots,\ e_n]$ называется линейно зависимой), если существует нетривиальная комбинация этих векторов, равная нулевому вектору.

Теорема. Первое и второе определения линейной зависимости в случае $n > 1$ равносильны.

Доказательство. Если выполнено первое определение, то один из векторов может быть выражен через остальные. Перенеся все слагаемые налево, получим нетривиальную комбинацию, равную нулевому вектору.

Если выполнено второе определение, существует нетривиальная комбинация, равная неулевому вектору. Оставив слева один из векторов, коэффициент при котором отличен от нуля, перенеся все остальные векторы направо и разделив равенство на ненулевой коэффициент при векторе, оставшемся слева, получим, что этот вектор линейно выражается через остальные.

Система векторов называется **линейно независимой**, если она не является линейно зависимой.
Отсюда получаем два равносильных определения линейной независимости.

Первое определение линейной независимости $n > 1$ векторов. Векторы $e_1,\ e_2,\ \ldots,\ e_n$ называются **линейно независимыми** (или так: система векторов $[e_1,\ e_2,\ \ldots,\ e_n]$ называется линейно независимой), если ни один из этих векторов не может быть линейно выражен через остальные (иными словами, ни один из них не является линейной комбинацией остальных).

Второе определение линейной независимости.Векторы $e_1,\ e_2,\ \ldots,\ e_n$ называются **линейно независимыми** (или так: система векторов $[e_1,\ e_2,\ \ldots,\ e_n]$ называется линейно независимой), если не существует нетривиальная комбинация этих векторов, равная нулевому вектору; иными словами, все нетривиальные комбинации не равны нулевому вектору. А поскольку тривиальная комбинация векторов равна нулевому вектору, получаем такую формулировку: только тривиальная комбинация равна нулевому вектору. На практике это работает так: предполагаем, что комбинация векторов равна нулевому вектору, и доказываем, что все коэффициенты равны нулю.

Очевидные утверждения для систем геометрических векторов: система из одного вектора линейно зависима тогда и только тогда, когда она состоит из нулевого вектора. Для двух векторов – когда векторы коллинеарны. Для трех векторова – куогда векторы компланарны.

Леммы о линейной зависимости.

Лемма 1. Пусть к линейно независимой системе добавили вектор v. Получившаяся система линейно зависима тогда и только тогда, когдак $v \in \mathcal{L}[e_1, \ldots, e_n]$

Лемма 2. Даны две системы – из k векторов и m векторов, причем $k < m$, и векторы второй системы лежат в линейной оболочке первой системы. Тогда вторая система линейно зависима.

Перед доказательством второй леммы докажем элементарное утверждение.

Утверждение. Если $e_k \in \mathcal{L}[e_1, e_2, \dots, e_{k-1}]$, то $\mathcal{L}[e_1, \dots, e_k] = \mathcal{L}[e_1, \dots, e_{k-1}]$

$\triangleleft \ \triangleright$

Благодаря этому Утверждению можно предположить, что векторы $v_1, \dots, v_k$ первой системы линейно независимы. По условию векторы $u_1, \dots, u_m$ второй системы линейно выражаются через векторы первой системы. Запишем это в матричном виде $U = CV$, где U и V – столбцы одноименных векторов, C – матрица $m \times k$.

Требуется доказать, что существует нетривиальная комбинация векторов второй системы, равная нулевому вектору. Это можно записать в виде$\Lambda U = 0$, где Λ – строка коэффициентов, которую нужно подобрать. Отсюда $\Lambda(CV) = 0$; $(\Lambda C)V = 0$, а так как столбец V состоит из линейно независимых векторов, $\Lambda C = 0$. Чтобы стало привычней, протранспонируем это равенство: $C^\top \Lambda^\top = 0$. То, что у нас получилось, является системой линейных однородных уравнений, в которой число неизвестных больше числа уравнений. Такая система всегда имеет ненулевые решения. Мы об этом аккуратно не говорили, будем говорить позже, сейчас же предлагается в это, почти очевидное, утверждение поверить. Геометрически этот факт означает, что если в n-мерном пространстве провести менее n гиперплоскостей, проходящих через начало координат, то все они имеют хотя бы общую прямую.

Размерность линейного пространства $\dim L$ – это такое натуральное число n, что в этом пространстве существует n линейно независимых векторов, но нет $n+1$ линейно независимых векторов. В этом случае линейное пространство называется **конечномерным** (точнее, n-мерным). Если для любого n можно найти n линейно независимых векторов, то пространство называется **бесконечномерным**. Будем считать, что размерность пространства, состоящего из одного вектора (естественно, нулевого), равна нулю.

Базис линейного пространства. Можно давать разные определения базиса. Я в качестве основного выбираю то, которое чаще всего используют при конкретной проверке того, что данная система векторов является базисом.

Базис Б $= [e_1,\ e_2,\ \dots,\ e_n]$ (иногда мы будем использовать угольчатые скобки $< e_1, \dots, e_n >$ – это система векторов (на всякий случай подчеркнем – *упорядоченная* система векторов), такая, что:
1) Эти векторы линейно независимы;
2) любой вектор этого пространства линейно выражается через эти векторы (такая система называется полной)

Равносильное определение:
1) векторы линейно независимы;
2) их количество совпадает с размерностью пространства.

Равносильное определение:
1) линейно независимы;
2) каждая система, содержащая эту в качестве подсистемы, линейно зависима.

Равносильное определение:
1) любой вектор линейно выражается через эти векторы;
2) если выкинуть хотя бы один вектор из этой системы, то первое условие уже не будет выполнено (то есть это *минимальная* полная система)

Равносильное определение:
1) любой вектор линейно выражается через эти векторы;
2) и не как-нибудь выражается, а единственным образом.

Задача слушателям – доказать равносильность этих условий.

Будем пользоваться основным определением.

Теорема. Каждый вектор линейного пространства единственным образом раскладывается по векторам базиса.

Коэффициенты разложения вектора v по базису Б называются **координатами** вектора. Записав их в виде столбца X, получаем, что

$$v = \text{Б}X.$$

Теорема. Линейные операции с векторами приводят к линейным операциям с координатами этих векторов.

Естественные (канонические) базисы в линейных пространствах $\mathbb{R}$; $\mathbb{R}^n$; $\mathbb{R}^{m\times n}$; P_n.

Любое ненулевое число образует базис в $\mathbb{R}$. Каноническим естественно называть базис, образованный единицей.

Канонический базис в $\mathbb{R}^n$: $\text{Б} = [(1,0,0\ldots,0);(0,1,0,\ldots,0);\ldots;(0,0,\ldots,1)]$.

Канонический базис в $\mathbb{R}^{m\times n}$ – чаще всего в качестве такового берут базис из матриц E_{ij} (такая матрица, напомню, состоит из единицы на (i,j)-м месте и нулей на остальных местах), упорядоченных *лексикографически* (то есть сначала берут матрицы с единицей в первой строчке и т.д.).

Канонический базис в P_n – это $[1,\ t,\ t^2,\ \ldots,\ t^n]$ (иногда их располагают в обратном порядке).

Ранг системы векторов.

Ранг $\mathrm{rg}\,(v_1,v_2,\ldots,v_n)$ системы векторов – сколько среди этих векторов можно выбрать линейно независимых (то есть если $\mathrm{rg}\,(v_1,v_2,\ldots,v_n)=r$, то среди этих векторов можно выбрать r линейно независимых, но любая система из большего числа векторов будет уже линейно зависимой).

Почти очевидным является следующее утверждение.

Теорема 1. $\mathrm{rg}\,(v_1,v_2,\ldots,v_n) = \dim\mathcal{L}[v_1,v_2,\ldots,v_n]$

Доказательство. Пусть ранг системы векторов равен r, тогда среди них есть r линейно независимых векторов (пусть для определенности это будут первые r векторов системы), а остальные линейно через них выражаются . Докажем, что $\text{Б}=[v_1,\ldots,v_r]$ является базисом линейной оболочки всей системы. То, что они принадлежат линейной оболочке всей системы, не обсуждается, как и их линейная независимость. Остается доказать, что все векторы из линейной оболочки линейно выражаются через эти векторы. Сначала выражаем их через векторы всей системы, пользуясь определением линейной оболочки, после чего выражаем $v_{r+1},\ldots,v_n$ через первые r векторов, пользуясь первой леммой (если к линейно независимой системе добавляем вектор, то получившаяся система линейно зависима титт,к добавленный вектор линейно выражается через векторы линейно независимой системы). Ради забавы постараемся сделать эту выкладку "красиво".

Обозначим набор $[v_1,\ldots,v_n]$ через $<v>_n$; набор первых r векторов – через $<v>_r$.

$v\in\mathcal{L}[v_1,\ldots,v_n]\Rightarrow v=\sum\limits_{i=1}^{n}\lambda_iv_i=<v>_n\Lambda$, где Λ – столбец коэффициентов.

Поскольку векторы системы, начиная с $r+1$-го, лежат в линейной оболочке первых r, можем написать $v_{r+1}=<v>_r\Lambda_{r+1},\ldots,\ v_n=<v>_r\Lambda_n$, откуда

$v=(<v>_r\cdot[E_r|\Lambda_{r+1}|\ldots|\Lambda_n])\Lambda=<v>_r([E_r|\Lambda_{r+1}|\ldots|\Lambda_n]\Lambda)$

Теорема 2. Если $\mathcal{L}[<u_i>]\subseteq\mathcal{L}[<v_j>]$, то $\mathrm{rg}\ <u_i>\leq\mathrm{rg}\ <v_j>$ (на самом деле можно было говорить не о линейной оболочке первой системы, а о каждом векторе первой системы). Теорема очевидна, так как ранг системы совпадает с размерносьтью линейной оболочки, а если одна линейная оболочка лежит внутри другой, то очевидно, что ее размерностиь не может быть больше.

Определение. Две системы векторов называются **эквивалентными**, если векторы каждой системы лежат в линейной оболочке второй. Ясно, что это равносильно тому, что линейные оболочки этих систем совпадают. Поэтому следующая теорема очевидна.

Теорема 3. Ранги эквивалентных систем равны.

Ранг матрицы

Минор k-го порядка прямоугольной матрицы – это определитель матрицы, составленной из элементов, стоящих на пересечении выбранных k строк и k столбцов этой матрицы.

Ранг матрицы – это наибольший из порядков ее миноров, отличных от нуля. Если все миноры равны нулю, то есть если матрица состоит из одних нулей, то ранг считается равным нулю. Ясно, что ранг не может быть больше ни числа строк, ни числа столбцов.

Базисный минор матрицы – это любой ненулевой минор наивысшего порядка.

Теорема 4 (теорема о базисном миноре).
1) Строки матрицы, проходящие через базисный минор, линейно независимы как элементы $\mathbf{R}^n$;
2) любая строка матрицы, не проходящая через базисный минор, линейно выражается через строки, проходящие через базисный минор.

Доказательство. Пусть $\mathrm{rg}\, A = r$, и D – базисный минор. Для простоты предположим, что он расположен в верхнем левом углу. Строки, проходящие через D, линейно независимы, так как иначе одна из них линейно выражалась бы через остальные, а тогда определитель равнялся бы нулю.

Остается доказать, что остальные строки линейно выражаются через эти. Применим не совсем тривиальный трюк. Добавим к базисному минору еще одну строчку и еще один столбец. При этом - внимание! добавляется новая строка, но столбец добавляется или новый или старый - это важно. Получившийся определитель по любому равен нулю, поскольку если добавлен новый столбец, определитель будет равен нулю, так как это будет минор, порядок которого больше ранга матрицы. Если же добавлен старый столбец, то определитель равен нулю как определитель с равными столбцами.

Разложим определитель по последнему столбцу. Принципиальным является то, что алгебраические дополнения получаются вычеркиванием последнего столбца и поэтому от него не зависят. Приравняв это разложение к нулю, получив выражение элемента в правом нижнем углу (а там в принципе может оказаться любой элемент добавленной строки) через элементы этого столбца, лежащие в базисном миноре, причем эти коэффициенты от этого элемента не зависят (при этом мы используем, что коэффициент при этом элементе является базисным минором и поэтому отличен от нуля).

Следствие 1 Ранг матрицы равен рангу системы ее строк.

Следствие 2 – Критерий равенства определителя нулю. Определитель квадратной матрицы равен нулю тогда и только тогда, когда ее строки линейно зависимы.

Следствие 3. Определитель отличен от нуля титт,к ее строки линейно независимы.

Теорема 5. Ранг матрицы не изменяется при элементарных преобразованиях строк.

Доказательство очевидно.

Теорема 6. Ранг ступенчатой матрицы равен числу ступенек.

Доказательство очевидно.

Теорема 7. Ранг матрицы не меняется при ее транспонировании.

Доказательство очевидно.

Следствие. Все свойства ранга матрицы, доказанные для строк, справедливы и для столбцов.

Теорема о ранге произведения матриц.

$$\mathrm{rg}\,(AB) \leq \min\{\mathrm{rg}\, A; \mathrm{rg}\, B\}$$

Доказательство. Проанализируем, что такое произведение матриц - строчки левой умножаются на столбцы правой. Если обратить внимание сначала только на первый столбец матрицы AB, то он получается из строк матрицы A, умноженных на фиксированный, первый столбец матрицы B. При этом все элементы первого столбца матрицы A умножаются на b_{11}, элементы второго столбца – на b_{21} и т. д. Поэтому первый столбец матрицы AB – это линейная комбинация столбцов матрицы A с коэффициентами, равными элементам первого столбца матрицы B. Поэтому первый столбец матрицы AB лежит в линейной оболочке столбцов матрицы A. То же справедливо и для других столбцов матрицы AB. Таким образом, все они лежат в линейной оболочке, имеющей размерность, равную рангу матрицы A. Поэтому ранг произведения не превосходит ранга A. Для доказательства второй части можно, например, протранспонировать произведение матриц.

Следствие. Если одна из двух матриц невырождена, то ранг произведения равен рангу второй матрицы.

Легко следует из того, что можно вернуться к первоначальной матрице, умножив произведение на матрицу, обратную к невырожденной матрице.

Теорема 2. Любую линейно независимую систему векторов конечномерного линейного пространства можно дополнить до его базиса.

Почти очевидная теорема.

Следствие. Любой ненулевой вектор может быть включен в базис.

Переход от базиса к базису. Пусть в конечномерном линейном пространстве L заданы два базиса $Б_1 = e =< e_1, \dots, e_n >$ и $Б_2 = f =< f_1, \dots, f_n >$. Раскладывая векторы второго базиса по векторам первого базиса и записывая полученные координаты по столбцам, получим так называемую *матрицу перехода* C от первого базиса ко второму базису. В матричном виде связь базисов записывается как

$$Б_2 = Б_1 \cdot C.$$

В других обозначениях имеем $f = eC$, или $< f_1, \dots, f_n >=< e_1 \dots, e_n > C_{e\to f}$, или $Б_2 = Б_1 C_{12}$.

Свойства матрицы перехода.

1) Матрица перехода от базиса к базису невырождена, то есть ее определитель отличен от нуля.

2) $C_{e\to f} = C^{-1}_{f\to e}$

3) Обозначим столбцы координат некоторого вектора x в базисах e и f через X_e и X_f соответственно (то есть $x = eX_e = fX_f$). Тогда

$$X_f = C_{f\to e} X_e = C^{-1}_{e\to f} X_e$$

(соответственно $X_e = C_{e\to f} X_f = C^{-1}_{f\to e}$).

Таким образом, столбцы матрицы перехода от e к f – это координаты векторов второго базиса в первом базисе, то строчки матрицы перехода – это коэффициенты разложения старых координат через новые координаты.

4) Пусть $Б_1 = e$, $Б_2 = f$, $Б_3 = g$ – три базиса линейного пространства L. Тогда

$$C_{e\to g} = C_{e\to f} C_{f\to g};$$

в других обозначениях –

$$C_{13} = C_{12} C_{23}.$$

5) Пусть L – линейное пространство размерности n. Квадратная матрица $C \in \mathbf{R}^{n\times n}$ является матрицей перехода от базиса e к некоторому базису тогда и только тогда, когда $\det C \neq 0$.